

# KISTI의 차세대 정보 서비스 연구

이미경<sup>\*</sup>, 정한민, 성원경  
한국과학기술정보연구원 정보서비스연구팀

## A Study on Next-Generation Information Service of KISTI

Mikyoung Lee<sup>\*</sup>, Hanmin Jung, Won-Kyung Sung  
Information Service Research Lab, KISTI  
E-mail : jerryis@kisti.re.kr, jhm@kisti.re.kr, wksung@kisti.re.kr

### 요 약

시맨틱 웹 기반의 차세대 정보 서비스인 OntoFrame2007은 시맨틱 웹과 웹 2.0 기술을 상호 보완하여 개발한 플랫폼으로 URI기반의 개체 중심적 통합 검색을 통해 Topic, Person, Event 기반의 개체 페이지를 구성한다. 각 페이지에는 Add-on 서비스를 통해 연구자 그룹, 연구자 네트워크, 통계 정보를 제공하며, 매쉬업 기술을 이용하여 매핑 서비스, 도서 서비스를 제공한다. OntoFrame2007은 서비스 설계 과정에서 Google, Google Scholar, CiteSeer, yesKisti 사이트와 비교하여 사용자의 요구사항 및 각 서비스의 장점을 최대한 반영시켰으며, 개발 단계마다 반복적 사용성 평가를 통해 사용자 중심의 서비스를 제공하도록 개선시켰다. 본 시스템에서는 URI를 통해 개체의 모호성을 제거하여 키워드 기반의 통합 검색의 한계를 극복하였으며, 사용자 질의어를 식별하여 단위 서비스로 구성된 최적 개체 페이지를 통해 정교한 검색 서비스를 제공해준다. 개체 페이지들은 키워드 검색 결과, 연구 동향 정보 및 연구자 그룹, 전문가 네트워크 및 지역별 연구 동향 등을 제공해준다.

### 1. 서론

한국과학기술정보연구원(KISTI)은 현재 과학기술정보 포털 사이트인 yesKISTI와 해외 학술정보 제공 서비스인 NDSL을 운영하며 논문, 특허, 연구 보고서 등 국내외 정보검색 및 원문 서비스를 제공하여 과학기술분야의 연구자 및 학생들에게

많은 도움을 주고 있다. 조만간 KISTI Service2.0을 통해 새로운 과학기술정보포털을 발표할 계획이다. 우리는 또한 시맨틱 웹 기반의 차세대 정보 서비스의 프로토타입을 개발 중에 있다. 본 논문에서는 현재 개발 중인 OntoFrame에 대해 설명하고자 한다.

2005년부터 연구·개발하고 있는 OntoFrame은 국

내외 과학기술기반 정보를 연구자의 연구 개발 전 주기에 제공하는 시맨틱 웹 기반의 정보 서비스 플랫폼이다. OntoFrame은 2005년 ‘연구자 그룹 및 전문가 추천’, ‘연구자 추적’, ‘연구맵’ 서비스를 시작으로 2006년 ‘연구자 네트워크’, ‘연구자 정보’, ‘연구 성과 맵’, ‘통계정보’, ‘성과정보’, ‘기관정보’ 서비스를 개발하였다[1][2]. 본 시스템은 국내·외 연구 소개를 통해 성과를 인정받고 있으며, 연구자에게 좀 더 실직적인 도움이 되는 서비스를 제공하기 위해 지속적으로 개선하고 있다. 본 논문에서는 2007년 개발된 OntoFrame2007 서비스에 대해 설명한다.

## 2. 기존 연구

### 2.1 NDSL

국가과학기술전자도서관(NDSL)은 “과학기술정보유통체계 구축-과학기술전자도서관 구축 중심” 과제를 중심으로 한국과학기술원(KAIST)에서 개발하여 2001년 5월 서비스를 개시하였다. 2006년 NDSL의 주관기관을 한국과학기술정보연구원(KISTI)으로 이관되어 운영되고 있다. NDSL은 국내 학계, 연구계, 산업계의 모든 연구자를 위한 해외 학술 저널 및 프로시딩 포탈로서 56,000여 종의 학술저널과 188,000여 종의 프로시딩을 서비스 하고 있다. 또한 대학, 연구소, 기업체, 병원 등 학술연구기능을 수행하는 각 기관의 학술저널 콘텐츠를 대폭 확충하기 위해 전자저널 공동구매컨소시엄(KESLI)을 운영하고 있으며, 이를 통해 도서관 협력망을 운영하고 있다[3].

NDSL은 서지-초록-원문이 연계된 원스톱 서비스를 지향하여 간단한 검색이나 저널 브라우징을 통해 원문 조회를 지원하고 전자적 이용 불가능한 논문은 원문 복사 서비스를 제공한다. NDSL은 국내외 협력기관들이 함께 협력 네트워크를 통해 정보자원 개발, 국내 보유 자원 공유, 콘텐츠 개발, 공동 연구 등을 해야하며 현재 KISTI에서는

NDSL을 국가과학기술 정보허브로 자리매김하기 위해 해외 학술저널 중심의 대용량 콘텐츠 고품질화, 국내 학술지, 보고서, 회의자료, 특허 등을 연계, 이용자 간 양방향 정보교류를 지원하는 서비스와 콘텐츠 개발에 주력하고 있다.

### 2.2 yesKISTI

KISTI에서 운영 중인 yesKISTI는 국내외 과학기술정보 포털 사이트로 국내외 논문, 연구보고서, 특허 등과 같은 텍스트 기반의 정보와 동향 분석, 과학기술인력, 세미나 동영상과 같은 멀티미디어 정보 등의 83,459,771건의 과학기술 DB와 웹 문서, 커뮤니티, 블로그의 정보를 한번에 통합 검색하는 기능을 제공한다[4].

그 외 해외과학기술동향, 과학향기, KISTI 발간물 등의 자료를 이용할 수 있고, KISTI가 제공하는 맞춤형서비스, 원문서비스 등의 다양한 서비스를 이용할 수 있다. 특히 RSS, 블로그와 커뮤니티 기능을 강화하여 산학연 연구자들의 지식교류의 장을 제공해주고 전자원문 통합, 전자원문 미리보기, 특허청 및 특허정보원과 협력한 국내외 특허정보의 전자원문서비스 및 웹사이트 실시간 모니터링 체제 등을 개발하여 운영 중이다. 현재 더욱 강력해진 yesKISTI2.0 서비스를 개발 중이며 2008년 오픈할 예정이다.

### 2.3 CiteSeer

CiteSeer는 과학논문 디지털도서관, 컴퓨터 과학 및 기술관련 논문 검색 서비스를 제공한다.

97년 NEC research institute에서 개발되어 현재는 Pennsylvania State University의 College of Information Science and Technology에 의해 호스팅되고 있다. CiteSeer는 ACI(Autonomous Citation Indexing) 메소드를 사용하여 인용문헌을 자동으로 인덱싱, 링킹할 수 있게 한 Citation Index, Search Engine, Digital library이다. 이 사이트의 특징은 유사도(%)를 체크해서 유사한 문헌을 찾는 Similar Documents와

Related Documents와 같이 co-Citation 등 관련성 있는 문헌들을 찾아주며, 최신 업데이트된 자료들도 함께 검색할 수 있도록 한다. 그리고 논문 전문 검색 및 온라인 출판사가 제공하는 citation link를 통해 각 DB내 통합 검색을 가능하게 하는 Reference Linking 기능을 제공해 주며, DBLP, ACM Digital library로 링크되어 현재 인덱스되어 있는 모든 논문의 전체 메타데이터를 볼 수 있는 장점이 있다. CiteSeer는 검색엔진으로서 디지털 도서관의 역할 뿐만 아니라 다른 디지털 도서관을 위한 리소스도 제공하고 있다[5].

### 3. OntoFrame2007 서비스

#### 3.1 서비스 개요

시맨틱 웹 기반의 차세대 정보 서비스인 OntoFrame2007은 시맨틱 웹과 웹 2.0 기술을 적용한 플랫폼이다. URI기반의 시맨틱 웹 기술 및 추론 서비스를 이용하여 개체 중심적 통합 검색을 제공하고 있으며, Add-on 서비스에서는 소셜 네트워크, 통계 정보 및 OpenAPI를 이용한 매쉬업 서비스를 선보였다.

서비스를 설계 과정에서 Google, Google Scholar, CiteSeer, yesKisti 사이트와 비교하여 사용자의 요구사항을 최대한 반영시켰으며, 개발 단계마다 사용자 평가를 통해 사용자 중심의 서비스를 제공하도록 개선시켰다. 본 시스템에서는 URI를 통해 개체의 모호성을 제거하여 키워드 기반의 통합 검색의 한계를 극복하였으며 사용자 질의어를 식별하여 단위 서비스로 구성된 최적 개체 페이지를 통해 정교한 검색 서비스를 제공해준다.

#### 3.2 URI기반 개체 중심적 통합 검색

일반적인 검색엔진에서는 키워드 기반 검색을 제공하기 때문에 원하는 정보를 정확히 찾는 데 어려움이 있다. 이것은 단절된 검색과 단순한 통합 검색의 한계 때문이다. OntoFrame에서는 이런 단

점을 극복하기 위해서 URI기반의 개체 중심적 통합 검색을 추구한다.

시스템의 구성은 그림 1과 같다. 검색 창에 질의어가 입력되면 키워드의 URI와 개체를 파악한 후, Topic, Person, Event와 같이 질의어의 개체에 적합한 페이지를 구성하게 된다. 페이지를 구성하기 위해서는 검색 및 추론 서비스가 이용되며 매쉬업을 통해 서비스가 제공된다.

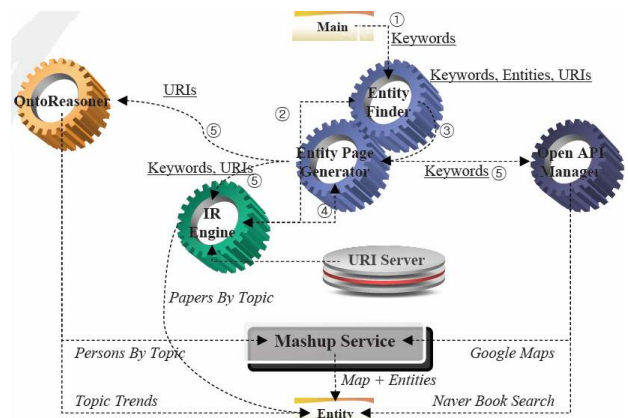


그림 1. OntoFrame workflow

그림 2는 “semantic web”이라는 질의어를 넣었을 때 제공되는 페이지이다. “semantic web”의 URI와 개체를 파악하여 질의어가 Topic으로 판단한 후 주제 페이지를 구성한다. 주제 페이지의 경우는 일반적인 검색 결과인 Search Results, 주제들의 연관 관계를 제공하는 Topic Trends, 주제어의 상·하위 관계를 제공하는 See Also, 주제별 전문가 순위를 나타내는 Persons by Topic, 주제별 전문가 기관을 나타내는 Institutions by Topic, 주제어를 가지는 논문들의 리스트를 나타내는 Papers by Topic 서비스로 구성된다. 오른쪽 Add-on 서비스에서는 “semantic web” 분야의 연구자 그룹과 연구자 그룹별 연구자 네트워크를 제공하는 Social Network, 주제별 통계정보, 주제와 관련된 학술대회 정보, OpenAPI를 이용한 지역별 연구기관과 연구자 정보와 책 정보를 제공한다.

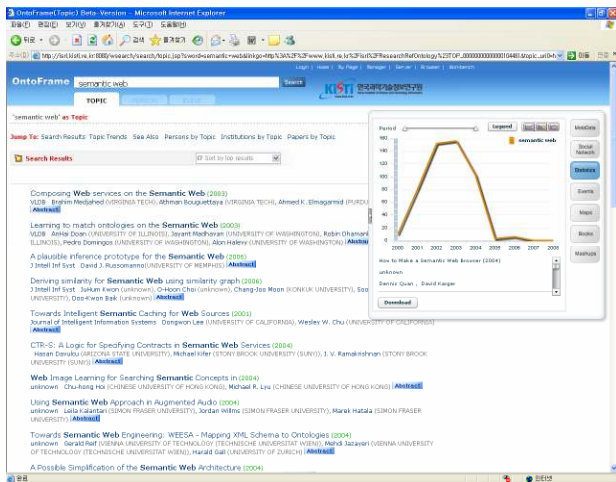


그림 2. Topic Page

그림 3은 OntoFrame2007에서 “steffen staab”라는 연구자를 검색한 결과이다. 사람 이름을 질의어로 입력하면 URI와 개체 검색을 통해 자동으로 연구자 페이지가 구성된다. 주제 페이지와 비슷하게 Search Results, 현재 연구자와 연구분야가 유사한 Similar Persons, 논문의 인용정보를 이용한 Persons in Citations, Cited By 서비스로 구성되며 Add-on 서비스에서는 연구자의 상세 정보, 연구자의 네트워크, 통계정보, 연구자의 저서 정보를 제공해준다.

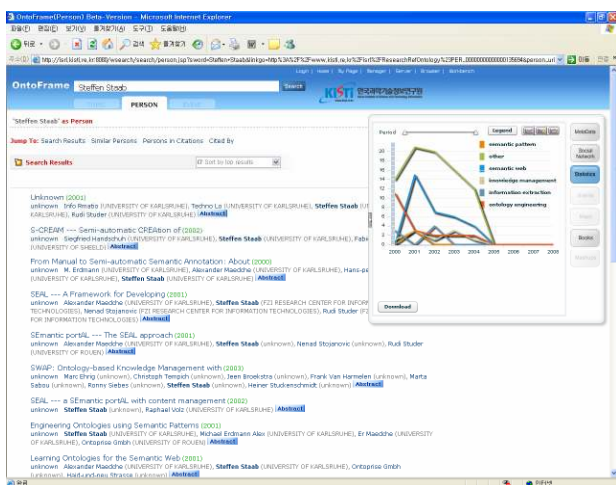


그림 3. Person page

현재 OntoFrame에 등록된 데이터는 논문 366,065편, 인력 983,677명, 주제 435,985건, 기관 17,820개, 위치 정보 451,763건이다. URI기반의 테

이터들이 URI서버에 등록되면서 자동으로 생성된 트리플은 추론 엔진에 의해 Article의 메타데이터에서 Person으로 전파되는 알고리즘을 가진다. 현재 각 컴포넌트는 검색엔진에 의한 결과와 추론엔진에 의한 결과로 나뉘져 있으며 2차적 정보를 제공하는 경우는 추론을 거쳐 나온 결과 값이다.

### 3.3 Add-on 서비스

개체 페이지의 Add-on 서비스에서 가장 눈에 띄는 서비스는 연구자 네트워크와 지역별 연구동향을 확인할 수 있는 맵 서비스이다.

그림 4는 Persons by Topic 컴포넌트에서 리스트에 나타나는 5명의 연구자들의 그룹을 나타낸 것이다. 연도의 변화에 따라 같이 연구하는 연구자 그룹의 모양은 바뀌게 된다. 그룹을 선택하면 그룹에 속한 연구자들간의 관계를 Co-author, Citation 등의 관계에 따라 네트워크를 보여준다. 연구자 페이지의 Social Network의 경우에는 선택된 연구자와 관련된 연구자들의 네트워크를 보여준다. 네트워크를 통해 현재 선택된 연구자와 밀접한 연구를 하고 있는 사람 및 과거에 함께 연구했던 사람들의 정보를 확인 할 수 있으며, 연구자 네트워크의 모양을 통해 연구자들간의 관계를 유추해 볼 수 있는 장점을 가진다.

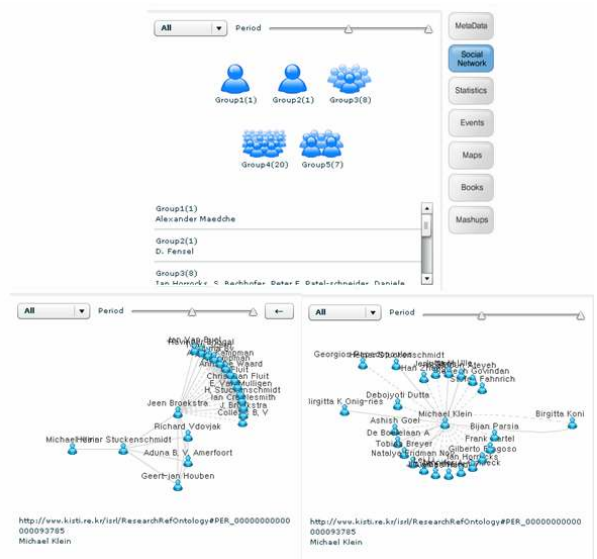


그림 4. Social Network

연구 맵 서비스의 경우, Google Maps OpenAPI와 매쉬업하여 서비스하고 있다. 실제 연구 기관의 GPS를 이용하여 연구자 위치를 파악하게 되고 추론 서비스 결과와 매쉬업하여 주제와 관련된 연구를 수행하는 지역 및 특정 연구자가 소속된 기관의 위치 정보를 파악할 수 있다.



그림 5. Maps

Naver Book OpenAPI를 이용하여 주제, 연구자와 관련된 책 정보를 제공하고 실제 구입할 수 있는 북 서비스도 제공한다.

#### 4. OntoFrame2007 사용성 평가

OntoFrame2007은 사용자 요구 사항을 반영하고 서비스의 사용성을 높이기 위해 설계와 구현 시 반복적 사용성 평가를 실시한 후, 그 결과를 반영하여 시스템을 구현하였다[6][7].

##### 4.1 사용자 인터뷰를 통한 사용성 평가

서비스 시안 작성 후, 시스템 구현에 들어가기 앞서 사용자 인터뷰를 통해 사용자의 의견을 반영하여 시스템을 설계하였다. 먼저 사용성 목표를 세운 후, OntoFrame 서비스 시안을 토대로 실시한 1차 사용성 평가는 사용자의 실제 연구 형태를 파악하고 서비스 시안이 사용자 요구에 적합한지를 확인해보았다. 1차 평가는 연구자 3명과 박사과정 학생 3명을 대상으로 하였으며 인터뷰는 학술연구 정보 서비스 이용 형태를 파악하고 OntoFrame 서비스 시안에 대해 설명한 후, 서비스 시안에 대한

의견을 수집하는 형태로 진행되었다.

사용자 인터뷰는 평균 1시간 내외로 진행하였으며 OntoFrame2007 서비스 시안, 학술연구정보 서비스 제공 웹사이트(Google, Google Scholar, CiteSeer 등)를 사용하여 인터뷰를 진행하였다.

이 결과를 통해 사용자들이 주로 이용하는 웹사이트들의 장점을 우리 서비스에 반영시키고 서비스 시안에 대한 사용자 의견을 토대로 단점을 수정하거나 삭제하고 장점을 부각시킬 수 있는 서비스를 추가하여 OntoFrame2007-a 버전을 구현하였다.

##### 4.2 비교 평가를 통한 사용성 평가

2차 사용성 평가는 OntoFrame2007-a 버전이 서비스 목표에 부합되는지 1차 사용성 평가 시 문제가 해결되었는지 판단하기 위해 실시하였다. 평가 방식은 과제를 주고 이것을 달성하기 위한 사용자들의 행동을 관찰하는 방법 및 사용자 인터뷰를 이용하였다. 비교 평가를 위해 학술연구정보 서비스인 Google Scholar, CiteSeer와 OntoFrame를 대상을 7개의 과제를 주고 수행작업을 통해 결과를 분석하였다. 단순 검색을 통해 해결할 수 있는 1차 과제와 정보 조합과 분석을 통해 해결할 수 있는 2차 과제로 나누어졌다. 2차 사용성 평가 대상자는 평소 CiteSeer과 Google Scholar를 통해 학술정보를 검색한 경험이 있는 박사과정 6명, 석사과정 3명을 대상으로 시행하였다.

단순 검색의 경우에는 Google Scholar가 목표 달성에 도움을 주는 반면, OntoFrame의 경우에는 해당 서비스가 제공됨에도 직관적이지 않은 메뉴 이름과 복잡한 로직으로 인해 사용자들이 쉽게 사용하지 못했다. 그리고 2차 과제의 경우에는 OntoFrame이 정보 조합과 분석이 필요한 환경에서 유용한 정보를 제공한다는 의견이 나왔다.

각 서비스 UI에 대한 의견과 OntoFrame에 대한 의견을 중심으로 검색 로직과 개체 페이지의 수정, 메뉴 이름들을 수정하였으며, UI는 사용자들이 선

호하는 Google형식의 검색 UI를 제공하도록 수정하여 OntoFrame2007-β 버전을 구현했다.

OntoFrame2007-β 버전을 개발한 후 시행한 3차 사용성 평가는 2차 사용성 평가와 같은 질문지를 통해 OntoFrame2007-β 버전에 대한 검사를 수행하였으며 인터뷰 대상자는 박사과정 학생 6명과 외국인 학생 6명을 대상으로 시행하였다. 3차 사용성 평가는 서비스 페이지의 사용성 향상 및 다국어 문화 지원에 초점을 맞췄으며 3차 평가의 결과는 즉시 반영할 수 있는 부분만 반영하고 나머지는 2008년 서비스에 반영할 예정이다.

사용성 평가 목표에 따라 서비스 시안을 대상으로 실시한 사용자 인터뷰 중심의 1차 사용성 평가와 서비스 프로토타입을 대상으로 실시한 학술연구정보 서비스간 비교 평가 중심의 2차 사용성 평가를 상이하게 구성함으로써 사용성 관련 문제를 다양한 관점에서 도출하였고 매 구현 단계마다 반복적으로 실시하여 학술연구정보 서비스의 사용성을 지속적으로 향상시키고 있다.

## 5. 결론

KISTI는 현재 OntoFrame이라는 시맨틱 웹 기반의 차세대 정보 서비스의 프로토타입을 개발하고 있다. 2005년부터 시작된 우리의 연구는 차세대 시맨틱 웹 기술과 웹 2.0 기술을 상호 보완하여 적용한 OntoFrame을 개발하였고 현재 OntoFrame2008 서비스를 개발 중에 있다. OntoFrame2007은 URI기반의 개체 중심적 통합 검색을 통해 Topic, Person, Event 기반의 개체 페이지를 구성하며, 각 페이지에는 Add-on 서비스를 통해 연구자 그룹, 연구자 네트워크, 통계 정보를 제공하며, 매핑 서비스 및 도서 서비스의 매쉬업 기술을 이용하는 서비스를 제공하고 있다. 본 시스템에서는 URI를 통해 개체의 모호성을 제거하여 키워드 기반의 통합 검색의 한계를 극복하였으며 사용자 질의어를 식별하여 단위 서비스로 구성된 최적 개체 페이지를 통해

정교한 검색 서비스를 제공해준다. OntoFrame 서비스는 개발 단계마다 반복적 사용성 평가를 통해 사용자 요구에 적합한 서비스를 구현하기 위해 노력하고 있다. 2008년 개발될 서비스에서는 다중 통합 검색과 확장된 Add-on 서비스를 통해 좀 더 강력한 서비스를 제공할 예정이다.

## [참고문헌]

- [1] OntoFrame-K: 연구자 간 협업 지원 서비스를 위한 시맨틱 웹 기반 정보유통플랫폼, 정한민, 이미경, 성원경, 박동인, 한국컴퓨터종합학술대회, 2006
- [2] 시맨틱 웹 기반 플랫폼상에서의 웹2.0 활용 서비스, 정한민, 이미경, 성원경, 한국정보처리학회지, 14권 4호, pp76-82, 2007
- [3] <http://www.ndsl.or.kr/eng/newindex.html>
- [4] <http://www.yeskisti.net/yesKISTI/index.jsp>
- [5] <http://citeseer.ist.psu.edu/>
- [6] 반복적 사용성 평가를 통한 학술연구정보 서비스 구현, 정한민, 이미경, 성원경, 제34회 한국정보과학회 추계학술대회, 2007
- [7] Development of an Academic Research Information Service through Repeated Usability Evaluations, H. Jung, M. Lee, S. Lee, W. Sung, The First International Conference on Advances in Computer-Human Interaction, 2008