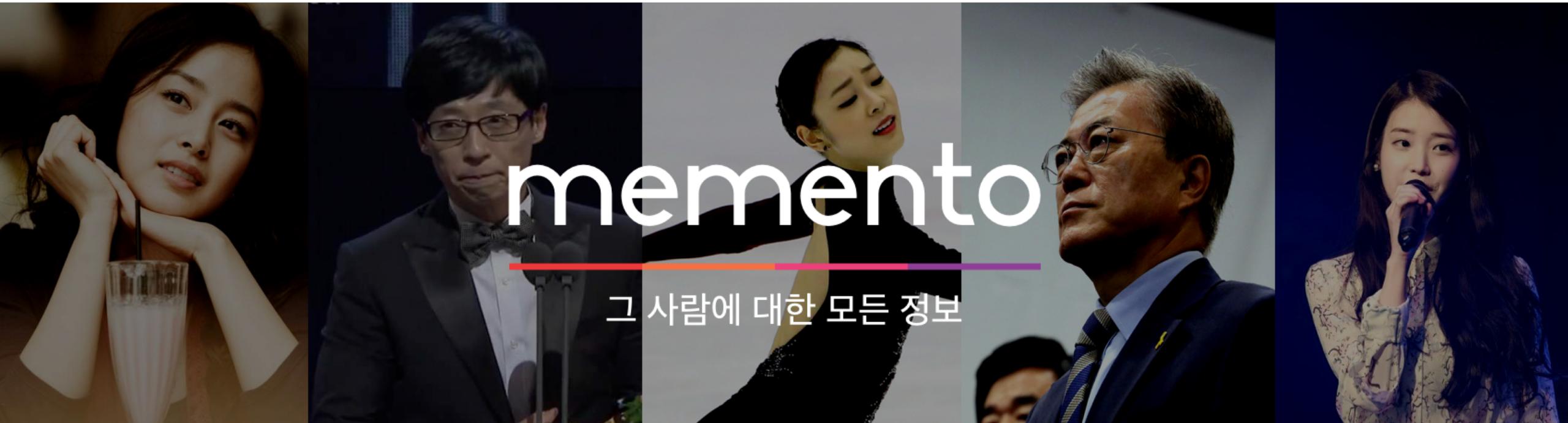




SW_Maestro 7기



memento

그 사람에 대한 모든 정보

이영수 백동환 배지운
멘토 최광선

팀 메멘토

2017년 6월 23일

목차





1) 서비스 컨셉

팬이나 미디어 소식을 알고 싶은 사람이

관심있는 유명인 “그”에 대해 알고 싶을 때

“그”가 살았던 인생을 들려 주는 서비스

얘가 이 사건 때문에 유명해졌구나

이 사건은 이렇게 진행되었네

이때 이 연기 대박이었는데, 다시 한번 봄까?

이 사람은 최근 뭐하고 지냈으려나

다른 사람도 이렇게 느꼈구나. 나만 그런 게 아니었네

맞아. 이때는 다들 이 사람 욕만 했지..



1) 서비스 컨셉

웹 상에 넘쳐나는 단편적인 정보를 모두 모아

그 사람의 **인생 이야기**를
그 때의 **분위기**를 반영하여

언제 어디서나 볼 수 있는
매거진으로 만들어주자

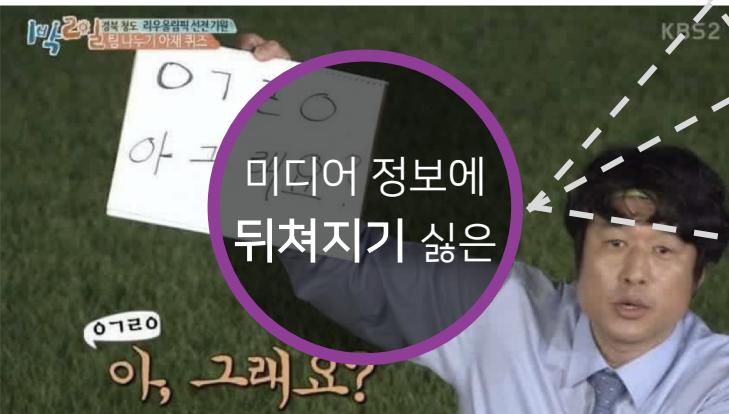
그것도 **기술**을 이용해서 **자동**으로!





2) 사용자 분석

페르소나



기능 및 사용자 경험

인물 매거진

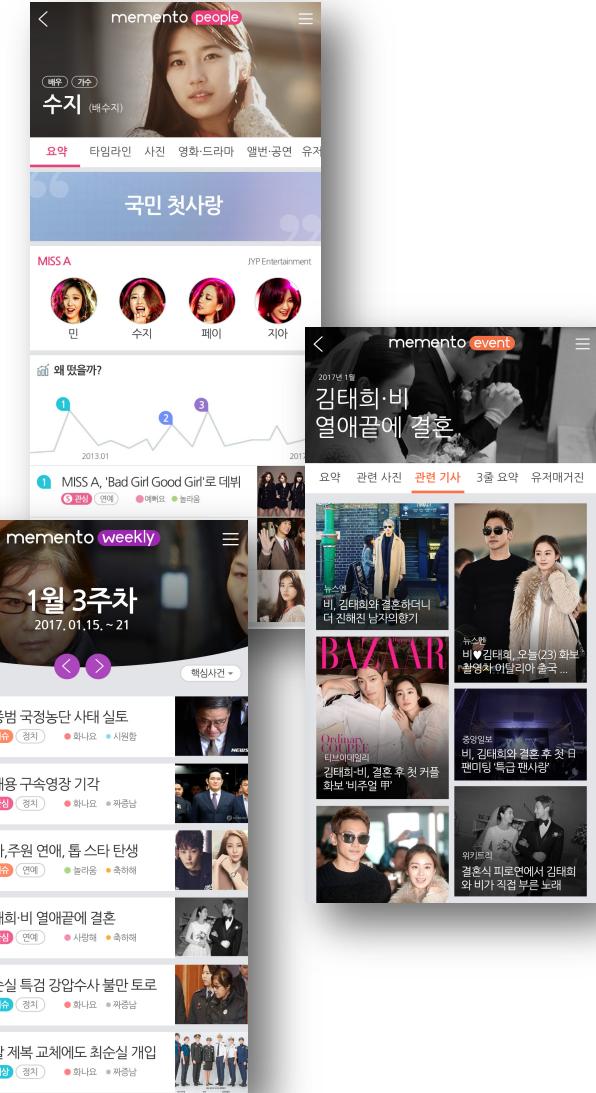
- 인물의 삶을 훑어 보는 경험
- 인물에 대한 팬심과 관심의 표현

사건 매거진

- 한 눈에 파악하는 사건
- 현장의 분위기 파악 및 공감

위클리 메멘토 & 홈

- 미디어 배경지식 제공
- 컨텐츠 추천





3) 기능 및 사용자 경험

인물 매거진

한마디로

왜 떴을까?

재미로 보는 능력치

김태희가 듣는 수업은 언제나 학생들로 꽉 차있었다
- 당시 서울대 재학생

김태희 대표사진 선택

요약 타임라인 사진 영화·드라마 유저매거진

2017 김태희-비 열애 끝에 결혼

2015 김태희, SBS 드라마 "용팔이" 열연

2012 김태희 타임라인

2009 KBS 드라마 "아이리스" 출연

2005 서울대학교 의류학과 졸업

1980 3월 29일 출생

액션 1.6 감정연기 4.0
배역 소화도 4.1 안정감 7.9
외모 5★
안정감 4★
배역소화도 4★
감정연기 4★
액션 4★

인물의 삶을 궤뚫어 보는 경험

- 요즘 자주 나오던데, 어떤 이유로 뜨게 된 거지?
- 얘가 어떤 사건들을 겪었더라?
- 대중들은 이 사람에 대해 어떻게 생각하고 있지?

인물에 대한 팬심과 관심의 표현

- 좋아하는 연예인의 장점을 알리자
더 예쁜 사진을 수집하거나 공유하자

3) 기능 및 사용자 경험

사건 매거진

The image shows two screenshots of the memento event app. Both screenshots feature a black and white photo of a couple in formal attire, with the text "2017년 1월 김태희·비 열애끝에 결혼" (January 2017, Kim Tae-hee and Bi end their romance with marriage).

Top Screenshot (Left):

- Header: memento event
- Date: 2017년 1월
- Title: 김태희·비 열애끝에 결혼
- Section: **요약** (Summary)
- Links: 관련 사진 (Related photos), 관련 기사 (Related articles), 3줄 요약 (3-line summary), 유저매거진 (User magazine).
- Section: **관련 인물** (Related people)
 - Kim Tae-hee (배우, Actress)
 - Bi (정지훈, 가수, 배우, Singer, Actor)
- Section: **3줄 요약** (3-line summary)
 - 1 배우 김태희와 가수
 - 2 비는 자신의 인스타
 - 3 결혼식은 소박하게
- Section: **이슈 등급** (Issue Level)
 - 이슈 등급
 - "세간의 관심"
- Section: **연관 키워드** (Related keywords)
 - 화보, 싸이, 5년, 비공개, 신혼여행, 편지, 최고의 선물, 연애, 결혼식

Bottom Screenshot (Right):

- Header: memento event
- Date: 2017년 1월
- Title: 김태희·비 열애끝에 결혼
- Section: **관련 기사 모아주기** (Related article collection)
- Links: 표지, 관련 사진, 관련 기사, 3줄 요약, 유저매거진.
- Section: **관련 인물** (Related people)
 - 김태희 (배우)
 - 비 (정지훈, 가수, 배우)
- Section: **3줄 요약** (3-line summary)
 - 1 배우 김태희와 가수
 - 2 비는 자신의 인스타
 - 3 결혼식은 소박하게
- Section: **사람들의 반응** (People's reactions)
 - 사랑해, 멋져요, 대단함, 무서움, 슬퍼요, 화가남, 부러움, 나빠요, 행복해, 재밌음, 놀라움, 공감됨.
- Section: **이슈 등급** (Issue Level)
 - 이슈 등급
 - "세간의 관심"
- Section: **연관 키워드** (Related keywords)
 - 화보, 싸이, 5년, 비공개, 신혼여행, 편지, 최고의 선물, 연애, 결혼식

한 눈에 파악하는 사건

- 볼만한 기사들이 뭐가 있을까?
- 사건을 간단하게 요약하자면?
- 이 사건과 관련된 게 무엇이 있고 뭐가 중요하지?

현장의 분위기 파악 및 공감

- 얼마나 이슈 되었던 사건이지?
- 사람들의 반응은 어땠지?

서비스 소개

3) 기능 및 사용자 경험

위클리 메멘토 & 홈

주간 핵심 사건

사건 탐색

볼만한 사건들

3년전 오늘

3년전 오늘

대화 시 필요한 미디어 지식 제공

- 꼭 알아야 하는 사건이 뭐가 있을까?
- 이 사람은 내가 몰랐는데, 유명한가 보네
- 요즘 친구들이 얘기하던 사건이 이거구나

인물 및 사건 중심의 컨텐츠 추천

- 심심한데 볼만한 건 없나?
- 아 이런 사건도 있었구나, 잘 몰랐던 건데

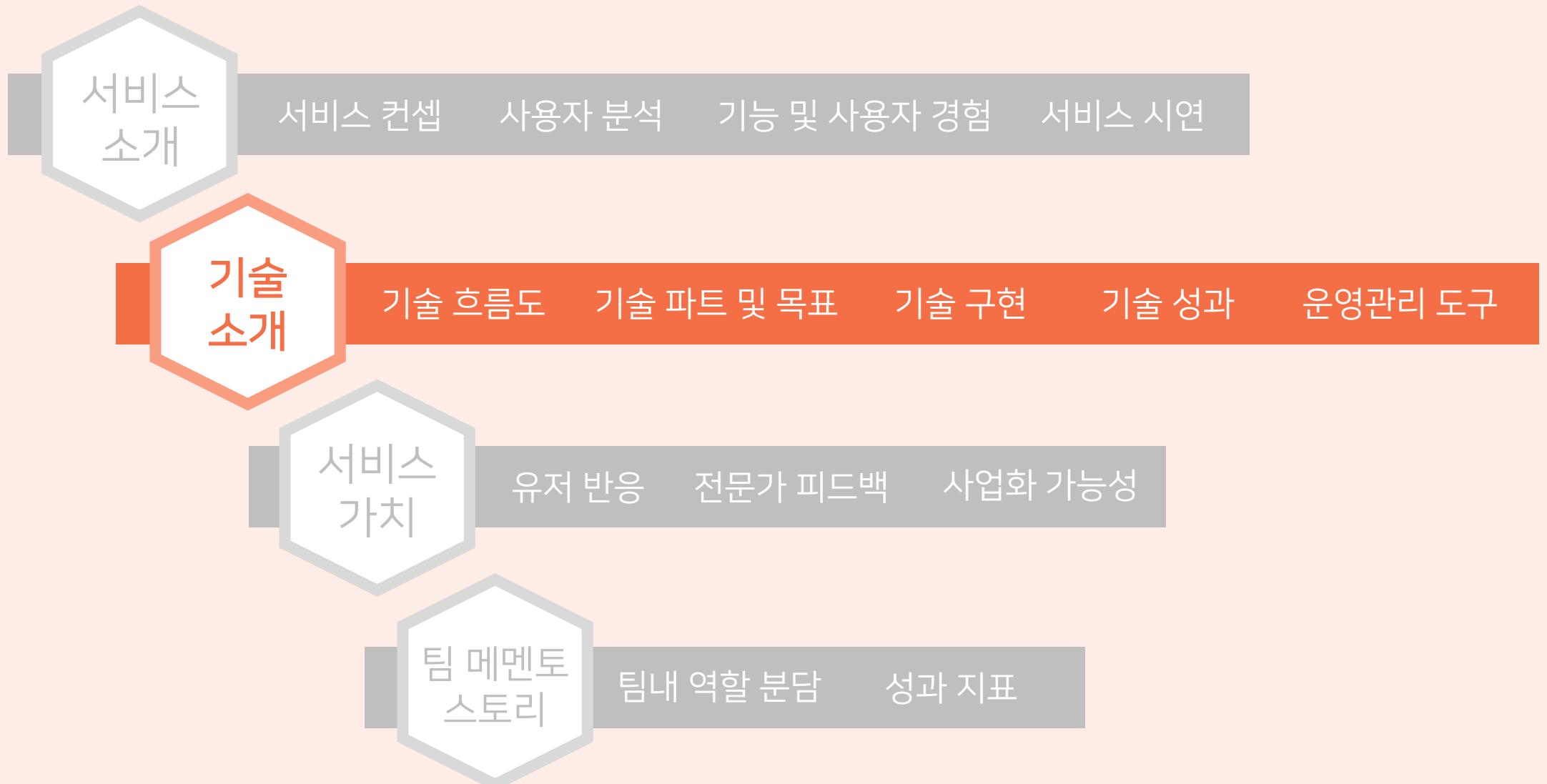


4) 서비스 시연

<https://memento.live>

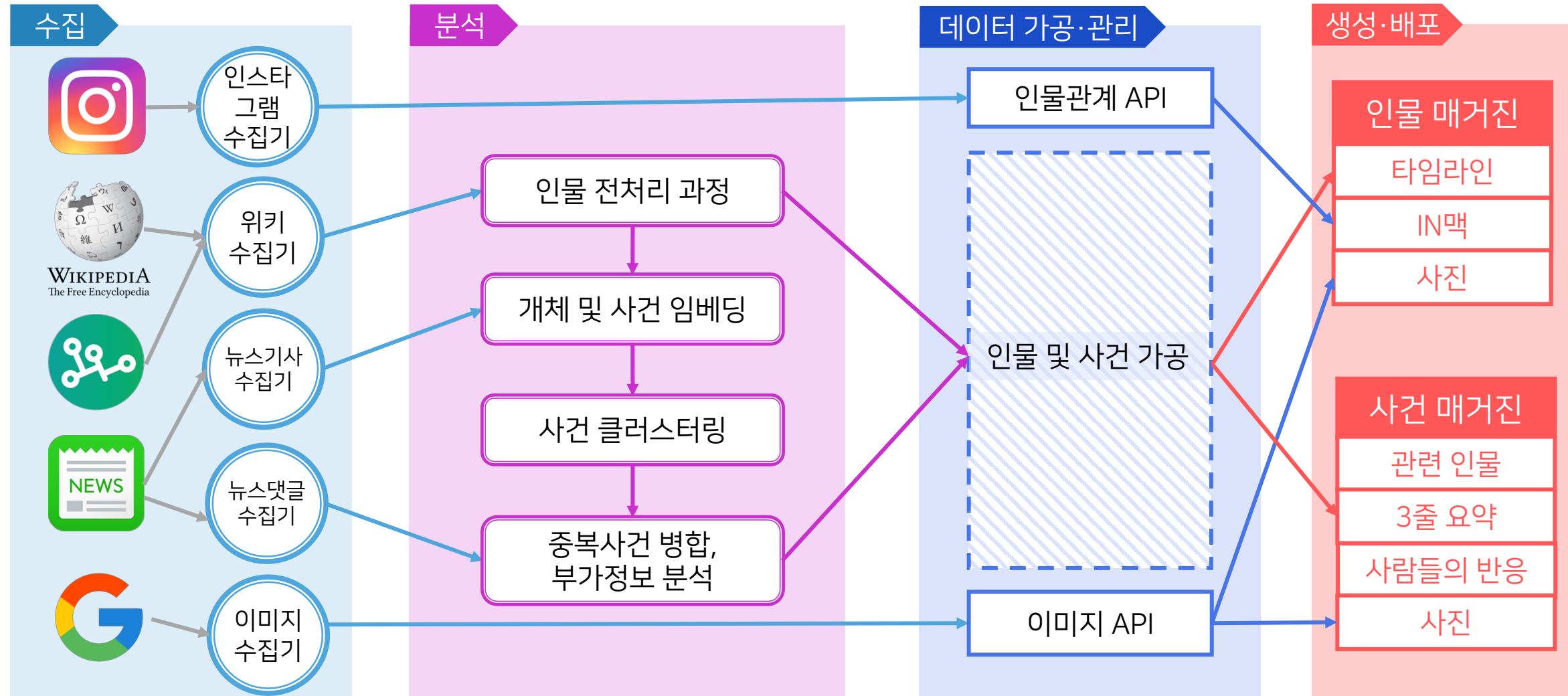
약 350명의 유명인을 대상으로 베타서비스 중
지속적으로 인물 추가 중







1) 기술 흐름도



2) 기술 파트 및 목표

수집

- ① 하루에 수백 명 이상의 유명인 중심 정보 수집
- ② 새로운 채널이 생겼을 때 기존 시스템 변경 없이 빠르게 대응할 수 있는 구조

분석

- ① 90% 이상의 개체 명 인식 & 동명 이인 분리
- ② 수천, 수 만개의 사건을 단시간 내에 처리

데이터 가공·관리

- ① 30년 이상의 인물·사건 정보를 가공·관리
- ② 확장성 있는 설계를 통한 신규 비정형 데이터 수용이 용이한 구조

생성·배포

- ① 서비스와 수집·분석 작업의 분리를 통한 부하 최소화
- ② 명료한 프로세스 설계 및 새벽시간 동안 생성·배포 완료



3) 기술 구현 | 수집

수집 > 분석 > 가공·관리 > 생성·배포

1) 하루에 수백 명 이상의 유명인 중심 정보 수집

사건, 사진, 인적정보, 인물 관계 등

2) 새로운 채널이 생겼을 때 기존 시스템 변경 없이 빠르게 대응할 수 있는 구조

서비스 기능 추가, 타 서비스 제휴 등

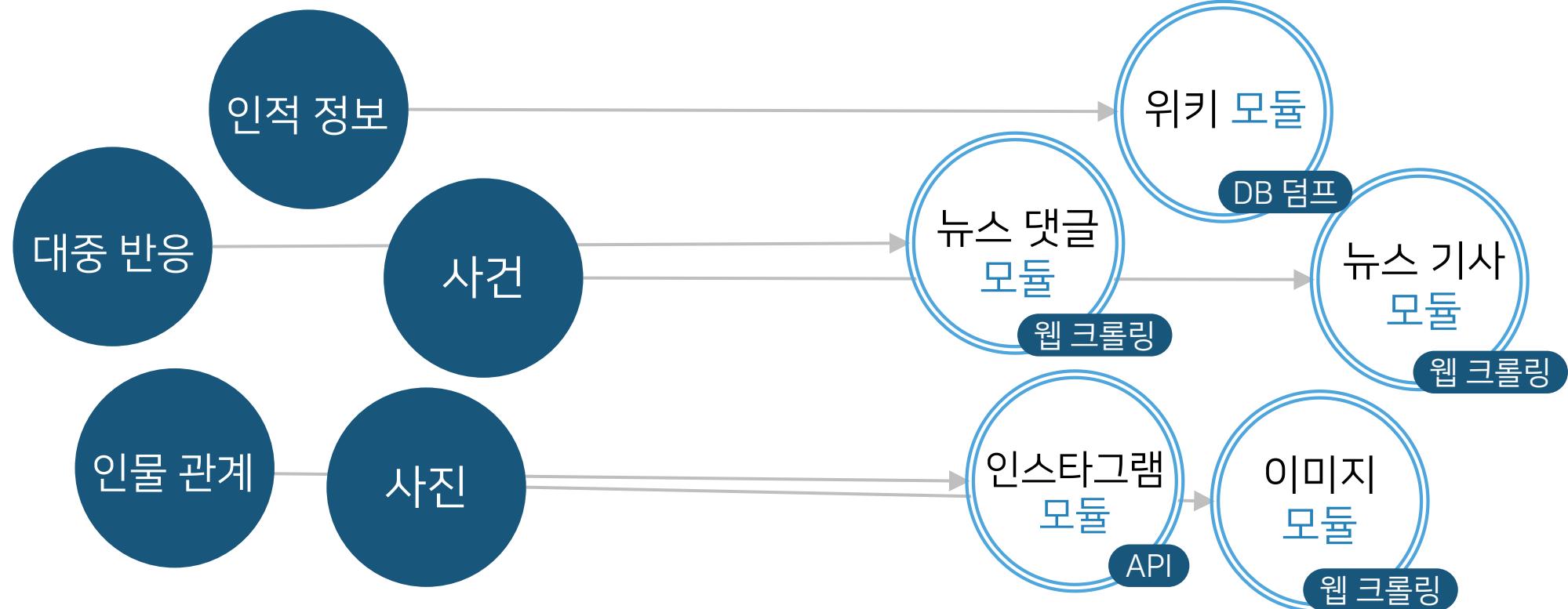
3) 기술 구현 | 수집

유명인 중심의 정보 수집

사건, 사진, 인적 정보, 인물 관계 등의 데이터

데이터 채널 별 수집기 분리

새로운 채널이 생겼을 때 확장 용이
용도에 맞는 내부 방식으로 빠르게 수집 가능



3) 기술 구현 | 수집

Part 1

가상환경으로 모듈 패키징



내부 구조



컨테이너 기술 Docker 활용

관리 프로그램 통신 및
수집 대상 탐색

수집
프로세스

수집
프로세스

수집
프로세스

수집
프로세스

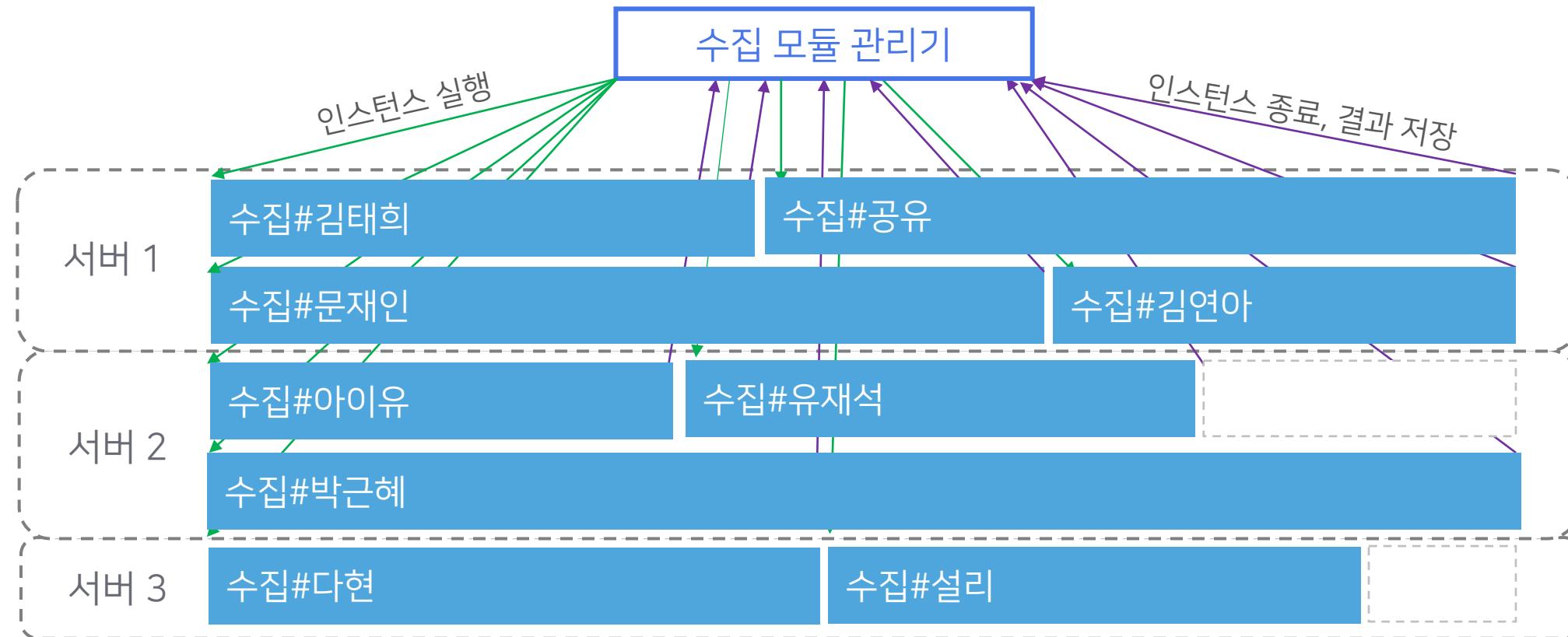
수집 결과 저장



3) 기술 구현 | 수집

다수 서버에서 동시에 수집 진행

수집속도 향상 및 순간대응 가능 (대선 등 데이터가 많이 생성되는 특수한 날)



관리기

: 모듈 스케줄링 및 오류관리

■

: 뉴스기사 수집기 Docker 인스턴스



3) 기술 구현 | 분석

수집 > 분석 > 가공·관리 > 생성·배포

1) 90% 이상의 개체 명 인식 & 동명 이인 분리

사람 이름, 추후 제품으로 확장

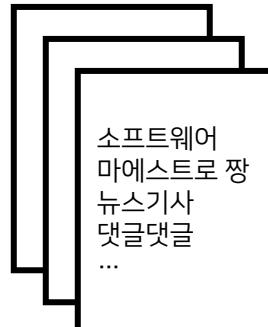
2) 수천, 수 만개의 사건을 단시간 내에 처리

하루동안 발생하는 모든 사건을 새벽 시간 동안 처리



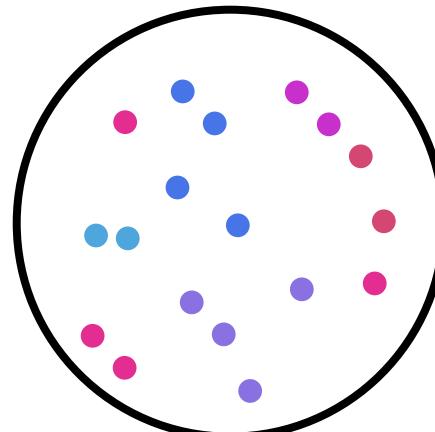
3) 기술 구현 | 분석

분석 프로세스



데이터의 특징을 잘 살려서
다루기 쉬운 숫자로 변환

Embedding

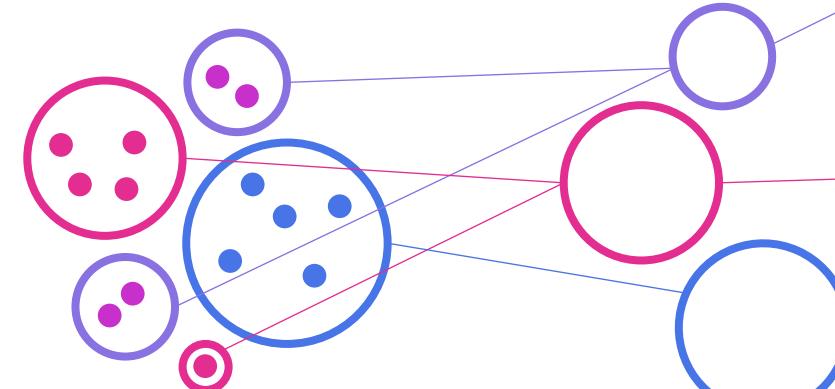


전처리 과정

형태소 분석
메타 데이터 분석
개체명 인식
동명 이인 분리

데이터의 포함된 인물과 사건
기반으로 유사한 데이터 모음

Clustering



중복 사건 제거
이벤트 등록

이벤트

이벤트

이벤트

3) 기술 구현 | 분석

전처리 과정

김태희 작가는 29일 오후 서울 마포구 상암동 MBC 상암 신사옥에서 진행된 '2014 MBC 방송연예대상'에 참석했다. 이날 '황금어장-라디오스타'로 작가상을 받은 **김태희** 작가는 "영광스러운 상을 받아 정말 기쁘다"고 말문을 열었다.

이어 그는 "그만둔 지 꽤 됐지만 '무한도전' PD님과 멤버들에게도 감사하다"며 "특히 고구마를 가지고 가다 넘어진 **노홍철**과 **정형돈**의 고백을 거절했을 때가 생각나 미안한 마음을 갖고 있다"고 말해 시상식장을 웃음바다로 만들었다.

한편 **김태희** 작가 수상 소감에서 언급된 '무한도전'의 메인 MC **유재석**은 같은 날 MBC 방송연예대상을 받았다.

개체명 인식

외래어가 혼재된 개체명도 인식 가능

여러 형태소로 분리되는 이름들도 묶음

한국어 개체명 인식을 위한 자질 추가

여러 오픈 소스도 양상블

90% 이상의 정확도



3) 기술 구현 | 분석

전처리 과정

김태희 작가는 29일 오후 서울 마포구 상암동 MBC 상암 신사옥에서 진행된 '2014 MBC 방송연예대상'에 참석했다. 이날 '황금어장-라디오스타'로 작가상을 받은 **김태희** 작가는 "영광스러운 상을 받아 정말 기쁘다"고 말문을 열었다.

이어 그는 "그만둔 지 꽤 됐지만 '무한도전' PD님과 멤버들에게도 감사하다"며 "특히 고구마를 가지고 가다 넘어진 노홍철과 정형돈의 고백을 거절했을 때가 생각나 미안한 마음을 갖고 있다"고 말해 시상식장을 웃음바다로 만들었다.

한편 **김태희** 작가 수상 소감에서 언급된 '무한도전'의 메인 MC 유재석은 같은 날 MBC 방송연예대상을 받았다.

작가, 서울, 마포구, 상암동, 황금어장, 라디오스타, 작가상, 수상, 소감, 무한도전

작가 김태희 연관 키워드

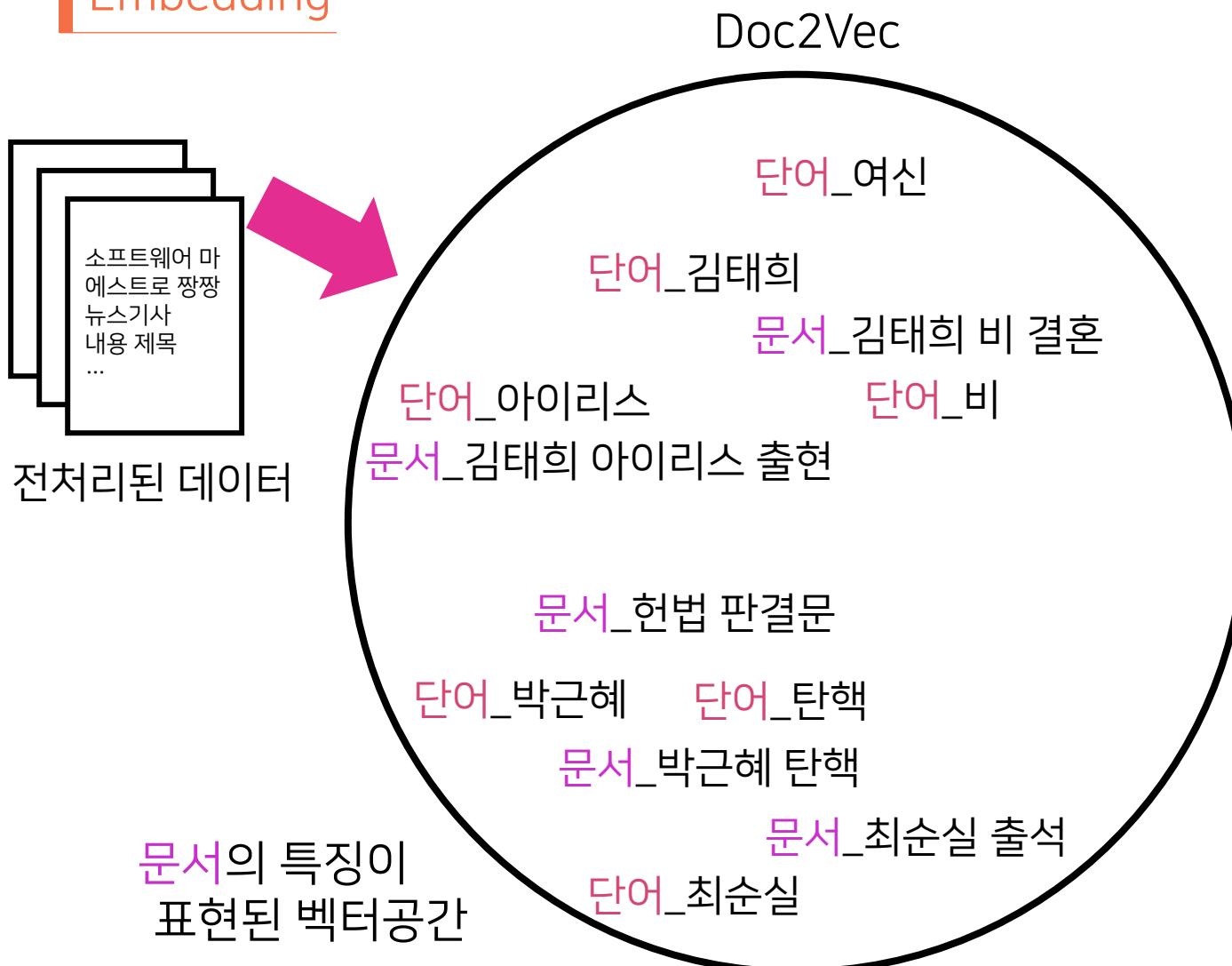
키워드	가중치
김태희	1500
드라마	599
미녀	550
작품	514
한국	513
서울대	492
데뷔	450
년대	450
이후	400
미모	400

배우 김태희 연관 키워드

같은 “**김태희**” 인가?

3) 기술 구현 | 분석

Embedding



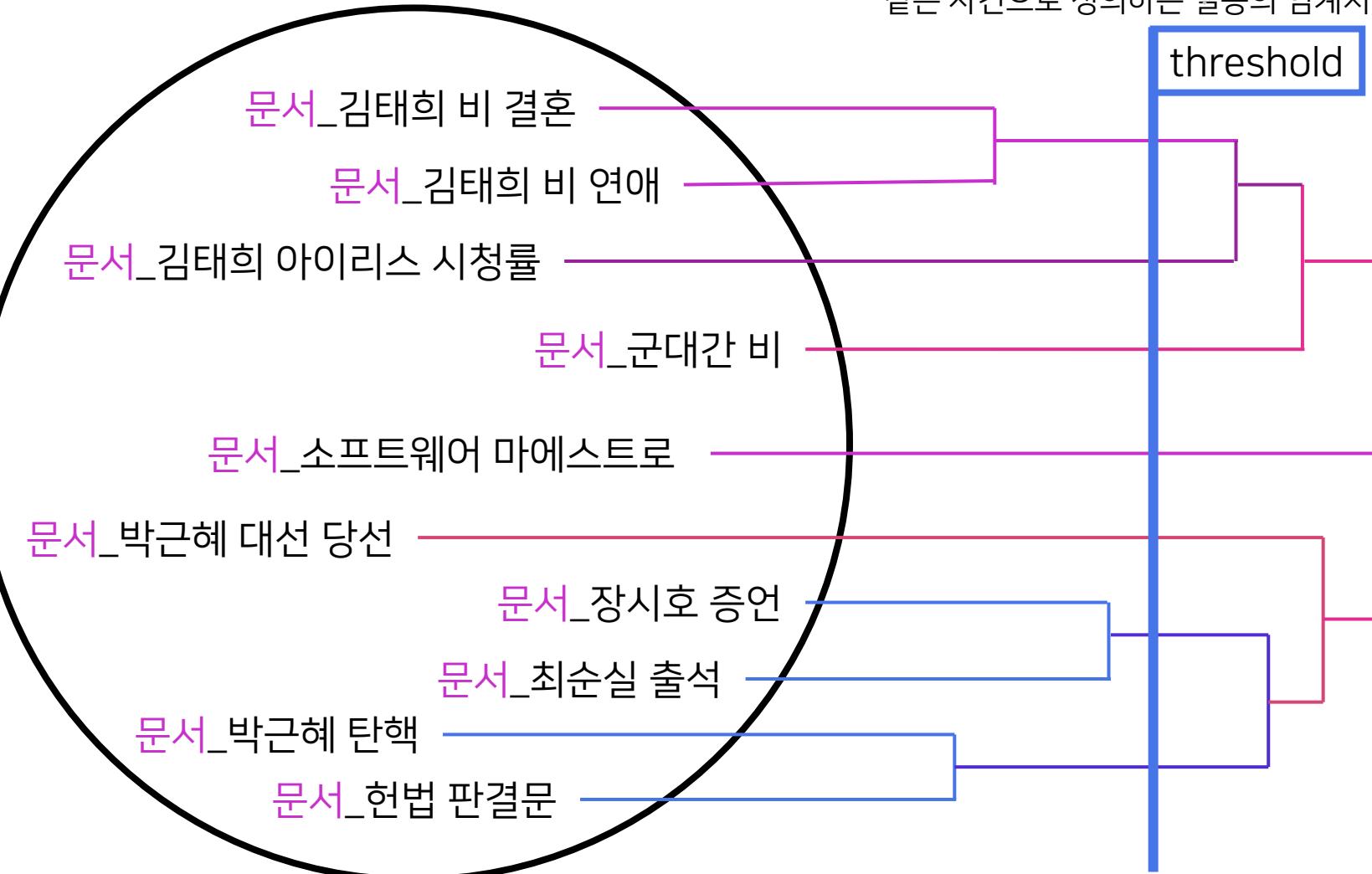
연관된 문서와 단어, 관계 파악

수학적 연산이 가능한 벡터 형태

- 날짜 단위로 세분화
분리할 수 있는 사건을 미리 분리하여 정확도 향상
- 형태소 분리 및 가중치
특정 형태소에 가중치를 부여해 정확도 향상
- 수집된 데이터의 특징을 살려 분석
기사의 메타데이터를 통해 rule-base로 미리 분리

3) 기술 구현 | 분석

Clustering



여러 문서를 모아 사건을 재구성

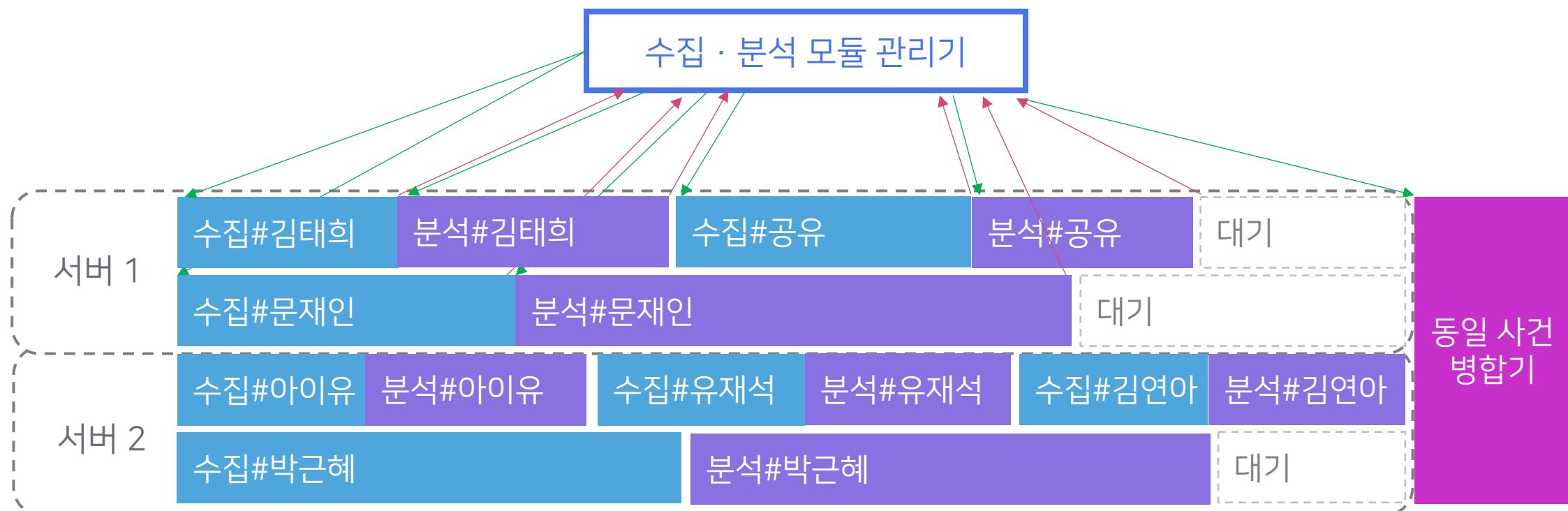
중복되거나 비슷한 문서를 한데 모아 하나의 사건으로 재구성하고 사건 사이의 관계를 파악

- 계속 발전해 나가는 구조
Doc2vec, TF-IDF와 교차 검증하여 threshold를 재조정, 최적의 값을 찾아내고 반영
- 인물마다 서로 다른 threshold
미리 구축된 인물 정보에서 인물을 구별할 만한 정보가 많이 있느냐에 따라서 엄격하게/느슨하게 분리

3) 기술 구현 | 분석

다수 서버에서 동시에 분석 진행

분석속도 향상 및 순간대응 가능 (대선 등 데이터가 많이 생성되는 특수한 날)





3) 기술 구현 | 데이터 가공·관리

수집 → 분석 → 가공·관리 → 생성·배포

1) 30년 이상의 인물·사건 정보를 가공·관리

수집·분석 과정을 거친 정보

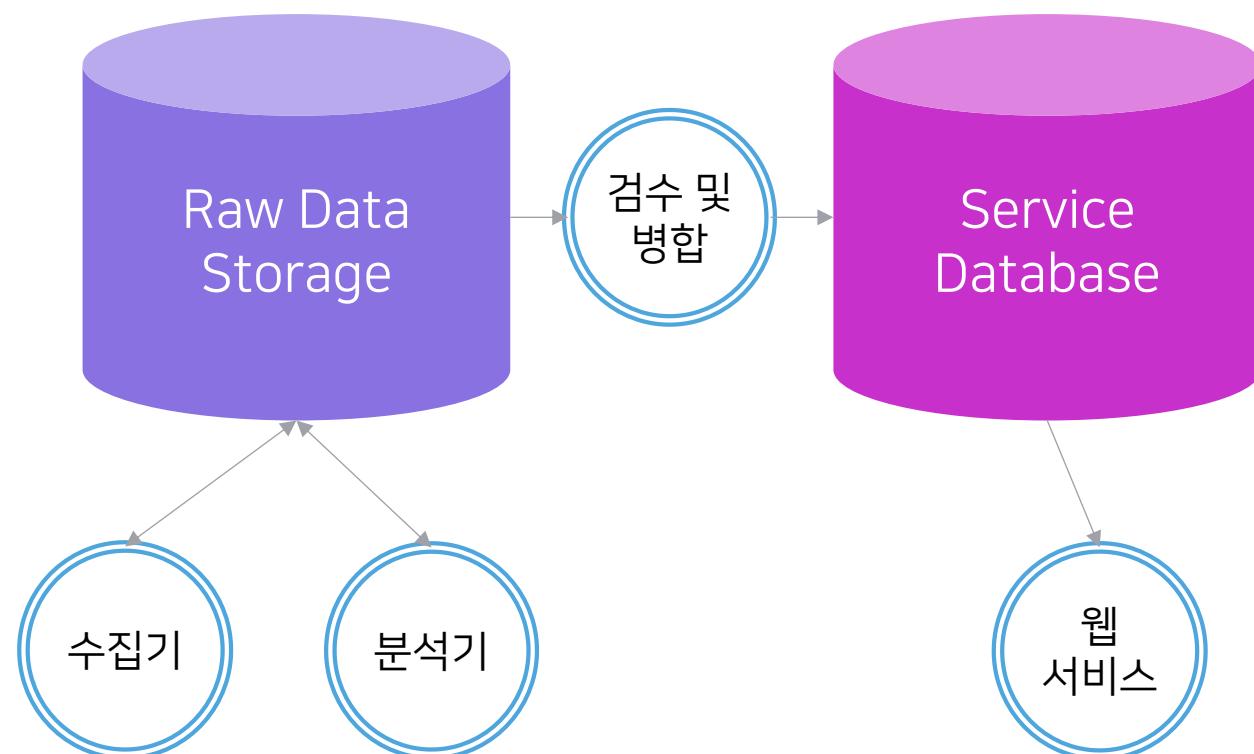
2) 확장성 있는 설계를 통한 신규 비정형 데이터 수용이 용이한 구조

비정형 + 관계형 융합

(인물-인물 관계, 인물-사건 관계 등)

3) 기술 구현 | 데이터 가공·관리

수집 및 분석 작업에서 사용되는 원본 데이터 스토리지와
실제 서비스에서 사용되는 데이터베이스 분리





3) 기술 구현 | 데이터 가공·관리



인물 사진, 뉴스 기사, 댓글, SNS 등의 원본 데이터

- Document Search Engine
비정형 데이터 위주, 관계 쿼리 최소화
- 스토리지 분할을 통한 수십 년간의 대량 데이터 효율적 저장 가능



“인물”과 “사건”에 대한 정리된 정보 (관계형)

- 서비스 확장이 가능한 데이터 모델 설계
인물 뿐만 아니라 기업/제품/미디어로 확장 가능
- 통합 API를 통한 데이터 가공·관리 + 문서화를 통한 커뮤니케이션 비용 감소



3) 기술 구현 | 생성·배포

수집 > 분석 > 가공·관리 > **생성·배포**

1) 서비스와 수집·분석 작업의 분리를 통한 **부하 최소화**

사용자가 폭증해도 서비스 정상 작동

2) **명료한 프로세스** 설계 및 새벽시간 동안 생성·배포 완료

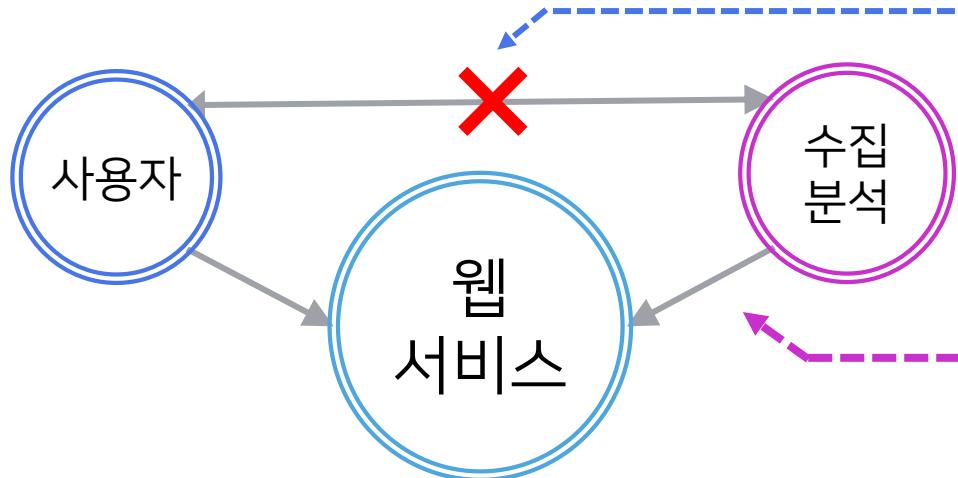
사용자 접속과 수집·분석 사이의 프로세스

매일 신규 인물/사건 매거진 발행



3) 기술 구현 | 생성·배포

정적 웹 기술 도입 이유



사용자 트래픽과 수집 · 분석계층은
비 동기적인 시점을 가져야함

일정 주기 마다 매거진을 만드는 구조

흩어진 정보를 모아 사건과 인물로
재구성해서 보여주는 일종의 **잡지**



서비스 운용 비용 ▼
캐시를 통한 접근속도 ▲



3) 기술 구현 | 생성·배포

정적 웹을 이용한 생성·배포 기술의 목표

- ① 생성 및 배포 속도가 빨라야 한다
- ② Fault Tolerant 해야 한다
- ③ 유지보수가 용이해야 한다
- ④ 프로세스가 명료해야 한다

코어와 서비스코드 분리



범용 라이브러리로 제작 및
오픈소스화

:직지(Jikji)

② Fault Tolerant

: 중간에 문제가 발생해도, 퍼블리시 작업 전체가 중단 되면 안된다



3) 기술 구현 | 생성·배포

생성 및 배포 작업 흐름도



4) 기술 성과

수집

- ① 유명인 300여명에 대한 국내에서 발행되는 모든 뉴스기사, 댓글, SNS 매일 수집 가능
- ② 모듈화를 통해 시스템 변경 없이 신규 채널 추가 가능

분석

- ① 구축한 인물 사전을 바탕으로 개체 명 인식 & 동명 이인 분리 가능
- ② 매일 발생하는 사건을 2시간 이내에 분석 가능

데이터 가공·관리

- ① 80GB 이상의 원본 데이터와 10만개의 가공된 사건 정보 관리 중
- ② 원본 스토리지 및 서비스 데이터베이스 분리를 통한 확장성 확보

생성·배포

- ① 정적 웹을 이용하여 사용자 트래픽 완벽 대응
- ② 프레임워크 설계를 통한 명료한 프로세스 설계 및 초당 100페이지 이상의 생성 속도 확보



5) 운영관리 도구

<https://monitor.memento.live>

자동화와 가상화를 바탕으로 사람의 큰 개입 없이 운영 가능
운영관리 도구를 통해 지속적 서비스 품질 향상 가능



1) 유저 반응

베타서비스 공개 후 유저 반응

메멘토 서비스가 만족스러운 이유

관련된 컨텐츠를 모아 정리해서 보여줘서 좋다



내가 좋아하는 사람의 정보를 잘 볼 수 있어서 좋다



새롭게 일어나는 수많은 사건을 정리해서 보여줘서 좋다



메멘토 서비스 사용 이유

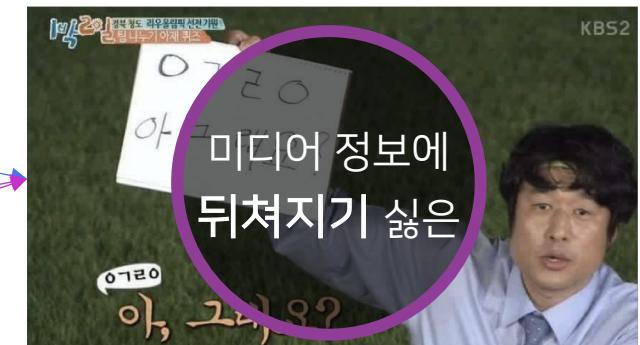
내가 좋아하는 인물을 검색하기 위해서



잘 모르는 인물을 알아보기 위해서



최근에 일어났던 중요한 사건들에 대해 알아보기 위해서

2) 전문가 피드백

초청한 전문가

- 소셜 마케팅 분석전문 스타트업 CEO
- 미디어 분석 전문가
- 전문 서비스 기획자
- 빅데이터 전문가
- 기계학습 전문 멘토

전문가 평가

“

기술을 바탕으로 유명인에 관한 정보를 정리해서 제공한다는 점

“새로운 시도”

- 김태완, Oracle

“

기업이나 정치인/공인들이 매우 좋아할 것

- 신석원, 루디엔 대표



3) 사업화 가능성

스타 중심의 마케팅

- 특정 스타 중심의 마케팅이 중요한 서비스 회사와 협력
- 정치인 및 공인 대상으로도 활용 가능
- 메뉴 혹은 위젯 형태로 팬들이 모여 있는 다양한 곳에 적용 가능



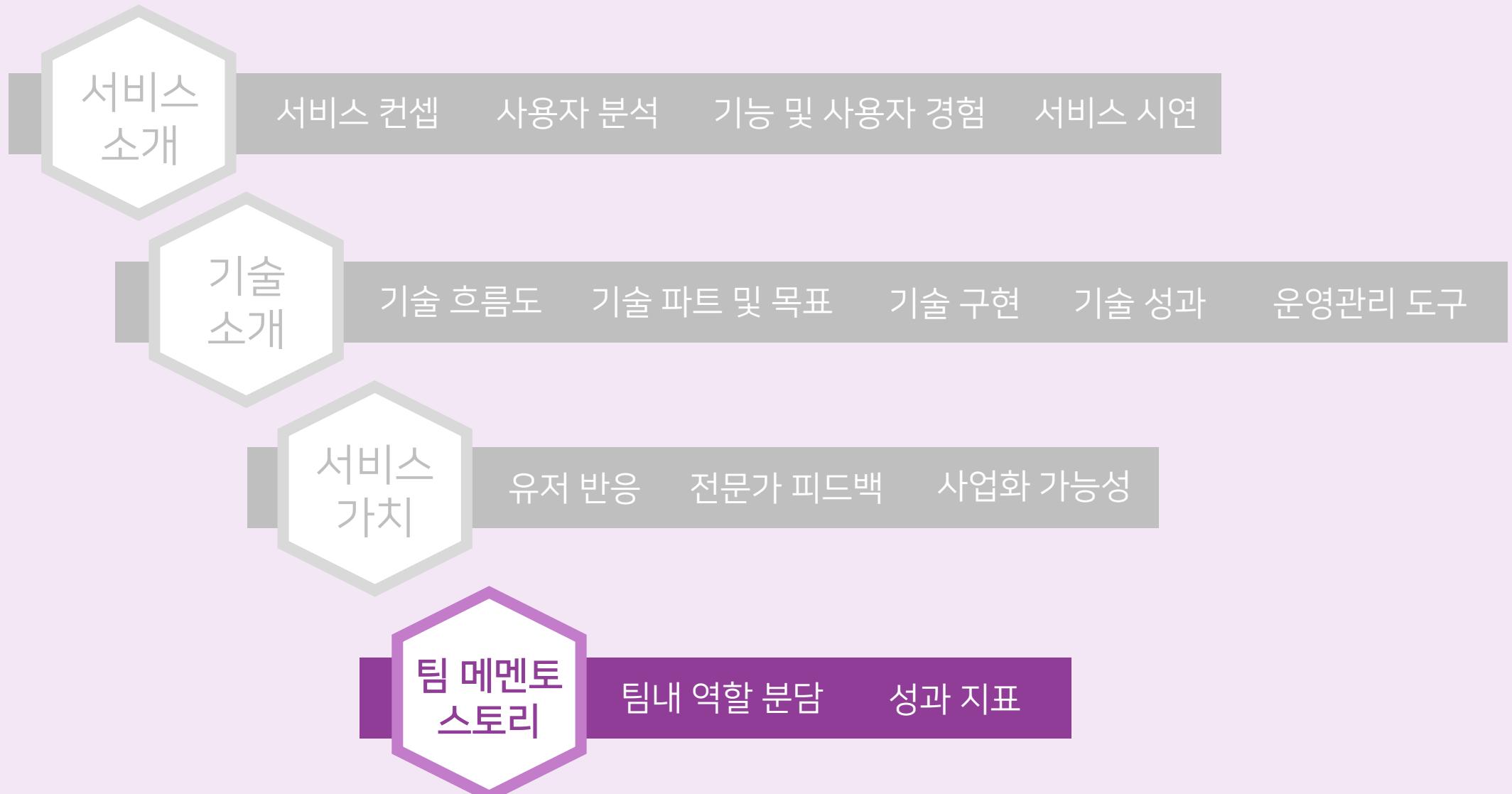
인물에서 제품으로 확장

- 제품이나 영화, 드라마 등의 미디어로도 확장 가능
- 제품의 정보를 제공하는 사이트(ex 쇼핑몰)에 위젯 제공 가능



현재 메멘토 서비스

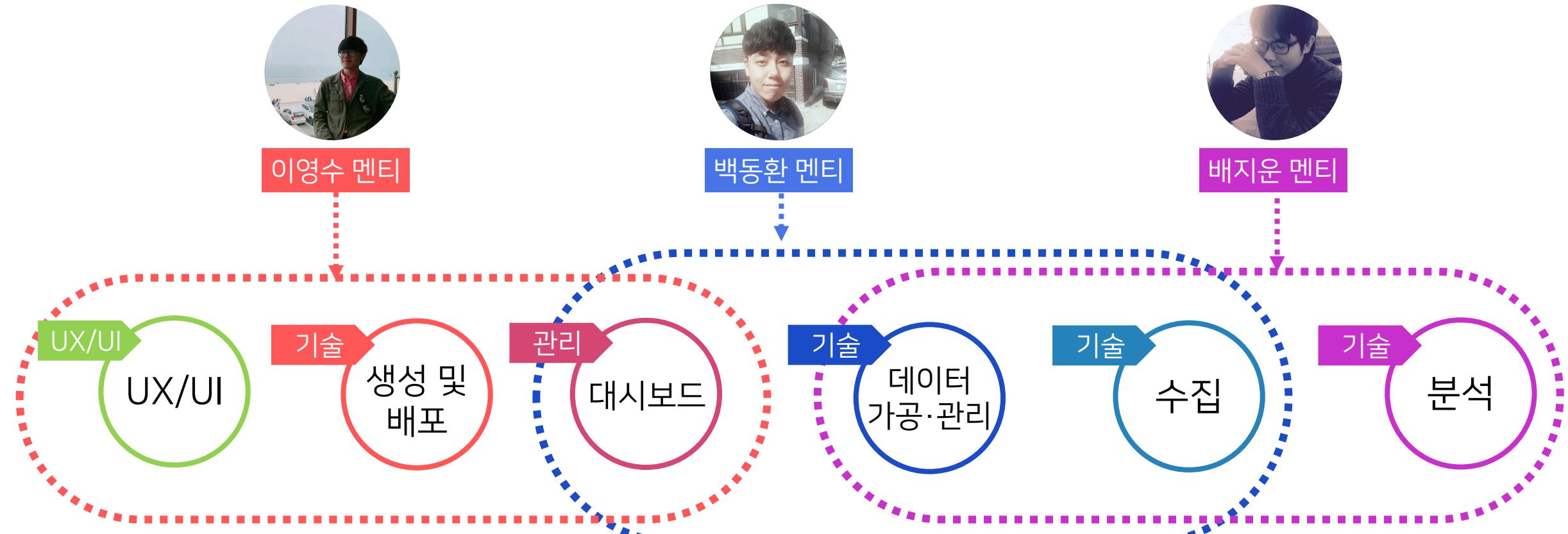
시범 운영 ➔ 기술 연구, 사용자 피드백 확보





1) 팀내 역할 분담

기존 자신의 전문 분야 역량 강화 + 공동 작업을 통한 기술 스펙트럼 확장





1) 팀내 역할 분담

기획/설계 과정은 공동으로 체계적 절차와 함께 진행





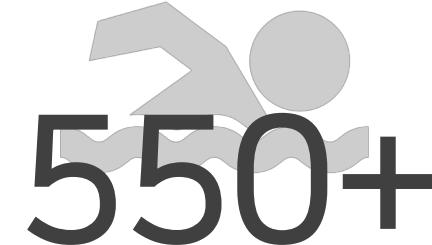
2) 성과 지표



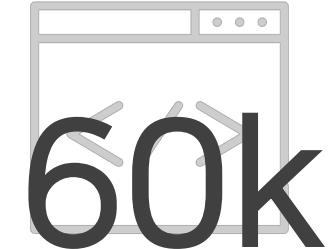
총 멘토링 횟수



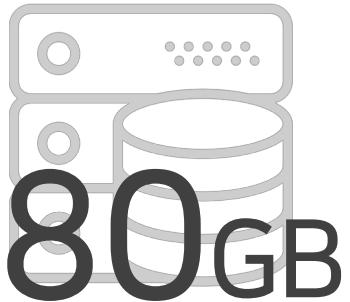
총 멘토링 시간



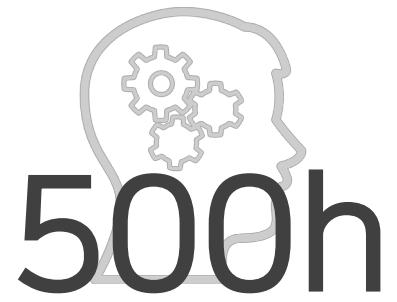
커밋 수



코드 라인 수



수집한 데이터 수



기계가 스스로
학습한 시간



읽은 논문 수



초빙한 외부
멘토 / 전문가

감사합니다



메멘토

그 사람에 대한 모든 정보