Journal of Destination Marketing & Management 3 (2014) 198–209

Contents lists available at ScienceDirect

Journal of Destination Marketing & Management

journal homepage: www.elsevier.com/locate/jdmm

Research Paper

Big data analytics for knowledge generation in tourism destinations – A case from Sweden

Matthias Fuchs [a,*], Wolfram Höpken [b], Maria Lexhagen [a]

[a] Mid-Sweden University, Östersund, Sweden
[b] University of Applied Sciences Ravensburg-Weingarten, Weingarten, Germany

ARTICLE INFO

Article history:
Received 26 June 2014
Accepted 22 August 2014
Available online 22 September 2014

Keywords:
Big data analytics
Tourism destination
Destination management information system
Business intelligence
Data mining
Online Analytical Processing (OLAP)

ABSTRACT

This paper presents a knowledge infrastructure which has recently been implemented as a genuine novelty at the leading Swedish mountain tourism destination, Åre. By applying a Business Intelligence approach, the Destination Management Information System Åre (DMIS-Åre) drives knowledge creation and application as a precondition for organizational learning at tourism destinations. Schianetz, Kavanagh, and Lockington's (2007) concept of the '*Learning Tourism Destination*' and the '*Knowledge Destination Framework*' introduced by Höpken, Fuchs, Keil, and Lexhagen (2011) build the theoretical fundament for the technical architecture of the presented Business Intelligence application.

After having introduced the development process of indicators measuring destination performance as well as customer behaviour and experience, the paper highlights how DMIS-Åre can be used by tourism managers to gain new knowledge about customer-based destination processes focused on pre- and post-travel phases, like "*Web-Navigation*", "*Booking*" and "*Feedback*". After a concluding discussion about the various components building the prototypically implemented BI-based DMIS infrastructure with data from destination stakeholders, the agenda of future research is sketched. The agenda considers, for instance, the application of real-time Business Intelligence to gain real-time knowledge on tourists' on-site behaviour at tourism destinations.

1. Introduction

Since the advent of the WWW, major parts of tourism information processes and transactions are handled electronically (Buhalis, 2006; Fuchs, Höpken, Föger, & Kunz, 2010; Fuchs, Scholochow, & Höpken, 2010). Thus, customers leave electronic traces during all travel-related activities, like searching and trip planning, reservation and booking, service consumption as well as feedback provision in community web-sites (e.g. social media platforms) or through online surveys (Fuchs & Höpken, 2011). Consequently, a huge amount of data on customer needs and behaviour as well as perception is stored in various knowledge sources at tourism destinations. For instance, web-search data is stored in web server log-files, while survey data is stored in data bases of destination suppliers (Fuchs, Höpken, & Lexhagen, 2015). However, in tourism destinations these valuable knowledge sources typically remain unused (Höpken, Fuchs, Keil, and Lexhagen, 2011; Pyo, 2005).

In order to remove this deficiency, this paper presents a knowledge infrastructure which has been prototypically implemented on the base of real data, as a genuine novelty at the leading Swedish mountain tourism destination Åre. By applying a Business Intelligence approach (Larose, 2005), the proposed Destination Management Information system Åre (DMIS-Åre) drives knowledge creation and application as a precondition for organizational learning at tourism destinations (Pyo, Uysal, & Chang, 2002). From a theoretical standpoint the paper is based on Schianetz, Kavanagh, and Lockington's (2007) concept of the '*Learning Tourism Destination*'. Following these authors, through the generation, management and intelligent access of relevant information, the knowledge level of tourism stakeholders can be significantly increased. In order to foster learning processes among destination stakeholders, the '*Knowledge Destination Framework*' (Höpken et al., 2011) builds the fundament for the technical architecture of the proposed DMIS infrastructure. After having discussed theoretical foundations of the '*Knowledge Destination*' in Section 2, the elements of the framework architecture are described in Section 3. Subsequently, before highlighting how DMIS-Åre can be used by destination managers and suppliers, the development process of indicators measuring destination performance as well as tourist behaviour and experience is thoroughly discussed in Sections 4 and 5. The paper concludes

* Corresponding author.
E-mail addresses: Matthias.Fuchs@miun.se (M. Fuchs), Wolfram.Hopken@hs-weingarten.de (W. Höpken), Maria.Lexhagen@miun.se (M. Lexhagen).

http://dx.doi.org/10.1016/j.jdmm.2014.08.002
2212-571X/

by providing a summary and by outlining limitations and the agenda for future research steps.

2. The knowledge destination

Following the knowledge-based view of the firm (Grant, 1996), an organization's value is limited by the amount of knowledge within it. Thus, the sustainable development of whole industries as well as (e.g. tourism) regions and destinations is related to available (and accessible) knowledge which is needed to (re-)configure 'resources', especially knowledge-based resources, to remain competitive (Back, Enkel, & V. Krogh, 2007). Resources are defined as 'the totality of assets, capabilities, organizational processes, information, and knowledge controlled by an organization that enables it to conceive of and implement strategies that improve efficiency and effectiveness' (Barney, 1991, p. 101). However, only if these resources are perceived as valuable and scarce by customers and difficult to imitate and to substitute by competitors, will they provide competitive advantages in the long-run (Barney, 1997). If these conditions are fulfilled, the entrepreneurial activity of combining and reconfiguring resources will be based upon core competencies which, in turn, need to be validated and renewed through continuous knowledge acquisition and learning processes. This, 'dynamic capability' is described in the literature as the 'ability to integrate, build and reconfigure internal and external competences to address changing environments' (Teece, Pisano, & Shuen, 1997, p. 516). Accordingly, organizational learning is operationalized by two core capabilities: by efficiently multiplying established processes and operations (i.e. the *replication capability*), and by continuously modifying existing resource configurations through the acquisition and development of new core-competencies (i.e. the *reconfiguration capability*). Replication capabilities are typically driven by firm-internal knowledge transfer and related codification processes. By contrast, reconfiguration capabilities are predominantly determined by the capacity to absorb external knowledge and by the ability to deduce generalizable cause-effect relationships from existing knowledge applicable to a wider range of strategic options (Back et al., 2007; Tajeddini, 2010). Finally, it is empirically shown that the reconfiguration capability is positively affected by a firm's 'proximity to the customer', thus indicating the crucial relevance of customer-based knowledge for learning and innovation processes (Burman, 2002).

In the context of tourism research, the knowledge-based school of thought regards tourism as a complex social phenomenon where *knowledge* is the essential basis for tourism development and competitiveness (Jafari, 2001). This platform of thought postulates that through the generation and intelligent application of knowledge (especially on customer needs) information asymmetries between destination stakeholders can be reduced (Hallin & Marnburg, 2008; Shaw & Williams, 2009). While tourism destinations are viewed as 'value networks of competencies that coordinate complex social stakeholder constellations and resource configurations to deliver and mediate co-created tourist experiences' (Coles, Hall, & Duval, 2006), this leads to an enhanced collaboration and innovation capacity, which, in turn fosters market cultivation processes and improves service effectiveness by using destination resources in a more efficient and sustainable way (Buckley, 2012).

However, particular approaches are needed that support stakeholder collaboration and learning processes on an organizational and destination level. Schianetz, Kavanagh, and Lockington (2007) propose the concept of the *Learning Tourism Destination* and define two major areas of knowledge: (1) an area where knowledge is created and, (2) an area where knowledge is applied and learning occurs. By acknowledging that organizational, community and individual learning are highly interlinked, Schianetz et al. (2007) suggest that 'the learning focus should be on the understanding of how a tourism destination functions, how market possibilities can be enhanced, the requirements for adaptation to changing environments, how to promote collective awareness of economic, social and environmental risks and impacts, and how risks can be minimized and/or countered' (Schianetz et al., 2007, p. 1486). Finally, Schianetz et al. (2007) argue that the implementation of a networked infrastructure that collects customer-based data and also applies and disseminates gained knowledge, is fundamental to foster knowledge exchange between different organizations and enable effective learning cycles (Fuchs, Abadzhiev, Svensson, Höpken, & Lexhagen, 2013). Thus, it is clear why information and communication technologies (ICTS) are playing such a crucial role in realizing the full potential of a knowledge destination (Buhalis, 2006). Indeed, *destination competitiveness* is affected by the extent to which knowledge creation and application is supported by ICT-based infrastructures and services (Shaw & Williams, 2009). Accordingly, the proposed *knowledge destination framework* focuses on the inclusion of the customer, and represents the framework for a prototypically implemented Web-based infrastructure that collects customer-based data and creates and disseminates knowledge among destination stakeholders (Höpken et al., 2011). Thereby, we argue that knowledge creation and acquisition processes at tourism destinations can be significantly enhanced by applying methods of *Business Intelligence* (*BI*). BI is an umbrella term which comprises: (1) data identification and preparation, (2) database modelling and the population of a data warehouse, and (3) the application of Online Analytical Processing (OLAP) and data mining (DM) techniques, respectively (Hastie, Tibshirani, & Friedman, 2009; Larose, 2005). DM comprises statistical and machine learning techniques for identifying trends and patterns in huge data sources, like (a) classification (for example artificial neural networks [ANN], decision tree analysis, rule induction, *K*-Nearest Neighbour techniques), (b) estimation, (c) prediction (such as multivariate statistics, ANN), (d) clustering (e.g. *k*-means, hierarchical, Kohonen Networks) and (e) association rules (particularly for market basket analyses).

Literature only recently emphasizes BI and DM for knowledge creation in travel & tourism (Fuchs & Höpken, 2009; Magnini, Honeycutt, & Hodge, 2003; Min, Min, & Emam, 2002; Morales & Wang, 2008; Palmer, Montano, & Sesé, 2006; Wong, Chen, Chung, & Kao, 2006), and only few BI studies exist for tourism destinations (Cho & Leung, 2002; Fuchs et al., 2013; Höpken, Fuchs, Höll, Keil, & Lexhagen, 2013; Höpken et al., 2011; Höpken, Fuchs, & Lexhagen, 2014; Pyo et al., 2002). According to our proposed framework, knowledge activities deal with extracting information from different customer- and supplier-based sources as well as with the generation of relevant knowledge which can be applied in the form of intelligent services for customers or destination stakeholders (Fuchs et al., 2013). Thus, the *knowledge destination framework* (Höpken et al., 2011) distinguishes between a *knowledge generation* and a *knowledge application layer* (Fig. 1).

The *knowledge generation layer*, through methods of information gathering, extraction and storage, makes knowledge sources accessible to stakeholders: for instance, on the customer side, knowledge is generated through feedback mechanisms, like (e.g. online) surveys and review platforms (Gräbner, Zanker, Fliedl, & Fuchs, 2012; Sidali, Fuchs, & Spiller, 2012). Moreover, tourists' information traces (web search) are made explicit through web-mining (Liu, 2008; Pitman, Zanker, Fuchs, & Lexhagen, 2010). Furthermore, knowledge about tourists' buying behaviour is generated through mining transaction data, while tourists' mobility behaviour may be traced by GPS/WLAN-based position tracking (Zanker, Jessenitschnig, & Fuchs, 2010). On the supply side, knowledge about products can be extracted from information sources (web-sites) in the form of product profiles and availability information (Pyo, 2005).

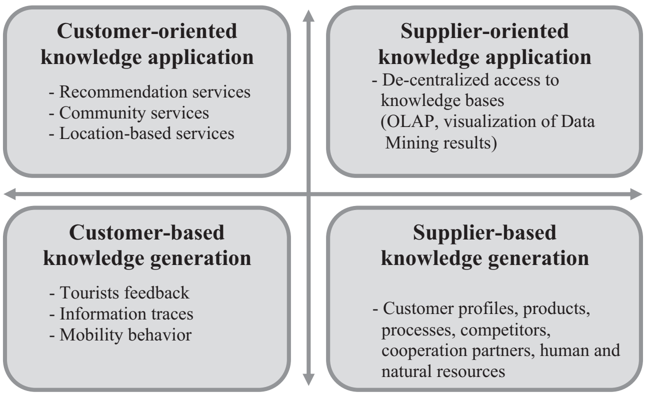


Fig. 1. The knowledge destination framework.
(*Source*: Adapted from Höpken et al. (2014)).

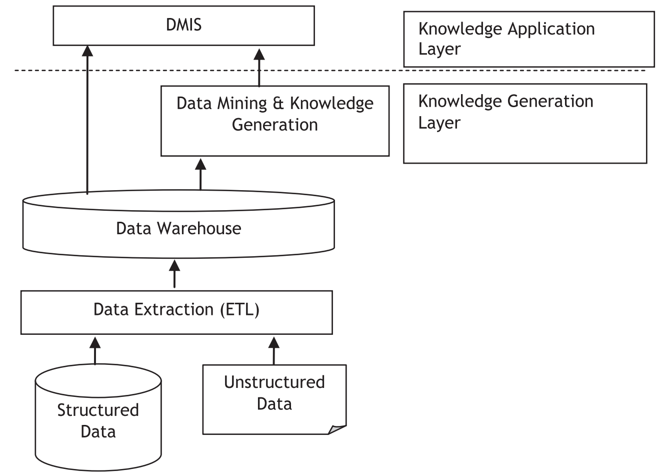


Fig. 2. The knowledge destination framework architecture.
(*Source*: Adapted from Höpken et al. (2011: 420)).

The *knowledge application layer* offers e-services that inform about supply elements and tourists' activities. For instance, on the customer side, intelligent location-based services adaptive to the user can guide tourists to the most attractive destination spots (Höpken, Fuchs, Zanker, & Beer, 2010; Rasinger, Fuchs, Höpken, & Beer, 2009). On the supply side, BI-based management information systems enable the decentralized generation of specific knowledge relevant to the destination management organization (DMO), and private/public destination stakeholders (Cho & Leung, 2002; Fuchs, Eybl, & Höpken, 2011; Höpken et al., 2011; Olmeda & Sheldon, 2002).

3. The knowledge destination framework architecture

Similar to Schianetz et al. (2007), the proposed architectural framework distinguishes between a knowledge creation and a knowledge application layer: the former comprises the various sources of *customer*-based data (e.g. web-search, booking, and feedback data, respectively), the technical components for data extraction, transformation and loading (ETL processes), a centralized data warehouse and data mining (Larose, 2005). The decentralized presentation and ad-hoc visualization of data mining models and underlying data rests on the knowledge application layer, the DMIS-Åre 'cockpit' (Höpken et al., 2011). Fig. 2 displays major components of the knowledge destination framework architecture which are described next (Fuchs et al., 2013; Höpken et al., 2011):

- *Data sources:* Customer-based data come either in the form of *explicit* tourists' feedback, provided knowingly and intentionally, like guest surveys and e-reviews, or in the form of *implicit* tourist's information traces, provided unknowingly and unintentionally, such as web-navigation data, online requests, booking and payment data, as well as GPS-based coverage of tourists' spatial movements. Data sources can be differentiated into *structured* data (e.g. transaction data, surveys, ratings) and *unstructured* data, composed by free text (e.g. e-reviews) and rich content from web 2.0 applications (e.g. YouTube.com) (Höpken et al., 2011).
- *Data extraction:* Different data sources require different techniques for the extraction, transformation and loading (ETL) of relevant information dependent on the *data format* at hand. Thus, the integration of heterogeneous data sources is typically done by extracting structured and semi-structured data (*html-documents*) by means of semantic, linguistic or constraint-based techniques of information integration. In contrast, unstructured data is extracted by wrappers or text mining (i.e. statistical language models, natural language processing approaches) (Höpken et al., 2011).
- *Data warehousing*: Data from different data sources are mapped into a homogeneous data format and stored in a central Data Warehouse that embraces all data relevant to destination stakeholders. Through a harmonization process it is possible to carry out a destination wide and all-stakeholder encompassing analysis approach. Thus, based on a tourism-ontology, individual data sources are transformed into a central data model and into a dimensional structure (Höpken et al., 2011).
- *Knowledge generation through data mining*: By employing methods of data mining interesting patterns and relationships in the data can be detected. However, only recently, data mining became important for tourism for its ability to discover unknown patterns in huge data bases, and, in contrast to most statistical methods, for its ability to also consider non-linear relationships in the analysed data (Fuchs et al., 2013; Fuchs & Höpken, 2009; Magnini et al., 2003; Olmeda & Sheldon, 2002). Further advantages of data mining compared to most statistical methods are the weak assumptions regarding data quality, as data can be incomplete, noisy, redundant, and dynamic (Larose, 2005). Although, the potential of data mining is not fully used in tourism yet, all major data mining techniques can be found in the literature (Höpken et al., 2014). For instance, descriptive/explorative analyses can be used in form of reports (OLAP) to visualize tourism arrivals per dimensions, such as time/season, travel type or customer origin (Fuchs & Weiermair, 2004; Wöber, 1998). Moreover, methods of *supervised learning*, like classification and estimation, are used to explain tourists' booking and cancellation behaviour (Morales & Wang, 2008) or to predict tourism demand (Chu, 2004; Law, 1998). As a method of *unsupervised learning*, clustering is typically applied to the task of customer segmentation or customer relationship management (Bloom, 2004). Finally, the topic of *web data mining* gained special attention in tourism where *web content mining* can analyse tourists' comments in blogs or review platforms especially in the form of opinion mining and sentiment detection (Gräbner et al., 2012; Kasper & Vela, 2011, 2012; Schmunk, Höpken, Fuchs, & Lexhagen, 2014). *Web usage mining* is dealing with the analysis of tourists' click- and search-behaviour when using tourism websites (Pitman et al., 2010). Finally, *web structure mining* discovers useful information from the hyperlink structure of a specific web domain or website,

such as communities of users or categories of similar websites, like content-rich authority sites or hubs (i.e. overview sites) (Baggio & Del Chiappa, 2014).

- Finally, the presentation of data mining models and the underlying data rest on the *knowledge application layer* (see Fig. 2): the *DMIS* '*cockpit*' acts as the human–computer interface and provides stakeholders decentralized access to instant analyses and ad-hoc visualizations of BI-based analysis results.

4. A business intelligence-based destination management information system

Designing and engineering a Business Intelligence-based destination management information system (DMIS) requires a profound understanding of the nature of knowledge behind management processes and an appropriate interpretation of the management objectives behind decision making at the level of tourism destinations (Hallin & Marnburg, 2008). According to the literature, knowledge relevant in a tourism destination context subsumes knowledge about market cultivation (e.g. how to attract valuable customers) and knowledge relevant for destination management, development, and planning (e.g. the development of new product–market combinations for valuable customer segments, training, private–public partnerships, etc.) (Bornhorst, Ritchie, & Sheehan, 2010; Pyo et al., 2002). Especially, *customer-based knowledge* is gained through customer segmentation techniques and service performance evaluation (Cho & Leung, 2002; Pyo, 2005; Ritchie & Ritchie, 2002). Thus, the knowledge sources considered in our presented BI application reflect tourists' search behaviour (i.e. Web navigation/search), tourists' booking behaviour, and tourists' feedback (i.e. feedback from all types of surveys and e-review platforms). Put differently, data collected, stored, analysed and visualized in DMIS-Åre include tourists' demographic and psychographic characteristics, buying motives and brand perceptions as well as customers' information usage and product consumption patterns, respectively (Fuchs et al., 2013; Höpken et al., 2011).

Since the effective use of a DMIS requires not only a sophisticated technology application, but particularly demands the establishment of organizational learning processes, it is crucial to integrate private and public stakeholders which define their specific and highly individual knowledge requirements. Thus, based on a literature review (Bornhorst et al., 2010; Chekalina & Fuchs, 2009; Chekalina, Fuchs, & Lexhagen, 2013; Dwyer & Kim, 2003; Gretzel & Fesenmaier, 2004; Pyo, 2005; Wang & Russo, 2007) and input from stakeholders of the Swedish mountain tourism destination, Åre, the following set of indicators has been defined (Fig. 3).

Economic performance indicators, comprise indicators, like bookings, overnights, prices, occupancy, sales; *Customer behaviour indicators* comprise measures for website navigation and search (e.g. page views, search terms), booking and consumption behaviour (e.g. booking channels, conversion rates, length of stay, cancellations, guest tracking), and customer profile (e.g. country of origin, age, gender, skiing travel behaviour, customer life time value, preferred type of accommodation and transportation, purpose of visit). Finally, *Customer perception and experience indicators* comprise indicators measuring perceived destination brand awareness (e.g. brand visibility, knowledge about the destination, information sources), destination value areas (e.g. skiing and non-skiing winter activities, summer activities and attractions, services and features, atmosphere, social interaction), value for money and customer satisfaction (e.g. functional and emotional value) and loyalty (i.e. cognitive, affective and conative loyalty) (Chekalina, Fuchs, & Lexhagen, 2013, 2014).

As already mentioned, through a business process oriented multidimensional data modelling approach these indicators are assigned to sequential destination processes, namely "*Web-Navigation*", "*Booking*" and "*Feedback*" (Höpken et al., 2013; Kimball, Ross, Thronthwaite, Mundy, & Becker, 2008). Each process is composed by the main variable(s) of analysis (i.e. *facts*) and their context (i.e. *dimensions*). By identifying 'common dimensions' across different business processes (i.e. *conformed dimensions*), this procedure allows DMIS to provide analyses *across* various processes, thus, to join so far disconnected and separately filed knowledge areas (Kimball et al., 2008), a condition which is considered as crucial for enhanced learning and creativity processes (Schianetz et al., 2007). Information extraction, transformation and loading (ETL) are based on the *Rapid Analytics Business*

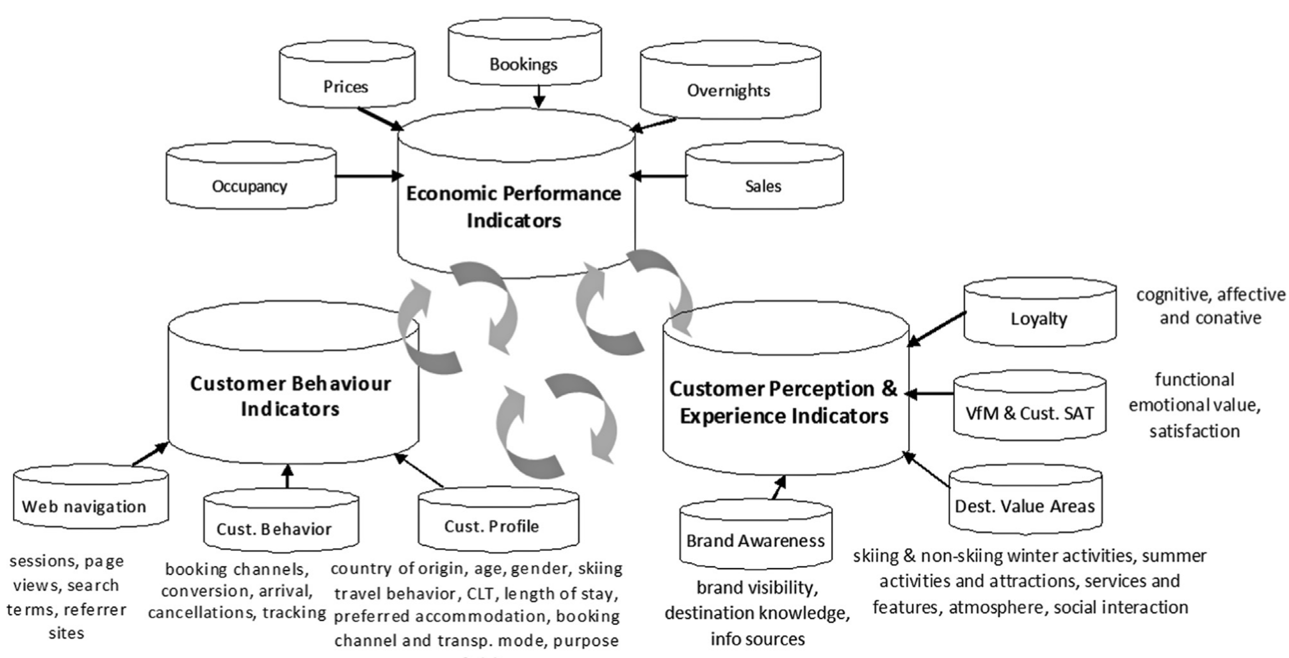


Fig. 3. Destination performance indicators by DMIS-Åre.

Intelligence server®, while the DMIS 'cockpit' is developed as a *html*-based web application (www.dmis-are.com). DMIS-Åre is technically fully validated, tested and implemented as a genuine novelty at the tourism destination Åre (Höpken et al., 2014). In its present form DMIS provides instant reports (*dashboards*) and OLAP analyses, thus, grants destination stakeholders real-time access to the data stored in the Data Warehouse.

Exemplarily for the business process "Web Navigation", Fig. 4 shows how destination managers and suppliers can generate new knowledge and trigger learning processes through the web-based DMIS cockpit. Customized dashboards and tables generate instant reports about customers' search and navigation behaviour on a supplier's website, comprising metrics, like the number of visits (sessions), unique visitors, page views, top 10 visited pages, and clicks/sessions grouped by date, and other pre-selected dimensions (see Fig. 4).

Fig. 5 shows the user dialog for executing OLAP analyses. OLAP enables users of DMIS-Åre to analyse one or more key indicator(s) (i.e. *fact*) on an aggregated level through grouping it by multiple and manually selectable grouping variables (i.e. *dimension attributes*), like customer profile and product type. Thus, the user selects the facts to be shown together with the appropriate aggregation function (e.g. *sum or average*), defines one or several dimension attributes the data should be grouped by and, finally, specifies constraints the data are filtered by, if necessary (Fuchs et al., 2013). The OLAP analysis in Fig. 5 again deals with customers' web navigation data: selected facts are the total amount of clicks and the average time between two clicks in seconds. Data is grouped by customer country, product area and day time. The example demonstrates the flexibility of the OLAP approach.

Like the Web Navigation process, also the Booking process pertains to the "pre-arrival phase" and includes data with respect to the booking stage of Åre guests. In the DMIS opening screen, a *cross business process analysis* is available: the relationship between web-sessions (Web Navigation process) and actual bookings (Booking process) over the variable time is graphically presented (Fig. 6). For the whole destination as well as for individual destination suppliers, the correlative patterns between

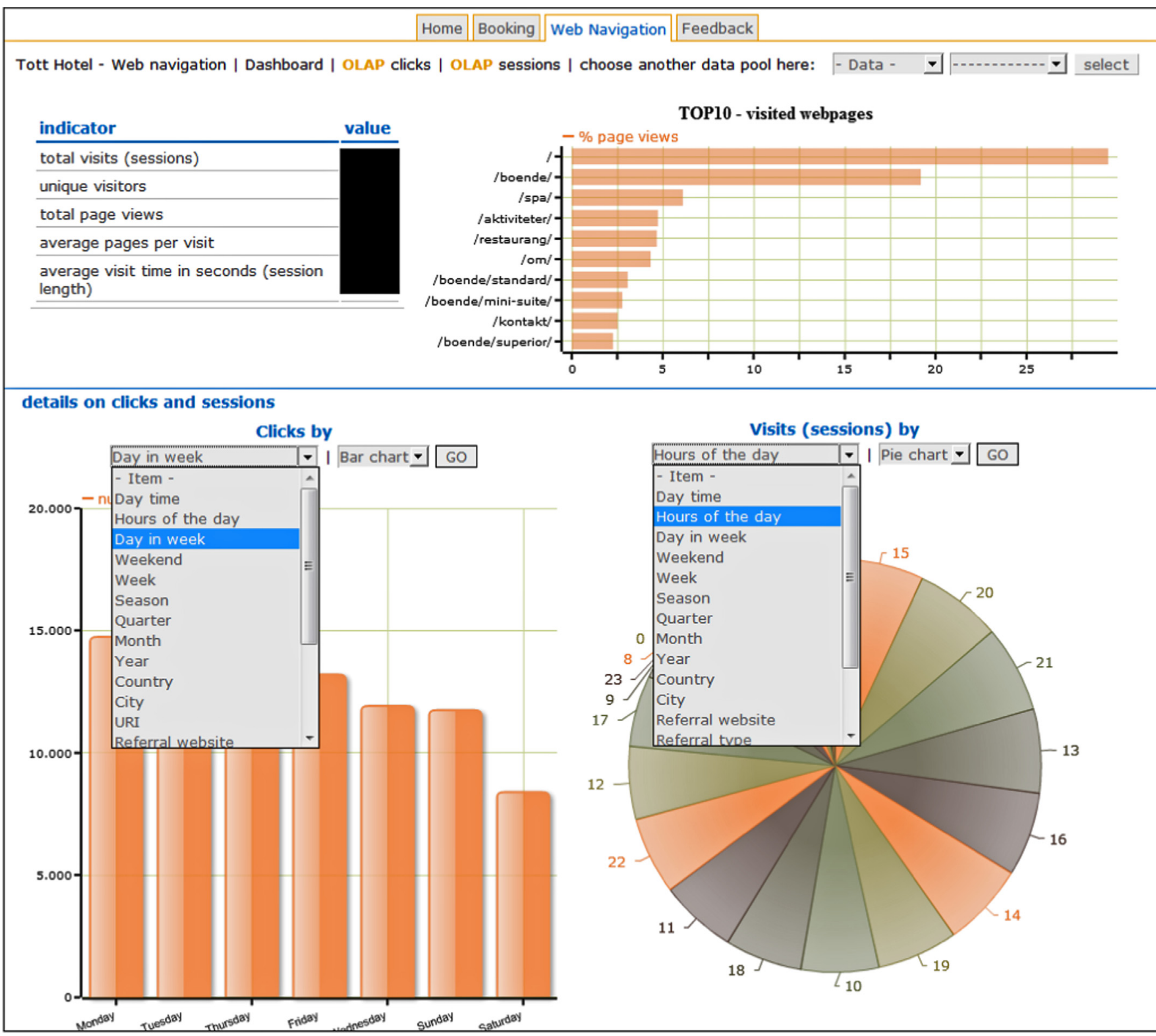
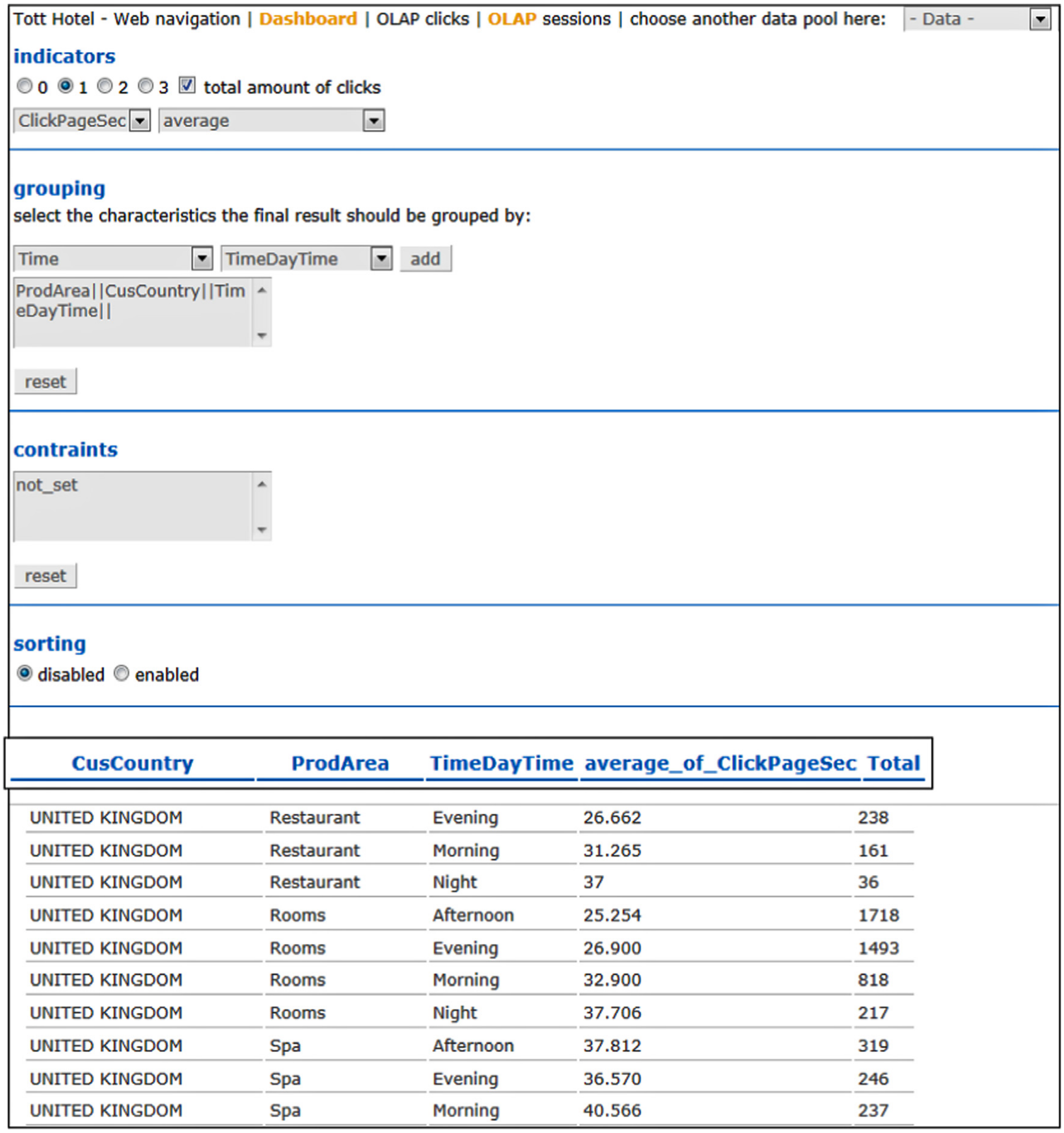


Fig. 4. DMIS-Åre dashboard: analysis of log-file data (process: Web Navigation).



CusCountry	ProdArea	TimeDayTime	average_of_ClickPageSec	Total
UNITED KINGDOM	Restaurant	Evening	26.662	238
UNITED KINGDOM	Restaurant	Morning	31.265	161
UNITED KINGDOM	Restaurant	Night	37	36
UNITED KINGDOM	Rooms	Afternoon	25.254	1718
UNITED KINGDOM	Rooms	Evening	26.900	1493
UNITED KINGDOM	Rooms	Morning	32.900	818
UNITED KINGDOM	Rooms	Night	37.706	217
UNITED KINGDOM	Spa	Afternoon	37.812	319
UNITED KINGDOM	Spa	Evening	36.570	246
UNITED KINGDOM	Spa	Morning	40.566	237

Fig. 5. DMIS-Åre OLAP: analysis of log-file data (process: Web Navigation).

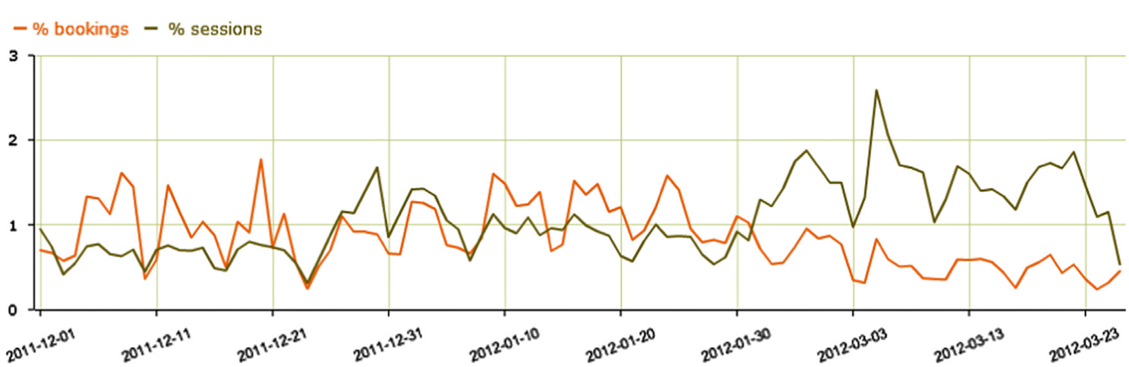
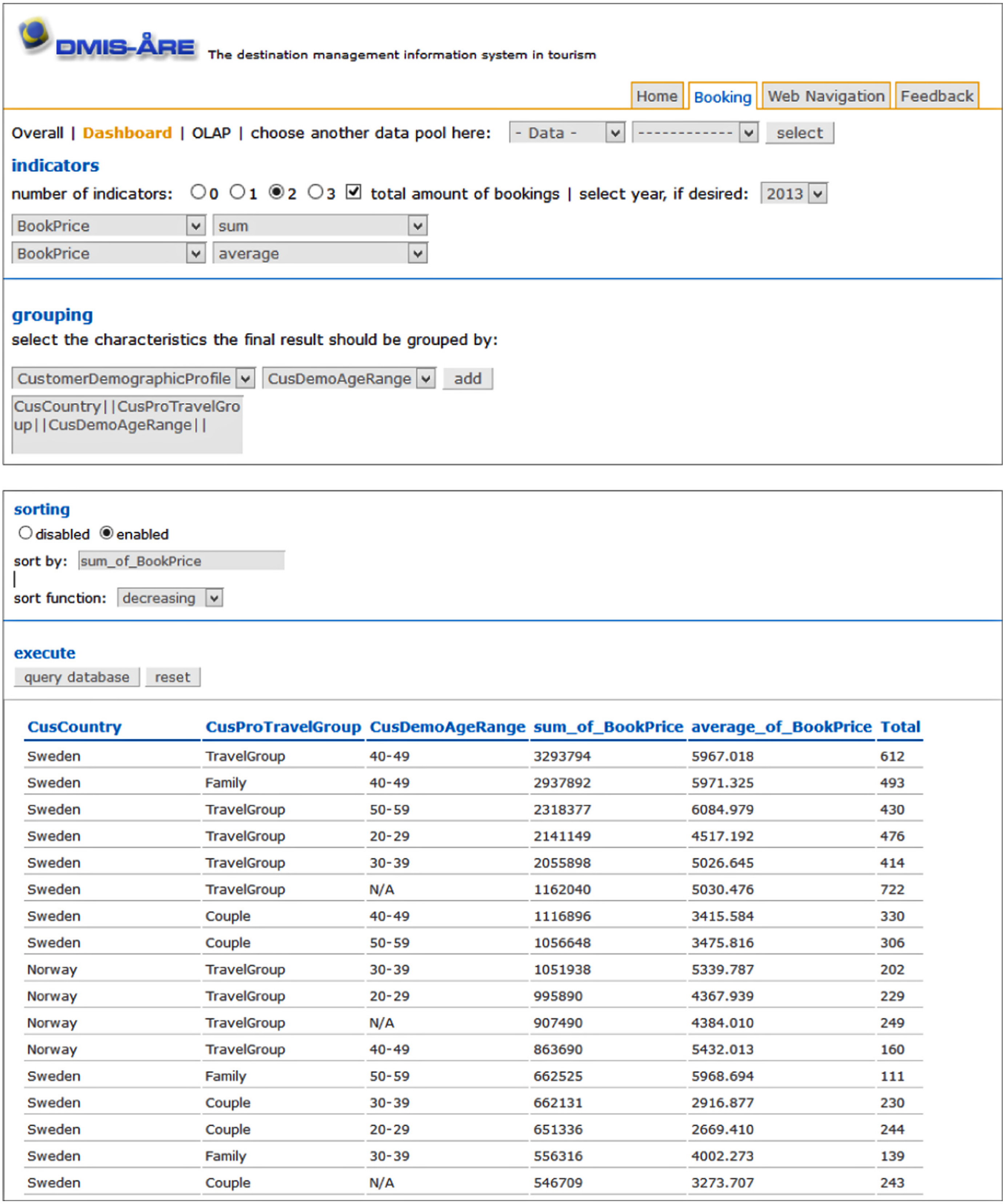


Fig. 6. DMIS-Åre across-business analysis (process: Booking). (Correction between Bookings and Sessions).

searching and booking can be recognized and is especially useful to forecast tourist arrivals from various sending countries.

The OLAP-based analysis displayed in Fig. 7 groups the sum and the average of booking prices from all destination suppliers, by customers' country, travel group, and age range. By using the "sorting function" for the fact "sum" (i.e. sales), the customer segment with the relatively strongest impact on sales generation can easily be identified – for the Åre case this is a Swedish travel group aged 40–49 and generates a sales portion amounting at 3,293,793 SEK (see Fig. 7).



CusCountry	CusProTravelGroup	CusDemoAgeRange	sum_of_BookPrice	average_of_BookPrice	Total
Sweden	TravelGroup	40-49	3293794	5967.018	612
Sweden	Family	40-49	2937892	5971.325	493
Sweden	TravelGroup	50-59	2318377	6084.979	430
Sweden	TravelGroup	20-29	2141149	4517.192	476
Sweden	TravelGroup	30-39	2055898	5026.645	414
Sweden	TravelGroup	N/A	1162040	5030.476	722
Sweden	Couple	40-49	1116896	3415.584	330
Sweden	Couple	50-59	1056648	3475.816	306
Norway	TravelGroup	30-39	1051938	5339.787	202
Norway	TravelGroup	20-29	995890	4367.939	229
Norway	TravelGroup	N/A	907490	4384.010	249
Norway	TravelGroup	40-49	863690	5432.013	160
Sweden	Family	50-59	662525	5968.694	111
Sweden	Couple	30-39	662131	2916.877	230
Sweden	Couple	20-29	651336	2669.410	244
Sweden	Family	30-39	556316	4002.273	139
Sweden	Couple	N/A	546709	3273.707	243

Fig. 7. DMIS-Åre OLAP: Booking price sum and average (process: Booking).

Finally, the business process "*Feedback*" refers to the tourists' *post-trip* phase. Currently, this process embraces the most comprehensive data input (Chekalina et al., 2014). More concretely, the feedback data comprise: (a) Åre Destination Brand Equity surveys for the winter season 2010/2011 and the summer season 2012 (Fuchs, Chekalina, & Lexhagen, 2011), (b) real-time feedback data from Åre guests during their stay provided by an electronic customer registration and survey tool (e-CRST) accessible via

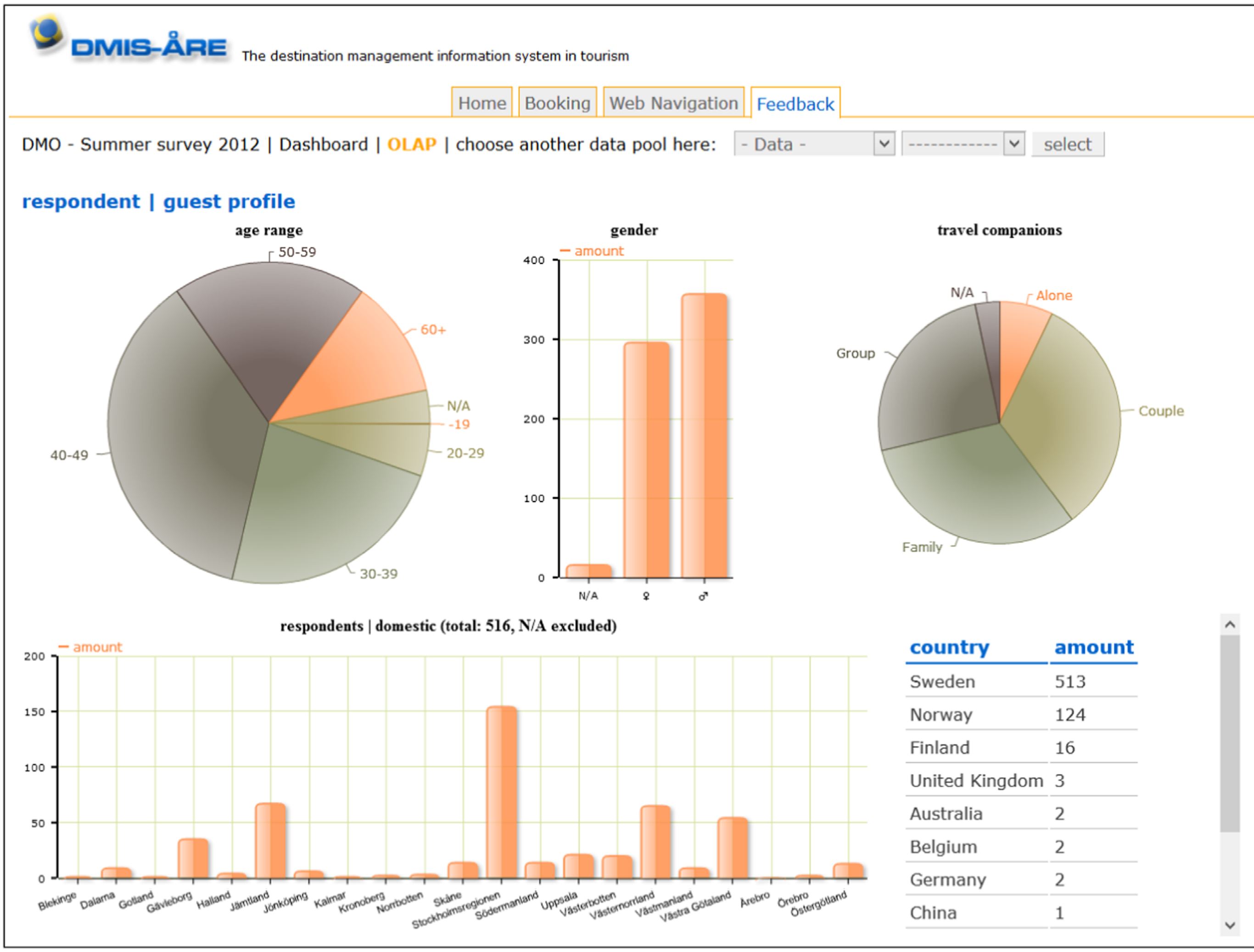


Fig. 8. DMIS-Åre dashboard: winter survey data (process: Feedback).

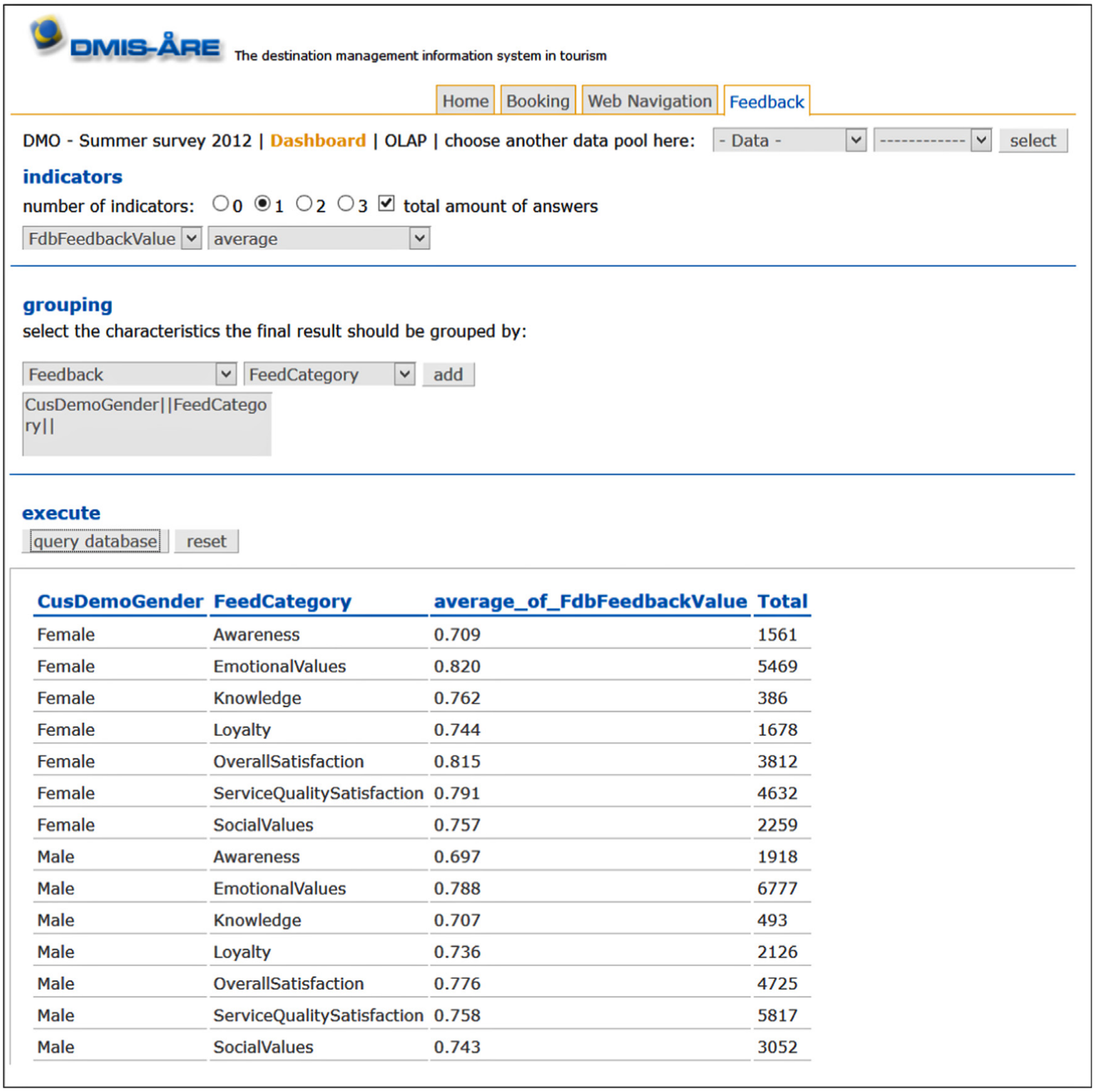
Quick Response Codes (Höpken et al., 2012), (c) user generated content (UGC) in the form of e-reviews automatically extracted from the major social media platforms *Booking.com* and *TripAdvisor.com* (Fuchs & Zanker, 2012; Gräbner et al., 2012; Schmunk et al., 2014), and finally, (d) customer feedback data based on individual surveys conducted by various destination stakeholders, such as the accommodation providers Copperhill Mountain Lodge and Tott Hotel Åre. The "Overall dashboard" aggregates feedback data from all data pools providing customer feedback at the destination, thus providing complex benchmarking functionalities.

For guest survey data, Fig. 8 shows how a destination supplier can apply knowledge and trigger learning processes through the web-based DMIS-'cockpit' and personally customized dashboards (Fuchs et al., 2013, p. 131).

The OLAP analysis in Fig. 9 again deals with customer feedback (i.e. survey) data. The selected fact is the feedback value aggregated as average normalized values (i.e. 0 = totally unsatisfied; 1 = totally satisfied). The data are grouped by the feedback category and gender.

As mentioned, also user generated content (UGC) in the form of ratings and e-reviews regarding the Swedish destination Åre extracted from social media platforms *TripAdvisor.com* and *Booking.com* are integrated into DMIS-Åre. First of all, text mining techniques are applied to semi-automatically extract single statements from each of the crawled reviews (Schmunk et al., 2014). Subsequently, by using machine-learning techniques, like Support Vector Machine, Naïve Bayes, and Nearest Neighbour (Hastie et al., 2009), as well as a dictionary-based approach (Liu, 2008), these statements are classified (1) as either "positive" or "negative" experiences, or "neutral" (i.e. "other comments or concerns") and (2) into the evaluated product area (i.e. "FoodBreakfast", "Rooms", "ServicePersonnel", "Wellness" or "Mixed"). From the proportion of classified positive and classified negative experiences (i.e. statements) an "average feedback value" is computed (i.e. a normalized value between 0 and 1). The dashboard in Fig. 10 displays the "average feedback value" grouped by (1) product area and project partners, (2) project partners alone, and (3) product areas and all major hotels in Åre

The OLAP analysis in Fig. 11 is again for customer feedback (i.e. UGC) data. The purpose of the analysis is to show the average feedback value of statements (i.e. 1 = positive experience; 0 = negative experience '?' = neutral) grouped by the specific statement text, accommodation suppliers, product areas and the social media platform where the feedback was given by Åre guests. Thus, the selected fact is the overall "feedback value" aggregated as average value. In addition, the "count function" is activated. Finally, the data is grouped by the dimensions (1) statement text-attribute "VarFeed", (2) accommodation-attribute "TraProLodgingSupplier", (3) product-attribute "ProdArea" and (4) social media platform-attribute: "ChaName". This final example again demonstrates the flexibility of the OLAP approach applied by DMIS-Åre (Fig. 11).



CusDemoGender	FeedCategory	average_of_FdbFeedbackValue	Total
Female	Awareness	0.709	1561
Female	EmotionalValues	0.820	5469
Female	Knowledge	0.762	386
Female	Loyalty	0.744	1678
Female	OverallSatisfaction	0.815	3812
Female	ServiceQualitySatisfaction	0.791	4632
Female	SocialValues	0.757	2259
Male	Awareness	0.697	1918
Male	EmotionalValues	0.788	6777
Male	Knowledge	0.707	493
Male	Loyalty	0.736	2126
Male	OverallSatisfaction	0.776	4725
Male	ServiceQualitySatisfaction	0.758	5817
Male	SocialValues	0.743	3052

Fig. 9. DMIS-Åre OLAP: summer survey data (process: Feedback).
(*Source*: Höpken et al. (2013)).

5. Conclusions and future research

To summarize, the presented research conducted in collaboration with core-stakeholders of the leading Swedish mountain tourism destination Åre addressed both the *generation of customer-based knowledge* within a tourism destination as well as the BI-based *supplier-oriented knowledge application* to support suppliers' decision making. The focus was on the *pre-trip* and *post-trip phases*. Accordingly, customer-based knowledge sources, like tourists' search (Web navigation), booking and feedback-behaviour (surveys, review platforms), allow the provision of an all-stakeholder encompassing *Destination Management Information System* (DMIS). All critical concepts, like the definition of industry's knowledge requirements, data extraction, data warehousing and user-interfaces, have been technically validated, tested and implemented as a genuine novelty at the tourism destination Åre. In more detail, the gained insights and research results comprise:

(a) *Business processes* (web navigation, booking, customer feedback), and *indicator sets* effectively measuring customer behaviour and perception defined on the basis of the literature and input from the industry (Bornhorst et al., 2010; Chekalina et al., 2014; Gretzel & Fesenmaier, 2004)
(b) The centralized *Destination Data Warehouse* (*DW*) built upon a uniquely defined data model that includes fact and dimension tables. Through *conformed* dimensions common to various business processes (e.g. customer profiles) this procedure enables DMIS to provide analyses *across* processes, thus, to join so far disconnected and separately filed knowledge areas (Kimball et al., 2008).
(c) The technical *DMIS architecture* based on the *Rapid Analytics BI* server-environment effectively supporting data extraction, transformation and loading (ETL) and data analysis (OLAP/Data Mining) (Höpken et al., 2013).
(d) The *DMIS*-'*cockpit*' implemented as *html*-based Web application (www.dmis-are.com) providing stakeholders with decentralized access to instant analyses and ad-hoc visualization of analysis results; and finally,
(e) A *business model* according to which the central DW is localized at the DMO, thereby offering destination stakeholders multiple benefits, like benchmarking and free access to valuable knowledge sources.

items	Copperhill Mountain Lodge	Diplomat Åregården Hotel	Fjällgården, Hotell	Holiday Club Hotell	Karolinen Hotell	STF Hotell Åre Torg	Tott Hotell	Åre Bed & Breakfast	Åre Continental Inn	AGO - Mitt i Åre	Granen & Residences, Hotel	Parkvi
Food/ Breakfast	0.467	0.704	1	0.545	0	0.500	0.667	1	1	?	?	?
Location	0.837	0.938	0.556	0.764	1	0.909	0.933	1	0.667	?	?	?
Mixed	0.569	0.680	0.667	0.731	1	0.625	0.895	1	0.571	1	?	1
Rooms	0.524	0.571	0	0.404	0.333	0.333	0.227	?	0.200	?	?	?
Service/ Personnel	0.522	0.750	0	0.600	?	0.500	0.750	1	1	?	?	?
Wellness	0.950	0.750	?	0.917	?	?	0.857	?	1	?	?	?

Fig. 10. DMIS-Åre Dashboard: UGC-based satisfaction score by hotels and product areas (process: Feedback).

In its present version, the DMIS prototype comprises web-search, booking and feedback data (e.g. survey-based, user-generated content) from the Destination Management Organization, Åre Destination AB, and the major destination operator, Ski Star Åre, operating cable cars and ski-lifts, but also offering accommodation and ski rentals. However, also small- and medium-sized accommodation suppliers, like Tott Hotel Åre and Copperhill Mountain Lodge Åre, are constantly providing their customer-based data to DMIS through a semi-automated process of extracting, loading and transforming data into the homogenous and centralized destination Data Warehouse (Fuchs et al., 2013; Höpken et al., 2013). Privacy issues are especially secured through a responsible data handling process: technically, sensitive customer data is stored to a minimal extent and access to such data is handled as restrictively as possible. Following this trust keeping mechanisms, each stakeholder can visualize only analysis results regarding its own data compared to aggregated, thus, fully anonymized data.

Although, in its present version, the DMIS prototype mainly considers customer-based data, it is planned in the course of future research to also integrate supplier-based data sources from the entire digital eco-system of the destination (Baggio & Del Chiappa, 2014). These data sources, primarily, include information on products, processes and collaboration partners extracted from sources (i.e. web-sites) in the form of product profiles and availability information (e.g. booking engines). Thus, valuable

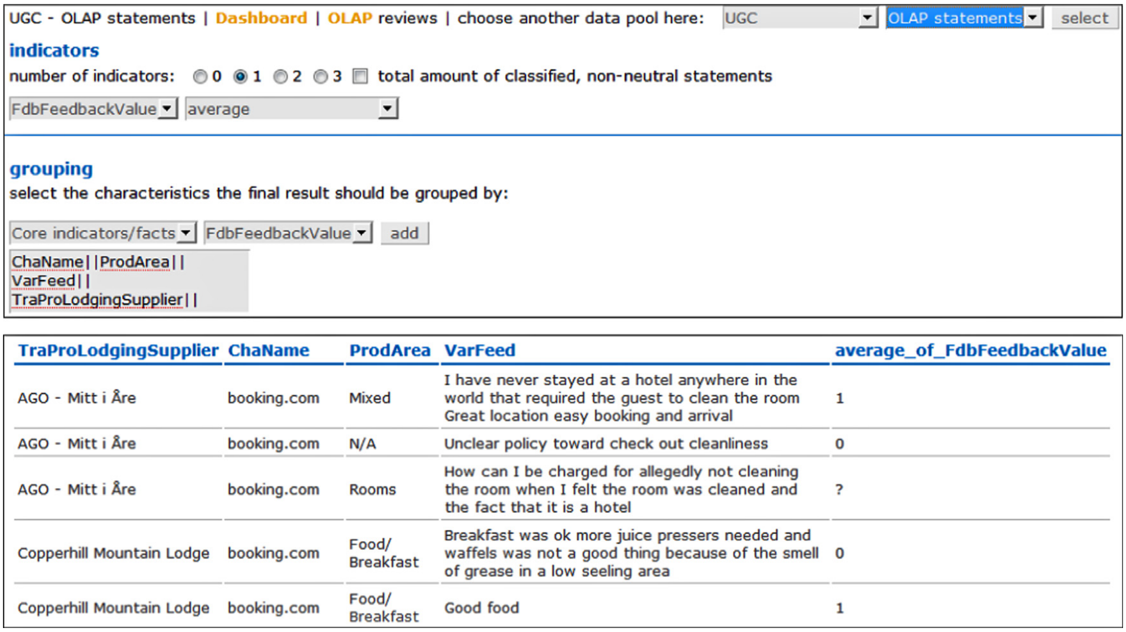


Fig. 11. DMIS-Åre OLAP – UGC Analyser (process: Feedback).

knowledge about suppliers' service potential (e.g. property status), the complementarity of destination offers (on the basis of market basket analyses), and their evaluation through tourists' feedback will be gained. Moreover, in the near future, the DMIS-'cockpit' will also provide data mining processes, like classification, clustering, or prediction executed by the *RapidMiner*® data mining software (Fuchs et al., 2013; Höpken et al., 2011).

A final future research goal comprises the application of real-time Business Intelligence (Gravic, 2013; Savvy Data, 2013) in order to gain real-time knowledge on tourists' on-site behaviour (Pike, Murdy, and Lings, 2011). Respective customer data can be gathered through QR Code-based electronic customer cards collecting, tourists' (GPS/WLAN-based) position and ad-hoc feedback (Höpken et al., 2012; Zanker, Jessenitschnig, & Fuchs, 2010). This valuable new knowledge serves as input for intelligent mobile (i.e. ubiquitous) end-user applications, capable of recommending tourists most promising matches with the actual destination offer, thus, enhancing tourists' quality of experience (Jannach, Zanker & Fuchs, 2014; Rasinger et al., 2009; Wang et al., 2012). Finally, at the supply side, this newly generated knowledge input may be applied by small and medium-sized destination suppliers to react on segment specific needs in real-time.

Acknowledgements

This research was financed by KK-Foundation project '*Engineering the Knowledge Destination*' (no. 20100260; Stockholm, Sweden). The authors would like to thank the managers Lars-Börje Eriksson (Åre Destination AB), Niclas Sjögren-Berg and Anna Wersén (Ski Star Åre), Peter Nilsson and Hans Ericsson (Tott Hotel Åre), and Pernilla Gravenfors (Copperhill Mountain Lodge Åre) for their excellent cooperation.

References

Back, A., Enkel, E., & V. Krogh, G. (2007). *Knowledge networks for business*. New York: Springer.
Baggio, R., & Del Chiappa, G. (2014). Real and virtual relationships in tourism digital ecosystems. *Journal of Information Technology & Tourism*, *14*(1), 3–19.
Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(*1*), 99–120.
Barney, J. B. (1997). *Gaining and sustaining competitive advantage*. New York: Addison-Wesley.
Bloom, J. Z. (2004). Tourist market segmentation with linear and non-linear techniques. *Tourism Management*, *25*(6), 723–733.
Bornhorst, T., Ritchie, J. R., & Sheehan, L. (2010). Determinants for DMO & destination success: An empirical examination. *Tourism Management*, *31*, 572–589.
Buckley, R. (2012). Sustainable tourism: Research and reality. *Annals of Tourism Research*, *39*(*2*), 528–546.
Buhalis, D. (2006). The impact of ICT on tourism competition. In: A. Paptheodorou (Ed.), *Corporate rivalry and market power: Competition issues in the tourism industry* (pp. 143–171). London: IB Tauris.
Burman, C. (2002). Wissensmanagement als Determinante des Unternehmenswerts. *Zeitschrift für Führung und Organisation*, *71*(*6*), 334–341.
Chekalina, T., & Fuchs, M. (2009). An assessment of survey-based brand metrics for destinations. In: L. Dioko, & L. Xiang (Eds.), *3rd International conference on destination branding & marketing, Macau* (pp. 130–141).
Chekalina, T., Fuchs, M., & Lexhagen, M. (2013). Determinants of the co-created destination experience. In: M. Kozak, L. Andreu, J. Gnoth, S. Lebe, & A. Fyall (Eds.), *Tourism marketing: On both sides of the counter* (pp. 57–79). Newcastle upon Tyne, UK: Cambridge Publishing.
Chekalina, T., Fuchs, M., & Lexhagen, M. (2014). A-value creation perspective on the customer-based brand equity model for tourism destinations. *Finnish Journal of Tourism Research*, *10*(1), 7–23.
Cho, V., & Leung, P. (2002). Knowledge discovery techniques in database marketing for the tourism industry. *Journal of Quality Assurance in Hospitality & Tourism*, *3*(*3*), 109–131.
Chu, F. L. (2004). Forecasting tourism demand: A cubic polynomial approach. *Tourism Management*, *25*(2), 209–218.
Coles, T., Hall, C. M., & Duval, D. T. (2006). Tourism and post-disciplinary enquiry. *Current Issues in Tourism*, *9*(*1*), 4–5.
Dwyer, L., & Kim, C. (2003). Destination competitiveness: Determinants and indicators. *Current Issues in Tourism Research*, *6*(*5*), 369–417.
Fuchs, M., Abadzhiev, A., Svensson, B., Höpken, W., & Lexhagen, M. (2013). A knowledge destination framework for tourism sustainability – A business intelligence application from Sweden. *Tourism – An Interdisciplinary Journal*, *61*(2), 121–148.
Fuchs, M., Chekalina, T., & Lexhagen, M. (2011). Destination brand equity modeling and measurement – A case from Sweden. In: R. H. Tsiotsou, & R. E. Goldsmith (Eds.), *Strategic marketing in tourism services* (pp. 95–116). Bingley, West Yorkshire, UK: Emerald.
Fuchs, M., Eybl, A., & Höpken, W. (2011). Successfully selling accommodation packages at online auctions – The case of eBay Austria. *Tourism Management*, *32*(5), 1166–1175.
Fuchs, M., & Höpken, W. (2009). Data Mining in Tourism (In German: „Data Mining im Tourismus"). *Praxis der Wirtschaftsinformatik*, *270*(*12*), 73–81.
Fuchs, M., & Höpken, W. (2011). E-business horizons in the tourism industry – Challenges for research and practice. In: K. Sidali, A. Spiller, & B. Schulze (Eds.), *Food, agriculture and tourisms – Linking local gastronomy and rural tourism* (pp. 140–160). Berlin: Springer.
Fuchs, M., Höpken, W., Föger, A., & Kunz, M. (2010). E-business readiness, intensity, and impact – An Austrian DMO study. *Journal of Travel Research*, *49*(2), 165–178.

Fuchs, M., Höpken, W., & Lexhagen, M. (2015). Applying business intelligence for knowledge generation in tourism destinations – A case from Sweden. In: Pechlaner H., & Smeral E. (Eds.), *Tourism and leisure – Current issues and perspectives of development in research and business* (pp. 161–174). Wiesbaden: Springer.

Fuchs, M., Scholochow, C. h., & Höpken, W. (2010). E-business adoption, use and value creation – An Austrian hotel study. *Journal of Information Technology and Tourism*, *11*(4), 267–284.

Fuchs, M., & Weiermair, K. (2004). Destination benchmarking – An indicator system's potential for exploring guest satisfaction. *Journal of Travel Research*, *42*(*3*), 212–225.

Fuchs, M., & Zanker, M. (2012). Multi-criteria ratings for recommender systems: An empirical analysis in the tourism domain. In: C. Huemer, & P. Lop (Eds.), *E-commerce and web technologies. Lecture notes in business information processing*, 123 (pp. 100–111). Heidelberg, London: Springer.

Gräbner, D., Zanker, M., Fliedl, G., & Fuchs, M. (2012). Classification of customer reviews based on sentiment analysis. In: M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and communication technologies in tourism* (pp. 460–470). Wien, New York: Springer.

Grant, R. M. (1996). Toward a knowledge-based view of the firm. *Strategic Management Journal*, *17*(*2*), 109–122.

Gravic (2013). *Shadowbase for real-time business intelligence*. ⟨www.gravic.com/shadowbase⟩; retrieved March 2013.

Gretzel, U., & Fesenmaier, D. R. (2004). Implementing a knowledge-based tourism marketing information system: The Illinois tourism network. *Journal of Information Technology* & *Tourism*, *6*(*2*), 245–255.

Hallin, C. A., & Marnburg, E. (2008). Knowledge management in the hospitality industry: A review of empirical research. *Tourism Management*, *29*(*2*), 366–381.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning – Data mining, inference and prediction* (2nd ed). New York: Springer.

Höpken, W., Deubele, Ph., Höll, G., Kuppe, J., Schorpp, D., Licones, R., et al. (2012). Digitalizing loyalty cards in tourism. In: M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and communication technologies in tourism* (pp. 272–283). New York: SpringerIn: M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and communication technologies in tourism* (pp. 272–283). New York: Springer.

Höpken, W., Fuchs, M., Höll, G., Keil, D., & Lexhagen, M. (2013). Multi-dimensional data modeling for a tourism destination data warehouse. In: L. Cantoni, & Ph. Xiang (Eds.), *Information and communication technologies in tourism 2013* (pp. 157–169). New York: Springer.

Höpken, W., Fuchs, M., Keil, D., & Lexhagen, M. (2011). The knowledge destination – A customer information-based destination management information system. In: R. Law, M. Fuchs, & F. Ricci (Eds.), *Information and communication technologies in tourism 2011* (pp. 417–429). New York: Springer.

Höpken, W., Fuchs, M., & Lexhagen, M. (2014). The knowledge destination – Applying methods of business intelligence to tourism. In: J. Wang (Ed.), *Encyclopedia of business analytics and optimization* (pp. 307–321). Pennsylvania: IGI Global Publisher.

Höpken, W., Fuchs, M., Zanker, M., & Beer, Th. (2010). Context-based adaptation of mobile applications in tourism. *Journal of Information Technology and Tourism*, *12*(2), 175–195.

Jafari, J. (2001). The scientification of tourism. In: V. Smith, & M. Brent (Eds.), *Hosts and guests revisited: Tourism issues of the 21st century* (pp. 28–41). New York: Cognizant Communication Corporation.

Jannach, D., Zanker, M., & Fuchs, M. (2014). Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. *Journal of Information Technology and Tourism*, *14*(2), 119–149.

Kasper, W., & Vela, M. (2011). Sentiment analysis for hotel reviews. In *Computational linguistics–Applications conference*, Katowice, Poland, 10,2011 (pp. 45–52).

Kasper, W., & Vela, M. (2012). Monitoring and summarization of hotel reviews. In: M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and communication technologies in tourism 2012* (pp. 471–482). New York: Springer.

Kimball, R., Ross, M., Thronthwaite, W., Mundy, J., & Becker, B. (2008). *The data warehouse lifecycle toolkit* (2nd ed). Indianapolis: Wiley & Sons.

Larose, D. T. (2005). *Discovering knowledge in data*. New Jersey: John Wiley & Sons.

Law, R. (1998). Room occupancy rate forecasting – A neural network approach. *International Journal of Contemporary Hospitality Management*, *10*(6), 234–239.

Liu, B. (2008). *Web data mining* (2nd ed.). New York: Springer.

Magnini, V., Honeycutt, E., Jr., & Hodge, S. (2003). Data mining for hotel firms: Use and limitations. *Cornell Hotel and Restaurant Administration Quarterly*, *44* (December), 94–105.

Min, H., Min, H., & Emam, A. (2002). A data mining approach to develop the profile of hotel customers. *International Journal of Contemporary Hospitality Management*, *14*(*6*), 274–285.

Morales, D. R., & Wang, J. (2008). Passenger name record data mining based cancellation forecasting for revenue management. *Innovative Applications of Operations Research*, *202*(2), 554–562.

Olmeda, I., & Sheldon, P. J. (2002). Data mining techniques and applications for tourism internet marketing. *Journal of Travel* & *Tourism Marketing*, *11*(*2–3*), 1–20.

Palmer, A., Montano, J. J., & Sesé, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism Management*, *27*(4), 781–790.

Pike, S., Murdy, S., & Lings, I. (2011). Visitor relationship orientation of destination marketing organizations. *Journal of Travel Research*, *50*, 443–453.

Pitman, A., Zanker, M., Fuchs, M., & Lexhagen, M. (2010). Web usage mining in tourism – A query term analysis and clustering approach. In: U. Gretzel, R. Law, & M. Fuchs (Eds.), *Information and communication technologies in tourism 2010* (pp. 393–403). New York: Springer.

Pyo, S. (2005). Knowledge map for tourist destinations. *Tourism Management*, *26*(*4*), 583–594.

Pyo, S., Uysal, m., & Chang, H. (2002). Knowledge discovery in databases for tourist destinations. *Journal of Travel Research*, *40*(4), 396–403.

Rasinger, J., Fuchs, M., Höpken, W., & Beer, Th. (2009). Building a mobile tourist guide based on tourists' on-site information needs. *Tourism Analysis*, *14*(4), 483–502.

Ritchie, R. J. B., & Ritchie, J. R. B. (2002). A framework for an industry supported destination marketing information system. *Tourism Management*, *23*(*2*), 439–454.

Savvy Data (2013). *Data Collection is Key to Customer Understanding and Personalization in Tourism*. ⟨www.eyefortravel.com/social-media-and-marketing/savvy-data-collection-key-customer-understanding-and-personalisation⟩; Retrieved January 2014.

Schianetz, K., Kavanagh, L., & Lockington, D. (2007). The learning tourism destination: The potential of a learning organization approach for improving the sustainability of tourism destinations. *Tourism Management*, *28*(3), 1485–1496.

Schmunk, S., Höpken, W., Fuchs, M., & Lexhagen, M. (2014). Sentiment analysis – Implementation and evaluation of methods for sentiment analysis with RapidMiner®. In: Ph. Xiang, & I. Tussyadiah (Eds.), *Information and communication technologies in tourism* (pp. 253–265). New York: Springer.

Shaw, G., & Williams, A. (2009). Knowledge transfer and management in tourism organisations. *Tourism Management*, *30*(*3*), 325–335.

Sidali, K. L., Fuchs, M., & Spiller, A. (2012). The effect of electronic reviews on consumer behaviour. In: M. Sigala, U. Gretzel, & R. Vangelis (Eds.), *Web 2.0 in travel, tourism and hospitality* (pp. 239–253). Surrey: Ashgate.

Tajeddini, K. (2010). Effect of customer and entrepreneurial orientation on innovativeness: Evidence from the hotel industry in Switzerland. *Tourism Management*, *31*, 221–231.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(*7*), 509–533.

Wang, D., Park, S., & Fesenmaier, D. R. (2012). The role of smartphones in mediating touristic experience. *Journal of Travel Research*, *51*(4), 371–387.

Wang, Y., & Russo, S. M. (2007). Conceptualizing and evaluating the functions of destination marketing systems. *Journal of Vacation Marketing*, *13*(*3*), 187–203.

Wöber, K. W. (1998). Global statistical sources – TourMIS: An adaptive distributed marketing information system for strategic decision support in national, regional or city tourist offices. *Pacific Tourism Review*, *2*(3), 273–286.

Wong, J.-Y., Chen, H.-J., Chung, P.-H., & Kao, N.-C. (2006). Identifying valuable travellers and their next foreign destination by the application of data mining techniques. *Asia Pacific Journal of Tourism Research*, *11*(4), 355–373.

Zanker, M., Jessenitschnig, M., & Fuchs, M. (2010). Automated semantic annotation of tourism resources based on geo-spatial data. *Journal of Information Technology and Tourism*, *11*(4), 341–354.