



# A new open source BI platform powered by Druid

Metatron Project

## Contents

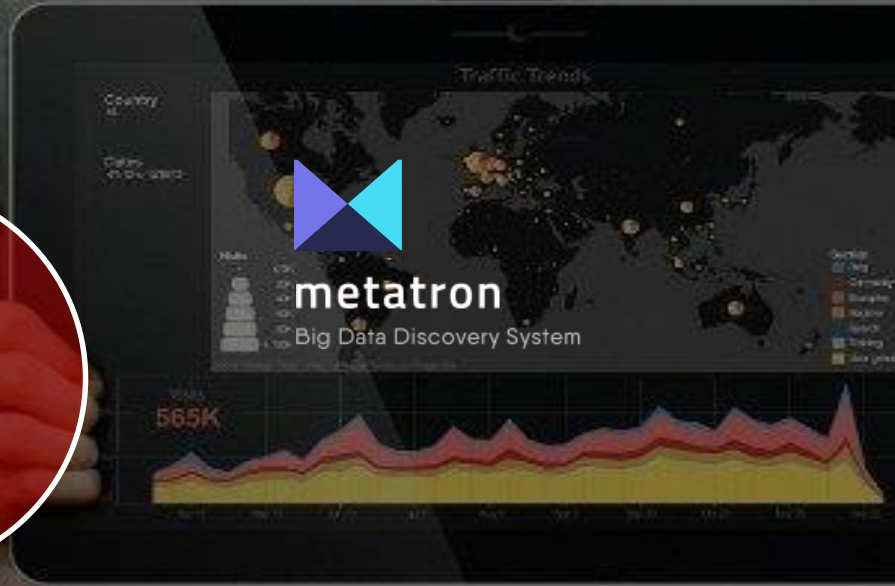
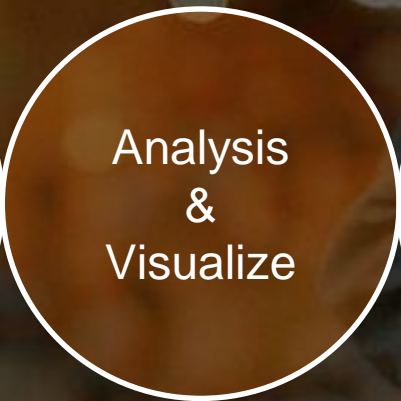
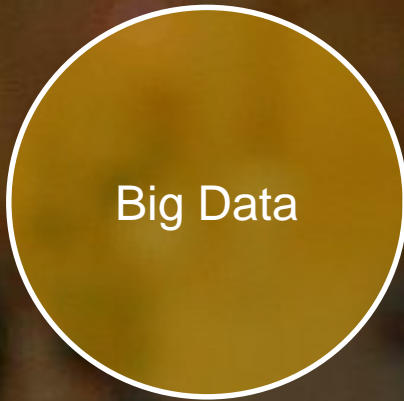
1. Overview
2. Metatron 2.0
3. Use case
4. Technology
5. Roadmap

metatron

# Overview



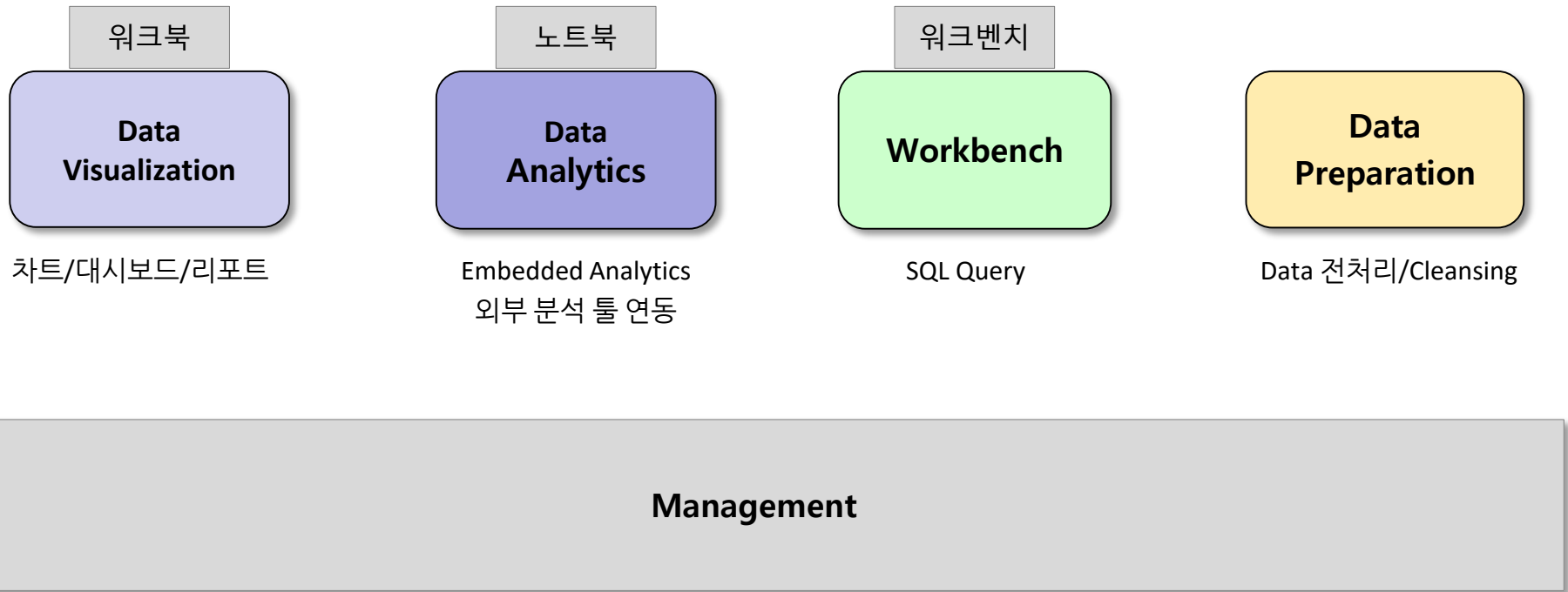
Self-Service Data Discovery & Analytics



*OLAP, 시각화, 머신 러닝이 융합하여 비전문가도 데이터로부터 상위 레벨의 가치를 손쉽게 얻을 수 있는 Intelligence 솔루션*

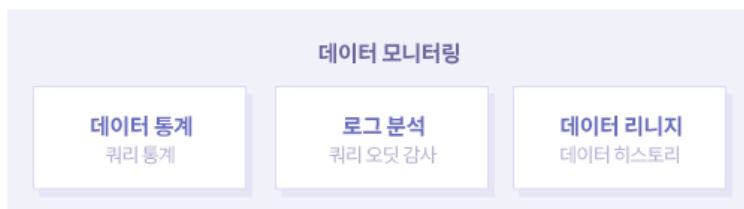
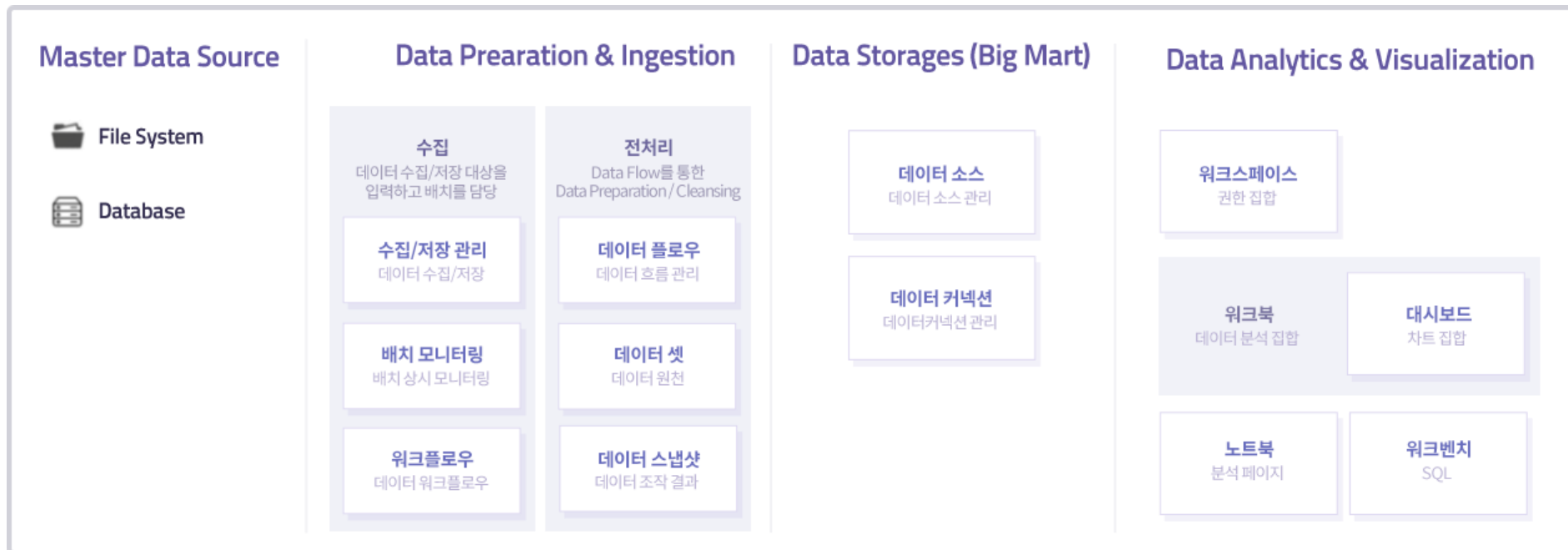
# Single Solution for Data Discovery

대용량 데이터 Preparation 부터 시각화 기반 데이터 탐색, 분석까지 한번에



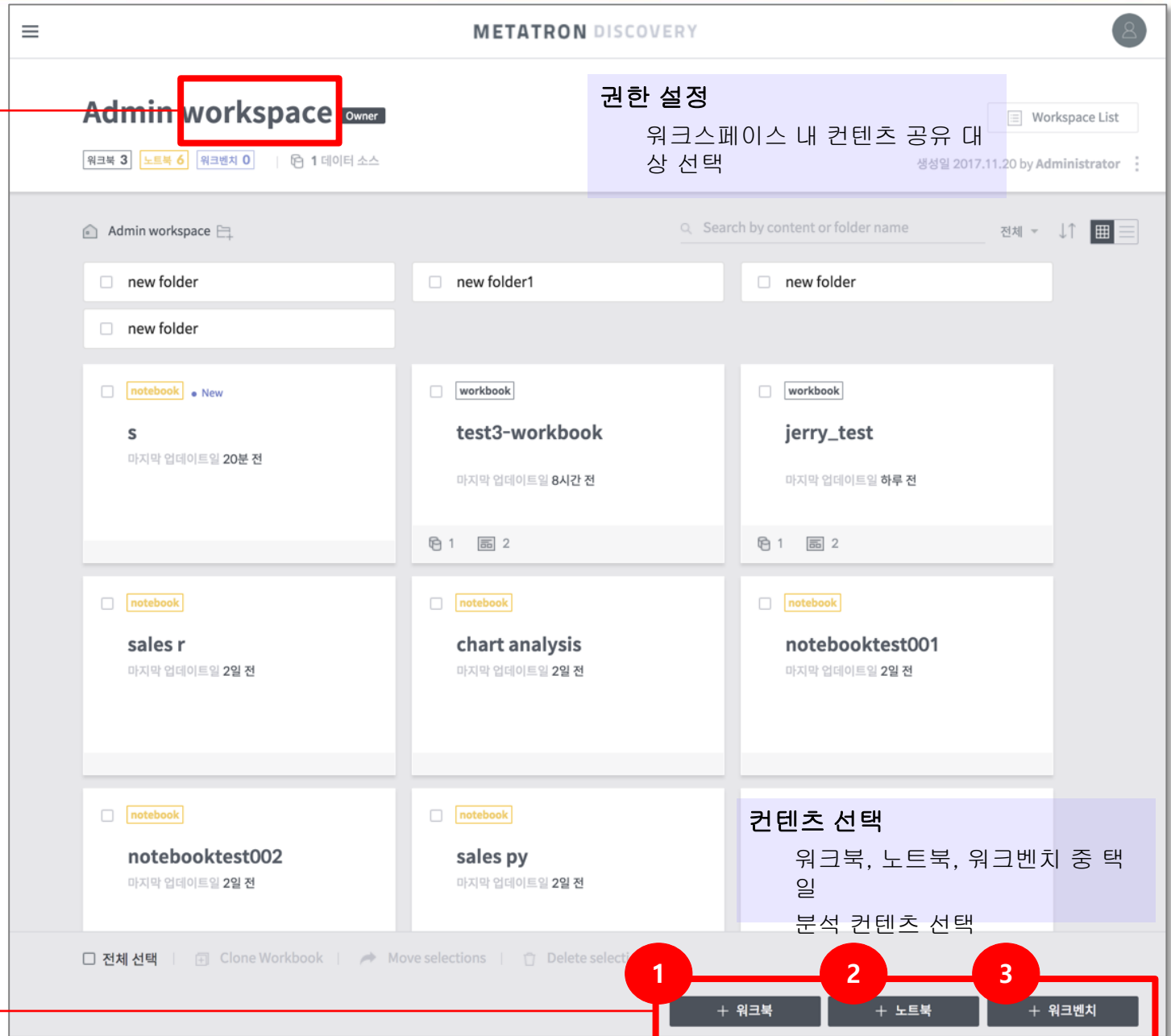
데이터 소스 관리, 데이터 리니지, 사용자/그룹 관리, 인증, 공유, 스케줄러, 서비스 모니터링 등

# [참고] metatron End-to-end 기능



# metatron 소개 - 워크스페이스

워크스페이스:  
분석 공간  
(개인, 공유)



1 워크북:  
시각화 기반 분석

2 노트북:  
ML 기반 고급 분석

3 워크벤치:  
SQL 기반 분석

# 워크스페이스 개요

개인 혹은 권한이 부여된 다수의 사용자 간의 공동 작업공간

개인  
워크스페이스

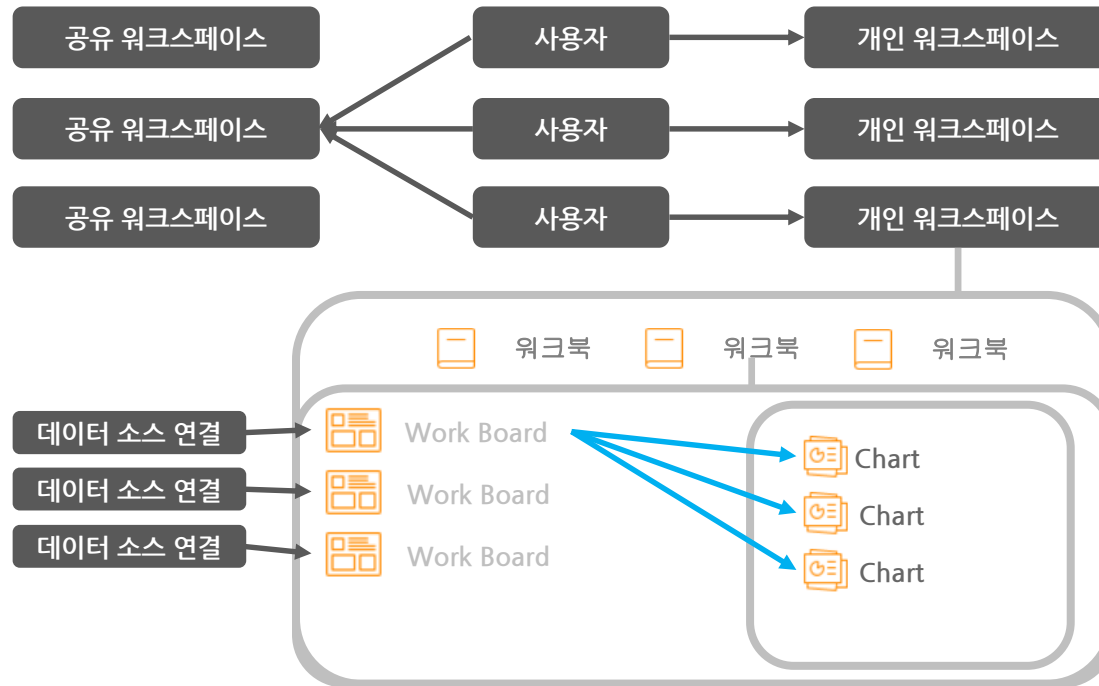
개인이 자신만의 공간에서  
작업하기 위해 기본적으로 주어진 공간

공유  
워크스페이스

다수의 권한을 가진 사용자에게  
공동 작업을 하고 관리하기 위한 공간

# METATRON 워크스페이스 전체 구성도

METATRON 간의 관계도



## ※ Metatron 기본 구조

- (Workspace) 워크북, 페이지, 워크보드를 작성 및 활용할 수 있는 작업공간(개인/공유 workspace)
- (Workbook) 여러 페이지차트와 워크보드를 포함하는 온라인 서류
- (Dashboard) 동일 워크북내 차트를 조합하여 만든 차트간 연동형 분석 보드
- (Chart) 데이터소스에서 조회한 데이터를 다양한 차트를 이용하여 분석하기 위한 기본 분석단위



# 1 워크북 – 시각화 기반 분석

## ■ 워크북을 통해 사용자는 보기 편하고 익숙한 PPT UX 환경에서 분석

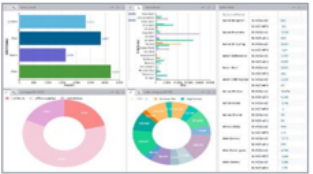
☰
METATRON DISCOVERY
👤

← 워크북 workbook
Copy of sale\_board 대시보드 설명을 입력해 주세요.
수정 2017.11.16 12:00:36 by Administrator
🖨️ 프레젠테이션 뷰
📄 대시보드 편집

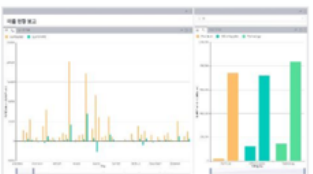
대시보드(3) | 댓글(0)

대시보드 모음  
다양한 대시보드를 표현


1 Copy of sale\_board



2 sales by location



3 sale\_board



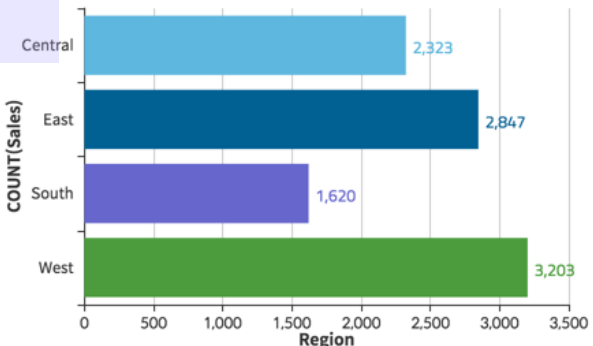
+ 대시보드 추가

데이터를 선택하세요.

**댓글 기능**  
워크북 내의 공유 내용에 대한 댓글 형식의 커뮤니케이션 수단 제공

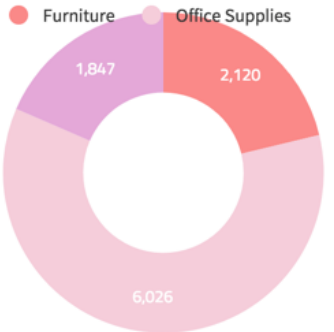
**프리젠테이션 뷰 기능 제공**  
리포팅 및 공유를 위한 프리젠테이션 기능 제공

**New Chart**



Region	Count(Sales)
Central	2,323
East	2,847
South	1,620
West	3,203

**category별 거래량**

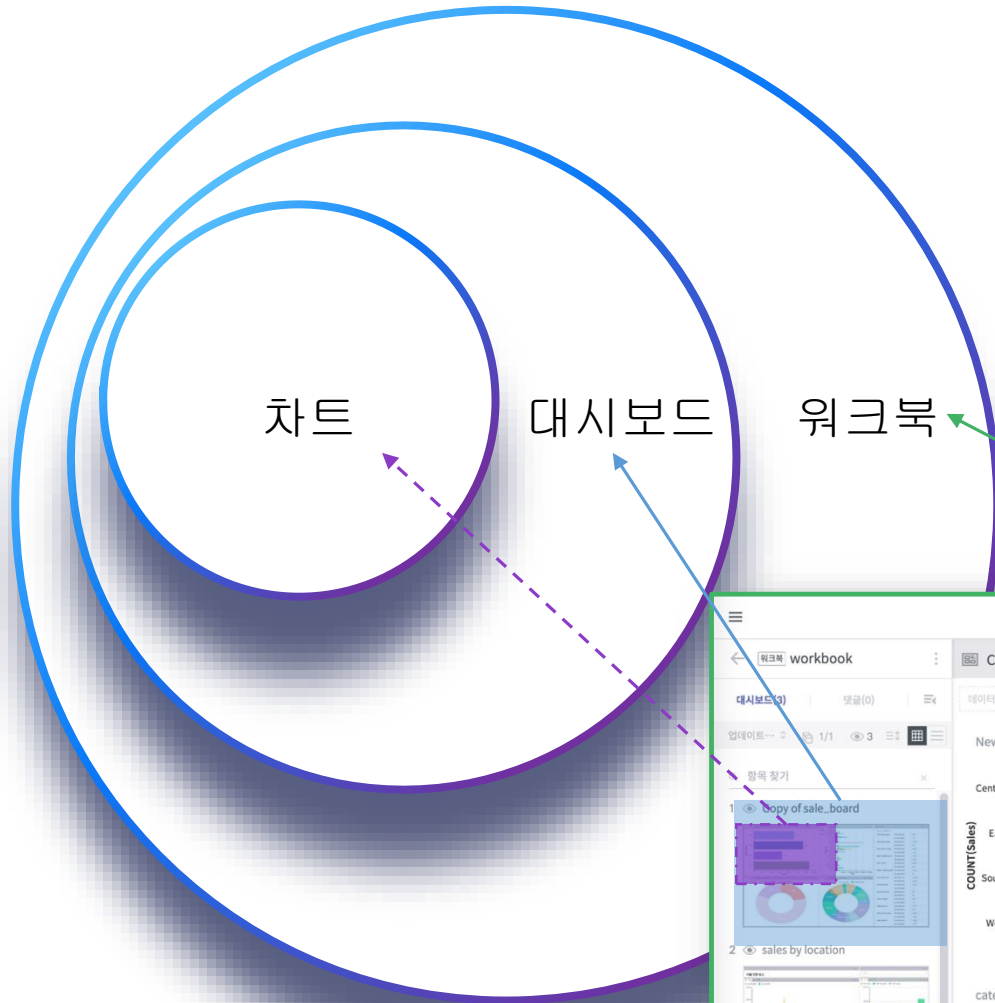


Category	Count(Sales)
Furniture	2,120
Office Supplies	1,847
Other	6,026

**프리젠테이션 뷰 기능 제공**  
리포팅 및 공유를 위한 프리젠테이션 기능 제공

CustomerName	SUM(Sales)	SUM(Profit)
Aaron Bergman	887	129
Aaron Hawkins	1,744	364
Aaron Smayling	3,051	-254
Adam Bellava...	7,757	2,055
Adam Hart	3,251	283
Adam Shilling...	3,255	63
Adrian Barton	14,476	5,445
Adrian Hane	1,736	-2
Adrian Shami	58	

# [참고] 워크북 계층



**METATRON DISCOVERY**

워크북 workbook | Copy of sale\_board | 대시보드 설명을 입력해 주세요. | 수정 2017.11.16 12:00:36 by Administrator | 프레젠테이션 뷰 | 대시보드 편집

대시보드(3) | 탭(0) | 업데이터: 1/1 | 3 | 항목 찾기

1 Copy of sale\_board

2 sales by location

3 sale\_board

+ 대시보드 추가

**New Chart**

Central: 2,323  
East: 2,847  
South: 1,620  
West: 3,203

category별 거래량

1/2 | Furniture: 1,847 | Office Supplies: 2,120 | 6,026

**New Chart**

SUM(Sales) by City

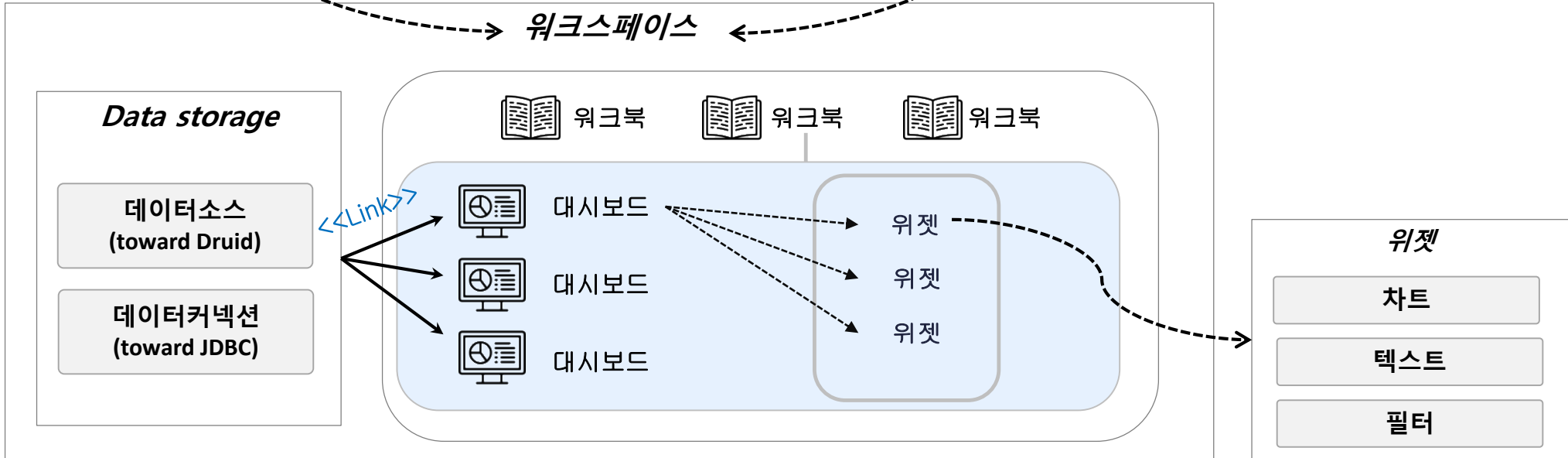
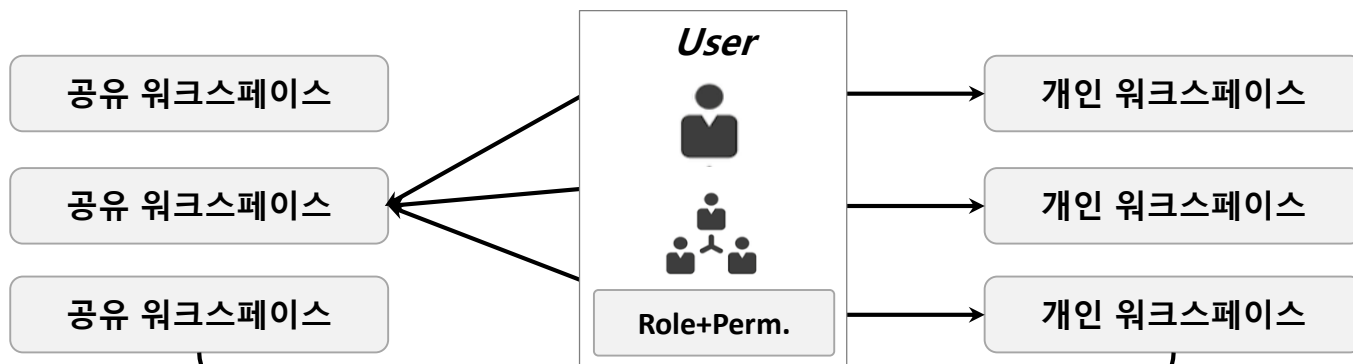
sub-category 별 판매

1/8 | Accessories: 59,148 | 401 | 107,538 | 203,428 | 114,879 | 328,454 | 149,530 | 189,243 | 1,170 | 78,475

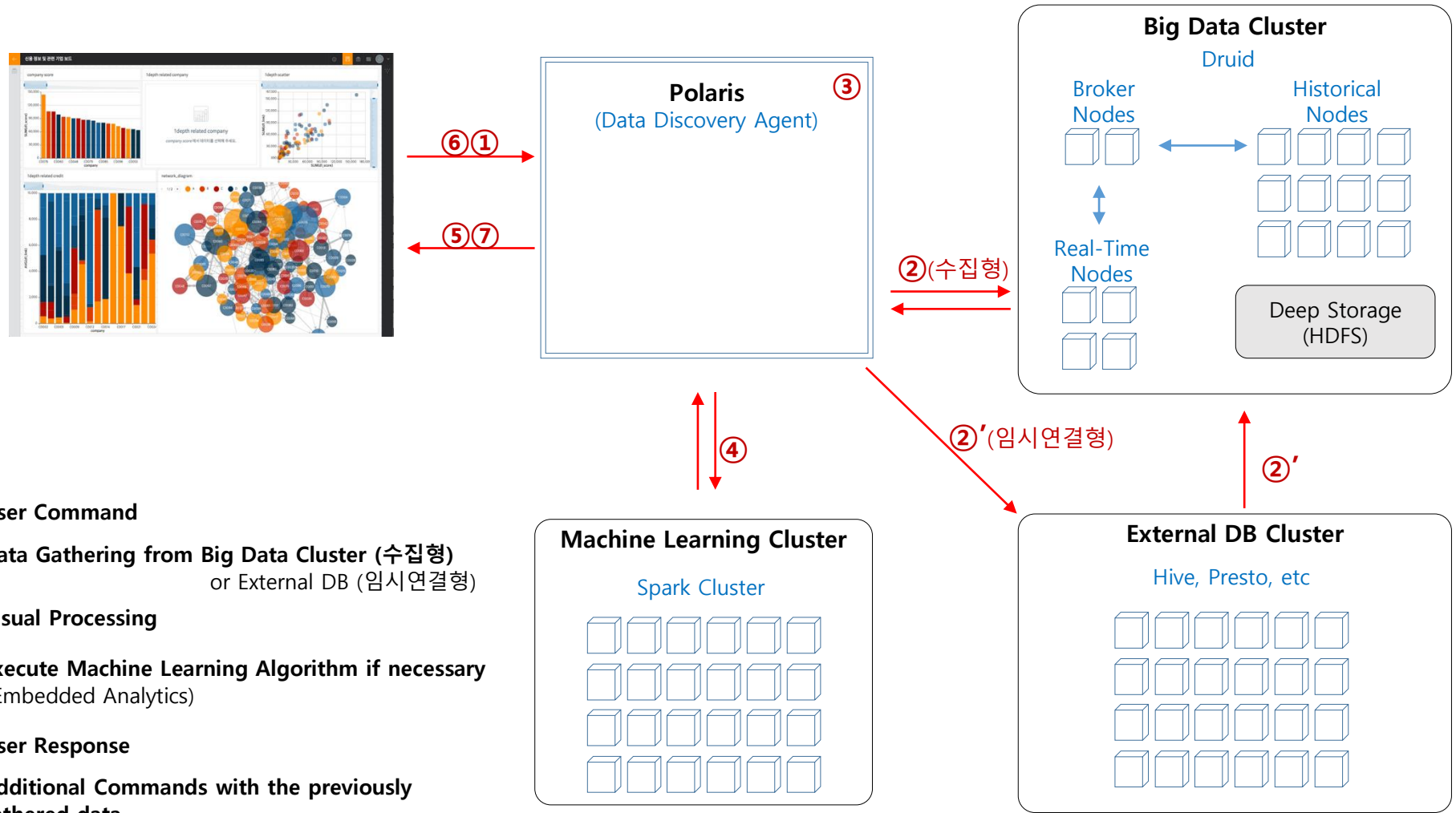
**월별 거래량**

CustomerName	SUM(Sales)	SUM(Profit)
Aaron Bergman	887	129
Aaron Hawkins	1,744	364
Aaron Smaying	3,051	-254
Adam Bellava...	7,757	2,055
Adam Hart	3,251	283
Adam Shilling...	3,255	63
Adrian Barton	14,476	5,445
Adrian Hane	1,736	-2
Adrian Shami	58	

# 워크북 구조



# 워크북 동작 (Data Visualization)



- ① User Command
- ② Data Gathering from Big Data Cluster (수집형) or External DB (임시연결형)
- ③ Visual Processing
- ④ Execute Machine Learning Algorithm if necessary (Embedded Analytics)
- ⑤ User Response
- ⑥ Additional Commands with the previously gathered data
- ⑦ User Response

## ■ 분석 대상 선택

- 데이터소스, 대시보드, 차트 중 택일
- 분석 대상 데이터 선택

Select a data type

데이터 소스

대시보드

차트

취소

## ■ 노트북 상세 정보 입력분석

- Server type 선택
- 개발언어 선택 (Jupyter : R / PYTHON, Zeppelin : Spark)
- 노트북명과 설명 입력

노트북 생성하기

Please complete notebook creation

차트 workbook > dashboard > PAGE\_1507784456280

Server type jupyter

Develop language R

이름  
이름을 입력하세요

## ■ Jupyter notebook 팝업

- 개발언어 분석 대상 데이터셋 로딩 코드 삽입
- 분석가가 R 코드를 작성 후 저장

jupyter 9f65563f-103b-4325-8d56-a1f8ee3e05ae (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Code

**1. load dataset**

In [ ]: `library(RMetis)`

In [ ]: `dataset <- charts.get(client.url('metatron.mcloud.sktelecom.com',80), 'c59b75cc-`

**2. analyze**

In [ ]:

## ■ 노트북 API 생성

- Return type (html, json) 선택
- notebook code  
`response.write(_user_object_)`

← Notebook sales analysis

Data type: CHART  
Data source: PAGE\_1507685416920  
Develop language: R  
Code: [Detail](#)

---

Name: sales analysis report  
Description:  
URL: <http://metatron.mcloud.sktelecom.com/ap>  
Return type: HTML  
API result: [Result](#)

Edit API  Delete API

### API information

Return Type  HTML  JSON

이름  
이름을 입력하세요

설명  
설명을 입력하세요

[Cancel](#) [Done](#)

### sales analysis report

james  
Thu Oct 12 15:41:21 2017

```
library(RMetis)

## Loading required package: httr

## Loading required package: jsonlite

## Loading required package: methods

dataset <- charts.get(client.url('localhost',4200), 'c59b75cc-5b1d-4c7a-b5fa-e7ec8e29c16c')
summary(dataset)

##   SUM(Sales)   Category
##   Min.   :719127 Length:3
##   1st Qu.:730566 Class :character
##   Median :742006 Mode  :character
##   Mean   :765785
##   3rd Qu.:789114
##   Max.   :836221

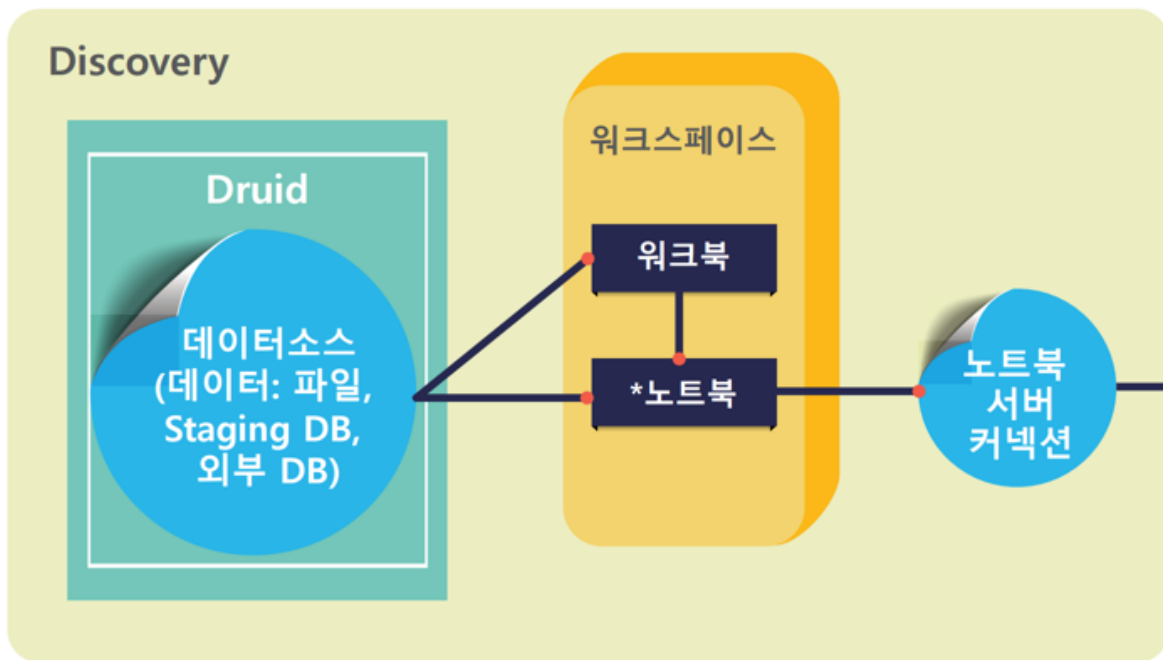
regression_model <- list(coefficients = 2.5, intercept = 0.0)
response.write(regression_model)

## {"coefficients":[2.5],"intercept":[0]}
```

## ■ 생성된 URL을 통해 RESTful 서비스 제공

- Result 클릭하여 조회 가능

```
{"coefficients":[2.5],"intercept":[0]}
```



[노트북 구조]

—●— 데이터 흐름

외부 분석 도구	사용 언어
Jupyter	R
	Python
Zeppelin	Spark

[외부 분석 도구와 사용 가능 언어 ]

**METATRON DISCOVERY**

워크벤치 Workbench for hive 설명을 입력하세요

Hive(1.2)

default

Table list

contract\_temp

contract\_orc\_part

abc1

employee

sample\_ingestion

sample\_ingestion\_time

```

1 SELECT
2 *
3 FROM
4 default.contract_orc_part;

```

히스토리  
쿼리 히스토리 조회 및 결과 리스트

결과 실행  
결과 전체 실행, 블록 실행

다이나믹 차트 생성  
다이나믹하게 데이터소스를 구성하고 차트 생성 테스트 수행 가능

데이터소스 생성 슛 패스 제공  
쿼리 결과를 바로 데이터소스로 인제스천

온라인 엑셀 기능  
온라인 엑셀 프로그램으로 연동을 통하여 결과 데이터를 표시

스키마 뷰  
테이블 정보, 데이터 프리뷰, 컬럼 스키마 조회

쿼리 1 - 결과 1 x

컬럼 데이터 검색

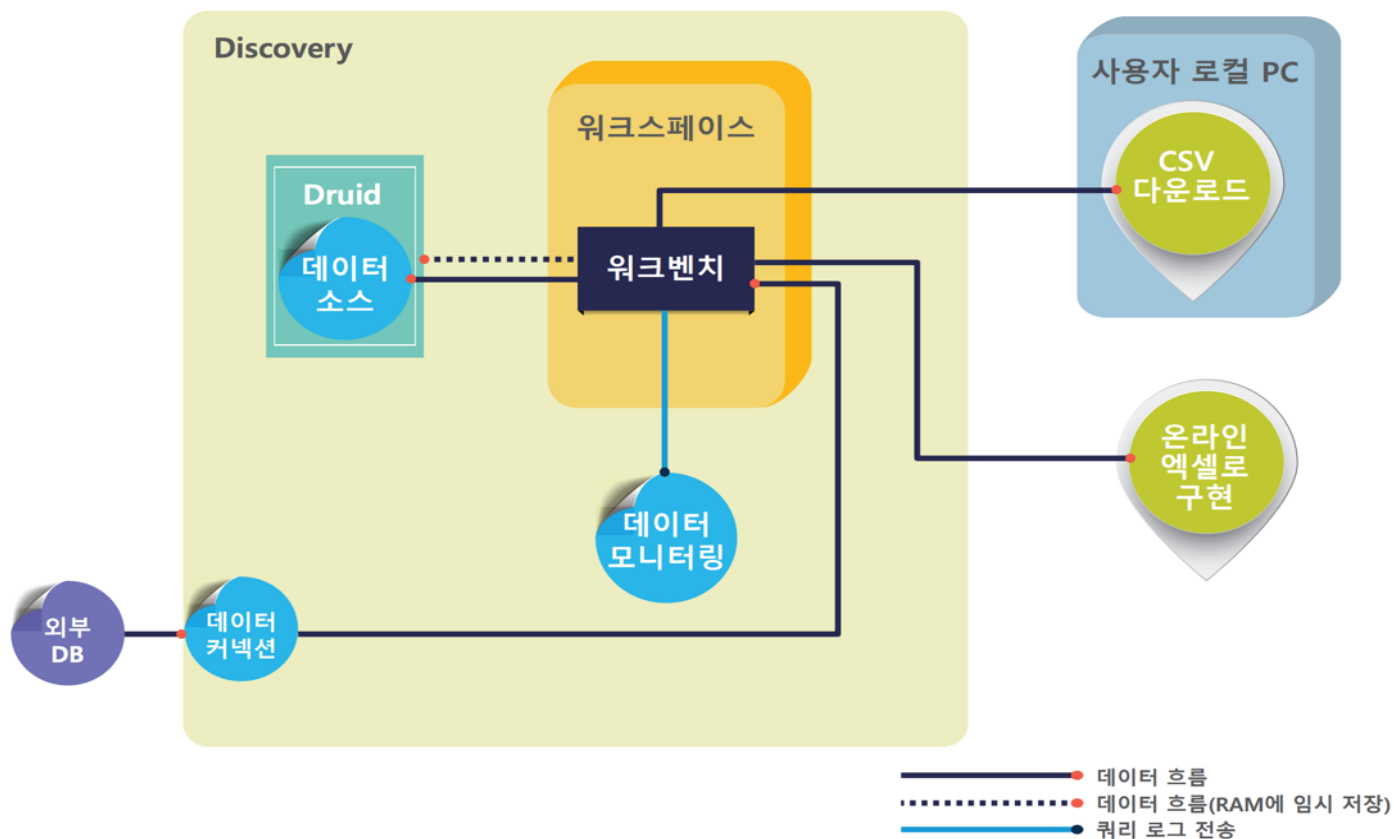
	contract_id	product_01_code	product_02_code	product_03_code	product_04
1	cid000000001	01	01	004	0002
2	cid000004158	04	04	004	0025
3	cid000005953	02	04	004	0034
4	cid000005982	01	02	004	0030
5	cid000006210	02	02	003	0019
6	cid000014008	02	04	005	0005
7	cid000016535	02	05	005	0014
8	cid000021993	02	03	003	0044
9	cid000022344	04	04	005	0041

37 Rows

< 이전 다음 >



## 워크벤치 구조 및 특징



- 여러 서버에 분산된 다양한 외부 데이터베이스(Oracle, MySQL, Hive, Presto, Tibero)를 한꺼번에 작업공간에 조회 가능
- 연동된 데이터베이스의 스키마를 자유자재로 탐색하여 테이블과 컬럼을 쉽게 조회 및 선택
- 쿼리 편집 도구가 내장되어 데이터의 조회, 추가, 삭제, 수정이 용이
- 쿼리 결과를 실시간으로 확인하고 로컬 파일로 다운로드 가능
- 출력된 데이터를 Druid 엔진으로 옮겨서 다양한 차트로 시각화
- 데이터 모니터링 메뉴에서 로그 분석 가능

*metatron 2.0*

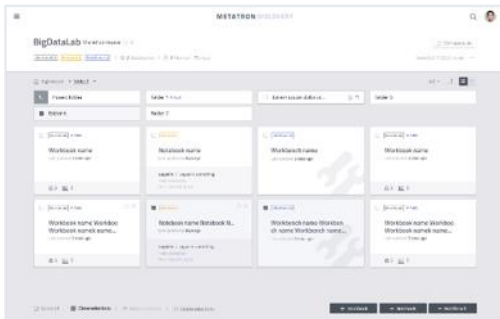


metatron

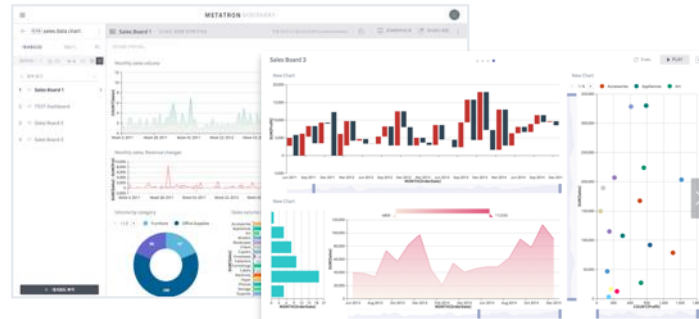
# V2.0 개선 - UX/UI

## metatron 2.0 Release

총 120장의 페이지의 디자인 개편을 통한 보다 개선된 UX 제공



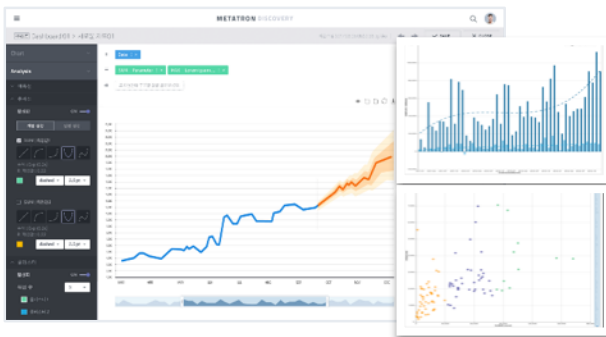
Workspace



Dashboard/Chart



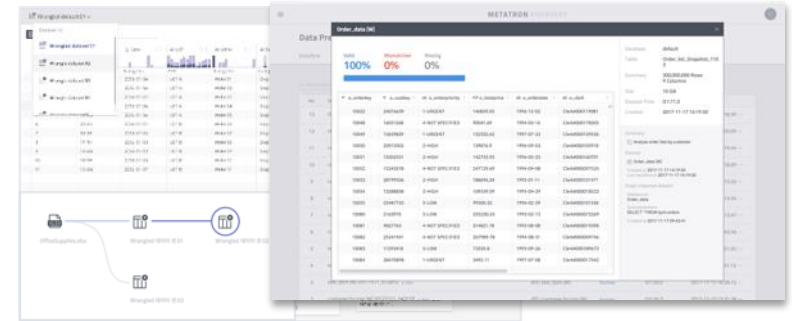
Workbench - SQL Explorer



Embedded Analytics



Data Lineage



Data Prep. - Data Transformer

## V2.0 개선 - Data Preparation

### ■ 데이터를 시각화, 분석하기 위한 지루하고 반복적인 ETL작업을 누구나 쉽게

- 서로 다른 데이터 소스(RDB, Hive, Log file, etc.)에 대해 일관성있는 추상화 레벨 제공
- 샘플 데이터로 전후 결과를 빠르게 보고 비교하면서 편집할 수 있음 (Step-by-step, back-and-forth)

The screenshot displays a data preparation tool interface. On the left, a table titled 'Customer\_data [W]' is visible with columns: #, Customer\_ID, ab c\_name, and ab c\_address. The main workspace shows a 'Join' configuration screen. It features two data sources: 'Master dataset' (Customer\_data [W]) and 'Dataset to join' (Nation\_code [W]). Both datasets have their respective columns selected. The 'Join type' is set to 'Inner'. Below this, a 'Join keys' section shows the mapping: 'Customer\_ID' from the master dataset is joined to 'n\_nationkey' from the dataset to join. On the right, a 'Rule list' panel contains several transformation rules, including 'rename', 'split', 'sort', 'settype', 'derive', and 'replace'. At the bottom, there are 'Show result', 'Cancel', and 'Add' buttons.

#	Customer_ID	ab c_name	ab c_address
1	Customer#...	Removed for p	
2	Customer#...	Removed for p	
3	Customer#...	Removed for p	
4	Customer#...	Removed for p	
5	Customer#...	Removed for p	
6	Customer#...	Removed for p	
7	Customer#...	Removed for p	
8	Customer#...	Removed for p	
9	Customer#...	Removed for p	
10	Customer#...	Removed for p	
11	Customer#...	Removed for p	
12	Customer#...	Removed for p	
13	Customer#...	Removed for p	
14	Customer#...	Removed for p	
15	Customer#...	Removed for p	
16	Customer#...	Removed for p	

#	Customer_ID	# c_nationkey	ab c_mktsegment
1		15	BUILDING
2		13	AUTOMOBILE
3		1	AUTOMOBILE
4		4	MACHINERY
5		3	HOUSEHOLD
6		20	AUTOMOBILE
7		18	AUTOMOBILE
8		17	BUILDING
9		8	FURNITURE
10		5	HOUSEHOLD

#	n_nationkey	ab n_name	ab r_name
0		ALGERIA	AFRICA
5		ETHIOPIA	AFRICA
14		KENYA	AFRICA
15		MOROCCO	AFRICA
16		MOZAMBIQUE	AFRICA
1		ARGENTINA	AMERICA
2		BRAZIL	AMERICA
3		CANADA	AMERICA
17		PERU	AMERICA
24		UNITED STATES	AMERICA

# V2.0 개선 - Data Preparation

데이터를 시각화, 분석하기 위한 ETL기능 제공

데이터 스냅샷 생성

온라인 엑셀 프로그램으로 연동을 통하여 결과 데이터를 표시

ETL

The screenshot displays the METATRON DISCOVERY interface with several key components:

- Customer\_data [W] Table:**

Customer_ID	c_name	c_address	c_nationkey	c_acctbal	c_mktsegment	c_com...	Tel_Region_Code	split_c_phone2	split_c_pl
1	Customer...	Removed for privac...	15	711.56	BUILDING	to the even, r...	25	989	741
2	Customer...	Removed for privac...	13	121.65	AUTOMOBILE	l accounts, bl...	23	768	687
3	Customer...	Removed for privac...	1	7498.12	AUTOMOBILE	deposits eat ...	11	719	748
4	Customer...	Removed for privac...	4	2866.83	MACHINERY	requests, fina...	14	128	190
5	Customer...	Removed for privac...	3	794.47	HOUSEHOLD	n accounts w...	13	750	942
6	Customer...	Removed for privac...	20	7638.57	AUTOMOBILE	tions, even d...	30	114	968
7	Customer...	Removed for privac...	18	9561.95	AUTOMOBILE	ainst the iron...	28	190	982
8	Customer...	Removed for privac...	17	6819.74	BUILDING	among the si...	27	147	574
9	Customer...	Removed for privac...	8	8324.07	FURNITURE	r theodolites ...	18	338	906
10	Customer...	Removed for privac...	5	2753.54	HOUSEHOLD	es regular de...	15	741	346
11	Customer...	Removed for privac...	23	-272.6	BUILDING	ckages, requ...	33	464	151
12	Customer...	Removed for privac...	13	3396.49	HOUSEHOLD	to the careful...	23	791	276
13	Customer...	Removed for privac...	3	3857.34	BUILDING	ounts sleep c...	13	761	547
14	Customer...	Removed for privac...	1	5266.3	FURNITURE	, iron pack...	11	845	129
15	Customer...	Removed for privac...	23	2788.52	HOUSEHOLD	platelets, reg...	33	687	542
16	Customer...	Removed for privac...	10	4681.03	FURNITURE	kly silent cou...	20	781	609
- Order\_data [W] Table:**

o_orderkey	o_custkey	o_orderpriority	o_totalprice	o_orderdate	o_clerk
10023	24076439	1-URGENT	140699.05	1996-12-02	ClerkK00019081
10048	16051268	4-NOT SPECIFIED	90041.69	1994-05-16	ClerkK000178505
10049	13659839	1-URGENT	132332.62	1997-07-23	ClerkK000129536
10050	20913502	2-HIGH	139876.9	1996-09-03	ClerkK000105918
10051	13202551	2-HIGH	142753.95	1996-05-25	ClerkK000160701
10052	15243578	4-NOT SPECIFIED	247129.69	1994-09-08	ClerkK000097525
10053	28799506	2-HIGH	188696.34	1992-01-11	ClerkK000101471
10054	13288838	2-HIGH	109539.09	1995-04-29	ClerkK000185233
10055	25487725	5-LOW	99503.32	1996-02-29	ClerkK000101558
10060	2163970	5-LOW	255230.55	1993-02-13	ClerkK000072269
10061	9827765	4-NOT SPECIFIED	314821.78	1993-08-08	ClerkK00019298
10062	25241941	4-NOT SPECIFIED	207869.78	1994-08-31	ClerkK000099746
10063	11293418	5-LOW	73035.8	1995-09-26	ClerkK000189673
- Rule List:**
  - rename rename col: c\_custkey to: 'C customer\_ID'
  - split split col: c\_phone on: '-' limit: 3 ignoreCase: true
  - sort sort order: Customer\_ID
  - settype settype col: split\_c\_phone1 type: Integer
  - rename rename col: split\_c\_phone1 to: 'Tel\_Region\_Code'
  - derive derive value: if(Tel\_region\_Code=20, 'East Side', 'West Side') as: 'region'
  - replace replace col: c\_address with: 'Removed for privacy' on: '/' global: true
- Customer\_data [W] Table (Detailed):**

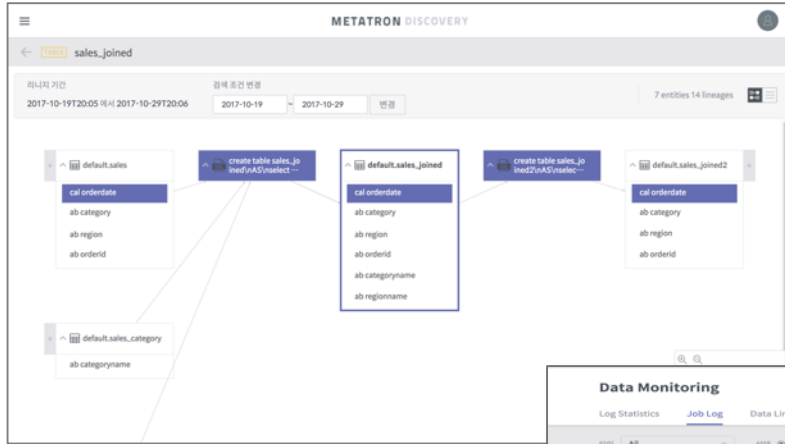
Customer_ID	c_name	c_address	c_nationkey	c_acctbal	c_mktsegment	c_com...	Tel_Region_Code	split_c_phone2	split_c_pl
1	Customer...	Removed for privac...	15	711.56	BUILDING	to the even, r...	25	989	741
2	Customer...	Removed for privac...	13	121.65	AUTOMOBILE	l accounts, bl...	23	768	687
3	Customer...	Removed for privac...	1	7498.12	AUTOMOBILE	deposits eat ...	11	719	748
4	Customer...	Removed for privac...	4	2866.83	MACHINERY	requests, fina...	14	128	190
5	Customer...	Removed for privac...	3	794.47	HOUSEHOLD	n accounts w...	13	750	942
6	Customer...	Removed for privac...	20	7638.57	AUTOMOBILE	tions, even d...	30	114	968
7	Customer...	Removed for privac...	18	9561.95	AUTOMOBILE	ainst the iron...	28	190	982
8	Customer...	Removed for privac...	17	6819.74	BUILDING	among the si...	27	147	574
9	Customer...	Removed for privac...	8	8324.07	FURNITURE	r theodolites ...	18	338	906
10	Customer...	Removed for privac...	5	2753.54	HOUSEHOLD	es regular de...	15	741	346
11	Customer...	Removed for privac...	23	-272.6	BUILDING	ckages, requ...	33	464	151
12	Customer...	Removed for privac...	13	3396.49	HOUSEHOLD	to the careful...	23	791	276
13	Customer...	Removed for privac...	3	3857.34	BUILDING	ounts sleep c...	13	761	547
14	Customer...	Removed for privac...	1	5266.3	FURNITURE	, iron pack...	11	845	129
15	Customer...	Removed for privac...	23	2788.52	HOUSEHOLD	platelets, reg...	33	687	542
16	Customer...	Removed for privac...	10	4681.03	FURNITURE	kly silent cou...	20	781	609
- Workflow Diagram:** Shows data flow from Customer\_data, Nation\_code, Region\_code, and Order\_data tables through various ETL processes to create a final Order\_data table.
- Summary:** Order\_data [W] table contains 10,000 rows and 12 columns.

데이터 플로우 생성

롤 적용

# V2.0 개선 - Data Management (Lineage & Job Log)

- Data Lineage : 쿼리 로그를 이용해 데이터 전후 관계를 관리
- Job Log : 워크벤치에서 발행된 질의 이력 및 다양한 통계 정보 관리



그래프 표출을 통한 컬럼 히스토리  
MDM의 일환으로 전체 데이터가 어떤 테이블, 어떤 컬럼들과 연결되어 실행되었는지 확인할 수 있는 기능 제공

**Data Monitoring**

Log Statistics | Job Log | Data Lineage

Performed Start Time: **전체** Today Last 7 days vvvv-MM-dd hh:mm vvvv-MM-dd hh:mm 작업

Job 또는 커넥터/이벤트/오류 검색

상태	Job name	Job ID	Application ID	Queue	사용자 (ID)	시작시간	작업시간
SUCCESS	select * from polaris_v2.book				admin	2018-02-20 15:00:40	2 ms
SUCCESS	select * from polaris_v2.book				admin	2018-02-20 13:51:08	2 ms
SUCCESS	select * from book				admin	2018-02-20 13:50:50	1 ms
SUCCESS	SELECT * FROM polaris_v2.book				admin	2018-02-20 13:36:30	2 ms
SUCCESS	SELECT * FROM polaris_v2.audit				admin	2018-02-20 13:36:12	32 ms
SUCCESS	SELECT * FROM polaris_v2.audit				admin		
SUCCESS	select * from polaris_v2.book				admin		
SUCCESS	select * from polaris_v2.book				admin		
SUCCESS	select * from polaris_v2.book				admin		
SUCCESS	select * from polaris_v2.analysis_binary				admin		
SUCCESS	select * from polaris_v2.analysis_binary				admin		
SUCCESS	select * from polaris_v2.book_notebook				admin		
SUCCESS	SHOW SCHEMAS				hadoop		

쿼리 이력 목록 조회 및 검색  
쿼리 이력 상세 조회

**METATRON DISCOVERY**

Recently performed on 2018-02-20 11:35:22 by admin

(select \* from polaris\_v2.book)

**Data Information**

상태: SUCCESS

Job name: select \* from polaris\_v2.book [Detail >](#)

시작시간: 2018-02-20 11:35:22

작업시간: 0 sec

User: admin

Connection: **작업** MySQL

Host: metatron-goc-a01

Port: 3306

JDBC URL: jdbc:mysql://metatron-goc-a01:3306/

**Query History**

Query date	User	작업시간	Result
2018-02-20 15:50:40	Administrator	2 ms	SUCCESS <a href="#">Detail &gt;</a>
2018-02-20 13:51:08	Administrator	2 ms	SUCCESS <a href="#">Detail &gt;</a>
2018-02-20 13:50:50	Administrator	1 ms	SUCCESS <a href="#">Detail &gt;</a>
2018-02-20 13:36:30	Administrator	2 ms	SUCCESS <a href="#">Detail &gt;</a>
2018-02-20 13:36:12	Administrator	32 ms	SUCCESS <a href="#">Detail &gt;</a>

# V2.0 개선 - Embedded Analytics





metatron

# Use case



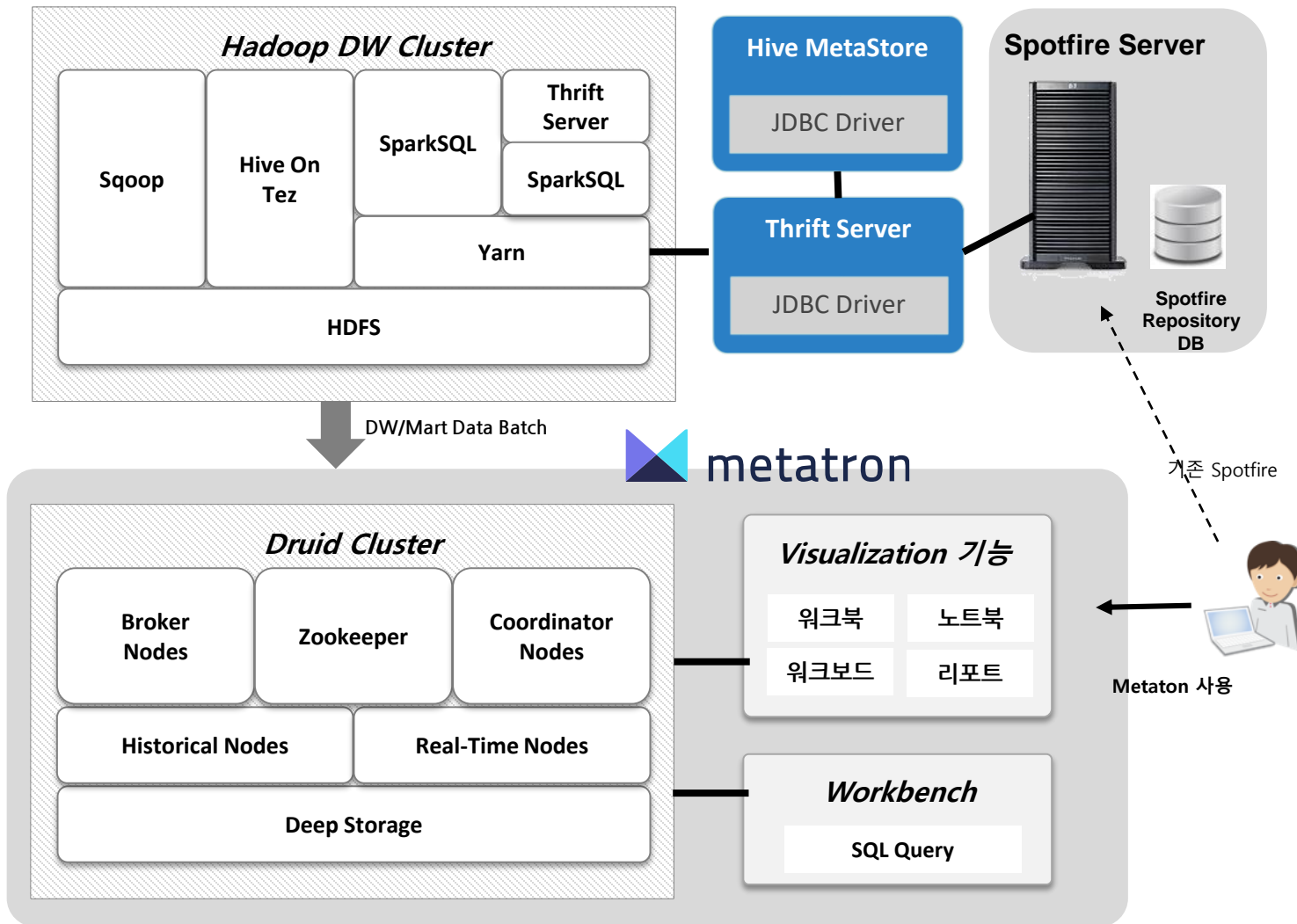


# Use case - SKT TANGO N/W Analytics

- ❖ '17년 3월 Metatron은 TANGO DW의 데이터 분석 환경으로 상용 적용되어 기존 Spotfire를 대체하고 있음

## 【 Spotfire 단점 】

- 메모리의 한계를 벗어나는 데이터의 분석이 어려움
  - 필요한 데이터를 Client PC 또는 Spotfire 서버로 조회하여 시각화  
(CELL 단위 분석만 하더라도 여러 가지 한계 발생)
- Data Mart 생성에 대한 유지보수 양의 증가
  - 기본적으로 조회된 데이터의 편집을 허용하지 않음
  - 따라서, 분석에 필요한 데이터를 Mart 로 구성해 뒤야 하며, 이에 따른 유지보수 어려움 증가



# [참고] SKT TANGO Analytics 적용 화면

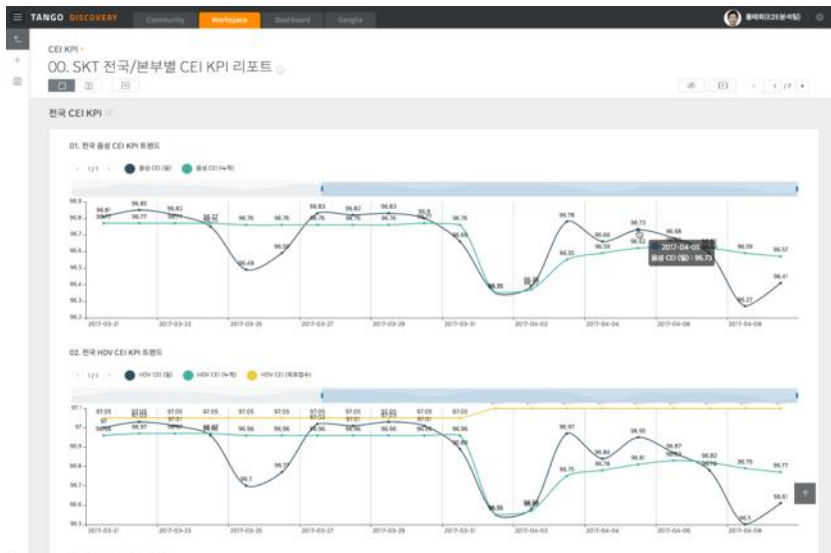
## 1. 전국 CEI 현황 대시보드



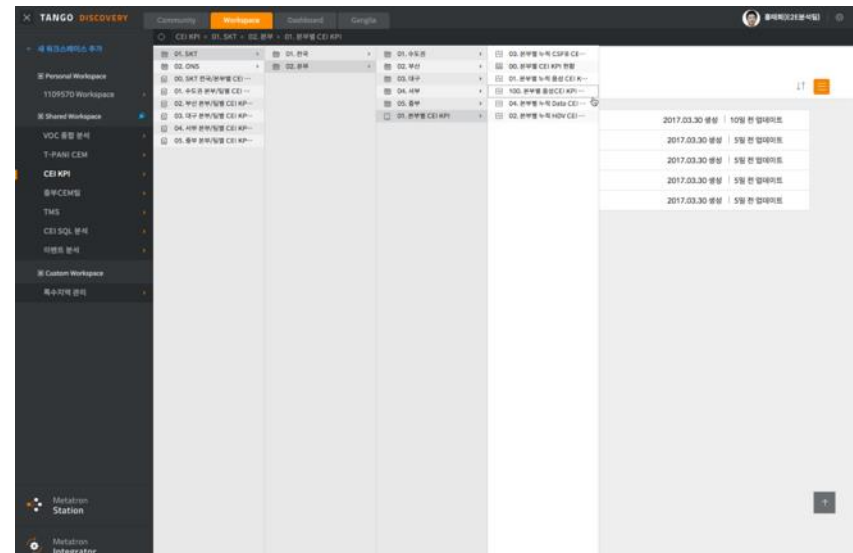
## 2. 본부별 CEI 현황 대시보드



## 3. CEI 결과 리포트

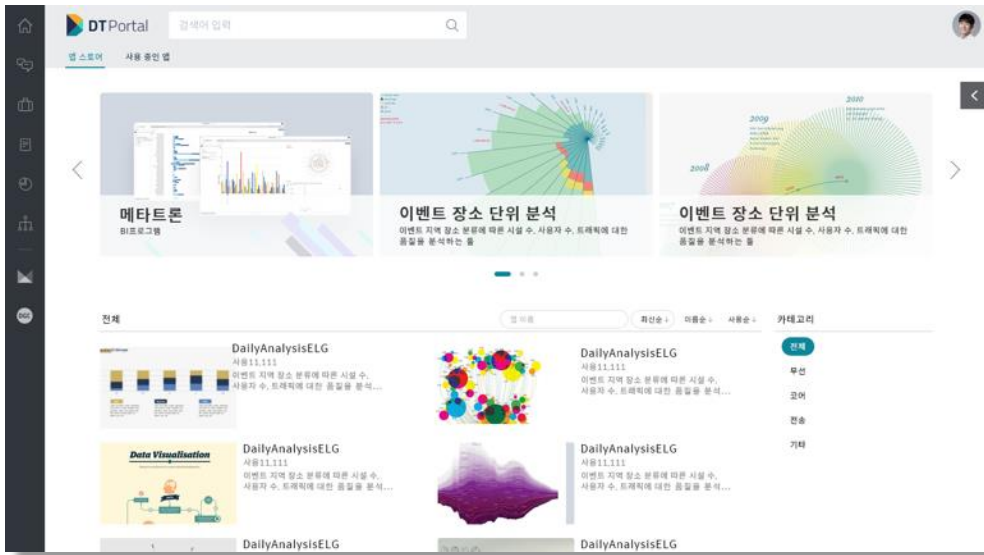


## 4. 다양한 Board를 위한 네비게이션

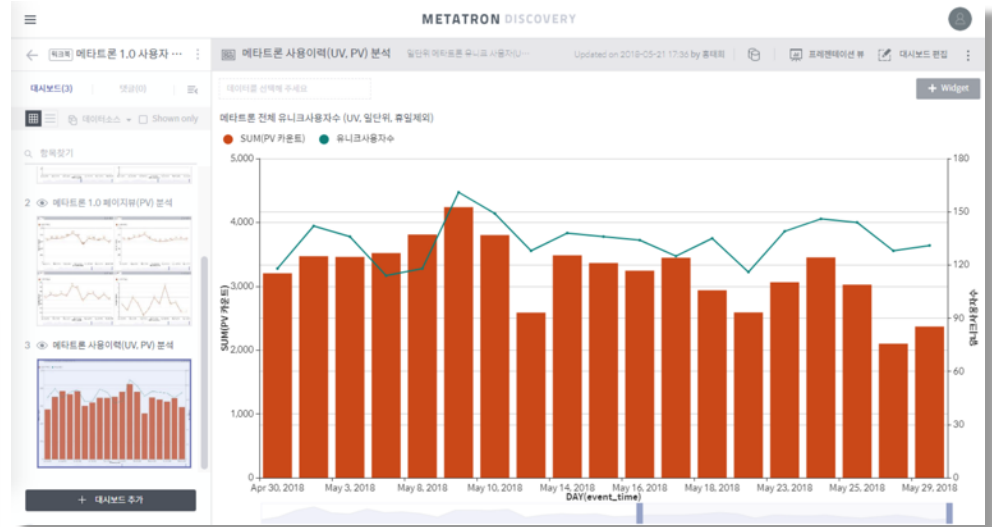


# [참고] SKT TANGO Analytics 2.0 적용 화면

## 포탈



## 사용자 현황

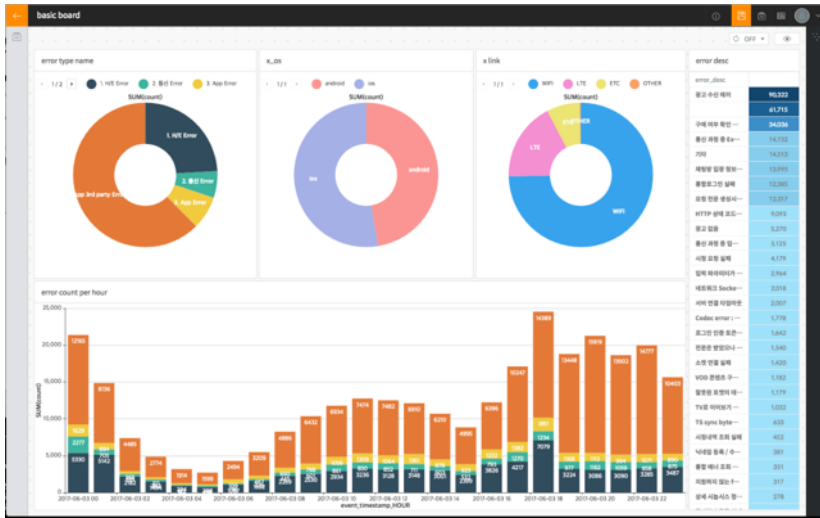


## 수도권 VoC 현황

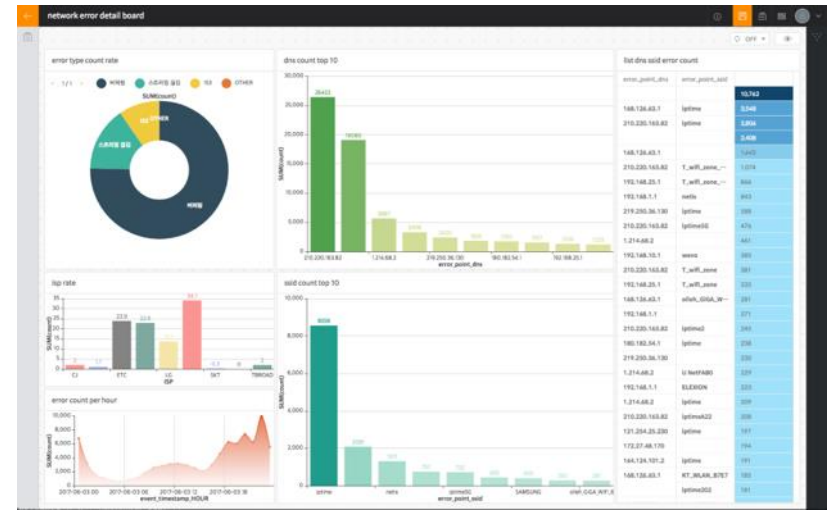


# 상용화 - SKB 옥수수 Log 분석

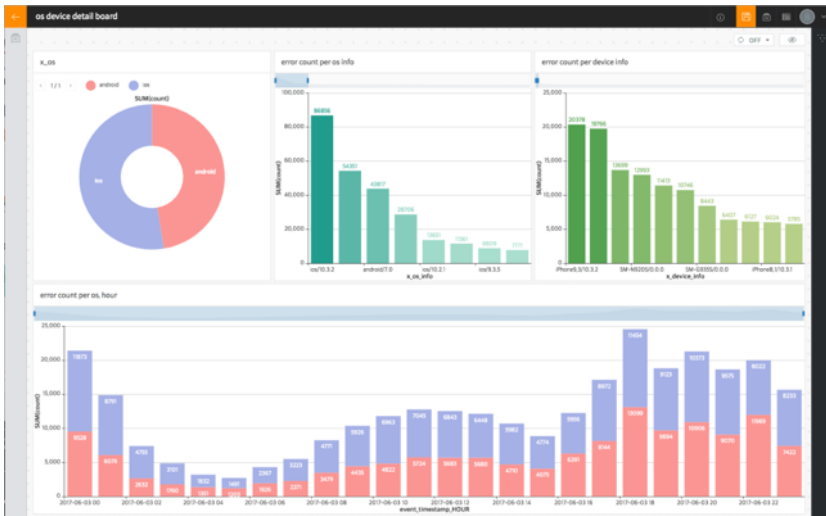
- Error monitoring



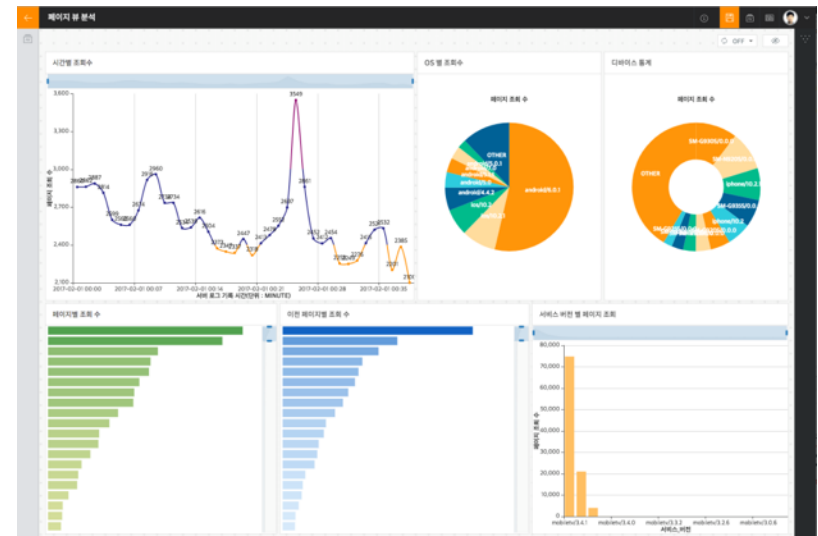
- Network error detail board



- OS device detail board




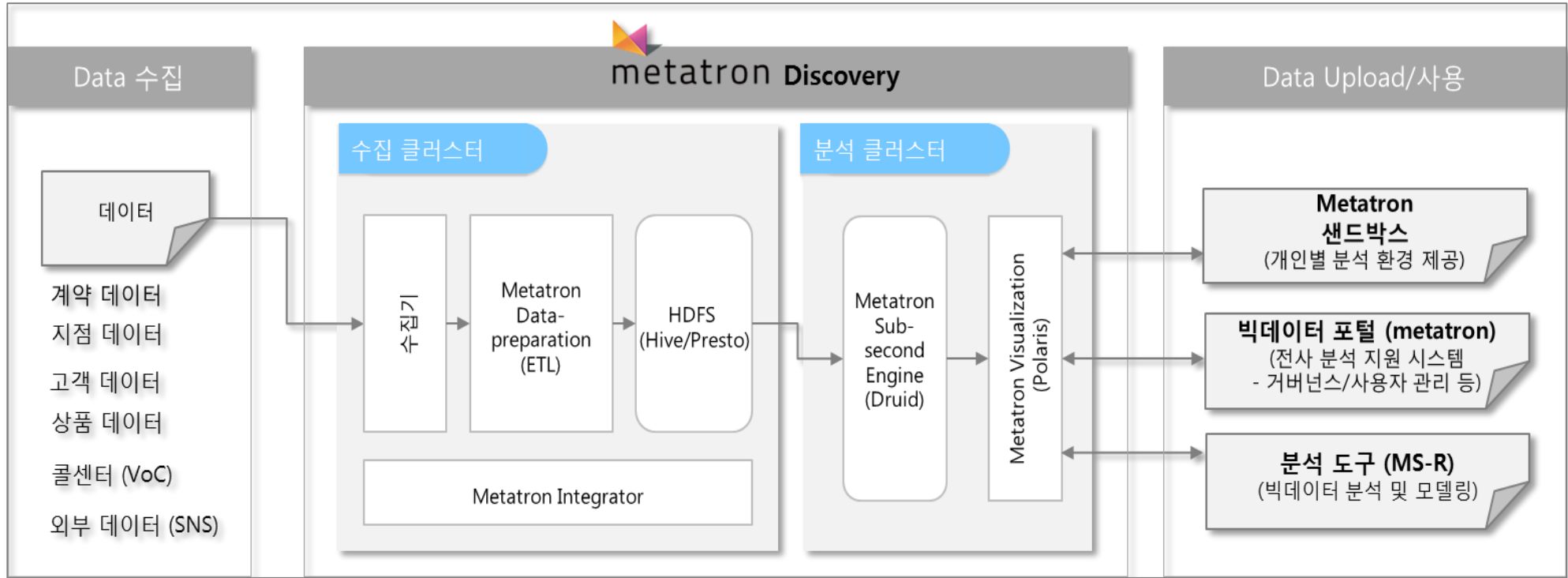
- 페이지 뷰 분석



# 상용화 – IBK 빅데이터 포털(영업 지원 시스템)

❖ '18년 5월 정식 오픈 했으며, 동시 사용자 100명 수준의 인프라 구성

- IBK 빅데이터 분석 인프라 



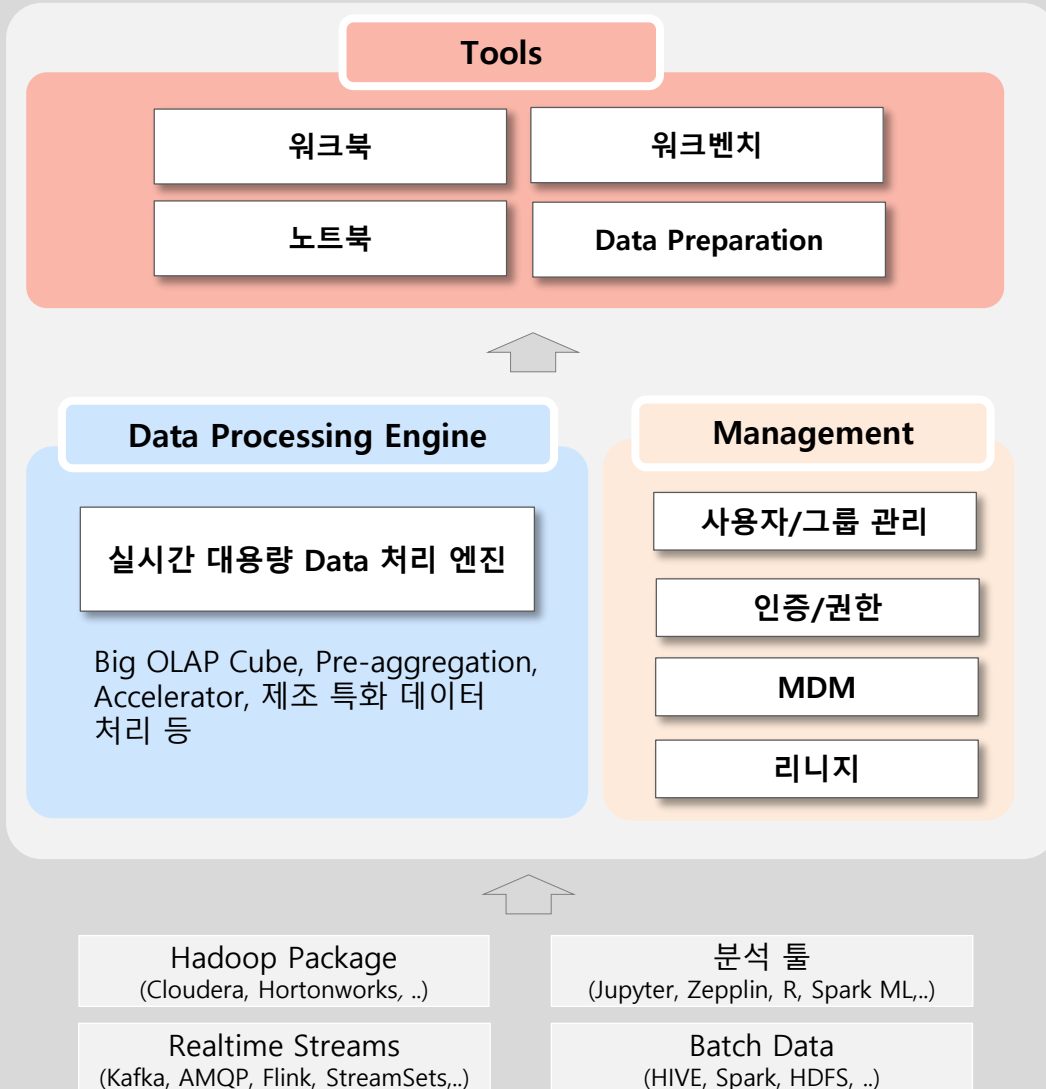
An aerial photograph of a dense city skyline, likely New York City, taken during the "blue hour" of sunset. The sky is a mix of soft pinks, oranges, and blues, with scattered white clouds. The city's skyscrapers are silhouetted against the bright horizon. In the center of the image, the word "metatron" is written in a clean, white, sans-serif font. It is framed by two horizontal white lines, one above and one below the text.

metatron

# Technology

# metatron 주요 특징

## Architecture



## Key Features

### Intuitive Analysis

직관적인 인터페이스로 빅데이터에 대해 데이터 준비 단계부터 분석 차트까지의 End-to-end 기능으로 쉽게 분석 가능

### Prepare Data by Yourself

시각화 기반 Data Preparation 기능으로 사용자가 직접 수정, 통합, 변환하여 데이터 준비

### Visualize & Share Your Insights

Massive Data의 빠른 시각화와 실시간 모니터링을 통해 새로운 Insight를 발굴하고, 동료들과 공유

### Notebook & Model Manager

다양한 분석 도구 연동을 지원하며, metatron에 저장된 데이터를 활용하여 분석하고, 동료들과 공유

### Big OLAP Cube

대용량 Fact 데이터에 다양한 Dimension을 결합하여 Big Mart를 생성함으로써 ETL 비용 감소, 속도 향상 및 스키마 변경 가능

### Sub-second Processing Engine

시간에 따라 In-memory, Local Storage, Deep Storage로 데이터를 이관하여 TB 이상의 대용량 데이터에 대해서도 빠르게 응답

### Enhanced Druid Engine

핵심 기능 고도화, 성능 향상 등 metatron에 특화된 Druid 탑재

### Load from Everywhere

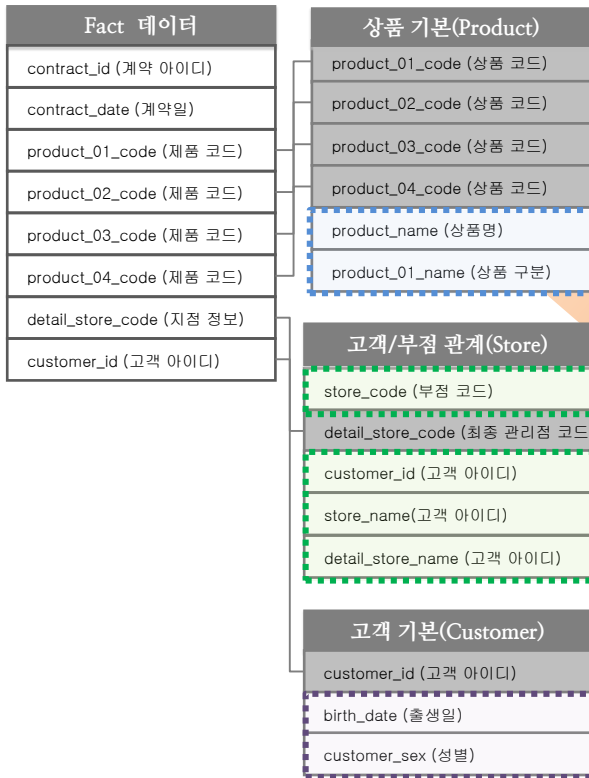
데이터를 metatron에 직접 넣는 방식(Insert)과 연결(Link)하는 방식을 지원하여 모든 데이터 소스 활용 가능



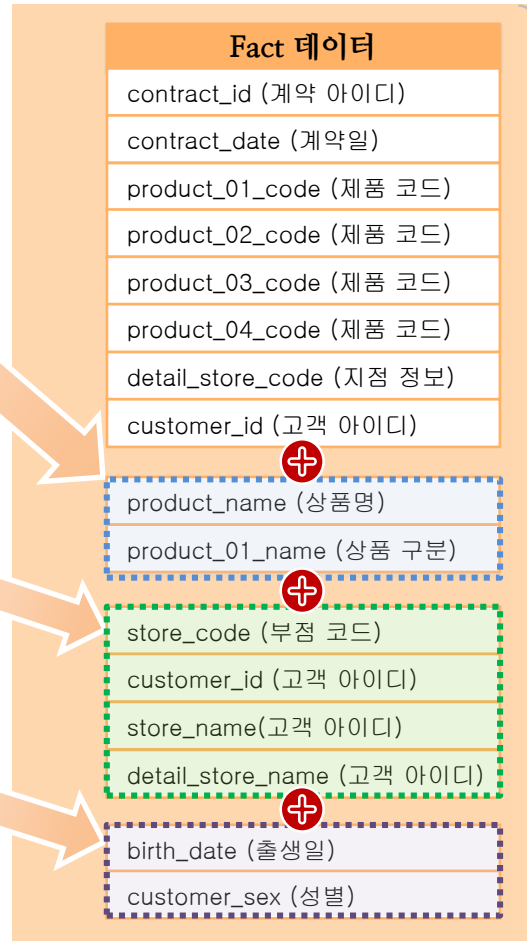
# 주요 특징 - Big OLAP Cube

❖ 대용량 Fact 데이터를 기준으로 다양한 Dimension 데이터를 결합하여 하나의 Big OLAP Cube(Mart) 생성

## 기존 방식



## Big OLAP Cube



조인하여  
Big OLAP  
Cube 생성

### 장점 ① ETL 비용 최소화

- Mart 의 생성을 최소화 함에 따라 Mart 생성을 위한 ETL 비용 감소
- Mart 데이터 수가 적어짐에 따라, 구조 변경에 따른 영향도를 최소화 할 수 있음
- Fact 데이터를 모두 저장하므로, 다양한 요구사항에 대응 가능

### 장점 ② 빠른 속도 및 스키마 변경 수용

- 분산 아키텍처 기반으로 큰 규모의 데이터도 저장 가능하며, 빠른 속도로 결과 출력 가능
- Dynamic 스키마 채용으로 스키마 변경 시에도 별도의 스키마 재정의가 필요치 않음

### 장점 ③ 실시간 처리

- 테이블을 원본 그대로 저장하므로 레코드 단위의 실시간 처리 가능

- ❖ Druid는 '12년부터 ImplyData社와 MetaMarkets社가 Open Source로 만든 Time Series에 최적화된 Columnar Store 기반의 데이터 처리 엔진
- ❖ SKT는 ImplyData社와 개발 협력('16~'17) 및 Druid Open Source 활동 중('16~)
- ❖ Metatron 상용 적용에 필요한 핵심 기능을 직접 개발하여 현재는 별도 Branch化하여 개발 중

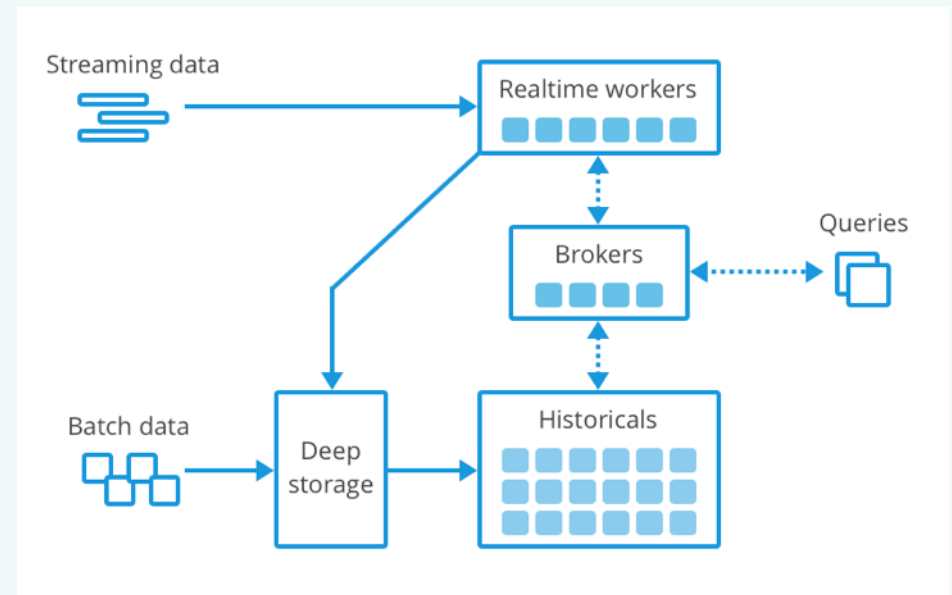
## ■ Open Eco-system 활동

- Druid 사용자 모임(meetup) 한국 Host
- Druid github, google groups 에 의견 개진 및 코드 기여 (300+)

## ■ 오픈 소스 단순 사용이 아닌 핵심 기능, 성능 향상 등 직접 개발 (별도 Branch)

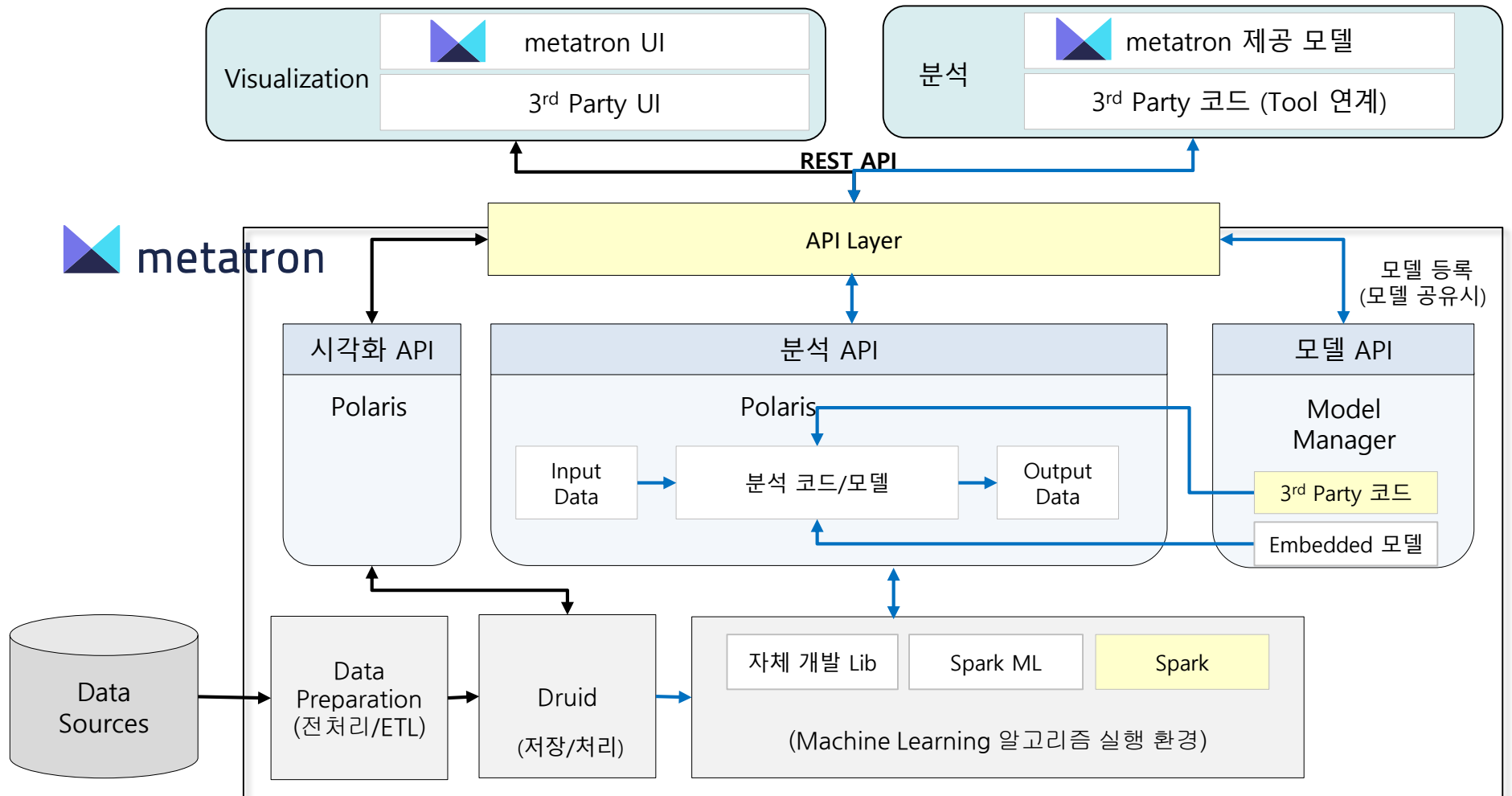
- 데이터 소스간 Join
- 검색 기능 (Lucene index, Spatial index)
- 쿼리 확장 (Sketch query)
- Query 통계 기능
- Window functions
- Virtual column map type
- Multi-valued metric
- ...

- 실시간과 배치 모두를 지원해 배치 분석 뿐 아니라 실시간 처리 가능
- 시간에 따라 in-memory, local-storage, deep storage로 데이터가 이관되어 TB 이상의 대용량 데이터 지원
- 각각의 기능(질의 처리, 저장, 색인)은 별도의 서버로 분리 가능하여 사용량 증가 시에도 Scale out 가능



## 주요 특징 - Notebook/Model Manager (외부 연동)

- ❖ Metataron은 고급 분석을 위한 분석 도구 연동 지원 (현재 Jupyter 통한 R, Python 연계)
- ❖ Jupyter에서 metataron에 저장된 데이터를 활용하여 분석하고, 개발한 분석 코드 및 실행 결과는 타인과 공유 (공유를 위해 Jupyter 분석 코드는 metataron API를 이용해 Model Manager에 등록되어 최신 데이터 연동)
- ❖ REST 기반의 API 를 통해 사용자 정의 UI 개발 가능



# Self-service Solution for Big Data Discovery

All-in-one analytics from easy big data preparation to fast visualization

metatron

[DOWNLOAD](#)

[LEARN MORE](#)

# Roadmap



## UNIFIED SOLUTION

Metatron is a single product for field users enabling data analysis through interactive web UI



## BIG DATA HANDLING

Metatron is powered by a massively scalable column store: Druid which enables us to create big OLAP cube!



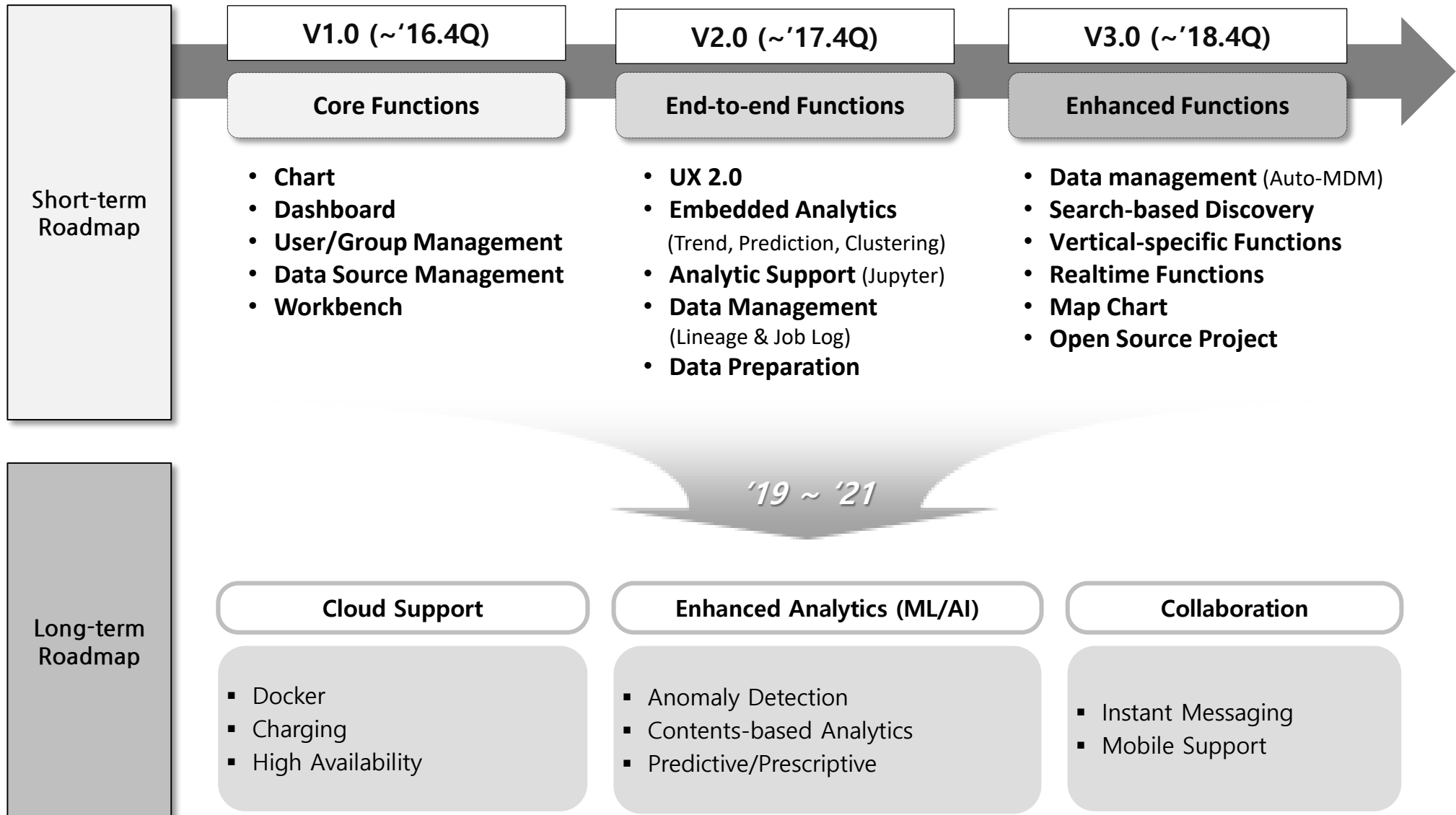
## POWERFUL MANAGEMENT

Manage data smart through all connectable data sources and visualized data monitoring features

## Feature Highlights



# Roadmap



# Roadmap – Automated MDM (Meta-Data Management)

- ❖ Hive, Presto 중심의 Meta 정보 관리
- ❖ Event Hook 기반의 SQL 모니터링을 통해 Meta 정보의 자동 업데이트 반영
- ❖ 향후 Advanced Search-based Discovery 에 활용

The image displays three overlapping screenshots of the METATRAN DISCOVERY web interface. The top-left screenshot shows the 'Explore Data' page with a search bar and a sidebar for 'Catalogs'. The middle screenshot shows a detailed view for '세일즈 데이터' (Sales Data) with tabs for Overview, Information, Columns, and Lineage. The bottom-right screenshot shows a 'Lineage' diagram illustrating data flow from 'Table 02' to 'NewTable'.

**Overview of Sales Data (세일즈 데이터)**

Tag	#
세일즈	11
CEI	1
Tech	1

Top User		
User	Last used	Used in
TammarnayTamma...	yyyy-mm-dd hh:mm:ss	Workspace name 001 > Dash...
Julia	yyyy-mm-dd hh:mm:ss	Workspace name 002 > Dash...
Jens	yyyy-mm-dd hh:mm:ss	Workspace name 003 > Dash...

Recently Updated		
User	Last used	Used in
TammarnayTamma...	yyyy-mm-dd hh:mm:ss	Information > Description
Julia	yyyy-mm-dd hh:mm:ss	Columns > City
Jens	yyyy-mm-dd hh:mm:ss	Columns > ORG_SORT_CD

**Lineage Diagram:**

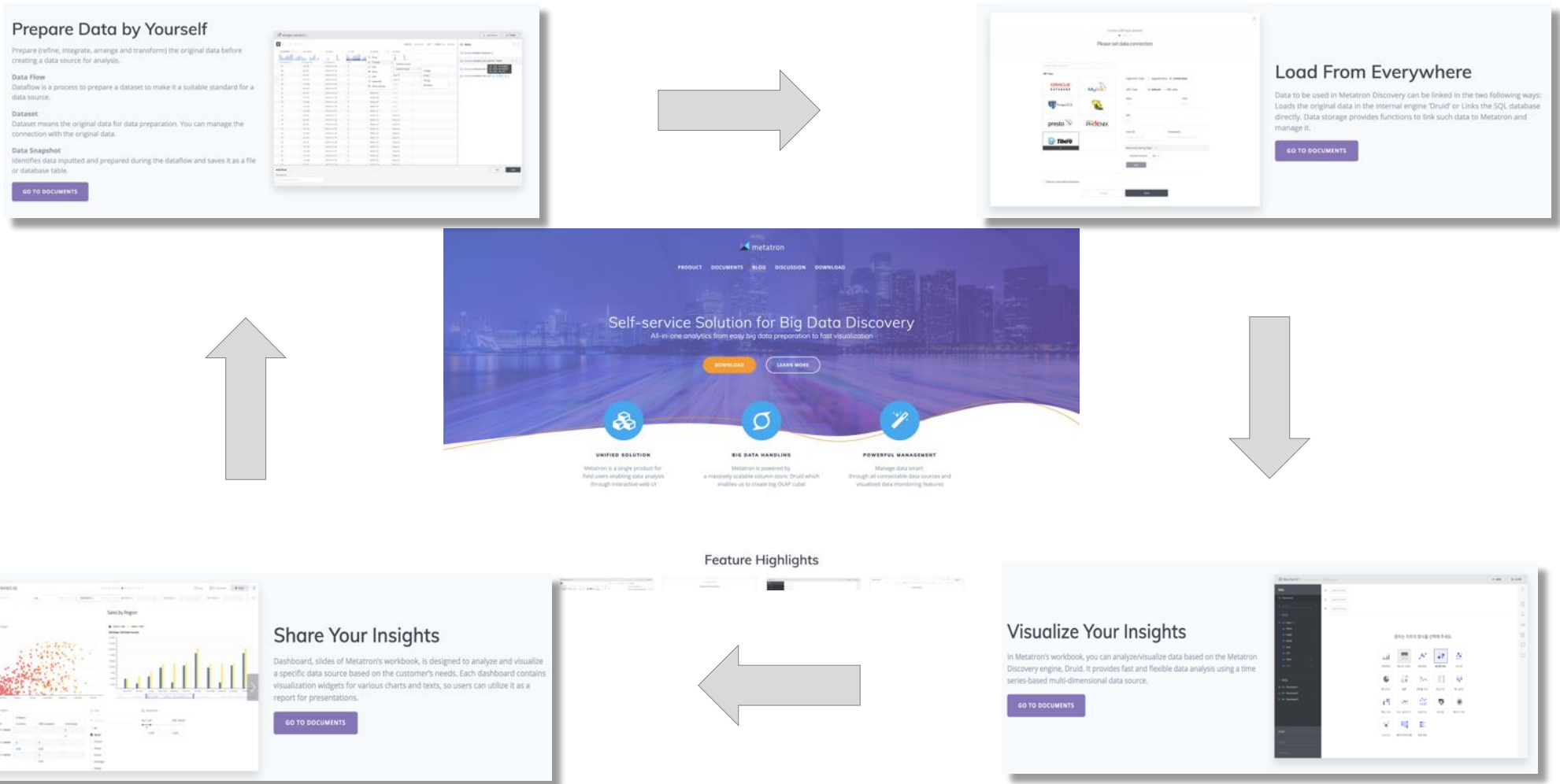
```
graph LR; Table02[Table 02] --> Select[Select * from Table where a = 3 Select * fr...]; Select --> NewTable[NewTable];
```

**Table 02 Columns:** OrderDate, Category, City, Country, CustomerName, OrderID, PostalCode, Quantity.

**NewTable Columns:** custom\_id, product02, code\_01\_int, code\_02\_int, code\_03\_int, code\_04\_int.

# Roadmap – Open source project

- ❖ <https://metatron.app>
- ❖ 7월경 오픈소스화 예정



## About us





Join us