To Graph or Not to Graph: The Missing Pieces for Knowledge Graphs in Sustainable Tourism

Stefan Neubig [1], Dominik Rebholz [2], Andreas Hein [3], Robert Keller [4], Helmut Krcmar [5]

Abstract: Sustainability is a critical challenge in modern tourism, exacerbated by climate change and globalization. Thanks to digitization, data-driven approaches constitute a key technology for addressing related issues, such as overtourism. However, the overarching complexity of the touristic data landscape, amplified by the interplay of diverse digital platform ecosystems, poses considerable challenges to both data owners and consumers. To mitigate such issues, knowledge graphs (KGs) have received significant attention. KGs focus on data quality by employing unified data models and continuous data refinements, making them well-suited for data-driven applications. Although promising, many challenges must be addressed to make KGs useful in practice. This paper overviews the state of the art of the field and identifies avenues for future research, explicitly focusing on touristic value and sustainability. Following our results, future research should focus on different areas, notably real-time knowledge graph population, distributed and parallelized processes, and ontologies for dynamic data types.

Keywords: Sustainable Tourism, Knowledge Graphs, Semantic Data Integration

1 Introduction

Sustainability is one of the key challenges of modern tourism in the current century. Battered by the adverse effects of climate change and globalization, overcoming challenges such as overtourism is even more critical today [Be20; Ca19]. During the worldwide Covid-19 pandemic, it became even more apparent that not only famous places such as Venice, Dubrovnik, and Barcelona are suffering from these problems [Ca19; GVT20], but they also increasingly apply to smaller destinations [KHL22]. Thanks to digitization, data-driven approaches are a key technology for tackling overtourism and

[1] Technical University of Munich, Boltzmannstr. 3, 85478 Garching & Outdooractive AG, Missener Str. 18, 87509 Immenstadt, stefan.neubig@tum.com, https://orcid.org/0000-0002-3794-7260
[2] University of Applied Sciences Kempten, WTZ, Bahnhofstraße 61, 87435 Kempten (Allgäu) & Fraunhofer Institute for Applied Information Technology FIT, Augsburg, dominik.rebholz@hs-kempten.de, https://orcid.org/0009-0003-8769-0365
[3] Technical University of Munich, Boltzmannstr. 3, 85478 Garching, andreas.hein@tum.de, https://orcid.org/0000-0001-9565-5840
[4] University of Applied Sciences Kempten, WTZ, Bahnhofstraße 61, 87435 Kempten (Allgäu) & Fraunhofer Institute for Applied Information Technology FIT, Augsburg, robert.keller@hs-kempten.de, https://orcid.org/0000-0001-7097-1724
[5] Technical University of Munich, Boltzmannstr. 3, 85478 Garching, helmut.krcmar@tum.de, https://orcid.org/0000-0002-2754-8493

related problems [Gr15; Li17; SR21; UN19]. Examples include the use of sensor technology to capture overtourism-relevant data (e.g., POI occupancy and visitor movement patterns [Bo22]) and the development of smart applications that make use of this data (e.g., raising visitor awareness and sustainable recommendations [Sc22]). Unfortunately, the complex nature of touristic data, resulting from both high connectivity as well as a wide range of stakeholders and intersecting digital platform ecosystems [He20], the creation of such applications is complicated. In practice, relevant data is often unavailable, located in isolated data silos, or suffering from uncertain data quality [SR21]. Consequently, this paper proposes using knowledge graphs (KGs) to mitigate these issues. We argue that KGs can effectively capture the complex nature of data in touristic ecosystems (including complex relationships among different entities), making them well-suited for supporting sustainable tourism by facilitating data integration and managing sustainability-related data.

However, most literature primarily focuses on the initial construction processes of KGs [TG23]. It often neglects crucial aspects of maintenance and lifecycle management (e.g., schema management, continuous data updates), a fundamental prerequisite for the widespread adoption of KG technology. Therefore, this paper presents what we consider the most pressing challenges in this regard. Our analysis draws on our previous work regarding the use of data and related service improvements in tourism [Ne22], our practical experiences with the touristic data landscape, and a systematic review of recent literature. More precisely, we aim to answer the following research question:

What are the key components and relationships in a knowledge graph lifecycle and maintenance model to support the management of touristic data?

We review recent literature to answer this research question and synthesize our findings to describe a conceptual perspective of an overall knowledge graph lifecycle model. By doing so, our paper contributes to both theory and practice. First, we provide an overview of the existing knowledge graph lifecycle and maintenance concepts. Second, we synthesize these concepts into a single process and support future research in positioning itself within the overall KG lifecycle. Third, we analyze the potentials and limitations of recent approaches concerning the heterogeneous touristic data landscape and identify the most promising avenues for future research. Besides, our findings support practitioners by evaluating whether knowledge graphs are a good choice for their use cases.

2 Background

Tourism is a complex domain, and so is its related data landscape. The tourism sector has also been an early adopter of semantic data models and has driven a widespread implementation of semantic web technologies. The following sections briefly overview touristic data types, data-driven use cases, and current data management challenges in tourism. Furthermore, we introduce the concept of knowledge graphs as a natural continuation of semantic web technology.

2.1 Touristic data landscape

In tourism, destinations are typically managed and promoted by destination management organizations (DMOs) that work with different stakeholders (e.g., government, businesses, local communities) and are responsible for brand development as well as the implementation of an appropriate tourism strategy [BLR05]. Recently, DMOs have also been increasingly faced with challenges regarding sustainability. In previous work [Ne22], we interviewed DMO representatives to gain insights into the data they collect, the insights they draw from it, and the related use cases they are pursuing to tackle touristic sustainability challenges. According to the identified use cases, there is a shift from static data to the increasing need for dynamic and real-time data, which is necessary for implementing use cases such as cross-DMO data sharing, or visitor guidance. Implementing these use cases at scale can significantly contribute to the sustainability of the overall ecosystem. For instance, the AIR research project [AIR23] focuses on implementing an AI-based recommender system aiming at a more sustainable distribution of visitors among hot and cold spots. Such efforts, however, strictly rely on the availability of real-time data and the broad participation of various DMOs. With many DMOs as data owners, however, data in the touristic ecosystem is highly fragmented and isolated [SR21]. Considering the German-speaking tourism area as an example, a couple of different approaches exist trying to defragment and integrate such data. Commercial service providers, which most DMOs use for managing large subsets of their data, recently initiated the so-called tourism tech alliance (TTA) [ND23], which aims to share data among these providers. Moreover, publicly funded data hubs, data spaces, and related research projects (e.g., [AIR23; Ba19; DAT23; DZT22]) have emerged to enable a standardized and open data exchange. However, since the overall practical adoption is slow and each approach above suffers from its restrictions, the broad availability and integration of touristic data is still in their infancy [SR21].

Another exacerbating problem that complicates data representation and integration lies in the nature of the data itself. Touristic data is heterogeneous [So16], and entities have highly dynamic relations to each other. For example, political and marketing regions co-exist and overlap. Unclear responsibilities lead to duplicates, such as creating multiple instances of the same point of interest (POI) (e.g., there are multiple instances of the Neuschwanstein castle in the Outdooractive platform [OA23]). Due to the increasing need for real-time data, future data management increasingly needs to cope with fast-arriving and frequently changing data types as well as with larger volumes of data. The next section introduces knowledge graphs as an enabling technology for standardized data exchange, even across the borders of a single DMOs. By utilizing knowledge graphs, data can be made broadly available, which allows the widespread implementation of use cases contributing to the sustainable development of the overall tourism ecosystem.

2.2 Knowledge graphs

Since a blog post by Google in 2012 [Si12], the term knowledge graph has received significant attention from companies and the research community. Many of its underlying ideas are borrowed from the semantic web domain [BL01], which aims at employing semantic technologies to create a web of data in which machines (i.e., intelligent agents [Ki21]) can "understand" the content of interlinked web pages. The emerging semantic web research community has since been developing various standards and technologies for this purpose, commonly referred to as the semantic web technology stack [Wc15]. The most notable concepts and technologies include ontologies [Gu09] (i.e., formal conceptualizations of real-world entities and relations), the resource description framework (RDF) [SR14] (i.e., a graph-based representation of facts in triple-form), the web ontology language [Hi12] (i.e., a family logic-based ontology languages), and the shape constraint language (SHACL) [KK17], which enforces specific structures in graph-based data models. Fig. 1 illustrates an exemplary touristic KG. The conceptual layer (i.e., ontological concepts) describes what general data types (i.e., tour, POI, category) exist and how they may be linked to each other (i.e., contains, has_category), the instance layer describes a possible manifestation of these concepts.

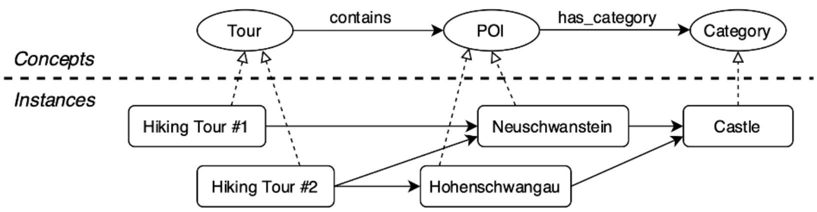

Fig. 1: Exemplary knowledge graph comprising tours, POIs, and POI categories

In a recent review, [Hi21] describes a focus shift within the research community from establishing a "web of data" to knowledge graphs, which represent a natural continuation of the development of the field with some shifts in emphasis, such as higher data quality, reduced openness, and industrial adoption. Considering the large number of related publications, however, the existing literature is surprisingly dividing on what a knowledge graph actually is. [Fä17], for example, define a knowledge graph as "any RDF graph" [Fä17]. In contrast, [VT17] provide a more restrictive definition by adding the requirement that RDF graphs must be linked to terminological knowledge (i.e., types) provided by an ontology. [Hi21] points out that knowledge graphs primarily differ from (possibly ontology-driven) RDF graphs in that they are more self-contained and usually benefit from better internal consistency. Following these definitions, knowledge graphs can be perceived as a new framing and emphasis on old (semantic web) concepts.

However, knowledge graphs have become relevant in various fields and are not limited to a semantic-web-based, technology-specific definition. Thus, [Ho22] provide a more general definition, describing a knowledge graph as "a graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities" [Ho22]. The authors explicitly state that their definition is meant to allow different graph models, not limited to RDF. However, [EW16] argue that such a general definition is not really helpful regarding implementing a KG-based system. Instead, while being technology-agnostic, the authors define a knowledge graph as a knowledge-based system that "acquires and integrates information into an ontology and applies a reasoner to derive new knowledge" [EW16]. In line with many of the discussed definitions, this paper considers knowledge graphs as graphs of real-world entities with an attached ontological schema, but without requiring a specific technology (i.e., a trade-off between flexibility and rigor). Similar to [Pa16], we do not expect the presence of a reasoner but require the underlying schema to capture ontological knowledge. Through the lens of sustainable tourism, this constraint is necessary due to the importance of cross-DMO data sharing and the related requirement of well-founded, standardized data models (e.g., [DZT22]).

3 Literature Review

We conducted a systematic literature review based on [WW02] to identify and analyze previous research. Following the interdisciplinary nature of our study, we chose Scopus as our primary literature database, given its comprehensive coverage across technical and domain-specific publications. To ensure capturing a wide range of relevant studies even beyond Scopus, we conducted a rich forward and backward search supported by tools like ResearchRabbit and Scite.ai. Our overall process (Fig. 1) consists of three steps. To retrieve articles that primarily focus on identifying literature on knowledge graph lifecycle and maintenance, we applied the search string *"knowledge graph" AND ("lifecycle" OR "maintenance")* to the title, abstract, and keyword fields. We deliberately chose not to use the keyword "tourism" in our search string to avoid limiting our search results to a narrow subset. After screening and pre-filtering the results by their titles and abstracts, we read and analyzed the remaining papers in-depth and complemented this process with a forward and backward search. We included any English article describing at least one potential phase that could be part of a KG lifecycle (e.g., creation, data updates, maintenance tasks). In contrast, we excluded papers that do not match our definition of a knowledge graph.

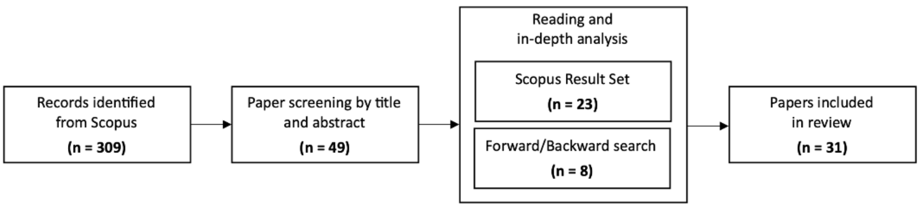

Fig. 2: Literature review methodology, including the number of hits at each step

4 Results

To obtain our overall results, we adhered to the open, axial, and selective coding guidelines presented by [WFW13]. More precisely, we followed an inductive approach to identify a set of key concepts (i.e., characteristics) related to creating or maintaining knowledge graphs. In a subsequent step, we grouped these characteristics to form a set of overarching categories (i.e., dimensions). The first category (data source, D1) involves the diverse data source types knowledge graphs are derived from and comprises structured data (e.g., relational databases), semi-structured data (e.g., documents such as JSON or XML), and unstructured data (e.g., free text or images). The second category (construction, D2) refers to the construction method of a (physical) knowledge graph, which may either follow a top-down or a bottom-up approach. In contrast, the third category (virtualization, D3) refers to knowledge graphs that do not follow any physical construction process but rely on virtualization. The fourth category (refinement, D4) refers to (optional) KG refinement steps (i.e., error detection or graph completion), which were present in most of the identified papers. Furthermore, we introduced a fifth category (tourism-specific, D5) to emphasize approaches developed explicitly for tourism-related use cases. However, only three publications ([Si21; SUF20; Si23]; all authors from the same institute) focused on a touristic use case. Fig. 4 depicts the final dimensions and characteristics.

<table>
<tr><th>Dimension</th><th colspan="6">Characteristics</th></tr>
<tr><td>D1: Data Source</td><td colspan="2">Unstructured (16)</td><td colspan="2">Semi-Structured (19)</td><td colspan="2">Structured (21)</td></tr>
<tr><td>D2: Construction</td><td colspan="3">Top-Down (18)</td><td colspan="3">Bottom-Up (11)</td></tr>
<tr><td>D3: Virtualization</td><td colspan="3">Fully Virtualized (3)</td><td colspan="3">Partly Virtualized (2)</td></tr>
<tr><td>D4: Refinement</td><td colspan="3">Error Detection (13)</td><td colspan="3">Graph Completion (10)</td></tr>
<tr><td>D5: Tourism-Specific</td><td colspan="3">Yes (3)</td><td colspan="3">No (28)</td></tr>
</table>

Fig. 3: Identified Dimensions and Characteristics during the coding phase
(Brackets indicate the number of representatives of a given characteristic)

4.1 KG lifecycle and maintenance process

Based on this coding, we synthesized different phases and activities to compile a comprehensive knowledge graph lifecycle model (Fig. 4), distinguishing between bottom-up and top-down construction iterations, as well as virtual knowledge graphs (VKGs). The bottom-up construction paradigm does not rely on a previously defined ontology but includes constructing such an ontology as an integral part of its process. In contrast, top-down approaches and VKGs use a previously defined ontology. Top-down construction methods are generally more complex, while VKGs suffer from the limitations of virtualization. Each type comprises various activities representing the fundamental common elements compiled from the analyzed literature. While some of these activities are considered required (i.e., are present in all investigated approaches), activities only present in a subset of the investigated approaches are considered optional (e.g., KG refinement methods). In the following, we describe the individual activities in more detail.

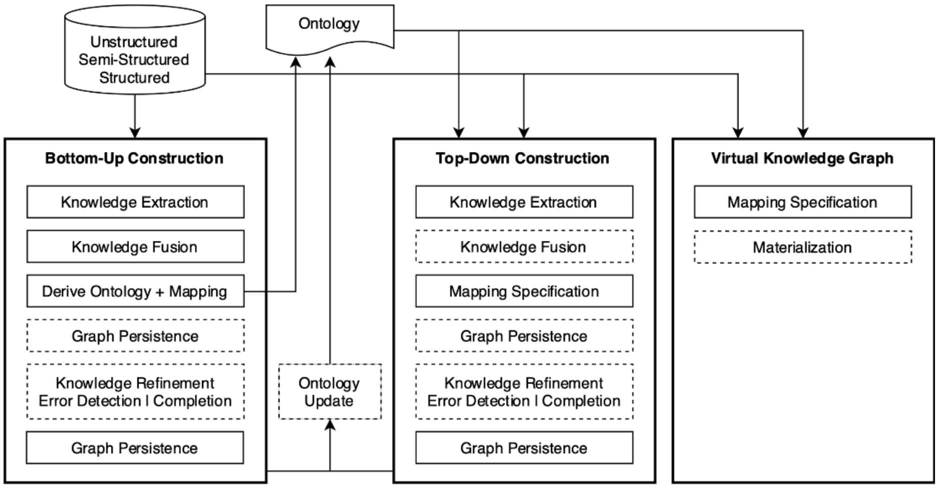

Fig. 4: Synthesized KG lifecycle model considering bottom-up and top-down construction processes, as well as virtual knowledge graphs
(Boxes with solid lines refer to required activities, dashed boxes refer to optional activities)

- *Knowledge Extraction*. Based on one or multiple data sources, the knowledge extraction activity aims at extracting entities, relations (linking the entities), and attributes [Ch22; La22; TG23; Wa21]. This activity is usually supported by various methods, including named-entity recognition (NER) or different natural language processing (NLP) methods. The reliable extraction of knowledge, especially from unstructured or semi-structured data, is a crucial part of knowledge graph creation and a very active field of research [Si23; TG23]. Note, however, that the extracted knowledge may still contain unconnected entities or isolated subgraphs (e.g., when different data sources are involved).

- *Knowledge Fusion*. In its most basic version, knowledge fusion refers to integrating the previously extracted data, thus, dissolving data silos that may have emerged due to using different data sources. Some authors extend this concept by employing error detection and error solving, such as removing duplicates and eliminating ambiguity [La22; TG23; Wa21], while others delay such activities to a dedicated knowledge refinement phase [Ch22; Pa16; SUF20]. Some authors primarily consider the knowledge fusion task for bottom-up approaches (e.g., [Ch22]) and assume that top-down approaches do not necessarily need this task due to the existence of an ontology that may sufficiently define appropriate links, even across data from different sources. The result of the knowledge fusion task is an integrated layer of knowledge without data silos.

- *Ontology and Related Mappings.* The bottom-up construction approach does not depend on a previously defined ontology, meaning its creation is integral to the process. Thus, in a bottom-up process, an ontology is created by analyzing the structure of the fused knowledge of the previous task and by deriving more general concepts [Ch22; TG23]. The ontology creation can be handcrafted or extracted in an automated way. In contrast, top-down construction approaches link the extracted data to the previously defined ontology based on a formal mapping specification [Fl22]. In the latter case, the main challenge lies in creating the mapping specification, which usually requires considerable manual effort. The result of this task is a layer of instances connected to an ontological schema.

- *Graph Persistence.* Knowledge graphs can be persisted in a range of different databases, including relational databases (RDBMS), triple stores (i.e., a specific type of an RDBMS), and graph databases [Si23; TG23]. While some authors do not consider the persistence of the graph at this stage (e.g., [Ch22]), other authors base the subsequent refinement task on capabilities specific to an individual (semantic) database system [Si23].
- *Knowledge Refinement.* Once a KG is readily established (persisted or not), it can be further refined and completed using various methods, including KG embeddings and link prediction [Ro21; WQW21; Zh20]. [Pa16] distinguishes approaches for graph completion from those for error detection. Traditionally, knowledge refinement tasks can be performed using logical reasoning or machine learning (ML), but even purely handcrafted approaches exist [Ji20]. Recent approaches [Ch20] leverage representational learning (i.e., learning low-dimensional feature vectors from raw graph data) to overcome scalability issues from a growing graph's increased data sparsity. [Ch20; WQW21] found such approaches to perform significantly better in terms of prediction performance. The result of the KG refinement task is persisted in an appropriate database (see graph persistence).

4.2 Virtual knowledge graphs

Even before the popularity of knowledge graphs, a sub-stream of research evolved focusing on the virtualization of ontologies and RDF graphs. So-called ontology-based data access (OBDA) approaches are often described as a triple $(\mathcal{O}, \mathcal{M}, \mathcal{S})$, where $\mathcal{M}$ is a mapping between a (primarily relational) source database with a schema $\mathcal{S}$ and a target ontology $\mathcal{O}$ [Xi18]. Later, with the increasing popularity of knowledge graphs, such systems were frequently referred to as virtual knowledge graphs (VKGs) [Xi19]. One advantage of VKGs is that data resides in its original databases and is always up to date. Although partly materialized approaches exist [SAM14; Vi18], these are usually limited to improving the performance of queries and do not change the overall behavior. Querying VKGs is usually based on query-rewriting techniques, that is, queries directed to the VKG (e.g., SPARQL [HS13]) are translated to queries specific to the source database. Thanks to the introduction of the OWL QL profile, such queries even support basic reasoning (e.g., hierarchical subclass relations) [KEK12]. VKGs, if fully virtualized, do not require any additional storage. On the other hand, VKGs are limited concerning their logical reasoning capabilities and put additional load on the original data sources.

4.3 Updates and maintenance

Updating the knowledge graph is required when the connected data sources change (e.g., new data is emitted, or additional data sources are added) or ontological changes are introduced. For VKGs, updates relate to the underlying mapping specifications and are only necessary if either the ontology or the database schema evolves; integrating new data added to the already mapped databases, however, does not require any action. For physical knowledge graphs, the update task involves performing a new iteration of the overall process [Ch20; Do21; TG23], including the extraction and (optional) fusion of new knowledge, a mapping to the previously defined or updated ontology, and possible employment of additional knowledge graph refinements. In contrast to the initial construction iteration, however, an ontology is guaranteed to exist. Therefore, subsequent iterations usually employ a top-down approach.

5 Discussion and Future Research Directions

As argued in the background section of this paper, a sustainable touristic ecosystem specifically benefits from a wide range of available and reliable data, particularly dynamic and real-time data, such as traffic, weather, or data specific to (local) overcrowding. To reflect such data in knowledge graphs, small and frequent data updates are therefore unavoidable. According to our findings, however, most literature mainly deals with the original creation of a knowledge graph but falls short in providing clear methodological or technical guidelines regarding its subsequent maintenance and data updates. Also, literature specifically addressing the tourism domain is relatively sparse.

In the following paragraphs, we outline what we consider the most pressing research directions focusing knowledge graphs to provide data to build a more sustainable touristic ecosystem.

- *Real-Time Knowledge Graph Population.* Although KGs are dynamic by nature [Ta19], many existing studies are limited to their initial construction phase [TG23]. Moreover, most literature in our study is quite vague about how to perform instance updates to keep a KG current or even requires the repetition of the complete, heavyweight lifecycle for any update, which results in an entire new release of the graph (e.g., [Do21]). While the latter seems reasonable for relatively static data and helps to ensure rigorous quality levels, a more lightweight sub-process (i.e., knowledge graph population) is required to update KG instances in (or close to) real-time based on data streams and frequent updates of data sources (e.g., sensors). Such a process could trade flexibility and consistency guarantees for better performance while being limited to a subset of instance-related operations (e.g., creating or updating instances). Accordingly, rigorous data processing is required upstream to reliably ensure the quality and consistency of data streams and to avoid quality gaps between the KG population and further graph refinement steps.

- *Distribution and Parallelization.* Depending on the use case, it may be necessary to synchronize KG population processes running in parallel. Current literature often describes the construction of KGs as a rather sequential process, not considering the presence of parallel input streams. However, considering real-world scenarios (e.g., a destination with distributed sensor sets, each driven by different owners), parallel KG population processes could rather be the default than the exception. In even more complex scenarios, preprocessing could not be limited to simple aggregations but could also include more sophisticated ML methods to train different models (e.g., to predict the occupancy of a given area). ML algorithms, however, usually need access to all relevant data to create reliable prediction models, which is likely to raise privacy concerns. As a possible solution, future research could investigate the integration of federated learning [Fe22] into the KG population process to enable the collaborative training of ML models without needing to share raw and confidential data.
- *Touristic Ontology for Dynamic Data.* Due to the rather static nature of past touristic data collections, current data models (e.g., [ABA19; DZT22; FF22]) still need to meet the new challenges of tourism data management. Thus, broadly applicable ontologies are necessary to reflect the nature of dynamic data types, including the online and offline behavior of individual visitors or groups, as well as transport and mobility. Such capabilities should be integrated into existing and accepted models to promote interoperability further and avoid becoming a semantic zoo.

6 Conclusion

This paper reviewed the current state of the art regarding creating and maintaining knowledge graphs. We focused on their use in sustainable touristic ecosystems and analyzed them in the context of the touristic data landscape. We identified the most urgent directions for future research to increase the practical relevance of KGs, including (i) a quality-preserving and scalable knowledge graph population, (ii) distributed data processing, and (iii) the enhancement of existing tourism ontologies. Like other work, our study suffers from certain limitations. First, our investigation is limited to a technical perspective of KG lifecycles, omitting related non-technical challenges. Second, we perceive touristic KGs as graphs with a pre-built schema (i.e., an ontology) and excluded KGs that do not satisfy this definition from our study. Third, the primary sources of literature review are limited to publications available in Scopus, which, while having a broad coverage of multidisciplinary literature, may miss out relevant publications. Fourth, due to its nature, the selection of papers is subjective to a certain degree. Regarding future work, we plan to further strengthen our research by incorporating additional literature and domain experts as well as by widening the scope of our review by further exploring additional needs of touristic data management. Furthermore, we will address the identified research gaps by focusing on advanced knowledge graph population processes, their synchronization, and the use of federated learning in combination with KG refinement approaches based on graph neural networks (GNNs) [KW17]. Finally, we aim to evaluate these processes based on real-world use cases addressing current touristic challenges such as overtourism and active visitor management.

Acknowledgements

This work is a result of a cooperation between Outdooractive AG, the Technical University of Munich, and the University of Applied Sciences Kempten. This study was supported by the AIR research project (67KI21005B), funded by the federal ministry for the environment, nature conversation, nuclear safety, and consumer protection.

Bibliography

[AIR23] AIR – AI-basierter Recommender für nachhaltigen Tourismus – Ein Verbundprojekt zur Entwicklung eines Digitalen Besuchermanagements in deutschen Tourismusdestinationen, https://air-tourism.de/, accessed 2023-03-26.

[ABA19] AlpineBits Alliance: AlpineBits. Open Data Exchange in the Alpine tourism, https://www.alpinebits.org, accessed 2019-08-30.

[BL01] Berners-Lee, T.; Hendler, J.; Lassila, O.: The Semantic Web. Scientific American 284, Scientific American Inc., No. 5, pp. 34–43, 2001.

[Ba19] Baecker, J.; Engert, M.; Müller, J.; Neubig, S.; Schaffer, N.; Kunz, I.; Schneider, A.; Krcmar, H.; et al.: BayernCloud Gesamtreferenzarchitektur, 2019.

[Be20] Benner, M.: Overcoming overtourism in Europe: Towards an institutional-behavioral research agenda. Zeitschrift für Wirtschaftsgeographie 64, pp. 74–87, 2020.

[BLR05] Blain, C.; Levy, S. E.; Ritchie, J. R. B.: Destination Branding: Insights and Practices From Destination Management Organizations. Journal of Travel Research, 2005.

[Bo22] Bollenbach, J.; Neubig, S.; Hein, A.; Keller, R.; Krcmar, H.: Using Machine Learning to Predict POI Occupancy to Reduce Overcrowding, GI, Bonn, 2022.

[Ch22] Chen, Q.; Li, Q.; Wu, J.; Mao, C.; Peng, G.; Wang, D.: Application of knowledge graph in power system fault diagnosis and disposal: A critical review and perspectives. Frontiers in Energy Research 10, 2022.

[Ca19] Capocchi, A.; Vallone, C.; Pierotti, M.; Amaduzzi, A.: Overtourism: A Literature Review to Assess Implications and Future Perspectives. Sustainability 11, 2019.

[Ch20] Chen, Z.; Wang, Y.; Zhao, B.; Cheng, J.; Zhao, X.; Duan, Z.: Knowledge Graph Completion: A Review. IEEE Access 8, pp. 192435–192456, 2020.

[DAT23] DATES project Tourism Data Space European Tourism Data Space, https://www.tourismdataspace-csa.eu/, accessed 2023-03-26.

[DZT22] Deutsche Zentrale für Tourismus e.V.: Domain Specifications Browser, https://open-data-germany.org/domain-specifications-browser/,accessed 2022-10-03.

[Do21] Dojchinovski, M.; Forberg, J.; Frey, J.; Hofer, M.; Streitmatter, D.; Yankov, K.: The DBpedia Technology Tutorial 3064, 2021.

[EW16] Ehrlinger, L.; Wöß, W.: Towards a Definition of Knowledge Graphs, p. 4, 2016.

[Fä17] Färber, M.; Bartscherer, F.; Menne, C.; Rettinger, A.: Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. In: Zaveri, A.; Kontokostas, D.; Hellmann, S.; Umbrich, J.; Zaveri, A.; Kontokostas, D.; Hellmann, S.; Umbrich, J. (eds.) Semantic Web 9, pp. 77–129, 2017.

[FF22] FIWARE Foundation: Smart Data Models, https://www.fiware.org/smart-data-models/, accessed 2022-10-03.

[Fl22] Flores, J.; Jamin, E.; Nadal, S.; Romero, O.: The Knowledge Graph Lifecycle in NTT DATA. In: Proceedings of the ISWC, 3254, 2022.

[Fe22] Fernández, J. D.; Menci, S. P.; Lee, C. M.; Rieger, A.; Fridgen, G.: Privacy-preserving federated learning for residential short-term load forecasting. Applied Energy 326, p. 119915, 2022.

[GVT20] Gowreesunkar, V. G.; Vo Thanh, T.: Between Overtourism and Under-Tourism: Impacts, Implications, and Probable Solutions. In: Séraphin, H.; Gladkikh, T.; Vo Thanh, T. (eds.): Overtourism. Springer International Publishing, pp. 45–68, 2020.

[Gu09] Guarino, N.; Oberle, D.; Staab, S.: What Is an Ontology? In: Staab, S.; Studer, R. (eds.): Handbook on Ontologies. Berlin, Heidelberg: Springer, pp. 1–17, 2009.

[Gr15] Gretzel, U.; Werthner, H.; Koo, C.; Lamsfus, C.: Conceptual foundations for understanding smart tourism ecosystems. Computers in Human Behavior 50, pp. 558–563, 2015.

[HS13] Harris, S.; Seaborne, A.: SPARQL 1.1 Query Language, 2013.

[Ho22] Hogan, A.; Blomqvist, E.; Cochez, M.; D'amato, C.; Melo, G. D.; Gutierrez, C.; Kirrane, S.; Gayo, J. E. L.; et al.: Knowledge Graphs. ACM Computing Surveys 54, pp. 1–37. 2022.

[Hi21] Hitzler, P.: A review of the semantic web field. Communications of the ACM 64, pp. 76–83, 2021.

[Hi12] Hitzler, P.; Krötzsch, M.; Parsia, B.; Patel-Schneider, P. F.; Rudolph, S.: OWL 2 Web Ontology Language Primer (Second Edition): W3C~Recommendation, 2012.

[He20] Hein, A.; Schreieck, M.; Riasanow, T.; Setzke, D. S.; Wiesche, M.; Böhm, M.; Krcmar, H.: Digital platform ecosystems. Electronic Markets 30, pp. 87–98, 2020.

[Ji20] Jia, Y.; Liu, D.; Sheng, Z.; Feng, L.; Liu, Y.; Guo, S.: EasyKG: An end-to-end knowledge graph construction system. Communications in Computer and Information Science 1157, 2020.

[Ki21] Kirrane, S.: Intelligent software web agents: A gap analysis. Journal of Web Semantics 71, p. 100659, 2021

[KW17] Kipf, T. N.; Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks, arXiv, 2017.

[KK17] Knublauch, H.; Kontokostas, D.: Shapes Constraint Language (SHACL), https://www.w3.org/TR/shacl/, accessed 2022-10-03.

[KHL22] Krajickova, A.; Hampl, F.; Lancosova, E.: Visitors' perception of overtourism impacts in a small destination. Anatolia 33, pp. 236–246, 2022.

[KEK12] Krötzsch, M.: OWL 2 Profiles: An Introduction to Lightweight Ontology Languages. In: Eiter, T.; Krennwallner, T. (eds.): Reasoning Web. Semantic Technologies for Advanced Query Answering: 8th International Summer School, Vienna, Austria. Berlin, Heidelberg: Springer, pp. 112–183, 2012.

[Li17] Li, Y.; Hu, C.; Huang, C.; Duan, L.: The concept of smart tourism in the context of tourism information services. Tourism Management 58, pp. 293–300, 2017.

[La22] Lai, F.; Zhao, Z.; Gao, D.; Luo, W.; Lou, C.; Han, S.; Ma, C.: Using Semantic Dependencies to Realize the Construction of Cloud Data Center Operation and Maintenance Knowledge Graph. In: Prognostics and Health Management Conference, pp. 236–240, 2022.

[Ne22] Neubig, S.; Baecker, J.; Gehring, N.; Hein, A.; Weking, J.; Krcmar, H.: Data-driven Initiatives of Destinations Supporting Sustainable Tourism. In: Proceedings of the Americas Conference on Information Systems, 2022.

[ND23] neusta destination.one: Tourism Tech Alliance, www.tourismtechalliance.org/, accessed 2023-05-19.

[OA23] Outdooractive, https://www.outdooractive.com, accessed 2023-05-20.

[Pa16] Paulheim, H.: Knowledge graph refinement: A survey of approaches and evaluation methods. In: Cimiano, P. (ed.) Semantic Web 8, pp. 489–508, 2016.

[Ro21] Rossi, A.; Barbosa, D.; Firmani, D.; Matinata, A.; Merialdo, P.: Knowledge Graph Embedding for Link Prediction: A Comparative Analysis. ACM Transactions on Knowledge Discovery from Data 15, 2021.

[Si21] Şimşek, U.; Angele, K.; Kärle, E.; Opdenplatz, J.; Sommer, D.; Umbrich, J.; Fensel, D.: Knowledge graph lifecycle: Building and maintaining knowledge graphs. In: Second International Workshop on Knowledge Graph Construction, 2021.

[SR14] Schreiber, G.; Raimond, Y.: RDF 1.1 Primer, 2014.

[So16] Soualah-Alila, F.; Coustaty, M.; Rempulski, N.; Doucet, A.: DataTourism: Designing an Architecture to Process Tourism Data. In: Inversini, A.; Schegg, R. (eds.): Information and Communication Technologies in Tourism 2016. Cham : Springer International Publishing, pp. 751–763, 2016.

[SR21] Schmücker, D.; Reif, J.: The Big Data Illusion. Zeitschrift für Tourismuswissenschaft 13, pp. 157–166, 2021.

[SAM14] Sequeda, J. F.; Arenas, M.; Miranker, D. P.: OBDA: Query Rewriting or Materialization? In Practice, Both! In: Mika, P.; Tudorache, T.; Bernstein, A.; Welty, C.; Knoblock, C.; Vrandečić, D.; Groth, P.; Noy, N.; et al. (eds.): The Semantic Web 8796. Cham: Springer International Publishing, pp. 535–551, 2014.

[Si12] Singhal, A.: Introducing the Knowledge Graph: things, not strings, https://blog.google/products/search/introducing-knowledge-graph-things-not/, accessed 2022-10-02.

[SUF20] Şimşek, U.; Umbrich, J.; Fensel, D.: Towards a knowledge graph lifecycle: A pipeline for the population of a commercial Knowledge Graph 2535, 2020.

[Si23] Şimşek, U.; Kärle, E.; Angele, K.; Huaman, E.; Opdenplatz, J.; Sommer, D.; Umbrich, J.; Fensel, D.: A Knowledge Graph Perspective on Knowledge Engineering. SN Computer Science 4, 2023.

[Sc22] Schmücker, D.; Keller, R.; Reif, J.; Schubert, J.; Sommer, G.: Digitales Besuchermanagement im Tourismus – Konzeptioneller Rahmen und Gestaltungsmöglichkeiten, 2022.

[TG23] Tamašauskaitė, G.; Groth, P.: Defining a Knowledge Graph Development Process Through a Systematic Review. ACM Transactions on Software Engineering and Methodology 32, pp. 1–40, 2023.

[Ta19] Tasnim, M.; Collarana, D.; Graux, D.; Orlandi, F.; Vidal, M.-E.: Summarizing Entity Temporal Evolution in Knowledge Graphs. In: Companion Proceedings of the 2019 World Wide Web Conference. San Francisco USA: ACM, pp. 961–965, 2019.

[UN19] UNWTO: Overtourism? New UNWTO report offers case studies to tackle challenges. Madrid, Spain, 2019.

[VT17] Villazon-Terrazas, B.; Garcia-Santa, N.; Ren, Y.; Faraotti, A.; Wu, H.; Zhao, Y.; Vetere, G.; Pan, J. Z.: Knowledge Graph Foundations. In: Pan, J. Z.; Vetere, G.; Gomez-Perez, J. M.; Wu, H. (eds.): Exploiting Linked Data and Knowledge Graphs in Large Organisations. Cham: Springer International Publishing, pp. 17–55, 2017.

[Vi18] Vidal, V.; Magalhães Pequeno, V.; Moura Arruda Júnior, N.; Casanova, M. A.: Publication and maintenance of RDB2RDF views externally materialized in enterprise knowledge graphs. International Journal of Web Information Systems 18, pp. 255–285, 2022.

[WQW21] Wang, M.; Qiu, L.; Wang, X.: A Survey on Knowledge Graph Embeddings for Link Prediction. Symmetry 13, p. 485, 2021.

[Wc15] W3C: Semantic Web, https://www.w3.org/standards/semanticweb/, accessed 2022-10-02.

[WW02] Webster, J.; Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26, pp. xiii--xxiii 2002.

[WFW13] Wolfswinkel, J. F.; Furtmueller, E.; Wilderom, C. P. M.: Using grounded theory as a method for rigorously reviewing literature. European Journal of Information Systems 22, Taylor & Francis, pp. 45–55, 2013.

[Wa21] Wang, J.; Wang, X.; Ma, C.; Kou, L.: A survey on the development status and application prospects of knowledge graph in smart grids. In: IET Generation, Transmission and Distribution 15/3, pp. 383–407, 2021.

[Xi18] Xiao, G.; Calvanese, D.; Kontchakov, R.; Lembo, D.; Poggi, A.; Rosati, R.; Zakharyaschev, M.: Ontology-Based Data Access: A Survey, 2018.

[Xi19] Xiao, G.; Ding, L.; Cogrel, B.; Calvanese, D.: Virtual Knowledge Graphs: An Overview of Systems and Use Cases. Data Intelligence 1, pp. 201–223, 2019.

[Zh20] Zhang, Z.; Cai, J.; Zhang, Y.; Wang, J.: Learning Hierarchy-Aware Knowledge Graph Embeddings for Link Prediction. In: Proceedings of the AAAI Conference on Artificial Intelligence vol. 34, pp. 3065–3072, 2020.