

EDI 의미사전 및 브리지의 개발

정용규
서울보건대학 컴퓨터정보과
ygjung@sh.ac.kr

Development of EDI Semantic Dictionary and Bridge

Yong Gyu Jung
Dept. of Computer Information, Seoul Health College

요 약

본 연구에서는 KIEC 라이브러리와 UBL 라이브러리 간에 의미를 기반으로 한 상호운용의 결과물과 연계가 되도록 하며 이를 확장하여 EDIFACT문서를 대상으로 XML-EDI 간의 상호운용이 가능하게 한다. 이를 위해서 우선 KEDIFACT 표준전자문서에 대해 그 사용실태를 조사하고 사용되고 있는 문서를 중심으로 의미를 기반으로 한 항목정의 및 분석을 하였다. 관련된 국내·외 표준을 준용하여 KEDIFACT 표준 전자문서를 대상으로 EDI 의미사전(Semantic Dictionary)을 구축하여 의미를 기반으로 한 EDI 콘텐츠를 구축하였다. 구축한 KEDIFACT 의미사전을 기반으로 기존에 구축되어있는 XML(KIEC, UBL) 의미사전과 의미기반의 비교·분석 작업을 통하여 EDI-XML 라이브러리 간 매핑이 가능하도록 브릿지를 구축하여 상호운용 기반을 마련하였다.

Key words : EDI, 의미사전, 브리지, 매핑, 라이브러리

1. 서론

국내 KEDIFACT 표준전자문서는 14개 업종 263종에 이른다. 이러한 전자문서를 대상으로 의미사전을 구축하게 되면 전자문서의 항목 상호간 의미를 기반으로 개념체계가 정립될 것으로 기대된다. 현재 국내뿐만 아니라 국제적으로도 이러한 시맨틱 표준화가 미흡한 상태이며, 시맨틱 표준관리 체계가 구축되면, 전자상거래 또는 e-비즈니스 시스템에서 사용되는 데이터간의 연동이 단일하게, 표준 체계를 통해서 통합적으로 이루어지므로 업종간, 국제간 전자상거래 시스템 상호운용 구현이 용이하게 된다. 이를 해결하기 위해서 ISO TC154의 워킹그룹에서 BSR 관련연구가 진행된 바 있으며, KIEC에서도 BSR기반의 사업간 DB 통합에 관한 연구를 산업기술평가원의 표준화기술개발과제로 추진하여 현재 KIEC 라이브러리와 UBL 라이브러리 간에 의미를 기반으로 한 상호운용이 가능한 상태이며, BSR을 통해 이기종 간의 전자문서를 변환할 수 있게 되었다. 그러나 EDI 의미 사전 구축은 기존의 기본의미등록기가 XML기반의 이기종 전자문서 간의 상호운용만을 지원하고 아직까지 업계에서 많이 사용되고 있는 EDI와의 상호운용

을 지원하지 않고 있다. 따라서 XML-EDI 간의 상호운용을 가능케 하기 위한 콘텐츠를 구축할 필요가 있다.

2. 표준현황 및 사용실태

각 산업분야별 전자문서는 1990년대 활발히 사용되어 왔다. 특히 무역과 관련된 통관, 상역, 물류부문의 전자문서는 현재도 활발히 사용되어 오고 있다. 국내에서는 국제표준의 흐름을 따라, UN/EDIFACT 표준전자문서(UNSM)가 사용되고 있다. 무역, 상역, 통관 및 물류업종에서는 국제교환문서가 많은 특성에 따라 UNSM의 Sub Set를 표준으로 정하고 이를 사용해 오고 있다. 다만, 의료부문은 국내 자체개발한 문서가 대부분이며 이는 업종의 특성과 관련이 있다고 볼 수 있다. KEDIFACT 표준전자문서는 14개 업종 263종에 이른다. 이러한 263종의 EDI 전자문서는 점차 XML형태로 바뀌어가고 있지만 아직도 많은 부분에서는 널리 사용하고 있다. UN/EDIFACT 표준은 그동안 많은 연구 활동이 이루어져서 공고한 표준이기 때문에, 이미 설계된 UNSM의 기본구조를 재설계할 필요 없이, 기존의 구조를 재사용하는 것이 효과적

이다. 전자문서 설계방식에 대한 분류는 XML이 도입된 이후에 공인된 라이브러리가 없는 상태에서 라이브러리를 자체개발 또는 EDIFACT를 활용한 라이브러리를 개발하여 설계한 방식을 XML/EDI 방식으로 구분하고, 한국전자거래진흥원의 전자문서 개발가이드라인 보급에 따라 설계한 방식을 Component 방식으로 구분하였다. XML/EDI 방식은 전자문서 개발 가이드라인이 보급되기 이전에 개발되었기 때문이다. 이러한 상황에서 EDI 전자문서를 대상으로 의미사전을 구축하게 되면 전자문서의 항목 상호간 의미를 기반으로 개념체계가 정립될 것으로 기대된다.

3. 의미사전 구축

EDI 의미사전은 EDI전자문서와 XML전자문서간의 표준의 상호호환성과 데이터의 상호운용성을 제공할 수 있다. 이런 측면에서 의미사전과 브리지가 구축되려면 두 파일(테이블)간의 연결고리를 할 수 있는 외래키가 있어야 하며, 의미사전의 매핑정보로 EDI의 식별키를 정해야함. 이러한 의미사전을 구축하기 위해서 UNTDED 엘리먼트 중 물류 및 무역 분야에 해당하는 정보를 선정하여 BSR에 추가로 등록하기 위한 작업을 진행하였다. UNTDED 엘리먼트 중 물류 및 무역 분야에 해당하는 BBIE를 선정하였다. BBIE를 의미적으로 유사한 것들을 그룹핑하여 ABIE 또는 ASBIE로 정의하였다. UNTDED에서 하나의 EntryName에 대한 정보를 TBG17 Form 형태에 맞게 변환하였으며, 한글 EntryName을 추가로 정의하였다. 또한 기존 KIEC와 UBL 브리지에 위탁 화물에 대한 정보가 존재하지 않아 브리지에 추가하였다. Dictionary에 명시된 ABIE, BBIE 등에 대해 현재 사용되고 있는 EDI 전자문서의 의미적 내용을 반영하였다. EDI는 동일 세그먼트에 대해 한정어를 추가로 정의하여 사용되고 있으므로, 이에 따라 BIE에 한정어를 추가하는 작업을 수행하였다.

의미사전은 기본적으로 TBG 17 Form으로 변환하여 작업하였으며, 영문 이외에 국내 사용자들의 편의를 위해 한글화 작업을 하였다. EntryName을 기반으로 Object Class Qualifier(Term), Property 등의 항목을 정의하였다. 또한 Dictionary 내 그룹핑 재확인, EDI 문서의 한정어 적용, KIEC, UBL 브리지 보완등을 통해 콘텐츠를 수정 보완하였다. 의미사전은 EDI 전자문서를 기반으로 XML 전자

문서를 만든 경우 반드시 TAG 제출용 XML 전송항목 테이블에 EDI 매핑정보를 추가토록 요구해야 한다. 이는 향후 EDI문서와 XML문서간의 전자문서 작성 시 혹은 데이터 변환 시 필요한 참조사전이 될 수 있기 때문이다. 의미사전(Semantic Dictionary)의 콘텐츠는 다음과 같이 구성하였다.

- Source : 콘텐츠를 개발한 기관의 명을 영문 약어로 표기한다.
- ID : 의미사전(Semantic Dictionary)을 유일하게 구분할 수 있는 식별자
- Entry Name : 의미사전(Semantic Dictionary)의 요소명으로 한글로 표기하며, 한개 이상의 BSC로 구성된다.
- Indicator : 의미사전(Semantic Dictionary)의 속성을 구분한다. 의미사전은 BIE(BBIE, ABIE, ASBIE), CC(BCC, ACC, ASCC)로 구성된다.
- Definition : 의미사전(Semantic Dictionary)에 대한 정의를 한글로 표기하며, 국립국어연구원의 표준국어대사전을 기준으로 한다.
- 그 외 Version, Object Class Term, Property Term, Data Type/Representation Term 등을 구성하였다.

기본 의미 등록기를 운영/관리하기 위해서는 ISO 11179-3에서 메타데이터에 대한 관리를 하기 위한 속성을 운영하듯이 기본 의미 등록기도 등록기 안에 등록 저장되는 콘텐츠를 운영/관리하기 위해서는 콘텐츠에 대한 관리 속성을 두어야 한다.

4. 브리지의 구축

브릿지를 기술하기 위한 문법으로는 대상 디렉토리의 구문규칙이나 문법에 맞춰져야 한다. 예를 들어 UN/EDIFACT나 ASC X12와 같은 디렉토리는 각자의 구문규칙이나 설계규칙에 맞춰서 브릿지를 기술하여야 하며, XML로 기술되어지는 XML 문서의 경우에는 XML의 문법에 맞게 기술되어야 할 것이다. 예를 들면, UN/EDIFACT와 ASC X12의 경우에 각각의 주요 구문 요소들, 즉, 세그먼트, 세그먼트 레벨, 복합 데이터의 경우, 각각의 구성 요소들은 '+'로 구별되며, 데이터 요소나 코드값에는 콜론 ':'이 사용된다.

예) 구매주문.발급.날짜 (PurchaseOrder.
Issue. Date)로 명명된 기본의미 단위의

UN/EDIFACT 동의성에 대한 브리지는

BGM+C002:1001/220+DTM+C507:2005/137:2380:2379/101 또는 102로 표시된다.

UNTDDED 엘리먼트를 중심으로 물류 및 무역 분야에 해당하는 BBIE를 선정하여, 의미적으로 유사한 것들을 그룹핑하여 ABIE 또는 ASBIE로 정의하였고, UNTDDED에서 하나의 EntryName에 대한 정보를 TBG 17 Form 형태에 맞게 변환하였다. 한글 EntryName을 추가로 정의하였고, 이를 기존의 KIEC 브리지에 위탁화물예에 대한 정보가 존재하지 않아 브리지에 추가하였다.

5. 결론

현재 국내뿐만 아니라 국제적으로도 이러한 시맨틱 표준화가 미흡한 상태이며, 시맨틱 표준관리 체계가 구축되면, 전자상거래 또는 e-비즈니스 시스템에서 사용되는 DB간의 연동이 단일하게, 표준 체계를 통해서 통합적으로 이루어지므로 업종간, 국제간 전자상거래 시스템 상호연동 구현이 용이하게 된다. 또한 서로 다른 언어간의 원활한 의사소통을 하기 위해서는 의미를 기준으로 하여 다양한 언어의 표현방법을 매핑하여 이에 대해 각 국가에서 공인하여 사용하여야 할 것이다. 전자상거래나 e비즈니스에서 사용되는 메타 데이터는 그렇게 많지는 않다. 전 세계적으로 많이 사용하고 있는 여러 Syntax를 Semantic을 기준으로 표준화하고 이를 상호연동하기 위해서는 시맨틱(의미)기반의 표준화와 이를 자유로이 등록/저장, 검색할 수 있는 등록기가 절실히 필요하다. 국내 KEDIFACT 표준전자문서는 14개 업종 263종에 이른다. 이러한 전자문서를 대상으로 의미사전을 구축하게 되면 전자문서의 항목 상호간 의미를 기반으로 개념체계가 정립될 것으로 기대된다. 이러한 관점에서 의미사전의 구축은 향후 표준 전자문서의 개발 및 이 기종간의 전자문서 및 데이터 변환에도 활용이 가능하며, 전자문서를 비롯한 데이터의 상호운용에도 활용될 것이다. 특히 구축된 의미사전을 지속적으로 확대하기 위해서 EDI 전자문서를 기반으로 XML전자문서를 만든 경우 반드시 TAG 제출용 XML 전송항목 테이블에 EDI 매핑정보를 추가토록 요구해야 하며, 이들 전송항목은 산업별 정보로 관리해야 하며 CC로 추가될 수 있도록 해야 한다.

따라서 이렇게 수집된 XML전송항목 테이블은 산업별 의미사전으로 구축될 수 있으며 향후 EDI문서와 XML문서간의 전자문서 작성시 혹은 데이터 변환시 필요한 참조사전이 될 수 있다. 또한 XML 전자문서 개발 자동화툴이 개발 완료되면 활용할 콘텐츠로 활용될 수 있다. EDI 의미사전은 EDI전자문서와 XML전자문서간의 표준 데이터를 제공할 수 있다

참고문헌

- e-Biz 표준화백서, 산업자원부, 2004
- e-비즈니스백서, 산업자원부, 2005
- ISO 11179 (데이터요소의 명세와 표준화)
- ISO 16668 (기본의미등록기)
- BSR 콘텐츠 (KIEC, UBL 라이브러리)
- ISO 7372 (UNTDDED, 무역데이터항목사전)
- ISO/DTS 15000 (eBXML CCTS) 2.01
- OASIS UBL 1.0
- 전자문서 표준코드집 제4판, 2004
- 정용규, UNEDIFACT메시지의 EDI보안 알고리즘, 한국정보과학회, 2004
- 정용규, UNCEFACT전자문서의 XML변환규칙, 한국정보과학회, 2002