

국방연구정보의 모바일 시맨틱 통합검색 시스템

유 동 희⁺ · 나 민 영^{††}

요 약

국방 분야의 연구정보들을 효과적으로 관리하기 위해서는 국방연구정보에 관한 메타데이터가 체계적으로 관리되어야 하며, 이를 기반으로 국방연구정보의 융합과 관리를 돕는 통합시스템이 구축되어야 한다. 이와 함께, 모바일 장비를 이용한 검색이 증가하고 있는 추세이므로 사용자들에게는 모바일 환경에서의 효과적인 통합 검색 서비스가 제공되어야 한다. 본 논문에서는 국방연구정보를 활용한 모바일 시맨틱 통합검색 시스템을 제안하고자 한다. 이를 위해, 더블린코어의 메타데이터와 주요 국방 관련 연구기관의 데이터베이스 정보를 분석하여 국방연구 온톨로지와 국방연구 규칙을 정의하였고, 정의된 온톨로지와 규칙이 국방연구정보의 시맨틱 통합검색에 활용되는 것을 보이기 위해 모바일 시맨틱 통합검색 시스템의 프로토타입을 구현하였다. 또한 모바일 환경에서의 시맨틱 통합검색을 지원하기 위한 트리플 기반 검색 서비스와 향후 활용 가능한 모바일 시맨틱 통합검색 서비스들을 제시하였다.

키워드 : 시맨틱 웹, 온톨로지, 국방연구정보, 모바일 시맨틱 검색, 모바일 장비

A Mobile Semantic Integrated Search System of National Defense Research Information

Yoo, Donghee⁺ · Ra, Minyoung^{††}

ABSTRACT

To effectively manage research information in the field of national defense, metadata about the information should be managed systematically, and an integrated system to help convergence and management of the information should be implemented based on the metadata. In addition, the system should provide the users with effective integrated search services in a mobile environment, because searching via the use of mobile devices is increasing. The objective of this paper is to suggest a MSISS (Mobile Semantic Integrated Search System), which satisfies the requirements for effective management of the national defense research information. Specifically, we defined national defense research ontologies and national defense research rules after analyzing the Dublin Core metadata and database information of the major military research institutions. We implemented a prototype system for MSISS to demonstrate the use of the ontologies and rules for semantic integrated searching of the military research information. We also presented a triple-based search service to support semantic integrated search in a mobile environment and suggested future mobile semantic integrated search services.

Keywords : Semantic Web, Ontology, National Defense Research Information, Mobile Semantic Search, Mobile Devices

1. 서 론

국방 분야에서는 정보 시스템이 도입된 이래 정보, 인사, 군수, 작전, 예산 등 전 기능에 걸쳐 정보화가 확대되고 있으며, 각 정보 시스템은 기획 문서, 연구 보고서, 기술 문서 등과 같은 군 관련 지식들을 체계적으로 관리하고 검색하는 기능을 제공하여 국방연구지식들의 활용도를 높여주고 있

다. 하지만, 현재 국방 관련 주요 기관별로 지식관리체계 시스템이 따로 운영되고 있기 때문에, 다음과 같은 문제들에 대한 개선의 필요성이 요구되고 있다[2]. 첫째, 메타데이터 관리에 관한 부분이다. 즉, 자료의 입력과 검색에는 메타데이터 항목이 활용되고는 있지만, 메타데이터 자체는 체계적으로 관리되지 않고 있다. 둘째, 정보 접근에 제한성이 발생한다. 그 이유는 특정 기관에서 생성된 연구지식의 경우 주로 해당 기관에만 등록되며, 타 기관에 함께 등록되는 비율은 상당히 낮아 분산된 지식관리체계에 저장된 연구지식들의 통합 검색이 제한된다. 셋째, 외부 기관의 자료를 등록할 경우 자료가 중복되어 저장되는 문제와 등록된 지식의 저작권과 관련된 문제 발생의 가능성도 언급되고 있다. 따라서

* 본 연구는 육군사관학교 화랑대 연구소의 2011년도 연구활동비 지원을 받아 연구되었음.

+ 정 회 원 : 육군사관학교 전자정보학과 조교수

†† 정 회 원 : 육군사관학교 전자정보학과 교수(교신저자)

논문접수 : 2011년 4월 12일

수정일 : 1차 2011년 6월 2일, 2차 2011년 7월 4일

심사완료 : 2011년 7월 4일

국방 분야에 관련된 다양한 연구지식들을 효과적으로 관리하기 위해서는 국방연구정보에 관한 메타데이터가 체계적으로 관리되어야 하고, 이를 기반으로 국방연구지식의 융합과 관리를 돕는 통합시스템에 관한 구축이 필요하다.

이와 같은 통합시스템이 구축되려면, 통합된 국방연구정보를 활용한 다양한 검색 서비스가 사용자에게 효과적으로 제공되어야 한다. 초창기 인터넷 검색이 주로 가정집 또는 사무실과 같은 정적인 공간에서 이루어진 검색이라면, 오늘날은 기존의 검색 공간 이외에 추가적으로 언제 어디서나 인터넷에 접속할 수 있는 동적 공간에서의 검색 즉 모바일 검색도 함께 증가하고 있는 추세이다[13, 22, 23]. 따라서 PDA, 휴대폰, 스마트폰 등과 같은 모바일 장비를 이용하는 사용자들이 원하는 정보를 정확하게 검색해 주는 검색 서비스의 중요성이 높아지고 있다.

본 연구에서는 시맨틱 웹 기술을 이용하여 분산 환경에 존재하는 국방연구정보들을 효과적으로 통합하고, 이를 활용하여 모바일 환경에서의 시맨틱 통합검색을 제공하는 모바일 시맨틱 통합검색 시스템을 제안하고자 한다. 여기에서 시맨틱 웹(Semantic Web)이란 웹 상의 정보를 컴퓨터가 이해할 수 있는 형태로 변환하여 인간의 개입 없이 컴퓨터가 자동으로 웹의 정보를 검색하고 관리하는 차세대 웹을 말하며[7], 시맨틱 검색이란 검색엔진이 검색 정보의 의미를 이해하여 사용자에게 보다 의미 있는 검색결과를 제공하는 검색 방법을 뜻한다. 이와 같이 시맨틱 웹 환경에서의 통합검색 시스템을 구현하기 위해서는 두 가지 요구사항이 필요하다[3]. 첫 번째는 통합된 연구정보가 컴퓨터가 이해할 수 있는 표준화된 언어로 표현되어야 한다. 두 번째는 표현된 정보에서 사용된 여러 개념들과 개념들 간의 관계의 의미(semantics)를 명확하게 정의한 온톨로지(Ontology)[10]가 정의되어 있어야 한다. 즉, 국방연구정보에 관한 다양한 개념들이 온톨로지에 정의되어 있고, 기관별 국방연구정보에 관한 메타데이터가 온톨로지에 정의된 개념을 활용하여 표준화된 형태로 표현될 경우, 시스템은 각 기관의 연구정보를 이해하고 추론하여 효과적으로 국방연구정보를 통합할 수 있고 다양한 시맨틱 통합검색에도 이용할 수 있다.

다음으로 모바일 장비에서 통합된 국방연구정보를 이용하여 모바일 시맨틱 통합검색을 수행하는 방법은 국방연구정보에 접근하는 방법에 따라 두 가지로 구분할 수 있다. 첫 번째는 모바일 장비는 검색 화면 기능만 지원하고 나머지 부분은 외부 통합 시스템의 통합검색 기능을 활용하는 방법이다. 두 번째는 모바일 장비가 포함하고 있는 로컬 저장소에 통합된 국방연구정보들을 저장한 후, 모바일 장비에서 직접 자신의 로컬 저장소에 시맨틱 통합검색을 실행하는 방법이다. 전자는 무선 통신망 접속이 가능한 곳에서만 검색 서비스가 지원되며, 후자는 특수 상황에 의해 무선 통신망 접속이 불가능하더라도 모바일의 로컬 저장소를 이용한 검색 서비스가 지원되는 방법이다. 본 연구에서는 군의 특수 상황을 반영하기에 적합한 후자의 접근 방법을 활용하고자 한다. 즉, 현재 군에서는 휴전선 전방과 같이 무선 통

신망의 접속이 불안정한 지역이 존재하며, 전쟁과 같은 비상 사태 때는 무선 통신망의 사용이 제한되는 특수 상황이 발생한다. 따라서 후자의 접근 방법을 통해 무선 통신망의 접속이 원활할 경우 서비스를 제공하는 외부 통합시스템으로부터 최신의 국방연구정보를 갱신하여 활용하고, 특수 상황에 의해 무선 통신망 접속이 불가능할 경우 모바일 장비의 로컬 저장소를 활용한 시맨틱 통합검색 방법들을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 모바일 시맨틱 통합검색 시스템의 핵심 기술인 시맨틱 웹 기술과 모바일 환경에서 시맨틱 웹 기술이 사용된 시스템들에 관한 문헌 연구를 실시하였다. 3장에서는 모바일 시맨틱 통합검색 시스템 구현에 사용된 시스템 아키텍처와 이를 구성하고 있는 구성요소들을 설명하였고, 4장에서는 통합된 국방연구정보들이 모바일 시맨틱 통합검색에 어떻게 활용되는가를 보여주었다. 또한 구현 기술에 대한 요약과 함께 기존 검색 방법과 본 논문에서 제안한 검색 방법의 성능을 평가하는 실험을 실시하였다. 마지막으로 5장에서는 연구 요약과 함께 향후 모바일 시맨틱 통합검색 시스템에 필요한 사항들을 언급하였다.

2. 관련 기술 조사

2.1 시맨틱 웹 기술

본 절에서는 모바일 시맨틱 통합검색 시스템 구현에 필요한 시맨틱 웹 기술을 조사하였다. 시맨틱 웹에서는 트리플(triple) 구조로 웹 상의 자원들의 의미 집합인 온톨로지와 온톨로지를 기반으로 정의되는 정보들을 표현한다. 트리플 구조는 <주어부(subject)-서술부(predicate)-목적부(object)>로 구성되며, 하나의 트리플 구조는 하나의 서술문을 나타낸다. 국방연구 온톨로지를 구축할 경우 트리플 구조를 기본으로 연구정보에 관한 개념들과 그들의 관계를 정의할 수 있는데, 대표적인 온톨로지 표현 언어에는 Web Ontology Language(OWL)[5, 18]가 있다. 예를 들어, OWL을 이용하여 보고서라는 개념을 클래스로 표현할 경우, <보고서(Report)-타입이다(rdf:type)-클래스(Class)>의 트리플 구조로 표현된다. 이러한 온톨로지에 국방 관련 전문 지식을 사용자가 정의한 규칙 형태로 표현하기 위해서는 Semantic Web Rule Language(SWRL)를 사용하면 된다. SWRL은 유사 Horn 규칙을 수용할 수 있도록 OWL을 확장시킨 언어로서 OWL에 정의된 클래스와 속성을 이용하여 규칙을 정의한다[11, 12]. 국방연구 온톨로지가 구축되면 Resource Description Framework(RDF)[14, 17]를 이용하여 다양한 시맨틱 정보를 트리플 구조로 표현할 수 있다. 예를 들어 RDF를 사용하면 '보고서 A는 연구보고서 클래스의 유형에 속하며, 홍길동에 의해 작성되었다'라는 정보를 두 개의 트리플 구조인 <보고서A-유형에속하다 연구보고서>와 <보고서A-작성되다 홍길동>으로 표현할 수 있다. 시맨틱 웹에서 기계가 정보의 의미를 이해한다는 것은 추론 엔진이 국

방연구 온톨로지와 시맨틱 정보의 논리 구조를 이해하여 새로운 정보를 추론하는 것을 의미한다. 이와 같이, 국방연구 온톨로지, 시맨틱 정보, 그리고 추론된 정보가 지식 베이스(knowledge base)에 저장되면, 트리플 구조에 대한 질의어인 SPARQL[19]을 활용하여 다양한 시맨틱 검색이 가능하게 된다.

2.2 모바일 환경에서의 시맨틱 웹 기술

최근 모바일 환경에서 RDF로 정의된 시맨틱 정보들을 모바일 장비에서 활용할 수 있게 도와주는 모바일 시맨틱 웹 프레임워크에 관한 연구들이 진행되고 있다. Mobile RDF[29]는 J2ME 환경에서 운영되는 시맨틱 웹 프레임워크로, RDF 처리에 필요한 기본 기능인 RDF 데이터 읽기 기능과 읽은 데이터들을 트리플 형태로 파싱하는 기능을 제공한다. 온톨로지 차원에서는 RDF Schema에서 제공되는 개념인 rdfs:List와 rdfs:Container로 표현된 정보의 처리를 지원한다. μ Jena(Micro Jena)[28]는 기존 시맨틱 웹 프레임워크로 널리 알려진 Jena[27]를 J2ME 환경에 이식(porting)하여 구현된 모바일 시맨틱 프레임워크로 온톨로지 매칭(matching)과 추론 기능을 제공하고 있다. 현재 이 시스템은 프로토타입 상태이며 N-Triple 형태로 작성된 RDF 데이터의 처리를 지원한다. Androjena[24] 또한 구글 안드로이드 환경에 Jena를 이식하여 사용하고 있으며 RDF 데이터 처리에 필요한 기본 기능을 제공한다. 현재 파일 형식의 RDF 저장과 SPARQL 질의 처리 기능이 개발 중에 있다.

시맨틱 웹 기술을 모바일 장비에 적용할 때 고려되어야 할 다양한 연구 주제들이 있다. i-MoCo[21]에서는 대용량 시맨틱 정보의 저장과 질의의 중요성을 언급하며, 애플리케이션

구현을 통해 대용량 컨퍼런스 정보의 트리플 저장과 질의 처리의 가능성을 보여 주었다. MobiSem[20]에서는 실시간으로 변경되는 시맨틱 정보를 모바일 장비의 로컬 저장소에 효과적으로 저장하는 방법을 제안하였으며, DBpedia Mobile[6]에서는 위치기반정보를 이용한 다양한 시맨틱 검색 방법들을 소개하고 있다.

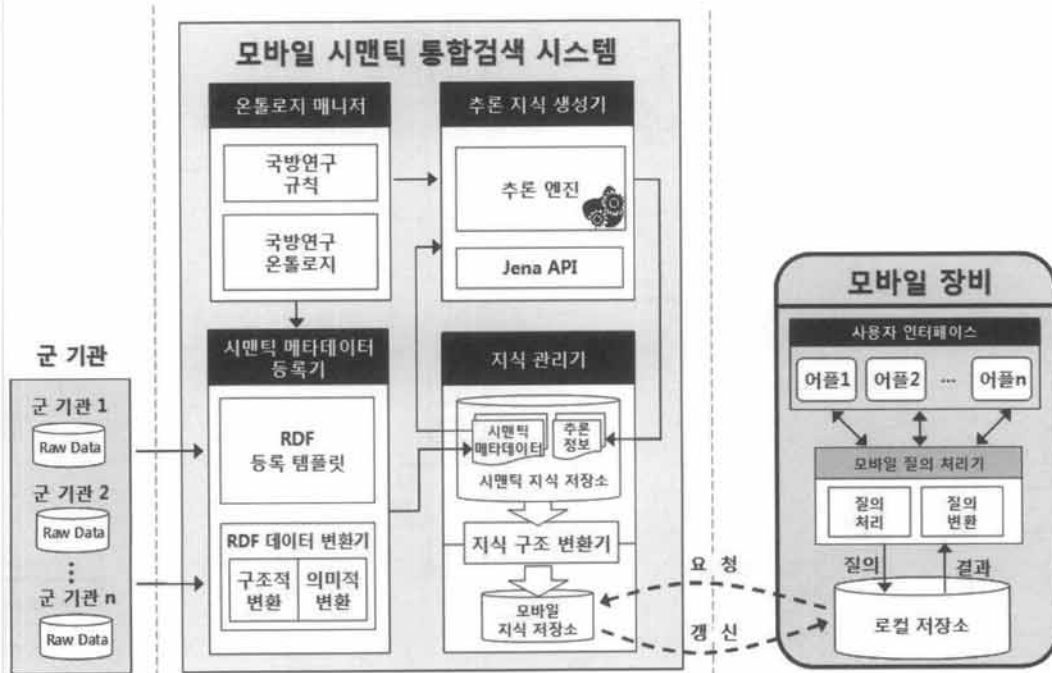
현재 모바일 환경에 시맨틱 웹 기술을 적용하는 시도는 초기 연구 단계라 할 수 있으며, 시맨틱 웹 기술을 이용하여 분산된 국방연구정보를 통합하고 통합된 국방연구정보를 모바일 환경에서 이용하는 연구는 아직까지 이루어지지 않고 있다. 이에 본 논문에서는 국방연구정보를 활용한 모바일 시맨틱 통합검색 시스템 구축에 관한 연구를 진행하였다.

3. 모바일 시맨틱 통합검색 시스템

3.1 시스템 아키텍처

본 절에서는 시맨틱 웹 기술을 기반으로 구현된 모바일 시맨틱 통합검색 시스템에 관한 시스템 아키텍처를 설명하고자 한다. (그림 1)에서 보는 바와 같이, 시스템 아키텍처는 1) 온톨로지 매니저, 2) 시맨틱 메타데이터 등록기, 3) 추론 지식 생성기, 4) 지식 관리기 그리고 5) 모바일 장비로 구성된다. 각각의 주요 기능을 요약하면 다음과 같다.

- 온톨로지 매니저: 국방연구 온톨로지와 국방연구 규칙에 관한 생성 및 관리 역할을 담당한다. 국방연구 온톨로지에는 국방연구정보에 대한 메타데이터와 정보 통합에 필요한 용어들이 정의되어 있으며, 국방연구 규칙에는 도메



(그림 1) 시스템 아키텍처

인 전문가가 정의한 규칙 정보가 정의되어 있다.

- 시맨틱 메타데이터 등록기: 업무 담당자로부터 입력된 국방연구정보를 자동으로 시맨틱 메타데이터 형태로 변환하는 RDF 등록 템플릿과 기존의 국방연구정보를 RDF 형식의 시맨틱 메타데이터로 변환해주는 RDF 데이터 변환기로 구성된다.
- 추론 지식 생성기: 국방연구 온톨로지와 국방연구 규칙을 이용하여 시맨틱 메타데이터로부터 새로운 지식 정보를 추론하고, 추론된 결과를 지식 관리기로 전달하는 기능을 제공한다.
- 지식 관리기: 시맨틱 메타데이터와 추론 정보를 시맨틱 지식 저장소에 저장하는 기능과 저장된 정보를 모바일 장비에서 운용 가능한 구조로 변환한 후 모바일 지식 저장소에 저장하는 지식 구조 변환기를 포함하고 있다.
- 모바일 장비: 모바일 시맨틱 통합검색을 지원해 주는 어플리케이션들과 어플리케이션으로부터 입력된 질의를 SQL 또는 SPARQL 형태로 변환하는 모바일 질의 처리기 그리고 국방연구정보가 저장된 로컬 저장소로 구성된다.

3.2 온톨로지 매니저

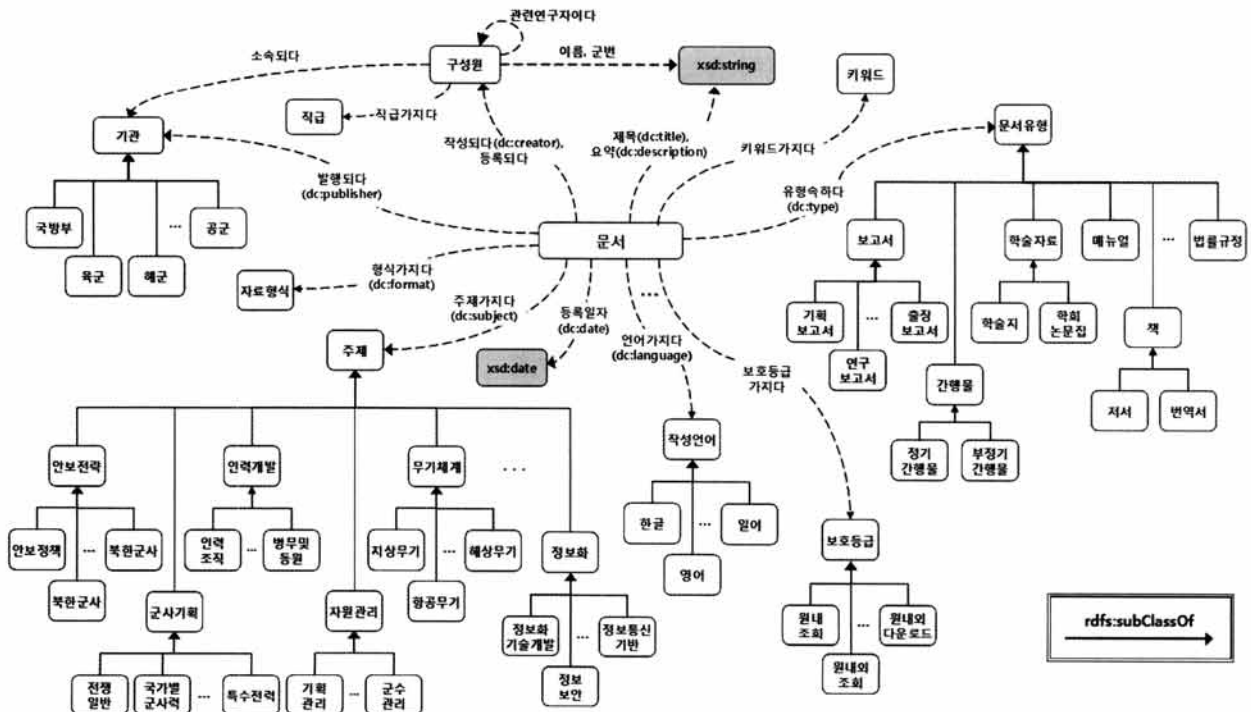
온톨로지 매니저에서 관리하는 국방연구 온톨로지는 연구 정보에 관한 다양한 개념들과 그들의 관계를 정의한 온톨로지를 뜻한다. 여기에서 국방연구 온톨로지를 구축하는 이유는 여러 군 연구기관에 분산되어 등록된 연구정보에 관한 메타데이터를 기계가 이해할 수 있는 형태로 변형한 뒤 통합 관리 및 시맨틱 통합검색에 활용할 수 있기 때문이다. 메타데이터의 추출은 국제 표준으로 널리 알려진 더블린코

어[26]의 요소를 선별하여 추출하였다. 본 연구에서는 <표 1>에서 보는 바와 같이, 15개의 더블린코어 요소 중 필수 항목으로 10개 요소를 선정하였으며, 부가 항목으로서는 10개의 요소를 선정하였다. 여기서 부가 항목이란 효과적인 데이터 관리를 위해 부가적으로 필요한 메타데이터를 의미한다.

<표 1> 추출된 메타데이터 정보

항목 특징	메타데이터
필수 항목	제목, 저자, 공동저자, 주제, 자료유형, 발행처, 언어, 발행일, 자료형식, 요약
불필요 항목	식별자, 출전, 관련자료, 내용범위, 이용조건
부가 항목	등록자, 등록일, 보존연한, 연구수행기관, 연구기간, 발주기관, 키워드, 공개범위, 보호등급, 첨부파일

추가적으로, 국방연구정보 통합에 필요한 개념들은 4개의 주요 군 연구기관의 데이터베이스에 저장되어있는 연구정보를 참조하여 정의하였다. 먼저 각 기관 별로 연구정보를 저장할 때 사용되는 개념들을 조사한 후, 동일한 특성을 지닌 개념들을 클래스로 구분하였고 계층 관계로 표현되는 클래스의 경우 클래스 계층도를 구성하였다. 예를 들어, 연구정보가 속할 주제에 대한 클래스는 4개 연구기관에서 사용되고 있는 '주제' 부분과 관련된 개념들을 조사한 뒤, 주제에 관한 클래스 계층도를 구성하였다. 효과적인 연구정보통합을 지원하기 위해 각 클래스들은 4개의 연구기관에서 사용된 개념들을 모두 포함할 수 있게 설계하였다. 이렇게 조사



(그림 2) 국방연구 온톨로지

된 메타데이터들과 개념들을 바탕으로 Protégé[30]를 이용하여 국방연구 온톨로지를 OWL로 구축하였으며, 더블링크어의 메타데이터들의 경우 Protégé에 import하여 정의하였다. 온톨로지 구축에 사용된 방법론으로는 국방 온톨로지 추출에 활용된바 있는 6단계 상향식 방법론[1]을 참고하였다. (그림 2)는 구축된 국방 온톨로지의 주요 개념들과 그들의 관계를 보여준다.

(그림 2)의 모서리가 둥근 직사각형은 국방연구 온톨로지의 클래스를 나타낸다. 주요 클래스에는 문서, 기관, 구성원, 주제, 문서유형, 작성언어, 보호등급 등이 있다. 클래스 간의 계층 관계는 실선 화살표인 rdfs:subClassOf 관계로 표현된다. 예를 들어, 클래스 주제의 경우 안보전략, 군사기획, 인력개발, 정보화 등의 하위 클래스를 포함하고 있다. 점선 화살표는 국방연구 온톨로지의 속성 정보를 의미하며, 크게 오브젝트속성(owl:ObjectProperty)과 데이터형속성(owl:DatatypeProperty)으로 구분된다. 오브젝트속성은 클래스간의 관계를 표현할 때 사용되며, 데이터형속성을 표현하기 위하여 XML 스키마에서 정의된 클래스인 xsd:string과 xsd:date를 이용하였다. 여기에서 속성의 이름 옆에 있는 접두어가 dc인 용어들은 더블링크어에서 추출한 메타데이터를 나타낸다.

온톨로지 매니저에는 국방연구 통합검색에 필요한 전문 지식이 규칙 형태로 정의된 국방연구 규칙을 포함하고 있다. 현재 국방연구에 관한 지식 정의와 함께 증가되는 지식 정보들을 효과적을 생성하고 관리하는 기술의 중요성도 강조되고 있다. 왜냐하면, 모바일 시맨틱 통합검색에 필요한 새로운 서비스에 관한 요구사항이 발생하게 되면, 해당 서비스를 지원하기 위한 새로운 전문 지식이 필요하기 때문이다. 온톨로지 매니저에서는 이러한 사항을 반영하여 시스템 관리자가 직접 관련 규칙들을 정의할 수 있는 기능을 제공하며, 이와 같은 국방연구 규칙들은 앞서 구현한 국방연구 온톨로지의 개념과 속성을 활용하여 SWRL의 형태로 정의된다. 여기에서 국방연구 규칙의 형태는 전제(antecedent) → 결과(consequent)로 표현되며, 전제 부분은 국방연구 온톨로지의 클래스와 속성으로 표현되는 하나 이상의 요소(atom)들의 논리곱을 구성되며, 결과 부분은 하나의 요소로 구성된다. 예를 들어, '문서 A를 작성한 저자가 여러 명 있다면 그들의 관계는 관련 연구자이다'라는 국방연구 규칙을 SWRL로 표현하면 다음과 같다.

$$\text{Document}(?a) \wedge \text{dc:creator}(?a, ?b) \wedge \text{dc:creator}(?a, ?c) \wedge \text{swrlb:notEqual}(?b, ?c) \rightarrow \text{relatedResearcher}(?b, ?c)$$

3.3 시맨틱 메타데이터 등록기

모바일 시맨틱 통합검색 시스템에서는 국방연구 온톨로지를 기반으로 표현된 RDF 형태의 국방연구정보가 필요하다. 본 연구에서는 이와 같은 형태의 국방연구정보를 생성하기 위해 RDF 등록 템플릿과 RDF 데이터 변환기를 구현하였다.

RDF 등록 템플릿은 연구 보고서에 관한 메타데이터 정보가 등록 템플릿 화면을 통해 입력되면 입력된 정보를 RDF 형태로 자동 변환하는 기능을 제공한다. 본 연구에서는 이렇게 RDF 형태로 변환된 메타데이터를 가리켜 시맨틱 메타데이터라 한다. RDF 등록 템플릿의 화면은 (그림 3)과 같으며, 연구 보고서에 관한 메타데이터 항목들은 국방연구 온톨로지에 정의된 개념들을 참조하여 구축하였다. 입력된 연구 보고서 정보에는 보고서의 제목, 등록자, 등록일자, 저자, 공동저자, 공개구분, 연구주제, 보고서 유형 등이 있다. 이외 '여기' 부분을 선택하여 보고서의 보존연한, 연구시작 및 종료일 등과 같은 세부 정보를 입력할 수 있다. 화면 하단에 있는 등록 버튼을 선택하면 보고서 정보의 입력이 완료된다. 이와 같은 과정을 통해 생성된 시맨틱 메타데이터는 지식 관리기의 시맨틱 지식 저장소에 전달된다. 여기에서 시맨틱 웹 기술에 익숙하지 않은 등록자라도 연구 보고서에 관한 정보를 손쉽게 등록할 수 있는데 이는 등록자의 사용 편의성을 높이기 위해 RDF 등록 템플릿을 기존의 메타데이터 등록 시스템들과 유사한 형태로 설계하였기 때문이다.

(그림 3) RDF 등록 템플릿 화면

RDF 데이터 변환기는 기존 연구기관들에 저장된 연구정보들을 시맨틱 메타데이터로 변환하는 작업을 도와준다. 기존 연구정보들은 관계형 데이터베이스(RDB)에 저장되어 있으므로, 시맨틱 메타데이터 형태의 변환을 위해서는 매핑(mapping) 작업이 필요하다. 매핑 방법은 매핑 과정을 기준으로 크게 정적인 방법과 동적인 방법으로 분류된다. 정적인 방법은 매핑 도구를 이용하여 RDB로부터 시맨틱 메타데이터를 물리적으로 생성하는 것이며, 동적인 방법은 특정 질의에 대한 응답으로 동적인 매핑을 하는 방법이다. 본 연구에서는 최근 데이터를 즉각 반영하지 못하는 단점은 있지만 질의를 포함한 추가적인 처리나 분석이 필요하지 않는 정적인 방법을 이용하였다. 정적인 방법을 지원하기 위한

매핑 도구로 RDBtoOnto[8]를 이용하였다. RDBtoOnto에서는 국방연구 온톨로지와 특정 연구기관의 데이터베이스 정보를 각각 불러들인 후, 사용자가 손쉽게 매핑할 수 있는 인터페이스를 제공한다. 이를 통해 RDB의 연구정보들은 시맨틱 메타데이터 형태로 구조가 변경되며, 각 개념들이 국방연구 온톨로지를 참조하기 때문에 의미적 변환 또한 동시에 이루어지게 된다. 그러나 매핑 도구가 RDB 안의 모든 정보들을 시맨틱 메타데이터로 한 번에 완벽하게 변환 시키는 것은 불가능하다. 따라서 매핑 도구가 지원해 주지 않는 변환 부분에 대해서는 추가적으로 XSLT[32]로 정의된 변환 규칙 프로그램을 작성하여, RDF 등록 템플릿을 통해 생성된 시맨틱 메타데이터와 RDF 데이터 변환기를 통해 변환된 시맨틱 메타데이터간의 상호운용을 가능하게 하였다.

3.4 추론 지식 생성기

추론 지식 생성기는 시맨틱 지식 저장소에 있는 국방연구 정보에 관한 시맨틱 메타데이터를 읽어 들인 뒤 국방연구 온톨로지와 국방연구 규칙을 참조하여 새로운 지식 정보를 추론하는 작업을 진행한다. 본 연구에서는 OWL과 SWRL의 추론을 지원하는 Bossam 추론 엔진[25]을 이용하였다. 추론 엔진으로부터 추론된 새로운 지식 정보는 시맨틱 지식 저장소의 추론 정보에 저장된다. 예를 들어, “문서 A의 저자는 홍길동과 장보고이다”라는 국방연구정보가 있을 경우, 추론 엔진은 국방연구 규칙에 정의된 relatedResearcher 속성에 관한 규칙을 참조하여, “홍길동과 장보고는 관련연구자이다”라는 새로운 연구정보를 추론하며, 추론된 정보는 자동으로 시맨틱 지식 저장소의 추론 정보 영역에 저장된다. 추론 지식 생성기와 지식 관리기간의 상호 작용 부분은 Jena 프레임워크를 활용하여 구현하였다.

3.5 지식 관리기

지식 관리기가 시맨틱 메타데이터 등록기와 추론 지식 생성기로부터 시맨틱 메타데이터와 추론 정보를 전달 받으면 각각의 정보들을 자동으로 시맨틱 지식 저장소에 저장된다. 여기에서 시맨틱 지식 저장소는 Jena의 하위 컴포넌트인 SDB[31]를 이용하여 구현하였다. SDB는 1) 트리플 형식의 RDF 정보들을 MySQL과 같은 RDB에 저장하는 기능과 2) SPARQL 인터페이스를 이용하여 MySQL에 저장된 트리플 정보를 검색하는 기능을 제공한다. 따라서 사용자들은 SDB를 이용하여 자신이 찾고자 하는 국방연구정보에 관한 조건들을 트리플 구조로 질의할 수 있다.

서론에서 언급한 바와 같이, 모바일 환경에서의 시맨틱 통합검색을 지원하기 위하여, 모바일 장비의 로컬 저장소에 시맨틱 지식 저장소의 정보들을 저장하여 활용하고자 한다. 이를 위해 우선적으로 MySQL을 기반으로 구현된 시맨틱 지식 저장소에 저장된 국방연구정보들은 모바일 지식 저장소인 SQLite3에 저장되어야 한다. 하지만 두 데이터베이스 시스템이 서로 다르기 때문에 이 둘의 호환성을 돕는 지식 구조 변환기가 필요하다. 본 연구에서 구현한 지식 구조 변

환기는 3단계의 변환 과정을 통해 기존 시맨틱 지식 저장소의 정보들을 모바일 지식 저장소에 저장하는 기능을 제공한다. 1단계 SQL 호환 기능(SQL compatibility function)에서는 시맨틱 지식 저장소의 정보들을 모바일 지식 저장소에 질의 가능한 SQL 형태로 추출한다. 2단계 SQL 변환 기능에서는 추출된 SQL의 일부 정보를 변경하는 기능으로, 테이블 속성에 정의된 데이터유형 부분을 모바일 지식 저장소에 사용 가능한 데이터유형으로 변경하는 작업 등이 여기에 포함된다. 마지막으로 3단계 SQL 실행 기능에서는 변환된 SQL 명령어를 실행하여 모바일 지식 저장소에 국방연구정보를 저장한다. 변환 과정에서 발생할 오류를 최소화하기 위하여 각 단계들은 시스템 관리자의 개입을 통해 진행된다. 이와 같은 변환이 완료되면 모바일 지식 저장소의 정보들은 모바일 장비의 로컬 저장소로 전송된다.

4. 모바일 시맨틱 통합검색

4.1 모바일 시맨틱 통합검색 서비스

모바일 장비에서 국방연구정보에 관한 시맨틱 통합검색 서비스를 지원하기 위해서는 사용자 인터페이스, 모바일 질의 처리기, 로컬 저장소가 필요하다(그림 1) 참고). 사용자 인터페이스는 모바일 시맨틱 통합검색 서비스를 제공하는 어플리케이션의 화면을 나타낸다. 사용자 인터페이스를 통해 사용자가 찾고자 하는 국방연구정보와 관련된 질의어가 입력되거나 선택되면, 질의 정보들은 모바일 질의 처리기를 통해 SQL 형식 또는 SPARQL 형식으로 변환되어 로컬 저장소에 전달된다. SQL 질의를 통해 질의 조건을 만족하는 결과 값들을 로컬 저장소의 테이블들로부터 검색할 수 있다. 그리고 SPARQL은 사용자가 찾고자 하는 국방연구정보에 관한 조건들을 트리플 구조로 질의할 수 있는 언어이며, 현재 모바일 장비의 로컬 저장소에 저장된 국방연구정보가 트리플 구조로 저장되어 있기 때문에 SPARQL을 이용한 검색 처리의 기반 환경을 제공해 준다.

(그림 4)는 SQL을 이용한 모바일 시맨틱 통합검색 서비스 중 하나인 트리플 기반 검색 서비스를 보여준다. 트리플 기반 검색 서비스란 로컬 저장소에 트리플 형태로 저장되어 있는 온톨로지와 시맨틱 메타데이터 정보들을 트리플 구조에 따라 설계된 검색 화면을 이용하여 관련 정보를 찾는 서비스를 말한다. 사용자는 트리플 기반 검색 서비스를 통해 국방연구정보에 관한 검색 조건을 트리플 구조로 질의할 수 있는데, (그림 4)는 트리플 구조 중 Subject 영역에 해당되는 정보를 찾기 위해 Property 영역과 Object 영역에 관련 검색어를 입력한 예를 보여준다. 여기에서 Property 영역에 해당되는 속성 정보들은 국방연구 온톨로지에 정의된 속성 정보이다(그림 4)의 b) 참고). 검색어 입력 후 Search 버튼을 누르면, Subject에 해당되는 각 연구 기관별 정보가 통합되어 검색된 것을 확인할 수 있다(그림 4)의 c) 참고). 이 결과는 연구보고서 클래스의 인스턴스에 해당되는 보고서 정보를 나타낸다.

련된 질의 정보들을 항목별로 구분하여 제시한다. 이와 같이 국방연구 온톨로지를 이용한 자동 추천기능을 통하여 사용자는 보다 정확하게 자신이 찾고자 하는 국방연구정보들을 검색할 수 있게 된다.

사용자가 서비스에 접근하는 가장 쉬운 방법 중 하나는 검색 단어가 아닌 자연어로 질의를 입력하는 것이다. (그림 5)의 b)와 같이, 의미기반 검색 서비스는 사용자가 자연어 질의를 입력한 후 변환 버튼을 누르면, 시스템에서는 국방연구 온톨로지와 로컬 저장소 정보를 참조하여 입력된 질의를 트리플 형태의 질의 조합으로 변환 시킨다. 만약 변환된 정보 중 중복된 의미를 지닌 정보가 발생할 경우 해당 정보의 의미를 구별할 수 있는 추가 선택 기준들을 사용자에게 제공하여, 사용자 선택에 따른 보다 정확한 정보 검색을 가능하게 해 준다. Lopez 등의 연구[16]에서 시도된 바 있듯이, 향후 의미기반 검색 서비스가 모바일 환경에 적용될 경우 사용자에게 좀 더 편리하고 만족도 높은 검색결과가 제공될 것으로 예상된다.

마지막으로 모바일 애플리케이션에서 제공되는 정형화된 검색 서비스 이외에, 사용자가 직접 SPARQL 질의를 작성하여 국방연구정보를 찾길 원한다면 (그림 5)의 c)와 같은 SPARQL 검색 서비스가 활용될 수 있다. 이 서비스는 사용자가 사전에 SPARQL 문법에 관한 지식을 학습해야 한다는 제약사항이 있지만, 사용자의 기호가 반영된 맞춤 검색(customized search)이 가능하다는 장점이 있다. Bizer 등의 연구[6]에서 활용된 것처럼 향후 모바일 환경에 시맨틱 웹 기술의 적용이 확산되면 그 활용도 또한 증가될 것으로 기대된다.

4.2 구현 기술 요약 및 성능 평가

지금까지 살펴본 국방연구정보를 활용한 모바일 시맨틱 통합검색 시스템은 현재 사용 가능한 시맨틱 웹 기술들과 오픈 소스 시스템들을 활용하여 그 프로토타입을 구현하였다. 구현에 사용된 시스템들을 요약하면 <표 2>와 같다.

<표 2> 시스템 구현에 사용된 기술

구성 요소	관련 기술	
모바일 장비	Google Android 2.2 platform, SQLite3, Java, Android Virtual Device	
시맨틱 국방연구 통합검색 시스템	온톨로지 매니저	RDF, OWL, SWRL, Protégé 4.0.2
	RDF 등록 템플릿	HTML, JSP, JavaScript, Java
	RDF 데이터 변환기	RDB2Onto, XSLT, Xalan-java processor, Java
	추론 지식 생성기	Bossam engine, Jena API, Java
	지식 관리기	SDB, Java, JDBC, MySQL, SQLite3, phpMyAdmin, SQLite manager
웹 환경 지원	Tomcat Server	

<표 3>은 프로토타입 시스템에서 사용된 데이터들의 트리플 정보를 온톨로지 부분과 시맨틱 메타데이터 부분으로 구분하고 있다. 온톨로지 부분은 국방연구 온톨로지와 국방연구 규칙을 구성하는 구성요소 별로 트리플 정보를 보여주고 있으며, 시맨틱 메타데이터 부분에서는 RDF 형태로 변환된 연구 보고서에 관한 트리플 정보를 설명하고 있다. 현재 모바일 장비의 로컬 저장소에는 7.2MB의 트리플 정보가 저장되어 있다.

<표 3> 시스템에서 사용된 트리플 정보

문서	구성요소	구성요소의 수	트리플 수
국방연구 온톨로지	클래스	81	213
	오브젝트 속성	22	66
	데이터형 속성	12	36
	개체	45	90
국방연구 규칙	규칙	15	326
시맨틱 메타데이터	연구 보고서	256	7822

연구 보고서 검색에 대한 검색 효과를 분석하기 위해 기존의 연구 보고서 검색 방법과 본 연구에서 제안한 트리플 기반 검색 방법을 비교하는 실험을 실시하였다. 현재 주요 연구기관들의 검색 시스템은 분산 환경에 있는 연구 보고서에 대한 통합 검색이 제한되기 때문에, 4개의 연구기관들 중 한 곳의 검색 시스템을 대표로 선택하여 비교 실험을 진행하였다. 실험에는 총 87개의 연구 보고서가 사용되었으며, 특정 검색어에 관한 정확률(precision)과 재현율(recall)을 평가하였다. 정확률은 검색 결과 중 검색어와 적합한 보고서를 포함하는 비율이며, 재현율은 전체의 적합한 보고서 중 적합하게 검색된 보고서의 비율을 나타낸다. <표 4>는 특정 검색어에 대한 각 검색 방법의 실험 결과를 보여준다. 실험 결과 트리플 기반 검색 방법에서 재현율이 높게 측정되었다. 특히 기존 검색 방법에서 검색되지 않은 특수전력과 같은 검색어는 국방연구 온톨로지에 관련 개념이 정의되어 있기 때문에 트리플 기반 검색 방법에서 찾을 수 있었다. 실험 결과 트리플 기반 검색 방법이 의미적으로 연관된 정보 검색에 있어서는 높은 성능을 보였지만, 사용자가 온톨로지의 구조를 사전에 알고 있어야 하고 검색 방법이 기존 검색 방법에 비해 어렵다는 점이 분석되었다.

<표 4> 기존 검색 방법과 트리플 기반 검색 방법 비교

검색어	기존 검색 방법		트리플 기반 검색 방법	
	정확률	재현율	정확률	재현율
안보전략	80.00%	44.44%	92.31%	66.67%
무기체계	100.00%	66.67%	100.00%	83.33%
학술자료	75.00%	42.86%	83.33%	71.43%
연구보고서	85.71%	54.55%	90.00%	81.82%
특수전력	0.00%	0.00%	100.00%	75.00%

5. 결 론

본 논문에서는 시맨틱 웹 기술을 이용하여 주요 국방 관련 연구기관에 분산되어 관리되고 있는 국방연구정보를 통합한 후, 통합된 국방연구정보를 활용한 모바일 시맨틱 통합검색 시스템을 제안하였다. 이를 위해, 더블링크어의 메타데이터와 4개의 주요 연구기관의 데이터베이스 정보를 분석하여 국방연구정보의 통합검색에 필요한 국방연구 온톨로지 및 국방연구 규칙을 구축하였고, 이들을 이용한 모바일 시맨틱 통합검색 시스템을 구현하였다. 또한 구현된 시스템에서는 모바일 환경에서의 시맨틱 통합검색을 지원하기 위한 트리플 기반 검색 서비스를 구현하였으며, 향후 SPARQL 질의엔진이 개발되면 활용 가능한 서비스들을 제시하였다. 이와 같은 과정들을 통하여 시맨틱 웹 기술이 국방연구정보의 통합과 모바일 시맨틱 통합검색에 얼마나 유용하게 적용되는가를 확인할 수 있었다.

그러나 현재의 모바일 시맨틱 통합검색 시스템에는 다음과 같은 추가적인 연구가 필요하다. 첫째, 앞서 언급한 바와 같이 모바일 장비의 로컬 저장소에 효과적인 SPARQL 질의를 지원해 주는 질의엔진이 개발되어야 보다 다양한 모바일 시맨틱 통합검색 서비스가 제공될 것이다. 둘째, 모바일 장비에서 제한된 용량의 로컬 저장소를 사용하고 있는데, 향후 트리플 정보가 증가할 경우 모바일 장비에서는 이러한 대용량의 트리플 정보를 효과적으로 처리할 수 있는 기능이 요구되며, 이와 동시에 모바일 시맨틱 통합검색 시스템 내에서 대용량 RDF 정보들을 보다 효율적으로 처리하여 검색 성능을 높일 수 있는 연구가 추가적으로 필요하다. 셋째, 현재 시스템에서는 모바일 장비에 어플리케이션이 설치되면 어플리케이션에 사용될 국방연구정보가 모바일 장비의 로컬 저장소에 함께 저장되도록 설계되었다. 따라서 새로운 국방연구정보가 통합시스템의 모바일 지식 저장소에 등록될 경우 실시간으로 변경된 내용을 모바일 장비에 전달해 주는 방법과 모바일 지식 저장소를 분산 환경에서 운용하여 여러 모바일 장비에 효과적으로 변경된 내용을 전달하는 방법에 관한 연구가 필요하다.

참 고 문 헌

- [1] 나민영, 양경용, "6단계 상향식 방법에 의한 국방 온톨로지 추출," 한국컴퓨터정보학회, 제14권 제6호, pp.17-26, 2009.
- [2] 나민영, 유동희, "정보융합 및 관리개선을 위한 시맨틱 웹 적용 연구," 국방부, 2010.
- [3] 유동희, "시맨틱 웹 애플리케이션의 시스템 프레임워크 및 도전 과제에 관한 연구," 한국컴퓨터정보학회, 제14권 제12호, pp. 255-266, 2009.
- [4] 유동희, 나민영, "군 병원을 위한 시맨틱 웹 기반 진료 의사결정 지원 시스템," 정보처리학회논문지 B, 제17-B권 제4호, pp. 317-326, 2010.
- [5] Antoniou, G. and van Harmelen, F., Web Ontology Language: OWL. In: S. Staab and R. Studer (eds.), Hand book on Ontologies, Springer Verlag, pp.67 - 92, 2004.
- [6] Becker, C. and Bizer, C., "DBpedia Mobile: A Location-Aware Semantic Web Client," In Proceedings of the Semantic Web Challenge at ISWC 2008.
- [7] Berners-Lee, T., Hendler, J. and Lassila, O., "The Semantic Web," Scientific American, Vol. 284, No.5, pp.34-43, 2001.
- [8] Cerbah, F., "Learning Highly Structured Semantic Repositories from Relational Databases: The RDBtoOnto Tool," In Proceedings of the 5th European Semantic Web Conference (ESWC 2008), Tenerife, Spain, June, 2008.
- [9] Golbreich, C. and Wallace, E. K., OWL 2 Web Ontology Language New Features and Rationale. W3C Recommendation 27 October 2009, Available at: <http://www.w3.org/TR/owl2-new-features/>
- [10] Gruber, T. R., "A Translation Approach to Portable Ontology Specifications," Knowledge Acquisition, Vol.5, No.2, pp. 199-220, 1993.
- [11] Horrocks, I., Patel-Schneider, P. F., Bechhofer, S. and Tsarkov, D., "OWL rules: A proposal and prototype implementation," Journal of Web Semantics, Vol.3, No.1, pp. 23-40, 2005.
- [12] Horrocks, I., Patel-Schneider, P.F., Boley, H., Tabet, S., Grosz, b. and Dean, M., SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission 21 May 2004, Available at: <http://www.w3.org/Submission/SWRL/>
- [13] Kamvar, M. and Baluja, S., "Deciphering trends in mobile search," Computer, Vol.40, No.8, pp.58 - 62, 2007.
- [14] Klyne, G. and Carroll, J. J., Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation 10 February 2004, Available at: <http://www.w3.org/TR/rdf-concepts/>
- [15] Le-Phuoc, D., Parreira, J. X., Reynolds, V. and Hauswirth, M., "RDF On the Go: An RDF Storage and Query Processor for Mobile Devices," In Posters and Demos of the ISWC 2010, Shanghai, China, 2010.
- [16] Lopez, V., Uren, V., Motta, E. and Pasin, M., "AquaLog: An ontology-driven question answering system for organizational semantic intranets," Journal of Web Semantics, Vol.5, No.2, pp.72-105, 2007.
- [17] McBride, B., The Resource Description Framework (RDF) and its Vocabulary Description Language RDFS. In: S. Staab and R. Studer (eds.), Handbook on Ontologies, pp.51 - 66, 2004.
- [18] McGuinness, D. L. and Harmelen, F. V., OWL Web Ontology Language Overview. W3C Recommendation 10 February 2004, Available at: <http://www.w3.org/TR/owl-features/>
- [19] Prud'hommeaux, E. and Seaborne, A., SPARQL Query Language for RDF. W3C Recommendation 15 January 2008, Available at: <http://www.w3.org/TR/rdf-sparql-query/>

[20] Schandl, B. and Zander, S., "Adaptive RDF Graph Replication for Mobile Semantic Web Applications," Ubiquitous Computing and Communication Journal (Special Issue on Managing Data with Mobile Devices), 2009.

[21] Weiss, C. Bernstein, S. and Boccuzzo, S., "i-MoCo: Mobile Conference Guide - Storing and querying huge amounts of Semantic Web data on the iPhone/iPod Touch," In ISWC Billion Triple Challenge, 2008.

[22] Mobile Marketing Association, Introduction to Mobile Search, 2008, Available at: <http://www.mmaglobal.com/mobilesearchintro.pdf>

[23] Latitude Digital Marketing Limited, Q3 Mobile PPC Report, 2010, Available at: <http://www.latitudegroup.com/>

[24] Androjena, Available at: <http://code.google.com/p/androjena/>

[25] Bossam, Available at: <http://bossam.wordpress.com/>

[26] Dublincore, Available at: <http://dublincore.org/>

[27] Jena, Available at: <http://jena.sourceforge.net/>

[28] Micro Jena, Available at: http://poseidon.ws.dei.polimi.it/ca/?page_id=59

[29] Mobile RDF, Available at: <http://www.hedenus.de/rdf/>

[30] Protégé, Available at: <http://protege.stanford.edu/>

[31] SDB, Available at: <http://openjena.org/SDB/>

[32] XSLT, Available at: <http://www.w3.org/TR/xslt>

유 동 희



e-mail : donghee.info@gmail.com

2002년 고려대학교 경영정보학과(학사)

2009년 고려대학교 경영학과(경영학박사)

2009년~현 재 육군사관학교 전자정보학과
조교수

관심분야: Semantic Web, Ontology, Linked Data, Enterprise Architecture, Decision Support System, Recommendation System 등

나 민 영



e-mail : myra@kma.ac.kr

1978년 육군사관학교

1983년 서울대학교 컴퓨터공학과

1986년 서울대학교 컴퓨터공학과(석사)

1990년 University of Florida, Dept. of
Computer and Information Sciences (Ph. D.)

1995년 미 IBM Watson 연구소 객원연구원

2003년 Georgia Institute of Technology 교환교수

1986년~현 재 육군사관학교 전자정보학과 교수

관심분야: Database Design, Distributed Database Systems, Semantic Web, Ontology, Bioinformatics Databases, Metadata Management, Data Integration 등