

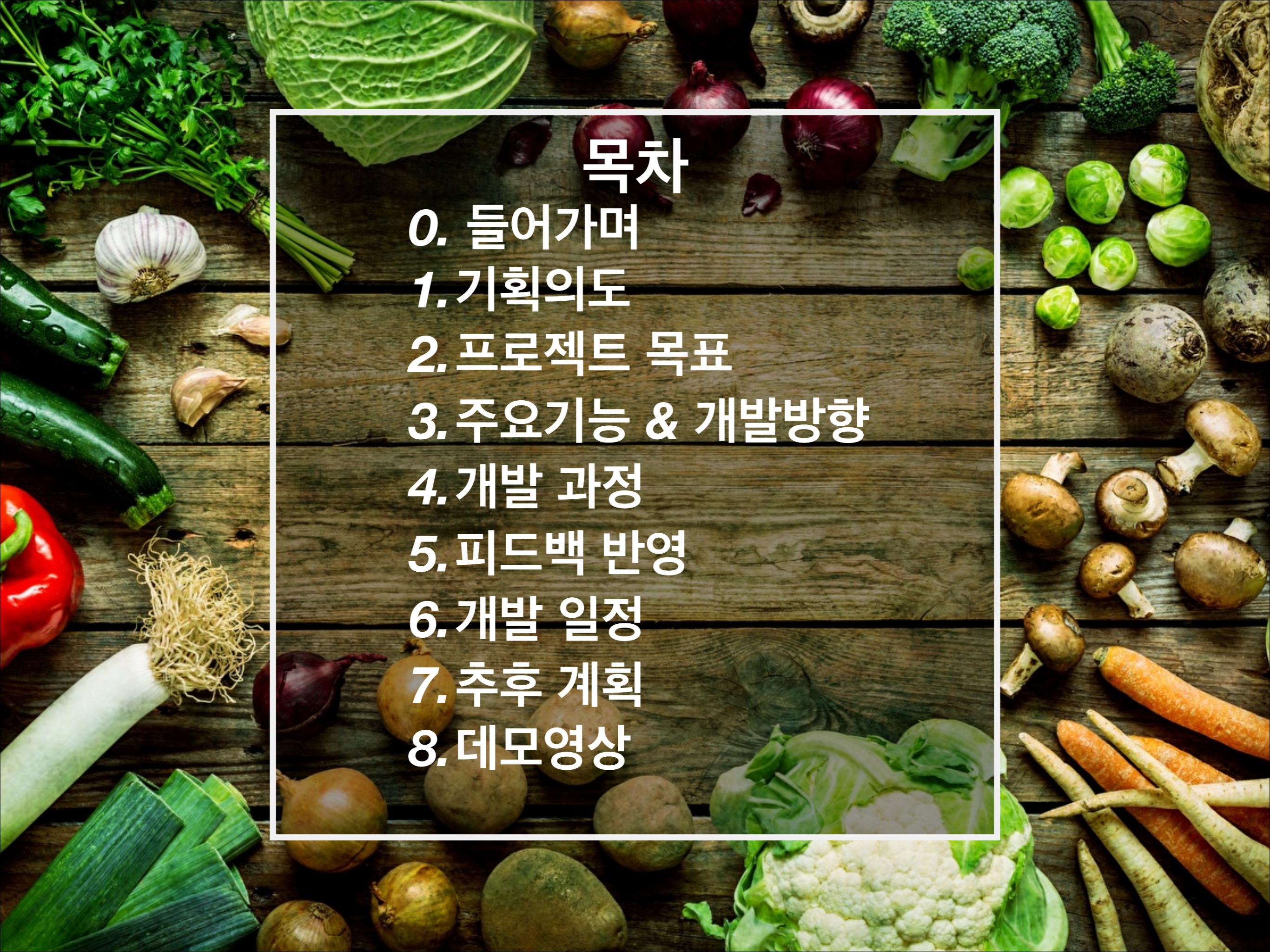


# Begin Vegan

---

팀명 : 디비딥

팀원 : 한지수, 박태준, 고동섭



# 목차

0. 들어가며
1. 기획의도
2. 프로젝트 목표
3. 주요기능 & 개발방향
4. 개발 과정
5. 피드백 반영
6. 개발 일정
7. 추후 계획
8. 데모영상

## 0. 들어가며

Travel Is Always Fun! However..

*Begin  
Vocational*



## 0. 들어가며

# 비건(vegan)이란?

의미1) 채식주의자의 단계 중 채소와 과일을 제외한 어떤 동물성 음식도 먹지 않는  
가장 높은 단계의 채식주의자.

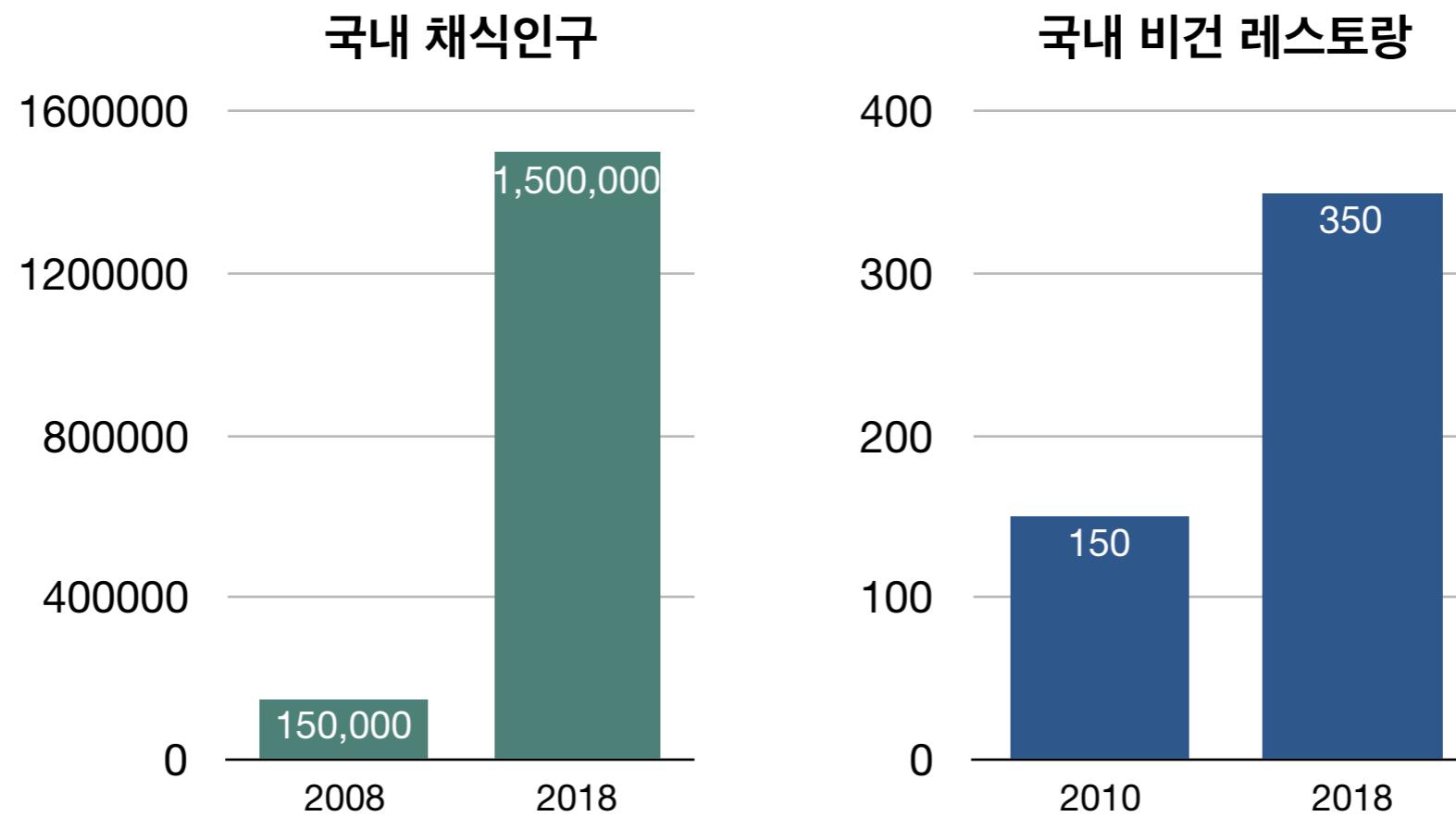
의미 2) 6가지 단계로 채식주의자를 나누는 것이 아닌 동물성 영양을 기준으로 비건/ 논비건 구분시 사용

**채식의 단계**

| 단계                   | 제한 음식 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| ① 플렉시테리언 Flexitarian | 닭고기   |
| ② 세미 semi            | 닭고기   |
| ③ 페스코 Pescatarian    | 닭고기   |
| ④ 락토오보 LactoOvo      | 우유    |
| ⑤ 락토 Lacto           | 우유    |
| ⑥ 비건 Vegan           | 우유    |

※ 식물이 살아가는 근간이 되는 뿌리나 줄기를 먹지 않고 열매, 잎, 곡식을 먹는 프루트채식Fruit.  
채식을 하되 익힌 것이나 가열한 것을 먹지 않고 생것으로 먹는 생채식Uncooked도 있다.

## 1. 기획의도



- 식물성 고기 브랜드 (비욘드 미트) 국내 출시 **한달만에 1만팩** 이상 팔려나감
- 식물성 대체 우유 브랜드 (아몬드 브리즈) **연평균 74% 성장**

출처: 뉴스룸 갈수록 커지는 비건시장 2019.1.30 보도

- 2018 트위터 **트렌드 1위** - “비거리니즘” 트윗 2천만건 (출처 : 트위터)
- 영국 이코노미스트 - ‘세계대전망 2019’ **올해는 비건의 해가 될 것** (출처 : 영국 이코노미스트)

## 1. 기획의도

(이화여자대학교 채식주의 동아리 ‘솔찬’에게 실시한 설문결과)

비건으로 생활하면서 불편한 점이 있나요?

1. 편의점이나 마트에서 성분을 확인하는데 들이는 시간이 오래 걸립니다...
2. 채식주의 식품 관련 정보를 얻는 곳이 너무 한정적입니다.

비건과 논비건 식품을 어떻게 구분하시나요?

1. 제품의 원재료 및 함량(성분표)을 보고 판단합니다.
2. 성분에 대해 문의하여 확인합니다.
3. SNS, 비건 커뮤니티, 어플리케이션 등을 참고합니다.

제품을 검색하면 채식주의 단계를 알려주는 앱이 있다면 어떨 것 같나요?

1. 정말 좋을 것 같아요!
2. 사용할 것 같습니다.

# 1. 기획의도

## 유사 서비스



타깃 고객

엄선



Is It Vegan

아이 엄마

비건

데이터베이스  
확보 방식

자체 데이터 수집

사용자가 작성

데이터 미확보 제품  
처리방식

없음

사용자가 비건 여부  
판단 후 직접 작성

국내 제품 데이터  
정확도

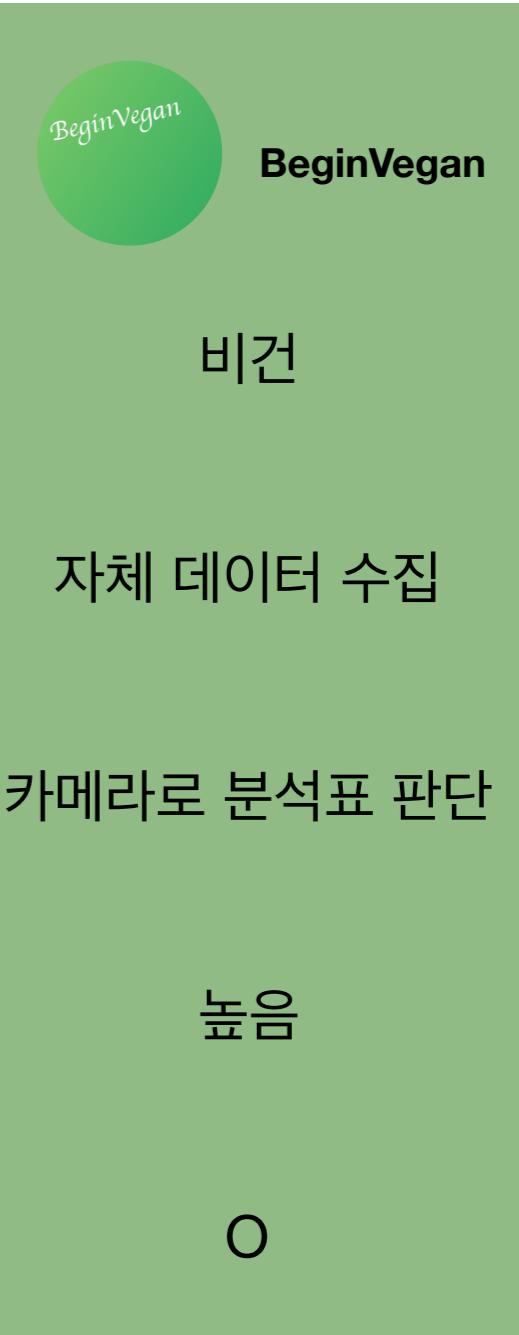
높음

매우 낮음

대체재 추천 기능

O

X



- 엄선 - 타겟층이 아이를 둔 부모, 비건 관련 기능이 없다
- Is It Vegan - 국내 식품 관련 데이터 거의 없다, 데이터 부정확함

# 1. 기획의도



## 2. 프로젝트 목표

사용자가 제품을 검색하거나  
원재료명 사진 촬영하여  
식품의 적합성 판단

상황 1

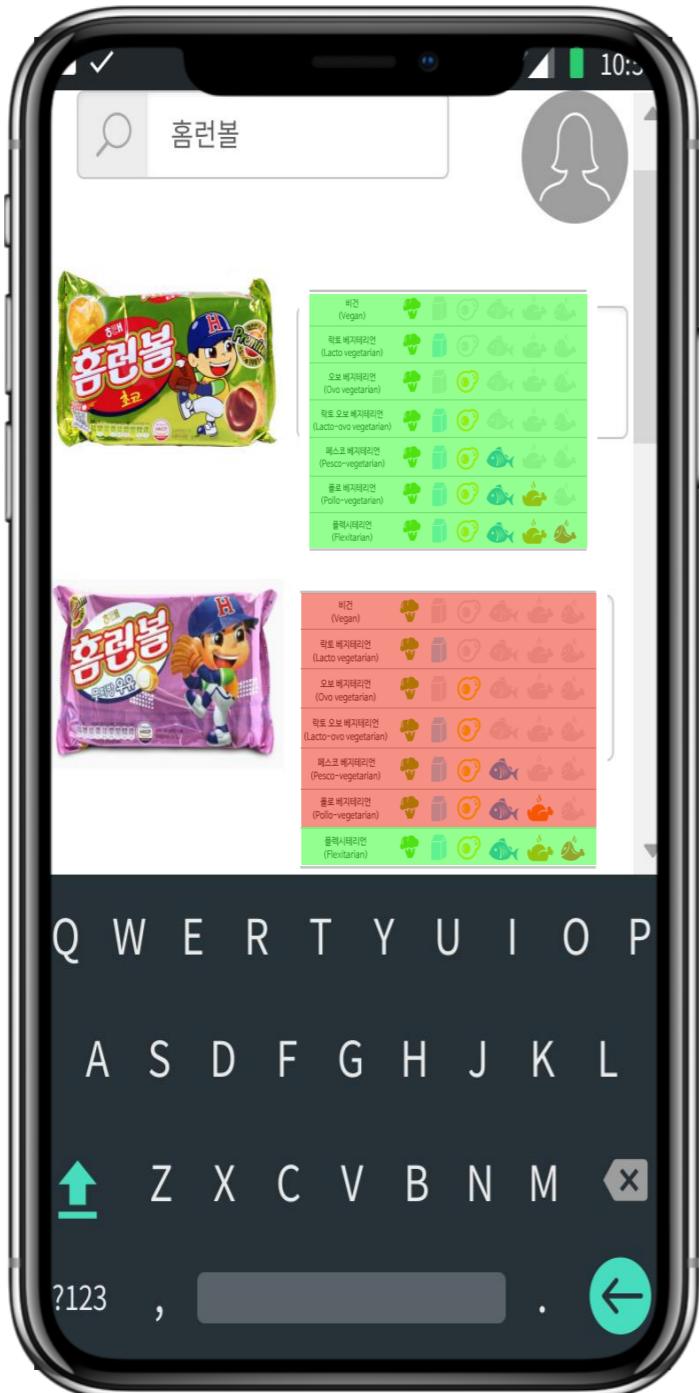


상황 2



### 3. 주요 기능 & 개발방향

#### 주요기능 1 - 검색 후 적합성 판별



### 3. 주요 기능 & 개발방향

#### 주요기능 2 - 사진 촬영 후 적합성 판별



### 3. 주요 기능 & 개발방향

검색

query 검색

사진 촬영 후 원재료 분석

이미지 처리(GCP+NLP)

실시간 통신

**Client iOS / Android (React Native)**

**Server (NGINX) / API ( Django )**

**NLP Engine**

**DB 구축 (Mysql)**

**데이터 구하기 (Storage)**

**Linux (AWS) / Ubuntu (AWS) / Mac OS / Window**



#### 4. 개발 과정

DB) SSG.com 이미지 > AWS S3(Storage) 크롤러

식품 회사의 데이터 제공 X



상품상세정보      고객상품평 (535)      상품 Q&A(31)      배송/반품/교환 안내

맛 선택하세요.

원재료명 및 함량  
아몬드액 95%, 정제수,  
전분(마국산), 식염,  
영양강화제, 산도조절제,  
유화제, 합성착향료(아몬드향)

영양성분 (1회 제공량 190ml)

열량	35 kcal	
탄수화물	3 g	1 %
당류	0.2 g	
단백질	1.2 g	2 %
지방	2.1 g	4 %
포화지방	0 g	0 %
트랜스지방	0 g	
콜레스테롤	0 mg	0 %
나트륨	125 mg	6 %
칼슘	230 ml	33 %
마그네슘	13 mg	4 %
비타민E	9.8 mgα-TE	89 %

%영양소기준치 : 1일 영양소기준치에 대한 비율



합계 13,900 원

장바구니

바로구매

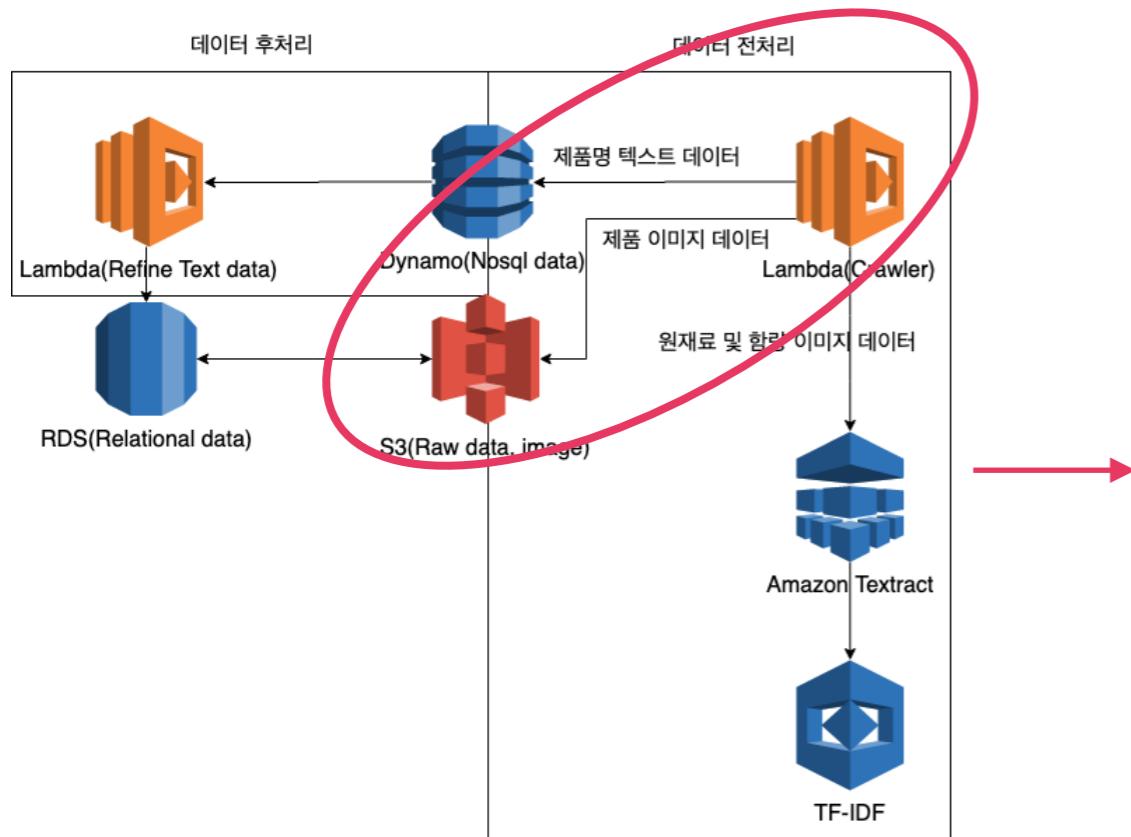


선물

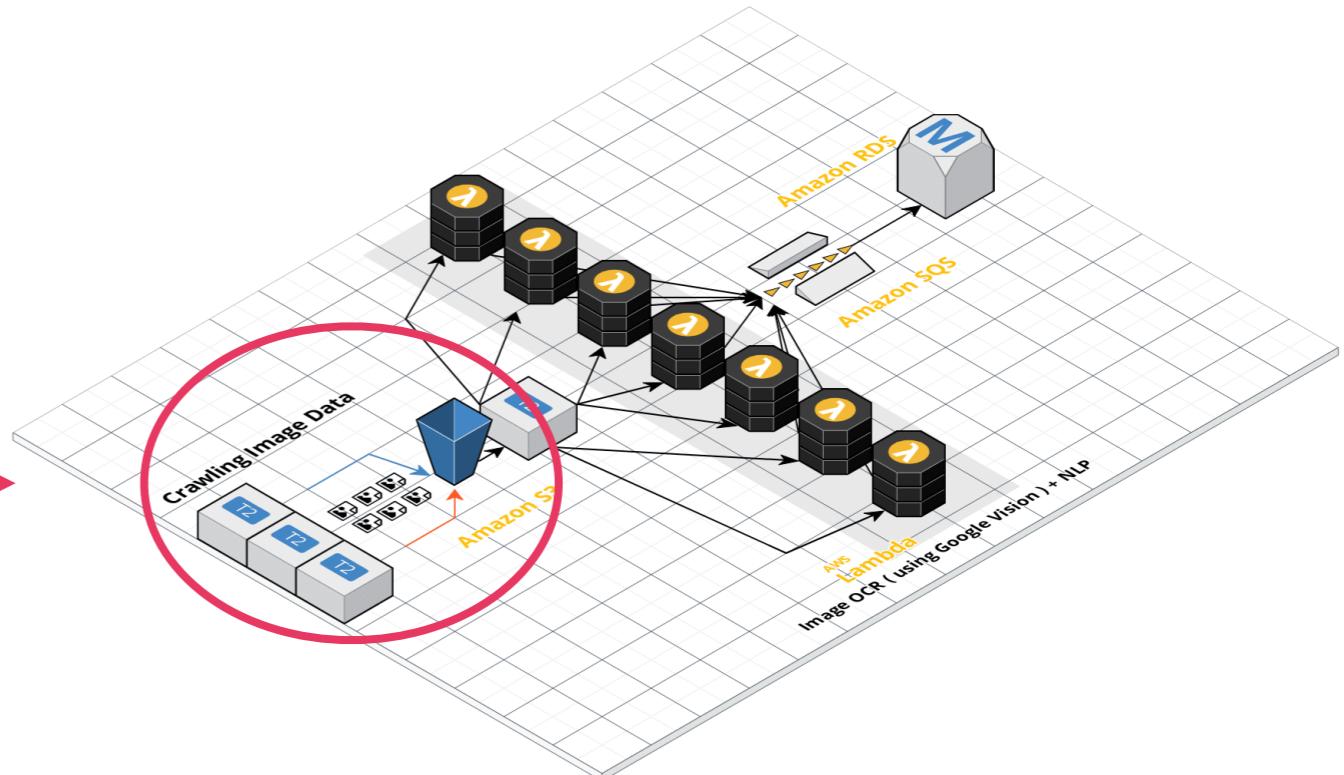
## 4. 개발 과정

### DB) SSG.com 이미지 > AWS S3(Storage) 크롤러

**Before**



**After**



## 4. 개발 과정

### DB) SSG.com 이미지 > AWS S3(Storage) 크롤러

#### 10초 이내로 재요청 반복 시 서버에서 거부

```
jjony@DESKTOP-GLU26QK:~/SProject/crawling$ python scrap.py  
http://www.ssg.com/disp/category.ssg?ctgId=6000052190&page=8  
<title class="notranslate">티포 에그쿠키 90g * 12개, 신세계적 쇼핑포털 SSG.COM</title>  
<title class="notranslate">파스키에 그릴레틴 통밀 161g x 3 팩, 신세계적 쇼핑포털 SSG.COM</title>  
<title>연속적인 새로고침 오류, 신세계적 쇼핑포털 SSG.COM</title>
```

매 10초마다 기다릴 경우 수만 건의 크롤링 데이터를 모두 수집 불가

#### 시도 1. Proxy 서버를 이용한 병렬 처리

이미 쇼핑몰에 차단된 IP주소 다수, 요청이 프록시 서버에서 거부되는 경우 빈번

#### 시도 2. AWS EC2를 이용한 병렬처리

크롤링 시간 단축

but...

+ 로컬에 데이터를 저장하는 데 용량 한계가 존재

---

EC2의 수집 데이터 AWS S3에 저장까지 자동화

## 4. 개발 과정

### DB) SSG.com 이미지 > AWS S3(Storage) 크롤러

#### RAW DATA 수집

pd_img/ 객체 76198개 - 10.2 GB	
img/ 객체 316635개 - 154.0 GB	
<input type="checkbox"/> 이름 ▾	마지막 수정 ▾
<input type="checkbox"/> (1+1) 성풍 적강낭콩 빨간콩 2.5kg , 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1+1구성) 대상웰라이프 홍삼 서향 진액 正, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1+1구성) 대상웰라이프 홍삼 자운 순액 正 , 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1+1구성) 대상웰라이프 홍삼 활력플러스, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (10포 추가 증정!) 100% 이란산 순수한 석류즙 2박스 (총 70포) (F20160314038), 신세계적 쇼핑포털 S...	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 눈 밝은 세상 루테인+지아잔틴+베타카로틴, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 올트라 더블업 힘(HIM) 쏘팔메토+옥타코사놀+아연, 신세계적 쇼핑포털 SSG....	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 종합비타민 미네랄 올트라 맨+ 올트라 우먼 , 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 종합비타민 미네랄 올트라 맨, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 종합비타민 미네랄 올트라 우먼 60정, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 하루 유산균 19 x 12ea, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 하루 한 알 관절건강 N-아세틸 글루코사민 x 6병, 신세계적 쇼핑포털 SSG.CO...	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 하루 한 알 눈건강 루테인 60캡슐 x 6병, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900
<input type="checkbox"/> (1년구성) 대상웰라이프 하루 한 알 여성건강 감마리놀렌산 x 6병, 신세계적 쇼핑포털 SSG.COM.jpg	8월 12, 2019 2:02:31 오후 GMT+0900

#### AWS S3

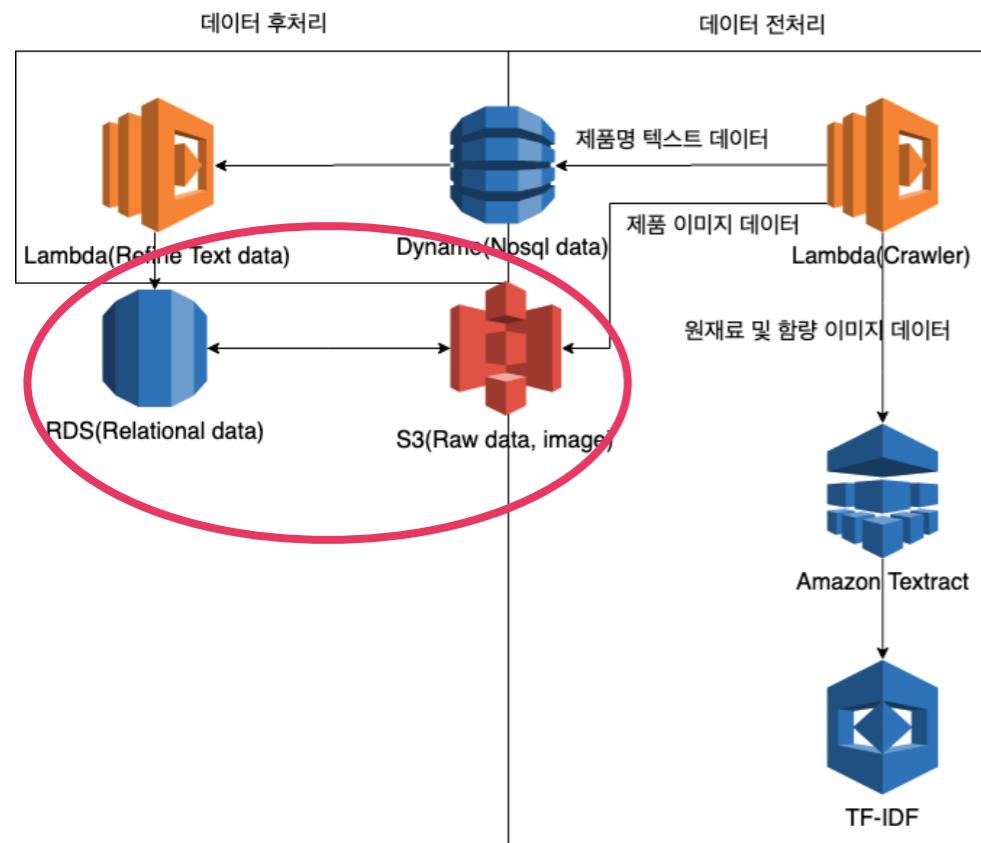
제품 이미지, 텍스트 정보  
각 76198 개 수집

원재료명 관련 이미지  
총 316635 개 수집

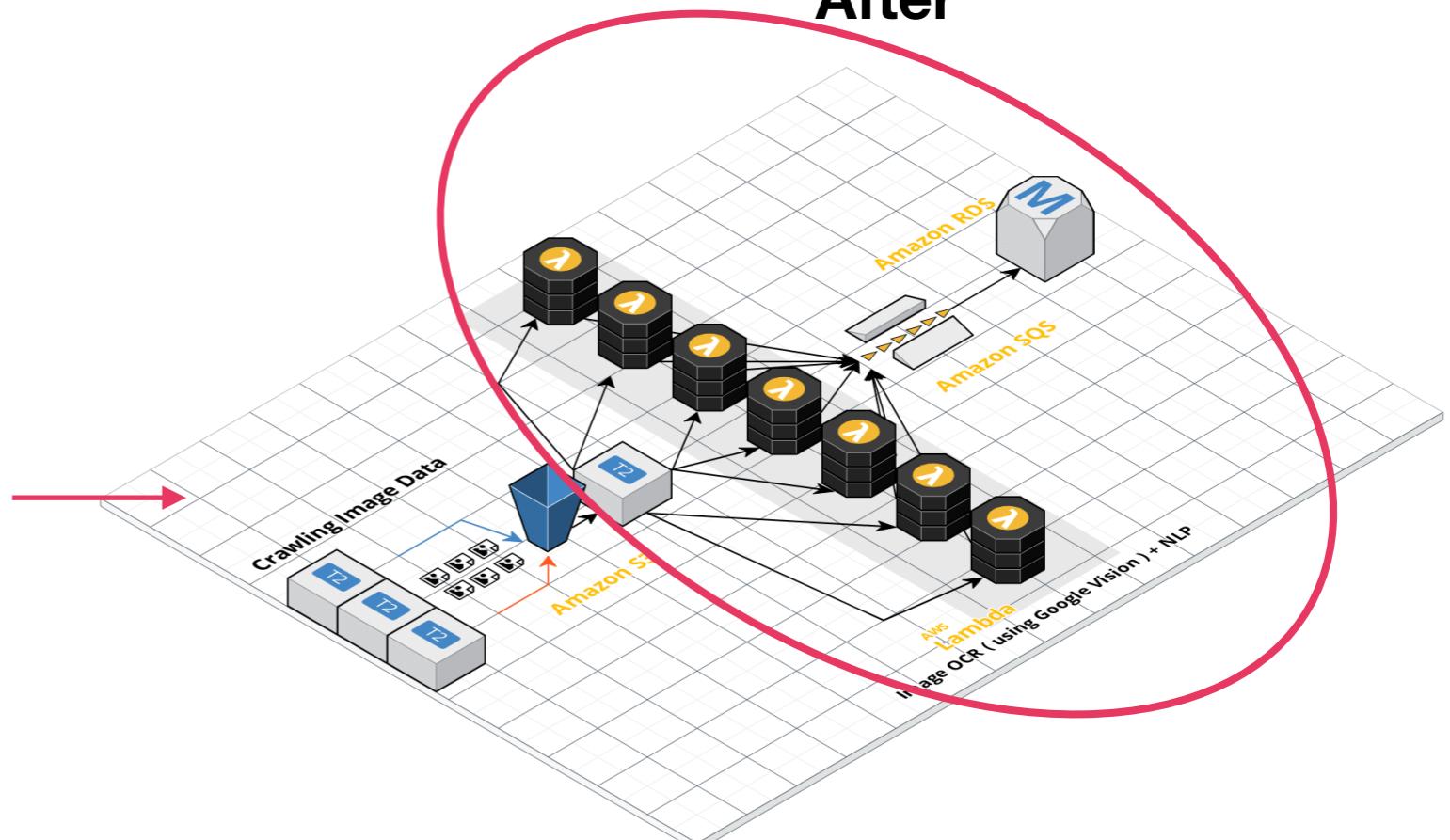
## 4. 개발 과정

### DB) AWS S3(Storage) > AWS RDS (Mysql)

**Before**



**After**



## 4. 개발 과정

**DB) AWS S3(Storage) > AWS RDS (Mysql)**

**문제 상황 : 직렬 처리로 31만장의 이미지에서 OCR 모델을 통해  
텍스트를 추출할 경우 처리시간이 5일 예상**

**시도1 : AWS Lambda를 통한 병렬 처리 직후 디비 적재**

**시도 1 문제 1 : Google Vision 분당 1800 리퀘스트 제한**

```
[ERROR] ResourceExhausted: 429 Quota exceeded for quota metric 'vision.googleapis.com/default_requests' and limit  
'DefaultRequestsPerMinutePerProject' of service 'vision.googleapis.com' for consumer 'project_number:828926346516'.
```

**시도 1 문제 2 : DB deadlock 발생**

```
[ERROR] OperationalError: (1213, 'Deadlock found when trying to get lock; try restarting transaction')
```

**시도 2 : AWS Lambda를 통한 병렬 처리(최대 람다 생성갯수 제한) +  
AWS SQS(큐) 를 이용한 DB deadlock 해결**

---

**예상 기간 4 ~ 5일 > 7시간 (분당 최대 개수 요청)**

## 4. 개발 과정

### DB) AWS S3(Storage) > AWS RDS (Mysql)

| 사나이  
매울신  
우리가 몰랐던 신라면 맛의 비밀?  
신라면 특유의 매콤하고 얼큰한 국물맛은 1년 넘게 연구된 농심만의  
특화된 고추와 표고버섯, 건파, 마늘 등을 첨가하여 만든 특별 스프와  
200여 차례가 넘는 실험과정을 통해 탄생된 쫄깃쫄깃하면서 부드러운  
면발이 조화를 이루어 매콤하고 개운한 신라면 맛을 내고 있습니다.

신라면의 특징  
신라면은 소고기를 주원료로 완전진공 농축 시킨 소고기 엑기스에  
각종 천연 양념을 잘 조화 시켜 만든, 전통 가정요리로부터 유래된  
얼큰한 소고기국 맛의 제품입니다.

신라면의 인기  
신라면은 1백 500여 종의 제품이 각축을 벌이고 있는 한국 라면시장에서  
25%의 점유율을 차지하고 있는 전무후무한 라면의 대명사입니다.

0%

콜레스테롤  
트랜스지방

|제품명 |신라면

|●면소액분(미국산, 호주산), 감자전분(독일산), 팜유(말레이시아산),  
|변성전분, 난각칼슘, 정제염, 야채조미추출물, 면류점가말칼리제(산도조절제),  
|혼합제(산도조절제), 올리고녹차풍미액, 비타민B2. ●스프류/소고기맛베이스, |  
|원재료명 |정제염, 매콤양념분말, 간장양념분말, 정백당, 볶음양념분, 조미소고기분말,  
|간장분말, 후추가루, 조미홍고추분말, 5-리보뉴클레오티드이나트륨, 호박산  
|이나트륨, 대두단백, 건파, 건청경채, 건표고버섯, 건당근, 건고추,  
밀, 대두, 되지고기, 계란, 쇠고기 함유  
|유통기한측면 표기일까지

품목 |1986036000721(안성), 200105495411(구미),  
|보고번호 | 197201540011(부산), 19760342001155(안양)  
|-이제품은 우유, 메밀, 땅콩, 고등어, 게, 새우, 토마토, 호두, 닭고기, 오징어, 조개류  
|(글, 전복, 홍합 포함)를 사용한 제품과 같은 제조시설에서 제조하고 있습니다. “본 제품은 |  
|소비자기본법에 의거 교환, 환불해드립니다. ●보관상 주의사항: 직사광선을 피하고 서늘  
|하고 건조한 곳에 보관하십시오. ●본 제품은 농산물 원료를 주로 사용하여  
|벌레의 영향을 받기 쉬우니 포장이 파손되지 않도록 보관에 유의해 주시고  
|개봉 후에는 즉시 조리해 드십시오. ●부정·불량식품 신고는 국번없이 1399 PP  
|식품유형] 유통면

Nutrition Facts

영양성분

LOLL

영양성분표

일일권장량 대비 %

1회 제공량 기준

열량

505kcal

탄수화물

24%

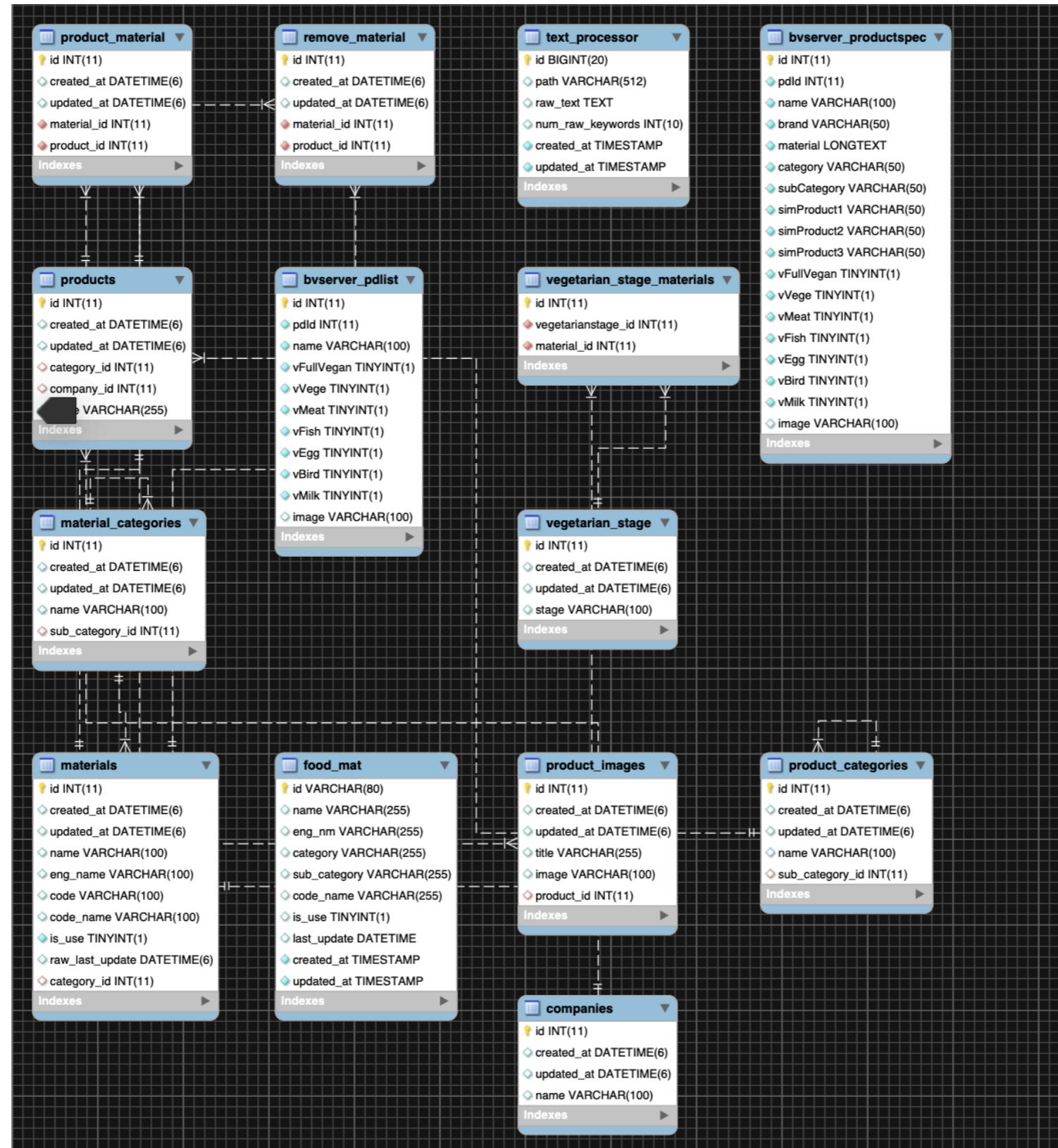
당류

```
mysql> SELECT COUNT(*) FROM text_processor;
+-----+
| COUNT(*) |
+-----+
| 14705 |
+-----+
1 row in set (0.01 sec)
```

**31만개의 이미지 처리 후  
원재료명이 들어간 이미지  
14705개의 텍스트 저장**

## 4. 개발 과정

### DB) 데이터 베이스 모델링 (ERD)



## 4. 개발 과정

### NLP) 원재료 추출 모델

if (   ) ↓  
developer



- 카카오에서 1년간 제작한 OCR도 여전히 오타 존재
- OCR만 6개월을 진행할 수는 없음

OCR에서 추출한 텍스트에서 오타교정과 원재료명 추출을 위한 자연어 처리 모델로 변경

## 4. 개발 과정

### NLP) 원재료 추출 모델

### GCP 텍스트 출력시 결과물

제품명	허니버터칩		
식품의 유형	과자(유탕처리제품)	내용량	138 g
원재료명	감자(국내산), 혼합식용유[팜올레인유(말레이시아산), 해바라기유(외국산(우크라이나, 스페인, 말레이시아 등)), 토코페롤(혼합형)], 복합조미식품[허니버터맛시즈닝{결정과당, 백설탕, 정제소금(국내산), 탈지분유(미국산), 버터혼합분말 (가공버터(호주산)), 아카시아꿀분말(아카시아꿀(국내산), 고매버터 (프랑스산), 합성감미료 (수크랄로스)}]		
유통기한	알레르기유발물질: 우유, 대두, 밀, 굴 함유 후면표기일까지   포장재질   폴리에틸렌(내면)		
제조원	F1: 해태가루비(주)/강원도 원주시 문막읍 문막공단길 154, F2: 해태가루비(주)문막 제2공장/강원도 원주시 문막읍 반산로 54		
유통전문판매원	해태제과식품(주)/서울특별시 용산구 한강대로 72길 3(남영동)		
품목보고번호	F1: 1995038332688, F2: 201603834281		
• 직사광선 및 습기찬 곳을 피하여 서늘하고 통풍이 잘 되는 곳에 보관하여 주십시오. • 본 제품은 공정거래위원회 고시에 의거 피해보상 • 반품 및 교환장소 본사 및 구입한 곳 • 부정·불량식품 신고는 국번없이 1399 ※이 제품은 계란 땅콩, 게, 새우, 돼지고기, 복숭아, 토마토, 호두, 닭고기, 쇠고기, 오징어, 조개류(전복, 홍합 포함), 아황산류를 사용한 제품과 같은 제조 시설에서 제조하고 있습니다. ※신선한 제품 공급을 위해 질소 충전 포장을 하였습니다.			

```
test_list1 = [치자청색소, 허니버터칩, 제품명, 식품의, 유형, 과제유탕처리제품, 원재료명, 감
재국내산, 혼합식용유[팜올레인유[말레이시아산], 해바라기유, 외국산, 우크라이나, 스페인, 말레이시
아, 등, 토코페롤, 혼합형, 복합조미식품, [허니버터맛시즈닝결정과당, 백설탕, 정제소금, 국
내산, 탈지분유미국산, 버터혼합분말, 가공버터(호주산), 아카시아꿀분말, 아카시아꿀, 국내산, 고
메버터, 프랑스산, 합성감미료, 수크랄로스], 알레르기유발물질, 우유, 대두, 밀, 굴, 함유
유통기한, 후면 표기일까지, 제조원, 내용량, 138g, LOLA, 0, TTT, 포장재질, 폴리에틸렌>
내면, F, 해태가루비, 주, 강원도, 원주시, 문막읍, 문막공단길, 154, F2, 해태가루비
주, 문막, 제2공장강원도, 원주시, 문막읍, 반산로, 54, 유통전문판매원, 해태제과식품, 주
서울특별시, 용산구, 한강대로, 72길, 3, 남영동, 품목보고번호, F1, 1995038332688, F2
201603834281, • 직사광선, 및, 습기찬, 곳을, 피하여, 서늘하고, 통풍이, 잘, 되는
곳에, 보관하여, 주십시오, • 본, 제품은, 공정거래위원회, 고시에, 의거, 피해보상, • 반품
및, 교환장소, 주본사, 및, 구입한, 곳, 부정불량식품, 신고는, 국번없이, 1399, ※이
제품은, 계란, 땅콩, 게, 새우, 돼지고기, 복숭아, 토마토, 호두, 닭고기, 쇠고기, 오징어,
장어, 조개류, 전복, 홍합, 포함, 아황산류를, 사용한, 제품과, 같은, 제조, 시설에서,
제조하고, 있습니다, ※신선한, 제품, 공급을, 위해, 질소, 충전, 포장을, 하였습니다]
```

### 원하는 출력 형태 :

치자청색소, 혼합식용유, 탐올레인유, 해바라기  
유, 토코페롤, 허니버터맛시즈닝, 결정과당, 백설탕,  
정제소금, 탈지분유, 버터혼합분말, 우유, 대  
두, 밀, 굴

어떻게 원하는 단어만 얻을 수 있을까?

## 4. 개발 과정

### NLP) 원재료 추출 모델

#### 시도 1. 필요없는 단어들을 거르기 (Stopword)

```
stop_words = [  
    '피해보상', '정', '의거', '표기', '생인전제', '식품유형', '처리',  
    '원료', '가공', '부분적', '고시', '재품은', '위생적', '본사', '폴리프로필렌',  
    '제품', '보고', '기한', '잘', '포장재질', '곳에', '식', '후면', '습', '구입한',  
    '환불', '신고는', '지정', '되는', '현', '슬기차', '시푸으혀', '이', '교환장소', '신고',  
    '인증', '식약처', '직사', '시', '프로필렌', '교환', '처', '주십시오', '마크', '사항',  
    '소비자', '부으', '포장', '곳', '부정', '파개', '재질', '번호', '므로', '관리',  
    '봉', '품목', '섭취전', '껍질', '작사광선', '전과', '물질', '건조', '유통기한', '부정',  
    '제조', '전', '번품', '약', '식품', '직시광선', '제품명', '식품안전관리인증', '등', '함위',  
    '품목보고번호', '개봉', '통풍이', '흡습', '주의', '폴리', '함', '서늘하고', '유형', '광선', '코',  
    '주의사항', '불량', '후', '유', '소비자', '식봉약', '전과정', '및', '전제', '기본법', '안전', '보관하여',  
    '일', '파자', '유통', '표기일', '곳을', '보관', '공정거래위원회', '유처리제품', '흡', '국번', '불량식품',  
    '안심', '생인', '섭취', '혼합', '홀힐제제', '제제', '미국산', '외국산', '려시아', '헝가리', '세', '세르', '비아',  
    '원재료', '원재료명', '명', '맛', '베이스', '조미', '분말', '함유', '과', '유탕', '유탕처리제품', '고객', '고객센터', '센터',  
    '수신자', '수신자부담', '부담', '품저', '영업소', '구입', '내면', '풍', '풍목보고번호', '목', '이상', '경우', '구입점포', '점포',  
    '전국', '지점', '고객상담실', '상담실', '반품', '서울', '수신자요금부담', '요금', '호남', '광주', '광주광역시', '광역시', '내용',  
    '내용량', '량', '영양', '정보', '1일', '영양성분', '성분', '기준치', '비율', '업', '업소명', '소명', '소재지', '주', '주호남',  
    '해결', '기준', '공정', '거래', '위원회', '보상', '수', '반', '반품', '장소', '제조원', '판매원', '품질', '보전', '공기', '충전',  
    '포상', '진', '치아', '치아손상', '손상', '우려', '있으', '비닐', '있습', '신선도', '보존', '질소', '질소충전', '하', '소비', '자',  
    '피해', '직사공선', '공선', '손상', '습도', '급격', '요', '전화', '문자', '금', '금처', '이스', '영양정보', '밀양', '밀양시', '유동기',  
    '동기', '표', '기일', '포장지질', '지질', '폴리에틸렌', '내용량당', '농업', '농업회사법인', '회사', '법인', '오리온', '오리온농협', '농협',  
    '주식', '주식회사', '전상', '도', '나트륨', '소', '재지', '백범', '길', '수화물', '원재료명목', '옥', '호주', '말레이시아', '말레이시아산', '산',  
    '스페인', '스페인산',  
]
```

문제 : 사람이 단어를 필터링하기엔 필터링할 단어가 너무 많다

### NLP) 원재료 추출 모델

#### 시도 2. Konlpy를 이용해 형태소를 분석 후 명사만 이용

```
[('해바라기', 'NNG'), ('유', 'NNG')]  
[('외국산', 'NNG')]  
[('우', 'NNG'), ('크', 'VA'), ('라', 'ECD'), ('이', 'NP'), ('나', 'JX')]  
[('스페인', 'NNG')]  
[('말레이시아', 'NNP')]  
[('등', 'NNG')]  
[('토코페롤', 'NNG')]  
[('혼합형', 'NNG')]  
[('복합', 'NNG'), ('조미', 'NNG'), ('식품', 'NNG')]  
[(['', 'SS')]  
(['하', 'VV')]  
(['니', 'ECD')]  
(['버터', 'NNG')]  
(['맛', 'NNG')]  
(['시', 'NNG')]  
(['즈닝', 'UN')]  
(['결정', 'NNG')]  
(['과', 'JC')]  
(['당', 'NNG')]  
[('백설탕', 'NNG')]  
[('정제', 'NNG'), ('소금', 'NNG')]  
[('국내산', 'NNG')]  
[('탈', 'NNG'), ('지분', 'NNG'), ('위', 'NNG'), ('미국산', 'NNG')]  
[('버터', 'NNG'), ('혼합', 'NNG'), ('분말', 'NNG')]  
[('가공', 'NNG'), ('버테', 'UN'), ('호주', 'NNG'), ('산', 'XSN')]  
[('아카시아꿀', 'NNG'), ('분말', 'NNG')]  
[('아카시아꿀', 'NNG')]  
[('국내산', 'NNG')]  
[('이', 'VCP'), ('고', 'ECE'), ('메', 'NNG'), ('버터', 'NNG')]  
[('프랑스', 'NNG'), ('산', 'XSN')]  
[('합성', 'NNG'), ('감미료', 'NNG')]  
[('수', 'NNB'), ('크랄', 'UN'), ('로스', 'NNG'), (']', 'SS')]  
[('알레르기', 'NNG'), ('유발', 'NNG'), ('물질', 'NNG')]  
[('온', 'NNG')]
```

**허니버터맛시즈닝결정과당 :**  
**허, 니, 버터, 맛, 시, 즈닝, 결정, 과, 당**

**탈지분위 :** 탈, 지분, 위  
**가공버테 :** 가공, 버테  
**수크랄로스] :** 수, 크랄, 로스, ]

**문제 :** 조금이라도 사전에 있는 단어와 어긋나면 제대로 작동불가

## 4. 개발 과정

### NLP) 원재료 추출 모델

#### 시도 3. FASTTEXT를 이용한 오타 교정 모델 구현

```
model = FastText(sentences = products, size = 100, window = 5, min_count = 5, workers = 4, sg = 0)
skipgram_model = fasttext.train_unsupervised(
    'mat_test.txt',
    model = 'skipgram')
cbow_model = fasttext.train_unsupervised(
    'mat_test.txt',
    model = 'cbow')
```

Word : 혼합식용위듬올레인유말레이시아산  
[-0.01311858 -0.01588668 -0.05187563 -0.01566839 0.03739906 -0.0017157  
0.04063955 -0.01416075 0.00072793 -0.05280482 0.01575475 0.00015604  
-0.02103458 -0.0049095 0.00371732 -0.01412302 0.01209693 0.0093779  
-0.00294398 0.00547642 0.00096039 0.02376367 0.00153844 0.00829637  
-0.01130184 -0.00448126 -0.025183 0.01664916 0.00529875 0.00258075  
-0.00787498 -0.01462745 -0.00177572 0.00226074 0.00566321 0.00526664  
-0.00761347 0.00604911 -0.00748946 -0.00566926 -0.00032247 0.00258158  
0.002117 -0.00366235 -0.01647162 -0.02116376 0.02358607 0.00741411  
0.01726332 -0.00167425 0.01423866 0.00270188 0.01044532 0.00529633  
0.00112527 0.00094998 0.01147553 -0.00890631 -0.00016149 0.00911648  
0.00856208 0.00893442 -0.01416403 0.00262038 -0.0106391 0.00574232  
0.00744074 -0.02008644 -0.00265141 0.01687132 0.00804999 -0.01410521  
0.01237786 0.00681428 -0.00666253 0.01255275 -0.01413474 0.01083147  
0.00988377 0.01207598 0.00480753 0.01291618 0.00940145 -0.01397438  
-0.01090315 -0.01948567 -0.00804586 -0.02186265 -0.01347922 0.02222546  
0.00063784 0.01224336 0.01105155 0.00932292 0.00057703 0.02046198  
0.00764262 0.00730082 0.005019 0.00299509]

불 : 특장점 피부 뷰티 뷰티뉴트리션 뉴트리션 -1 유지 오 오메기3 매기 비타민  
향신 북 북대서양 대서양 부 부원료 여성 생각 캡슐 원료명 힘량 정제어유  
량 섭취방법 방법 테이트 피 너 츠 즈오일 오일 디 디얼알파토코페롤 얼  
겐 레시틴 캡슐기제 기제 젤라틴 우내 우내피 글리세린 알루미늄 호일 염화  
액 카카오 카카오색소 색소 종 대두 쇠고기 기능 기능정보 헐 혈중 중  
형성 상쾌 포 건강기능식품유통 전문 전문판매원 성장 빌달 유해 소로 세포  
구 오창읍 창 읍 각리 1길 씨 씨제이제일제당 제이 제일 제일제당 제당  
화물 단백질 지방 우수 기술력 바탕 엄격 품질관리 영 합 제조번호 영양소기  
네스 네스카페 카페 크레 크레마 마 라 라떼 때 커피 롯데 롯데네슬레코리  
봉로 72번길 번 맥 표시 일자 식물성 식물성크림물엿 크림 물엿 식물성유지경화  
랜드 랜드산 조절제 혼합분유 분유 혼합분유위탈자 위 탈지 문 문유 소비자상  
스 프랑스산 혼합분위탈자 불 위 덱스트린 유당 카 카제인 제인 식물성유지  
식품위생법 위생법 한글 한글표시사항 표시 보관방법 방법 개봉후 냉장 냉장보관  
폐리 틀레인 탄산수 수입업 수입업소명 마트 5678 서울특별 서울특별시 성동  
지 탄수화물 당류 지방 트랜스 트랜스지방 포화 포화지방 콜레스테롤 단백질 드  
채 채주스 주스 휴 휴럴 럼 충북 청주 청주시 청원 청원구 구 오 오  
방법 컵 1컵 희석 기호 만시 착즙 착즙액 액 베트남 함량 비타민 내포  
확인 산지 생육 생육환경 환경 채취 채취시기 시기 색상 차이 변질 내용물  
정당 ]

문장 기반의 NLP인 FASTTEXT  
: 데이터셋이 잘 맞지 않음

∴ 형태소 분석 모델보다  
낮은 정확도 도출

## 4. 개발 과정

### NLP) 원재료 추출 모델

#### 시도 4. WhiteWord를 이용한 문장 필터링

오리지널노니주스  
물 한방울 섞이지 않은 노니 원액  
check point  
100% 노니원액  
NFC 오리지널 착즙주스  
Original 100  
NONI  
노니주스  
ORIGIN  
100%  
Original 100%  
FRESH REAL JUICE  
노니주스  
NON  
ORIGINAL NONI 100 %  
아방을 섞이지 않은 노니 원액  
노니주스  
0%  
100kcal)  
노니  
란 무엇인가요?  
노니란, 겉면이 유통불통한 감자 모양을 지닌  
열대과일로 해파극, 치즈과일, 인도뽕나무 등  
세계 각 지역에서 다양한 이름으로 불리며  
주로 주스, 분말, 차 등으로 가공하여 먹습니다.  
#인도뽕나무  
#해파극 #치즈과일

출력 텍스트를 줄 단위로  
편집

100% 노니원액  
  
아방을 섞이지 않은 노니 원액  
  
노니

WhiteWord를 구성하여  
WhiteWord가 포함된 문장  
필터링

노니원액  
  
노니 원액  
  
노니

남은 문장에서  
원재료 코퍼스를 이용해  
오태교정

원재료명:  
노니  
노니원액

형태소로 분리후  
원재료명일 확률  
높은 명사만  
저장

## 4. 개발 과정

### NLP) 원재료 추출 모델

#### 시도 4. whiteWord를 이용한 문장 필터링

물 YES 기 소스 설탕 간장 정제수 밀 정제소금 설탕 과당시럽 정제수 전분 카라멜색소 식초 정제수 초산 밀가루 대두어 두유 포도당 설탕 밀 정제수 코지 코지 옥수수전분 흰후추분말 정제소금 닭고기 대두 밀 함유

제품명 맛있는 숯불 닭꼬치(간장소스 맛)  
축산물  
| 가공품의 양념육(살균제품)  
| 수입원 |(주)이마트 02)380-567  
서울특별시 성동구 뚝섬로377(성수동2가  
|| 3. 에어프라이어 - 별도의 해동없이 180° °C에서  
| 제조사 CPF (Thailand) Public  
Company Limited (EST. 42)  
뒷면 표시일까지 (년/월/일순)  
500g(1,284kcal)  
섭취하십시오.  
해동후재 냉동시키지  
마시기 바랍니다.  
이 제품은 계란, 우유를 사용한 제품과 같은 제조  
시설에서 제조하고 있습니다.  
량식품 신고는 구분없이 1300  
| 전새도 기다리살), 글레이징소스[설탕, 간장(정제수)|  
| 본제품은 공정거래위원회고시 소비자 분쟁해결  
밀, 정제소금, 설탕), 과당시럽, 정제수, 변  
선분 카라멜색소, 식초(정제수, 초산), 마리네 |  
|| 0반품 및 교환장소: 구입체  
트정제수, 밀가루, 대두유, 포도당수화물, 설탕  
제(L글루탐산나트륨), 트레할로스, 폴  
| 영양정보  
산나트륨, 간장(정제수, 대두, 밀, 정제소금,  
밀, 정제수), 코지(코지, 옥수수전분(유전  
|| |나트륨  
옥수수포함 가능성 있음), 감자전분) | 탄수화물  
| 흰후추분말, 정제소금]  
기준에 의거 교환 또는 환불받을 수 있습니다  
총 내용량 600g  
100g 당 214kcal  
기준치에 대한 비율  
닭고기, 대두, 밀 함유  
나포상-폴리에틸렌 비닐류  
1일 영양성분 기준치에 대한 비율(%)는  
2,000kcal 기준이므로 개인의 필요 열량에  
따라 다를 수 있습니다.

약 60%의 정확도

일부 의도치 않은 원재료 파생

-오타 교정 중 발생하는 오류

-필요 없는 단어도 원재료 코퍼스와  
비교하는 과정에서 발생

## 4. 개발 과정

### NLP) 원재료 추출 모델

#### 시도 5. 딥러닝 & 머신러닝 도입

제품명 맛있는 숯불 닭꼬치(간장소스 맛)  
축산물  
| 가공품의 양념육(살균제품)  
| 수입원 |(주)이마트 02)380-567  
서울특별시 성동구 뚝섬로377(성수동2개  
|| 3.에어프라이어 - 별도의 해동없이 180° °C 에서  
| 제조사 CPF (Thailand) Public  
Company Limited (EST. 42)  
뒷면 표시일까지 (년/월/일순)  
500g(1,284kcal)  
섭취하십시오.  
해동후재 냉동시키지  
마시기 바랍니다.  
이 제품은 계란, 우유를 사용한 제품과 같은 제조  
시설에서 제조하고 있습니다.  
당신은 구매합니다. 1399  
| 전라도 기다리살), 글레이징 소스[설탕, 간장(정제수)]  
| 본제품은 공정거래위원회 고시 소비자 분쟁해결  
밀, 정제소금, 설탕), 과당시럽, 정제수, 변  
선분 카라멜색소, 식초(정제수, 초산), 마리네 |  
|| 0반품 및 교환장소: 구입체  
트정제수, 밀가루, 대두유, 포도당수화물, 설탕  
제(글루탐산나트륨), 트레할로스, 폴  
| 영양정보  
산나트륨, 간장(정제수, 대두, 밀, 정제소금,  
밀, 정제수), 코지(코지, 옥수수전분(유전  
|| |나트륨  
옥수수포함 가능성 있음), 감자전분) | 탄수화물  
| 흰후추분말, 정제소금]  
기준에 의거 교환 또는 환불받을 수 있습니다  
총 내용량 600g  
100g 당 214kcal  
기준치에 대한 비율  
닭고기, 대두, 밀 함유  
나포장-폴리에틸원 비닐류  
1일 영양성분 기준치에 대한 비율(%)는  
2,000kcal 기준이므로 개인의 필요 열량에  
따라 다를 수 있습니다.

#### 일부 의도치 않은 원재료 파생

- 오타 교정 중 발생하는 오류
- 필요 없는 단어도 원재료 코퍼스와 비교하는 과정에서 발생
- 필요 없는 단어는 원재료가 없는 문장의 단어

원재료가 있는 문장을  
구분할 수 있다면?

## 4. 개발 과정

### NLP) 원재료 추출 모델

#### 시도 5. 딥러닝 & 머신러닝 도입

##### Training Set

- 이미지 추출 텍스트의 한 문장
- 해당 문장의 원재료명 포함 여부

"상품 상세정보  
식품의 유형  
과·채주스  
70 mL x 50포  
내용량  
품목보고번호  
20040275056-317  
원재료명 및 함량  
석류과즙 100% (13 Brix 이상, 정제수, 석류농축액 / 이란산)  
내포장재질  
반품 및 교환처  
권장섭취량 및 섭취방법  
폴리에틸렌(PE)  
구입처 또는 고객상담실  
성인 1일 1~3회, 1회 1포씩 음용하십시오  
① 전자렌지에 직접 넣어 데우지 마십시오.", 1  
② 개봉 시 내용물이 흘러 나올 수 있으므로 주의하십시오., 0  
③ 포장이 변형, 팽창, 손상되었거나 내용물이 변질되었을 경우  
음용하지 마십시오.", 0  
④ 제품의 개봉 또는 섭취 시 포장재에 의해 상처를 입을 수  
있으니 주의하시기 바랍니다.", 0  
⑤ 임산부, 수유부, 특이체질, 알레르기 체질이신 분은 성분을  
확인 후 섭취하여 주십시오.", 0  
⑥ 이 제품은 우유, 메밀, 대두, 밀, 계, 새우, 복숭아, 토마토 등을  
사용한 제품과 같은 제조시설에서 제조하고 있습니다.", 0

-> 원재료명을 포함한 문장 : 1

-> 원재료명을 미포함한 문장 : 0

## 시도 5. 딥러닝 & 머신러닝 도입

[  
[ "/#/SW",  
"인도/NNP",  
"뽀/VV",  
"ㅁ/ETN",  
"나/JC",  
"무/NNP",  
"/#/SW",  
"해/NNG",  
"파/NNP",  
"ㅋ/NNG",  
"/#/SW",  
"치즈/NNP",  
"과일/NNP",  
"ㅣ/SW",  
"짜/VV",  
"아/EC",  
"노니주스/NNP",  
"를/JK0",  
"만나/VV",  
"아/EC",  
"보/VX",  
"ㅅ/EP",  
"어요/EF",  
"!/SF"  
]  
0  
],

- 각 문장의 형태소를 나눠 json 형태로 저장
  - 데이터에서 가장 많이 나타난 형태소 10000개를 추출, 벡터화
  - 벡터화된 데이터를 여러 모델에서 실험
  - **LogisticRegression 모델**에서 가장 높은 성능 측정되었음

## 4. 개발 과정

### NLP) 원재료 추출 모델

#### 결과 5. 딥러닝 & 머신러닝 도입

테이스티홈의 데리야끼 닭꼬치 달콤 짭쪼름한 소스에 부드러운 닭다리살의 식감! 데리야끼 닭꼬치 제품명 닭꼬치 100g (50gx2) 중량 및 구성  
가열 양념육(살균제품) 축산물가공품의 유형 제조사(주)다누립푸드 포장상태 냉동 냉동보관(-18°C 이하) 보관방법 유통기한 제조일로부터 1년  
성분표시 닭고기 77.08%(정육(브라질산) 꼬치만발데리소스 21.04% 물엿(옥수수전분 100%(옥수수(미국산/브라질산/우크라이나산등)), 정제수, 산분  
해간장(정제수, 탈지대두(인도산/중국산/미국산), 천일염(호주산), L-글루타민산나트륨(향미증진제), 고과당, 카라멜), 백설탕, 미향(정제수, 기타과  
당, 포도당, 레몬식초(레몬 100%(이스라엘산)), 정제소금(국내산), 주정), 우스타소스(정제수, 설탕, 발효식초(주정, 발효영양원, 정제소  
금(국내산), 혼합간장(탈지대두(외국산)), 정제소금(국내산), 토마토페이스트(중국산)), 변성전분, 냉동마늘(중국산), 양파(국내산), 정제소금(국내  
산), 카라멜색소, 잔탄검, 스모크향 대두, 밀, 쇠고기, 토마토 함유 데리야끼 닭꼬치 100% 닭다리살을 사용하여 쫄깃하고 부드러운 식감을 자랑합  
니다.

[['감', '천일염', '카라멜', '잔탄검', '소스', '옥수수전분', '정제수', '스모크향', '닭다리살', '백설탕', '포도당', '물엿', '토마토', '산  
분해간장', '물', 'L-글루타민산나트륨', '닭고기', '변성전분', '냉동마늘', '데리야끼', '기타과당', '발효영양원', '탈지대두', '양념육', '  
혼합간장', '고과당', '옥수수', '발효식초', '정제소금', '함유', '레몬', '제조', '미향', '설탕', '주정', '대두', '카라멜색소', '토마토페  
이스트', '레몬농축과즙', '양파']]

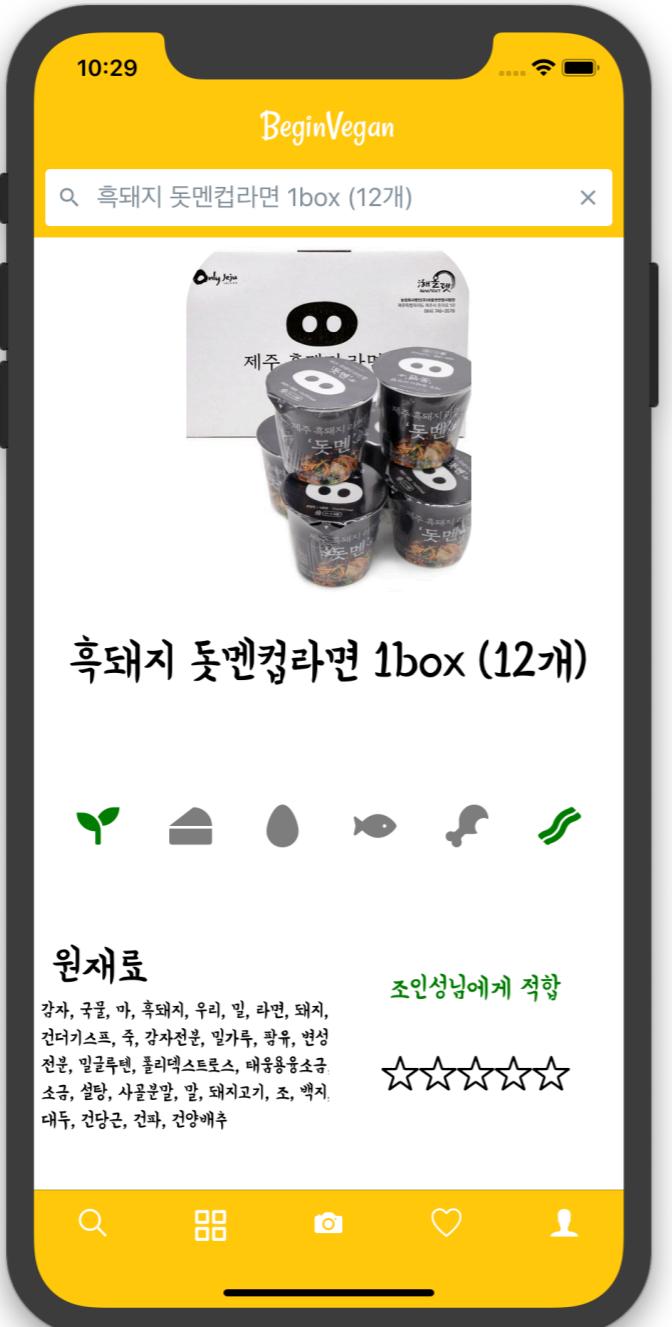
부정, 불량식품 신고는 국번없이 1399 [전새도 기다리살], 글레이징소스 [설탕, 간장(정제수)], 본제품은 공정거래  
위원회고시 소비자 분쟁해결 밀, 정제소금, 설탕), 과당시럽, 정제수, 변선분 카라멜색소, 식초(정제수, 초산),  
마리네 [0반품 및 교환장소: 구입체 트정제수, 밀가루, 대두유, 포도당수화물, 설탕 증진제(L글루탐산나트륨)  
, 트레할로스, 폴 | 영양정보 리인산나트륨, 간장(정제수, 대두, 밀, 정제소금, 주정당밀, 정제수), 코지(코지, 옥  
수수전분(유전), 나트륨 자변형 옥수수 포함 가능성이 있음), 감자전분) | 탄수화물 | 흰후추분말, 정제소금] 기준에  
의거 교환 또는 환불받을 수 있습니다.

[['정제소금', '트레할로스', '밀가루', '감자전분', '나트륨', '소스', '대두', '인산나트륨', '과당시럽', '당  
밀', '정제수', '옥수수전분', '설탕', '옥수수', '국', '주정', '초산', '식초', '흰후추분말', '포도당', '카  
라멜색소', '간장', '두유', '코지']]

- Logistic Regression 모델로 필터링한 문장을 Whiteword로 필터링
- 95% ~ 100% 원재료를 추출, 1~2 개의 미포함 원재료를 포함 오류
- 약 91%의 원재료 추출 성능 달성

## 4. 개발 과정

### 추천 알고리즘 구현



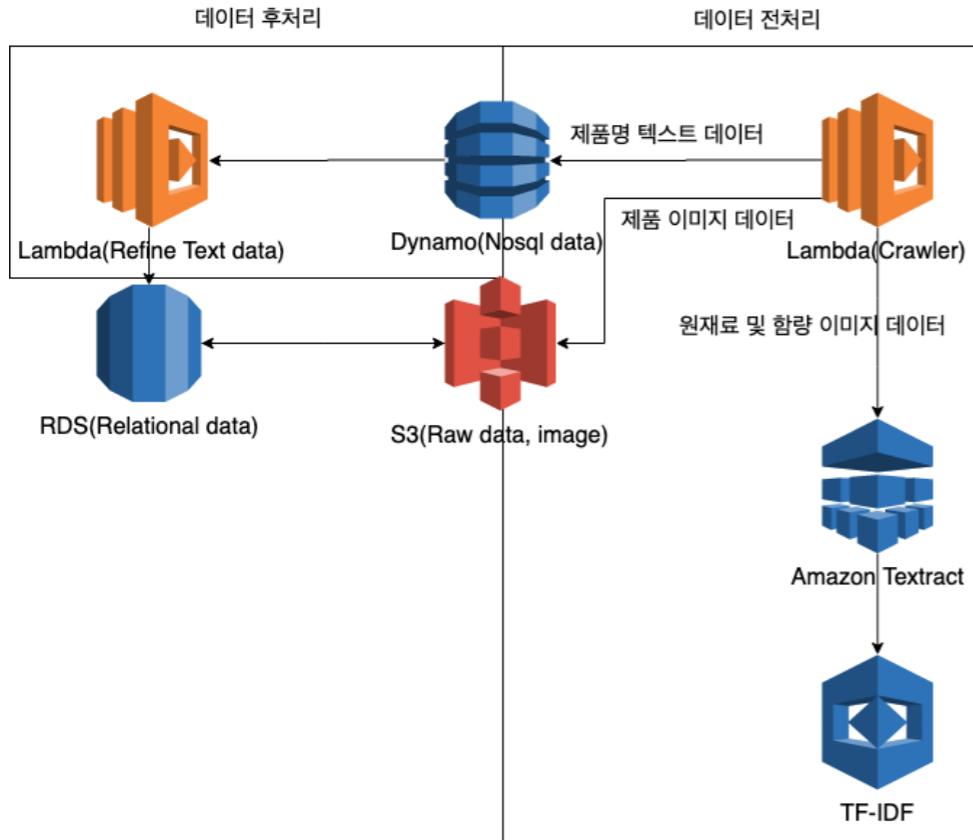
- K-Nearest Neighbor Algorithm 이용
- 사용자가 평가한 제품 평점을 기반으로 작동
- 유사도 중 높은 값을 가지면서  
비건 단계에 맞는 제품 제공

- 추천 알고리즘을 위해  
사용자 평가 시스템 도입

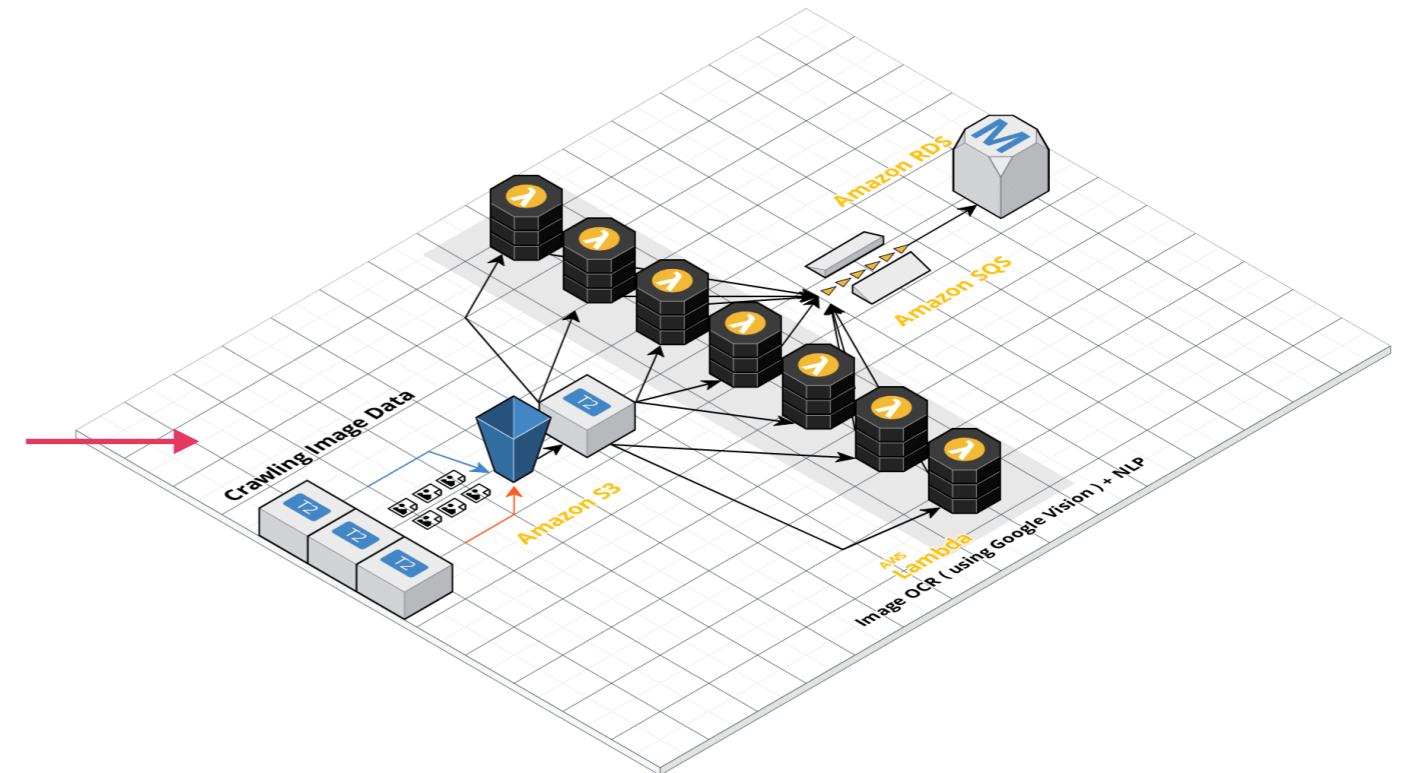
## 4. 개발 과정

# Data Engineering 아키텍처

## Before



## After



## 4. 개발 과정

# App - Server(API) 제작(사용자 플랫폼 : 어플리케이션)

## 제품상세페이지 API

### Product S Pec Api

GET /BeginVegan/productSpec/?title=%EB%9D%BC%EB%A9%B4%EC%9D%98%20%EC%9B%90%EC%

```
HTTP 200 OK
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

{
  "Img": [
    {
      "id": 18684,
      "product": {
        "id": 18690,
        "category": {
          "category": "유탕면류"
        },
        "company": null,
        "created_at": "2019-08-12T05:47:09Z",
        "updated_at": "2019-08-17T14:20:07Z",
        "name": "라면의 원조 컵 불닭볶음면 70g30개"
      },
      "created_at": "2019-08-12T05:47:09Z",
      "updated_at": "2019-08-12T07:29:43Z",
      "title": "라면의 원조 컵 불닭볶음면 70g30개",
      "image": "https://beginveganscrapdata.s3.ap-northeast-2.amazonaws.com/pd_img/76259.jpg"
    }
  ],
  "Mat": [
    {
      "material": {
        "id": 11153,
        "created_at": "2019-08-15T11:18:43Z",
        "updated_at": "2019-08-15T12:31:21Z",
        "name": "소맥분",
        "eng_name": "Wheat",
        "code_name": "식품원료(A코드)",
        "is_use": 0,
        "raw_last_update": "2018-03-30T17:36:14Z",
        "vcategory": "식물",
        "sub_category": "씨앗",
        "category": "식물"
      }
    }
  ]
}
```

## 제품검색 API

### Product List Api

GET /BeginVegan/pdList/?name=%EC%82%AC%EA%B3%BC

```
HTTP 200 OK
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

{
  "count": 982,
  "next": "http://localhost:8000/BeginVegan/pdList/?name=%EC%82%AC%EA%B3%BC&page=2",
  "previous": null,
  "results": [
    {
      "id": 76259,
      "created_at": "2019-08-12T06:25:10Z",
      "updated_at": "2019-08-12T07:42:49Z",
      "title": "힐사이드 유기농 하드캔디 사과맛",
      "image": "https://beginveganscrapdata.s3.ap-northeast-2.amazonaws.com/pd_img/76265.jpg"
    },
    {
      "id": 76147,
      "created_at": "2019-08-12T06:25:06Z",
      "updated_at": "2019-08-12T07:42:47Z",
      "title": "흙살림 저탄소 사과 2.5kg",
      "image": "https://beginveganscrapdata.s3.ap-northeast-2.amazonaws.com/pd_img/76153.jpg"
    },
    {
      "id": 75905,
      "created_at": "2019-08-12T06:24:56Z",
      "updated_at": "2019-08-12T07:42:41Z",
      "title": "후디스 순사과 15g",
      "image": "https://beginveganscrapdata.s3.ap-northeast-2.amazonaws.com/pd_img/75911.jpg"
    },
    {
      "id": 75872,
      "created_at": "2019-08-12T06:24:54Z",
      "updated_at": "2019-08-12T07:42:40Z",
      "title": "황토농원 곤약젤리 뜬사과",
      "image": "https://beginveganscrapdata.s3.ap-northeast-2.amazonaws.com/pd_img/75878.jpg"
    }
  ]
}
```

요청 링크 : [BeginVegan/productSpec](#)

변수 : title (제품 명칭)

반환 JSON : 제품명, 제품 이미지, 카테고리  
원재료, 원재료의 비건단계

[BeginVegan/pdList](#)

name (검색어) , page

검색어가 포함된 제품명, 제품 이미지

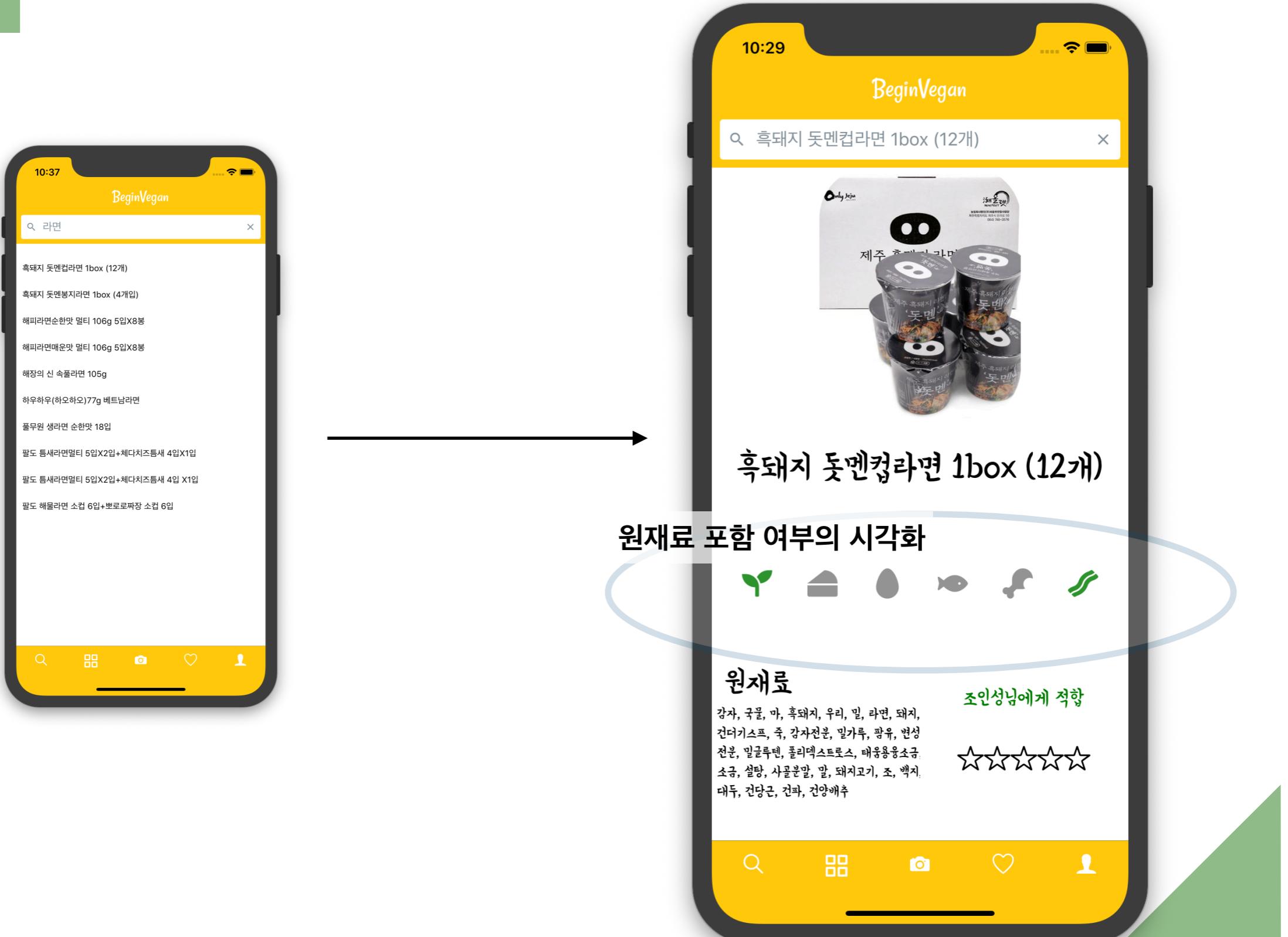
## 4. 개발 과정

### App - Frontend ) 제품의 검색



## 4. 개발 과정

### App - Frontend ) 제품의 검색



## 4. 개발 과정

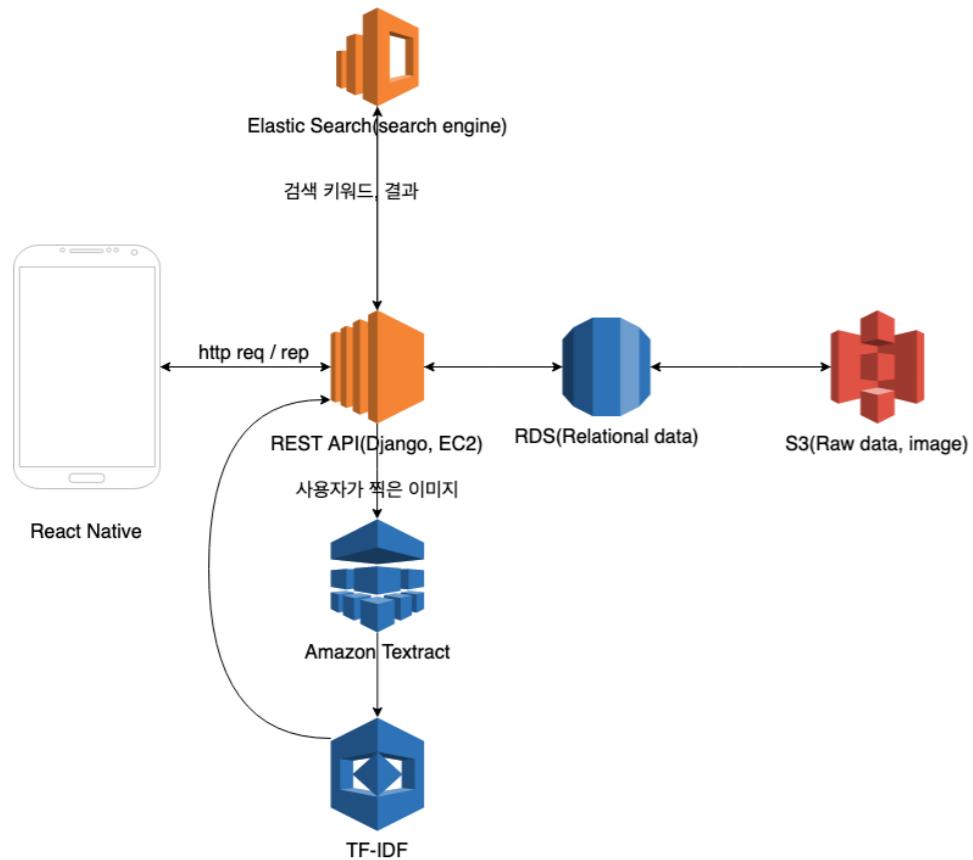
### App - Frontend ) 제품 촬영



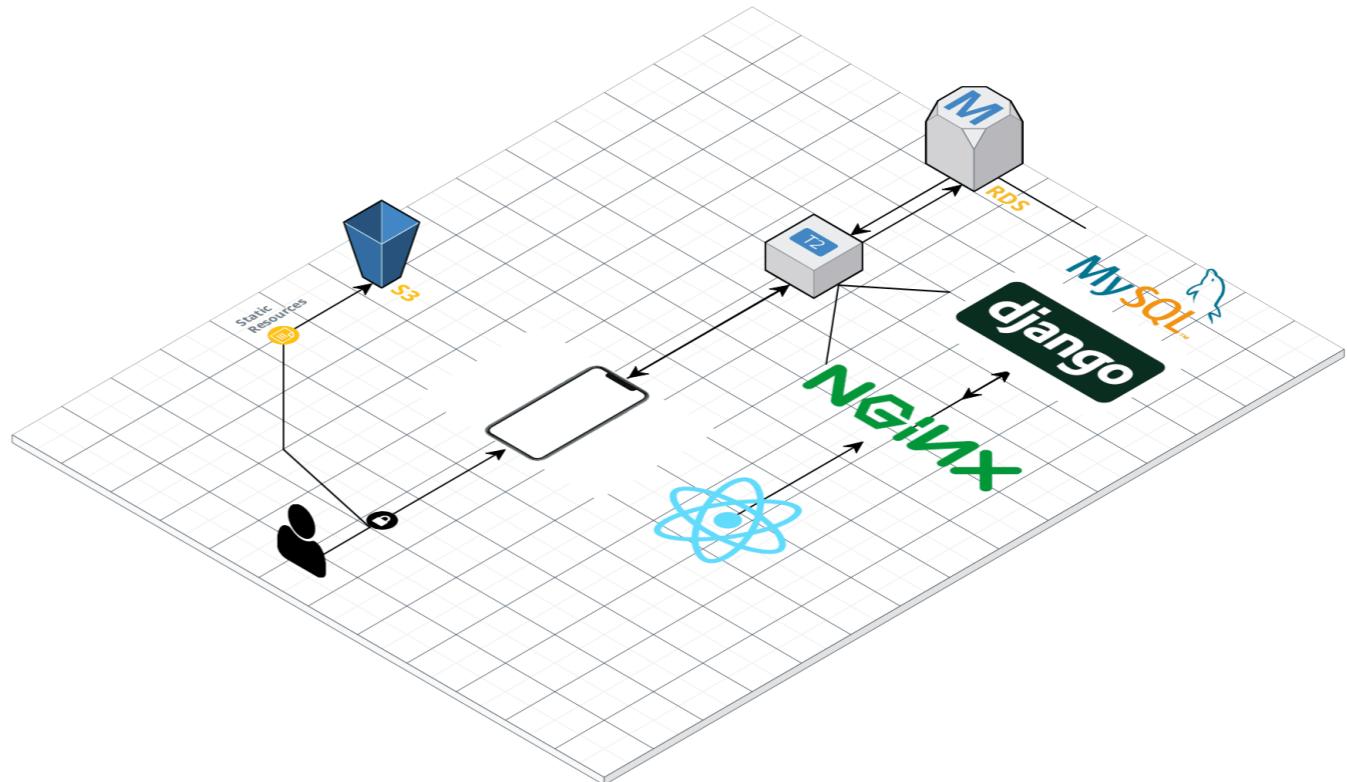
## 4. 개발 과정

# Begin Vegan App 아키텍처

### Before



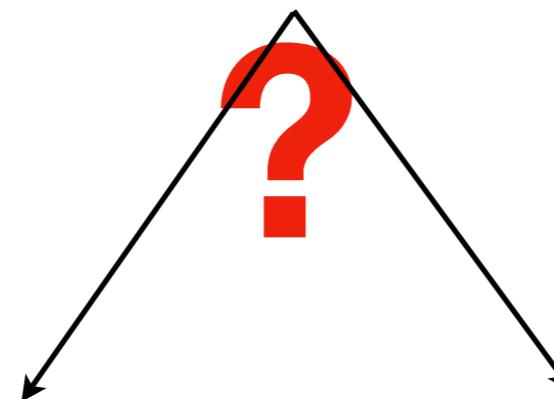
### After



## 5. 피드백 반영 :발단

비건이 아니라면 알 수 없는 문제 존재

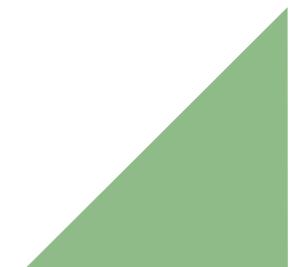
### 닭고기 소스



닭고기에  
곁들이는  
소스

닭고기가  
들어간  
소스

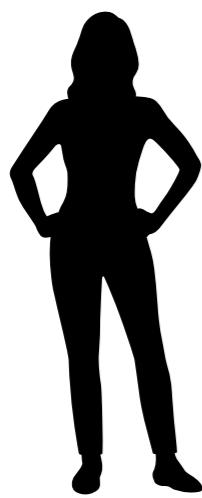
인터뷰를 통해 해결 시도



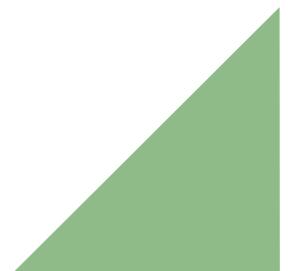
## 5. 피드백 반영

### 비건들의 피드백

이런 어플리케이션이 나온다면 정말 좋아요.  
꼭 사용할 것 같아요.  
다만 몇가지 **문제점**을 해결해주세요.



비건 5명 인터뷰 진행



## 5. 피드백 반영

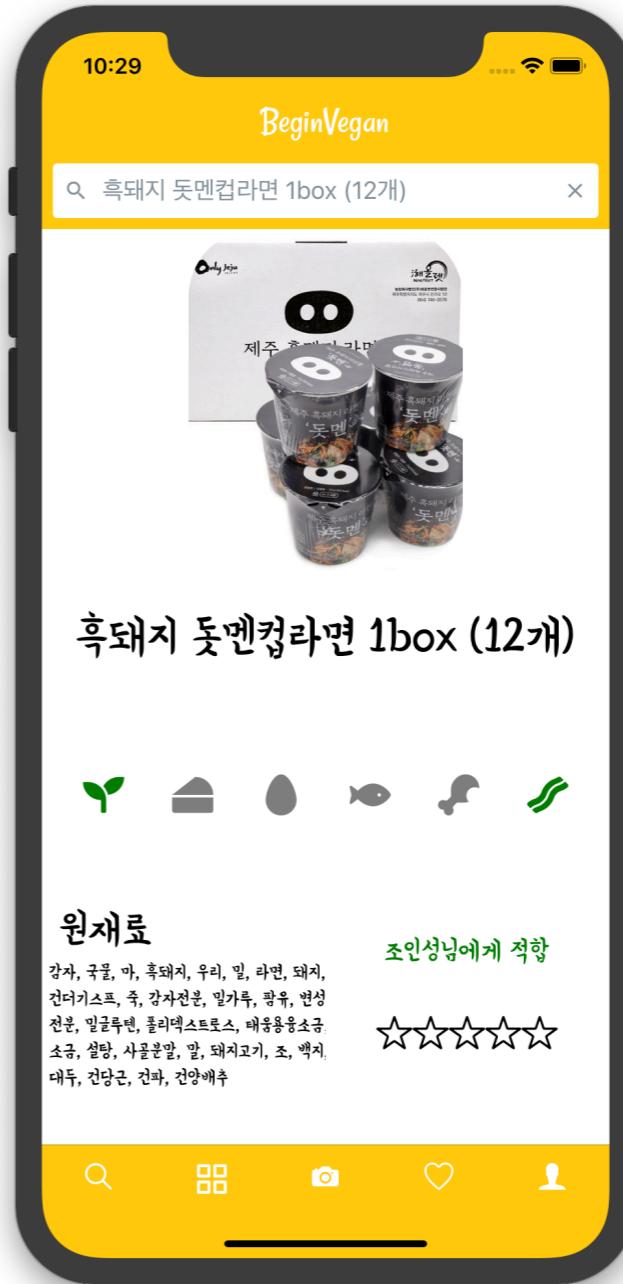
채식 단계를 명시하는 것은 바람직하지 않습니다.



비건 단계를  
단순한 시각화로 변경하여 표시

## 5. 피드백 반영

비건은 건강 이상으로 환경, 평화 또한 의미합니다.  
더 이상 녹색 어플리케이션은 싫어요.



요구에 따라 색상 변경

## 5. 피드백 반영

### 제조 시설의 특이사항을 표기해주세요.

| 제조일포함, · 이 제품은 메밀, 땅콩, 대두, 게, 사우, 돼지고기, 복숭아, 토마토, 호두, 닭고기, 쇠고기, 오징어, 조개류(굴), 야황산류를 사용한 제품과 같은 제조시설에서 생산하였습니다. ● 소비자기본법에 의한 피해보상  
● 제품교환장소: 유통전문판매원 및 구입처 · 소비자상담실: 080-739-8572 수신자부담).제품명, 알레르기 유발물질 | 밀게란우유함유,

| 0 섭취자의 신체 상하에 따라 반응에 차이가 있을 수 있으니 고객상담실로 문의 하시기 바랍니다. 0본, 호두, 닭고기 쇠고기 오징어, 조기류물, 전복, 홍합 포함, 짓을 사용한 제조시설에서 제조하셨습니다., '본 제품은 주 원료인 홍삼의 상육환경, 수학시기 등에 따라 맛과 색상의 차이가 있을 수 있습니다., |제조원,제조번호 :,

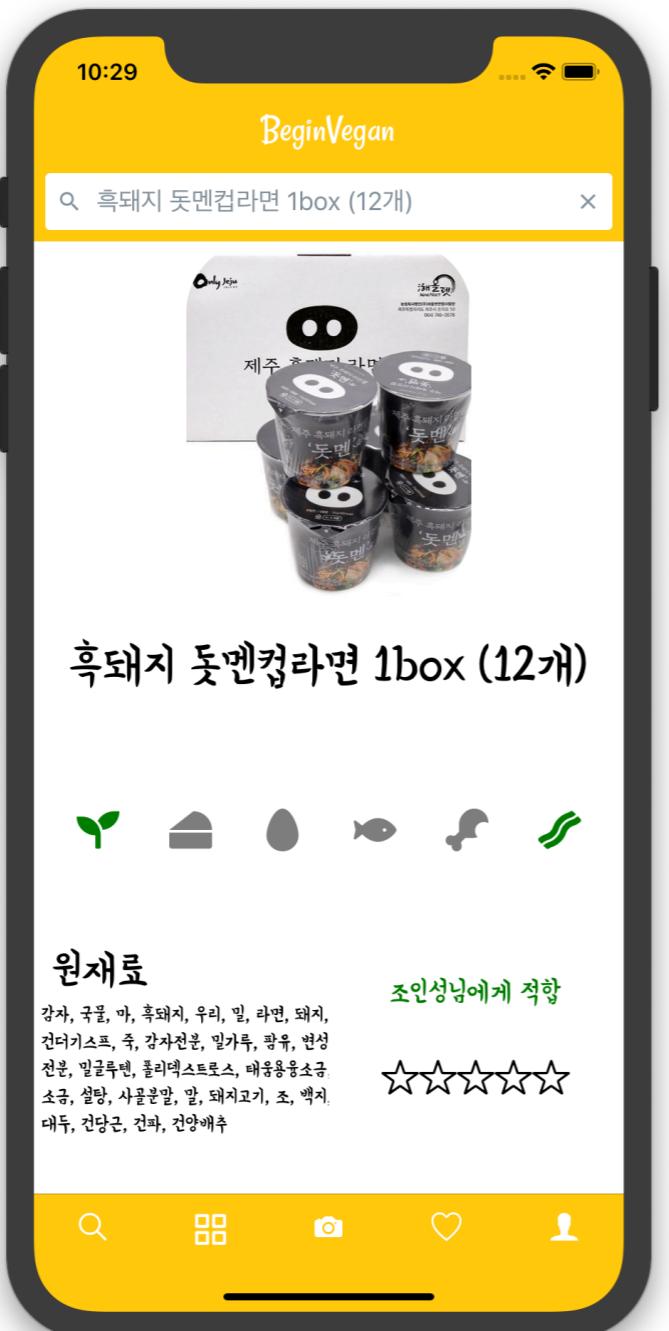
| -유유나 두유 등의 음료 및 기타 요리에 첨가하여 드셔도 좋습니다., 소량으로 효율적인 영양분을 섭취할 수 있습니다.,제품명 : 핑거루트추출분말, 중량 : 제품에 별도표기,제조원 : 제품에 별도표기,주의사항 : 본 제품은 땅콩, 호두, 대두, 우유, 밀 등을 사용한 제품과 같은 소분 시설에서 포장하고 있습니다.,본 제품은 공정거래위원회 고시 소비자 분쟁해결기준에 의거 교환 또는 보상받을 수 있습니다.,

| 제품정보,제조및판매원 | 푸드어셈블 / 부산광역시 남구 신선로337번길 83-1 (용당동),유통 기 한 | 제조일로부터 7일,[원재료명 |외]27.9%,설탕(국내산)18.6,소스류(볶음요리기본소스)[콩기름(대두유),정제소금,,제조및판매원 | 푸드어셈블 / 부산광역시 남구 신선로337번길 83-1 (용당동),유통 기 한 | 제조일로부터 12개월,[외]27.9%,고춧가루,식초,설탕(국내산) 18.6,소스류(볶음요리기본소스)[콩기름(,※ 이 제품은 난류, 메밀, 땅콩, 고등어, 게, 새우, 복숭아, 호두, 아황산류,,닭고기, 오징어, 조개류(굴, 전복, 홍합 포함)를 사용한 제품과 같은 제조,시설에서 제조하고 있습니다.,※ 제품 수령 후 바로 냉장보관 해주시고 개봉한 제품은 변질의 우려가 있으니,※ 고기 일부에 뼈가 혼입되었을 수 있으니 섭취 시 주의하십시오.,※ 본 제품은 소비자분쟁해결기준에 의거 교환 또는 보상 받을 수 있습니다.,

### 제조 시설의 특이사항 데이터베이스에 저장

## 5. 피드백 반영

분류가 애매한 제품은 나에게 맞게 알려주세요.



- K-Nearest Neighbor Algorithm 이용
- 사용자가 평가한 제품 평점을 기반으로 작동
- 유사도 중 높은 값을 가지면서  
비건 단계에 맞는 제품 제공

-**추천 알고리즘을 위해  
사용자 평가 시스템 도입**

## 6. 개발 일정

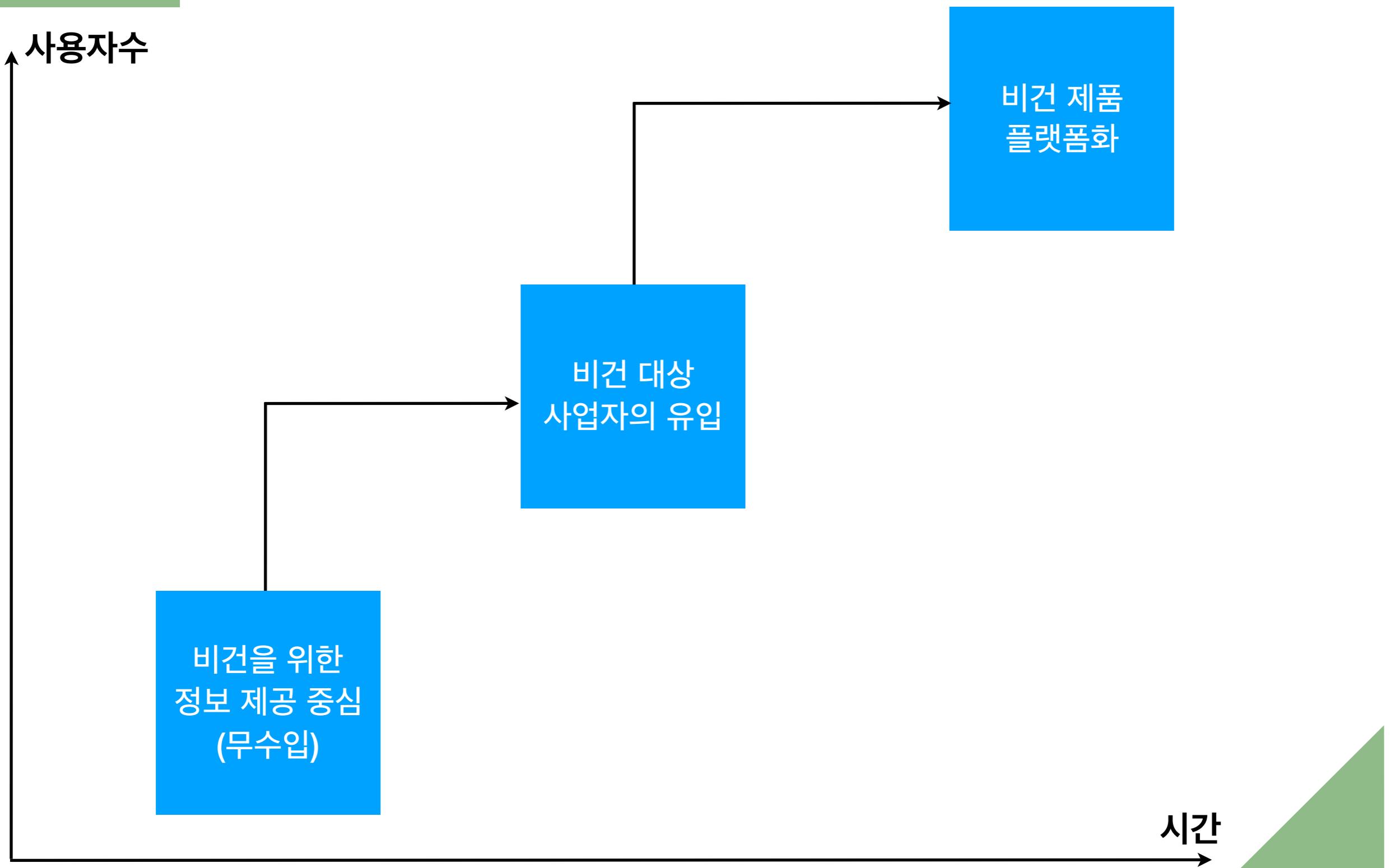
### 프로젝트 일정

기능명	구분	추진내용	진척도	프로젝트 기간					
				6월	7월	8월	9월	10월	11월
데이터 베이스	계획	멘토링을 통한 설계 개선	100%						
	분석	구성도, 개념도 정의	100%						
	설계	Django를 이용해 DB 설계	100%						
	개발	데이터베이스 프레임 구축, 데이터 입력 및 연동	100%						
	테스트	입력 데이터 검수 및 테이블 간 연결 타당성 검토	100%						
NLP	계획	원자료명 추출 설계 성능 목표 설정 학습을 위한 데이터셋 정의, 확보	100%						
	분석	탐색적 데이터 분석, 데이터 모델 구축	100%						
	검증	확보한 데이터셋 검증, 데이터 모델 성능 검증	80%						
	테스트	OCR에서 추출한 텍스트에 모델 직접 적용	100%						
카테고리 및 제품추천 알고리즘	계획	추천 방식, 알고리즘 계획	100%						
	분석	UX에 맞는 제품 구성	90%						
	설계	멘토링을 통해 설계 개선	90%						
	개발	제품 추천 시스템 개발	100%						
	테스트	어플리케이션 작동 검사	100%						
APP	계획	어플리케이션을 위한 언어 도구 선정	100%						
	분석	UI 구성	100%						
	설계	멘토링을 통해 설계 개선	100%						
	개발	API 개발, Front 개발, 디자인 개선	80%						
	테스트	어플리케이션 작동 검사	80%						

11월 22일

## 7. 추후 계획: 사업화

### 사업 Stage 모델



## 6. 추후 계획 : 사업화

### 사업 기대효과 및 수익 모델

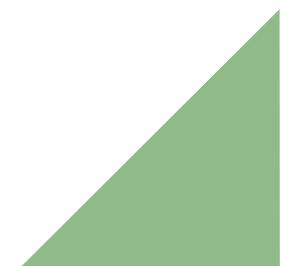


비거리 음식점 실천하는데 편의성 증가 및 진입장벽 해소



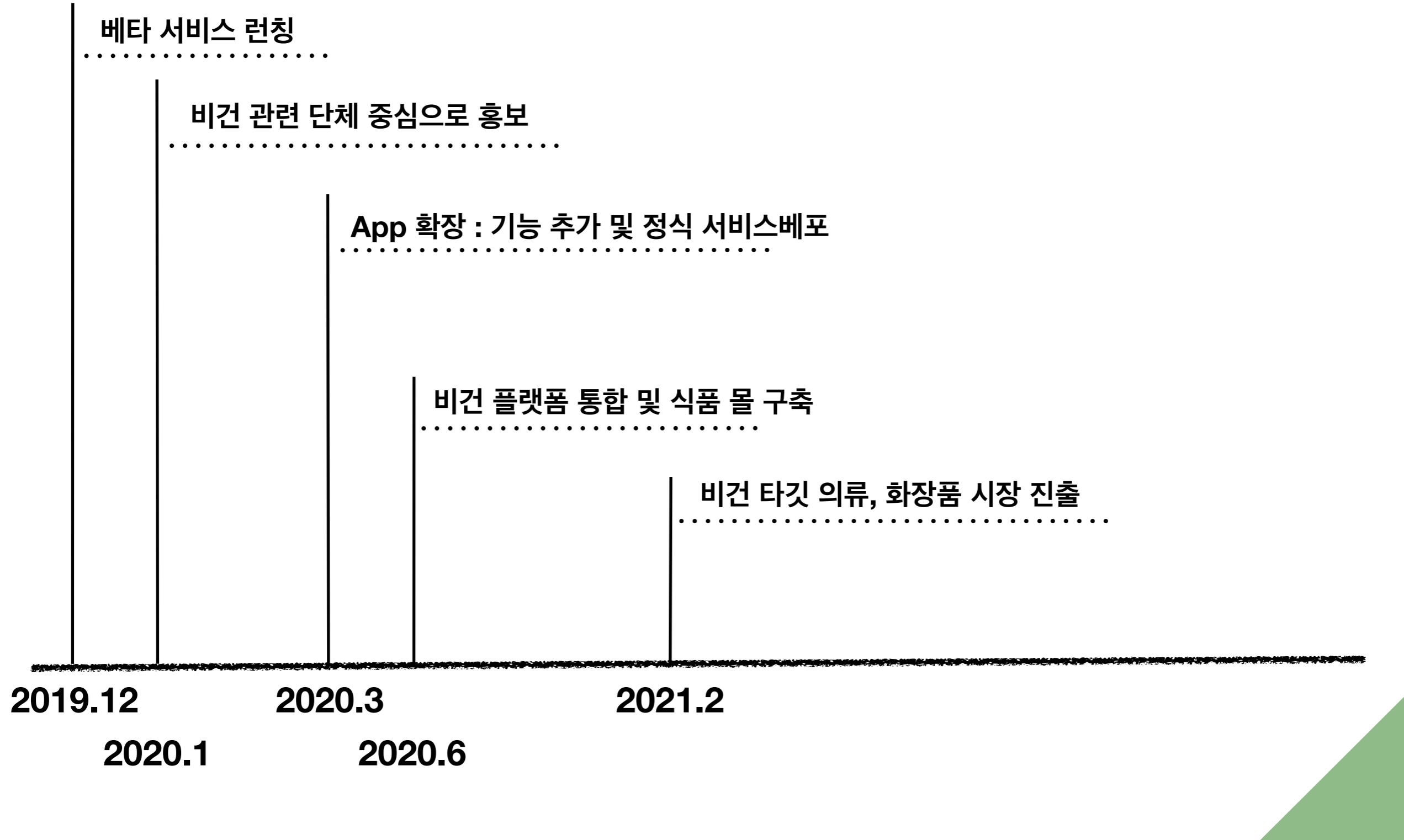
현재 플랫폼 내 채식주의 온라인 식품 몰 구축  
비건 모바일 플랫폼 구축을 위한 인프라 확보 :  

- 사용자 기록을 바탕으로 비건 관련 데이터 확보
- 식품 정보에 대한 독자적인 데이터베이스 확보
- 비건이라는 특정 계층 고객과 사업자의 중개자 역할



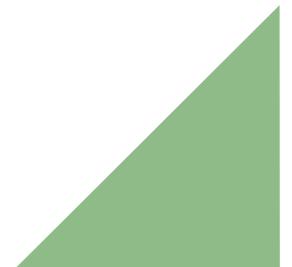
## 6. 추후 계획 : 사업화 & 기능 추가

### 프로젝트 추진방향



## 7. 데모영상

시연 영상



10:28

BeginVegan

🔍 나에게 맞는 제품은 뭐가 있을까?

카테고리로 찾기



과일



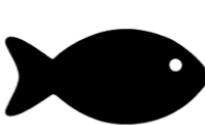
채소



쌀/잡곡



정육/계란류



수산물/해산물



유제품/냉장/냉동/간편식



과자/초콜릿/씨리얼/빵

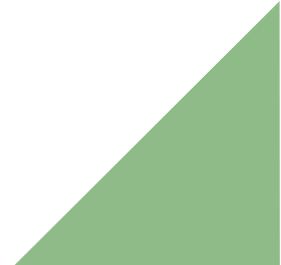
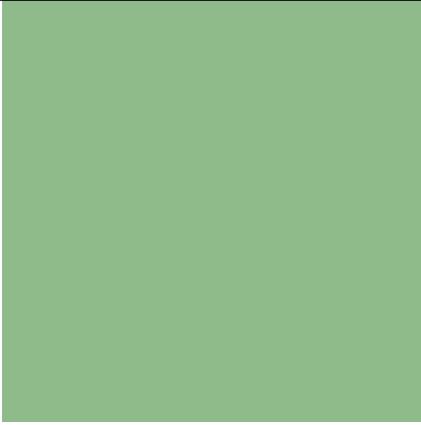


라면/통조림/조미료/장류



음료/주류/커피/분유





**감사합니다.**

## \* 과제기획심의 평가의견 반영 결과

개선요구사항	평가의견 반영결과
비건이란 타겟 계층을 위한 앱이어서 관련 사람들에게 유용하여 활용도 높을 것을 판단	감사합니다.
채식주의자라는 시장으로 조금 더 확장하면 훨씬 시장이 확대될 수 있을 것 같음	채식주의자까지 고객으로 확보하기 위하여 명확한 명칭 사용 대신 식품 함유 원재료를 이미지로 대체하였음
비건단계에 따른 음식추천 방식에 대한 개선이 필요해 보임	추후 추천 시스템 구현 시 멘토링을 통해 개선할 계획
ocr, data crawling ,Elastic search, 딥러닝 , AWS 서버등 기술습득 및 개발에 많은 시간이 소요되며 가능한 추진 일정 필요	개발 기간 단축을 위하여 OCR모델은 GCP Vision API를 사용하여 대체. 직접 OCR 모델 구현은 베타서비스 편성 이후로 계획
DB정보 획득에 필요한 명확한 구축 리스트 필요	DB 모델링을 통해 수집해야할 정보를 명확히 정리하였음
서비스 개념 및 목적에 맞는 플로우 재정립 필요	멘토링을 통해 플로우를 재정립하였고 이후 플로우를 따라 화면구성을 하였음
수집하는 데이터의 범위를 한정해서 하는 것이 좋을 것 같음	최대한 많은 데이터를 확보하였으나 중요 데이터를 우선으로 정확도 개선을 하는 중
비건에 대한 음식에 시작하여 전체 음식 건강 앱으로 발전	그렇게 하기 위해 데이터 확보 다각화를 진행 중
향후의 비건 라이선스, 비건 마크 등 기술외적인 우수한 비즈니스 방안을 잘 활용하면 성공 가능성 이 높을 것으로 사료됨	일정 수 이상의 고객을 확보한 뒤 사업 진행방향 중 하나로 고려하고 있음.
본 시스템을 바탕 향후 당뇨병환자, 건강 다이어트 관련 앱으로 지속 개발	현재 초기 목표였던 비건을 타겟으로 우선개발한 뒤 꾸준한 업데이트를 통해 개선하도록 하겠음
결과물의 기능 중 원료표기사진을 통해 비건 단계 혹은 비건 여부를 OpenAPI로 확대검토	개발 완료 후 팀 협의를 거쳐 진행하겠음

## \* 중간평가 평가의견 반영 결과

개선요구사항	평가의견 반영결과
<ul style="list-style-type: none"><li>◦ 데이터와 실제 성분의 정합성, 정확도를 높이고 고객의 신뢰도를 높이는 것이 본 프로젝트의 선결 과제이며 향후 데이터 통계에 관한 페이지 등에서도 개선 필요.</li></ul>	<p>식품 정보에 대한 <u>비건들의 의견</u>을 반영하기 위해 고객층인 <u>비건</u>을 대상으로 인터뷰를 진행하였고, 이를 통해 까다로운 식품에 대한 정보 제공을 해결할 수 있었음.</p> <p>또한 데이터 신뢰도 향상을 위해 원재료 추출 <u>NLP 모델의 정확도 향상을</u> 위해 노력함.</p>
<ul style="list-style-type: none"><li>◦ 서비스 이용자가 많아질 경우 이 <u>앱의 사용목적은 무엇이 될 것인지</u> 고민. 결국 검색 및 판단 보다는 신뢰를 쌓은 서비스에서 추천하는 제품을 구매하는 수요가 늘게 되어 이 기술의 목적이 고객 유입정도로 머무를 수 있음. <u>비건을 대상으로 프로젝트 프로토타입을</u> 진행하지만, <u>성분표를 인식하여 DB를 구축하고 이미지를 분석하여 결론을 제공하는 요소기술의 개발 및 판매에 대한 부분을 부각해도 좋을 것</u></li></ul>	<p>중간발표 이후 <u>Kakao</u> 컨퍼런스에 참여하여 한국어 OCR 모델 발표에 참석하였고, 실질적인 구현에 어려움이 있음을 알고 개발 방향을 선회하였음. <u>Kakao if</u> 2019에서 보인 OCR 모델의 경우 1년간의 개발에 몰두하였으나 모바일 환경에서는 낮은 성능을 보임. 그리하여 우리 팀은 OCR 직접 기술구현 과제 대신 <u>NLP 모델의 고도화를</u> 꾀하였음.</p> <p>정체성에 관해서, 팀원들의 토의, <u>비건</u> 인터뷰를 통해 비록 최종 목적이 모든 사람들의 식품 정보 제공 어플리케이션으로 발전하더라도, 초기 단계에는 <u>비건들을</u> 고객으로 설정하여 중요 고객 확보를 하고자 결정하였음. 이로써 <u>비건을 위한 데이터 기반, 데이터 서비스 또는 정보 제공 플랫폼</u>으로 나아갈 것이며, <u>비건을 타겟으로 하는 사업과 비건의 중개자 역할을 할 수 있을 것</u>이라 예상하였음.</p>
<ul style="list-style-type: none"><li>◦ 프로젝트명은 유통식품의 비전 등급 판별 <u>앱</u> 개발이지만, 현재까지 구현한 사항은 등급 판별을 위한 <u>개발적 요소가 없어 보임</u>. 프로젝트 일정에 비해 개발 <u>해야될</u> 부분이 많이 남아있어 프로젝트 진척관리 필요. OCR인식 개발 부분 미진하고, <u>OCR인식율이 정확도가 떨어짐에 따라 OCR대신 바코드 인식을 통한 원재료 분석이 더 효율성이 있어 보임</u>.</li></ul>	<p>등급 판별은 <u>비건의 기준이 애매하다는 문제점</u>에 봉착해있었으며, 이를 해결하기 위해 <u>비건들과의 인터뷰를</u> 진행하였음. 또한 해당 식품이 <u>비건들의 단계와 맞지 않을 경우 비슷하지만 해당 단계와 맞는 추천 서비스를 개발하였음</u>. 이는 추후 유저 데이터를 확보 시 유저의 개개인에게 맞는 추천 알고리즘으로 개선할 수 있음.</p> <p>OCR 직접 개발의 어려움에 따라 <u>NLP 모델 고도화로</u> 방향을 선회하였음. 바코드 인식 또한 좋은 방법이나 기존에 확보한 텍스트 데이터를 활용하는 것이 더 효율적이라 판단하여 이를 자연어 처리로 원재료 추출을 결정하였음.</p> <p>바코드 정보는 대한상공회의소 유통물류진흥원에서 보유중이나, 데이터 제공을 위해선 기관과 협의가 필요하기에 소형 프로젝트 팀이 확보하기엔 한계가 있다고 판단하였음</p>

<p>식품 포장재에 기재된 성분 구성표를 촬영한 이미지에서 채식주의 식품으로 적절한가의 판단에 대한 문자인식의 기술요소 부분 진척은 아직은 많이 부족해 보임.</p> <p>문자 인식이나 자연어 처리와 같은 핵심적인 기술 요소는 현 수강생들이 단기간에 적응하기에는 어려운 주제임</p>	<p>OCR 직접 개발의 어려움에 따라 Vision API에서 추출된 텍스트에서 원재료를 추출하는 NLP 모델로 대체하였음</p>
<ul style="list-style-type: none"> <li>○ 향후 제품화를 고려한다면 신뢰할 수 있는 양질의 데이터를 확보하는 것을 고려</li> <li>○ 사용자 편의성을 위해 키워드 검색뿐이라 이미지 매칭부문도 고려</li> </ul>	<p>양질의 데이터 확보를 위해 문장을 필터링하는 딥러닝 모델을 개선하였음.</p> <p>데이터의 신뢰성을 확보하기 위해 비건과 인터뷰를 진행하였음.</p>
<ul style="list-style-type: none"> <li>○ 공동구매를 통한 유통마진의 최소화 전략은 사업자 측면의 유통구조를 가져야 하므로 매우 추상적인 사업화 방안임.</li> <li>○ 비건 제품의 판단에 따른 제품과 장소에 대한 이용자 간의 공유에 대한 고려가 미흡하므로 향후 사업화 및 이용자 확대의 한계가 있을 것임.</li> <li>○ 대량의 크롤링 데이터의 구축 및 정보 추출에 대한 접근은 양호하나 사용자가 촬영하는 식품정보 데이터에 대해 동일한 정보추출 엔진을 사용하는 점은 시스템의 효율성 측면에서 다소 아쉬움.</li> </ul>	<p>공동구매와 같은 방안은 바람직하지 못함을 인지하였음. 이용자 확보와 데이터 신뢰성을 높이기 위해 비건 인터뷰를 진행하였음.</p> <p>직접 촬영을 통해 데이터를 축적하는 것은 사업 초기단계엔 매우 부적절하다고 생각하였음. 데이터는 쌓아놓되 일정량을 초과해 검증을 할 단계가 되면 효율성을 높이는 방안을 고려하겠음.</p>

