Business intelligence and big data in hospitality and tourism: a systematic literature review

Article

Accepted Version

Tables

Table 1: Major components and functional subdomains within the Business Intelligence umbrella

Major Concept	Acronym	Short Explanation	Reference
Decision Support System	DSS	Computer-based information system that supports decision making resulting in ranking, sorting, or choosing from among alternatives	Burstein and Holsapple, 2008; Sauter, 2011
Data Warehousing	DW	Central data repository system of integrated data from one or multiple sources that stores current and historical data in one single place and format	Kimball et al., 2008
Online analytical processing	OLAP	Provides multi-dimensional analytical queries encompassing data warehousing and reporting. Supports the operations of consolidation, drill-down, slicing and dicing	Kimball et al., 2008
Data Mining	DM	Discovers correlations and patterns in (usually large) data sets involving methods of machine learning, statistics and mathematical modelling	Larose, 2005; Rud, 2009
Business Intelligence	BI	Umbrella term comprising the domains of DW, OLAP and DM	Kimball and Ross, 2016
Descriptive Analytics	-	Uses data aggregation (e.g. sums, averages, percent, changes, etc.) and data mining to provide insight into the past to answer: “What has happened”?	Williams, 2016
Predictive Analytics	-	Uses statistical and DM models to forecast the future and answer: “What could happen?”	Dedić and Stanier, 2016
Prescriptive Analytics	-	Uses machine learning and computational modelling to advice on optimal outcomes and answers: “What should we do?”	Dedić and Stanier, 2016; Williams, 2016
Big Data	BD	Data sets that are so large or complex that traditional data processing application software is inadequate to deal with them. Includes challenges, as data extraction, storage, analytics, visualization, querying, updating and information privacy	Erl et al., 2015

Table 2. Business Intelligence works in Hospitality and Tourism (selected works)

Article (author & title)	Research topic	Type of paper (conceptual/ empirical)	Source(s) of data	Type of data & size	Data collection methods	Data analysis techniques	Data reporting and visualization
Amadio and Procaccino (2016). Competitive analysis of online reviews using exploratory text mining	Text-based online review analysis using exploratory text mining techniques and visual analytics for SWOT analysis, applied to the hotel industry	Empirical	Online reviews from TripAdvisor	Unstructured	Manually (one-time)	Latent Dirichlet Allocation (LDA) topic model, random forest classification	Dashboard, SWOT analysis
Arbelaitz et al. (2013). Web usage and content mining to extract knowledge for modelling the users of the Bidasoa Turismo website and to adapt it	Combined web usage and content mining to generate user navigation profiles and semantically enriched user interest profiles as input to website optimization and marketing	Empirical	Web page content and web server log files of Bidasoa Turismo website	Structured and unstructured	Automatically	PAM (Partitioning Around Medoids) clustering with Edit Distance sequence alignment method, SPADE (Sequential Pattern Discovery using Equivalence classes), Latent Dirichlet Allocation (LDA) topic model	N/A
Ashiabor et al. (2007). Logit models for forecasting nationwide intercity travel demand in the United States	Nested mixed logit models to estimate market share of automobile and commercial air transportation	Empirical	American Travel Survey	Structured	One-time collection of existing secondary data	Nested and mixed logit models	Market share plots for 5 income groups

Carrasco et al. (2013). A multidimensional data model using the fuzzy model based on the semantic translation	Fuzzy Model based multidimensional data model to solve Opinion Aggregation when integrating heterogeneous information (including unstructured data)	Empirical	Web pages for data extraction are Atrapalo, Booking, eDreams, Expedia, TripAdvisor	Structured and unstructured	Automatically (periodically)	Explorative Data Analysis (EDA)	Dashboard, On-Line Analytical Processing (OLAP)
Chen and Tsai (2016). Data mining framework based on rough set theory to improve location selection decisions: A case study of a restaurant chain	Data mining framework based on rough set theory (RST) to support location selection decisions	Empirical	Survey of 33 directly-managed stores of a restaurant chain	Structured	Manually	Rough set theory (RST)	N/A
Chiang (2015). Applying data mining with a new model on customer relationship management systems: A case of airline industry in Taiwan	Mining high-value family travelers for CRM systems of online airlines and travel agencies to identify decision rules for discovering market segments	Empirical	Customer survey	Structured	Manually (one-time)	RFM model, Analytic hierarchy process (AHP), C5.0 decision tree	N/A
Dursun and Caber (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis	Profiling hotel customers by recency, frequency and monetary (RFM) indicators	Empirical	CRM system	Structured	Manually (one-time)	RFM model, self organizing map (SOM), k-means clustering	Self organizing map (SOM)
Fuchs et al. (2013). A knowledge destination framework for tourism sustainability: A business intelligence application from Sweden	A Destination Management Information system focusing on Online-Analytical Processing (OLAP) to measure proportion of tourists with smallest ecological footprint	Empirical	Customer feedback data (survey-based)	Structured and unstructured	Manually (one-time) and automatically (periodically)	Explorative Data Analysis (EDA)	Html-based web application, dashboards, On-Line Analytical Processing (OLAP)
Fuchs et al. (2014). Big data analytics for knowledge generation in tourism	BI-based knowledge infrastructure implemented at the Swedish mountain	Empirical	Web search, booking and feedback data (e.g., survey-based,	Structured and unstructured	Manually (one-time) and	Explorative Data Analysis (EDA) and machine	Html-based web application, dashboards, On-Line Analytical

destinations - A case from Sweden	destination, Åre and examples of use by tourism managers		user-generated content)		automatically (periodically)	learning (Support Vector Machine, Naïve Bayes, Nearest Neighbor)	Processing (OLAP)
Holland et al. (2016). The role and impact of comparison websites on the consumer search process in the US and German airline markets	Examines how consumers search for airline tickets based on a comparative analysis of the US and German markets	Empirical	Click-stream panel data by ComScore	Structured	Manually (one-time)	Consideration set theory	N/A
Höpken et al. (2015). Business intelligence for cross-process knowledge extraction at tourism destinations	A novel approach for BI-based cross-process knowledge extraction and decision support for tourism destinations	Empirical and conceptual	Web search, booking and feedback data (e.g., survey-based, user-generated content)	Structured and unstructured	manually (one-time) and automatically (periodically)	Explorative Data Analysis (EDA) and data mining techniques (Decision Trees, Association Rule Mining); Multi-dimensional data warehouse modelling	Html-based web application, dashboards, On-Line Analytical Processing (OLAP)
Hsieh (2011). Employing a recommendation expert system based on mental accounting and artificial neural networks into mining business intelligence for study abroad's P/S recommendations*	A recommendation Expert System for travel agencies based on mental accounting and artificial neural networks	Empirical	Online (student) survey about travel motivations and final decision making	Structured	Manually (one-time)	Back propagation neural networks	N/A
Hsieh (2009). Applying an expert system into constructing customer's value expansion and prediction model based on AI techniques in leisure industry	An Expert System platform addresses customer's value analysis based on artificial intelligence	Empirical	Online customer survey	Structured	Automatically (periodically)	Self-organizing feature map neural network for cluster analysis	N/A
Kisilevich et al. (2013). A GIS-based decision support system for hotel room rate estimation and temporal price prediction: The hotel brokers' context	A tool that assists travel intermediaries to acquire missing strategic information about hotels to leverage profitable deals. The GIS-based DSS	Empirical	OpenStreetMap data (public), Private data by a hotel brokerage static: names of hotels, internal IDs, location coordinates,	Structured	Automatically (periodically) and manually (one-time)	Multi-dimensional scaling (MDS); Voronoi tessellation partitioning;	MDS component for exploratory data Analysis and

	estimates room rates using hotel and location characteristics		hotel facilities, room amenities, hotel categories. Dynamic: room prices for one night customers received during their search, date of search, date of order			additive regression with isotonic regression; Locally Weighted Learning with Linear Regression; LibSVM nu-SVR; Multilayer Perceptron (ANN)	Graphs to visualize price estimation results
Köseoglu et al. (2016). Competitive intelligence practices in hotels	Assessment of awareness and knowledge about competitive intelligence efforts in the hotel industry	Empirical	23 hoteliers' knowledge and awareness about competitive intelligence	Unstructured	In-depth interview	N/A	N/A
Kwok and Yu (2016). Taxonomy of Facebook messages in business-to-consumer communications: What really works?	Combines machine learning and human intelligence to analyze Facebook messages initiated by hospitality companies	Empirical	2,654 Facebook messages initiated by 26 hospitality companies	Unstructured and structured	Automatically/ manually	Machine learning (support vector machines)	Taxonomy of Facebook message types
Li et al. (2015). Identifying emerging hotel preferences using Emerging Pattern Mining technique	Identification of emergent hotel features by extracting frequent keywords from online reviews	Empirical	118,000 online reviews from TripAdvisor	Unstructured	Automatically (one-time)	Unsupervised feature extraction by frequent keywords, emerging pattern mining (EPM)	N/A
Lu and Zhang (2015). Imputing trip purposes for long-distance travel	Machine learning methods estimate trip purposes for long-distance passenger travel	Empirical	A passively collected long-distance trip dataset is simulated from the 1995 American Travel Survey	Structured	Manually (one-time)	Decision tree and meta-learning	Confusion matrices from trip purpose imputation models

Marine-Roig and Anton Clavé (2015). Tourism analytics with massive user-generated content: A case study of Barcelona	Studying online image of Barcelona as transmitted via social media through the analysis of more than 100,000 relevant travel blogs and online travel reviews	Empirical	Heterogeneous including the travel blogs, webpages, travelogues and travel reviews about Barcelona Heterogeneous including the travel blogs, webpages, travelogues and travel reviews about Barcelona (250,000 pages)	Unstructured	Data was extracted automatically through Offline Explorer Enterprise	Pre-processing: web content mining, language detection, user's hometown, cleaning, debugging. Processing: parser settings and categorizations through Site Content Analyzer	Tables created through word count
Pope et al. (2009). Conceptual framework for collecting online airline pricing data: Challenges, opportunities, and preliminary results	Challenges and opportunities to collect large volumes of data from airline websites and travel agencies are discussed. Research questions are highlighted that can be investigated with this type of data.	Conceptual	N/A	N/A	N/A	N/A	N/A
Ritchie and Ritchie (2002). A framework for an industry supported destination marketing information system	Guidelines for the establishment of a comprehensive destination marketing information system (DMIS)	Empirical	Industry stakeholders' knowledge needs and current use of research & intelligence (Inter-)National Travel Survey	Primary survey data from 68 individuals Secondary data (travel surveys)	Semi-structured interview	N/A	N/A
Rossetti et al. (2016). Analyzing user reviews in tourism with topic models	A description of the topic model method with application focus on the tourism domain	Empirical	Yelp Data set Challenge; TripAdvisor Dataset	Structured and unstructured	Yelp is existing data set; TripAdvisor automatically collected by crawler	K-Nearest Neighbor User Based (KNN-UB), K-Nearest Neighbor Item Based	Illustrative examples for selected Topics related to multi-criteria dimensions

						(KNN-IB), Probabilistic Matrix Factorization (PMF)	
Sánchez-Franco et al. (2016). Online Customer Service Reviews in Urban Hotels: A Data Mining Approach	Extraction of features from hotel reviews and analysis of their relationship with guests' hotel rating in the online travel agencies environment	Empirical	19,318 hotel reviews from 2014 to 2015 from booking.com	Structured and unstructured	Automatically (one-time)	Pathfinder network scaling, principal component analysis (PCA), linear mixed-effects regression	N/A
Snavely et al. (2008). Modeling the world from internet photo collections	Presents algorithms and results as a step towards 3D modeling of the world's well-photographed sites, cities, and landscapes from Internet imagery	Empirical	Flickr	Large sets of image data	Automatically downloaded from Flickr	Keypoint detection (SIFT keypoint detector) and matching (by approximate nearest neighbors (ANN) kd-tree); Structure for motion (to recover camera parameters); geo-registration (by digital elevation maps)	reconstructed scenes and photo connectivity graphs for 11 sites
Solnet et al. (2016). An untapped gold mine? Exploring the potential of market basket analysis to grow hotel revenue	Market Basket Analysis to identify and predict the purchasing behavior of customers based on their expenditure patterns in order to determine the most attractive additional products and services	Empirical	56,906 guest sales records from a luxury hotel group in Australia from 2009 to 2014	Structured	Automatically (one-time)	multivariate logit model	N/A
Sun et al. (2016). Chinese Customers' Evaluation of Travel Website Quality: A Decision-Tree Analysis	Identification of critical attributes that influence quality levels of a customer's travel	Empirical	Survey data from 25 individuals	Structured	Manually	Attention-interest-desire-action (AIDA) model, C4.5 decision tree	N/A

	agency's website experience						
Tseng and Won (2016). Integrating multiple recommendation schemes for designing sales force support system: A travel agency example	Proposes a design of sales force support (SFS) system with business intelligence methodologies	Empirical	N/A	N/A	N/A	Explorative data analysis (EDA) and data mining (e.g. sequential pattern discovery)	Dashboards, On-Line Analytical Processing (OLAP)
Versichele et al. (2014). Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent	Spatiotemporal tourism behavior by mining of association rules in tourist attraction visits based on Bluetooth tracking methodology	Empirical	17,496 Bluetooth devices being detected over 235,597 time intervals by 15 Bluetooth sensors in Ghent in 2015	Structured	Automatically	A-priori association rule mining	Visit pattern maps
Wu et al. (2010). Data mining for hotel occupancy rate: An independent component analysis approach	Identification of major factors determining the hotel occupancy rate and incorporation of these factors to decompose hotel occupancy rates and examine the effect of each factor on the hotel occupancy rate	Empirical	Monthly hotel occupancy rate time series for each district of Hong Kong from January 1996 to May 2009	Structured	Manually	Independent component analysis (ICA)	N/A
Zhang and Huang (2015). Mining tourist motive for marketing development via twice-learning	Application of twice-learning framework to predict tourists' travel motives from tourists' external and internal features, useful for targeted marketing strategy development	Empirical	On-site surveys in Nanjing, China, fromOctober to November 2012 with 121 responses	Structured	Manually	Twice-learning framework, neural networks, C4.5 decision tree, Naïve Bayes	N/A
Zhu et al. (2016). Get into the spirit of a location by mining user-generated travelogues	Location information extraction from user-generated travelogues, examining contents and structures of travelogues, as well as their interplay	Empirical	80,384 travelogues related to tourist destinations in the United States	Unstructured	Manually	Gazetteer-based location detection, semantic correlation detection by natural language	N/A

						parsing techniques, location concept network by PLSA	

Table 3. Big Data works in Hospitality and Tourism (selected works; in "type of data and size" an asterisk indicates large quantities of data, > 100 000 records)

Article (author and title)	Research topic	Type of paper (conceptual/ empirical)	Sources of data	Type of data and size	Data collection methods	Data analysis techniques	Data reporting and visualization
Buhalis and Foerste (2015). SoCoMo marketing for travel and tourism: Empowering co-creation of value	Proposes social context mobile (SoCoMo) marketing as a new framework that enables marketers to increase value for all stakeholders at the destination	Conceptual	N/A	N/A	N/A	N/A	N/A
Carter (2016). Where are the enslaved?: TripAdvisor and the narrative landscapes of southern plantation museums	Explores what visitors learn about the history of the enslaved on two tours (Laura and Oak Valley) and how they participate in the narrative construction of the plantation	Empirical	TripAdvisor visitor reviews (Laura and the Oak Alley museums, USA)	Unstructured	Web (manual) scraping	Word frequency and words associations in reviews	Standard tables
Dolnicar and Ring (2014). Tourism marketing research: Past, present and future	Critical review of tourism marketing research	Literature review	N/A	N/A	N/A	N/A	N/A
Fuchs et al. (2014). Big data analytics for knowledge generation in tourism destinations - A case from Sweden	Presents a knowledge infrastructure implemented at the Swedish mountain tourism destination, Åre and examples of use by tourism managers	Empirical	Web search, booking and feedback data (e.g., survey-based, user-generated content)	Structured and unstructured	Data Warehouse (DW) including Facts and Dimensions Tables	On-Line Analytical Processing (OLAP); Support Vector Machines (SVM), Naïve Bayes (NB) and	Html-based web application

						K-Nearest Neighbor (KNN)	
García-Pablos et al. (2016). Automatic analysis of textual hotel reviews	Describes OpeNER, a NLP platform applied to the hospitality domain to automatically process customer-generated textual content	Empirical	Online reviews from Zoover and HolidayCheck	Unstructured	Web crawler	Natural Language Processing: Named Entity Recognition, Sentiment Analysis and Opinion Mining	Standard tables
García-Palomares et al. (2015). Identification of tourist hot spots based on social networks: A comparative analysis of european metropolises using photo-sharing services and GIS	Use of photo-sharing services for identifying and analyzing the main tourist attractions in eight major European cities	Empirical	Panoramio photos	Unstructured (*)	Panoramio website API + ArcGIS	Density graphs, spatial autocorrelation	Standard tables and Anselin Local Moran's I graph
Gretzel et al. (2015). Smart tourism: Foundations and developments	Defines smart tourism, sheds light on current smart tourism trends, and lays out its technological and business foundations	Conceptual	N/A	N/A	N/A	N/A	N/A
Gunter and Önder (2016). Forecasting city arrivals with google analytics	10 Google Analytics website traffic indicators from the Viennese DMO website are used to predict actual tourist arrivals to Vienna	Empirical	Google analytics variables collected on a monthly basis over the period August 2008-October 2014	Structured	Simple access to Google analytics	VAR model class: BVAR, FAVAR, BFAVAR.	Basic tables of descriptive statistics

He et al. (2016). Travel-package recommendations leveraging social influence of different relationship types	Develops a probabilistic topic model leveraging individual travel history and social influence of co travelers to capture personal interests and propose a recommendation method to utilize the proposed model	Empirical	Structured travel records on travel packages	Structured	Access to a private company database	Biggs sampling	Basic tables of descriptive statistics
Jackson (2016). Prediction, explanation and big(ger) data: A middle way to measuring and modelling the perceived success of a volunteer tourism sustainability campaign based on 'nudging'	Uses 'automatic linear modelling' that can cope with big data and presents the results as visualizations	Empirical	Structured (responses from questionnaire)	Structured	Survey	Automatic linear modelling and preparation through IBM SPSS	Basic tables of descriptive statistics and graphs stemming from automatic linear modelling (IBM SPSS)
Kong and Song (2016). A study on customer feedback of tourism service using social big data	Design of an analysis model for the top Korean travel agency to help the company improve customer satisfaction and service quality	Empirical	Internal sources (emails, counselling data, bulletin information, after use comments/ evaluations) and external sources (Twitter, Facebook, OnlineNews, Blog, Community)	Mostly unstructured (e.g., data from emails, social media networks) and several structured (e.g., bulleting info)	BuzzMonitoring (Types and proportion of keywords from the extracted data are digitized to analyze incidents and phenomena)	BuzzMonitoring including the following modules: NLP, data clustering, text summarization, sentiment analysis, structure data joinder.	No table nor graphs stemming from the Buzz Monitoring

Law et al. (2011). Identifying changes and trends in Hong Kong outbound tourism	Trends in Hong Kong outbound tourism in terms of Future trip intentions Travel destinations Motivation to travel.	Empirical	Tourism behavior survey data	Survey questionnaire	Historical domestic Surveys	Data mining, association rules, contrast set mining	Tables
Mariani et al. (2016). Facebook as a destination marketing tool: Evidence from Italian regional destination management organizations	Explores how Italian regional Destination Management Organisations (DMOs) strategically employ Facebook to promote and market their destinations, and improves on the current metrics for capturing user engagement	Empirical	Overall number of Facebook posts posted on the official Italian regional DMOs' Facebook pages	Structured (*)	Data extractor based on Facebook APIs	Data parser and analyzer calculating per post statistics	Tables created through data analyzer module. Graphs created through the data visualizer module
Marine-Roig and Anton Clavé (2015). Tourism analytics with massive user-generated content: A case study of Barcelona	Studying the online image of Barcelona as transmitted via social media through the analysis of more than 100,000 relevant travel blogs and online travel reviews (OTRs) written in English	Empirical	Heterogeneous including the travel blogs, webpages, travelogues and travel reviews about Barcelona (250,000 pages)	Unstructured (*)	Data was extracted through Offline Explorer Enterprise (OEE).	Pre-processing: web content mining, language detection, user's hometown, cleaning, debugging. Processing: parser settings and categorizations through Site Content Analyzer	Tables created through word count
Mi et al. (2014). A new method for evaluating tour online review based on grey 2-tuple linguistic	Evaluation of online reviews	Empirical	Reviews from tourism website	Unstructured	Web crawler	Grey 2-tuple linguistic evaluation (expert evaluations)	Standard tables

Mocanu et al. (2013). The twitter of babel: Mapping world languages through microblogging platforms	Survey on worldwide linguistic indicators and trends through	Empirical	Large-scale dataset of geotagged tweets	Structured and unstructured (*)	Twitter API	Language detection (Google Chromium Compact Language Detector) - Geographical analyses	Maps + charts and tables
Noguchi et al. (2016). Advanced, high-performance big data technology and trial validation	Presents technology for analyzing data and location data	Conceptual, application design, case study	Smartphone application	Structured: app’s usage logs, location data, and individual	From app logs	Clustering analysis	Maps + charts and tables
Orellana et al (2012). Exploring visitor movement patterns in natural recreational areas	Explores the properties of the collective movement of visitors in recreational natural areas	Empirical	Global Positioning System tracking data	Structured (GPS data)	Recording of GPS data	Kernel-density function classification and Generalized Sequential Patterns	GIS + standard tables
Paldino et al. (2015). Urban magnetism through the lens of geo-tagged photography	Tastes of individuals, and what attracts them to live in a particular city or spend vacation there.	Empirical	Geo-tagged photos	Structured: metadata from photos (*)	Flickr API	Identification of resident, tourist and unknown, statistical analysis, network analysis (origin/destinatio n)	Maps + charts and tables

Park et al. (2016). Using twitter data for cruise tourism marketing and research	Analysis social media data on cruise tourism	Empirical	Tweets containing search words	Unstructured/ structured	Twitter API and Web scraping	Word frequency, content analysis, and network analysis	Charts, tables, network diagrams
Raun et al. (2016). Measuring tourism destinations using mobile tracking data	Measure space-time tracking data to analyze, monitor and compare destinations based on data describing actual visits	Empirical	Anonymized roaming data of the foreign mobile phone call detail records	Structured	From telcom operator	Statistical analyses, ArcGIS for spatial analyses, binary logistic regression	Charts, tables, maps
Shi et al. (2016). Applying semantic web and big data techniques to construct a balance model referring to stakeholders of tourism intangible cultural heritage	Apply semantic web and big data techniques to help collect data, and implement platform and questionnaire design to construct stakeholder balance model for tourism intangible cultural heritage	Empirical	Questionnaire + User Generated Content (reviews)	Structured/ unstructured	Unknown UGC	Structural equation model, path analysis	Charts, tables,
Su et al. (2016). Characterizing geographical preferences of international tourists and the local influential factors in china using geo-tagged photos on social media	Characterize geographical preferences of international tourists and quantify local influential factors of tourists' destination preferences across time and space and origins	Empirical	Metadata online geotagged photos	Structured	Flickr API	Statistical and spatial analyses	Charts, tables, maps

Sun et al. (2016). Internet of things and big data analytics for smart and connected communities	Integration of Internet of Things (IoT) and big data analytics for smart connected communities	Conceptual + case study	Design of an IoT system personal sensors, open data, and participatory sensing to enhance the services in the area of tourism and cultural heritage with a Context-Aware Recommendation System	N/A	N/A	N/A	N/A
Supak et al. (2015). Geospatial analytics for federally managed tourism destinations and their demand markets	Examine the general geospatial demand for overnight recreation on federal lands prior to the 2008 recession and the specific geospatial demand for national park regions	Empirical	National Recreation Reservation Service reservations database	Structured (*)	DB access	Statistical and spatial analyses	Charts, tables, maps
Tang et al. (2016). Spatial network of urban tourist flow in Xi'an based on microblog big data	Study related to spatial network of tourist flows and its structure in the urban areas	Empirical	Geotagged microblog posts	Structured (*)	API from Sina Microblog.	Statistical, spatial and network analyses	Charts, tables, maps, network diagrams

Wang et al. (2015). Revenue management: Progress, challenges, and research prospects	Discuss evolution and future developments of revenue management and use of big data analytics	Conceptual	N/A	N/A	N/A	N/A	N/A
Wood et al. (2013). Using social media to quantify nature-based tourism and recreation	Online posted photos are used to estimate visitation rates and travelers' origins. compare to empirical data showing that crowd-sourced information can serve as reliable proxy for empirical visitation rates	Empirical	Empirical datasets that quantified visitation to 836 sites in 31 countries around the world + Flickr metadata	Structured (*)	Dataset + Flickr API	Statistical and spatial analyse	Charts, tables, maps
Yang et al. (2014). Predicting hotel demand using destination marketing organization's web traffic data	Demonstrate the value of website traffic data in predicting demand for hotel rooms at a destination, and potentially future revenue and performance	Empirical	Website traffic data and local hotel room demand data	Structured	Google analytics + standard data	Statistical and time series forecasts	Charts, tables

Yang et al. (2016). The big data analysis of land use evolution and its ecological security responses in silver beach of china by the clustering of spatial patterns	Use remote-sensing images to analyze the land use evolution and to evaluate its ecological security	Empirical	Landsat satellite high-definition images	Pictures + metadata	Landsat DB (*)	Land use temporal statistical analysis	Charts, tables, maps