

빅데이터 분산을 통한 소셜 미디어 분석 서비스

TRENDFLOW

부울경 특화 2반 E205

박상민 권혁근 김보경 김수민 김진호 문석환

목차

01 기획 배경

02 UCC

03 주요 서비스

04 특징점

05 시연

06 기술 스택

07 기대 효과

08 팀원 소개

09 Q&A



김수민 (28, 여)
삼성전자 마케터

일하는 데에
세 가지 애로사항이
있어요.



김수민 (28, 여)
삼성전자 마케터

첫 번째,



막대한 양의 정보

전체적인 동향 파악에
시간이 많이 소요됨



김수민 (28, 여)
삼성전자 마케터

두 번째,



부족한 일관성

개인의 판단에 의해
긍, 부정 평가가 결정됨



김수민 (28, 여)
삼성전자 마케터

세 번째,



실시간 파악 애로

진행중인 전략에 대한
고객의 반응 측정 어려움

빅데이터 분산을 통한 소셜 미디어 분석 서비스

TRENDFLOW

TRENDFLOW

소셜 분석

유튜브 분석

비교 분석

키워드를 입력하세요

오늘의 HOT 키워드



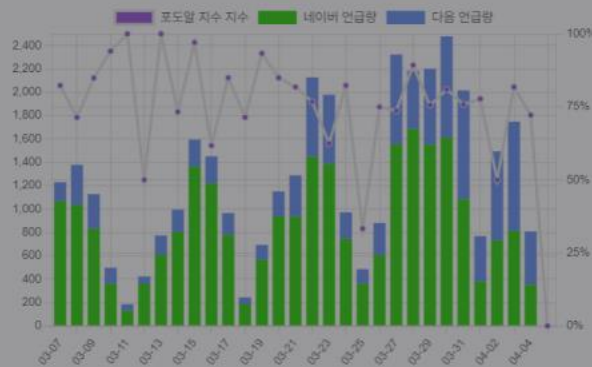
1	기업	-
2	시장	-
3	AI	-
4	미국	-
5	대표	-
6	중국	-
7	은행	-
8	기술	-

이번 주 HOT 키워드



1	시장	-
2	기업	-
3	AI	-
4	미국	-
5	전자	-
6	사업	-
7	지난해	-
8	중국	-

기업 일간분석



포도알 지수

전날에 비해
50%p 증가

언급량

775건

연관 키워드

1	기업	-
2	AI	-
3	투자	-
4	미국	▲ 2
5	시장	▼ 1
6	사업	▼ 1
7	반도체	▲ 2
8	기술	-

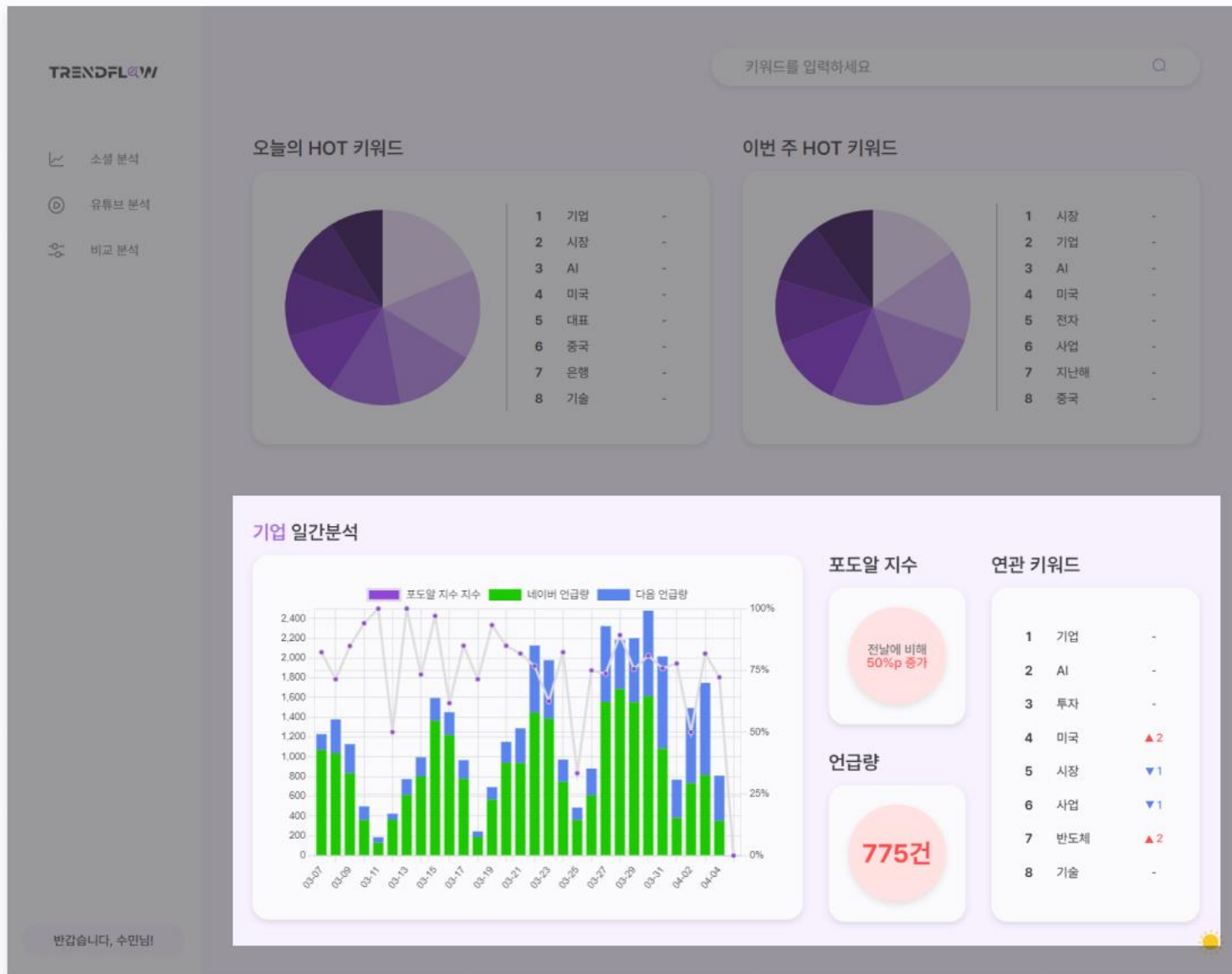
반갑습니다, 수민님!



Main

HOT 키워드

키워드의 언급량 비율과 순위를
일간, 주간 기준으로 살펴볼 수 있다.



Main

키워드 일간 분석

언급량 및 긍정 지수의 변화량을 시각화한 자료로 확인할 수 있다.

등록된 키워드와 함께 검색된 연관 키워드들을 조회할 수 있다.

TRENDFL@W

소셜 분석

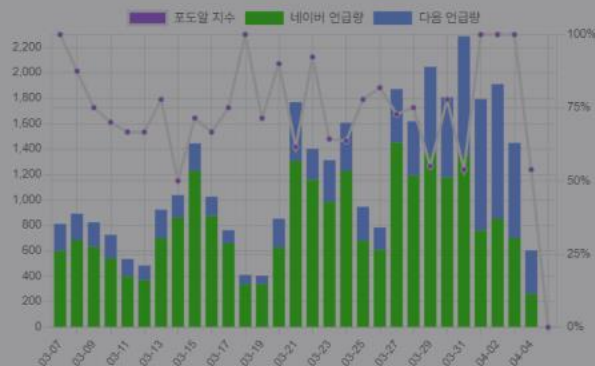
유튜브 분석

비교 분석

미국 소셜 분석 리포트 ☆

미국

2023-03-07 ~ 2023-04-05



연관 키워드



Social

워드 클라우드

연관 키워드를 워드클라우드로
한눈에 확인할 수 있다.

금부정 추이



관련 콘텐츠

뉴스 블로그 유튜브



'꽃 판매 호조' 정희선 "더 열심히 하겠다"

4일 서울모빌리티쇼 현장 방문...KG모빌-벤츠-BMW 등 주요 전시
관 관람 정희선 현대자동차그룹 회장 2023.3.3/뉴스1 © News...

2023-04-04



잘 나가는 김주형, 우즈-메릴로이와 마스터스 대회 연습

기사내용 요약김주형, 2023 마스터스 토너먼트 연습 리운드우즈
최연소 기록 넣었던 김주형, 함께 연습 [오거스타(미국)]=AP/뉴...

2023-04-04



"4월에도 최저가 할인"은 계속된다"

이마트 최저가 '리미티드' 60종 곧 공개롯데마트 한달 내내 '반값족
발' 등 행사홈플러스 인기상품 50% 할인전 '영광'이마트-롯데마...

2023-04-04

받겠습니다, 수민님!



Social

공부정 차이

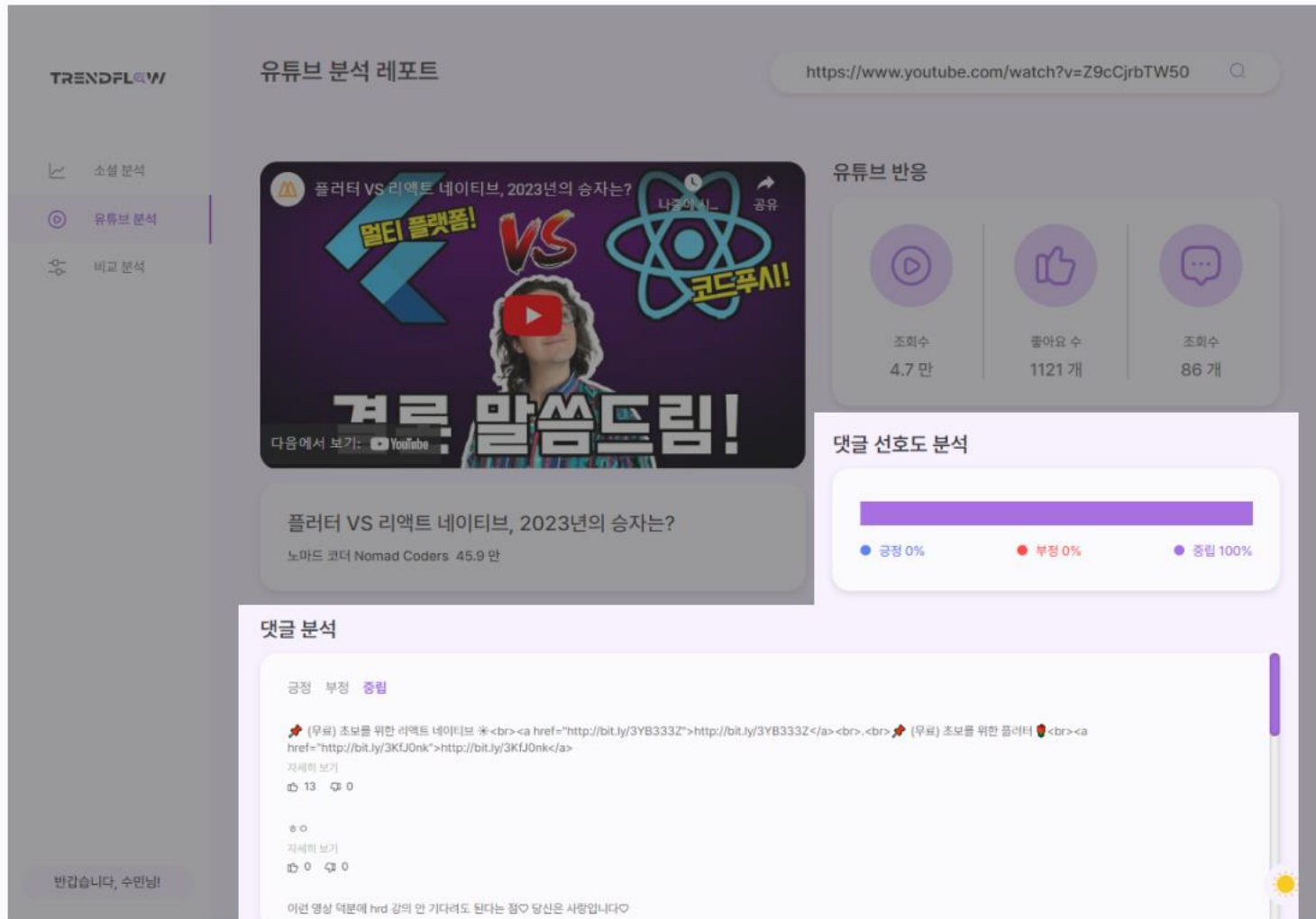
금부정 지수 변화 추이를
일간 꺾은선 그래프로 확인할 수 있다.



Social

관련 콘텐츠 조회

키워드와 관련된
기사, 블로그, 유튜브를 확인할 수 있다.



Youtube

댓글 및 댓글 선호도 분석

영상에 달린 댓글의 반응을
시각화된 자료로 확인할 수 있다.
선호도 별 댓글을 확인할 수 있다.

기획 배경

UCC

주요 서비스

특장점

시연

기술 스택

기대 효과

팀원 소개



150만 개의 데이터



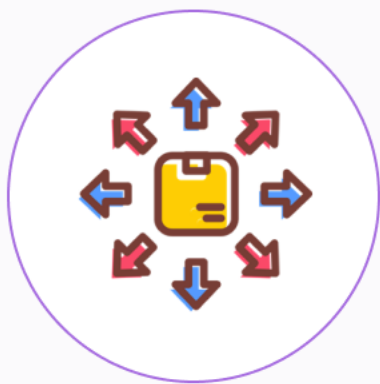
MSA 구축



캐시 서버 활용

빅데이터 분산을 통한 소셜 미디어 분석 서비스

TRENDFLOW



01
분산 처리



03
redis



02
MSA



04
CDD

YouTube

kafka

Streaming Data

Spark
Streaming

kafka

NAVER

Daum

hadoop
HDFS

Batch Data

APACHE
Spark

hadoop
Map Reduce

MySQL

spring
boot

Data Source

Unstructured Data

Analytics

Server

스케줄링/모니터링



기획 배경

UCC

주요 서비스

특장점

시연

기술 스택

기대 효과

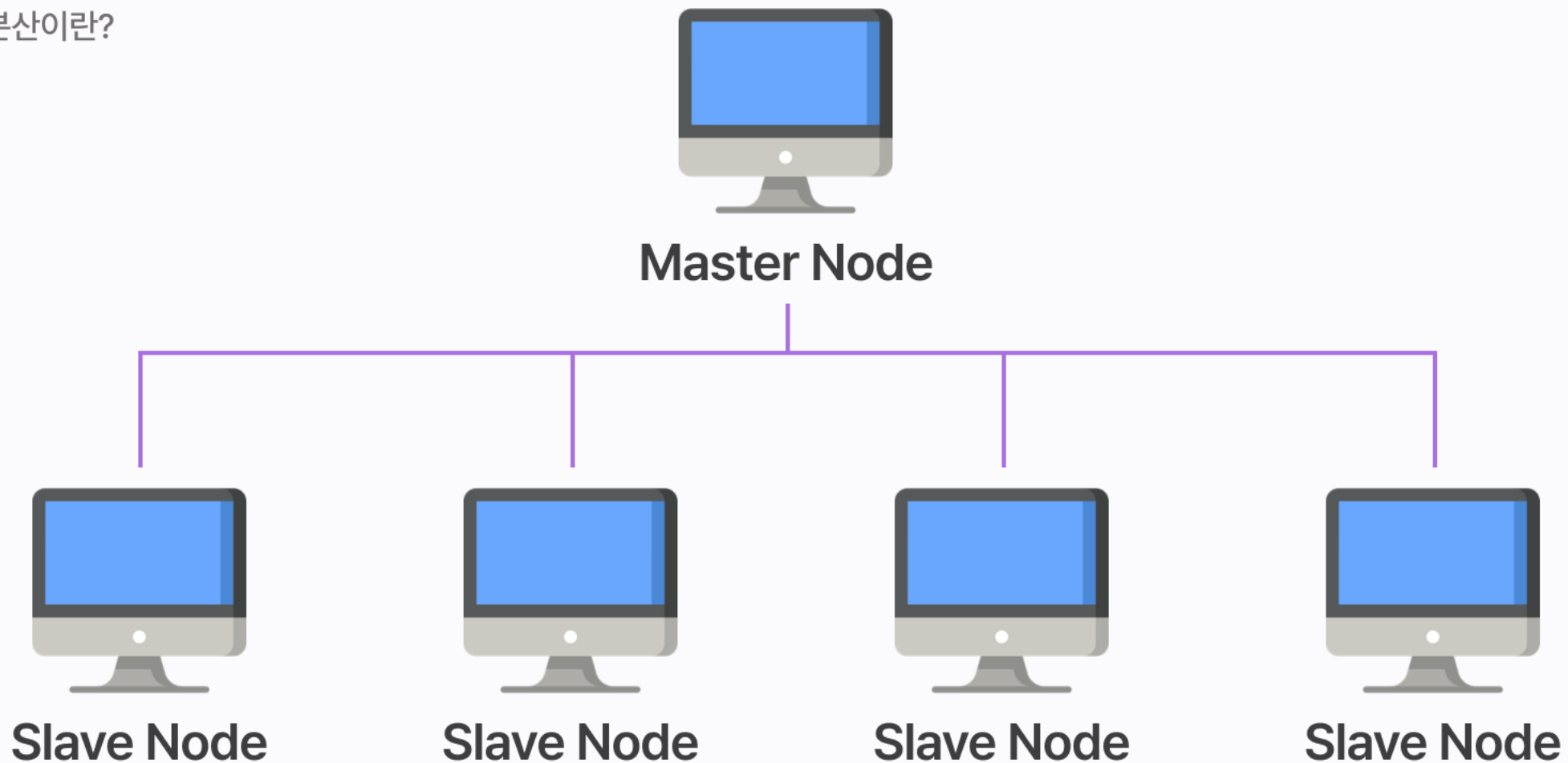
팀원 소개

분산 처리

분산이란?

✍️ 분산 처리

분산이란?



YouTube

kafka

Streaming Data

Spark Streaming

kafka

NAVER

Daum

hadoop HDFS

Batch Data

APACHE Spark

hadoop Map Reduce

MySQL

spring boot

Data Source

Unstructured Data

Analytics

Server

스케줄링/모니터링





분산 처리

배치 데이터 프로세스



Data Source



긍부정 추이 분석





분산 처리

Map Reduce

삼성은 이번 모델에서 5G 기술을 적용하여 더욱 빠른 인터넷 속도와 안정적인 네트워크 연결을 제공하고, 새로운 디자인과 고화질 카메라 등 다양한 기능을 탑재하여 고객들의 만족도를 높이고 있다.

또한 삼성전자는 최근 인공지능 기술 개발에 적극적으로 나서고 있다.

삼성전자의 경영 철학은 '사람 중심 경영'이다. 이 회사는 고객, 직원, 지역사회, 환경 등 모든 이해관계자들을 중심에 두고 경영을 하고 있으며, 이를 통해 글로벌 기업으로 성장하고 있다.

인공지능 기술을 적용한 새로운 제품 출시를 통해 다양한 분야에서 혁신적인 변화를 이끌고 있다.

특히 삼성의 인공지능 스피커인 '갤럭시홈'은 음성인식 기능을 바탕으로 일상생활에서의 편리함을 제공하며, 시장에서 큰 인기를 끌고 있다.



분산 처리

Map Reduce

삼성은 이번 모델에서 5G 기술을 적용하여 더욱 빠른 인터넷 속도와 안정적인 네트워크 연결을 제공하고, 새로운 디자인과 고화질 카메라 등 다양한 기능을 탑재하여 고객들의 만족도를 높이고 있다.

또한 삼성전자는 최근 인공지능 기술 개발에 적극적으로 나서고 있다.

삼성전자의 경영 철학은 '사람 중심 경영'이다. 이 회사는 고객, 직원, 지역사회, 환경 등 모든 이해관계자들을 중심에 두고 경영을 하고 있으며, 이를 통해 글로벌 기업으로 성장하고 있다.

인공지능 기술을 적용한 새로운 제품 출시를 통해 다양한 분야에서 혁신적인 변화를 이끌고 있다.

특히 삼성의 인공지능 스피커인 '갤럭시홈'은 음성인식 기능을 바탕으로 일상생활에서의 편리함을 제공하며, 시장에서 큰 인기를 끌고 있다.

인공지능

기술

적용

제품

출시

분야

혁신

변화

전자

개발

음성인식

삼성

기능

편리함

고객

만족도



✍️ 분산 처리

Map Reduce

연관 키워드



관련 콘텐츠

뉴스 블로그 유튜브



삼성SDI 주가 시동 거나...증권가 "증설 모멘텀 시작"

기사내용 요약증권사를 목표주가 상향'보수적 कैपा' 저평가 해소하
나 [서울=뉴시스]삼성SDI는 '인터배터리 2023'에 참가해 사물 ...

2023-04-06



롯데하이마트, '이사·혼수 트렌드 페어' 진행

롯데하이마트가 최근 엔데믹으로 결혼 수요가 늘어남에 따라 오는 30일까지 혼수 가전 할인 행사를 진행한다. 통계청 조사결과 지...

2023-04-06



[산업 리그테이블] ①힘들었던 2022년 '숫자로 나타났...

삼성·SK·LG·현대차·한화그룹 비금융 상장계열사 분석 /그래픽=비즈워치 지난해 한국 산업계는 생존을 위한 사투를 벌였다. 코로나...

2023-04-06

✍️ 분산 처리
긍부정 추이 분석



머신러닝

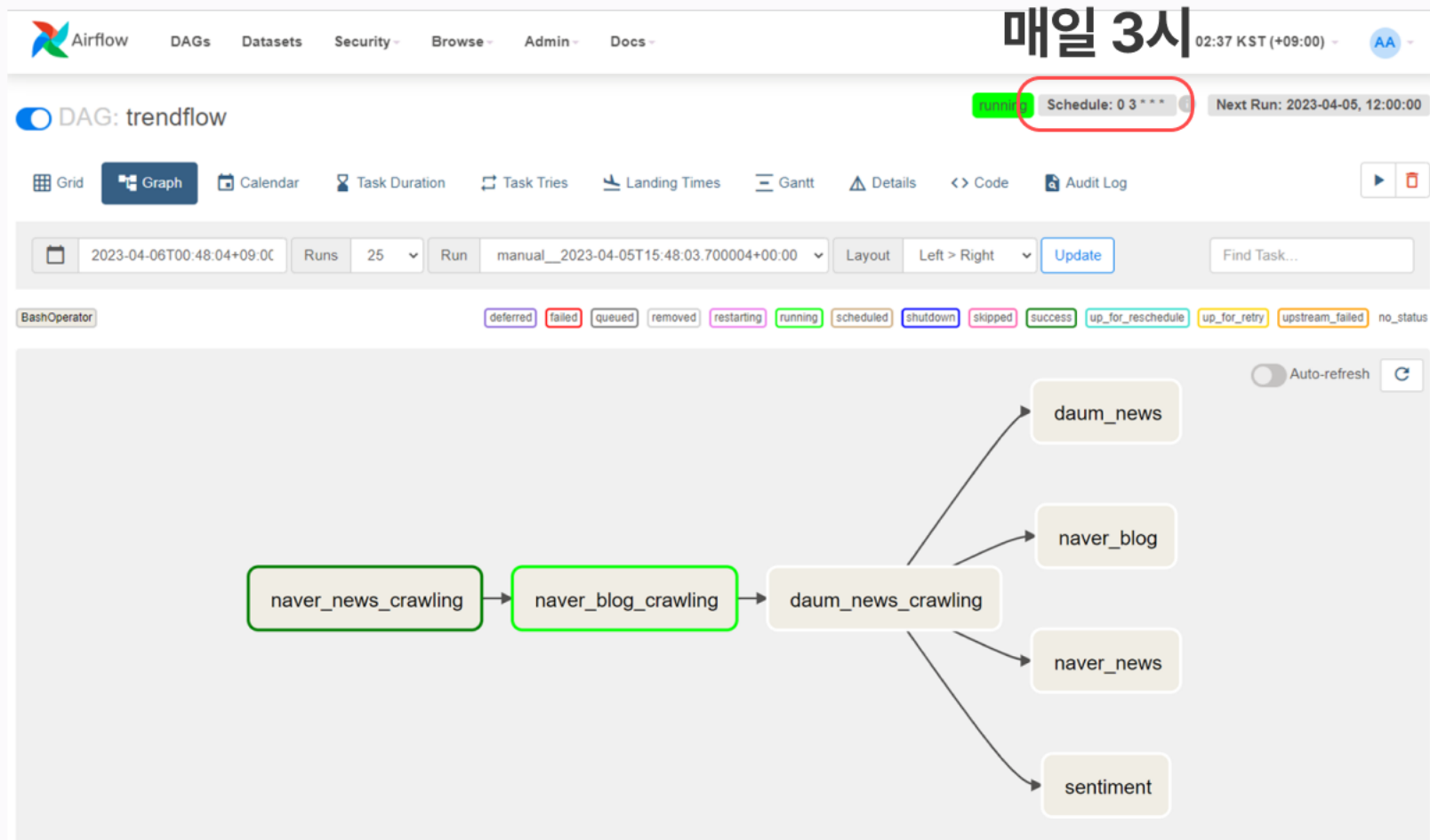


분산처리 기술



분산 처리

배치 데이터 프로세스





분산 처리

배치 데이터 프로세스

Airflow DAG: trendflow

매일 3시 02:37 KST (+09:00)

Schedule: 0 3 * * * (circled in red)

Next Run: 2023-04-05, 12:00:00

Grid Graph Calendar Task Duration Task Tries Landing Times Gantt Details <> Code Audit Log

2023-04-06T00:48:04+09:00 Runs 25 Run manual__2023-04-05T15:48:03.700004+00:00 Layout Left > Right Update Find Task...

BashOperator

deferred failed queued removed restarting running scheduled shutdown skipped success up_for_reschedule up_for_retry upstream_failed no_status

Auto-refresh

데이터 수집

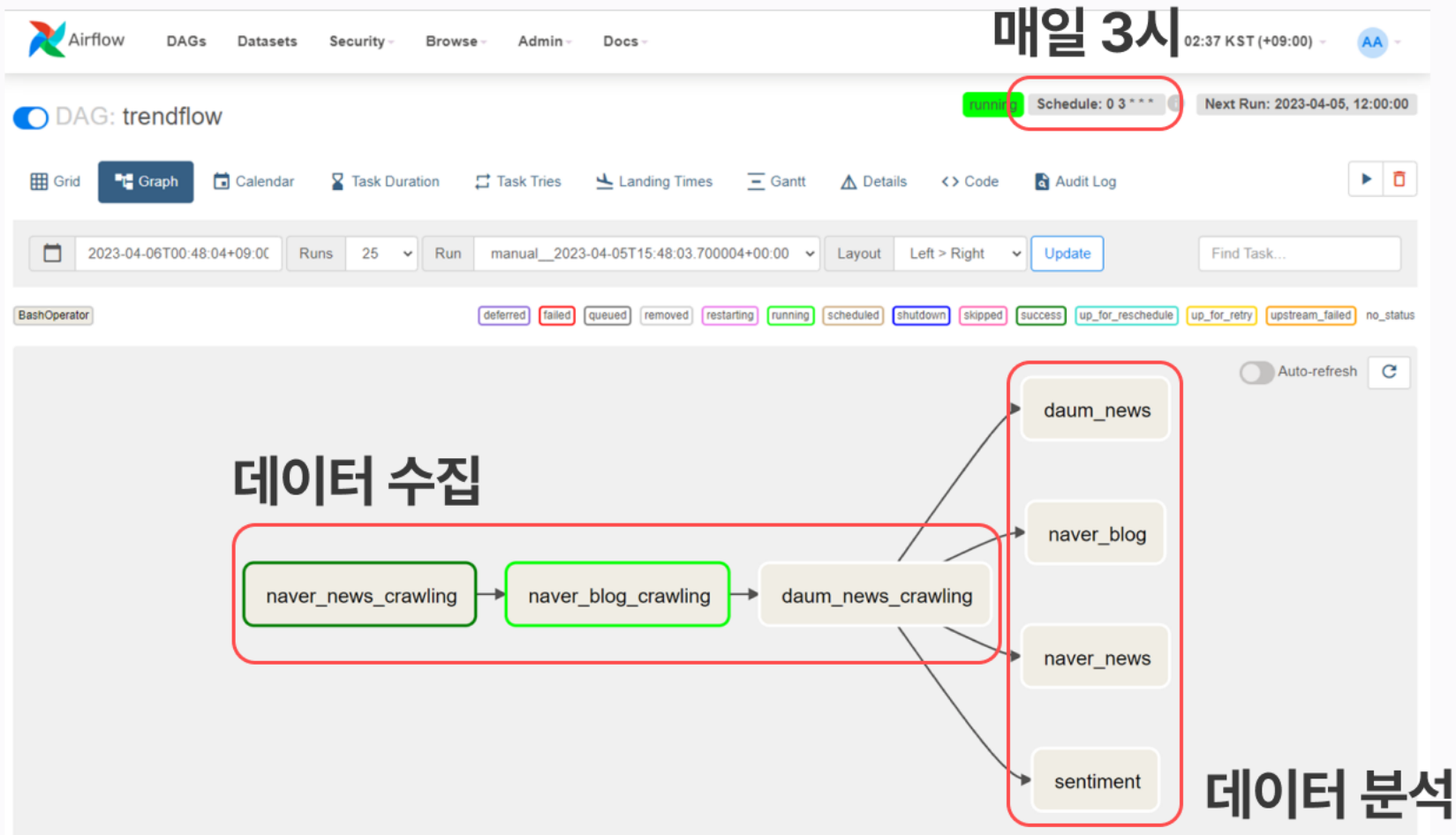
```
graph LR; naver_news_crawling --> naver_blog_crawling; naver_blog_crawling --> daum_news_crawling; daum_news_crawling --> daum_news; daum_news_crawling --> naver_blog; daum_news_crawling --> naver_news; daum_news_crawling --> sentiment;
```

The diagram shows a data collection process. It starts with 'naver_news_crawling', which flows into 'naver_blog_crawling', which then flows into 'daum_news_crawling'. From 'daum_news_crawling', the process branches out to 'daum_news', 'naver_blog', 'naver_news', and 'sentiment'. The first three crawling tasks are highlighted with a green border, and the entire sequence is enclosed in a red box.



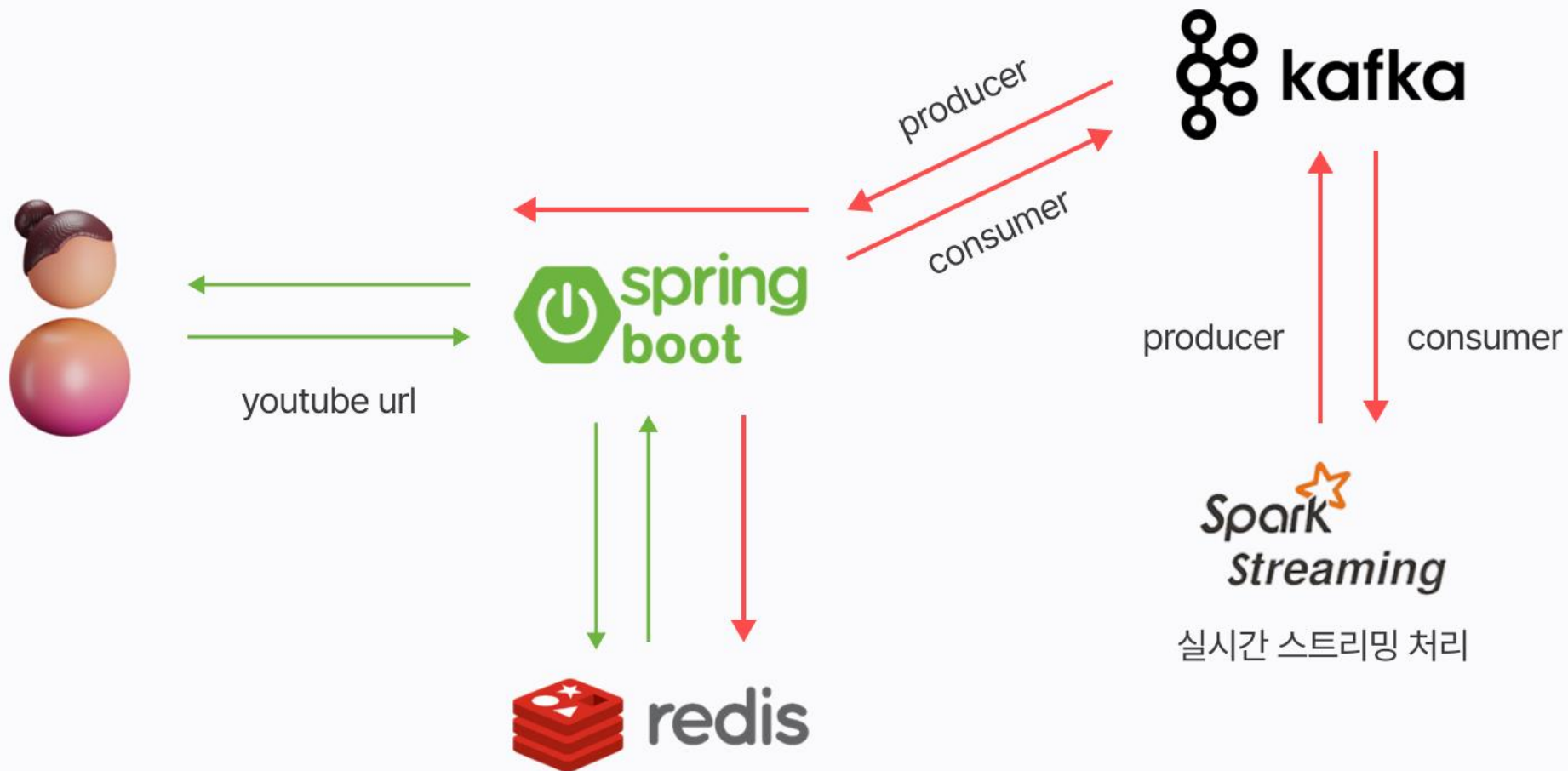
분산 처리

배치 데이터 프로세스



👉 분산 처리

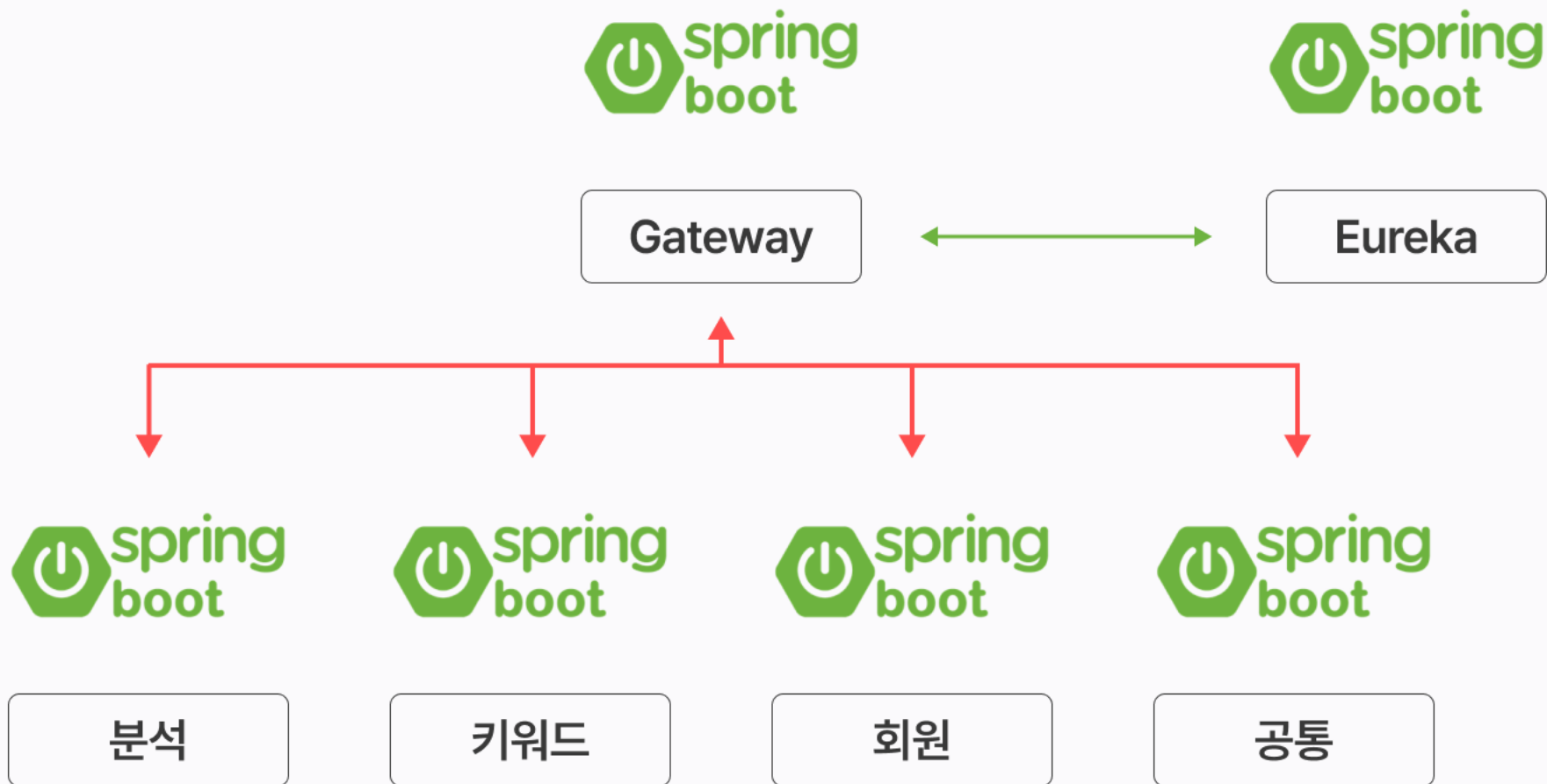
스트리밍 데이터 프로세스





MSA

MSA란?





로드 밸런싱

Eureka를 통해
많은 요청을 분산



피해 최소화

서버 다운 시
피해 최소화

Redis

Redis란?



01

주로 **캐시 DB** 용도로 사용되며,
다른 DB의 부하를 막아주는 역할

02

TRENDFLOW에서는
사용자의 정보를 저장하는 세션 캐시와
데이터 분석 결과를 저장하는 캐시로
이중화하여 구축

기획 배경

UCC

주요 서비스

특장점

시연

기술 스택

기대 효과

팀원 소개



CDD

CDD란?

Component Driven Development

PRESENTED BY



기획 배경

UCC

주요 서비스

특장점

시연

기술 스택

기대 효과

팀원 소개



CDD

CDD란?



디자인 시스템



아토믹 디자인 패턴



스토리북

Design System

Colors

				
BRAND 100 #BF1FF	BRAND 200 #E4CDFC	BRAND 300 #C79BF3	BRAND 400 #A86FE1	BRAND 500 #9147D8
				
RED 100 #FE2E2			RED 400 #FE4C4C	
				
BLUE 100 #DDE7FF			BLUE 400 #5B84EE	
				
WHITE 100 #BFAFC	WHITE 200 #F2F3F3	WHITE 300 #D0D0D0E	WHITE 400 #C0BFC1	
				
BLACK 100 #9B9B9B	BLACK 200 #716E73	BLACK 300 #3E3D3E	BLACK 400 #383838	

Typography

데스크탑 헤드라인1 Desktop Headline1

H1 2.25rem

데스크탑 헤드라인2 Desktop Headline2

H2 2rem

데스크탑 헤드라인3 Desktop Headline3

H3 1.75rem

데스크탑 헤드라인4 Desktop Headline4

H4 1.5rem

데스크탑 큰 글꼴 Desktop Large Text

lg 1.25rem

데스크탑 기본 글꼴 Desktop Basic Text

base 1rem



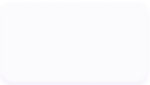
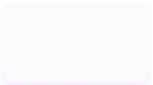
데스크탑 작은 글꼴 Desktop Small Text

sm 0.875rem

데스크탑 더 작은 글꼴 Desktop XSmall Text

xs 0.75rem

Shadow

	
BLACK SM 0px 3px 8px rgba(0, 0, 0, 0.1)	BLACK LG 0px 5px 10px rgba(0, 0, 0, 0.1)
	
BRAND SM 0px 3px 8px rgba(0, 0, 0, 0.1)	BRAND LG 0px 5px 10px rgba(0, 0, 0, 0.1)

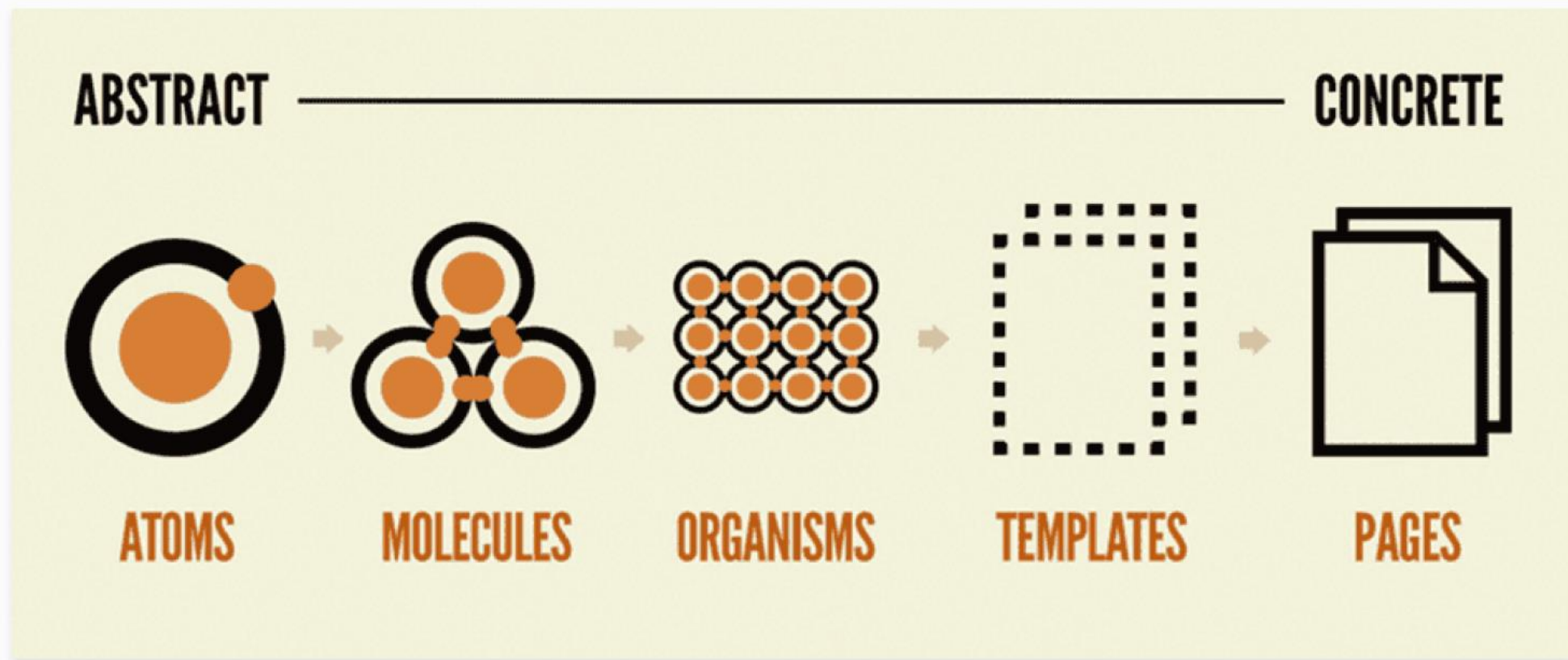
Radius

				
XS 8px	SM 12px	MD 16px	LG 24px	ROUND 50%

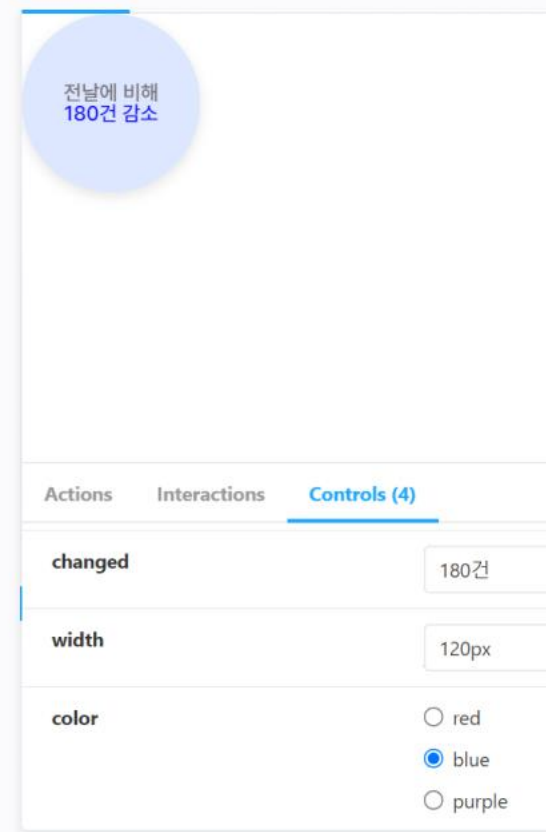
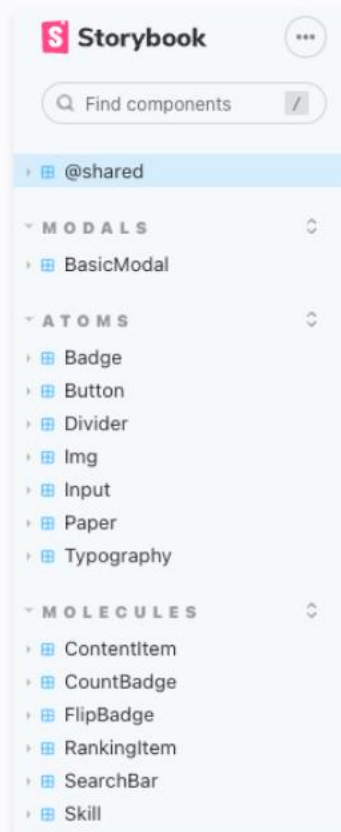
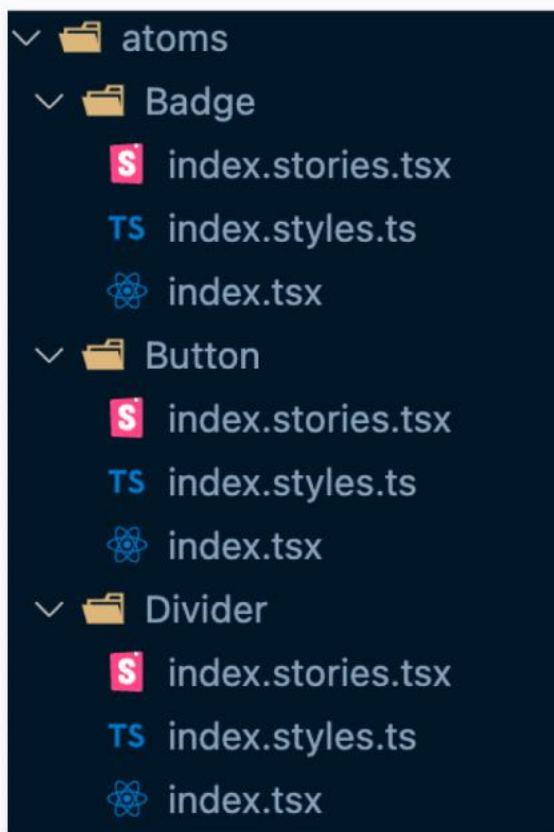


CDD

아토믹 디자인 패턴



CDD 스토리북



기획 배경

UCC

주요 서비스

특장점

시연

기술 스택

기대 효과

팀원 소개

Backend



Frontend



기획 배경

UCC

주요 서비스

특장점

시연

기술 스택

기대 효과

팀원 소개



기업

마케팅 전략 수립
경쟁사와의 차별화
문제점 개선



소비자

평판 파악
구매에 대한 의사 결정

기획 배경

UCC

주요 서비스

특장점

시연

기술 스택

기대 효과

팀원 소개



팀장 / Back
박상민



Back
권혁근



Back
김보경



Front
김수민



Front
문석환



Front
김진호

감사합니다.

부울경 특화 2반 E205

박상민 권혁근 김보경 김수민 김진호 문석환