

[솔루션 발굴] N-SearchPlus

**한국소프트웨어진흥원
공개SW기술지원센터**

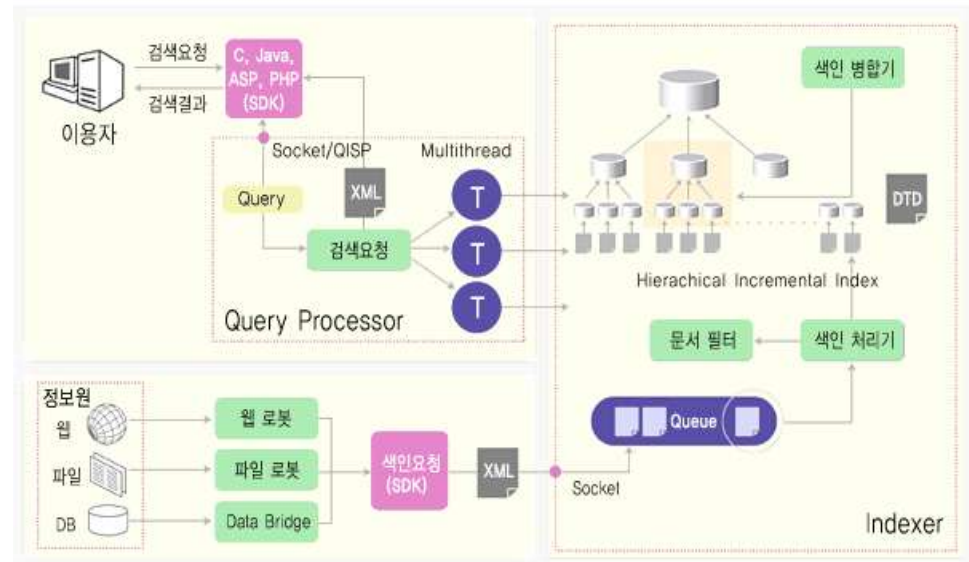
<Revision 정보>

일자	VERSION	변경내역	작성자
2007. 7. 30.	0.1	초기 작성	김용규

기업 / 제품 정보

회사명*	(주)다음기술	웹사이트*	http://www.ntechs.com/
주소*	서울시 강남구 논현동 86-6 마루빌딩 5층		
연락처*	02-541-0840	E-MAIL*	webmaster@ntechs.com
솔루션 및 서비스명	N-SearchPlus		
솔루션 및 서비스 설명	0 제품 소개 - N-SearchPlus는 (주)다음기술에서 개발한 순수 국산 전문(全文:Full Text) 검색엔진이다. 분산 구조로 설계되어 있어서 사용자와 데이터의 증가에 유연하게 대처가 가능하며, Hierachical Incremental Indexing 방식으로 고속 색인과 색인 중 검색이 아주 빠르다. - 660405-1029430 색인 복구기능이 있어서 색인 중 시스템 이상이 발생하더라도 시스템이 정상 복구되면 문제가 발생한 시점부터 색인을 계속해준다. - 유니코드 기반의 색인 구조를 갖고 있으며 한글 형태소 분석기, n_gram 방식의 색인을 동시에 지원하여 다양한 사용자 요구사항 및 다국어 환경에 효과적인 대처가 가능하다. - JAVA, C, Active X 등의 API를 제공하여 다양한 웹 프로그램 및 응용 프로그램과 쉽게 연동이 가능하며, 기초적인 XML 지식만 있으면 색인 구조를 설계하여 문서를 색인 할 수 있다. - 다양한 형식의 문서에 대한 문서필터를 기본으로 제공하여 텍스트 문서 뿐 아니라 각종 편집기 문서의 내용에 대해서도 자동으로 색인을 색원한다. - 웹 서비스뿐 아니라 문서관리시스템(EDMS), 지식관리시스템(KMS) 등 응용 프로그램과의 효과적인 통합을 염두에 두고 설계 및 개발된 솔루션이다.		

0 구성도



0 기능적 특징

- 필드/다중필드 검색 지원
- 최고 성능의 형태소 분석기 지원
- 형태소 분석을 통한 자연어 검색
- 웹 기반의 관리도구 / 시소러스 관리기
- 시소러스를 통한 질의어 확장
- Wildcard 검색 지원 (전방, 후방, 전후방)
- Relevance Ranking 지원
- 인접도(near) 검색 지원
- 문서 필터 지원 (PDF, HWP, 오피스 문서 등)
- 지식검색 지원

0 구조적 특징

- 전문(全文:Full Text) 색인 및 검색
- Hierarchical Incremental 실시간 색인 및 색인 중 검색 지원
- DATE에 대해 구조적으로 기본 정렬 지원
- 확장이 용이한 XML/Socket 기반의 분산구조 지원
- 단일 소스 코드로 멀티 플랫폼 지원 (윈도우, 유닉스, 리눅스)
- 응용 프로그램 통합에 용이한 질의 구조 (QISP 프로토콜)
- 데이터 및 색인의 압축을 통한 저장 공간 절약
- 멀티쓰레드 질의 처리 서버로 사용자의 질의 처리 코딩 불필요
- 형태소 분석과 n_gram 방식 색인 동시 지원
- 검색 결과 캐쉬 기능으로 검색 성능 향상
- 다양한 개발 언어에 대한 인터페이스 지원 (C, Java, Active X)
- 관련 정보검색 솔루션 연동 (웹 로봇, 자동 분류, 메일링 엔진)

0 서치플러스 세부기능 - 색인

	실시간 색인	Hierachical Incremental Indexing 방식으로 문서를 실시간으로 고속 색인하며, 자동으로 텍스트나 편집기 문서에서 색인어를 추출한다.
	전문색인	원문의 일부분이 아닌 전체 문서를 대상으로 색인한다.
	형태소 분석기	국내 최고 성능의 한글 형태소 분석기를 채용하여 색인과 검색의 질을 높여준다.
	Stemming/Stopwords 처리	어근처리나 불용어 처리, 불필요한 단어제거 등을 통해 색인 크기를 줄일 수 있다.
	색인중 검색기능	색인 도중에도 성능에 거의 영향 없이 빠른 검색이 가능하다
	필드단위의 색인	색인 분할 및 필드단위로 색인 및 검색이 가능하다.
	다국어 지원	Unicode 기반으로 설계되어 다국어 처리가 가능하다.
0 서치플러스 세부기능 - 검색		
	전문(全文)검색	검색이 원문 일부가 아닌 전체 문서를 대상으로 한다.
	불린(Boolean) 검색	불린 기호를 이용 좀더 구체적인 질의어를 만들 수 있다.
	중요도 검색	질의어와 관련성이 높은 문서 순서로 검색 결과를 정렬할 수 있다.
	영문 대소문자 구분	대소문자 구분 없이 결과를 찾아준다.
	자동 Stemming	형태소 분석기를 이용한 Stemming 처리로 검색 정확도를 좀더 높여준다.
	시소러스 적용	유사어, 동의어 사전을 구축하여 질의어 확장 검색을 지원한다.
	NEAR 검색 지원	문서 내에서 단어 사이의 인접도에 따른 검색을 지원한다.
	순쉬운 검색 화면 구축	최소한의 API 만으로 쉽고 빠르게 개발이 가능하다.
	순쉬운 결과 화면 구축	결과를 XML 형태로 제공하여 개발자가 쉽고 빠르게 검색결과 페이지를 꾸밀 수 있다.
	XML Parser 제공	검색결과 XML을 Parsing 해서 사용하기 위한 인터페이스를 제공한다.
	제한검색 지원	검색 결과를 여러 가지 조건으로 제한할 수 있어서 보다 정확한 결과를 찾아낼 수 있다.
	결과 내 검색	검색한 결과 중에서 다시 검색을 수행할 수 있다.
	기본 정렬 기능	검색결과를 날짜(DATE)나 중요도(Relevance)로 정렬하는 기능을 구조적으로 기본 제공하여 별도의 처리 없이 빠른 결과를

		제공한다.
	필드 검색	원하는 필드 혹은 필드 조합을 대상으로 문서를 검색할 수 있다.
	Wildcard 검색	전방일치(A*), 후방일치(*A)와 같은 Wildcard 검색을 지원한다.
	결과 정렬	결과를 정의된 모든 필드 값을 기준으로 정렬할 수 있다.
	검색결과 캐싱	최근 질의된 결과를 검색엔진 내부에 저장함으로써 동일한 검색이나 결과내 검색시 좀더 빠르게 결과를 돌려준다.
	구문검색	구문 단위로 검색을 할 수 있습니다.
	Library API	C, JAVA, Active X, PHP 등 다양한 개발 인터페이스를 제공한다.
	검색결과 하이라이트	검색결과에서 질의어를 하이라이트 시켜주는 API를 제공한다.
시스템운영환경	O/S (지원가능)	Linux, Unix, Windows 기반 사용 가능
	Database	Oracle, MS-SQL, Mysql 등 가능
	H/W 최소사양	HDD 1G이상, CPU P3 500Mhz 이상, 최소128MB RAM
	기타 App	Java, JSP, ASP, PHP, C
기타환경		

첨부 #1 :

- 제품소개서 & 참조사이트

<http://www.ntechs.com/>