

의미사전 활용을 위한 EDI 브리지의 구축

정 용 규
을지대학교 의료산업학부
ygjung@eulji.ac.kr

Construction of EDI Bridges to Utilize Semantic Dictionary

Yong-Gyu JUNG
School of Medical Industry, Eulji University

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 KEDIFACT 표준전자문서에 대해 구성항목을 분석하고, 사용되고 있는 문서를 중심으로 의미를 기반으로 한 항목정의의 하였다. 관련된 국내·외 표준을 준용하여, 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.

1. 서론

전자상거래나 e비즈니스에서 사용되는 메타 데이터는 그렇게 많지는 않다. 전 세계적으로 많이 사용하고 있는 여러 Syntax를 Semantic 기준으로 표준화하고 이를 상호연동하기 위해서는 의미기반의 표준화와 이를 자유로이 등록/저장, 검색할 수 있는 등록기가 절실히 필요하다. 특히나 인터넷 기반의 전자상거래나 eAi, 웹서비스, 기타 부문에서 주요한 데이터 표현수단으로서 XML을 활용하고 있는데, 하나의 동일한 개념에 대해서 여러 가지 다양하게 표현하고 있으며, 비록 네임 스페이스를 이용하더라도 이는 구분을 위한 수단일 뿐, 근본적인 해결방법이 아니다. 이러한 혼란을 방지하기 위해서는 데이터의 시맨틱을 기본 키로 하여 여러 다양한 Syntax를 호환할 수 있는 레지스트리가 필요하다. 이러한 레지스트리는 전자상거래 뿐만 아니라 e비즈니스, 시맨틱 웹, 웹서비스 등에서도 선결 구비조건이다.

국내 KEDIFACT 표준전자문서는 14개 업종 263종에 이른다. 이러한 전자문서를 대상으로 의미사전을 구축하게 되면 전자문서의 항목 상호간 의미를 기반으로 개념 체계가 정립될 것으로 기대된다. 현재 국내뿐만 아니라 국제적으로도 이러한 시맨틱 표준화가 미흡한 상태이며, 시맨틱 표준관리 체계가 구축되면, 전자상거래 또는 e비즈니스 시스템에서 사용되는 데이터간의 연동이 단일하게, 표준 체계를 통해서 통합적으로 이루어지므로 업종간, 국제간 전자상거래 시스템 상호운용 구현이 용이하게 된다. 이를 해결하기 위해서 ISO TC154의 워킹그

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

본 연구에서는 BSR기반의 사업간 DB 통합에 관한 연구로 추진하여온 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 표준전자문서 263종에 대해 그 사용실태를 조사하고 사용되고 있는 문서를 중심으로 의미를 기반으로 한 항목정의 및 분석을 하였다. 대상 EDI 문서로는 14개 업종 KEDIFACT 표준전자문서 263종을 대상으로 하되 현재 사용되고 있는 문서를 중심으로 대상문서를 정하였다. 그리고 관련된 국내·외 표준을 준용하여 KEDIFACT 표준 전자문서를 대상으로 EDI 의미사전(Semantic Dictionary)을 구축하여 의미를 기반으로 한 EDI 콘텐츠 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. 이와 같은 방으로 상호운용의 기반을 마련하기 위한 EDI 의미사전을 구축하였고 브릿지를 구축하였다. 이를 위해서 우선 전자문서표준코드집을 기반으로 KEDIFACT 라이브러리 항목 도출하였고, KEC 표준으로 등록되어 있는 263종의 EDI 표준전자문서에 대해 그 사용현황을 조사하고 관련 국내·외 표준을 참고하여 현재 많이 사용되고 있는 항목을 중심으로 의미사전을 작성하였다. 또한 EDI 브릿지 구축을 위해서 KIEC 및 UBL 라이브러리, 브릿지에 대한 분석하였고, 구축된 KEDIFACT 의미사전을 KIEC 사전, UBL 사전과 비교·분석하고 동일한 의미의 항목들을 정리하였다. 또한 EDI - XML 간 매핑이 가능하도록 EDI 브릿지 작성하였다.

2. 브릿지

브릿지는 기본의미 단위가 의미적인 교환을 위해서 대상 디렉토리에서 어떻게 표현되는지를 보여주는 방법이다. 또한 브릿지는 두 개, 혹은 그 이상의 디렉토리에서 동일한 정보 단위사이의 상호 참조나 지도를 보기 위한 수단을 제공할 것이다. 특정한 기본의미 단위를 하나의 축으로 사용함으로써, 브릿지는 하나의 디렉토리 안에서, 하나의 정보 단위로부터 다른 디렉토리에 있는 동일한 정보 단위로 연결할 수 있도록 해준다. 기본의미 단위와 디렉토리 내의 관련된 정보 단위 사이의 브릿지는 양방향성을 지닌다. 즉, 기본의미 단위에서 브릿지로, 또한 브릿지에서 기본의미 단위로 변환 될 수 있다.

정보 단위는 종종 코드나 데이터 요소, 또는 일반적인 데이터 요소의 결합, 수식어가 될 수 있지만, EDIFACT 디렉토리 구문이나 XML 구문으로 통제되는 상위 구조가 될 수도 있다. 하지만, 이러한 정보 단위는 충분한 정보와 함께 브릿지를 확인해서, 가능한 간단한 방법으로 관계를 명확히 하며, 디렉토리의 구문상 구조를 완전하게 반영하도록 한다. 디렉토리의 본질 때문에, 하나 이상의 디렉토리로부터 선택된 정보 단위는 동일한 개념을 표현하는 정도에 따라 다양하게 변할 수 있다. 이런 변화는 개념의 정확한 동의성이나, 부분적이거나 부정확한 동의성, 또는 전혀 동일하지 않은 상황의 정도로 간주되기도 한다.

브릿지란 의미사전(Semantic Dictionary)와 다양한 이기종 디렉토리(예를 들어 KIEC 라이브러리, UBL 라이브러리 등)에서 그와 동일한 의미 사이를 연결한다. 브릿지는 의미사전이 목적교환을 위해 주어진 디렉토리에서 어떻게 보여주는 방법이다. 또한 브릿지는 두개, 혹은 그 이상의 디렉토리에서 동일한 정보 단위사이의

상호 참조나 지도를 보기위한 수단을 제공한다. 특정한 의미사전을 하나의 축으로 사용함으로써, 브릿지는 하나의 디렉토리 안에서, 하나의 정보 단위로부터 다른 디렉토리에 있는 동일한 정보 단위로 연결할 수 있도록 해준다. 브릿지는 두개의 서로 다른 디렉토리 사이의 직접적인 연결고리는 아니다. 의미사전과 디렉토리 내의 관련된 정보 단위 사이의 브릿지는 두 가지 방향 즉, 의미사전에서 브릿지로, 또한 브릿지에서 의미사전으로 가는 방향 모두와 관련이 있다.

- 1) 모든 브릿지는 하나의 디렉토리에서 의미사전과 개념적으로 동일성을 가져야 한다.
- 2) 모든 브릿지는 적절할 경우, 필요한 실행 협약이나 실행 지침사항을 인지할 것이다.
- 3) 각각의 브릿지는 개별적으로 구별된다.
- 4) 의미사전에서 디렉토리 엔트리까지 브릿지에 대한 다양한 선택이 가능할 때, 우선적인 방법이 적절하거나 가능할 경우, 상세하게 명시될 것이다.
- 5) 브릿지의 개발은 BSR 생산팀/BSR 운용기관과 함께 적절한 디렉토리 에이전시(브릿지 에이전시)의 책임이다.
- 6) 브릿지를 정의할 때, 의미사전에서 하나의 디렉토리 동의성에 이르기까지 각각 가능한 브릿지는 비교, 변환/이동, 새로운 디자인의 3가지로 이루어진다.

3. 브릿지 구축

3.1 개요

브릿지가 유용해지기 위해서는 오직 정확한 동의성만이 구축되어야 한다. 하지만, 일부 브릿지가 동일하지 않을 경우, 실제 관계에 있어서, 브릿지 정보와 함께 주석이 함께 제시되기도 한다. 브릿지의 개념 동의성은 브릿지에 의해 연결되는 개념의 동의성 정도에 따라 정의된다. 즉, 이는 참조 디렉토리에서 관련된 정보 단위에 의해 정의된 개념과 함께 기본의미 단위에 의해 정의된 개념의 동의성 정도를 말한다. 기본의미 단위와 디렉토리 내의 관련 정보 단위 사이의 브릿지는 다음과 같이 동일한 개념 등급 가운데 하나를 가지며, 이를 기술하여야 한다.

- 정확한 동의: 기본의미 단위와 의미 및 범위가 동일하다.
- 부정확한 동의성: 기본의미 단위처럼 동일한 일반적인 개념을 표현하고 있지만, 개념들의 의미는 정확하게

동일하지는 않다.

- 부분적인 동의성: 기본의미 단위가 참조 디렉토리에서 정확하게 동일한 개념으로 부합될 수는 없지만, 브리지는 광의나 협의의 의미와 함께 기본의미 단위를 선택함으로써 정의될 수 있다.

- 브리지 사용법: BSU에서 어떤 동일한 디렉토리까지 하나 이상의 브리지가 있는 경우, 우선적으로 선호되는 방법이 이용된다.

UN/EDIFACT와 같은 애플리케이션이 메시지의 여러 가지 다른 섹션(헤더, 세부사항, 요약)에서 동일한 속성을 실행한다는 것을 인식하는 한편, 이런 실행이 별개의 BSU를 초래하지는 않는다는 것을 알아야 한다. 각각의 실행은 별개의 브리지로 기록될 것이다. 일반적으로 이런 상황은 메시지에서 헤더나 라인 아이템 단계에서 정보를 전달할 때 적용된다. 이런 사실은 라인 아이템과의 관계에서 적절한 개념일 때에만, BSU 규격과 브리지 모두에서 기록될 것이다. 이러한 차이점이 반드시 별개의 기본 의미 단위를 중복시키지는 않을 것이다. 필요하다면, 이렇게 서로 다른 표기(표현)는 동일한 기본의미 단위에 대한 다양한 브리지 규격처럼 다루어질 것이다.

3.2 브리지 구축과정

브리지를 기술하기 위한 문법으로는 대상 디렉토리의 구문규칙이나 문법에 맞춰져야 한다. 예를 들어 UN/EDIFACT나 ASC X12와 같은 디렉토리는 각각의 구문규칙이나 설계규칙에 맞춰서 브리지를 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 브리지에서는

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

특히 물류 및 무역 분야에 해당하는 정보를 선정하여 BSR에 추가로 등록하기 위한 작업을 진행하였고, 그 결과 만들어진 의미사전을 기준으로 브리지를 구성하였다.

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

[표] 브리지 구성항목

구분	항목명	정의
Source	산업군	해당 문서(엘리먼트)가 사용되는 산업군 명시
Document Name	문서명	해당 문서의 전자문서명을 명기(영문, 국문)
Document Code	문서코드	해당문서의 전자문서 코드를 기입
Document Description	문서설명	해당 문서의 설명을 기입
Group	그룹	해당 정보(엘리먼트)가 위치한 세그먼트가 속한 그룹명을 명기
Segment	세그먼트	해당 정보(세그먼트)가 위치한 세그먼트의 이름을 명기
Tag	태그코드	해당 정보를 표현하기 위해 사용된 한정어(태그) 코드
Tag Description	태그설명	해당 정보를 표현하기 위해 사용된 한정어(태그)에 대해 그 의미를 명기(영문, 국문)
Qualifier	한정어	해당 태그에 사용된 한정어 코드를 기입
Qualifier Description	한정어 설명	해당 태그에 사용된 한정어를 설명
EntryName (H)	한글명	의미사전(Semantic Dictionary)의 요소명을 영문으로 기입
EntryName (E)	영문명	의미사전(Semantic Dictionary)의 요소명을 한글로 기입
Indicator	항목속성	항목 속성을 구분하는 것으로 BIE(BBIE, ABIE, ASBIE), CC(BCC, ACC, ASCC)로 구성된다.

3.3 브리지 구축결과

브리지는 EDI전자문서에서 의미사전과 서로 의미를 매핑할 수 있도록 위에서 설명한 원칙과 과정을 통해 만들어졌으며 결과물은 엑셀 파일로 별도로 제작되어 부록으로 첨부하였다. 의미사전 콘텐츠의 활용도를 높이기 위해 자주 사용되는 해운항만청 제출서류인 20종의 EDI전자문서를 선정하여 EDI 의미사전을 만들고 이 의미사전의 Entry Name을 매핑이 가능한 엑셀파일간의 브리지로 구성하였다. 의미사전과 브리지 2개의 엑셀파일을 만들고 이 엑셀파일간의 연결을 위해 양 파일에 Entry Name을 중복해서 아래와 같이 구성하였다.

2	Group	Segment	Tag	TagDescription	Qualifier	QualifierDescription	EntryName(H)	EntryName(E)	Indicator
1152		ATT	##	코드목록관리기관코드	6AA	외항수상부	기관명	Administration Name	BBIE
1153		TDT	##	운송편차 환경어	20	구운송	Transport Means Identifier	Transport Means Identifier	BCC
1154		DTM	##	입항당해년도 환경어	247	Year of occurrence	운항 도착 연도 일자	Shipping Arrival Year Date	BBIE
1155		DTM	##	시간 형태 환경어	602	CCYY	(시간 형태 식별자)	(DateTime Format Identifier)	(NEW)
1156		QTY	##	입항회수 환경어	6A1	입항회수	운항 도착 횟수 수량	Shipping Arrival Quantity	BBIE
1157		DOC	##	B/L 환경어	705	Bill of lading	운송 식별자	Document Identifier	BCC
1158	1	NAD	##	하역회사 환경어	UP	Unloading party	하역 당사자	Discharge Party	ASBIE
1159	1	NAD	##	하역회사 환경어	5DU	하역회사	하역 당사자	Discharge Party	ASBIE
1160	1	NAD	##	코드목록관리기관코드	6AA	외항수상부	기관명	Administration Name	BBIE
1161	2	NAD	##	국내과주 환경어	GO	국내과주	과물 국내 과주 식별자	Cargo Domestic Cargo O	BBIE
1162	2	NAD	##	사업자등록번호 혹은 주민등록	58	사업자등록번호	사업 식별자	Person Identifier	BBIE
1163	2	NAD	##	사업자등록번호 혹은 주민등록	174	주민등록번호	사업 식별자	Person Identifier	BBIE

[그림] 브리지 구축결과 (일부)

4. 결론

표준 전자문서의 개발 및 이 기종간의 전자문서 및 데이터 변환에 활용이 가능하다. 이는 현재 KIEC 라이브러리, UBL 라이브러리 두개의 브릿지를 개발, 적용하여 프레임워크 간의 전자문서 개발이 가능하다. 또한 전자문서를 비롯한 데이터의 상호운용분야로는 단일한 표준 시맨틱 DB를 구축함으로써, 한 업체 내부의 사업간 DB연동이 가능하며 업체간 N-N간의 DB연동 구현이 용이하며 시스템 연동 구현 가능하다. XML과 관련된 시장의 확대가 XML과 관련된 비즈니스 표준화가 미비하게 기인하며 이는 BSR을 이용한 DB통합 및 연동 그리고 전자데이터의 상호운용성 확대는 XML과 관련된 시장의 확대를 용이하게 할 수 있을 것으로 판단된다.

국내 KEDIFACT 표준전자문서는 14개 업종 263종에 이른다. 이러한 전자문서를 대상으로 의미사전을 구축하게 되면 전자문서의 항목 상호간 의미를 기반으로 개념체계가 정립될 것으로 기대된다. 특히 전자거래진흥원에서 개발한 KIEC라이브러리와 UBL라이브러리간의 EDIFACT문서의 상호운용성도 확보될 것이다. 사업자간 DB연동이나 전자문서 상호연동에서 1-1이나 1-N 매핑은 간단한 매핑 규칙의 설정에 의해서 구현이 가능하지만, 인터넷이라는 개방형 기반에서 N-N의 거래나 시스템 연동을 위해서는 반드시 의미에 기반한 메타데이터 표준화를 통해서 상호운용이 이루어져야 한다. 이러한 관점에서 의미사전의 구축은 향후 표준 전자문서의 개발 및 이 기종간의 전자문서 및 데이터 변환에도 활용이 가능하며, 전자문서를 비롯한 데이터의 상호운용에도 활용될 것이다.

참고 문헌

- [1] ISO 11179 (데이터요소의 명세와 표준화)
- [2] ISO 16668 (기본의미등록기)

- [3] ISO 7372 (UNTDDED, 무역데이터항목사전)
- [4] ISO/DTS 15000 (eXML CCTS) 2.01
- [5] OASIS UBL 1.0
- [6] 전자문서 표준코드집 제4판, 2004
- [7] 정용규, UNEDIFACT메시지의 EDI보안 알고리즘, 한국정보과학회, 2004
- [8] 정용규, UNCEFACT전자문서의 XML변환규칙, 한국정보과학회, 2002
- [9] EDIFACT 메시지 구현 지침, 한국전자거래진흥원, 1998
- [10] EDIFACT 메시지 설계 지침, 한국전자거래진흥원, 1998
- [11] e-Biz 표준화백서, 산업자원부, 2004
- [12] e-비즈니스백서, 산업자원부, 2005
- [13] BSR 컨텐츠 (KIEC, UBL 라이브러리)
- [14] <http://www.iso.ch>