

## 시맨틱 자원을 이용한 KOSEN 통합검색 서비스

### Unified Search Service using Semantic Resource in KOSEN

한희준, 윤정선

한국과학기술정보연구원

Hee-Jun Han, Jungsun Yoon

Korea Institute of Science and Technology  
Information

#### 요약

검색기능은 사용자가 원하는 정보를 효과적으로 탐색하고 이용가능하게 하는 중요한 서비스 요소이다. 본 논문에서는 과학기술분야 전문분석자료, 수집자료, 동향정보를 제공하는 KOSEN(한민족과학기술자네트워크)의 통합검색 서비스 구축에 대해 논한다. 어휘지능망, 연구주제망 등 시맨틱 자원을 활용하여 사용자의 검색 편의성 및 효율성을 최대화 하였으며, 빠른 검색속도와 다양한 검색 부가기능을 제공함으로써 정보의 획득 및 이용의 만족도를 증가시키고자 FAST 검색엔진을 적용한 특징이 있다.

## I. 서론

웹을 통해서 방대한 양의 정보를 공유하고 탐색하기 위해 검색 기술은 진화를 거듭한다. 사용자의 정보 획득의 노력을 줄이고자 키워드 매칭 기반의 검색은 결과 문서들의 랭킹 기법, 사용자 질의분석 및 로그기반의 개인화 검색 기법, 다양한 언어자원 및 온톨로지 기반의 시맨틱 기술을 적용한 지능형 검색으로 탈바꿈을 하고 있다[1]. KOSEN(한민족과학기술자네트워크) 서비스는 국내외 과학기술자의 온라인 사이버 커뮤니티를 구축하여 연구정보, 동향, 첨단기술정보, 지식정보를 교류하고 연구자 네트워크 이용해 고부가가치의 분석자료를 생성하며 이를 확산시키고 있다. 다양한 자료를 제공하는 게시판 형태의 서비스에서 대부분 사용자의 첫 행동은 검색 실행이다. 본 논문에서는 기존의 키워드 기반의 검색 환경을 개선하여 사용자가 원하는 정보를 쉽고 효과적으로 찾을 수 있도록 FAST 검색엔진을 적용하고 다양한 시맨틱 언어 자원을 활용한 KOSEN 통합검색 서비스에 대해 논한다. 2장에서는 검색엔진 설계 및 검색 기술을 설명하고 3장에서 서비스를 소개하며 4장에서 결론을 맺는다.

## II. 관련 기술

### 1. 검색엔진

#### 1.1 색인설계

통합검색 대상이 되는 KOSEN 서비스의 자료는 분석자료, 수집자료, 행사정보, 논문모집공고, 채용정보, WHATIS, GTB(글로벌동향브리핑), 사업공고, 동영상자료와 해외출장자료이다. FAST 색인기의 JDBC Connector는 오라클 DB에 접근하여 데이터를 읽어들이고, 문서처리기는 언어치리, 형태소분석 등의 과정을 처리한다. 자료는 검색엔진에서 각각의 collection 단위로 정의되며 이는 통합검색을 위해 하나의 뷰(view)로 묶인다. 각 자료는 공통적으로 제목, 요약, 작성자 필드를 가지고 있으며 이는 검색의 대상이 되는 필드이다.

#### 1.2 검색기능

FAST 검색엔진은 미리 정의한 필드를 기준으로 빠른



### III. KOSEN 통합검색서비스

빠른 속도를 보장하고 다양한 검색 부가기능을 구현하여 사용자 검색 편의성 및 효율성을 고려하고, 언어자원을 활용한 시맨틱 기술을 접목한 KOSEN 통합검색서비스는 약 26만여건의 자료를 제공한다.

본 서비스의 특징은 언어자원을 활용하여 검색 효율성을 최대화 하였으며, 검색결과를 하나의 표로 요약하여 제시하고 이를 통해 조건검색 및 세부탐색을 용이하도록 한 점이다. 그림 3은 서비스에서 제공하는 다양한 기능을 보여주는 화면이다.



▶▶ 그림 3. KOSEN 통합검색 화면

### IV. 결론

본 논문에서는 과학기술분야 전문분석자료, 수집자료, 동향정보를 제공하는 KOSEN(한민족과학기술자네트워크)의 통합검색 서비스 구축에 대해 논하였다. 사용자의 검색 편의성 및 효율성을 최대화하기 위해 어휘기능, 연구주제망 등 언어자원을 활용하여 시맨틱 기법을 적용하였고, 빠른 검색속도와 다양한 검색 부가기능을 제

공한다. 향후 개인의 검색성향 및 행태를 분석한 개인화 검색의 접목이 필요하다.

### ■ 참고 문헌 ■

- [1] G Antoniou, F Van Harmelen, 2004. A Semantic Web Primer. MIT Press
- [2] 임지희, 최호섭, 옥철영, “어휘망(U-WIN)의 구분 관계 자동구축”, 정보과학회논문지, v.35, no.10, pp.627-635, 2008