

제품 데이터의 전자거래 활용을 위한 웹 기반 지식관리

박상우*, 윤홍규*, 유상봉*, 김철환**

Web-based Knowledge Management for Using Product Data in E-Commerce

S.W. Park, H.K. Yoon, S.B. Yoo, C.W. Kim

Abstract

As the networks (i.e., intranet and internet) proliferate all over the world, it is inevitable to move some (or all) of the enterprise activities into virtual spaces. Differently from business data, product data have complex semantics and thus are not properly exchanged among different application programs. Even though some neutral formats of product data have been developed by standard organizations, translating them among various application programs still needs the comprehensive understanding of the complex semantics. Recently, it is widely recognized that capturing more knowledge is the next step to overcome the current difficulties on sharing product data. In this paper, we present Web-based knowledge management that facilitates seamless sharing of product data among various application programs in virtual enterprises. Three types of knowledge are managed by the knowledge management system - metadata, ontology, and mapping. In this environment, we consider both business applications (e.g., ERP, SCM, and EDI System) and engineering applications (e.g., CAD and CAM system).

Key Words : Knowledge Management, E-Commerce, Product Data, Web, Virtual Enterprise.

* 인하대학교 자동화공학과

** 국방대학교 무기체계과

1. 서론

인터넷과 인트라넷 등의 네트워크가 전 세계적으로 활성화됨에 따라, 기업의 활동들 중 일부 또는 전부의 가상공간으로의 이동이 필수적으로 되고 있다. 제품들의 전 생명주기에 걸친 정보를 제공하기 위해 가상 기업들은 크게 두 종류의 데이터 (비즈니스 데이터, 제품 데이터)를 관리해야 한다. 대부분의 현대 기업들은 비즈니스 데이터를 다루는데 있어서 충분한 경험을 가지고 있으나, 제품 데이터의 관리 측면에 있어서는 그렇지 못하다. CAD/CAM 데이터를 위시한 제품 데이터는 설계와 여러 생산 프로세스들에 의해서만 관리되어 왔으나, 가상 기업들에 있어서는 제품 데이터가 제품의 전 생명 주기에 걸쳐서, 각 프로세스의 다음 단계에 이용되도록 해야 할 필요가 있다 (예, 웹 기반의 제품 목록, 제품 설명서) [9].

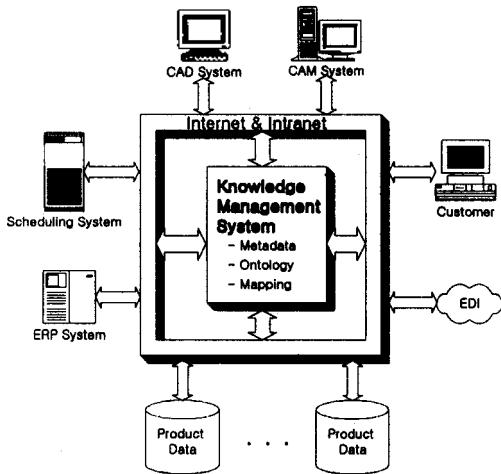
비즈니스 데이터와는 달리, 제품 데이터는 복잡한 의미상의 연결 구조를 가지며 서로 다른 응용 프로그램들 사이에 제대로 교환되지 않는다 [3, 16]. 비록 몇몇 제품 데이터의 중립 표준들이 표준화 기구들에 의해서 개발되어 왔으나, 다양한 응용 프로그램들간에 그들을 번역하는 것은 여러 데이터가 갖는 의미의 연결에 대한 종합적인 이해가 요구된다. 최근에는 제품 데이터의 공유상의 현 문제점들을 극복하기 위한 방안으로서 더욱 다양한 지식을 확보하는 방법론이 널리 인식되고 있다 [14]. 본 논문에서는, 가상 기업들이 필요로 하는 다양한 응용 프로그램들간에 제품 데이터의 원활한 공유를 촉진시킬 수 있기 위한 웹 기반의 지식 관

리를 제안한다.

기존의 많은 지식 기반 시스템들은 if-then 규칙들과 추론 엔진들에 기초하여 운영된다 [19]. 지식 기반 시스템들의 추론 기능은 제품 구성이나 기업 내의 자원 분배 등의 간단한 문제의 영역에는 잘 적용될 수 있으나, 동적인 스케줄링과 진단등과 같은 실질적인 대부분의 문제들에 대해서는 성공적으로 적용되기 힘들다. 최근 네트워크를 통해서 웹이 다량의 이질적 정보에 대한 표준 인터페이스를 제공할 수 있게 됨에 따라, 사용자들은 이러한 데이터의 활용에 여러 측면에서의 도움이 필요하게 되었다. 그 중, 가장 주목할 만한 컴퓨팅 환경의 변화는 여러 가지 종류의 지식 표현들에 대한 능동적 처리를 위해 온톨로지(ontology), 메타데이터(metadata), 에이전트(agent)같은 기술적 접근 방법들이 제시되었다는 점이다 [5, 15, 17].

본 논문은 이종의 응용 시스템들 사이에서 제품 데이터의 공유를 위해 필수적인, 다양하게 표현된 지식의 관리 측면에 중점을 두기로 한다. <그림 1>에 나타난 가상 기업의 구성은 인터넷이나 인트라넷을 통해 분산 환경에서의 제품 데이터가 여러 응용들에 의해서 공유되는 모습을 표현하고 있으며, 세 가지의 지식 관리 방법에 의해서 정보들이 관리된다. 이러한 환경에서, 우리는 비즈니스 응용(ERP, 스케줄링, EDI 시스템 등)과 공학 응용(CAD/CAM 시스템) 모두를 고려할 수 있고, 사용자들이나 고객들은 웹 브라우저를 통해서 지식 관리 시스템에 직접적으로 접근할 수도 있다.

가상 기업 환경에서, 제품 데이터는 웹 서버를 포함하는 중앙 집중식 데이터베이스



<그림 1> 전자 거래 환경의 구성도

또는 분산 데이터베이스들에 보관될 수 있는데, 대부분의 상용 PDM (Product Data Mangement) 시스템들은 Oracle 이나 DB2 같은 DBMS (database management system) 들에 기반을 두고 있다. 그리고, 이러한 DBMS 들은 데이터의 무결성 관리를 위해서 동시성 제어, 크래쉬 복구, 질의 처리 등과 같은 기능들을 제공한다. 그렇지만, 중앙 집중식 데이터베이스는 트랜잭션의 특정 순간에 네트워크 트래픽의 오버헤드 증가로 인해 병목 현상을 야기할 수 있다. 더 심각한 경우는 대부분의 설계자들이 자신의 워크스테이션 상에서 작업을 하다가 일정 시간 간격, 즉 주기적으로 중앙 데이터베이스를 업데이트 하게 되는데, 이런 경우 중앙 데이터베이스에서 관리되는 데이터가 올바르게 갱신되지 못할 수도 있다. 분산 데이터베이스 시스템의 경우는 방금 위에서 언급한 몇몇 문제를 해결하기는 하나, 분산 데이터베이스들에 저장된 제품 데이터는 어떠한 중앙 제어도 받지 않게 된다는 문제가 있다. 또한, 불필

요한 검색 프로세스들의 오버헤드로 인해 사용자와 네트워크에 부담을 주게 된다. 본 논문에서 표현되는 지식들은 제품 데이터의 저장 형태가 중앙 집중식인지 분산 데이터베이스인지에 무관하게, 사용자들이 필요로 하는 제품 데이터를 효과적으로 검색하고 사용할 수 있도록 도와준다.

본 연구에서는 표준 제품 모델로서 STEP [10]을 사용하는데, 이는 우리에게 다음과 같은 몇 가지 이점을 준다. 첫째로, STEP 은 자동차, 선박, 기계설계, 공정계획, 유한 요소 해석등과 같은 다양한 산업 분야에서 공유될 수 있는 개방형 정보 모델을 정의한다는 점이다. 상용 CAD 공급자들은 주로 자신의 시스템에만 최적화된 데이터 포맷을 사용하기 때문에, 서로 다른 상용 시스템 (예, AutoCAD, Catia)을 사용하는 사용자는 그들의 제품 데이터를 상호 교환할 수 없다. 둘째로, STEP 은 강력한 정보 모델 언어인 EXPRESS [11]를 제공하는데, 이 언어는 복잡한 공학 객체들의 모델링에 중점을 두고 STEP 을 위해서 개발되었다. 셋째로, STEP 은 SDAI (standard data access interface)라는 응용 프로그램 인터페이스를 제공하는데, 이것은 물리적 파일, 워킹 폼, 그리고 데이터베이스 관리 시스템과 같이 서로 다른 형식으로 저장된 STEP 데이터에 대한 접근을 가능케 해준다. SDAI 를 사용하면, 임의의 CAD 시스템이 특정 DBMS 의 관련 함수를 모르더라도 그 데이터베이스에 STEP 데이터를 저장하거나 인출할 수 있다. 결과적으로, 임의의 파일 시스템이나 데이터베이스가 저장소 시스템으로서 쓰일 수 있다. 이 인터페이스는 제조 산업들이 급속

한 하드웨어와 소프트웨어 기술들의 발전에 대응하여 CAD/CAM 시스템에 투자하는 것을 촉진시킬 수 있다.

본 논문의 구성은 다음과 같다: 섹션 2는 관련 연구를 기술하고, 섹션 3은 본 논문에서 다룬 지식의 세가지 종류를 기술한다. 섹션 4에서는 지식의 각 종류가 제품 데이터를 공유함에 있어서 어떻게 적용되는지를 다루며, 마지막으로 섹션 5는 결론과 향후 과제를 기술한다.

2. 관련 연구

기존의 지식 기반 시스템에서 지식은 추론 엔진에 의해서 사고되고 문제 해결에 적용될 수 있는, if-then 규칙들의 형태로 표현되었다. 전문가 시스템들과 같은 대부분의 이러한 지식 기반 시스템들은 시스템의 구조에 대한 불충분한 이해, 지식의 획득 등에 문제를 가지고 있으며, 한정된 분야의 완전한 솔루션에 초점을 맞춘다 [19]. 그렇지만, 인터넷이 전세계의 연산 자원들을 서로 연결함에 따라서, 다양한 지식 획득에 대한 요구는 급속하게 증가하고 있다. 이러한 요구들이 다양한 응용 분야에서의 발생된 지식들에 대한 새로운 인코딩 방법들과 관련된 능동적 리서치와 그 응용들을 낳게 되었다.

(a) 메타데이터

다양한 메타데이터 표준들이 응용 도메인에서의 자원을 기술하기 위해 개발되어 왔다. GILS (Government Information Locator Service), FGDC (Federal Data Geographic Committee), MARC (Machine Readable Card),

DC (Dublin Core), CIMI (Consortium for the interchange of Museum Information) 등이 그 예이다. 이러한 표준들은 각기 다른 의미와 문법을 사용하기 때문에 상호간의 운용성이 부족하다. RDF (Resource Description Framework)은 W3C에서 제안한 표준으로서, 다양한 응용의 영역들에 존재하는 메타데이터를 표현하는 중립 메커니즘을 제공한다 [13]. XML의 문법을 이용하므로, RDF는 웹상에서 다양한 메타데이터를 표현할 수 있다.

(b) 온톨로지

온톨로지는 어떤 도메인을 기술하는 용어들의 구조화된 집합을 제공하는데, 이들은 지식의 재사용을 촉진시키고 시스템 개발자들에게 상위 수준의 플랫폼을 제공한다. 스탠포드의 온톨로지 라이브러리는 모듈화된 기초 온톨로지들을 보유하고 있다. NIST는 그들 자신의 형식적이고 명확한 정의들을 통해 제조 관련 개념들과 용어들의 온톨로지를 개발했다 [18]. 인터넷 검색에 대한 응용으로서, OntoSeek는 직업별관과 제품 목록에 대한 언어 온톨로지를 연관시킨 구조화된 내용의 표현들을 이용하여, 내용 기반 검색의 정확도와 회수율을 증가시킨 바 있다 [8].

(c) 데이터 매핑

제품 데이터는 다양한 형식으로 정의되어 왔기 때문에, 많은 독점적인 또는 대중적인 매핑 프로그램들이 개발되어왔으며, 최근에는 다른 형식으로 저장된 제품 데이터를 서로 변환하는 방법들을 개발하는데 많은 노력이 기울여지고 있다. Product Data Markup Language (PDML)은 매핑 사양을 통

해 어휘들이 관련된 분야에서 JEDMICS 같은 행정 시스템이나 PDM 같은 상용 시스템들 간에 제품 정보의 교환을 지원하기 위해 설계된 XML (Extensible Markup Language) 이다 [2]. PBDH 프로젝트는 데이터 정의들과 모델들을 통해서 STEP 과 EDIFACT 사이를 연결을 시도하였고, EXPRESS-X 는 EXPRESS 언어에 의한 스키마에 기반을 둔 데이터 사이의 관계가 규정되도록 해 주는 구조화된 데이터 매핑 언어이다 [7].

(d) 에이전트 기반의 지식 공유

에이전트 기반의 지식 공유에 관한 연구들은 소프트웨어 에이전트들간의 협동 작업을 가능하게 하는데, 일반적으로 ACL (Agent Communication Language)과 같은 전처리 언어가 에이전트들간의 통신과 협력을 위해서 사용된다. SHARE 프로젝트에선 KQML (Knowledge Query Manipulation Language)가 설계 팀들에게 설계 프로세스의 정보를 공유를 지원하도록 사용되기도 하였다. Case 는 설계자들간에 통신과 협력을 위해서 VWL (Virtual Workspace Language) 을 사용한 바 있다.

위에 언급된 것들이 지식 표현의 전부는 아니다. 예로서, 모델 기반의 지식 획득과 기타 문제 해결 방법론들은 포함되지 않는다. 본 논문에서는 다양한 지식의 형태 중에 제품 데이터의 공유를 위한 메타데이터, 온톨로지 그리고 매핑 관계를 제시한다. 본 논문의 또 다른 주요 특성은 모든 인터페이스들이 웹에서 사용 가능하도록 설계된다는 점이다. 사용자 인터페이스들과 다른 모듈들은 웹 문서들로부터 다운로드될 수 있는 애플릿으로 구성했고, 메타데이터는 RDF

(Resource Description Framework)로 모델링하여 XML 로 인코드되도록 하였다. 이 W3C 표준들의 사용은 앞서 제안한 접근 방법들이 향후에도 다양한 관련 응용에 적용될 수 있도록 하기 위함이다.

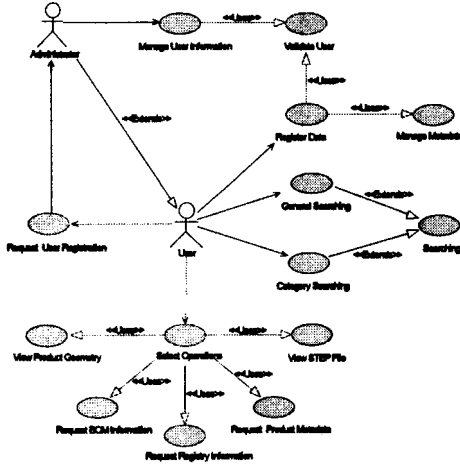
3. 제품 데이터 공유를 위한 지식

가상 기업 상에서 다양한 응용 프로그램 간에 제품 데이터의 공유를 활성화시키기 위해서는 제품데이터에 대한 지식이 적절히 발견되고 관리되어야만 한다. 지식은 지식 베이스에 저장되고 관리된다. 본 논문에 제시되는 지식 베이스는 메타데이터, 온톨로지, 그리고 데이터 매핑의 세가지 지식 타입들을 포함한다. 이번 섹션에서는 예제와 더불어 앞서 언급한 지식들을 소개한다.

3.1 제품 데이터의 지도 역할을 수행하는 메타데이터

메타데이터는 '데이터의 데이터'를 의미한다. 오늘날의 가상 기업에서 일반적으로 설계는 전세계에 걸쳐 분산된 다수의 엔지니어들에 의해 수행되는데, 이러한 환경에서는 특정 설계 정보를 적절히 찾아내는 것이 쉽지 않다. 메타데이터는 사용자들이 관심을 갖는 정보가 검색되기 용이하도록 제품 데이터 자체에 관한 정보를 제공한다. 메타데이터에 포함되는 정보의 예로서 제품의 이름, 설계자의 이름, 제품의 특성, 데이터의 위치, 데이터의 포맷 등이 될 수 있다.

메타데이터는 지식 베이스의 핵심 부분이긴 하나, 특정 응용에 종속적인 경우가



<그림 2> Use Case 다이어그램

대부분이므로 메타데이터의 내용에 관한 표준은 없는 상태이다. 지식 베이스를 구축하는 첫 단계는 메타데이터의 내용을 결정하는 것이다.

<그림 2>는 유도된 use case 다이어그램으로서, 지식 베이스는 두 종류의 사용자를 갖는다. 먼저 관리자는 등록되어 있는 사용자를 등록 해제 시키거나, 새로운 사용자를 등록하는 역할을 한다. 일단 사용자가 등록되면, 그는 다른 오퍼레이션들을 수행하기 위해 지식 베이스와 상호 작용을 하게 된다. 설계 정보의 등록, 메타 데이터의 질의를 통한 관련 데이터 검색, 선택된 항목에 대한 연산의 수행 등이 이러한 상호 작용에 포함되며, 메타 데이터를 이용한 전역적인 검색 기능은 선택된 카테고리에 대해서만 가능하도록 하였다. 각 use case 의 구체적인 기능 사항은 다음과 같다 :

- (a) **Manage user Information** : 본 use case 는 사용자들로부터의 요구에 기초하여 관리자에 의해서 수행된다.
- (b) **Register Data** : 본 use case 는 STEP 파일들이나 문서들을 등록한다. 사용자는 새로운 설계를 등록할 때, 그에 관련된 부가적인 정보들 (사용자 식별자, 이름, 전자 우편 주소, 설계 데이터에 대한 묘사, 데이터가 위치한 URL 등)도 입력을 하게 된다.
- (c) **General Search** : 사용자들은 지식 베이스에 저장된 모든 메타데이터에 대해 검색을 수행할 수 있다. 이때, 사용자가 검색하기 원하는 제품의 이름, 설계자 이름, 도면 이름, 회사 이름, 설계 데이터에 대한 묘사, 부품 이름 같은 정보가 검색을 위한 키워드로서 사용될 수 있다.
- (d) **Category Search** : 검색의 효율성을 향상시키기 위해서, 사용자들은 키워드의 특정 카테고리에 대한 검색을 할 수 있는데, 이 경우 선택될 수 있는 범주는 설계, 레지스트리, 제품, 문서, 사용자, 승인 중의 하나가 된다.
- (e) **Select Operations** : 사용자가 특정 항목을 선택하면, 오퍼레이션이 수행되고 그 결과가 항목에 적용되게 되며, “View Product Geometry”, “Request BOM Information”, “Request Registry Information”, “Request Product Metadata”, “View STEP file” 등이 선택 가능한 오퍼레이션들에 포함된다.

이러한 use case 다이어그램으로부터 우리는 지식 베이스에 의해 관리되어야 할 때

타데이터에 대한 필요성을 갖게 된다. 본 논문에서는 제품데이터는 STEP 응용프로토콜(Application Protocol 203:configuration - controlled design)으로 표현되어 있다고 가정하고, 메타데이터를 다음 여섯가지의 카테고리들로 분류하였다.

- (a) Design : STEP 물리적 파일들에 관련된 정보로서, 파일의 헤더 정보들과 사용자가 파일들을 등록할 때에 제공되는 정보들로서 구성된다.
- (b) Registry : 등록자에 관련된 정보로서 식별자, 이름, 전자 우편 주소, 등록 날짜를 포함한다.
- (c) Part : STEP 파일에 포함된 파트들의 정보로서 STEP 은 어셈블리 모델을 정의할 수 있기 때문에, 하나의 STEP 파일에는

여러 파트들이 포함될 수 있다.

- (d) Document : 식별자, 이름, 기타 사항, 형식 등과 같은 문서의 정보..
- (e) Person : 식별자, 이름, 역할, 고용주 등과 같은 인사 정보.
- (f) Approval : 승인자, 날짜, 상태 등과 같이 파트들의 승인에 관련된 정보.

<표 1>은 메타데이터의 전 목록이며, 카테고리 이름은 "Category Search"에서 사용된다. <표 1>에 정의된 메타데이터는 RDF (Resource Description Framework)에 의해 모델되었고, STEP 파일들로부터 추출된 것들이다. 메타데이터의 RDF model 은 지면 제약상 본 논문에서 거론하지 않는다.

<표 1> 메타데이터의 카테고리들과 요소들

카테고리	요소 이름	설 명	카테고리	요소 이름	설 명
Design	fileName	파일의 이름	Part	PartID	부품의 ID
	fileDesc	파일의 정보		PartName	부품의 이름
	preprocessor	파일을 생성하는 처리기		PartDesc	부품에 대한 정보
	schemaName	스키마의 이름		Level	부품 구성상의 수준
	url	URL		Quantity	수량
	fileSize	파일의 크기		AssociatedDoc	관련 문서
Registry	registrariID	등록자의 ID		AssociatedPerson	관련 작업자
	registrarName	등록자의 이름		ContainedBy	본 부품을 포함하는 부품
	registrarEmail	전자 우편 주소		Contains	본 부품이 포함하는 부품
	registryDate	등록일		ApprovalInfo	승인 정보
Document	docID	문서의 ID	Approval	ApprovalStatus	부품에 대한 승인 상태
	docName	문서의 이름		ApprovedBy	승인자
	docDesc	문서의 정보		ApprovalType	승인의 형태
	docType	문서의 타입		ApprovalDate	승인 날짜
Person	personID	작업자의 ID			
	personName	작업자의 이름			
	employer	고용주			
	personRole	역할			

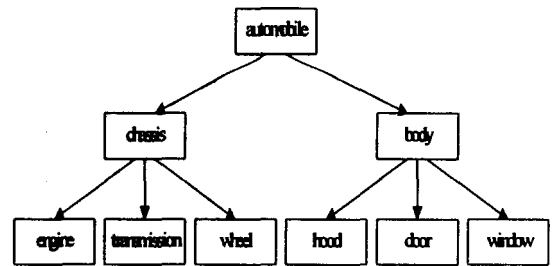
3.2 제품 데이터의 의미론을 보완하는 온톨로지

정보 기술 관점에서, 온톨로지는 특정 영역에서의 지식의 상호작용과 엔티티들의 작업 모델이다. 넓은 의미에서, 이전 서브섹션에서 표현된 메타 데이터는 온톨로지에 속한다고 볼 수 있다. 우리는 온톨로지를 제품 데이터에 포함된 어휘의 의미와 관계로서 사용한다. 제품들은 설계자가 직면해 있을 수 있는 다양한 환경에서 설계되기 때문에, 같은 개념 또는 같은 용어로 사용될 수 있는 서로 다른 용어들은 또한 다른 개념들로서 존재할 수 있다. 예로서, 'resource'는 워크 플로우 시스템과 공정 계획 시스템에서 조금씩 다른 개념을 나타낸다 [18]. 워크 플로우 시스템에서, 자원은 필요한 결정들을 내리는데 사용되는 정보로서 간주될 수 있고, 공정 계획 시스템에서는 주어진 작업을 수행하는 사람 또는 로봇으로 볼 수 있다. 엄밀하게 제품 데이터에 사용된 어휘들을 이해하기 위해서 우리는 그것들이 사용되는 컨텍스트(환경)를 고려해야만 한다.

표준화된 제품 모델에서 정의된 키워드들 이외에, 엔지니어들은 같은 의미를 표현하는 여러 가지 동의어를 선택할 수 있다. 간단한 동의어 목록의 예로서, [car, auto, automobile, motocar]를 들 수 있다. 일반적으로 제품 설계는 제품에 관한 간략한 묘사들을 지니게 되는데, 이러한 묘사들에는 이름, 용도, 사양, 제품의 특성 등이 포함될 수 있다. 묘사는 자연어로 기술되기 때문에, 동의어 목록에 있는 어휘들 중 임의의 것이 선택될 수 있다. 내용검색이나 연관 검색시에,

우리는 이러한 동의어 목록들을 사용함으로써 검색 결과를 탁월하게 향상시킬 수 있다.

어휘간의 또 다른 일반적인 관계 중의 하나는 part-of 관계이다. 엔티티 A가 다른 엔티티 B에 포함되는 경우, 엔티티 A와 B는 part-of 관계를 갖게 된다. <그림 3>은 이러한 관계의 예를 보여준다. 자동차는 새시(차대)와 몸체로 이루어지고, 새시는 엔진, 바퀴로 구성되며, 몸체는 지붕, 문, 창문을 요소로 한다. 또한 part-of 관계는 검색 결과를 향상시키기 위한 방편으로 사용될 수 있다. 예로서, 사용자가 전동 장치의 제품 데이터를 요구할 때, 검색 엔진은 새시와 자동차에 관련된 제품 정보를 반환하도록 할 수 있다.



<그림 3> part-of 관계의 예

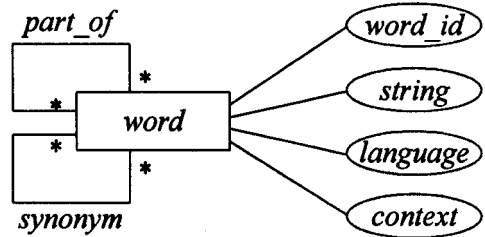
서로 다른 나라에서 작업하는 엔지니어들도 다른 언어들을 사용할 수 있다. 예로서, "Korea"라는 나라는 한국어로는 "한국"으로 쓰여지고, 중국어로는 "韓國"으로 쓰여진다. 사람 이름이나 지리명 같은 고유어는 원어로 쓰여질 수 있기 때문에 이에 대한 해석 관계는 다른 언어를 쓰는 엔지니어들이 외국으로부터 전송된 제품 정보들을 이해하고 검색할 수 있도록 하는데 도움을 줄 수 있다. 최근, Windows NT와 STEP과

같은 정보 시스템들은 다중-바이트 문자 코드를 인코드할 수 있도록 유니코드를 사용한다.

앞서 온톨로지의 필요성을 컨텍스트, 동의어, part-of 관계, 번역 또는 해석 관계로서 요약해 보았는데, 이러한 각 필요성은 온톨로지에 다음과 같이 포함될 수 있다:

- (a) 컨텍스트 : 이것은 그 단어가 사용되는 응용 도메인을 표현한다. 응용 도메인은 STEP 에 정의된 응용 프로토콜 사이에서 선택된다. 응용 프로토콜들의 몇가지 예로서, *associative drafting*, *configuration-controlled design*, *ship structure*, *plant spatial configuration* 등이 있다.
- (b) 동의어 : 이것은 여러 개의 단어 개체들 사이의 동의어를 표현한다. 동의어 관계가 정의되는 경우, 그 단어들의 해당 컨텍스트가 고려되어야 한다.
- (c) Part-of 관계 : 이것은 단어 개체들간의 방향성을 부여하는 관계이다.
- (d) 해석 관계 : 이것은 다른 언어 사이에서 존재하는 동의어 관계로 간주될 수 있다. 단어 개체는 그 속성으로서 언어의 이름을 갖는다.

온톨로지를 위한 데이터 모델은 <그림 4>와 같다. 엔티티 *word* 는 두 개의 다-대-다 관계 (i.e., *part-of* 와 *synonym*)와 네 개의 속성 (i.e., *word-id*, *language*, *spelling*, *context*)을 갖는다. 데이터 모델은 관계형 데이터베이스 시스템이나 객체 지향형 데이터 베이스 시스템을 사용하여 구현될 수 있다.

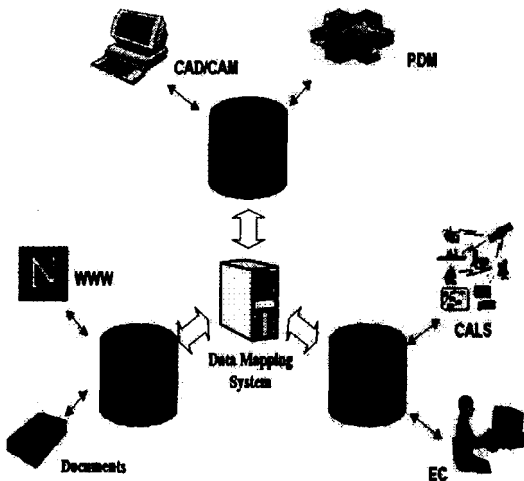


<그림 4> 온톨로지의 데이터 모델

3.3 서로 다른 포맷들간의 매핑 관계

오늘날 정보 기술은 대부분의 기업 활동들 (예, 마케팅, 제품 설계, 현장 프로세스, 품질 관리, 재고, 판매, 자원 계획, 서비스 등)에 적용되고 있으며, 그들은 네트워크를 통해 서로 연결되어 있다. 이런 상황에서, 서로 다른 응용 시스템들간의 데이터 교환에 대한 필요성은 기업에서 최고 우선 순위 중의 하나이다. 자동화의 초기 국면에서부터, 데이터 교환을 위한 많은 노력들이 기업 (예, General Electric, Boeing) 독자적으로 또는 표준 기구들 (예, ANSI, ISO)에 의해서 시작되어 왔다 [6]. 정보 기술의 방대한 사양들 중에서, <그림 5>에 나타나 있는 세계의 표준들이 가상 기업 환경에 중요한 역할을 담당한다. 제품 데이터를 위한 STEP, 문서에 대해서는 SGML/XML, EC (Electronic commerce)에 대해서는 EDI 가 바로 그것들이다.

기업 활동들에 포함되는 데이터는 비즈니스 데이터와 공학 데이터의 혼합으로서 생각할 수 있다. 비즈니스 데이터는 정수들이나 문자들이며 관리 활동에 관계된다. 공학 데이터는 형상, 위상, 제조 정보나 서비스 매뉴얼 같이 제품들의 사양을 정의한다.



<그림 5> 가상 기업 환경에서의 세가지 주요 데이터 교환

일반적으로 공학 데이터는 실수를 포함하는 복잡한 데이터 타입들로써 존재하며, 공학 활동들에 관계된다. 최근에, 기업 전체가 네트워크를 통해 통합됨에 따라 비즈니스 활동들과 공학 활동들사이의 데이터 교환은 일반적인 관계가 되고 있다. 다음은 서로 다른 응용 프로그램들간의 데이터 교환의 몇가지 예들이다.

- (a) 서로 다른 CAD 포맷들 : IGES, AutoCAD, STEP, ACIS, Parasolid
- (b) 제품 데이터와 웹 문서들 : STEP, SGML, XML
- (c) 제품 데이터와 ERP (Enterprise Resource Planning) : STEP, SAP
- (d) 제품 데이터와 EDI (Electronic Data Interchange) : STEP, EDIFACT
- (e) 제품 데이터와 데이터베이스 시스템 : STEP, Oracle, Objectstore

위에 열거된 목록은 기업 전체의 활동들에 포함되는 데이터 교환 방법들의 극히 일부이다. 많은 회사들은 자사의 특정 필요성을 위해 고유의 솔루션을 개발해 왔으며, 서로 다른 포맷들 간의 번역 관계는 회사의 자산을 높이는 데 유용하며, 일반적으로 프로그램이나 문서 또는 전문가의 지식에 포함되어 있게 된다. 실제로 대부분의 경우에 이러한 관계는 C나 Fortran 같은 프로그래밍 언어에 의해서 구현되는데, 실행 프로그램들에 포함되어 있는 지식은 이해하기 힘들고 재사용이 어렵다. 고수준에서 이러한 관계를 표현하는 한가지 방법은 스키마 관계이다.

스키마 관계는 스키마 수준에서 표현되기 때문에, 데이터 수준에서의 관계보다 이해와 관리가 용이하다.

다음 서브섹션에서는 제품 데이터와 EDI 간의 스키마 관계를 소개하기로 한다.

EDIFACT(EDI for Administration Commerce and Transport)는 미국에 의해서 개발되어 온, 주요 국제 EDI 표준 중의 하나이다. EDIFACT 로 정의된 다양한 메시지들 중에서 PRODAT, CONDRA, CONDR0 메시지들은 STEP 으로 정의된 제품 데이터와 관련이 있다. GOSET 은 STEP 과 EDIFACT 간의 관계를 조사해 왔고, 몇몇 STEP 엔티티들이 EDIFACT 요소들 (<표 2> 참조)로 매핑될 수 있다는 것을 알아냈다 [1]. 현재 이 매핑은 STEP 과 EDIFACT 에 의해 다뤄지는 모든 정보 도메인의 아주 일부분이다. 공통 비즈니스와 공학 개념들을 ISO Central Register BSR (Basic Semantic Registry)로 정의

<표 2> STEP 엔티티들과 EDIFACT 요소들과의 매핑 관계의 예

product.id	LIN+++C212:7140/<value>
product.name	IMD+7077/F++C273:::7008<value>
product.description	IMD++C273:::7008/<value>
product.category.name	PGI+5379/11+C288:::5388/<value>
product.category.description	PGI+5379/11+C288:::5388/<value>
product concept.id	RFF+C506:1154/<value>:1153/*
product concept.description	FTX
document.type	DOC+C002:1001/<relevant code>
document.type.product data type	BGM+C002:::1000<value>
document.id	BGM+++C106:1004/<value>
document.name	FTX+4451(ABX=Document name and documentary requirements.)++C108:4440

하기 위한 많은 노력들이 수행되고 있다. 스키마 매핑 지식을 이용한 STEP 과 EDIFACT 간의 데이터 번역 시스템의 구현은 다음 섹션에서 설명된다.

4. 웹 기반의 지식 인터페이스

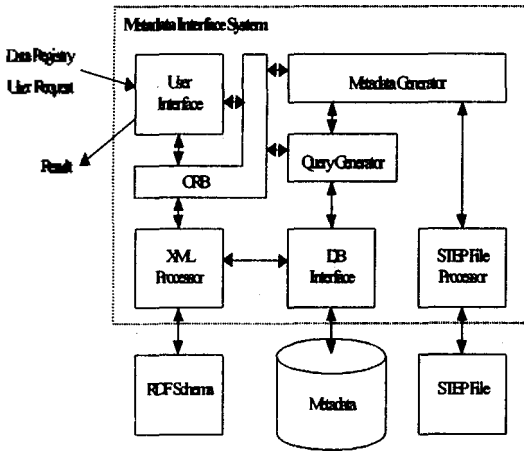
4.1 메타데이터의 웹 인터페이스

메타데이터 서버에 있는 지식 기반 시스템은 섹션 3.1 에서 설명한 바와 같이 메타데이터를 관리한다. 사용자 인터페이스 모듈은 자바 에이전트로 구현되었고 CORBA ORB 를 통해서 지식 기반 시스템과 통신한다. ORB 인터페이스는 이질적인 자바 가상머신들 상에서 보다 동적인 사용자 인터페이스를 가능하게 해 준다. 기존의 CGI 오퍼레이션들은 HTML 폼과 결합되어야만 하기 때문에, 유연성이 떨어지고 사용자 인터페

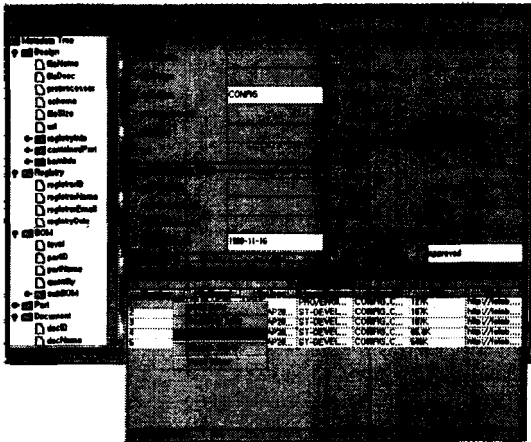
이스 측면에서 임의의 변화에 대해 지루한 업데이트 작업을 요구한다.

메타데이터는 관계형 데이터베이스 시스템에 저장 및 관리되며 XML 파일로서도 저장될 수 있다. 본 논문에서는 데이터베이스 시스템을 택했는데, 그 이유는 빠른 인덱스 구조와 동시성 제어, 시스템 오류 회복, 쿼리 처리와 같은 데이터 관리 도구들 때문이다. 메타데이터의 크기가 증가함에 따라 XML 파일 처리의 성능은 비례적으로 저하된다. 또한 XML 파일들은 사용자들에게 결과를 반환하기 위해 사용된다.

메타데이터는 사용자들이 설계 데이터나 문서들을 등록할 때에 생성된다. 이전 섹션에서 설명한 바와 같이, 메타데이터의 대부분은 STEP 파일로부터 자동으로 추출된다. 메타데이터 스키마는 STEP 응용 프로토콜들에 기초하므로, 주어진 STEP 파일들로부터 메타데이터를 추출해 내는 것은 그다지



<그림 6> 메타데이터 인터페이스의 구조



<그림 7> 메타데이터에 대한 사용자 인터페이스

어려운 작업이 아니다. <그림 6>에 있는 메타데이터 생성기는 메타데이터와 데이터베이스 인터페이스들을 추출하기 위해서 STEP 파일 처리와 상호 작용하게 된다. 데이터 검색은 일반 검색, 카테고리 검색, 또는 XQL 패턴로서 수행될 수 있다. 카테고리 검색은 <표 1>에 정의된 바와 같이 메타데이터 스키마에 기초한다. 메타데

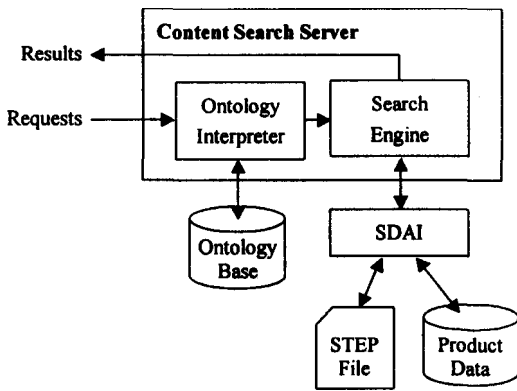
이터에 관한 사용자 인터페이스는 <그림 7>에 나타나 있다.

4.2 온톨로지 사용을 통한 내용 검색

제품 데이터는 표준 중립 포맷들 (예, STEP, IGES)이나 벤더-종속적인 데이터 포맷을 이용하여 보관될 수 있다. 표준 중립 포맷들은 ISO 나 ANSI 같은 표준 기구들에 의해서 개발되어 왔다. 벤더-종속적인 파일 포맷들은 상용 시스템들 (예, ACIS, AutoCAD)에서 사용되나, 그들의 구조는 개방되지 않기 때문에 서로 다른 시스템들간에 데이터 교환을 하는 목적으로 사용하기에는 부적절하다. 여기서 구현된 내용 검색 시스템은 STEP 에서 제공하는 개방형 구조 (예, Standard Data Access Interface [SDAI])와 EXPRESS 데이터 모델을 이용하여 구성했다. 이 시스템은 파일 이름들 뿐만 아니라 제품 데이터에 정의된 엔티 개체들을 검색할 수 있다. 추가적으로, 사용자들은 제품들의 특성 (이름, 재질, 기능 등)을 가지고 제품 데이터를 검색할 수 있다.

그렇지만, 이전 섹션에서 언급한 것처럼 단어들은 서로 다른 응용 도메인에서는 각기 다른 개념들로서 사용될 수 있으며, 이것은 주어진 쿼리들에 대해 올바르게 않은 제품 정보가 선택될 소지가 있다. 따라서, 검색의 정확성을 향상시키기 위해서는 주어진 키워드들의 컨텍스트를 고려해야 할 필요가 있다. 우리는 제품 구조들을 통합함으로써 내용 검색 기능을 강화시킬 수 있다. 이전 섹션에서 언급했던 바와 같이 컨텍스트와 제품 구조는 온톨로지에 정의된다. 주

어진 키워드들의 온톨로지를 고려함으로써, 설계 데이터베이스에 저장되어 있는 설계 지식의 사용을 향상시킬 수 있다. <그림 8>은 온톨로지 시스템을 갖는 검색 엔진의 구조를 표현한 것이다.



<그림 8> 온톨로지 해석기와 결합된 내용 검색 서버의 구조

STEP에서 모든 제품 모델은 객체 지향 정보 모델링 언어인 EXPRESS에 의해서 정의되며 [11], 파일 시스템이나 관계형 데이터베이스 시스템 또는 객체 지향형 데이터베이스 시스템과 같이 다양한 저장 시스템들에 저장된다. 서로 이질적인 저장 시스템들에 대해, 같은 메소드들로서 인터페이스할 수 있기 위해 SDAI가 정의되었다. 본 연구에서는 네트워크가 연결되어 있는 환경에서 분산 객체들에 접근하기 위해서 SDAI의 IDL 바인딩을 사용한다. 제품 데이터는 STEP 파일들이나 데이터베이스에 저장된다. 이번 섹션에 제시된 내용 검색 알고리즘은 온톨로지와 함께 전개된다. 주어진 키워드

가 검색되기 이전에, 키워드를 포함할 수 있는 단어들 이 온톨로지 저장소로부터 part-of 관계를 사용하여 수집된다. 그리고, 수집된 단어들 각각에 대해 같은 컨텍스트에서의 동의어들이 선택된다. 이 절차는 다음 알고리즘과 같이 정리될 수 있다.

알고리즘 1. 온톨로지 전개를 통한 내용 검색

- 입력: 요구 (해당 컨텍스트에서의 키워드)
- 출력: 제품 데이터의 목록
- 절차:

1. 비어 있는 두개의 리스트, key_list 와 onto_list를 초기화한다.
2. 주어진 키워드를 key_list에 입력한다.
 1. 온톨로지 저장소로부터 part-of 관계를 사용하여, 주어진 키워드를 포함하는 모든 단어들을 추가한다.
 2. key_list에 있는 각 단어에 대해서, 키워드와 그에 관련된 동의어들을 onto_list에 저장한다. 검색된 동의어들 중에서, 주어진 키워드가 존재하는 컨텍스트와 상응하는 것들만 선택한다.
 3. onto_list에 저장된 각각의 단어들에 대해 내용 검색을 실행하고 결과를 반환한다.

예로서, 자동차 설계 데이터를 관리하는 PDM 시스템에서 키워드와 컨텍스트가 (chassis, AP214)로서 주어진다고 가정을 해보자. AP214는 자동차 제품 데이터를 위한 STEP 응용 프로토콜의 이름이다. 알고리즘 1의 2 단계에서 key_list는 현재 [(chassis, AP214)]가 된다. 우리가 <그림 3>에서와 같이 part-of 관계를 갖는다고 가정을 한다면,

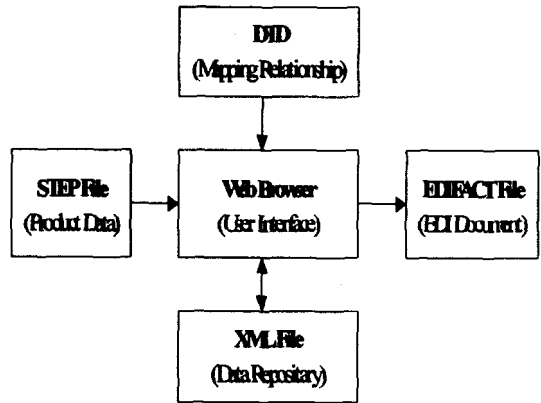
3 단계에서의 key_list 는 [(chassis, AP214), (automobile, AP214)]가 된다. [car, auto, automobile, motorcar] 또는 [chassis, frame]와 같은 동의어 목록들을 사용하여, onto_list 는 [(chassis, AP214), (frame, AP214), (automobile, AP214), (car, AP214), (auto, AP214), (motocar, AP214)]가 된다. 이 onto_list 가 검색 엔진으로 전달되고, 모든 단어/컨텍스트 쌍에 대해 내용 검색이 수행된다.

본 섹션에 서술된 내용 검색은 사용자들 또는 응용 시스템들이 파일 이름이나 도면 번호와 같이 특정 식별자에 관한 정보가 없더라도 현재 존재하는 PDM 시스템으로부터 필요한 제품이나 부품 정보들에 대한 검색이 가능하게 해준다. 이것은 현재 가지고 있는 설계 정보의 이용을 향상시키게 되는데, 이 결과로서 설계 소요 시간과 비용을 절감과 설계의 품질 향상 효과를 얻을 수 있다. 대부분의 상용 CAD/CAM 시스템들에 의해서 사용되는, 벡터-중속적인 데이터 포맷들은 비개방 구조이므로, 사용자가 같은 인터페이스 매소드들을 통해 내용을 검색할 수 없다. 본 섹션에 설명된 내용 검색 시스템은 개방형 구조와 STEP의 인터페이스 매소드들을 이용한다.

4.3 제품 데이터의 EDI 로의 번역

제조 회사들이 보관하는 많은 EDI 문서들이 제품 데이터와 관련된 공학 정보를 포함한다. 몇가지 예로서 제품 이름, 설계자, 사양, 묘사, 부품 목록 등이 있다. 문서들의 개수와 관련 제품 데이터는 방대하고 그들의 포맷들은 자주 변하기 때문에 각 경우

를 위한 번역기를 구현하는 것은 비실용적이다. 이번 서브섹션에서는 STEP 제품 데이터를 EDIFACT 문서들로 번역하기 위해 스키마 수준의 매핑 관계를 DTD (Document Type Definition)로 정의하는 방법을 다룬다.



<그림 9> 제품 데이터를 각 EDI 문서로 번역하는 흐름도

일단 매핑 관계가 DTD 로 정의되면, 사용자 인터페이스로서 웹 브라우저가 사용될 수 있다 (<그림 9> 참조). 번역 프로그램들과 비교하여, 스키마 수준의 매핑 관계는 정의하기가 쉽고, 차후 변화에도 대처가 용이하다.

매핑 관계는 각 EDI 문서에 대해서 DTD 로 정의된다. 예로서, 다음 DTD 는 XML 요소 중 document 와 id를 정의한다.

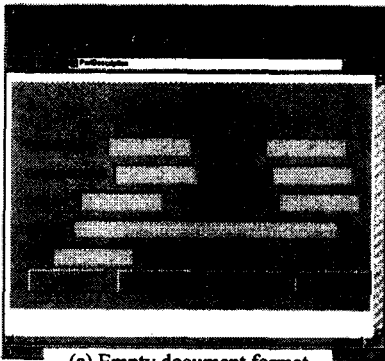
```
<ELEMENT document (id, name, type,
                                description)>
<ATTRIBUTE document id CDATA #Required>
<ELEMENT id (#PCDATA)>
<ATTLIST id
step_schema CDATA #FIXED '10303-203'
```

```
step_entity CDATA #FIXED 'document.id'
prefix CDATA #FIXED 'BGM++'
d_type CDATA #FIXED 'an..35'
suffix CDATA #FIXED "" >
```

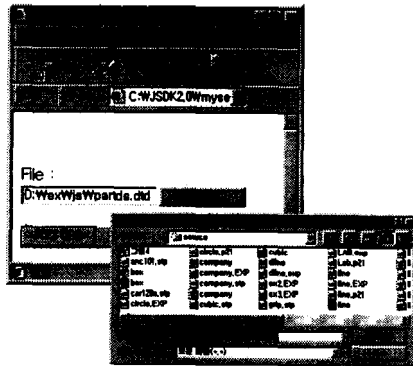
위의 DTD 에서 요소 id 는 다섯 가지 속성들 (i.e., step_schema, step_entity, prefix, d_type, suffix)을 지니고 있다. step_schema 와 step_entity 는 id 의 값이 관련되는 STEP 데

이터로부터 스키마와 엔티티 이름을 명시한다. 다른 세 개의 속성들은 EDIFACT 문서들로의 매핑을 규정한다. 데이터가 EDIFACT 문서로 번역될 때, id 는 “BGM++”와 같은 접두사와 “”와 같은 접미사를 갖게 되며, 이것은 35 alpha-numeric 문자들 까지 가질 수 있다. STEP 파일에 다음과 같은 데이터가 존재한다고 가정해 보자 :

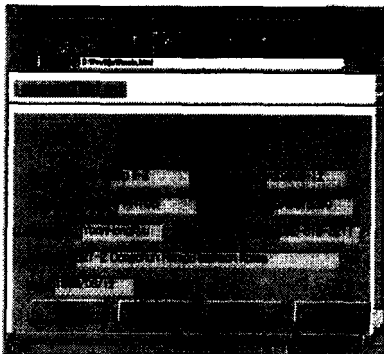
```
#500=DOCUMENT('AS990812','PARTS');
```



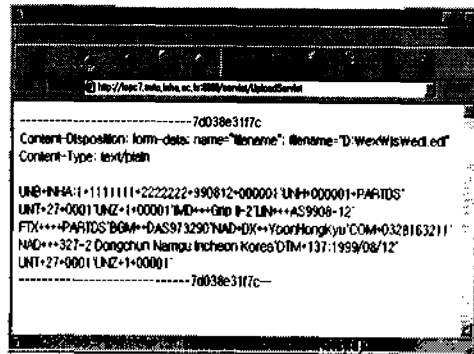
(a) Empty document format



(b) Select a STEP file



(c) Import product data



(d) Generated EDIFACT document

<그림 10> 제품 데이터의 EDI 문서로의 변환 예

위에 기술된 DTD 에서 정의된 매핑 관계에 의해서 다음과 같은 XML 파일이 생성된다:

```
<?xml version=1.0?>
<id> AS990812 </id>
```

그리고, 위의 데이터로부터 번역된 EDIFACT 파일의 일부는 다음과 같다:

```
BGM++AS990812'
```

STEP 데이터를 EDIFACT 문서로 번역하는 순서는 <그림 10>의 예제에서 보여지는 바와 같다. <그림 10-(a)>는 비어있는 문서 포맷을 보여준다. 사용자가 "Import" 버튼을 클릭하게 되면, <그림 10-(b)>에서와 같이 STEP 파일을 선택하거나 입력할 수 있다. <그림 10-(c)>는 제품 데이터 중 일부가 DTD 에 정의된 매핑 관계에 의해서 번역된 모습이다. <그림 10-(d)>는 <그림 10-(c)>에서 "ToEDI" 버튼을 눌렀을 경우 생성되는 EDIFACT 문서를 보여준다.

5. 결론

인터넷의 보급과 업무 전산화의 급속한 진행으로 각 기업체에서 많은 디지털 데이터를 보유하고 있으나, 이들을 적절히 활용하기 위한 지식관리는 아직 초보적인 단계이다. 기존 기식기반 시스템들은 주로 if-then 형태의 규칙과 추론 엔진을 이용하여 전문가들의 역할을 대신하는 용도로 많이 사용되었으나, 최근 웹 환경에서 전자거래를 지원하기 위한 많은 연구들이 진행되고 있다. 본 논문에서는 이러한 지식관리 기법으로서

메타데이터, 온톨로지, 그리고 데이터 매핑을 이용하여 제품데이터를 전자거래에 활용하는 방법을 기술하였다.

메타데이터는 데이터에 대한 정보로서 인터넷 환경에서 제품데이터에 대한 맵(map)으로서의 역할을 수행한다. 메타데이터는 웹 브라우저에 계층 구조로 이루어진 데이터 맵과 검색 기능을 제공함으로써 사용자에게 인터넷 상에 분산된 제품 데이터에 대한 정보를 제공한다. 온톨로지는 제품데이터에 포함된 용어들 간의 관계를 제공한다. 본 연구에서는 컨텍스트, 포함관계, 그리고 국내의 동의어를 고려하여 제품데이터의 내용검색에 적용하였다. 매핑 관계는 제품데이터의 다양한 표현 형식 간의 변환 관계를 표시하며, 스키마 수준의 변환 관계를 지식베이스에 의하여 관리하고 해석기에 의한 자동 변환시스템의 생성에 대하여 기술하였다.

본 논문에서 기술된 기식들은 웹 환경의 전자거래 구현에 직접 활용될 수 있다. 그러나 최근 새로운 CALS 및 웹 표준들이 계속 개발되고 있어서 이러한 신기술을 반영하기 위한 연구가 계속되어야겠다. 특히 메타데이터와 관련하여 개발되고 있는 XMI (XML Metadata Interchange) 표준은 서로 다른 사용자나 프로그램 간에 메타데이터를 교환하여 데이터의 활용을 촉진하는 역할을 한다. 그리고 본 연구에서 제안된 지식을 기존의 지식 표현 방식인 KIF (Knowledge Interchange Format) 등으로 표현하여 활용하는 방안도 연구되어야겠다.

참고 문헌

- [1] [Alain Bezos, 1999] "CEN/ISSS PBDH Project", Association GOSET, France.
- [2] [William C. Burkett, 1999] "Product Data Markup Language", <http://www.pdit.com/pdml>.
- [3] [D. Carter and B Baker, 1992] "Concurrent Engineering: The Product Development Environment for the 1990s", Addison-Wesley publishing company, Reading, Massachusetts.
- [4] [M. Case and S. Lu, 1996] "The Discourse Model for Collaborative Engineering Design," Computer-Aided Design, Vol. 28, No. 5, pp. 333 - 345.
- [5] [Y. Chen, C Liao, and B. Prasad, 1998] "A Systematic Approach of Virtual Enterprising through Knowledge Management Techniques", Concurrent Engineering: Research and Applications, Vol. 6, No. 3, Technomic Publishing.
- [6] [D. Chorafas and S. Legg, 1988] The Engineering Database, Butterworths.
- [7] [Peter Denno, 1999] "EXPRESS-X Language Reference Manual", ISO TC184/SC4/WG11 N088.
- [8] [Nicola Guarino, Claudio Masolo, and Guido Vetere, 1999] "OntoSeek: Content-Based Access to the Web," IEEE Intelligent Systems, pp. 70 - 80.
- [9] [Martin Hardwick, David Spooner, Tom Rando and K. C. Morris, 1996] "Sharing Manufacturing Information in Virtual Enterprises", Communications of the ACM, vol. 39, no. 2, pages 46-54.
- [10] [ISO, 1994] "Industrial Automation Systems and Integration Product Data Representation and Exchange Part 1: Overview and Fundamental Principles", Geneva.
- [11] [ISO, 1994] "Industrial Automation Systems and Integration Product Data Representation and Exchange Part 11: Description Methods: The EXPRESS Language Reference Manual", Geneva.
- [12] [ISO, 1988] "Industrial Automation Systems and Integration Product Data Representation and Exchange Part 22: Implementation Methods: Standard Data Access Interface Specification (Final Draft International Standard)", Geneva.
- [13] [Ora Lassila and Ralph Swick, 1999] "Resource Description Framework (RDF) Model and Syntax Specification", W3C Recommendation, <http://www.w3.org/TR/REC-rdf-syntax>.
- [14] [Jay Liebowitz, 1999] "Knowledge Management Handbook", CRC Press.
- [15] [N. Matta, O. Corby, and B. Prasad, 1998] "A Generic Library of Knowledge Components to Manage Conflicts in CE Tasks", Concurrent Engineering: Research and Applications, Vol. 6, No. 4, Technomic Publishing.
- [16] [A. McKay, M. Bloor, and A. de Pennington, 1996] "A Framework for Product Data", IEEE Transactions on Knowledge and Data Engineering, Vol. 8, No. 5.
- [17] [Daniel E. O'Leary, 1998] "Using AI in Knowledge Management: Knowledge Bases and Ontologies," IEEE Intelligent Systems, pp. 34 - 39.

- [18] [Craig Schlenoff, Rob Ivester, and Amy Knutilla, 1998] "A Robust Process Ontology for Manufacturing Systems Integration," Proceedings of 2nd International Conference on Engineering Design and Automation, Maui, Hawaii.
- [19] [Bill Swartout, 1996] "Future Directions in Knowledge Based Systems," ACM Computing Surveys, 28(4es).
- [20] [G. Toye, M. Cutkosky, J. Tenenbaum, and J. Glicksman, 1993] "SHARE: A Methodology and Environment for Collaborative Product Development," Proceedings of Second Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, pp. 33 ~ 47.
- [21] <http://ontolingua.stanford.edu>

저자소개

박상우 (sangwoo@ispcl.auto.inha.ac.kr)

인하대학교 자동차공학과 학사

현재 인하대학교 자동차공학과 석사과정

윤홍규 (reddog@ispcl.auto.inha.ac.kr)

인하대학교 자동차공학과 학사

현재 인하대학교 자동차공학과 석사과정

유상봉 (syoo@inha.ac.kr)

서울대학교 제어계측공학과 학사

미국 ARIZONA 주립대학교 전기및컴퓨터공학과 석사

미국 PURDUE 대학교 전기및컴퓨터공학과 박사

AT&T Bell 연구소 연구원

삼성전자 컴퓨터부문 선임연구원

현재 인하대학교 자동차공학과 부교수

김철환 (cwkin3@unitel.co.kr)

육군사관학교 병기공학 이학사

미국 PURDUE 대학원 재료공학과 석사

미국 PURDUE 대학원 재료공학과 박사

한양대학교 행정대학원 행정학과 석사

미국 해군대학원 객원교수

현재 국방대학교 무기체계과 교수