Hotel Websites, Web 2.0, Web 3.0 and Online Direct Marketing: The Case of Austria

Ioannis Stavrakantonakis, Ioan Toma, Anna Fensel and Dieter Fensel

Abstract Direct communication with customers in order to increase sales has become one of the most important marketing methods used by small, medium and large hotels alike. With the rapid development of ICT technologies, including the Internet, Web, and recently Web 2.0 and 3.0, the number of channels in which hotels can interact directly with customers has grown even larger. Being visible on all these channels and using these technologies has now become a requirement if effective marketing and massive direct sales are to be achieved. In this chapter, we perform a rigorous empirical analysis of the advances towards the employment of Web 2.0 and 3.0 technologies in the tourism domain. We begin by presenting our methodology, including criteria and evaluation metrics, and follow by analysing the uptake of Web 2.0 and 3.0 technologies for Austrian hotels. As this chapter demonstrates, despite the benefits of new Web technology for online marketing, the hotels in Austria are not using these technologies and do not follow the online developments. Since employing their use is a relatively cheap undertaking, a severe competence gap seems to emerge either directly in the touristic service industry, or in the industry providing them with their on-line presence.

Keywords Social Web · Semantic Web · Semantic annotations · Austrian hotel business

I. Stavrakantonakis (✉) · I. Toma · A. Fensel · D. Fensel
Semantic Technology Institute (STI), Innsbruck University of Innsbruck, Innsbruck, Austria
e-mail: ioannis.stavrakantonakis@sti2.at

I. Toma
e-mail: ioan.toma@sti2.at

A. Fensel
e-mail: anna.fensel@sti2.at

D. Fensel
e-mail: dieter.fensel@sti2.at

1 Introduction

Tourism is one of the core sectors of the Austrian economy, accounting for around
9 % of Austria's GDP (Wikipedia 2013). In 2007, Austria ranked 9th worldwide in
international tourism receipts, with a total of 18.9 billion US\$ (Wikipedia 2013),
and a total of more than 121 million overnight stays (Statistik Austria 2013). By
2012, an 8 % increase can be noticed with a little over 131 million overnight stays
(Statistik Austria 2013). Tourism has comparable influence on the economic and
social situation of Austria. Therefore, possessing all of the required key compe-
tencies necessary for running this economic sector is of great strategic importance
for Austria. Losing a cornerstone of this process could lead to a significant loss of
income and revenue.

The rapid advance of ICT technologies, and their increasing importance in the
tourism domain, has led to an exponential growth in online communication
opportunities. This includes the websites of the touristic service providers, other
websites (e.g., the website of the Tourist Board, Wikipedia, etc.), email, email
lists, chat, instant messaging, news, message boards, internet fora, blogs, mi-
croblogs, podcasts, photo and video sharing, collaborative tagging, social net-
works, mobile platforms, recommendation sites, booking channels, destination
management websites, search engines and meta-search engines, rich snippets,
semantic annotations, etc. In addition, high visibility and ranking in popular search
engines such as Google, Yahoo or Bing is also a key goal that any online presence
must not neglect. Known as Search Engine Optimization (SEO), this goal can be
achieved mainly by having accurate, relevant, up-to-date annotations in the for-
mats and vocabularies that search engines understand and use as part of their
internal search mechanisms.

All channels mentioned above are important means of implementing successful
online marketing and driving sales. These are no longer an option for touristic
service providers, but a necessity. Maintaining competence and competitiveness in
online marketing for tourist service providers, coupled with ubiquitous access,
interaction and book ability of the services they provide through mobile devices,
may therefore be key to the future prosperity of Austria. This chapter analyses the
online presence and marketing of a large, representative set of Austrian touristic
service providers ranging from small to large hotels and hotel chains. We perform
an empirical analysis of the usage pattern of the Web, particularly, Web 2.0—
Social Web—and Web 3.0—Semantic Web—technology for their online pres-
ence. We begin by introducing our work through research questions and presenting
related studies (Sect. 2). Next, we introduce the methodology setup for the analysis
(Sect. 3). We then analyse the online presence of touristic service providers based
on our results (Sect. 4), and finally conclude in Sect. 5.

2 Related Work and Research Questions

The application of internet technologies in the tourism domain has been the topic of several scientific studies which focus specifically on the adoption of classical internet/web technologies by hotels. For example, Murphy et al. (2006) evaluates the internet adoption by analysing 200 Swiss hotel websites and email responses. Automatic techniques based on complementary multivariate and artificial neural networks help in the evaluation and classification process. Chan and Law (2006) focuses on the usability of websites, effectiveness of interface, amount of information, ease of navigation and user friendliness of a dozen hotels from Hong Kong. Schegg et al. (2002) explore a small slice of the soaring travel marketplaces, investigating how 125 Swiss hotels use Web-based marketing tools. The study focuses on classical Web technologies and finds that most hotel websites broadcast static information but provide limited transactional functions (e.g., rates, offers, etc.). A more recent study coming from the same group (Schegg et al. 2008) analyses the application of Web 2.0 technologies for more than 3,000 tourism businesses. At the time of the study, the authors concluded that most of the tourism enterprises were in the early stages of applying Web 2.0 concepts to their businesses. Scaglione et al. (2013) revisit the results of Schegg et al. (2008), showing the evolution of this adoption over the period 2008–2012. The set of criteria used in Schegg et al. (2008) is adopted and complemented in light of the technological progress that occured during the period previously mentioned. A total of 4,700 tourism websites in Europe have been analysed by primarily looking at the general use of Web 2.0 and Web 3.0 technologies. The authors conclude that while for some techniques the take-off phase is finished, for other new techniques, such as RDF, the tourism sector is still at the very beginning of the adoption process. The level of infiltration of Web 2.0 technologies in the tourism domain is the theme of several studies such as Ayeh et al. (2012) for the Hong Kong hospitality industry, or Stankov et al. (2010) and Shao et al. (2012) for destination marketing organizations (DMOs).

Other studies such as Grüter et al. (2013) focus more on analyzing the multimedia aspects of the hotels' online presence and the potential benefits of using these multimedia applications. Using the Swiss hotel business as an example, the study shows that while pictures appear on nearly every website, the use of videos and three-dimensional (3D) presentation formats is much less common. They stress that advances in multimedia formats such as videos and 3D presentations on hotel websites could have a positive impact on brand awareness, image and confidence in the hotel on the part of the customer.

Despite a large volume of studies available in the literature, none of them provide an extensive analysis of uptake of the Web technologies (Web 1.0, 2.0 and 3.0) in the tourism domain considering: (1) a large number of criteria from these areas and (2) including thousands of hotels, as in our study. Therefore, the research questions of this study are formulated as follows:

1. *To what extent do hotels in Austria exploit the Web 2.0 and 3.0 solutions?*

2. *Is the Content Management System related to the integration of Web 3.0 technologies?*
3. *Is there any correlation between the hotels' star rating with the usage of Web 2.0 and 3.0 technologies?*

In this respect, by investigating these questions, we aim to identify the adoption of Semantic Web annotations and Social Web channels and discover correlations between these, and the hotels' infrastructure. These results will help to observe and analyse any correlation between the current situation in online direct marketing and the extent that hotels are using the various technologies and means of communication on the Web through their websites. If the results indicate vast room for improvement, it would imply that online direct marketing could also bring better results and return on investment (ROI) by improving the situation and strengthening the online weaknesses.

3 Methodology

In this section, we introduce the methodology that we followed which comprises of: (a) the dataset specification as well as any technical decisions; and (b) the selection and evaluation criteria for the analysis.

We began by collecting data from various sources (Google Places, TripAdvisor) about the hotels that exist in Austria. Our dataset of hotels includes URLs, star ratings and geo-coordinates. The sample was selected randomly. Over 2,000 hotels were selected (i.e., 2,155 to be precise), and 75 % of the hotels selected have 3–5 stars. As shown in Fig. 1, they are distributed within the Austrian borders.

Our research workflow continues with the specification of the criteria with which we used to evaluate the dataset. Thus, our criteria can be divided into two main categories, namely the Web 2.0 direction and the Web 3.0 direction. In our case, the Web 2.0 represents the media that is mainly described as the 'Social Web'. Furthermore, the Web 3.0 represents the Semantic Web technologies that a hotel could apply on its website in order to gain more visibility on the web. Particularly, we focus on the existence of semantic annotations that could bring a hotel's website into a better position in search results of the major search engines (e.g., Google, Bing, Yahoo!, etc.) and with a richer presentation among the results by exploiting the opportunities that exist in the user interfaces of the search engines, such as in the case of Google with the Google Rich Snippets (Steiner et al. 2010).

The next step towards gathering the hotel websites' data was the implementation of a Web crawler that would access the hotels' websites and extract information about the criteria that we have specified and which are demonstrated in Sect. 3.1. Therefore, we built a Web crawler in Python based on a well-known open source high-level screen scraping and web crawling framework, i.e., Scrapy
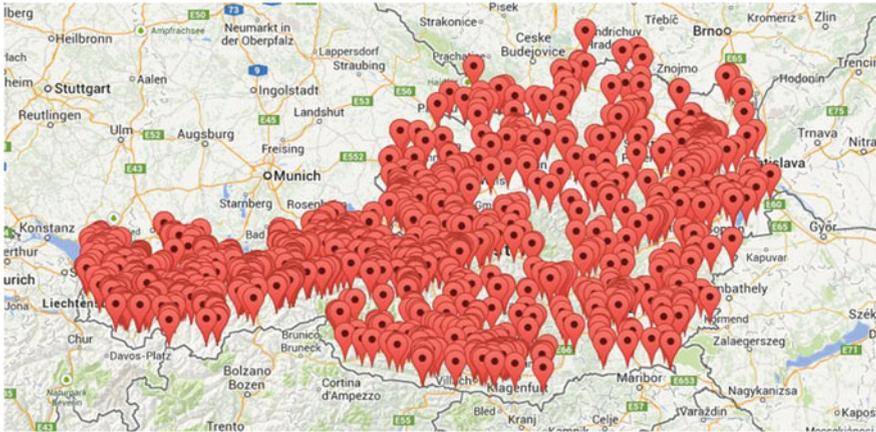
Fig. 1 Geographic distribution of hotels in our dataset

(www.scrapy.org). For the storage of the crawled data, we preferred to use a NoSQL database, namely the MongoDB (www.mongodb.org), and store JSON objects with the distilled information from the various pages. It is worth mentioning that we set a limitation of depth for the crawler, i.e. three levels from the starting page, in order to achieve better performance. However, this limitation cannot affect the accuracy of our results as our criteria are met at the very first pages of the websites.

3.1 Criteria

The aim of our Web analysis is to unveil insights hidden at the HTML mark-up of the hotel websites by crawling the URLs of our dataset. We applied the criteria during the crawling process, rather than downloading the whole website locally. Our criteria cover the Web 2.0 major channels and the Web 3.0 technologies, namely the formats and vocabularies for semantic annotations, as shown in Tables 1 and 2, respectively.

Table 1 presents the Web 2.0 dimensions that we looked into during our analysis of the collected data, while Table 2 presents the Web 3.0 formats and vocabularies from which hotels could benefit.

As we were interested in the uptake of the Web 3.0, we focused on the existence or nonexistence of technologies used for adding semantic annotations (i.e., machine-readable metadata) to websites. In this respect, we distinguish between formats and vocabularies within our criteria. Formats refer to the technologies with which a web developer could add semantic annotations to a website. As shown in Table 2, there are three main formats, as broadly adopted options, namely microformats, RDFa and microdata.

Table 1 Web 2.0 criteria

Channel	Description	Crawling method in HTML
CMS	Content management systems are used to publish and modify the content of websites without dealing with the HTML code	Within the HTML head tag: < meta name = "generator" content = "*CMS Name*">
Facebook	Facebook offers social plugins and widgets for websites, adding a Social Web layer to the website	iframes, links, official embed HTML code
Flickr	Photo sharing web application. Mostly used by hotels to share high-quality photos from events and places	iframes, links, official embed HTML code
Foursquare	Location-based Social Network (LBSN) used by tourism businesses to keep in touch with guests during their stay	iframes, links, official embed HTML code
Google+	G+ is meant to be the social spine of the Google products (Seeking Alpha 2012), as it adds a social and sharing layer at every Google product	iframes, links, official embed HTML code
Instagram	Photo sharing mobile application. Mostly used by individuals to share experiences at real time	iframes, links, official embed HTML code
Pinterest	Pinterest is a web application used to collect and organize online entities (text, photos, links, etc.) about a topic, like places	iframes, links, official embed HTML code
RSS	Really simple syndication (RSS) keeps subscribed users of the RSS feed updated about content changes in a website	RSS links
Twitter	The most popular social network for micro-blogging that could help businesses connect with customers in real time	Widget, links, official embed HTML code
Vimeo, YouTube	Video sharing platforms. Hotels could benefit by sharing videos about their premises and events while also embedding those in their website	Widget, links, official embed HTML code
Tripadvisor, HolidayCheck	Review sites, where guests publish their experiences are considered to have very important impact to the conversion rates (lookers to bookers)	Widget, links, official embed HTML code

Table 2 Web 3.0 criteria

	Description	Crawling method in HTML
Formats		
Microformats v1	Semantic annotations re-using HTML tags in order to add metadata to the content	Vocabulary terms in HTML class attributes
Microformats v2	New version of the microformats standard. Introduces prefixes for which class names are used for microformats. Improves implementation and disambiguates the annotations from simple class names	Vocabulary terms in HTML class attributes
Microdata	An attempt to enable developers to annotate with semantics HTML items in a much simpler and effective way than microformats and RDFa	HTML tag attributes: itemscopre, itemtype, itemprop
RDFa	Resource description framework in attributes proposes attribute level extensions to XHTML in order to enable semantics to the presented web content	HTML tag attributes of the specification (e.g., vocab)
Vocabularies		
Schema.org	Joint effort of major search engines (Google, Bing, Yahoo!, Yandex) to provide a shared collection of schemas about various domains (http://schema.org)	HTML tag attribute: itemtype = "http://schema.org/"
Open Graph protocol	Enables any web page to become a rich object in a social graph. Mainly used by Facebook (http://ogp.me)	og: namespace (e.g., og:title)

Microformats (abbreviated as μF) are conventions used to describe a specific type of information on a web page (e.g., people, organizations, locations, etc.). In general, microformats overload the class attribute in the HTML tags to assign descriptive names to entities. They can be realised as format and vocabulary combined. The second version of microformats add prefixes to the terms in order to understand which class names are used by microformats (e.g., *u-photo* is used to annotate the URL of a photo).

Resource Description Framework in Attributes (*RDFa*) provide a set of mark-up attributes to augment Web page content with semantic annotations. RDFa are based on attributes by re-using HTML tags and defining namespaces in the XHTML to assign types and names to entities and properties. None of the attributes introduced or used by RDFa have any effect on the rendering of the web pages.

Microdata specification is similar to microformats, but introduces new HTML tag attributes (i.e. itemscope, itemprop, itemtype, etc.) that can host terms from any vocabulary. It is supported by schema.org and is part of the HTML 5 specification. In comparison to the aforementioned formats, we could say it combines ease of use, effectiveness and flexibility, all of which make it a great option for semantic annotations.

The results from our analysis, presented in Sect. 4, aim to answer our research questions.

4 Analysis Results

Following the methodology presented in Sect. 3, we collected data from more than 2,000 hotels in Austria, and from which we have gained some useful insight that will answer the research questions which we posed at the beginning of this chapter. The distribution of CMS systems is shown in Fig. 2 and Table 3. The percentages of Table 3 refer to the occurrences of the second column divided by the total number of websites with a CMS system (i.e., 946 out of 2,155 websites).

Studying the data in Table 3, it is evident that there is great diversity in the CMS decisions by the web development agencies. It is worth mentioning that we have information about the type of the CMS system for roughly 44 % of the dataset. The main reason is the way that this piece of information is extracted. Thus, we are able to capture only those CMS systems that specify their name explicitly within the HTML metadata fields. However, according to our knowledge, all the major CMS systems do so, which means that for the rest of the hotels we could say that they are not using any of the well-known and well-designed CMS systems or they do not use a CMS system at all.

The most popular CMS seems to be TYPO3 with 44 % frequency among the CMS systems retrieved for our dataset. In second place is the well-known Joomla! CMS and third is WordPress, which is mostly used for blogging. It is worth mentioning the fact that we observed the usage of systems, like Adobe GoLive,[1] which have not been available for the last 5 years. This fact implies that there are a significant number of hotels that have stopped investing in their Web presence as far as their website is concerned. It is common practice to update the website from time to time in order to keep the design and functionality aligned to the current paradigms and competitors' level.

Figure 2 offers interesting insight into the diversity. Thus, 27 % of hotels use other CMS systems rather than one of the popular (i.e., Drupal, Joomla!, TYPO3, WordPress, MS FrontPage). We observed 87 different CMS systems within the aforementioned portion of 27 %.

Regarding the adoption of Web 2.0 technologies, we measured the usage of the major Social Web channels (i.e., Facebook, Flickr, Google+, HolidayCheck, Instagram, Pinterest, RSS, Tripadvisor, Twitter, Vimeo, YouTube) within the websites of the hotels. We found that 53 % of the hotels in the dataset exploit the opportunities offered by the Web 2.0 by adding links on their websites to one channel or more of their Social Web profiles. However, this does not imply that the remaining 47 % of hotels are not present on the observed channels. It only highlights the fact that they have not connected those channels with their website. Thus, it can be the case that a hotel maintains a Facebook page but does not point the website users to it.

---

[1] *Adobe stopped development and sales of Adobe GoLive 9 software on April 28, 2008*, https://www.adobe.com/products/golive/.

Table 3 Distribution of CMS systems

CMS	Occurrences	Percentage (%)	Comments
TYPO3	424	44.82	
Joomla!	131	13.84	
WordPress	68	7.18	
Microsoft FrontPage	56	5.91	
Drupal	15	1.58	Drupal 7 only
Condeon	25	2.64	
Web to date	24	2.53	
ez Pubish	26	1.69	
CMS Contenindo	12	1.27	
Adobe GoLive	8	0.84	Unsupported since 2008

Fig. 2 Distribution of content management systems (CMS)
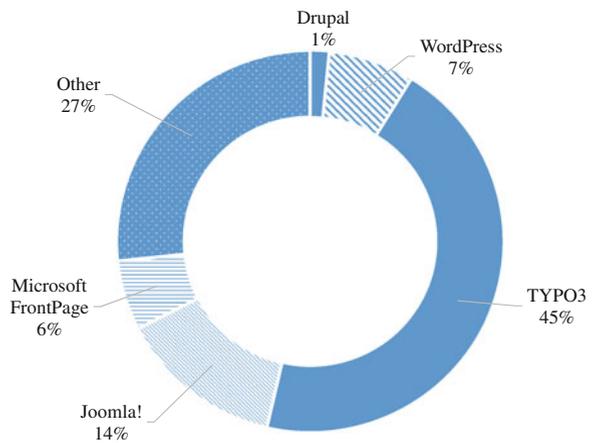
Furthermore, 68 % of the active Web 2.0 hotel websites (roughly 1,150), link their fan page or account to Facebook, while 48 % link to HolidayCheck. The potential of Twitter and Google+ seems to be realised by 14–15 % of the hotels in the dataset, while 25 % exploit YouTube and RSS as shown in Fig. 3. We measured more than 600 hotels linking to at least one review site.

The next step includes the analysis of the Web 3.0 (also known as Semantic Web) technologies uptake by the hotels in our dataset. The results in this dimension prove that most of the hotels completely ignore the existence of technologies that could enrich the website content with high level metadata and give machine readable meaning to the presented information. As shown in the pie chart of Fig. 4, only 5 % of websites employ some Semantic Web technologies, while the rest seem to ignore the potential of adding semantics to their websites.

Other than not being aware of the potential benefits and advantages the Web 3.0 can offer, another possible reason for lack of uptake could be the delays in the alignment of web development agencies with state of the art technologies in the field, or the hesitation to adopt Semantic Web technologies due to their
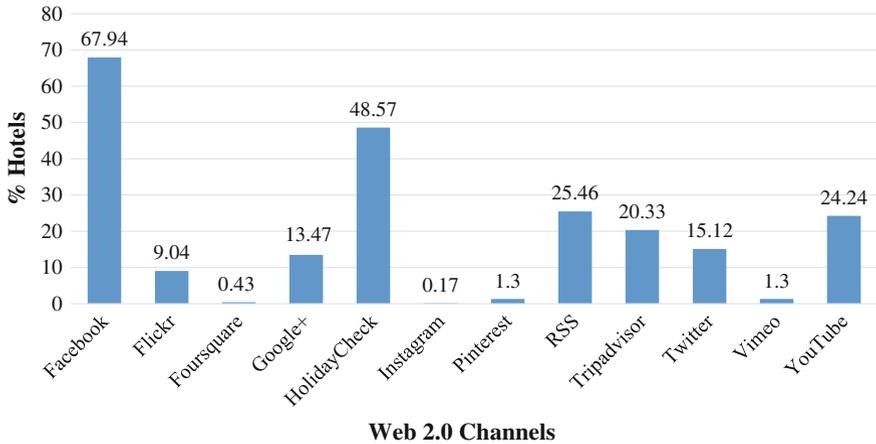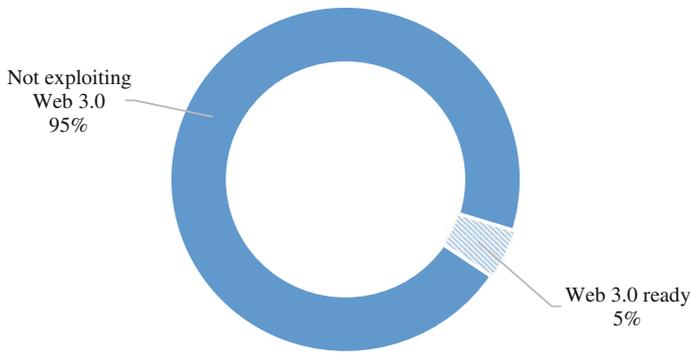
Fig. 3 Usage of Web 2.0 channels

Fig. 4 Uptake of Web 3.0 technologies in the Austrian accommodation sector

complexity. However, semantic annotation technologies are already mainstream and widely applied on the Web.

Moreover, it is interesting to examine the distribution of the CMS systems in conjunction with the use of semantics as described in the second research question presented in Sect. 2. In this respect, Fig. 5 depicts the usage of CMS systems among the hotels which adopt Web 3.0 technologies, i.e. 5 % of the hotels in our dataset according to Fig. 4. The presented figures are related to the number of hotels that use a CMS, i.e. as mentioned earlier, 44 % of the hotels in our dataset.

In the case of "Web 3.0 ready" hotel websites, WordPress seems to hold the first position, while TYPO3 is in second. Thus, the distribution of Web 3.0 integration within CMSs does not follow the CMS usage distribution. It is remarkable that we did not observe any occurrence of Drupal 7 within this subset of hotels, as
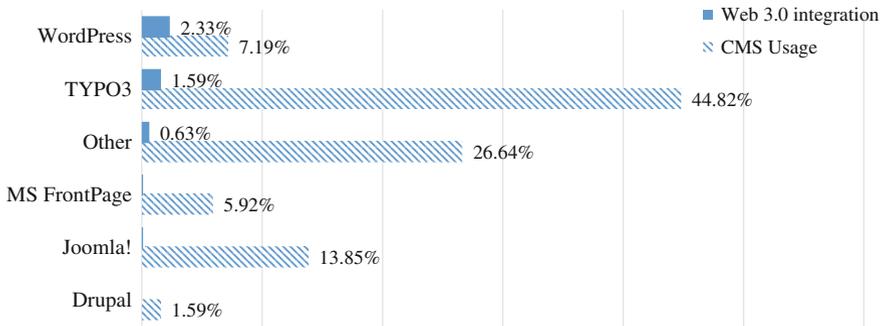
Fig. 5 Correlation of CMS systems and Web 3.0 integration

Fig. 6 Correlation of hotels applying Web 2.0–3.0 with number of stars
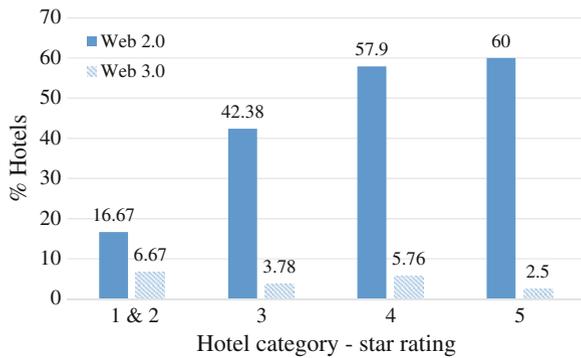
Drupal 7 is meant to be the most semantically friendly and supportive CMS at the time of writing, and to the best of our knowledge (Clark 2011).

Last but not least, we examined the correlation between the number of stars and the adoption of Web 2.0 and Web 3.0 technologies (3rd research question). Therefore, the diagram of Fig. 6 depicts the correlation among these variables. Hotels with one and two stars seem to be less responsive to the development of new web technologies and paradigms, like Web 2.0 and 3.0. This result can be expected, as the development of the onsite quality of services is of highest importance and takes priority over online presence, especially considering the limited budget available in most of the cases.

Furthermore, Fig. 6 demonstrates a proportional relationship between the number of stars and the percentage of hotels per category that link their websites with their Web 2.0 channels. Specifically, 42 % of the websites of three-star hotels make use of Web 2.0 technologies including RSS feeds, video and photo sharing services, like Flickr and YouTube, and Social Web channels, like Facebook and Twitter. While, an average of 59 % of four and five-star hotels make use of the Web 2.0 channels, the corresponding figures for the Web 3.0 adoption do not exceed 6 % of the websites in each category, which are not significant enough to

compare. The same situation more or less is observed in the remaining cases, meaning hotels rated with one and two stars.

Concluding from the data in Fig. 6, we can deduce that semantic annotations have a long way to go in order to bring the hotel websites to their full potential in terms of visibility on the Web search engines.

5 Conclusion

This chapter analyses the use of the new Web 2.0 and Web 3.0 technologies on a set of hotel websites in Austria. The aim of this work has been to explore the uptake of the related technologies, in order to identify the status quo of direct online marketing in the Austrian tourism domain, as online visibility is the cornerstone of advertising in today's world. In this respect, our approach has included crawling and extracting the information on the current use of modern technologies, such as Social Media (Web 2.0) and Semantic Web markup (Web 3.0), on the hotels' websites.

The analysis outcome has shown the current sparse usage of such technologies by the hotels' websites. The slow technology take-up is hindered by the technical (e.g., difficult integration due to the usage of heterogeneous CMSs within in the sector) and educational factors (e.g., knowledge about the new technologies and understanding of their advantages). Thus, there is large space for improvement regarding the visibility of the hotels on the current Web ecosystem. Failure to act on improving the use of the latest Web paradigms would keep the visibility of the hotels moderately low in search engine algorithms, as they are becoming more and more sophisticated. This fact prevents online direct marketing to be monetised via direct bookings and limits the impact of Online Travel Agencies (OTA) on the budget of the hoteliers.

Future steps have already been considered, including the expansion of the geographical region in which our criteria and crawling are applied in order to cover the whole of Europe. Furthermore, the criteria will be re-worked in order not to overlook any means of dissemination and added value within the website.

References

Ayeh, J. K., Leung, D., Au, N., & Law, R (2012). Perceptions and strategies of hospitality and tourism practitioners on social media: An exploratory study. In M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and Communication Technologies in Tourism 2012* (pp. 1–12). New York: Springer.

Chan, S., & Law, R (2006) Automatic website evaluations: The case of hotels in Hong Kong. *Information Technology and Tourism, 8*(3–4), 255–269.

Clark, L. (2011, April 12). *The Semantic Web, Linked Data and Drupal, Part 1: Expose your data using RDF*. Retrieved from IBM-developerWorks: http://www.ibm.com/developerworks/library/wa-rdf/

Grüter, M., Schneider, V., & Myrach, T. (2013). Virtual experience on hotel websites: A web analysis. *Information and Communication Technologies in Tourism, 2013*, 400–411.

Murphy, J., Schegg, R., & Olaru, D. (2006). Investigating the evolution of hotel internet adoption. *Information Technology and Tourism, 8*, 161–177.

Scaglione, M., Schegg, R., & Trabichet, J.-P. (2013). Analysing the penetration of Web 2.0 in different tourism sectors from 2008 to 2012. *Information and Communication Technologies in Tourism, 2013*, 280–289.

Schegg, R., Liebrich, A., Scaglione, M., & Ahmad, S. F. (2008). An exploratory field study of Web 2.0 in Tourism. *Information and Communication Technologies in Tourism, 2008*, 152–163.

Schegg, R., Steiner, T., Frey, S., & Murphy, J. (2002). Benchmarks of web site design and marketing by Swiss hotels. *Information Technology and Tourism, 5*, 73–89.

Seeking Alpha. (2012). *Google call transcript.* Retrieved from http://seekingalpha.com/article/495351

Shao, J., Rodríguez, M. A., & Gretzel, U. (2012). Riding the social media wave: Strategies of DMOs who successfully engage in social media marketing. *Social Media in Travel, Tourism and Hospitality* (pp. 87–98) Farnham: Ashgate.

Stankov, U., Lazić, L., & Dragićević, V. (2010). The extent of use of basic Facebook user-generated content by the national tourism organizations in Europe. *European Journal of Tourism Research, 3*(2), 105–113.

Statistik Austria. (2013). *Ankunfts- und Nächtigungsstatistik.* Retrieved from http://www.statistik.at/web_de/statistiken/tourismus/beherbergung/ankuenfte_naechtigungen/index.html

Steiner, T., Troncy, R., & Hausenblas, M. (2010). How Google is using linked data today and vision for tomorrow. *FIA 2010, Linked Data in the Future Internet Assembly, Ghent, Belgium*, December 15–17, 2010.

Wikipedia. (2013). *Tourism in Austria.* Retrieved from http://en.wikipedia.org/wiki/Tourism_in_Austria