



## > Semantic Web Use Cases and Case Studies

### Case Study: CRUZAR – An application of semantic matchmaking for eTourism in the city of Zaragoza

---

Maria Jesús Fernández, City Council of Zaragoza and Antonio Campos, CTIC Foundation, Spain

August 2008



#### General Description

The web is a big showcase for cities that want to build their tourism industry. Nowadays, many tourists plan their trips in advance using the information that is available on web pages. Cities compete against each other to offer the most attractive and complete information and services through the tourism section of their web sites. However, this often leads to information-bloated and multimedia-rich web sites which are similar to digital versions of printed brochures. Everyone receives the same information, regardless of their interests. This is unlike when they visit a tourism office, and receive customized information and recommendations based on their profile and desires.

CRUZAR is a web application that uses expert knowledge (in the form of rules and ontologies) and a comprehensive repository of relevant data (instances) to build a custom route for each visitor profile. While some cities have a few predefined routes for the most common profiles (such as “Visit Zaragoza in just three days”), CRUZAR can potentially generate an infinite number of custom routes and it offers a much closer fit for each visitor's profile.

CRUZAR, is available at the web site of the city of Zaragoza (Spain). There are a number of reasons that make this city an excellent test bed for such a project. In the first place, Zaragoza has a high density of POIs (Points of Interest). Zaragoza is one of the biggest cities in Spain, and it enjoys a very dynamic cultural agenda, as well as frequent top-level sport events. Finally, the city council has extensive databases with all the aforesaid information, including content in five languages.

#### Technical details of the solution

The first challenge of our use case in the eTourism domain was to collect the required data from existing relational databases which are used to feed the content of the Official Website of Zaragoza. This data was split across the following four information silos:

- The CMS database which feeds the city council web site with pertinent information for tourists visiting Zaragoza. This database stores structured descriptions about the monuments and historical buildings of the city (more than 2.000 resources), restaurants, accommodation, green spaces, shopping areas and other relevant points of interest.
- A database which contains up-to-date information about upcoming cultural events and leisure activities.
- The city council web site which mainly displays photographs of the area.
- The IDEZar, which is a Geographic Information System hosted by the University of Zaragoza. It is designed to use REST web services to fetch maps as raster images and to compute the shortest path between two geo-referenced points of the city.

All these data sources are created and maintained by the Website Office of the City Council of Zaragoza. The information contained in these databases is transformed into RDF data using specific adapters. This process takes place regularly every time the databases are updated.

#### Representing knowledge of the domain

An ontology (see [Figure 1](#)) is used to organize the RDF data. The CRUZAR ontology captures information about three types of domain entities: 1) Zaragoza's tourism resources, mainly events and POIs, 2) user profiles to capture the visitors' preferences and their context, and 3) the route configuration. The conceptual structure of CRUZAR is based on the upper-ontology DOLCE.

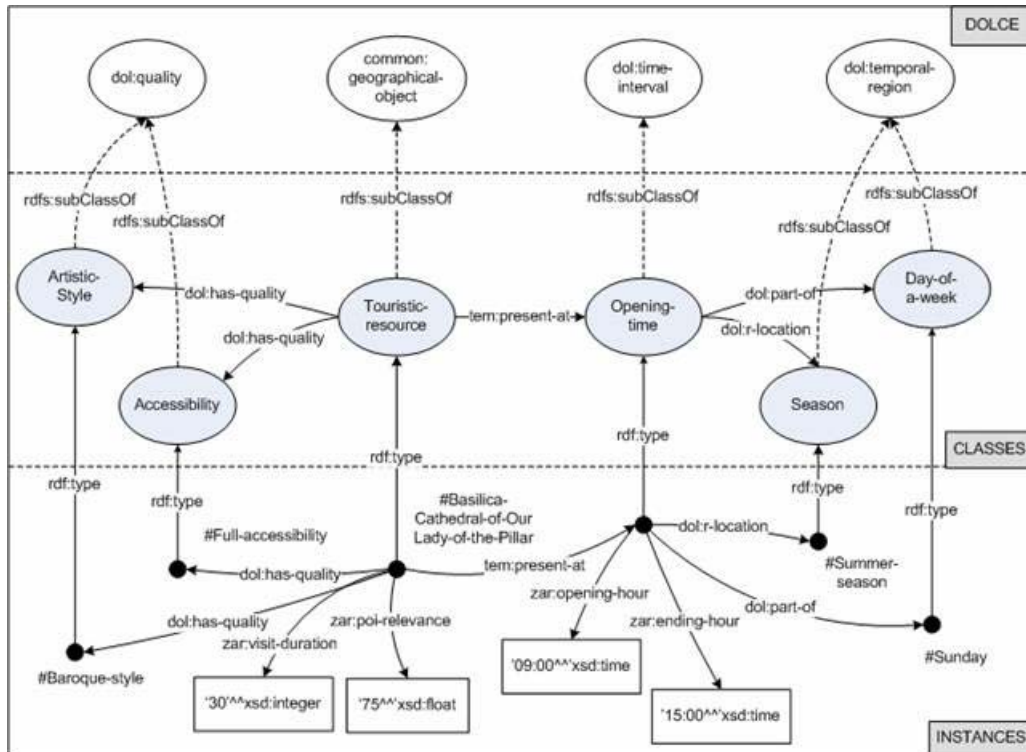


Figure 1: This diagram shows how the ontology describe Zaragoza's resources

Events and POIs are defined in terms of their intrinsic features: position, artistic style or date. Conversely, visitors' profiles contain information on their preferences and their trip: arrival date, composition of the group, preferred activities, etc. In order to match the local information with the preferences, a shared vocabulary is needed. The central concept of this intermediate vocabulary is "Interest". Visitors' preferences are translated to a set of "Interests", and POIs and events can attract people with certain "interests". This translation is captured as production rules, which are executed using the Jena rule engine. These rules are simple enough to be easily understood by the domain experts.

#### POIs ranking

All the POIs in Zaragoza are dynamically ranked to reflect their "subjective interest" according to the profile of each visitor. At the end of the matchmaking process, a numerical score is assigned to all POIs to quantify their anticipated level of interest. Initially every POI has a static score or relevance which was decided by the experts of the domain ("objective interest"). The semantic matchmaking process is executed individually for each POI, and its output is a calculated score for the resource ("subjective interest"). The value of this score depends on how many of the visitor's interests (derived from their profile) are fulfilled by each POI.

#### Route Planning

After all the candidate POIs have been sorted by their subjective interest, a planner algorithm is run in order to create the route. The main driving force of the algorithm is to balance the quantity and quality (interestingness) of the selected POIs and the distance. Visitors can influence this equilibrium. For instance, those who declare that they "don't mind walking" will get less condensed routes and longer distances between POIs. On the other hand, minimizing the distance between POIs can be important for physically handicapped visitors.

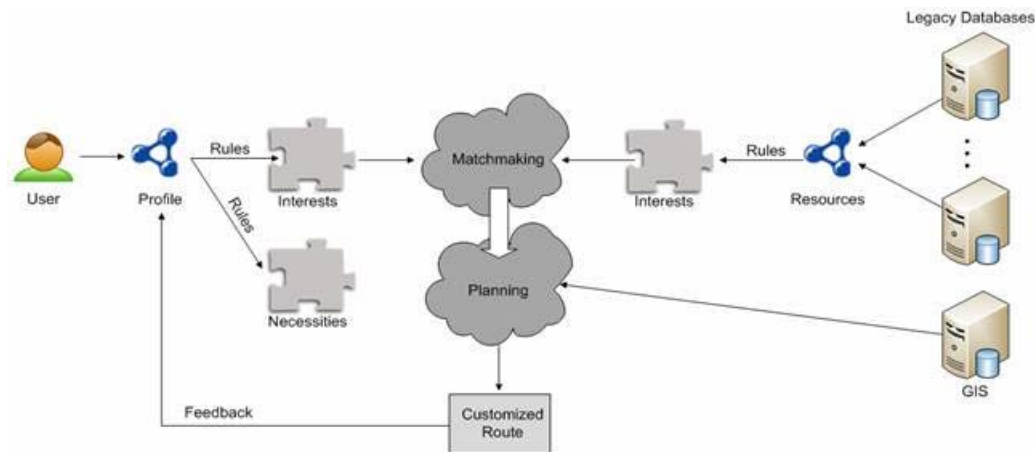


Figure 2: System architecture

### Route customization

The route proposed by the system is offered to the user using an accessible, information-rich interface that includes: the sequence of selected POIs, a tentative timetable for each day, a map highlighting the POIs, suggestions of other interesting places near the route, and two sets of recommended restaurants near the last POI of the route (one set proposing lunch and the other proposing dinner options). Complementary activities, such as events (music concerts, sport events, etc.) and shopping, are also suggested. Users can interact with the generated route in a number of ways. They can “ban” individual POIs, either because they do not like them or because they have already seen them during a previous visit to the city. There is also a “not now” option for each POI. When the “not now” button is clicked, the system modifies the plan so that a particular POI is moved to a different time slot. For visitors particularly interested in a certain POI, there is also an option to manually pick the POI and force its inclusion in the route, even if it was given a low score by the matchmaker.

### Key Benefits of Using Semantic Web Technology

CRUZAR is a semantic-based application created for the city of Zaragoza, which is already deployed for the city council web servers.

Semantic web technologies are put into practice:

- to integrate and to organize data from different sources
- to represent and to transform user profiles and tourism resources
- to capture all the information about the generated routes and their constraints.

CRUZAR implements a matchmaking algorithm between objects that are described in RDF, and it pipes the results to a planner algorithm. The result is a customized tourism route across the city, that is selected for each visitor profile. Although all the information was already available on the web, CRUZAR is able to merge it within a single user interface. Moreover, at the same time, it offers an innovative service for visitors to plan their trip in advance, exploiting expert knowledge. These features are often used as important examples to illustrate the promises of the Semantic Web.

Both the ontology (<http://www.zaragoza.es/ciudad/turismo/es/visitar/ontologia.htm>) and the RDF instances (<http://www.zaragoza.es/turruta/resource/list>) are available on the web. The CRUZAR service can be accessed online, and two screenshots are also available (see [first](#) and [second](#) screenshots).