MAPPING FERATEL CONTENT TO SCHEMA.ORG IN THE TOURISMUSVERBAND INNSBRUCK PILOT

Ioan Toma, Corneliu Valentin Stanciu, Ioannis Stavrakantonakis
STI Innsbruck, University of Innsbruck,
Technikerstraße 21a, 6020 Innsbruck, Austria
firstname.lastname@sti2.at

2014-03-13

Semantic Technology Institute Innsbruck

Contents

1. Introduction

Having a good online marketing strategy results into higher online visibility and ultimately into increased sales. In order to achieve this goal it is however essential that multiple communication channels and technologies are properly used. In particular having annotations on the website that can be understood by search engines busts the online visibility and increases the chances that the website is in the search engines results to a relevant query. As part of our pilot with Tourismusverband Innsbruck (TVb)[1] we have annotated part of their website content with schema.org[2]. Some of this content is actually provided by feratel media technologies AG[3].

This document introduces the mapping of feratel content to schema.org in the TVb pilot. The reminder of this document is structured as follows. Section 2 briefly introduces schema.org, the vocabulary used for annotations. Section 3 describes which subset of schema.org was selected in order to annotate the content. Section 4 describes the actual implementation of the mappings and its limitations due to constraints we can't influence. Finally, Section 5 concludes the document and gives an outlook on how the mappings should be properly implemented.

2. schema.org

The main search engines, including Google, Yahoo!, Bing, announced in 2011 schema.org, a joint effort to create and support a common set of schemas for structured data markup on web pages (Google, 2011). Using schema.org, webmasters can markup their pages in ways recognized by major search providers. This brings a lot of benefits, enabling search engines to proper interpret content and, therefore, increasing the likelihood that the web pages are included in the search results for a related query. In a nutshell schema.org provide a rich vocabulary for talking about common things in the web that are of interest to search engines, such as people, places, reviews, recipes, offers and events. Schema.org is intended to help site owners and developers learn about structured data and improve how their sites appear in major search engines, as well as provide a one stop source for webmasters looking to add mark-up to their pages. One of the main benefits of schema.org lies in its simplicity. It includes schemas for a large number of concepts and domains, such as creative works (e.g. movies, music, TV, shows), events, places, products, organizations, lodging businesses, reviews, etc. Therefore, schema.org intends to be the de-facto source of vocabulary terms at the description of content in the Web.

3. Mappings

schema.org is a very large vocabulary counting hundreds of terms from multiple domains (see https://schema.org/docs/full.html for the full specification). Of course not all of the schema.org terms are relevant for the tourism domain. For the TVb website in particular the relevant schema.org terms are those that belong to the categories: Hotels, Food and Drink Establishments, Events, Trips, Place of Interest and News (Akbar, et al., 2014). As mentioned before only part of the content on TVb website is actually coming for feratel. More precisely this includes: Hotels, Apartments, Camping, Restaurants, Bars

---

[1] http://www.innsbruck.info/
[2] http://schema.org/
[3] http://www.feratel.at/

or Pubs, Cafes, Events and Sightseeing. Table 1 shows the mappings between content types coming from feratel to those in schema.org.

Table 1: Mapping feratel content to schema.org for TVb website

No	Concept in feratel	Type in schema.org
1	Hotel	schema:Hotel (rdfs:subClassOf schema:LodgingBusiness)
2	Apartment	schema:LodgingBusiness (rdfs:subClassOf schema:LocalBusiness)
3	Camping	schema:LodgingBusiness (rdfs:subClassOf schema:LocalBusiness)
4	Restaurant	schema:Restaurant (rdfs:subClassOf schema:FoodEstablishment)
5	Bar, Pub	schema:BarOrPub (rdfs:subClassOf schema:FoodEstablishment)
6	Café	schema:CafeOrCoffeeShop (rdfs:subClassOf schema:FoodEstablishment)
7	Event	schema:Event (rdfs:subClassOf schema:Thing)
8	Sightseeing	schema:TouristAttraction (rdfs:subClassOf schema:Place)

4. Implementation

Before we dive into the actual implementation of the mappings let us first shortly describe how the integration of TVb website with feratel is realized. This will help us understand later on the source of limitations of our current implementation. Figure 1 depicts the TVb website – feratel integration.
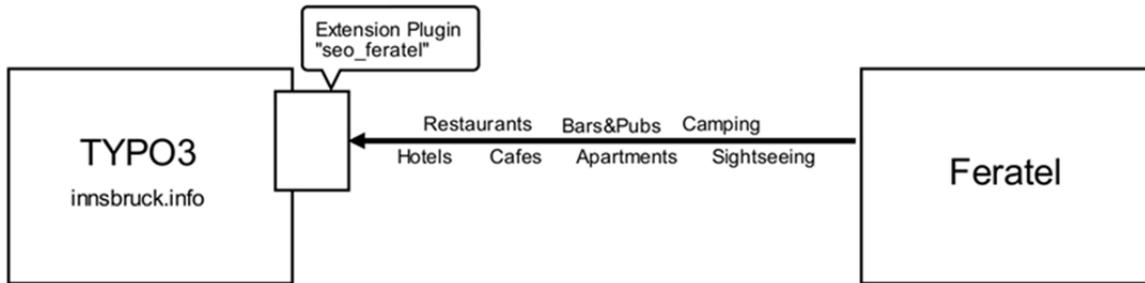
Figure 1: *TVb website – feratel integration – seo_feratel.*

The integration is implemented as a Typo3 extension plugin, referred as seo_feratel in the previous figure. TVb web site is built using the Typo3 content management system, and the Typo3 extensions seo_feratel is responsible for getting periodically content from feratel and shown it into the Typo3 website. The plugin is actually using a html template to structure the content - Hotels, Restaurants, etc. as shown in Figure 2.
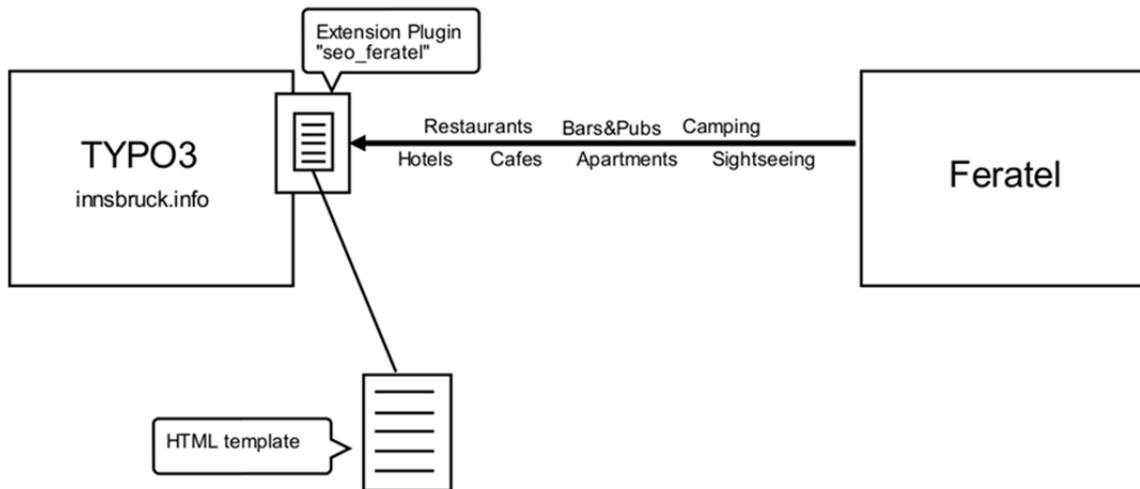

Figure 2: *TVb website – feratel integration - HTML template.*

In order to insert annotations according to *schema.org* into the TVb website we extended the HTML template as shown in Figure 3. Using the modified HTML template we insert annotations inside the HTML tags for:

- Hotels, Apartments and Campings using the following properties: Name, Email, URL, Map, PostalAddress (including streetAddress, addressCountry, postalCode, addressLocality, telephone and

faxNumber). For the following properties Logo, Image, Geo, Description, PriceRange, PaymentsMethods and OpeningHoursSpecification annotations are not created as this cannot be done by extending the HTML template, but rather requires changes in the code of the seo_feratel extension plugin.

- Restaurants, Cafes, Bars&Pubs, and Sightseeing using the following properties: Name, Map, PostalAddress (including streetAddress, addressCountry, postalCode, addressLocality, telephone and faxNumber). For the following properties Logo, URL, Email, Image, Geo, Description, PriceRange, PaymentsMethods and OpeningHoursSpecification annotations are not created as this cannot be done by extending the HTML template, but rather requires changes in the code of the seo_feratel extension plugin.
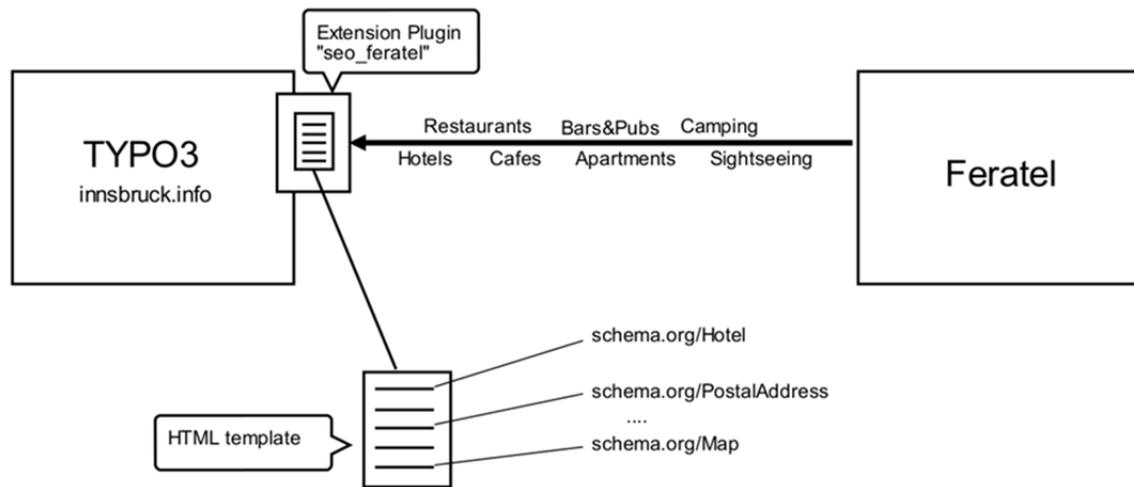

Figure 3: *TVb website – feratel integration - HTML template modified to insert schema.org annotations*

Using the modified HTML template in order to insert schema.org annotations has however a set of strong limitations. The seo_feratel plugin used the same HTML template for Hotel, Apartment and Camping and therefore is no possible to distinguish between them. To coop with this issue we use just the common top-level class of the three types, namely schema.org:LodgingBusiness but this leads to losing semantics. We have the same situation for Restaurants, Cafes, Bars&Pubs and Sightseeing where again only one HTML template is used. The solution here again was to use the top-level class i.e. schema.org:Plane with the same risk of losing semantics.

5. Conclusion and outlook

As mentioned before, our current implementation based on modified HTML templates has a set of strong limitations. This is due to the various limitations we described in Section 4. A much cleaner and scalable solution would be to have the annotation inserted not at the destination i.e. various destination web site (including TVb Innsbruck) but rather at the source i.e. feratel.
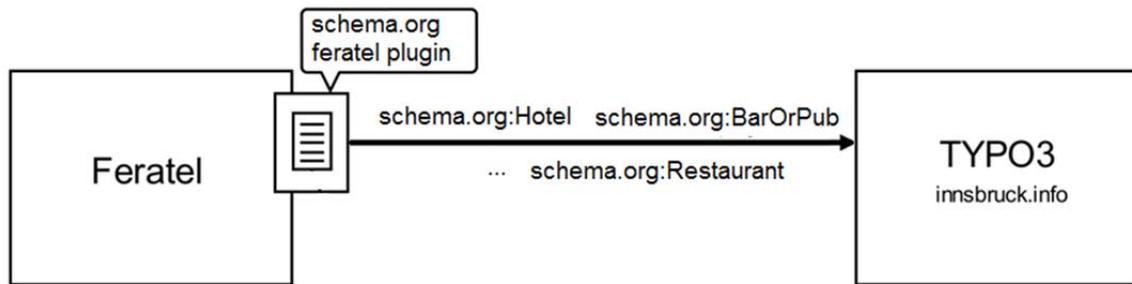
Figure 4: *schema.org feratel plugin*

This can be realized by developing a feratel plugin, which annotates content according to the mappings defined in Section 3. Figure 4 depicts the envisioned solution which can be implemented as a cooperation between feratel and STI Innsbruck.

References

Akbar, Z., Lasierra, N., & Tymaniuk, S (2014). Tourismusverband Ontology. January, 2014. Available at: http://oc.sti2.at/results/white-papers/tourismusverband-ontology

Google. (2011, June 02). Introducing schema.org: Search engines come together for a richer web. Retrieved from Webmaster Centeral Blog: http://googlewebmastercentral.blogspot.co.at/2011/06/introducing-schemaorg-search-engines.html