

추론기반 과학기술 정보유통 인터페이스 설계

구남양^o 김평 이승우 강인수 정한민 성원경
한국과학기술정보연구원

{nak0519^o, pyung, swlee, dbaisk, jhm, wksung}@kisti.re.kr

(The Interface Design for Distribution of Science and Technology Information based on Inference)

Namang Kuh^o, Pyung Kim, Seungwoo Lee, Insu Kang, Hanmin Jung, Wonkyung Sung
Korea Institute of Science and Technology Information

요 약

연구개발에 필요한 정보의 생성 주기가 빨라지고 그 양이 방대해지는 것은 물론, 다양한 주제와 분야의 연구가 진행되면서 연구자간의 연구 교류와 연구 성과물에 대한 공유의 필요성이 증대되고 있다. 연구자 각 개인들이 컴퓨터에 소장하고 있는 지식정보는 연구자에 의해 생성되거나 또는 핵심자원으로 선별·분석된 것이므로, 개인이 소장하고 있는 지식정보 자원의 효율적인 공유와 유통은 연구자들의 연구를 촉진시키고, 결과적으로 해당 분야의 연구협력의 기회를 증진시킬 수 있다. 본 연구에서는 URI 기반 과학기술 온톨로지로 표현된 성과물을 대상으로 연구자, 기관, 지역 간의 연구 활동은 물론 다양한 관점에서 성과물 정보를 제공함으로써 연구자들의 연구 활동을 지원하는 정보유통 인터페이스를 설계하였다.

1. 서 론

최근 국가 간의 기술경쟁이 치열해짐에 따라, 국가의 경쟁력 확보를 위해서는 연구자간의 교류 및 연구 성과물에 대한 공유의 활성화가 필요하다. 이를 효율적으로 지원하기 위하여 한국과학기술정보연구원(KISTI)에서는 2004년부터 “지식기반 정보유통 플랫폼 개발 연구” 사업의 일환으로 “지식기반 정보유통 플랫폼(OntoFrame-K^o)(이하 “OFK”로 표기)”을 개발 중에 있다[1]. OFK는 과학기술 연구·개발자들의 협력 연구를 가상공간에서 지원하기 위한 연구·개발 인프라로서, 연구자간 연구 현황은 물론, 기관별, 지역별 연구 성과물 정보의 유통과 관리를 목표로 하고 있다. 이를 위해 IT 분야 시소러스, 시소러스 개념어와 분야 분류체계 매핑 작업, 또한 URI 기반의 성과물 정보를 과학기술 온톨로지로 표현하고, 표현된 온톨로지를 기반으로 한 추론을 통해 과학기술 정보의 효율적인 유통을 위한 추론 에이전트를 개발하고 있다.

다양한 공유·유통체계의 개발과 더불어 정확한 정보의 선별과 제공을 위한 기술적 방안에 대한 필요성도 강조되고 있다. 90년대 중반에 실시된 미국 과학재단의 조사에 의하면 연구·개발자들은 정보의 폭증으로 인해 연구·개발 시간의 40% 이상을 정보 수집 및 분석에 소비하고 있으며 연구·개발의 생산성 향상을 위해서는 무엇보다도 정보획득 과정의 개선이 필요하게 되었다[2]. 그러나 기존의 키워드 또는 정보의 형태적 유사성의 의한 정보의 수집과 제공은 이용자가 일정한 곳에 저장되어 있는 정보 자원을 이용자가 원하는 키워드에 의해 수집된 정보를 제공하는 것으로, 폭증하는 대량의 정보자원에서 수요자가 요구하는 정보만을 효율적으로 선별·제시하지 못하는 기술적 한계를 드러내고 있다. 이에 대한 해결 방안으로서 최근 시멘틱 웹의 추론 기반 검색기술이 대두되고 있다. 정보자원의 의미를 기준으로 정보를 관리하고 제공이 가능한 시멘틱 웹(semantic web)과 관련하여 추론

서비스는 여러 곳에 흩어진 정보를 이용자가 원하는 키워드에 의해 수집된 정보의 의미를 개념으로 정의하고 개념 간의 관계성을 표현함으로써 정보를 공유할 수 있도록 하며, 이들 정보의 의미를 서로 연계하여 이용자에게 보다 효율성 있는 정보를 제공할 수 있다. 이러한 시멘틱 웹의 비전을 현실화하기 위해서는 온톨로지를 기반으로 의미 정보로 구성된 시멘틱 문서들에 대한 추론을 통해 사용자가 필요로 하는 지식을 검색하기 위한 기술이 필수적으로 요구된다[3].

본 논문은 1장에서 본 연구의 필요성과 목적 소개, 2장에서 인터페이스의 세부 구성요소, 3장에서 세부 구성요소별 상호 정보연계 방안, 4장에서 결론 및 향후 연구 방향에 대해서 기술하였다.

2. 인터페이스 설계 방향

본 연구의 목적은 URI기반 온톨로지로 표현된 연구개발 성과물(보고서, 특허, 논문)을 기반으로 추론 서비스를 통해 연구자, 기관, 지역 간의 연구 동향은 물론, 연구자간의 효율적인 정보 공유를 위한 사용자 인터페이스를 설계하는 것을 목표로 한다. OFK는 다음과 같이 6개의 세부 인터페이스로 구성되어 있다.

- 연구자 네트워크 : 주제와 분야별 전문가 순위와 연구자간 성과물 중심으로 공동 연구자 관계 및 공동 연구 정도를 제시
- 연구자 정보 : URI 기반 연구자에 대한 검색과 연구자의 주제 및 분야 제시
- 연구 성과맵 : 연구가 수행되는 지역의 연구 활성화 정도나 지역별 기관이나 연구자 분포 현황 제시
- 통계정보 : 연구자에 대한 성과물 통계정보 제시
- 성과정보 : 유형별 성과물 정보 제시
- 기관 정보 : 기관의 연구현황 및 성과물 정보 제시

OFK에서 서비스를 위한 정보자원은 다음과 같다.

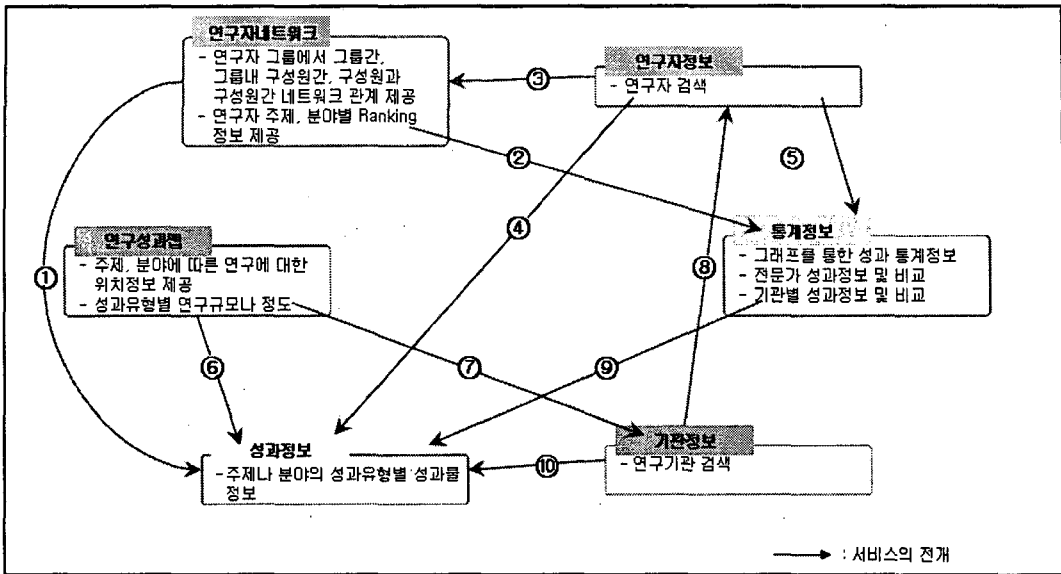


그림 1 서비스 구성요소 및 서비스의 연계

- URI 기반 기관, 부서, 연구자 정보 : 기관, 부서 연구자 식별에 필요한 기본 정보.
- 범용 시소러스 및 IT분야 시소러스 : 개념어별로 분야 분류코드와 매핑되어 있음.
- IT 분야 분류체계 : 3단계 분류체제로 구성됨
- 성과물 온톨로지 : 보고서, 특허, 논문과 저자, 저자의 소속기관, 참여율 등이 URI 기반으로 표기되어 있음.

- ⑦ 연구성과맵을 통해 선택된 지역에 위치한 기관 리스트 제공
- ⑧ 기관 검색을 통해 선택된 기관의 연구자 리스트 제공
- ⑨ 통계정보에서 선택된 연구자의 해당 연도에 대한 성과정보 제공
- ⑩ 기관 검색을 통해 선택된 기관의 성과정보 제공

3. 추론 인터페이스 세부 구성요소

인터페이스의 추론 서비스와 관련하여 성과물 유형(전체, 보고서, 특허, 논문), 연도(최근 몇 년간의 연구수행기간), 주제 및 분야(주제 또는 분야별 다중 선택)로 제한하여 정보를 제공할 수 있다. 다음에서는 인터페이스를 구성하는 6개 세부 인터페이스별로 각각의 정보 생성 방법과 역할에 대해서 설명하도록 한다. 각 세부 인터페이스에 사용되는 정보는 과학기술 온톨로지와 추론을 통해 얻어진 결과를 기반으로 정보를 제공한다.

모든 정보는 성과물 온톨로지를 중심으로 생성된다. 성과물 원문은 시소러스를 참조하여 주제 및 분야 코드가 자동으로 할당되고, 이를 통해 성과물의 주제/분야가 결정되며, 연구자는 연구자가 참여한 성과물, 기관은 기관의 연구자가 참여한 성과물로, 지역은 지역에 속한 기관의 성과물을 기준으로 특성을 부여받는다. 즉 성과물과 연구자, 기관, 지역 정보는 상호 밀접한 관계를 가진다.

그림1은 인터페이스를 구성하는 세부 인터페이스의 역할 및 세부 인터페이스간 정보 연계를 통한 서비스의 흐름을 보여주고 있다.

- ① 연구자 네트워크에서 선택된 연구자의 성과정보를 성과물 유형에 따라 제공
- ② 연구자 네트워크에서 선택된 연구자의 성과물 통계정보 제공 (다중 선택시 연구자간 성과물 통계 비교)
- ③ 연구자 검색을 통해 선택된 연구자를 중심으로 공동 연구 현황 및 공동 연구 정도 제공
- ④ 연구자 검색을 통해 선택된 연구자의 성과 유형별 성과정보 제공
- ⑤ 연구자 검색을 통해 선택된 연구자의 성과물 통계정보 제공 (다중 선택시 연구자간 성과물 통계 비교)
- ⑥ 연구성과맵을 통해 선택된 지역의 성과정보 제공

3.1 연구자 네트워크

연구자 네트워크에서는 연구자가 작성한 성과물 정보를 기반으로 연구자의 연구 주제 또는 분야를 선정하고, 성과물의 중요도와 참여율에 따라 연구자별 가중치를 계산하여 전문가 순위를 결정한다. 또한 성과물에 참여한 공동 연구자 정보를 통해 연구자간 상호 연구자 관계를 알 수 있도록 연구자 네트워크를 구축한다. 연구자 네트워크를 통해 주제 또는 분야별 전문가와 연구자간 상호 협력 관계를 파악하고, 이를 통해 연구자간 협업 연구를 증진시킬 수 있다. 연구자 네트워크를 생성하는 단계는 다음과 같다.

- ① 성과물 주제 또는 분야 정보로 연구자의 주제와 분야 정보 산정
- ② 성과물에 대한 연구자별 가중치 계산

- ③ 가중치의 합으로 전문가 순위 선정
- ④ 성과물의 공동 연구자 정보로 연구자 그룹을 생성하고 연구자간 공동 연구에 따른 가중치 계산

연구자 네트워크는 공동 연구자 그룹의 수에 따라 그룹별 전문가 수만 보여주는 전문가 네트워크와 그룹 구성원의 상호 협력관계와 그 정도를 보여주는 연구자 네트워크로 구분된다.

3.2 연구자 정보

연구자의 URI 또는 이름 검색을 통해 특정 연구자를 선택할 수 있으며, 선택된 연구자의 생성한 성과물에 따라 연구 주제 및 분야, 연구자 네트워크, 연구 주제 또는 분야별 관련 성과물 정보와 소속기관 정보를 제공한다. 연구자 정보를 생성하는 단계는 다음과 같다.

- ① 연구자 이름 또는 URI 검색
- ② 검색 결과로 연구자별 소속기관과 URI를 제공하여 특정 연구자를 선정
- ③ 선정된 연구자의 성과물 분석을 통해 연구자의 연구 주제 및 분야 정보 및 특정 그룹과 랭킹정보 제공
- ④ 주제 또는 분야에 따라 유형별, 연도별 성과물 리스트 제공

3.3 기관 정보

기관의 URI 또는 명칭 검색을 통해 특정 기관을 선택할 수 있으며, 선택된 기관이 보유한 성과물에 따라 연구 주제와 분야, 연구 주제와 분야별 관련 성과물 정보와 소속 연구자 정보를 제공한다. 기관 정보를 생성하는 단계는 다음과 같다.

- ① 기관 명칭 또는 URI 검색
- ② 검색 결과로 기관 명칭과 URI, 지역 위치 정보를 제공하여 특정 기관을 선정
- ③ 선정된 기관의 성과물 분석을 통해 기관의 연구 주제 및 분야 정보를 제공
- ④ 주제 또는 분야에 따라 유형별, 연도별 성과물 리스트 제공

3.4 연구 성과맵

연구 성과맵은 연구의 주제나 분야, 성과물의 유형, 연도별 연구에 대한 우리나라 각 지역별 활성화 정도나 성과의 규모에 대한 정보를 제공한다. 또한 지역별로 전문가, 전문가 정보를 제공함으로써 지역별 연구개발 현황은 물론 특성화된 지역, 기관 검색에 유용하게 사용될 수 있다. 연구 성과맵을 생성하는 단계는 다음과 같다.

- ① 연구 성과물의 참여 연구자별 소속기관(성과물 생성 당시)을 이용하여 성과물의 소속기관, 성과물의 생성 지역을 판별
- ② 지역별 연구 현황 정보를 맵을 이용하여 제시
 - ✓ 서울 지역은 구 단위로 정보 제공
 - ✓ 타 지역은 광역시도 단위로 제공
- ③ 지역 선택 (복수 시도도 선택가능)
 - ✓ 지역을 선택하여 해당 지역에 속한 전문기관 또는 전문가 리스트를 볼 수 있음
 - ✓ 선정된 연구자의 성과물 분석을 통해 연구자의 연

구 주제 및 분야 정보를 제공

- ④ 지역별 성과물 정보 제공

3.5 통계정보

통계정보는 연구자의 성과물 통계, 기관의 성과물 통계정보, 지역별 성과물 통계정보를 제공함으로써, 연구자간(최대 3인)의 비교를 가능하게 해준다. 통계정보를 생성하는 단계는 다음과 같다.

- ① 연구자별 성과물 리스트를 추출
- ② 성과물의 생성 연도에 따라 그래프 생성
 - ✓ 단일 연구자의 경우 성과물 유형별 그래프 생성
 - ✓ 다중 연구자의 경우 전체 성과물로 그래프 생성
- ③ 그래프를 클릭하는 경우 성과물 리스트 제공
 - ✓ 해당 연구자의 특정 연도 및 전체 연도의 성과물 정보

3.6 성과정보

성과정보는 연구자, 기관, 지역별 성과물 리스트를 제공하는 것으로 연구자 정보, 기관 정보, 연구 성과맵, 통계정보에서 성과물 리스트 보기를 선택 시 성과물 유형별로 성과물 정보를 제공한다. 성과물에 대한 메타데이터 검색 및 원문에 대한 검색은 별도의 정보관리 인터페이스를 통해 제공된다[4].

4. 결론

KISTI에서는 과학기술 문헌로지와 추론을 통해 연구자 네트워크, 연구자 정보, 기관 정보, 연구 성과맵, 통계정보, 성과정보, 기관정보를 제공하는 시스템을 개발중이고, 이를 위해 사용자 인터페이스를 설계하였다. 인터페이스는 주제 또는 분야별 전문가/기관/지역 정보를 제공하는 것은 물론 연구자 공동 연구 정도와 연구자간의 관계, 기관/지역별 성과물 리스트와 통계 정보를 제공함으로써 연구자간의 연구 교류 및 연구 성과물 공유 활성화를 촉진할 수 있다. 그러나 시스템에 사용되는 성과물이 제한되어 있고, 성과물 정보에서 추론 가능한 정보만을 대상으로 인터페이스를 구성하였기 때문에, 보다 정확한 정보를 생성하기 위해서는 성과물 유형의 확대 및 추론의 정확성을 높이기 위한 연구가 계속되어야 할 것이다.

참고문헌

- [1] 정한민 외 2, 연구자 간 협업 지원 서비스를 위한 지식 베이스 설계, 제17회 한글 및 한국어정보처리학회, 2005.
- [2] Bielawski Lary 외 1명, Electronic Document Management System, Prentice Hall PTR, 1997.
- [3] 김우주, "웹 환경에서의 지식교환/유통을 위한 지식 표현 및 추론기술 연구(한국전산원)", 연세대학교, pp. 1-7, 2004.
- [4] 구남양외 1명, 효율적 정보자원 공유 및 유통을 위한 P2P 기반 플랫폼 설계, 2005년 12회 한국정보관리학회 학술대회, 2005.