

Data-driven Initiatives of Destinations Supporting Sustainable Tourism

Completed Research

Stefan Neubig

Outdooractive AG

stefan.neubig@outdooractive.com

Nicolas Gehring

Technical University of Munich

n.j.gehring@tum.de

Jörg Weking

Technical University of Munich

joerg.weking@tum.de

Julius Baecker

Technical University of Munich

julius.baecker@tum.de

Andreas Hein

Technical University of Munich

andreas.hein@tum.de

Helmut Krcmar

Technical University of Munich

helmut.krcmar@tum.de

Abstract

Digitization has arrived at the heart of global tourism. However, due to challenges such as climate change and overtourism, touristic destinations not only need to become more digital but also more sustainable. Although ever-increasing quantities of touristic data provide a growing basis for innovative digital solutions, destinations vary widely in driving sustainable, data-driven initiatives due to uncertain targets and challenging data acquisition. Despite its rising importance, research still lacks a thorough understanding of how destinations can use data to become more sustainable. We build on interviews with ten destination management organizations (DMOs) to develop a taxonomy of their data-driven initiatives for sustainable tourism to address this issue. Our findings contribute to research by bridging the gap between available data sources and important targets of data-driven initiatives from the unique perspective of touristic destinations. For practice, we provide a structured overview for positioning own ambitions of DMOs.

Keywords

Sustainable Tourism, Taxonomy, Big Data, Destination Management Organizations (DMOs)

Introduction

Tourism has become increasingly globalized, resulting in increased tourist flows and the diversification of destinations. At the same time, digitization provides an ever-increasing variety of data, often referred to as big data, which is expected to greatly impact the tourism industry (Samara et al. 2020; Yallop and Seraphin 2020). The availability of big data offers novel opportunities to address challenges such as overtourism. Caused by low-cost airlines, inexpensive accommodations, or overexposure due to social media, overtourism is harmful to the environment and negatively affects tourism acceptance among local inhabitants. Hence, overtourism is not only a growing concern of touristic destinations but also gained attention within the research community (Capocchi et al. 2019).

Overtourism is also at the heart of the development of sustainable tourism (i.e., developing touristic destinations while ensuring socio-economic, cultural, and environmental sustainability criteria), driven by so-called destination management organizations (DMOs). DMOs are responsible, among other things, for the marketing and infrastructural development of one or multiple incorporated destinations. Focusing on sustainability, they ensure the long-term competitiveness of a destination, while taking care of the preservation of its natural environment, its cultural heritage, as well as the needs of other involved stakeholders (e.g., tourism-related enterprises). As a result, managing sustainable tourism is a complex

task, which includes (but is not limited to) the need of intelligent visitor guidance (e.g., promoting alternative places to distribute visitors), service personalization, product and service optimization, as well as infrastructure adjustments and mobility services. With the adaption of digital technology, however, DMOs may be able to address these challenges, while, at the same time, decreasing costs, streamlining operations, and improving productivity and efficiency (Ivanov 2019).

While previous research has mainly focused on literature and case studies to provide a better understanding of the use of data in tourism (Li et al. 2018), this work aims at investigating data initiatives of DMOs (in the following only referred to as *data initiatives*) that are currently being performed or planned and seek to support the development of sustainable tourism. We focus our work on understanding the structure of such initiatives, that is (i) which data is available, (ii) how it can be acquired, (iii) what insights can be derived from it, and (iv) what use cases are involved in the sustainable development of a touristic destination. As an illustrating example, a beach destination could exploit real-time parking lot data (available data), collected by physical sensors (acquisition), to derive the number of beach visitors (insight). Afterwards, this information could be used for a more effective visitor guidance, by informing arriving visitors early enough and in an automatized way, so that they can navigate to an alternative location. Therefore, to describe the structure of such data initiatives, we aim at developing a conceptually grounded taxonomy. More precisely, we seek to answer the following research question:

Which conceptually grounded and empirically validated characteristics structure data-driven initiatives for sustainable tourism?

We follow a structured taxonomy development method based on qualitative interviews to reveal the characteristics of data-driven initiatives for sustainable tourism. Thereby, we first draw on expert interviews with ten DMOs from Germany to gain rich information on DMO initiatives. We then apply iterative cycles for taxonomy development using qualitative coding techniques to develop the desired taxonomy for structuring data-driven initiatives for sustainable tourism. The resulting taxonomy contributes to research by bridging the gap between available and needed data sources and corresponding targets of data-driven DMO initiatives. We increase the understanding on the current state and planned ambitions of such initiatives. DMOs can use our findings for positioning own initiatives or planning new ambitions systematically. Our illustrative case insights further support practical applicability.

Related Work

As in many research fields, technologies like big data and artificial intelligence (AI) are seen as key driver in the current and future of tourism. Especially the application of big data and AI technologies in different fields have been addressed extensively in previous research (Ivanov 2019). Concerning overtourism and associated visitor guidance, tourism demand forecasting is of great relevance. Song et al. (2019) review 211 publications between 1968 and 2018 to investigate how technologies for tourism demand forecasting changed over time. They conclude that despite proposed forecasting models have become more complex and also have improved largely with respect to generalization, there is still no single method that can be applied for all kinds of situations, thus, research is still ongoing. As a sidenote, however, they also mention that previous research mainly focusses on international tourism demand, while there is little research with respect to domestic tourism. A different area of data use in tourism is sentiment analysis, that is, processing free text by the means of natural language processing to extract sentiment scores. Alaei et al. (2019) review existing approaches for sentiment analysis and emphasize its importance based on the grown volume of data, which makes it impossible to review online feedback manually. Thelwall (2019) even considers sentiment analysis a key component of tourism big data research and illustrates its value based on a range of different examples from social media content. Another popular application of AI in tourism are chatbots and smart assistants. Fensel et al. (2020) leverage integrated touristic data and describe how knowledge graph technology can be used in order to create such technologies. Another direction of AI in tourism that has been addressed are robots, however, most of the existing studies are purely conceptual (Ivanov 2019).

Besides investigating the application of concrete technologies, previous research has also addresses the consequences of big data and AI in tourism at a more general level. Yallop and Seraphin (2020) review opportunities and risks regarding the use of big data in tourism and hospitality organizations. After investigating both academic and industry databases and reports, they conclude that big data can be regarded as a beneficial technology, however, comes at the expense of certain risks, including ethical,

privacy, and security concerns. Among other things, they highlight the beneficial use of big data technologies for revenue management and exemplify the use and combination of internal and external data in this context. Samara et al. (2020) review the benefits and overall role of big data and artificial intelligence in the tourism domain. In contrast to Yallop and Seraphin (2020), however, they do not focus on disadvantages or risks, but on the ability of data-driven approaches to overcome challenges and highlight the importance of adopting such technologies to stay competitive. Ivanov (2019) focusses on the impact of robots, artificial intelligence and service automation and emphasizes the emergence of two large segments, that is, high-tech tourism companies that are mainly offering robot-delivered services and high-touch companies mainly relying on human interaction. Despite of the mentioned research, however, there are still plenty of research gaps in the topic of understanding the impact of big data and AI on companies in the tourism domain, especially when it comes to the topic of sustainability (Ivanov 2019).

Moreover, sharing data across the borders of touristic destinations (both domestically and across national borders) has been identified as an important research direction. Open data initiatives have emerged and address the need of exchanging data among different DMOs and service providers. Şimşek et al. (2019) describe how they adapt the well-known *schema.org* ontology to create more expressive semantic data models for the tourism domain. Besides, the AlpineBits Alliance (2019) is an initiative located in South Tyrol that aims at creating a set of standards specifically targeting the alpine areas. Within the scope of the BayernCloud project, a multi-domain reference architecture (Baecker et al. 2019) has been developed, targeting the exchange of (open) data in its touristic fragment.

While research on how (big) data can be used to sustain the innovation process of tourism destinations has in general received little consideration (Del Vecchio et al. 2018), this particularly holds for the on-site perspective of individual DMOs. For this purpose, we introduce a taxonomy that considers the specifics of data initiatives performed by individual DMOs taking into account the nature of available data, the on-site data collection, the associated insights, and the specific targets, DMOs seek to achieve.

Research Methodology

This study follows a two-stage research approach. First, our data collection draws on qualitative interviews with representatives from a total of ten DMOs. Second, we develop a taxonomy in iterative cycles based on the guidelines proposed by Nickerson et al. (2013).

Data Collection

For data collection, we conducted a total of ten qualitative interviews with German DMOs to gain rich information on data-driven initiatives of touristic destinations (see Table 1). We selected our interview partners to cover a wide range of touristic focuses (Hiking (H), Biking (B), Relaxing (R), Swimming (S), Winter sports (WS), Sightseeing (SI)) and environmental properties (Alpine (A), Lakes (L), Nature (N), Historical buildings (HB), Sea and Beach (SB)) distributed all over Germany. Our focus on qualitative case data enables the generation of novel theories in this little-examined field by providing deep insights into the topic under investigation (Eisenhardt 1989). Therefore, we applied semi-structured interviews to increase the explorative character of the interviews in our qualitative data collection (Myers and Newman 2007). The interviews focused on data initiatives of the DMOs that are currently running as well as initiatives that are planned, but also expert assessments of what should be done in the future. To avoid active listening (McCracken 1988), we phrased the questions in an unobtrusive and non-directive manner. Furthermore, we conducted all interviews in the respondent's language to avoid ambiguity of language (Myers and Newman 2007).

We divided all interviews into five parts: (1) a personal introduction of the interviewer and organizational matters; (2) personal introduction of the interviewee including the current position in the DMO; (3) questions on the status quo on the use of data and analysis in the DMO; (4) questions on the use cases of data and analysis in the DMO; (5) questions on further potentials relating to the use of data and where there is a lack of data and how the yet non-existing information could be used. The open questions centered on understanding how DMOs already use data and analysis to assist DMO initiatives, what further potential lies in the application of data and analysis, and which challenges concern a lack of data and information for DMOs. Alongside the interviews, publicly available information from the web presences of the DMOs provided additional insights and supported data triangulation (Yin 2003).

ID	Position	Touristic Focus of DMO											Interview Duration
		H	B	R	S	WS	SI	A	L	N	HB	SB	
IP-1	Deputy Managing Director	X	X	X	X	X	X	X	X	X	X		46 min
IP-2	Project Manager					X		X					48 min
IP-3	Divisional Management “sustainable recreation and tourism”	X				X		X	X	X			32 min
IP-4	Digital Marketing - Systems, Campaigns	X	X	X	X	X	X	X	X	X	X		62 min
IP-5	Authorized officer, Head of Strategic Development	X	X	X	X	X	X	X	X	X	X		48 min
IP-6	Product Management Information and Communication Technology / Online Marketing	X	X	X	X	X	X	X	X		X		61 min
IP-7	Deputy manager at a Tourist Information	X	X		X		X		X	X	X		55 min
IP-8	Managing Director, responsible for finances, personnel, training, projects, public relations	X	X	X	X			X		X			47 min
IP-9	Tourism Director	X	X	X		X	X	X		X	X		45 min
IP-10	Deputy Board Member Online Marketing	X	X	X	X		X			X	X	X	73 min

Table 1. Overview of Interview Partners (IP)

Data Analysis

For data analysis, we applied the taxonomy development approach proposed by Nickerson et al. (2013) based on standards for qualitative research (Mayring and Fenzl 2014; Miles and Huberman 1994). Following Nickerson et al. (2013), we describe a taxonomy T as a set of n dimensions D_i ($1 \leq i \leq n$), each of which consisting of $m_i \geq 2$ characteristics C_{ij} ($2 \leq j \leq m_i$). To acknowledge the complex nature of data-driven initiatives (Hunke et al. 2019), however, we allow the selection of multiple characteristics C_{ij}, C_{ik} ($j \neq k$) within a single dimension D_i , since data-related dimensions tend to occur in combined forms in practice (Baecker et al. 2021). In addition, this modification ensures to keep the taxonomy both concise and comprehensive, which are two key attributes of useful taxonomies (Nickerson et al. 2013).

For taxonomy development, we first specified a meta characteristic that serves “as the basis for the choice of characteristics in the taxonomy” (Nickerson et al. 2013). We, therefore, chose as our meta-characteristic the “data requirements and analysis objectives of existing and prospective data-driven initiatives of tourism destinations”. Second, we agreed on the eight objective and five subjective ending conditions for the taxonomy development process suggested by Nickerson et al. (2013). Based on the interview transcripts, we applied empirical-to-conceptual cycles with our DMO cases. We conducted a structured analysis for qualitative data following our chosen meta-characteristic to code the case information into common characteristics (Nickerson et al. 2013). Open and axial coding supported the analysis of our transcripts (see Table 2) to identify the characteristics and dimensions of the taxonomy (Strauss and Corbin 1990). We iteratively classified the instances to the current state of the taxonomy and added further dimensions and characteristics until all cases were included. Finally, we checked the ending conditions by Nickerson et al. (2013). We concluded on 15 dimensions and 65 characteristics, which we further aggregated into four meta-dimensions for readability.

Interview statement	Open Code	Axial Code
“[It is] in quite a lot of destination really just the <u>most urgent issue</u> : How can I <u>direct visitors</u> ? How can I <u>steer visitors</u> ? And that's where we are really at the very beginning in terms of the data we need for this”. (IP-4)	Visitor Guidance	Use Case Targets

Table 2. Illustration of the Coding Process

Results

The resulting taxonomy of our work is illustrated in Table 1. Each line i corresponds to a single dimension D_i , while each column j describes the respective characteristic C_{ij} . Four meta dimensions (Available Data, Data Acquisition, Gained Insights, and Use Cases) have been added for better readability.

	Dimension	Characteristics							
Available Data	Volatility	Static Data				Dynamic Data			
	Origin	Internal		External		Public		Open	
	Categories	Touristic Entities	User Interaction	Destination Status	Count & Movement	Customer	Nature & Environment	Social Data	
	Format	Unstructured			Semi-Structured		Structured		
Data Acquisition	Automatization	Manual			Semi-Automatic		Automatic		
	Collection Space	Virtual Space				Physical Space			
	Collection Object	Feedback	Subscription	Booking	Service	Sensor	Access Control	Tracker	
	Cost Factor	Free		Personnel		Technology		Data	
Gained Insights	Perspective	Online				Offline			
	Unit	Individual		Group		Location		Time	
	Type	Qualitative				Quantitative			
	Reference	Visitor	Motivation	Impact	Mood	(Movement) Pattern	Occupancy	Distribution	
Use Cases	Approach	Informing		Sharing		Monitoring		Forecasting	Benchmarking
	Stakeholder	Visitor		Local Inhabitant		Service Provider		DMO	
	Target	Visitor Information			Tourism Acceptance			Product & Service Optimization	
		Visitor Guidance			Improved Marketing			Infrastructure Adjustment	
		Efficiency Gain			Personalization			Collective Growth	

Table 1. Taxonomy to classify data initiatives of DMOs

In the following, we briefly summarize each (meta) dimension separately. For the sake of a better understanding, we underpin selected dimensions with exemplary quotes from the conducted interviews, which we translated into English.

Available Data

D₁: Volatility: Data types relating to tourism initiatives of DMOs can roughly be characterized as static or dynamic. Static data refers to data that is unlikely to change often (e.g., existing points of interest), while dynamic data refers to data that is subject to frequent changes (e.g., the number of people at a certain place).

D₂: Origin: During our interviews, we identified four levels of data origin. Internal data refers to data that is available to the DMO performing the classified initiative only, that is, data that is owned by this DMO (e.g., web analytics data from its website). In contrast, external data is obtained from other organizations, including service providers or other DMOs (e.g., by voluntary data sharing). This data origin is important for two reasons: first, DMOs may not have direct access to data, which resides within the scope of larger apps. IP-9 notes that especially outdoor and fitness tracking apps are great data suppliers. IP-4 adds,

The visitor will not have our app installed. We're not fooling ourselves, we're too small.

A second yet evenly important reason for using external data is that most DMOs do not have the necessary expertise in its maintenance. IP-1 states that data should always be maintained by its specific experts and underlines the need for creating and maintaining data points at their origin region. Besides the more generic characteristics of internal and external data, we distinguish further between data from public authorities

and open data, which we consider a special case of external data. While data of public authorities (e.g., official booking numbers of a destination) can be obtained by DMOs directly, it is not available for other organizations or individuals. In contrast, open data refers to data published under an open license, that is it can be obtained and used by everyone.

D₃: Categories: Besides volatility and origin, we also distinguish between different touristic data categories. The sum of all data that is available within this dimension can be seen as the digital representation of a destination. Characteristics of this dimension (i.e., categories) refer to different groups of data types, namely, touristic entities (e.g., events, points of interest), user interaction (e.g., web tracking), destination status (e.g., infrastructure, construction sites), count and movement (e.g., individual or aggregated visitor positions, frequented ways), customer (e.g., reviews, ratings), nature and environment (e.g., forestry data, water temperature), and social data (e.g., public holidays).

D₄: Format: A rather technical yet important distinction of data is given by its format. More precisely, data can be unstructured (e.g., textbooks), semi-structured (e.g., data on websites), or structured (e.g., data originating from APIs or databases). Depending on its format, data must be processed differently. Together with the origin (D_2), this dimension indicates the overall variety of the available data.

Data Acquisition

D₅: Automatization: In this dimension, initiatives are classified with respect to the effort of their data collection part. Manual data collection processes are completely performed by humans and can be time-consuming and tedious (e.g., counting visitors manually at a specific place). In contrast, semi-automatic approaches refer to data collection processes that involve both human interaction and some degree of automatization (e.g., a sensor that collects data automatically but requires manual synchronization to a computer). On the other hand, automatic approaches do not require any kind of human interaction (e.g., sensors that are fully connected to a cloud). While automatization certainly reduces the effort of data collection, IP-2 admits that gathering data about guests has mainly been done based on surveys and online questionnaires. Other DMO representatives add that even data that already exists digitally is collected manually (e.g., approximates of visitor counts from webcams, car traffic from online maps applications).

D₆: Collection Space: Data can be collected in both the virtual and the physical space. The virtual space includes data collection methods within digital services such as website analytics. Accordingly, physical space methods comprise all data collection methods that collect data in the physical world. As a result of our interviews, data collection methods in the virtual space seem to be more widespread nowadays, even for destinations that also use data collection methods in the physical space. As an example, IP-10 notes that, despite using quite progressive methods to collect data within the physical space, they still consider their web analytic data as the linchpin of their data acquisition. It is important to note, however, that collection methods in the virtual space according to our definition are not limited to collect data about virtual objects. One prominent example is given by digital platforms of the outdoor and fitness domains that allow to collect both data about virtual customer behavior (e.g., web analytics) as well as data about physical movement patterns (e.g., customer-recorded tracks).

D₇: Collection Object: While the collection space (represented by D_6) refers to the space where data is collected, the collection object refers to the object that performs the actual data collection. Examples that have already been mentioned include (online and offline) services or sensors. Other collection objects include bookings that usually contain a lot of demographic data.

So I know exactly who the guest is. Family of four with two children, including their ages, all of this data is known. 4-star hotel, half-pension, for example. (IP-4)

Besides, other collection object characteristics include visitor feedback (e.g., based on surveys after a stay (IP-8)), subscriptions (e.g., apps but also newsletters (IP-10)), trackers, which involves both online (e.g., on websites) and offline trackers (e.g., WiFi and Bluetooth beacon trackers (IP-6)), as well as access controls. One prominent example for visitor data that can be captured by the means of access controls is a guest card, which allows creating movement profiles (IP-9).

D₈: Cost Factor: Costs of data acquisition can involve several types. Although some destinations report that volunteers help to collect data (e.g., about infrastructural conditions (IP-9)), the acquisition of data usually involves further costs. Even when relying on internal or open data only, costs usually include

technology costs (e.g., purchasing and maintaining hardware) and salary expenses, as collected data must be integrated into the existing systems, which can involve serious effort (see automatization (D_5)). Using external data, however, can lead to significant data fees.

Gained Insights

D_9 : Perspective: While the collection space dimension (D_6) explains where a data collection activity is performed (virtual or physical space), this dimension refers to whether the generated insights from the collected data describe online or offline states. To clarify, data can be collected in the virtual space (e.g., outdoor and fitness apps) but still refer to tours that were conducted by visitors offline. In the same fashion, surveys performed within the physical space can ask questions regarding online behavior (e.g., the use of outdoor and fitness apps).

D_{10} : Unit: Gained insights may refer to individuals, groups of people, locations, or time. This is important to differentiate as we can directly conclude the kind of granularity under investigation. Insights referring to individuals contain personal information and thus can be used for personalization use cases. On the other hand, insights about groups can help to optimize services and products (D_{15}), as it rather captures knowledge about aggregated guest needs. Two other important characteristics of this dimension are location and time. As an example, the occupancy of locations usually depends on the time (e.g., day of week).

D_{11} : Type: Besides the perspective and unit, insights can be further distinguished between qualitative and quantitative insights. While qualitative insights may contain information such as the motivation of visitors, examples for quantitative insights include occupancies of certain locations and (statistical) distribution of visitors. One conclusion of our interviews is that qualitative data seems harder to track than quantitative data, especially, when it comes to day tourists (although this group is generally harder to track than overnight guests (IP-7)). IP-10 notes,

As far as day tourists are concerned, no we actually have no way to get qualitative data.

D_{12} : Reference: The reference dimension captures possible objects and properties gained insights may relate to. Insights may refer to visitors (both individuals and groups), their (online and offline) movement patterns, (statistical) distributions or occupancies, as we already outlined in several examples above. However, depending on what data is acquired, obtained insights can also relate to the motivations or intents of visitors. IP-3 underlines the importance of visitor's motivations, asking

Why is one chosen? Are you elected because you are located inland or is it a conscious vacation decision?

Other characteristics to classify insights include mood (e.g., mood of visitors or local inhabitants) and impact (i.e., getting insights about the impact of previously performed actions, such as marketing campaigns).

Use Cases

D_{13} : Approach: Use cases as identified during our interviews involve different actions that our taxonomy characterizes in the approach dimension. Informing refers to the action of passing information onto a stakeholder (D_{14}), sharing refers to the action of sharing data among different stakeholders. In the data origin dimension (D_2), we already highlighted the importance of sharing data. Besides the aforementioned reasons of availability and data expertise, one important use case is reducing the maintenance effort that may originate from doubled work, such as the redundant creation of points of interest. As I8 notes,

The vision is basically that you only have to create a POI once. if I now want to make use of this dataset on my website, then I take it when it is free there. (I8)

This points out the importance of open data that is approached by different cross-organizational initiatives and was introduced in the Background section. However, also smaller cross-organizational initiatives turned out to be helpful. IP-1 describes a content network in their region consisting of 50 co-located websites that share data with each other. Besides informing and sharing, monitoring refers to the action of describing a destination's current state (e.g., with respect to the occupancy of places), while forecasting

refers to predicting a destination's future state. Finally, benchmarking refers to the comparison of different stakeholders (e.g., the performance of two DMOs).

D₁₄: Stakeholder: Next to the approach dimension, the stakeholder dimension explains what stakeholders are involved in the planned use cases. According to our interviews, stakeholders involve visitors, local inhabitants, touristic service providers, and other DMOs.

D₁₅: Target: Following the results of the conducted interviews, we identified nine targets. By informing visitors (visitor information), guests can be sensitized to the specific conditions of a certain region, such as the animals that live there and the behaviors that are necessary to protect them. Tourism acceptance is an important touristic challenge that appeared as a consequence of overtourism and refers to the mood of local inhabitants. Visitor guidance aims at distributing guests with respect to time and location and is a possible counter measure to overtourism and a way to improve both tourism acceptance and environmental protection. One way of implementing visitor guidance is recommending suitable alternatives that fits a guest's needs, thus, personalization is another important target. On a more general level, both product and service optimization, as well as improved marketing measures are further targets and can even support visitor guidance by targeting marketing campaigns towards less crowded places. Infrastructure adjustments (e.g., strengthening public transport) can help reduce car traffic in favor of more sustainable transportation means. Efficiency gain (e.g., automatization of data collection) frees personnel and allows to focus more on the sustainability-related tasks described above. Finally, collective growth (e.g., cross-DMO benchmarking) fosters learning processes and allows collaboratively shaping a sustainable future together.

Discussion

One compelling result of our interviews is that many DMOs, while being convinced about the usefulness of the use of data to support the development of their destination, are still struggling in one or multiple stages of their data initiatives. Challenges range from knowledge about which data to collect and how, its processing to derive appropriate insights, as well as concrete use cases that really help them improving their destination in a sustainable way. Moreover, data initiatives as discussed within this paper are quite complex. They differ from traditional business intelligence processes in that they need to address particular sustainability issues with a wide range of conflicting objectives (e.g., revenue as well as environmental protection) and stakeholders (e.g., visitors, local inhabitants as well as touristic enterprises). Furthermore, they must be considered on a cross-destination basis, as many of the underlying challenges can only be solved through cooperation (e.g., overtourism). The proposed taxonomy can be used to classify data initiatives of DMOs and is helpful at any stage, including planning, performing, and reflecting in retrospective, as it incorporates the challenges mentioned above and sheds light on the data types, acquisition methods, insights and use cases that may help to guide a destination's development in a more sustainable direction.

One important finding is the relevance of economic factors, even when targeting other sustainability goals as main targets, which is also reflected by the target dimension (*D₁₅*). On the one side, economic success can be seen as a precondition as data purchases and data collection technology are considerably expensive, as mentioned by IP-9 and IP-10, respectively. If these investments did not pay off economically, probably few DMOs would take this step. On the other side, however, both may even go hand in hand. As an illustrating example, IP-10 describes how they use humans to manually count visitors and send the results via text messaging to administrators who, in turn, add this data to the destination's website. On the other side, IP-10 also reports that they are using advanced technologies like sensors and image recognition technology to collect visitor numbers. They are also planning to continue their automatization efforts, which is likely to result in both a larger and better data base that can be used for improved visitor guidance as well as reduced personnel costs and a gain in efficiency. Furthermore, the more stable a data base is the better DMOs can address other economic goals, including improved marketing and service optimization.

Contribution

Investigating the data-driven initiatives of DMOs for promoting sustainable tourism has implications for theory and practice. Our developed taxonomy contributes to research by strengthening the important perspective of DMOs within the scientific discourse on the development of sustainable tourism. We thereby provide details on available data, data acquisition, analytical insights, and use cases of data-driven DMO

initiatives which create a novel view in the digital development of touristic destinations. Thus, our taxonomy paves the way for addressing important future directions in the context of sustainable tourism development of DMOs.

For practice, our findings provide valuable orientation for DMOs. First, the taxonomy helps DMOs to position the status quo of their data-driven initiatives in the light of other destinations. It further inspires potential use cases on which to increase focus. Finally, the taxonomy supports practical exchange by providing a structured overview of relevant characteristics as well as a consistent terminology.

Limitations

Like other studies, our work has some limitations. As of the nature of taxonomies, no taxonomy is ever complete or perfect. Although we applied rigorous ending conditions as proposed by Nickerson et al. (2013), it is still possible that more dimensions or characteristics pop up when including additional DMOs. Furthermore, our study focuses on DMOs from Germany. While we address a variety of touristic focuses (such as hiking, swimming, sightseeing, or winter sports) and environmental properties (e.g., alpine, lakes, nature, beach), future research may add the perspective of destinations from other countries and cultures.

Avenues for Future Research

During our study, we also identified fruitful avenues for future research. First, as a consequence of our limitation to the German tourism area, we welcome any contribution to our taxonomy that adds the perspectives of other countries and cultures to create evidence on the transfer of the taxonomy to other touristic destinations. Second, the conducted interviews shed also light on aspects that are worth investigating, including challenges that are involved in data initiatives. Many DMOs reported that there exist serious problems with accessing data that is principally available.

I'll say the data is there. But they are not yet accessible to us (IP-1).

These data are available. Getting at it is certainly difficult. (IP-2).

This can be due to the costs of proprietary data or because the needed data is not sold by its owner (e.g., traffic from online maps). The same applies to data that could potentially be collected but requires expensive hardware that DMOs are reluctant to buy. We consider reasoning over such associated challenges an important direction for future research. Third, as of the nature of taxonomies, they are limited to specific classification tasks and thus are unable to express certain relationships within the described dimensions and characteristics, such as – in our case – the relation between data types and ways to acquire that data, or the relation between these data types and the targets, which may help DMOs to decide which data to collect. Making this decision is a serious challenge for DMOs as we found during our interviews. Therefore, we conclude that future research should address this challenge and propose our taxonomy as a basis for quantitative analysis and classification of data initiatives of DMOs to detect interrelationships as mentioned above. Such an investigation could result in an expressive ontology (Guarino et al. 2009) which uses our taxonomy at its core (Nickerson et al. 2013).

Conclusion

In this study, we developed a taxonomy that identifies and classifies data-driven initiatives of DMOs targeting the development of sustainable tourism. Based on interviews with ten DMOs, we derived a total of 15 dimensions and 65 characteristics that increase the understanding of the current state and future ambitions of DMOs in this global field of action. Our taxonomy offers a generalized structure incorporating a variety of different DMO cases and expert assessments on existing and prospective data-driven initiatives of tourism destinations. We, therefore, see fruitful avenues for DMOs in the development of data-driven initiatives to advance sustainable tourism.

Acknowledgements

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

REFERENCES

- Alaei, A. R., Becken, S., and Stantic, B. 2019. "Sentiment Analysis in Tourism: Capitalizing on Big Data," *Journal of Travel Research* (58:2), pp. 175–191.
- AlpineBits Alliance. 2019. "AlpineBits. Open Data Exchange in the Alpine Tourism." (<https://www.alpinebits.org>, accessed August 30, 2019).
- Baecker, J., Böttcher, T. P., and Weking, J. 2021. "How Companies Create Value From Data—A Taxonomy on Data, Approaches, and Resulting Business Value," *Twenty-Ninth European Conference on Information Systems (ECIS 2021)*.
- Baecker, J., Engert, M., Müller, J., Neubig, S., Schaffer, N., Kunz, I., Schneider, A., Krcmar, H., and Eckert, C. 2019. "BayernCloud Gesamtreferenzarchitektur."
- Capocchi, A., Vallone, C., Pierotti, M., and Amaduzzi, A. 2019. "Overtourism: A Literature Review to Assess Implications and Future Perspectives," *Sustainability* (11:12), p. 3303.
- Del Vecchio, P., Mele, G., Ndou, V., and Secundo, G. 2018. "Open Innovation and Social Big Data for Sustainability: Evidence from the Tourism Industry," *Sustainability* (10:9), Multidisciplinary Digital Publishing Institute, p. 3215.
- Eisenhardt, K. M. 1989. "Building Theories from Case Study Research," *Academy of Management Review*, pp. 532–550.
- Fensel, D., Şimşek, U., Angele, K., Huaman, E., Kärle, E., Panasiuk, O., Toma, I., Umbrich, J., and Wahler, A. 2020. "Why We Need Knowledge Graphs: Applications," in *Knowledge Graphs: Methodology, Tools and Selected Use Cases*, D. Fensel, U. Şimşek, K. Angele, E. Huaman, E. Kärle, O. Panasiuk, I. Toma, J. Umbrich, and A. Wahler (eds.), Cham: Springer International Publishing, pp. 95–112.
- Guarino, N., Oberle, D., and Staab, S. 2009. "What Is an Ontology?," in *Handbook on Ontologies*, S. Staab and R. Studer (eds.), Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–17.
- Hunke, F., Engel, C., Schüritz, R., and Ebel, P. 2019. "Understanding the Anatomy of Analytics-Based Services – A Taxonomy to Conceptualize the Use of Data and Analytics in Services," in *ECIS 2019 Proceedings . 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. Research Papers*, Stockholm/Uppsala: AIS eLibrary (AISEL), Paper:-25.
- Ivanov, S. 2019. "Ultimate Transformation: How Will Automation Technologies Disrupt the Travel, Tourism and Hospitality Industries?," *Zeitschrift Für Tourismuswissenschaft* (11:1), pp. 25–43.
- Li, J., Xu, L., Tang, L., Wang, S., and Li, L. 2018. "Big Data in Tourism Research: A Literature Review," *Tourism Management* (68), pp. 301–323.
- Mayring, P., and Fenzl, T. 2014. "Qualitative Inhaltsanalyse," in *Handbuch Methoden Der Empirischen Sozialforschung*, N. Baur and J. Blasius (eds.), Wiesbaden: Springer Fachmedien Wiesbaden, pp. 543–556.
- McCracken, G. 1988. *The Long Interview*, Sage.
- Miles, M. B., and Huberman, A. M. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, sage.
- Myers, M. D., and Newman, M. 2007. "The Qualitative Interview in IS Research: Examining the Craft," *Information and Organization*, pp. 2–26.
- Nickerson, R. C., Varshney, U., and Muntermann, J. 2013. "A Method for Taxonomy Development and Its Application in Information Systems," *European Journal of Information Systems* (22:3), pp. 336–359.
- Samara, D., Magnisalis, I., and Peristeras, V. 2020. "Artificial Intelligence and Big Data in Tourism: A Systematic Literature Review," *Journal of Hospitality and Tourism Technology* (11:2), Emerald Publishing Limited, pp. 343–367.
- Şimşek, U., Angele, K., Kärle, E., Panasiuk, O., and Fensel, D. 2019. *A Formal Approach for Customization of Schema.Org Based on SHACL*, pp. 1–16.
- Song, H., Qiu, R. T. R., and Park, J. 2019. "A Review of Research on Tourism Demand Forecasting: Launching the Annals of Tourism Research Curated Collection on Tourism Demand Forecasting," *Annals of Tourism Research* (75), pp. 338–362.
- Strauss, A., and Corbin, J. 1990. *Basics of Qualitative Research*, Sage publications.
- Thelwall, M. 2019. "Sentiment Analysis for Tourism," in *Big Data and Innovation in Tourism, Travel, and Hospitality*, M. Sigala, R. Rahimi, and M. Thelwall (eds.), Singapore: Springer Singapore, pp. 87–104.
- Yallop, A., and Seraphin, H. 2020. "Big Data and Analytics in Tourism and Hospitality: Opportunities and Risks," *Journal of Tourism Futures* (6:3), Emerald Publishing Limited, pp. 257–262.
- Yin, R. 2003. "Case Study Research: Design and Methods," *Sage Publications, Inc*, p. 11.