FP7-SME-1
Project no. 262289

HARMOSEARCH

Harmonised Semantic Meta-Search in
Distributed Heterogeneous Databases

D3.3 V1.0

Extension to the Harmonise Ontology for metadata representation

Due date of deliverable: 2012-11-30
Actual submission date: 2012-12-05

Start date of project: 2010-12-01   Duration: 27 month

Project funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	X
PP	Restricted to other participants (including the Commission Services)	
RE	Restricted to a group specified by the Consortium (including the Commission Services)	
CO	Confidential, only for members of the Consortium (including the Commission Services)	

PROJECT ACRONYM: **HARMOSEARCH**

Project Title: Harmonised Semantic Meta-Search in Distributed Heterogeneous Databases

Grant Agreement: 262289

Starting date: December 2010   **Ending date:** February 2013

Deliverable Number: D3.3, Version 1.0

Title of the Deliverable: Extension to the Harmonise Ontology for metadata representation

Lead Beneficiary: PROMOTER

Task/WP related to the Deliverable: WP 3, Task 3.3

Type (Internal or Restricted or Public): Public

Author(s): Claudio Prandoni, Marlis Valentini, Albert Rainer, Inés Matres, David Faveur

Partner(s) Contributing: PROMOTER, TU-WIEN, SPK, AFIDIUM

Contractual Date of Delivery to the CEC: November 30 2012

Actual Date of Delivery to the CEC: December 05 2012

PROJECT CO-ORDINATOR

Company name: [X+O]

Name of representative: Manfred Hackl

Address: Siebensterngasse 4/22, A-1070 Vienna, Austria

Phone number: +43-676-842755-100

Fax number: +43-676-842755-599

E-mail: manfred.hackl@xpluso.com

Project WEB site address: www.harmosearch.org

TABLE OF CONTENTS

1 INTRODUCTION

1.1 PURPOSE OF THE DOCUMENT

This document reports about the extension to the HarmoNET Tourism Ontology, aka Harmonise Ontology, that was needed to address the requirements of the project.

The extension includes:

- the integration with the ontology for the registry model
- the extension to express date in a more machine computable way
- the extension of the already existing reference lists
- the definition of a model for the accommodation offers

1.2 RELATIONSHIP WITH OTHER DOCUMENTS

The description of the registry ontology is already provided in D3.2. This document reports on how these new registry concepts are related to the HarmoNET ontology and on how the integration is implemented within the application.

Another input to this document is the *HTO User Manual v4002* and the associated rdfs file, to which one should refer for a detailed understanding of the concepts and relationships.

1.3 STRUCTURE OF THE DOCUMENT

Chapter 2 clarifies how the registry ontology reuses some HarmoNET ontology concepts to describe the data that the HarmoSearch participants provide, and how the integration is implemented.

Chapter 3 presents the proposed extension to express dates in a more machine computable way.

Chapter 4 summarizes the reasons, and the proposed solution, for extending the already existing reference lists.

Chapter 5 describes how the Simple Knowledge Organisation System (SKOS), an RDF vocabulary for representing thesauri, taxonomies and classification schemes, can be helpful to represent and map Harmonise reference lists.

Chapter 6 reports about a possible future extension of the HarmoNET ontology to deal with accommodation prices and offers.

2 REGISTRY MODEL

2.1 CONTEXT

Within the HarmoSearch application, the registry has two primary functions:

- to facilitate the searches by restricting the processes only to relevant data providers;

- to allow to include external services, at run time, into the process workflow.

In order to realize these functions, the registry stores and manages knowledge about the participants[1], i.e. about the data and the services they provide.

2.2 REGISTRY ONTOLOGY

In order to allow some reasoning on it, this information is stored in terms of owl triples within the Fuseki module, and inside of the Registry Core component (for architecture details please refer to *D5.2 Registry component*).

The data model is therefore formalized as an ontology, that has three main parts:

- the HarmoSearch participant and its description (see Figure 1)

- the data which a participant provides, and its description (see Figure 2)

- the service which a participant provide, and its description (see Figure 3)

---

[1] **HarmoSearch participant:** A single, unique entity, representing a real user (normally a company) of the HarmoSearch system. Each HarmoSearch participant can operate several *data providers*, which represent a queryable view of a data source (e.g., a specific event data query interface exposed by a provider). Furthermore, each HarmoSearch participant can provide several (external) services to be used in a HarmoSearch workflow.
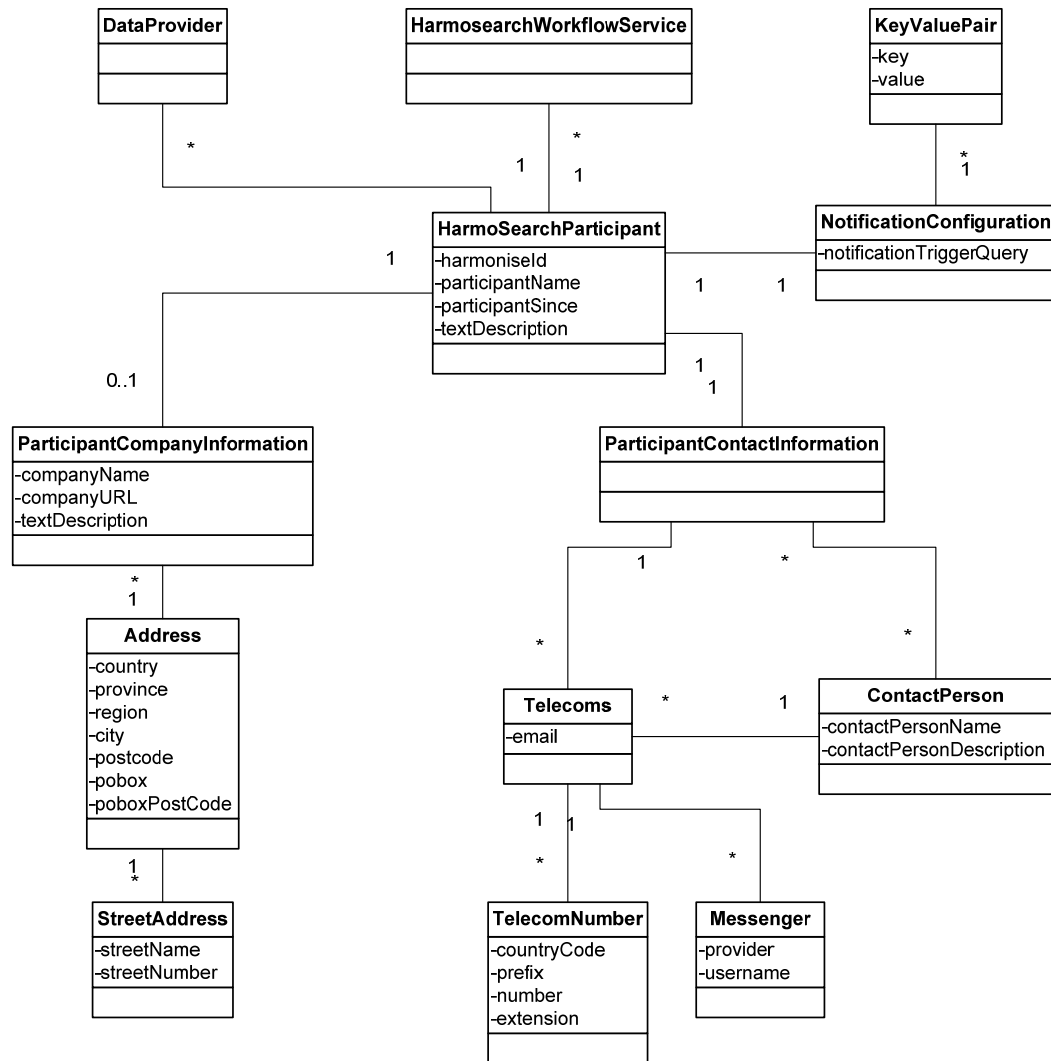
Figure 1: Concepts describing a HarmoSearch participant
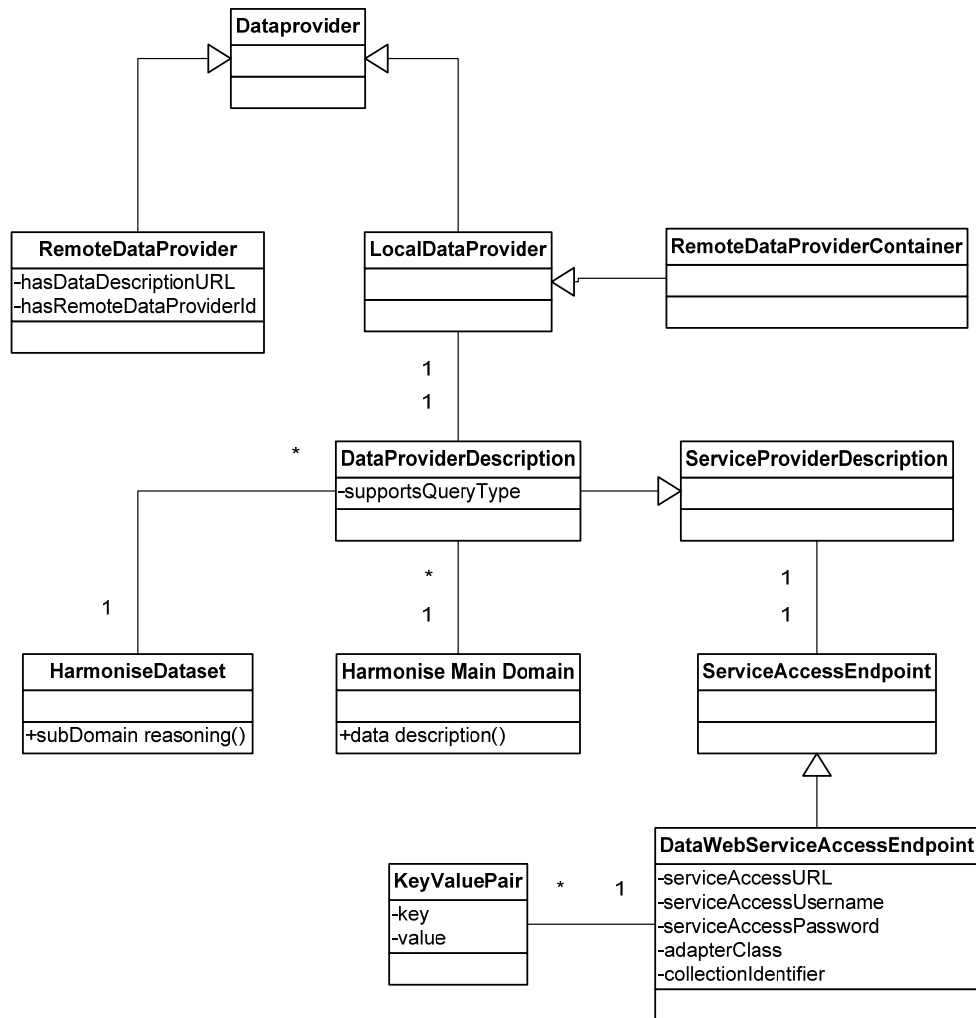
Figure 2: Concepts describing a HarmoSearch data provider
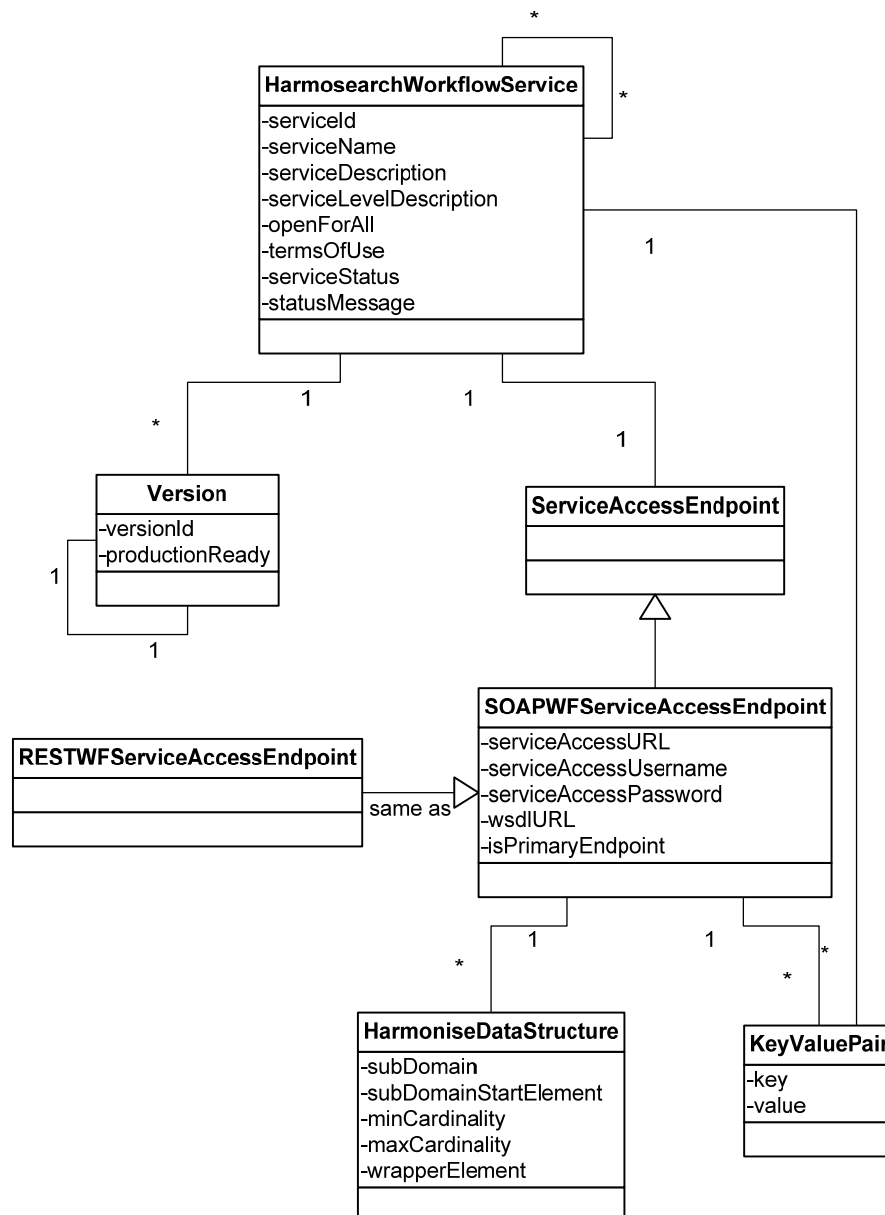
Figure 3: Concepts describing a HarmoSearch workflow service

For more details on the data model please see *D3.2 Ontology for the registry model*.

2.3 ONTOLOGY INTEGRATION

The registry model is integrated with the HTO ontology, as the DataProviderDescription reuses concepts from the HTO ontology to document subdomain, category, and location of the events (or possibly accommodations) that a participant provides, as showed in the following ontology extract, and in Figure 4.

```xml
<owl:ObjectProperty rdf:about="#descriptionOfProvidedData">
      <rdfs:comment
          >This concept contains the actual description of the data
offered by the data provider. It contains one of the start elements of
the main subdomains of the Harmonise ontology. Currently these are the
base concepts 'Event', 'Accommodation', 'Attraction' and 'Gastro'. In
order to describe the offered data, the appropriate concept from the
Harmonise ontology is used.</rdfs:comment>
      <rdfs:domain rdf:resource="#DataProviderDescription"/>
      <rdfs:range>
         <owl:Class>
            <owl:unionOf rdf:parseType="Collection">
              <rdf:Description rdf:about="&HTOv4002;Accommodation"/>
              <rdf:Description rdf:about="&HTOv4002;Attraction"/>
              <rdf:Description rdf:about="&HTOv4002;Event"/>
              <rdf:Description rdf:about="&HTOv4002;Gastro"/>
            </owl:unionOf>
         </owl:Class>
      </rdfs:range>
   </owl:ObjectProperty>
```

The figure below reports graphically an extract of the triples stored in the registry. The nodes are classes and individuals that are part of the description of the Biella data provider, and rely both on the registry ontology and on the HTO ontology.
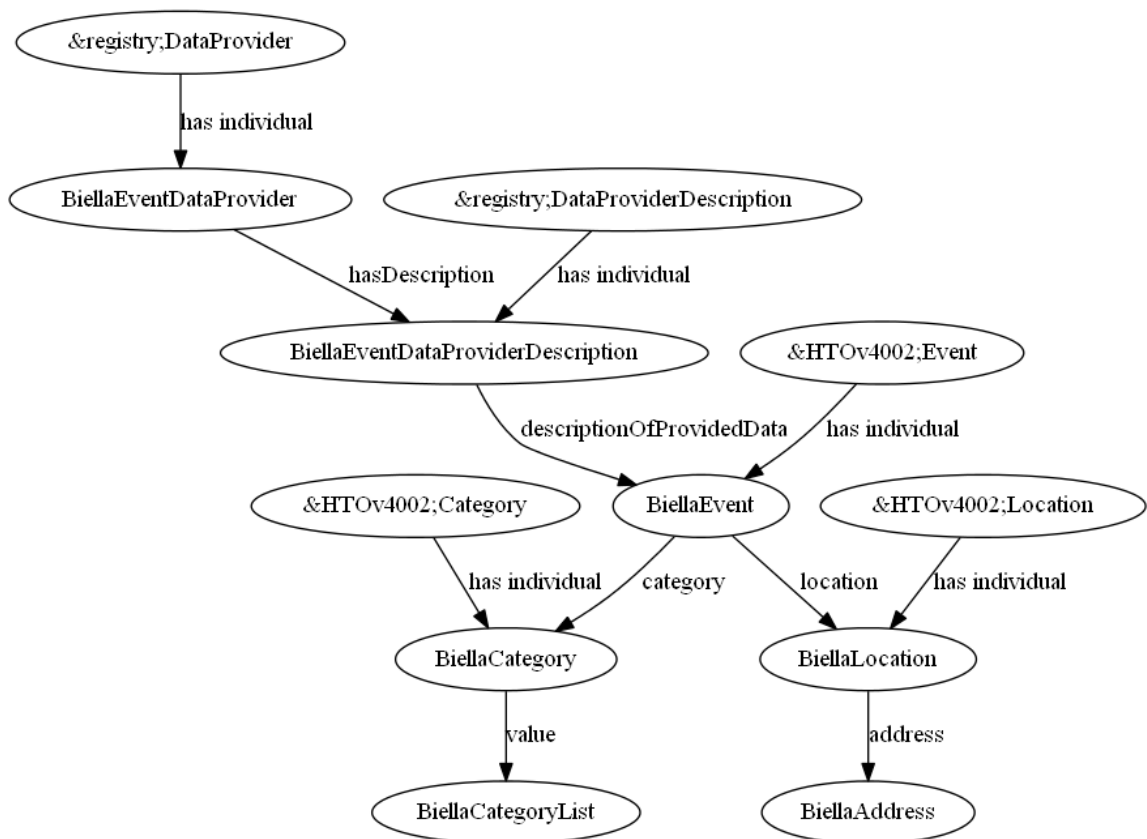
Figure 4: Ontology classes and individuals describing a data provider

The integration of the two ontologies is implemented by importing the HTO into the registry ontology, adding the HTO schema with a different namespace to the containing schema.

```
<!DOCTYPE rdf:RDF [
    <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
    <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
    <!ENTITY owl2xml "http://www.w3.org/2006/12/owl2-xml#" >
    <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
    <!ENTITY HTOv4002 " http://www.harmonet.org/ontology/HTOv4002#" >
    <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
    <!ENTITY registry " http://www.harmonet.org/ontology/registry.owl #"
>
]>
<rdf:RDF xmlns=" http://www.harmonet.org/ontology/registry.owl#"
    xml:base=" http://www.harmonet.org/ontology/registry.owl"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:owl2xml="http://www.w3.org/2006/12/owl2-xml#"
    xmlns:HTOv4002="http://www.harmonet.org/ontology/HTOv4002#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:registry=" http://www.harmonet.org/ontology/registry.owl #">
    <owl:Ontology rdf:about="">
        <owl:imports rdf:resource="file:/C:/registry/HTOv4002.rdfs"/>
    </owl:Ontology>
```

3 EXTENSION TO EXPRESS DATES

In order to allow automatic reasoning over the Harmonise ontology, it is proposed to introduce a *dateString* property to the class *Date*, as a more practical alternative to the aggregation of the properties *day*, *month* and *year*. The date string would be expressed as ISO:8601 in the extended form YYYY-MM-DD (e.g. 2012-05-01).

- *dateString(single): Date → literal (string)*

The actual *dateText* property is already designed to express the name of a particular day (e.g. Christmas), being chosen from a list of enumerated values.

This extension should be used being aware of the risk with respect to having the different pieces of information separated semantically into different data elements (day, month and year). In particular, if the day or the month was missing in a concrete source format, then there would be a loss of the still available information.

4 REFERENCE LISTS

4.1 OBJECTIVE

The goal of reviewing the reference lists in the actual version of the HarmoNET tourism ontology (HTO) v4.0.0.2 is to check if all the requirements of the HarmoSearch application are covered. To do this, the inputs have been the *HTO User Manual v4002* and the associated rdfs file.

Some standards of tourism terminology have been identified and it was analysed how the current lists of enumerated values in HTO are related to these standards, and if there is the need for further lists of enumerated values.

4.2 ORIGINS OF THE HARMONET REFERENCE LISTS

Setting up the HarmoNET tourism ontology (HTO) has been done by aiming at having the consensus of an international consortium of tourism key players, who agreed on a set of relevant concepts in the tourism domain. Among the organizations that became members of the Tourism Harmonisation Network (THN) set up by the Harmonise project, there were National Tourist Boards such as TourInFrance (France), SIGRT (Portugal), Finish Tourist Board and Spain Tourist Board, world tourism organizations such as WTO (World Tourism Organisation), IFITT (International Federation for IT and Tourism), and tourism standard organizations like OTA/TTI (Open Travel Alliance/Travel Technology Initiative), and systems such as WhatsOnWhen and TIScover. The standards that have been considered when the ontology was created are:

(1) The *Thesaurus on Tourism & Leisure Activities* from WTO (World Travel Organisation)

   i.   It is reported that this has been the most important source

(2) The *OpenTravel Specifications* from OTA/TTI (Open Travel Alliance/Travel Technology Initiative)

The HarmoNET tourism ontology has been subject of validation and approval of the Tourism Harmonisation Network and following evaluation is reported.
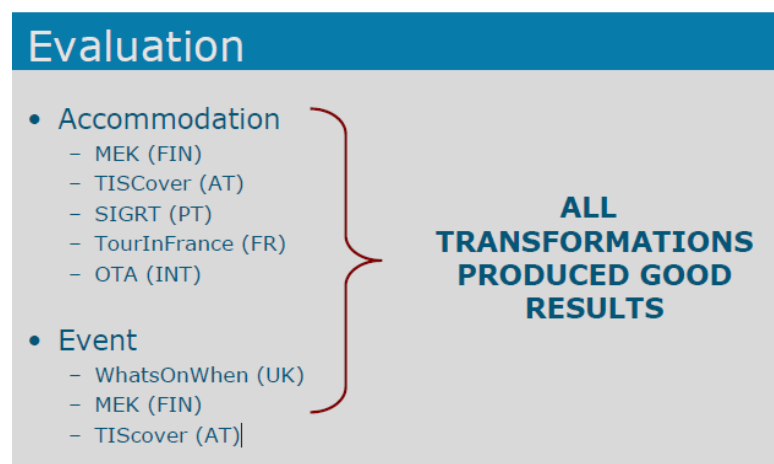
The reason for reviewing the ontology is to check how well it fits for the needs of the customers within the HarmoSearch application. As a first step, we have checked if there are still other standards that have not been taken into consideration and see how they are related to the HarmoNET tourism ontology. The following are considered.

(3) ISO 18513:2003: Tourism services — Hotels and other types of tourism accommodation — Terminology

(4) The *XFT schema* (eXchange For Travel)[2]

The validation against XFT was done with help of the partner AFIDIUM, who is member of the XFT association and contributed to its development.

The ISO standard was analysed by trying to map all the encountered terms into the HarmoNET tourism ontology, and identifying if an extension is needed.

The first step however was to identify the reference lists that are needed in HarmonSearch.

4.3 SELECTION OF REFERENCE LISTS FOR HARMOSEARCH

The following are the selected lists of enumerated values that must be implemented in the application. In the *HTO User Manual v4002* these lists are extensively reported with all their possible values.

(1) Event Category

(2) Event Profile

(3) Accommodation Type

(4) Accommodation Building Type

(5) Accommodation Profile

(6) Accommodation Facility

(7) Accommodation Services

(8) Accommodation Location Environments

(9) Accommodation and Unit Profile Person Types

(10) Accommodation and Unit Profile Meal Types

(11) Unit Type

(12) Unit Profile

(13) Unit Facility

(14) Extent

(15) Timeline Season

(16) PriceRange Discount

(17) Support Agency Payment Modes

---

[2]   http://www.exchangefortravel.org/

4.3.1 Considerations on the accommodation rating

There was, in the beginning, in HarmoSearch the intention to include some concept of rating, either or both to allow a user to search by accommodation rating and to have the results with this information attached.

However, as the aim of the HarmoSearch project is to integrate highly heterogeneous tourism data sources, this results to be a quiet complex task. The reason is that accommodation ratings may include awards or grading evaluations having very different origins. It may be an award from a body in a specific sector, like camping; it may come from a tourism body that covers a geographical area; it may be a categorization that was done on a voluntary basis; it may be the collection of client feedbacks. All these examples may be expressed in a number of ways. And one single hotel may have several accommodation ratings.

The issue is that mapping accommodation ratings between the different formats of the data providers makes no sense. However, it is possible to express the various awards in terms of one ontology, the HTO, where the Award has been modeled by a flexible concept. It is defined as follows:

Award

> *awardingBody : string*
>
> *awardRange : floating point integer*
>
> *awardAchieved : floating point integer*
>
> *dateAwardAchieved : Date*

Award Commentary

This aggregated concept provides information about the award or standard allocated to a tourism offering such as an accommodation or a restaurant. The aggregated concept reflects the reality that there are many awarding bodies, including the AA, the ADAC, the national tourism organisations, etc. The awardRange specifies the maximal value an offering can achieve. Usual gradings like '+' and '-' are represented by the values '.3' and '.7'.

For instance, in Germany the *Deutsche Hotel- und Gaststättenverband (DEHOGA)* offers a national standardised classification system called ″Deutsche Hotelklassifizierung″ (German Hotel Classification). The classification is voluntarily, and divides the accommodations into five star categories. The *awardRange* in this case would be 5, a five-star hotel would have an *awardAchieved* of 5 and a one-star accommodation would have an *awardAchieved* of 1. Furthermore, 3+ would then correspond an *awardAchieved* of 3.3, and 3- to an *awardAchieved* of 2.7.

In general, this concept should reflect the reality that different awards are not comparable and you cannot just arithmetically map from one award to another.

As a conclusion, it was decided that the accommodation ratings do not have an enumerated list of values in the application. Just inside of the data mappings each

provider must specify how its grading values are translated into a HTO Award instance.

4.4 INTEGRATION APPROACH OF THE STANDARDS

The approach in defining the enumerated lists within the HarmoNET tourism ontology (HTO) was to aim at covering all the possible concepts, rather than to adopt one particular standard. This is true for many features, like *Accomodation type, Unit type, Event category* etc. but not for the following features where a precise and unique standard has been chosen:

- Telecommunications Country Codes (ITU E.164 standard)

- Country Codes (ISO 3166 two-letter (or 'A2') country codes)

- Languages (ISO 639-1 two-letter codes)

- Days of the week (ISO 8601 both alternative with numbers and with names are allowed)

- URLs (A URL must begin with http: or https:. A URL must end with one of the ICANN values (.com, .org, etc.) or with an ISO 3166 two-letter country code (.de, .ie, etc.).)

- Currencies (ISO 4217)

- Distance units (m, km)

- File sizes (Bytes, KBytes, Mbytes, GBytes)

For future developments of the current HarmoSearch application it is envisaged to implement the enumerated values by supporting a set of selected standards, giving the possibility to chose among one of the alternative lists of values. This would be practical for users that already conform to some standard who would need to map just the concepts (e.g. the term corresponding to the HTO *Accommodation Type*) into the HarmoNET interchange format, but not the enumerated values. Each selected standard would already be mapped inside the application into every other (selected) standard.

However, as a first implementation, there will be only one list of values for each feature. The idea is to have a sufficient coverage of terms that allow to map into HarmoNET all possible values coming from the different data providers without loss of information. It may however be that one term is a specification of another (e.g. *double room* and *deluxe double room*, *hostel* and *youth hostel*, *hut* and *mountain hut*).

4.5 VALUES TO BE ADDED

The analysis of HTO and of the selected standards has shown that the current enumerated values already have a very good coverage of the terminology available in various standards. In some cases, the terms are the same inside different standards and in HTO, but if they are not, it is quite clear how they correspond to the HTO terms.

However, with respect to the domain of accommodations, we propose to add a small number of terms that have been encountered in the ISO 18513. They are the following:

HTO: Unit type

- Twin room (room with sleeping facilities for two persons in separate beds)

  o   It would be a specification of the already existing *double room*

  o   It would be a simpler alternative to a *double room* having as its profile *bed type - single*

- Family room (room with sleeping facilities for three or more persons, at least two of which are suitable for adults)

  o   It would be a specification of the already existing *multiple-bed room*

  o   It would be a simpler alternative to a *multiple-bed room* having as one of its profiles *suitable for - families*

HTO: Accommodation and Unit Profile Meal Types

- Continental breakfast (breakfast containing at least bread, butter, marmalade and/or jam and a hot drink)

- Expanded breakfast (continental breakfast supplemented by a greater variety of bread, marmalade and/or jam, cold drinks and cheese and/or cold meat)

- Buffet breakfast (self-service with a free choice of at least an expanded breakfast)

- Full breakfast (expanded breakfast supplemented by additional hot and cold foods)

4.6 REFERENCE LISTS FOR LOCATIONS

Besides the already existing enumerated values, some additional reference lists were considered to be necessary in HarmoSearch, namely for identifying, in a unique way, the event locations, like museums, art galleries, castles, train stations or any other possible place hosting some event. Having predefined values for the location of exhibitions, spectacles, music and sport events, allows to categorize the events by location and to do search & retrieval by location. This is a requirement for some data providers, and becomes a requirement for HarmoSearch to support the exchange (import or push) of data.

As a first solution, it was decided to adopt the lists of values used by euromuse.net[3], one of the data providers inside of the HarmoSearch project.

The portal www.euromuse.net observes museums not only as locations, but as organisers of exhibition events. This is in order to be able to present travelling exhibitions correctly. An example to clarify this: it is possible that the Victoria and Albert museum in London organises an exhibition in the Bundeskunsthalle in Bonn, Germany. Remaining in this case the museum organiser the Victoria and Albert, but

---

[3]   euromuse.net is a museum portal for exhibitions in Europe using Harmonise, where participating museums are sending their data via Harmonise to the central database, and where interested users can search for the cultural offer in Europe.

being the venue the Bundeskunsthalle, or possibly both will be organisers. In addition to this distinction, euromuse.net gives identifiers additionally to umbrella organisations that host museums' complexes (such as the case of the Staatliche Museen zu Berlin, with 17 museums and 20 venues; or the Universalmuseum Joanneum in Graz, with 14 Museums in 9 Venues).

It is not to be expected, that data providers will organise their event data in such manner, nor that they provide events with the identity values euromuse.net has. In the case that a museum provider has a simple organisation (one museum in one location), it would be sufficient that their event data is characterised, with at least one identifier (be it local or standard). In case information from one single provider with event information from different museums, it is essential that they provide a local reference list of these museums and venues (minimal differentiation, in order to perform a mapping). The way these reference lists are provided is handled in this document (see Chapter 5: **Fehler! Verweisquelle konnte nicht gefunden werden.**).

Towards a location standard for museums: the Museum ISIL. In Germany an initiative just ended which distributed all museums with a unique Identifier (Museum ISIL). The Institute for Museum research distributed this identifier to ca. 6.500 museums registered in the census of Museums of Germany under coordination of the national ISIL Agency (the State Library in Berlin). In order to sound the use of this identifier in different countries we conducted a survey to the ISIL agencies of those countries which most populate euromuse.net (Germany, Belgium, Austria, Finland, Italy and Switzerland). The result of this survey evidenced that the ISIL Identifier for museums is not sufficiently used in European countries (fully implemented only in Germany, partly in Austria and Switzerland). Nevertheless, this questionnaire evidences also that museums are in a period of change and this identifier might be more extended in the near future (initiatives have began in Italy to distribute museums this identifier and Belgium showed interest in this action).

You can find a summary of the ISIL Identifier as well as the results of this survey in the **Fehler! Verweisquelle konnte nicht gefunden werden.**.

4.7 IMPLEMENTATION

Once the lists of enumerated values needed in the HarmoSeach application are selected (see par. 4.3), they must be formally defined either or both in XSD (XML Schema Definition) and in RDFs.

4.7.1 Definition of reference lists within RDFS artifacts

In the article *Representing Specified Values in OWL: "value partitions" and "value sets"*[4] it is reported how an enumerated list of values can be modeled in OWL. It describes two possible patterns: as enumerations of individuals, and as partitions of classes.

"The first is simple and intuitive but has limitations. The second is more complex but is more flexible. Some classifiers also work more reliably with the second pattern than with the first.

---

[4]   http://www.w3.org/TR/swbp-specified-values/

As an example, in the first approach, the class *Health_Value* would be considered as an enumeration of the individuals *good_health*, *medium_health*, and *poor_health*. Values are sets of individuals.[...] This assumes that a value is just a unique symbol [...] the values will all need to be asserted to be different from each other. In OWL, any two individuals might represent the same thing unless there is an axiom that says, explicitly, that they are different. [...] If we did not include the differentFrom axiom in the example, then it would be possible that *good_health* and *poor_health* were the same thing, so that it would be possible to have a person who was both in good health and poor health simultaneously.

The approach is shown diagrammatically in the figure below."
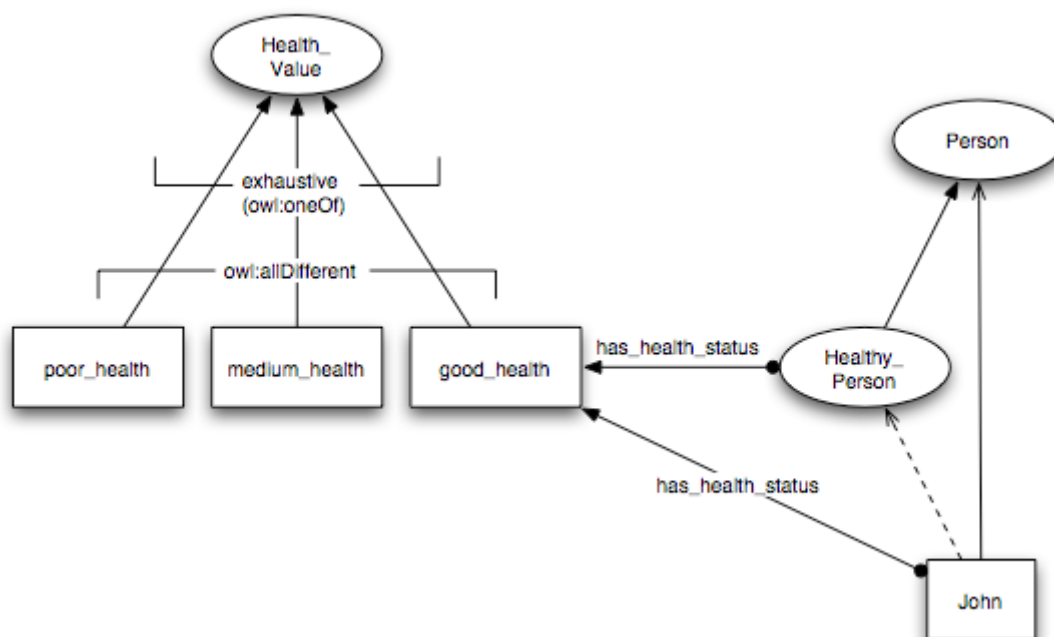
"In the second approach we consider the feature as a class representing a continuous space that is partitioned by the values in the collection of values. To say that "John is in good health" is to say that his health is inside the *Good_health_values* partition of the *Health_value* feature. Theoretically, there is an individual health value, *Johns_health*, but all we know about it is that it lies some place in the *Good_health_value* partition.[...]"
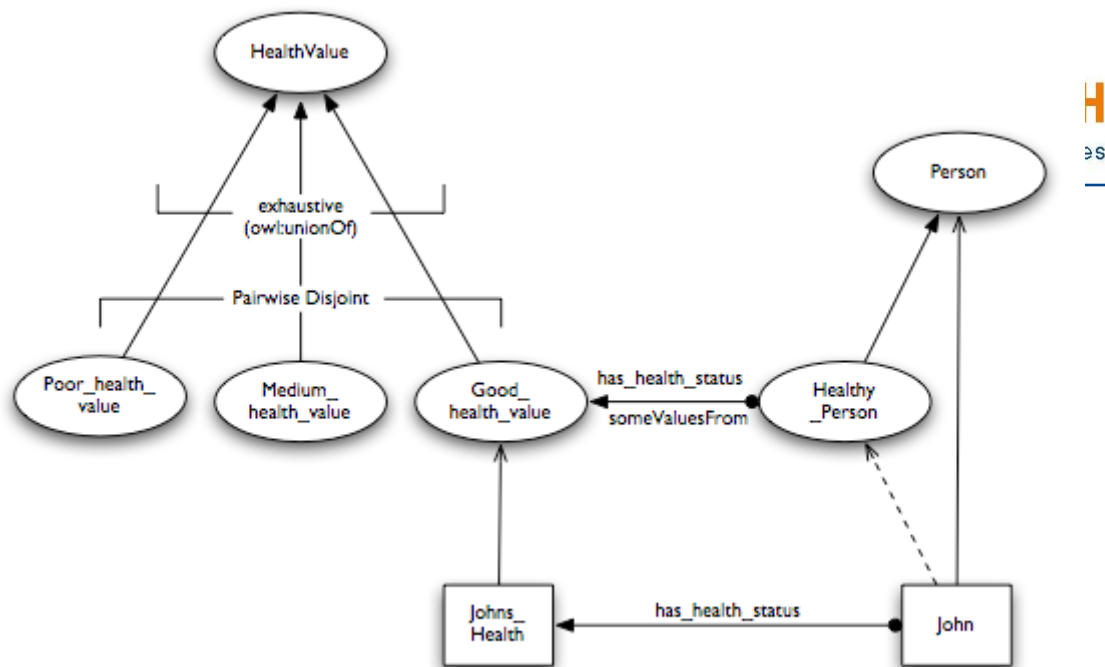
4.7.2 Considerations using the two patterns

The considerations done in the article about the first approach, and that are applicable in HarmoSearch are:

(1) […]

(2) "Many people find this the more intuitive approach".

(3) "There is no possibility of further subpartitioning of values. This is because OWL supports only equality or difference between individuals. It does not allow individuals to have partial overlaps. It is not possible, as it is for classes, to say that one individual is equivalent to the union (disjunction) of two other individuals."

    i.   In the HarmoSearch reference lists that are object of analysis we do not need the partial overlap; in some cases it would make sense (e.g. 'double room' has an overlap with 'twin room'), but it is sufficient to have just a list of values as we do not use any reasoning when applying the mappings

(4) "There is no way to represent alternative partitionings of the same feature space. Because individuals cannot overlap, if Health_Value is defined as equivalent to enumeration of one list of distinct values, it cannot also be equivalent to a different list of distinct values. To do so would cause the reasoner to indicate a contradiction. […]"

    i.   Neither this feature is required in our application.
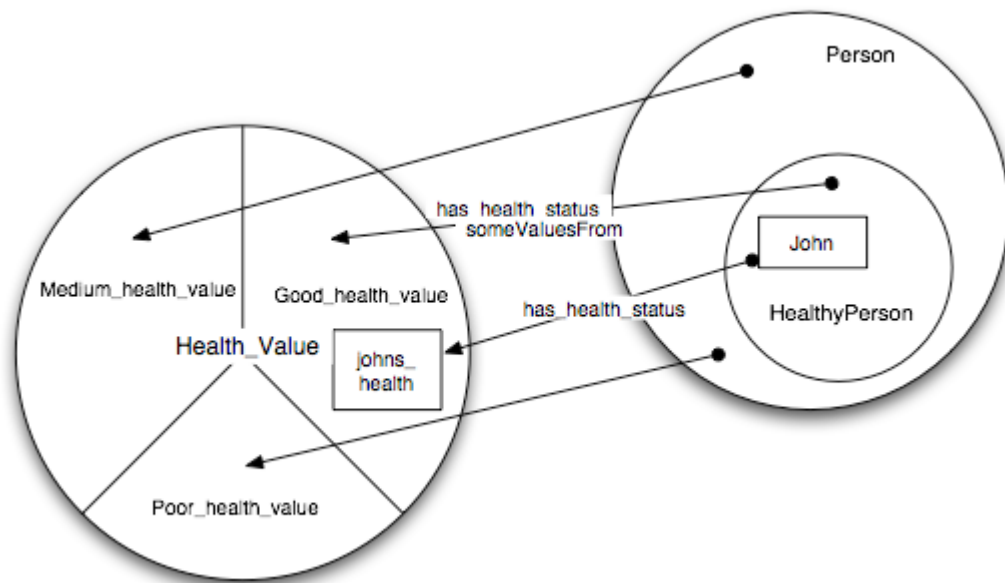
Figure 5: An adapted Venn diagram showing the use of partitioning classes to represent lists of values

The considerations about the second approach, and that are applicable in HarmoSearch are:

(1)   The result is in OWL-DL and classifies correctly using either FaCT or Racer - and almost certainly any other reasoner that handles any reasonable subset of OWL-DL.[…]

   i.   Not needed in our application

(2)   The values can be further subpartitioned, e.g. Good_health_value might be split into Moderately_good_health_value and Robust_good_health_value, simply by subdividing the Good_health_value partition.

   i.   Could make sense for some of our enumerated values, but the HarmoNET tourism ontology already has an own approach of addressing two-level list of enumerated values

(3)   There can be several alternative partitionings of the same feature space.

   i.   Could make sense maybe to address the integration of several overlapping standards (to be analysed); but is not the approach that we adopted in HarmoSearch

(4)   […]

(5)   The use of classes for values seems unintuitive to many people who come from the database and frame communities where value sets are usually enumerated lists of symbols.

From the conclusions above, we propose to implement the enumerated list of values as a set of individuals. In RDFS Frames, the format in which the HarmoNET tourism ontology is available now, this will imply that the enumerated values will be coded

within a separate RDF file. This is a good feature, that allows to keep separate the concepts of the ontology from the enumerated values that are more subject to evolution in time and differences from one country to another.

4.7.3 Definition of reference lists within XSD artifacts

Likewise for RDFs, also with XSD it is possible to define separately the main elements, corresponding to the classes and properties of the HTO ontology, and the enumerated values. Here below an extract of the current version of HTO XSD is reported.

```xml
<xs:complexType name="Accommodation">
   <xs:annotation>
      <xs:documentation>The Accommodation aggregated concept is the
primary accommodation entity. An accommodation is in one location and
is managed by an organiser. It may physically consist of a number of
buildings in proximity to one another. However, if an accommodation is
spread across a number of locations then these are treated as multiple
accommodations.
An accommodation may consist of a hotel, or a bed and breakfast, or a
pension, or any other form of accommodation which provides private-
room, catered services to guests. The self-catering and camping sectors
are excluded from the scope of the accommodation concept. However, the
accommodation ontology is structured in such a way as to facilitate the
broadening of the sub domain in the future.
An accommodation is made up of a number of units (often bedrooms). The
accommodation has facilities which are available on a per-accommodation
basis (accommodation-level facilities), such as a swimming pool or car
park. The units also have facilities, which are available on a per-unit
basis (unit-level facilities), such as hair-dryers.
</xs:documentation>
      </xs:annotation>
            <xs:sequence>
                  <xs:element name="id" type="hto:IDComponent"
minOccurs="0"/>
                  <xs:element name="name" type="hto:MultiLanguageText"
minOccurs="0"/>
                  <xs:element name="accommodationType"
type="hto:ListValue" minOccurs="0" maxOccurs="unbounded"/>
                  <xs:element name="description" type="hto:Description"
minOccurs="0"/>
                  <xs:element name="schedule" type="hto:Timeline"
minOccurs="0" maxOccurs="unbounded"/>
                  <xs:element name="legalIdentifier"
type="hto:LegalIdentifier" minOccurs="0" maxOccurs="unbounded"/>
                  <xs:element name="profile" type="hto:Profile"
minOccurs="0"/>
                  <xs:element name="award" type="hto:Award"
minOccurs="0" maxOccurs="unbounded"/>
                  <xs:element name="facility" type="hto:Facility"
minOccurs="0" maxOccurs="unbounded"/>
                  <xs:element name="service" type="hto:Service"
minOccurs="0" maxOccurs="unbounded"/>
                  <xs:element name="location" type="hto:Location"
minOccurs="0"/>
                  <xs:element name="organiser" type="hto:Organisation"
minOccurs="0" maxOccurs="unbounded"/>
                  <xs:element name="reservation"
```

```
type="hto:SupportAgency" minOccurs="0" maxOccurs="unbounded"/>
                <xs:element name="documentation"
type="hto:MultiMediaItem" minOccurs="0" maxOccurs="unbounded"/>
                <xs:element name="price" type="hto:Price"
minOccurs="0"/>
                <xs:element name="unit" type="hto:Unit" minOccurs="0"
maxOccurs="unbounded"/>
                <xs:element name="relatedTo" type="hto:Link"
minOccurs="0" maxOccurs="unbounded">
                    <xs:annotation>
                        <xs:documentation>Sister
accommodations</xs:documentation>
                    </xs:annotation>
                </xs:element>
                <xs:element name="buildingType" type="hto:ListValue"
minOccurs="0" maxOccurs="unbounded"/>
                <xs:element name="changed" type="hto:PointOfTime"
minOccurs="0"/>
            </xs:sequence>
        </xs:complexType>
```

The enumerated values would be defined as follows.

```
    <xs:element name="unit_facility">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="air conditioning"/>
                <xs:element ref="baby monitor"/>
                <xs:element ref="balcony"/>
                <xs:element ref="barbecue"/>
                <xs:element ref="bath"/>
                […]
                <xs:element ref="wc and shower/bath"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>
```

The main XSD would then include the reference lists through one of the following two options:

```
<xs:import namespace="http://<HarmoSearch reference lists location>" />

<xs:include schemaLocation="EnumLists.xsd" />
```

The import allows to add multiple schemas with different target namespace to the containing schema, while the include allows to add multiple schemas in the target namespace of the containing schema.

5 SKOSIFICATION OF HARMONISE REFERENCE LISTS

The Simple Knowledge Organization System[5] (SKOS) is a vocabulary for representing knowledge organization systems, such as thesauri, taxonomies, and classification schemes. These representations are machine-readable and can be exchanged between applications and published on the World Wide Web.

Skosication denotes the process of generating a SKOS representation from an existing terminology. Domain experts, especially in the cultural heritage sector, are in general more familiar with a SKOS representation of thesauri than the more technical representation in XSLT. Proprietary data can be transformed to a corresponding SKOS representation in order to allow better reasoning and interlinking.

5.1 CONCEPTS

The key element of the SKOS vocabulary is the element "concept". A concept is a unit of thought and exists in the mind as an abstract entity which is independent of the terms used to label them[6].

A concept in SKOS is represented by the class skos:Concept that asserts that a given resource is a concept. This is done in two steps:

1. creating (or reusing) a Uniform Resource Identifier (URI) to uniquely identify the concept, for instance,
   <http://www.harmonet.org/architecture>.
   This makes the private value "architecture" uniquely identifiable and separable from other private "architecture" values.

2. asserting in RDF, using the property rdf:type, that the resource identified by this URI is of type skos:Concept, for instance,
   <http://www.harmonet.org/architecture>
   <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
   <http://www.w3.org/2004/02/skos/core#Concept> .

A shortened version is provided by the Turtle notation which would reduce the example above (with proper namespace for the prefixes hto, rdf, and skos) to:
hto:architecture rdf:type skos:Concept.

5.2 LABELS

Labels serve as the means to refer to concepts in natural language. SKOS provides three types of labels that can be attached to concepts, skos:prefLabel, skos:altLabel, and skos:hiddenLabel. Each of these labels can be augmented with an additional language tag that restricts the label to a particular language.

The preferred label (skos:prefLabel) property serves as the main reference to a concept for any given language and there should be only one such type of label for a particular language. A preferred label that has no language tag is a means to have a

---

[5]   SKOS Specification at: http://www.w3.org/2009/08/skos-reference/skos.html

[6]   Glossary of terms relating to thesauri and other forms of structured vocabulary at: http://www.willpowerinfo.co.uk/glossary.htm

domain specific identifier. For instance, it can serve for database primary keys such as "113", see below Section 5.5.1.

The alternative label (skos:altLabel) property may be used to denote synonyms or other kinds of alternative references to a concept such as acronyms. Any concept may have any number of alternative labels for a particular language.

Finally, hidden labels (skos:hiddenLabel) are used to express unwanted labels like commonly misspelled terms and may be used by lookup-systems in order to find references even for such cases.

Going on with this example above a SKOS document would look like the one shown on Figure 6.

```
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix hto:   <http://www.harmonet.org/> .
hto:architecture rdf:type skos:Concept;
 skos:prefLabel "architecture"@en ;
 skos:prefLabel "Architektur"@de .
```

Figure 6: Sample SKOS document in Turtle notation

5.3 COLLECTIONS OF CONCEPTS

SKOS allows defining meaningful groupings or "collections" of concepts. In relation to the HTO reference lists this feature is used to create groups that correspond to these lists. For instance, the HTO ontology introduces a list of possible categories of (art)   exhibitions:
exhibition: ancient world, applied arts, archaeology, architecture, art history,…

The corresponding representation using SKOS is straightforward; the group name (exhibition) is used to denote the collection while the (skosified) values make up the content of the collection. Collections are created using the skos:Collection element that can have any number of members of type skos:Concept, see Figure 7.

```
hto:exhibition rdf:type skos:Collection ;
 skos:member
  hto:ancient_world,
  hto:applied_arts,
  hto:archaeology,
  hto:architecture,
  hto:art_history,
  ...
```

Figure 7: A SKOS collection with members

5.4 MAPPING PROPERTIES

SKOS matching properties serve to relate concepts pairwise to each other in such a way that one concept refers, to a certain degree, to the same abstract entity as the other concept.

The SKOS standard defines skos:closeMatch, skos:exactMatch, skos:broadMatch, skos:narrowMatch, and skos:relatedMatch which are used to state mapping (alignment) links between SKOS concepts.

The properties skos:broadMatch and skos:narrowMatch are used to state a hierarchical mapping link between two concepts. These two properties are the inverse of each other, so, if <A> skos:broadMatch <B> holds than <B> skos:narrowMatch <A> follows.

The property skos:relatedMatch is used to state an associative mapping link between two concepts.

The property skos:closeMatch is used to link two concepts that are sufficiently similar that they can be used interchangeably in some information retrieval applications.

The property skos:exactMatch indicates a high degree of confidence that the concepts can be used interchangeably across a wide range of information retrieval applications. skos:exactMatch is a transitive property which means that if <A> skos:exactMatch <B> and <B> skos:exactMatch <C> hold then the logical implication <A> skos:exactMatch <C> follows.

In the HTO system we advocate to use the skos:exactMatch property. Main reason is that we do not have sufficient domain knowledge to express the degree of similarity between concepts. In addition, an exact match it is consistent with inverse relations, that is, if <A> skos:exactMatch <B> holds than it follows that <B> skos:exactMatch <A> holds, too. This feature makes it easy to translate from the source domain to the target domain via the HTO ontology. Future implementations may reconsider this proposal and may support also other types of matching relations.

5.5 IMPLEMENTATION

The HTO ontology covers a wide range of concepts that are commonly used in the tourism domain. As a limitation, these concepts are only available as simple strings that even may contain arbitrary whitespace and some special characters. Examples are the concepts "café theatre" that contains whitespace and the acute accent "´" or "b&b" (bed and breakfast) that contains the ampersand character "&". Such concepts cannot be used directly, neither in XSD enumeration lists as element names nor in Semantic Web applications as classes, even if they are augmented with a proper namespace. As a consequence, the concepts have to be encoded accordingly. Within the HarmoSearch project a simple tool has been created that performs these encodings. As a result, the examples above are transformed to a skosified version as shown on Figure 8.

```
hto:caf%C3%A9_theatre rdf:type skos:Concept;
 skos:prefLabel "café theatre"@en .
hto:b%26b rdf:type skos:Concept;
 skos:prefLabel "b&b"@en .
```

Figure 8: Url-encoded concepts of values with special characters.

Special characters are replaced by their URL-encoded values while for whitespace we use the underline character "_" in order to enhance readability for humans.

5.5.1 Data provider domain

Data providers that maintain their own domain specific values have to create mappings of these values to the corresponding values taken from the intermediate language (here the HarmoNET ontology). Typically, such values have a meaning only in the context of the data provider and often the representation is a reference to some database primary key value as in the following example taken from two organisations from the cultural and museum domain[7].

```
113 = 28. Remark: Category Musik maps to category Musik.
```

The keys (113 respectively 28) have obviously a meaning only in the corresponding context of the involved organisations. In order to make these concepts usable for HarmoSearch they have to be augmented with a proper namespace to maintain references and to avoid misinterpretations since these keys may have a different meaning in the domain of other data providers. In order to overcome this problem we can use the same approach as has been outlined for the skosification of HarmoNET value lists. Following the two step procedure from above a minimal SKOS representation will result in the creation of an URI to uniquely identify the concept and the concept assertion.

```
<http://www.smb.museum/113>  rdf:type skos:Concept.
```

```
<http://www.euromuse.net/28> rdf:type skos:Concept.
```

In addition, labels may be associated with these concepts in order to support retrieval systems and give meaning to human readers. Finally, mapping between these two concepts can be achieved by asserting an exact match using the skos:exactMatch property:

```
<http://www.smb.museum/113> skos:exactMatch <http://www.euromuse.net/28>.
```

Note that this kind of mapping does not employ the intermediate language approach but maps directly from the source to the target organisation and vice versa. In order to use HTO (or any other standard) it is necessary to assert mappings from source to target via the concepts provided by the employed standard. This is realised with two mappings and builds on the transitiveness of the exact match property.

```
<http://www.smb.museum/113> skos:exactMatch <http://www.harmonet.org/music>.
```

```
<http://www.harmonet.org/music> skos:exactMatch <http://www.euromuse.net/28>.
```

Since the inverse of exact match also holds, the two mappings could be defined also in a way that both organisations make references to a concept taken from the intermediate language as in the following example:

```
<http://www.smb.museum/113> skos:exactMatch <http://www.harmonet.org/music>.
```

```
<http://www.euromuse.net/28> skos:exactMatch <http://www.harmonet.org/music>.
```

---

[7] Database one: Staatliche Museen zu Berlin - Preußischer Kulturbesitz, http://www.smb.museum Database two: http://www.euromuse.net

5.5.2 XSLT and RDF

In HarmoSearch, data from data providers is obtained in some kind of XML format and transformed via an XSLT processor and the intermediate HTO representation to the format and domain of the receiver. At some point of this process, the translation of the (skosified) values has to be performed, too. This is realised by sending a SPARQL[8] query from the XSLT processor as an HTTP GET request to a component capable to process the query and to return matching substitutions for the given domain specific values. The returned result set (in XML format) is then used by the XSLT processor to construct the output document. The query is constructed from the input data like the one shown on Figure 9.

```
PREFIX skos: http://www.w3.org/2004/02/skos/core#
SELECT ?concept
from <http://localhost:8080/demo/mapping.ttl>
WHERE {
 http://www.smb.museum/113 skos:exactMatch ?concept .
 }
```

Figure 9: Simple SPARQL query for matching concepts.

---

[8]   SPARQL specification at: http://www.w3.org/TR/rdf-sparql-query/

6 MODELLING THE OFFERS

6.1 OBJECTIVE

The HarmoSearch application is a portal that provides search & retrieval, and data exchange functionalities. The focus of the actual implementation is the domain of events, covering museum exhibitions, spectacles, and music and sport happenings.

This chapter summarizes the analysis of possible extensions to the HTO ontology that would be needed when dealing with accommodations. In fact, it is envisaged, as a future extension to the HarmoSearch portal, to include accommodation-specific implementations of the current search & retrieval services.

A final decision has not yet been taken on how to model the prices and the offers, nevertheless the current status of conclusions is reported for future reuse.

6.2 OVERVIEW OF THE ACCOMMODATION MODEL

In the HarmoNET tourism ontology (HTO) the modeling of accommodations relies on two main concepts: the *Accommodation* and the *Unit*.

> *The Accommodation aggregated concept is the primary accommodation entity. An accommodation is in one location and is managed by an organizer. It may physically consist of a number of buildings in proximity to one another. However, if an accommodation is spread across a number of locations then these are treated as multiple accommodations.*

> *An accommodation may consist of a hotel, or a bed and breakfast, or a pension, or any other form of accommodation which provides private-room, catered services to guests. The self-catering and camping sectors are excluded from the scope of the accommodation concept. However, the accommodation ontology is structured in such a way as to facilitate the broadening of the sub domain in the future.*

> *An accommodation is made up of a number of units (often bedrooms). The accommodation has facilities which are available on a per-accommodation basis (accommodation-level facilities), such as a swimming pool or car park. The units also have facilities, which are available on a per-unit basis (unit-level facilities), such as hair-dryers.*

> *The unit is the offering element which is acquired by an end user. Typically, this will be a room in a hotel, a pitch in a camping-site, a holiday apartment, etc. For communal accommodation (gite, hostel), the unit is a bed.*

> *Each unit has its own pricing model, which reflects its type and its facilities. Each unit has its own profile – this collection of attribute-value pairs describes the relevant properties of the unit. For a camping pitch, this may include its terrain, its area and the availability of an electricity point. For a suite, it may include the number of rooms, whether a four-poster bed is available, etc.*

The properties that an Accommodation may have are the following:

Accommodation

>   *id (single): IDComponent*

name (single): MultiLanguageText

accommodationType (multiple): ListValue (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

description (single): Description

schedule (multiple): Timeline

legalIdentifier (single): LegalIdentifier

profile (single): Profile (see Appendix - Accommodation Profile*)*

award (multiple): Award

facility (multiple): Facility (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

service (multiple): Service (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

location (single): Location

organiser (multiple): Organisation

reservation (multiple): SupportAgency

documentation (multiple): MultiMediaItem

price (single): Price

unit (multiple): Unit

relatedTo (multiple): Link (sister accommodations)

buildingType (multiple): ListValue (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

changed (single): PointOfTime

The properties that a Unit may have are the following:

Unit

id (single): IDComponent

name (single): MultiLanguageText

unitType (multiple): ListValue (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

telecoms (multiple): Telecoms

description (single): Description

unitProfile (single): Profile (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

facility (multiple): Facility (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

service (multiple): Service (see Appendix - **Fehler! Verweisquelle konnte nicht gefunden werden.***)*

documentation (multiple): MultiMediaItem

price (single): Price

Here below an overview of the relationships that affect *Accommodation* and *Unit*.
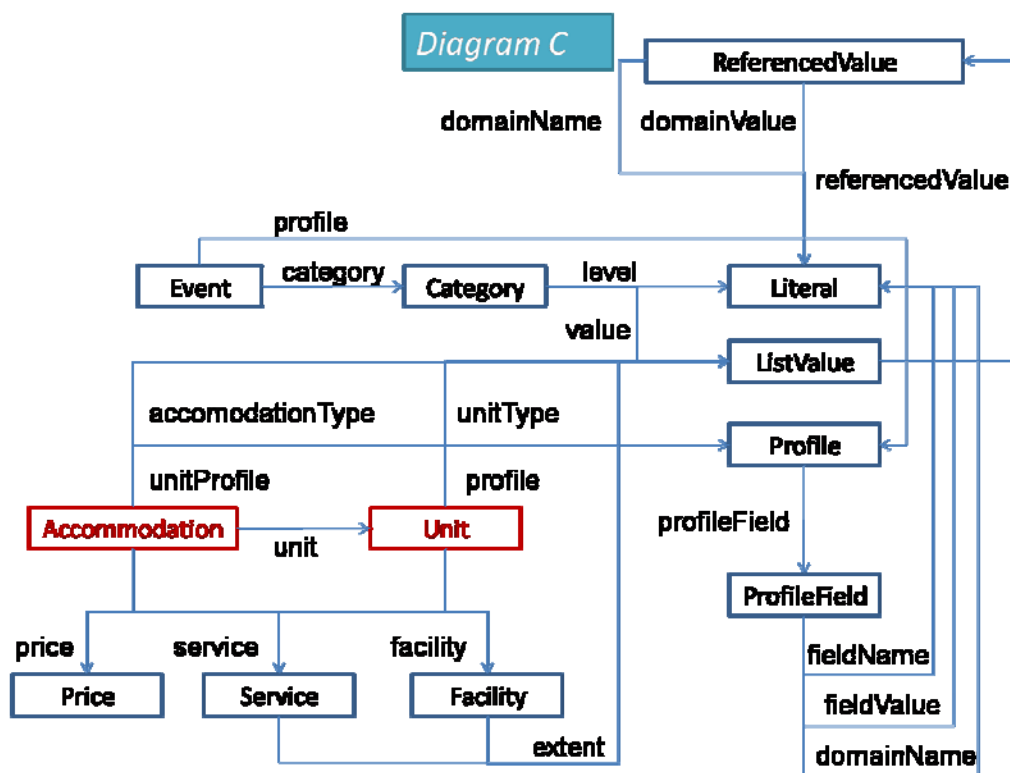
Figure 10: Accommodation and Unit centric overview of the HarmoNET concepts
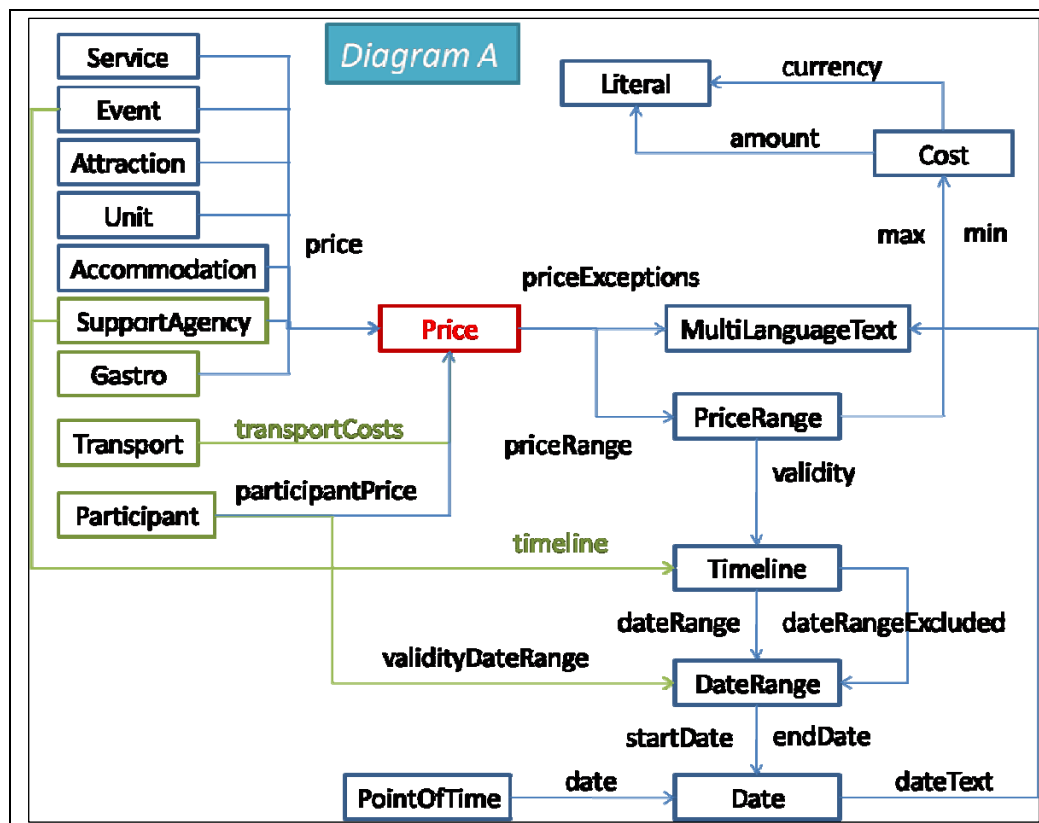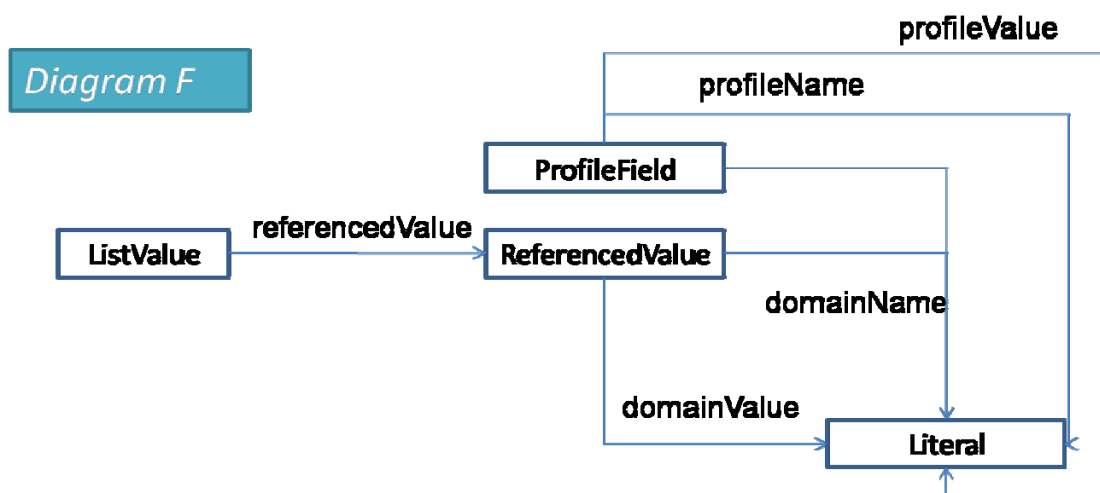
Figure 11: Price centric overview of the HarmoNET concepts

- The **ListValue** aggregated concept contains a value of a value list defined globally within the complete HarmoNET network as well as a value of a domain specific value list together with the name of the domain. The HarmoNET-wide unique value enables a mapping between domain specific value lists while the domain specific value avoids a loss of information in case a precise mapping is not possible.
- The **ReferencedValue** aggregated concept contains a value and the name of the domain the value is defined in.

Figure 12: ListValue concept diagram

To clarify, the following diagram shows the use of enumerated lists of values (i.e. the ListValue class). The ovals represent individuals of the HTO ontology. The diagram expresses that a given house providing accommodations is classified as having accommodation type "b&b" (bed and breakfast) in terms of HTO, and for completion, also the classification from the data provider is kept, where the term "garnì" was used.
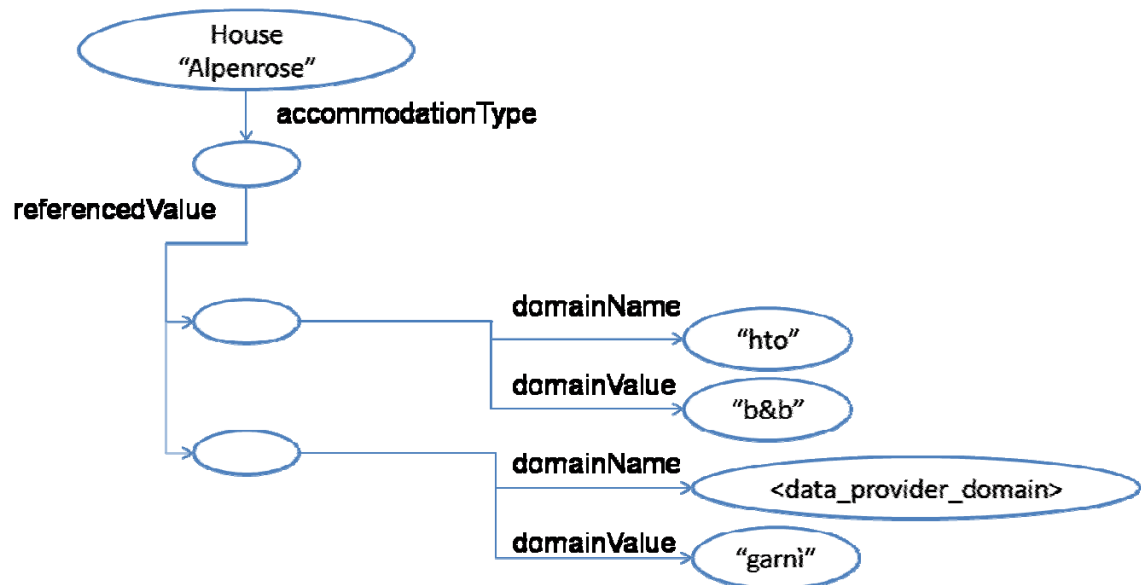

Figure 13: Example – Representation of enumerated values in the HTO ontology

6.3 EXAMPLES

Here below an example of how the HarmoNET classes and properties are used to build up an example of accommodation. It describes the features of an accommodation that a b&b house offers, including price description.

The properties have lower-case initials. The classes used are only *Accommodation* and *Unit*. The individuals have all capital letters (e.g. ACC1).

Example:

Double room with shower in a Bed&Breakfast with balcony, on the first floor with a view onto the valley and internet connection in the room. The b&b has 25 beds in total, the name is "Alpenrose", it is classified with two stars as from the "Deutsche Hotelklassifizierung (from DEHOGA)". It has car park, drying room and dogs are allowed.

The price varies between 30 EUR and 40 EUR (a discount for children under 7 years), is valid from 6.1.2012 until 3.2.2012, and applies per person and per night with a minimum of stay of three days, with full occupancy (two persons) of the room. This price is not applicable for older (frequent) clients, and for children with less than 3 years, who do not pay.

The cost for internet access is varies from 5 EUR to 10 EUR (a discount for unemployed). It applies to the Wi-fi access for 24 hours per person, starting from 12am. It is not applicable for conferences.

Accommodation ACC1

- o **name: "Alpenrose"**
- o **accomodationType:**
  - ▪ **LIST_VAL1**
    - • **referencedValue: REF_VAL1**
      - o **domainName: "hto"**
      - o **domainValue: "b&b"**
    - • **referencedValue: REF_VAL2**
      - o **domainName: "__sourceDomain__"**
      - o **domainValue: " garnì"**
- o **profile: PR_ACC1**
  - ▪ **profileField: PR1_FL_ACC1**
    - • **domainName: "hto"**
    - • **fieldName: "dogs allowed"**
  - ▪ **profileField: PR2_FL_ACC1**
    - • **domainName: "hto"**
    - • **fieldName: "number of beds"**
    - • **fieldValue: "25"**
- o **Award: AWR2**
  - ▪ **AwardingBody: "Deutsche Hotelklassifizierung (DEHOGA)"**
  - ▪ **AwardRange: 0.4**
  - ▪ **AwardAchieved: 0.3**
- o **facility: FAC_ACC1**
  - ▪ **facilityName: LIST_VAL2**
    - • **referencedValue: REF_VAL3**
      - o **domainName: "hto"**
      - o **domainValue: "car park"**
  - ▪ **facilityName: LIST_VAL3**
    - • **referencedValue: REF_VAL4**
      - o **domainName: "hto"**
      - o **domainValue: "drying room"**

- o **service**
  - ▪ **serviceName: LIST_VAL4**
    - • **referencedValue: REF_VAL5**
      - o **domainName: "hto"**
      - o **domainValue: "internet"**

  - ▪ **Price: PRICE1**
    - • **priceRange: PR_RANGE1**
      - o **min: COST1**
        - • **amount: "5"**
        - • **currency: "EUR"**
      - o **max: COST2**
        - • **amount: "10"**
        - • **currency: "EUR"**
      - o **discount: LIST_VAL5**
        - • **referencedValue: REF_VAL6**
          - • **domainName: "hto"**
          - • **domainValue: "unemployed"**
      - o **criterion: CRIT1**
        - • **languageText: LAN_TXT1**
          - • **language: "ENG"**
          - • **text: "Wi-fi access for 24 hours per person, starting from 12am"**
    - • **priceExceptions: PRC_EXCETP1**
      - o **languageText: LAN_TXT2:**
        - • **language: "ENG"**
        - • **text: "not applicable for conferences"**
  - o **unit UNIT1**

- **Unit UNIT1**
  - o **id:"Number 3"**
  - o **unitType: LIST_VAL6**
    - ▪ **referencedValue: REF_VAL7**
      - • **domainName: "hto"**
      - • **domainValue: "double room"**

- referencedValue: REF_VAL8
    - domainName: "ISO:18513"
    - domainValue: "twin room"
- unitProfile: PR_UNIT1
    - profileField: PR1_FL_UNIT1
        - domainName: "hto"
        - fieldName: "floor-1"
    - profileField: PR2_FL_UNIT1
        - domainName: "hto"
        - fieldName: "view-valley"
- facility: FAC_ACC1
    - facilityName: LIST_VAL7
        - referencedValue: REF_VAL9
            - domainName: "hto"
            - domainValue: "balcony"
    - facilityName: LIST_VAL8
        - referencedValue: REF_VAL10
            - domainName: "hto"
            - domainValue: "wc and shower"
    - facilityName: LIST_VAL9
        - referencedValue: REF_VAL11
            - domainName: "hto"
            - domainValue: "internet access"
- price PRICE2
    - priceRange: PR_RANGE2
        - min: COST3
            - amount: "30"
            - currency: "EUR"
        - max: COST4
            - amount: "40"
            - currency: "EUR"
        - validity: VALID1
            - dateRange: D_RANGE1
                - startDate: DATE1

- • *dateString[9]:* "2012-01-06"
    - • **endDate: DATE2**
        - • *dateString:* "2012-02-03"
    - • **discount: LIST_VAL10:**
        - o **referencedValue: REF_VAL12**
            - • **domainName: "hto"**
            - • **domainValue: "children under 7"**
    - • **criterion: "per person and per night with a minimum of stay of three days, with full occupancy (two persons) of the room"**
    - ▪ **priceExceptions: PR_EXCPT2**
        - • **languageText: LAN_TXT3:**
            - o **language: "ENG"**
            - o **text: "old (frequent) clients"**
        - • **languageText: LAN_TXT4:**
            - o **language: "ITA"**
            - o **text: "clienti di vecchia data"**
        - • **languageText: LAN_TXT5:**
            - o **language: "ENG"**
            - o **text: "children under 3 years do not pay"**

Other example accommodations that can be expressed are the following:

- ▪ Family room with wc/bath in a comfort hotel (three stars), with satellite tv, full board, swimming pool, view on the lake, 3th floor, with a given price per unit (room).

- ▪ Accommodation in a bunk bed of a mountain hut, incl. breakfast, with a given price per person.

- ▪ Bed in dormitory of youth hostel, with a given price only for sleeping (breakfast excluded).

To see all the possible values for Accommodation and Unit Type, Profiles, Facilities and Services please refer to the *HTO User Manual v4002*.

---

[9]   New property proposed, as more practical alternative to the aggregation of the properties *day, month* and *year*. The date string will be expressed as ISO:8601 in the extended form YYYY-MM-DD.

6.4 MODELLING ACCOMMODATION OFFERS

An accommodation offer is a particularly advantageous pricing model that applies under well defined conditions. For example, it may refer to a stay of two persons in a double room for one week from Saturday to Saturday in a given period. It may include further details, like facilities and services included in the offer package. Finally, it is regulated by some terms and conditions, delivery terms and a cancellation policy. An accommodation offer includes, conceptually, every information that makes it possible for an interested customer just to say "yes" and he is able to conclude the contract. Therefore, the following properties would be required at a minimum to model an offer:

Offer

    value (a pricing model)

    description

    validity (e.g. price valid today only)

    terms & conditions

    delivery terms

    cancellation policy

    forms of discount

In the HarmoSearch application however there is no support foreseen to conclude the reservation transaction. The goal is just to exchange (register/publish and retrieve) information on accommodation offers. Term & conditions, cancellation policy, delivery terms apply like they do for the accommodations retrieved through a metasearch, and are out of scope of the application. All what might be needed when dealing with accommodations inside of HarmoSearch is a model of prices.

Prices within the accommodation domain are already considered in the HarmoNET Tourism Ontology. It must be checked if this model is already sufficient to express what could be needed in the future extensions of the HarmoSearch application, when addressing the scenarios of accommodation search.

6.4.1 Prices in HarmoNET ontology

*The **Price** aggregated concept specifies the price of a tourism service as a price range and price exceptions. The Price concept is intended to specify a basic price of a tourism service but not a precise price structure, depending on person type, booking period, length of stay, etc.*

Price

    priceRange (multiple): PriceRange

    priceException (multiple): MultilanguageText

*The **PriceRange** aggregated concept includes the information needed to ascertain the cost of access to, or participation in, a tourism offering. This includes support for two levels of pricing (Max and Min), and a description of the criterion that controls which price is to be quoted (e.g. the days of the week when lower prices apply). It*

also allows a validity period to be established, during which the price applies, and zero or more discounts which may be available (e.g. for students, older people, groups, etc.).

> *PriceRange*
>
>> *name (single): MultiLanguageText*
>>
>> *max (single): Cost*
>>
>> *min (single): Cost*
>>
>> *criterion (single): MultiLanguageText*
>>
>> *validity (multiple): Timeline*
>>
>> *discount (multiple): ListValue (see **Fehler! Verweisquelle konnte nicht gefunden werden.**)*

The price applies to units (e.g. a room, a bed in a dormitory) and to services (e.g. babysitter, internet access, shuttle service). It is part of the specification of the price, in particular within the property *criterion*, if the price is per room, per person, and per night. Each unit or service has only one price, which may be the aggregation of several price ranges (depending on period of validity for instance).

6.5 IMPLEMENTATION

The following features that are specific for an accommodation offer have been identified:

(1)   to express a concrete and precise price (an offer), not a vague indication of a price range

(2)   to express possibly an end price for the full stay, not only per person and per night

(3)   to express the validity in which the price applies

(4)   to express the number of persons

(5)   to express duration of stay

(6)   to express the amount of a discount (not only "discount for under-3")

(7)   to express the amount of a price rise (e.g. for an added bed or cot)

The analysis showed that the price model of the HTO is not sufficient to express the required information. In order to extend it, two main alternatives were identified:

(1)   The introduction of a new concept that represents a concrete price

> This concept could be a subclass of hto:Price, thus inheriting several properties of the more general current concept of price. Figure 14 gives an overview of this solution (the new classes and properties are in green). Beside the specialization of price, also the concept of occupation would be added. An occupation would refer to the stay of a given number of persons

for given days, and to a particular unit (e.g. room) or accommodation (e.g. a whole holiday house).

The apparent advantage of the solution is that it would theoretically be possible to model the concepts and relationships in the most intuitive way. However, the model is needed to map from and towards real travel data exchange languages, which are quite complex in order to be expressive enough. Thus, creating a new model would require a very good analysis of the main travel data exchange formats.

(2)   *The adoption of the XFT[10] model of price*

The second solution would be to adopt and benefit from an already existing model, which could be XFT (eXchange For Travel).

The new concrete price concept could coexist, or substitute in some circumstances, the existing hto:Price (Figure 15). It would be necessary to understand if also some other related concept, like the XFT Segment, Product or Offer, would have to be adopted and integrated into HTO (Figure 16).

As reported at the beginning, a final decision was not taken, because the focus of the actual implementation is on events. Future work will, however, be the extension of the portal with accommodation services, where the concept of price and of offer play a central role.
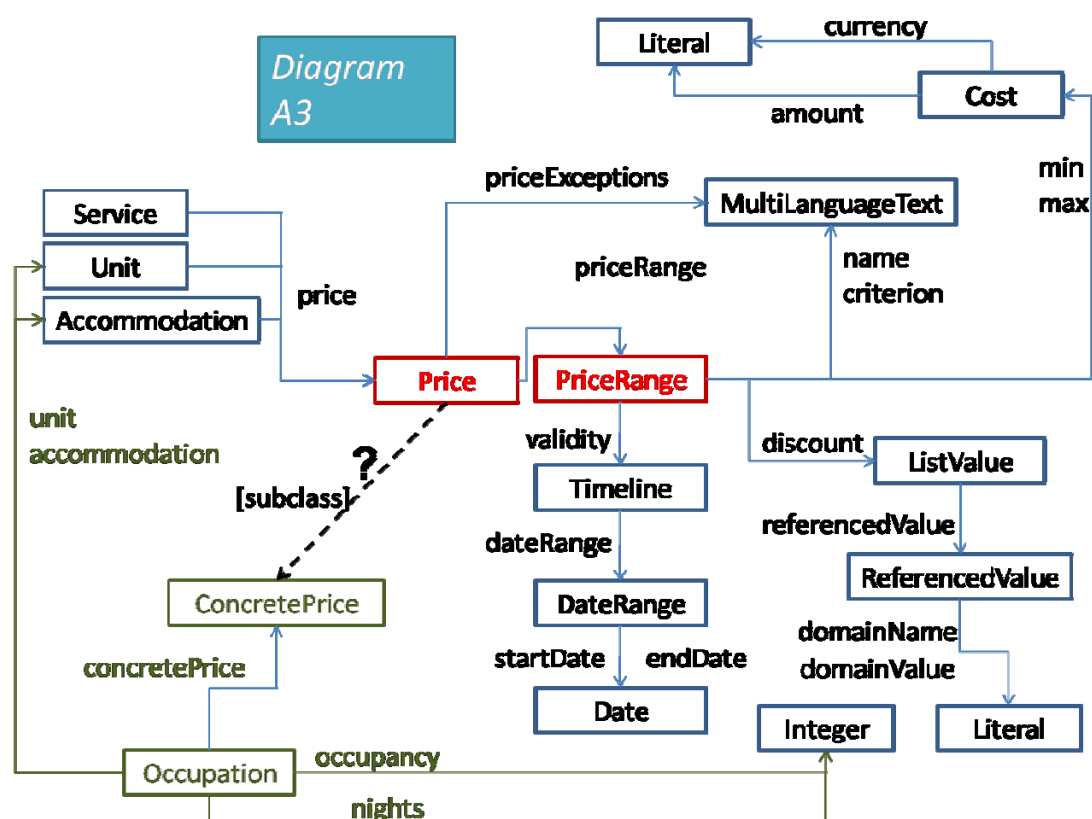
---

[10]   http://www.exchangefortravel.org/

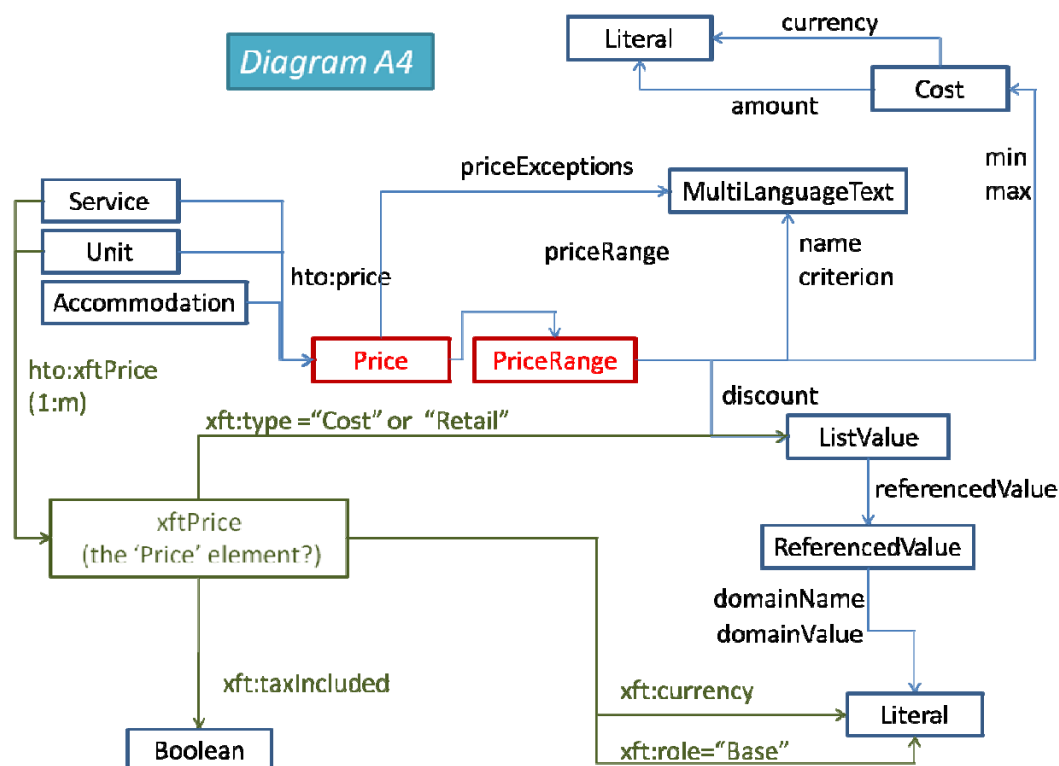*Figure 14: Concrete price as subclass of hto:Price*

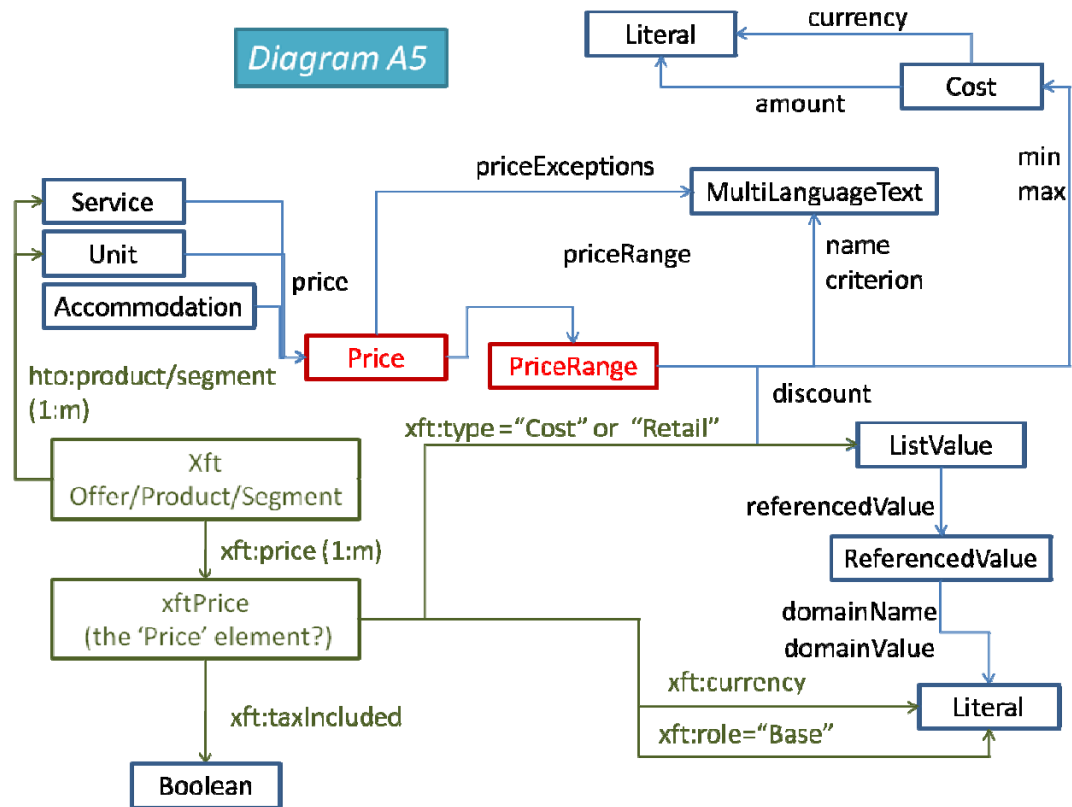*Figure 15: Integration with XFT price concept (variant 1)*

Figure 16: Integration with XFT price concept (variant 2)

7 CONCLUSIONS

An extension of the HarmoNET ontology was needed in order to implement the application scenarios of the HarmoSearch project. This extension included:

  (1)   the integration with the registry data model (chapter 2)

  (2)   adding a property to allow to express dates also as strings (as ISO:8601 standard) (chapter 3)

  (3)   adding further values in the existing reference lists (chapter 4):

       i.   adding two values to the list of Unit types (paragraph 4.5)

       ii.   adding four values to the list of Profile Meal Types (see par. 4.5)

  (4)   adding reference lists for event locations (section 4.6)

Finally, an extension of the HTO ontology for modelling accommodation offers was considered, and the preliminary outcomes of the analysis are reported in chapter 6.

ANNEX: ISIL – INTERNATIONAL STANDARD IDENTIFIER FOR LIBRARIES AND RELATED ORGANIZATIONS

DEFINITION

ISO 15511:2011 specifies the International Standard identifier for libraries and related organizations (archives, museums) with a minimum impact on already existing systems.

An ISIL identifies an organization, i.e. a library, an archive, a museum or a related organization, or one of its subordinate units, which is responsible for an action or service in an informational environment (e.g. creation of machine-readable information).

THE MUSEUM ISIL

Between museums, but also among museums and other memorial institutions (libraries, archives…) the collaboration and transfer of electronically data is growing; be it in the framework of collaborative projects or other form of mutual handle. Consequently, it makes sense to provide museums with an international unequivocal identifier which is for its similarities based on the ISIL system structure.

The Museum-ISIL differs slightly from the Library ISIL. Its structure remains nevertheless simple:

Country code (i.e. DE for Germany), followed by a hyphen (-); the particle "MUS" and again hyphen; the ISIL is completed by a 6-cypher number.[11]

METHODOLOGY

A survey was conducted to estimate the use and consciousness of the ISIL number for Museums. We addressed the international ISIL Agency as well as national ISIL agencies of countries that collaborate with www.euromuse.net closely (Austria, Finland, Italy, Switzerland and Germany).

Questionnaire

- Can you tell us how expanded is the use of the ISIL Identifier for Museums in your country/region?

- Is there an institution which is specifically in charge of procuring the ISIL number to museums?

- Could you estimate if museums are aware and using this number for electronic data exchange, e.g. EUROPEANA

SUMMARY

- The tendency is that only museums that hold a library and/or an archive have the ISIL number, and use the same ISIL number of the library (Austria).

---

[11]   http://www.museen-in-deutschland.de/index.php?t=isil

- ▪ Some countries are starting to systematically provide with ISIL for Museums (Switzerland, Germany and Italy). It is as of today work in progress.

- ▪ Not all European countries have chosen to adapt to the ISO 15511 standard, so this solution will not be feasible in countries like Spain, Greece, Portugal, Ireland and many East European countries.

ANSWERS

Germany

All Registered museums at the Institute for Museum Research have an ISIL Number. This does not mean that all museums will use it to qualify their electronically treated data. But they can.

Belgium

There are two establishments that attribute the codes. The Royal Library of Belgium attributes the codes to all libraries and documentation centres, and the National Archive attribute them to all archives.

Museums can also apply for an ISIL-code, but have not yet done so (*"we do not consider museums as our main target group"*). The few museums that are represented in our database have requested an ISIL-code for the library which is part of the museum.

Switzerland

All Museums with a library and or an archive are currently receiving the same ISIL number for the museum. From the 1700 registered cultural heritage institutions, it is estimated that 300 museums have this number.

Currently other museums are interested in receiving an ISIL number.

Finland

In principle the National Library (national ISIL Agency) is willing to assign ISILs to museums, too, but so far no museum has requested one. Currently all our ISILs are assigned to libraries or publishers.

Italy

In Italy there is no ISIL Agency specific for museums. The Cultural Ministry is in charge of a project (MuseiD) that will be in charge of procuring ISIL to some museums. The service of 'procuring' ISIL for museums (5.000 will receive ISIL) will be hosted in the ICCU, the Instito Centrale per il Catalogo Unico, which is the national ISIL Agency.

8 LIST OF FIGURES