

Ontologie pro e-Commerci

Pavel Jisl

30. dubna 2003

1 Úvod

Ve své práci se zabývám nejen porovnáním, ale také diskutuji možnosti převodu datových formátů, používaných v oblasti e-Commerce. Pro podrobný popis jsem vybral dva z běžně používaných formátů - *cXML* [1] a *xCBL* [2]. Jejich podrobný popis, včetně definic typů dokumentů, lze nalézt na příslušných stránkách.

Jazyk *cXML* vytváří a udržuje firma *Ariba Inc.*¹ a jedná se o otevřené řešení, založené na jazyku XML [5]. Definice je volně šířena ve tvaru DTD² a dodává se k němu také elektronická publikace - *cXML Users Guide* [3]. Druhým jazykem je *xCBL* (literatura na stránce [4]), který je vyvíjen firmou *CommerceOne*³ a je založen na standardu *EDI/FACT*⁴, ale využívá zápisu v jazyce XML. Jeho definice lze získat v mnoha různých formátech (kromě DTD také např. ve formátu vhodném pro server BizTalk⁵ firmy Microsoft).

2 Přístup k problému

Jednotlivé formáty jsou dostatečně využívány, mají podporu u velkých společností, ale při jejich používání vzniká problém naznačený v obrázku 1. Pokud *product catalog 1* bude zapsán např. v *cXML* a *product catalog 2* v *xCBL*, je nutné navrhnout *translation server*. Do té doby mezi sebou tito zákazníci nebudou moci komunikovat.

Nejjednodušším přístupem je navrhnout XSL⁶ transformaci (podrobně popsáno v [6]). Ta umožňuje přetransformovat vstupní XML dokument do

¹<http://www.ariba.com>

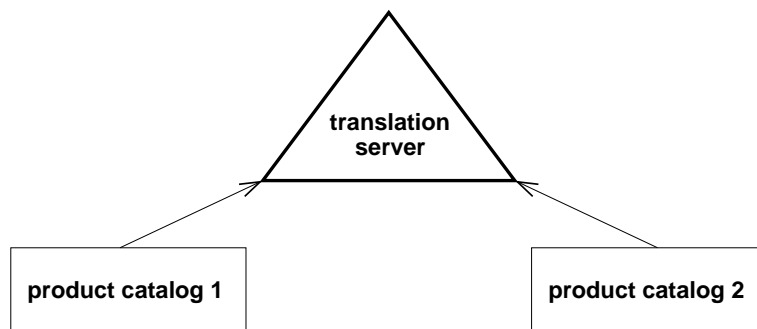
²Document type definition

³<http://www.commerceone.com>

⁴<http://www.unece.org/trade/untdid/welcome.htm>

⁵<http://www.biztalk.org/>

⁶eXtensible Stylesheet Language



Obrázek 1: Schéma transformace různých XML dokumentů

výstupního XML dokumentu pomocí šablon, které je nutné ručně navrhnout. Pokud jsou však požadovány převody mezi obsáhlými formáty dokumentů, může být jejich příprava velmi časově náročná. Hlavním problémem je však obousměrný převod. Pokud je požadován, je nutné navrhnout pro převod mezi dvěma formáty dvě různé šablony. Pak časová náročnost tohoto vývoje roste kvadraticky - $O(n^2)$ (např. pro 6 různých datových formátů je potřeba 36 transformačních šablon).

Pro zjednodušení těchto převodů lze použít metodu, založenou na ontologiích. Ty byly navrženy pro vytváření a sdílení bází znalostí. Základním prvkem ontologie je **class** (třída), což je soubor reálných entit, které mají nadefinované vztahy mezi dalšími entitami. Tyto vztahy jsou nazývané **inheritance** (dědičnost). Jednotlivé třídy pak obsahují soubor **instancí** a **properties** (vlastností). *Instance* jsou entity, příslušející k určité třídě a *properties* jsou pak funkce, která přiřazuje instanci jiné třídě. Tento popis si lze představit jako stromovou strukturu. Bližší informace o ontologiích lze nalézt na stránce [10] a jejich využití v e-Commerce v [8].

Přístup k problému za použití ontologií spočívá v průzkumu definic typu dokumentu (DTD nebo XML-Schema⁷) a rozpoznání smyslu (sémantiky) jednotlivých částí dokumentu. Sémantiky se pak zapíše ve tvaru ontologií, za použití tříd, dědičností a vlastností, čímž získáme informativní model, jak je podrobně popsáno v [9]. Pro zápis sémantiky je výhodné použít jazyk *RDF*⁸ [7]. Ten umožňuje přímé zapsání ontologií a umožňuje definovat také vztahy *instance of* a *subclass of*.

Získaný informativní model vyjadřuje obecné vlastnosti dokumentů a je na nich nezávislý. Na tento model mapujeme jednotlivé elementy z datových formátů a, případně, pokud je třeba např. převádět jednotky, přiřazujeme

⁷<http://www.w3.org/TR/xmlschema-0>

⁸Resource Description Framework

transformační funkce. Výsledkem pak je model, ve kterém jsou vyjádřeny obecné vztahy mezi jednotlivými formáty.

3 Závěr

V úvodu práce byl požadavek k vyjádření shodných informací a převodu datových formátů. Pokud získáme informativní model s namapovanými elementy jednotlivých formátů, lze převod uskutečnit vygenerováním XSL šablony a následným provedením XSL transformace. Vzhledem k tomu, že veškeré informace jsou generovány dynamicky, není zde nutnost ručně přepisovat několik šablon, ale stačí pouze upravit model.

Reference

- [1] Ariba Inc. *cXML - Commerce XML resource* [online]. <<http://xml.cxml.org/current/cXMLUsersGuide.pdf>>.
- [2] Commerce One. *xCBL - XML Common Business Library* [online]. <<http://www.xcbl.org/xcbl40/documentation.html>>.
- [3] Ariba Inc. *cXML Users Guide* [online]. <<http://xml.cxml.org/current/cXMLUsersGuide.pdf>>.
- [4] Commerce One. *xCBL Users Guide* [online]. <<http://www.xcbl.org/xcbl40/documentation.html>>.
- [5] W3 Consortium. *extensible markup language* [online]. <<http://www.w3.org/XML/>>.
- [6] W3 Consortium. *The extensible stylesheet language*. <<http://www.w3.org/Style/XSL/>>.
- [7] W3 Consortium. *The resource description framework*. <<http://www.w3.org/Metadata/>>.
- [8] Fensel D. *Ontologies: Silver bullet for knowledge management and electronic commerce*. <citeseer.nj.nec.com/413498.html>.
- [9] Omalayenko B. and Fensel D. *Rdft: A mapping meta-ontology for business integration* [online]. <citeseer.nj.nec.com/508272.html>.
- [10] Ontology.Org. *Ontology working group*. <<http://www.ontology.org/>>.