Scientific
Research
Publishing

Linked Data Based Framework for Tourism Decision Support System: Case Study of Chinese Tourists in Switzerland

Zhan Liu[1], Anne Le Calvé[1], Fabian Cretton[1], Nicole Glassey Balet[1], Maria Sokhn[1], Nicolas Délétroz[2]

[1]Institute of Information Systems, University of Applied Sciences and Arts Western Switzerland (HES-SO Valais-Wallis), Sierre, Switzerland
[2]Institute of Tourism, University of Applied Sciences and Arts Western Switzerland (HES-SO Valais-Wallis), Sierre, Switzerland
Email: zhan.liu@hevs.ch, anne.lecalve@hevs.ch, fabian.cretton@hevs.ch, nicole.glassey@hevs.ch, maria.sokhn@hevs.ch, nicolas.deletroz@hevs.ch

Abstract

Switzerland is one of the most desirable European destinations for Chinese tourists; therefore, a better understanding of Chinese tourists is essential for successful business practices. In China, the largest and leading social media platform—Sina Weibo, a hybrid of Twitter and Facebook—has more than 600 million users. Weibo's great market penetration suggests that tourism operators and markets need to understand how to build effective and sustainable communications on Chinese social media platforms. In order to offer a better decision support platform to tourism destination managers as well as Chinese tourists, we proposed a framework using linked data on Sina Weibo. Linked Data is a term referring to using the Internet to connect related data. We will show how it can be used and how ontology can be designed to include the users' context (e.g., GPS locations). Our framework will provide a good theoretical foundation for further understand Chinese tourists' expectation, experiences, behaviors and new trends in Switzerland.

1. Introduction

Switzerland, as one of the most desirable European destinations for Chinese tourists, is strongly positioned as a steep and consistent quantitative growth market. At the end of 2013, there were a total of nearly 1 million overnight stays in Switzerland for Chinese travelers. China ranked 6th in terms of total overnight volume, 5th in turnover and 3rd in daily expenditure (330 CHF) [1]. Chinese visitors spent a total of 3 billion Swiss francs on

luxury goods and souvenirs in Switzerland in 2012. These data indicate unquestionable opportunities for economic growth for Switzerland, particularly for its tourist goods industry such as watches and other luxury products that are largely dependent on foreign tourists. To keep track of the current developments in the industry and adapt the communication and strategies, it is vital to maintain the competitive advantage of Switzerland as a preferred traveling destination and make economic contributions to other sectors (e.g., hotel industry, watch retail industry). A better understanding of Chinese tourists in Switzerland is therefore essential for successful business practices.

With the rising popularity of traveling websites (e.g., TripAdvisor) and social networking sites (e.g., YouTube, Facebook) that contain information submitted by tourists, the ways that travelers search for and evaluate travel information are changing [2]. Indeed, Chinese tourists now increasingly want to travel by themselves or in small groups of friends to visit Switzerland [1]. When deciding the place and direction, they often turn to online user-generated content available through social media to obtain information prior to their vacations, believing other travelers' experiences to be trustworthy and useful [3]. As a result, recommendations of other travelers who have prior experience are not only the most preferred source, but also the most influential source for travel decision making.

In China, the largest and leading social media platform—Sina Weibo, a hybrid of Twitter and Facebook—now has over 600 million users and is currently regularly used by more than 140 million users. Weibo's great market penetration suggests tourism operators and markets need to understand how to build effective and sustainable communications on Chinese social media platforms. By analyzing Chinese consumers' Weibo activities, it is possible to learn about their characteristics, interests, preferences and concerns, as well as their expectations and experiences in Switzerland.

However, the tourism decision making is difficult, both for tourists as well as tourism destination managers, due to the complexity of the large number of information within this industry. As for Chinese consumers abroad, the issues to be addressed relate to food, accommodation, transportation, language, shopping, entertainment, weather, etc. For tourism destination managers, the massive nature and complexity of data also make it difficult to make decisions. Furthermore, Chinese people have changed significantly of their ways and consumption behaviors recently, leading to the dynamic changes of tourism data. Using traditional ways to collect data from group tourists and industry statistics to estimate the volume of demand might be problematic because it takes long time and may not capture the individual trips. Therefore, it is necessary to identify up-to-date changes, and make timely management decisions to better attract Chinese tourists and provide better services.

Linked Data is a term referring to using the Internet to connect related data. This term is often used to describe a recommended best practice for exposing, sharing, and connecting data, information, and knowledge on the Semantic Web using URIs and RDF. It is an important factor for improving the quality of natural language processing (NLP), such as user's interpretation from social media platforms. Many NLP tasks can greatly benefit from making use of this wealth of knowledge available on the Web in a structured form such as Linked Data. For tourism destination mangers, they want information presented in a format that effectively assists them in making decisions. We believe that semantic technology based linked data can help to identify Chinese tourists' behavior and their consumption trends, which in turn help managers to better manage the tourism destination.

Nevertheless, although several researchers have addressed the perceptions of Chinese tourists about certain destinations to check their intention to visit, we argue that these studies are somehow subjective because the participants in these studies do not know or have never been to the destination. However, there are far fewer studies which have empirically examined real Chinese tourists' attitudes and experiences (e.g., on social media platforms). Even fewer have analyzed the positioning of Switzerland as a destination for Chinese tourists. Research on linked data with behavior analysis in this field is also notably lacking.

The goal of this research is to study decision support system by adopting a linked data approach. We intend to build this framework on a Chinese social media platform with a large-scale analysis with linked data and explore Chinese tourists' attitude and behavior towards Switzerland. We selected Sina Weibo, which has played an important role as a powerful media channel to extend personal networks and informal information sources, as our use case to describe this framework. Future research should use real data to validate it.

2. Background and Literature Review

The study mainly draws upon the literature relative to user generated content and, in particular, the influence on

tourism of such content from social media. In the next section, we will first present the development and analysis of the Chinese tourism market in Switzerland and then discuss the related work of user generated content and tourism.

2.1. Chinese Tourism in Switzerland: Descriptive Analysis

Tourism has been considered as one of the most important economic activities in the world, both for developing and developed countries. People all over the globe are more willing to spend on traveling. As a consequence, the tourism industry has accumulated extravagant economic contributions. According to the World Tourism Organization (WTO) in 2013, about 9% of the world's Gross Domestic Product (GDP) is generated by the tourism sector. International tourist arrivals worldwide reached a record 1087 million, and these are expected to grow to 1.8 billion by 2030. China leaped to first place in the ranking by international tourism expenditure in 2012. In 2013, it consolidated its position as the number one tourism source market in the world, spending US$129 billion on international tourism.

Compared with other counties, Switzerland is strongly positioned as a steep and consistent quantitative growth market. In 2013, the overnight stay of mainland Chinese tourists in Switzerland reached 994,456; an increase of 19% as compared to the year 2012. It is expected that there will be a high market potential for Chinese tourists in Switzerland in the following years due to a very high GDP growth and the high share of young and first-time visitors. Taking advantage of increasingly fewer restrictions on travel visas, as well as an appreciating currency, Chinese consumers will continue to accumulate economic contributions for Switzerland. In 2012, Chinese visitors spent a total of 3 billion francs on luxury goods and souvenirs in Switzerland.

While many Chinese tourists still travel in Switzerland in organized groups due to the linguistic barriers, this trend is changing. Individual trips are expected to increase in the next few years, particularly among experienced tourists. The reason is probably because long-distance Chinese tourists who belong to a middle-high social class not only have greater purchasing power, but they are also better educated and have better knowledge of other languages, as reported by the [4]. In addition, the majority of Chinese travelers in Switzerland tend to be young: 69.7% of Chinese travelers are less than 35 years old [1].

The main reasons for Chinese tourists to visit Switzerland are the following [1]: nature, special natural attraction, easy access, pleasant climate/weather and lakes/rivers. In short, Chinese travelers in general seek new cultures and natural experience in Switzerland. However, in addition to technical reports by tourism organizations and officials, few studies have examined Chinese consumer's behaviors in Switzerland. Therefore, the focus of this study is on the behavior analysis of Chinese consumers in Switzerland in general.

2.2. User Generated Content

In a travel context, consumers often spread word of mouth (WOM) due to the fact that traveling is associated with a product—experience. Research has indicated that, due to the lack of commercial self-interest associated with WOM recommendations, consumers tend to trust and be more influenced by this type of information than by more commercial sources such as travel agents or accommodation operators ([2] [5] [6]). This is probably because the provider of such information is not generally aiming to make a financial gain from sharing their experiences and views with others. [6] documented that prospective tourists often rely heavily on WOM from friends, family and other peer groups, particularly when planning travel to a destination not previously visited.

As in the online environment, recommendation sources can come from other unknown consumers, and there is far more information available than from traditional WOM. Today, many websites allow users to add their own content. Such content is known as User Generated Content (UGC). Individuals can learn more about tourist destinations, conduct detailed trip planning and seek out the best price or value through the Internet via UGC [7]. [6] also documented that the majority of US tourists use the medium to search for travel information.

However, the quality of the information and the expertise from these sources sometimes varies tremendously. UGC has become one of the most influential information sources for traveling decision making. [8] found that a higher percentage of tourists turned to UGC when visiting a destination for the first time, as well as visiting an international destination. [9] proved that when tourists to a well-known destination decide where to spend their summer holiday, they tend to use social media. Therefore, Switzerland, as a well-known destination with a relatively high share of first-time Chinese visitors, is quite relevant in this research.

The flourishing promise of UGC is now being hyper-realized with social media. In China, the most prevailing

social media platform is Sina Weibo. Launched on August 28, 2009, Sina Weibo is growing at an astonishing rate. By the end of 2013, its registered users had reached 600 million with 143.8 million active monthly users. More than 5000 companies and 2700 media organizations in China use Sina Weibo. The current volume of weibos (microblogging messages on Sina Weibo, like tweets in Twitter), is over 2.8 billion per month. By providing an unprecedented and simple way of communicating, Sina Weibo allows people to publicly express themselves in real time, interact with others, as well as be heard publicly and be exposed to the rich ideas, cultures and experiences of the broader world. Traditional media uses Weibo as a distribution channel for their headline news. Government agencies and officials use Weibo as an official communication channel for disseminating timely information and gauging public opinion to improve public services. Companies use Weibo as a key marketing tool to seize market opportunities and boost their presence, thus maximizing their business values in China.

Sina Weibo users tend to be young—the vast majority is under the age of 32—and better educated, wealthier than the general population. Many use Weibo to keep up with current events. The demographic characteristics are similar to the majority of individual tourists to Switzerland who want to decide by themselves about the place and directions. Therefore understanding Chinese tourists' expectations and experiences with regard to Swiss tourist products will provide interesting insight.

Research on characteristics of user behavior on the Sina Weibo is available. For example, [10] collected data and identified users and their behavior on Weibo in general, and found that a one-way relationship is the most common form of communication between Weibo users. [11] analyzed and compared user behaviors, including textural features, topics and sentiment polarity of micro-posts on two different microblogging platforms: Sina Weibo and Twitter. [12] discussed the posting and reposting characteristics of some hot events. However, so far there exists little knowledge about Chinese tourist's expectations and experiences in Switzerland from Weibo user behavior. [13] conducted a very basic analysis of pictorial tourism-related content on Sina Weibo, regarding "Switzerland Travel" as a tourism destination, in attempt to understand the most important aspects of Chinese tourists in Switzerland. However, we argue that there are strong bias problems in their study. The reasons are twofold: First, the data used in the study was collected and analyzed through human interpretation, and as a result, only 440 pictures were included in the final analysis, which is problematic. In June 2014, querying Weibo's search engine using the keywords "Switzerland Travel" returns almost 2 million results on Weibo. Analyzing and interpreting such huge amount of information by any human endeavor may be rather futile. Second and more importantly, the keyword "Switzerland and travel" fail to serve as a good representation because Weibo users may post text messages without explicitly mentioning the word "Switzerland" or "travel" but still refer to the object of travel in Switzerland. For example, a Weibo likes "I am enjoying fondue in Valais" and "I am in front of the Olympic museum in Lausanne" is highly related to traveling experience in Switzerland. In addition, by the time the data was collected in June 2012, monthly active users were less than 80 million, and daily active users were 3.5 million. Both numbers almost doubled in early 2014. It is, therefore, necessary to conduct new studies to understand Chinese consumer behavior in Switzerland. In this paper, we intend to close this research gap.
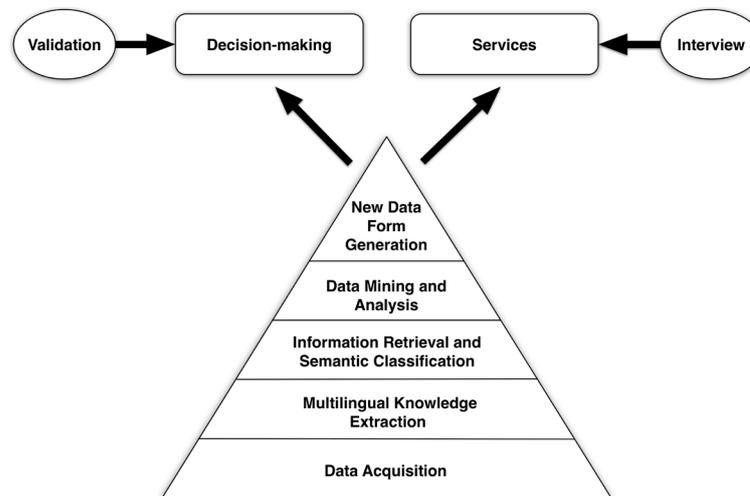
3. Framework of Tourism Decision Support System

This paper is part of linked data based tourism decision support system research program. The main objective of the current stage is to propose a framework towards a tourism decision making support system platform. Such framework is expected to identify Chinese tourists' expectations and their real experiences in Switzerland, and therefore understand their consumption patterns to support tourist manager's decision and to provide better services to Chinese tourists.

In this section, we propose a framework of linked data based tourism decision support system. **Figure 1** describes the proposed platform, including data acquisition, multilingual knowledge extraction, and information retrieval, semantic classification, data mining and analysis, new data form generation, as well as decision support system and services, and interview validation.

3.1. Dataset Acquisition

As of the preliminary stage, the dataset can be collected via the Sina Weibo Open API, which has been open since July, 2010. Sina Weibo represents the convergence of user text commentaries with photos, video, and mu-

Figure 1. Framework for tourism decision support system.

sic in a simple and user-friendly way, allowing users to create, distribute, discover, share and communicate Chinese language content. Developers or websites can create creative applications to access the information on Weibo through the Open API platform. Two types of datasets can be considered here: user profiles and micro blogs (weibos). The profiles contain the user's demographic information, such as login name, real name, location, gender, date of birth, education and profession, as well as other information such as uses preference and network, including preferred tags (e.g., travel, movies, food, etc.), the number of micro blogs, and the list of followers and followings. For the purpose of our research, our attention would focus on the content related to Switzerland, including their text messaging and location information. However, using the subject "Switzerland and travel" as the only keyword to search the Weibo database will not capture enough important information, because in many cases the user's micro blog only contains information related to Switzerland, such as the name of a city in Switzerland, but does not explicitly mention the word "Switzerland" or "travel" in the text message. Therefore, to increase the validity of this research, it is necessary to build a multiple language Switzerland based ontology as the first step before collecting data from Weibo. The design of this ontology will focus on its class and property for the administrative units and the most evidence units of Switzerland, such as the points of interest, famous buildings, and the most visited places, in which the triple data are linked to other data from existing ontologies, for instance, Geonames, DBPedia, Word Net, and data.admin.ch. Moreover, user location information could also help to provide additional information on Swiss related data. By building this ontology, other domains such as luxury products, food, education, job opportunities, economic and commercial authorities may also possibly be included in future research.

3.2. Multilingual Knowledge Extraction

To match the business needs of this research, it is necessary to build an ontology-based system that includes two languages-Chinese and English, representing over 99% of micro blogs. The objective is to create the knowledge from unstructured sources, such as user's text. Therefore, two main methods are applied to achieve the objective: text mining and multilingual natural language knowledge extraction. Text mining is a domain of natural language processing (NLP), a computer science field where machines extract knowledge from text written in natural language. In most NLP tasks, machine learning tools have been successfully used for years. The downsides of this approach are the amount of human work to annotate existing corpus to train the system, and the limitation of the system that becomes efficient only for the domain it has been trained for. Recently, knowledge based techniques have yielded comparatively good results, and even better results in some cases, as demonstrated by [14] for word-sense disambiguation or [15] for text classification. In this approach, the knowledge from multilingual natural language (*i.e.* Chinese and English) is described in a structured manner, for instance using ontologies, and the NLP algorithms [16] exploit that knowledge to perform their tasks on the texts. Moreover, data-driven clustering algorithms will be used to correctly identify the similarity-based measure. That is to say, these algorithms will capture cases when different words are used to describe the same concept, it requires to

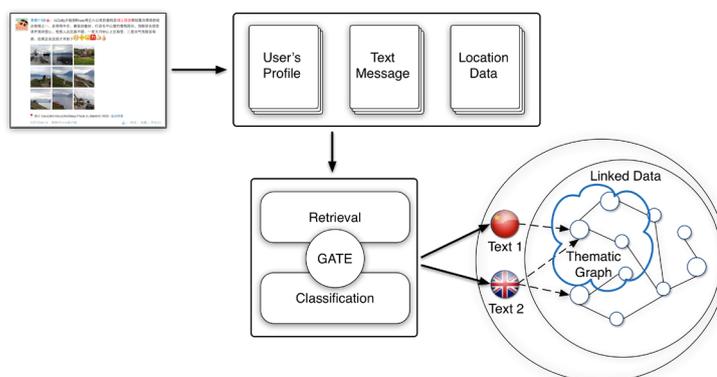generate a schema based on the source data.

3.3. Information Retrieval and Semantic Classification

Information retrieval and ontology-based classification techniques are used to calculate frequency of certain text, attempting to identify the most relevant elements of a corpus to satisfy the information retrieval. An ontology represents knowledge as a hierarchy of concepts within a domain, using a shared vocabulary to denote the types, properties and interrelationships of those concepts [17]. According to [18], the ontology-based model includes a set of linked vocabularies that would disambiguate meanings of words from natural language text sentences, overcoming the problem of having a word which may have multiple meanings depending on the applicable context. Full-text information retrieval, providing a simple match between a search keyword and that word in the indexed texts, is quite limited when it comes to lexical semantics. Therefore we intend to take one step further to develop a system that can analyze whole sentences that contain unstructured text from Weibo, and use a multi-lingual ontology reflecting the domain knowledge for a semantic classification of the user's interpretations of Weibo. Traditional NLP and information retrieval technics both have to deal with lexical semantic issues, for which semantic web technologies provide very interesting features. GATE (General Architecture for Text Engineering) is a multilingual engine for natural language processing. It includes ANNIE (A Nearly-New Information Extraction System), which we hope use in our study. ANNIE has some important relevant modules for our study for instance, tokenizer, gazetteer, sentence splitter, part of speech tagger and disambiguation [19]. As shown in **Figure 2**, to store the results of analysis retrieval and semantic classification through ANNIE, we will develop an NLP ontology, where each text submitted is linked with triple data from a different ontology. According to [20], the graph of this ontology includes the domain entities organized into categories and interconnected by relationships, as well as instances and linkages among them. Moreover, the user's unstructured text is converted into a thematic graph of entities. The ontological classification of the entities in the graph is then analyzed to determine the overall categorization of the thematic graph both for English and Chinese.

3.4. Data Mining and Analysis

In this step, Fayyad *et al.* [21]'s six common classes of tasks in data mining analysis can be applied:
- Anomaly detection: the identification of unusual data records, that might be interesting or data errors that require further investigation. In this task, we identify the collected Weibo dataset which is not in conformity with an expected pattern.
- Association rule learning: searches for relationships between variables. We define association rule learning for discovering interesting relations between user variables in Weibo databases.
- Clustering: is the common descriptive task to identify a finite set of categories or clusters to describe the data. This task is used to combine the objects in the same group, similar to the data retrieval method that we introduced in the previous section.
- Classification: is the task of generalizing a known structure to apply to new data. In this step, we identify and map the user's natural language in Weibo's interpretation as a data item into one of several predefined

Figure 2. Process of data retrieval and semantic classification.

classes with the linked data.

- Regression: is learning a function that maps a data item to a real-valued prediction variable. This task is used to predict Chinese consumer behavior in Switzerland, such as purchasing power, desired service, interesting place products, etc.
- Summarization: providing a more compact representation of the dataset, including visualization and report generation for the final exploratory data analysis in this study.

3.5. New Data Form Generation

This is an important step to delivering visualized and understandable data to the DSS and services. For example, Chinese tourists' preferred destinations and accommodations can be displayed on a map with the frequencies and consumption information. Firms or service providers can also use business model canvas as a visual chart and useful tool to describe and design new or develop existing business models to illustrate potential trade-offs before make any tourist related decisions.

3.6. DSS and Services

Based on new data form analysis, the final step is to build the user interface to communicate with the tourist destination managers and service providers. All data that has been collected, classified, analyzed and visualized from the massive and complex data in Sina Weibo in previous steps should be easily interacted with users. The final tourism decision support system will help the tourism destination managers set up the concept of tourism process, and provide them with important basis in forecasting tourism situations and guiding Chinese tourists in a reasonable and reliable manner. Users can set his/her preferences of the category and format of the data. For example, a restaurant manager, who wants to introduce new menu items to attract more Chinese tourists, would probably be interested in Chinese tourists' preferences in terms of food taste and price category. He can select the relevant data in his city and display the data in a distributed pie chart.

3.7. Validation: Interviews

In order to study in more depth and provide additional evidence for our results, we also include the validation stages in our framework. The validation stage mainly consists of individual interviews and focus group interviews. More specifically, individual interviews with key person such as tourist agency managers and potential partners can help to understand Chinese consumers' behavior and trend, and focus group interviews with Chinese consumers can be conducted to explore their motivations and expectation (for consumers who have not visited Switzerland), and experiences (for consumers who have visited Switzerland).

Individual interviews. To capture Chinese consumers' behavior and trend, a survey procedure that focus on the depth and richness of interview data should be designed to encourage free expressions of interviewees' thoughts and feelings. The main target groups of the individual interviews are managers in tourist agencies and potential partners in tourism or related industries,

Focus group interviews. Since Chinese consumers' motivations and expectation (for consumers who have not visited Switzerland), and experiences (for consumers who have visited Switzerland) of visiting Switzerland have not been studied extensively, to elicit an in-depth understanding of the issues and provide a solid base for further relevant instrument construction, a focus group method is thought to be appropriate as it can offer researchers opportunities to garner substantive information about participants' thoughts and feelings on the topic of interest in a relatively short period of time [22] [23]. The focus group respondents are selected in a way that maximized the views of people with a wide range of occupation, gender, age and education.

4. Conclusions and Future Research

This paper proposed a framework towards a tourism decision making support system platform by adopting linked data approach. It described a use case of the usage of tourism data from Sina Weibo on Chinese tourists in Switzerland, and thus provided evidence on how semantic web technology-based linked data can be used to manage tourism data to support decision making.

This study filled an important gap in the literature on Chinese consumer behavior in Switzerland based on social media platforms. To date, little is known about the tourism-related UGC from the Chinese social media

platform Sina Weibo. This research work would serve a first step toward a better understanding in this regards.

Future research should follow the proposed framework by employing real tourism data in Sina Weibo. With a large amount of text messages in Sina Weibo and intensive analyses, the results will allow a comprehensive view of how Chinese tourists perceive Switzerland and their preferences regarding products and services. Such understanding includes identifying Chinese tourists' favorite visiting place, determining factors that influence travel decisions, measuring satisfactions of tourism activities performance and recognizing new opportunities.

Acknowledgements

References

[1] Switzerland Tourism (2013) Research Report China: Market Analysis and Insight. http://www.stnet.ch/files/?id=63853

[2] Cox, C., Burgess, S., Sellitto, C. and Buultjens, J. (2009) The Role of User-Generated Content in Tourists' Travel Planning Behavior. *Journal of Hospitality Marketing and Management*, **18**, 743-764. http://dx.doi.org/10.1080/19368620903235753

[3] Yoo, K.-H., Lee, Y., Gretzel, U. and Fesenmaier, D.R. (2009) Trust in Travel-Related Consumer Generated Media. In: Hopken, W., Gretzel, U. and Law, R., Eds., *Information and Communication Technologies in Tourism* 2009, Springer Verlag, 49-59. http://dx.doi.org/10.1007/978-3-211-93971-0_5

[4] China Tourism Academy (CTA) (2013) Annual Report of China Outbound Tourism Development 2013. Tourism Education Press, Beijing.

[5] Burgess, S., Sellitto, C., Cox, C. and Buultjens, J. (2009) User-Generated Content (UGC) in Tourism: Benefits and Concerns of Online Consumers. *Proceedings of* 17*th European Conference on Information System*, Verona, 8-10 June 2009.

[6] Litvin, S.W., Goldsmit, R.E. and Pan, B. (2008) Electronic Word-of-Mouth in Hospitality and Tourism Management. *Tourism Management*, **29**, 458-468. http://dx.doi.org/10.1016/j.tourman.2007.05.011

[7] Pan, B., MaClaurin, T. and Crotts, J.C. (2007) Travel Blogs and the Implications for Destination Marketing. *Journal of Travel Research*, **46**, 35-45. http://dx.doi.org/10.1177/0047287507302378

[8] Simms, A. (2012) Online User-Generated Content for Travel Planning–Different for Different Kinds of Ttrips? *E-Review of Tourism Research*, **10**, 76-85.

[9] Jacobsen, J.Kr.S. and Munar, A.M. (2012) Tourist Information Search and Destination Choice in A Digital Age. *Tourism Management Perspectives*, **1**, 39-47. http://dx.doi.org/10.1016/j.tmp.2011.12.005

[10] Guo, Z., Li, Z., Tu, H. and Li, L. (2012) Characterizing User Behavior in Weibo. *Proceedings of the Third FTRA International Conference on Mobile*, *Ubiquitous*, *and Intelligent Computing*, Vancouver, 26-28 June 2012, 60-65. http://dx.doi.org/10.1109/MUSIC.2012.18

[11] Gao, Q., Abel, F., Houben, G.J. and Yu, Y. (2012) A Comparative Study of Users' Micro-Blogging Behavior on Sina Weibo and Twitter. *User Modeling, Adaptation, and Personalization*, **7379**, 88-101. http://dx.doi.org/10.1007/978-3-642-31454-4_8

[12] Guan, W., Gao, H., Yang, M., Li, Y., Ma, H., Qian, W., Cao, Z. and Yang, X. (2014) Analyzing User Behavior of the Micro-Blogging Website Sina Weibo during Hot Social Events. *Physica A*, **395**, 340-351. http://dx.doi.org/10.1016/j.physa.2013.09.059

[13] Hu, T., Kalbaska, N. and Cantoni, L. (2013) A Picturesque Online Representation of Switzerland: A Research Conducted on Chinese Social Network Weibo. http://www.sinoptic.ch/textes/recherche/2013/201304_USI_swiss-images-cn.pdf

[14] Navigli, R. (2009) Word Sense Disambiguation: A Survey. *ACM Computing Surveys*, **41**, 1-69. http://dx.doi.org/10.1145/1459352.1459355

[15] De Maat, E., Krabben, K. and Winkels, R. (2010) Machine Learning versus Knowledge Based Classification of Legal Texts. *Proceedings of the* 2010 *Conference on Legal Knowledge and Information Systems*, IOS Press, Amsterdam, 2010, 87-96.

[16] Hellmann, S., Lehmann, J., Auer, S. and Brummer, M. (2013) Integrating NLP Using Linked Data. *Proceedings of the* 12*th International Semantic Web Conference*, 21-25 October 2013, Sydney. http://dx.doi.org/10.1007/978-3-642-41338-4_7

[17] Gruber, T.R. (1993) A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, **5**, 199-220.

http://dx.doi.org/10.1006/knac.1993.1008

[18] Pan, X.S. (2002) A Context-Based Free Text Interpreter. Master's Thesis, California Polytechnic State University San Luis Obispo.

[19] Cunningham, H., Wilks, Y. and Gaizauskas, R. (1996) GATE—A General Architecture for Text Engineering. *Proceedings of the* 16*th International Conference on Computational Linguistics*, Copenhagen.

[20] Janik, M. and Kochut, K. (2008) Training-Less Ontology-Based Text Categorization. *Proceedings of the ECIR Workshop on Exploiting Semantic Annotations in Information Retrieval* (*ESAIR*'08), Glasgow, 30 March 2008, 3-17.

[21] Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996) From Data Mining to Knowledge Discovery: An Overview. In: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R., Eds., *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, Cambridge.

[22] Kitzinger, J. (1995) Qualitative Research. Introducing Focus Groups. *British Medical Journal*, **311,** 299-302. http://dx.doi.org/10.1136/bmj.311.7000.299

[23] Powell, R.A. and Single, H.M. (1996) Focus Groups. *International Journal for Quality in Health Care*, **8**, 499-504. http://dx.doi.org/10.1093/intqhc/8.5.499