

웹 드라이브(e드라이브)를 통한 문서 통합 본문 검색 어플리케이션 제작

프로젝트 조원 : 안형욱, 이승윤

1. 서론

1.1. 프로젝트 배경

클라우드 컴퓨팅은 4차 산업혁명 시대의 기반 인프라이다. 클라우드 컴퓨팅은 언제 어디서나 필요한 만큼의 컴퓨팅 자원을 필요한 시간만큼 인터넷을 통하여 활용할 수 있는 컴퓨팅 방식이다.

이러한 서비스를 통해 현재 대다수의 사람들이 클라우드 컴퓨팅을 활용한 서비스를 이용하고 있으며, 사회의 공동체적인 목적으로 사용하는 것 뿐만 아니라 개인적인 공간으로도 활용하여 이용하고 있다.

클라우드 스토리지 서비스에 대한 수요가 많아지는 만큼, 이 서비스를 제공해주는 플랫폼의 종류도 다양하게 증가하고 있는 추세이다. 국내 서비스 플랫폼의 대표적인 예로는 네이버 클라우드(Naver), 클라우드 베리, 삼성 클라우드 등이 있으며, 해외 서비스의 경우 구글 드라이브, 원(One)드라이브, 드롭박스(dropbox) 등이 있다.

개인의 경우 이러한 플랫폼이 제공하는 클라우드 스토리지를 이용하여 언제 어디서든 인터넷만 연결되어 있는 환경이라면 마치 컴퓨터의 하드디스크에 저장되어 있는 파일을 꺼내 보는 것처럼 사용 할 수 있으며, 해당 파일을 각 디바이스로 다운로드하여 볼 수 있다.

하지만 각 클라우드 서비스마다 제공하는 뷰어시스템도 다르며, 대부분 서비스의 경우 간단한 파일 탐색기정도의 기능을 하고 있지만, 클라우드 스토리지만을 위한 검색 및 뷰어 서비스를 제공하는 기능을 가진 모듈은 거의 찾아볼 수 없다.

1.2. 프로젝트 목표

앞서 서술한 것과 같이 클라우드 스토리지를 위한 검색 엔진은 거의 없는 실황이다. 개인 단위의 클라우드 스토리지 사용면에서는 미미한 활용을 보여 줄 수

있으나, 기업이나 단체 단위의 클라우드 스토리지 사용면에서는 스토리지에 넣는 문서 파일의 양이나 종류, 리스트 수가 직접 눈으로 보고 찾기에는 무리가 있을 것이라고 판별된다. 따라서 이를 위한 검색엔진 UI를 따로 구현하여 마치 인터넷상의 검색엔진과 같이 해당 키워드를 입력하여 검색하면, 그에 따른 자료 및 미리보기도 같이 결과물에 나옴과 동시에, 웹 드라이브의 순기능인 다운로드까지 원활하게 진행 될 수 있도록 구현하는 것이 목표이다. 뿐만 아니라 현재 제공중인 검색엔진으로서의 기능을 일부 추가하여 사용자가 이를 사용하는데 불편함이 없을 뿐 아니라 해당 기능을 이용함으로써 시간단축이 될 수 있도록 방향을 잡고 구현할 것이다.

2. 참고자료 / 조사

2.1. 오픈 소스 및 개발 프레임워크

2.1.1. Django

<https://www.djangoproject.com/>

Django는 파이썬으로 만들어진 무료 오픈소스 웹 애플리케이션 프레임워크로, 이를 활용해 백엔드 서버를 구축해 회원가입/로그인부터 검색 Query를 전송하는 등의 api를 구현할 예정이다.

2.1.2. Elasticsearch

<https://www.elastic.co/kr/elasticsearch/>

텍스트, 숫자, 위치 기반 정보, 정형 및 비정형 데이터 등 모든 유형의 데이터를 위한 무료 검색 및 분석 엔진이다. 이를 사용해 e드라이브 내에서 저장한 문서 암호화 및 Document Filter, 본문 내용 및 이미지 추출 등의 작업을 수행한다.

2.1.3. React.js

<https://reactjs.org/>

사용자 인터페이스를 만들기 위한 자바스크립트 라이브러리로 컴포넌트를 기반으로 UI를 설계할 수 있다. 다른 라이브러리들과도 함께 사용할 수 있다는 점에서 확장이 용이하고 이를 활용해 eDrive 서비스의 UI를 구성할 예정이다.

2.1.4. Redux

<https://redux.js.org/>

전역 상태 관리를 위한 자바스크립트 라이브러리로 React 컴포넌트의 state를 전역적으로 관리하고, redux-thunk, redux-saga 등의 다른 middleware를 함께 사용할 수 있다.

2.2. 유사 서비스

2.2.1. 구글

<https://www.google.com/>

대표적인 검색 플랫폼으로, 구글의 UI/UX를 참고하여 프로젝트를 구현할 계획이다.

2.2.2. 네이버

<https://www.naver.com/>

한국의 대표적인 검색 플랫폼으로, 이 또한 참고사항으로 이용 가능하다.

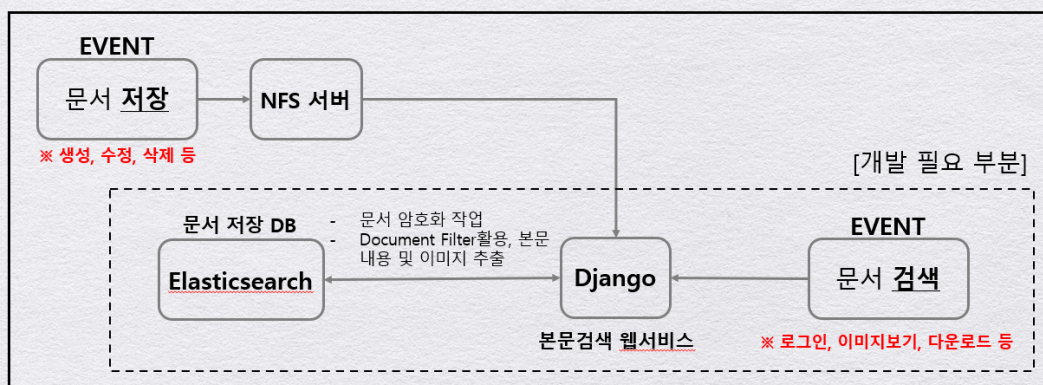
3. 프로젝트 내용

3.1. 시나리오

- 본문 검색 예상 구조

E드라이브 본문검색

e드라이브를 사용하며 저장한 문서의 내용을 base로 검색을 할 수 있는 기능



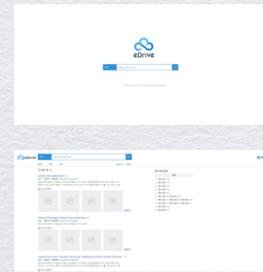
- 본문 검색 개발 상세

E드라이브 본문검색 개발 상세

e드라이브를 사용하며 저장한 문서의 내용을 base로 검색을 할 수 있는 기능

Django (웹서버 구축)

- I. NFS 연동
 - a. 문서 작업 관련 이벤트 받기
- II. Elasticsearch 구축 및 연동
 - a. NFS에서 받은 이벤트 update 하기
 - b. 검색 Query 전송하기
- III. Publishing (UI 디자인)
- IV. (추가) 특허 관련 검색 (광개토연구소)



3.2. 유저 인터페이스 요구사항

- 인트로 화면 구성 (아래 사진들은 샘플 이미지들입니다.)

인증 받은 사용자 (로그인 완료)만 접속 가능



© 2019 GAEASOFT Co., LTD. All Rights Reserved.

- 본문 검색 내용 화면 구성

전체 개인 부서 고급검색 설정

검색결과 총 3개

Sample Document.docx 파일 다운로드

문서 내 이미지

Sample Document Sample Document.docx 파일 상세보기

문서 내 이미지 Thumbnail 형식으로 출력

Sample Document Sample Document Sample Document Sample Document...

광개토연구소 (일단 보류)

특허 관련 검색

특허

1 특허 관련 기사

2 특허 관련 기사

3 특허 관련 기사

4 특허 관련 기사

5 특허 관련 기사

6 특허 관련 기사 특허 관련 기사

7 특허 관련 기사 특허 관련 기사 특허 관련 기사

8 특허 관련 기사 특허 관련 기사 특허 관련 기사

9 특허 관련 기사

10 특허 관련 기사

- 본문 검색 내용 화면 구성 (자세히 보기)

전체 개인 부서 고급검색 설정

검색결과 총 3개

Sample Document.docx

문서 내 이미지

파일 정보

문서명 Sample Document.docx

작성자 홍길동

파일 경로 가나다/라마바/사아자/차카타/파...

파일 생성일 2019.07.01

파일 수정일 2019.07.05

기본검색

고급검색

단어/문장 검색

최초 작성자

최종 수정자

기간검색

시작일 - 종료일

검색 닫기 초기화

다양한 검색 방식 추가 예정 (현재 자세한 기능은 미정)

이 외의 파일 정보 추가 가능 (현재 미정)

- 본문 검색 내용 화면 구성(이미지 더보기)



3.3. 서비스 품질 요구사항

검색 결과를 도출하여 화면에 출력하는데 시간이 오래 걸리지 않아야 하며, 이미지파일 또한 미리보기를 하는데 시간 절약 및 훼손이 없도록 해야 한다. 또한 카테고리별 검색을 통해 검색 이외의 결과물이 나오지 않도록 해야 할 것이며, 별도의 가이드 없이 원활하게 이용 할 수 있도록 해야 할것이다.

4. 결론 및 기대효과

가장 많이 쓰는 서비스 중의 하나인 검색 웹사이트를 구현해보면서 평소 웹서비스를 사용해보면서 느꼈던 불편한 점들을 개선시키는 경험을 할 수 있다. 또 4차산업혁명의 핵심 기술인 클라우드 서비스를 이해하고 적용해보며 현업에서 어떻게 클라우드 서비스가 활용되는지를 이해한다.

서비스를 사용하는 실 사용자를 고려하여 기획, 설계, 구현 및 배포의 단계를 거치며 전체적인 개발 프로세스를 경험할 수 있고, 이를 통해 훗날 실무에서 활용 가능한 개발 기술을 획득한다.