# Ksearch: 한국형 검색엔진

홍성진

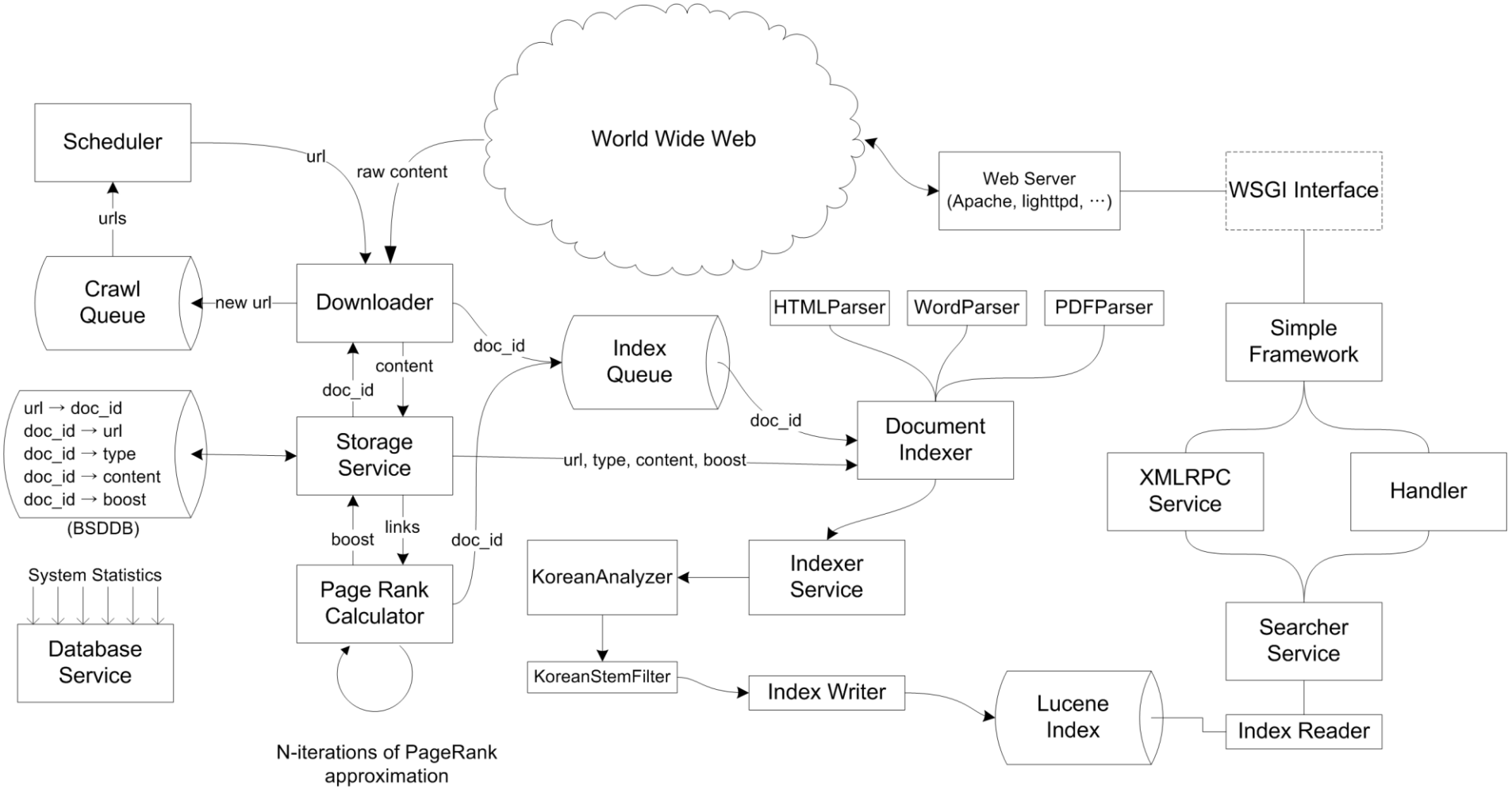
# Project Environment

- 5명의 개발자
- Version Control - Subversion
- Issue Tracking - Trac
- Code Review – ReviewBoard
- Continuous Integration – BuildBot

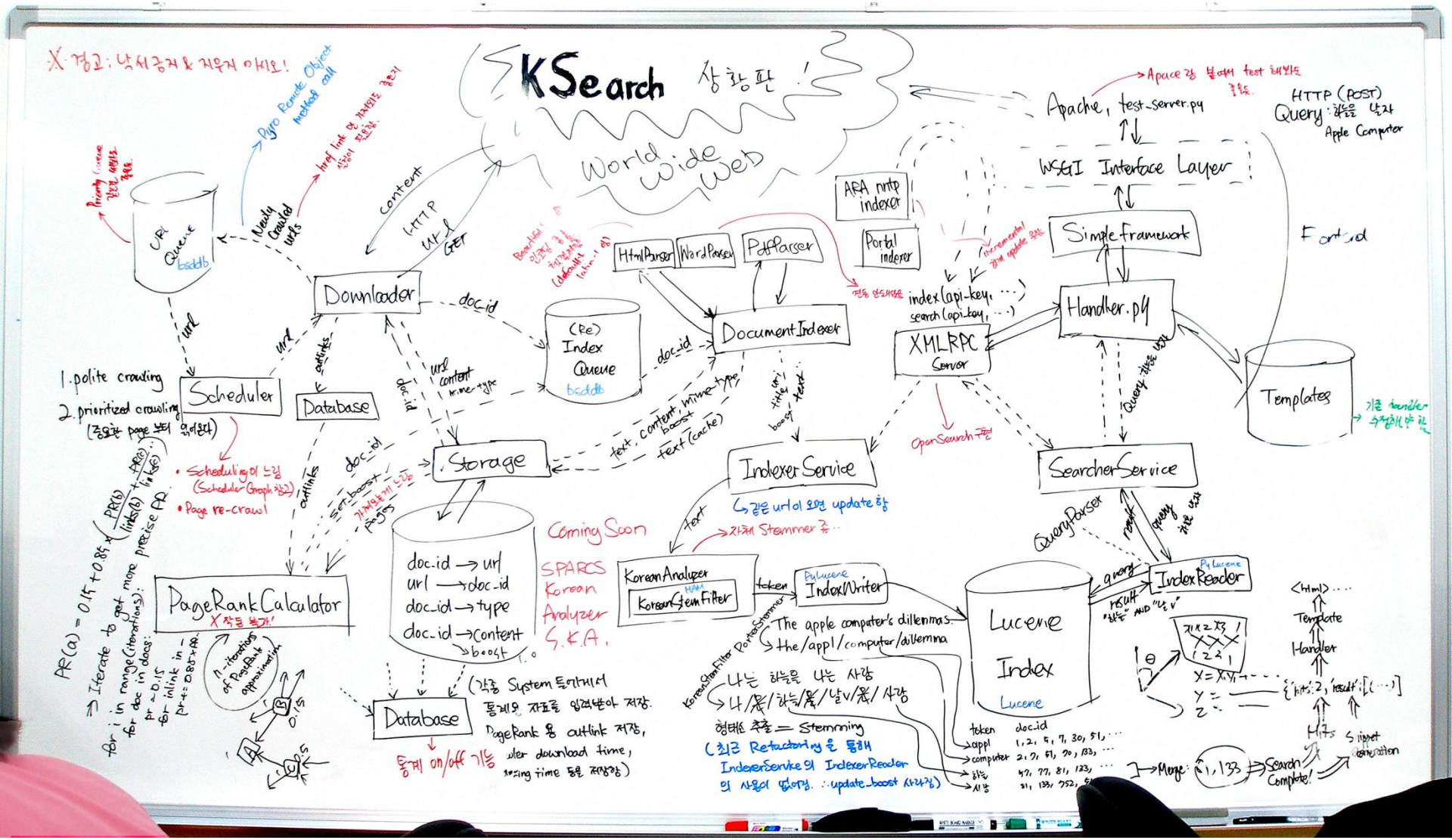
# KSearch Dependency

- Berkeley DB 4.6
- PyLucene
- Pyro
- Django (Template engine only)
- Eventlet
- Xpdf, antiword
- matplotlib, pygooglechart

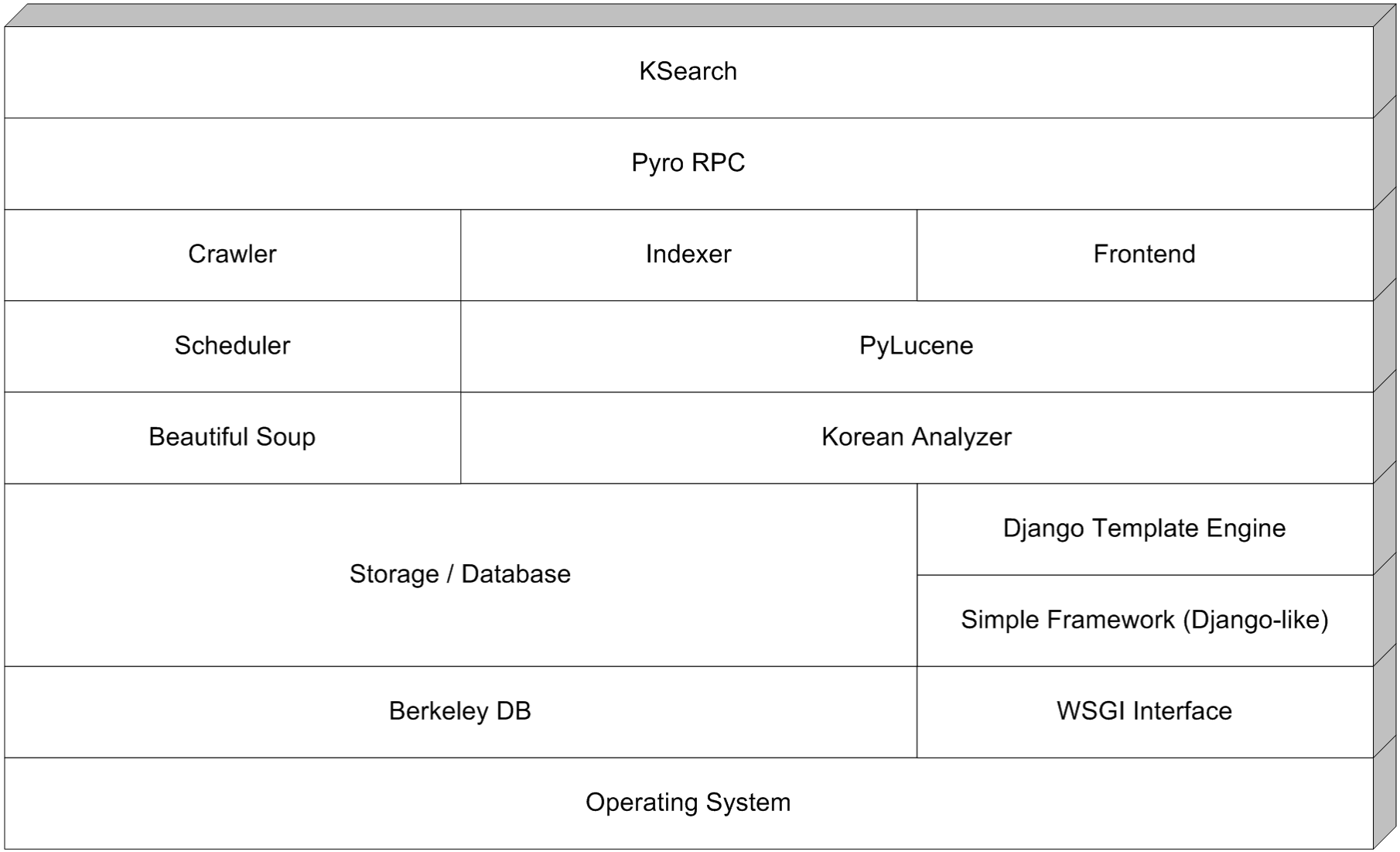
# KSearch Architecture



# KSearch Management



# Ksearch Stack



# History

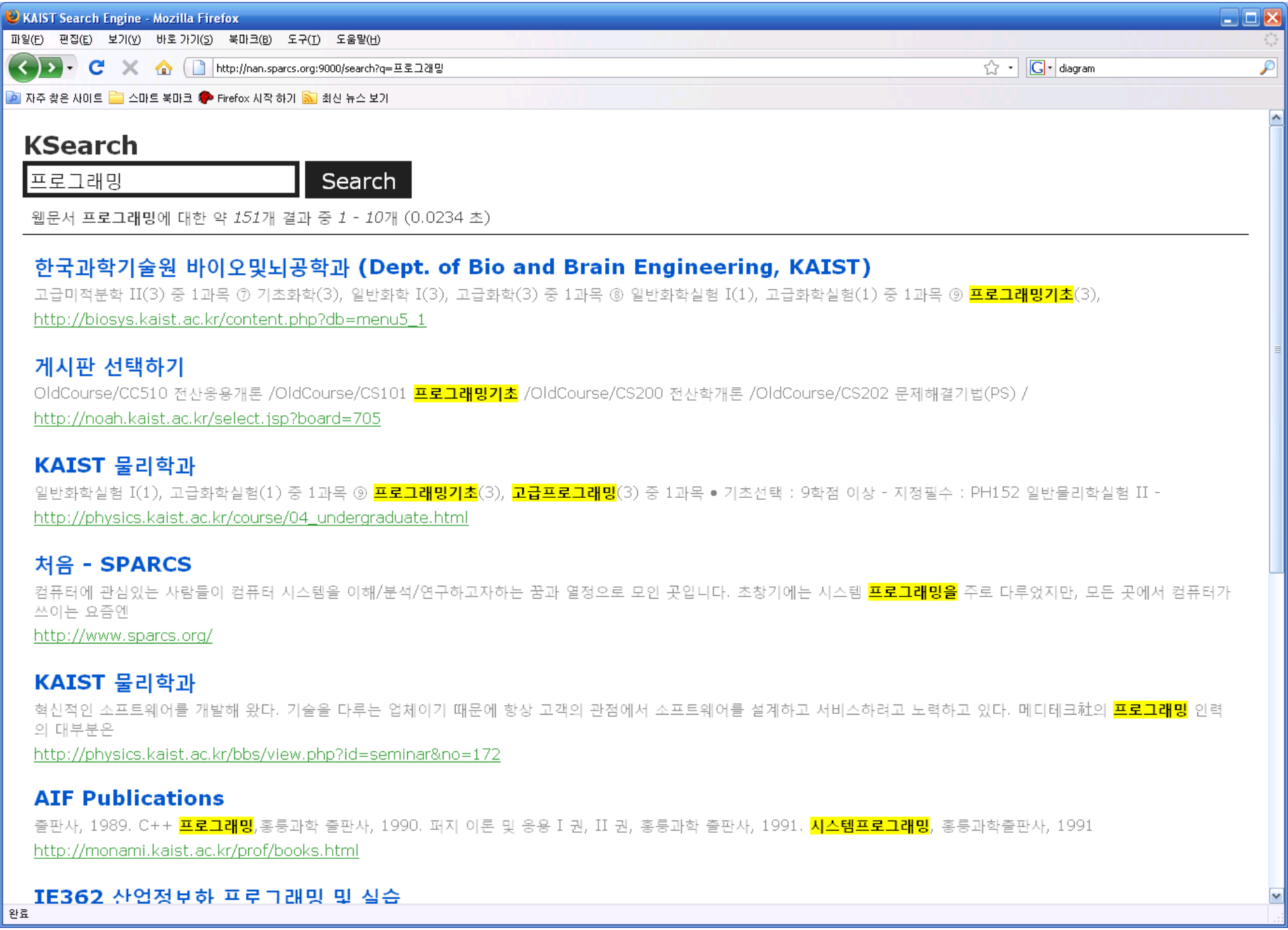
- 2007년 10월 KSearch 초기 멤버 결성
- 2007년 11월
  - 최초의 실험적 크롤러 완성
  - 크롤러-인덱서-프론트엔드 동작 시작
  - 첫 공개 테스트용 서비스 오픈
- 2008년 1월 학교측에서 지원을 받음
- 2008년 7월 개발 다시 시작
- 2008년 11월 KSearch 첫 오픈 베타 서비스 (예정)

# KSearch

Search

Feedback | Statistics | Now

A proud service of SPARCS.
Copyright © 2008, KSearch Team.


KAIST Search Engine - Mozilla Firefox
파일(F) 편집(E) 보기(V) 바로 가기(S) 북마크(B) 도구(T) 도움말(H)
http://nan.sparcs.org:9000/search?q=프로그래밍
diagram
자주 찾은 사이트 스마트 북마크 Firefox 시작 하기 최신 뉴스 보기
KSearch
프로그래밍
Search
웹문서 프로그래밍에 대한 약 151개 결과 중 1 - 10개 (0.0234 초)
한국과학기술원 바이오및뇌공학과 (Dept. of Bio and Brain Engineering, KAIST)
고급미적분학 II(3) 중 1과목 ⑦ 기초화학(3), 일반화학 I(3), 고급화학(3) 중 1과목 ⑧ 일반화학실험 I(1), 고급화학실험(1) 중 1과목 ⑨ 프로그래밍기초(3),
http://biosys.kaist.ac.kr/content.php?db=menu5_1
게시판 선택하기
OldCourse/CC510 전산응용개론 /OldCourse/CS101 프로그래밍기초 /OldCourse/CS200 전산학개론 /OldCourse/CS202 문제해결기법(PS) /
http://noah.kaist.ac.kr/select.jsp?board=705
KAIST 물리학과
일반화학실험 I(1), 고급화학실험(1) 중 1과목 ⑨ 프로그래밍기초(3), 고급프로그래밍(3) 중 1과목 • 기초선택 : 9학점 이상 - 지정필수 : PH152 일반물리학실험 II -
http://physics.kaist.ac.kr/course/04_undergraduate.html
처음 - SPARCS
컴퓨터에 관심있는 사람들이 컴퓨터 시스템을 이해/분석/연구하고자하는 꿈과 열정으로 모인 곳입니다. 초창기에는 시스템 프로그래밍을 주로 다루었지만, 모든 곳에서 컴퓨터가 쓰이는 요즘엔
http://www.sparcs.org/
KAIST 물리학과
혁신적인 소프트웨어를 개발해 왔다. 기술을 다루는 업체이기 때문에 항상 고객의 관점에서 소프트웨어를 설계하고 서비스하려고 노력하고 있다. 메디테크社의 프로그래밍 인력의 대부분은
http://physics.kaist.ac.kr/bbs/view.php?id=seminar&no=172
AIF Publications
출판사, 1989. C++ 프로그래밍,홍릉과학 출판사, 1990. 퍼지 이론 및 응용 I 권, II 권, 홍릉과학 출판사, 1991. 시스템프로그래밍, 홍릉과학출판사, 1991
http://monami.kaist.ac.kr/prof/books.html
IE362 산업정보화 프로그래밍 및 실습
완료

# Future of KSearch

- Thrift
  - Lightweight RPC System from Facebook
  - Language Independent
- Machine Learning
  - Learning from clicks
  - Spam filtering
- Distributed Service

# Retrospect

- 파이썬 좋아요!
  - 6000줄에 검색엔진을
- 오픈소스 좋아요!
  - 누군가 이미 고민하고 코딩 해놨다
- "시스템" 디버깅 어려워요!
  - 리셋할 수 없는 경우에는 더더욱
  - "인터넷" 이라는 수없이 많은 경우의 수