

Realize your vision

SAMSUNG SDS Brightics AI v3.8



AGENDA

- I. AI/Analytics Trend
- II. Pain Point
- III. Why Brightics ?
- IV. Brightics 주요기능

I. AI/Analytics Trend

“ AI/Analytics Trend ”



COVID-19 이후
데이터 분석
시장의 변화



데이터 거버넌스의
중요성 증가



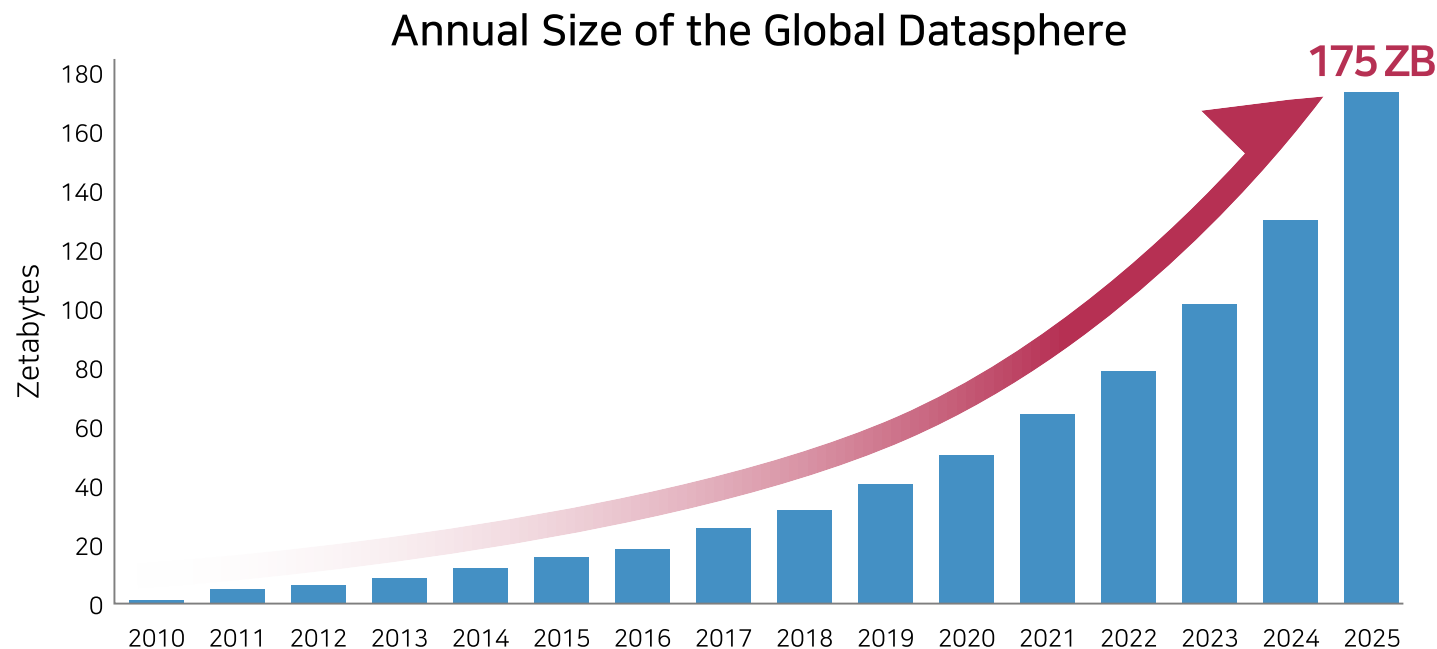
DW를 결합한
Data Lake의
확산 가속화



MLOps를 통한
모델 개발 및
운영 최적화

① COVID-19 이후 데이터 분석 시장의 변화

COVID-19로 인해 시장의 변화는 매우 빠르고, 경험하지 못한 환경을 만들고 있습니다.
급격한 변화에 대응하기 위해 새로운 AI 기술의 개발과 보다 빠른 속도의 분석이 요구되고 있습니다.



Gartner®

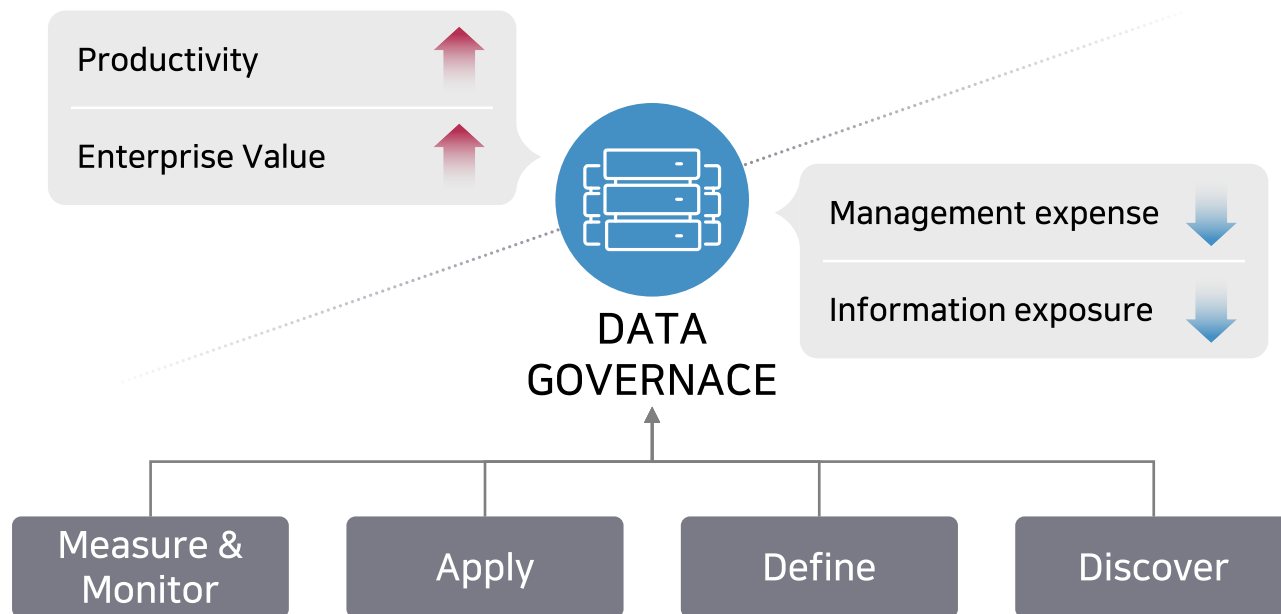
전례없는 시장 변화에 직면하여
데이터 및 분석 리더는 혁신을 가속화하고
COVID-19 이후 세계로 가는 새로운
경로를 구축하기 위해 처리 및 액세스
측면에서 분석의 속도와 규모가 계속
증가해야 합니다.

Gartner analyst, VP Rita Sallam

* Source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018

② 데이터 거버넌스의 중요성 증가

데이터 세계와 분석 세계가 분석 세계를 중심으로 통합되고,
데이터 분석은 매우 다양하고 복잡한 프로세스로 구성되어 End-To-End 관점에서의 관리가 필수적입니다.
이러한 상황을 해결하기 위한 **데이터 거버넌스의 중요성이 증가하고 있습니다.**



Market Insight Report

데이터 거버넌스 시장은
2020년 21억달러에서
2025년에는 57억달러로 전망기간 중
22.3%의 연평균 성장률(CAGR)을
기록할 전망이다.

MarketInsightReport, 2020

* Source : TTA(Telecommunications Technology Association)

③ DW를 결합한 Data Lake의 확산 가속화

정형 데이터 기반의 DW와 반/비정형 데이터 기반의 Data Lake를 따로 구축하지 않고
Data Lake를 중심으로 통합하여 Insight를 도출하는 방향으로 시장은 진화하고 있습니다.
Data Lake를 성공적으로 구축/활용하기 위한 고도화 된 솔루션은 필수적입니다.



확장성 있는 아키텍처 지원

조직이 Digital Transformation을 추구함에 따라 데이터의 생산량은 기하 급수적으로 증가하며 이를 빠르고 무결하게 지원할 수 있는 아키텍처가 필요합니다.



장기적인 데이터 관리 및 거버넌스 지원

데이터 Provenance, 메타 데이터 관리 뿐만 아니라
데이터 Lifecycle 측면까지 관리되어야 합니다.



낮은 비용과 높은 신뢰성을 지원

불필요한 데이터까지 수집하여 정제하는 과정에서 저장 및 관리 비용은
급속히 증가가 예상됩니다.

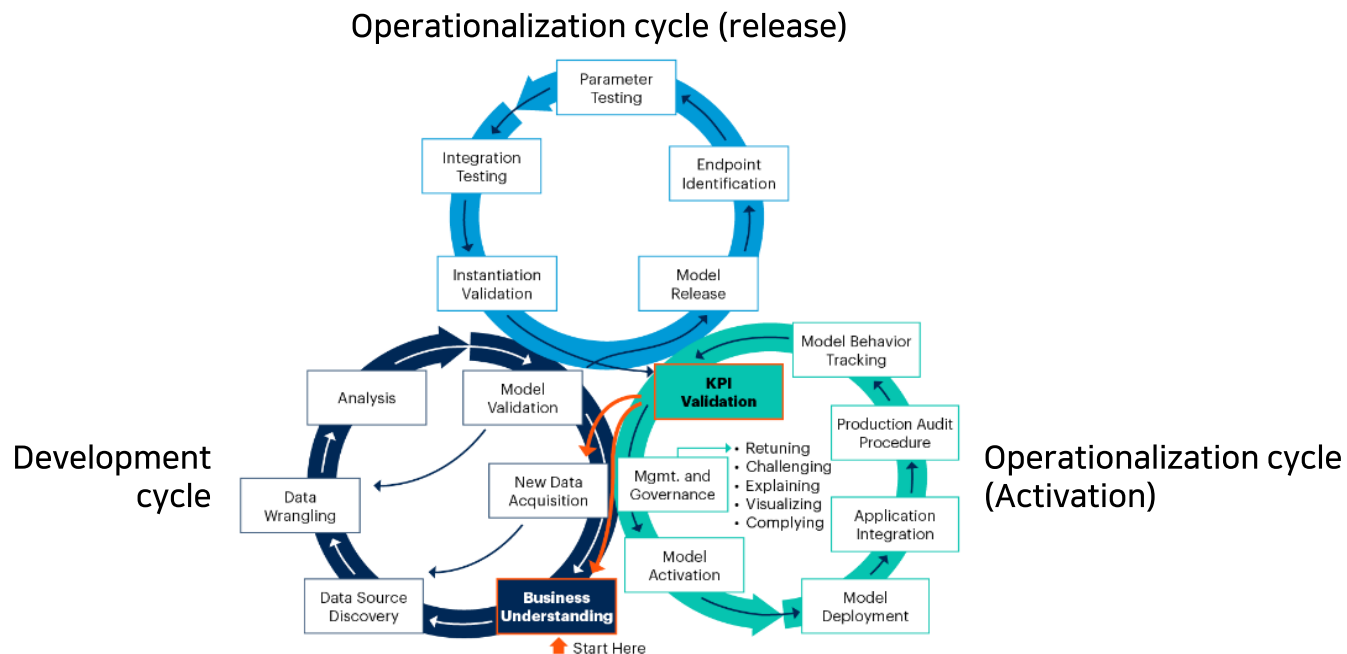
Market Insight Report

Data Lakes Market은
2019년에 374억 달러로 평가되었으며
2025년에는 1,760억 달러에 도달하여
2020년 ~ 2025년 기간동안
연간 29.9% 성장할 것으로 예측됩니다.

MarketInsightReport, 2020

④ MLOps를 통한 모델 개발 및 운영 최적화

AI 개발/서비스 환경은 Data Scientist를 위한 AI 개발 환경 제공 중심에서 개발 모델 최적화 및 운영 관리 등 모든 단계를 연결하여 지속적으로 고도화하는 MLOps로 발전중입니다. MLOps 도입으로 **AI 개발 시간을 단축하고 AI 운영을 최적화** 할 수 있습니다.



Gartner®

기업에서 DSML¹의 사용은
지난 2년 동안 3배로 증가했습니다.
IT 리더는 AI 생산성을 최적화하고,
데이터 및 분석 리더는 AI 운영을
가속화 하기 위해 AI 오케스트레이션
플랫폼을 고안해야 합니다.

Gartner, '20.12월

* Source: Gartner 725627_C

¹ DSML: Data Science & Machine Learning

II. Pain Point

Digital Transformation에 대한 고민

1

복잡하고 다양한 **필요 데이터**를
어떻게 **실시간으로 수집**할지?

3

디지털 혁신을 위해 수집된
데이터를 어떻게 **통합, 활용**할지?



2

AI, 자동화로 얻을 수 있는
비즈니스 효과는 무엇인지?

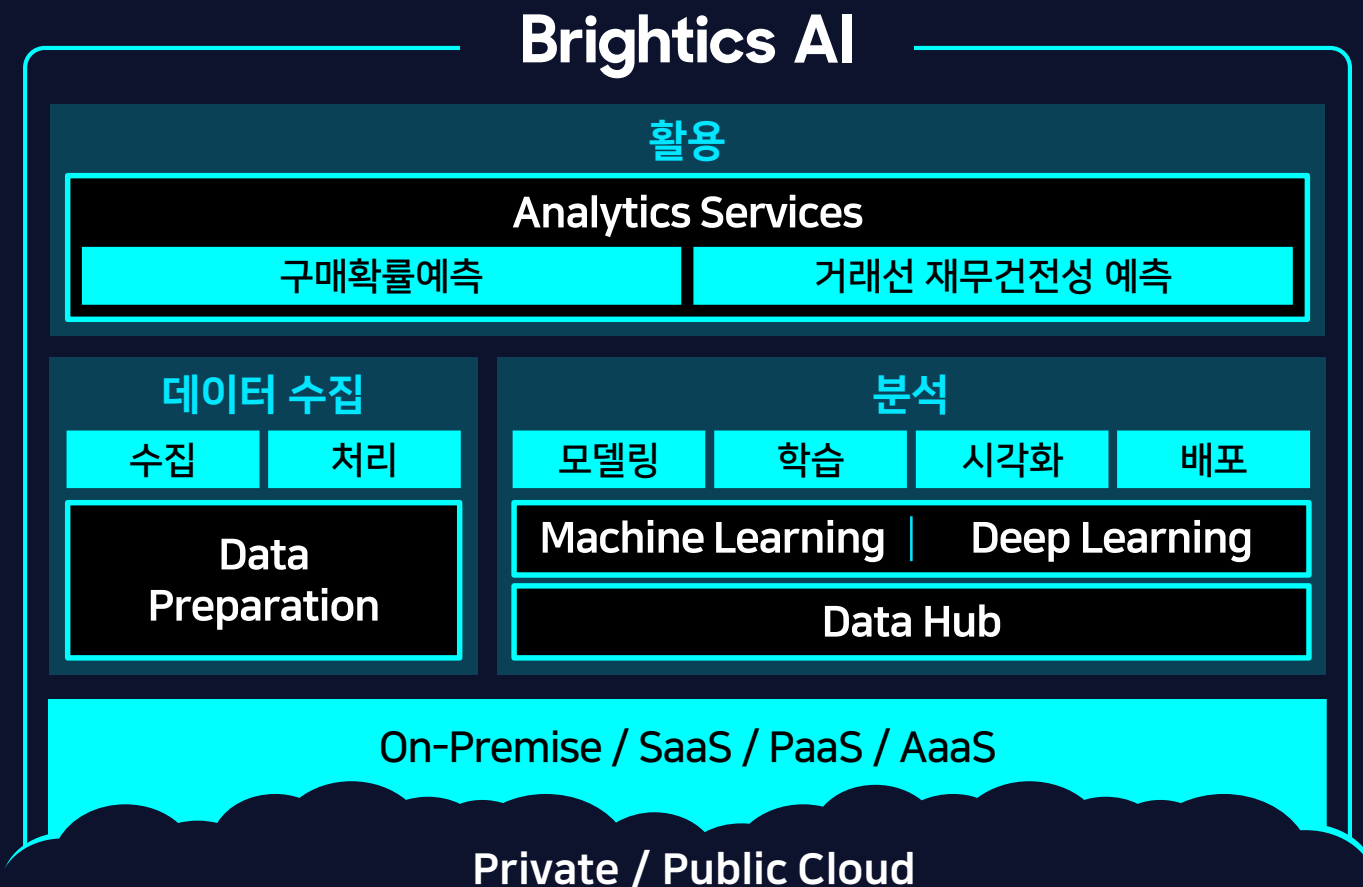
4

회사에 **디지털 혁신**을 실행 할
역량이나 전문가가 있는가?

III. Why Brightics ?

Brightics AI는

흩어져 있는 방대한 양의 정보를 수집, 처리하여 AI로 빠르게 분석하고, 이해하기 쉽게 시각화해 주는 '통합 AI 플랫폼'으로 국내 최초로 Gartner, Forrester 에 동시 등재되었습니다.



Gartner
Magic Quadrant 등재



Forrester Wave 등재



GS 1등급 인증 획득



CIO 100 Awards 수상
'16년 ~ '20년 연속



Brightics AI 관련 특허



Brightics AI는 분산된 데이터를 자동화 처리기능을 통해 직관적으로 수집하고 관리 할 수 있습니다.

- ① Ingest Accelerator 기능으로 데이터 고속 수집이 가능합니다.
- ② 다양한 데이터 유형과 데이터 소스로부터 데이터 수집이 가능합니다.
- ③ 실시간 데이터 처리 및 분석이 가능합니다
- ④ Auto Schema Builder로 Data 파싱 자동화를 지원 합니다.
- ⑤ Visual Query로 데이터 전처리를 쉽고 편리하게 합니다.

Brightics AI는 비즈니스 요구에 최적화된 Hadoop 환경을 편리하게 구성하고 관리합니다.

- ❶ 안정적인 Data 분석을 위한 통합 하둡 배포판을 제공합니다
- ❷ 편리한 Web Installer 제공합니다.
- ❸ 서비스별 Configuration 최적값을 추천합니다.
- ❹ 시스템 자원의 통합 모니터링 Dashboard를 제공합니다
- ❺ 서비스 장애 Alert 기능을 제공합니다.



Brightics AI는 누구나 쉽고 빠르게 AI 분석 서비스를 개발하고 활용할 수 있습니다.

- ① 시각화된 분석환경으로 쉽고 직관적으로 활용할 수 있습니다.
- ② 자동화된 분석 가이드만 따르면 분석이 어렵지 않습니다.
- ③ Container 기반의 개인화된 맞춤형 분석 환경을 제공합니다.
- ④ Edge Analytics 기술로 분석 처리속도를 향상시킵니다.
- ⑤ 분석모델의 결과를 손쉽게 리포트로 제작 할 수 있습니다.

Brightics AI는 데이터 라벨링 자동화와 최적모델 자동추천으로 딥러닝 개발 생산성을 극대화 합니다.

- ① 높은 정확도의 이미지 데이터 전처리 자동화를 지원합니다.
- ② 글로벌 수준의 확장된 자동 라벨링 기능을 지원합니다.
- ③ 통합 라벨 매니지먼트 기능을 통해 라벨링 협업이 가능합니다.
- ④ 학습 데이터에 최적의 딥러닝 모델을 자동 추천합니다.

Brightics AI는

권위있는 국내/외 어워드에서 꾸준한 수상 이력 보유하고 있습니다.

CIO 100 Awards

2017~2020

IDG社 주관으로 1987년 시작,
매년 1회 IT기술을 통해 높은 사업가치를 창출한
100개 회사를 수상자로 선정

물산 '20년

건설안전사고 예방을 위한 Digital Transformation

전기 '19년

AI 기반 MES 구축을 통한 Intelligent Factory 구현

전자 독일 '18년

AI기반 가전 수요량 예측 시스템 구축

물산 '17년

Cloud 기반 반도체 제조 Big Data 분석 통합 Infra 구축

임팩테크

2019

과학기술정보통신부에서
주최하고,
한국경제 전문사와
한국정보통신 진흥협회가
공동 주관하는 ICT 분야
국내 최고 권위의 상



과학기술정보통신부
장관상 수상

UX 디자인 어워드

2018

독일 베를린 국제 디자인
센터에서 진행하며
제품/App.분야에 특화된
UX전문 어워드



IDEA 디자인 어워드

2017

미국 산업디자이너협회가
주관하는 국제 디자인상으로,
레드닷 디자인 어워드,
iF 디자인어워드와 함께
세계 3대 디자인상



Brightics AI는

전자/제조, 공공, 금융 서비스 등 다양한 기업에서 활용하고 있습니다.



전자/제조

SAMSUNG

SAMSUNG ELECTRO-MECHANICS

SAMSUNG SDI

SAMSUNG DISPLAY

AMORE PACIFIC



공공



기획재정부



과학기술정보통신부



Incheon Airport
인천국제공항공사



경찰청
KOREAN NATIONAL POLICE AGENCY



소방청
National Fire Agency



LH 한국토지주택공사



KEA Korea Electronics Association
한국전자정보통신산업진흥회



수원시
SUWON CITY



유통/서비스/플랜트

SAMSUNG C&T



GOM
& COMPANY

SAMSUNG ENGINEERING

searcheese



KAIST



UNIST
ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY



금융



한국은행
THE BANK OF KOREA

SAMSUNG LIFE INSURANCE

SAMSUNG CARD

NH농협생명

Brightics AI는 다양한 생태계 조성 활동을 수행하고 있습니다.



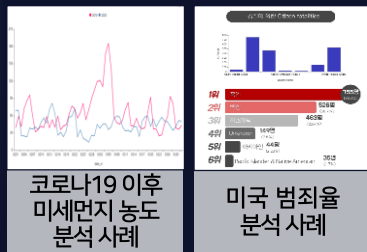
Brightics
Academy

10개 대학교



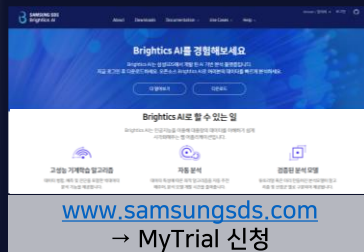
Brightics
Supporters

20개 대학교



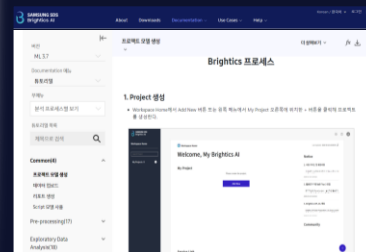
Trial

무료체험



Tutorial

200개
시나리오



Brightics TV

사용법/사례
영상



CSR

중·고등학생
데이터분석 교육



IV. Brightics 주요기능

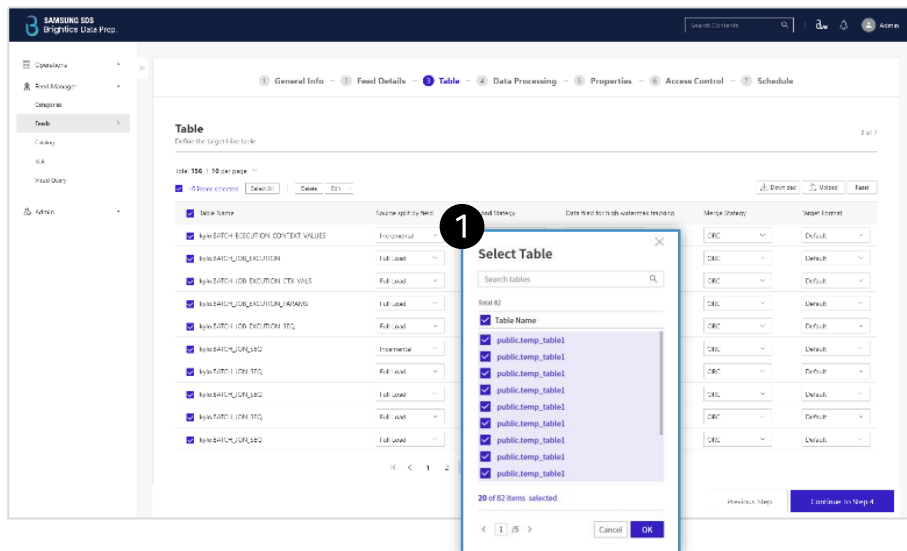
1 Ingest Accelerator 기능으로 데이터 고속 수집이 가능합니다.

대용량, 대규모 테이블을 고속 수집할 수 있는 Ingest Accelerator 기능을 제공합니다.
이를 통해 Apache Sqoop 대비 데이터 수집 속도를 30% 개선하였습니다.

대량 테이블 고속 수집

✔ Multi Table Ingestion Feed 생성

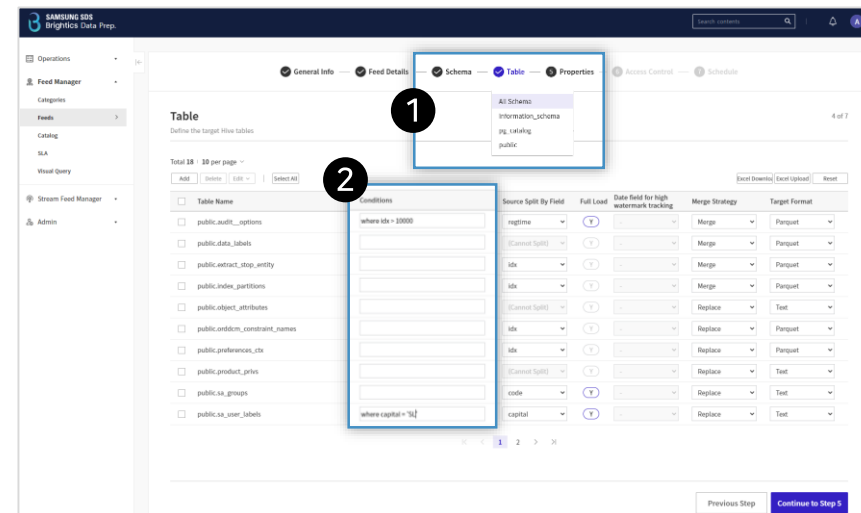
- 1 여러 개의 테이블을 하나의 Feed를 통해 대량 데이터 수집 가능



조건에 맞는 맞춤형 데이터 수집제공

✔ Data Table 조회 및 수집

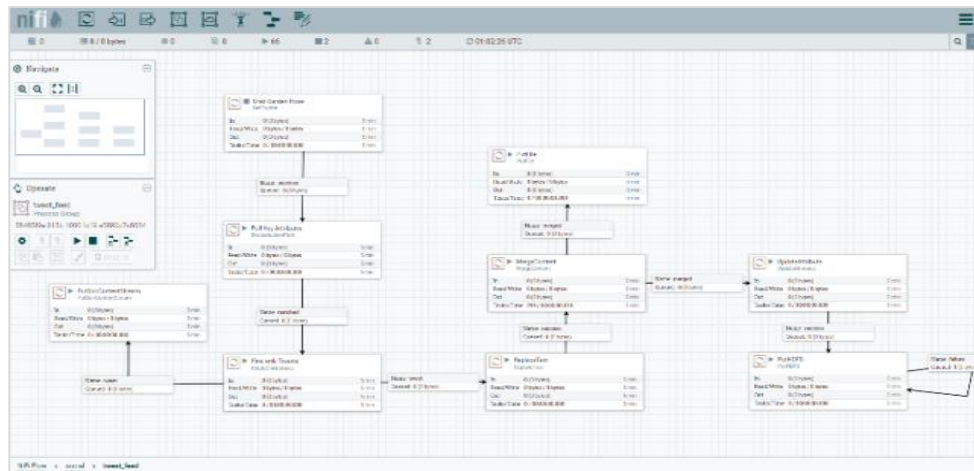
- 1 DB의 Schema 단위로 Table 조회 및 수집 가능
- 2 데이터 수집 시 조건 (Where절) 추가 가능



2 다양한 데이터 유형과 데이터 소스로부터 데이터 수집이 가능합니다.

데이터 소스 51종으로부터 데이터 수집이 가능하고, 정형뿐만 아니라 비정형데이터의 수집도 가능합니다.
템플릿 기능을 활용하여 데이터 수집을 쉽고 편리하게 할 수 있습니다.

1 데이터 파이프라인 설계



- 수집 데이터 소스 51종 지원
(File, FTP, Kafka, Oracle, GreenPlum, HDFS/Hive, S3/Azure 외 다양한 정형 / 비정형 소스 지원 및 사용자 정의 커넥터 개발 용이)

2 작업 유형별 Template 등록

Name	Feed Count	Last Updated	Export Template
Advanced Ingest(No Savepoint)	2	Jan 3, 2020 15:44:48	Export
Data Ingest	8	Jan 6, 2020 12:47:42	Export
Data Transformation	2	Jan 6, 2020 12:24:33	Export
GPDB Data Transformation	2	Jan 6, 2020 11:50:19	Export
Semi-Structured Data Ingest	2	Jan 6, 2020 13:09:24	Export

Data Ingest / Transformation / Stream Ingest 등

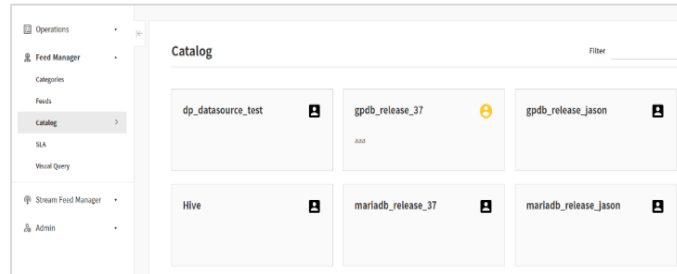
3 Template 세부사항 설정

- 데이터 소스 / 타겟 스키마 지정
- 데이터 병합 규칙 설정
- 배치 스케줄

3 데이터 가상화를 통해 데이터 수집 전, 탐색/변환이 가능합니다.

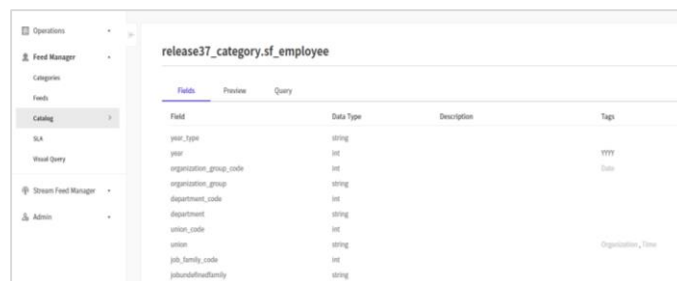
이기종 소스시스템의 데이터를 수집 전에 미리 탐색하고, 병합 및 변환하여 수집하고자 하는 데이터를 미리 확인해 볼 수 있는 기능을 제공합니다.

외부 / 내부 데이터 소스 통합 관리



- Oracle, GPDB 등 RDB 데이터 소스 등록

데이터 카탈로그



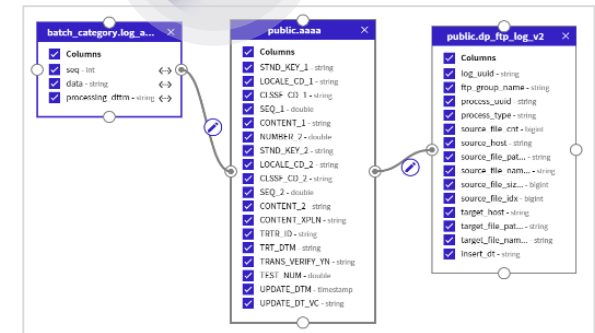
- 데이터 카탈로그를 통해 Preview, Query 가능

외부 / 내부 데이터 병합 결과 미리 보기

id	order_id	order_date	delivery_date	delivery_type	customer_id	store
1	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
2	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
3	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
4	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
5	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
6	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
25	US-2013-02021020-4	1300221200000	1300280400000	Level 1	SL-2022082	Store 3
47	US-2013-02021020-4	1300280400000	1307420400000	Standard	KS-2014036	Store 1
48	US-2013-02021020-4	1300280400000	1307420400000	Standard	KS-2014036	Store 1
63	US-2013-02021020-4	1307420400000	1307420400000	Level 2	LC-1000000	Store 3
66	US-2013-02021020-4	1307420400000	1307420400000	Standard	KS-2014036	Store 1
73	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
74	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
75	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
76	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
89	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
90	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
91	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
92	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
106	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14
112	MO-2013-080101020-4	1301762000000	1302104000000	Standard	KM-1057028	Store 14

- 병합 결과 정제, 변환, 가공

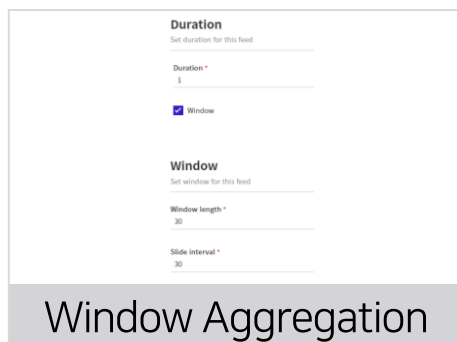
- 외부/내부 데이터 병합 (ex. Oracle + Hive Table)



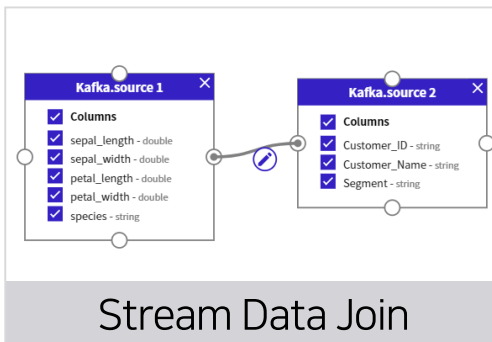
4 실시간 데이터 처리 및 분석이 가능합니다.

정해진 Window Size로 Stream Data를 Block화 하여 Aggregation 할 수 있습니다.
실시간으로 테이블간 조인이 가능하며, 대시보드를 통한 실시간 모니터링 기능을 제공합니다.

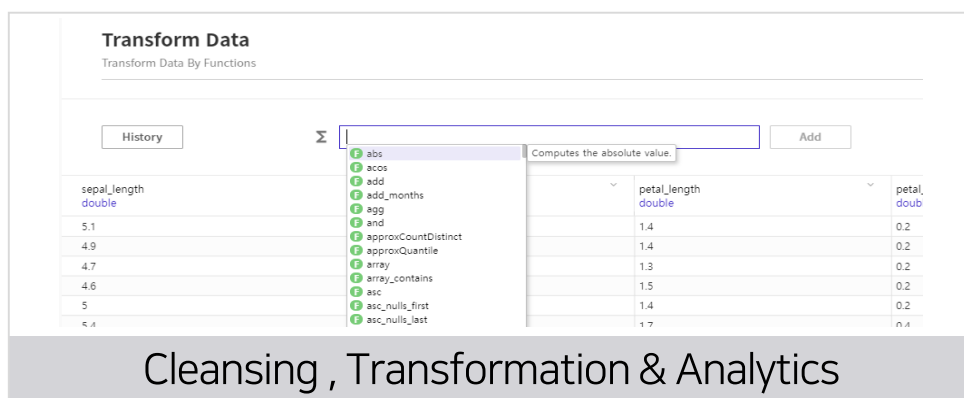
① 실시간 데이터 수집 / 처리 / 분석



Window Aggregation



Stream Data Join



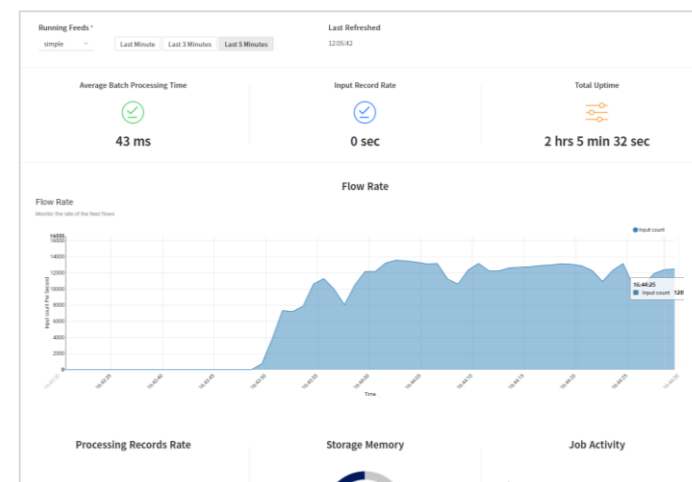
Cleansing , Transformation & Analytics

② Spark Streaming 기반 분산 처리

※ 200,000 Record/sec/node
(2 Stream data Join, Window and Aggregation)



③ 실시간 모니터링



5 Auto Schema Builder로 Data 파싱 자동화를 지원 합니다.

반정형 로그 데이터를 데이터 영역 인지 기술과 패턴 추천 기술을 통해 자동으로 정형 데이터로 변환할 수 있습니다.

2 데이터 유형 자동 파악

sample_type3.txt
File Size 2.08 KB | Row Count 25

Estimated Data Type
Fixed Length 18 % Delimiter 0 % Key-Value 12 % Report 68%

3 Label/Table 영역 인지/분리

4 정형 데이터로 자동 파싱
- 컬럼 인지 / 패턴 매칭 기술

5 Column/Row 보정 UI 제공

1 Load sample text log data

Data

Valid Label 9 Header 1 Table 10 Invalid Different 0 No-Target 0 Outlier 0

Label

Table

Columns 8 Rows 10 Valid 10 Invalid 0

	Time	Step_No.	[M04]_NF3_SET	[M04]_NF3_FLOW	[M04]_NF3_VALVE	[M04]_NF3_PRESS	[M04]_NF3_TEMP
1	[12:33:43.453]	1	0	0.000	0.000		
2	[12:33:43.765]	1	0	1.000	0.000		
3	[12:33:44.406]	1	0	0.000	0.000		4000
4	[12:33:44.781]	1	0	1.000	0.000		4000
5	[12:33:45.265]	1	0	1.000	0.000		4000
6	[12:33:47.265]	1	0	1.000	0.000		4000
7	[12:33:47.266]	1	0	1.000	0.000		3000
8	[12:33:47.267]	1	0	1.000	0.000		2000
9	[12:33:47.268]	1	0	1.000	0.000		4000
10	[12:33:47.269]	1	0	1.000	0.000		4000

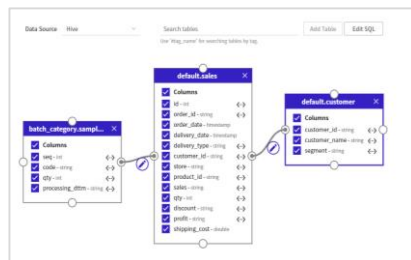
Change Column Name
[M04]_NF3_VALVE

Duplicate Column
Remove Columns
Merge Columns
Substring Data
Replace Data
Apply Regular-Expression

6 Visual Query는 데이터 전처리를 쉽고 편리하게 합니다.

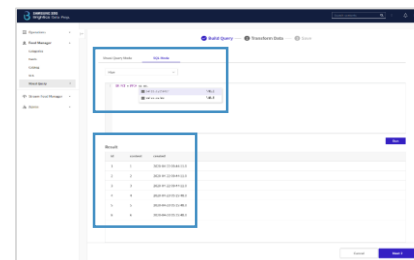
SQL에 능숙하지 않은 비전문가도 GUI를 통해 테이블간 조인, 데이터 필터링이 가능하며 Visual Transform을 통해 분석하기 좋은 형태로 데이터를 가공할 수 있습니다.

Visual Join



- 대규모 데이터 결합
- SQL 편집 기능 제공
- 이기종 소스 데이터 Virtual 조인 (ex. Oracle + Hive Table)

SQL모드 강화



- SQL모드를 통하여 SQL Query를 활용한 데이터 조회 가능
- Table 검색 및 자동 완성 (예약어, Table, Column)
- SQL Query를 통해 조회된 결과를 확인하고 데이터 추출 가능

Visual Transform

- Excel-like 한 직관적인 UX
 - 쉽고 빠르게 데이터 정제/변환
- Preview
 - 샘플 데이터로 예상 결과물 미리 확인
- 수백 종의 데이터 변환 함수 (한글 전처리 등) 제공
- Filtered Histogram

함수 시각화/히스토리 기능 제공

- 함수, 히스토리, 히스토그램
- 함수 시각적 편의성 증대
 - 예제 제공/결과 통계값 제공
 - 함수 히스토리 기능 제공

7 Semantic Matching은 Data 정제 자동화를 지원합니다.

유사 데이터셋 자동탐지 및 표준값 추천을 통해 테이블간의 관련성을 자동 탐지해 줄 뿐만 아니라 데이터 정제를 효과적으로 할 수 있습니다.

유사데이터 자동 추출 & 표준 값 추천

1 Group similar data in columns

2 Refine to standard values

Row Count	Values	Recommended Value
4 (100.00%)	4	Galaxy J3 2017
×	1 (25.00%)	Galaxy J3_2017
×	1 (25.00%)	Galaxy_J3_2017
×	1 (25.00%)	galaxy_j3_2017

- Provides various matching algorithms (K-Means, Minhash-LSH, String Edit Distance, Metaphone)

조인키 자동 탐지

1 Recommend Table & Column

2 Easily Join Data

- Algorithms : Jaccard Similarity

Columns	Value Matching(%)
sample_trans seq	100.0%

Relative Statistic		Numeric Stats	
Null	0%	MIN	1
Empty	%	Max	3
Unique	100%	Mean	2
Duplicates	0%	Standard Deviation	0.8165
		Variance	0.6667
		Sum	6

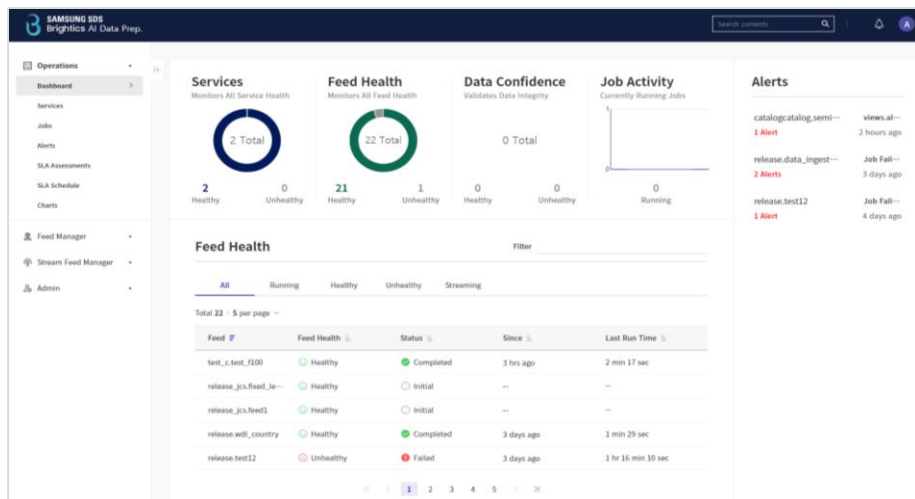
Value	Count
3	1
2	1
1	1

Setting

8 데이터 수집 및 전처리 Job에 대한 모니터링 기능을 제공합니다.

SLA기반으로 데이터 관리를 할 수 있도록 대쉬보드를 제공하며, 오류 발생 시 즉각적으로 대응할 수 있도록 Push방식의 Alert 기능을 제공합니다.

Dashboard & Job 모니터링



- Dashboard를 통해 Feed/Job Monitoring 등을 사용자가 직관적으로 파악하도록 시각화 기능 제공
- Service 상태, Job 건수 및 실행 상태, 데이터 신뢰도 등을 파악

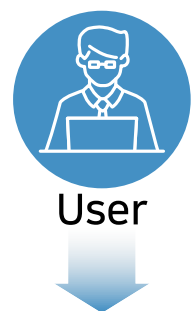
데이터 SLA

SLA Assessments		
All	Failure	Warning
Total 1771 5 per page		
Name	Result	Created
release.data_ingest_filesystem_manual_um...	Success	Jan 6, 2020 15:25:00
release.data_ingest_filesystem - Feed Failur...	Success	Jan 6, 2020 15:25:00
catalogcatalog.semi_fix_0106 - Feed Process...	Warning	Jan 6, 2020 15:25:00
release.data_ingest_filesystem_manual_um...	Success	Jan 6, 2020 15:20:00
release.data_ingest_filesystem - Feed Failur...	Success	Jan 6, 2020 15:20:00

- 수집 오류가 발생하거나 정해진 시간안에 Job이 완료되지 않는 경우 Alert

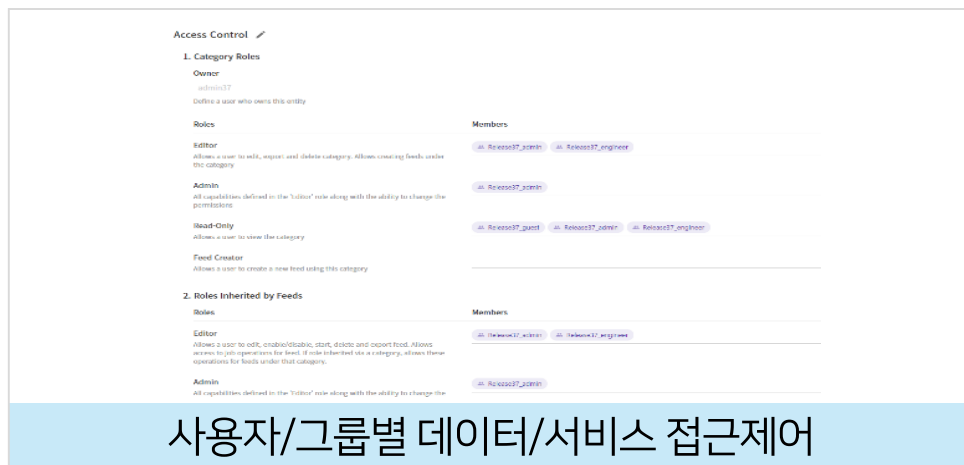
9 사용자에 대한 권한관리 및 데이터별 권한관리를 제공합니다.

사용자 Level 및 사용자 그룹을 통해 접근 권한을 관리할 수 있고, 데이터 작업공간/테이블/컬럼별로 데이터 접근제어(Access Control) 체계를 제공합니다.

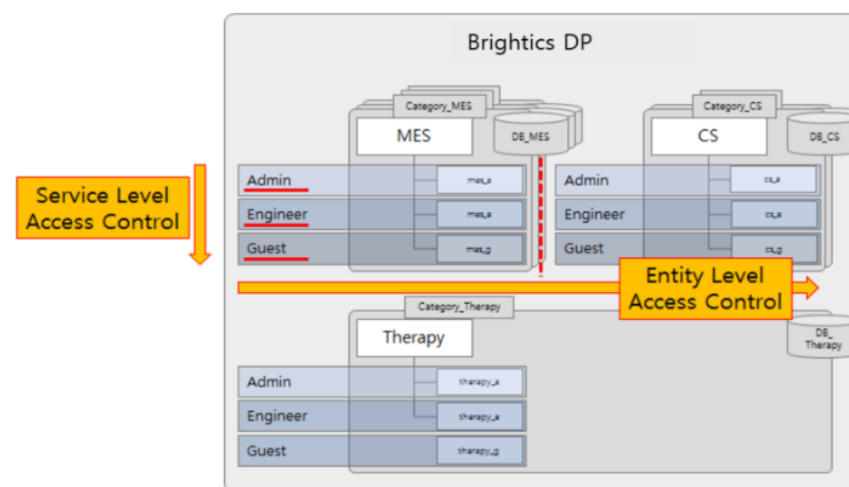


User

Brightics DP : Access Control



사용자/그룹별 데이터/서비스 접근제어



✓ Service Level Access Control

Project 내 Role별 접근 제어로 하나의 Project 안에서 Role에 따라 메뉴 접근부터 특정 화면에서의 행위(신규 생성, 수정, 삭제 등)를 제어할 수 있습니다.

✓ Entity Level Access Control

작업 공간 별 접근 제어로 기본적으로 다른 Project은 작업 공간이 분리된다. 또한 개별 Entity(Template, Category, Feed, Datasource)에 관련된 작업 권한을 지정할 수 있습니다.

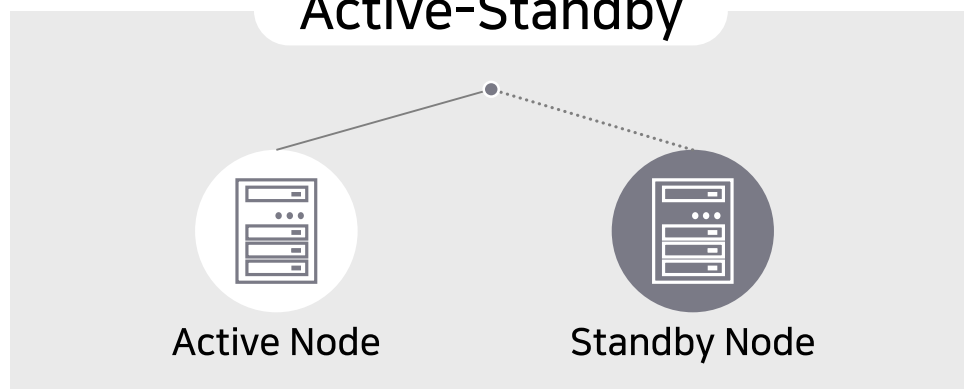
10 HA 구성 기반의 Active - Active Clustering 구조를 지원합니다.

Active-Active Clustering 구조 확보를 통해 High Availability 제공합니다.

Zero Downtime에 가까운 성능 및 Cluster 구조의 Scalability 확보를 통한 고성능 구조를 제공합니다.

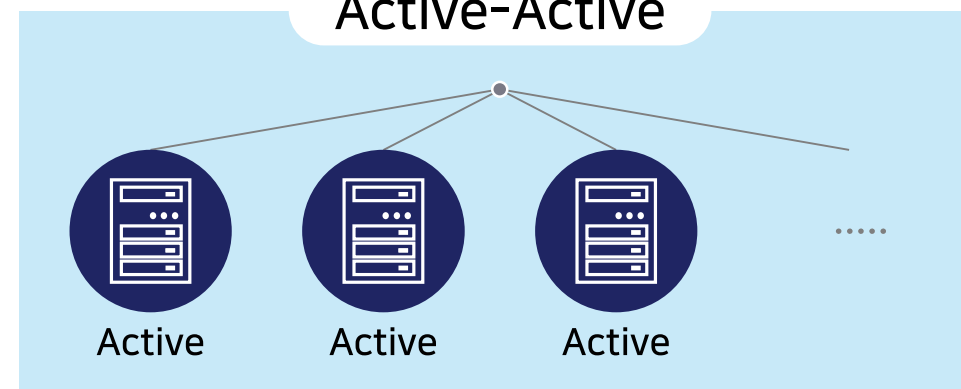
고가용성 확보를 통한 DP 안전성 및 확장성 확보

Active-Standby



- ✔ Active-Standby
- ✔ 장애 복구 시간 약 5분 소요
- ✔ Scale-Out 불가
- ✔ 아키텍처 유연성 ↓

Active-Active



- ✔ Active-Active Clustering
- ✔ 장애 복구 시간 약 20초 소요
- ✔ Scale-Out 가능
- ✔ 아키텍처 유연성 ↑

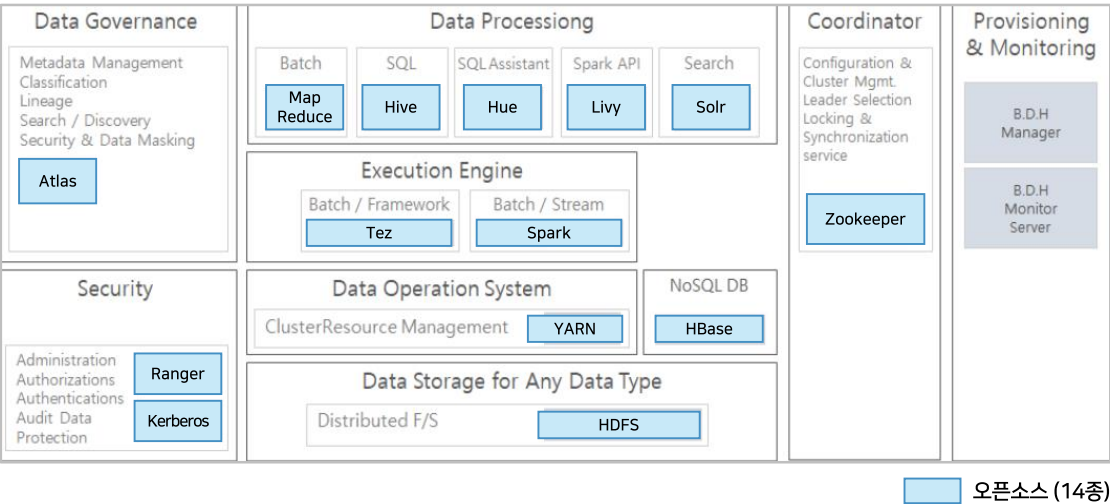
장애 시 복구 시간 약 15배 감소 및 확장성 확보

1 안정적인 Data 분석을 위한 통합 하둡 배포판을 제공합니다.

최신 하둡 기반 Ecosystem 14종의 상호 호환성을 확보하고 검증 완료된 버전으로 패키징하였습니다.
이를 통해 대용량 데이터 저장, 처리, 관리할 수 있는 Framework을 제공합니다.

Brightics Data Hub 구성

- ✓ 데이터를 효율적으로 관리할 수 있도록 컴포넌트를 구성
- ✓ Brightics Hadoop Manager 제공함으로써
통합 모니터링 강화



Brightics Data Hub 특징점

- ✓ 최신 하둡 기반 Ecosystem 14종 Packing
 - 엔터프라이즈용으로 개발된 하둡 오픈소스 플랫폼
 - 14종의 중요한 오픈소스 프로젝트와 Hadoop 통합
- ✓ 대용량 데이터 저장, 처리, 관리할 수 있는 Framework.
 - HDFS와 Hbase를 이용하여 다양한 데이터 저장 가능
 - Yarn을 이용하여 분산 어플리케이션 리소스 관리
 - TEZ, Spark를 사용한 데이터 처리 성능 향상 가능
 - Atlas, Ranger를 이용한 Data Governance, 보안관리
- ✓ Brightics AI와 연동 검증
 - Brightics ML/DP 등과 완벽한 연계 검증

2 편리한 Web Installer 제공합니다.

Web UI 기반으로 서비스별 쉬운 설치가 가능하고, 설치 진행상황과 설치 후 서비스들의 상태를 모니터링 할 수 있습니다.

Web UI 기반의 쉬운 설치

Install

Host Info

Service List

Deployment Status

Configurations

Service List

Total 16

7 items selected

Service Name	Version	Mode	Description	External
☐ Atlas	2.0.0		Atlas is a data governance tool for data governance and metadata management on enterprise Hadoop clusters.	
☒ Base	1.0		Brightics base package (Java / Scala / Hadoop / Host-Monitoring agent).	
☒ HBase	3.0.0		HBase is a distributed column-oriented database built on top of the Hadoop file system.	
☒ HDFS	3.2.0-hc1.1		HDFS is a distributed file system that handles large data sets running on commodity hardware.	
☐ Hive	2.8.6		Hive is a data warehouse infrastructure tool to process structured data in Hadoop.	
☐ Hue	4.6.0		Hue is an open source SQL assistant for querying Databases & Data Warehouses and collaborating. Its goal is to make self service data querying more widespread in organizations.	
☐ Kerberos	5-1.17.1		Kerberos is an authentication mechanism used for strong authentication of Hadoop users and services.	
☐ Livy	0.5.0		Livy is a service that enables easy interaction with a Spark cluster over a REST interface.	
☒ MonitoringServer	3.7.2		Brightics Hadoop service monitoring server. (required)	
☒ PostgreSQL	9.4.14		PostgreSQL is a powerful, open source object-relational database system.	☐
☒ Ranger	1.2.0-hc1		Ranger is a framework to enable, monitor and manage comprehensive data security across the Hadoop platform.	
☐ Solr	7.7.1-hc1		Solr is an open source search platform built upon a Java library called Lucene.	
☐ Spark	2.8.0		Spark is an open source engine developed specifically for handling large-scale data processing and analytics.	
☐ Tez	0.8.1		Tez is a distributed execution framework that works on computations represented as dataflow graphs.	
☒ YARN	3.2.0-hc1.1		Yarn is the resource management and job scheduling technology in the open source Hadoop distributed processing framework.	
☒ Zookeeper	3.4.11		Zookeeper is a distributed, open-source coordination service for distributed applications.	

설치 진행상황

Install

Host Info

Service List

Deployment Status

Configurations

Services

Atlas

Component

Hosts

Atlas

dp03

Base

Component

Hosts

Java

dp03

Scala

dp03

Hosts

dp03

Service

Components

Atlas

Atlas

Base

Java

Scala

HadoopFS

HostAgent

HBase

HBaseMaster

HBaseRegionServer

HemNode

DataNode

설치 진행 현황 모니터링

Install Progress

Host Info

Service List

Deployment Status

Configurations

All

Base

HBase

HDFS

MonitoringServer

PostgreSQL

YARN

Zookeeper

Hosts

Progress

Status

HBase Environment setting

100%

Completed

Service

Progress

Status

Base

100%

Completed

HBase

0%

Installing

HDFS

0%

Installing

MonitoringServer

10%

Installing

PostgreSQL

100%

Completed

YARN

0%

Installing

Zookeeper

0%

Installing

Each step may require several minutes to install.

Console Log

3 서비스별 Configuration 최적값을 추천합니다.

서비스별로 Configuration을 수정하거나 삭제할 수 있고, Version History를 제공합니다.
 사용자가 특정 설정 값을 수정할 경우 관련 되어 있는 다른 설정 값을 자동으로 추천 합니다. (YARN, Hive, HDFS, Spark 등)

서비스별 Configuration 버전 및 History 제공

Dashboard

Installer

Services

Base

HBase

HDFS

Hive

Hue

MonitoringServer

PostgreSQL

Tez

YARN

Zookeeper

Hive

Summary

Config

Output

Installed Date

Service Status

HiveServer2 web

http://dp03:10002

2020-08-05

Running

Hive

Summary

Config

Config Version

Version 2

Version 1

Properties

enable-atlas-hook

meta.datasource.password

metastore.monitoring-port

metastore.uris.port

metastore.warehouse.dir

server2.monitoring-port

Search versions

2020-08-26 18:19:39

2020-08-26 18:14:34

30600

9083

/warehouse/tablespace/managed/hive

30610

Configuration을 수정하면 새로운 Version 생성됨

Configuration 최적값 추천

Config 설정값 변경 시 영향있는 설정값 변경 제한

Hive

Summary

Config

Config Version

Version 1

Properties

enable-atlas-hook

meta.datasource.password

metastore.monitoring-port

metastore.uris.port

metastore.warehouse.dir

metastore.warehouse.external.dir

server2.authentication

server2.monitoring-port

server2.thrift.http.port

server2.thrift.port

server2.webui.binary.port

false

brightics

30600

9083

/warehouse/tablespace/managed/hive

/warehouse/tablespace/external/hive

NOSASL

30610

10001

10000

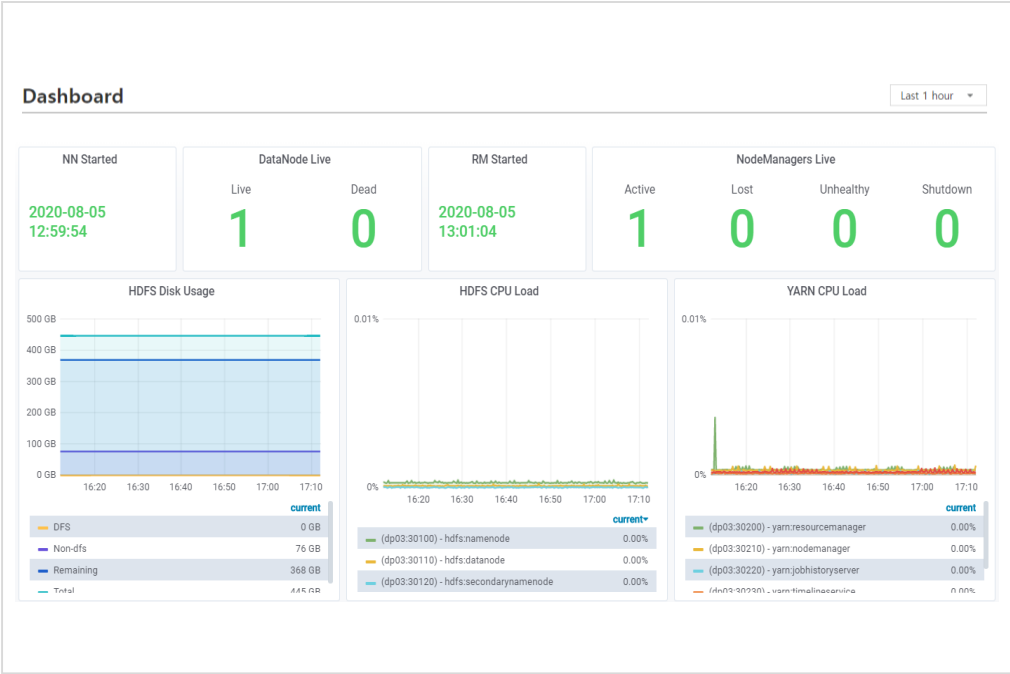
10004

Reset All

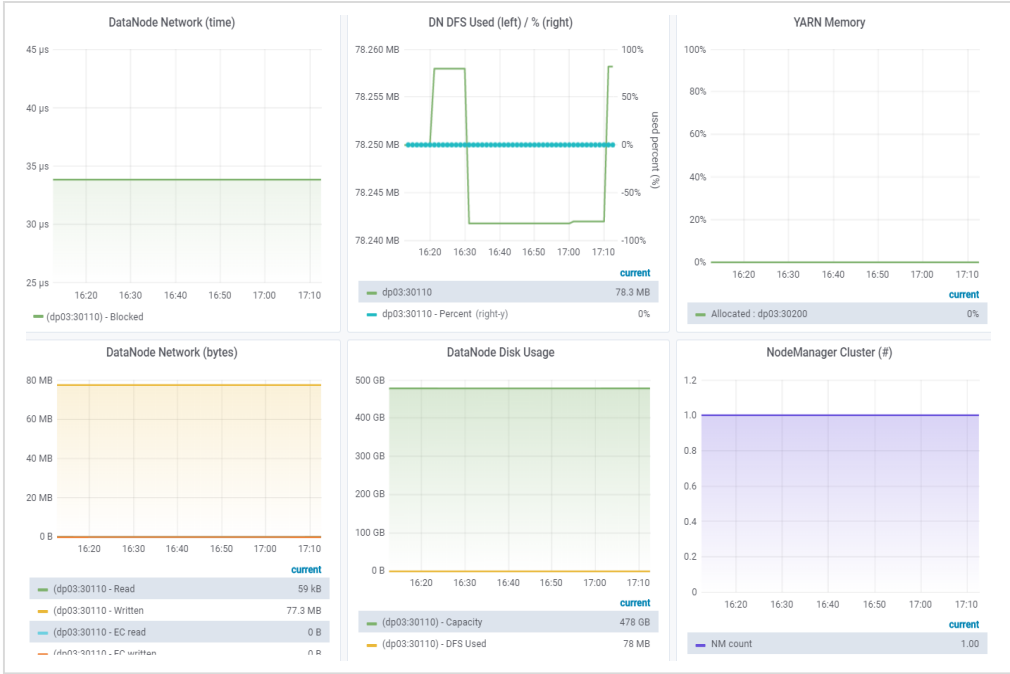
4 시스템 자원의 통합 모니터링 Dashboard를 제공합니다.

시스템 관리자가 대시보드/모니터링 화면을 통하여 시스템의 상태를 확인할 수 있는 다양한 차트를 제공합니다. Hadoop 클러스터 모니터링 Dashboard를 통해 주요 Node 정보 및 Metric 정보를 모니터링을 제공합니다.

상태 모니터링 대시 보드 제공



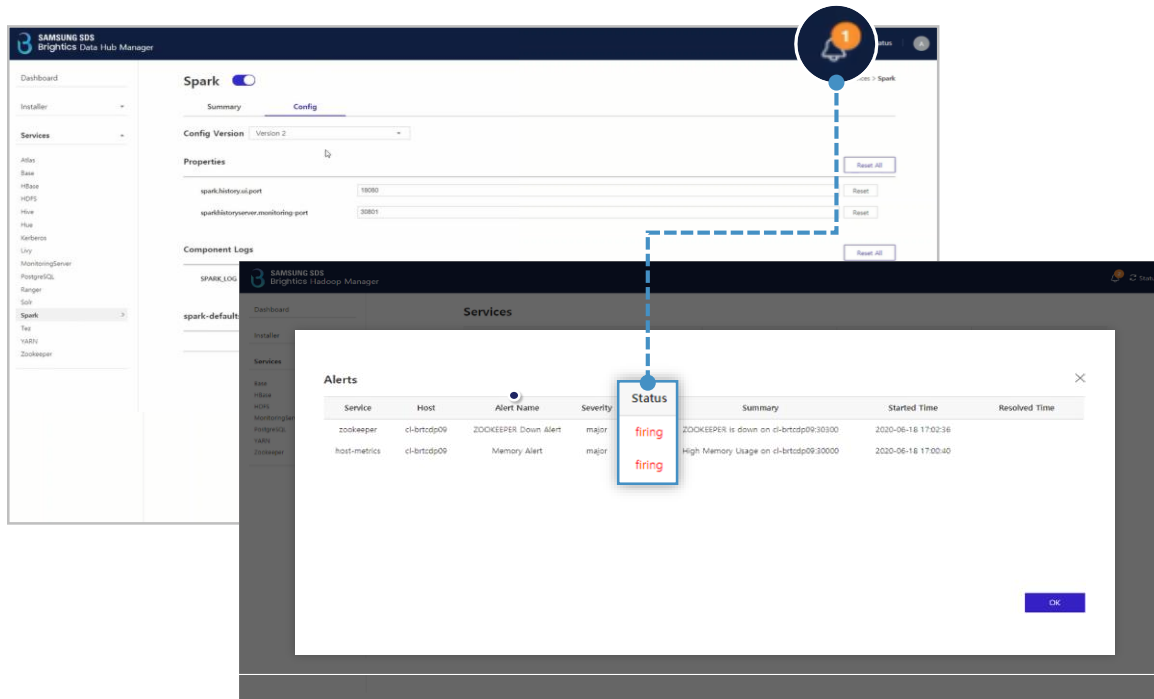
다양한 시스템 리소스 차트 제공



5 서비스 장애 Alert 기능을 제공합니다.

시스템 및 서비스 이상 동작을 감지하여 Warning 메시지를 Push 방식으로 제공합니다.

Brightics Data Hub 구성



Alert이 발생하면 종 모양 아이콘에 숫자가 표시되며
Alert 아이콘을 클릭하여 Alert History 조회

Brightics Data Hub 특징점

- ✓ Alert 종류
 - 서비스 내 Component 단위로 Up / Down Alert 제공
 - Node/HDFS 관련 4종 Alert 추가 제공 (CPU / MEMORY / DISK / STORAGE)
- ✓ 실시간 Push
 - Alert이 발생하면 화면에 실시간으로 Push가 되어 사용자가 즉각 인지할 수 있게 지원

6 Cluster 리소스 관리와 Data Node 추가를 화면에서 할 수 있습니다.

GUI에서 Hadoop Cluster가 사용할 용량을 할당할 수 있어 자원관리를 효율적이고 간편하게 할 수 있으며, Scale-Out을 위한 Data Node 추가는 BDH Manager 화면에서 쉽게 작업 가능합니다.

Hadoop 리소스 관리를 위한 Yarn Queue Manager

YARN Queue Manager

Config Version

Version 2

Edited from service config version 1

Scheduler Properties

maximum-applications

10000

Reset

maximum-am-resource-percent

20

Reset

node-locality-delay

40

Reset

resource-calculator

DefaultResourceCalculator

Reset

queue-mappings

Reset

queue-mappings-override.enable

☐

Reset

Queue

Add

Expand All

Collapse All

root

default

llap

root Capacity

capacity

100

Reset

maximum-capacity

100

Reset

root Properties

state

☒

Reset

user-limit-factor

1.0

Reset

minimum-user-limit-percent

100

Reset

maximum-applications

inherited

Reset

maximum-am-resource-percent

inherited

Reset

priority

0

Reset

maximum-allocation-mb

inherited

Reset

maximum-allocation-vcors

inherited

Reset

Data Node 추가 (Scale-Out)를 위한 BDH Manager

Scale Out

1 Host Info

2 Deployment Status

Identifcation

The Account must have full sudo privileges. The Account info of all hosts must be the same.

ID

brightics

Group

brightics

Identification Type

☒ Password

☐ SSH Private Key

Upload

HA (High Availability)

☐

Host Info

This is not IP Address, it should be Host Name.

Total 3

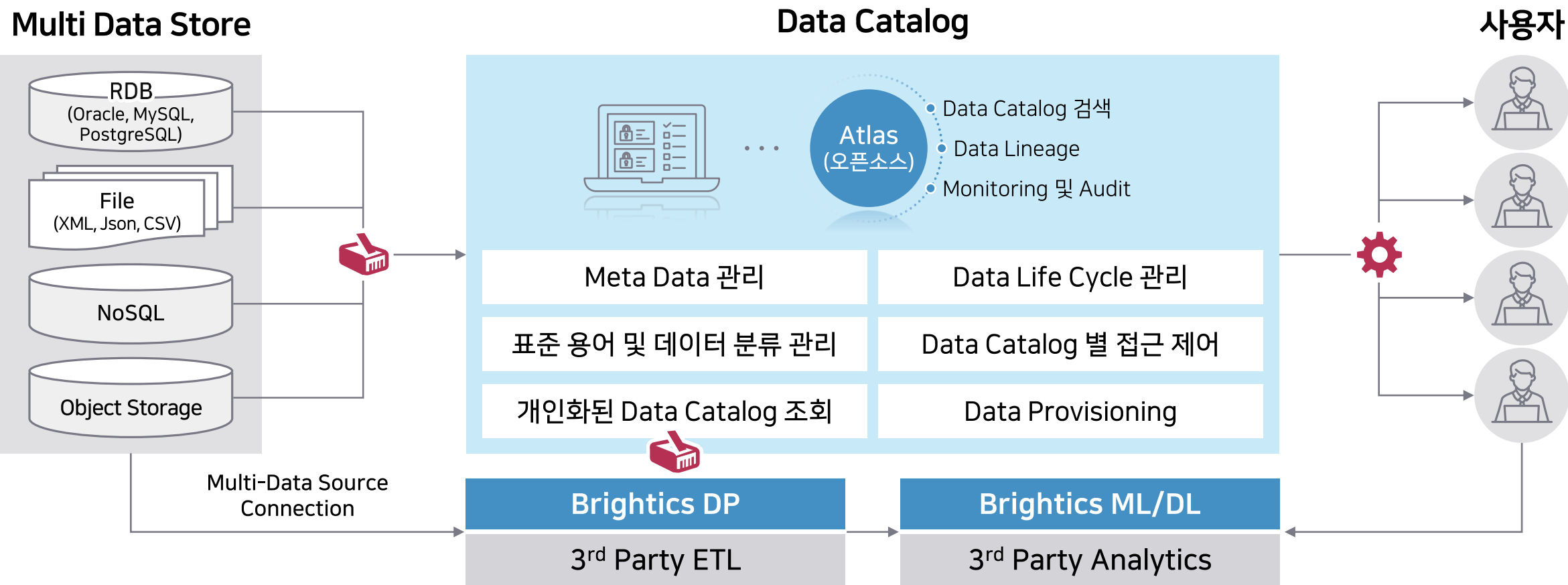
Add

Delete

Host Name	Port	Master	Sub Master	Slave
dp03	22	☒	☒	☒
dp04	22	☐	☐	☒
dp05	22	☐	☐	☒

7 Data Catalog를 통해 조직의 데이터 분석을 가속화합니다.

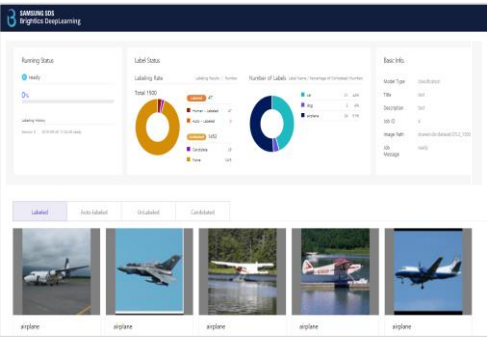
전사적 관점의 데이터 관리를 위한 Meta 데이터 관리, Profile 자동 생성, 데이터별 관한 관리 기능을 제공합니다. 이를 통해 데이터를 효율적으로 탐색하고 활용하여 데이터 분석을 가속화 할 수 있습니다.



1 강력한 정형/비정형 분석 기능을 제공합니다.

정형데이터는 물론, 이미지, Text, file 등 비정형/반정형 데이터에 대한 분석을 지원합니다.
 다양한 소스데이터를 연계하여 분석에 활용 가능한 데이터로 변환 후, ML/DL 을 연동한 AI 모델링을 지원합니다.

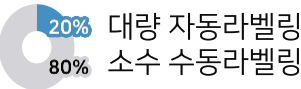
비정형 데이터 처리 및 정제



Data Augmentation

총 48종 기본/랜덤 처리기능
(Crop, Resize, Rotate 등)

Auto-Labeling



Text Processing

- 형태소 분석
 - 유의어 변환
 - 임베딩 기술
 - 키워드 추출
 - Bag Of Words

- Term-Document Matrix
 - Document-Term Matrix
 - Word Count
 - Text Parser
 - (문장부호, 공백제거 등)

정형/비정형 분석 함수 활용

Classification

- Decision Tree
- XGB
- Random Forest
- KNN 등

Clustering

- K-Means
- Hierarchical
- Gaussian Mixture 등

Recommendation

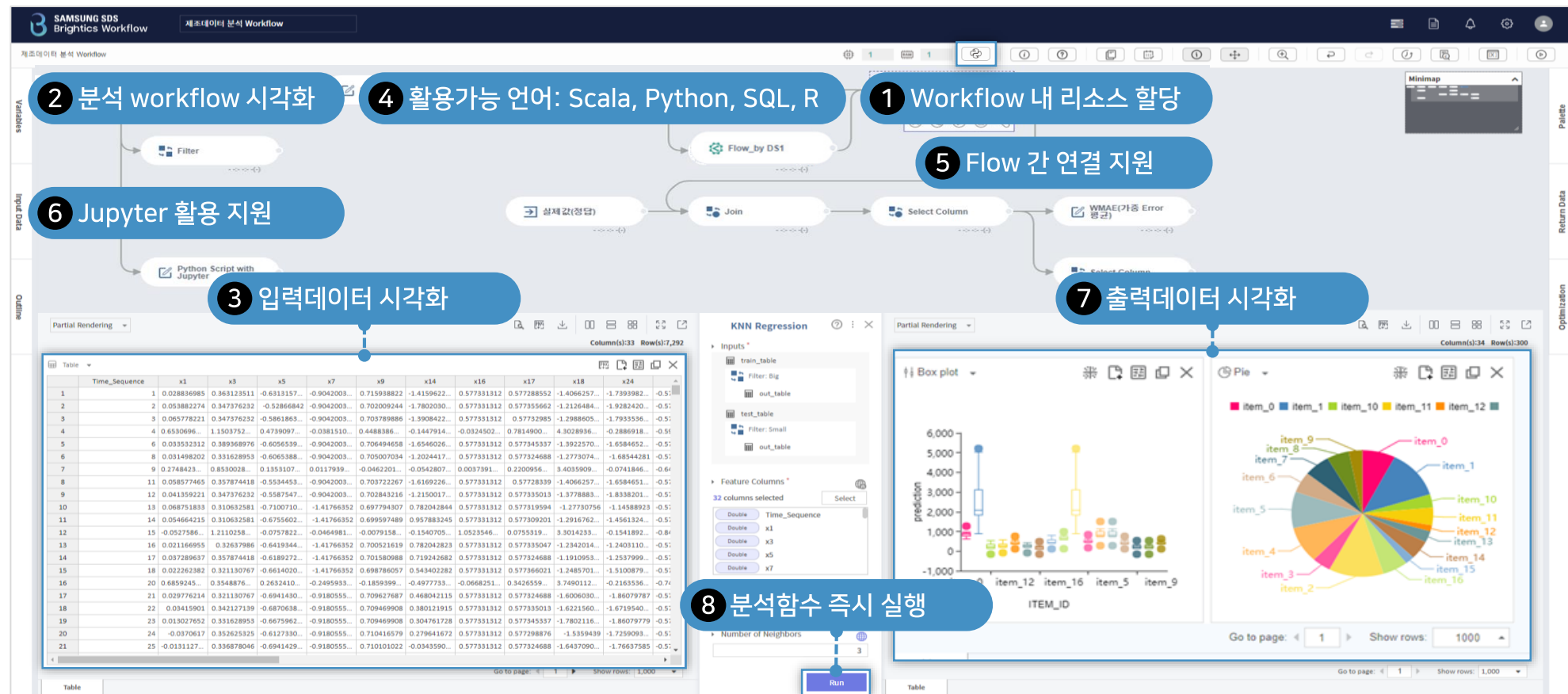
- Association
- ALS
- Collaborative 등

Deep Learning

- Alexnet
- VGG
- Inception V3 등

2 시각화된 분석환경으로 쉽고 직관적으로 활용할 수 있습니다.

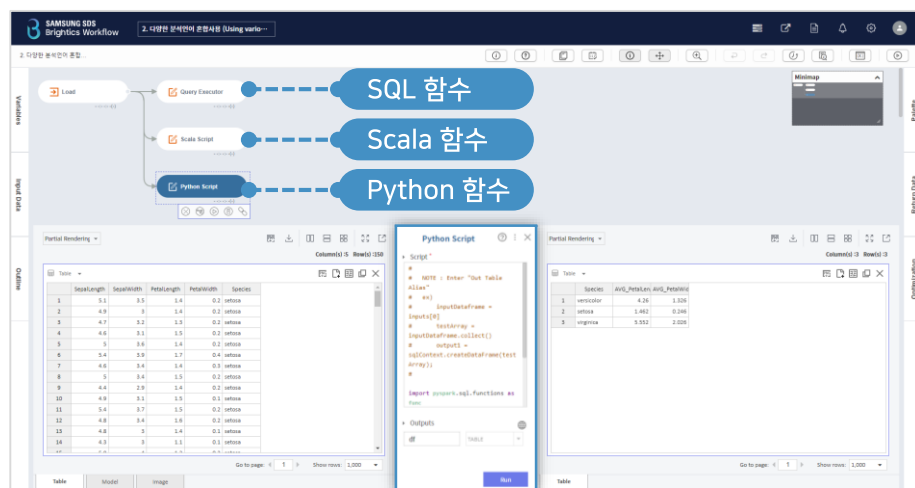
Citizen DS들도 쉽고 편리하게 분석 단계를 구성할 수 있도록, Drag&Drop을 통해 workflow를 제공합니다. 입/출력 데이터와 적용된 알고리즘을 한눈에 직관적으로 확인할 수 있습니다.



3 통합 플랫폼 제공으로 강력한 협업/공유가 가능합니다.

Scala(대용량), Python, R, SQL 등 다양한 언어를 수용하여 하나의 시각화된 분석환경에서 활용 가능합니다. DS들이 각각의 언어로 만든 모델들을 하나의 Flow로 모아 한번에 실행시킬 수 있을 뿐 아니라, 리포트 및 Analytics App을 통해 유기적으로 공유가 가능합니다.

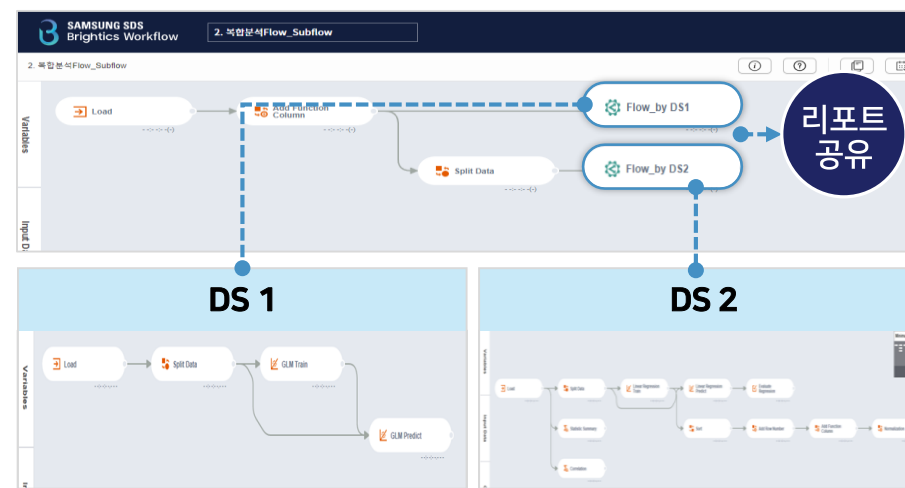
다양한 분석 언어 활용



- Scala, Python, R, SQL을 활용해 동일한 기능을 다양한 언어로 만들 수 있습니다.
- 다양한 언어로 만든 기능을 통합 flow로 연결해 관리할 수 있습니다.



DS간의 협업/공유 가능



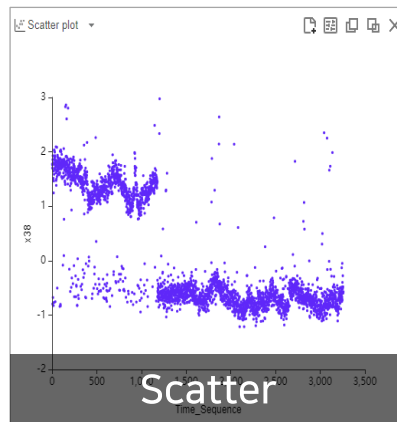
- DS들이 각각의 선호하는 언어로 만든 모델과 결과값을 하나의 Flow로 모을 수 있을 뿐 아니라 리포트 및 Analytics App에서 활용 가능합니다.

4 다양한 시각화 차트를 제공합니다.

입/출력 데이터를 쉽고 빠르게 파악할 수 있도록 Scatter Plot, Heat map, Bar 차트 등 다양하고 활용도가 높은 33종의 시각화 차트를 제공합니다.

type	code	source	year	toe
발견	regen	태양광	2005	14399
발견	regen	태양광	2006	31022
발견	regen	태양광	2007	71279
발견	regen	태양광	2008	284315
발견	regen	태양광	2009	566191
발견	regen	태양광	2010	772801
발견	regen	태양광	2011	917198
발견	regen	태양광	2012	1103227
발견	regen	태양광	2013	1605182
발견	regen	태양광	2014	2556300
발견	regen	태양광	2015	3979159
발견	regen	풍력	2005	129888
발견	regen	풍력	2006	238911
발견	regen	풍력	2007	375641
발견	regen	풍력	2008	436034
발견	regen	풍력	2009	685353
발견	regen	풍력	2010	816950
발견	regen	풍력	2011	862884

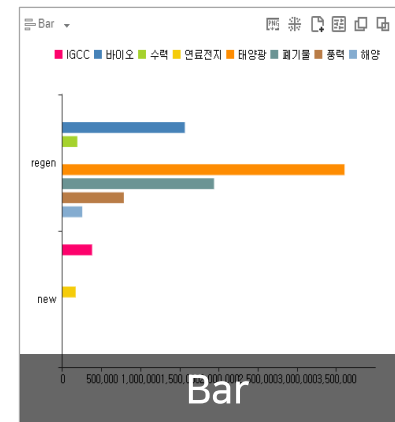
Table



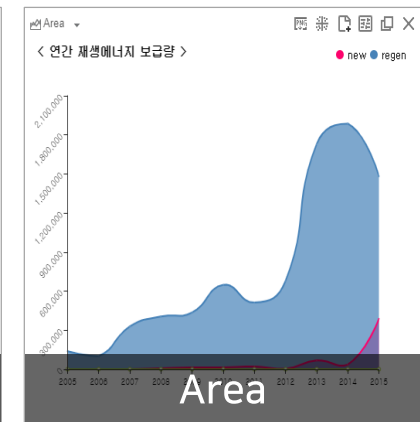
Scatter



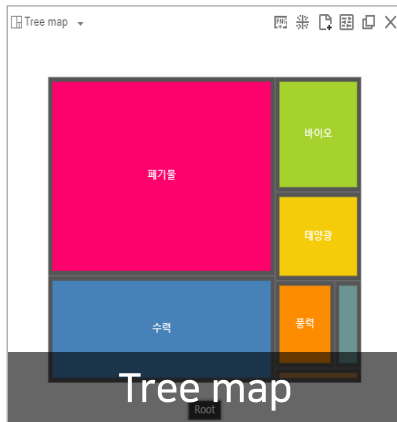
Pie



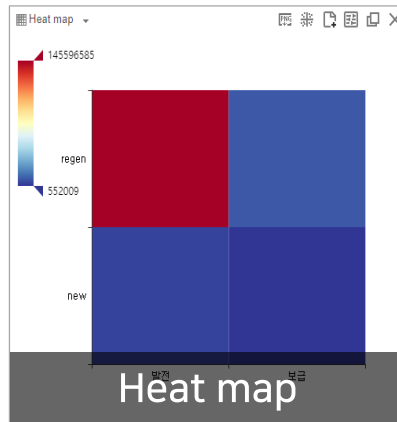
Bar



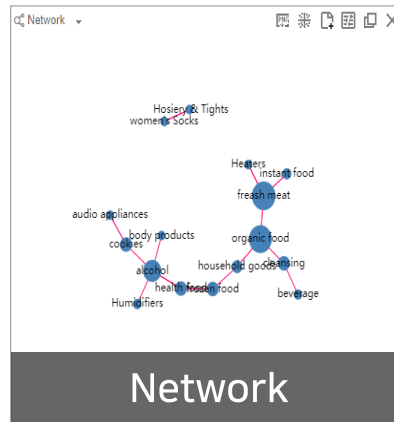
Area



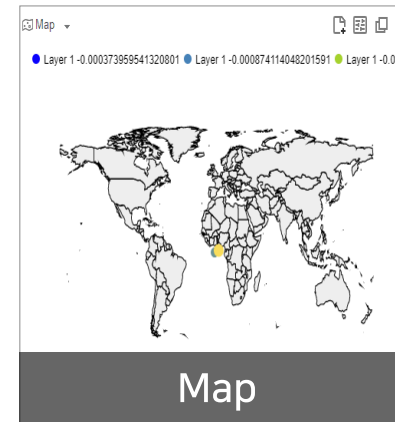
Tree map



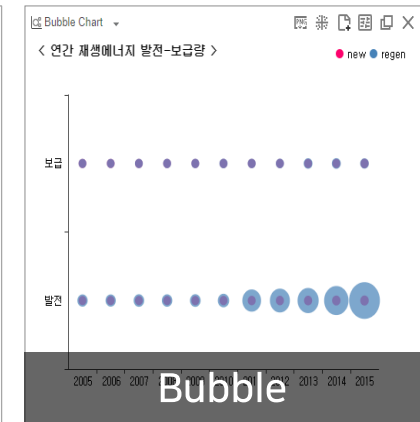
Heat map



Network



Map



Bubble

5 자동화된 분석으로 최적의 분석 모델을 찾을 수 있습니다.

분석 환경과 데이터에 최적의 알고리즘 추천으로 현장 전문가들도 쉽게 모델링이 가능합니다.
최적의 모델 및 학습 인자를 찾기 위해 소요되는 시행착오를 Brightics를 통해서 최소화 할 수 있습니다.

추천 알고리즘과 적합한 파라미터 값 제시

No.	regressor	params	metrics
1	Random Forest	numTrees:29,maxDepth:6,featureSubsetStrategy:all,subsamplingRate:1.0,maxBins...	0.8369521480724534
2	Random Forest	numTrees:16,maxDepth:7,featureSubsetStrategy:sqrt,subsamplingRate:1.0,maxBin...	0.8164499375153857
3	Decision Tree	maxBins:45,impurity:variance,minInfoGain:0.0,minInstancesPerNode:1,maxDepth:8	0.7695649265437133
4	Decision Tree	maxBins:94,impurity:variance,minInfoGain:0.0,minInstancesPerNode:1,maxDepth:6	0.72835193130518
5	Random Forest	numTrees:26,maxDepth:4,featureSubsetStrategy:log2,subsamplingRate:0.7,maxBi...	0.6985091260102894
6	Decision Tree	maxBins:97,impurity:variance,minInfoGain:0.0,minInstancesPerNode:1,maxDepth:5	0.6975167596938445
7	Random Forest	numTrees:18,maxDepth:3,featureSubsetStrategy:all,subsamplingRate:1.0,maxBins...	0.667804478833853
8	Random Forest	numTrees:20,maxDepth:3,featureSubsetStrategy:log2,subsamplingRate:0.7,maxBi...	0.6458478070499631
9	Decision Tree	maxBins:69,impurity:variance,minInfoGain:0.0,minInstancesPerNode:1,maxDepth:4	0.6348005586584676
10	Linear Regression	solver:l-bfgs,elasticNetParam:0.1972,regParam:0.09,standardization:true	0.5956223340066424
11	Linear Regression	solver:auto,elasticNetParam:0.8164,regParam:0.0176,standardization:true	0.5955694374848154
12	Linear Regression	solver:l-bfgs,elasticNetParam:0.0056,regParam:0.0506,standardization:true	0.5955243624931023
13	Linear Regression	solver:normal,elasticNetParam:0.272,regParam:0.0044,standardization:true	0.5954464549310082
14	Linear Regression	solver:normal,elasticNetParam:0.3398,regParam:0.083,standardization:false	0.5954355970065653
15	Decision Tree	maxBins:93,impurity:variance,minInfoGain:0.0,minInstancesPerNode:1,maxDepth:2	0.4247173613278097

추천 알고리즘의 순위

- Auto Data Cleansing
- Auto Feature Selection
- Auto Regression
- Auto Decision Tree
- Auto Classification
- Auto Time Series

20가지 autoML 함수를 제공합니다.

6 자동화된 분석 가이드만 따르면 분석이 어렵지 않습니다.

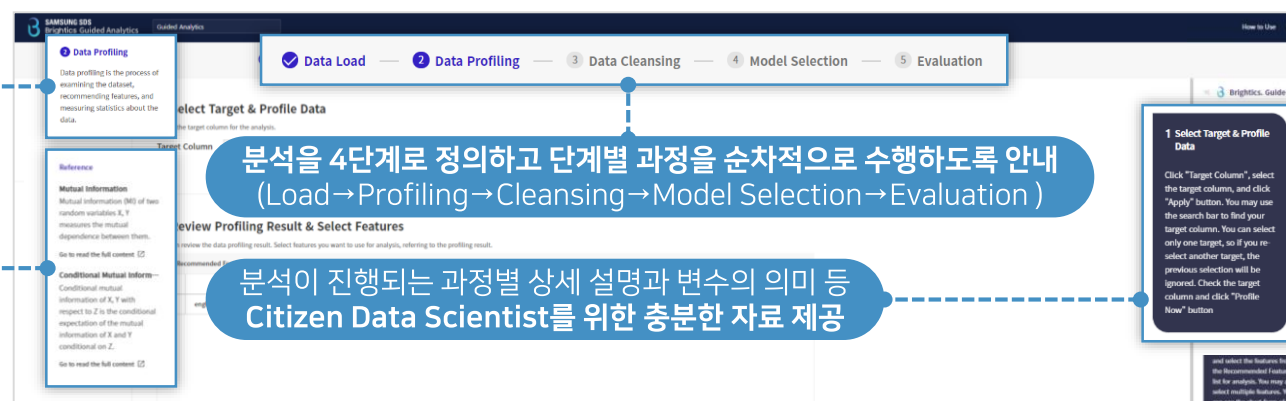
분석의 단계를 가이드대로 하나씩 따라가다 보면, 어려운 분석도 손쉽게 해낼 수 있습니다.

각 단계에서 수행하는 분석의 각 과정을 상세히 설명하고, 친절히 가이드하며 뛰어난 분석 결과를 제공합니다.

Guided Analytics

현재 진행중인 단계에 대한 소개와 목적 안내

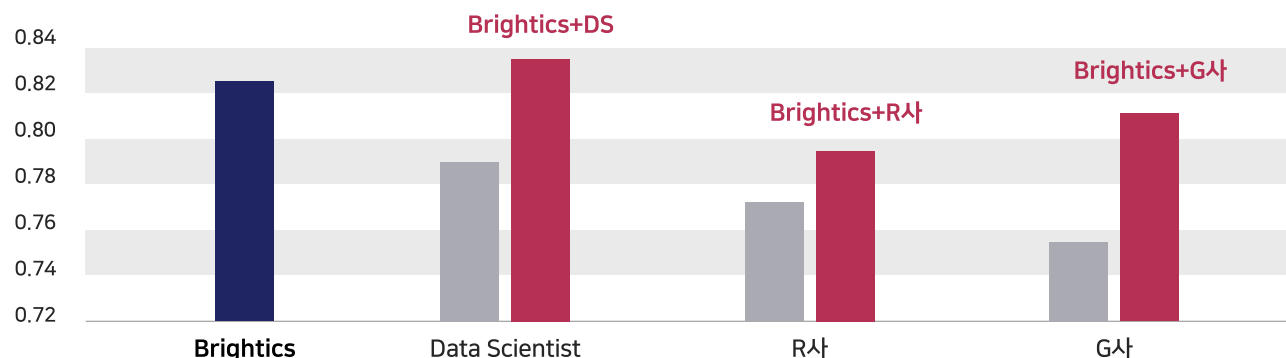
분석 용어에 대한 참조 자료 제공



에너지 수요 예측 분석 모델을 통한 경쟁사 비교

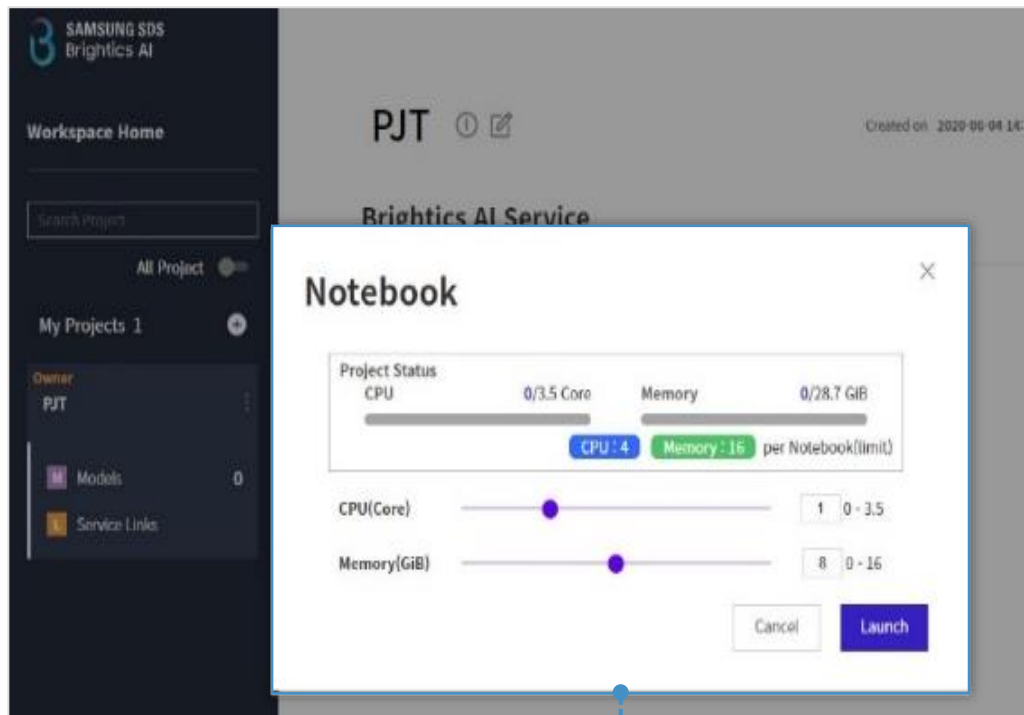
Brightics, DS :
RandomForest 추천
(G사 : Gradient Boosted Trees)

Brightics 추천 Feature +
각 비교 제품의 추천 Parameter 값 조합 시
성능 향상

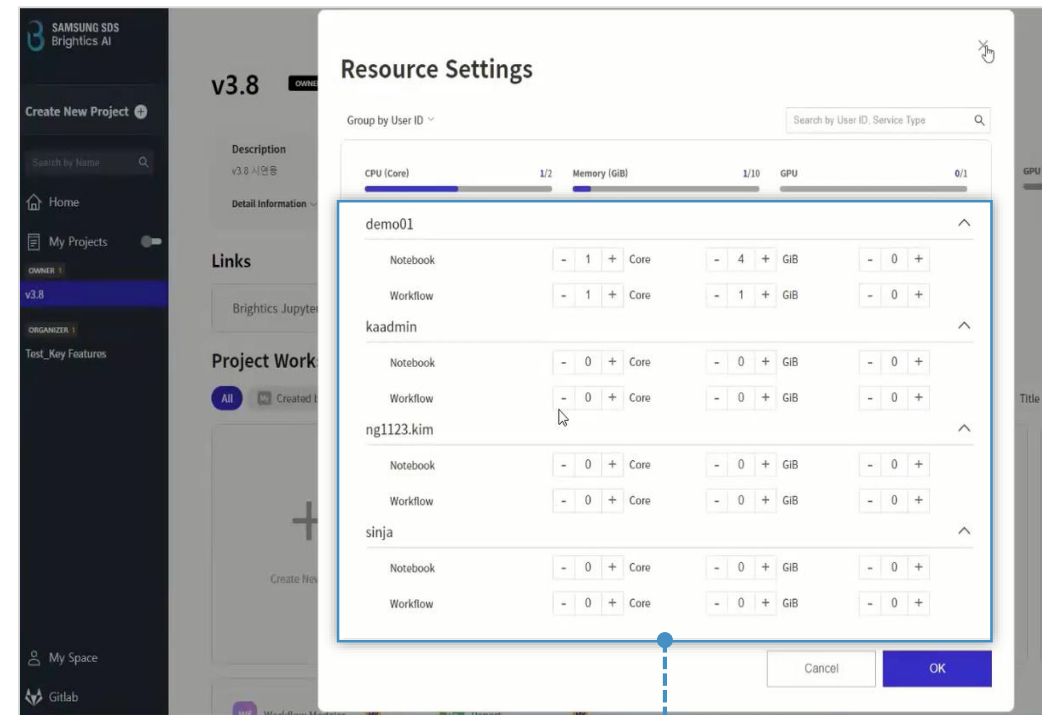


8 자유롭게 개인별 리소스를 할당하고 관리할 수 있습니다.

분석 프로젝트/개인별 컴퓨팅 자원을 UI 기반으로 손쉽게 할당하고 사용할 수 있으며, 관리자는 지속적인 모니터링 통해 체계적이고 효율적인 리소스 관리가 가능합니다.



프로젝트/개인 별 리소스 부여



개인별 리소스 사용현황 모니터링

9 Workflow 내 Jupyter 모델 연동으로 타분석 모델간 협업이 가능합니다.

Professional DS가 Jupyter Lab을 통해 만든 모델을 Brightics Workflow 내에서 즉시 확인할 수 있습니다.
Jupyter Lab에서 분석가별 개인 분석환경을 설정하고 필요에 따라 손쉽게 변경/재구성할 수 있습니다.

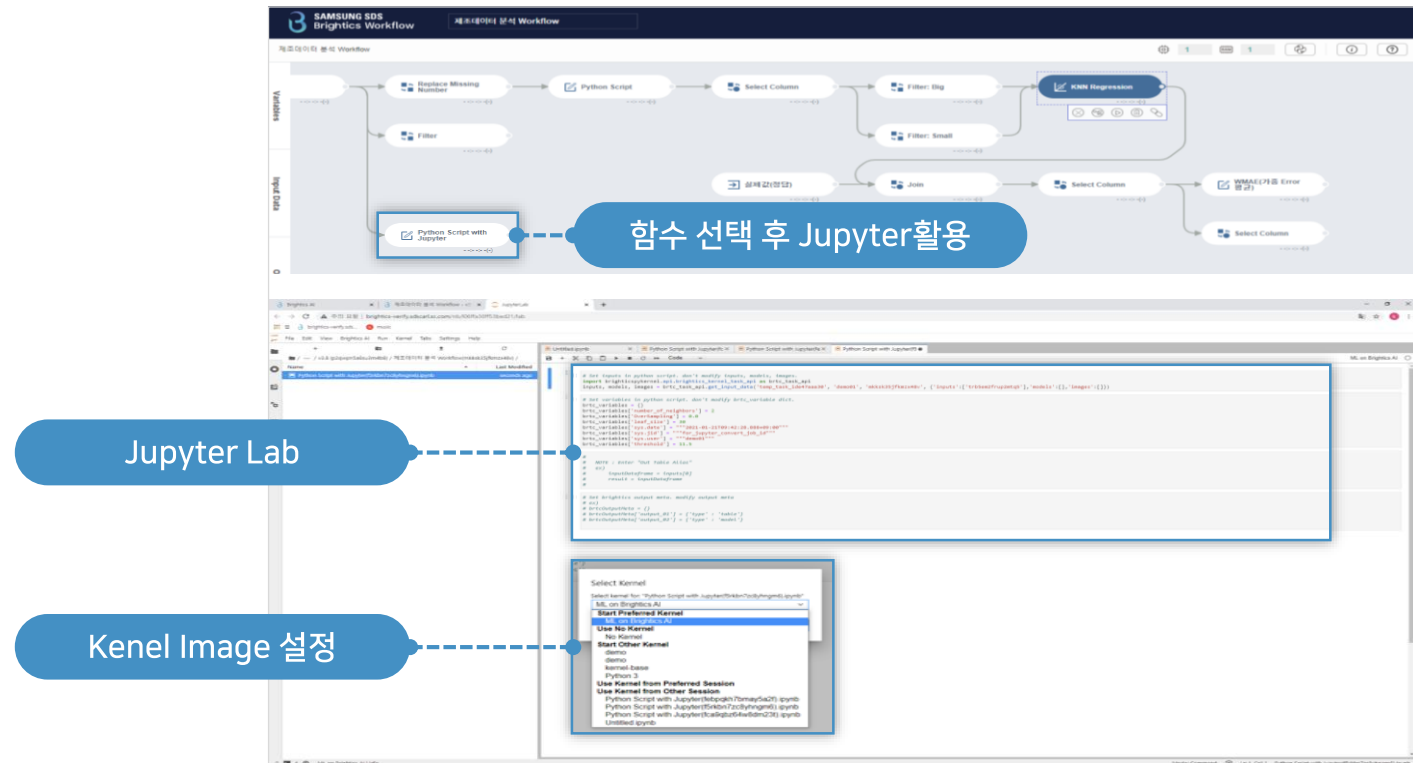
Workflow 내 Jupyter 연동

Workflow 내 Jupyter 연동

- Python Script with Jupyter 함수 제공하여 Workflow 내 Script 작성 가능
- Jupyter Lab을 활용하여 만든 분석모델 Workflow 연동 지원

손쉬운 개인 분석환경 설정

- Jupyter Lab 내에서 전문 분석가가 필요한 개인 분석환경 설정 가능
- 개인별/프로젝트 별 등록되어 있는 Kernel Image 선택하여 필요한 개인 분석환경 구성
- 프로젝트별 공유되어 있는 Shared Kernel Image 설정하여 공동 분석환경 즉시 구성 가능



10 여러 곳의 서버로 모델을 배포하고 관리할 수 있습니다.

분석 모델의 배포 기능 추가로 다양한 환경에 동일한 모델을 적용할 수 있습니다.
중앙 집중형 모델 관리 및 모니터링을 통한 의사결정을 지원합니다.

Model Manager

Configuration

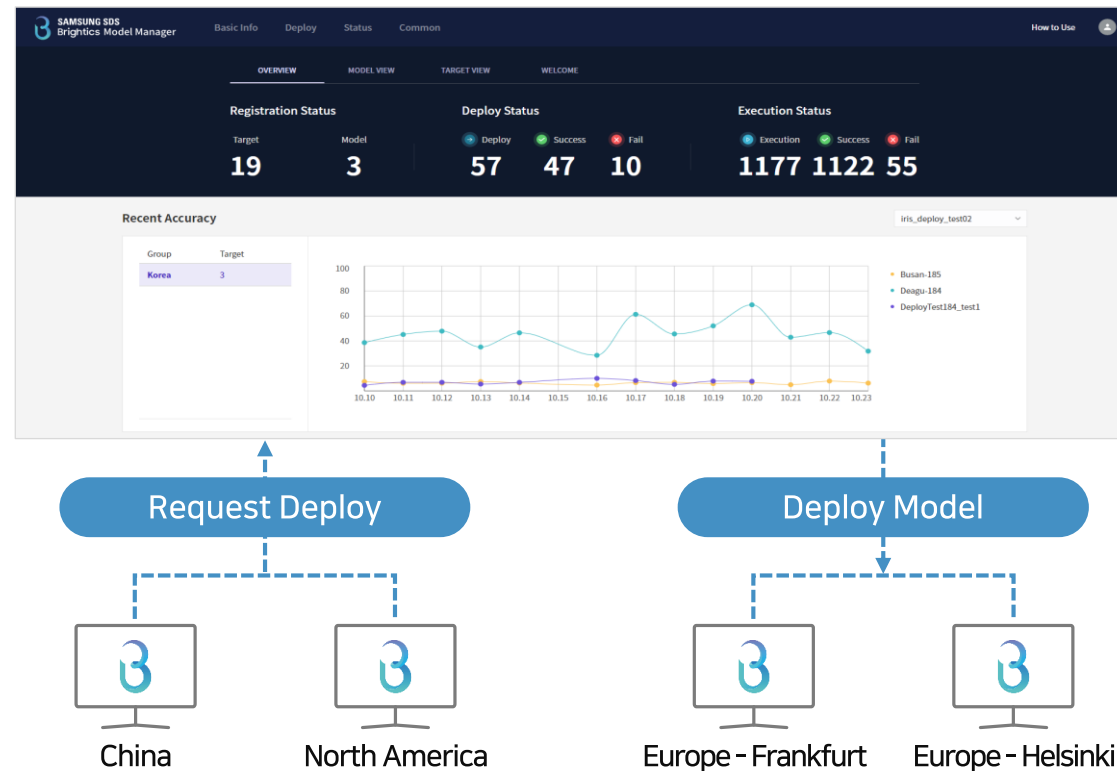
- 배포 Target에 대한 System Management 수행

Deploy Management

- Deploy Type (Batch, API)
- Run Configuration (Batch Cycle, API URL)
- Result Configuration (Accuracy Formula)
- Upgrade Model
- Deploy Version Management

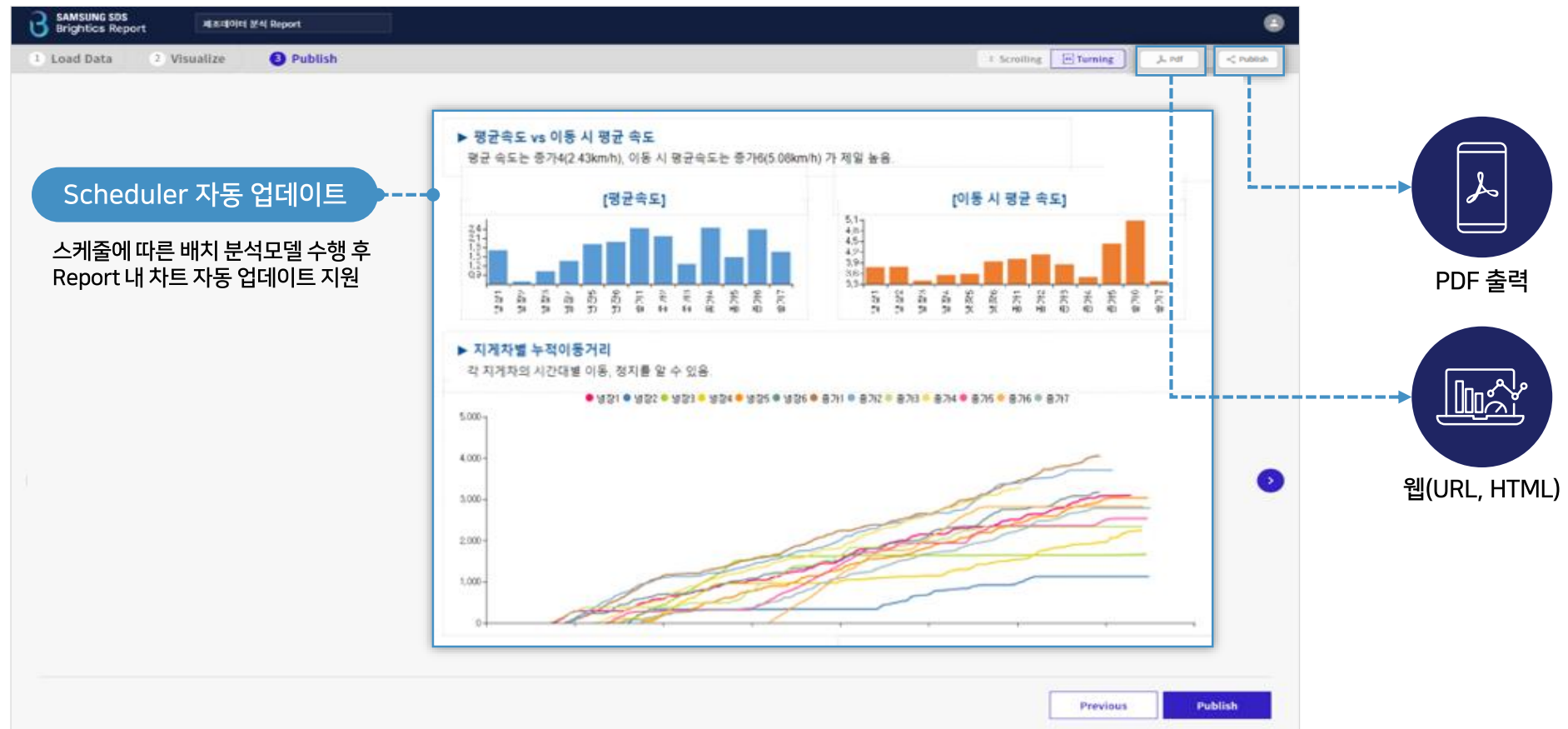
Monitoring

- 배포된 분석 모델의 현황 제공
- 모델의 정확도 추이 모니터링 제공
- 모델 History View 제공



11 분석모델의 결과를 손쉽게 리포트로 제작 할 수 있습니다.

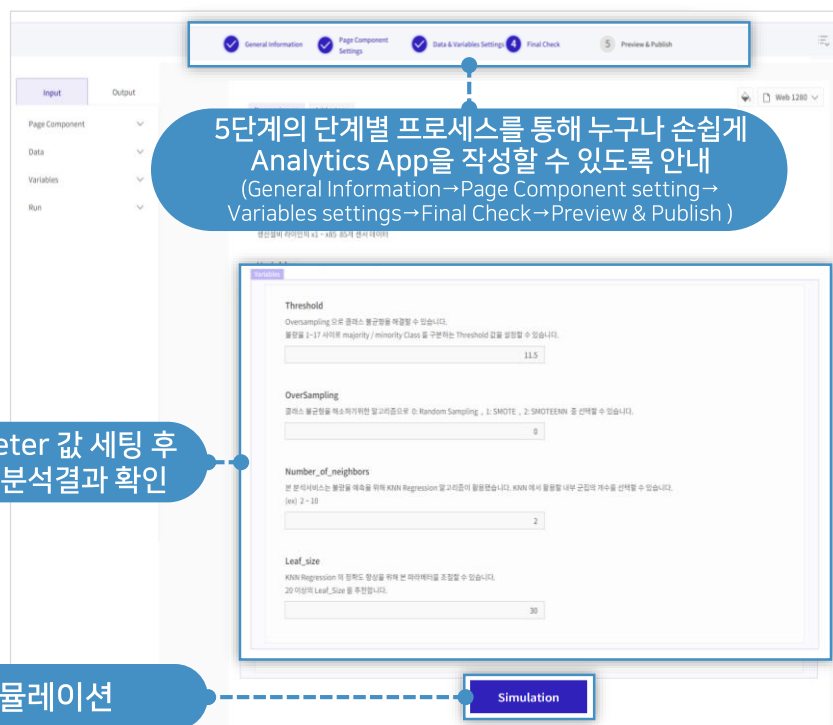
간편한 리포트 편집 기능과 웹/모바일을 통한 분석 리포트 배포 기능을 제공합니다.
스케줄에 따른 배치 분석모델 수행 후 Report 내 차트 자동 업데이트가 가능합니다.



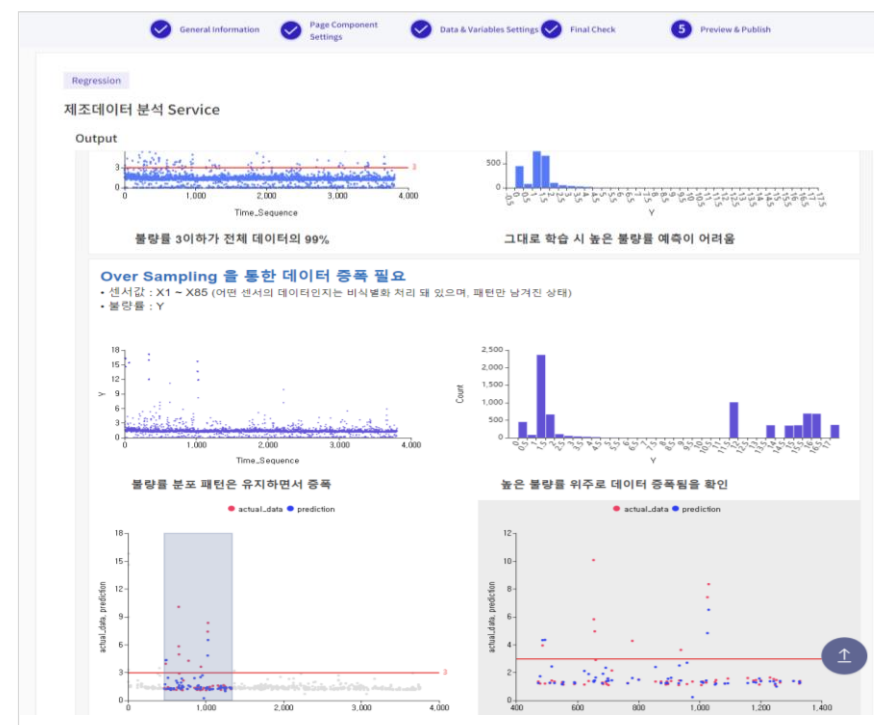
12 Analytics App을 통해 언제든지 필요한 분석 결과를 확인할 수 있습니다.

분석 모델에 대해 필요한 Parameter 값을 설정하여 언제 어디서든 사용자가 원하는 분석 모델 결과값을 Analytics App을 통해 확인할 수 있습니다.

Parameter값 조정으로 시뮬레이션 조건 설정



설정값에 따른 시뮬레이션 결과

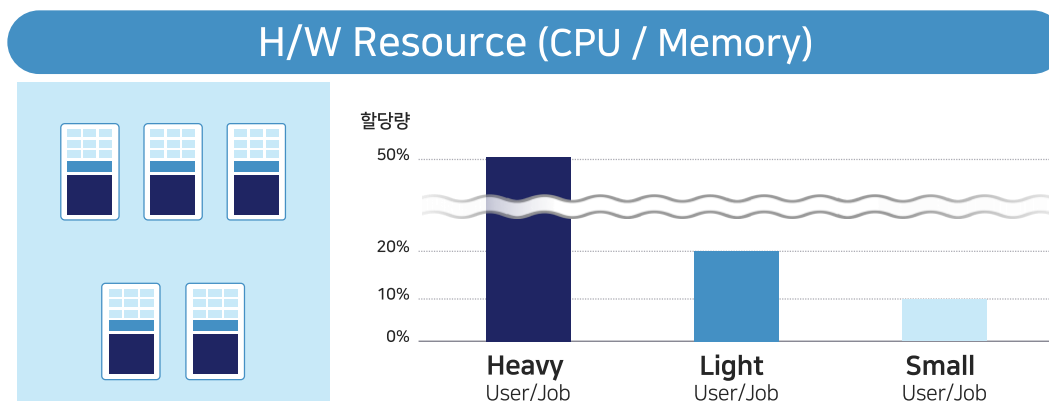


13 효율적인 리소스 관리 기능을 제공합니다.

리소스의 효율적 관리를 통해 안정적인 시스템 환경을 제공합니다.

사용자가 Agent List에서 Job의 Status를 확인해, 최적의 리소스 활용이 가능한 Agent를 선택할 수 있습니다.

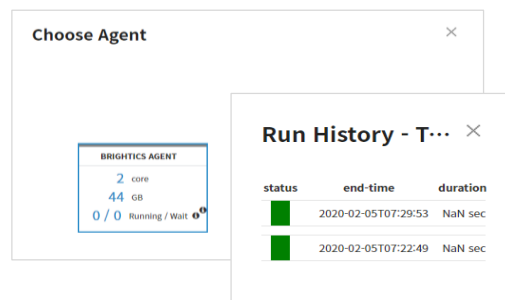
리소스 관리



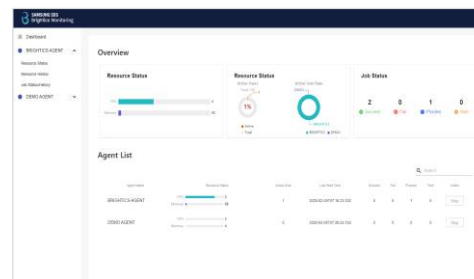
✔ 유연한 운영 환경 제공

- Hadoop/Spark의 전체 리소스를 작업별/개인별로 유동적 할당하여 높은 안정성과 유연성을 제공합니다.

Agent 관리



Agent 별 Job Status 관리



Resource Status 관리

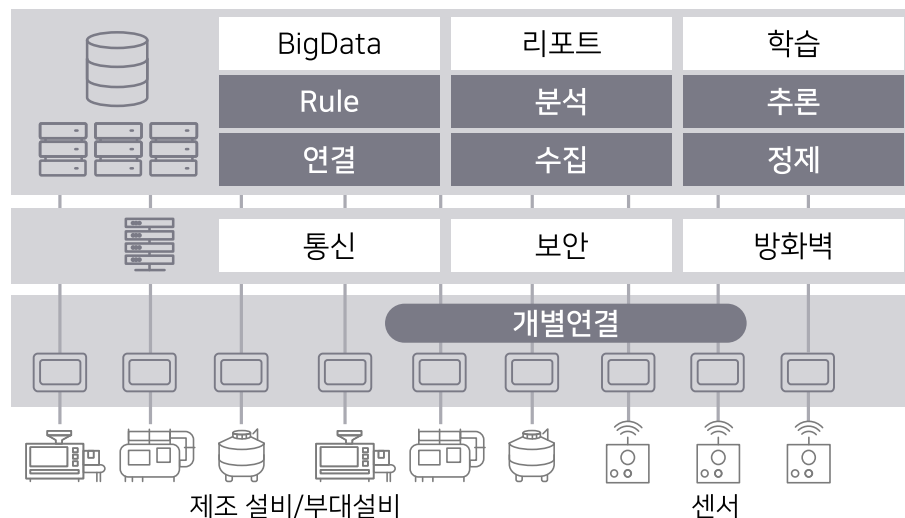
✔ User가 Agent를 선택

- Agent List에서 Agent별 Job/리소스 Status 확인 가능
- 최적의 효율을 기대하는 Agent를 직접 선택해 Job 할당

14 Edge Analytics 기술로 분석 처리속도를 향상시킵니다.

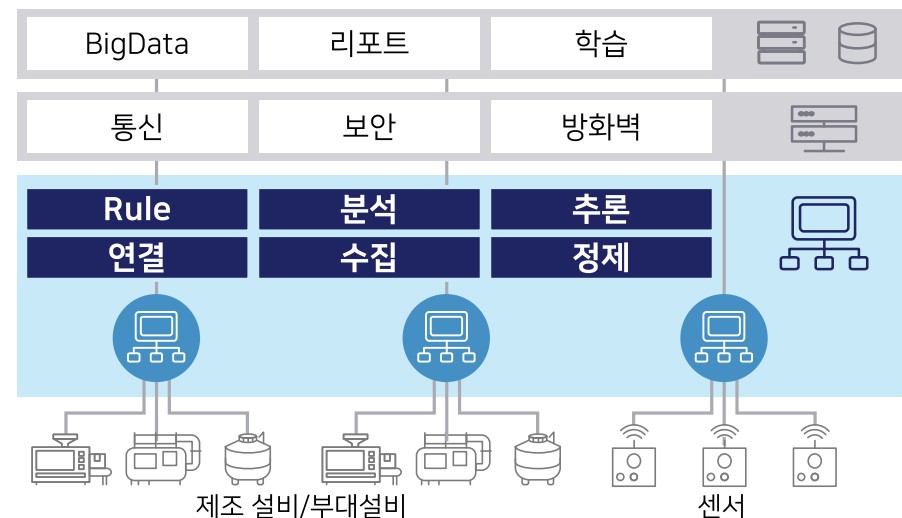
Brightics Analytics Edge를 통해 디바이스 연결 기반의 다양한 데이터를 실시간으로 수집/처리/분석하여 데이터 활용 효율을 높이고 Edge Manager를 통해 모니터링 및 관리 기능도 제공합니다.

기존 방식



- ✓ 현장 설비와 센서를 개별적으로 연결
- ✓ 서버 단에서 데이터 정제와 분석 수행
- ➡ 구축 비용 등의 이슈로 중요설비에 국한하여 활용

Edge Analytics



- ✓ Edge AI Analytics 기반의 실시간 이상감지 및 예지보전
- ✓ 대용량 데이터의 사전 정제로 네트워크 및 시스템 H/W 비용 감소
- ✓ 전설비/센서로 확장하여 체계적인 통합 관제
- ✓ Edge Manager 통해 분석 수행을 위한 관리 UI 지원 및 Edge App/Module 모니터링/관리 기능 제공

15 검증된 BP 기반의 Pre-Built를 활용할 수 있습니다.

현장에 적용되어 이미 검증된 분석 모델을 바로 활용해 보실 수 있습니다

업종별, 분석모델 유형별로 제공되는 Pre-Built 모델을 통해 빠르게 Insight를 확보 할 수 있습니다.

집계형, 예측형, 처방형분석 제공

✔ 머신러닝, 딥러닝, 강화학습 지원



다양한 Pre-Built 분석모델 적용범위

✔ 산업별, 프로세스별 분석모델 110개



분석프로세스 별 분석모델 Tutorial 제공

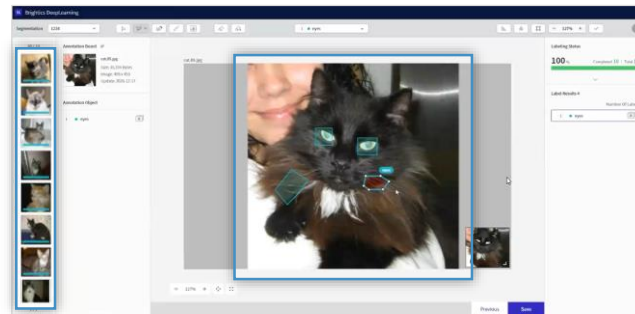
Common	Pre-processing	EDA
Clustering	Regression	Classification
Recommendation	Time Series	Text Analysis

1 높은 정확도의 이미지 데이터 전처리 자동화를 지원합니다.

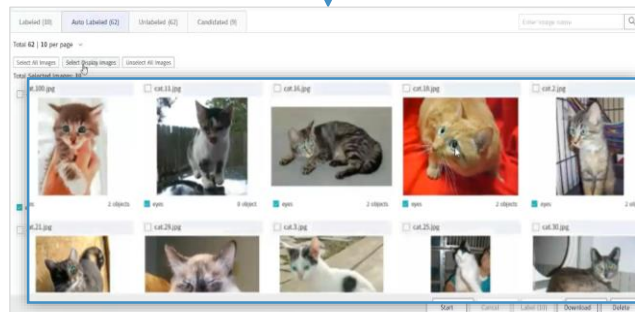
현존하는 Auto Labeling 기술 중 세계 최고 자동화율 수준을 달성하였으며,
업계 최고 수준인 98.1%의 자동 라벨링 정확도를 제공하여 AI 모델 개발 생산성을 가속화 합니다.

Active Learning 기술로 전체 이미지 중
20% 수동 라벨링 후 80%의 이미지 자동 라벨링하여
시간 70~80% 절감

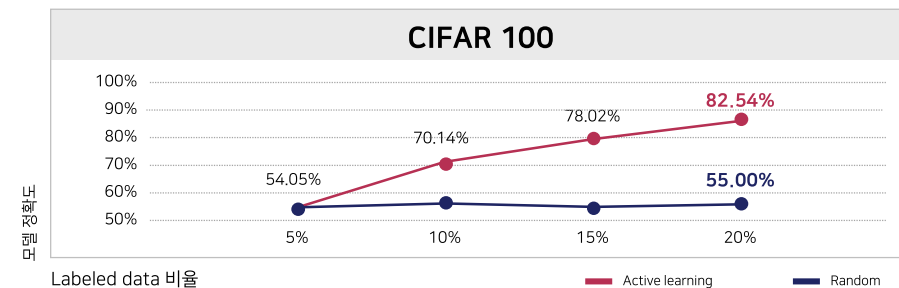
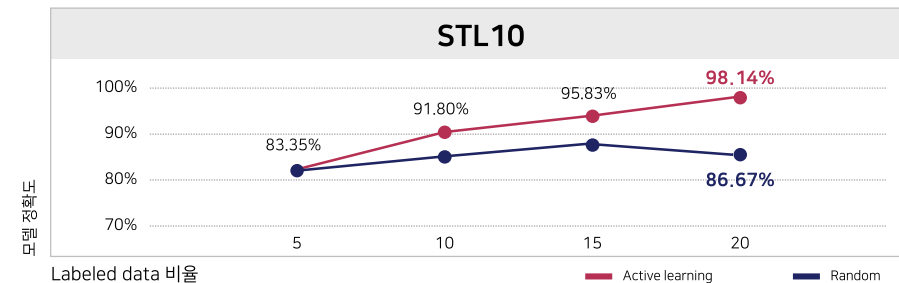
- 학습효과가 큰 이미지
20% 선별하여
수동 라벨링 진행



- 나머지 80% 데이터를
자동 라벨링 함



공개데이터(STL10) 80%에 대해
Classification Auto Labeling 결과
98.1%의 높은 정확도 달성

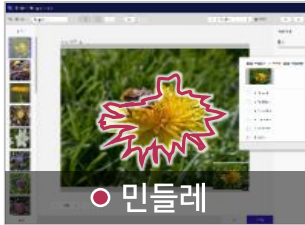


- A사 제조 데이터, Detection Auto Labeling 정확도 80.63% 달성

2 글로벌 수준의 확장된 자동 라벨링 기능을 지원합니다.

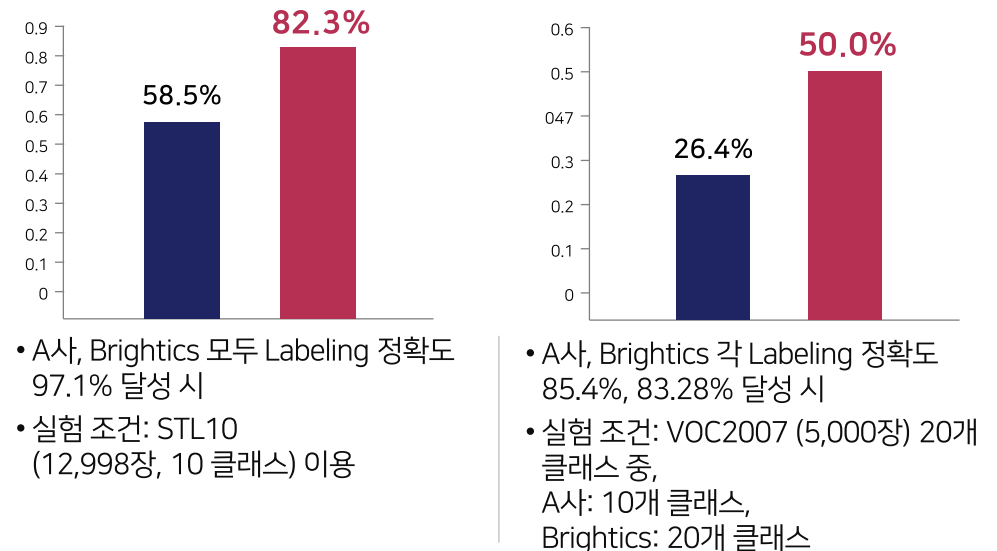
이미지 분석 주요 분야인 Classification, Detection, Segmentation 3개 분야에 대해 타사 대비 높은 자동화율의 Auto Labeling 기능을 제공합니다.

Classification, Detection, Segmentation 3개 분야 지원

Classification (분류)	Detection (검출)	Segmentation (분할)
		
• 이미지 안의 객체의 종류를 분류하여 라벨링	• Object를 Boxing하여 라벨링	• 픽셀단위로 라벨링

이미지 자동 라벨링 3개 분야 확장으로
자동 라벨링을 통한 이미지 분석을 활용할 수 있는 산업 분야가 확대됨

A사 대비 최대 1.8배 높은 자동화율 제공



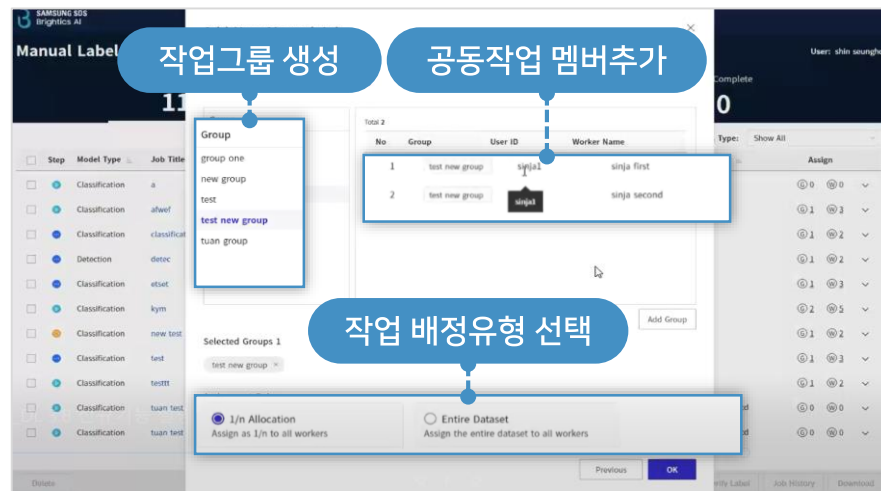
Classification (분류)
Brightics 자동화율 1.4배

Detection (검출)
Brightics 자동화율 1.8배

3 통합 라벨 매니지먼트 기능을 통해 라벨링 협업이 가능합니다.

이미지 분석 사전작업을 위한 라벨 매니지먼트 및 작업공유 기능을 통해 라벨링 협업이 가능하여 보다 빠르고 효율적인 작업이 가능합니다.

이미지 라벨링 협업 가능



- 이미지 분석 사전 작업을 위한 라벨링 작업에 많은 시간이 소요되는 문제 해결
- 라벨 작업을 생성하고 함께 작업할 멤버를 직접 구성하여 필요 시 언제든지 라벨링 작업 협업 가능



라벨링 공동작업 관리기능 강화



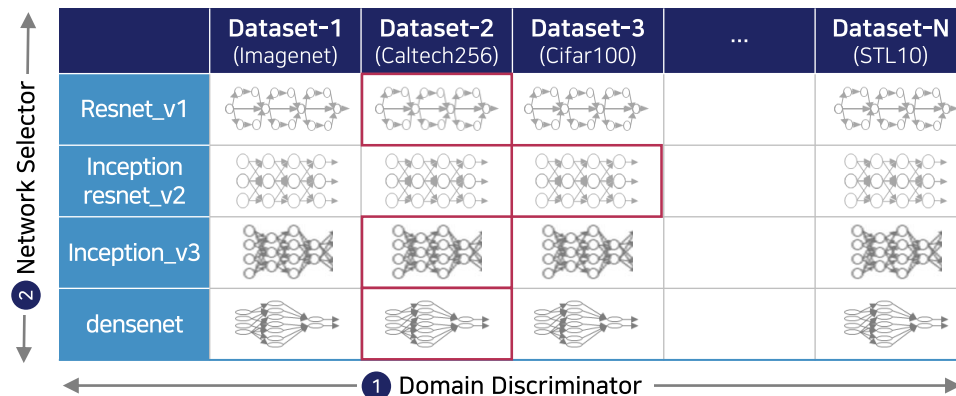
- 관리자 권한 부여를 통해 모델별 라벨링 공동작업 진행현황 수시 모니터링 가능
- 할당된 개인 worker별 작업 진행현황 확인가능

4 학습 데이터에 최적의 딥러닝 모델을 자동 추천합니다.

세계 최고 수준의 Pre-trained Model 기반의 모델 추천 기술을 세계 최초로 상용화하였습니다. (핵심특허 출원)
전체 Network Search가 없는 획기적인 입력데이터 기반 최적모델 선정방식으로 IT자원 효율을 극대화하였습니다.

학습 데이터에 최적인 딥러닝 모델 자동 추천

- 非숙련 SI개발자들이 딥러닝 모델 선정에 많은 시간을 소비하는 문제 해결
- Pre-trained 모델을 활용해서, 모델 선정속도 최소 2~8배 향상



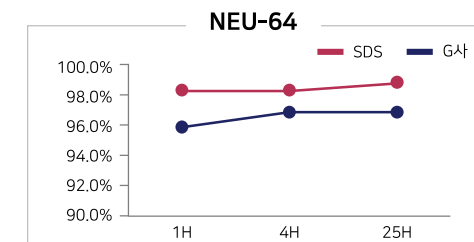
- 1 입력데이터와 가장 유사한 사전학습 데이터 탐색
- 2 입력 데이터로 학습 시 성능이 가장 높게 예상되는 모델 추천

대용량 학습 데이터에도 G사 AutoML 보다 정확도는 높고, IT자원 소비 효율은 9배 우수

- Neural Network Architecture Search 방식 대비 적은 IT자원 소비
- 제조업에 강한 Pre-trained model 28개 보유
- ICIP 2019 논문 채택

NEU-64-CLS

	SDS BMP	G사 AutoML
1H	98.06%	96.26%
4H	98.20%	97.20%
25H	98.60%	97.40%



*NEU-64-CLS dataset : Surface defect classification, 9 classes, Factory dataset과 유사한 형태의 public dataset

SAMSUNG SDS
Realize your vision