2000

An Approach to Enable Interoperability in Electronic Tourism Markets

Marion Kaukal
Vienna University of Economics and Business Administration, kaukal@wi.wu-wien.ac.at

Hannes Werthner
Vienna University of Economics and Business Administration, werthner@wi.wu-wien.ac.at

Wolfram Hoepken
START Amadeus GmbH, wolfram.hoepken@start.de

An Approach to Enable Interoperability in Electronic Tourism Markets

Marion Kaukal*, Wolfram Höpken**, Hannes Werthner*
* Department of Information Systems, Vienna University of Economics and Business Administration
{kaukal, werthner}@wi.wu-wien.ac.at
** START AMADEUS GmbH, Frankfurt am Main Wolfram.Hoepken@start.de

Abstract **The exchange of semantically consistent service descriptions is an important issue for flexible integration facilities for electronic commerce. Currently there is a lack of semantic consistency on the Web, burdening arbitrary market relationships. Several standardization initiatives have addressed this issue before, but nonetheless, the setup and maintenance costs have been too high. Furthermore, too rigorous standardization is not appropriate. As tourism markets are particularly heterogeneous, there is a high demand for flexible, but consistent data schemes for distributed service descriptions on the Web. A mediated tourism market could solve the integration problem by providing global data schemes that are individually extendable. In this paper we propose an approach based on the Resource Description Framework (RDF) and eXtensible Markup Language (XML) Namespaces, which are promising technologies that could be used for addressing the interoperability issues, which remain, however, hard problems.**

I. Introduction

In order to cope with the enormous amounts of information on the web, automated processing capabilities are a necessity. This represents a serious challenge since online information is mostly unstructured and ambiguous.

In this paper we concentrate on automated integration facilities for suppliers of tourism services. The focus is on a solution for the interoperability problem between suppliers that are using different data schemes for describing the same domain. Additionally, issues like individual extensions and flexible relationships between "chunks" of information will be covered by an integration concept for an electronic tourism market.

We describe the concept of a mediator which is operating for the domain of hotel accommodation and which is responsible for integrating the data schemes of different accommodation suppliers. Mediators can take the required responsibility for integrating data resources by means of rigorously defined terms and relationships in a specific domain of interest. According to Wiederhold and Genesereth [1] those *ontologies* should be defined domain-dependent, proposing a set of interacting mediators, each focusing on its own domain.

In section II we describe features of the tourism industry, including the respective value chain transformation initiated by the employment of new media. New potentials for flexible integration of tourism services are outlined. Section III focuses on the notion ontology which refers to a specific domain terminology, its formalization is indispensable for automated communication on the WWW. Thereafter emerging technologies for ontology construction and integration on the WWW are presented and finally used to provide an example how hotel information in an electronic tourism market could be modeled.

II. Electronic TOURISM Markets AND Integration ISSUES

Figure 1 shows a simplified model of the typical tourism value chain. The first stage are the *suppliers* of tourism services and components, like flights, accommodation, catering, or entertainment. At an intermediate level, *Tour Operators* are responsible for aggregating a number of components and to offer these as packaged services, i.e. they are performing integration tasks. The *travel agencies* mainly act as retailer and are the intermediaries between the tour operators and the potential customer [2]. Many tourism services are composed of various components that originate from different suppliers. For example, tennis camps integrate services of hotels and tennis courts. This package is aggregated by a tour operator, the travel agency collects the different packages and is responsible for informing the potential customers and for the booking services.

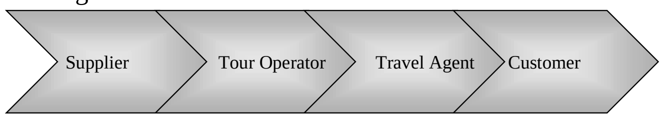

Fig. 1: The (traditional) Tourism Value Chain

Communication between the travel agencies and the tour operators as well as between tour operators and huge service suppliers such as airlines or internationally acting hotel chains, is mainly conducted by CRS/GDS (computerized reservation systems / global distribution systems) over proprietary networks. Traditional suppliers, within destination, mostly small and medium sized enterprises (SMEs), are excluded from this electronic communication. CRS/GDS are based on a rather old technology, developed in the 1960s, which is hard to maintain and to upgrade and not flexible enough to allow ad-hoc integration of new participants and services. Furthermore, it is not accessible by the end customer. For informing the individual customer the travel agencies use their own and informal vocabulary. The customers are mainly informed by telephone and face-to-face meetings.

Web based tourism markets are no 1:1 counterparts of the traditional ones. Both, structure and players change. Many suppliers like hotels and restaurants offer online information directly to the customers (although mainly in an unstructured manner) by the use of WWW interfaces, that are not forced to exclusively rely on person-to-person interaction any more. The importance of traditional travel agencies is probably declining in the online market and new types of intermediaries like Expedia, Travelocity and Tiscover are entering the tourism value chain. Additionally, the configuration of the value chain is becoming more ad-hoc and dynamic, allowing the different parties of the chain to occupy more flexible roles. Tourism suppliers can sell their services to the customers either directly or indirectly by the use of an intermediary service. For example, a hotel can offer its rooms and services over an individual web site or it can become part of an open electronic tourism market by supplying its products to some intermediary.

Furthermore, web markets will be more open than traditional tourism markets, i.e., new participants may enter the

search

market more easily and also new services can be easily included in the market offers. Traditional EDI arrangements required complex configuration efforts for new market participants. Ideally, the set up for newcomers of web based markets should be carried out automatically without the need for any human intervention.

A. *The Heterogeneity of Tourism Service Descriptions*

Main challenges for web based tourism markets are the search for and the comparison of tourism information as well as the composition of service packages, i.e., the integration tasks. This is a hard task due to the heterogeneous tourism suppliers as they make use of varying ways to present their offers. Both aspects, the "syntactic", as well as the "semantics", make the information-gathering and the comparison services time-consuming and cumbersome. A classification of possible heterogeneities is given in the following overview:

- One term may ambiguously describe different services
  - An *overnight* may include breakfast or not, depending on the country
- One concept may be described by different terms
  - departure time, pick-up time
- The attribute values may be different
  - a hotel category may be described by stars, crowns, numbers etc.
- Structuring levels may differ
  - price and currency may be separated attributes vs. price and currency are provided together
- The annotated information for the "same" basic product may differ
  - a room price may be provided per person and night, an apartment price may be provided per apartment and night
- The entity itself can be complex
  - is a 2-bed-room for one, two or three persons?
- The representation format may vary
  - HTML, XML, dynamic database query, …
- The handling of the user interface may differ depending on the interaction facility

B. *Interoperability*

An electronic market is an open and distributed system, based on the interoperability of the participating systems [3]. Interoperability on the WWW is based on the Hypertext Transfer Protocol (HTTP) that provides the standard protocol definition for the exchange of hypertext documents. In order to allow content-based integration among the participants, the integration services should be expanded to the content of the messages passed between HTTP clients and servers.

Accordingly, what is needed first for an integration approach is a common understanding about the domain of interest, i.e., the terms and relationships that are used to unambiguously describe a tourism service, which is possibly composed of several basic components. The assigning of semantics, i.e. the meaning, to the single terms of a language is a critical issue for integration facilities. In the easiest way all participants of a marketplace are using unified syntax and semantic.

Therefore, data models referring to the same domain are subject to standardization. For the tourism domain several standardization initiatives try to model a tourism market and provide a standard for describing information and service interaction, for example, by analyzing huge amount of tourism data material or conducting expert interviews. A service description for an electronic market characterizes an offered service by describing all relevant information for accessing the service. Service descriptions typically contain the structure of the service, its semantics, its relationships to other services, as well as processing information. Service descriptions cover descriptions of tourism products like *hotel* or *flight* as well as tourism entities like *booking*, *search request* or *customer data*. Within the field of tourism, a lot of different standardization initiatives have evolved in the past. Examples are the OpenTravel Alliance (OTA), United Nations rules for Electronic Data Interchange for Administration, Commerce and Transport – Travel Tourism & Leisure (UN/EDIFACT TT&L), XML/EDI, Tourism Information Norm for the German Tourism (TIN), Hospitality Industry Technology Integration Standards (HITIS), omnis-online, International Air Transport Association (IATA), Travel Technology Initiative (TTI), and Association of Car Rental Industry System Standards (ACRISS). At the International Federation for Information Technology and Tourism (IFITT) a special interest group is currently working on a reference model for the unification of different existing electronic tourism markets. See also section III.C for a more detailed description of XML/EDI and OTA standards.

The objective of standardized service descriptions for electronic markets is the interoperability of different systems and, consequently, the seamless exchange of information. Such a unified service description would be appropriate, but is not easy achievable, since the individual parties involved may have some objections against a "total" standard due to competitive reasons. In addition, they are internally deploying many heterogeneous data models especially constructed for their own needs. Thus, a broad standardization of service descriptions, enabling a global interoperability between tourism information systems has not been successful up to now.

III. Ontologies

Ontologies are a mechanism for "knowledge sharing", the sharing of information over a specific domain of interest. They provide a common unambiguous description of a vocabulary used for describing the objects and relationships therein. For communication purposes ontologies enable a shared understanding between people with different needs and viewpoints arising from their particular contexts. Thereby it serves as interlingua for the translation between different languages and representations and supports the inter-linking of multiple distributed systems.

There are many different definitions for an Ontology (see [4] for an overview). We primarily agree with Uschold and Gruninger who define an ontology as the term used to refer to the shared understanding of some domain of interest which may be used as a unifying framework that serves as base for communication, interoperability and systems engineering [5]. For our concerns communication and interoperability are the deployment areas of ontologies.

search

The degree of formality by which the vocabulary of an ontology is built and the meaning of terms are specified, varies from informal definitions expressed in natural language over a taxonomy of terms that are arranged into a generalization-specialization hierarchy to definitions stated in a formal language based on first-order logic with a rigorously defined syntax and semantics. The use of ontologies ranges from informal requirements such as a glossary for shared understanding among human users to more formal requirements such as interoperability among software tools [5].

In electronic markets there is a strong need for highly formal ontologies. Semantics, i.e. the specification of the meaning of definitions, is an important issue for information provision on the WWW. There are different purposes for the need of semantics: Semantics can be defined formally in order to allow machine-understandable information. On the other hand, semantics is also important for unambiguous human web processing. But for this objective also natural language descriptions of the language parts are sufficient, which, however have to be retranslated according to a specific cultural context. Note, that in personal communication ambiguities can be resolved much simpler than in any kind of automated interaction.

Semantics determines the facts in the world to which the sentences of a language refer [6]. Semantic information serves the descriptive information behind the data itself. It must be defined uniquely within the ontology which means that two different descriptions should not be applicable for one semantic purpose and one term should not define varying semantics. This can be realized by a semantic data model or data scheme which structures the information, documents the meaning, usage and relationships and imposes integrity constraints [7]. For defining semantics a metalanguage is used that assigns the meanings to the terms and relationships of a language. The formal specification of semantics is subject to ontological engineering.

Over the last two decades, the knowledge representation and object-oriented database communities have developed a number of languages that may be used for the expression of formal ontologies (see [8]). The most important language is the Knowledge Interchange Format (KIF) [16]. KIF is a prefix version of first order predicate calculus with extensions to support non-monotonic reasoning and definitions. The language description includes both a specification for its syntax and one for its semantics. It is currently subject to ANSI standardization. Ontolingua [17] is a popular project at the Stanford University and provides an ontology development environment and a representation language that is based on KIF extended with object-oriented and advanced syntax constructs. Other examples of traditional knowledge representation languages are CYCL [20] and LOOM [21]. The disadvantage of these languages is that they are not very popular outside the AI community.

For ontology construction suitable for WWW based electronic commerce new languages and techniques have emerged which will be presented in section III.B.

A. Mediated Architectures

In a market an obvious conflict exists: on one hand, descriptions of services should not differ, and at the same time, the suppliers should not be forced to agree on one standard service configuration. Thus, in order to relate heterogeneous information on the web there must be a way of exposing information from different systems. These systems may use a variety of internal data models which implies the requirement for some generic concept of data at a low level that is in common between each system [9].

Most promising implementation solutions for markets with heterogeneous participants are based on the concept of mediated architectures that allow interoperation among semantically distinct ontologies. Mediators allocate information from diverse sources and are responsible for providing integrated information that is automatically composed without the need to physically integrate the base data sources. They are responsible for accessing and retrieving relevant data from multiple heterogeneous sources, abstracting and transforming retrieved data into common representation and semantics, integrating the homogenized data according to matching keys. In addition integrated data can be abstracted to increase the information density of the final result [1].

Wiederhold and Genesereth [1] are proposing such a mediated architecture which extends the traditional client/server-architecture with an intermediate layer, which is showed in figure 2. The *information layer* adapts heterogeneous data sources and transforms them in queryable object-relational tables. The *mediation layer* is responsible for the integration of the various data sources of the information layer and delivers them in the appropriate format to the *customer application layer* which contains the user programs and additional tools for browsing, navigating and querying mediated data sources.

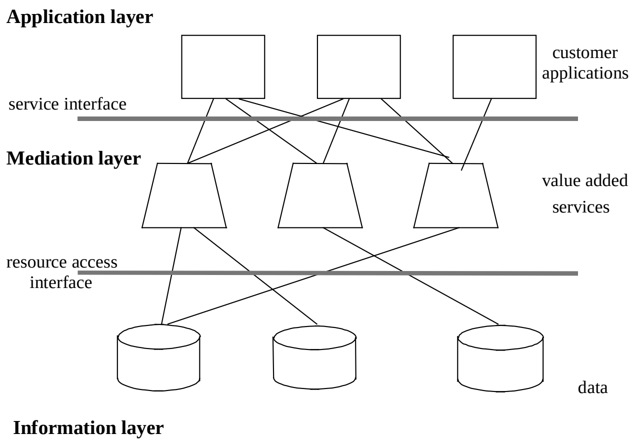

Fig. 2: Three layered mediation architecture ([2] following [1])

Among others, the mediation services include the mapping of domain terminologies, the integration of material from diverse sources, the assessment of the quality of the obtained information and the aggregation of information into higher levels to obtain enhanced information.

Mediation represents a general architectural concept, its focus lies in the tackling of the problem of semantic interoperation. This task is highly domain dependent, no single general mediator will be able to cover all topics of any application field. Such a domain should be limited by its ability to define and maintain internal consistency [1], i.e., identical semantics for the terms used. Even in the domain of tourism we can expect a set of different mediators, since many applications will need to cover different topics [2].

B. *Ontology Construction for the WWW*

The emergence of the eXtensible Markup Language (XML) [13] that enables content tagging, provides the basic possibility for automated and content based processing of information on the web. It is a description language for highly structured information. XML combines the simplicity of Markup Languages with the computability of EDI standards. A XML Document Type Definition (DTD) allows very elementary schema definitions. It defines the sets of elements and their attributes as well as element relationships, the elements are named by element tags. Therewith any (business) markup language can be specified. Each XML document that is defined by the means of a DTD can be validated.

For the exchange of strongly typed business data the traditional DTD-model has to be enhanced, as XML DTDs do not provide strong typing. In addition, XML DTDs do not support inheritance and do not allow to specify sequences and alternatives. Solving this problem initiatives emerged that suggest more advanced schema definition requirements. Among others, Commerce One's Schema for Object-oriented XML (SOX) [24] and the Document Definition Markup Language (DDML) [25] are currently considered by W3C's XML Schema Working Group [26].

The incremental and decentralized development of Web applications also requires documents containing an ad-hoc mixture of features from multiple sources. The XML Namespace facility [21] allows the respective vocabulary mix-in by associating element and attribute names used in XML documents with namespaces for the belonging data schemes identified by a Universal Resource Identifier (URI) [15] reference.

However, XML only includes implicit semantics by the use of "meaningful" XML tags. It is mainly usable for structured business documents like orders and invoices, as the DTDs or advanced schema definitions allow mainly the syntax specifications. Such documents are, for example, the traditional EDI messages with a clear syntax explicitly defined and validable. Thus, XML mainly provides syntactic interoperability.

For approaching semantics on the web additional mechanisms emerged, that use metadata. Metadata refers to information that describes other information sources. Historically, it was mainly deployed for libraries and similar domains. For example, records in online library catalogs consist of metadata about books, including information such as publisher, publication dates, and author names. Metadata is suitable for providing information to manage and retrieve resources requiring more structure than is present in the data itself.

Metadata for WWW resources is considered by the W3C Metadata Activity [22] which developed the Resource Description Framework (RDF) [18]. It is based on directed labeled graphs where the nodes are referring to a resource, each resource described by a set of properties. This is called an RDF Description. Each of these properties has a property type and a property value. Figure 3 shows a part of an example hotel RDF description. The hotel web site (the resource) is represented by a node in a graph. The arcs are delineating properties (category, owner), property values can be either other resources or literals that are shown as rectangles ("A", "Thomas Berger"). Each resource that is identifiable by an URI can be described by RDF.

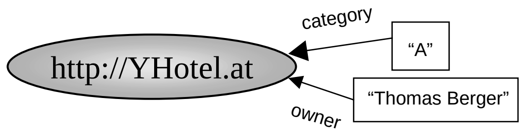

Fig. 3: RDF Node and Arc Diagram

RDF provides a framework in which independent communities can develop metadata vocabularies that fit their specific needs, and they can share these vocabularies with other communities. Thus, it is suitable for ontology purposes. Complementary, in order to share vocabularies, the meaning of the terms must be declared in detail. Therefore *RDF Schemes* [19] define the meanings, characteristics, and relationships of a set of properties. The RDF Schema establishes vocabularies using object-oriented constructs and allows validation of RDF descriptions. A specific schema is a collection of classes. The classes are organized in a hierarchy, and offer extensibility through subclass refinement. The RDF schema specification language is less expressive than full predicate calculus languages like KIF, but simpler to implement [19].

In comparison to XML that assures interoperability between applications that exchange data, RDF enables interoperability between applications that exchange metadata. Of course the distinction between data and metadata is not always strict. By the use of RDF, metadata for any web resource (no matter if image, sound file or video), can be expressed. Furthermore, the content of the resource itself does not need to be strongly typed and structured.

Thus, RDF provides mechanisms to define information about information that is contained in web documents. Moreover, the relationships between the information can be described implicitly. It provides the possibility to express information about a WWW resource or about the relation between two or more WWW resources apart from the objects themselves. Information modeled by the means of RDF can be serialized and exchanged via XML using a specific XML grammar. RDF itself allows each metadata document to define which vocabulary is used by assigning each schema a namespace, accessible by a URI.

In the following, two XML-based standard descriptions of tourism markets are described. Afterwards we show an example of a RDF data model for a tourism market which allows the integration of the different product descriptions by the use of a standard market place schema. This can also be individually extended by each participating party. Thus, it allows both, integration as well as extensibility.

C. *XML-based Tourism Standards*

In this section we present two example protocol definitions: OpenTravel and XML/EDI. OpenTravel concentrates on service description like accommodation, whereas the XML/EDI objective is more on process descriptions like booking. They are both attempts to define more flexible specifications of tourism services, instead of defining fixed standards.

a. *OpenTravel*

The *OpenTravel* standard is an initiative of the OpenTravel Alliance (OTA) comprising of members of different travel industry sectors like hotel, air, car and travel agencies [10]. Its objective is to develop an industry-wide information standard enabling the information exchange between all industry sectors. OTA is expected to publish its first document in December 1999.

The basic approach of the OpenTravel standard is to define a vocabulary of travel-related terms and a syntax for composing those terms to service descriptions. Thus, the standardization does not take place on the level of complete service descriptions but on the level of basic data elements. Instead of defining a completely fixed structure, service descriptions can then be composed of this basic data elements in a flexible way.

In order to enable a flexible and efficient communication between different tourism sectors, specific vocabularies for different sectors can be used. Thus, instead of a central, fixed vocabulary, which has to be used by all participants, different specific vocabularies can be defined and integrated into one global vocabulary. The mapping between different vocabularies is done within this global vocabulary.

Beyond this sector-specific vocabularies communication partners can define private vocabularies and use them bilaterally. This can be done by extending the common vocabulary by new, private terms, used exclusively for the communication between those communication partners.

```
<!ELEMENT HOTEL (NAME, ADDRESS, CATEGORY, ROOM+)>
  <!ELEMENT NAME(#PCDATA)>
  <!ELEMENT ADDRESS(STREET, CITY, STATE, ZIP)>
    <!ELEMENT STREET(#PCDATA)>
    <!ELEMENT CITY(#PCDATA)>
    <!ELEMENT STATE(#PCDATA)>
    <!ELEMENT ZIP(#PCDATA)>
  <!ELEMENT CATEGORY(#PCDATA)>
  <!ELEMENT ROOM(TYPE, PRICE)>
    <!ELEMENT TYPE(#PCDATA)>
    <!ELEMENT PRICE(AMOUNT, CURRENCY)>
  <!ELEMENT AMOUNT(#PCDATA)>
  <!ELEMENT CURRENCY(#PCDATA)>
```

```
<?xml version="1.0"?>
<HOTEL>
   <NAME>Cityhotel</NAME>
   <ADDRESS>
      <STREET>54 Chambers Street</STREET>
      <CITY>New York</CITY>
      <STATE>New York</STATE>
      <ZIP>10007</ZIP>
   </ADDRESS>
   <CATEGORY>4</CATEGORY>
   <ROOM>
      <TYPE>single</TYPE>
      <PRICE>
         <AMOUNT>140</AMOUNT>
         <CURRENCY>USD</CURRENCY>
      </PRICE>
   </ROOM>
   <ROOM>
      <TYPE>double</TYPE>
      <PRICE>
         <AMOUNT>180</AMOUNT>
         <CURRENCY>USD</CURRENCY>
      </PRICE>
   </ROOM>
</HOTEL>
```

Fig. 4: XML DTD and an example XML Document for a hotel description

The OpenTravel standard uses XML for describing the defined vocabulary, i.e. the common vocabulary is represented by a set of XML tags, setting up an XML namespace. The communication between participants takes place via XML documents. The example in figure 4 illustrates, how an XML element describing a *hotel* would look like. On the right-hand side an example XML Document can be seen, the belonging DTD is presented in the left part of the figure. The DTD states that, for example, each `hotel` element consists of one `name`, one `address` and one `category` element and one or more (+) `room` elements. The `name` element for example is an elementary data and defined by the SGML construct `#PCDATA` which refers to document text.

b. XML/EDI

EDI (Electronic Data Interchange) aims at facilitating the electronic exchange of business data between communication partners by defining syntax rules for messages to be interchanged. The communication partners have to agree upon a common set of messages. Messages can be structured as groups of segments, which consist of data elements.

EDIFACT TT&L (Electronic Data Interchange For Administration, Commerce and Transport – Travel Tourism & Leisure) focuses on electronic data interchange within the tourism domain by defining a set of standard messages needed for trading tourism services (e.g. messages for information, availability and reservation requests and responses) [11]. The standard EDI data elements are collected in central dictionaries enabling the look up of their meaning and definition.

However, a crucial drawback of EDI standards is the necessity to commit to fixed message formats between the communication partners.

Therefore, the basic approach behind the “extension” XML/EDI is to express EDI mechanisms using XML syntax and in this way enhance EDI mechanisms by the flexibility and extensibility of XML [12]. The segments of EDI messages are replaced by XML elements and analogous EDI messages by XML documents. Since the use of XML alone would not solve all past EDI problems, three additional components are designed for full dynamic electronic commerce: process templates, software agents and repositories. *Process templates* contain processing information and therefore define what happens to the data. *Software agents* interpret the process templates and interact with EDI data definitions and user applications. *Repositories* provide the syntactic and semantic information needed for the execution of EDI transactions. Repositories contain standard XML DTDs and standard EDI segments and elements. The communication partners can automatically retrieve the content of a repository. Based on this information a mapping between a specific model and the global repository can be achieved.

IV. A RDF ONTOLOGY FOR A TOURISM MARKET

The initiatives described in section III.C lack the ease of configuration as well as flexibility and the possibility to define advanced semantics. Furthermore, individualization is not covered properly. In the following we propose an approach based on RDF focused to overcome these limitations.

The main part is a mediator responsible for keeping a global core ontology and for supporting the mapping of the participants’ individual data schemes as well as for the integration of additional subdomain ontologies in order to dynamically compose packaged products. Additionally, it provides a decentralized extension mechanism for the individualization of the ontology description.

We are assuming that the hotel suppliers have already agreed to some formal or informal language for presenting their information internally or externally. In order to support the data supplier already applying a common standard data schema like TIN [23] before entering the market, basic mapping facilities can be offered by the mediator.

```
<rdf:RDF
xmlns:rdf=”http://www.w3.org/1999/02/22-rdf-syntax-ns#”
xmlns:rdfs=”http://www.w3.org/TR/1999/PR-rdf-schema-19990303#”>
...
<rdf:Description ID=“HotelRoom”>
  <rdf:type resource=“ http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Class” />
  <rdfs:comment>This class describes a standard hotel room</rdfs:comment>
  <rdfs:subClassOf rdf:resource=”http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Resource”>
</ rdf:Description >

<rdf:Description ID=“bedNumber”>
  <rdf:type resource=“ http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Property” />
 <rdfs:label xml:lang=“en”>bedNumber</rdfs:label>
 <rdfs:label xml:lang=“de”>bettenAnzahl</rdfs:label>
  <rdfs:comment>The number of beds in a hotel room</rdfs:comment>
  <rdfs:range rdf:resource=“http://www.datatypes.org/useful_types#Integer_One_to_50” />
  <rdfs:domain rdf:resource=“#HotelRoom”/>
</rdf:Description>

<rdf:Description ID=“roomView”>
  <rdf:type resource=“ http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Property” />
  <rdfs:comment>The view out of a hotel room</rdfs:comment>
  <rdfs:range rdf:resource=“#RoomView” />
  <rdfs:domain rdf:resource=“#Hotelroom”/>
</ rdf:Description >

<rdf:Description ID=“RoomView”>
  <rdf:type resource=“ http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Class” />
  <RoomView rdf:ID=“Mountainview”/>
  <RoomView rdf:ID=“Streetside”/>
  <RoomView rdf:ID=“Seaside”/>
  <RoomView rdf:ID=“Other”/>
</ rdf:Description >
.....
```

Global RDF Schema

```
...
<rdf:Description ID=“wwwAccess”>
  <rdf:type resource=“ http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Property” />
  <rdfs:comment>The availability of access to the WWW in a hotel room  </rdfs:comment>
  <rdfs:domain rdf:resource=”http://www.w3.org/examples/hotelschema#HotelRoom”>
  <rdfs:range rdf:resource=” http://www.datatypes.org/useful_types#Boolean”>
</ rdf:Description >
...
```

Individual RDF Schema extension

Fig. 5: Hotel Market Schemes

A metalanguage is used, that provides the possibility to export the semantics of the information to the mediator. This metalanguage is referred to as hotel ontology, which describes the terms and relationships used to express metainformation about the hotel domain. The metalanguage should provide the possibility to define the configuration of a hotel service product, for example, the description of hotel rooms.

The approach is based on the use of RDF schemes and the mapping of every individual core data model to one centrally provided RDF schema. The local instances are not forced to use this global schema but to define once how their own data can be mapped with the global model, i.e., the URI where the resources can be found. Thereby they can be supported by a mapping tool provided by the mediator.

Thus, data and data schemes are distributed over the network. The intermediary server provides the descriptive information, i.e. the vocabulary and the according semantic descriptions. The tourism suppliers provide the data themselves or the URI where the server can access the up-to-date data like the number of vacant hotel rooms.

As RDF allows flexible integration of different vocabularies by the means of XML namespaces, individual suppliers are allowed to extend the core data model by individual descriptive RDF information. Accordingly, the mediator can allow individual extensions of data schemes, but can also assure global integrity.

In the RDF schemes in figure 5, a hotel room on a hotel market is described using RDF XML syntax. The presented schema describes a respective class and its properties. Each property is in the domain of a class and has a range that constrains the property values. Within the range, the property data types are provided by URI references to schemes, describing the permitted data values. RDF classes and properties may also contain comments as well as labels, which can be defined language-dependent.

In the example two namespaces are referred to: RDF by the identifier `rdf`: and RDF Schema by the identifier `rdfs`:. The specified schema itself establishes a new namespace.

The class `HotelRoom` is described by the properties `roomView` and `bedNumber`. The `bedNumber` property is applicable to any instance of `HotelRoom` and its value has to correspond to the `Integer_One_to_50` description, which is provided at the datatypes.org server by the belonging RDF definition. For declaring constraints associated with classes and properties the RDF Schema properties `range` and `domain` are used. The `domain` property states on which class resources a property may be used. Complementary, the `range` property expresses that the value of a property should be a resource of a designated class.

The values for the `roomView` property must be one of `{Mountainview, Streetside, Seaside, Other}`. Thus, both a `roomView` property and a class `RoomView` is defined. The `rdfs:range` is used to state that a `roomView` property only makes sense if it has a value which is an instance of the class `RoomView`.

There is another, individual schema that extends the main schema by the use of a property called `wwwAccess`, which describes the available WWW access in a hotel room and therewith extends the basic schema. The range of this property is a `boolean`.

The concrete RDF model is presented as a graph in figure 6. It makes use of the namespace facility and imports both schemes by identifying the associated URIs and assigning identifiers. In our example the basic hotel schema is identified by the prefix `hotel:`, the extension is identified by the prefix `hotelA:`. The schema identifier for RDF itself is referred to by `rdf:`.

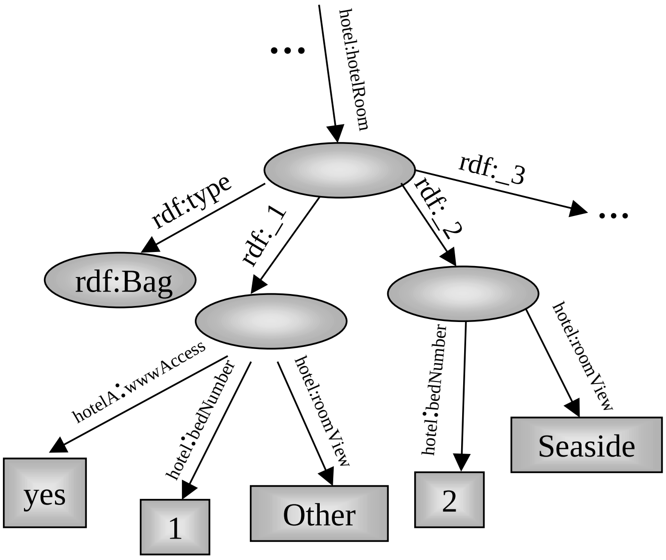

Fig. 6: Hotel RDF Model

In the example, there exists a bag, i.e., an unordered sequence of hotel rooms. This RDF construct is defined by the property `rdf:type` and the resource `rdf:Bag`. The membership properties are named `rdf:_1`, `rdf:_2`, `rdf:_3`, etc. The second hotelroom resource in the bag is described by the properties `roomView` and `bedNumbers`, which are provided by the core schema (`hotel:`), the first resource is additionally described by the property `wwwAccess`, defined in the extended data schema (`hotelA:`). This RDF model can be a part of a more comprehensive hotel market description and may be serialized to XML syntax.

The mediator is responsible for integrating the different hotel descriptions into one global data schema as well as defining the interfaces for combining services with other domains like, for example, a tennis court ontology. The integration of ontologies of different domains can be conducted by another metalanguage that may also deal with business rules.

V. CONCLUSIONS

With the emergence of new WWW technologies like XML and RDF new possibilities for information integration are given. They enable more flexibility as well as ad-hoc integration possibilities that have been main obstacles for a full potential of web based e-commerce so far. In this context we outlined the possibilities of RDF. RDF schema as well as XML schema work is not yet finalized.

Mediated architectures may form the base for new intermediaries, which will, however, also offer new services based on such a flexible infrastructure. In our future work we will focus on the definition of a metamodel within an international cooperation and on the definition of mapping rules between the global and the respective local schemes.

VI. REFERENCES

[1] Wiederhold, G., Genesereth, M.: The Conceptual Basis for Mediation Services. IEEE Expert Intelligent Systems 12/5, 38 - 47, 1997.

[2] Werthner, H., Klein, S.: Information Technology and Tourism - A Challenging Relationship. Springer. Vienna. 1999.

[3] Merz, M.: Elektronische Dienstemärkte. Modelle und Mechanismen des Electronic Commerce. Springer. Berlin. 1999. (in German)

[4] Guarino, N.: Ontologies and knowledge-bases - towards a terminological clarification, in Towards Very Large Knowledge Bases - Knowledge Building and Knowledge Sharing. IOS Press. Amsterdam. 1995.

[5] Uschold, M., Gruninger, M.: 1996, Ontologies: Principles, Methods and Applications, Knowledge Engineering Review, 11/2.

[6] Russell, S., Norvig P.: Aritificial Intelligence. Prentice Hall. 1995.

[7] Swick, R.R., Thompson, H.S.: The Cambridge Communiqué, W3C Note, http://www.w3.org/TR/1999/NOTE-schema-arch-19991007, 1999-10-07

[8] McEntire, R., Karp, P., Abernethy, N., Olken, F., Kent, R.E., DeJongh, M., Tarczy-Hornoch, P. et.al.: An Evaluation of Ontology Exchange Languages for Bioinformatics. ftp://smi.stanford.edu/pub/bio-ontology/OntologyExchange.doc , 1999-08-01

[9] Berners-Lee, T., Connolly, D., Swick R. R.: Web Architecture: Describing and Exchanging Data. W3C Note, http://www.w3.org/1999/04/WebData, 1999-06-07

[10] Lanyon, N., Cunningham, P.: The OpenTravel Initiative – An e-commerce strategy to enable the seamless exchange of travel information. Version 1.0, http://www.xoutech.com/xou_oti.pdf, 1998.

[11] United Nations Directories for Electronic Data Interchange for Administration, Commerce and Transport, http://www.unece.org/trade/untdid.

[12] Bryan, M.: Guidelines for using XML for Electronic Data Interchange, http://www.geocities.com/WallStreet/Floor/5815/guide.htm, 1998.

[13] Bray, T., Paoli, J., Sperberg-Mc Queen, C.M.: Extensible Markup Language (XML) 1.0, W3C Recommendation. http://www.w3.org/TR/1998/REC-xml-19980210, 1998-02-10

[14] Bray, T., Hollander, D., Layman, A.: Namespaces in XML, W3C Recommendation. http://www.w3.org/TR/1999/REC-xml-names-19990114

[15] Berners-Lee, T., Fielding, R., Irvine, U.C., Masinter, L.: Uniform Resource Identifiers (URI): Generic Syntax. Internet Engineering Task Force Network Working Group. RFC 2396.. http://www.ics.uci.edu/pub/ietf/uri/rfc2396.txt, August 1998.

[16] The Knowledge Interchange Format (KIF), http://logic.stanford.edu/kif/kif.html

[17] Fikes, R., Farquhar A.: Distributed Repositories of Highly Expressive Reusable Ontologies. IEEE Intelligent Systems. March/April 1999. 73-79.

[18] Lassila, O., Swick, R. R.: Resource Description Framework (RDF) Model and Syntax Specification. W3C

search

Recommendation, http://www.w3.org/TR/REC-rdf-syntax/, 1999-02-22

[19] Brickley, D., Guha, R. V.: Resource Description Framework (RDF) Schema Specification. W3C Proposed Recommendation, http://www.w3.org/TR/PR-rdf-schema/, 1999-03-03

[20] CycL: The CYC Representation Language, http://www.cyc.com/tech.html#cycl

[21] The LOOM Project, http://www.isi.edu/isd/LOOM/LOOM-HOME.html

[22] The W3C Metadata Activity: http://www.w3.org/Metadata/

[23] Die Touristische-Informations-Norm (TIN) für den deutschen Fremdenverkehr. Bonn/Munich. May 1992. (in German)

[24] Davidson, A., Fuchs, M., Hedin, M., Jain, M., Koistinen, J., Lloyd, C., Malnoey, M., Schwarzhof, K.: Schema for Object-Oriented XML 2.0. W3C Note, http://www.w3.org/TR/NOTE-SOX, 1999-07-30

[25] Document Definition Markup Language (DDML) Specification, Version 1.0, W3C Note, http://www.w3.org/TR/NOTE-ddml, 1999-01-19

[26] Malhotra, A., Maloney, M.: XML Schema Requirements. W3C Note, http://www.w3.org/TR/NOTE-xml-schema-req, 1999-02-15.