

## 6단계 상향식 방법에 의한 국방 온톨로지 추출

나민영\*, 양경용\*

### Extraction of Military Ontology Using Six-Step Bottom-up Approach

Minyoung Ra\*, Kyungyong Yang\*

#### 요약

국방 분야에서 기존에 구축된 정보시스템은 대부분 대용량의 정보처리 등 단순한 정보처리 중심이었기 때문에 정보의 지적인 수행 능력이나 지식의 표현기술에 대한 능력은 부족한 실정이다. 따라서 기존의 국방정보시스템이 효율적 정보검색 및 관리를 가능케 하는 지능형 국방정보시스템으로 변화되기 위해서는 지식 구축의 핵심이 되는 국방 온톨로지 구축이 필수적이다. 본 논문에서는 6단계 상향식 방법을 이용한 온톨로지 추출 방안을 기술하고, 이를 국방교육훈련 분야에 적용하여 국방 온톨로지 프로토타입을 추출한 후 대표적인 온톨로지 개발 툴인 Protégé를 이용하여 구현하였다.

#### Abstract

In national defense, established information systems are mainly based on simple information processing, such as mass data query. They have thus lacked intelligent ability of information and knowledge representation ability. We therefore need the research about the construction of military ontology which is the main topic for knowledge construction. Military ontology can help us develop the intelligent national defense information system which can search and manage information efficiently. In this paper, we present the six-step bottom-up approach for military ontology extraction, then we apply this approach to one of military domain, called national defense educational training, and finally implement it using Protégé which is one of the most useful ontology development tool.

▶ Keyword : 국방 온톨로지(Military Ontology), 온톨로지 추출(Ontology Extraction), 상향식 방법 (Bottom-up Approach)

• 제1저자 : 나민영

• 투고일 : 2009. 04. 22, 심사일 : 2009. 05. 12, 게재확정일 : 2009. 06. 02.

\* 육군사관학교 전자정보학과 교수

※ 본 논문은 육군사관학교 화랑대연구소의 2009년도 연구활동비 지원을 받아 연구되었음.

## 1. 서론

국방분야에서는 인사, 군수, 작전, 정보, 예산 등 전 분야에 걸쳐 정보화를 확대해 오고 있으며 각 기능별 혹은 단위업무별로 정보시스템을 이미 상당부분 구축하여 사용중에 있다. 국방 업무를 원활히 수행하기 위해서는 이렇게 다양하게 개발된 정보시스템 및 다양한 데이터 소스로부터 사용자 위주의 의미 기반 정보 검색이 효과적으로 제공될 수 있어야 한다. 그러나 현재 구축된 정보시스템을 살펴보면 이들은 대용량의 정보처리와 정보의 디지털화 및 응용 등 단순히 정보의 저장과 가공이 중심이었기 때문에 정보의 지적인 수행 능력이나 지식의 표현기술에 대한 능력은 부족한 실정이다. 따라서 우리 군에서도 과학화된 군, 경제적인 군, 그리고 비전 2020을 위해서는 지식의 관리가 필요하게 되었다. 지식이란 경험, 상황, 분석 및 설명 등이 결합된 것으로써 정보로부터 얻어진다. 군에서의 각 구성원은 나름대로의 지식을 가지고 있으며 또한 이를 발전시키고 있다. 그러나 문제는 이런 지식들이 체계적으로 정리되거나 저장, 공유되지 못하고 개인의 머릿속에만 남는 경우가 많다는 점이다. 또한 개인의 지식들은 대부분 비구조적이기 때문에 효과적 관리와 공유에 있어 많은 시간과 자원을 요구한다는 점이다. 따라서 전투력 향상 및 효율적인 군 운용을 위해서는 이러한 지식들을 어떻게 효율적으로 관리하고 유통시키며 재활용되도록 하느냐가 중요한 관건이 된 것이다.

이러한 국방 지식관리를 위해서는 기존의 정보시스템으로부터 지식관리시스템으로의 변화가 요구된다. 즉, 정보의 검색과 추출, 통합관리 등에 대한 새로운 기술적 접근이 필요하다. 특히 정보통합은 군에서 아주 중요한 기술로서 다양한 포맷의 지식들과 서비스들의 이질성을 극복하면서 상호운용성을 갖도록 연동해야 한다. 이를 위해서 바로 온톨로지 기술이 필요하다. 온톨로지 기술이 확보되면 정보와 지식의 홍수 속에서도 군에 꼭 필요한 유용한 지식을 효율적으로 발굴하고 관리할 수가 있게 된다. 그러나 국방 온톨로지에 관한 연구는 우리나라에서는 외국에 비해 아직 초보적인 단계로서 군수/조달 분야나 군사정보시스템에서 온톨로지 적용방안을 다루는 등 이제 연구가 시작되고 있는 실정이다[1,2].

온톨로지란 특정 분야에서 사용되는 어휘들의 개념화된 모음이라고 할 수 있다. 온톨로지는 그동안 자연어의 기계 번역과 인공지능 분야에서 활용되어 왔으나, 최근에는 시맨틱 웹이라는 새로운 세대의 웹 환경에서 인터넷 자원을 관리하는데 온톨로지가 제안되어 각광을 받고 있다[3]. 이러한 온톨로지는 특히 컴퓨터관련 분야에서 시맨틱 상호운용성, 표준화, 커

뮤니케이션, 지식 관리 및 검색, 그리고 기타 응용 등의 목적을 달성하기 위해 사용되고 있다[4,5,6].

이러한 온톨로지의 확보를 위해 데이터베이스(DB) 스키마로부터 온톨로지를 추출하려는 시도가 늘고 있다. 이러한 시도에는 크게 나누어 ER 모델로부터 시작해서 관계형 DB 스키마를 역공학적으로 도출해낸 다음 이로부터 온톨로지를 추출해내고자 하는 접근 방법과 관계형 스키마로부터 직접 온톨로지를 추출해내고자 하는 접근방법이 있다. [7]은 ER 모델로부터 온톨로지를 얻는 OntER이라는 tool을 제안하였고, [8]은 ER 모델로부터 OWL 온톨로지를 얻을 수 있는 ER2WO를 제안하였다. 관계형 모델 분야에서는 [9]는 관계형 모델 스키마로부터 OWL 온톨로지로 바꾸어주는 매핑 규칙을 제안하였으며, [10]은 ER2WO로부터 얻은 관계형 스키마로부터 매핑규칙을 이용하여 자동적으로 온톨로지를 생성해내는 tool인 D2OMapper를 제시하였다. 이러한 기법들은 대부분 데이터가 이미 관계형 데이터베이스로 구축되어 있는 경우에 유용하다. 그러나 국방분야는 관계형 데이터베이스만 존재하는 것이 아니고 파일시스템 등 다양한 형태의 데이터 소스가 존재할뿐더러 이들이 통합되는 광범위한 분야에 대한 온톨로지가 필요한 경우도 있다. 또한 아직 데이터베이스가 구축되지 않고 이제 시작하려는 분야도 있을 수 있다. 이러한 경우에는 앞에서 언급한 매핑기법을 적용하기가 곤란하다. 따라서 본 연구에서는 관계형 DB 스키마로부터의 매핑기법을 적용하기가 곤란한 상황에서 적용할 수 있는 기법인 상향식 방법에 의한 온톨로지 추출 기법을 다룬다. 본 논문에서 제시하는 6단계 상향식(Bottom-up) 온톨로지 추출방법은 [11]에서 언급된 반복적 클래스 계층 개발 방법을 확장하여 제안한 것이다. [11]의 방법에서는 온톨로지를 이용한 추론과 같은 지식처리시 사용될 논리구조분석 단계를 제외하면 온톨로지 개념화 모델링을 위해서 핵심어휘 추출, 계층구조형성, 개념클래스/인스턴스 정의 등의 3단계를 제시하였다. 이는 단순하면서도 매우 강력한 방법이나 클래스들간의 상호관계에 대한 정의가 동의어, 유사어 및 상위하위계층만 다루는 등 미흡한 면이 있었다. 따라서 본 논문에서는 이를 확장하여 온톨로지 추출을 위한 단계를 6단계로 구체화하고 상호관계 5가지, 즉 IS-A, Instance-Of, Part-Of, Syn-Of, Property를 명세하여 체계화하였다. 이와 같은 6단계 방법을 기존의 3단계 방법과 비교해보면 6단계 방법에서는 DB 모델링 이론에 근거한 속성관계 도출 등을 자세히 기술할 수가 있다.

본 논문의 구성은 다음과 같다. 2장에서는 온톨로지 구축을 위한 핵심기술인 XML, RDF, OWL을 다루고, 3장에서는 6단계 상향식 온톨로지 추출기법을 기술한다. 4장에서는

3장에서 제시된 기법을 국방교육훈련 분야에 적용하여 국방 온톨로지 프로토타입의 추출 및 Protégé를 이용한 구현을 다루고, 5장에서 결론을 맺는다.

## II. 온톨로지 구축 기술

온톨로지 구축을 위한 핵심 요소기술로서는 그 기반이 되는 XML, 웹 자원을 서술하는 RDF, 그리고 웹 온톨로지 언어인 OWL을 들 수 있다.

### 1. 기반기술 : XML

XML은 SGML(Standard Generalized Markup Language)의 서브셋으로 그 목적은 웹에서 구조화된 문서를 전송 가능하도록 하기 위한 것이다[12,13]. XML은 데이터의 구조를 정의하는데 초점을 맞추고 있으며 데이터의 표현은 XSL(Extensible Stylesheet Language) 스타일 시트가 담당한다. 즉, 문서와 프리젠테이션의 독립을 가진 구조이며 XML 문서에는 데이터만 기술하는 것이 아니라 구조적인 정보를 포함하고 있다.

XML을 사용하는 목적 중 하나는 서로 다른 기관이나 부서, 조직간의 정보를 공유하는 것이라고 할 수 있다. 따라서 어느 한 조직만을 위해 XML 문서를 설계하는 것보다는 다른 조직 즉, 다른 조직의 정보체계와 상호운용성을 고려해야 한다. 예를 들어, A 정보체계에서는 "이름"을 "부대이름"이라는 의미로 사용하고, B 정보체계에서는 "간부이름"이라는 의미로 사용하고 있다면, 두 정보체계는 같은 "이름"이라는 요소를 사용하지만 의미상으로는 다른 용도로 사용하고 있기 때문에 두 정보체계를 상호 운용할 경우 문제가 발생할 소지가 있다. 그래서 이와 같은 문제를 방지하기 위해서 XML에서는 이름공간(namespace)이라는 개념을 적용하고 있는데 이름공간이란 관련 있는 요소와 속성들의 집합으로 URI(Uniform Resource Identifier)를 통해 구별하고 있다. 그러므로 이름공간은 고유한 값을 의미하며, 요소와 속성 이름을 특정 조직의 범위를 넘어서까지 고유하게 함으로써 문서를 결합하거나 혹은 여러 개의 문서를 동시에 처리할 때 서로 다른 XML 문서 형식의 요소와 속성을 구별할 수 있는 수단을 제공하여 상호 충돌을 방지한다. 이름공간을 사용하기 위한 선언은 XML 문서 내에서 이름공간을 요소나 속성의 접두어로 사용하여 다른 요소나 속성과 구별하게 한다.

선언 : <(접두어):Element(xmlns:(접두어)=URI)\*>

XML 문서의 구조를 명시적으로 선언하기 위해서는 DTD와 XML 스키마를 사용한다. DTD(Document Type Definition)는 XML 문서를 개발하는데 사용되는 규칙들의 집합이고 XML 스키마는 DTD의 대안으로 등장한 XML 문서의 구조 기술 언어이다.

### 2. 웹자원 서술 : RDF/RDF스키마

웹자원을 트리플(triple)형태로 서술하는 RDF(Resource Description Framework)는 1999년에 W3C 권고안으로 확정되었다. RDF는 웹상의 리소스에 관한 정보를 명확하고 논리적으로 표현하는 데이터 모델이며 이를 기술하기 위한 언어체계이다[4,14]. RDF의 표현은 주어 역할을 하는 주어부, 서술어의 역할을 하는 부분으로 자원간의 관계를 표현하는 술어부, 그리고 목적어의 역할을 하는 목적부의 트리플 형태의 서술문으로 표현된다. RDF에서 사용하는 모든 자원의 이름은 식별이 가능해야 하므로 URI참조에 의해 표현된다. 이러한 트리플은 RDF의 가장 기본적인 기술단위로서 이의 그래프 표현은 그림 1과 같다.

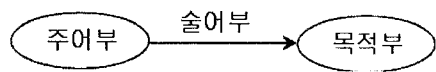


그림 1. RDF 그래프 표현  
Fig. 1. RDF Graph Expression

이러한 RDF 서술문의 세 가지의 개념을 이용하여 "3중대의 중대장은 홍길동이다."라는 문장을 그래프로 표현해보면 그림 2와 같다.

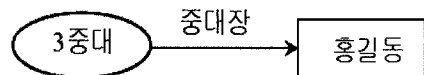


그림 2. RDF 그래프 표현 예  
Fig. 2. An Example of RDF Graph Expression

이 그래프는 각 이름이 붙어있는 노드와 이를 연결하는 아크(arc)로 구성되어 있는 방향성 그래프이다. 아크는 자원(서술문의 주어부)에서 출발하여 값(서술문의 목적부)로 향한다. RDF 그래프에서 목적부는 URI에 의해 참조되는 리소스뿐만 아니라 문자열(literal)이 되는 경우도 있다. 목적부가 리터럴인 경우는 그림 2에서와 같이 사각형으로 표현된다. 그림 2의 RDF 그래프를 컴퓨터가 이해할 수 있도록 XML을 이용하여 나타내면 다음과 같다.

```
<rdf:Description rdf:about="&mydomain:3중대">
  <mydomain:중대장>
    홍길동
  </mydomain:중대장>
</rdf:Description>
```

RDF는 사용자가 기존의 어휘 체계를 사용하여 자원들을 설명할 수 있게 해주는 언어이지만 속성들의 도메인을 제한하거나 비슷한 자원들을 묶어서 클래스로 표현하는 등의 기능이 없어 이를 해결하기 위해 RDF 스키마가 제안되었다. RDF 스키마는 속성, 메소드를 가지고 있고 클래스와 상위클래스, 하위클래스의 개념이 있고 각각 상위클래스로부터 속성과 메소드를 상속받아 사용할 수 있다.

### 3. 웹 온톨로지 언어 : OWL

OWL(Web Ontology Language)은 RDF 및 RDF 스키마의 제한된 표현력에서 오는 한계를 보완하기 위해 개발된 온톨로지 언어이다[14].

OWL에서 클래스는 그 의미에 따라서 계층을 형성하게 된다. OWL에서 클래스의 계층을 형성하는 간단한 예는 다음과 같다.

```
<owl:Class rdf:ID="CommandCadet">
  <rdfs:subClassOf rdf:resource="#cadet"/>
</owl:Class>
```

위 예에서 보듯이 하위 클래스는 rdfs:subClassOf를 통해 상위 클래스와 연결이 된다. 이 표현은 앞서 다룬 RDF 스키마에서 사용되는 표현이다. 이와 같이 OWL에서는 추가적인 요소가 없이 동일한 의미로 사용되는 표현은 RDF 스키마와 동일하게 사용한다.

또한 OWL에서는 RDF 스키마에서의 부족한 표현력을 향

상시키기 위해 몇가지 추가적인 어휘를 사용한다. 예를 들면, TransitiveProperty, SymmetricProperty, FunctionalProperty, onProperty, allValuesFrom[12] 등의 세부적인 속성을 나타낼 수 있는 어휘를 사용하여 속성의 표현력을 높이고 있다.

## III. 상향식 온톨로지 추출 방법

6단계 상향식 온톨로지 추출 방법의 각 단계는 다음과 같다.

### 단계 1. 온톨로지 도메인/범위 결정

온톨로지가 추출되어질 도메인과 범위를 결정한다.

### 단계 2. 핵심 용어 추출

결정된 도메인 분야에서 사용되는 핵심 용어를 도출한다. 이러한 용어는 MDR(Metadata Registry)에 정의되어 있는 용어를 사용하면 차후 상호운용성에 유용하다.

단계 3. 중복용어/의미모호 용어 정리 및 동의어/유사어 식별이 단계에서는 중복되거나 의미가 모호한 용어를 정리하고 동의어/유사어를 식별해낸다.

### 단계 4. 클래스 정의

정리된 핵심용어로부터 bottom-up 방법을 적용하여 클래스들을 도출해낸다.

### 단계 5. 계층구조 형성

핵심 용어로부터 도출된 클래스들을 그룹화하여 클래스 계층을 형성한다.

### 단계 6. 계층구조 세부 정의

클래스간 관계 및 속성을 정의하여 온톨로지를 완성한다. 이 때 동의어/유사어 관계도 정의한다. 온톨로지에서 각 개념들은 상호관계에 의해 서로 연결된다. 이 과정은 전문 지식이 많이 필요한 과정으로 좋은 온톨로지 구축을 위해서는 관계 정보가 빠짐없이 정의되어야 한다. 본 논문에서는 이러한 상호관계로 다음의 5가지를 사용한다.

- IS-A : 이 관계는 특수화(specialization)를 나타내는 데 사용된다. 만일 Cj가 Ci의 한 종류라면 Cj로 표현된 개념은 Ci로 표현된 개념의 특수화라고 한다.
- Instance-Of : Instance-Of 관계는 개념의 실체화이

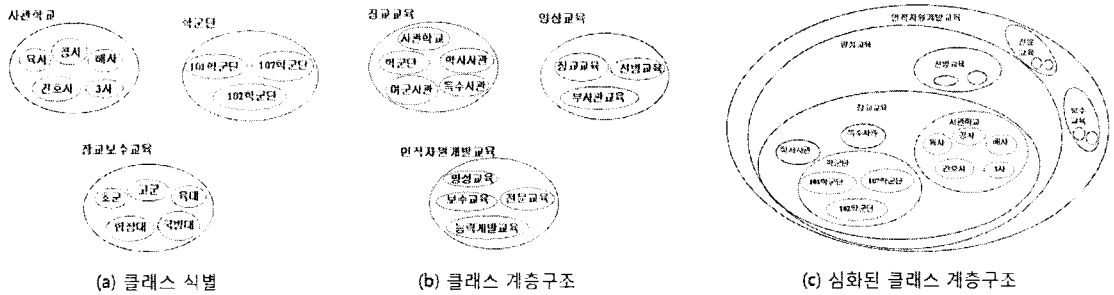


그림 3. 국방 온톨로지 추출 과정  
Fig 3. Extraction Process of Military Ontology

다. 만일 개념  $C_j$ 가 개념  $C_i$ 의 하나의 예라면 그들간의 관계는 Instance-Of 관계에 해당된다.

- Part-Of : 만일  $C_i$ 가  $C_j$ 를 부분으로서 가지고 있거나 또는  $C_j$ 가  $C_i$ 의 일부분이라면,  $C_j$ 에 의해 표현된 개념은  $C_i$ 로 표현된 개념의 Part-Of이다.
- Syn-Of:  $C_i$ 와  $C_j$ 가 동의어를 나타낸다.
- Property:  $C_i$ 와  $C_j$ 가 특정 속성관계를 맺고 있음을 나타낸다. 여기서  $C_i$ 와  $C_j$ 는 외부키에 의한 참조 속성 개념을 만족할 수 있도록 한다. 이를 위하여 속성은 외부키 관계가 적용될 수 있는 엔티티에만 적용한다. 외부키는 상대방 클래스  $C_j$ 의 키 애트리뷰트가  $C_i$  클래스에 있어야 하므로 이러한 관계가 성립될 수 있도록 각 클래스는 다음 사항을 만족해야 한다.
  - $C_i$ 와  $C_j$ 는 부모 자식 관계가 아니어야 한다.
  - $C_i$ 와  $C_j$ 는 그 클래스의 인스턴스들이 전부 속성 관계에 참여하는 전체참여(Total Participation) 관계를 갖는다.

#### IV. 국방교육훈련분야 온톨로지 추출

이 장에서는 앞 장에서 기술한 6단계 상향식 기법을 이용하여 국방교육훈련 분야에 대해 온톨로지를 추출한다.

##### 1. 국방 온톨로지 추출

이 절에서는 국방 온톨로지를 추출해가는 과정을 단계별로 예를 들어 다룬다.

##### 단계 1. 국방 온톨로지 도메인/범위 결정

본 논문에서의 국방 온톨로지는 그 도메인을 국방교육훈련 분야로 하여 온톨로지를 구축한다.

##### 단계 2. 국방교육훈련 핵심 용어 추출

국방교육훈련 분야에서 사용되는 핵심 용어를 도출한다. 도출된 용어의 예는 아래와 같다.

국방이론, 군사행정, 국방교육훈련, 군사시설장비, 신병훈련, 장교훈련, 장교훈련, 사관학교, 육사, 해사, 공사, 3사, 부사관훈련, 장교재훈련, 국방대학, 간호사, ROTC, 107학군단, 양성교육, 보수교육, 전문교육, 능력개발교육, 개인훈련, 부대훈련, 실무능력향상교육, 야간학위교육, 장교교육, 장교보수교육, 전문학위교육, 신병교육, 부사관교육, 여군사관, 국외군사교육, 군가족지원교육, 사관학교, 장교교육, 직무향상교육, 전직지원교육, 전문학위교육, 사병교육, 연합훈련, 합동훈련, 전술훈련, 학사사관, 학군단, ROTC, 학군교, 사단신교대, 육군훈련소, 주특기훈련, 기본훈련, 체력단련, 준사관교육, 공중강습훈련, 도하훈련, 연합훈련, 공지합동훈련, 대전차공격훈련, 수류탄투척훈련, 야외종합훈련, 과학화대대전투훈련, BCTP, KCTC, 전투지휘훈련, 소부대전투기술훈련, 삼사, 신병교육대, 학군후보생, 간부교육, 부사관보수교육, 초군반, 고군반, 육군대학, 합참대, 국방대, 초급반, 고급반, 101학군단, 102학군단, 특수사관, 국내석사과정, 국내박사과정, 국외석사과정, 국외박사과정, 야간석사과정

##### 단계 3. 중복용어/의미모호 용어 정리 및 동의어/유사어 식별

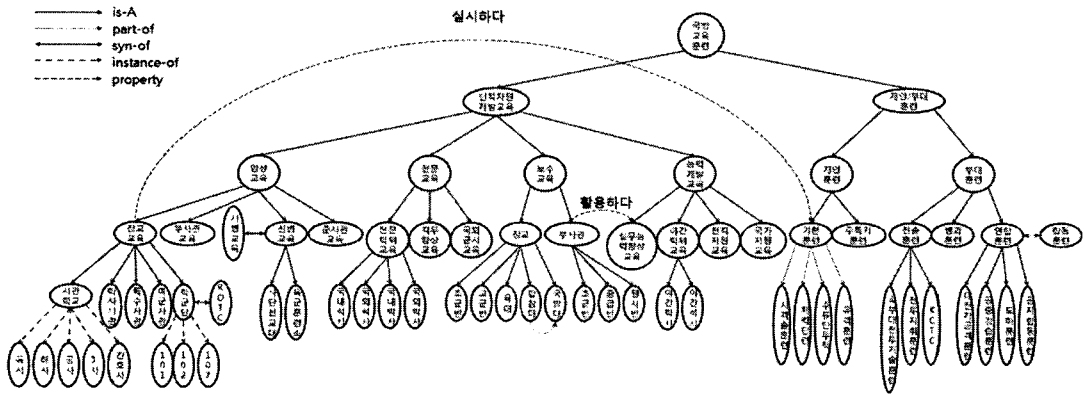


그림 4. 국방교육훈련 온톨로지  
Fig. 4. National Defense Educational Training(NDET) Ontology

위의 예에서 의미가 모호한 용어는 분별해내어 모호성이 제거된 용어로 통일하여 사용토록 하였다. 예를 들어 다음과 같은 용어 정리를 수행하였다.

- 신병교육, 신병훈련 ==> 신병교육으로 통일
- 장교교육, 장교훈련 ==> 장교교육으로 통일
- 부사관교육, 부사관훈련 ==> 부사관교육으로 통일
- 장교재훈련, 장교보수교육 ==> 장교보수교육으로 통일

또한 다음과 같은 동의어/유사어를 식별해 내었다.  
(삼사, 3사), (ROTC, 학군단)  
(사병교육, 신병교육), (연합작전, 합동작전)

**단계 4. 클래스 정의**

위의 예에서 상향식 방법을 적용하여 도출된 클래스들은 다음과 같다.

- 육사, 해사, 공사, 3사, 간호사 ==> 사관학교
- 101학군단, 102학군단, 107학군단 ==> 학군단
- 초군반, 고군반, 육대, 합참대, 국방대 ==> 장교보수교육
- 초급반, 고급반 ==> 부사관보수교육
- 사관학교, 학군교, 학사사관, 특수사관, 여군사관 ==> 장교교육
- 전문학위교육, 직무향상교육, 군외군사교육 ==> 전문교육
- 실무능력향상교육, 야간학위교육, 전직지원교육, 군가족지원교육 ==> 능력개발교육
- 장교보수교육, 부사관보수교육 ==> 보수교육

장교교육, 신병교육, 부사관교육 ==> 양성교육  
양성교육, 보수교육, 전문교육, 능력개발교육==>인적자원개발교육

이와 같은 클래스들을 이해하기 쉽게 그림으로 나타내면 그림 3의 (a)와 같다.

**단계 5. 계층구조 형성**

클래스 계층 형성과정 및 계층은 그림 3의 (b), (c)에서 보는 바와 같다.

**단계 6. 계층구조 세부 정의**

클래스간 관계 및 속성을 정의하여 완성된 온톨로지를 그림으로 나타내면 그림 4와 같다. 이 그림에는 앞에서 설명한 5가지 관계가 모두 나타나 있다. 즉, '학군단'과 'ROTC'가 동의어로, '101학군단', '102학군단', '107학군단'이 인스턴스로, 그리고 '기본훈련'의 부분개념으로 '사격술훈련', '체력훈련', '수류탄투척', '유격훈련'이 표현되어 있고, 기본 계층구조는 IS-A 관계를 나타낼 수 있다. 속성 관계를 살펴보면 '장교교육' 클래스와 '기본훈련' 클래스 간에는 부모 자식 관계도 아니고 또한 모든 장교교육에서는 모든 기본훈련을 받게 되므로 전체참여로 판단되어 '실시하다'라는 속성이 추출되었다. 또한 '부사관보수교육' 클래스와 '실무능력향상교육' 클래스 간에는 '활용하다'라는 속성이 추가되었음을 알 수 있다.

2. Protégé를 이용한 구현

본 논문에서는 그림 4의 온톨로지를 Protégé를 이용하여 구현하였다. 구현방법은 먼저 클래스와 계층구조를 입력한 후 관계를 입력하여 완성하였다. Protégé(15)는 스탠포드 대학교에 의해 개발된 온톨로지 개발 툴로서 현재 세계적으로 많은 사용자들이 이를 사용하고 있다.

Protégé를 이용하면 클래스를 그래픽 유저 인터페이스를 통해서 손쉽게 설정해 줄 수 있다. 이때 클래스는 우리가 익숙한 트리구조로 보여주기 때문에 확인이나 수정도 손쉽게 할 수 있다. 또한 만들어진 온톨로지를 RDF/XML 소스 코드로 자동으로 변환 제공가능하다. 따라서 온톨로지를 직접 소스 코드로 바꾸는데 필요한 과정을 완전 자동화 할 수 있다. 그림 5는 그림 4의 온톨로지를 Protégé를 사용하여 구축한 결과에 대한 XML 소스를 보여준다. 이 그림에서 보는 바와 같이 XML 소스 코드는 RDF와 OWL 어휘를 사용하고 있음을 알 수 있다.

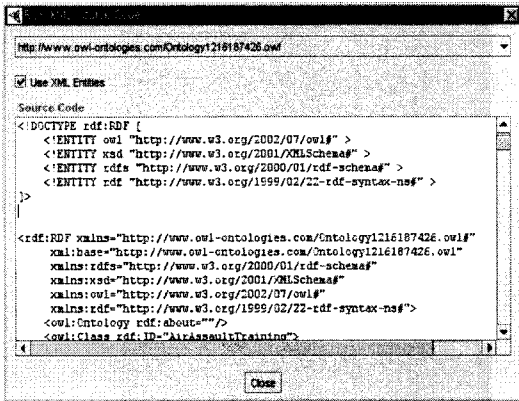


그림 5. XML 변환  
Fig. 5. Translation into XML

Protégé에서는 만들어진 온톨로지를 보기 쉽게 그래프의 형태로 표현해주는 플러그인이 많이 사용되는데 그 중에서 클래스간의 관계를 보여주는 OWLViz와 특정 개체를 중심으로 연관된 클래스와 인스턴스들을 확장 단계별로 거미줄처럼 보여주는 TGViz를 사용하여 국방교육훈련 온톨로지의 구성을 도식화해본다.

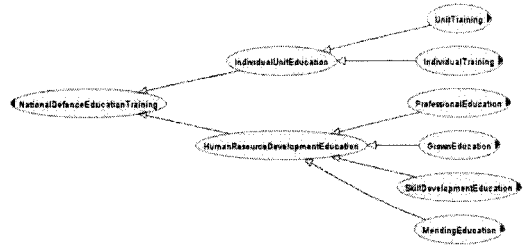


그림 6. 국방교육훈련부터 2단계 계층만을 나타낸 그래프  
Fig. 6. 2 Level Graph From NDET Using OWLViz

그림 6은 OWLViz를 사용하여 국방교육훈련부터 2단계 계층까지의 클래스를 나타낸 그래프이다. 이와 같이 OWLViz를 이용하면 클래스 간의 상하관계를 한눈에 파악할 수가 있다.

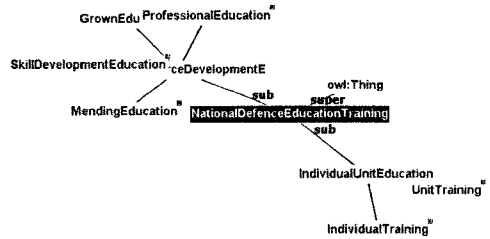


그림 7. 국방교육훈련을 중심으로 2단계 확장까지 나타낸 결과  
Fig. 7. 2 Level Extension From NDET Using TGViz

그림 7은 TGViz를 사용하여 국방교육훈련을 중심으로 2단계 확장까지의 클래스의 이름과 속성을 나타낸 것이다. 이 때 나타내고 싶은 확장 단계나 확장의 기준이 되는 개념은 자유롭게 설정 가능하다. 따라서 특정 개념을 기준으로 그 개념이 다른 개념들과 어떻게 연관되어 있는가를 알아보기 쉽게 도식적으로 나타낼 수 있다. 즉 컴퓨터가 쉽게 이해하도록 만들어진 온톨로지상의 개념적 인접성을, 사람이 이해하기 쉽게 공간적 인접성으로 전환하여 보여줄 수가 있다.

3. 분석

이렇게 국방 온톨로지를 구축함으로써 얻을 수 있는 이점은 각 클래스간의 속성을 통해서 클래스간의 연관성을 빠르게 파악할 수 있기 때문에, '기본훈련을 교육과정에 포함하고 있는 훈련에는 무엇이 있는가?' 와 같은 질문에 빠르고 명확하게 답할 수가 있다. 다음 그림 8은 이를 보여주고 있다.



그림 8. 기본훈련을 중심으로 1단계 확장한 결과  
Fig. 8. 1 Level Extension From BasicTraining

그림 8은 기본훈련(BasicTraining)과 1단계 연관을 맺고 있는 클래스들을 나타내어 보여준다. 여기서 우리는 이 클래스를 훈련과정에 포함하고 있는 클래스들, 즉 beConsistOf 라는 속성으로 나타내어지는 클래스에는 학사장교(GraduateOfficers), 학군단(ReserveOfficers), 특수사관(SpecialOfficers)이 있다는 것을 알 수 있다. 그림 9는 Protégé의 질의(query) 기능을 이용한 질의 수행 결과를 보여주는 그림으로, 질의를 통해서도 동일한 결과를 얻을 수 있음을 알 수 있다.

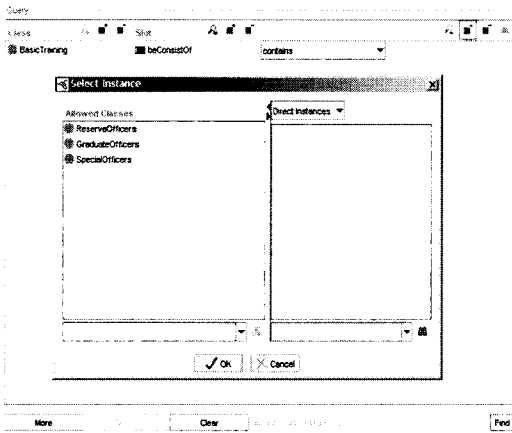


그림 9. 질의 실행 결과  
Fig. 9. Query Result

본 논문이 기여한 바를 요약하면 다음과 같다.

- 국방분야에 대한 온톨로지의 추출 : 국방분야의 한 도메인에 대하여 온톨로지를 추출해 봄으로써 시맨틱 웹 기술을 국방분야에 적용하는 선도역할을 수행하였다.
- 상향식 온톨로지 추출방법 상세화 : 기존의 3단계 방법

을 6단계로 확장하여 구체적인 절차를 제시하였고 기존의 방법에서 취약했던 속성관계 추출을 자세히 다루었다.

본 논문에서 제시된 6단계 상향식 방법을 기존의 3단계 방법[11]과 비교해보면 기존의 3단계 방법에서는 클래스간의 속성 관계를 자세히 나타내는 것이 제한되었던데 반해 본 논문에서 제시된 6단계 방법은 절차를 상세화하였고 특히 속성 관계 도출을 DB 모델링 이론에 근거하여 자세히 설명하였다. 이러한 속성 관계는 차후 데이터 저장시 효율적인 DB 구축에 활용될 수 있다.

다른 상향식 온톨로지 생성 방법인 Cyc[16]은 다수의 지식을 규칙화하여 저장한 후 이를 기반으로 연관된 지식을 도출하는 방법이다. 이는 전통적인 상향식 방법을 따르고 상당한 양의 지식을 포함하고는 있지만 대부분이 상식적인 수준으로 간단한 지식사전 수준의 속성 관계를 얻어내기에는 좋지만 좀 더 세분화된 온톨로지의 추출은 힘들다. 이에 반해 본 논문에서 제시된 방법은 개념화 단계를 충분히 거쳤기 때문에 연관성 깊고 세부적인 온톨로지의 추출이 가능하다.

또한 본 논문에서 제시된 방법을 스키마변환에 의한 온톨로지 생성 방법과 비교해보면 본 논문에서 제시된 6단계 상향식 방법에 의한 온톨로지 추출 기법은 데이터베이스가 구축되어 있지 않은 자료에도 적용가능하고 비교적 단순하며 적용이 용이하다. 그러나 DB 스키마로부터 매핑에 의한 방법보다 개념적인 접근방식으로 클래스 간의 관계 파악이 개발자 지식에 의존해야 하는 등 온톨로지를 구축하는 전문가의 직관과 능력에 크게 의존적인 점이 단점으로 지적될 수 있다. 이를 요약하여 표로 나타내면 표 1과 같다. 이 표에서는 여러 방법들의 상대적 성능 관계를 상, 중, 하로 비교 표현하였다.

표 1. 온톨로지 추출방안 비교

Table 1. Comparison of Approaches for Ontology Extraction

구분	상향식 3단계	상향식 6단계	Cyc	DB스키마 변환
입력	DB스키마 없어도 가능	DB스키마 없어도 가능	DB스키마 없어도가능	DB 스키마
상호 관계 표현성	하	중	중	상
온톨로지 추출 용이성	상	상	중	중



## V. 결론 및 향후 연구방향

국방 분야에서 그동안 구축된 정보시스템들의 효율적 정보 검색 및 관리를 가능케 하기 위해서는 무엇보다도 더 특정분야의 개념구조를 명시적으로 표현하는 지식 구축의 핵심이 되는 국방 온톨로지 구축에 관한 연구가 필수적이다.

본 논문에서는 먼저 온톨로지 구축을 위한 핵심 기술로서 XML, RDF 및 OWL을 조사한 다음, 6단계 상향식 방법을 이용한 온톨로지 추출 기법을 기술하고, 이를 국방교육훈련 분야에 적용하여 국방 온톨로지 프로토타입을 추출한 후 대표적인 온톨로지 개발 틀인 Protégé를 이용하여 구현하였다.

이렇게 구축된 국방교육훈련 온톨로지를 이용하면 그 중에서 특정한 혹은 관심이 있는 어떠한 훈련 개념을 가지고 그 훈련개념과 유사한 다른 훈련이라든지, 그 훈련이 속해있는 상위개념의 훈련 혹은 하위 개념의 훈련은 무엇인지를 쉽게 파악할 수가 있다. 이러한 특징은 정보통합시 유용하게 사용될 수 있다[17,18]. 또한 유사한 개념이나 동일개념 등의 파악이 매우 용이하다. 뿐만 아니라 사용자가 실제로 찾아서 파악해야 하는 지식적인 추론과정이 온톨로지 안에 담겨져 있어서 단순한 검색이 아닌 지식적 추론 검색도 적용 가능하다 [19,20].

앞으로의 연구 방향으로서 제시된 6단계 상향식 방법과 DB 스키마로부터 매핑에 의한 방법을 혼합하여 온톨로지를 추출하는 방안인 혼합방안에 관한 연구와 국방 온톨로지 통합 기법 및 국방 온톨로지 기반 정보 검색 기법에 관한 연구 등이 계속되어야 할 것이다.

## 참고문헌

- [1] 김성환, "국방 군수 표준 온톨로지 구축에 관한 연구", 서울대 대학원 석사학위논문, 2006년.
- [2] 송세현, 신석철, 김민구, "군사정보시스템을 위한 온톨로지 기반 사용자 프로파일", 한국정보과학회 2008 가을학술발표논문집 제35권 제2호(C), 223-227쪽, 2008년 10월.
- [3] L. Ma, et al., "Effective and Efficient Semantic Web Data Management over DB2", SIGMOD'08, pp.1183-1193, 2008.
- [4] 노상규, 박진수, "인터넷 진화의 열쇠 온톨로지", 가스토이, 2007년.
- [5] 안후영, 박영호, "관계형 데이터베이스 기반 온톨로지 기

- 법을 활용한 모바일 패선정보 제안", 한국컴퓨터정보학회 논문지, 제12권, 제6호, 207-212쪽, 2007년 12월.
- [6] 신병호, 최이권, 이상범, 정준영, "온톨로지 기반의 프로세스 소프트웨어 개념모델 설계", 한국컴퓨터정보학회 논문지, 제13권, 제2호, 1-9쪽, 2008년 3월.
- [7] J. Trinkunas, O. Vasileas, "Building Ontologies from Relational Databases Using Reverse Engineering Methods", CompSysTech'07, 2007.
- [8] Z. Xu, X. Cao, Y. Dong, and W. Su, "Formal Approach and automated Tool for Translating ER Schemata into OWL Ontologies", Advances in Knowledge Discovery and Data Mining, LNAI 3056, Springer, pp.464-475, 2004.
- [9] M. Li, X. Du, S. Wang, "Learning Ontology from Relational Database", Proceedings of the 4th International conference on machine Learning and Cybernetics, pp.3410-3415, 2005.
- [10] Z. Xu, S. Zhang, Y. Dong, "Mapping between Relational Database Schema and OWL Ontology for Deep Annotation", International Conference on Web Intelligence(WI'06), 2006.
- [11] 이현실, "합성 온톨로지 기반의 한의학 처방 지식관리 시스템", 한국학술정보, 2006년.
- [12] 이호경, 송순원, "Opening XML", 구민사, 2002년.
- [13] 홍성용, "XML 원리와 응용", 한빛미디어, 2003년.
- [14] K. Masahide 지음, 황석영, 양해술 공역, "RDF/OWL 입문", 홍릉과학출판사, 2008년.
- [15] Protégé, <http://protege.stanford.edu>
- [16] Cyc, <http://cyc.com/>
- [17] D. Dou and P. LePendu, "Ontology-Based Integration of Relational Databases", SAC'06, pp. 461-466, 2006.
- [18] N. F. Noy, "Semantic Integration : A Survey of Ontology-Based Approaches", SIGMOD RECORD Vol 33, No 4, pp.65-70, 2004.
- [19] L. Khan, D. McLeod, E. Hovy, "Retrieval Effectiveness of an Ontology-based Model for Information Selection", VLDB Journal 13, pp. 71-85, 2004.
- [20] B. Schueler, S. Sizov, S. Staab, "Querying for Meta Knowledge", WWW'08, pp.625-634, 2008.

## 저자 소개



### 나 민 영

육군사관학교 졸업  
서울대학교 컴퓨터공학과 졸업  
서울대학교 대학원 컴퓨터공학과 졸업,  
M.S.

University of Florida,  
Dept. of Computer and  
Information Sciences,  
Ph.D.

미 IBM Watson 연구소 연구원  
Georgia Institute of Technology  
교환교수

현재: 육군사관학교 전자정보학과 교수



### 양 경 응

KAIST 졸업  
서울대학교 대학원 컴퓨터공학과  
졸업, M.S.

현재: 육군사관학교 전자정보학과 강사