

## 제품 정보의 검색에 온톨로지 활용 방법

최무라\*, 유상봉\*\*

### A Method of Applying Ontology for Product Information Search

Choi, M. R.\* and Yoo, S. B.\*\*

#### 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, exchanging 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 utilize the ontology concept in order to facilitate information search for product data in the internet environment. A prototype of search system implemented using the ontology for automobile product data is presented.

**Key words** : Design information, Search, Ontology, Database, Product data

#### 1. 서 론

최근 인터넷, 인트라넷 등의 네트워크가 활성화 됨에 따라, 기업의 활동들 중 일부 또는 전부가 가상공간으로 이동함이 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 의해서 주로 관리되어 왔으나, 가상 기업들에 있어서는 제품 데이터가 제품의 전 생명 주기에 걸쳐서 활용될 수 있어야 한다(예, 웹 기반의 제품 목록, 제품 설명서 등)<sup>1,2)</sup>.

비즈니스 데이터와는 달리, 제품 데이터는 복잡한 의미상의 연결 구조를 가지며 서로 다른 응용 프로그램들 사이에 제대로 교환되지 않는다<sup>3,4)</sup>. 마록 몇몇 제품

데이터의 중립 표준들이 표준화 기구들에 의해서 개발되어 왔으나, 다양한 응용 프로그램이 이들을 효과적으로 교환하기 위해서는 이들 데이터가 갖는 의미에 대한 종합적인 이해가 요구된다. 최근에는 제품 데이터의 공유에 있어서의 이러한 문제점들을 극복하기 위한 방안으로서 더욱 다양한 지식을 활용하는 방안들이 연구되고 있다<sup>5,7)</sup>. 본 논문에서는, 가상 기업들이 필요로 하는 다양한 응용 프로그램들간에 제품 데이터의 원활한 공유를 촉진시킬 수 있기 위해 온톨로지를 이용한 제품 데이터 검색 방법을 제안한다.

온톨로지란 말은 철학의 존재론 또는 인식론에서 나온 말로서, 스탠포드 대학의 로직 그룹은 지식공유를 위한 프로젝트에서 온톨로지를 문제영역에서 존재하는 구체적 혹은 추상적인 엔티티로 정의 하였고<sup>8)</sup>, 스탠포드의 Tom Gruber는 간단히 개념화에 대한 명세서라 하였다<sup>9)</sup>. 온톨로지는 사물(비행기, 사자, 바다 등)이나 행위(자다, 생각하다, 쓰다 등)의 상속, 계층구조로 개념을 분류하며 개념들 사이를 'is\_a', 'part\_of', 'have\_a' 등 다양한 관계들으로써 연결해 준다. Ontology.org에서는 온톨로지의 주 목적을 서로 다른 시스템, 정보구조, 응용분야에 있는 컴퓨터 시스템간에 정보교환을 가능케 하는데 있다고 하였다<sup>10)</sup>. 이는 온톨로지가

\*학생회원, 인하대학교 자동차공학과  
\*\*정회원, 인하대학교 전자전기컴퓨터공학과  
- 논문투고일: 2001. 03. 06  
- 심사완료일: 2001. 07. 19

용어들의 정의와 구조 관계표현으로 서로 다른 시스템에서도 부리한 해석 없이도 적용시킬 수 있음을 의미한다. 최근에는 객체 모델링이나 XML에 포함시켜 전자상거래 분야에 도입되어, 에이전트들 사이에 정확한 의미 전달을 위해 활용되고 있다.

가상기업 중 자동차, 선박, 또는 항공 산업등과 같이 많은 부품들을 취급하는 분야에서는, 관련된 기관이나 단체 혹은 개인이 각각의 부품들을 어떻게 효율적으로 검색하고 활용하는 것이 중요한 과제가 되고 있다. 이러한 산업에서는 부품의 종류가 다양하고 상호 관계가 복잡하여 원하는 부품을 효과적으로 검색하는 것이 용이하지 않다.

본 논문에서는 자동차 부품 관련 용어를 대상으로 온톨로지를 구축하여, 부품의 상, 하위 및 관계에 의한 구조적 검색이 가능하도록 하였다. 또한 검색된 부품들에 대하여 제품 데이터 파일로의 연계를 통하여 부품들의 직접적 활용이 가능하게 하여준다. 주어진 키워드 간의 'AND' 또는 'OR' 관계에 의한 일반 검색이 가능하도록 하였고, 'class', 'component', 'slot' 등과 같은 다양한 관계를 통하여 각 부품들간의 기능적, 물리적, 개념적 검색이 가능하도록 하였다. 본 논문의 구성은 다음과 같다: 2장은 제품 정보분야에서 활용된 기존 온톨로지 연구를 기술하고, 3장은 본 논문에서 예로 사용한 자동차 부품 간에 존재하는 온톨로지를 설명한다. 4장에서는 3장에서 기술된 온톨로지를 검색에 활용하기 위하여 관계형 데이터베이스를 이용하여 구현한 프로토타입과 실행 예를 설명하고, 마지막으로 5장은 결론과 향후 과제를 기술한다.

## 2. 관련 연구

온톨로지의 개념은 언어 영역에서 유래하였고 그 결과 온톨로지가 단순히 언어 사전적인 측면이 중시되어 왔다. 그러나 온톨로지의 영역이 공학적인 측면으로 옮겨오면서 인공지능분야를 비롯해서 전자상거래 및 여러 가지 정보교환 프로토콜에까지 다른 분야의 전문화된 지식을 공유하려는 연구가 많이 진행되었거나 진행 중에 있다. 국내외의 대표적인 연구 사례는 다음과 같다.

미국의 NIST(National Institute of Standards and Technology)에서는 제조분야에 증가하는 복잡도와 다양한 소프트웨어간의 효과적인 정보교환을 위하여 제조분야의 개념이나 용어에 대하여 공식적이고 정확한 정의를 내리고 체계적인 분류(taxonomy)와 온톨로지 에 대한 연구를 수행하였다<sup>[13]</sup>. 서로 다른 기능을 하는

제조분야에서는 같은 의미에 서로 다른 용어를 사용하거나 같은 용어를 서로 다른 의미로 사용할 수 있기 때문이다. 이 연구에서는 기존의 온톨로지 시스템들을 분석해서 어떤 것이 제조분야에서의 용어의 개념을 모델링하는데 적합한지 분석하였다.

미시간대학교 도서관에서는 온톨로지에 기반하여 도서목록 메타데이터를 구조화 하고 그 관계들을 정의하였다<sup>[12]</sup>. 기존 도서목록으로 사용되는 카탈로그는 대부분 자연어로 쓰여있고, 이들 카탈로그는 도서의 검색에 도움이 되지만 컴퓨터 프로그램이 직접 활용하기는 어렵다. 미시간 대학의 온톨로지 그룹은 도서 목록들의 관계에 대한 양식적인 개념화를 통하여 이들 개념들의 아주 정교한 구조를 정의하였다.

한국과학기술연구원 CAD/CAM연구센터에서는 Active Design Support(ADS)라는 디자인 지식 관리 프레임워크를 통해 디자이너에게 필요한 디자인 지식을 제공하고 지난 제품개발 동안에 발생한 디자인 어려나 성공적인 디자인을 통해 합리적인 디자인을 이끌어 내도록 하는 연구를 진행했다<sup>[13]</sup>. 이런 목적을 위해서 제품디자인 과정 속에 있는 다양한 디자인 레포트, 문서 또는 메모 같은 지식 항목들의 분석을 통해 온톨로지 구조를 정의한 다음 이것을 ADS프레임워크의 정보 모델링의 기반으로 하였다. 즉 설계문서의 체계적 분석과 분류를 통하여 지식을 정량적으로 표현하게 함으로써, 설계자의 정보요구에 효과적으로 대응하기 위하여 온톨로지라는 개념을 사용하여 설계지식을 분류하고 새로운 설계지식의 유추를 가능케 하였다.

이 외에도 일반적인 부품들간의 구조적인 관계에 중점으로 만들어진 스탠포드 대학의 Ontolingua 서버<sup>[14]</sup>가 있고, Brigham Young에서는 온톨로지의 개념적 측면에서 접근하여 자동차 광고나 판매영역으로의 적용을 시도하였다<sup>[15]</sup>. 인터넷 검색에 대한 응용으로서, OntoSeek는 직업별란과 제품 목록에 대한 언어 온톨로지를 연관시킨 구조화된 내용의 표현들을 이용하여, 내용 기반 검색의 정확도와 검색효과를 증가시켰다<sup>[16]</sup>. 본 연구에서는 온톨로지 개념을 자동차 부품용어에 적용하여 네트워크 상에서의 제품 데이터 검색에 활용하는 방법과 프로토타입으로 제시하였다.

## 3. 제품 정보에서의 온톨로지

정보 기술 관점에서, 온톨로지는 특정 영역에서의 지식의 상호작용과 엔티티들의 작업 모델이다. 본 연구에서는 온톨로지를 제품 데이터에 포함된 어휘의 의미와 관계로서 사용한다. 제품들은 다양한 환경에서 서

로 다른 설계자에 의해 설계되기 때문에, 같은 용어가 서로 다른 개념들로서 사용될 수 있다. 예로서, '자원(resource)'는 워크플로우 시스템과 공정 계획 시스템에서 조금씩 다른 개념을 나타낸다<sup>[17]</sup>. 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 환경(context)을 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', 그리고 'motorcar' 등을 들 수 있다. 일반적으로 제품 설계는 제품에 관한 간략한 설명(description)들을 지니게 되는데, 이러한 설명들에는 이름, 용도, 사양, 제품의 특성 등이 포함될 수 있다. 제품의 설명은 자연어로 기술되기 때문에, 동의어 목록에 있는 어휘들 중 임의의 것이 선택될 수 있다. 내용검색이나 연관 검색 시에, 우리는 이러한 동의어 목록들을 사용함으로써 검색 효과를 향상시킬 수 있다.

STEP의 파트 214(Core data for Automotive mechanical design processes)<sup>[18]</sup>는 자동차의 기계 설계 과정에 수반되는 핵심 데이터를 정하는 응용프로토폴이다. 즉 차체부, 동력 전달부, 새시, 내장부의 설계과정에서의 다양한 개발 단계의 핵심 데이터들을 지원한다. 각각의 부분들은 여러 구성요소로 이루어져 있다. Fig. 1은 STEP 파트 214에서 사용한 구성 요소들 간의 'sub\_component' 관계를 도시한 것이다. 이와 같이 방향을 가지는 관계는 제품 데이터의 검색에 활용할 경우 두가지 방법으로 활용할 수 있다. 즉, 주어진 용어가 포함하는 부품을 검색하는 방법과 주어진 용어를

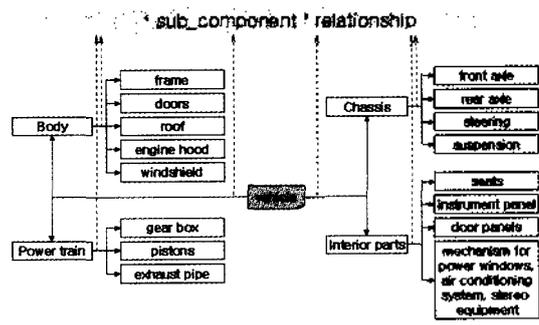


Fig. 1. Sub\_component relationship in STEP AP214.

포함하는 부품을 검색하는 방법이다. 예를 들어, '새시'가 검색어로 주어졌을 경우 그에 포함되는 '차축' 등을 함께 검색하거나, '새시'를 포함하는 '자동차'에 관련된 제품 데이터를 모두 검색하는 것이다.

서로 다른 나라에서 작업하는 엔지니어들도 다른 언어들을 사용할 수 있다. 예로서, 'Korea'는 한국어로는 '한국'으로 씌여지고, 중국어로는 '韓國'으로 씌여진다. 사람 이름이나 지리명 같은 고유어는 원어로 씌여질 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 인코딩할 수 있도록 유니코드를 사용한다. 이상에서 설명한 용어들 간의 관계 외에 자동차 제품 데이터의 용어들 간의 관계는 Table 1과 같이 정리될 수 있다.

Table 1에서 'translation' 관계는 서로 다른 언어간의 관계를 말한다. 예를 들면, '자동차'는 영어로 'car', 'vehicle', 'automobile'등으로도 표현된다. 'abbreviation'은 약어 관계이다. 예를 들면, 'spring'을 'SPG'와 같이 표현하는 것이다. 'synonym' 관계는 서로 다른 기관 또는 그룹 간에 발생한다. 예를 들면, 일반적인 자동차 부품 카타로그에서의 'BLOCK ASSY\_CYLINDER'이 AP214 파일에서는 'Block Assay Cylinder'처럼 다른 용어로 표현되기도 한다. 'class' 관계는 일반적인 일반화 또는 특별화 관계이다. 예를 들면, 자동차를 일반적인 용도별로 승용차, 버스, 트럭등과 같이 나누는 방법이다. 'component' 관계는 부품과 구성품 간의 관계이

Table 1. Relationship in product data terms

Relation	Subrelation
translation	
abbreviation	
synonym	
class	super_class of
	sub_class of
component	has sub_component
	sub_component of
slot	has part_slot
	subpart_slot of
connect	connected_components
	connects_components
replacement	equivalent
	temporary
synchronous repair	

다. 'slot' 관계는 부품이 체결되어 하나의 특정 기능을 갖는 것을 말한다. 예를 들면, 실린더 헤드와 실린더 블록은 'slot' 관계를 갖는다. 'connect' 관계는 부품들 사이의 연결을 나타낸다. 'replacement' 관계는 부품과 대체 가능한 부품들 간에 존재한다. 대처하는데 있어서 완벽하게 호환 가능한 'equivalent'와 일시적으로 대체 가능한 'temporary'로 구분한다. 'synchronous repair' 관계는 부품 교체 시 동시에 교체할 필요가 있는 부품들 간의 관계이다.

#### 4. 제품 정보 검색 시스템

##### 4.1 전체 시스템 구조

Fig. 2는 온톨로지를 이용한 제품 데이터 검색 시스템의 구조를 나타낸다. 기본적으로 클라이언트측은 자바 애플릿이기 때문에, 인터넷이 가능한 환경이면 어디에서든지 가능하다. Web Server측은 Solaris 2.6에서의 OrbixWeb3.0을 CORBA로 사용하였고, 온톨로지 데이터 베이스인 Onto\_Auto DB는 MS\_SQL 7.0을 사용하였다. 개발 툴로서는 JDK 1.2.2를 이용하였다.

웹 브라우저와 서버는 HTTP 프로토콜을 사용하고 Corba 프로그램은 OMG의 IIOP(Internet Inter-ORB Protocol)를 사용한다. 사용자는 애플릿을 통하여 사용자 등록과 질의를 수행하고, 검색 엔진은 주어진 검색어와 검색 옵션을 바탕으로 갖고 온톨로지 용어 DB에 질의한다. 주어진 검색 결과에 대하여 사용자는 STEP 파일 등의 관련된 파일을 추가로 요청할 수 있다.

##### 4.2 온톨로지 DB 구조

자동차 관련 용어들의 의미, 구조, 관계분석 결과를 토대로 수집된 자료들과 설정한 관계들을 이용해 데이터베이스를 구축하였다. 구축된 데이터베이스는 온톨로지 사전으로서의 역할을 하며 다른 응용프로그램의 지원역할을 한다. 예를 들면 자동차 검색엔진에 장착되어 단순히 일치되는 문자열과 그 의미만을 반환하는 것이 아닌, 검색하고자 하는 용어가 가지는 구조적인 의미와 다양하게 관계된 다른 용어들을 나열함으로써 정확한 검색을 가능하게 해준다. 또한 검색된, 관계 있는 용어들을 통하여 검색자가 원하는 정보를 추론할 수 있게 해주고 일차로 얻은 정보들을 가지고 다른 데이터 베이스로 이차 검색을 시도해 의도한 목적에 빠르게 근접하게 해준다. 이미 설정된 관계들을 통해 검색된 결과들 속에는 자동차 분야에서 많이 필요로 하는 정보를 용어와 관련해서 인출하므로 효과적인 검색을 수행한다.

본 논문에서는 온톨로지 중 단어들 간의 관계를 저장하여 검색에 활용한다. Fig. 3은 자동차 용어간의 관계를 저장하기 위한 E-R(Entity-Relationship) 모델을 보여준다. Fig. 3에 도시된 E-R 모델을 엔티티 "word"를 중심으로 Table 1에 포함된 9개의 관계로 이루어진다. 이러한 정보모델을 데이터 베이스에 저장하기 위하여 엔티티와 각 관계에 대응하는 테이블을 정의한다.

온톨로지 데이터베이스는 모든 용어들이 저장되는 word 테이블을 중심으로 용어들 간의 관계를 저장하는 10개의 테이블로 구성된다. Table 2는 word테이블의 속성들과 각 속성의 타입과 의미를 설명한다.

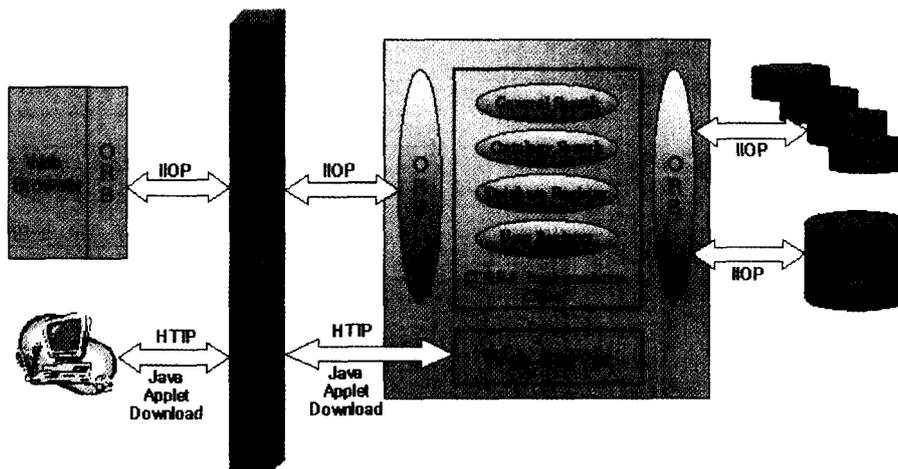


Fig. 2. Structure of search system based on ontology.

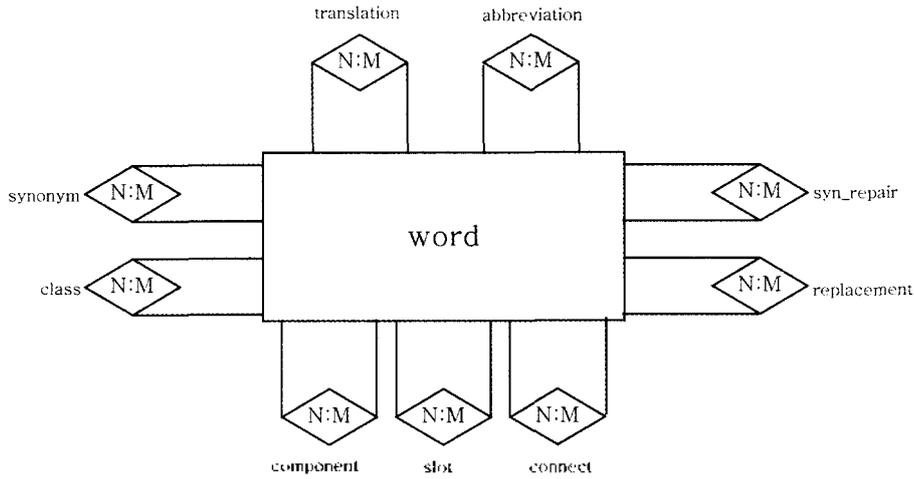


Fig. 3. Entity-relationship model to store ontology.

Table 2. Attributes of table "word"

Attribute	Type	Key	Meaning
Word_id	Int	Primary	용어의 고유 id
Name	Varchar		용어 이름
Description	Varchar		용어에 대한 간단한 설명
Location	Varchar		포함된 파일의 위치를 나타낸다.

Fig. 3에 포함된 관계(relationship)들은 모두 M:N 관계로서 각각 별도의 테이블을 생성하여 저장한다. 그 중 한 예로서 component 테이블의 정의는 Table 3과 같다.

검색하려는 용어에 대한 여러 가지 관계들을 갖는 용어들을 데이터베이스에 질의하기 위해서는 일관적인 질의문이 요구된다. 즉 단순한 용어검색을 비롯하여 상위계층 검색에 사용되는 질의문, 하위계층 검색 시 필요한 질의문, 소유관계 검색 시 필요한 질의문, 관리기관검색, 파일 및 이미지 검색, 동의어검색 등 사용자가 원하는 관계를 질의하기 위해서는 질의 템플릿을 만들어 요구하는 질의의 용어만 바꾸며 일관된 형

Table 3. Attributes of table "component"

Attribute	Type	Key	Meaning
Compo_id	Int	Primary	Component관계를 나타내는 고유 id
Word_id	Int	F	관계를 가지는 주체 Word_id
SubWord_id	Int	F	관계를 가지는 상대 Word_id

태의 질의를 처리할 수 있다.

4.3 제품 데이터의 검색 예

사용자는 등록을 통해 등록된 사용자임을 먼저 확인 후 데이터 베이스 등록과 검색을 가능하게 하여준다. 그리고 데이터 베이스 등록 인터페이스를 통해 외부 데이터베이스를 등록하여 검색능력을 확장 시켜준다. 검색 방법에 있어 두 가지 인터페이스를 제공하는데 일반 검색과 온톨로지 이용 검색이다. 일반 검색은 단순히 'Keyword'의 'And'와 'Or' 기능만 이용하여 검색하는 것이고, 온톨로지 검색은 원하는 관계설정 옵션에 따라 검색을 하는 것이다. Fig. 4는 온톨로지 검색 인터페이스로서, 검색어로 'CYLINDER BLOCK'을 입

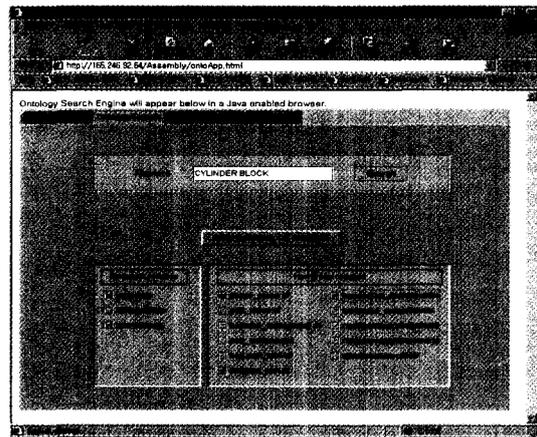
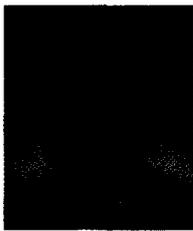


Fig. 4. User interface of search system.



12. Weinstein, P., Alloway, G. and Ahroheim, J., *Ontology Based Metadata*, University of Michigan Digital Library (UMDL), 1998.
13. 하성도 외 7인, 설계지식 체계화에 관한 연구, 한국과학기술연구원, 2000.
14. <http://ontolingua.stanford.edu/>.
15. Embley, D. W., Jiang, Y. and Ng, Y., "Record-Boundary Discovery in Web Documents," *Proceedings of ACM SIGMOD Conference*, pp. 467-478, 1999.
16. Guarino, N., Masolo, C. and Vetere, G., "OntoSeek: Content-Based Access to the Web," *IEEE Intelligent Systems*, pp. 70-80, 1999.
17. Schlenoff, C., Ivester, R. and Knutilla, A., "A Robust Process Ontology for Manufacturing Systems Integration," *Proceedings of 2<sup>nd</sup> International Conference on Engineering Design and Automation*, Maui, Hawaii, 1998.
18. ISO TC184/SC4, *Part214-Application Protocol: Core data for Automotive mechanical design processes*, 1998.



**최 무 라**

1998년 한양대학교 제어계측공학과 졸업 (공학사)  
 2001년 인하대학교 대학원 자동화공학과 졸업(공학석사)  
 2001년-현재 (주)쓰리에프 연구원  
 관심분야: 객체지향데이터 베이스, XML, ERP



**유 상 봉**

1982년 서울대학교 제어계측공학과 학사  
 1986년 Arizona 주립대학교 전기및컴퓨터공학과 석사  
 1990년 Purdue 대학교 전기및컴퓨터공학과 박사  
 1989년 AT&T Bell 연구소 연구원  
 1990년 삼성전자 컴퓨터부문 선임연구원  
 1992년-현재 인하대학교 전자전기컴퓨터공학부 교수  
 관심분야: 지식 및 데이터 베이스, STEP, Web Data처리