

017 데이터 사일로

Data Silo

조직 내 데이터가 부서나 시스템 단위로 고립되어 공유되지 않는 현상

- 데이터 간 연계가 단절되어 통합 분석이나 협업이 어려워지는 구조적 문제
- AI와 데이터 기반 의사결정을 저해하는 주요 장애 요인

데이터 사일로 개요

‘데이터 사일로’는 곡물이나 사료를 외부와 차단된 저장탑에 따로 보관하던 사일로(silo) 구조에서 유래한 말로, 조직 내 데이터가 부서나 시스템 단위로 고립되어 공유되지 않는 상태를 뜻합니다. 이는 데이터가 부서 중심으로 관리되고 시스템 간 호환이 부족하거나, 보안 규정과 조직 문화가 분리되어 있을 때 발생합니다. 예를 들어 마케팅 부서와 영업 부서가 서로 다른 고객 데이터를 사용하면 전체 고객 행동을 통합적으로 분석하기 어렵습니다. 이처럼 데이터가 사일로화되면 정보의 흐름이 단절되고, 조직 전체의 데이터 활용성과 의사결정 효율이 저하됩니다.



데이터 사일로의 원인과 문제점

데이터 사일로는 기술적·조직적 요인이 함께 작용해 형성됩니다. 기술적으로는 서로 다른 포맷, 데이터베이스 구조, 전송 프로토콜 등이 통합을 어렵게 만들고, 조직적으로는 부서 간 이해관계나 관리 권한이 협업을 제한합니다. 이러한 분리 구조는 데이터 중복, 불일치, 품질 저하를 초래하며, 전체 시스템의 효율성을 떨어뜨립니다. 특히 AI 개발에서는 학습 데이터의 다양성과 규모가 제한되어 모델 성능이 저하되거나 편향이 심화될 수 있습니다.

데이터 사일로에 대한 대응 방법

데이터 사일로를 해소하기 위해서는 기술적 통합과 조직적 협업이 병행되어야 합니다. 기술적으로는 클라우드 기반 데이터 레이크나 통합 플랫폼을 구축해 데이터를 중앙에서 관리하고, 표준화된 메타데이터 체계를 도입해 이질적인 데이터를 연결할 수 있습니다. 조직적으로는 부서 간 데이터 공유 정책과 협업 프로세스를 마련해, 데이터의 소유보다 활용 중심의 문화를 조성해야 합니다. 이러한 개선을 통해 조직은 데이터 자산을 통합적으로 분석·활용할 수 있고, AI 학습의 품질과 정확성도 높일 수 있습니다. 즉, 데이터 사일로의 해소는 AI 시대의 데이터 개방성과 협력 생태계 구축을 위한 필수 조건입니다.