

테크놀로지 인텔리전스에 기반한 R&D 전략 수립 지원 서비스 설계

이미경^o 정한민 김평 성원경
한국과학기술정보연구원 정보기술연구실

jerryis@kisti.re.kr, jhm@kisti.re.kr, pyung@kisti.re.kr, wksung@kisti.re.kr

Design of Technology Intelligence- based Service for Establishing R&D Strategy

Mikyoung Lee^o Hanmin Jung Pyung Kim Won- Kyung Sung
Korea Institute of Science and Technology Information

정보의 과부하로 인해 단순한 검색만으로는 의사 결정에 필요한 정보를 찾기 어려워지고 있다. 기존의 예측분석 도구들이 있지만 소규모 데이터 분석에 국한되어 있는 기능적인 한계가 있어서 대용량 전문 특화 자원에 대해 다양한 분석 기능을 제공해주는 서비스는 미비한 실정이다. 우리는 연구자들의 연구 생산성을 극대화하고 체계적인 연구개발을 수행할 수 있도록 지원해주는 테크놀로지 인텔리전스 기반의 R&D 전략 수립 지원 서비스를 개발하고자 한다. 우리의 서비스는 연구 개발을 수행하는 의사결정권자들을 주요 타겟으로 정하고 이들이 연구 개발 전략을 수립하거나 의사를 결정할 때 필요한 정보를 제공하기 위해 특허, 논문, 보고서 등의 기술 자료를 대상으로 분석된 정보를 제공한다. 테크놀로지 인텔리전스에 기반한 R&D 전략 수립 지원 서비스(TIS)는 텍스트로부터 유의미한 개체와 개체 간 관계를 추출하여 정보 활용 가치를 증대시키는 동시에 이 결과를 시맨틱 서비스 플랫폼 상에서 메타데이터와 융합하여 분석적 가치를 제고하는 의사 결정을 지원한다. 요소 기술을 포함하는 정보기술(Technology), 국가·기관·연구자 등의 연구 주체(Research Agents), 논문·특허 등의 연구 성과(Research Results)를 경쟁과 협력 관계 중심으로 다각적 관점에서 분석하며, 연구 성과, 기술, 주체 간의 발견, 조합, 비교, 예측을 주요 요소로 제공하는 테크놀로지 인텔리전스 서비스를 보여준다.

TIS는 OntoFrame기반의 서비스로 시맨틱 웹 기술과 텍스트 마이닝 기술을 융합하여 사용하고 있다. OntoFrame은 시맨틱 웹 기반의 서비스 플랫폼으로 사용자가 필요한 정보를 유용하게 검색하여 지식을 발견하는 것을 목적으로 개발되었다[1][2]. 또한 텍스트 내에 존재하는 핵심개체와 관계 정보를 추출하기 위해서 SINDI(Scientific INtelligent DIscovey)라는 대용량의 과학기술문서에서 중요한 의미를 갖는 핵심 개체를 인식하고 인식된 개체들 간의 상호작용을 기술하는 연관 관계를 추출하여 고차원적인 전문지식을 생성하는 과학기술지식처리 아키텍처를 사용한다[3]. TIS에서는 연구자(Person), 기관(Insitution), 국가(Nation), 성과물(Article, Patent), 기술정보(Topic) 등의 17개의 클래스와 94개의 프로퍼티로 구성된 연구 성과 온톨로지를 모델링하여 사용하고 있다. TIS의 데이터는 Green Technology분야의 34만건의 논문과 30만건의 특허 데이터를 중심으로 3억개에 달하는 RDF 트리플 형태의 시맨틱 지식을 사용한다.

TIS는 연구 성과물을 중심으로 기술과 연구 주체를 추출하고 서로 간의 발견, 조합, 비교 서비스를 제공한다. 모든 서비스는 논문과 특허 데이터의 비율에 따라 연구 관점과 사업 관점으로 서로 다른 분석 결과를 제공한다. 서비스는 크게 기술·주체 복합 서비스, 기술 중심 서비스, 연구 주체 중심 서비스로 구분할 수 있다.

기술 · 주체 복합 서비스의 대표적인 서비스인 기술/주체 맵은 해당 기술의 주요 연구 주체들은 어떤 연관된 기술에 집중하고 있는지, 연구주체들간의 경쟁/협력 관계는 어떤지, 연구 성과는 논문과 특허 중 어느 곳에 집중하고 있는지를 비교하고 분석된 결과를 한눈에 파악할 수 있다. 기술 중심 서비스의 대표 서비스인 기술 네트워크 서비스는 기술간의 요소 기술, 유사기술, 경쟁 기술 등의 주요 관계를 시각화하여 보여준다. 기술의 현재 성숙도 위치를 파악하고 기술과 연구 주체의 신규 진입 시기를 파악하는 서비스 또한 제공되고 있다. 주체 중심 서비스의 대표서비스인 연구자 그룹 서비스는 연구 주체들간의 협력 그룹을 통해 해당 기술을 연구하는 주체들의 경쟁/협력 관계를 파악할 수 있고, 해당 기술이 독보적인 대표 주체에 의해 시장이 주도되는지 서로 협력하여 균형 있게 발전하는지 등의 기술에 대한 시장의 흐름을 파악할 수 있다. 그리고 연구주체 협력 네트워크를 통해 좀 더 자세한 협력관계를 제공한다.

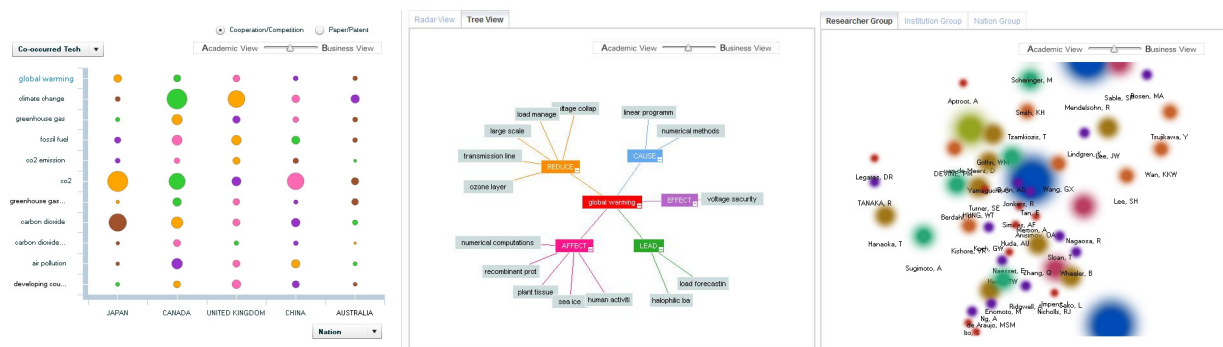


그림 1. 기술/주체 맵(a), 기술관계네트워크(b), 연구주체 그룹 서비스(c)

TIS는 기존의 분석 도구들과 차별화를 위해 빠르고 자동화된 지식 구축, 의사 결정에 도움이 되는 효과적인 서비스 개발, 외부 오픈 데이터와 연계한 정보 접근성을 향상시키는 것을 목표로 서비스를 개발하였다. 우리의 서비스는 시맨틱 웹 기술을 기반으로 구성된 서비스로 정보의 연계가 쉬우며, 정보 서비스를 뛰어 넘는 지식 서비스의 개발이 가능하다. 논문을 특허와 같이 이질적인 정보를 융합함으로써 다관점의(Multifaceted) 정보 제시를 가능하게 하였다. 논문을 중심으로 한 연구 관점, 특허를 중심으로 한 사업 관점 등 다른 성격을 가지는 연구 성과들의 구성 비율을 조절하여 다양한 조건에 대한 분석 정보를 확인할 수 있게 만들었다. 기존의 데스크탑 기반의 특허/논문 분석 지원도구의 경우는 확보된 데이터 셋을 중심으로 분석 정보가 가능하며, 정해진 입력 형태, 제한적인 데이터의 크기에 대해서만 분석이 가능한 한계를 가지고 있으나, 우리의 서비스는 무제한의 데이터를 대상으로 사용자가 검색한 개체와 연관된 동향 분석 정보를 원스톱으로 제공하는 장점을 가진다. 그리고 텍스트 마이닝을 통해 유의미한 개체를 자동으로 파악하고 LOD와 같은 외부 데이터 셋을 활용하여 정보 접근성을 향상시켰다. 향후, 우리는 기술과 주체의 관점에서 다양한 조합을 통해 서비스를 연계 · 융합하는 서비스를 추가하고 분석 결과에 대한 예측 기능을 추가할 예정이다.

참고문헌

[1] "Use Case: OntoFrame 2008-A Semantic Portal Service of Academic Research Information", 2009. <http://www.w3.org/2001/sw/sweo/public/UseCases/OntoFrame/>

[2] "Case Study: Integrated, Connected Search Service for Technical Standards Information", 2010, <http://www.w3.org/2001/sw/sweo/public/UseCases/Kisti/>

[3]"그리드 기반의 고성능 과학기술지식처리 프레임워크 개발", 정창후, 최윤수, 최성필, 윤화목, 한국콘텐츠학회논문지, 2009.