

과학기술정보 개인화 서비스 설계 및 구현

한희준*, 최성필**

*한국과학기술정보연구원 융합서비스센터

**경기대학교 문헌정보학과

e-mail:hhj@kisti.re.kr

A Study on Design and Implement of S&T Information Personalization Service

Heejun Han*, Sungpil Choi**

*Convergence Service Center, Korea Institute of Science and Technology Information

**Dept of Library and Information Science, Kyonggi University

요 약

방대한 정보를 사용자에게 제공하기 위해 검색 엔진은 다양한 알고리즘을 통해 사용자마다의 최적화된 정보를 구성한다. 과제, 논문, 특허, 연구보고서 등 과학기술정보를 서비스 하는 주체 역시 나름의 검색 알고리즘으로 정보를 제공하지만, 질의어와 문서간의 적합도만을 측정하여 검색 결과를 제시할 뿐 사용자의 관심 분야나 요구를 반영하지 않고 있다. 특히 관심 분야에 적합한 과학기술정보를 사용자가 접근하기 쉽게 제공하는 것은 매우 중요하다. 본 논문에서는 사용자 관심분야를 서비스 이용행태로부터 결정하여 이를 과학기술정보 개인화에 반영하는 서비스에 대해 제안하였다. 이를 위해 실시간 관심분야 추적, 관심 태그 클라우드 제공, 관심분야 기반 추천정보 제공, 검색 결과 개인화 네 가지 기능으로 구성된 과학기술정보 개인화 서비스를 설계하고 구현하였다.

1. 서론

국내 과학기술정보서비스 NTIS는 과제, 논문, 특허, 연구보고서, 기술이전, 시설장비 정보 등을 제공하고 연구자, 학생이 폭넓게 사용하는 지식 포털이다. 500만 건 이상의 과학기술정보에 대해 동일한 질의어에 같은 검색결과를 동일한 형태 또는 순위로 제공하는 것은 관심분야가 제각각인 사용자의 요구를 만족시키기 어렵다. 따라서 본 연구에서는 사용자 관심분야를 국가과학기술표준분류 기반으로 실시간으로 추적하고, 이를 이용해 과학기술정보를 개인화하는 서비스를 구현하고자 한다. 여기서 개인화란 사용자가 의도하는 바를 파악하여 검색결과 또는 기능적으로 제공하는 정보를 최적화시켜 주는 것을 의미하며, 검색결과란 입력한 질의어에 적합한 정보뿐만 아니라 입력 의도와 관련된 추천 및 제안 정보도 포함한다. 개인화 서비스와 관련하여 최근까지 이용자 프로파일 분석, 질의어 확장, 협업 필터링, 유사문서 클러스터링 기법 등을 활용한 검색결과 재조정과 정보 추천에 대한 연구는 많이 수행되어 왔다. 하지만 개인의 검색 의도를 추적하여 적용한 정보 개인화에 대한 연구는 부족하다. 본 연구에서는 개인 관심분야 추적 및 정보 개인화 프로세스를 정의한 후 과학기술정보서비스 사용자의 관심분야를 반영한 정보를 제공하기 위한 서비스 기능적 구성과 설계를 제안하며 이를 구현하였다. 2장에서는 개인화 서비스를 위한 DB테이블 설계에 대해 논하고 3장에서 과학기술정보개인화 서비스 구성 및 기능에 대해 제안하며 4장에서 결론을 맺는다.

2. 개인화 서비스를 위한 DB 설계

정보서비스는 검색 대상 데이터베이스와 데이터를 색인하여 검색결과를 제공하는 검색 엔진, 검색된 결과를 사용자에게 제공하는 사용자 인터페이스의 구현이 필요하다. 본 연구에서는 NTIS에서 제공하는 정보와 동일한 데이터를 활용한다. 이를 위해 NTIS에서 제공하는 과학기술정보를 모두 검색할 수 있도록 제공하는 NTIS 공개 API(OpenAPI)를 이용하여 검색 기능을 구현하였다. 검색 API를 활용하기 때문에 별도의 과학기술정보 관리를 위한 DB는 필요 없으나, 사용자정보, 이용행태별 관심정도 점수, 사용자 관심분야 등의 정보관리를 위해 데이터베이스에 대한 설계가 필요하다. 표 1은 서비스를 위해 설계된 6개 DB 테이블과 각 테이블 별 필드의 구성을 나타낸다.

<표 1> 과학기술정보 개인화서비스 DB 테이블

테이블명	필드명	데이터 타입	설명
P_USER_INFO	USER_ID	varchar2(20)	사용자 아이디
	USER_NM	varchar2(20)	사용자 이름
	USER_EMAIL	varchar2(20)	사용자 이메일
	ORG_NM	varchar2(20)	소속기관명
	REGIST_DT	date	가입일시
P_INTEREST_CODE	CODE	varchar2(2)	이용행태 코드
	CODE_NM	varchar2(50)	이용행태
	CODE_WEIGHT	number	관심정도 점수
P_USER_INTEREST	SEQ	number	시퀀스 번호
	USER_ID	varchar2(20)	사용자 아이디
	SCIENCE_NM	varchar2(200)	관심분야명
	SCORE	number	관심정도 점수
	DAY_SCORE	number	관심정보 점수 일일 누적

P_USER_SCRAP	TOTAL_SCORE	number	관심정보 점수 전체 누적 이용행태 발생일시
	REGIST_DT	date	
	USER_ID	varchar2(20)	사용자 아이디
	DOC_ID	varchar2(50)	문서 고유번호
	DOC_TYPE	varchar2(10)	문서 종류 (예. 과제, 논문 등)
P_USER_TAG	DOC_TITLE	varchar2(500)	문서 제목
	REGIST_DT	date	스크랩 일시
	USER_ID	varchar2(20)	사용자 아이디
P_USER_TAG	TAG	varchar2(200)	태그 키워드
	TAG_COUNT	number	태그 키워드 출현 빈도
P_USER_LOG	SEQ	number	시퀀스 번호
	USER_ID	varchar2(20)	사용자 아이디
	ACT_FLAG	varchar2(2)	이용행태 코드
	ACT_SCORE	number	관심정도 점수
	DB_TYPE	varchar2(10)	이용 DB
	USER_QUERY	varchar2(200)	질의어
ACT_DT	date	로그일시	

3. 개인화 서비스 구성 및 기능 구현

과학기술정보 검색 공개API를 이용하여 정보를 검색하고, 제공된 정보에 대한 사용자 이용행태를 분석하여 관심분야를 추적하며 결정된 개인 관심분야를 다시 정보를 개인화하는데 활용하는 일련의 선순환 프로세스를 구현하는 개인화 서비스는 그림 1과 같은 구성도를 가진다.

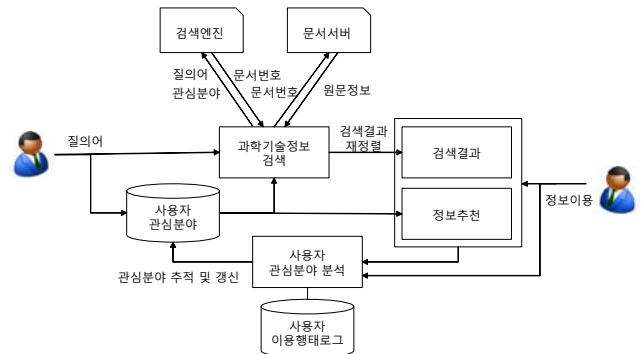


(그림 1) 서비스 구성도

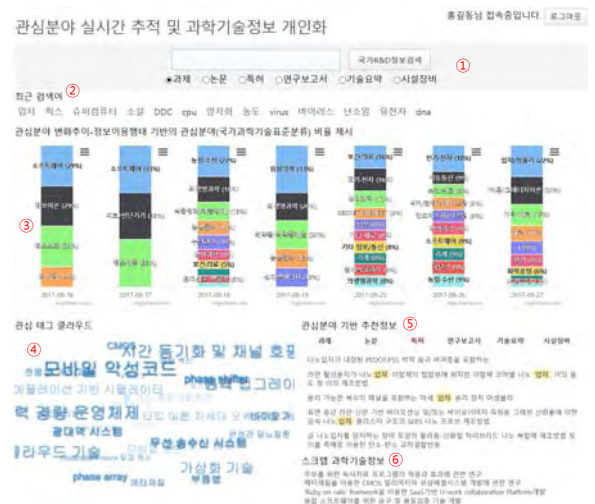
서버측은 회원정보를 관리하는 ‘사용자정보 관리’, 과학기술정보를 이용하는 이용 행태별 로그를 분석하여 관리하는 ‘이용행태 분석’, 분석된 결과를 이용해 국가과학기술 표준분류에 관심도 점수를 계산하여 관리하는 ‘개인 관심분야정보 관리’, 스크랩 정보와 개인 관심 키워드를 저장하는 ‘스크랩정보 관리’와 ‘키워드태그 관리’, NTIS 공개 API를 이용하여 검색결과로 메타데이터와 원문정보를 가져오는 ‘과학기술정보 검색’, 개인 관심분야에 따라 검색결과를 재순위화하는 ‘검색결과 재정렬’ 부분으로 구성된다. 클라이언트 측은 개인화된 과학기술정보를 제공하는 ‘검색결과 제공’, 정보 이용 행태 기반의 개인 관심분야 비율을 제시하는 ‘개인 관심분야 변화추이’, 검색 기능을 수행하지 않고도 관심분야 기반의 정보를 제공하는 ‘과학기술정보 추천’, ‘스크랩 정보 제공’, 사용자가 자주 이용하는 과학기

술정보가 지닌 키워드를 빈도순으로 저장하여 제공하는 ‘관심 태그 클라우드’, 검색 히스토리를 통해 검색을 바로 수행하도록 도와주는 ‘최근 검색어’ 기능으로 구성된다.

사용자에게 개인화된 과학기술정보를 제공하기 위해 가장 중요한 요소는 이용행태 분석을 통한 개인 관심분야정보 추적 및 갱신, 관심분야를 적용한 정보의 개인화이다. 그림 2는 개인 관심분야를 추적하고 정보를 개인화하는 프로세스 흐름도를 나타내며 그림 3은 과학기술정보 개인화 서비스 기능화면이다..



(그림 2) 정보 개인화 프로세스 흐름도



(그림 3) 과학기술정보 개인화 서비스 기능화면

4. 결론

방대한 정보 안에서 개인 관심분야에 적합한 정보만을 최적화하여 제공하기 위해 과학기술정보를 이용하여 개인화 서비스 구현을 위한 설계와 기능 구성에 대해 제안하였다. 향후 제안된 서비스와 NTIS와의 개인화 정보 제공 성능적 측면에서의 평가가 이루어져야 할 것이다.

참고문헌

- [1] 윤성희, “질의기반 사용자 프로파일을 이용하는 개인화 웹 검색”, 한국산학기술학회논문지, 제17권 제2호, pp.690-696, 2016.
- [2] 국가과학기술지식정보서비스(NTIS), www.ntis.go.kr