

차세대성장동력산업을 위한 전문정보 수집시스템에 관한 연구

박영욱*·최현규**

I. 서론

인터넷을 이용하는 사람들 중에는 전문정보의 습득을 위해 RSS 서비스를 활용하는 경우가 많다. 하지만, 정작 필요한 웹사이트는 RSS 서비스를 실시하지 않거나, 이용 대상 웹사이트가 수십 개에 이를 때는 전문정보 수집시스템을 활용하여 구축한 데이터베이스를 이용하는 것이 더욱 효과적일 것이다. 또한, 지식기반 정보화 사회에서 우리나라가 추진하는 차세대 성장동력산업의 성공을 위해서는 정확한 정보의 신속한 입수는 필수이므로, 이를 위한 전문정보 수집시스템이 필요한 실정이다. 따라서, 본 고에서는 차세대 성장동력산업을 위한 전문정보 수집시스템 구축에 관하여 고찰하였다

II. 차세대성장동력산업을 위한 전문정보 수집시스템의 필요성

1. 전문정보의 중요성

1) 기술정보 활동과 차세대성장동력산업

산업구조가 첨단화, 지식산업화 됨에 따라 기술정보는 연구개발의 성과를 좌우하는 핵심요소가 되었고, 기술정보의 수요형태도 심층정보 위주로 전문화, 고도화, 다양화가 이루어지는 추세다. 미국국립과학재단(NSF)의 조사결과에 따르면, 미국화학공학 연구자의 연구활동시 기술정보의 입수와 전달에 50.9%의 시간을 사용하여 다른 연구개발 활동에 비해 많은 시간을 기술정보에 투자하는 것으로 밝혀졌다.

차세대성장동력산업은 국민소득 2만 달러 시대에 도달하기 위해 국가적으로 추진하는 전략 산업으로, 선정된 10개의 산업 중에는 선진국과 비교하여 우리나라가 앞선 분야도 있지만, 거의가 열세인 실정이다. 따라서, 연구개발자들에게 종합적이고, 체계적인 방법으로 선진 기술정보를 제공하는 것은 차세대성장동력산업이 성공할 수 있는 단초가 될 것이다.

* 박영욱, 한국과학기술정보연구원 연구원, 042-828-5057, ywpark@kisti.re.kr

** 최현규, 한국과학기술정보연구원 책임연구원, 02-3299-6026, hkchoi@kisti.re.kr

2) 기술동향이 중요한 차세대성장동력산업

한국과학기술정보연구원(이하 "KISTI")의 조사¹⁾에 의하면, 프로젝트 수행과 기술의 상업화 과정에서 연구자들이 많이 필요로 하는 정보유형은 기술동향과 시장정보인 것으로 나타났다. 이것은 연구개발이 곧 상품개발이라는 차세대성장동력산업 관련 연구자들에게도 그대로 적용될 것이다. <표1>을 보면, 차세대성장동력산업은 산업별로 2~4개의 하위 연구개발 분야를 선정하였는데, 모두 40개로 “제품”이라는 명칭으로 상업화를 최종 목표로 한다.

<표 1> 차세대성장동력 산업분야별 전략제품(2004년 기준)

산업분야	전략제품	산업분야	전략제품
디지털TV방송	DTV 수상기	차세대이동통신	휴대인터넷/4세대 이동통신시스템
	DMB 단말기		차세대 융합유대 단말기
디스플레이	LCD		유비쿼터스 센서네트워크
	PDP		텔레매틱스서비스 고도화시스템
	OLED	차량탑재형 텔레매틱스	
지능형로봇	자동차제조용로봇	지능형홈네트워크	홈플랫폼
	청소방법용 로봇		미들웨어
	IT기반 지능형 로봇		지능형 정보가전
	네트워크기반 휴머노이드		홈네트워킹
미래형자동차	하이브리드자동차	디지털콘텐츠 / SW솔루션	콘텐츠 창작
	연료전지자동차		콘텐츠 보호유통
	지능형자동차		콘텐츠 제작
차세대전지	이차전지		미들웨어
	연료전지		기반소프트웨어
차세대반도체	SoC		응용소프트웨어
	반도체장비	이종장기생산용 복제되지	
	나노반도체	분석/진단용 바이오칩	
	SiC 반도체	약물전달시스템	
	IT SoC IP	세포치료제	
	IT SoC 설계용 SoC CAD	바이오신약	

1) 이준영 외, “국내 연구자들의 소속기관과 정보이용행태의 상관관계 분석(2004)”

2. 전문정보 수집시스템의 필요성

1) 미약한 기술동향 및 시장정보의 국내정보공급체계

<표2>는 정보유형별 이용도와 정보추구경로 이용도간 피어슨 상관계수를 나타낸 것이다. 문헌정보를 이용하기 위한 전문정보 서비스 체계는 잘 갖춰진 반면(.346), 차세대성장동력산업 관련 연구자들이 많이 필요로 하는 기술동향 및 시장정보는 상대적으로 그 기반이 미약한 것으로 드러났다.

<표 2> 정보유형별 이용도와 정보추구경로 이용도간 피어슨 상관계수 (KISTI, 2004년)

정보유형	정보추구경로				
	전문정보 서비스	웹검색서비스	관련기관/ 연구소 홈페이지	행사/전시회/ 세미나	외부전문기관
과학기술문헌정보	.346	.090	.156	.262	.097
생산기술정보	.078	.074	.115	.118	.168
특허정보정보	.216	.094	.061	.160	.188
기술동향	.139	.185	.206	.195	.111
시장동향	.088	.183	.189	.141	.184
정책정보	.118	.118	.258	.204	.195

정보이용자들은 기술동향 정보를 입수하기 위해 관련기관이나 연구소 홈페이지를 많이 활용하고 (상관계수 0.206), 다음으로 외부 행사에 참여하거나(.195), 일반 웹검색(.185)을 이용한다. 시장정보의 경우, 1순위는 기술동향과 마찬가지로 관련기관 홈페이지를 이용하는 것이지만, 2순위는 외부 전문기관에 의뢰하는 것으로 나타났다.

2) 차세대성장동력산업을 위한 전문정보 제공 웹사이트의 강화

KISTI는 차세대성장동력산업을 위한 전문정보를 제공하고자 특화된 웹사이트(이하 "Next10", <http://next10.yeskisti.net>)를 구축, 운영하고 있다. 현재 Nex10은 기술&시장, 정책동향, 해외동향, 학술논문, 연구보고, 특허정보, 행사정보 등의 다양한 정보를 제공하고 있지만, 정보의 출처가 KISTI 내로 한정되어 있다는 단점이 있다. 또한, 이용자 만족도 조사¹⁾ 결과, 연구 개발자들은 KISTI 보유정보 뿐만 아니라, 여러 개의 웹사이트를 통해 정보활동을 하는 것으로 나타났다. 따라서, 원스톱 서비스를 지향하는 Next10이 이용자들에게 폭넓은 정보를 제공하기 위해서는 정보의 출처를 더욱 다양화하여야 하고, 이를 위해 전문정보 수집시스템을 적용할 필요가 있다.

1) 차세대 성장동력사업단 실무자 및 연구개발자들을 대상으로 2006년 4월19일 실시

Ⅲ. 차세대성장동력산업을 위한 전문정보 수집시스템의 구축

1. 전문정보 수집시스템의 개요

1) 기능

전문정보 수집시스템은, 인터넷에 산재한 다양한 정보를 수집 및 분류, 데이터베이스화하여 이용자들이 관심분야의 정보를 쉽게 입수할 수 있도록 도와주는 시스템이다. 이용자가 정보 출처 사이트에 직접 접속하지 않아도 된다는 점에서 RSS¹⁾ 서비스와 비슷하지만, RSS 서비스를 제공하지 않는 웹사이트에서도 정보입수가 가능하고, 미리 설정된 키워드에 따라 분류가 가능한 장점이 있다. 따라서, RSS 서비스보다는 적용범위가 넓고, 분류 기능을 통한 맞춤형 정보 제공이 가능하다는 특징이 있다.

2) Next10의 수집시스템 이용

<그림1>은 차세대성장동력산업을 위해 전문정보를 제공하는 Next10이 수집시스템을 적용했을 경우를 보여준다. 파란색 화살표는 정보의 흐름을 나타낸다. 국내외에서 수집된 정보는 자료유형별, 10대 산업별로 나뉘어진 뒤, Next10에서 최종 서비스된다. 이때, 관리자는 수집시스템에 접속하여 수집된 정보를 재검증하거나, 시스템의 오류사항을 확인, 수정할 필요가 있다. 만약, 이러한 절차를 생략한다면 Next10은 가비지(garbage)성 데이터로 인해 이용자들에게 혼란을 야기할 수 있기 때문이다.



<그림 1> 수집시스템과 정보의 흐름

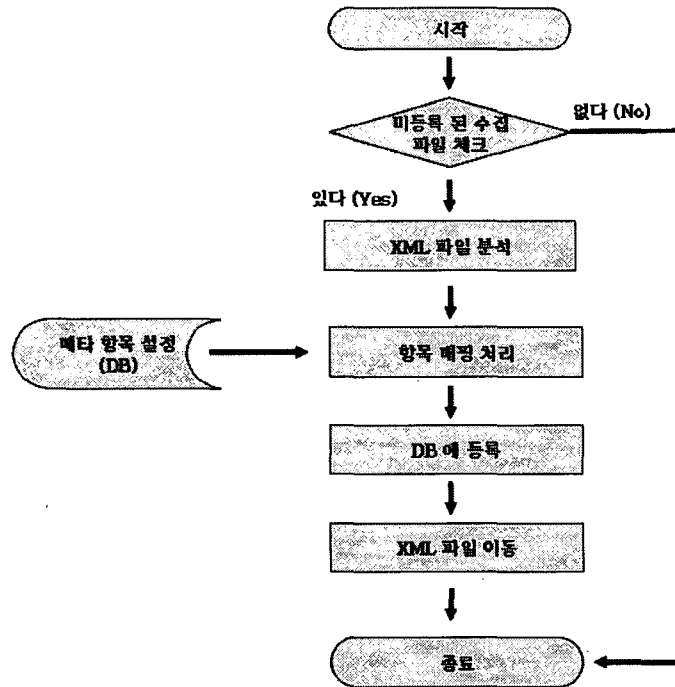
1) Rich Site Summary 또는, Really Simple Syndication

2. 수집시스템 알고리즘

차세대성장동력산업 관련 전문정보가 수집되어 정보이용자에게 전달되기까지는 수집, 데이터 베이스(DB) 등록, 분류, 서비스의 4단계의 과정을 거쳐야 한다. 본 절에서는 수집, 데이터, 그리고, DB등록에 대한 알고리즘을 설명하였다..

1) 수집

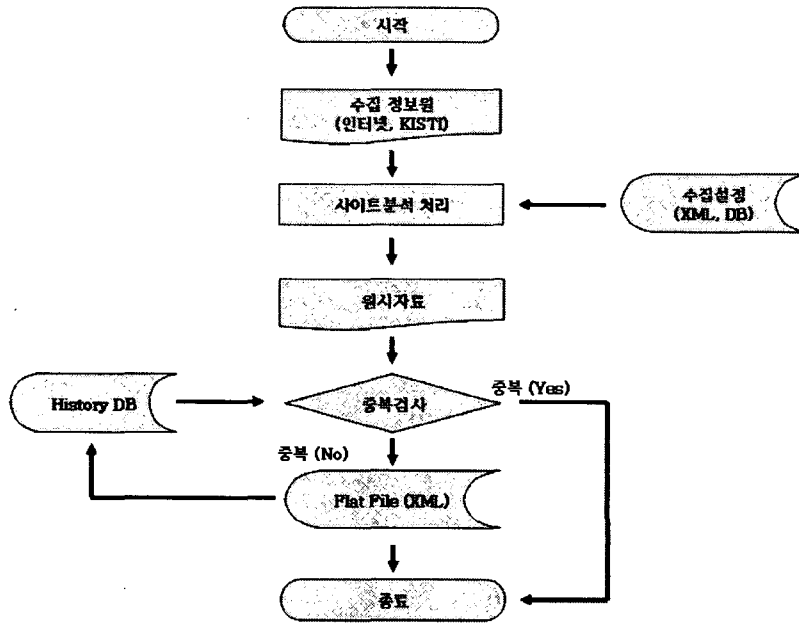
수집단계는 KISTI 또는 인터넷 웹사이트의 정보를 수집시스템의 검색로봇을 통해 입수하는 과정이다. 검색로봇은 수집설정의 조건에 따라 해당 웹사이트를 방문하여 신규정보를 검색한다. 만약, 신규정보가 존재하면 XML(eXtensible Markup Language) 파일 형태로 수집하여, DB등록 단계로 가게 된다.



<그림 2> 수집시스템의 정보수집 과정

2) DB 등록

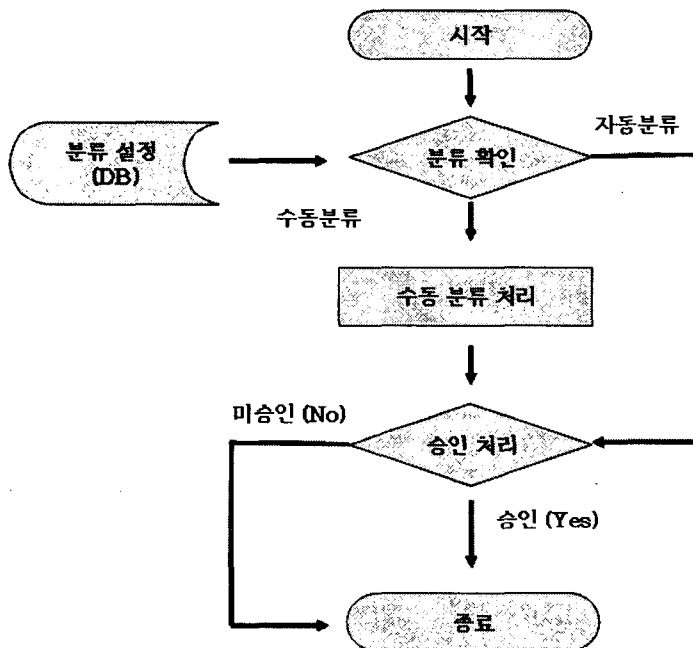
수집단계를 마친 정보를 무조건적으로 DB에 등록하지 않고, DB등록 단계를 설정한 이유는, 수집대상 사이트에서 제공하는 정보의 성격과 홈페이지 제작 형태가 다양하기 때문에, 이에 따른 메타정보의 수집설정 조건이 사이트별로 달라지기 때문이다.



<그림 3> 수집시스템의 DB등록 과정

3) 분류

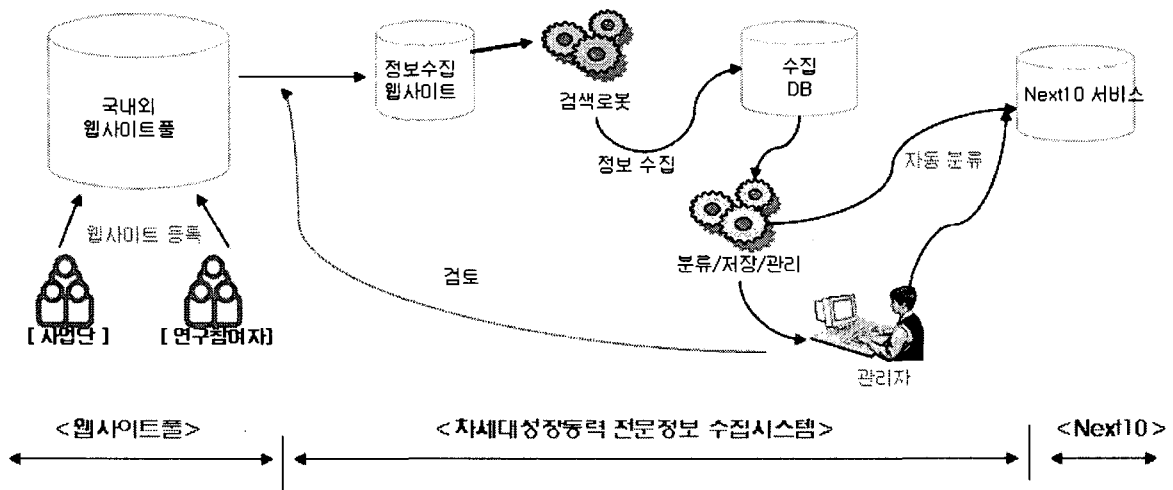
분류는 시스템에 의해 자동으로 수행하거나, 관리자가 직접 분류할 수 있는 수동분류 기능이 있어야 한다. DB에 등록된 정보를 10 개의 차세대성장동력산업, 또는 40 개의 제품 분야로 분류하기 위한 분류키워드가 미리 설정되어야 하고, 향후 기술변화에 부응하기 위해 설정값은 수시 변경이 가능해야 한다. 또한, 기술의 융복합화를 반영하기 위한 중복 분류도 가능해야 할 것이다.



<그림 4> 수집시스템의 분류 과정

3. 정보수집 웹사이트의 선정

정보수집 웹사이트는 차세대성장동력 전문정보를 제공하는 웹사이트로서, 수집시스템에 등록이 되어야만 이곳에서 생산되는 정보를 입수할 수 있다. 정보의 바다, 인터넷에는 무수히 많은 웹사이트가 존재하므로, 차세대성장동력산업과 관련된 정보를 제공하는 웹사이트를 신중히 선정하여 이용자들이 양질의 정보서비스를 받을 수 있어야 한다. 이를 위해 차세대성장동력 사업단과의 협력을 바탕으로 웹사이트 풀(pool)을 만들 것을 제안한다. 웹사이트 풀에서 전문정보 수집시스템에 등록하기 위한 웹사이트의 선별기준으로, 첫째 기술동향 및 시장정보의 제공 여부, 둘째 단순 기사성 정보(1)가 아닌 실질적으로 도움이 될 만한 정보를 제공하는지를 살펴봐야 할 것이다. <그림 5>는 웹사이트 풀을 첨부했을 때의 차세대성장동력 전문정보 수집시스템의 개략도이다.



<그림 5> 차세대성장동력 전문정보 수집시스템

IV. 결론

21세기 지식기반 정보화 사회에서 우리나라가 추진하는 차세대성장동력산업이 성공하기 위해서는 정보의 신속한 입수는 중요한 요소가 되었다. 하지만, 다양한 정보출처의 존재는 연구자들에게는 검색활동으로 인한 부담이 되므로, 필요 정보를 한 곳에서 제공해주는 서비스가 필요하다. 따라서, 차세대성장동력산업을 위한 전문정보 수집시스템을 구축한다면, 다음과 같은 효과를 기대할 수 있다. 첫째, 정보출처의 다양화로 이용자들에게 폭넓은 지식정보를 제공한다. 정보수집 웹사이트는 한 사람이 선정한 것이 아니고, 여러 사람이 추천한 다수의 웹사이트로 구성되므로, 연구의 시너지효과를 기대할 수 있다. 둘째, 체계적이고 효율적인 방법으로 정보가 수집되고, 관리되므로 정보의 정확성이 향상된다. 전문정보 수집시스템은 검색로봇을 통해 미리 설정된 조건에 따라 정보를 수집하고, 수집된 정보는 키워드에 따라 10개 산업 및 40개 제품 분야로 분류되므로, 더욱 정확한 정보의 제공이 가능해진다. 마지막으로, 이용자는 다양한 정보를 원스톱으로 제공받으므로 편리하게 정보를 이용할 수 있다.

1) 단순 기사성 정보는 포털사이트에서 조차 쉽게 검색되므로, Next10은 그러한 정보제공을 지양한다.

참고문헌

- 이준영 외(2004), “국내 연구자들의 소속기관과 정보이용행태의 상관관계 분석”, 2004KOSTI 워크샵
- 박영욱 외(2005), “차세대성장동력산업 정보서비스를 위한 기반 분석”, 2005 추계기술혁신학회
미래선도기술네트워크 홈페이지, <http://next10.yeskisti.net>
- 권이남 외(2005) “RSS(Really Simple Syndication) 응용기술”
차세대성장동력산업 웹사이트 설문조사 결과 (2006.4)
- 조현양 (1994), “국가 과학기술 정보정책과 출연(연)의 기술정보활동”, 「94한국데이터베이스 학회
지」, pp25-47