

품목분류코드 기반의 온톨로지 모델

김동규*, 손지성**, 백두권**

*고려대학교 컴퓨터정보통신대학원 소프트웨어공학과

**고려대학교 컴퓨터전파통신공학과

e-mail : {caref, redfunky07, baikdk}@korea.ac.kr

Ontology Model based on Harmonized Commodity Description and Coding System Codes

Dong Kyoo Kim*, Jiseong Son** and, Doo-Kwon Baik**

*Dept. of Computer and Information Technology, Korea University

**Dept. of Computer and Radio Communications Engineering, Korea University

요 약

FTA(Free Trade Agreement)는 체결국들의 자유로운 물자 이동을 기반으로 무역의 활성화와 외국 자본의 직접투자를 통해 경제성장을 가져오는 중요한 협정이다. 그러나 FTA 협정의 체결조건에 따라 품목별 세율 등의 정보가 달라진다. 국제통일상품분류체계(Harmonized Commodity Description and Coding System)코드도 마찬가지로 국제 공동코드 영역인 6 자리까지 동일 하나, 7 ~ 10 자리까지는 각 국가에 의해 운영관리 되어, 국가별로 품목분류 방식이 달라진다. 또한, 품목분류코드가 갱신 될 때 마다 연관 정보를 사용자에게 제공하고 있다. 그러나 우리나라와 FTA 체결국간 세부품목분류 방식이 달라 품목분류방식의 의미에 따라 담당자가 한 건씩 비교하고 맵핑하여, 많은 시간과 인력이 소모된다. 따라서, 의미를 갖는 품목분류코드를 맵핑하면서 발생하는 문제들이 있다. 이를 해결하기 위해, 이 논문에서는 품목분류코드 기반의 온톨로지 모델을 제안한다.

1. 서론

세계각국들은 경제활동의 긍정적 효과를 주는 FTA (Free Trade Agreement)에 많은 노력을 하고 있다.[1] FTA 를 통해 관세장벽이 철폐되어 자원배분의 효율성을 높이고, 국내 총 생산과 소득증가로 거시적인 경제성장 효과, 기술전파 가속화 등의 긍정적 효과를 통해 경제성장의 기초가 된다.[2] 그러나, 무역업 종사자의 경우 FTA 체결국가가 늘어날 때마다 체결 국가별로 다른 보호품목과 그에 따른 세율이나 원산지 판정기준 정보 등을 확인하기 위해서는 우리나라의 품목분류코드와 FTA 체결국가의 품목분류코드를 찾아 협정세율, 원산지 판정기준 정보 등을 확인하여 납부할 세액 등을 파악한다.[3]

이러한 정보를 수집하고 가공한 서비스를 제공하기 위해 관세청에서는 FTA 품목분류시스템을 구축하여 무역업무 종사자 및 관련 민원인에게 정보를 제공하고 있으나, 기초자료 처리과정에서 다음과 같은 문제점을 가지고 있다.

품목분류코드는 10 자리로 구성되고, 그 중 6 자리까지 국제 표준으로 동일하게 사용하나, 총 10 자리 중 6 자리를 제외한 4 자리는 국가별로 자유롭게 정의하여 사용한다. 따라서, 각 국가의 품목분류코드는 품목을 분류하는 규칙이나 관습에 따라 구분하는 방식이 달라지고, 규격형태에 따라 분류를 서로 달리하여 우리나라 품목분류코드만 이용하여 다른 국가들의 세부

품목을 찾기 어렵다는 문제점이 있다.

또한, 세부 품목을 분류하는 방법이 국가별로 다르고, 품목별로 관세율을 적용하는 방식 역시 다르므로 수출입업자나 무역업체에서는 각 품목을 우리나라에서 사용하는 품목분류가 해당국가에서는 어떤 방식으로 적용되는지 찾아야 하는 어려움이 있다. 예를 들어, 독일의 DMB 휴대전화는 한국과 달리 휴대전화기 아닌 TV 수상기로 분류되어 휴대전화에 붙지 않는 14%의 관세를 적용 받는다. 이와 같이, 국가별로 품목분류는 국가의 문화와 환경에 많은 영향을 받는다. 그러므로 명확하지 못한 품목분류정보는 무역업 종사자나 무역업체의 의사결정시 잘못된 판단을 야기시킬 수 있다.

또한, 품목분류검색 시 품목분류코드나 품명 등, 미리 지정된 범위에서 검색기능을 제공하여 사용자가 원하는 정보가 검색 조건에 존재 하지 않을 경우 검색하기 어렵다. 또한 품목분류코드를 모르거나 정확한 품명을 모르는 경우 품목정보결과를 이용해 업무 진행 중 의사결정에 참고되는 정보로 활용이 어렵다.

따라서, 품목분류정보의 FTA 체결 국가 간 일관성 있는 정보의 연계 작업이 어렵고, FTA 체결국가의 증가에 따라 품목분류정보의 확장과 재활용이 어려운 문제들을 해결하기 위해 본 논문에서는 품목분류정보 기반의 온톨로지를 구축한다.

본 논문의 구성은 다음과 같다. 2 장에서는 온톨로지 구축을 위한 온톨로지 구축 기법의 관련연구를 기

술하고, 3 장에서는 제안모델을 기술한다. 4 장에서는 제안 모델을 구현하고, 평가내용을 기술한다. 마지막 5 장에서는 결론 및 향후 연구 내용에 대하여 기술한다.

2. 관련연구

이 장에서는 품목분류정보의 온톨로지 구현을 위한 품목분류정보와 웹 온톨로지에 관해 소개한다.

2.1 품목분류정보

우리나라는 CCCN(관세 품목 분류표)의 4 자리코드를 1987 년까지 사용하고, 1988 년부터 국제통일상품분류체계(Harmonized Commodity Description and Coding System) 인 HS 를 사용하고 있다. HS 를 채택하는 국가는 모두 동일한 HS-code 를 적용해야 한다. 단, 동협약의 6 단위코드 체계까지는 변경하지 않는 범위에서 이를 확장하여 사용하는 것을 허용하고 품목분류의 일관성을 뒷받침하기 위한 품목분류 지침서가 존재한다. 그러나 이러한 장치에도 불구하고 품목분류는 전 세계가 완벽한 일관성을 이루기가 힘들다. 국제거래 상품을 HS 에 모두 나열할 수 없는 한계와 국가마다 문화와 관습, 산업상황, 무역여건에서 생기는 상품에 대한 인식차이가 원인이다. 즉, HS 협약이 일관된 원칙과 기준을 기반으로 작성되었지만 체결국간의 HS 해석에 차이를 보임으로써 같은 상품을 두고도 서로 다른 품목분류가 되는 경우가 적지 않다.[4]

2.2 웹 온톨로지와 OWL(Web Ontology Language)

온톨로지는 “공유된 개념에 대한 정형화된 명시적인 명세”라고 Gruber 는 정의 하였다.[5] 즉, 일상생활의 정보를 일정한 형태로 정형화하여 컴퓨터가 이해하고, 구조화된 지식 정보를 말한다. 온톨로지가 일상생활의 정보를 정형화하여 표현한다면 웹 온톨로지는 웹 상의 지식을 표현하고, 공유하며, 재사용할 수 있도록 하는 온톨로지이다. [6, 7]

인터넷 기반으로 개발된 XML 은 인터넷 기술이 적용 되면서 빠르게 발전하였다.[8] 또한, 의미처리를 위한 구조 저장과 어떤 형태의 정보도 개념 모델링이 가능하여 온톨로지 언어의 기반이 되고 있다.[9]

XML 의 기본 구조가 트리 형태로 구현되어 충분한 의미관계를 기술하기 어려워 W3C 에서는 의미 네트워크 기반의 정보 모델링과 그래프 형태의 표현 구조를 갖는 RDF 를 개발 하였다. RDF 는 정보 자원의 속성과 관계를 표현하는 모델로 자원(Resources), 속성(property), 기술문(Statement)의 3 가지 요소로 개념 구조를 모델화하는 온톨로지 언어이다.[10]

RDF 는 속성과 속성값의 관계만을 정의하고 있어, 개념 구조 정의에 필수적인 속성의 기술 체계나 속성과 자원간의 관계 정의를 위한 방법도 제공되지 않는다. 이러한 문제점을 해결하고자, RDF 스키마를 만들어 보다 명확한 의미관계 표현이 가능해 졌다. 그러나, RDF 스키마 역시 요소 정의, 역 관계, 합집합 관계, 교집합 관계, 논리 관계와 카디널리티 정의 등과 같은 의미 정의가 충분하지 않아 의미 표현에 한계가

있다. 이러한 이유로 RDF 와 RDF 스키마 모델링 요소를 확장하여 DAML+OIL 이 설계 되었다.

DAML+OIL 은 RDF/S 의 모든 개념을 정의하는 기능이 있고, 자원간의 제한과 논리 체계를 서술하는 방법을 제공한다. DAML+OIL 은 기술 논리 기반의 추론 메커니즘을 가지고 있는 범용 온톨로지 언어로서 추론 시스템 개발에 많은 연구가 있었다.[11] 이러한 노력의 결과로 W3C 가 2004 년 2 월에 DAML+OIL 를 확장하고, 의미관계를 보완하기 위한 목적으로 개발된 온톨로지 언어 OWL(Web Ontology Language)[12]을 온톨로지 기술 표준언어로 발표되었다.[13]



(그림 1) 온톨로지 언어의 발전단계

3. 제안방법

이 장에서는 이 논문에서 제시하는 문제점인 품목분류정보의 FTA 체결 국가간 연계에 따른 문제점을 해결하기 위해 Ontology Development 101[14]을 이용하여 HS-Ontology 라는 구축 방법을 제안한다.

3.1 HS-Ontology

본 논문에서 제시하는 HS-Ontology 구축기법은 구조화된 정보를 온톨로지로 구현할 때 최소한의 가이드와 규칙을 제시하여 온톨로지 구축 시 최소한의 통일된 규칙을 갖도록 하여, 품목분류정보의 온톨로지 구축방법을 통일 하도록 한다.

HS-Ontology 는 Ontology Development 101 을 구축기법을 기초로 (그림 2)과 같이 설계 하였다.



(그림 2) HS-Ontology 구현 절차

HS-Ontology 는 Ontology development 101 를 기초로 설계 하였으나, Domain 을 정의하거나 기존 온톨로지 활용방안을 생략하고, 클래스 비정규화, 속성 정의 규칙을 추가하여 기존 구현기법과 차이가 있다. 따라서, 본 논문에서 클래스 비정규화와 속성 정의만 정리하여 제안하였다.

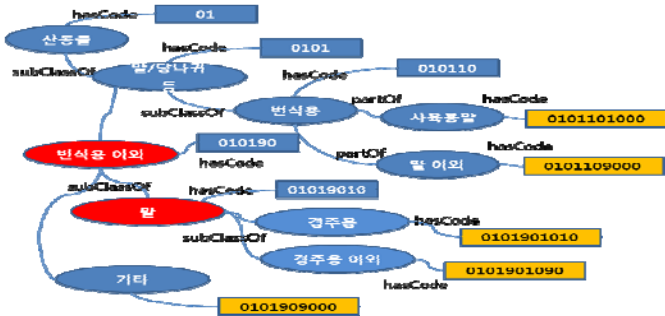
3.2 클래스 설계 및 비정규화

품목분류코드는 01 ~ 97 분류로 구분되고, 각 분류는 세부 분류에서 실사용 되는 10 자리의 완성된 코드를 갖는다. 또한 계층적 분류 형태의 품목분류코드는 추이적인(transitive) 성질 『P(X, Y) ∧ P(Y, Z) 이면, P(X, Z) 이다.』 [15]을 이용한 클래스 비정규화 순서를 <표 1>과 같은 과정을 적용하여 클래스를 간단하고 명확하게 재설계 한다.

<표 1> 품목분류 클래스 비정규화 순서

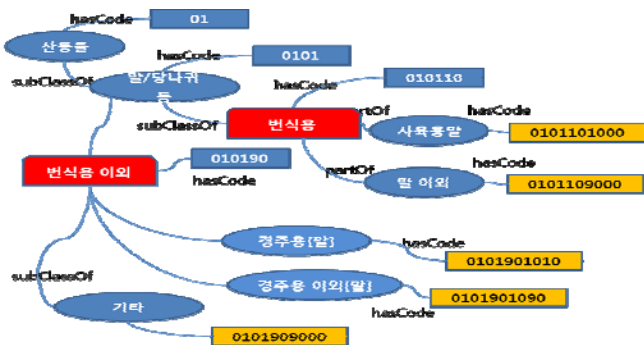
- 1) 분류 단위에 따라 기초클래스를 도출한다.
- 2) 기초 클래스가 갖는 Literal 이 최종 품목분류 코드를 갖지 않는 클래스를 클래스에서 제외 시킨다.
- 3) 이때, 제외된 클래스가 한 사이클 내에서 연속으로 여러 개가 발생할 경우 최 상위 단위의 클래스는 임시저장하고, 나머지 클래스는 삭제한다.
- 4) 이때, 삭제된 클래스 명은 하위 클래스에 상속한다.

예를 들어, 현재 운영되는 품목분류코드를 설계규칙에 따라 기초 클래스를 정의하면(그림 1)과 같이 표현할 수 있다.



(그림 3) 품목분류 기초 클래스 설계 예

이때 붉은 색으로 표시된, 비정규화 대상이 되는 『번식용 이외』, 『말』을 이용하여 클래스 비정규화를 설명하면, 두 클래스 모두 최종품목분류코드를 갖지 않으므로, 비정규화 대상이 되고, 번식용 이외와 말은 클래스 대상에서 제외 된다. 따라서, 정규화 규칙 4)와 같이 『말』은 하위클래스에 그 이름이 상속되고 『번식용 이외』는 임시 저장하게 된다. (그림 4)는 『번식용』과 『번식용 이외』는 임시보관 형태로 표현하고, 말이라는 클래스를 상속한 형태를 나타낸다.



(그림 4) 품목분류 기초 클래스 비정규화 적용 후

3.3 속성설계

온톨로지 구현에 사용되는 속성은 클래스와 클래스 관계를 정의하는 Object Property 와 클래스와 인스턴스 관계를 정의하는 Data Property 가 있다.

클래스 정규화 규칙을 적용 후<표 2>와 같이 품목분류 속성 규칙에 따라 속성을 설계한다.

<표 2> 품목분류의 속성 정의 방법

클래스정규화 과정에서, 임시 저장된 클래스는 Object Property 로 변경한다.

따라서, (그림 2)에서 나타난 『번식용』과 『번식용 이외』는 클래스에서 속성으로 변경 적용하고, 그 외 속성들은 Ontology Development 101 을 이용하여 속성을 설계한다. 단 속성 설계 시 기존의 규칙이 존재할 경우 활용하여 적용한다.

<표 3> 품목분류의 속성 정의 예

구분	속성명	Domain	Range	Inverse
Object	hasSubClass	산동물	말/당나귀..	subClassOf
Object	hasBreed	말/당나귀..	사육용 말	breedOf
Object	hasBreed	말/당나귀..	사육용 말외	breedOf
Object	hasNotBreedOf	말/당나귀..	경주말	notBreedOf

구분	속성명	도메인	타입	설명
Data	indonesiaCdOf	사육용	String	인도네시아분류코드
Data	koreaCdOf	말	String	대한민국분류코드

속성을 정리하면 <표 3>와 같이 정리가 가능하다. 품목분류 중 『산동물』을 중심으로 일부 속성을 예로 제시하였다.

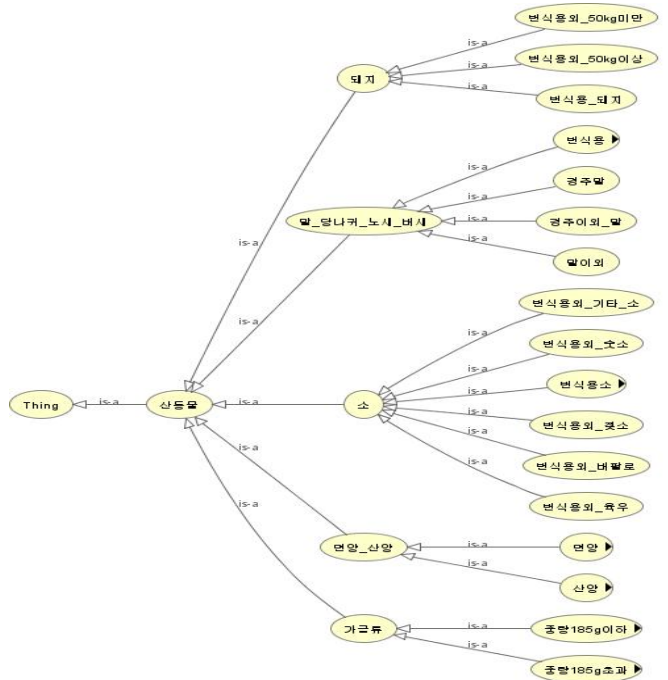
4. 구현 및 평가

이 장에서는 제안 방법을 이용하여 품목분류 온톨로지 모델을 구현하고, 평가하며, 제시된 문제 중심으로 해결 방법을 검증한다.

4.1 품목분류 온톨로지 모델 구현

앞서 제안한 구현기법을 이용하여 품목분류 온톨로지 스키마를 Protégé 4.0.2 를 사용하여 온톨로지 모델을 구축하고, OWL Viz 를 통해 (그림 5)와 같이 표현하였다.

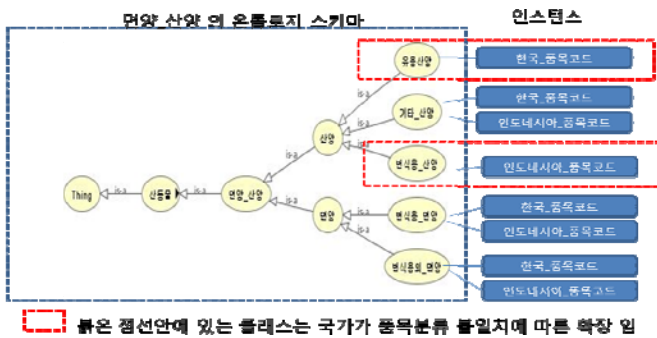
현재 구현된 스키마는 기존 구축 기법에서 명시되지 않던 클래스 간의 정규화를 통해 클래스를 보다 명확하고, 간결하게 온톨로지 모델을 구현하고, 온톨로지 모델을 통해 FTA 체결국가 간 맵핑 시 발생하는 문제와 코드 불일치를 개선하는데 그 목적이 있다.



(그림 5) 품목분류 온톨로지 모델

4.2 품목분류 온톨로지

품목분류 코드의 불일치는 국가별 문화나 환경에 의해 발생할 수 밖에 없다. 따라서 이 문제의 해결책으로 온톨로지를 제시하는 가장 큰 이유는 확장성에 있다. FTA 체결국가가 늘어날 때 마다 하나의 완성된 온톨로지 모델을 기반으로 품목분류코드를 확장 하여, (그림 6)와 같이 한국에 없는 번식용 산양을 추가하거나, 한국과 인도네시아에 있는 기타산양의 클래스가 한국에서 표현하는 코드와 인도네시아에서 표현하는 코드정보를 모두 갖는 인스턴스를 정의하여 하나의 온톨로지 모델을 통해 증가되는 FTA 체결국의 품목분류정보를 적용하고 구현이 가능하다.



(그림 6) 한-인도네시아 품목분류 모델

또한 분류방식의 불일치 문제는 한국의 DMB 휴대전화와 독일의 DMB 휴대전화를 Object Property 에서 Equivalent classes 가 되도록 관계를 맺어주어 문제를 해결하였다. 이렇게 한번 설계된 모델을 통해 다른 국가의 품목분류정보의 속성을 통해 관계를 맺어줌으로 시스템구조의 변경 없이 구현이 가능하고, 기존 시스템의 변경이 일어나지 않아 독립적이라고 할 수 있다. 예를 들어 기존 품목분류정보 맵핑 작업의 경우 DMB 전화기를 휴대전화와 TV 수상기로 맵핑을 위해 다른 분류 영역이므로 예외 발생에 따른 관계 table 을 따로 구현해야 하고 그로 인한 프로그램 수정이 불가피해짐으로 프로그램에 종속적이고, 기존 맵핑 작업의 경우 체결국가가 증가 될 때 마다 이미 매핑된 정보가 있어도 매번 처음부터 새롭게 만들어야 하는 재사용이 어려운 문제를 가지고 있으나, 온톨로지 모델을 통해 적용 할 경우 필요한 부분에 클래스를 추가하거나 인스턴스를 추가해주어 재사용이 가능한 장점을 통해 기존 품목분류코드의 맵핑방식보다 타당함을 보인다.

따라서, 온톨로지 구현 방법과 기존 맵핑 방식을 정성적으로 비교해 보면 <표 4>과 같이 나타낼 수 있다.

<표 4> 제안모델의 정성적 평가표

	맵핑 방식	온톨로지 구현
시스템 의존도	종속적	독립적
프로그램 의존도	종속적	독립적
관계에 따른 구조변경 여부	높음	낮음
재 사용성	낮음	높음
체결국 증가에 따른 확장성	낮음	높음

5. 결론 및 향후 연구

본 논문에서는 FTA 체결국가간 품목분류정보 연계의 문제점을 정의하고, 계층적 분류 형태의 정보를 온톨로지 구현하기 위한 HS-Ontology 라는 구현 기법을 제안하였다. 또한 제안 방법을 통해 품목분류 온톨로지 모델을 구현하고, 제안된 모델을 기존 시스템과의 비교 평가를 기술하였다. 또한, 온톨로지 구현을 통해 기존 품목분류코드를 가공하는 과정에서 발생하는 FTA 체결국가 간의 코드정보의 연계 및 확장 문제점들을 해결하였다.

향후 본 연구를 기초로 품목분류 온톨로지 시스템을 구현하고 일반화하는 후속연구와 추론 엔진을 적용하여 정확하고 다양한 정보 제공을 위한 연구가 필요하다.

참고문헌

- [1] 관세청 FTA 포탈 <http://fta.customs.go.kr>
- [2] 박순찬 외 “FTA 의 득과 실: 사례분석 중심으로”, 서울, 대외경제정책연구원, 2004
- [3] 김재식, “품목분류 왜곡을 통한 농산물의 관세회피 사례연구”, 관세학회, 2008.8 pp 130
- [4] 김창길 외 “HS 품목분류표와 한국분류지침 운용에 관한 연구”, 한국관세학회, 2006. 5 pp113~129
- [5] Gruber, T., “A Translation approach to portable ontologies”, Knowledge Acquisition, Vol. 5, No. 2, pp.199-220, 1993.
- [6] 옥철영. “한국어정보처리와 온톨로지”, 한국어정보처리연구회, 동계 튜토리얼 자료집. 2004.
- [7] 이현실. “온톨로지 기반 한의학 처방 지식관리시스템 설계에 관한 연구”, 중앙대 대학원, p183, 2003
- [8] Newton, Katy. “Putting the ICLD on the Semantic Web”. 2002
- [9] 한성국, 정영식, 주수종, 이현창. “XML 워크샵”, 2004.
- [10] Lannella, Renato. “An Idiot’s guide to the Resource Description Framework” The New Review of Information Networking, vol.4, 1999.
- [11] Gomez-Perez, A. State of Art in Ontologies from the Semantic Web perspective. Esperanto Services IST-2001-34373, 2002
- [12] <http://www.w3.org/2004/OWL/>
- [13] 이현실, 한성국 “OWL 을 이용한 온톨로지 기반의 목록시스템 설계 연구”, 정보관리학회, 2004,6 pp249~267.
- [14] Noy, N, F., McGuinness, D.L. “Ontology Development 101: a guide to crating your first ontology”, Stanford Knowledge Systems Labo
- [15] Kanzaki Masahide, “시맨틱 웹을 위한 RDF/OWL 입문”, 2008, pp.121~122