e-Review of Tourism Research (eRTR), Vol. 10, No. 2, 2012
http://ertr.tamu.edu

Special Issue – ENTER 2012 Idea Exchange

Challenges and potential of the Semantic Web for tourism

Nina Mistilis
Australian School of Business
University of New South Wales

Dimitrios Buhalis
International Centre for Tourism & Hospitality Research
Bournemouth University UK

Abstract

The paper explores tourism challenges and potential of the Semantic Web from a theoretical and industry perspective. It first examines tourism business networks and explores a main theme of network interoperability - data standards- followed by technology deficiencies of Web 1.0 and 2.0 and Semantic Web solutions. It then explicates Semantic opportunities and challenges for tourism, including an industry perspective through a qualitative approach. Industry leaders considered that the new Web era was imminent and heralded benefits for supply and demand side interoperability, although management and technical challenges could impede progress and delay realisation.

Keywords: tourism networks; Web 1.0 and Web 2.0 deficiencies; semantic web

1 Introduction

Tourism by definition is a network business, as it relies on a number of stakeholders for its ability to develop and deliver products and services (Dale, 2003; Novelli, Schmitz and Spencer, 2006; Robertson, 2011). The term *business network* refers to 'a collection of inter firm relationships, including alliances, long-term buyer-supplier relationships, and informal collaborations. Firms increasingly enter business networks because of greater process modularity and interdependencies made possible by investments in ICT' (Kaufman, Li and van Heck 2010, p 115). Information and Communication Technology (ICT) supports these networks and is therefore a key to tourism network transactions (Poon,1994, Law, Leung and Buhalis, 2009; Mistilis and D'Ambra, 2008).

The technology evolution from Web 1.0, to Web 2.0, to Web 3.0 continues, following Tim Berners-Lee's inventing the World Wide Web in 1989. Web 1.0 consisted of static web pages enabling global information sharing but not interactivity. It was 'primarily a one-way publishing medium' and information-centric (Murugesan, 2010 p 2). Web 2.0, often called the social or community web, linked people and users, and with user generated content capability (Murugesan, 2010): it is people-centric. Introduced in 2001, the third stage of web development Web 3.0 is the Semantic Web, where meaning of content is recognised and understood by computers, enabling *machine to machine* interaction. This capability makes data interoperable and capable of manipulation by computer programs, bringing 'structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users' (Berners-Lee, Hendler and Lassila, 2001).

The Semantic Web is defined as a "mesh of information linked up in such a way as to be easily processable by machines, on a global scale" (Siau & Tian, 2004). It uses 'common and minimal language to enable large quantities of existing data to be analyzed and processed' (Gutierrez, Hurtado, Mendelzon and Perez 2011, p.250). The evolution of these technologies from non-interactive/static *information-centric* capability to *people-centric* capability to *machine and interoperability-centric* capability suggest that tourism organisations experience continuous change and the uncertainty that has defined it.

Most studies have examined the Semantic Web and individual tourism components (Wang, Zeng and Tang, 2011; Xiang, Gretzel and Fesenmaier, 2009; Cardoso and Lange, 2007) or tourism industry sectors (Bilbao, Lejarazu and Herrero, 2010; Lamsfus, Grün, Alzua-Sorzaba and Werthner, 2009; Kanellopoulos and Panagopoulos, 2008): this research examines the *whole system*. It analyses interoperability deficiencies of

Web1.0 and 2.0 technologies for tourism networks, focussing on data standards, followed by Semantic provision of technical solutions. Finally it explicates Semantic challenges and potential, including reference to industry perspectives through a qualitative approach.

2 Theory/issues

In their recent review article, Zhang, Song and Huang (2009) point out that tourism products are service networks – value added chains of various individual components that collectively form service networks. Tourism network relationships comprise 'a matrix of government agencies...private-sector actors, from a broad array of industries including lodging, bars, and restaurants, transportation … (Robertson, 2011). Based on worldwide cooperation between stakeholders (Fodor and Werthner 2005), network interoperability is enhanced by ICT which provides evolution through innovation (Hjalager, 2010; Williams and Shaw, 2011).

With increased tourism traffic, complex visitor itineraries and greater demand for online supplier capability, **t**he online tourism domain continues to grow in size and complexity (Buhalis and O'Connor, 2005; Xiang, Gretel and Fesenmaier 2009). Yet in spite of ongoing significant Web 1.0 and 2.0 developments to achieve integration, increased global tourism and its accompanying information flows are outpacing them (Jones, Ivezic, and Gruninger 2001). An extensive literature review revealed four main interoperability problems - data standards, search engines, dynamic packaging and mobile device functionality. Here the scope is limited to data standards.

There is a *lack of data standards* used to define customer wishes and correctly interpret common entities, concepts and relationships between actors (Bilbao, Lejarazu, and Herrero, 2010). Most tourism data are neither standardized nor immediately e processable. Its data exchange capability exists largely between *people and computers* with little *machine-to-machine* interaction capability. There is no global system for reading and manipulating data in order to process it easily across various actors, technologies and borders and supply the product (Jones, Ivezic, and Gruninger 2001). Global Distribution Systems still dominate the marketplace.

The lack of data standards calls for enhanced interoperability in the tourism business environment (Siau and Tian, 2004). There is a clear need for the next generation's technology infrastructure to support those critical processes through its machine-centric capability: the Semantic Web can resolve this (Jones, Ivezic and Gruninger, 2001) and enhance network interoperability, so critical for tourism.

2.1 Interoperability and the Semantic Web in Tourism

The need to develop standard ontologies to advance interoperability and enable web services has led to various initiatives (Dogac et al, 2004). The Resource Description Framework (RDF) is the basic language of the Semantic Web. It uses URIs (Unified Resource Identifiers) to represent data which can be held in database and/or interchanged; XML is the language used for syntax and structure. Web 3.0 technologies support network interoperability, facilitated by uniform data standards.

Formal tourism ontologies are already developed, or developing, in projects such as Harmonise http://cordis.europa.eu/fetch?CALLER=PROJ_ICT&ACTION=D&CAT=PROJ&RCN=57155; Mondeca (www.mondeca.com),

OnTour (http://ontour.deri.org/ontology/ontour-02.owl; Prantner, Ding, Luger, Yan & Herzog, 2007);

SATINE (http://www.srdc.metu.edu.tr/webpage/projects/satine/index.html; Bilbao, Lejarazu, and Herrero, 2010),

Hi-Touch (http://cordis.europa.eu/fetch?CALLER=PROJ_IST&ACTION=D&DOC=38&CAT=PROJ&QUERY=012f2a12d9ba:1b62:7a1ef058&RCN=63604; Zhang, 2009) and

Open Travel Alliance, APEX (Convention Industry Council, 2004) and TTI Travel Technology Initiative (http://tti.org/).

3 Method

The research is exploratory and uses two main approaches. First, main themes for tourism and the Semantic Web were identified through the theoretical lens of a literature review, one theme (data standards) was analysed and Semantic Web tourism initiatives identified. Second, a qualitative approach was applied to seek opinions

from industry leaders regarding Semantic Web and tourism, their challenges and potential in the business environment focussing on one main interoperability problem associated with networks - the data standards theme. This dataset is ideal for studying rapid, frequent change (Brown and Eisenhardt, 1997) and the data collection being less structured is appropriate in this very exploratory study (Ritchie and Lewis, 2003).

Respondents were carefully chosen based on recognition of their key leadership roles in tourism, technology and industry member organisations, in and beyond the United Kingdom and held a top level position. The total of ten respondents is adequate for this research (Stebbins, 2001; Rouse, 2004). Five interviews were completed at the Tourism Distribution Summit event held in London 10-11 May 2011 in a quiet comfortable space, the others were by phone (three) or Skype (two).

The open-ended questionnaire was used as there was insufficient basis to identify categories of industry interest; it allowed respondents to express opinions on three main topics: Impact of the Semantic Web, Tourism Business Networks with the Semantic Web and The Future. Examples include *What do you expect to be the impact of the Semantic Web/Web3.0? What are the Business/ICT challenges in preparation for the Semantic Web/Web 3.0 environment for your business over the next 5 years?* The answers were recorded and then transcribed and scanned to identify main themes and then subjected to a content and thematic analysis. The themes were then consolidated into various sub topics to facilitate clarity of analysis; only the data standards theme is reported here.

4 Results

Respondent comments on the data standards theme were grouped into three subtopics - potential benefits and opportunities of the Semantic Web for tourism supply and demand; concerns and doubts surrounding Semantic technology and the way ahead:

4.1 The Semantic Web benefits and opportunities for supply and demand

Respondents recognised the importance of enhanced global interoperability in more easily processing B2B and B2C data and benefits and opportunities of B2B and B2C interactions. The Semantic Web would facilitate interoperability in B2B network as:
It will make easy collaboration within airline alliances.... (airline respondent).
In B2C interoperability, it would better enable business management and reaching customers as:
It enables you to categorise common interests/ behaviours (travel agency respondent).

They were equally aware of wider potential benefits of demand, in customer interpretation of needs and desires through credible information and search results:
The consumer will find it *easier to trust information if standardised* (technology respondent).

However they raised consumer-related implications, as heightened consumer expectations regarding information would further shift the balance of power:
Technology 'trains' consumers ..semantic Web will 'train' consumers also for example in the expectations that information they receive is relevant (tour operator respondent).

Respondents tacitly looked to the semantic web for a better tourism trading environment. Enhanced product creation, description and communication with semantic tools benefited B2C interactions, with greater efficiency in matching tourism supply and demand. This in turn would reduce dependency on third party intermediaries. With enhanced interoperability, B2B networking would improve – especially critical for new entrants to alliances. As well, clearer information, greater trust in the information and easier Web navigation would collectively facilitate consumer choice.

4.2 Semantic web technology - concerns and doubts

However, the complexity of developing standards, the many start-up company failures and intricacies required to achieve correct language translation and interpretation, appear to suggest insurmountable challenges to the introduction of the semantic web:
multilingual interpretation also produce major challenges, for example will Web3.0 recognise that de Hage means The Hague (tour operator respondent) and

The Semantic Web is NOT going to emerge as data are unclassified and unstructured; It is too complicated and cannot emerge technically (technology respondent)

Most expressing these doubts were technology respondents. The perceived necessity for development of data standards underlines their inherent importance in Semantic implementation. Some obstacles expressed in the road to developed data standards are the level of willingness of tourism organisations to embrace them and their leadership and ability drive their development. Even if these obstacles were overcome, problems posed by the subtleties of meaning in language would continue to hinder progress. Still management and industry respondents expressed more positive outlooks, whilst technology respondents remained sceptical.

4.3 The way forward - challenges and optimism

Respondents were frustrated with the state of supply side interoperability, raising widespread management and technology concerns about the way forward. They repeatedly pointed out the practical challenges of organisation resource allocation, coding, standards and inadequate technology systems:
If you want to be successful you will have to allocate resources (industry organisation respondent).
The scarcity of guiding signposts did not auger well for management decisions perhaps indicating collective uncertainty and anxiety:
How do you monetise this? Much work is needed prior to monetising this (industry organisation respondent).
That groups and industry sector organisations are collectively taking initiatives to move forward was indicated by an industry organisation respondent who also pointed to the continuing evolution of the tourism and technology relationship:
Different industry sectors need to do this by creating landmarks...This is no different to what has gone on before

Given many prescient comments, there seemed to be no doubt that the new Web era is imminent and heralds benefits for supply and demand side interoperability. Although management and technical challenges can impede progress and delay realisation in the short or medium term, the tourism industry has a history of adopting technology (Poon, 1994) so may well do so in the Semantic era.

5 Conclusions

In terms of Semantic challenges and potential for tourism, there are several key requirements which need to be introduced, the foremost being data standards, if wide adoption is to be realised. There are already some initiatives for ontologies to provide common concepts and language: they need to be extended. Innovative applications need further development beyond their early stage: further research is needed to examine what succeeds and why. There are challenges to be resolved regarding management ability to respond to the evolving business environment; there are also unknown consumer responses. The complex and somewhat inadequate semantic technologies seemed to cast doubts on its imminent introduction.

There is a change in the tourism power structure of networks which the consumers may/not drive; if not, then new intermediaries will dominate – both the literature and qualitative date were ambivalent on this. Who will benefit? Those capable will network with enhanced distribution, maximising network interoperability. Who will lose? Static systems that are left out of networks will. There are management challenges as the nature of the relationship of tourism and technology continues to evolve. The question remains - will they be realised?

Even so, pressure from tourism industry sector organisations, enterprise managers and customers may drive development of semantic technology and encourage a greater level of adoption by enterprises. From a historic perspective this is just another era unfolding in the ongoing relationship of technology and tourism.

References

Berners-Lee, T., Hendler, J. and Lassila, O. (2001) The Semantic Web. *Scientific American* (284): 34 - 43

Brown, S. and Eisenhardt, K. (1997) The Art of Continuous Change: Linking Complexity Theory and Time-Paced Evolution in Relentlessly Shifting Organizations, *Administrative Science Quarterly* (42): 1-34

Buhalis, D. and O'Connor, P. (2005) Information Communication Technology Revolutionizing Tourism *Tourism Recreation Research* 30(3), 7-16

Buhalis, D. and Law, R. (2008) Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism Management* 29 (2008) 609–623

Cardoso, J. and Lange, (2007) A framework for assessing strategies and technologies for dynamic packaging applications in e-tourism. *Information Technology & Tourism*,

Dale, C. (2003). "The competitive networks of tourism e-mediaries: New strategies, new advantages." Journal of Vacation Marketing 9(2): 109.

Dogac A, Kabak Y, Laleci G, Sinir S, Yildz A, Kirbas S, Gurcan Y. (2004) Semantically Enriched Web Services for the Travel Industry, *SIGMOD Record*, 33 (3) 21-27.

Fodor, O. & Werthner, H. (2004-5) Harmonise: A Step Toward an Interoperable E-Tourism Marketplace, *International Journal of Electronic Commerce*, 9 (2), 11-39.

Gutierrez, C., Hurtado, C. Mendelzon, A. and Perez, J. (2011) "Foundations of Semantic Web databases." *Journal of Computer and System Sciences 77*(3): 520-541.

Hjalager, A-M (2010) A review of innovation research in tourism.*Tourism Management* 31 (2010) 1–12

Jones, A., Ivezic, N. & Gruninger, M. (2001) Toward Self-Integrating Software Applications for Supply Chain Management. *Information Systems Frontiers*.Vol 3 Issue 4 pp 403- 412 Springer Netherlands

Kanellopoulos, D. and Panagopoulos, A. (2008) Exploiting tourism destinations' knowledge in an RDF-based P2P network *Journal of Network and Computer Applications* 31. 179–200

Kauffman, K., Li, T. and van Heck, E. (2010) Business Network-Based Value Creation in Electronic Commerce *International Journal of Electronic Commerce* / Fall, Vol. 15, No. 1, pp. 113–143.

Lamsfus, C., Grün, Alzua-Sorzabal, A. and Werthner, H. (2009) Context-based semantic matchmaking to enhance tourists' experiences. *European Journal of Informatics Professional*. XI. 2. April pp 14-21

Law, Leung and Buhalis, D. (2009) 'Information Technology Applications In Hospitality and Tourism: A Review Of Publications From 2005 TO 2007', *Journal of Travel & Tourism Marketing*, 26: 5, 599 — 62: DOI: 10.1080/10548400903163160

Mistilis, N. and D'Ambra, J. (2008) 'The Visitor Experience and Perception of Information Quality at the Sydney Visitor Information Centre'. *The Journal of Travel & Tourism Marketing*. Vol 24 issue 1. pp 35 – 46

Murugesan, S. (2010) Web X.0: A Road Map, *Handbook of Research on Web 2.0, 3.0 and X.0: Technologies, Business and Social Applications,"* Murugesan, S. (Editor), Chapter 1 pp1-11. ISBN13: 9781605663845,

Novelli, M., Schmitz, B. and Spencer, T. (2006) Networks, clusters and innovation in tourism: A UK experience *Tourism Management* 27, 1141–1152

Poon, A. (1994) The 'new tourism' revolution. *Tourism Management Volume* 15 (2) 91-92

Prantner, K., Ding, Y.m, Luger, M., Yan, Z. and Herzog, C. (2007) Tourism Ontology and Semantic Management System: State-Of-The-Arts Analysis. IADIS International Conference WWW/Internet http://www.iadis.net/dl/final_uploads/200712C070.pdf

Ritchie, J. and Lewis, J. (2003) *Qualitative Research Practice.* Sage. London

Robertson, P. (2011) An Assessment of Collaborative Governance in a Network for Sustainable Tourism: The Case of Rede Turis. *International Journal of Public Administration*,Volume 34, Issue 5, Pages 279 - 290 DOI: 10.1080/01900692.2010.550078

Siau, K & Tian, Y. (2004). Supply Chains Integration: Architecture and Enabling Technologies. *The Journal of Computer Information Systems*, 44(3), 67-72

Stebbins, R. (2001) *Exploratory Research in the Social Sciences.* Thousand Oaks: SAGE Publications, Inc..

Wang, W., Zeng, G. and Tang, D. (2011) Bayesian intelligent semantic mashup for tourism. *Concurrency and Computation: Practice and Experience*, 23: 850–862. doi: 10.1002/cpe.1676

Williams, A. and Shaw. (2011) 'Internationalization and innovation in **tourism**'. *Annals of* ***Tourism*** *Research*, 38 (1), pp. 27-51

Xiang, Z., Gretzel, U. and Fesenmaier, D. (2009) Representation of Tourism on the Internet. *Journal of Travel Research*, Vol. 47 Issue 4, p440-453

Zhang, X. (2009) A Semantic Grid Oriented to E-Tourism, *CloudCom 2009, LNCS 5931*, pp. 485-496, M.G. Jaatun, G. Zhao, and C. Rong (Eds.), Springer-Verlag Berlin Heidelberg

Zhang, Song, H. and Huang (2009) Tourism supply chain management: A new research agenda. *Tourism Management*, Volume 30, Issue 3, Pp 345-358