

### Portal Service Group 🔽

### 한글날 도서관 휴관 (Hangul Day Library Closure)

English	clone	로그이웃	Mobile
	CIUITE	노그이ㅈ	III MUUIK



한글날 도서관 휴관 검색

포탈 검색시스템은 포탈, 전화번호, 홈페이지, 공문서식, 원규집, IT서비스안내, 기타 웹 사이트 등을 대상으로 서비스 합니다.상세보기



## 도서관 휴관 안내

# Library Closure Information

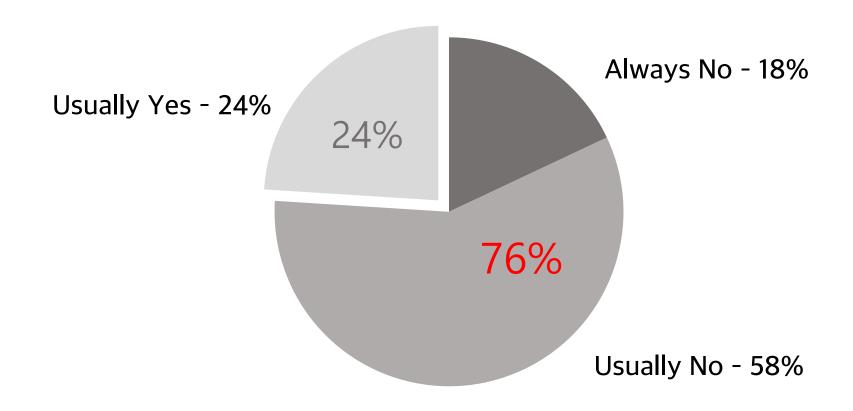
휴관일: 2018. 10. 9.(화)

사 유:한글날

대 상:학술문화관(도서관), 교양분관, 문지도서관

※ 교양분관 열람실은 24시간 개방합니다.

# Have you found looking information easily with KAIST Portal Search?



From 136 KAIST people

### xIS: extended Infomation Search

TEAM CREAMBEARS

Hyunwoo Kang, Jaemin Hong, Jiseok Kim

# Problem

It is hard to find looking articles in KAIST Portal using its search engine.

**Definition**What Is a Portal Article?

**R&D Notices** 

Recruting

Newsletter

Leadership/Intern/Counseling

IT Service

Today Notice

### **Portal Article**

GSC/USC Notice

International Community

Lecture/Academic/Paper

Dormitory

Scholarship/Welfare

Seminar/Events

Student Notice

Sports & Health/Clinic

Academic courses

Manuals/Forms



## Background

### Mechanism

of General Search Engines

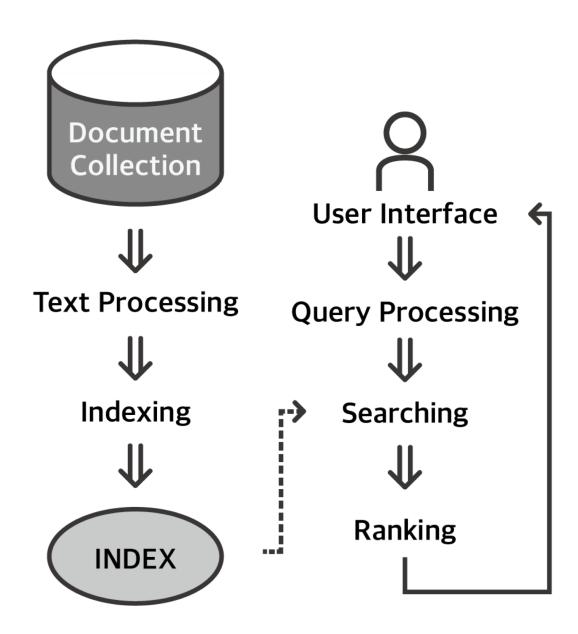
&

### Characteristics

of Portal Articles

### Mechanism

Search Engine





### Background

Mechanism

of General Search Engines

&

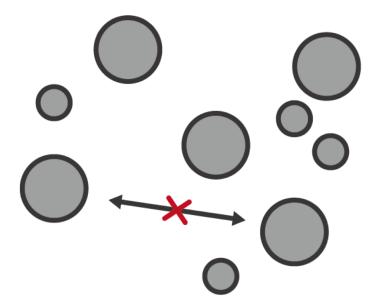
**Characteristics** 

of Portal Articles

Portal Articles

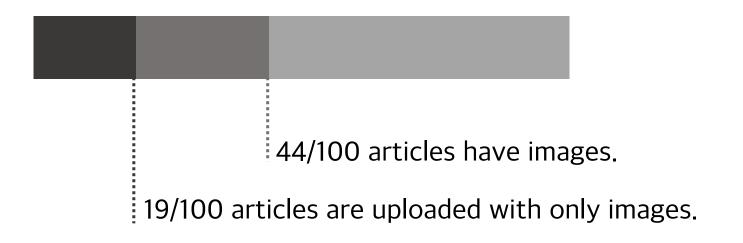
### 1. Very rare hyperlinks

Only 1/100 article has a link to another article.



Portal Articles

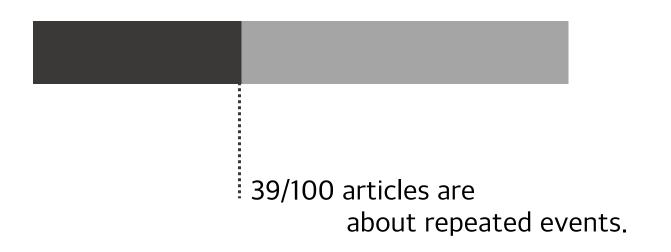
### 2. Posters and screenshots

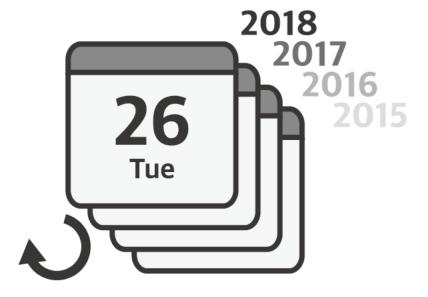




Portal Articles

### 3. Repeated events





Portal Articles

### 4. Unfamiliar official words

KAIST-POSTECH 학생대제전 VS 카포전

KAIST Clinic VS 파팔라도

학위 수여식 VS 졸업식

생활관 VS 기숙사

. . .

제목	2018 <u>KAIST-POSTECH 학생대제전</u> 우승기 전달식 개최 안내		
작성자(소속)	조재영(학생지원팀)	작성일(조회수)	2018.09.28 20:35:53(1156)
첨부파일			

2018 <u>KAIST-POSTECH 학생대제전</u> 종합우승을 기념하고 고생한 학생들에 대한 격려와 KAIST 학생들에게 자긍심 고취를 위하여 다음과 같이 우승기 전달식을 개최 하오니 학생들의 많은 참석을 부탁드립니다.

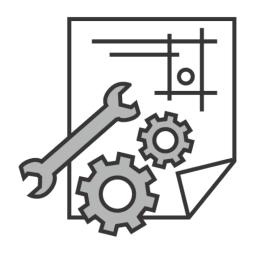
- 다 음 -

□ 행사개요

○ 행사명 : 2018 제17회 <u>KAIST-POSTECH 학생대제전</u> 우승기 전달식

○ 일 시 : 2018. 10. 1(월), 12:05 ○ 장 소 : 창의학습관 1층, 로비

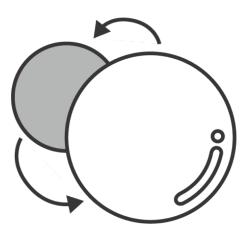
### **Possible Solutions**



Make an entirely new system



Customize an existing search engine

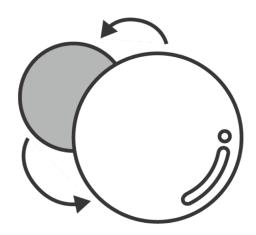


Replace with an existing search engine

# **Possible Solutions**

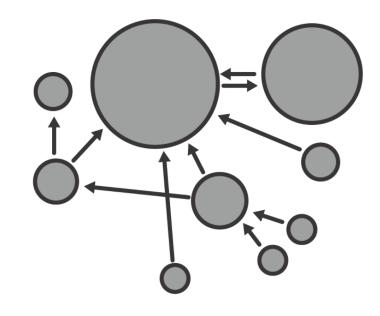


Customize an existing search engine

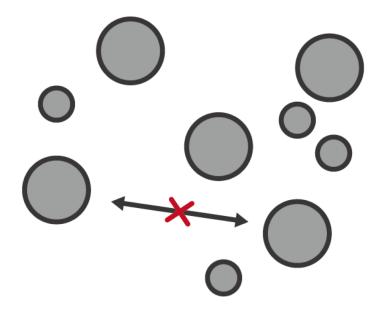


Replace with an existing search engine

# Why Not Google?

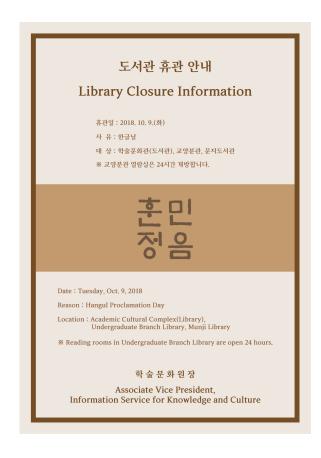


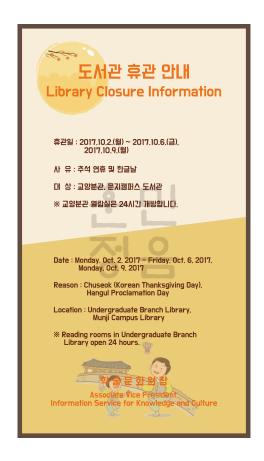
PageRank Algorithm

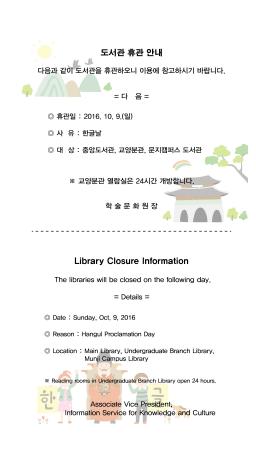


**Portal Articles** 

### **How about Others?**

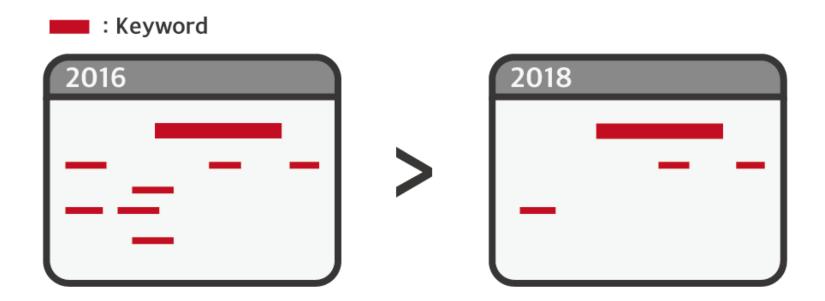






### **How about Others?**

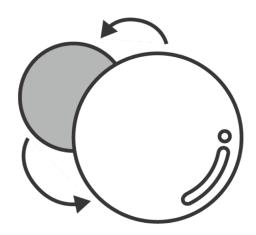
**TF-IDF**: Term Frequency-Inverse Document Frequency



### **Possible Solutions**



Customize an existing search engine



Replace with existing search engine

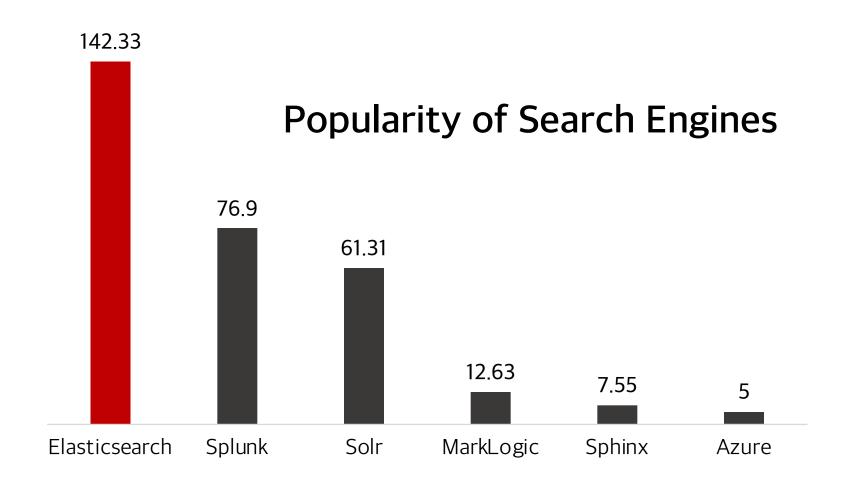
### The Possible Solution



Customize an existing search engine

# elasticsearch

# Why Elasticsearch?



# Why Elasticsearch?

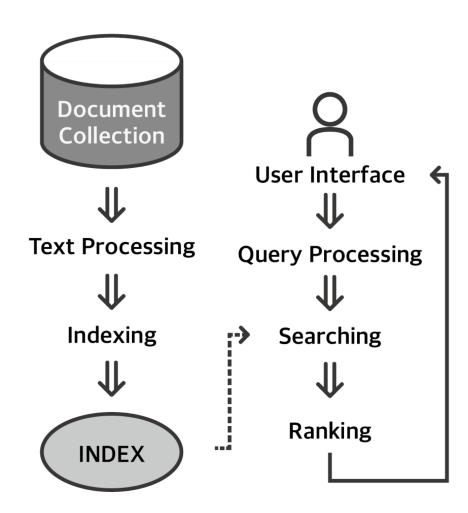


### Elasticsearch is customizable

Document format

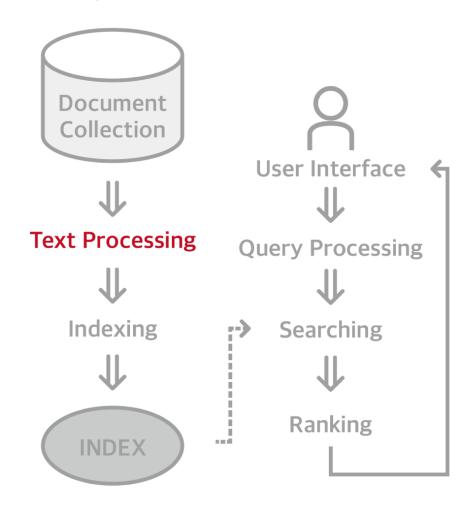
Ranking function

• •



Customize some steps to deal with characteristics of Portal articles properly

**Text Processing** 

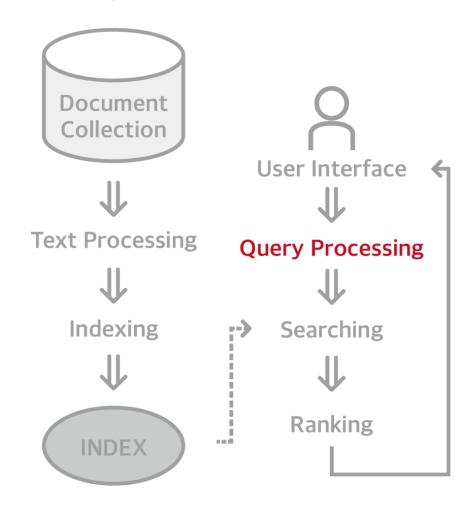


Tokenizing & stemming

Extracting texts from posters

Collecting metadata

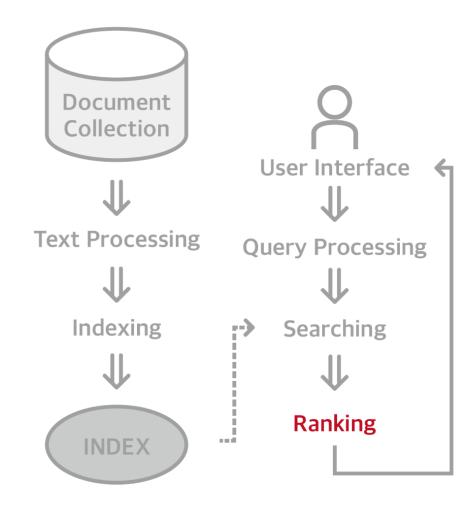
**Query Processing** 



### Handling synonyms

Analyzing user's search behaviors

Ranking



TF-IDF

Considering metadata

- Creation date
- Hits

### **Evaluation**



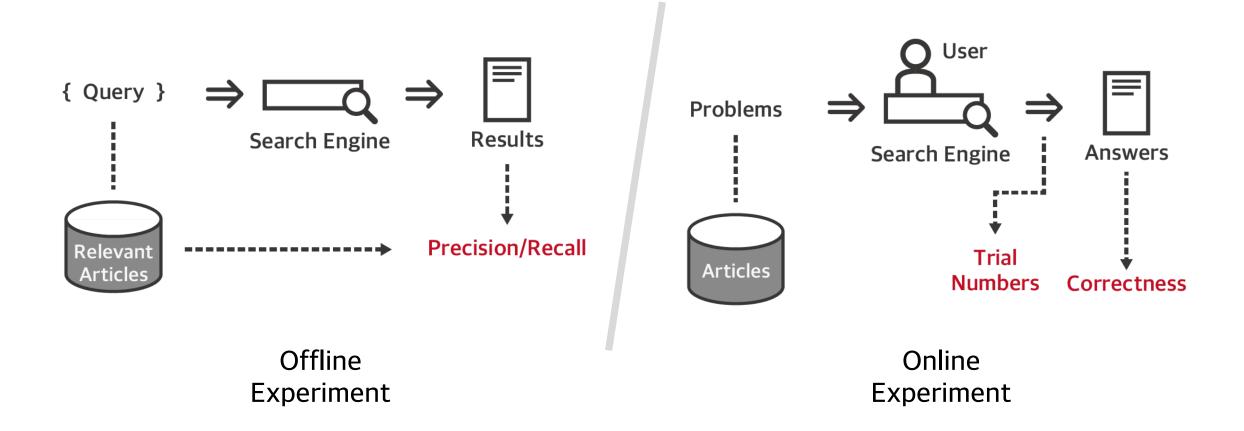








### **Evaluation**



- 검색어와 무관한 정보가 나오는 경우가 많다. 교환학생 프로그램 검색하면 최신정보가 안나오고 오래된게 먼 136 KAIST people responded to our survey, and 99 among them answered 'write down experience that you couldn't find information easily using Portal Search' |사항 등을 찾을 때 너무 오래된 자료들이 정확도가 높게 표시되는 경우가 잦음