

# AN ANALYSIS OF B2B CATALOGUE INTEGRATION PROBLEMS

## *Content and Document Integration*

Borys Omelayenko, Dieter Fensel

*Division of Mathematics and computer Science, Vrije Universiteit,*

*De Boelelaan 1081a, 1081hv, Amsterdam, The Netherlands*

*Email: {borys, dieter}@cs.vu.nl*

*URL: www.cs.vu.nl/~{borys, dieter}*

Key words: B2B Catalogue Integration, Content Integration, Document Integration

Abstract: Content Management becomes a cornerstone of successful B2B electronic commerce. The B2B players use different document standards to represent their business documents, and different content standards to specify the products. Thousands of the players meet together at B2B marketplaces, and the marketplaces must be able to integrate numerous document and content standards. The large number of the standards and their significant complexity make the integration problems non-trivial and require development of special integration architecture. In the present paper we discuss the tasks and the problems which occur during the content and document integration, and survey possible solutions and available techniques.

## 1. INTRODUCTION

Nowadays e-business includes thousands of companies, which provide a tremendous number of products for electronic markets. Unlike the traditional markets, the electronic markets allow the participants to easily compare the offers and perform business negotiation faster due to electronic information interchange between the market participants provided by the Business-to-Business (B2B) marketplaces.

Forecasts for the dollar value of B2B EC in the US range between \$600 billion to \$2.8 trillion for 2003 (cf. (U.S. Department of Commerce, 2000)). Electronic marketplaces for B2B electronic commerce bring together many online suppliers and buyers, which participate in the business interactions (cf. (Fensel, 2001) for an overview of the field). The (U.S. Department of Commerce, 2000) estimates around 800 B2B marketplaces in early 2000 and other studies estimate around 10,000 B2B marketplaces in the very near future. However, the marketplaces have to deal with the problem of heterogeneity in *product*, *catalogue*, and *document* description standards of their customers. Effective and efficient management of different description styles become a key task for these marketplaces. In addition, a number of serious mapping problems, which have to be solved to make the B2B area working. The B2B e-commerce includes an

interaction between three different types of standards:

- Different standards for describing exchangeable business documents like purchase orders (e.g. XML Common Business Library xCBL<sup>1</sup> and Commerce XML cXML<sup>2</sup>).
- Content standards, which specify the products with the hierarchy of product categories and the attributes of each category (e.g. UN/SPSC<sup>3</sup>, ecl@ss<sup>4</sup>, and UCEC<sup>5</sup>).
- Product catalogues, which specify the products according to some content standard and which are referenced by the exchangeable business documents (e.g. Ariba<sup>6</sup> and CommerceOne<sup>7</sup>).

From the technical point of view the content management task for the B2B area includes two major types of information mapping: mapping business documents and aligning content standards.

<sup>1</sup> [www.xcbl.org](http://www.xcbl.org)

<sup>2</sup> [www.cxml.org](http://www.cxml.org)

<sup>3</sup> [www.unspsc.org](http://www.unspsc.org)

<sup>4</sup> [www.eclass.de](http://www.eclass.de)

<sup>5</sup> [www.ucec.org](http://www.ucec.org)

<sup>6</sup> [www.ariba.com](http://www.ariba.com)

<sup>7</sup> [www.commerceone.com/solutions/business/content.html](http://www.commerceone.com/solutions/business/content.html)

In the paper we discuss how the above standards must be integrated to facilitate B2B e-commerce. We discuss the content standards in Section 2, and their integration problems in Section 3, followed by the document integration task discussed in Section 4. We finalize the paper with the future research directions and conclusions in Section 5.

## 2. THE CONTENT STANDARDS

The content standards provide a hierarchy of product descriptions and define the subclass-of relationship between the product categories. Each product from a product catalogue has an attached link to a certain product category, which describes the product. The content standards may be classified into ‘horizontal’ standards and ‘vertical’ standards (cf. Figure 1). As shown in the figure, the horizontal standards provide a high-level classification of *all* possible products and cover many domains. Each of the vertical standards provides a deep and narrow classification of a certain limited domain, e.g. the domain of IT devices. Normally, a vertical standard expands several bottom-level categories of a horizontal standard.

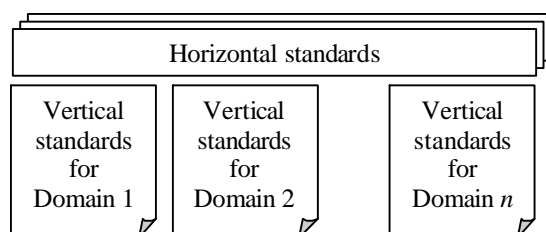


Figure 1. The relation between the horizontal and vertical standards

The most well-known horizontal standard UN/SPSC has a 5-level classification scheme with more than 12,000 categories. It is not descriptive, that is it contains no attributes for the products but only the hierarchy of product names. Consequently, the next initiative, UCEC provides an extension of the UN/SPSC standard with attributes. It uses only four top levels of the UN/SPSC classification hierarchy and provides a couple of attributes for each category, for example, the category [44-12-15-05] ‘Specialty envelopes’ has six attributes: type, length, width, weight, colour, and composition. Another horizontal standard ecl@ss supports the flow of products and information along the supply chain of an industrial enterprise and is mainly used in Germany. It provides the attributes for each of more than 12,700 categories tailored to the needs of industrial customers and their suppliers.

There exist a number of ‘vertical’ standards. ISO<sup>8</sup> provides many product coding standards specific for certain branches of industry. RosettaNet<sup>9</sup> contains a catalogue of IT products with a categorization, attributes and mapping of each category to UN/SPSC. Large manufacturers tend to develop special product cataloguing schemes, and usually these schemes are reflected in the product coding system used by the company. Finally, the company may have its own focus and therefore require specific product classification scheme.

Hence, the products may be classified in the suppliers’ catalogue according to some certain content standard, which may differ from the content standard used by the buyer to specify its needs. In consequence, a B2B mediation system must be able to reclassify a product, already classified once, according to another classification schemas.

## 3. ALIGNING THE CONTENT STANDARDS

To reclassify the products a B2B marketplace must be able to perform three different types of mappings:

- Aligning two horizontal standards.
- Linking a horizontal standard to a vertical standard.
- Aligning two vertical standards.

These mappings provide different problems as discussed in the rest of this section.

### 3.1 Aligning Horizontal Standards

The horizontal standards provide general hierarchies of products, and we can expect many equivalence mappings to appear between them. Sometimes the content standard providers publish these mappings as a part of the standard. For example, UN/SPSC provides direct one-to-one mapping between UN/SPSC and other ‘horizontal’ standards: CPV, NAICS, and SIC<sup>10</sup>.

Despite these published mappings, a number of problems arise in aligning the horizontal standards:

<sup>8</sup> [www.iso.ch](http://www.iso.ch)

<sup>9</sup> [www.rosettanet.org](http://www.rosettanet.org)

<sup>10</sup> [www.unspsc.org/crosswalk.html](http://www.unspsc.org/crosswalk.html)

- Only few horizontal standards have officially published mappings, and most of the standards are not related one to each other.
- The standards significantly differ in their classifications because of the absence of a consensus scheme for classifying *all* products.
- The standards differ in the granularity level in the classifications of each particular group of products. Hence, very often the published mappings list the concepts with different generality level as equivalent.
- The equivalence of the categories is not evident from their descriptions, e.g. NAICS<sup>11</sup> code [39] ‘Miscellaneous Manufacturing Industries’ is mapped to the UN/SPSC code [73] ‘Industrial Production and Manufacturing Services’.

An example of aligning two horizontal standards is shown in Figure 2. In general, aligning two horizontal standards has the following properties:

- It is based on the published official mappings.
- It contains additional mappings created by the user, which extend the set of official mappings.
- It contains multiple mappings if the pair of standards has overlapping categories.

---

‘Rollerball pens’ are subclassed in UN/SPSC in the following way:

[44] Office Equipment and Accessories and Supplies

[4412] Office supplies

[441217] Writing instruments

**[44121701] Rollerball pens**

[44121702] Pen or pencil sets

...

[44121711] Felt pen

ecl@ss classifies rollerball pens as Writing material follows:

[24] Communication technology

[24-11] Office Supplies

[24-11-01] Writing and drawing materials

**[24-11-01-01] Writing material**

[24-11-01-02] Drawing material

---

Figure 2. The example of mapping two horizontal standards

### 3.2 Linking Vertical and Horizontal Standards

Linking the vertical and horizontal standards differs from the previous case. The vertical

standards provide a deep and narrow classification opposite to wide and shallow horizontal standards. Normally a vertical standard expands one or few nodes of a horizontal standard.

For example, consider the mapping of the RosettaNet standard for the electronic component and IT supply chain to UN/SPSC. The mapping links only 136 UN/SPSC elements out of more than 12,000, most of which belong to the bottom level in the UN/SPSC hierarchy, to 445 categories and 2660 attributes of RosettaNet.

The vertical standards are very precise in describing the items they are focused. The same time they are even shallower than the horizontal standards in describing the things, which lay beyond their focus.

Aligning the vertical and horizontal standards requires:

- Mapping a relatively small number of top-level vertical concepts with more general concepts of the horizontal standard.
- Mapping the concepts, which are outside the focus of the vertical standard to the correspondent concepts in the horizontal standard. In this case the vertical concepts may have the same granularity level as the horizontal concepts.

This linking is simpler than the horizontal aligning and it has an evident top-down structure. Hence, technologically it can be treated as a light version of the horizontal mapping and it can be performed with the same technologies.

### 3.3 Aligning Vertical Standards

Aligning vertical standards requires linking their categories in a similar way as it is done for the horizontal standards. In addition the vertical standards have an extensive set of attributes, which can be even larger than the set of classes, as it is in the RosettaNet case. Each attribute can be represented with:

- Attribute name, e.g. ‘Screen size’.
- Name abbreviation, e.g. ‘ScreenSize’, which is a valid identifier produced from the attribute name.
- Attribute value type (e.g. string, integer, float, etc.). The type may be an enumerated type represented with a list of possible attribute values, e.g. the currency type actually contains a list of possible currencies.
- Attribute value format, which defines the way to interpret the attribute value. For example, ‘YYYY-MM-DD’ denotes that the date is

<sup>11</sup> www.naics.com

represented in a year-month-day format, e.g. 2001-02-16.

- The scale for the values, e.g. 'm' stands for meters as a length measure.
- Attribute domain(s), or the set of categories to which the attribute can be applied.

Hence, aligning the vertical standards requires:

- Mapping attribute names and attribute abbreviations.
- Transforming attribute types, e.g. transforming an integer value into a corresponding string representation.
- Mapping the list of possible values for the enumerating types.
- Mapping different value display formats.
- Transforming between the unit scales, e.g. translation of the length in meters into the length in feet.
- Mapping attribute domains: the list of the categories from the source standard, to which the attribute applies, must be translated into the list of the categories from the target standard. This translation exploits previously defined mappings

of the categories.

A number of potentially useful aligning techniques have been developed. The database community provides a number of approaches for database schema integration (Poulovassilis&Brien, 1998), (Batini, Lenzerini, Navathe, 1986). The knowledge engineering community provides a tool support with Protégé (Grosso et al., 2000), Chimaera (McGuinness et al., 2000), and PROMPT (Noy&Musen, 2000); and inference-based integration (cf. (Chalupsky, 2000)). The achievements from these areas must be combined together to solve the aligning problems of content management.

#### 4. DOCUMENT INTEGRATION

The standards for B2B document exchange contain a large number of different documents to be exchanged between the market participants. For example, the xCBL standard proposes a document infrastructure described with 594 XML DTD's.

```
<IndexItem>
  <ItemDetail>
    <UnitPrice>
      <Money currency="USD">1000</Money>
    </UnitPrice>
    <Description xml:lang="en"> Armada M700 PIII 500 12GB
  </Description>
  <UnitOfMeasure>EA</UnitOfMeasure>
  <Classification domain="SPSC"> C43171801</Classification>
  <ManufacturerPartID>140141-002</ManufacturerPartID>
  <ManufacturerName>Compaq</ManufacturerName>
  <URL>http://www.compaq.com</URL>
</ItemDetail>
<IndexItemDetail>
  <LeadTime>10</LeadTime>
  <ExpirationDate>2000-06-01</ExpirationDate>
  <EffectiveDate>2000-01-01</EffectiveDate>
  <SearchGroupData>
    <Name xml:lang="en">Notebook</Name>
    <SearchDataElement name="Processor Speed" value="500MHZ"/>
  </SearchGroupData>
  <TerritoryAvailable>USA</TerritoryAvailable>
</IndexItemDetail>
</IndexItem>
```

```
<CatalogSchema>
  <SchemaVersion>1.0</SchemaVersion>
  <SchemaStandard>UNSPSC</SchemaStandard>
  </CatalogSchema>
  <Product Type="Good" SchemaCategoryRef="C43171801">
    <ProductID>140141-002</ProductID>
    <Manufacturer>Compaq</Manufacturer>
    <CountryOfOrigin>
      <Country><CountryCoded>US</CountryCoded></Country>
    </CountryOfOrigin>
    <ShortDescription xml:lang="en">Armada M700 PIII 500 12GB
  </ShortDescription>
  <LongDescription xml:lang="en">This light, ...
</LongDescription>
  <ObjectAttribute>
    <AttributeID>Processor Speed</AttributeID>
    <AttributeValue>500MHZ</AttributeValue>
  </ObjectAttribute>
  <ProductVendorData PartnerRef="Acme_Laptops">
    <VendorPartNumber>12345</VendorPartNumber>
  </ProductVendorData>
  <ProductPrice>
    <Amount>1000</Amount>
    <Currency>
      <CurrencyCoded>USD</CurrencyCoded>
    </Currency>
  </ProductPrice>
</ProductVendorData>
</Product>
```

(a) cXML

(b) xCBL

Figure 3. Two fragments of product catalogues

Other standards define a document infrastructure of a similar complexity (see (Li, 2000) for a comparison). Let us consider a fragment of a product catalogue as defined by the cXML and xCBL formats and presented in Figure 3 (a) and (b) respectively.

Both catalogs contain two parts: static catalog information and dynamic information. The static part contains the descriptions, which are not updated frequently, such as product name and description, its UN/SPSC code and a manufacturer. The dynamic part contains the descriptions, which can be updated very often and will be sent to the user on request. However, different concepts are regarded as dynamic in xCBL and cXML. According to the xCBL format, product attributes are present in the static part, while the price is regarded as a dynamic part, requested from the vendor (and it appears in the ProductVendorData section). This scenario assumes that the user accesses full descriptions of all the products, while the price can change in time.

cXML expects another implementation scenario partially targeted to the needs of B2B website development. It assumes that the user browses through the descriptions of the first interest, such as product name, content classification, and the price. Detail product information, such as LeadTime, ExpirationDate and other is available on request (and regarded as a dynamic part and appears in the IndexItemDetail section) together with product attributes, which are treated as the static part in the previous case.

In the rest of this section we discuss the direct mapping of different representations which provides a partial solution to the integration task. Then we discuss a multi-layered framework, which eliminates some of the problems of the direct mapping approach.

## 4.1 Single-Layer Integration

The documents represented in Figure 3 represent the same information, however several differences occur in the document formats (see (Omelayenko&Fensel, 2001) for relevant discussion and examples):

- Different terminologies are used, i.e. the tag names used to denote semantically equivalent elements (these differences are called naming conflicts in the database schema integration area). For example, the price is marked up in xCBL with the tag Amount, while cXML uses the tag Money.
- The standards can use either XML attributes or XML elements to represent semantically equivalent information. For example, xCBL uses

the XML element CurrencyCoded to encode the currency, while cXML uses the XML attribute currency; in both cases the currencies are denoted in the same way.

- Different value formats and encoding conventions may be used. For example, as shown in Figure 3, the reference to UN/SPSC is encoded with the attribute domain in cXML, value domain="SPSC". At the same time the xCBL standard encodes the same link with the element SchemaStandard, value "UNSPSC". Hence, the values of the attributes and elements must be translated in addition to the element mapping.
- Different scales may be used for the values. For example, the price in US dollars has to be scaled for comparison to the price in DM. Unlike the representational differences shown in the previous example, the latter require the scales to be properly verified and timely updated.
- Different natural languages may be used in the *tag values*, as marked up with the XML xml:lang attribute.
- In addition to a multi-lingual tag values we can expect that some national document standards may use other languages in the *tag names*.

The standards are often represented in XML (see (Li, 2000) for a survey) and this tendency dominates. The W3C<sup>12</sup> consortium provides the standard architecture for XML document integration with the XSL-T (Clark, 1999) and XPath (Clark&DeRose, 1999) languages.

Hence, the document integration task can be principally resolved by means of the XSL-T language. This requires development of a set of XSL-T rules able to translate one XML serialization to another one. Direct document transformation with XSL-T rules is discussed in (Omelayenko&Fensel, 2001) and appeared to be a partial solution, and have raised a number of problems. The problems arise from the fact that this approach mixes several independent tasks in a single batch of XSL-T rules:

- Aligning different terminologies.
- Aligning the granularity level of the representations and performing necessary attribute splits with XPath expressions. Very often, this splitting is guided by ad-hoc rules, which split based on the element values. For example, one standard may store a street name and a house number address components in a

<sup>12</sup> www.w3c.org

single element, while another standard may allocate two separate elements for them.

- Transformation the attribute values.
- Restoring necessary formatting according to the target document standard.

The problems of the single-layer integration appear because two tasks run together with a single bungle of transformation rules: syntactical translations between different XML representations and semantical mapping between the terminology and granularity level of the representations. Naturally, these two types of transformations belong to different layers of representation.

The layered approach for information representation on the Web was proposed in (Melnik&Decker, 2000), where three layers, syntax layer, object layer, and semantic layer are proposed for information modelling on the Web. The syntax layer provides a way of serializing information content into a sequence of characters according to some standard, e.g. XML. The purpose of the object layer is to offer an object-oriented view on the information with the normalized data models of standardized triples. Finally, the semantic layer provides a conceptual model for the information. We use this partitioning to base our integration architecture on it.

Multi-layered integration provides a solution for these problems, as discussed in the next section.

## 4.2 Multi-Layer Integration

In the multi-layer approach (see (Omelayenko&Fensel, submitted) for a detailed discussion) we separate three layers of information representation, which are *Syntax layer*, *Data Models layer*, and *Ontology layer*.

**The Syntax layer** corresponds to the instance documents represented with their XML serialization. The serialization specifies the XML elements and attributes used, and their order. Even semantically equal documents may differ in their serialization.

**The Data models layer** serves as a bridge between the Ontology layer and the Syntax layer. On this layer the representations are abstracted from the differences imposed by the Syntax layer and the products are represented by object-property-object triples, where the attributes stand for products' attributes. The normalization is done according to the corresponding ontology which specifies the semantics of the elements at the granularity required for successful element splitting.

The terminology used on this layer is defined by the corresponding ontology and generally must coincide with the one used on the Syntax layer.

However, the former might be more detailed than the latter, e.g. the XML serialization may allocate only one element for street name and house number, while the ontology must allocate two separate elements. We assume that different terminologies must be aligned on the Ontology layer rather than on the Data Models layer.

The most suitable language candidate to encode the triples on this layer is RDF (Lassila&Swick, 1999), a W3C standard for describing of machine-processable semantics of data also represented with the object-attribute-value triples. Another possible candidate is Simple Object Access Protocol (SOAP)<sup>13</sup>.

**The Ontology layer** corresponds to the document ontologies used to represent the products. We assume that this layer specifies the documents in a detail level, sufficient enough to specify the transformations between the catalogues with one-to-one mapping rules. In addition, the ontology contains the elements specified as optional and possibly absent in the XML serialization and, therefore, helps in aligning them. Despite the fact we sometimes reference to this layer throughout the paper, we do not discuss further possible ontology mismatches or integration problems, which may arise on this layer, see for example (Klein, submitted).

## 4.3 The Integration Process

As we mentioned before, the difficulties of the single-layered representations are caused by several integration tasks running together. Therefore, we use a 'divide-and-conquer' approach to decompose these tasks into several subtasks, each of which is performed separately.

The decomposition is performed in a similar way to the structure of heuristic classification proposed in (Clancey, 1985). Heuristic classification assumes that the classification is performed on a layer of abstract structures, and the input data must be first abstracted, i.e. translated from some particular format into the abstract structure; after the classification it must again be refined from the abstract structure to specific solutions.

To realize this strategy (see Figure 4) we have developed a multi-layered approach (Omelayenko&Fensel, submitted) which assumes that the integration is performed at least via two

<sup>13</sup> [www.w3.org/TR/SOAP/](http://www.w3.org/TR/SOAP/). See (Haustein, 2001) for a comparison of SOAP with RDF

layers: the syntax layer of the actual XML documents and the layer of the normalized data models for the catalogues. Accordingly, the integration process passes through three steps: the translation of the source XML catalogue into its normalized data model on the *abstraction step*, the translation between a pair of data models of different catalogues on the *transformation step*, and the translation from the data model back into XML according to the target XML format on the *refinement step*.

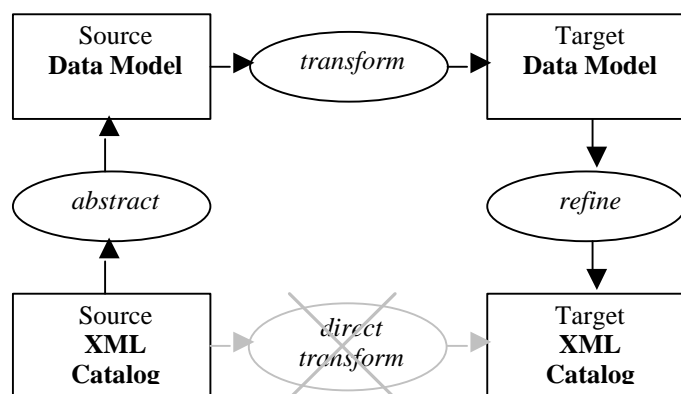


Figure 4. The model for data transformation

On the abstraction step the XML catalogues are translated into their normalized data models encoded with the RDF triples. This requires the following transformations:

- The translation of each XML element or XML attribute, which refers to a product feature into an RDF property with the same name (however, the latter may have another namespace).
- The split of a single XML element into two or more RDF triples, if this is required for the integration.
- The combination of multi-file descriptions into a single file, if required.
- Inclusion of the optional XML elements in the RDF triples; the values of the elements are filled with the default value.

All inter-catalogue mappings are performed on the layer of the normalized RDF data models. We assume that all necessary element splits have been performed during the abstraction stage and necessary element merges will be done on the refinement stage. Hence, only two types of mappings may appear between the attributes of the two catalogues: one-to-one mapping and many-to-many mapping. The latter requires no attribute splits or merges and can be easily expressed with a set of one-to-one mappings. The XSL-T rules for this layer

may look like the following (from cXML element Money to xCBL element ProductPriceAmount):

```

<xsl:for-each select="rdf:Description">
  <SchemaVersion>1.0</SchemaVersion>
  <ProductPriceAmount>
    <xsl:value-of select="Money"/></ProductPriceAmount>
    <ProductPriceCurrency>
      <xsl:value-of select="currency"/></ProductPriceCurrency>
    ...
</xsl:element>
  
```

During the refinement step all syntactical restrictions required by the target format are restored, and the necessary many-to-one transformations are performed. The rules must be able to perform the following transformations, if required by the target standard:

- Each RDF triple is translated into a corresponding XML element, XML attribute, or non-XML entity for a non-XML catalogue.
- The target XML elements are created in a proper order.
- One or more RDF triples have to be merged into a single XML element, if required.
- The XML representation may be partitioned into several files, if required.

In consequence, only one-to-one and many-to-one rules occur on this stage.

## 5. CONCLUSIONS

In the paper we discussed two problems, which are quite important for the B2B area: content integration and document integration. Each of the problems can be solved with an ad-hoc solution. However, given the very large amount of required mappings of content standards (more than 12,000 concepts plus several times more attributes) and the large amount of documents (400 documents and 5 different standards already require around 100,000 mappings) and document standards, this approach does not scale up to the actual needs of effective and efficient B2B e-Commerce.

Therefore, we developed a conceptual model for the mapping process with two main contributions:

- Dividing the overall mapping process into different steps.
- Identifying different layer that represent different aspects of the overall mapping process.

Both aspects help to reduce the overall complexity of the process and allow reusing simple rule patterns to actually define the mappings

Currently we define a simple rule pattern language on top of XSL-T customized to the specific integration needs of electronic commerce. Instead of defining transformation directly in XSL-T transformation by hand, they should be derivable from selecting and instantiating mappings defined at a more intuitive level. We are aiming on transforming a complex programming task into a simple plug-and-play process.

The Ontology layer for document integration has to be elaborated to handle the necessary information about element splits and terminology correspondence. Further elaboration of the integration techniques requires ontology aligning to guide the transformations on the lower layers of the model. Based on a unifying ontological representation layer we will overcome the exponential explosions in the number of required mappings.

## ACKNOWLEDGEMENT

The authors would like to thank Ellen Schulten for her helpful consultancy and discussions, and Volodymyr Zykov for his discussions.

## REFERENCES

- (Batini, Lenzerini, Navathe, 1986) Batini, C., Lenzerini, M., Navathe, S., 1986, 'A comparative analysis of methodologies for database schema integration', *ACM Computing Surveys*, 18(4), p. 323-364.
- (Chalupsky, 2000) Chalupsky, H., 2000, 'OntoMorph: A Translation System for Symbolic Knowledge', *In Proceedings of the Seventh International Conference on Knowledge Representation and Reasoning (KR-2000)*, Breckenridge, Colorado, USA, April 12-15.
- (Clancey, 1985) Clancey, W., 1985, 'Heuristic Classification', *Artificial Intelligence*, 27, p. 289-351.
- (Clark&DeRose, 1999) Clark, J. and DeRose, S., 1999, 'XML Path Language (XPath)', *W3C Recommendation*; available online at <http://www.w3.org/TR/xpath>
- (Clark, 1999) Clark, J., 1999, 'XSL Transformations (XSL-T)', *W3C Recommendation*; available online at <http://www.w3.org/TR/xslt/>
- (Fensel, 2001) Fensel, D., 2001, *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*. Springer-Verlag, Berlin.
- (Grosso et al., 2000) Grosso, W., Eriksson, H., Ferguson, R., Gennari, J., Tu, S., and Musen, M., 1999, 'Knowledge modeling at the millennium (the design and evolution of Protégé-2000)', *In Proceedings of the Twelfth Banff Workshop on Knowledge Acquisition, Modeling, and Management*, Voyager Inn, Banff, Alberta, Canada, October 16-21.
- (Haustein, 2001) Haustein, S., 2001, 'Semantic Web Languages: RDF vs. SOAP Serialization', *In Proceedings of the Workshop on the Semantic Web - SemWeb'2001 at the 10-th WWW Conference*, Hong Kong, May 1.
- (Klein, submitted) Klein, M. Combining ontologies: an analysis of problems and solutions, submitted; available online at <http://www.cs.vu.nl/~mcaklein/>
- (Lassila&Swick, 1999) Lassila, O. and Swick, R., 1999, 'Resource Description Frame-work (RDF) Model and Syntax Specification', *W3C Recommendation*, February; available online at <http://www.w3.org/TR/REC-rdf-syntax/>
- (Li, 2000) Li, H., 2000, 'XML and Industrial Standards for Electronic Commerce', *Knowledge and Information Systems*, 2(4), p. 487-497.
- (McGuinness et al., 2000) McGuinness, D., Fikes, R., Rice, J., Wilder, S., 2000, 'An Environment for Merging and Testing Large Ontologies', *In Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR-2000)*, Breckenridge, Colorado, April 12-15.
- (Melnik&Decker, 2000) Melnik, S. and Decker, S., 2000, 'A Layered Approach to Information Modeling and Interoperability on the Web', *In Proceedings of the Workshop on the Semantic Web at the Fourth European Conference on Research and Advanced Technology for Digital Libraries (ECDL-2000)*, Lisbon, Portugal, September 21.
- (Noy&Musen, 2000) Noy, N. and Musen, M., 2000, 'PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment', *In Proceedings of the 17-th National Conference on Artificial Intelligence (AAAI-2000)*, Austin, Texas, July 30 – August 3.
- (Omelayenko&Fensel, 2001) Omelayenko, B. and Fensel, D., 2001, 'An Analysis of the Integration Problems of XML-Based Catalogues for B2B Electronic Commerce', *In Proceedings of 9th IFIP 2.6 Working Conference on Database Semantics*, Hong Kong, April 25-28.
- (Omelayenko&Fensel, submitted) Omelayenko, B., and Fensel, D., 'A Two-Layered Integration Approach for Product Catalogs in B2B E-commerce', submitted; available online at <http://www.cs.vu.nl/~borys/papers/>
- (Poulovassilis&Brien, 1998) Poulovassilis, A. and Brien, P., 1998, 'A General Formal Framework for Schema Transformation', *Data & Knowledge Engineering* 28, p. 47-71.
- (U.S. Department of Commerce, 2000) U.S. Department of Commerce, 2000, *Digital Economy 2000, White paper*, June.