HARMOSEARCH
the future of information services

FP7-SME-1
Project no. 262289

HARMOSEARCH

Harmonised Semantic Meta-Search in
Distributed Heterogeneous Databases

SEVENTH FRAMEWORK PROGRAMME

D2.1 final

Use Case Specification

Due date of deliverable: 2011-01-31
Actual submission date: 2011-01-31

Start date of project: 2010-12-01 Duration: 24 month

Project funded by the European Comission 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)	

HARMOSEARCH
the future of information services

PROJECT ACRONYM: **HARMOSEARCH**

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

Grant Agreement: 262289

Starting date: December 2010 **Ending date:** November 2012

Deliverable Number: D2.1 final

Title of the Deliverable: Use Case Specification

Task / WP related to the Deliverable: WP 2, Task 2.1

Type (Internal or Restricted or Public): Public

Author(s): Adriano Venturini, Albert Rainer, Beatrix Lehmann, Bernhard Schrittwieser, Christoph Herzog, Claudio Prandoni, David Faveur, Manfred Hackl, Martin Navratil, Thomas Motal, Thorsten Siegmann

Partner(s) Contributing: [X+O], eCTRL, EC3 Networks, Museumsmedien, Afidium, TU-Wien, CPR, HarmoNET, SPK

Contractual Date of Delivery to the CEC: January 31st 2011

Actual Date of Delivery to the CEC: January 31st 2011

PROJECT CO-ORDINATOR

Company name: [X+O]

Name of representative: Manfred Hackl

Address: HAMBURGERSTRASSE, 10/7, A-1050 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 defines the functionalities which should be supported by the new components which will be developed within HarmoSearch project through a set of use cases and scenarios. These use cases refer to the three major components which will be implemented in this project, which are the Metasearch component, the Semantic Registry and the Mapping tool.

1.2 DEFINITIONS OF TERMS AND ABBREVIATIONS

Harmonise: name of the existing technological solution. The current version is Harmonise 2.0, which includes the Harmonise Ontology, Harmonise Service Center and the Harmonise Portal.

Harmonise Platform: name identifying the whole set of Harmonise components

Metasearch: one of the major functions to be implemented in this project and the name of the component which will support it. It provides distributed search capabilities to the integrated data sources.

Semantic Registry: component to be developed within this project which will contain semantic profile information about the services available within the Harmonise networks.

Mapping tool: the mapping tool is a standalone application that supports a user with little technical knowledge in creating visually the necessary mapping definitions from the data model of a Harmonise participant to the one of Harmonise and vice-versa. It consists of a graphical User Interface to show and manipulate mappings, a pluggable set of algorithms to support automatic mappings, a generator to create mapping artifacts, and an infrastructure in order to manage a mapping project.

1.3 RELATIONSHIP WITH OTHER DOCUMENTS

The input of this document is the description of the work of this project proposal. This document will then be input for the deliverable D 2.2 System Architecture.

1.4 STRUCTURE OF THE DOCUMENT

This document is structured in the following main sections:

- Chapter 2: provides the context of the system, its major actors and supported scenarios.
- Chapter 3: collects the main scenarios which will be enabled by the components developed in this project.
- Chapter 4: describes the use cases for supporting the metasearch scenario.
- Chapter 5: describes the use cases for the import scenario.
- Chapter 6: describes the use case for supporting the publishing and subscribe scenario.
- Chapter 7: describes the use case for supporting the data modification scenario.

- Chapter 8: describes the use cases for supporting the marketing intelligence scenario.
- Chapter 9: classifies the use cases considering their relationship with the major components which will be implemented within this project and in relationship with the Harmonise platform.

2 OVERVIEW

Goal of the HarmoSearch project is to leverage the use of an existing mediation and harmonisation service for the European tourism market, called Harmonise, by adding new components of clear market value addressing specific user needs. The current version of Harmonise, Harmonise 2.0, is an online service to exchange data with other partners without the need to change their local data schema.

To better capture these needs, this document describes a set of scenarios which address the typical situations which occur when on the one hand there are organisations exposing their content to the network and on the other hand there are users consuming the content.

In particular, these scenarios capture the needs of a general framework for content provisioning and consuming where:

- Some actors of the tourist industry provide specific content to be searched from members of the Harmonise network.
- Some actors provide innovative services adding value to the offered contents.
- Other actors consume contents and services which the service providers have made available through the platform.

HarmoSearch will enable these actors (Harmonise participants) to easily provide and consume contents.

From the scenarios were derived the use cases which provide a more detailed view (but at a more abstract level) of the desired functionalities.

These will allow collecting the necessary requirements which will drive the development of a new set of enabling components:

- To enable organisations to query contents offered by other organisations (Metasearch Process)
- To enable organisations to register their services and semantically describe the content they provide (Service Registry)
- To enable organisations to easily define how their data map with respect to the common Harmonise ontology (Mapping Tool).

The use cases are grouped by scenarios instead of following the more technical view of the enabling components in order not to miss functionalities which are necessary to support the desired scenarios.

How the use cases are then grouped considering the specific components is described in the table of the Chapter 9.

Thus, this document collects five major scenarios, which are able to capture the needs that tourism industry players have when exchanging content and services.

These scenarios are the:

- **Metasearch Scenario**. A service consumer, which could be a tourism portal for example, wants to offer to its users search functionalities over data offered by other service providers. In this scenario the most typical (and

probably one of the most complex) situation where the searched contents are bookable products.

- **Data Import Scenario**. In this scenario, a service consumer wants to perform batch imports of content offered by different content providers. Data may be stored in the repository of the service consumer.
- **Publishing and Subscription Scenario**. Service providers want to offer their data or services to other organisations which are interested in consuming them. This scenario describes the functionalities which are needed to enable these actors from one side to offer their data or services and from the other side to consume them.
- **Data Modification Scenario**. Content exchanged by organisations needs to be modified or improved. For example, a character set transformation could be needed, or translation services which could be offered by other organisations.
- **Market Intelligence Scenario**. By monitoring the transferred data or exploiting the search services to monitor content offered by data provider organisations, it is possible, by performing statistical analysis, to capture market trends or offer value added services. This scenario details the opportunities retained within this project.

While reading these scenarios, it is important to keep in mind that the goal of HarmoSearch is not to implement all the advanced services introduced in these scenarios, but to develop a set of new components that can be added to the existing ones to provide a platform where actors of the tourism industry can propose their services to other players, i.e. the so called Harmonise platform.

The following diagram shows the actors which the system will handle and summarises the scenarios described within the document.

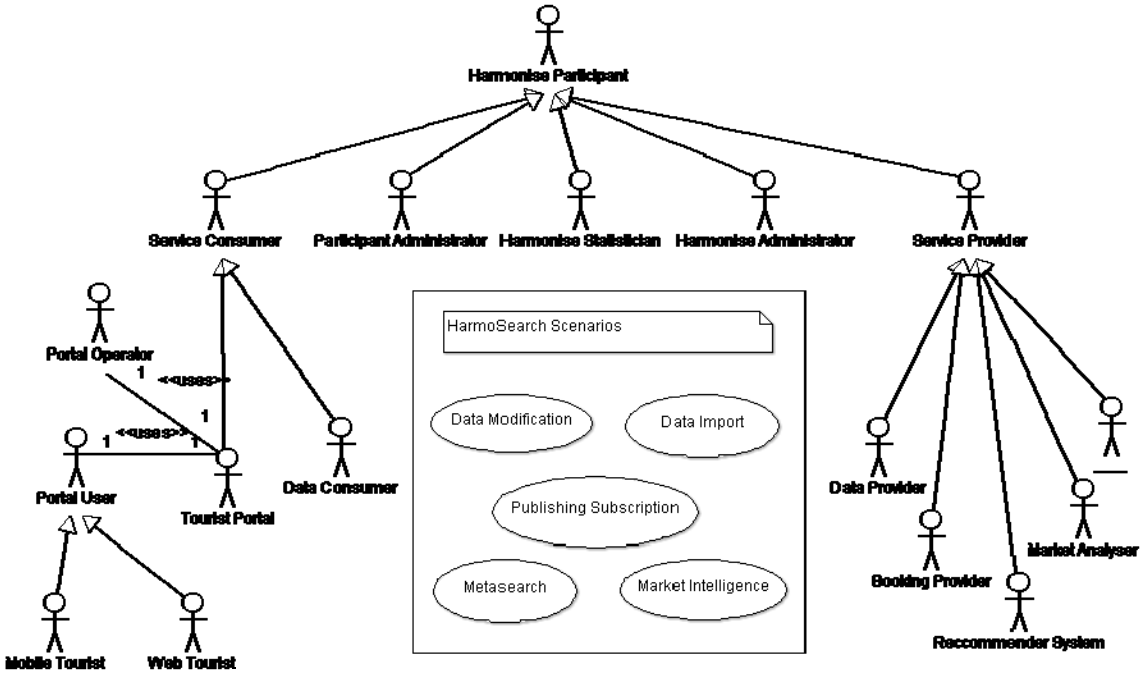

Figure 1: Actors and Scenarios

The diagram shows the basic framework addressed by HarmoSearch:

- Each actor in the network is called Harmonise Participant.
- One subgroup corresponds to the service providers, showing the most important ones considering the competences and business areas of the organisations involved in the project. Some organisations act as Data Providers, others as Booking Providers. Then there are value added services which could be offered via the Harmonise platform, like systems which offer personalised recommendations to the consumers of the data (recommender system) and Market Analysers, i.e. tools which monitor the network to capture and identify market trends and offer value added information (Market Intelligence).
- Service consumers are the organisations which exploit the network to access content and get value added services.
- Final users (like the Tourists or Operators of the organisations of the network) are shown as well in the diagram, since they are often the real final consumers of the offered services.
- Finally the central box contains the identified scenarios which are supported by HarmoSearch, which will be described in detail in the next section.

The complexity of the system leads to a system architecture where additional services could be easily plugged in. Some of them (the ones considered critical for the project from the business point of view) could become part of the Harmonise environment, others could be offered by third parties, and others could be developed by the Harmonise members themselves. This will be addressed in the deliverable D 2.2 Architectural Design.

3 APPLICATION SCENARIOS

3.1 METASEARCH SCENARIO

3.1.1 Title

Search multiple sources for bookable accommodation

3.1.2 Synopsis

A tourism portal wants to provide users of the portal with an accommodation metasearch. Users are presented with suitable hotels as well as price quotes and availability information according to their search criteria. Also a “deep link” to the (external) booking platform for each item must be provided.

3.1.3 Description

The operators of a tourism portal want to provide an accommodation metasearch to the portal users. Therefore, a solution is required which can take the users’ needs as input and provide information and prices for available accommodations in the region or country the portal represents. A possibility for the user to go to a booking page in order to book a specific hotel is also required.

A user’s search query is entered on a form on the portal and can encompass, e.g. stay date range, required accommodation type, number of persons, geographic location, etc.

A user will expect results almost instantly; therefore speed is of importance for the portal. Getting first search results within 10 seconds is considered sufficient. Search results will be displayed to the user as soon as they come in.

The portal operators want to have as many accommodations searched as possible. However, a fee being paid by the accommodation provider in order to be listed on the portal is envisaged. Therefore a possibility to control which accommodations (from which source) are presented to the user is desired. In order to charge for search, the portal also records how many accommodations from which providers were received, which were viewed in detail by the portal user and on which the user clicked the “book” deep link.

Since many accommodations are selected in this way, it is desirable to have the possibility to highlight the most relevant choices for a given user exploiting specific recommendation algorithms, which could be plugged in the process flow. Items could be ranked exploiting user profile information, which have been acquired on the tourism portal through specific forms, or implicitly derived by analyzing the user navigation on the portal (or any other means).

In addition, the portal could select only some items to be recommended to the user (suggested for you). To be able to reason on the items, for each item should be available structured information, like the type of accommodation, amenities, geo coordinates, geographical areas, type of rooms. This scenario can further be generalised for other types of products, like events, attractions, or other devices, like for example a mobile scenario.

This scenario is closely related to the scenario *"Importing accommodation information to the country portal"* – in order to speed up the display of search results, static information about the accommodations is imported to the portal regularly.

3.1.4 Actors

Accommodation Provider

- **Motivation**: Accommodation providers contributing to the metasearch process want to increase their visibility and attract customers.
- **Expectation:** They expect to be found when suitable search criteria are applied and want to be listed on the first page of the search results.
- **Use Cases:**
- MS-1
  - , Harmonise registration and setup
- MS-2
  - , Searching for bookable items
- MS-2a
  - , Modify Results
- MS-2b
  - , Ranking items based on the user profile
- MS-3
  - , Recommending bookable items

Tourism Portal

- **Motivation:** The operators of the portal want to improve the usefulness of the portal, increase the number of bookings in general and, possibly, make money by charging fees to the hotels to be found.
- **Expectation:** The portal operators expect the metasearch to be always available and produce results of a constant quality. It should also operate with minimum maintenance cost – especially regarding portal personnel.
- **Use Cases:**
- MS-1
  - , Harmonise registration and setup
- MS-2
  - , Searching for bookable items
- MS-2a
  - , Modify Results
- MS-2b
  - , Ranking items based on the user profile

- MS-3
  - , Recommending bookable items

3.2 IMPORT SCENARIO

3.2.1 Title

Importing static accommodation information to a tourism portal

3.2.2 Synopsis

The operators of a tourism portal want to import static accommodation data regularly from several specified sources. The data quality must be high enough to allow an import into the content management system.

3.2.3 Description

The operator of a tourism portal wants to import static data about accommodation in the portal's country into the portal's CMS. The goal is to have the data stored locally in order to be able to provide details about accommodation for list views, in a "featured accommodation" area, for search results retrieved from the accommodation metasearch, etc.

Therefore, the imported data must encompass basic information like a picture, short and long description, location, type of accommodation, etc. But also more detailed information – especially accommodation amenities – are necessary for the import to be useful, e.g. for automatically generated lists. A specific example would be a list of accommodations in a given region which offer fishing services.

The vocabulary used has to confirm 100% to the CMS, otherwise an import is impossible. Also, many fields in the CMS are mandatory; therefore the accommodation to be imported has to have all the required data.

The data must be imported only from specific data sources with whom a contract is established. These data sources are responsible to deliver data in sufficient quality to be automatically integrated into the portal. Especially all the data elements which are mandatory for the CMS have to be delivered (e.g. small picture, short description, etc.).

The import should be triggered (i.e. controlled) by the portal in order to select times of low server load for the import. Also the import should use delta updates (only new/changed data) when offered by the data providers. Furthermore, an indication of outdated accommodations which have to be deleted in the CMS must be provided. It should be possible to import data from different sources at different times in order to reduce the CMS workload.

On the CMS itself, there shall be no duplicates, even if the same accommodation is delivered by more than one of the sources. The duplicate detection is to be implemented on the portal in a semi-automatic process. For each accommodation an "authoritative source" is set from which further updates are gathered. This happens once when a new accommodation is received from any source. In order to minimise manual work, updates on accommodation data have to be possible without triggering the duplicate detection again.

This scenario is closely related to the scenario *"Search multiple sources for bookable accommodation"* – the imported "static" information is used to display details on search results.

3.2.4 Actors

Accommodation Provider

- **Motivation**: Accommodation providers want to be more present on tourism portals and thus attract customers.
- **Expectation**: They expect the information they send to be displayed correctly on the portal and that their accommodations automatically appear in the appropriate lists (e.g. of cottages offering fishing services).
- **Use Cases:**
- IMPORT-1
  - , Harmonise registration and setup
- IMPORT-2
  - , Batch transfer of static data

Portal Operator

- **Motivation**: The company operating the portal wants to improve the usefulness of the portal and increase the number of bookings in general.
- **Expectation**: The portal operators expect the data to be automatically imported, where a minimum of manual maintenance is required.
- **Use Cases:**
- IMPORT-1
  - , Harmonise registration and setup
- IMPORT-2
  - , Batch transfer of static data

3.3 PUBLISHING AND SUBSCRIPTION SCENARIO

3.3.1 Title

Publishing and subscription scenario

3.3.2 Synopsis

Data providers want to easily upload and publish once or on a regular basis their data (row or enhanced) and want to define to whom individual data may be pushed. Similarly data consumers need to be able to easily subscribe to raw or enhanced data from different providers and once or regularly download those data. The portal also needs to implement an alert and negotiation service to facilitate subscriptions to new data or new consumers.

3.3.3 Description

Basic data upload features allows the data provider to upload a file and select the recipient. A publishing feature allows the data provider to upload his data in the Harmonise network (and possibly transform the data according to services he would have subscribed). Data may be uploaded once manually or regularly scheduled. Data provider can then define the data consumers that will be allowed to download the data (through data subscription and negotiation), associate profiles to the data so as to allow alerts to be generated by the portal to non subscribers.

Data consumers need not only to be able to download files pushed by data providers, but they also must be able to select the content to import, possibly grouping data from different providers, using differential input or even possibly using web services instead of files. They also need alerts when subscribed data is updated or when new data corresponding to their profiles is published. They must then be able to negotiate data subscription through the portal. They can also use the same mechanism to subscribe to portal services (e.g. data transformation, market intelligence …). Some reference or transposing features may even be directly provided by the portal itself.

The platform must also provide participant profiles to store preferences, filters, access information about subscribed services and data as well as a way to charge and receive payment for the subscribed services.

This scenario is closely related to the scenarios *"Data transformation"* and *"Market Intelligence"* – in order to improve the data quality and usage feedback as well as the positioning of the portal itself.

3.3.4 Actors

Data provider

- **Motivation:** The data provider wants to publish and enhance its content to the widest possible appropriate audience with minimal effort, negotiating via the portal new data and new consumers and analysing data usage and users (e.g. Market intelligence).
- **Expectation:** The data provider expects to become the prime source of information for its data and be able to analyse its usage.
- **Use Cases:**
- PS-1
  - , Data Hosting
- PS-1a
  - , Use of Cross Standard Interoperability
- PS-1b
  - , Sending Multimedia Content
- PS-2
  - , Data Publishing
- PS-3

- , Data Enrichment

- PS-4
  - , Consumer Association

Data consumer

- **Motivation:** The data consumer wants to increase the quantity and quality of data that he can use as well as build relation between data so as to increase their value, visibility and business. He therefore requires alerts when subscribed data is changed or when interesting data are made available through the portal. The data consumer then requires the possibility to subscribe via the portal and easily define the way to receive the data.
- **Expectation:** The data consumer expects to get relevant data as soon they are made available and to be informed about new data so to always be up to date.
- **Use Cases:**
- PS-5
  - , Data Download
- PS-6
  - , Data Subscription
- PS-7
  - , Alert Definition

Service provider

- **Motivation:** This actor offers services to improve the usefulness of the Harmonise network by providing value added services for data publishing and subscription (on top of the data transformation and market intelligence scenario) and needs to maximise the number of users and of services so as to become the obvious data aggregator.
- **Expectation:** This actor expects to become the obvious data aggregator for certain types of content.
- **Use Cases:**
- PS-8
  - , Portal Service Subscription
- PS-9
  - , Negotiation
- PS-10
  - , Payment
- PS-11
  - , Profile Management

3.4 DATA MODIFICATION SCENARIO

3.4.1 Title

Data Modification

3.4.2 Synopsis

Harmonise is able to modify data according to users' needs on different levels and with different granularity.

3.4.3 Description

A data consumer collects information from several partners from all over Europe. Data is transferred to data consumer via Harmonise. The data provided by the partners often does not fulfil the consumer's needs exactly, which leads to very time consuming improvement actions for the data consumer. Furthermore, the same partner may also transfer their data to other portals that will have different requirements. A single data consumer therefore cannot influence the quality of the content provided nor impose its requirement to the partners. By using Harmonise, the data consumer wants to lower its editorial effort in restructuring data according to its needs by using a third party's service for automatic transformations.

Unexpected deviations of content received can be e.g. (1) "File provided instead of text" or (2) "data not completely encoded in Unicode/UTF-8" or (3) "Text is longer than the expected limit of 400 chars" or (4) "Content Language is German instead of English" or (5) media is not in the expected format and size. Harmonise ensures data quality by controlling, correcting, completing and transforming the content according to the recipient's needs or may indicate deviations from expected data format or content. Harmonise should offer the necessary data modification services to impact both the conceptual and content level and provide additional assisted services (like direct data input) to guarantee the success of service.

A data provider has structured its data according to its needs and receives requests from several partners using its own query interface. In order to become a participant as data provider in the Harmonise network, appropriate mappings are necessary to transform queries and data. Such mappings are created using the mapping tool and are registered at Harmonise.

Service providers may offer additional services that can improve the quality of raw data from data providers in various ways. Examples are simple ranking of result items based on a single criterion, decision support using recommender systems or natural language translation of content.

Examples:

(1) **Format**: Sender is providing information in a file like <link>http://example.com/information.pdf</link> but Recipient needs text <text>This is the text inside the PDF-file</text>. Harmonise reads the PDF and transforms its content to text information

(2) **Content structure**: Most of a text sent by Sender is encoded in UTF-8/Unicode (like agreed), but some chars are not. Recipient expects data fully in UTF-8/Unicode. Harmonise identifies the wrongly encoded chars correctly and transforms them to the needed UTF-8/Unicode.

(3) **Content length**: Recipient defines a special max./min. length for the content. Sender sends something too long/too short. Harmonise allows to define individual transformation rules and considers those in transformation.

(4) **Content language**: Recipient expects information in English, but Sender offers German language. Harmonise translates information to the language needed and/or informs Recipient about aberrations.

3.4.4 Actors

Data provider

- **Motivation:** The data provider wants to deliver its content to the widest possible audience with minimal effort
- **Expectation:** The data provider expects to receive requests in an appropriate format from Harmonise
- **Use Cases:**
- DM-1
  - , Provider mapping

Data consumer

- **Motivation:** The data consumer wants to have access to data with no/little effort needed for data integration.
- **Expectation:** The data consumer expects to receive data in a suitable form and content.
- **Use Cases:**
- DM-2
  - , Consumer mapping
- DM-3
  - , Data Cleansing
- DM-4
  - , Decision support
- DM-5
  - , Data Enrichment
- DM-6
  - , Data Translation

Service provider

- **Motivation:** A service provider wants to improve the usefulness of Harmonise by providing additional value added services to Harmonise participants.
- **Expectation:** The service provider expects appropriate usage according to predefined terms.
- **Use Cases:**

- DM-3
  - , Data Cleansing
- DM-4
  - , Decision support
- DM-5
  - , Data Enrichment
- DM-6
  - , Data Translation

3.5 MARKET INTELLIGENCE SCENARIO

3.5.1 Title

Market Intelligence Scenario

3.5.2 Synopsis

The HarmoSearch-solution shall allow to access data for analytical processing of market information and for statistics of usage and content. This data comes either from the data sources directly (e.g. querying accommodation prices) or from transactions processed via HarmoSearch (e.g. number and volume of transactions)

3.5.3 Description

The system shall allow collection of data from different sources, which is used as input for analytical processing. This shall allow to make a deeper statistical analysis of accommodation prices and to calculate forecasts of market price developments. The analysis itself is done by one of the data consumers, who is able to identify duplicates and define the query-intervals himself. (Just to mention: It could be of interest that this partner is also able to provide the identification of duplicates as a service to other users…)

In particular the system shall collect:

- Prices of definable European accommodations in a given region in a certain period from various sources (booking providers, accommodation providers). This collection shall include basic descriptions of the accommodations (category, facilities, location).
- In addition to the accommodation prices, also periodic and “unperiodic” (irregular) events shall be collected including description and prices.
- Furthermore rating sites shall be screened for comments and ratings of accommodations and events collected according to the points stated above. (text mining)

The end users in this scenario are companies analysing the data collected.

The second part of the scenario is not based on active data exchange, but on getting statistical information about the exchange of data in the network. The solution shall be able to display data usage per type of transaction, per axis specified in profile (product, geographical data, category, status …) in array and graphical format. For

these statistics user specific information as well as a global indicator compared to all users or users of certain types (same area, same type of data ...) will be returned. It should also define different views at a later stage (geographical usage, user typology ...).

It should be possible to display ranking information in array and graphical format (global for the user, per product, per category, per geography). Ranking would correspond to position within result lists, number of times it was displayed/clicked and can be clicked for more information or other types of services.

3.5.4 Actors

Market Analyser

- **Motivation:** Wants to gain specific data from specific data provider to allow comparison.
- **Expectation:** To get comparable data back based on the same query string
- **Use Cases:**
- MI-1
  - , Submit Ad Hoc Request
- MI-2
  - , Execute Interval Request

Harmonise Statistician

- **Motivation:** Wants to get statistics on the usage of the system.
- **Expectation:** To get useful statistics about usage of own system as well as in comparison with select other users (including ranking).
- **Use Cases:**
- MI-3
  - , Analyse Data
- MI-4
  - , Manage Notifications
- MI-5
  - , Log System Activity

4 USE CASES FOR METASEARCH SCENARIO

4.1 USE CASE DIAGRAM

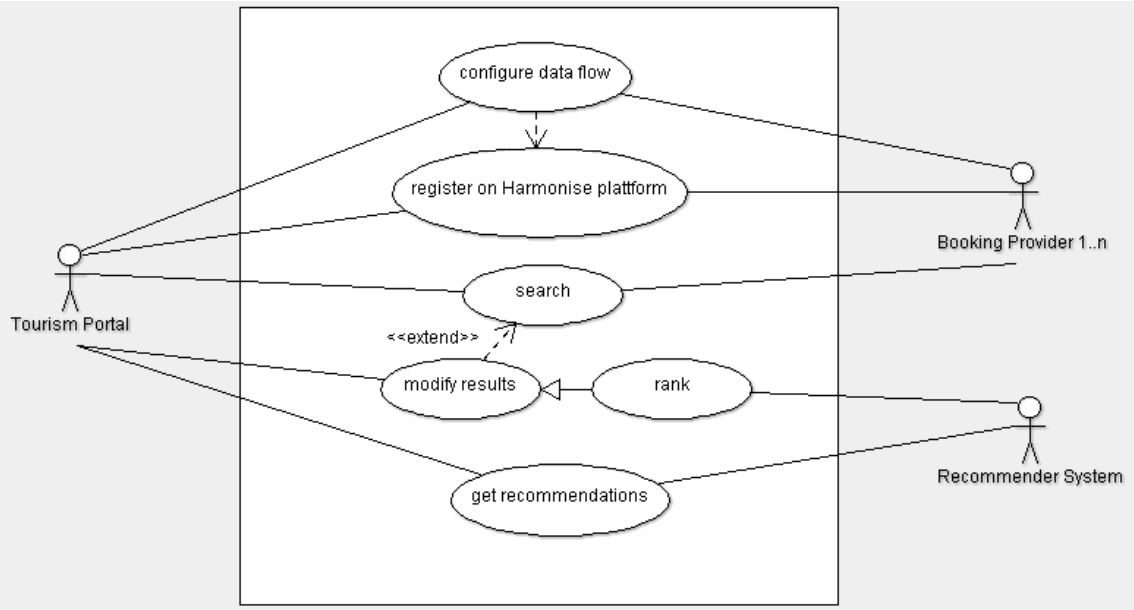

Figure 2: Use case diagram for the metasearch scenario

4.2 METASEARCH USE CASE – HARMONISE REGISTRATION AND SETUP

Use Case Id:

- MS-1

Use case goal:

The booking providers and the tourism portal register on Harmonise and set up the data flow.

Actors:

- Booking providers (data providers dealing with bookable data items, e.g. accommodation)
- Tourism Portal

Use case summary:

The booking provider and the tourism portal register on the Harmonise platform, creating an account each and set it up (i.e. uploading the mappings for data translation and query translation). Furthermore, the booking providers also specify what kind of data they can supply (i.e. what kind of search queries they can answer).

After setting up the account, all partners with whom a data flow can be established are configured on the Harmonise platform. This means that for the portal the partners to whom a query is sent are specified and for the data providers those partners which are allowed to query are configured.

Use case diagram:

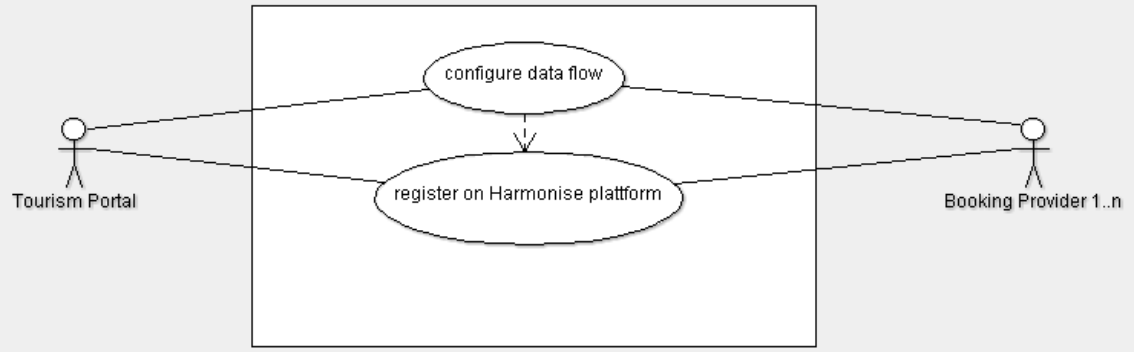

Figure 3: Use case diagram for Harmonise registration and setup

Prerequisites:

- In order to set up the data flow (see basic flow of events, item #3), the data provider needs to implement a connector to Harmonise which can take a search query from Harmonise as input and deliver search results back to Harmonise.

Basic flow of events:

(1) The Harmonise participant (tourism portal or booking provider) creates a new account on the Harmonise platform as dictated by the Harmonise registration process.

(2) The Harmonise participant configures the newly created account as specified by Harmonise guidelines (i.e. uploading the mapping, etc.).

(3) The Harmonise participant searches for the partners with whom a data exchange is desired and a contract is established. The prerequisite of this step is that all partners have completed step 2.

(4) The Harmonise participant configures the data flow on Harmonise

i. The tourism portal configures which data providers (with whom a contract is established) are queried by search queries sent to Harmonise.

ii. The booking providers configure from which sources they accept queries. This can include and exclude specific partners or a broader selection like “all data consumers from my country”.

Alternative paths:

(2a) During setup the booking provider can also specify which kind of search query can be answered (i.e. queries for accommodation from a specific country only).

4.3 METASEARCH USE CASE – BOOKABLE ITEMS SEARCH

Use Case Id:

- MS-2

Use case goal:

The goal is to allow a tourism portal to offer its users the ability to search for bookable data items offered by specific booking providers.

Actors:

- Tourism Portal
- Booking providers

Use case summary:

A user on the tourism portal enters specific search criteria. These include the type of bookable item to be found and specific criteria for the kind of item. For example in case of accommodation search this can include the number of persons and rooms, stay date and duration as well as specific criteria like number of stars or specific amenities (e.g. air conditioning).

From the tourism portal's side this results in a search query for bookable data in the "query language" (i.e. the terms) of the tourism portal.

This query is sent to the Harmonise system and distributed to the configured booking providers with whom a contract (between the tourism portal and the accommodation provider) is established. Harmonise translates the query from the tourism portal's terms to the terms of the booking providers.

Booking providers take the query and deliver the corresponding results (available bookable items, booking price and "deep link" for booking at the provider's site) back to Harmonise. Harmonise translates the search results into the terms of the tourism portal and delivers the results to the portal in an asynchronous way.

The tourism portal displays the results to its users, where duplicates are handled on the portal. Therefore, each bookable item must have a unique ID for each booking provider.

Use case diagram:

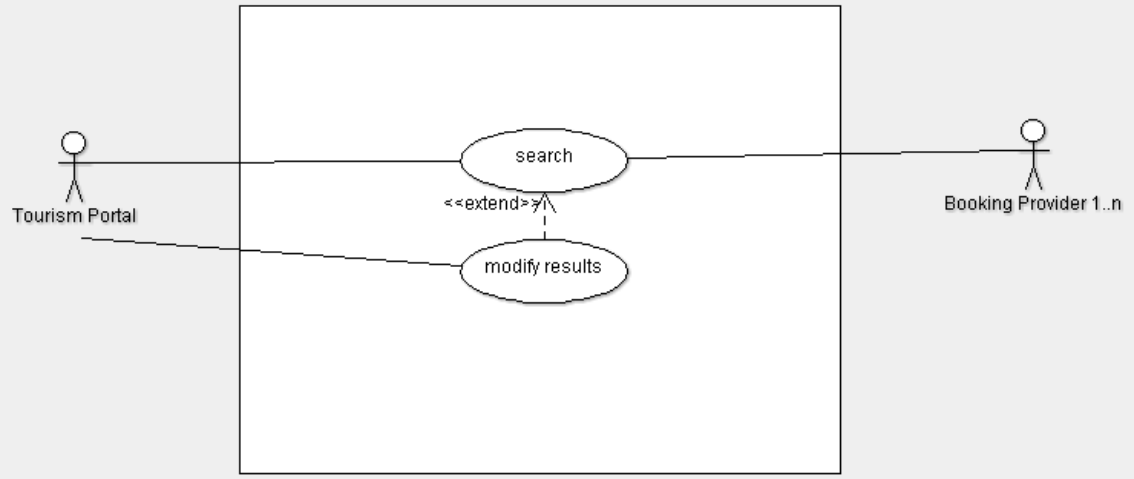

Figure 4: Use case diagram for the search for bookable data items

Basic flow of events:

(1) User of the tourism portal enters search criteria thus generating a query in the terms of the tourism portal. The portal sends the query to Harmonise.

(2) Harmonise translates the query into the local format of the configured booking providers.

(3) Harmonise sends the query to the configured data providers in their respective local formats.

(4) Booking providers execute the query on their local systems and send the results back to Harmonise as answer to the query sent by Harmonise. This is "live" data including availability and price information.

(5) Harmonise translates the search results from the data providers' local formats to the format of the tourism portal.

(6) Harmonise sends the translated search results to the tourism portal as answers to the portal's original query as soon as they are received from the booking providers (i.e. asynchronously).

(7) The tourism portal displays the search results to its users, including availability and price information and a link for booking the requested item on an external booking platform.

Alternative paths:

The user of the tourism portal does not want to wait for further search results and stops the search on the tourism portal. In this case the portal should be able to send a "stop" message to Harmonise in order to stop further answers from arriving. Harmonise should in this case either notify the data providers that the query no longer needs to be answered or simply discard any further search results received from the data providers.

4.4 EXTENDED METASEARCH USE CASE – MODIFY RESULTS

Use Case Id:

- MS-2a

Use case summary:

This use case extends the use case MS-2, where a search on Harmonise has been conducted by a tourism portal. Harmonise offers a possibility to paginate and sort the search results for the portal.

Basic flow of events:

(1) A search is conducted – see use case MS-2

(2) The results are delivered to the tourism portal in an asynchronous way. In the search query the portal can specify the sort order for (partial) results, e.g. by price, which are delivered by Harmonise in this sort order.

(3) After the search is stopped, the tourism portal can also ask for a specific part of the search results, i.e. search results number 21-30, ordered by the result order specified in the query. This effectively handles pagination for the tourism portal. Any request for specific search results has to stop the (asynchronous) query if it hasn't stopped already, since otherwise new results will modify the pagination in an unforeseeable way.

4.5 EXTENDED METASEARCH USE CASE – RANKING ITEMS BASED ON THE USER PROFILE

Use Case Id:

- MS-2b

Use case summary:

As extension of the MS-2 use case, the tourist portal wants to show to the user items ranked in a personalised way exploiting a recommender system which could be provided by a service provider. Items retrieved from a query are then ranked exploiting a given recommendation algorithm provided by the Recommender System.

Use case diagram:

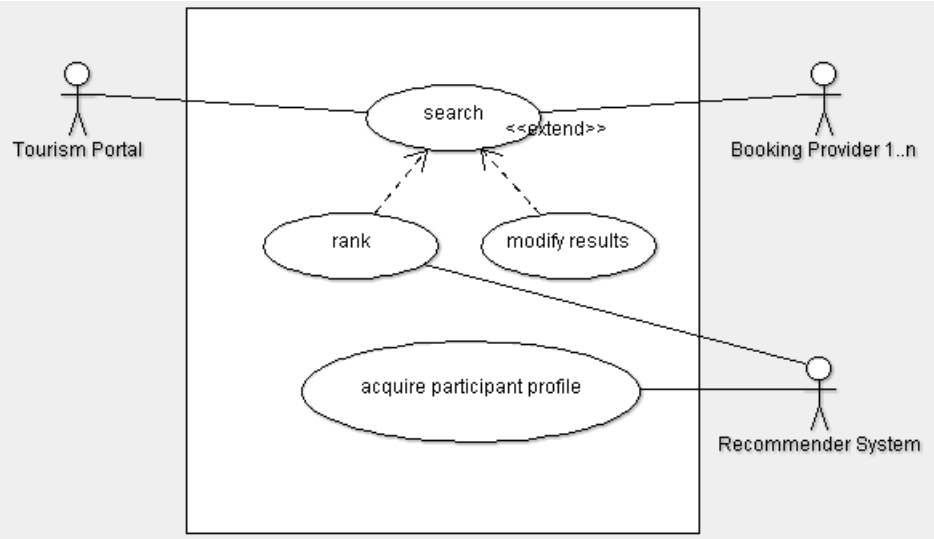

Figure 5: Use case diagram for user specific search result ranking

Basic flow of events:

(1) A search is conducted – see use case MS-2 and its extension MS-2a. In the search query has been specified as sort criteria the one provided by a specific recommender system selected in the Tourism Portal configuration preferences. Possibly, additional profile information of the user actually performing the query have been specified in the search query.

(2) The Harmonise system asks the selected recommender system to sort the result set.

(3) The recommender system acquires from Harmonise the semantic profile of the querying tourism portal.

(4) The recommender exploiting the semantic profile of the querying tourism portal and possibly additional profile information of the querying user sorts the result set and sends it back to the Harmonise Portal.

(5) An additional deep link allows the recommender system to provide additional feedback on the recommendation to the tourism portal.

4.6 METASEARCH USE CASE – ITEMS RECOMMENDATIONS

Use Case Id:

- MS-3

Use case summary:

The user of a Tourism Portal desires to get recommendations about items related to a specific topic of interest. The portal provides contextual information (for example geo-coordinates or specific theme of interest), some constraints and receives backs items best fitting the provided information. The user could be interacting also through a mobile device, thus on a mobile portal or with a mobile application.

Use case diagram:

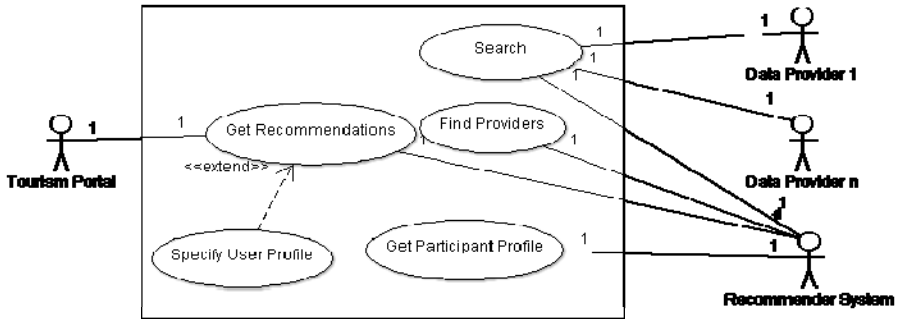

Figure 6: Use case diagram for item recommendation

Basic flow of events:

(1) The recommendation service, a specific service registered in the Harmonise Network, is invoked by the tourism portal.

(2) The tourism portal specifies contextual and specific user profile information.

(3) Harmonise portal invokes the recommender system, since it is this system which provides this specific functionality.

(4) The recommender system acquires profile information of the Tourist Portal, acquires from the semantic registry information about possible content providers exposing the desired content.

(5) The recommender system queries those providers, selects the most suited items and returns back the results to the tourism portal.

5 USE CASES FOR IMPORT SCENARIO

5.1 USE CASE DIAGRAM

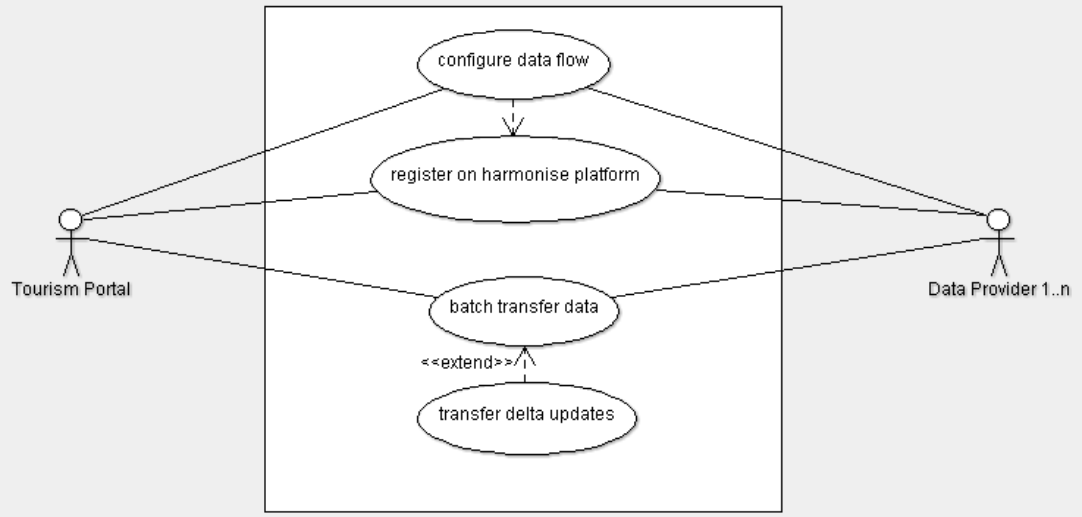

Figure 7: Use case diagram for the import scenario

5.2 IMPORT USE CASE – HARMONISE REGISTRATION AND SETUP

Use Case Id:

- IMPORT-1

Use case goal:

The data providers and the tourism portal register on Harmonise and setup the data flow. This use case is very similar to the use case MS-1, though it does differ in details.

Actors:

- Data providers (providing static information about the data items of interest)
- Tourism Portal

Use case summary:

The data providers and the tourism portal register on the Harmonise platform, creating an account each and setting it up (i.e. uploading the mappings for data translation and query translation). Furthermore, the data providers also specify what kind of data they can supply (i.e. what kind of search queries they can answer).

After setting up the account, all partners with whom a data flow can be established are configured on the Harmonise platform. In this specific case the setup incorporates the information about what data is to be retrieved (e.g. static accommodation data) and what the requirements are for the transfer (i.e. does not need to be delivered quickly).

Use case diagram:

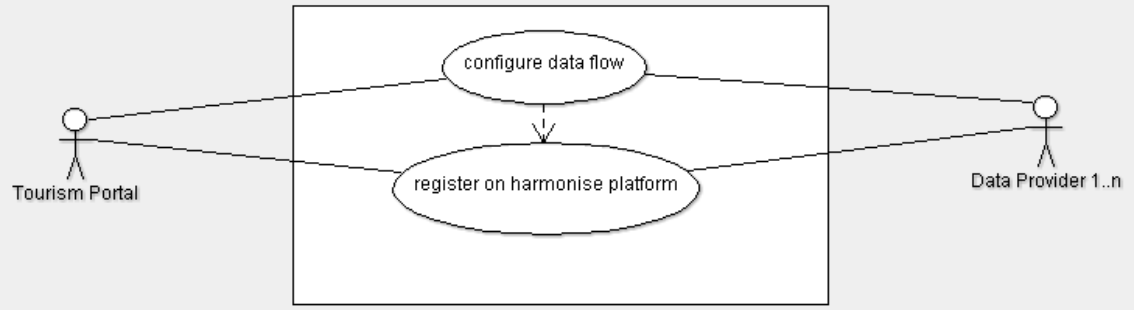

Figure 8: Use case diagram for Harmonise registration and setup

Prerequisites:

- In order to set up the data flow (see basic flow of events, item #3), the data provider needs to implement a connector to Harmonise which can send all available (static) data for a specific search query to Harmonise. This search query could limit data items to a specific country or requires having a specific property, but cannot ask for dynamic criteria like, e.g. for availabilities on a given date or date-range.
- The data provider can implement the possibility to send only data which has been created or changed since a given date. In this case also indications for data which has been deleted since this date have to be provided.

Basic flow of events:

Note: events 1 to 3 are identical to the corresponding events from use case MS-1.

(1) The Harmonise user (tourism portal or accommodation provider) creates a new account on Harmonise as dictated by the Harmonise registration process.

(2) The Harmonise user configures the newly created account as specified by Harmonise guidelines (i.e. uploading the mapping, etc.).

(3) The Harmonise user searches for the partners with whom a data exchange is desired and a contract is established. The prerequisite of this step is that all partners have completed step 2.

(4) The Harmonise user configures the data flow on Harmonise

i. The tourism portal configures the accommodation providers to whom the query triggering the data transfer is to be sent.

ii. The accommodation (data) provider sets up the tourism portal as allowed querying for batch transfer of static accommodation data. This especially takes relaxed requirements for response times into account.

Alternative paths:

(2a) During setup the accommodation provider (data provider) can also specify whether the query can encompass a date that indicates which accommodations are of interest. All accommodations changed or created since this date are to be transferred.

5.3 IMPORT USE CASE – BATCH TRANSFER OF STATIC DATA

Use Case Id:

- IMPORT-2

Use case goal:

The goal of the use case is to transfer static information (e.g. in case of accommodation: accommodation name, description, location, pictures, amenities, etc.) from the data providers to the portal in regular intervals.

Actors:

- Tourism Portal
- Data providers

Use case summary:

The tourism portal sends a specific query to Harmonise in regular intervals asking for batch transfer of static data of a predefined kind. The time at which this batch import is triggered depends on the load of the tourism portal's servers. The query can encompass specific constraints (e.g. "accommodations from Finland"). In addition the query can contain a time when the last update was done, therefore indicating that only changes since this time are of relevance (a so-called *delta update*). Each data item must have an ID which is unique within the scope of the data provider.

Harmonise translates the query to the local data format of the data providers and delivers it to the configured partners. The partners respond to the query with sending the corresponding data, taking the possibility of a delta update into account if possible. The responses are sent back to the tourism portal through Harmonise. Finally the tourism portal imports the data into its CMS and triggers further processing (e.g. local duplicate detection).

Use case diagram:

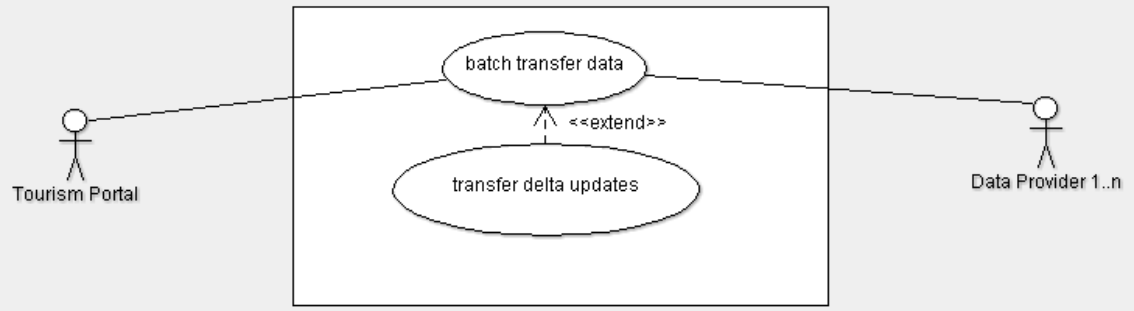

Figure 9 Use cases diagram for the import of static data

Prerequisites:

- All partners must have completed the steps described in use case IMPORT-1.

Basic flow of events:

(1) Tourism portal sends a query to Harmonise asking for batch transfer of static data from previously configured sources. The query is sent in the local format of the tourism portal and can encompass a date for delta updates.

(2) Harmonise translates the query to the local format of the data providers. It then distributes the query to the previously configured partners.

(3) The data provider answers the query by sending the relevant data to Harmonise in the provider's local format and in a batch.

(4) Harmonise translates the responses into the local format of the tourism portal and delivers it.

(5) The tourism portal receives the response and imports it into its CMS. Further data processing on the tourism portal is triggered (e.g. duplicate detection).

Extensions for use case "Transfer Delta Updates":

(1a) The query sent indicates that a delta update is preferred and states the date for which the delta update should be sent.

(3a) The data provider sends only the data items which have been changed or added since the indicated date. For data which has been deleted since this date a corresponding delete message must be sent. The data is also sent in a batch.

6 USE CASES FOR PUBLISHING SUBSCRIBE SCENARIO

6.1 USE CASE DIAGRAM

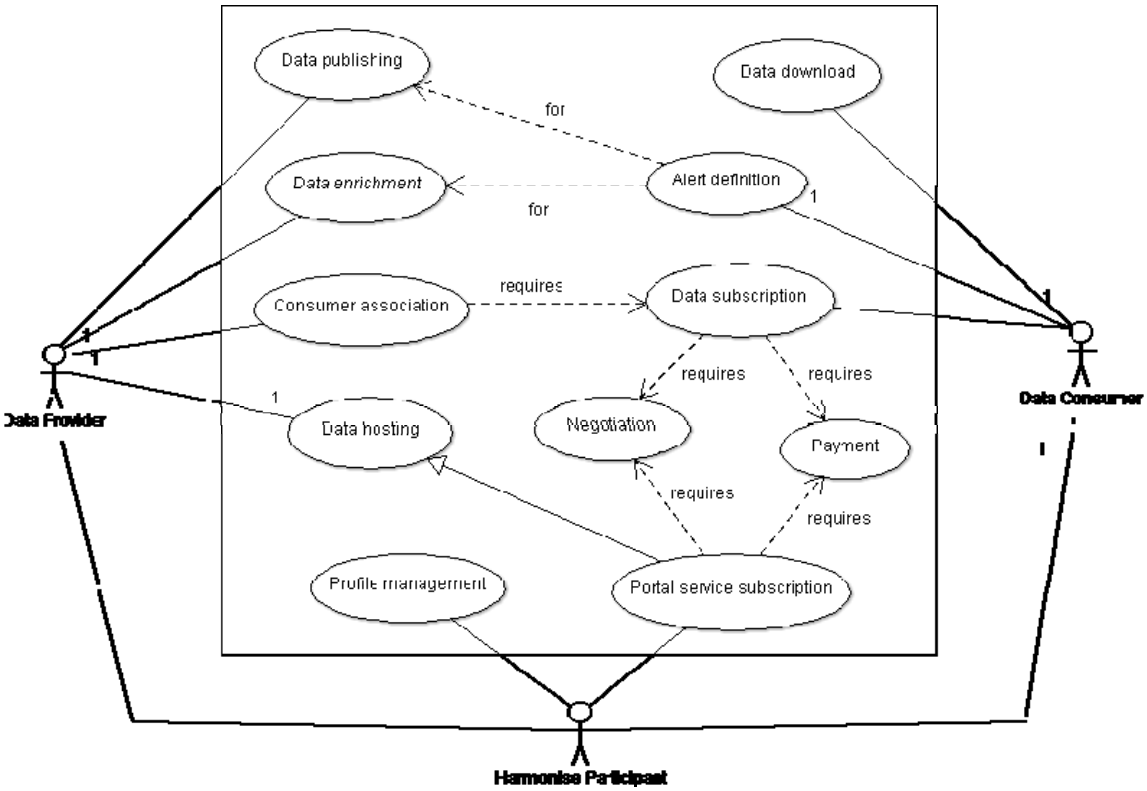

Figure 10: Publish and Subscribe Scenario – Use Cases overview

6.2 DATA HOSTING USE CASE

Use Case Id:

- PS-1

Use case goal:

This is to allow data providers to easily upload once or on a regular basis their data in the Harmonise portal.

Actors:

- Data provider

Use case summary:

A data provider wants to upload his data in the Harmonise network. Data may be uploaded once manually or regularly scheduled. A data provider wants to program regular batch uploads (for static data, for price and availability data, for booking synchronization …) or allow retrieval via web services, at certain times, on events... Each batch may contain one or several files. There may be several batches of different frequencies. User wants to be able to define batches, events.

Use case diagram:

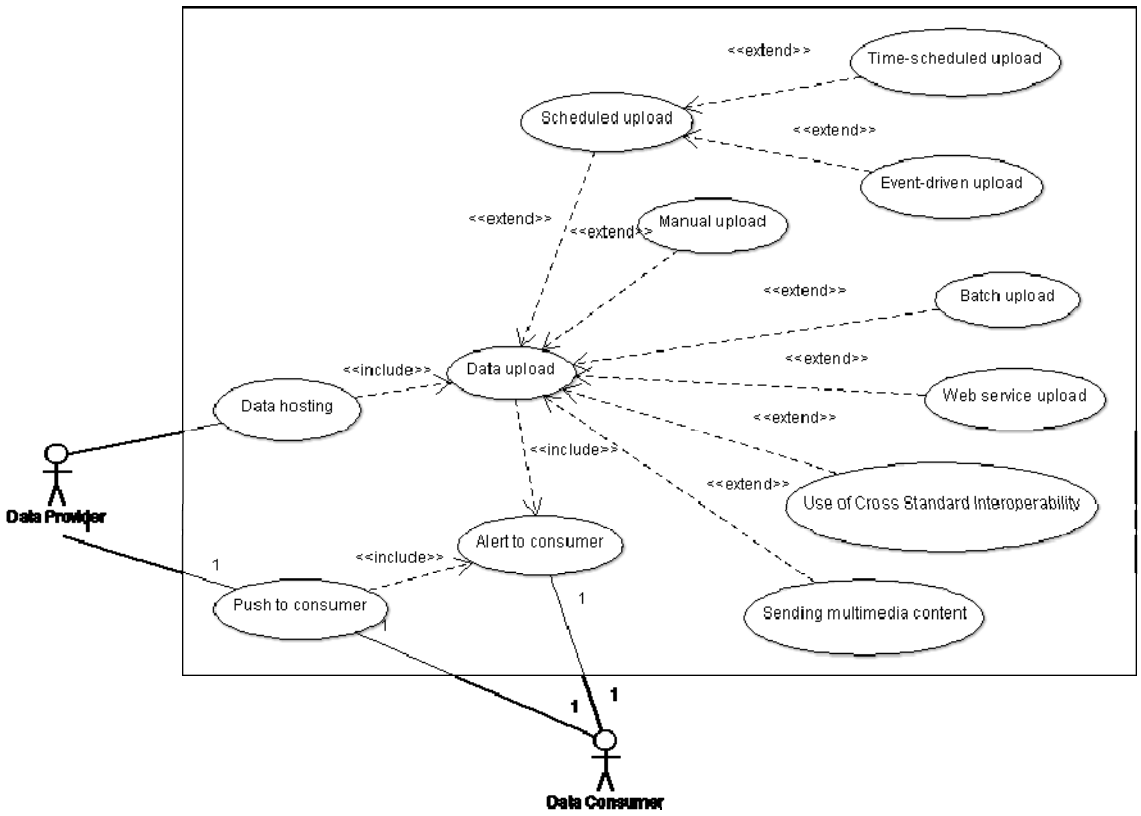

Figure 11: Data Hosting Use Case

Basic flow of events:

(1) Data provider selects the data files to be uploaded.

(2) Data provider defines a batch program.

(3) Data provider configures the scheduler in order to regularly launch the batch program at the specified time.

(4) Data are sent to the Harmonise portal.

(5) Harmonise system transforms the data according to Harmonise ontology using the mapping provided by the data provider.

(6) Data are stored in the portal.

(7) If the alert service is enabled, an alert is sent to the consumers which have subscribed to this kind of data profile.

Alternative paths:

(2a) Data provider configures a web service to automatically retrieve the data to be uploaded.

(3a) Data provider defines an event as a trigger for the batch/web service upload.

(3b) Data provider uploads once manually his data on the portal.

(6a) Data provider may want to push data directly to one or more consumers instead of uploading them to the portal. In this case data are transformed according to the recipient's data structure before delivery using the mapping provided by the data consumer.

6.3 EXTENDED DATA HOSTING USE CASE – USE OF CROSS STANDARD INTEROPERABILITY

Use Case Id:

- PS-1a

Use case summary:

This use case extends the use case PS-1, where a data provider wants to upload his data in the Harmonise portal. When a data provider sends meta-data or categories (e.g. location), this data is often stored in different categories on the recipients side. The Harmonise service centre shall allow storing mappings between reference lists which other mappings, used by customers, can refer to. This shall allow e.g. to translate from region names or geo-codes to postal codes by referring to these reference lists - thus a user can have a region translated to a geo-code.

Basic flow of events:

(1) Data provider selects the data files to be uploaded. This data files must be compliant to one of the standard supported by HarmoSearch.

(2) Data provider uploads his data on the portal (manually or scheduled) specifying the standard to which they are compliant to.

(3) Data are sent to the Harmonise portal.

(4) Harmonise system transforms the data using the pre-configured mapping from the standard specified by the data provider to the Harmonise ontology.

(5) Data are stored in the portal.

(6) If the alert service is enabled, an alert is sent to the consumers which are subscribed to this kind of data profile.

Alternative paths:

(5a) Data provider may want to push data directly to one or more consumers instead of uploading them to the portal. In this case data are transformed according to the recipient's data structure before delivery, possibly using a pre-configured mapping if the consumer data structure is compliant with one of the standard supported by HarmoSearch.

6.4 EXTENDED DATA HOSTING USE CASE – SENDING MULTIMEDIA CONTENT

Use Case Id:

- PS-1b

Use case summary:

This use case extends the use case PS-1, where a data provider wants to upload his data in the Harmonise portal. When a data provider sends some data, he may want to add multimedia files to the content transferred (included as zipped package or encoded, not linked).

Basic flow of events:

(1) Data provider selects the data files to be uploaded. This data files include also multimedia content.

(2) Data provider uploads his data on the portal (manually or scheduled) as a single package which includes textual data and multimedia files.

(3) Data are sent to the Harmonise portal.

(4) Harmonise system un-packages the file and transforms the textual files using the mapping provided by the data provider. Multimedia files are transferred without reconciliation.

(5) Data are stored in the portal.

(6) If the alert service is enabled, an alert is sent to the consumers which are subscribed to this kind of data profile.

Alternative paths:

(5a) Data provider may want to push data directly to one or more consumers instead of uploading them to the portal. In this case data are transformed according to the recipient's data structure before delivery, using the mapping provided by the data consumer. Multimedia files are transferred without reconciliation.

6.5 DATA PUBLISHING USE CASE

Use Case Id:

- PS-2

Use case goal:

This is to allow data providers to easily publish once or on a regular basis their data in the Harmonise network.

Actors:

- Data provider

Use case summary:

A data provider wants to publish data in the Harmonise network. This use case is similar to the Data Hosting (PS-1), but in this case what is uploaded and stored in the Harmonise portal are not the real data but some meta-data which are useful to describe what kind of data the data provider has to offer.

Use case diagram:

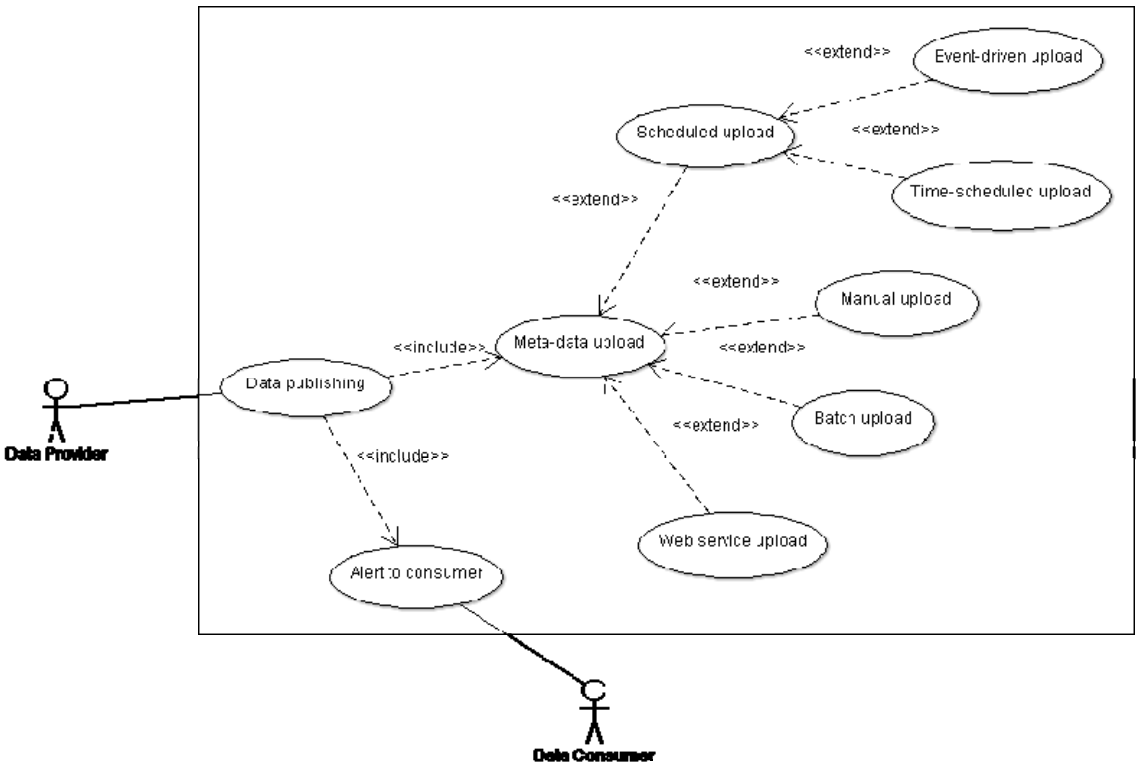

Figure 12 Data Publishing Use Case

Basic flow of events:

(1) Data provider selects the data files to be uploaded.

(2) Data provider defines a batch program.

(3) Data provider configures the scheduler in order to regularly launch the batch program at the specified time.

(4) Data are sent to the Harmonise portal.

(5) Harmonise system transforms the data according to Harmonise ontology using the mapping provided by the data provider.

(6) Data are stored in the portal.

(7) If the alert service is enabled, an alert is sent to the consumers which are subscribed to this kind of data profile.

Alternative paths:

(4a) Data provider configures a web service to automatically retrieve the data to be uploaded.

(5a) Data provider defines an event as a trigger for the batch/web service upload.

(4b) Data provider uploads once manually his data on the portal.

(7a) Data provider may want to push data directly to one or more consumers instead of uploading them to the portal. In this case data are transformed according to the recipient's data structure before delivery using the mapping provided by the data consumer.

6.6 DATA ENRICHMENT USE CASE

Use Case Id:

- PS-3

Use case goal:

To allow data providers to transform their data according to services they would have subscribed and to associate profiles to the data so as to allow alert to be generated by the portal to non subscribers.

Actors:

- Data provider

Use case summary:

Data provider must be able to update data and to manually or automatically associate additional information to its data to create data profiles to be used to facilitate subscription, to link data together (for cross selling or up selling features, to create logical links...). Automatic enrichment can be performed via information already present in the data itself (like product name, location information, categories and codes...).

Use case diagram:

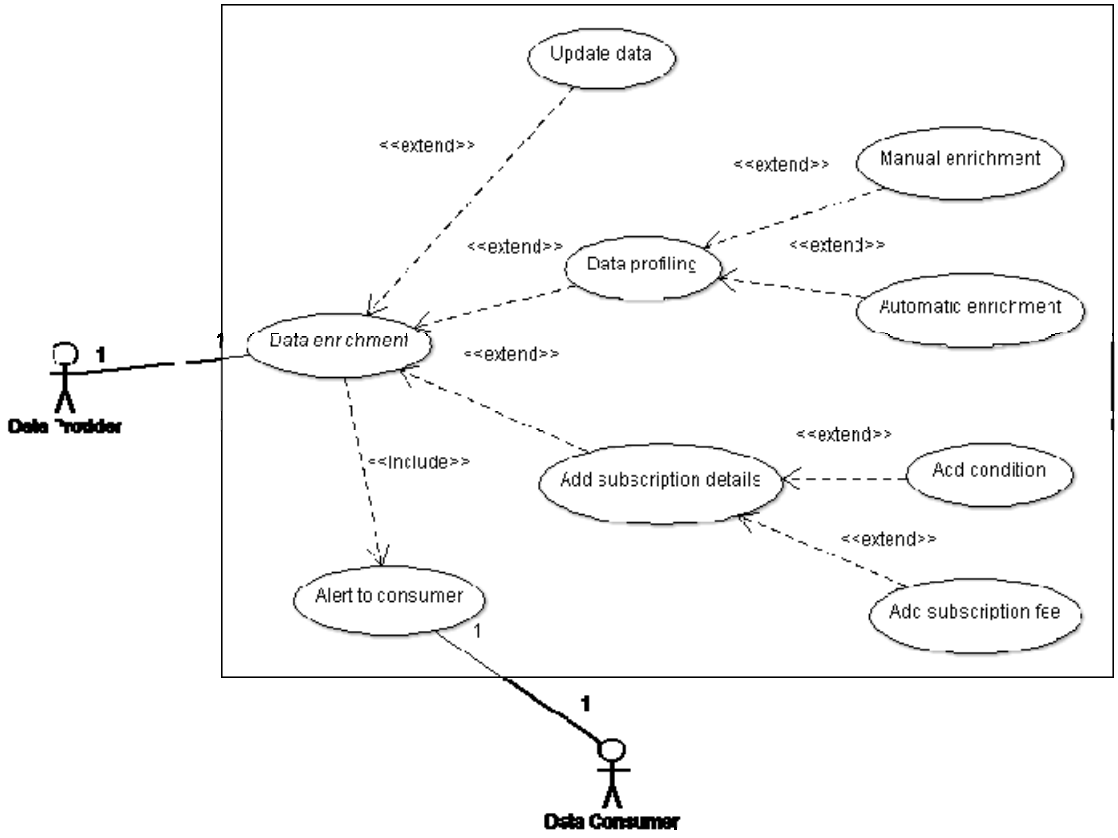

Figure 13: Data Enrichment Use Case

Basic flow of events:

(1) Data provider accesses to the data he uploaded on Harmonise portal.

(2) Data provider updates data information.

(3) Data provider associates one or more tags to his data, in order to profile the information and to link data together.

(4) Data provider associates subscription details (conditions, fees) to his data, possibly indicating whether they are negotiable or not.

(5) If the alert service is enabled, an alert is sent to the consumers which are subscribed to the updated data.

Alternative paths:

(3a) Data provider can automate data enrichment by extracting tag information from specific fields already present within the data.

6.7 CONSUMER ASSOCIATION USE CASE

Use Case Id:

- PS-4

Use case goal:

To allow data providers to define the consumers that will be allowed to download their data.

Actors:

- Data provider
- Data consumer

Use case summary:

Data provider wants to be able to define to whom its data will be distributed, either per batch or per data unit. Association may be performed by direct association of data consumers with data (with or without negotiation) or through data information or data profile mapping (like product name, location information, categories and codes…).

Use case diagram:

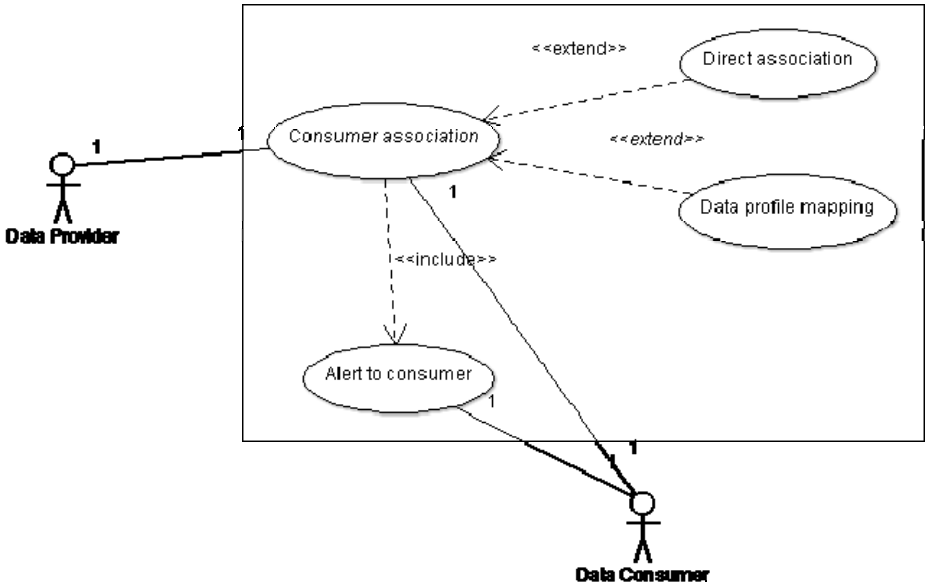

Figure 14: Consumer Association Use Case

Basic flow of events:

(1) Data provides receives a subscription request by a consumer for some of his data.

(2) Data provider accesses to the data he uploaded on Harmonise portal.

(3) Data provider associates the data to the consumer who subscribed to them.

(4) If the alert service is enabled, an alert is sent to the consumer.

Alternative paths:

(1a) Data Provider can make his data public without requiring any subscription.

(3a) Data provider can associate consumers to data profiles. In this case he links a consumer to a tag and the consumer can access and use every data enriched with that tag.

6.8 DATA DOWNLOAD USE CASE

Use Case Id:

- PS-5

Use case goal:

This is to allow consumers to download data pushed by data providers once or regularly.

Actors:

- Data consumer

Use case summary:

A data consumer wants to schedule and define download content and methodology (batch, web service, at fixed time, on events…), selecting the content to import, possibly grouping data from different providers, using differential input or even possibly using web services instead of files. A consumer batch may correspond to

one owner batch or aggregates several. Consumer shall be able to define data filters to limit the amount of data to be received from the different data sources he has subscribed to. Schedules are different from owner batch. In some cases, alert of data present can be sent to consumer. Alert on data changes could also be sent to data consumers for them to be able to refresh their data between batches.

Use case diagram:

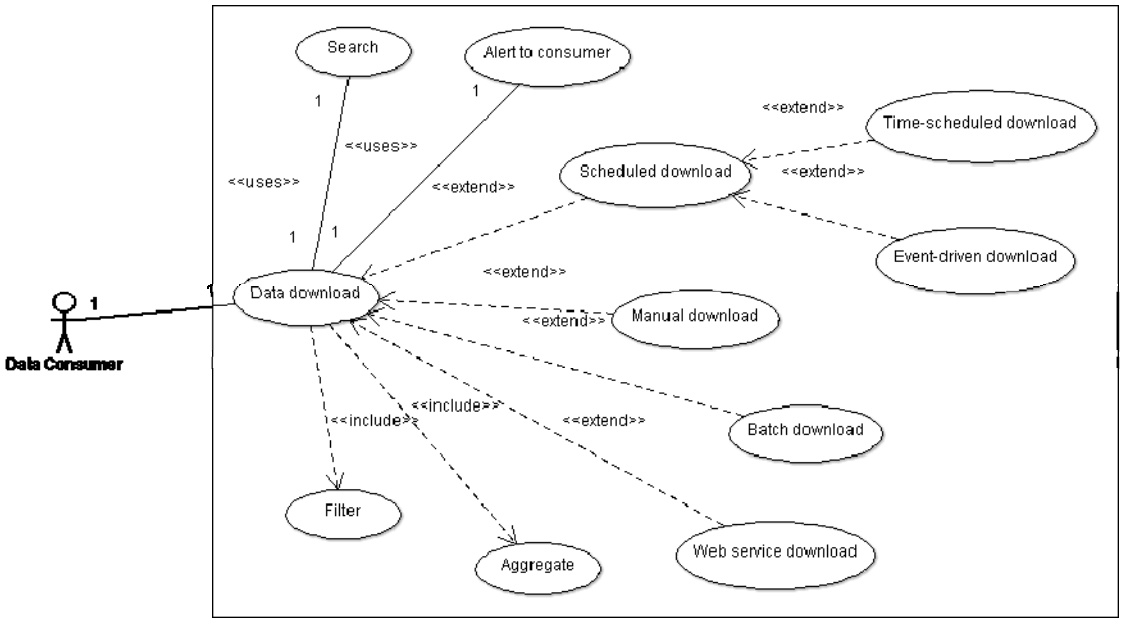

Figure 15: Data Download Use Case

Basic flow of events:

(1) Data consumer searches the data he is interested in through a query on Harmonise portal.

(2) Data consumer selects the data to be downloaded as one or several batches, possibly grouping data from different providers.

(3) Data consumer defines a filter to limit the amount of data to be received.

(4) Harmonise gets the data, transforms them according to the consumer's data structure and delivers them to the consumer.

Alternative paths:

(1a) If the alert service is enabled, the system alerts the data consumer that new data matching the consumer criteria or that an update to already downloaded data is available.

(1b) Harmonise scheduler starts a download process, as previously configured by the consumer, on a specified time/every specified period.

(1c) Harmonise event handler starts a download process, as previously configured by the consumer, when one or more specified events happen.

(3a) Data consumer's criteria for filtering data is defined in a file.

(4a) Harmonise system automatically pushes the downloaded data to the consumer system using a web service previously configured by the consumer.

6.9 DATA SUBSCRIPTION USE CASE

Use Case Id:

- PS-6

Use case goal:

This is to allow consumers to easily subscribe to row or enhanced data from different providers.

Actors:

- Data consumer
- Data provider

Use case summary:

Data consumers need to be able to subscribe or unsubscribe to data. They can find relevant data because they are invited by data providers, because of alerts they receive or because of manual searches. Once data have been selected, subscription may occur directly or after negotiation has taken place. Data consumers need also alerts when subscribed data is updated or when new data corresponding to their profiles is published.

Use case diagram:

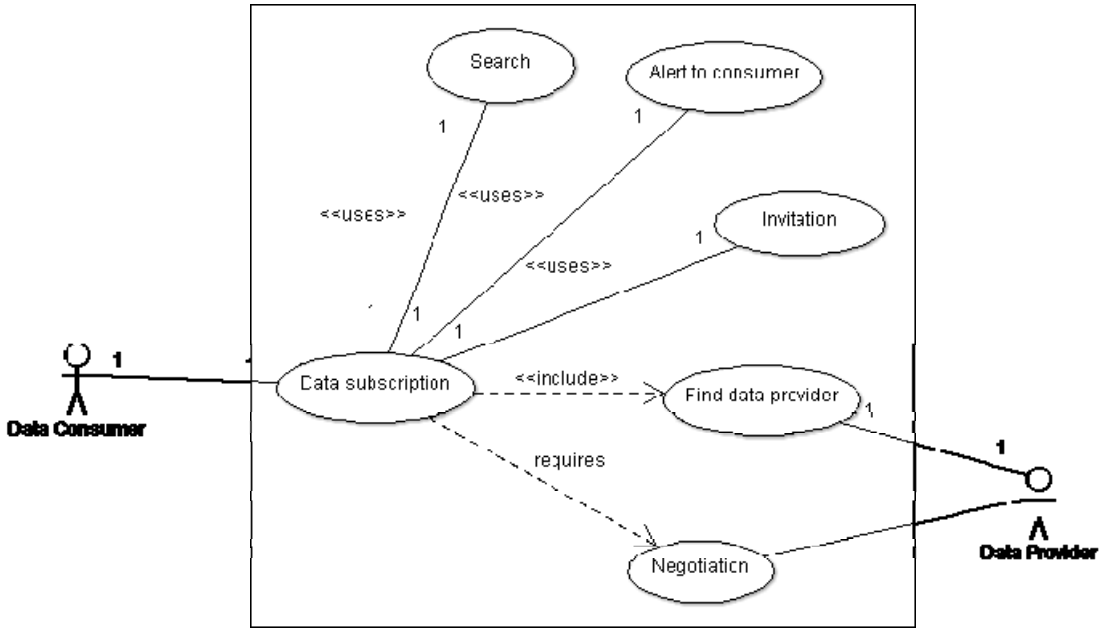

Figure 16: Data Subscription Use Case

Basic flow of events:

(1) Data consumer selects the data or data profile he wants to subscribe to after being invited by a data provider or through an alert or a manual search.

(2) Data consumer searches for the (best) data provider.

(3) Data consumer starts the negotiation process with the data provider selected that may end up with a contract or financial agreement.

(4) Data consumer subscribes to the requested data or data profile and starts to use them.

(5) If the alert service is enabled, the consumer receives an alert on data update or when new data corresponding to the data profile the consumer has subscribed to is available.

Alternative paths:

(1a) If the data consumer wants to unsubscribe to a data or data profile, he accesses the subscription management section of his profile on the Harmonise portal. In this section the data consumer can also modify other subscription parameters such as alerts, triggers, data filters, etc.

(3a) Some data or data profile may be free of charge and may not require a negotiation.

(4a) Negotiation between data consumer and provider fails and the consumer cannot access and use the data.

6.10 ALERT DEFINITION USE CASE

Use Case Id:

- PS-7

Use case goal:

To allow consumers to define alerts and triggers in order to be notified if relevant events occur or if new services or data are available which may be of interest for him according to his profile.

Actors:

- Service Consumer

Use case summary:

Service Consumers should be able to define alerts using a set of keywords or data profiles so as to receive notifications in case those keywords or data are available. This may lead to a negotiation or a subscription. Moreover consumers can store search results or queries with trigger information. In case certain information in the result set change or in case new result batch queries, a notification may be sent to the user (by mail, SMS …).

Use case diagram:

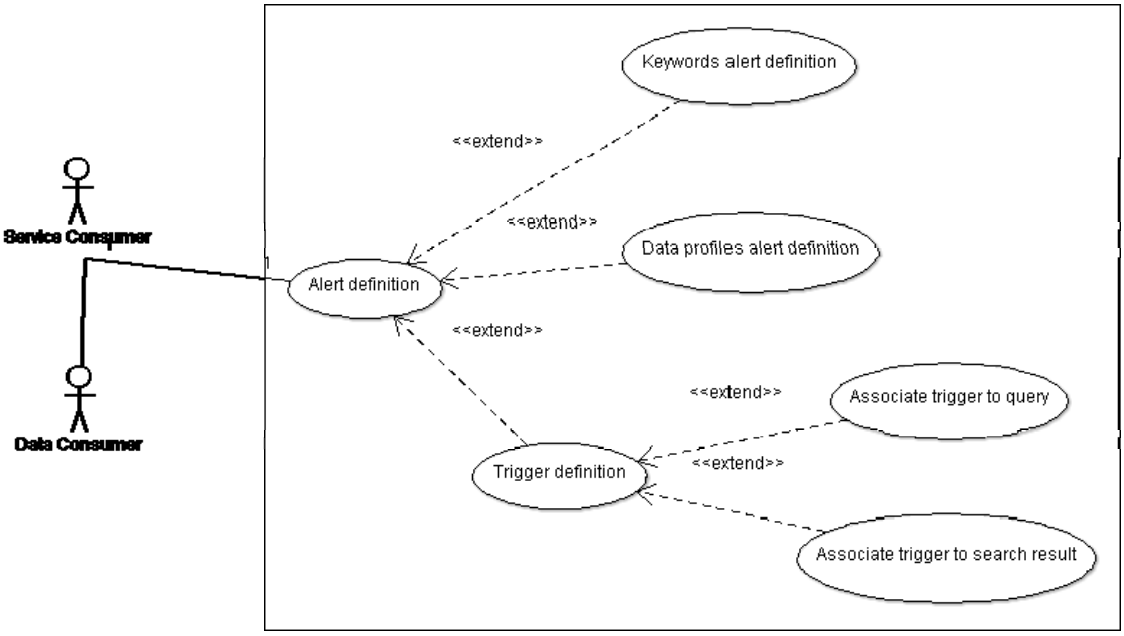

Figure 17: Alert Definition Use Case

Basic flow of events:

(1) Service Consumer accesses the alert definition section on the Harmonise portal.

(2) Service Consumer specifies which kind of alerts he wants to receive. He can define alerts for the availability of new data or updates corresponding to the data profiles he has subscribed to. Or he can define alerts using a set of keywords or tags so as to receive alerts in case those keywords or tags are present in unsubscribed sources.

(3) Service Consumer can modify the alerts he has defined at any time on the Harmonise portal.

Alternative paths:

(2a) Service Consumer specifies which kind of triggers he wants to activate. He can define triggers associated to search results or queries to receive a notification in case certain information in the result set changes or in case new results patch queries.

6.11 PORTAL SERVICE SUBSCRIPTION USE CASE

Use Case Id:

- PS-8

Use case goal:

This is to allow users to easily subscribe to portal services.

Actors:

- Service consumer
- Service provider

Use case summary:

All types of users must be able to subscribe or unsubscribe to the services provided by the portal for part or all of their data. Example of services will be booking, items recommendation, data hosting, data modification, statistical and market analyses, etc. Some services may be free of charge and others may require payments.

Use case diagram:

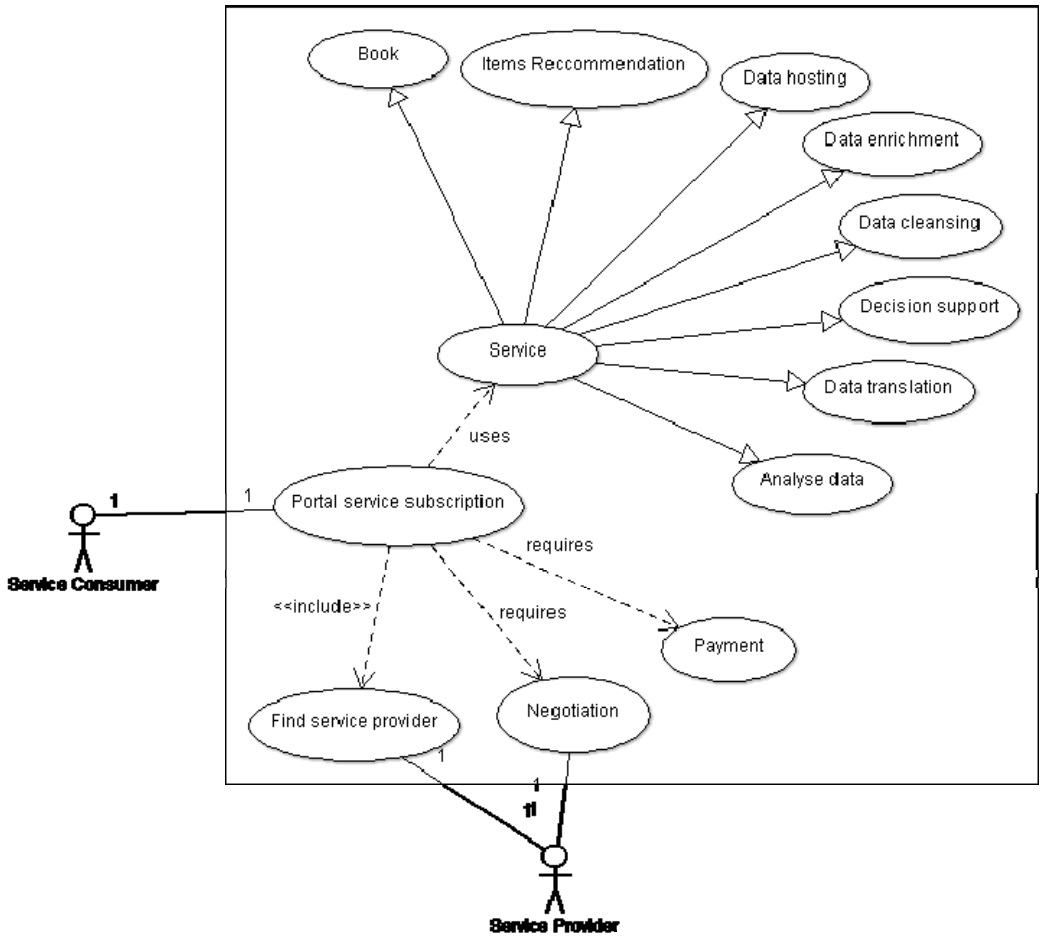

Figure 18: Portal Service Subscription Use Case

Basic flow of events:

(1) Service consumer selects the service he wants to subscribe to.

(2) Service consumer searches for the (best) service provider.

(3) Service consumer starts the negotiation process with the provider selected that may end up with a contract or financial agreement.

(4) Service consumer subscribes to the requested service.

(5) Service consumer may set up a complex service or a service flow by adding additional services and chaining them together.

Alternative paths:

(1a) If the data consumer wants to unsubscribe to a service, he accesses the subscription management section of his profile on the Harmonise portal. In this section the service consumer can also modify other subscription parameters.

(3a) Some services may be free of charge and may not require a negotiation.

(4a) Negotiation between service consumer and provider fails and the consumer cannot access and use the service.

6.12 NEGOTIATION USE CASE

Use Case Id:

- PS-9

Use case goal:

This is to allow negotiating services and data subscription through the portal in order to facilitate subscriptions to new services/data or to new consumers.

Actors:

- Service consumer
- Service provider

Use case summary:

Service and data consumers shall be able to search for specific services or data outside the ones they have subscribed to. If they are interested in certain services/data, they shall be able to initiate negotiation with the service or data provider in order to gain access to them. This may end up by being able to use the service/data after validation by the service/data provider or after a contract and financial agreement has been settled.

Use case diagram:

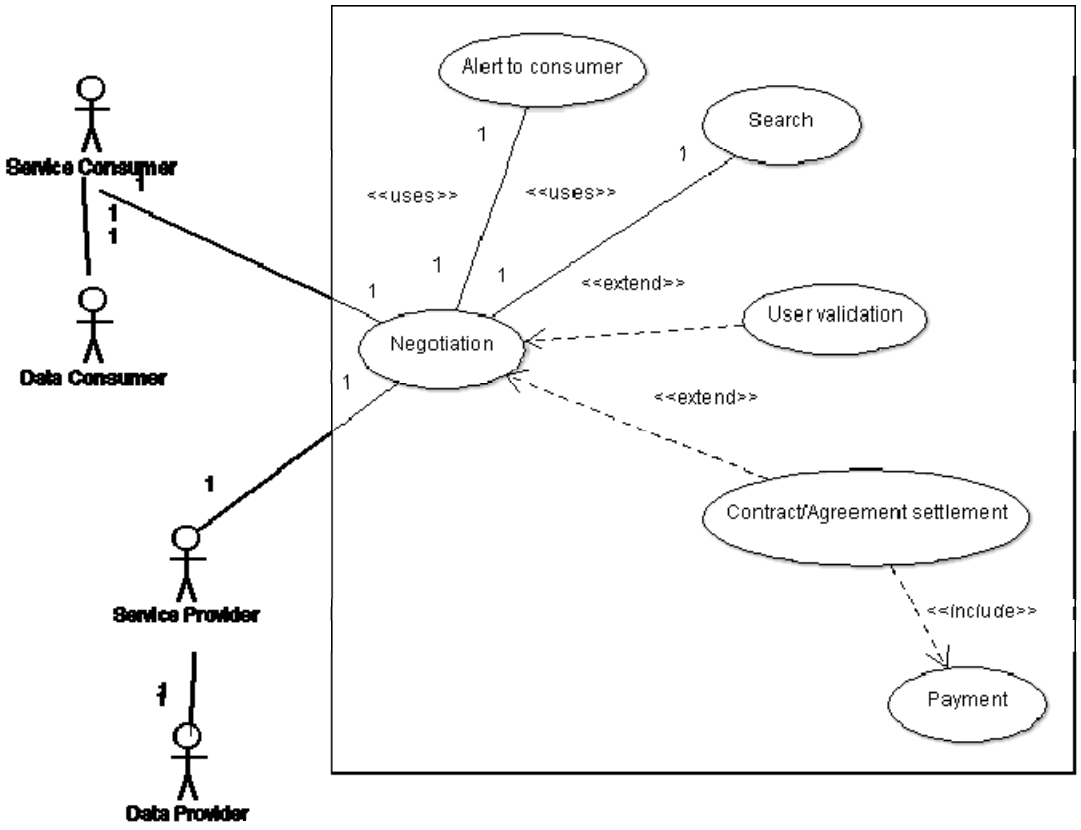

Figure 19: Negotiation Use Case

Basic flow of events:

(1) Service consumer selects the service or data he wants to subscribe to through an alert or a manual search.

(2) Service consumer searches for the (best) service provider.

(3) Service consumer, through the Harmonise portal, sends a negotiation request to the service provider.

(4) Service provider replies to the consumer's request with an offer and the conditions for the usage of service.

(5) Service consumer evaluates the provider's offer and eventually they sign a contract or financial agreement.

(6) Service consumer gains access to the service or data.

Alternative paths:

(4a) Service consumer is able to use the service after a validation by the provider.

(6a) Negotiation between service consumer and provider fails and the consumer cannot access and use the service or data.

6.13 PAYMENT USE CASE

Use Case Id:

- PS-10

Use case goal:

This is to allow providing a way to charge and receive payment for the subscribed services.

Actors:

- Service consumer
- Service provider

Use case summary:

The platform must be able to charge, invoice and take payment for platform services as well as for the reconciliation mechanism between service consumers and providers for the data requiring payment to be used.

Use case diagram:

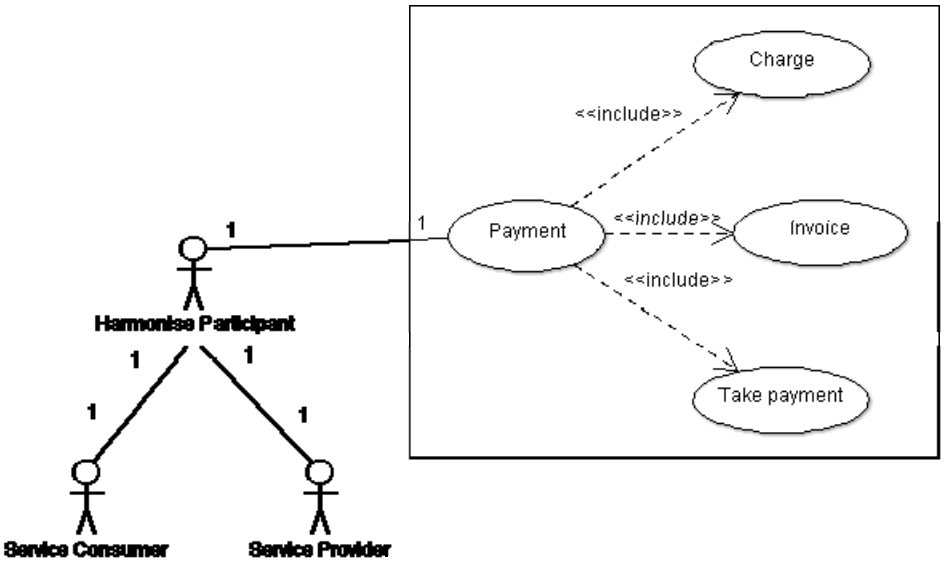

Figure 20: Payment Use Case

Basic flow of events:

(1) Service consumer has been charged for a service or data which requires payment.

(2) Service provider makes an invoice to the consumer.

(3) Service consumer give its payment details to the provider.

(4) Service provider takes the payment and gives back the receipt if requested.

6.14 PROFILE MANAGEMENT USE CASE

Use Case Id:

- PS-11

Use case goal:

This is to allow providing and managing user profiles to store preferences, filters, access to subscribed services and data.

Actors:

- Harmonise participant

Use case summary:

All types of users must be able to have and maintain a user profile containing identity information as well as data profiles, filters, subscriptions, invoice…

Use case diagram:

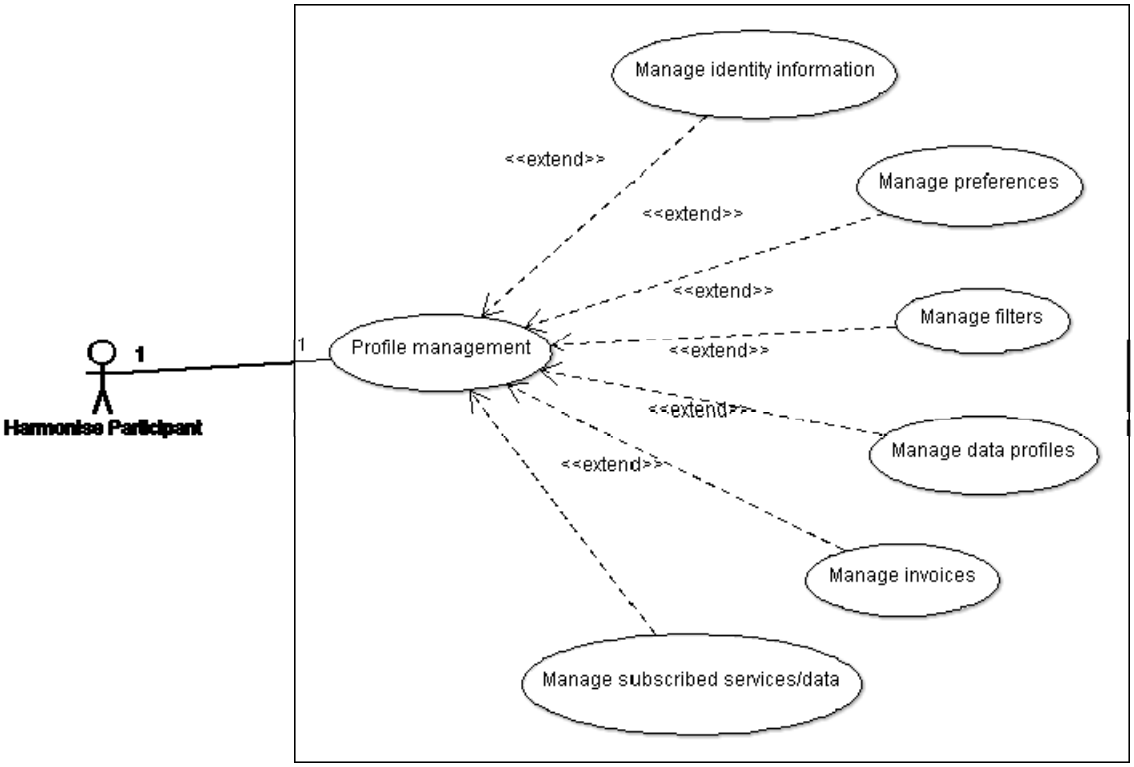

Figure 21: Profile Management Use Case

Basic flow of events:

(1) Harmonise participant accesses the profile management section of the Harmonise portal.

(2) Harmonise participant can view and manage his personal identity information.

(3) Harmonise participant can view and manage his preferences, data filters, data profiles.

(4) Harmonise participant can view and manage the services or data he is subscribed to.

Alternative paths:

(1a) The Harmonise system collects information on the participant's behaviour when he is using the portal content in order to profile his preferences and adds this information to the user profile.

7 USE CASES FOR DATA MODIFICATION

Data modification denotes all kinds of transformation from given source data to desired target data, possibly involving more than one transformation step. A modification can be performed at different levels of the original source data. For instance, decision support such as ranking of search results changes the order of result items in the result set but leaves the content of each item as it was before the transformation was applied. Data mapping, data enrichment, and data cleansing may or may not change the overall structure, but will change the structure and the content of source data items. Finally, language translation from a particular natural language to another language and other forms of data conversions such as currency conversion will leave the global data structure unchanged, but will perform modification on the content.

A Harmonise participant should be able to register his/her own data transformation. In addition, various services may be provided that can perform data modification on source data on behalf of a consumer.

7.1 USE CASE DIAGRAM

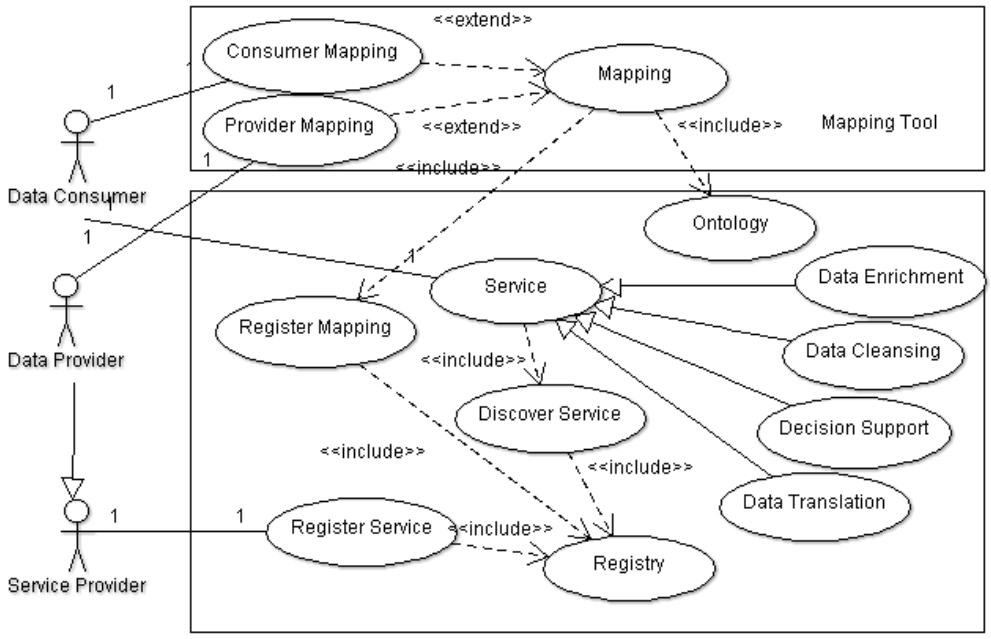

Figure 22: Data modification Use cases

7.2 PROVIDER MAPPING

Use case id:

- DM-1

Use case goal:

This is to allow data providers to define query and data mappings and to register these mappings.

Actors:

- Data provider

Use case summary:

Each Data provider expects requests to his services in a format suitable to his needs. A request either arrives at Harmonise system in the format of the data consumer or is already stored there as a request formulated in an intermediate standardised format (see WP4 Query Mapping). In the former case a transformation of the request is required using the Harmonise ontology and a transformation function (such as an XSLT artifact). In addition, responses are in the proprietary data format of the data provider and have to be transformed to the commonly agreed intermediate format. In order to do the transformation, the data provider has to offer appropriate transformation functions.

Basic flow of events:

(1) Data Provider specifies how a query on his data is mapped to a commonly agreed query language using the Harmonise Ontology (cf. WP3 Semantic concept and ontology and WP4 Query mapping) and the Harmonise Mapping tool (cf. WP6 Automatic mapping tool).

(2) Data Provider specifies how his data is mapped to an intermediate format using the Harmonise Ontology and the Harmonise Mapping tool.

(3) Mapping tool creates mapping.

(4) Data Provider registers the created mapping at Harmonise portal (cf. WP5 Semantic registry for meta search).

7.3 CONSUMER MAPPING

Use case id:

- DM-2

Use case goal:

This is to allow data consumers to define query and data mappings and to register these mappings.

Actors:

- Data consumer

Use case summary:

Each Data consumer expects data to be in a format suitable to his needs. The data arrives at Harmonise system in the format of the data provider and is transformed there to the intermediate standardised format. In order to perform a second transformation from intermediate format to the proprietary format of the consumer, the consumer has to provide a suitable transformation function.

Basic flow of events:

(1) Data Consumer specifies how data is mapped from an intermediate standardised format to his proprietary format using the Harmonise Ontology and the Harmonise Mapping tool.

(2) Mapping tool creates mapping.

(3) Data Consumer registers the created mapping at Harmonise portal.

7.4 DATA CLEANSING

Use case id:

- DM-3

Use case goal:

This is to allow data consumers to apply additional filter/cleanse services on raw data.

Actors:

- Data consumer
- Service provider

Use case summary:

Data collected from several data providers may contain unwanted or redundant elements for a particular data consumer (unwanted characters, duplicates…). Harmonise system may provide additional services that can filter or cleanse data.

Basic flow of events:

(1) Data consumer registers for additional filter/cleanse services that should be applied to raw data. Note that in general additional services may be applied in a sequence.

(2) Harmonise system invokes the additional filter/cleanse service with raw data provided by the data provider.

(3) Harmonise system delivers filtered/cleansed data to consumer.

7.5 DECISION SUPPORT

Use case id:

- DM-4

Use case goal:

To allow data consumers to apply additional decision support services on data such as ranking or recommender systems.

Actors:

- Data consumer
- Service provider

Use case summary:

In general, search is performed either to become aware of all relevant elements for further processing (as for instance in market research, see Use Case 3, Market Intelligence) or for decision making (see Use Case 1, Accommodation Meta Search). In the latter case, decision support such as ranking of result elements according to some relevance criteria may be provided as additional service.

Basic flow of events:

Same as before with Data Cleansing but use additional decision support services.

7.6 DATA ENRICHMENT

Use case id:

- DM-5

Use case goal:

To allow data consumers to apply additional data enrichment services on data.

Actors:

- Data consumer
- Service provider

Use case summary:

Data consumer may want that raw data is augmented with additional input. For instance, accommodation search data may be enriched with data provided from additional services such as rating agencies or user ratings, weather forecast services, or additional images. Data enrichment may change the structure of the data, for instance, by adding additional XML-elements to a result item.

Basic flow of events:

Same as before with Data Cleansing but uses additional data enrichment services.

7.7 DATA TRANSLATION

Use case id:

- DM-6

Use case goal:

To allow data consumers to apply additional translation services on data.

Actors:

- Data consumer
- Service provider

Use case summary:

While the former data modification Use cases in general do not change the contents of data elements (purging a result set from unwanted characters changes the content not substantially), translation typically involves a drastic change of content. Examples are translation of textual descriptions to a different natural language, or conversion of price information from a particular currency to another currency. Additional services may provide such functionality.

Basic flow of events:

Same as before with Data Cleansing but uses additional data enrichment services.

8 USE CASES FOR MARKET INTELLIGENCE SCENARIO

8.1 USE CASE DIAGRAM

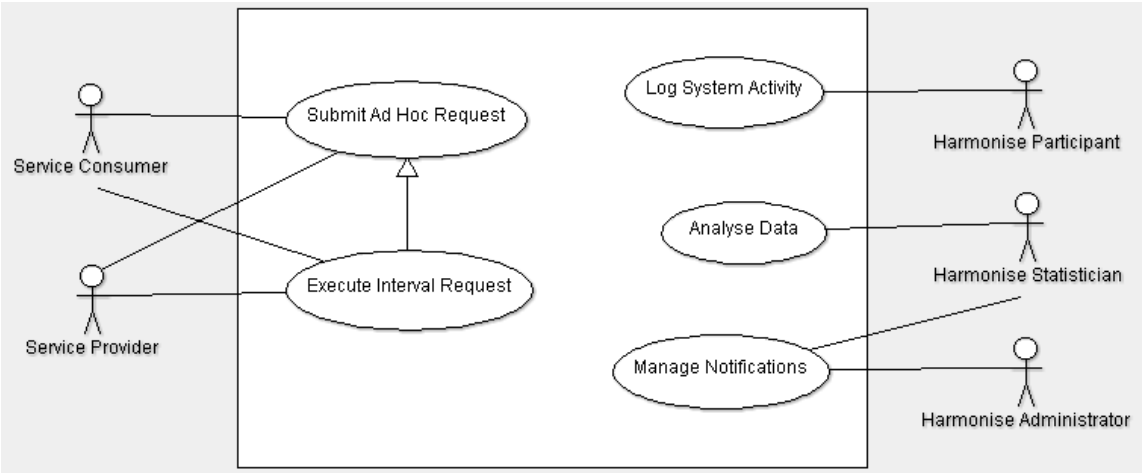

Figure 23: Market Intelligence Use Case Overview

8.2 SUBMIT AD HOC REQUEST

Use case id:

- MI-1

Use case goal:

A Service Consumer/Service Provider is able to submit an ad-hoc search query collecting filtered data about accommodations and events used for analytical processing of market information conducted by a Market Analyser.

Actors:

- Service Consumer
- Service Provider
- Market Analyser

Use case summary:

A Harmonise Participant (Service Consumer respectively Service Provider) should be able to formulate and send an ad-hoc search query collecting data about accommodations and events. The final result set serves as a basis for statistical analysis which will be conducted by external Market Analysers. In order to narrow the search focus additional filter criteria such as region or period of time may be specified by the Harmonise Participant. The Harmonise system executes the search query and returns a corresponding data set, accordingly. Thereby, the data set is constrained to the following information: price, object details, customer comments and customer ratings. In the following the data set can be forwarded by the Harmonise Participant to a Market Analyser who is capable of using the data set as a basis for statistical and analytical purposes. Compared to a regular search request the described ad-hoc request is not time-critical. It is most likely that the time-

consuming execution of the data acquisition will be executed during time periods outside the office hours (e.g. night time or weekends).

Use case diagram:

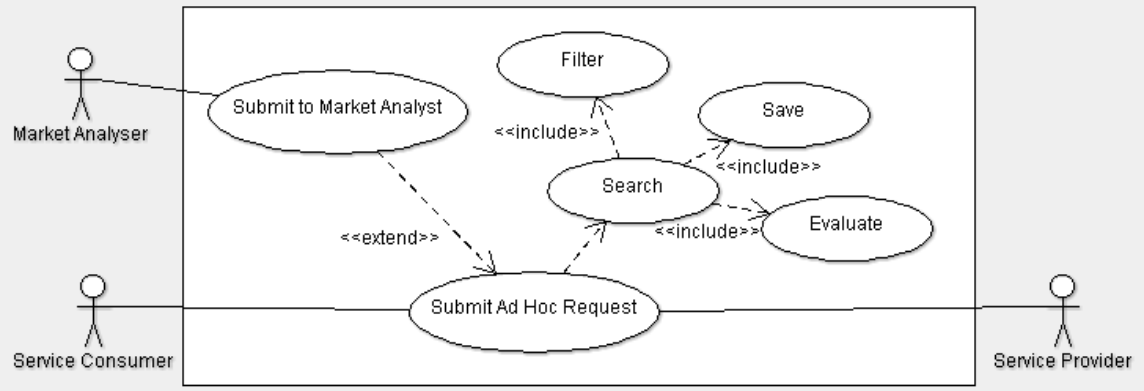

Figure 24: Ad-hoc use request use case diagram

Basic flow of events:

(1) The Service Consumer/Service Provider specifies a query and specific filter criteria – e.g. programmatically.

(2) The Service Consumer/Service Provider saves the query for further search requests at Harmonise system.

(3) The query is submitted to Harmonise system triggering an ad hoc search.

(4) Based on the query information Harmonise system searches among potential data providers.

(5) The Harmonise system filters the gathered information and prepares the result set accordingly.

(6) The Harmonise system evaluates the final result set against structural integrity (e.g. evaluation schema).

(7) The final result set will be submitted to the Service Consumer/Service Provider.

(8) In order to conduct statistical analysis among the acquired data the Service Consumer/Service Provider may forward the result set to a Market Analyst.

Alternative paths:

(1a) The Service Consumer/Service Provider loads an already saved query – e.g. from his participants profile.

(2a) The Service Consumer/Service Provider does not save the query. In this case the Service Consumer/Service Provider would directly proceed with case 3.

8.3 EXECUTE INTERVAL REQUEST

Use case id:

- MI-2

Use case goal:

A Service Consumer/Service Provider is able to create a new or modify an existing search interval submitting a specific search query to the Harmonise system.

Actors:

- Service Consumer
- Service Provider

Use case summary:

A Harmonise Participant (Service Consumer respectively Service Provider) should be able to (i) create a new or (ii) modify an existing interval. Creating a new interval allows the user to specify a certain time period wherein the system executes a specific search request. Thereby, Execute Interval Request inherits the behaviour of Submit Ad Hoc Request (see Figure 25: Interval request use case diagram). Thus, it is out of the scope here to go into more detail regarding the execution of the search request (see Figure 24: Ad-hoc use request use case diagram). Once created, the interval may be saved for the sake of reuse. Consequently, the Harmonise Participant is free to load or modify an existing interval. Regarding the former the Harmonise Participant may change the interval's configuration – e.g. period of time, search query.

Use case diagram:

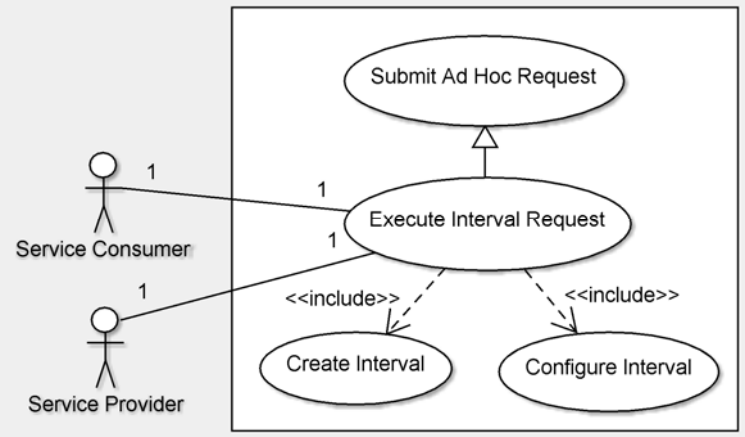

Figure 25: Interval request use case diagram

Basic flow of events:

(1) The Service Consumer/Service Provider creates a new interval request.

(2) The Service Consumer/Service Provider defines the interval – e.g. regular intervals for execution, daytime specifying when the request should be executed, etc.

(3) The Service Consumer/Service Provider specifies a query and specific filter criteria – e.g. programmatically.

(4) The Service Consumer/Service Provider saves the query for further search requests to his Harmonise Profile.

(5) The query is assigned to the interval request.

(6) The interval request is executed.

Note: The remaining steps are congruent with step 3-10 of the submit ad-hoc request use case (see Submit Ad Hoc Request).

Alternative paths:

(1a) The Harmonise Participant loads an existing interval request – e.g. from his participant profile.

i. The loaded interval is executed and sends a request to the Harmonise system. In this case the Service Consumer/Service Provider directly proceeds with step 6.

ii. The Service Consumer/Service Provider changes the configuration of the loaded interval request. In this case the Service Consumer/Service Provider directly proceeds with step 2.

(2a) The Service Consumer/Service Provider alters specific properties of the loaded interval request and executes the request afterwards. In this case the Service Consumer/Service Provider directly proceeds with step 6.

(3a) The Service Consumer/Service Provider loads an already saved query – e.g. from his participant profile. In this case the Service Consumer/Service Provider directly proceeds with step 4.

8.4 ANALYSE DATA

Use case id:

- MI-3

Use case goal:

The Harmonise system enables a Harmonise Statistician to gather and analyse log data statistically.

Actors:

- Harmonise Statistician
- Harmonise Administrator

Use case summary:

A Harmonise Statistician shall be able to perform statistical analysis among log data documenting how Harmonise Participants use the system. The statistical analyses are managed and configured via a status cockpit which is part of the Harmonise system. The status cockpit helps the Harmonise Statistician to interact with the system supporting the following major functionalities: (i) Enable the execution and visualization of statistical analysis and (ii) manage notifications. The former allows the Harmonise Statistician to conduct specific methods on the log data, which serves as a basis for retrieving ranked statistics on system usage as well as user group specific usage. The latter will be addressed in use case Manage Notifications and will not be further detailed. Regarding the former, the Harmonise Statistician may narrow the scope of the consumed data by defining certain search criteria (e.g. product, category, geography). In the following the status cockpit shall be able to render the final statistics in a human readable format, which could be either graphically (e.g. pie chart or bar chart) or in tabular form.

Use case diagram:

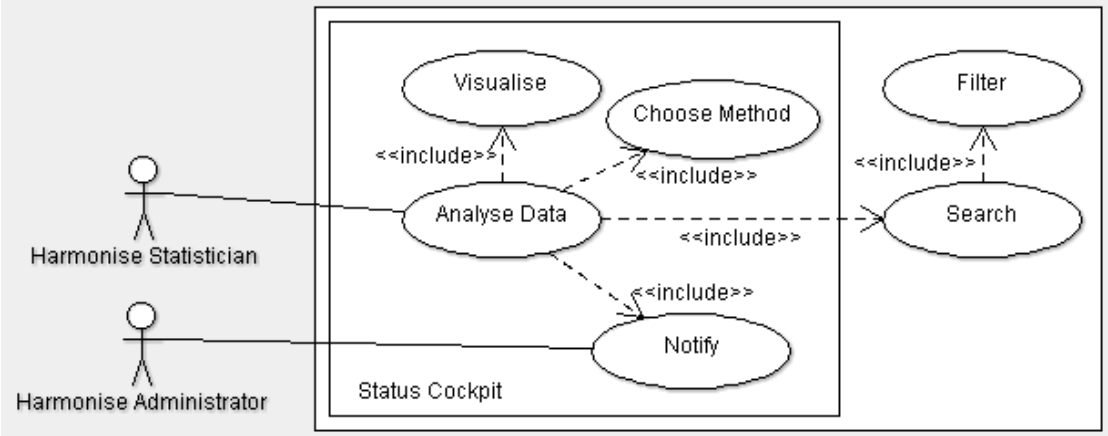

Figure 26: Analyse data use case diagram

Basic flow of events:

(1) Enter the status cockpit.

(2) Choose the statistical method of choice (e.g. global, user group specific).

(3) Create a search query with certain filter criteria.

(4) Submit the query to the Harmonise system.

(5) The Harmonise system queries the log data and returns the results accordingly.

(6) In case the results meet a certain notification criteria the status cockpit sends a notification to authorised Harmonise staff or Harmonise Administrator.

(7) After receiving the results the Harmonise Statistician chooses an appropriate visualization (e.g. graphically or tabular form).

Alternative paths:

(6a) No failure occurred. In this case step 6 can be ignored.

8.5 MANAGE NOTIFICATIONS

Use case id:

- MI-4

Use case goal:

The Harmonise Statistician/Harmonise Administrator should be able to monitor certain events and activities within the Harmonise system.

Actors:

- Harmonise Statistician
- Harmonise Administrator

Use case summary:

A Harmonise Statistician or Harmonise Administrator shall be able to perform statistical analysis among log data documenting how customers use the system. The statistical analyses are managed and configured via a status cockpit which is part of the Harmonise system. The status cockpit helps the Harmonise Statistician to

interact with the system. Its purpose is twofold: (i) to enable the execution of statistical analysis and (ii) to manage notifications. The former has already been described in use case Analyse Data. The latter triggers when the system should inform the Harmonise Statistician or the Harmonise Administrator about certain circumstances. The notification mechanism would allow a seamless monitoring specific to the needs of Harmonise. For example it is crucial to react as soon as possible in case of reoccurring transaction failures or services.

Use case diagram:

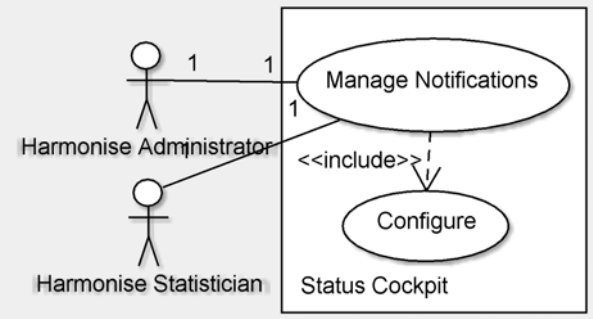

Figure 27: Manage notifications use case diagram

Basic flow of events:

(1) Enter the status cockpit.

(2) Set the notification options accordingly (e.g. notify on transaction failure).

(3) Save changes.

Alternative paths:

(2a) The Harmonise Statistician may change the type of notification (e.g. Twitter, Email, and SMS).

(2b) Define member authorised to receive certain notifications.

8.6 LOG SYSTEM ACTIVITY

Use case id:

- MI-5

Use case goal:

The Harmonise system shall be able to log certain activities conducted by Harmonise Participants. The collected data serves thereby as a basis for system-specific analyses.

Actors:

- Harmonise Participant

Use case summary:

The Harmonise system shall be able to recognise specific activities which are executed by a Harmonise Participant. A specific activity in the context of this use case is seen from a general point of view addressing all activities which may be of interest for statistical as well as system analysis – e.g. recognition of transactions between Data Providers and Data Consumers. However, the logging mechanism

should not be restricted to user triggered activities. In addition it is also considered to log the system's behaviour regarding critical events such exceptions or environmental related data (e.g. number of online Harmonise Participants, etc.). In addition, the log-data will also be used as input for other use cases (e.g. Analyse Data or Manage Notifications). Consequently, the logging mechanism has to be very flexible regarding the recognition of system activities as well as the extraction of meta-information for a particular activity. Thus, the Harmonise system has to be able (i) to evaluate the relevance of an occurring activity, (ii) to extract the activity's meta-data and (iii) to write the log-data to a log data source (e.g. database or log-file).

Use case diagram:

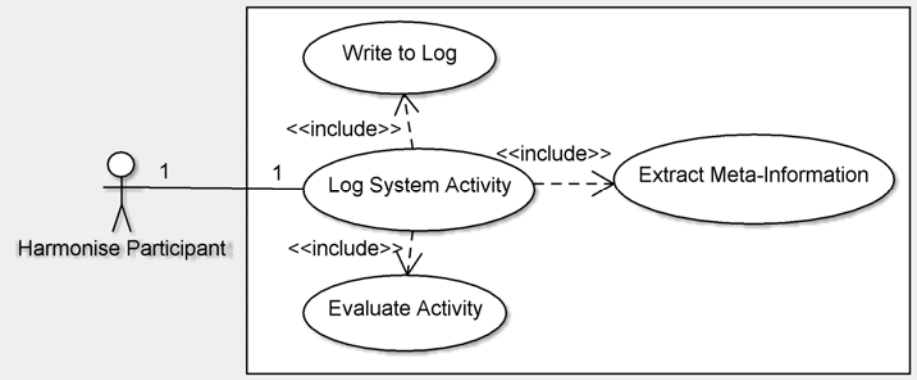

Figure 28: Log system activity use case diagram

Basic flow of events:

(1) A Harmonise Participant conducts an activity which is sensitive for logging.

(2) The Harmonise system recognises the activity.

(3) The Harmonise system evaluates if the activity should be logged or not.

(4) The Harmonise system decides (depending on the type of activity) which data should be logged.

(5) The Harmonise system extracts potential logging-data.

(6) The Harmonise system writes the meta-data to the log data source (e.g. database or log-file).

Alternative paths:

(1a) The Harmonise Participant conducts an activity which should not be logged.

(1b) The Harmonise system triggers a specific activity which shall be logged (e.g. critical system error occurred).

9 CLASSIFICATION OF THE USE CASES

This section classifies the use cases described in the previous section considering the functionality areas they belong to.

The major components foreseen in the Description of Work are the following:

Metasearch: This component will enable searches across different individual search components of heterogeneous websites and aggregate the results in a unified list. Partners will use this functionality to search heterogeneous data sources using a uniform interface and a consistent query language.

Semantic Registry: This component is a kind of intelligent index for identifying data sources based on a semantic understanding of the search intention and knowledge about existing partners in the network and of the data they can provide. This shall allow getting fast and reliable results even when thousands of partners can be searched.

Mapping Tool: The mapping tool is a standalone application. It supports a user in visually creating the necessary mapping definitions from the data model of a Harmonise participant to the one of Harmonise and vice-versa, with little technical knowledge. It consists in a graphical User Interface to show and manipulate mappings, a pluggable set of algorithms to support automatic mappings, a generator to create mapping artifacts, and an infrastructure in order to manage a mapping project.

The table below summarises the component(s) which the use cases belong to: query submission and execution, service representation within the registry and aided support to the mapping tool.

In addition, the impact on existing components of the Harmonise platform were considered.

Finally, each use case is classified according to the relevance with the HarmoSearch project. Three levels were identified:

- **Core:** Core services are built-in functions of the Harmonise platform and are consistent with the HarmoSearch DoW. Use cases classified as core are part of the proposal and have to be fully supported and implemented within this project.
- **External Service:** External services are additional functionalities which can be plugged in the Harmonise platform. They give additional value for a particular user (a Harmonise Participant) and an external service may be suitable only for this particular user while another user would prefer to have a similar but slightly different service. Harmonise platform offers to the participants the possibility to register and subscribe to external services. Use cases classified as external services are the ones that will be implemented via integration of external services. Thus HarmoSearch should enable the possibility to integrate these services but is not intended to implement the necessary technology to fully support the described functionality.
- **Out of scope:** A couple of use cases have been classified out of scope of the HarmoSearch project, since they are not part of the agreed functionalities to

be implemented. Anyhow, the architecture of the system should be designed in such a way that the integration of these functionalities will be possible in the future.

Use Case id	Use Case title	Meta-search	Semantic Registry	Map-ping Tool	Other Components	Relevance to HarmoSearch: CORE, EXTERNAL SERVICE, OUT
MS-1	Harmonise Registration and Setup	X	X	X	X	CORE
MS-2	Bookable Items Search	X	X			CORE
MS-3	Items Recommendation	X	X			EXTERNAL SERVICE
IMPORT-1	Harmonise Registration and Setup		X	X	X	CORE
IMPORT-2	Batch transfer of static data	X				CORE
PS-1	Data Hosting	X	X			EXTERNAL SERVICE
PS-2	Data Publishing	X	X			CORE
PS-3	Data Enrichment	X				EXTERNAL SERVICE
PS-4	Consumer Association	X	X			CORE
PS-5	Data Download	X				EXTERNAL SERVICE
PS-6	Data Subscription	X	X			CORE
PS-7	Alert	X				CORE

Use Case id	Use Case title	Meta-search	Semantic Registry	Map-ping Tool	Other Components	Relevance to HarmoSearch: CORE, EXTERNAL SERVICE, OUT
	Definition					
PS-8	Portal Service Subscription		X			CORE
PS-9	Negotiation				X	OUT
PS-10	Payment				X	OUT
PS-11	Profile Management	X	X	X	X	CORE
DM-1	Provider Mapping			X		CORE
DM-2	Consumer Mapping			X		CORE
DM-3	Data Cleansing	X		X		EXTERNAL SERVICE
DM-4	Decision support	X				EXTERNAL SERVICE
DM-5	Data Enrichment	X		X		EXTERNAL SERVICE
DM-6	Data Translation	X		X		EXTERNAL SERVICE
Mi-1	Submit Ad Hoc Request	X				EXTERNAL SERVICE
MI-2	Execute Interval Request	X				EXTERNAL SERVICE
MI-3	Analyse Data				X	EXTERNAL SERVICE
MI-4	Manage Notifications				X	EXTERNAL SERVICE

Use Case id	Use Case title	Meta-search	Semantic Registry	Map-ping Tool	Other Compon ents	Relevance to HarmoSearch: CORE, EXTERNAL SERVICE, OUT
MI-5	Log System Activity	X				CORE

10 LIST OF FIGURES