

Article

Knowledge Graphs for Online Marketing and Sales of Touristic Services

Anna Fensel ^{1,*} , Zaenal Akbar ^{1,2} , Elias Kärle ¹ , Christoph Blank ³, Patrick Pixner ⁴ and Andreas Gruber ⁵

¹ Semantic Technology Institute (STI) Innsbruck, Department of Computer Science, University of Innsbruck, 6020 Innsbruck, Austria; zaenal.akbar@lipi.go.id (Z.A.); elias.kaerle@uibk.ac.at (E.K.)

² Indonesian Institute of Sciences, Jakarta 12710, Indonesia

³ Kognitiv Corporation, 1070 Vienna, Austria; cb@weavs.io

⁴ m-Pulso GmbH, 6020 Innsbruck, Austria; patrick.pixner@m-pulso.com

⁵ Redlink GmbH, 5020 Salzburg, Austria; andreas.gruber@redlink.co

* Correspondence: anna.fensel@uibk.ac.at; Tel.: +43-512-507-96884

Received: 6 March 2020; Accepted: 28 April 2020; Published: 4 May 2020

Abstract: Direct online marketing and sales are nowadays an essential part of almost any business that addresses an end consumer, such as in tourism. On the downside, the data and content required for such marketing and sales are typically distributed, and difficult to identify and use, especially for small and medium enterprises. Further, a combination of content management and semantics for automated online marketing and sales is becoming practically feasible now, especially with the global adoption of knowledge graphs. A design and feasibility pilot of a solution implementing semantic content and data value chain for online direct marketing and sales, basing on knowledge graphs, and efficiently addressing multiple channels and stakeholders, is provided and evaluated with the end-users. The implementation is shown to be suitable for the use on the Web, social media and mobile channels. The proof of concept addresses the tourism sector, exploring, in particular, the case of touristic service packaging, and is applicable globally. The typically encountered challenges, particularly, the ones related to data quality, are identified, and the ways to overcome them are discussed. The paper advances the knowledge of employment of knowledge graphs in online marketing and sales, and showcases its related innovative practical application, co-created by the industry providing marketing and sales solutions for Austria, one of the world's leading touristic regions.

Keywords: knowledge graphs; online marketing; online sales; touristic service packaging; semantic technology; social media; mobile apps

1. Introduction

As end-consumers tend to spend more time online, in social networks, mobile apps, etc., and at the same time, the costs for using online booking and marketing platforms increase, online direct marketing and sales are becoming increasingly important, and even indispensable, for addressing potential customers. On the downside, the challenge is that the production of high-quality, attractive and engaging content, such as images, video clips, texts, for online channels requires creativity, worktime, skills and equipment, which most, especially smaller, businesses do not possess. Furthermore, the typical businesses do not have the content and data available for supporting their marketing; and they typically need to research:

- which contents have already been published by other similar businesses online (e.g., because they want to be unique in their communication and avoid posting the same kind of contents many others have already posted),

- how they compare to their competitors (e.g., because they want to refine their unique selling proposition),
- which marketing strategies have been effective or less effective so far in their region (e.g., because they want to improve their customer relationship management).

Obviously, the quality of marketing and sales activities can be substantially improved if the marketing resources, both content and data, would be interoperable and shareable by the businesses located within the same region, while, when needed, being also efficiently deliverable beyond the region, globally, with the required level of reputation and quality. While the quantity of produced data is increasing, methodologies to manage their semantic structures, however, are currently only elaborated at the level of enterprises [1], but not across enterprises on the web, as it brings further challenges. Further, as the regions and regional businesses are building their data and content management infrastructures that are multi-channel, supporting the social media and mobile use, as well as the content and data-sharing scenarios [2], the need for the relevant interoperability and communication infrastructural technology increases.

In short, the companies and businesses need to present themselves on the Web, and best practice approaches for this are not identified. Further, it is typically unclear which data exact management challenges are to be encountered in such multi-channel direct online marketing systems. Our research aims in this work are to identify a reference solution for an online marketing system that needs to combine multiple sources of the data, and to identify the typical data management challenges that are appearing with the deployment of such a system. The reference implementation description of a multi-channel direct online marketing system and the identified maintenance data management challenges contribute to the modern systems implementing knowledge graphs, and set up best practices for the academia and industry.

To address this, from a methodological perspective, we conceptualize and prototype a technology ecosystem and infrastructure to interoperate across different marketing and sales data and content resources using Knowledge Graphs, built with Linked Data [3] and semantic technologies [4]. Knowledge Graphs are among the fastest-growing fields in computer science, with the vast majority of companies adopting or wanting to adopt them. Creation of the Knowledge Graphs is made with various means, ranging from the “classical” manual engineering to techniques supported by machine learning, see for example, the related developments of Google’s Knowledge Vault [5] and at Alibaba [6]. In our use case, Knowledge Graphs enhance representation structures and interoperability of distributed marketing and sales resources for allowing meaningful searches, efficient information dissemination and consumption. The approach combines and extends solutions for data acquisition and management, extraction and integration from the Web, data and content interlinking, semantic annotation and enrichment, and is implemented in the area of tourism services marketing. Further, we evaluate the prototype with the end-users, who understand both the technology side, as well as the end-user side. Based on this, we draw conclusions regarding the challenges in adoption of the technology, and outline the approaches to overcome the most common of them.

As in many service-oriented businesses nowadays, the touristic service consumers want individualized experiences [7] and thus no longer want the “one-size fits all” touristic packages, as, for example, produced in a generic way by travel agencies. This imposes large challenges on data management, and has been a motivation for the TourPack project (On-demand Data-driven Production of Touristic Service Packages (TourPack): <http://tourpack.sti2.at>), in whose settings we develop our approach. The aim of the TourPack project [8] has been to design and prototype a production system that creates “on-demand” touristic packages catering to the individual touristic service consumer needs and preferences—applying the smart usage of the open and proprietary data for the information integration and service composition, and eventually, improving the multi-stakeholder data-driven production processes of touristic service offer, or, more specifically, to build a linked data-empowered system for touristic service packaging. Integrating information from multiple sources and systems employing linked data as a global information integration platform, and mining from the depths of

the “closed” data, the touristic service package production system is to be able to create the most optimal travel experience for the traveler. Further, the service packages are to be efficiently published and made bookable to the end consumers via intelligently selected most suitable communication and booking channels: especially the online channels with rapidly growing user audiences, such as the social media and the mobile apps. The TourPack project has delivered the settings for our work, and in this paper we demonstrate its outcomes, specifically, addressing the data management and marketing aspect of this system.

This paper is structured as follows. The addressed problem and typical tourism domain user scenario examples are in Section 2. In Section 3, the current state-of-the-art is presented. Our approach to the solution and architecture is found in Section 4. Section 5 contains implementation details, and Section 6 describes the evaluation. Section 7 concludes and summarizes the paper.

2. Description of Problem and User Scenario

The internet, web-based communication and booking channels are becoming increasingly important in today’s competitive world. Organizations of all sizes, commercial and not-for-profit, regularly face the challenge of communicating with their stakeholders using a multitude of contents, as well as of channels, e.g., websites, videos, PR activities, events, email, forums, online presentations, social media, mobile applications, and (more recently) structured data. The need to be present at different communication channels directly, and the impacts the channels are bringing on the marketing and the organizational processes have been investigated in practice [9,10]. Hence, management of the marketing- and booking- related data and being able to combine it on the fly and spread the message over various channels became a basic requirement for a marketer. The importance of marketing over multiple channels is also recognized in research, and approaches that involve multiple stakeholders of the marketing process are being investigated [11]. Naturally, the touristic service, marketing would extend to heterogeneous types of services (e.g., in food and drink, wellness, shopping, culture), and in the same personalized manner would deliver packaged offers for experiences matching the availability and expectations to the end-users.

For tourists, the linked data-based service system for direct marketing would be crucial when finding and consuming the most relevant services on the fly. A typical end-user scenario, reflecting a hotel’s and a hotel guest’s every-day business, is as follows: In the afternoon, a guest is up for some outdoor activities. He/she remembers seeing a “Sports” section in the app. Quickly browsing through the outdoor and sport activities in the app he/she finds a place he/she wants to go to. Since it is the guest’s first visit in this town he/she has no idea how to find this place. So he/she clicks the “Details” button in the description of the place and finds a direct link to OpenStreetMap which directs him/her to the desired place.

3. State of the Art and Current Knowledge

Addressing our showcase domain of tourism, data-centric information channels for Austria currently provide machine-processable information such as mountain bike routes or public transport schedules, and restaurants with specific food preferences. Online services allow the booking of hotels (solutions are provided by companies like Kognitiv, Easybooking), ski passes (e.g., provided by SkiData), concert tickets (e.g., provided by oeticket), etc. Making touristic services easy to publish for the service providers and easy to find and book for the tourists are the key challenges for the production of a complete online service tourist offer package. Considering the abundance and variety of travel services and the restricted time travelers typically possess on vacations or on business trips, the touristic service search, selection and combination require a lot of effort from the service consumer. The methods to access and process the applicable structured Web data are currently evolving, while most of them focusing on the scenarios of consumption of linked data from semantic repositories [12] and not yet investigating the use of all available distributed data to a full extent. In addition to addressing the

holistic solution of the set research question (in Section 1), our solution advances the state of the art in the individual state of the art cases, as described in the paragraphs below.

Overall within the touristic domain, service composition is a typically addressed problem, requiring ongoing adaptation and co-development of new technology and approaches e.g., for touristic planning [13]. The creation of travel information services with spatial, social and temporal dimensions can be built already on Web 2.0 data, namely, as travel mashups [14] or a broker platform [15]. With the appearance of Linked Data and the global use of schema.org, the input coming from data processing components e.g., making decisions on yield management would be more and more present in service composition approaches.

The availability of the data is a basic enabler for the sufficiently high data quality, in particular, of the customer data, and in addition, the sharing of the data is essential for the modern marketing [16]. There is also too much of marketing data and information available to choose from for the consumer [17]. Our approach aims to leverage on all the richness of the available data, and at the same time employ the most relevant for the consumer data, information and content sources. The data is either available explicitly, or can be derived with the available algorithms e.g., for mining of user-profiles from social media, and the marketing content recommended based on this profile data [18]. The possibility to generate high-quality user profile data from online interactions has been demonstrated, in particular via a touristic website [19]. Another advantage of our approach is being able to analyze the data and content usage across multiple channels, which is essential for tracking and better understanding the customer journey [20]. Our design's aim is also to achieve the best choice outcome while minimizing the cost of the decision-making process, which is seen as essential for customer satisfaction [21].

Furthermore, consumers are more and more interested in communication via different (and multiple) channels. The ability to answer customer demands wherever they are, and using the channel and device of their choice, will make a huge impact in their experience and consequently in the business. The fact that customers want access to all the services [22] creates the necessity of an integrated strategy. Mobile services must be integrated in the business process, not seen as a separate endeavor. The latter, in particular, are a valuable source of personalized access, customer and business data, which should also be employed in a connected manner [23].

To demonstrate the importance of the mobile experience, Google [24] took a deeper look at users' expectations and reactions towards their site experiences on mobile devices. Most interestingly, 61% of people said that they would quickly move onto another site if they did not find what they were looking for right away on a mobile site. The bottom line is that without a mobile-friendly site (that could be extended to mobile access to services) one will be driving users to the competitors. Having a great mobile site is no longer just about making a few more sales. It has become a critical component of building strong brands, nurturing lasting customer relationships, and making mobile work.

Regarding the mobile experience, many customers prefer interactions via online channels rather than face to face, a fact that is currently supported by the increased number of mobile devices within the customer's reach. A combination of various online platforms and mobile applications including social networks would increase the opportunity for consumers to purchase goods and services online [25]. An appropriate mobile strategy integrated into the online multichannel world will also benefit the management and customer service for the tourism business. Further, the technology would implement the known good practices of communications with the customers via the mobile phone channels [26]. Given the requirements described above, our solution is integrating both mobile and social media interaction modes, in a holistic manner.

4. Approach and Enabling Technology

Our approach to the service marketing and integration aims to close the gaps of the approaches introduced above and create solutions for the dynamic integration of touristic web services on the fly that will enable the creation of enhanced integrated services. Consequently, based on linked data,

the service offers will be distributed via the most appropriate channels. We have defined a system architecture (Figure 1), and data management requirements regarding the components of the system.

As visible in Figure 1, the solution takes in the closed data from touristic service providers as well as linked (open and closed) data, and forms valid touristic service packages. The information processing and representation are facilitated by schema.org-supporting plug-ins that implement Knowledge Graphs in content management systems, such as, e.g., in Wordpress and Drupal. After this, these packages are disseminated: both in an individual (tailored to a user profile) way, or via social media, pre-selected to tailor the typical audience of the touristic service provider, or, the kind of audience this service provider wants to receive.

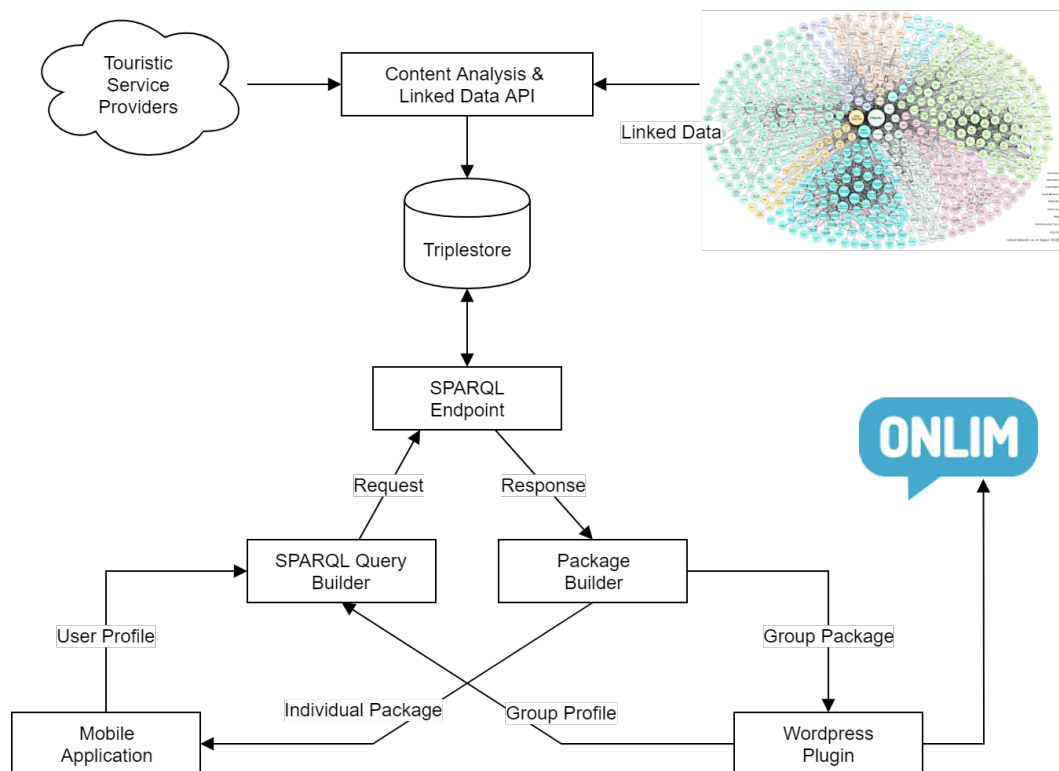


Figure 1. TourPack linked data-based content management architecture.

There are two main components involved in the definition of the relevant packages. First, SPARQL Query Builder is responsible to construct a query to be executed on the knowledge graph stored in a Triplestore through a SPARQL endpoint. The query will be constructed based on the provided profile. For an individual package, a user profile will be collected through a mobile application, while for a group package, a profile will be collected through a CMS (e.g., WordPress) plugin. Technically, a profile consists of a list of types of services or names of service providers or keywords which are relevant to users. The second main component is Package Builder which is responsible for providing the relevant package to users. A list of services will be obtained from the knowledge graph through the SPARQL Query Builder component. The services will be ranked according to the defined profile. For example, if a user requested restaurants closer to his/her hotel, then the list of obtained restaurants will be ordered based on the distance from the hostel to every restaurant. If the user requested outdoor activities that are open until late in the night, then the providers will be listed based on closed time.

We support automatic generation, clustering and packaging of semantically annotated touristic service offers from a variety of sources. The Application Programming Interfaces (APIs) and components perform information extraction, clustering and publishing in order to:

- obtain the extracted data in a linked data format, (semi-)automatically associating metadata;
- generate service representations in a linked data format according to ontological models;

- interlink, cluster, package and disseminate services in an automatic way;
- provide a semantic service and an online interface for easy publishing and access to the above-mentioned functionalities.

The effort already has run pilots, such as with the Touristic Association of Innsbruck, and already implemented semantic dissemination support by implementing schema.org support on their website (Website of Touristic Association of Innsbruck: <http://www.innsbruck.info>) and publishing the touristic data of the Innsbruck region as linked open data [27]. In cooperation with the province of Salzburg (SalzburgerLand), the touristic data of Salzburg are published in a linked open data format with schema.org, and are usable (SalzburgerLand Data Hub: <http://data.salzburgerland.com>) in particularly as the corresponding Knowledge Graph (Press release of the region of Salzburg on the Knowledge Graph: <https://newsroom.salzburgerland.com/daten/knowledge-graph/>), and the appropriate mobile app support exists. We are deploying our solution also with direct touristic service providers: starting with hotels, and extending to further touristic services. Including relevant touristic services in the platform, we have made a classification of all types of touristic service providers. These are extracted from the websites of Austrian regional touristic associations. The complete list of these types, accompanied by real-life examples of Innsbruck's touristic service providers is presented in Table 1. Herewith, the service providers of all these types have a straightforward marketing contents provisioning via our system. As schema.org became a de-facto standard supported by Google, Yandex, Bing and Yandex, our solution can be applied globally with increasing ease.

Table 1. Types and examples of touristic service providers (from Innsbruck).

Types of Touristic Service Provider	Example in Innsbruck	Types of Touristic Service Provider	Example in Innsbruck
Apartments	Riedz Apartments	Mountain restaurants and lounges	Höttinger Alm
Bars and pubs	Zappa	Payment solutions	CP Solutions
Beaches	Baggersee	Rentals	Hertz Autovermietung
Cafes	Katzung	Restaurants	Restaurant Burkia
Camp sites	Camping Kranebitter Hof	Shopping	Sillpark
City cards	Innsbruck Card	Sights	Goldenes Dachl
Cyber-cafes	Pub 7	Ski lifts	Bergbahnen Obergurgl
Events	Friska Viljor Concert, Weekender	Ski schools	Skischule Obergurgl
Exhibitions	Galerie Rhomberg, Heaven to Hell	Spas	Aqua Dome Längenfeld
Farm holidays	Sillerhof	Swimming pools	Höttinger Hallenbad
Hotels / rooms	Schwarzer Adler	Tours	Stadtrundgang Innsbruck
Insurances	UNIQA		

5. Implementation

For the technical format for information collection, processing and modeling, we heavily rely on linked data and the de facto standard schema.org, and are involved in co-defining its extensions, particularly at W3C. We use schema.org for modeling and communication of touristic service packages, including its actionable components for the booking part (see Figure 2 for our implementation).

For the knowledge repository construction, the datasets from Innsbruck Tourism, Kognitiv Hotels and Offers and Salzburg Tourism were put into a semantic format, which is used for linking data in the tourism area. Through the Redlink platform (Redlink Semantic Platform: <https://redlink.co/semantic-platform>), we provide these data in a cloud and linked data -empowered data repository for storing the semantically structured data.

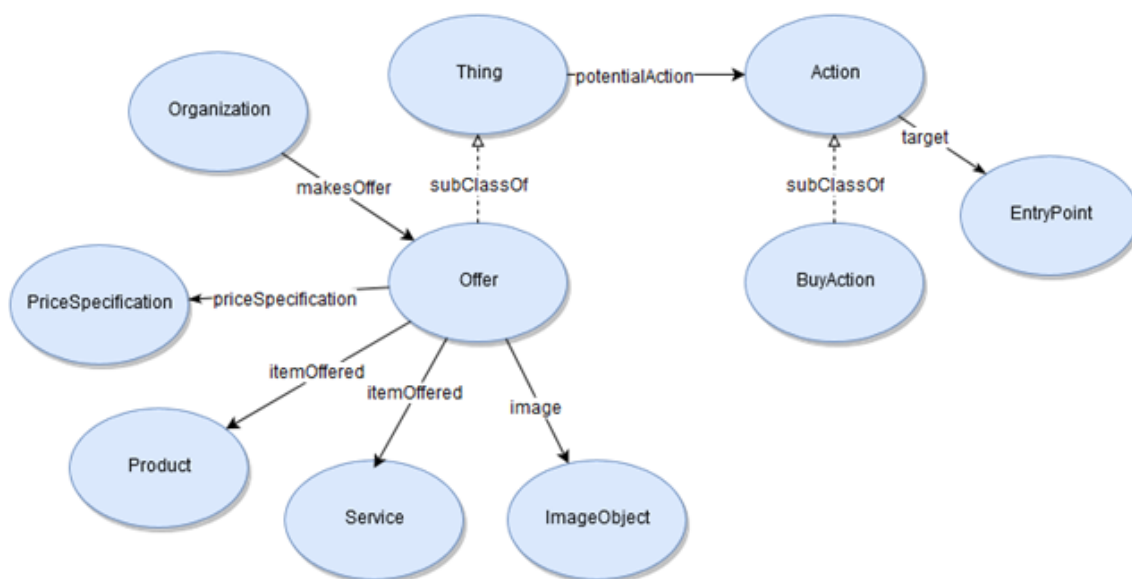


Figure 2. Visualisation of the TourPack schema.org model of a touristic service package.

Redlink’s Content Analysis solution offers fact extraction, topic classification, and fact linking from textual and media documents in different languages. Users are able to send text and get back facts and entities that have been identified automatically. In addition, the service provides references to existing datasets (public datasets like Dbpedia and Freebase, as well as custom user-created vocabularies like a product database) about the entities identified in the content. Users can create their own custom configurations with some simple steps, e.g., for adding advanced natural language processing features or commercial datasets.

Based on open source technology, Apache Marmotta (Apache Marmotta: <http://marmotta.apache.org>), we have extended the querying capabilities towards some requirements (e.g., implementing the GeoSPARQL extension [28]). The data can be accessed by the partners via interfaces using REST/Web Service access and via SPARQL. Through the Redlink platform, it can be enriched via annotation engines and thus linked to a greater set of entities available in the LOD world. In addition to this, several datasets from “Open Government Data” in Austria for the respective regions, have been added. The resulting dataset has amounted to 2200 lodgings, 9834 locations, and 2114 offers—all available for construction of the package.

Rules for automatic information dissemination, particularly on social media, in the context of Online Communication and Marketing Tool (ONLIM) (<http://onlim.com>) component, have been modeled, and the ontology to automate the online communication has been developed and published online [29]. ONLIM is an online tool based on semantic technologies that aim to facilitate managing various information dissemination channels (such as chatbot systems, social media platforms) by means of publishing posts and tracking the feedback given by other users. ONLIM supports several social media platforms such as Facebook, Twitter, YouTube, LinkedIn, Xing and Flickr. It also allows users to schedule their posts to enable more effective social media management and marketing. In contrast to other similar tools, ONLIM also has an automatic post generation feature that creates posts for publication in social media from external sources. Content, data and services that are annotated with schema.org can be converted to the posts by ONLIM automatically. In addition, these posts can be automatically forwarded to all appropriate channels to disseminate the information about products and services, as well as their packages. The modeled rules for information dissemination perform content adaptation, e.g., addressing that the content is expected in various modes at various channels, as well as communication adaptation e.g., posts should be more frequent on some platforms rather than on the others.

We furthermore have compiled a list of third-party touristic services and their possible uses within the TourPack platform. Additionally, various APIs available for TourPack have also been mutually introduced to identify possible connection points for the TourPack platform implementation.

Further, we had made an analysis and visualization of the data collected from various machine and human sources, and forming the packages out of this data. The packages have been created in a way to take into account the different types of touristic service provided (e.g., a user would likely need one dinner and one hotel room and not two or more, etc., see the created detailed split in Table 1), the time-based availability, user profile and further relevant criteria.

Research on data quality, and particularly, the data in schema.org format, has been another one of our major occupations. The schema.org data provides a very large source of open semantically-structured data, which is not always of a sufficiently high quality to generate the packages directly applicable for the bookings, e.g., it can state that there are more hotels in Tyrol, an Austrian province, than in Austria [30].

For the presentation layer, the user interfaces are being rendered in two forms: social media platforms postings and mobile app postings. In order to define meaningful pilot applications, several possible use case scenarios have been defined to fulfill the requirements of the project consortium and focus on the target user groups. Along with the definition, several application mockups have been created iteratively using state of the art technology. These interfaces include web-based, as well as mobile, use-cases and were carefully evaluated by stakeholders and project partners.

The social media postings are automatically generated from the data using the ONLIM component, and are content-wise addressed to audiences which the touristic service provider already has on its social media platforms e.g., as subscribers, or as the ones it wants to attract. An example of such automatically generated touristic services packages is displayed in Figure 3: the offered social posts are displayed on the right-hand side, and can be scheduled in a calendar. On the other hand, mobile apps can generate personalized messages, and technically can be implemented with varying mobile platforms and varying user interfaces. The key possibility here is to deliver a personalized user experience to the user. The implemented iOS-based example of the mobile app client, as described up to now, is depicted in Figure 4. It allows us to search by date and select various services, and group them into individualized packages.

The screenshot displays the ONLIM web interface. On the left is a dark blue sidebar with navigation icons for 'NEW POST', 'CALENDAR', 'YOUR ACCOUNTS', 'INCOMING CHANNELS', and 'STATISTICS'. The main content area features a calendar for January 2016, with dates from 28 to 31. Above the calendar are tabs for 'Detailed', 'Compact', and 'Month', along with a date range selector set to '01 Jan - 31 Jan 2016'. On the right side, there is a 'Suggested' panel showing two social media posts from 'Rotes Wildschwein'. The first post is about 'Harakiri Skipiste' and includes a photo of a snowy mountain slope. The second post is about a New Year's Eve event in Mayrhofen. The interface also shows 'Drafts: 0' and 'Suggested: 918' at the top right.

Figure 3. An example automatically-generated social media posts—hotel offer suggestions in Online Communication and Marketing Tool (ONLIM).

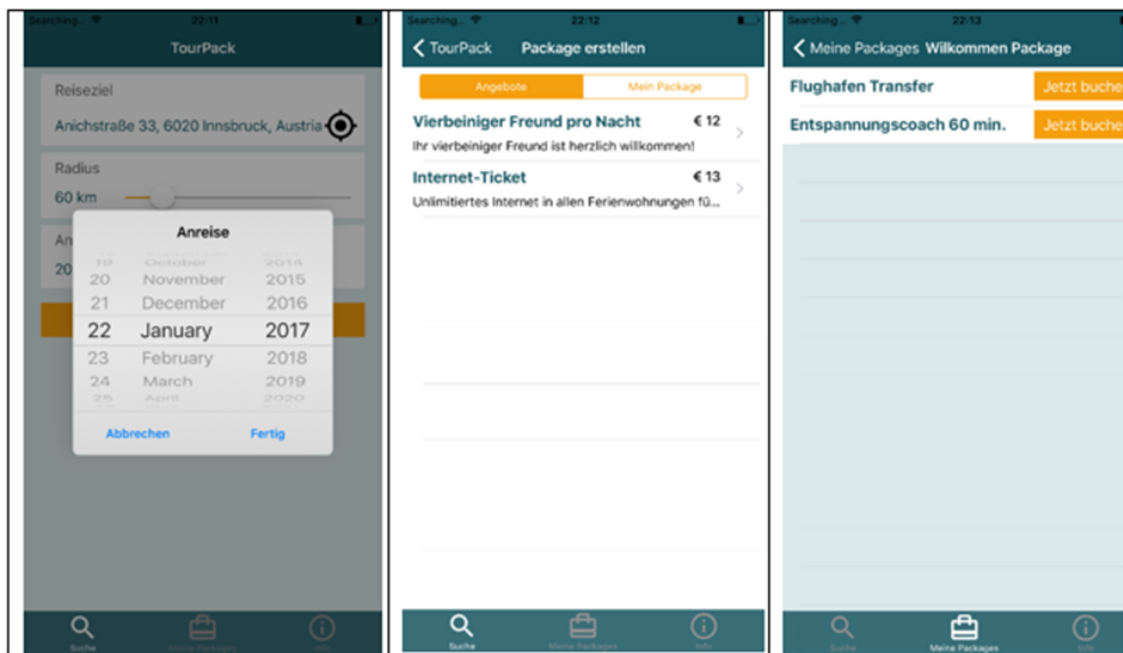


Figure 4. TourPack iOS mobile client implementation.

6. Evaluation

We have conducted field trial user studies with the set up and procedure described as follows. The goals of the evaluation study were:

- Assess the technical and business feasibility of the approach and implementation,
- Identify the key challenges observed by the human evaluators w.r.t. the linked data usage, as well as the system in general.

Evaluation set up has been, correspondingly, technical and empirical. From the technical stand-point, the system is implementable, as it has been implemented with the technology applied. The end-consumer related business use cases are also available and are implemented by companies Kognitiv and m-Pulso with their pilot customers. The use cases are as follows:

- During the course of the project Kognitiv analyzed the possibility to package accommodation related services during the accommodation booking process. While it became feasible from a technology perspective the main blocker to realize/offer a true package is legal and business-related packaging requires a travel agency license in order to combine the different offerings into a single offer for the booking guest. However, to offer clients added value, the technology developed as part of TourPack was incorporated into the Kognitiv extranet and its booking engine in a more loosely coupled way. Namely, with the TourPack technology, the hotelier can select available context-relevant information and offerings that he or she wants to promote during the booking process, and this increases sales and revenues, which is positive. Figure 5 displays the visualization of the services that are possible to package with the booking of the accommodation for the user.
- As m-Pulso already has a number of customers in the touristic sector (hotels, spas, ski resorts, and others), the delivered contributions are directly relevant for its product portfolio. The main goal of most customers in this area is to generate more bookings and to strengthen their customer loyalty. While the existing product achieved very good results in the second area, the generation of more bookings is still a critical issue. The developed packaging service developed fits this need perfectly and will be rolled out to several pilot customers to be evaluated within larger user groups in a production environment.

Karte und Anfahrtsbeschreibung
Koordinaten der Unterkunft: 47.1216577N, 10.2489476E

Sehenswürdigkeiten und Empfehlungen

Konzerthaus, Vienna
 The Konzerthaus is a concert hall located in Vienna, Austria, which opened in 1913. It is situated in the third district just at the edge of the first district in Vienna. Since it was founded it has always tried to emphasise both tradition and innovative musical styles. In 1890, the first ideas for a Haus für Musikfeste (House for music festivals) came about. The idea of the new multi-purpose [...]

Mozarteum University of Salzburg
 Universität Mozarteum Salzburg, also known simply as Mozarteum Salzburg, is a university in Salzburg, Austria which specializes in music and the dramatic arts. It is famous as one of the most respected universities in Europe for music studies and was named after Salzburg native Wolfgang Amadeus Mozart.

Stornierung

Es ist etwas wichtiges vorgefallen und Sie können nicht bei uns bleiben? Sie können die Stornooptionen sehen, indem Sie auf den folgenden Link klicken:
https://ibe.seekda.com/dsr/DEMO_HOLGER1977?skd-language-code=de&skd-booking-id=DSRBC1NWS29YJ&skd-booking-pin=923711#f/skd-ds/skd-bookings

Figure 5. Packages integration within Kognitiv booking engine (guests' view).

The human evaluators have been students at the University of Innsbruck at the Master study level, who have been taking "Semantic Web" course at the University of Innsbruck, but not possessing prior experience of the TourPack project. The TourPack system has been presented and explained to them during the course, also the related presentations and papers have been provided for the study at home. At the end of the course, the following question was posed to the students: "Consider a touristic service packaging application, creating bookable touristic service packages basing on schema.org data and other (linked) open data. Name at least six main challenges that you expect to appear in building and deployment of such an application". There has been 33 responding students to this question: 32 male and 1 female, mostly Austrian, but also some from abroad. The students have provided their replies in writing, indicating up to six challenges each, in a free form, providing their own ideas and descriptions.

In Table 2, we present the type challenges most frequently identified by the students, summarized from their texts with human manual post-processing. We indicate how many mentions each of them has had, as well as the corresponding detailed remarks. This table provides an overview of the most important issues to address in future research and development, in order to make the usage of distributed linked data and schema.org data broader.

Table 2. Commonly identified challenges and their reported frequency.

Group of Challenges	Mentions	Remarks, Examples
Handling heterogeneity and quality of external data	35	The mentions contain the issues of incomplete data, incorrect data, inconsistent data, and challenges with combining different data sources. Needs for data clearing, verification and interlinking are emphasized.
Generating correct semantic annotations	27	The mentions indicate the difficulties connected with the need to generate semantic annotations: finding the vocabularies and terms to use (linked data vocabularies and schema.org), producing correct annotations, choosing the right language dialects.
Service-related challenges	20	The mention lists the issues associated with the development of service technology applications: usable service annotations, frameworks and deployment choices, coordination of communication, selection and discovery of services for the composition, quality of service annotation and service interoperability.
Ontology and data versioning	11	The mentions refer to the characteristics of dynamics of the schema and data, and the necessity to cope with the change: updating the data according to the newest ontology and real-world developments, synchronizing different versions of the datasets, dealing with the outdated data.
Data management issues	10	The mentions refer to the challenges of set-up and management of data infrastructures: for data acquisition, storage, querying.
Take up issues	8	Take up challenges includes the need to for the service to gain the critical user mass, as well as the possible hindering factors due to the early stage of the semantic technology use in the sector (annotations, query interfaces may be lacking or not enough broadly spread).
Missing items in vocabularies and schema.org	6	These mentions point out that some needed for the semantic annotation concepts and properties may be not available in the available ontologies.
Security and misuse	6	Aspects of transfer of the correct data, safe payments, authentication, authorization.
User interface issues	6	The mentions refer to the complexity of design of the efficient user interface, and connecting it correctly to the changing in time APIs, data, contents, and services.
Pricing and payment issues	4	The aspects of correct setting of the price and distribution of the payments, under the complex conditions of international settings and multiple stakeholders.
Internationalization	3	Support of different natural languages is a necessity.
Legal factors, governance	2	Any other law-related aspects of the execution of the TourPack application.
Hardware and mobile issues	2	The requirements for efficient lower layer technical infrastructures are mentioned, such as computational power, mobile communication/network availability.

7. Conclusions

We have presented an approach to disseminating touristic marketing content, particularly service packages, based on Knowledge Graphs as an enabler for tourism businesses and tourists. The businesses can be productive by providing new experiences and finding new direct dissemination and booking channels, while leveraging their own existing touristic contents.

The approach has been implemented and evaluated. The results indicate the general feasibility of the solution, and a need for further improvement of the state of the art for enabling such applications. Three types of challenges have been most frequently mentioned in the feedback (see Table 2), and deserve further development, namely: (1) handling heterogeneity and quality of external data, (2) generating correct semantic annotations, (3) service-related challenges. A possible approach to resolving these challenges is described below.

7.1. Handling Heterogeneity and Quality of External Data

An approach to handle heterogeneity challenges is to represent data from different sources uniformly. First, information from those sources will be analyzed and modeled into a uniform representation, namely vocabulary, which contains classes (including their properties) and how they are related to each other. Then, data format and structure from every source will be inspected and aligned with the vocabulary to produce a data mapping. This mapping is then consumed by a wrapper that is responsible to generate relevant annotations. This approach has been applied to annotate a variety of tourism-related information in the regions of Mayrhofen, Seefeld, and Fügen [31]. A set of 1,6 million triples was generated by March 2017 and accounted for numerous topical information: accommodations and offers, events, infrastructures, organization, press-release including blog posting, information related to the region, and ski-areas. The generated annotations have been successfully identified by search engines, namely Google, indicated by rich snippets produced in its search results.

A solution to ensure high-quality annotations is by providing a guideline on how to enter data into a system correctly such that complete annotations could be generated. First, essential properties of the vocabulary will be categorized as “required” or “recommended”. Every required property needs to be associated with a correct value and therefore must be provided through the relevant data entry field. For example, an event must have a postal address value for its location, and the guideline will dictate that this address cannot be empty whenever a new event is entered into the system. The same case will be provided for every recommended property, where the guideline will recommend possible values to ensure the generated annotations will not raise errors or warnings messages.

On the level of linked data and vocabulary selection, solutions for automated assistance in the selection of the most “appropriate” vocabulary terms to annotate are being developed [32].

7.2. Generating Correct Semantic Annotations

While schema.org became a very practically important development around linked data on the Web, the practitioners still have difficulties with producing correct schema.org annotations. Hence, the data quality of structured content on the web suffers. With the work on semantify.it [33], a platform that supports users in the schema.org annotation creation and publication process is provided. Based on domain-related template files (Domain Specifications) semantify.it provides form-like annotation editors and saves, if the user wants, the resulting annotation files in a database to then be published to websites over a JavaScript snippet or through a plugin for content management systems. Besides that, semantify.it offers a validation functionality to check for semantic validity against predefined Domain Validation files as described in [34].

7.3. Service-Related Challenges

Automatic service packaging is, as mentioned before, highly dependent on available (web) services. To use a service in an automated way the semantics of the service, of its endpoint, parameters and return values, has to be known. Schema.org offers a means to annotate web services called schema.org actions. The vocabulary features properties for the status of the action, the executing agent the target endpoint, containing properties for input parameters, the result of the action and many more. If annotated web services can be found, the determination of their relevance and the combination of several web services becomes much easier or feasible in the first place. Contrary to the distribution of schema.org amongst static data, web services are not yet sufficiently annotated with schema.org.

Further, with content and data, there are a few particularities hampering their potential (re-)use in the marketing data value chain. Namely, selection and reuse of the content and data is possible only manually, unless the license information is published in a machine-readable form. Realization of machine-readable, semantic representations of data licensing, in particular, is still not fully solved, and semantic formats for licensing data, as well as the tools processing data license annotations,

are under-defined or non-existent. Semantic standards for licenses are being developed right now, and include efforts such as ODRL (Open Digital Rights Language (ODRL): <https://www.w3.org/ns/odrl/2/>), and a project that is deriving extended semantic models and tools for data licensing is DALICC (Data Licenses Clearance Center (DALICC): <https://dalicc.net>) [35]. Future work would also include models for content crowdsourcing, customer engagement, as well as quality evaluation metrics. Scalability deployment aspects related to hardware and networks are to be researched.

Author Contributions: Conceptualization: A.F., Z.A., E.K., C.B., P.P. and A.G.; Data curation: A.F. and A.G.; Funding acquisition: A.F., P.P. and A.G.; Investigation: A.F., Z.A., E.K., C.B., P.P. and A.G.; Methodology: A.F.; Project administration: A.F., C.B., P.P. and A.G.; Software: Z.A., E.K., C.B., P.P. and A.G.; Supervision: A.F., P.P. and A.G.; Visualization: E.K. and P.P.; Writing—original draft: A.F., Z.A., C.B., E.K., P.P. and A.G.; Writing, review & editing: A.F. and Z.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially funded by the Austrian Research Promotion Agency (FFG) in the project TourPack within the program “Future ICT”, as well as in the project WordLiftNG within the Eureka, Eurostars Programme.

Acknowledgments: The authors gratefully acknowledge useful input from all TourPack project participants, and of the STI Innsbruck’s start-up ONLIM, specializing on online direct marketing (chatbots and social media).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pistofidis, P.; Emmanouilidis, C.; Papadopoulos, A.; Botsaris, P.N. Management of linked knowledge in industrial maintenance. *Ind. Manag. Data Syst.* **2016**, *116*, 1741–1758, doi:10.1108/IMDS-10-2015-0409. [CrossRef]
2. Pierdicca, R.; Paolanti, M.; Frontoni, E. eTourism: ICT and its role for tourism management. *J. Hosp. Tour. Technol.* **2019**, *10*, 90–106, doi:10.1108/JHTT-07-2017-0043. [CrossRef]
3. Bizer, C.; Heath, T.; Berners-Lee, T. Linked data: The story so far. In *Semantic Services, Interoperability and Web Applications: Emerging Concepts*; IGI Global: Hershey, PA, USA, 2011; pp. 205–227.
4. Domingue, J.; Fensel, D.; Hendler, J.A. *Handbook of Semantic Web Technologies*; Springer-Verlag: Berlin, Heidelberg, 2011.
5. Dong, X.; Gabilovich, E.; Heitz, G.; Horn, W.; Lao, N.; Murphy, K.; Strohmman, T.; Sun, S.; Zhang, W. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In Proceedings of the 20th ACM SIGKDD International Conference On Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 601–610.
6. Zhang, W.; Paudel, B.; Zhang, W.; Bernstein, A.; Chen, H. Interaction embeddings for prediction and explanation in knowledge graphs. In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, Melbourne, Australia, 11–15 February 2019; pp. 96–104.
7. Cabanas-Abascal, A.; Rodríguez-González, A.; Casado-Lumbreras, C.; Fernández-González, J.; Jiménez-López, D. POST-VIA: Develop individualized marketing strategies for tourists. In *Electronic Business and Marketing*; Springer: Berlin, Heidelberg, 2013; pp. 29–42.
8. Fensel, A.; Kärle, E.; Toma, I. TourPack: Packaging and Disseminating Touristic Services with Linked Data and Semantics. In Proceedings of the 1st International Workshop on Semantic Technologies (IWOST), Changchun, China, 9–12 March 2015; pp. 43–54.
9. Piercy, N. Positive and negative cross-channel shopping behaviour. *Mark. Intell. Plan.* **2012**, *30*, 83–104, doi:10.1108/02634501211193930. [CrossRef]
10. Fensel, A.; Toma, I.; García, J.M.; Stavrakantonakis, I.; Fensel, D. Enabling customers engagement and collaboration for small and medium-sized enterprises in ubiquitous multi-channel ecosystems. *Comput. Ind.* **2014**, *65*, 891–904, Special Issue: New Trends on E-Procurement Applying Semantic Technologies, doi:10.1016/j.compind.2014.02.001. [CrossRef]
11. Payne, E.M.; Peltier, J.W.; Barger, V.A. Omni-channel marketing, integrated marketing communications and consumer engagement: A research agenda. *J. Res. Interact. Market.* **2017**, *11*, 185–197, doi:10.1108/JRIM-08-2016-0091. [CrossRef]

12. Luczak-Roesch, M.; Hollink, L.; Berendt, B. Current directions for Usage Analysis and the Web of Data: The diverse ecosystem of Web of Data access mechanisms. In Proceedings of the 25th International Conference Companion on World Wide Web, Montreal, QC, Canada, 11–15 April 2016; pp. 885–887.
13. Di Penta, M.; Esposito, R.; Villani, M.L.; Codato, R.; Colombo, M.; Di Nitto, E. WS Binder: A framework to enable dynamic binding of composite web services. In Proceedings of the 2006 International Workshop on Service-Oriented Software Engineering, Shanghai, China, 27–28 May 2006; pp. 74–80.
14. Cano, A.E.; Dadzie, A.S.; Ciravegna, F. Travel mashups. In *Semantic Mashups*; Springer: Berlin, Heidelberg, 2013; pp. 321–347.
15. Casillo, M.; Clarizia, F.; Colace, F.; Lombardi, M.; Pascale, F.; Santaniello, D. An approach for recommending contextualized services in e-tourism. *Information* **2019**, *10*, 180. [[CrossRef](#)]
16. Peltier, J.W.; Zahay, D.; Lehmann, D.R. Organizational learning and CRM success: A model for linking organizational practices, customer data quality, and performance. *J. Interact. Market.* **2013**, *27*, 1–13. [[CrossRef](#)]
17. Broilo, P.L.; Espartel, L.B.; Basso, K. Pre-purchase information search: Too many sources to choose. *J. Res. Interact. Market.* **2016**, *10*, 193–211, doi:10.1108/JRIM-07-2015-0048. [[CrossRef](#)]
18. Piao, G.; Breslin, J.G. Exploring dynamics and semantics of user interests for user modeling on Twitter for link recommendations. In Proceedings of the 12th International Conference on Semantic Systems, Leipzig, Germany, 12–15 September 2016; pp. 81–88.
19. Arbelaitz, O.; Gurrutxaga, I.; Lojo, A.; Muguerza, J.; Pérez, J.M.; Perona, I. Web usage and content mining to extract knowledge for modelling the users of the Bidasoa Turismo website and to adapt it. *Expert Syst. Appl.* **2013**, *40*, 7478–7491. [[CrossRef](#)]
20. Wolny, J.; Charoensuksai, N. Mapping customer journeys in multichannel decision-making. *J. Direct Data Digital Market. Pract.* **2014**, *15*, 317–326. [[CrossRef](#)]
21. Wang, Q.; Shukla, P. Linking sources of consumer confusion to decision satisfaction: The role of choice goals. *Psychol. Mark.* **2013**, *30*, 295–304. [[CrossRef](#)]
22. Gaffney, A. Improving customer satisfaction through more effective ICT strategies. *Alcatel-Lucent Enrich. Commun.* **2007**, *1*, 1–7.
23. Stone, M.D.; Woodcock, N.D. Interactive, direct and digital marketing: A future that depends on better use of business intelligence. *J. Res. Interact. Mark.* **2014**, *8*, 4–17, doi:10.1108/JRIM-07-2013-0046. [[CrossRef](#)]
24. Google. What Users Want Most From Mobile Sites Today. Google Survey Conducted by Sterling Research and Smith Geiger. Available online: <https://www.thinkwithgoogle.com/advertising-channels/mobile-marketing/what-users-want-most-from-mobile-sites-today/> (accessed on 28 April 2020).
25. Shpak, N.; Kuzmin, O.; Dvulit, Z.; Onysenko, T.; Sroka, W. Digitalization of the Marketing Activities of Enterprises: Case Study. *Information* **2020**, *11*, 109. [[CrossRef](#)]
26. Persaud, A.; Azhar, I. Innovative mobile marketing via smartphones: Are consumers ready? *Market. Intell. Plan.* **2012**, *30*, 418–443, doi:10.1108/02634501211231883. [[CrossRef](#)]
27. Toma, I.; Stanciu, C.; Fensel, A.; Stavrakantonakis, I.; Fensel, D. Improving the Online Visibility of Touristic Service Providers by Using Semantic Annotations. In *The Semantic Web: ESWC 2014 Satellite Events*; Springer International Publishing: Cham, Switzerland, 2014; pp. 259–262.
28. Battle, R.; Kolas, D. Enabling the geospatial Semantic Web with Parliament and GeoSPARQL. *Semant. Web* **2012**, *3*, 355–370, doi:10.3233/SW-2012-0065. [[CrossRef](#)]
29. Akbar, Z.; Fensel, A.; Fensel, D. An Ontology-Based Coordination and Integration of Multi-Channel Online Communication. *Int. J. Metadata Semant. Ontol.* **2017**, *12*, 219–231, doi:10.1504/IJMSO.2017.093646. [[CrossRef](#)]
30. Kärle, E.; Fensel, A.; Toma, I.; Fensel, D. Why Are There More Hotels in Tyrol than in Austria? Analyzing Schema.org Usage in the Hotel Domain. In *Information and Communication Technologies in Tourism 2016*; Springer International Publishing: Cham, Switzerland, 2016; pp. 99–112.
31. Akbar, Z.; Kärle, E.; Panasiuk, O.; Şimşek, U.; Toma, I.; Fensel, D. Complete Semantics to Empower Touristic Service Providers. In *On the Move to Meaningful Internet Systems*; Springer International Publishing: Cham, Switzerland, 2017; pp. 353–370.

32. Stavrakantonakis, I.; Fensel, A.; Fensel, D. Linked Open Vocabulary Ranking and Terms Discovery. In Proceedings of the 12th International Conference on Semantic Systems, Leipzig, Germany, 12–15 September 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 1–8, doi:10.1145/2993318.2993338. [[CrossRef](#)]
33. Kärle, E.; Şimşek, U.; Fensel, D. Semantify.it, a Platform for Creation, Publication and Distribution of Semantic Annotations. In Proceedings of the 11th Eleventh International Conference on Advances in Semantic Processing (SEMAPRO), Barcelona, Spain, 12–16 November 2017; pp. 22–30.
34. Şimşek, U.; Kärle, E.; Holzkecht, O.; Fensel, D. Domain Specific Semantic Validation of Schema.org Annotations. In *Proceedings of the International Andrei Ershov Memorial Conference on Perspectives of System Informatics*; Springer International Publishing: Cham, Switzerland, 2018; pp. 417–429.
35. Pellegrini, T.; Mireles, V.; Steyskal, S.; Panasiuk, O.; Fensel, A.; Kirrane, S. Automated rights clearance using semantic web technologies: The DALICC framework. In *Semantic Applications*; Springer Vieweg: Berlin, Heidelberg, 2018; pp. 203–218.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).