



AI·DATA  
AI & DATA ANALYTICS CENTER

DP 360 Catalog

메가존클라우드 DataPlatform360 Catalog

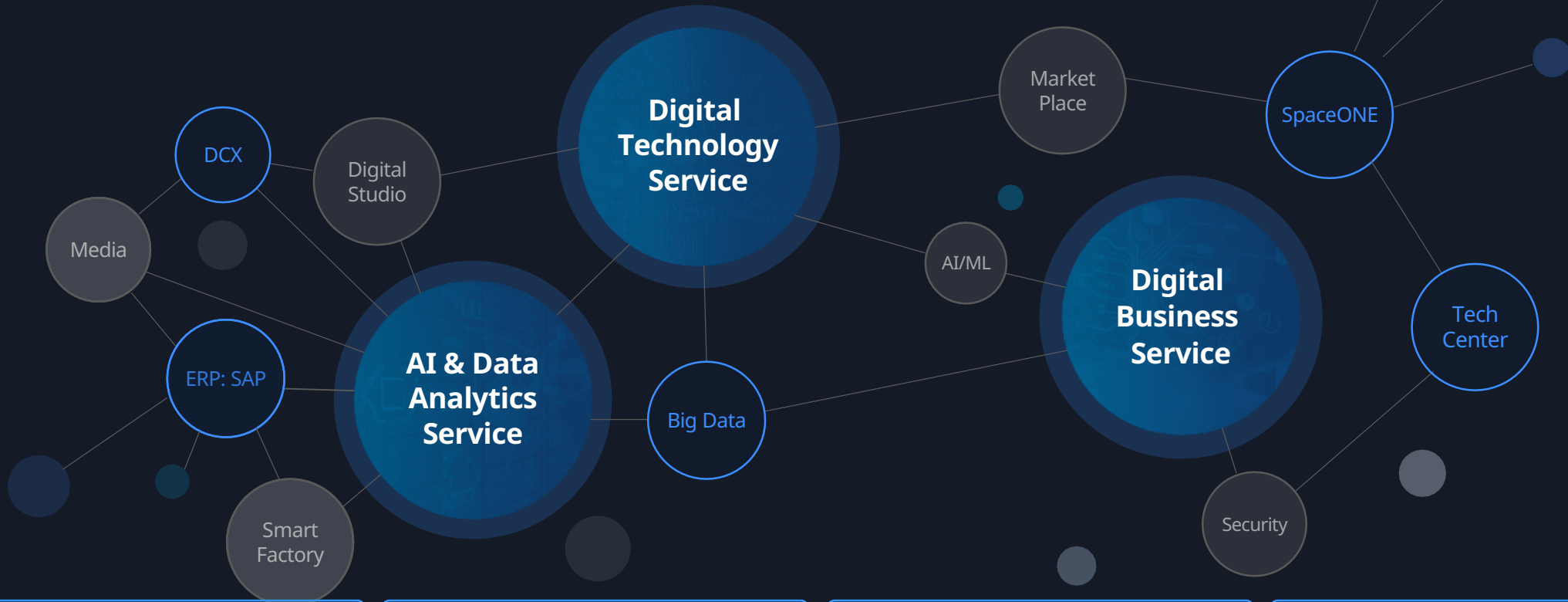
Winter 2024

AI & 데이터분석센터, 메가존클라우드



# 메가존클라우드 비즈니스 & 서비스

메가존클라우드는 기업 고객의 클라우드 도입부터 디지털 트랜스포메이션을 위한 컨설팅, 구축 및 운영 관리를 포함한 다양한 서비스를 통해 기업 고객의 성공적인 디지털 전환과 최적화 지원



**AI & Data Analytics**  
 데이터 분석 플랫폼 컨설팅 및 구축  
 AI/ ML 데이터 분석 컨설팅 및 구축

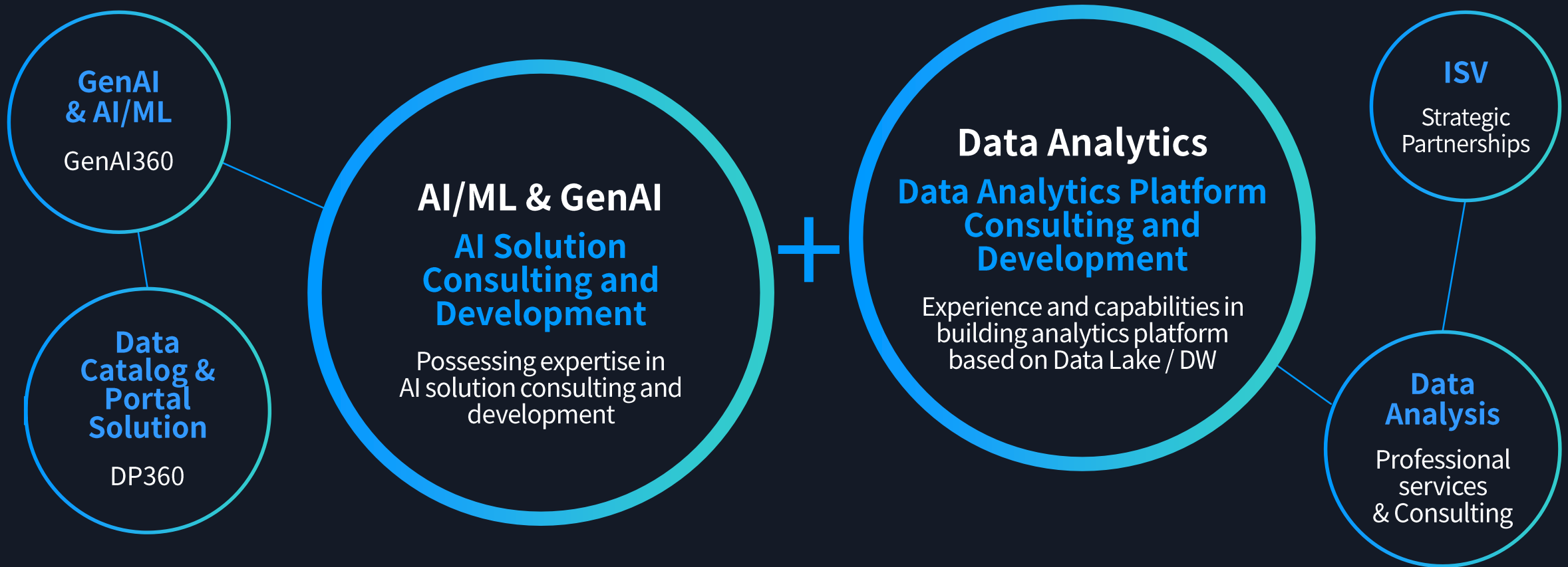
**Infrastructure**  
 디지털 혁신을 위한 최적화된 인프라

**Platform**  
 디지털 혁신 가속화를 위한  
 애플리케이션 개발/ 운영/ 관리 플랫폼

**Application**  
 신속한 디지털 비즈니스 의사결정 및 관리를  
 위한 비즈니스 애플리케이션

# AI & Data Analytics Center (ADC)

AI & Data Analytics Center (ADC) 는 엔터프라이즈 AX를 위한 AI & 데이터 분석 전문 조직으로,  
기업이 **실제적인 성과**를 이룰 수 있도록 **“성공적인 AI 경험”** 제공



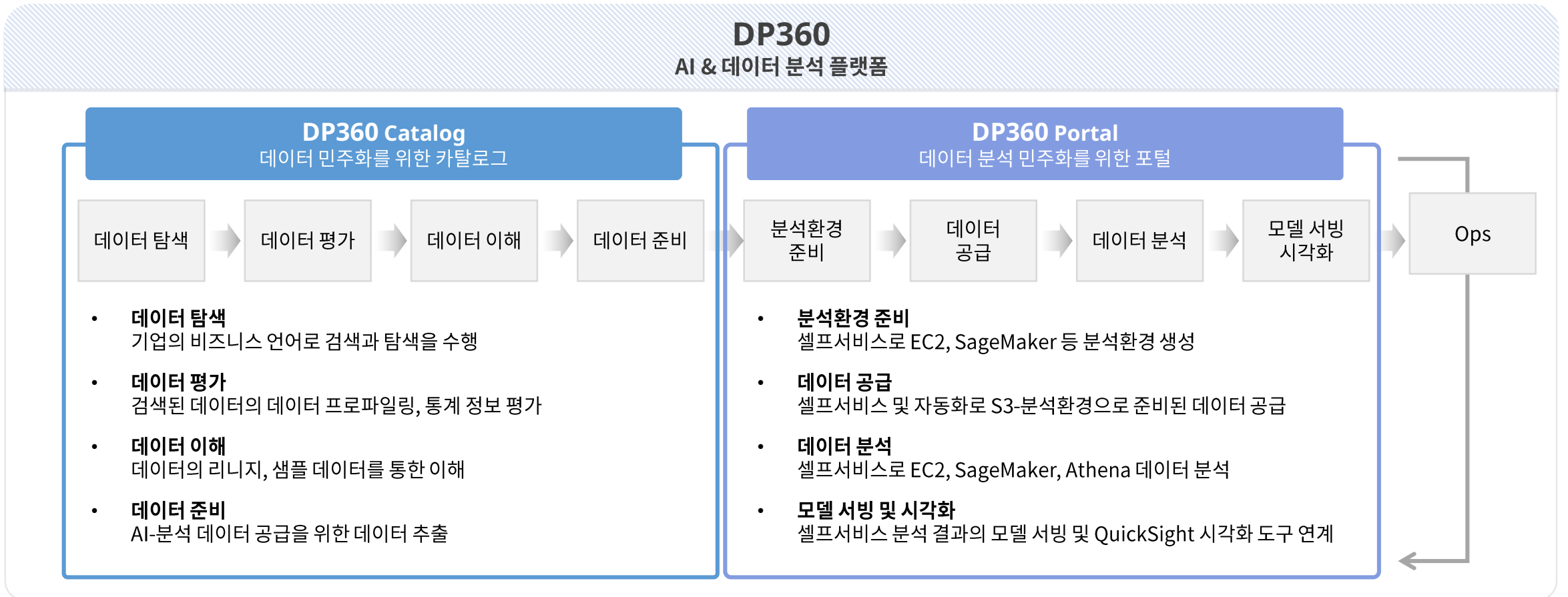
# 1. DP360, DataPlatform360

AI & Data Analytics Platform



## DP360 (Data Platform 360)

‘데이터 자산화 - 가치 탐색 - 가치 평가 - AI 분석 활용’ 등 기업의 데이터 혁신을 위한 프로세스 전반을 지원하는 AWS 기반의 엔터프라이즈 AI & 데이터 분석 플랫폼



## 2. DP360 Catalog

데이터 민주화, 강력한 데이터 거버넌스를 위한 차세대 데이터 카탈로그



## 데이터 민주화, 강력한 데이터 거버넌스를 위한 차세대 데이터 카탈로그

The image shows the DP360 Catalog interface with several callout boxes highlighting key features:

- 데이터 인지**: 분석 또는 비즈니스 문제 해결을 위한 접근 가능한 모든 데이터에 대한 정보 제공
- 데이터 이해**: 접근 가능한 모든 데이터의 의미, 데이터의 생성 경로와 사용법 등을 이해
- 최신 데이터 스택**: 최신 데이터 스택
- 데이터 발견**: 분석에 필요한 데이터 간 조합을 통한 빠른 탐색
- 데이터 분석**: 대량의 데이터간의 관계와 다차원 분석의 한계를 극복한 분석 데이터의 이해
- 분석 결과 배포**: 의사결정에 필요한 데이터, 프로세스 개선을 위한 데이터 등의 분석 결과 배포

### 데이터 디스커버리

- 수집된 데이터를 다양하게 검색 및 탐색할 수 있는 환경 제공

### 데이터 거버넌스

- 데이터 자산 활용을 위한 데이터 소유, 보안 및 접근 권한 제어

### 생성형 AI

- 생성형 AI 기술 활용 검색 데이터 설명, 자동 태깅, rule-based 개인정보 탐지 제공

### 데이터 분류 및 태깅

- 데이터 분류 및 태깅을 통한 데이터 구조화
- 생성형 AI를 활용한 자동 태깅 및 데이터 설명

### 데이터 품질 모니터링

- 데이터 프로파일링 통계 및 품질 관리

### 데이터 시각화

- 수집된 데이터의 분포도 등 데이터 현황을 시각적으로 표현

사용자의 빠르고 쉬운 데이터 탐색 및 이해를 위해 생성형AI를 활용한 편리한 기능 제공



생성형 AI  
기술 활용

검색 데이터 셋 설명 제공

태깅 및 데이터 설명 생성

rule-base 개인정보 탐지

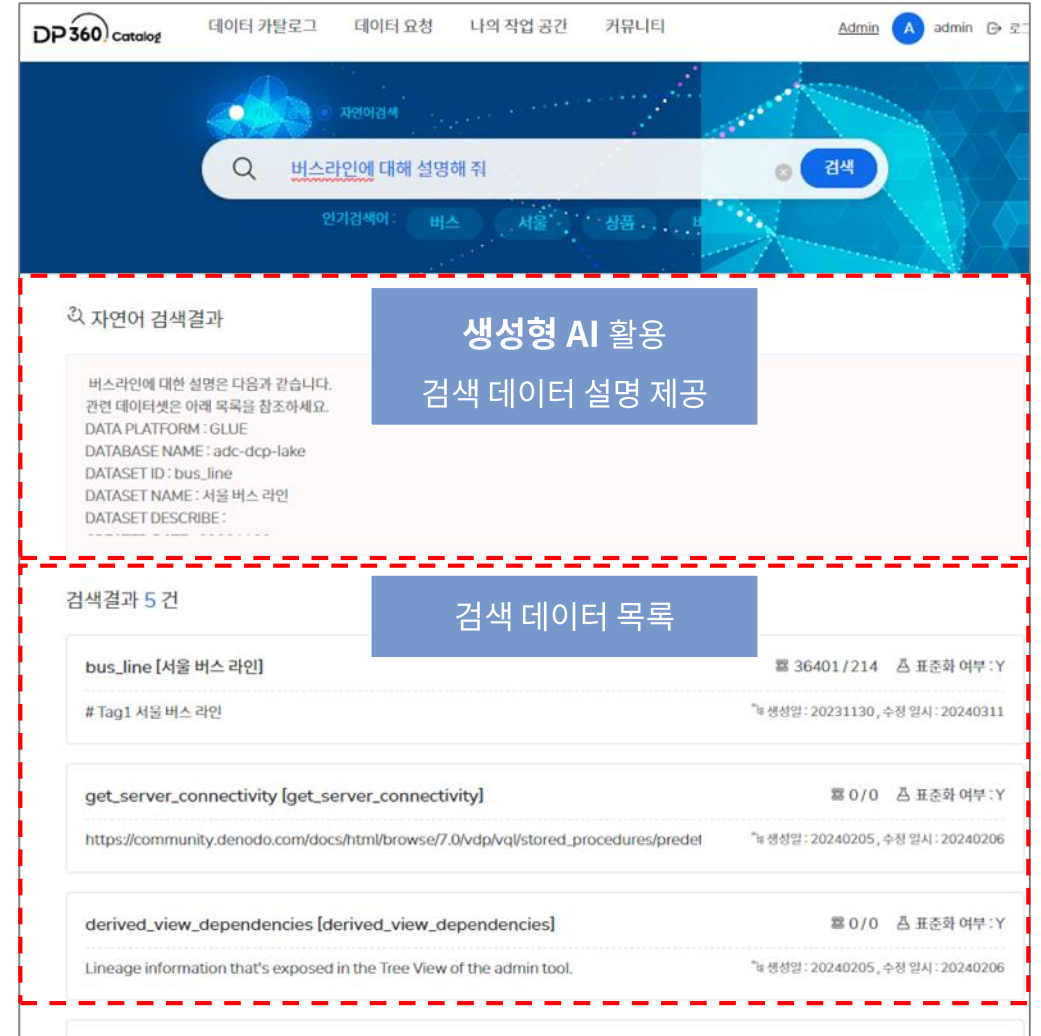
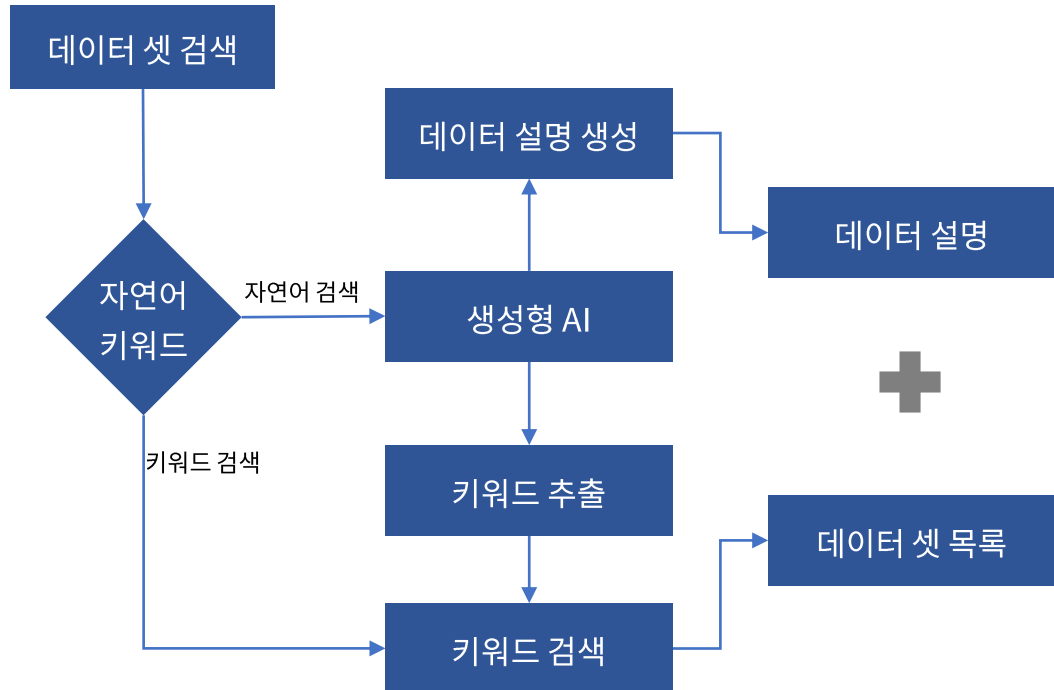


빠르고 쉬운  
데이터 탐색과 이해



# DP360 Catalog + GenAI > 검색 데이터셋 설명 제공

## 검색 데이터셋 설명 및 연관 데이터셋 제공



## AI 태깅 / AI 설명으로 편리한 데이터 셋 검색 정보 생성

데이터셋 컬럼 정보 샘플 데이터 추출조건 분포도

데이터셋: kangju\_liquid\_company      데이터명: 광주시 국제 물류      저장

관계사	MZC	소유자	입력해 주세요	<span style="float: right;">검색</span>
UUID	dbc6c070-0c29-4413-87bb-679148aa8c5d	사용여부	<span style="color: blue;">사용중</span>	
플랫폼	GLUE	DB명	adc-dcp-lake	
* 데이터셋 ID	kangju_liquid_company	* 데이터셋명	광주시 국제 물류	
* 물리DB명	adc-dcp-lake	표준화 여부	<span style="color: blue;">표준화</span>	
용량	549	건수	8	
생성일	2023-11-30	수정일	2024-07-22	

**태깅 및 설명 정보 수정 및 반영**

Tag: #GLUE #adc-dcp-lake #kangju\_liquid\_company #경기도광주시\_국제물류주선업체현황 #20231130 #20240311

설명: 이 데이터셋은 경기도 광주시에 위치한 국제물류주선업체들의 현황 정보를 담고 있습니다. 데이터베이스 이름은 adc-dcp-lake이며, 데이터셋 ID는 kangju\_liquid\_company입니다. 데이터셋은 2023년 11월 30일에 생성되었고, 가장 최근에는 2024년 3월 11일에 업데이트되었습니다.

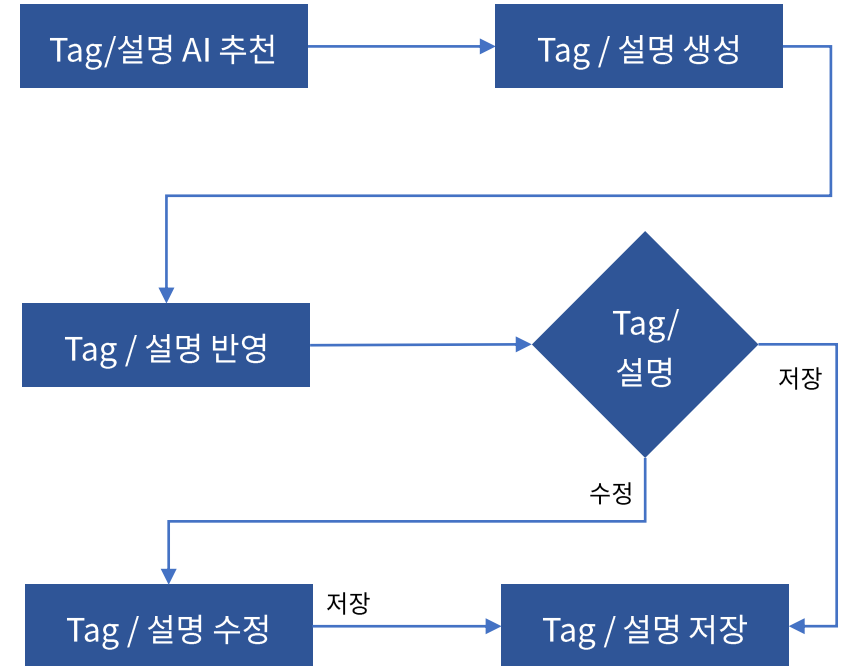
Tag / 설명 AI 추천    버튼 클릭 시, AI를 사용하여 가장 적합한 태그와 설명을 추천합니다. 추천된 내용을 반영 버튼을 통해 반영할 수 있습니다.

**태깅 및 설명 정보 생성**

AI Tag: #GLUE #adc-dcp-lake #kangju\_liquid\_company #경기도광주시\_국제물류주선업체현황 #20231130 #20240311      Tag 반영

AI 설명: 이 데이터셋은 경기도 광주시에 소재한 국제물류주선업체들의 현황 정보를 담고 있습니다. 데이터베이스 이름은 adc-dcp-lake이며, 데이터셋 ID는 kangju\_liquid\_company입니다. 데이터셋은 2023년 11월 30일에 생성되었고, 2024년 3월 11일에 최근 업데이트되었습니다. 이 데이터셋을 활용하면 광주시 지역 내 국제물류주선업체의 분포와 현황을 파악할 수 있을 것입니다.      설명 반영

## Tag / 설명 생성 및 저장



데이터 탐색 편의성 제공을 위한 생성형 AI 기술 활용

## rule-based 개인정보 탐지로 안전한 데이터 탐색 환경 제공

①

정보보안등급	정보보안분류명	정보보안컬럼목록	포맷
높음	주민등록번호	rrn,jumin_number,registration_number,jumin_no,ssn,citizen_id,unique_identifier	^d{2}(0 1-9){1}0-2)(0 1-9)\d{3}1)-[1-4]\d{6}\$
높음	여권번호	passport_number,passport_no,passport_id,no,number,pp_no	시A-기1 2)\d{6}\$
높음			
중간			
중간			
중간			
낮음			
낮음			

② 개인정보규칙관리 변경

정보보안등급:  높음  중간  낮음

정보보안분류: 주민등록번호

정보보안컬럼목록: rrn X jumin\_number X registration\_number X jumin\_no X ssn X citizen\_id X unique\_identifier X +

포맷: ^d{2}(0|1-9){1}0-2)(0|1-9)\d{3}1)-[1-4]\d{6}\$

③ 러셋 권한 - 개인정보 컬럼 조회/접근 권한 설정

데이터셋 권한

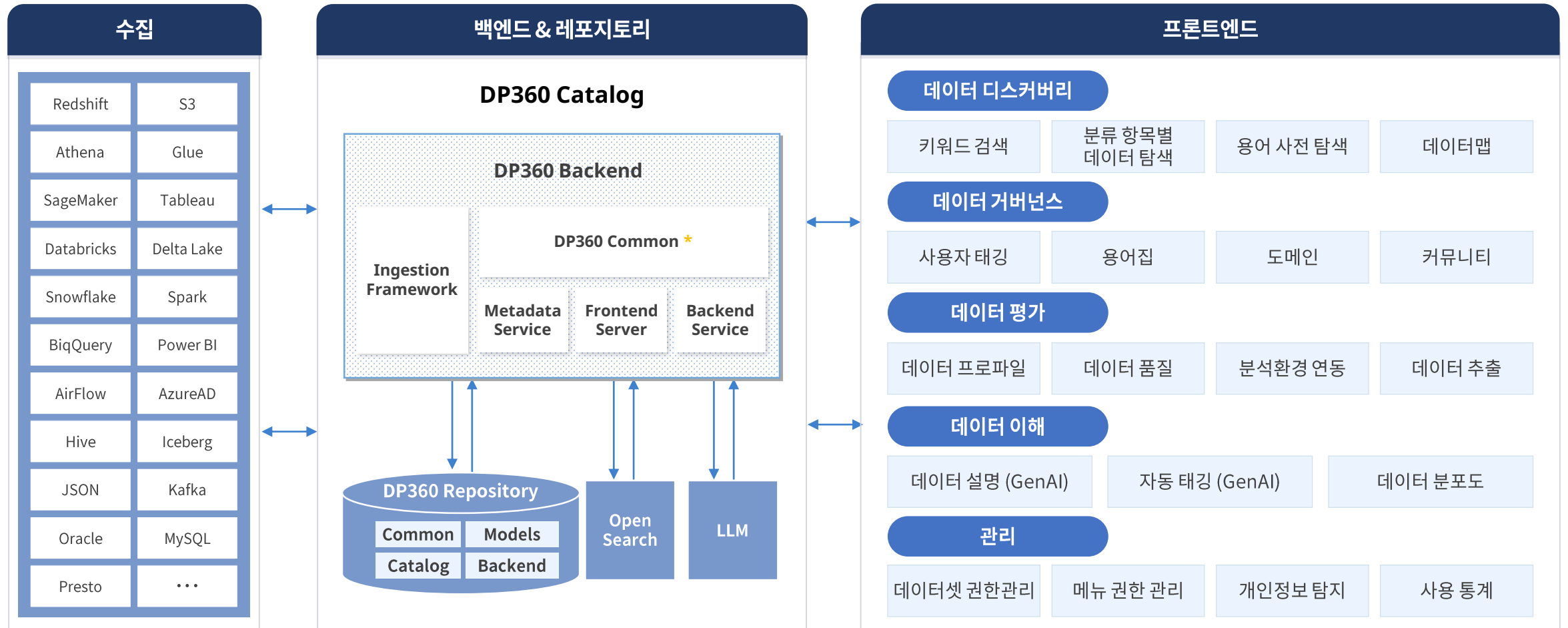
데이터셋 권한 ID	데이터셋 권한 명	데이터셋명	컬럼명	데이터유형	개인정보등급	개인정보분류	조회	접근
Admin-Full	전체 권한	dcp_table1	address	string	중간	집주소	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TEST	테스트	dcp_table1	email	string	낮음	이메일 주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table2	address	string	중간	집주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table2	email	string	낮음	이메일 주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table3	address	string	중간	집주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table3	email	string	낮음	이메일 주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table4	address	string	중간	집주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table4	email	string	낮음	이메일 주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table5	address	string	중간	집주소	<input type="checkbox"/>	<input type="checkbox"/>
		dcp_table5	email	string	낮음	이메일 주소	<input type="checkbox"/>	<input type="checkbox"/>
		heartrate_device	name	string	중간	이름	<input type="checkbox"/>	<input type="checkbox"/>

### 개인 정보 탐지 설정

- ① 정보보안대상 컬럼 설정
  - 컬럼별 보안 등급 설정
  - 정보보안 분류 설정 및 컬럼 목록 설정
  - 탐지 rule 설정
- ② 개인정보 관리 Rule 설정
  - 컬럼별 보안 등급 설정
  - 정보보안 분류 설정 및 컬럼 목록 설정
  - 탐지 rule 설정
- ③ 권한별 Rule 적용 설정
  - 데이터 셋 권한별 조회 및 접근 권한 설정

# DP360 Catalog 아키텍처

## 엔드투엔드 데이터 검색, 데이터 디스커버리 및 데이터 거버넌스 지원



\* DP360의 Common 모듈을 DP360 Portal과 공유

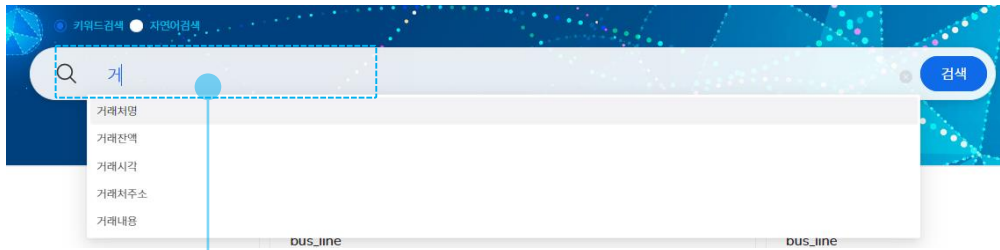
## 데이터 카탈로그 메인

The screenshot shows the DP360 Catalog main interface. At the top, there's a navigation bar with 'Admin', 'admin', and '로그아웃' links. Below it is a search bar with the text '키워드/자연어 검색' and a search button. The search bar contains the text '검색어를 입력하세요.' Below the search bar are tabs for '상품', '상품명', '품명', '버스', and '비'. The main content area is divided into three columns: '즐거찾기' (Discover), '최근에 본 데이터 셋' (Recently Viewed), and '인기 데이터 셋' (Popular). Each column lists various data sets with their names and brief descriptions.

### 주요 메뉴

- **Catalog 탐색**
  - 수집 시스템별 데이터 셋 탐색
  - 트리구조 및 데이터 맵기 기반 탐색 환경 제공
- **Data 검색**
  - 데이터 셋 키워드 검색 제공
  - 자연어 검색 제공
- **즐거찾기 데이터 셋**
  - 사용자별 자주 사용하는 데이터 셋의 즐겨찾기 기능 제공
- **최근 검색 데이터 셋**
  - 최근에 검색, 참고된 데이터 셋 목록 제공
- **인기 데이터 셋**
  - 조직내/사내에서 자주 사용되는 데이터셋 목록 제공

## 데이터셋 검색 편의 기능



데이터셋 검색  
자연어/키워드

자연어 검색 결과  
생성형 AI + 키워드 검색

자연어 검색결과

5. 순환버스 (노란색) 특정 지역 내에서 순환 운행하는 노선으로, 주로 도심 내 소규모 구역을 연결합니다.  
6. 마을버스 (노란색) 작은 버스로, 주로 주거 지역과 가까운 지역이나 간선버스 정류장 등을 연결하는 역할을 합니다.  
서울시 버스 노선은 사용자의 편의를 위해 정기적으로 조정되며, 노선 정보는 서울시 대중교통 포털이나 서울시 버스정보시스템과 같은 온라인 플랫폼을 통해 확인할 수 있습니다. 이러한 시스템들은 실시간 버스 위치 정보, 도착 예정 시간, 노선 검색 기능 등을 제공하여 이용자들이 보다 편리하게 버스를 이용할 수 있도록 돕습니다. 관련 데이터셋은 아래 목록을 참조하세요.  
- DATA PLATFORM : GLUE, DATABASE NAME : adc-dcp-lake, DATASET ID : bus\_line, DATASET NAME : 서울 버스 라인  
- DATA PLATFORM : GLUE, DATABASE NAME : adc-dcp-lake, DATASET ID : seoul\_trans\_utility, DATASET NAME : 서울교통공사\_기계시설정보

검색결과 5 건

bus_line [서울 버스 라인]
# Tag1 서울 버스 라인
seoul_sub_rent_info [서울 지하철가 임대정보]
서울교통공사 지하철가 임대 정보 # Tag1 서울 교통 지하 상가 임대

키워드 검색 결과

검색결과 7 건

daegu\_bridge [대구광역시 교량제원 현황]

# Tag1 대구 교량

kangju\_liquid\_company [경기도 광주시\_국제물류추진업체 현황]

# Tag1 광주 주류

sr\_monthly\_usage [SR 월별역별차차인원]

# Tag1 SR 차차인원

데이터 셋 상세 조회

데이터셋 ID	bus_line	용량	36,401
데이터셋 명	서울 버스 라인	건수	214
데이터 플랫폼	GLUE	데이터 베이스	adc-dcp-lake
		관계사	MZC
		소유자	admin
		생성일	2023-11-30
Tag	서울 버스 라인	수정일	2024-03-21

설명: 서울의 버스 노선에 대한 데이터입니다.

관련 정보 >

샘플 조회 >

분포도 (X: line\_nm, Y: trans\_type) >

관련 정보

순서 ↓	컬럼 영문명 ↓	컬럼 한글명 ↓	데이터 타입 ↓	데이터 길이 ↓
1	trans_type	trans_type	string	0
3	line_nm	line_nm	string	0
4	station_id	station_id	string	0
5	station_nm	station_nm	string	0
6	station_count	station_count	bigint	0

샘플 조회

trans_type ↓	line_nm ↓	station_id ↓
세종시(일반)	540번(장군면사무소~장군면사무소)	6800133
세종시(일반)	540번(장군면사무소~장군면사무소)	6800134
세종시(일반)	540번(장군면사무소~장군면사무소)	6800124
세종시(일반)	550번(산성동종점~조치원역)	6801913
세종시(일반)	550번(산성동종점~조치원역)	6800655
세종시(일반)	550번(산성동종점~조치원역)	6800656

## 목록 기반 데이터 탐색

데이터 카탈로그 > 목록

카테고리명 또는 데이터셋명을 입력하세요.

- MZC
  - SALESFORCE
  - DATABRICKS
  - GLUE
    - ADC-DCP
      - 서울 버스 라인
      - 대구광역시 교량제한 현황
      - 경기도 광주시\_국제물류주선업체 현황
      - 서울 지하철개업예정정보
      - 서울교통공사\_기게시설정보
      - SR 광역역별역사인원
      - 국도교통부 철도특별사법경찰대\_철도방법용 영상 감시설비
    - CATE-1-2
    - DENODO

데이터셋 ID	bus_line	용량	36,401
데이터셋명	서울 버스 라인	건수	214
데이터 플랫폼	GLUE	데이터 베이스	adc-dcp-lake
설명	서울의 버스 노선에 대한 데이터입니다.	관계사	MZC
Tag	서울 버스 라인	소유자	admin
		생성일	2023-11-30
		수정일	2024-03-21

트리구조 데이터 셋 탐색

컬럼 정보 > 샘플 조회 > 분포도 (X: line\_nm, Y: trans\_type)

컬럼 정보

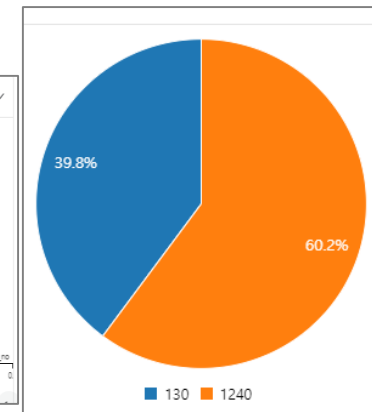
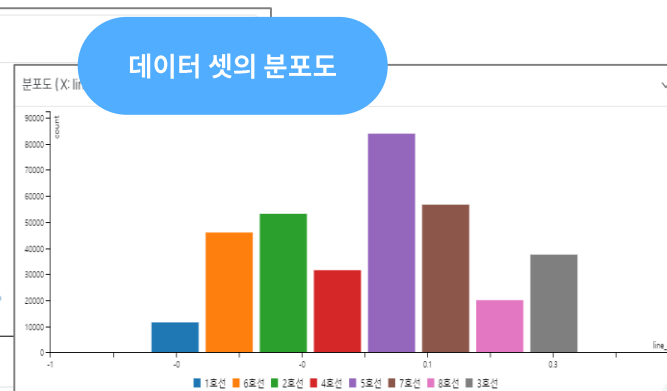
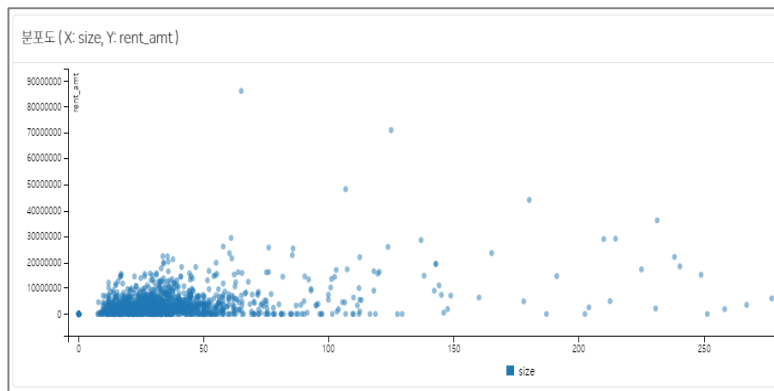
순서	컬럼 영문명	컬럼 한글명	데이터 타입	데이터 길이
1	trans_type	trans_type	string	0
3	line_nm	line_nm	string	0
4	station_id	station_id	string	0
5	station_nm	station_nm	string	0
6	station_count	station_count	bigint	0

컬럼 정보 조회

샘플 조회

trans_type	line_nm	station_id
세종시(일반)	540번(장군면사무소~장군면사무소)	6800133
세종시(일반)	540번(장군면사무소~장군면사무소)	6800134
세종시(일반)	540번(장군면사무소~장군면사무소)	6800124
세종시(일반)	550번(신성동종점~조치원역)	6801913
세종시(일반)	550번(신성동종점~조치원역)	6800655
세종시(일반)	550번(신성동종점~조치원역)	6800656

샘플 데이터 조회



## 맵 기반 데이터 탐색

데이터 카탈로그 > 맵

카테고리명 또는 데이터셋 이름을 입력하세요.

- MZC
  - SALESFORCE
  - DATABRICKS
  - GLUE
    - ADC-DCP
    - CATE-1-2
    - DENODO

맵 기반 데이터 셋 탐색

데이터셋 ID	bus_line	용량	36,401
데이터셋명	서울 버스 라인	건수	214
데이터 플랫폼	GLUE	데이터베이스	adc-dcp-lake
설명	서울의 버스 노선에 대한 데이터입니다.	관계사	MZC
		소유자	admin
		생성일	2023-11-30
Tag	서울 버스 라인	수정일	2024-03-21

데이터 문의 | 데이터추출

맵 기반 데이터 정보

데이터 카탈로그 > 통계

3 주간 사용자  
전주 대비 비율

3 월간 사용자  
전월 대비 비율

234 데이터셋 수집 현황  
소유자:테그명, 분야:설명명, 도메인명

주간/월간 사용자 현황 및 수집 현황

실 접속 현황 및 검색 현황

주간 상위 검색어

검색어	횟수
상용	35
버스	29
서울	28
비	23
상품명	23

데이터 플랫폼 현황

주간 최다 조회 데이터셋

데이터셋	조회수
서울 버스 라인	90
국도교통부 철도특별사법경찰대_철도법범용 영상감시시스템비	80
서울 지하철가 입대정보	44
AllActivatedContracts	26
dcp_table1	25

검색 상위 현황



# DP360 Catalog 스크린샷 > ④ 나의 작업 공간

## 데이터 추출 요청 / 분석 환경 관리

나의 작업 공간 > 데이터활용이력

데이터활용이력 분석환경관리

Jupyter Notebook

보유 Jupyter Notebook  
Notebook Type: ml.t3.medium / 5 GiB  
Notebook 상태: Stopped

**노트북 정보**

등록일 2024-02-21 to 2024-03-21 데이터셋 전체 검색어 입력 상태 전체 검색

데이터셋 ID	데이터셋 명	플랫폼	DB 명	추출일시	상태	다운로드
bus_line	서울 버스 라인	GLUE	adc-dcp-lake	2024-03-21 13:28:28	SUCCEEDED	다운로드
dcp_table2	dcp_table2	DATABRICKS	dcp_db	2024-03-19 00:00:39	SUCCESS	다운로드
dcp_table5	dcp_table5	DATABRICKS	dcp_db	2024-03-18 23:58:33	FAILED	다운로드
bus_line	서울 버스 라인	GLUE	adc-dcp-lake	2024-03-18 00:59:18	SUCCEEDED	다운로드
dcp_table1	dcp_table1	DATABRICKS	dcp_db			다운로드
dcp_table3	dcp_table3	DATABRICKS	dcp_db			다운로드
seoul_trans_utility	서울교통공사 기계차량정보	GLUE	adc-dcp-lake			다운로드

**데이터 추출 요청 내용 및 데이터 다운로드**

데이터요청 > 데이터문의

기간 2024-02-20 to 2024-03-21 검색어 입력 신규 조회

No.	제목	작성일	상태
1.0	서울 버스 라인의 데이터	2024-03-21	

**데이터 자동 추출 / 권한 요청**

데이터요청 > 자동추출요청

요청일 2024-02-21 to 2024-03-21 데이터셋 전체 검색어 입력 상태 전체 검색

데이터셋 ID	데이터셋 명	플랫폼	DB 명	요청일	추출주기	추출시간	상태
---------	--------	-----	------	-----	------	------	----

데이터요청 > 데이터권한요청

요청일 2024-02-21 to 2024-03-21 검색어 입력 상태 검색어 주석요

데이터셋 ID	데이터셋 명	플랫폼	DB 명	요청일자	상태
---------	--------	-----	------	------	----

나의 작업 공간 > 분석환경관리

데이터활용이력 분석환경관리

Jupyter Notebook

보유 Jupyter Notebook  
Notebook Type: ml.t3.medium / 5 GiB  
Notebook 상태: Stopped

**분석 노트북 타입 및 사용기간 조회**

신청일 2024-02-21 to 2024-03-21 상태 전체 인스턴스타입 전체 검색

신청일	사용기간	상태	기존 노트북타입	기존 용량	신규 노트북타입	신규 용량
2024-03-18	2024-03-18 ~ 2024-04-30	신청	ml.t3.medium	5	ml.t3.medium	10

노트북 타입 변경 신청

분석환경 업그레이드

현재 노트북 타입 ml.t3.medium 현재 용량 5 GiB

신규 노트북 타입

- ml.t3.medium 2vCPU 4GiB(Memory)
- ml.t3.large 2vCPU 8GiB(Memory)
- ml.t3.xlarge 4vCPU 16GiB(Memory)

신규 용량 입력해 주세요 GiB 사용 기간 2024-03-2 to 2024-03

취소 신청

## 도메인 / 단어 / 용어 / 동의어,유의어 관리

### 표준 도메인 그룹/분류 관리

도메인그룹명 입력해 주세요 [검색]

도메인 그룹

No.	삭제	사제	도메인그룹명	도메인그룹설명
1	<input type="checkbox"/>	<input type="checkbox"/>	금액	입력해 주세요
2	<input type="checkbox"/>	<input checked="" type="checkbox"/>	납액시간	입력해 주세요
3	<input type="checkbox"/>	<input type="checkbox"/>	내용	입력해 주세요
4	<input type="checkbox"/>	<input type="checkbox"/>	명칭	입력해 주세요
5	<input type="checkbox"/>	<input type="checkbox"/>	번호	입력해 주세요
6	<input type="checkbox"/>	<input type="checkbox"/>	수량	입력해 주세요
7	<input type="checkbox"/>	<input type="checkbox"/>	일	입력해 주세요
8	<input type="checkbox"/>	<input type="checkbox"/>	코드	입력해 주세요

**도메인 그룹 정의**

### 표준 단어 관리

단어정보 입력해 주세요 [검색] [등록]

No	단어약어명	단어영문명	단어한글명	단어설명	도메인분류명	동의어목록	금칙어목록	비전	차수	공공	최종수정자	최종수정일시	
1	ABER	Abbreviation	약어		-				1	6	Y	system	2024-02-29 02:01
2	ABL	Abolition	폐지		-								
3	ABND	Abandonment	포기		-								
4	ABSN	Absence	부재	不在. 그곳에 있지 않음	-								
5	ABSTR	Abstract	적요	摘要. 요점을 뽑아 적음. 또는 그 기록	-								
6	ACAVSR	Academic Adviser	지도교수	指導教授. 자기가 직접 가르치는 학생이나, 학업 따위의 문제가 상담...	-								
7	ACBG	Academic Background	학력	學府. 학교를 다닌 경력	-								
8	ACOND	Actual Condition	실태	實態. 있는 그대로의 상태 또는 실재의 모양	-								

**단어 관리**

### 표준 단어 정보 변경

단어약어명: ABSN

단어영문명: Absence

도메인분류명: ::: 선택해 주세요 :::

동의어목록 (선택표로구분): 결석, 빠짐, 없음, 부재중, 불참, 불출석, 결근, 비어있음, 공백, 공허

금칙어목록 (선택표로구분): 입력해 주세요

단어 설명: 不在. 그곳에 있지 않음

등록일시: 2024-02-29 02:05:26

최종수정일시: 2024-02-29 02:05:26

**동의어목록**

### 표준 도메인 관리

도메인 그룹/분류 선택

도메인 그룹/분류 관리해 주세요

도메인 관리

No	도메인명	그룹명	분류명	대리자유형	길이	소수점	저장형식	비고
1	가액N10	금액	가액	NUMERIC	10	-	9999999999	9,99
2	가액I15	금액	가액	NUMERIC	15	-	9999999999999999	999,99
3	금액N13	금액	금액	NUMERIC	13	-	9999999999999	9,999

**도메인 그룹, 유형 정의**

### 표준 용어 관리

용어정보 입력해 주세요

No	용어영문약어명	용어한글명	도메인명	표준단어	표준단어영문약어	표준단어영문명
1	ABER_NM	약어명	영V300	약어, 명	ABER_NM	Abbreviation, Name
2	ABL_YMD	폐지일자	연월일C8	폐지, 일자	ABL_YMD	Abolition, YEARMonthDay
3	ABL_YN	폐지여부	여부C1	폐지, 여부	ABL_YN	Abolition, YessoNo
4	ABND_NOCS	포기건수				Abandonment, TheNumbe
5	ABND_RSN	포기사유				Abandonment, Reason
6	ABND_YMD	포기일자				Abandonment, YEARMont
7	ABSN_RSN	부재사유	내용V4000	부재, 사유	ABSN_RSN	Absence, Reason

**용어 관리**

단어 설명: 不在. 그곳에 있지 않음

등록일시: 2024-02-29 02:05:26

최종수정일시: 2024-02-29 02:05:26

오픈서치 동의어/유의어 배치

MZC OSSM 오픈서치동의어/유의어배치 수동배치 00011\*10 (데이터베이스 개입/방화벽/백업하는 용어/의미/단어/시를 등록하지 않는 배치입니다.)

MZC OSSM 오픈서치동의어/유의어배치 수동배치 00011\*10 (표준어-단어 동의어/유의어를 오픈서치에 적용하는 배치입니다.)

**도메인 그룹/분류**

**도메인 세부정보 등록 및 수정**

## 시스템 / 데이터베이스 / 테이블 / 컬럼 표준화 관리

**표준 정보 시스템 관리**

입력해 주세요 검색 등록

No	표준 데이터베이스 관리														
1	정보시스템 <input type="text"/> DB검색 <input type="text"/> <span style="float: right;">검색 등록</span>														
2															
3	No	물리DB명	논리DB명	DBMS종류	DBMS버전	구축일자	테이블수	최종수정자	최종수정일시						
4	1	표준 테이블 관리													
	2	정보시스템 전체 데이터베이스 <input type="text"/> <span style="float: right;">검색 등록</span>													
	3	No	정보시스템	물리DB명	테이블명	테이블한글명	소유자	테이블유형	보전기간	데이터건수	발생주기				
	1	표준 컬럼 관리													
	2	정보시스템 전체 데이터베이스 테이블 <input type="text"/> <span style="float: right;">검색 등록</span>													
		No	정보시스템	물리DB명	테이블명	컬럼명	컬럼한글명	데이터타입	데이터길이	NOT NULL	PK여부	FK정보	개인정보여부	암호화여부	공
		1	GenAi 구축 프로젝트	ADC	TEST1	CD_GROUP_ID	코드그룹아이디	VARCHAR	1000	Y	Y		Y	Y	
		2	GenAi 구축 프로젝트	ADC	TEST1	tb_col_test	코드설명	NUMBER	10	Y	pk_tb_col_test	fk_tb_col_test	Y	Y	

1

## 카탈로그 설정 및 데이터 셋 매핑 관리

### 카테고리 데이터셋 매핑

카테고리 목록

카테고리 이름을 입력하세요.

- ▼ MZC
  - SALESFORCE
  - > GLUE
  - DENODO

### 데이터셋 목록

관계사: 관계사를 선택하세요. 데이터셋

설명: 설명을 입력하세요. Tag

생성일: 시작 날짜 to

### 카테고리 관리

카테고리 목록

카테고리 이름을 입력하세요.

- ▼ MZC
  - SALESFORCE
  - ▼ GLUE
    - ADC-DCP
    - CATE-1-2
    - DENODO

상위 카테고리 ID: MZC

카테고리 ID: SALESFORCE

카테고리 설명: SALESFORCE

사용여부:  사용

정렬순서	3
카테고리명	SALESFORCE
관계사	메가존클라우드
생성일	2024-02-19
수정일	2024-02-19
노출여부	<input checked="" type="checkbox"/> 노출

매핑여부	데이터셋 ID	데이터셋 명	플랫폼	상위 카테고리 ID	카테고리 ID	카테고리 설명	생성일
<input checked="" type="checkbox"/>	AllActivatedContracts	AllActivatedContracts	SALESFORCE				
<input checked="" type="checkbox"/>	AllContracts	AllContracts	SALESFORCE				
<input checked="" type="checkbox"/>	AllDraftContracts	AllDraftContracts	SALESFORCE	Contract	메가존클라우드		2024-02-13
<input checked="" type="checkbox"/>	AllInApprovalContracts	AllInApprovalContracts	SALESFORCE	Contract	메가존클라우드		2024-02-13
<input checked="" type="checkbox"/>	MyActivatedContracts	MyActivatedContracts	SALESFORCE	Contract	메가존클라우드		2024-02-13
<input checked="" type="checkbox"/>	MyDraftContracts	MyDraftContracts	SALESFORCE	Contract	메가존클라우드		2024-02-13
<input checked="" type="checkbox"/>	MyInApprovalContracts	MyInApprovalContracts	SALESFORCE	Contract	메가존클라우드		2024-02-13
<input type="checkbox"/>	all_base_views	all_base_views	DENODO	ewdsvc	메가존클라우드		2024-02-05
<input type="checkbox"/>	all_catalog_sales_bigd	all_catalog_sales_bigd	DENODO	ewdstd	메가존클라우드		2024-02-02

카테고리 별 데이터 셋 매핑

카테고리 관리

# DP360 Catalog 스크린샷 > ⑦ 메타 데이터 관리

메타 데이터의 데이터 셋, 컬럼정보, 샘플데이터, 추출조건, 데이터 분포 등을 관리

메타데이터관리 **메타 데이터 관리**

관계사 **메가존클라우드** 데이터셋 **선택해 주세요** 입력해 주세요 **검색**

**데이터 셋** 조건 분포도

관계사 **메가존클라우드** 데이터 셋 : bus\_line 데이터 명 : 서울 버스 라인 **저장**

관계사 **메가존클라우드** MZC 데이터셋 **컬럼 정보** **컬럼 정보** 소유자 admin **검색**

관계사 **메가존클라우드** 플랫폼 데이터 셋 : bus\_line 데이터 명 : 서울 버스 라인 표준화여부: Y **저장**

관계사 **메가존클라우드** \* 데이터셋 ID \* 물리DB명 용량 생성일 Tag 설명

물리DB명	사용여부	데이터관리자
ewdsvc	사용	
ewdsvc	사용	
	사용	admin
	사용	

데이터셋 : bus\_line 컬럼명 pk 여부 조건 순서 컬럼 설명 데이터 타입 데이터 길이

순서	컬럼 영문명	컬럼 한글명	pk 여부	조건 순서	컬럼 설명	데이터 타입	데이터 길이
1	데이터셋	컬럼 정보					
3	데이터 셋 : bus_line	데이터 명 : 서울 버스 라인					
4	데이터 셋 : bus_line	데이터 명 : 서울 버스 라인					
5							
6	trans_type	line_nm	station_id	station_nm	gps_y		
7	세종시(일반)	데이터셋	컬럼 정보	샘플 데이터	추출조건		
8	세종시(일반)	데이터 셋 : bus_line	데이터 명 : 서울 버스 라인				
9	세종시(일반)						
10	세종시(일반)						
11	세종시(일반)						

데이터셋 : bus\_line **샘플 데이터** 샘플 다운로드 샘플 업로드 **저장**

**추출 조건** **저장**

컬럼 영문명	컬럼 한글명	조건 순서	데이터 타입	추출옵션	쿼리
station_order	station_order	1	string	단일쿼리	쿼리를 입력해주세요.
station_id	station_id	2	string	복수쿼리	쿼리를 입력해주세요.

**데이터셋** 컬럼 정보 샘플 데이터 **추출조건**

데이터셋 : bus\_line **저장**

46 >

**저장**

## 배치 실행 스케줄링 관리 및 실행 이력 관리

### 배치스케줄링 관리

조회
저장

관계사

메가존클라우드

배치구분

사용자데이터추출

실행주기

매월

실행일시

20 일

14:58:00

Cron표현식

0 58 14 20 \*\*

비교 사용자 데이터 추출 배치입니다.

회사코드	배치코드	배치명	실행주기	사용여부	실행일시	CRON 표현식	비고
MZC	UIDET	사용자데이터추출	매월		20일 14:58:00	0 58 14 20 **	사용자 데이터 추출 배치입니다.

### 배치 실행 이력 관리

배치 스케줄링 이력 모니터링

관계사

메가존클라우드

배치구분

전체보기

검색

회사코드	배치구분	배치유형	실행일시(월/일/일시)	배치시작일시	배치종료일시	소요시간(초)	성공여부	실패/성공 내용
MZC	SalesForce데이터추출	수동배치	2024-02-21 04:58:00	2024-02-21 04:58:01.008	2024-02-21 04:58:19.044	18.036	성공	SalesForce 메타데이터 추출 배치 실행 결과 처리-
MZC	Denodo메타데이터추출	수동배치	2024-02-06 01:10:36	2024-02-06 01:10:36.314	2024-02-06 01:24:44.326	848.012	성공	Denodo 메타데이터 추출 배치 실행 결과 처리수(T
MZC	사용자데이터추출	수동배치	2023-12-22 02:46:28	2023-12-22 02:46:28.778	2023-12-22 02:46:28.835	0.056	성공	사용자 데이터 추출 배치를 실행 합니다.

1

표준화 된  
데이터 관리

명확한  
데이터 이해

데이터  
혁신을 통한  
DT 가속화

## 검색 및 디스커버리



- 데이터 스키마, 필드값 태그, 사용정보

## 거버넌스



- 사용자/그룹별 권한 설정, 접근 제어

## 데이터 프로파일링



- 쿼리 및 데이터 파이프라인 실행의 데이터 프로파일링 및 평가

## 데이터 품질



- 메타 데이터 표준화, 데이터 프로파일

## 컴플라이언스 및 분류



- 규정 준수, 다양한 데이터 셋 및 요소 분류, 보호 데이터 분류

## 비즈니스 용어



- 도메인, 용어사전, 용어그룹



# Talk to DATA CREW!

상세한 내용 또는 궁금하신 점이 있으세요?  
데이터 크루에게 문의 주십시오!

[AskData@megazone.com](mailto:AskData@megazone.com)

<https://AIDATA.megazone.com>





# Thank you

