

Data Analytics

Data Catalog

데이터 자산의 메타데이터를 수집하고 통합 관리하는 서비스

Data Catalog는 기업 내 모든 데이터 자산의 메타데이터를 수집하여 통합 관리함으로써 데이터 분석가, 데이터 사이언티스트, 개발자 등 다양한 사용자들이 데이터를 더욱 효과적으로 활용할 수 있도록 지원하는 서비스입니다. 데이터 수집과 통합을 자동으로 수행하여 메타데이터를 항상 최신 상태로 유지할 수 있으며, 강력한 검색과 데이터 분류를 통해 더욱 효율적으로 데이터 자산을 활용할 수 있습니다.

메타데이터 자동 수집

다양한 DBMS 및 Hive의 데이터 소스에서 메타데이터를 수집하는 Data Crawler 기능을 제공합니다. 자동화된 Data Crawler로 많은 양의 데이터를 빠르고 효율적으로 처리할 수 있으며, 수집한 메타데이터가 매핑된 Catalog를 통해 의사 결정을 내리는 데 도움이 되는 새로운 인사이트를 얻을 수 있습니다.

데이터 품질

데이터의 출처와 흐름을 시각화하여 제공합니다. 한눈에 보여주는 데이터 계보는 데이터의 출처나 데이터에 일어난 변화, 데이터 파이프라인 내에서의 최종 목적지에 대한 정보를 제공하여 데이터의 일관성과 품질을 보장하며, 오류의 원인을 추적할 수 있습니다.

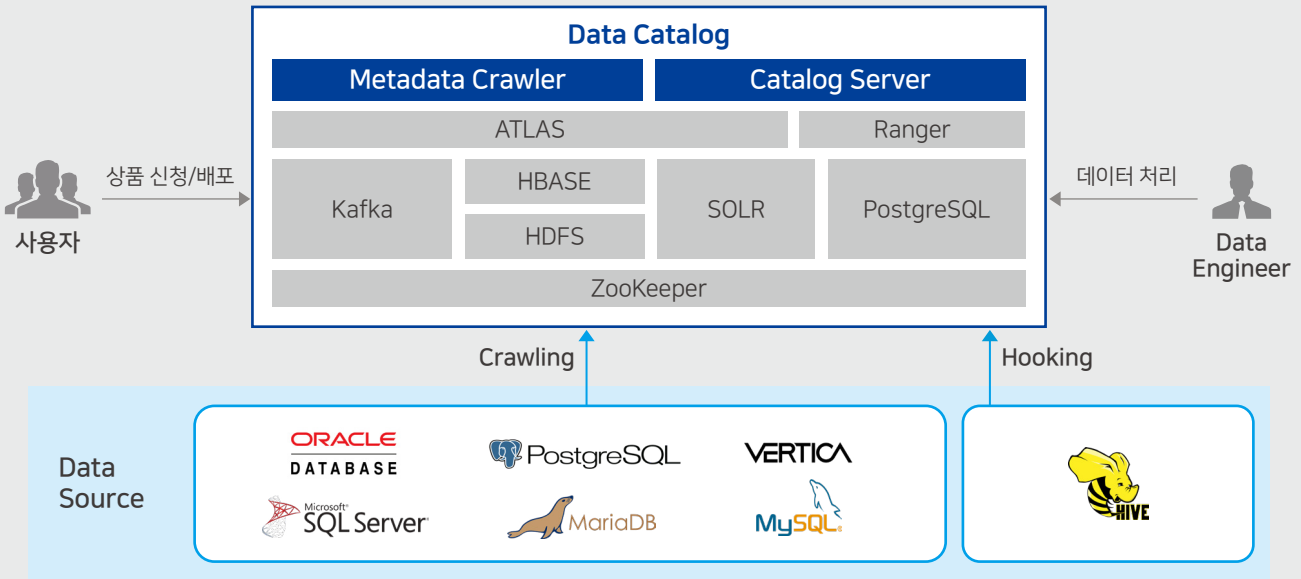
편리한 통합 검색

여러 가지 선택 옵션을 통해 한번의 검색으로 원하는 데이터를 쉽고 빠르게 찾을 수 있는 강력한 통합 검색 기능을 제공합니다. 또한 메타데이터의 접근 권한과 사용자가 등록된 비즈니스 메타 정보를 확인함으로써 조직 내부에서 데이터 자산을 더욱 효율적으로 활용할 수 있습니다.

데이터 분류

데이터의 비즈니스 용어를 설정하고 관리할 수 있는 용어집을 제공하여 데이터에 대한 잠재적 이해 관계자가 데이터에 쉽게 액세스하고 이해할 수 있습니다. 비즈니스 용어는 데이터의 의미, 다른 데이터와의 관계, 데이터 출처 및 형식을 포함하며, 이를 통해 사용자들은 데이터의 분류체계를 이해하고 데이터를 명확히 식별할 수 있습니다.

서비스 구성도



주요 기능




- 메타데이터 자동 수집
 - 메타 크롤러 : 데이터 소스의 DB, 스키마, 테이블, 칼럼 등 메타 정보 수집
 - 리니지 크롤러 : 데이터 소스의 History 정보 수집
 - 샘플 크롤러 : 메타데이터의 샘플 데이터 수집
- 데이터 계보 확인
 - 데이터의 흐름을 시각화하여 제공
 - 테이블 및 스키마 변경 이력 관리
- 통합 검색
 - 메타데이터, 테이블명, 태그 등의 조건으로 데이터 검색
 - Table Summary, Columns, Lineage 조회 등 테이블 상세 조회
 - Role, Owner, Classification, Terms 등 필터별 검색
- 데이터 분류
 - 자산의 주요 특성 식별
 - 데이터 보호를 위한 그룹화
 - Tag Policy를 통한 메타데이터 접근제어 기능 제공

요금 기준

- 과금
 - Data Catalog가 사용하는 VM 자원에 따른 과금
 - VM 자원 비용, 스토리지 비용은 별도 과금

FOR MORE INFORMATION

SAMSUNG SDS

 www.samsungsds.com / cloud.samsungsds.com
 contact.sds@samsung.com / scp_sales@samsung.com
 youtube.com/samsungsds

