

Forbes insights

지능형 데이터
카탈로그:
디지털 변혁을
선도하다

공동 발행인:



Informatica

정보에 대한 신뢰는 오늘날 데이터 중심 비즈니스과는 떼려야 뗄 수 없는 가치입니다. 사람을 통해서든 아니면 인지 시스템을 통해서든 관계없이 기업 전반에 걸친 모든 부분의 정보를 통합하고 활용함으로써 임원진은 의사 결정 시 거의 지체할 필요가 없는 디지털 비즈니스를 구축할 기회를 얻습니다. 실시간 또는 거의 실시간으로 의사 결정이 이루어지는 상황에서 정보 탄탄함과 이러한 정보가 기업, 고객, 시장 및 나머지 환경에 제공하는 뷰가 얼마나 광범위한지에 따라 신뢰성의 수준이 결정됩니다.

다양한 정보 유형을 식별하고 통합하는 것은 필수적인 요소이며 성공적인 디지털 변혁 전략에 있어 핵심입니다. 조직은 제품 및 서비스를 개선하고, 고객에게 더 가까이 다가가고, 경쟁업체를 따돌리고, 혁신가가 되기 위해 빠르게 이동해야 하는 부담감을 그 어느 때보다 더 많이 떠안게 되었습니다. 성공적인 디지털 변혁 전략은 새로운 사고방식과 새로운 기술에 기반하여 과도하게 경쟁하는 경제 상황에서 성공에 이르기 위한 길을 마련해 줍니다. 모든 디지털 전략에는 분석적인 힘이 필요한데, 이 힘은 데이터 목록을 만들고 데이터의 가치를 평가하며 필요한 시기에 필요한 장소에서 데이터를 기업 전반에서 사용하도록 만들 수 있어야 얻을 수 있습니다.

오늘날의 과제는 기업의 비즈니스와 기술 측면이 단절되었다는 데 있습니다. 이 단절은 데이터 중심 디지털 변혁을 이루기 위한 시도를 방해하고 저해합니다. 이러한 단절에는 여러 가지 형태가 있습니다.

- 현업 부서 사용자는 데이터가 어디에 있는지, 데이터가 어떤 내용인지를 알지 못하며 심지어 조언을 요청할 대상조차 알지 못하기 때문에 데이터를 통합하고 활용할 수 없습니다.
- 비즈니스 및 IT 팀은 다양한 언어로 말하기 때문에 서로 의견이 맞지 않습니다.
- 현업 부서 사용자는 상충하는 데이터 지점과 다양한 소스에 대해 동의하지 못하는 경우가 자주 있습니다.
- 데이터를 필요로 하는 모든 사람이 사용할 수 있고 액세스할 수 있게 해주는 데이터 '민주화'는 아직 걸음마 단계에 있습니다.

오늘날 많은 조직에서 이러한 단절이 발생하고 있는 원인은 포괄적인 데이터 거버넌스의 부족, IT 중심의 신뢰할 수 있는 데이터를 찾는 방법의 결여, 엔터프라이즈 데이터 환경 및 관련된 애플리케이션에 대한 일반적인 가시성 부족 때문입니다. 그 결과, 적절한 정보를 찾는 것은 정보를 바탕으로 하는 분석보다 보물찾기에 가깝습니다. 비즈니스와 IT 모두 데이터 자산을 활용하여 완전한 잠재력을 활용할 때가 온 것입니다.

지능형 데이터 카탈로그는 데이터 중심 기업 및 디지털 기업을 뒷받침하는 메타데이터를 관리하며 이에 대한 액세스를 제공합니다. 지능형 데이터 카탈로그는 궁극적으로 통찰력 있는 의사 결정을 지원하는 분석 엔진의 기능을 촉진하는 중요한 데이터를 안내합니다. 이 생각의 리더십(thought leadership) 백서에서는 새롭게 등장한 지능형 데이터 카탈로그의 역할과 기업에서 데이터 자산을 조직하고 발견할 수 있도록 지원하는 데 있어 인공 지능 및 기계 학습의 역할이 점점 중요해지고 있는 상황에 대해 살펴보겠습니다.

Boston Consulting Group의 파트너 및 관리 이사인 Shervin Khodabandeh는 이렇게 말합니다. "대부분의 조직은 액세스 권한을 지닌 데이터 중 일부만 사용하고 있으며 내 경험상으로는 데이터 중 5% 미만을 사용하고 있으면서 테라바이트의 데이터를 수집 및 저장하고 있습니다. 예를 들어, 많은 공익사업 기업은 데이터에 숨겨져 있는 인텔리전스가 비즈니스를 변혁할 것이라는 기대감을 갖고 스마트 계량기를 설치하고 있습니다. 그러나 대부분은 아직 이러한 정보를 의미 있게 활용하지 못하고 있습니다. 이와 유사하게 많은 소매업체는 고객 거래 및 행동에 대한 데이터를 온라인으로 수집하고 있지만 더 개인화된 고객 경험을 생성하기 위해 이러한 정보를 아주 조금만 활용하고 있을 뿐입니다."

Shervin Khodabandeh는 이것은 모든 산업에서 공통적인 문제라고 말하며 다음과 같이 덧붙였습니다. "너무 많은 데이터와 너무 많은 변수가 있으며 이 데이터로 어떤 일을 수행할지, 가장 잘 활용할 방법은 무엇인지가 공통적인 문제입니다. 따라서 대부분의 조직은 데이터 조각에만 중점을 두고 있는데 이러한 데이터 조각을 직관적으로 이해하고 경험하며 나머지를 메타데이터로 다루고 있습니다. 실제로는 메타데이터와 같은 것은 없습니다. 모두 데이터일 뿐입니다. 오늘의 데이터는 어제의 메타데이터였던 겁니다."

소개: 지능형 데이터 카탈로그의 역할

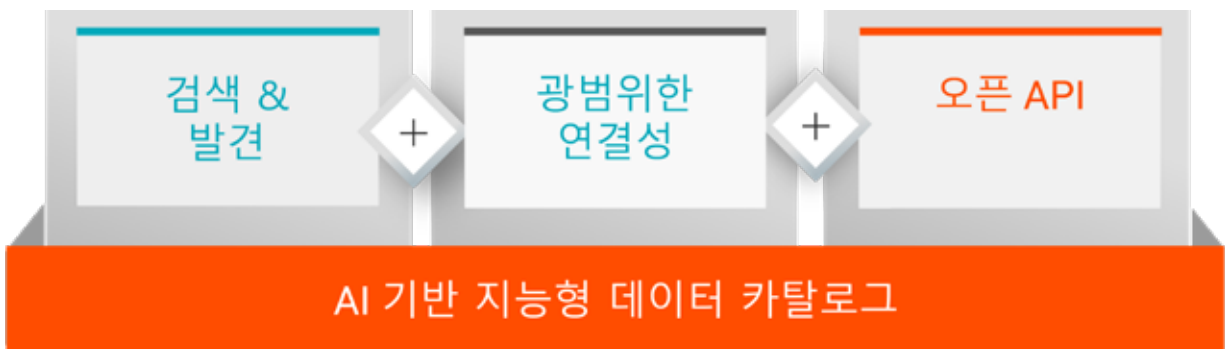
수년 동안 데이터 통합은 조직이 달성하기 어려운 목표였으며 사일로, 부서별 애플리케이션 및 외부 소스에는 항상 다양한 형태의 데이터가 숨겨져 있었습니다. 기업의 모든 영역에 있는 데이터는 물론 그 영역 밖에 있는 데이터까지 활용하는 디지털 기업의 등장으로, 상황이 점점 더 복잡해지고 있습니다. 데이터 소스와 자산은 규모가 폭발적으로 증가하고 있어 데이터를 추적하고 발견하기 위해서는 더욱 자동화된 방식이 필요해졌습니다. 지능형 데이터 카탈로그는 이러한 요구사항을 충족시켜 줍니다.

지능형 데이터 카탈로그 정의

지능형 데이터 카탈로그는 기존의 데이터 카탈로그, 메타데이터 리포지토리 및 데이터 레이크가 채우지 못했던 지점에서 출발합니다. 이 새로운 세대의 카탈로그는 데이터 자산의 디렉토리를 제공할 뿐만 아니라 기업이 점점 증가하고 있는 정보의 저장소를 가장 잘 활용할 수 있도록 자동화 기능을 통합합니다.

지능형 데이터 카탈로그는 메타데이터 톨과 리포지토리의 진화된 버전이라고 할 수 있습니다. TDWI의 Philip Russom에 따르면 이 새로운 카탈로그가 "후미진 사일로에서부터 풍부하고 포괄적인 기능을 제공하는 중앙 집중식 교차 플랫폼 시설에 이르는 메타데이터 관리"를 담당한다고 합니다. "온 프레미스나 클라우드, 기업 내부, 인터넷 또는 파트너 회사와 관계없이 모든 소스에서 추출한 메타데이터를 상상해 보십시오. 그렇게 축적된 메타데이터는 개선을 거쳐 기업 전반과 기업을 넘어 공유되며 유례없는 일관성, 생산성, 신뢰성 및 거버넌스를 제공하고 있습니다."¹

지능형 데이터 카탈로그는 AI의 지원을 바탕으로 기술 메타데이터, 비즈니스 컨텍스트, 사용자 주석, 관계, 데이터 품질 및 데이터 활용을 포함하는 통합된 메타데이터 뷰를 제공하여 비즈니스 및 IT 사용자가 기업 데이터 자산의 완전한 잠재력을 실현할 수 있도록 돕습니다. 또한 이러한 카탈로그는 직원의 데이터 큐레이션, 거버넌스 및 관리 작업을 지원합니다.



¹ The Data Catalog's Role in the Digital Enterprise(디지털 엔터프라이즈에서 데이터 카탈로그의 역할), 2017년 TDWI, Philip Russom
<https://tdwi.org/research/2017/11/ta-all-informatica-the-data-catalogs-role-in-the-digital-enterprise.aspx>

조직에 지능형 데이터 카탈로그가 필요한 이유

인간이 대용량의 데이터 저장소, 데이터 구조 및 소스를 효율적으로 관리하기 어려운 수준까지 데이터가 폭발적으로 증가하고 있습니다. 또한 데이터 웨어하우스와 심지어 데이터 레이크도 갈수록 증가하는 데이터 소비자 계층에 적합하지 않게 되었습니다. 간단히 말해, 오늘날 지식 기반 작업자와 의사 결정자는 데이터 소스를 확인하려고 시도하거나 이를 추적할 시간이 없습니다. 이러한 자산을 신뢰하고 정보를 바탕으로 의사 결정을 내릴 수 있도록 하려면 어떻게 해야 할까요? 이들이 가져오는 데이터에 비즈니스 및 기술 컨텍스트를 어떻게 제공하시겠습니까? 질문은 누구에게 물어야 할까요? 데이터를 어떻게 민주화하시겠습니까?

이러한 문제는 Marlette Funding이라는 온라인 소비자 금융 서비스 회사의 CTO 겸 CSO인 Brian Conneen가 제기했습니다. 그는 "우리는 데이터가 부족한 적이 없으며 모든 종류의 소스에서 가져온 데이터를 확보하고 있습니다. 우리는 고객이 제공하는 내용에서 데이터를 확보하고, 웹 상호 작용을 측정하는 데이터 분석 및 툴도 보유하고 있으며, 애플리케이션 및 검증 프로세스의 일부로 신용 기관의 데이터도 확보하고 있고, 성능 데이터와 같이 애플리케이션 자체에서 데이터를 수집하고 있습니다."고 말했습니다. Conneen의 팀은 의사 결정자와 분석가가 이러한 리소스가 무엇이고 어디에 있는지 파악할 수 있도록 지원하기 위해 20~30개의 다양한 데이터 소스에 액세스하여 데이터 카탈로그를 유지 관리합니다.

이전의 데이터 관리 및 통합 솔루션 세대, 특히 데이터 웨어하우스는 오늘날 빠르게 변화하는 디지털 기업에게는 너무 느리고 번거로웠습니다. Conneen은 이렇게 덧붙입니다. "데이터가 존재하는 곳에서 데이터를 찾아야 합니다. 통합 솔루션을 통해 모든 데이터를 과도하게 엔지니어링하려고 하지 마십시오. 정말 원하는 데이터 소스를 가지고 있다고 하더라도 데이터 웨어하우스에 보관하는 데에만 3개월이 걸리기도 합니다. 3개월이나 지나면 그 데이터는 더 이상 유용하지 않죠. 단순히 3개월의 시간과 노력만 더해진 것뿐입니다."

데이터 민주화 및 의사 결정자에게 유용한 데이터로 만들기 위한 과제:

- **사일로에 있지만 의사 결정자가 알지 못하고 보이지 않는 데이터, 특히 비정형 데이터.** IDC의 추정에 따르면 비정형 데이터의 양이 모든 디지털 데이터의 90%에 이른다고 합니다. 또한 상위 조직인 IDG의 연구에서는 정형 데이터가 여전히 우선시 된다는 점을 보여줍니다. IT 전문가의 83%는 정형 데이터 이니셔티브가 조직에서 중요한 우선순위라고 했으며 이에 비해 비정형 데이터 프로젝트가 중요하다고 한 전문가는 43%였습니다. 이러한 결과는 '많은 조직이 비정형 데이터를 제대로 활용하여 얻을 수 있는 상당한 비즈니스 가치를 도출할 기회를 놓치고 있다'는 점을 보여줍니다.²
- **수많은 데이터 소스.** 오늘날 데이터 중심 조직들은 IoT 장치, 엔터프라이즈 시스템 및 소셜 미디어와 같은 다양한 소스에서 데이터를 가져옵니다. TDWI에 따르면 기술 메타데이터는 데이터 구조, 구성 요소 및 데이터 유형을 문서화하는데, 이것은 데이터 추출 및 로드, 기타 컴퓨터화된 프로세스와 높은 수준의 기술 인터페이스에서 기본적인 것이라고 합니다.
- **다수의 플랫폼에 존재하는 데이터.** 조직이 더 디지털화되는 방향으로 발전함에 따라 조직은 데이터를 저장하고 처리하기 위해 온 프레미스 및 클라우드 시스템 모두를 활용합니다. 그 결과 데이터 관리는 여러 위치에 저장되어 있는 데이터를 활용하는 새로운 차원으로 나아가게 되었습니다. 새로운 소스를 식별하고 잠금을 설정하는 것과 더불어 기업은 여러 시스템 유형에서 제공되는 데이터가 실행 및 액세스가 가능하며 안전한지 확인해야 합니다.
- **데이터 분석을 온전히 활용하지 못함.** 새로운 혁신과 방식을 가지고 시장에 진출하는 기업은 이를 실현하기 위해 데이터 분석을 활용할 것입니다. 오늘날의 기업은 대부분 다양한 소스에서 많은 양의 데이터를 확보하고 있지만 그 데이터를 경쟁 우위를 달성하는 데 활용하지 못하지 못하는 기업들이 많습니다.

² "Solving the Unstructured Data Challenge(비정형 데이터 과제 해결)", 2015년 6월 25일, Jaikumar Vijayan, CIO.
<https://www.cio.com/article/2941015/big-data/solving-the-unstructured-data-challenge.html>

- **시스템이 복잡하며 데이터 소스와 시스템 간의 상호 작용을 더 잘 이해해야 함.** 기업이 클라우드와 온 프레미스 시스템 모두에서 자산을 활용하기 위해 하이브리드 컴퓨팅으로 전환하고 있는 시대에는 데이터를 다양한 소스에서 가져와 분석합니다.
- **데이터 출처에 대한 이해 부족.** 데이터 신뢰를 확보하는 데 있어 최대 과제 중 한 가지는 사용자가 데이터 출처, 데이터가 생성된 방식과 데이터를 사용하기 위한 적절한 권한을 이해하지 못한다는 점입니다.
- **실시간 기능을 활용하려면 제대로 조사된 데이터 소스를 필요할 때 사용할 수 있어야 함.** 실시간 상호 작용 특히 인공 지능 및 기계 학습 인터페이스를 통한 상호 작용의 필요성이 점점 더 대두됨에 따라 더 신뢰할 수 있고 투명한 데이터가 필요합니다.
- **규정 준수 의무를 이행하려면 데이터 검색 및 검증이 필요함.** 적절한 정보를 찾고 계속되는 보고 목적에 맞는 지 검증하는 데 기업 담당자의 시간과 리소스가 상당히 소비되고 있습니다.
- **사용자는 엔터프라이즈 데이터를 검색할 때 Google과 같은 의미 검색을 필요로 하며 기대함.** 소비자에서 직원, 고객에 이르기까지 모두가 무언가를 문의할 때 Google이나 Twitter와 같은 소셜 미디어 사이트에서처럼 즉각적인 결과를 제공하는 환경에 완벽하게 익숙해졌습니다. 이러한 부분은 내부 기업 정보를 제공과 관련해서도 새로운 표준을 수립하게 만듭니다.

인공 지능으로 향상된 지능형 데이터 카탈로그

새롭게 등장한 지능형 데이터 카탈로그 세대는 단순히 메타데이터를 관리, 태그 지정 및 저장하는 것이 아니라 인공 지능(AI) 및 기계 학습을 적용하여 데이터를 더 깊이 있고 폭넓게 이해할 수 있도록 해줍니다.

AI를 지능형 데이터 카탈로그에 활용하면 다음과 같은 이점이 있습니다.

- 사용자가 비정형 데이터 내에서 엔티티를 검색하도록 지원
- 데이터 간의 매우 복잡한 관계성 발견
- 소스에서 데이터를 활용하는 애플리케이션에 이르는 엔드 투 엔드 데이터 분석을 통해 데이터 계보(Lineage)에 대한 이해 제공

기계 학습은 의미 추론 및 추천 기능도 지원합니다. 몇 가지 예로는 복잡한 데이터(예: 머신 데이터)의 구조 발견, 유사하고 중복된 데이터 세트 식별, 데이터 세트 전반에서 복합 엔티티 발견, 활용도, 품질 또는 출처, 다른 요소에 기초하여 대체 또는 유사한 데이터 세트 추천이 있습니다.

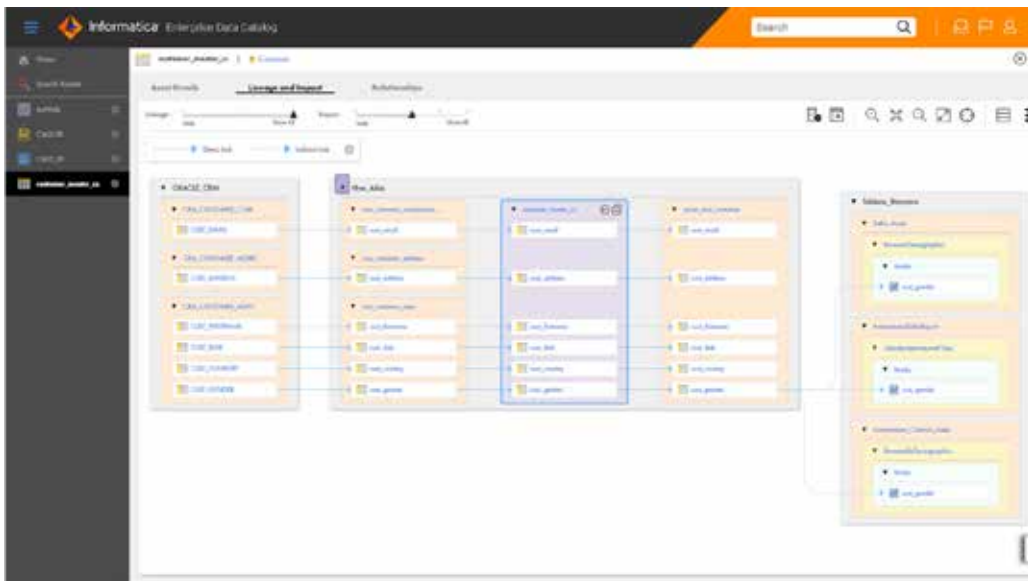
지능형 데이터 카탈로그가 미치는 영향

오늘날 많은 기업에서 필요로 하는 지능형 데이터 카탈로그는 디지털 변혁을 위한 노력의 토대가 됩니다. 지능형 데이터 카탈로그는 적절한 데이터 소스를 찾아낼 뿐만 아니라 데이터 검색과 연결된 워크플로우를 관리하는 데 도움을 줍니다.

데이터 자산 및 관계에 대한 종합적인 뷰



데이터 계보(LINEAGE) 및 영향 분석 이해



Cisco의 직원 및 전략, 계획, 디지털 변혁 선임 리더인 Flavio A. Zanetti는 42개국에서 800명의 엔지니어로 구성된 회사의 팀이 고객에게 서비스를 제공하는 데 필요한 정보에 즉시 액세스할 뿐만 아니라 참여를 처리하는 방식에 대해서도 이해하는 것이 중요하다고 말하며 다음과 같이 덧붙였습니다. "우리는 회사 전체에서 다양한 데이터 인스턴스, 다양한 데이터베이스, 다양한 데이터 레이크를 보유하고 있습니다. 이 모든 요소를 같은 장소에 통합하여 콘텐츠에 가장 적합한 위치와 엔지니어가 찾고 있는 최적의 소스를 찾아야 합니다."

Zanetti는 팀에서 이미 데이터 카탈로그 활용을 통해 얻을 수 있는 다양한 이점을 확인했다고 말합니다. 그는 주로 문서 및 데이터를 찾는 데 걸리는 시간을 줄여 엔지니어가 개발, 데모, 제안, 고객 참여 및 토론과 같은 높은 수준의 활동에 더 많은 시간을 할애하도록 하였습니다. "그 결과 위치 및 볼륨에 따라 다양하지만 지금까지 적어도 15%에서 20%까지 생산성이 향상된 것을 확인했습니다."라고 Zanetti는 덧붙였습니다.

더 이상 몇 안 되는 분석가와 관리자만 데이터를 활용하는 것이 아니라 조직의 거의 모든 사람이 데이터를 활용하여 의사 결정을 하고 있습니다. 완전히 '민주화된' 데이터 기능을 활용하는 데이터 중심 조직은 안전한 방식으로 기업 전반에서 적절한 데이터를 사용 가능하게 설정할 뿐만 아니라 주요 규정을 담당하는 팀의 관리를 받습니다. 예를 들면 다음과 같습니다.

- 데이터 과학자 및 분석가는 지능형 데이터 카탈로그와 연결된 셀프서비스 분석을 활용하여 관련된 데이터 자산을 찾고, 계보를 시각화하며 정보에 대한 360도 뷰를 확보할 수 있습니다.
- 지능형 데이터 카탈로그를 활용하는 데이터 관리자 및 데이터 소유자는 비즈니스 글로서리를 기술 개체에 연결하고 기술 계보에 대해 비즈니스를 확인하며, 중요한 데이터 요소의 규정 준수를 충족하는 데 도움을 주는 데이터 거버넌스를 채택할 수 있습니다.
- 데이터 아키텍트 및 개발자는 지능형 데이터 카탈로그를 활용하여 데이터 자산 관리를 채택하고, 열 수준(column-level) 계보 및 변경 영향을 분석하며 변환 로직을 볼 수 있습니다.
- 거버넌스 및 보안 관리자.
- 데이터 소비자.

데이터 카탈로그가 중요한 역할을 하는 7가지 주요 활용 사례가 있습니다.



1. 데이터 검색 및 이해. 비정형 데이터는 많은 새로운 경쟁 통찰력을 발견할 수 있는 열쇠로 간주되지만 이러한 데이터는 완전히 숨겨져 있지 않다고 하더라도 찾기가 어려운 게 많은 기업들의 현실입니다. 한 보고서에 따르면 비정형 데이터 자산을 제한적으로 이해하거나 이해하지 못하는 조직의 비율이 78%라고 합니다.³ AI의 지원을 받는 지능형 데이터 카탈로그는 사용자 및 애플리케이션이 기업 전반과 기업을 넘어 소스에서 신뢰할 수 있는 데이터를 신속하게 찾고 검색할 수 있도록 해줍니다. 그 결과 모든 데이터 저장소, 플랫폼 및 서비스 전반에서 통합된 뷰를 확보할 수 있으며, 분류 및 소비는 도메인 및 엔티티의 자동화된 추론뿐만 아니라 검색 추천의 지원을 받습니다.

또한 지능형 데이터 카탈로그는 데이터 도메인, 사용자 및 관련된 데이터 세트를 포함하는 모든 적절한 데이터에 대해 오랜 노력 끝에 얻은 360도 뷰와 함께 엔드 투 엔드 계보, 자동 스캐닝 및 데이터 태그 지정을 제공합니다. 이를 바탕으로 사용량, 데이터 구조의 추론, 데이터 중복 방지에 기초하여 클라우드 소싱 메타데이터의 이점을 활용할 수 있습니다.



2. 데이터 거버넌스. 성공적인 데이터 중심 기업에는 효과적인 데이터 거버넌스도 필요합니다. 데이터 카탈로그가 기업과 함께 확장되고 성장함에 따라 데이터를 안전하게 보호하고 비즈니스 정책 및 우선순위에 맞게 조정하는 것이 중요합니다. 규정 준수 의무에는 데이터 자산의 투명성 및 관리를 위한 요건도 포함되어 있습니다. 데이터 카탈로그는 기업 내에서 구축된 데이터 자산에 자동으로 규칙 및 정책을 적용하면서 조정 작업을 제공합니다. 메타데이터가 관련 사일로, 애플리케이션 또는 기능과 밀접하게 연관이 있는 경우 지능형 데이터 카탈로그는 데이터 소스에 대한 정보를 중앙 집중식 엔터프라이즈 환경으로 가져옵니다.

³ "Five Steps to Unstructured Data Security(비정형 데이터 보안을 위한 5가지 단계)", 2017년 1월 25일, ITProPortal의 Michael Fimin. <https://www.itproportal.com/features/five-steps-to-unstructured-data-security/>

또한 지능형 데이터 카탈로그는 데이터 자산이 얼마나 크고 다루기 어렵게 증가하고 있는지에 상관없이 규정 준수 및 데이터 조사에 대한 프로세스, 규칙 및 정책을 자동화함으로써 더 원활한 거버넌스를 지원합니다. 또한 Russom은 지능형 데이터 카탈로그는 데이터가 메타데이터 및 카탈로그 패턴의 다른 기능을 통해 처리됨에 따라 데이터의 계보를 자동으로 레코딩할 수 있다고 말하며 다음과 같이 덧붙였습니다. "계보는 글로서리 내에서 비즈니스 용어 간의 링크를 사용하여 엔티티 및 특성 수준으로 세분화되어야 합니다. 이러한 상세한 정보를 통해 데이터 계보는 데이터 출처 및 영향 분석에 대한 신속하고 깊이 있는 통찰이 가능해집니다."



3. 협업. 지능형 데이터 카탈로그를 통해 제공된 자동화된 거버넌스 기능으로 데이터 프로세스가 기업 전반에 걸쳐 팀들에게 공개되고, 사용자가 권한 수준에 기초하여 파일에 대한 액세스 권한을 자동으로 부여받은 덕분에 협업의 능력이 올라갑니다. 예를 들어, 사용자는 데이터 세트에 태그를 지정하고 의견을 남기고 '좋아요'를 누르거나 다른 프로젝트 팀원에게 이를 추천할 수 있습니다. 또한 사용량 통계 및 메타데이터는 널리 쓰이고 관련성 높은 데이터 세트를 찾는 데 도움을 줄 수 있어 사용자가 데이터 세트를 안전한 방식으로 공유할 수 있습니다. 민감한 데이터의 식별을 어렵게 만들기 위해 데이터 마스킹을 적용하여 권한 있는 사용자 간에 개인적으로 식별 가능한 정보 또는 기타 민감한 데이터 유형을 노출하는 위험 없이 데이터를 공유할 수 있습니다.

데이터 카탈로그는 우수한 커뮤니케이션을 촉진할 뿐만 아니라 기업 전반에서 중복된 작업을 없애는 데 도움이 됩니다. "누군가가 질문이 있거나 호주에서 도움이 필요한 경우, 동일한 엔지니어가 영국에서 동일한 문제에 대답을 할 수 있습니다. 동일한 질문을 리서치하느라 시간을 보내고 있다면 우리 모두 리소스를 낭비하고 있는 것입니다."라고 Zanetti는 말했습니다. Cisco의 영업 팀은 이제 제안, 명세 및 단일 장소에 있는 기타 모든 관련 데이터에 대한 액세스를 제공하여 사용자가 누가 어떤 일을 수행 중이며 그에 따라 정보를 공유하고 있는지 확인할 수 있도록 지원합니다.



4. 실시간 기업. IoT 및 고객 요구사항이 발생할 때 이를 해결하는 능력이 중시되는 오늘날의 경제 상황에서 통찰력을 실시간 또는 거의 실시간으로 처리하고 전달하는 역량은 중요합니다. 기존 접근 방식에서는 사용자 또는 분석가가 모든 가능한 데이터 소스 전체를 조사해야 했습니다. Conneen은 "분석가가 데이터를 알아내는 데 이들이 걸렸습니다. 하지만 지능형 데이터 카탈로그를 활용하면 동일한 답변에 도달하는 시간을 15분까지 줄일 수 있습니다. 거의 실시간으로 실시간 데이터를 확인할 수 있는 것입니다."라고 말했습니다. 동시에 기존의 데이터 환경 및 데이터 웨어하우스에는 상당한 직원 리소스뿐만 아니라 데이터를 효과적으로 통합하기 위한 시간도 필요했습니다. Conneen은 이렇게 덧붙입니다. "이러한 데이터를 단일한 대형 데이터 웨어하우스에 통합하려고 시도했다면 지금까지 통합하는 일에만 시간을 보냈을 것입니다. 핵심은 데이터 위치와 데이터를 보다 신속하게 활용할 수 있도록 데이터를 카탈로그화하는 것입니다."



5. 셀프서비스 분석. 오늘날 성공적인 데이터 중심의 디지털 기업은 강력한 셀프서비스 윤리를 갖고 있습니다. 즉, 최종 사용자가 필요로 하는 데이터에 필요한 때에 액세스할 수 있도록 권한을 부여하기 위해 적극적으로 노력하고 있습니다. 지능형 데이터 카탈로그는 현업 부서 사용자가 필요한 데이터를 쉽게 식별하고 찾도록 도와줍니다. 이것은 중역실에서부터 현업 부서 사용자에게 이르는 비즈니스 사용자들이 회사에서 어떤 데이터 자산을 보유하고 있는지, 이 데이터 자산이 어디에 있는지 알지 못하기 때문에 중요합니다.

이전에는 데이터 자산에 대한 지식을 공유하려면 직접 구두로 전달하거나 기타 오프라인 수단을 거쳐야 했습니다. 따라서 적절한 데이터를 검색하거나 추적하는 데 많은 시간을 소비했습니다. 데이터 소스와 볼륨이 지속적으로 증가하고 발전함에 따라 최종 사용자 기반이 확대되고 있어 IT 부서에서 사용자에게 업데이트된 정보를 제공하는 것이 어려워졌습니다. 대부분의 시간을 데이터 검색에 할애해 왔으며 오히려 분석하는 시간은 적었습니다.

따라서 AI 중심 데이터 카탈로그의 지원을 받는 셀프서비스 환경에서 사용자와 팀이 검색 대신 분석에 더 많은 시간을 할애하도록 지원하는 것이 이상적입니다. "우리는 이중 데이터 소스를 활용하고 있으며, SQL 유형 구문을 이러한 데이터 소스 위에 레이어하는 기술도 사용하고 있습니다. 덕분에 셀프서비스를 제공하여 사용자 대시보드에 이러한 데이터 소스를 노출할 수 있습니다. 우리가 사용자에게 '이것은 애플리케이션 데이터를 위한 데이터 세트입니다.' '이것은 웹 성능 데이터를 위한 데이터 세트입니다.' 또는 '이것은 고객 상호 작용 데이터를 위한 데이터 세트입니다.'라고 알려 주기 때문에 사용자는 데이터의 다양한 출처를 실제로 알 필요가 없습니다. 사용자는 보고 및 대시보드를 분석하고 찾을 수 있는지만 알면 되는 것입니다."라고 Conneen은 말했습니다.



6. 클라우드로의 여정. 데이터 저장소는 규모가 기하급수적으로 증가할 뿐만 아니라 기업 장벽 밖의 소스에도 의존하고 있습니다. 비즈니스 쿼리는 점점 더 많은 무리의 클라우드 기반 소스를 대상으로 실행되고 있어 기업 내에서 원활하게 데이터를 검색 및 관리하기 위한 기능도 필요합니다. 데이터 소스 및 데이터 대상이 모두 클라우드에 있기도 하고, 하이브리드 시스템의 일부인 경우도 있습니다. 지능형 데이터 카탈로그는 고객, 시장, 운영 및 기타 중요한 환경에 필요한 필수적인 360도 뷰의 일부로 검색하기 위해 모든 외부 데이터 소스를 엮니다.



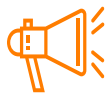
7. 조직적인 전환 관리. 조직이 성장하거나 인수 합병을 통해 진화함에 따라, 비즈니스의 새로운 일부가 된 데이터 자산에 불확실성이 생기거나 현업 부서 간에 데이터가 중복될 수 있습니다. 데이터 카탈로그는 데이터 자산의 출처에 관계없이 데이터 자산을 추적하는 역할을 수행합니다.

지능형 데이터 카탈로그를 최대한 활용하는 방법

빠르게 성장하고 있는 디지털 경제에서 경쟁하기 위해 기업은 신뢰할 수 있고 액세스 가능한 데이터를 필요로 합니다. 지능형 데이터 카탈로그는 데이터를 사용하고 향후 기회와 과제에 대비하는 데 있어 핵심적인 툴입니다. 다음은 기업 내에서 지능형 데이터 카탈로그를 시작하고 최대한 활용하기 위한 몇 가지 권장 사항입니다.



1. 현업 부서와 긴밀하게 협력하여 중요한 데이터 소스를 식별합니다. 현업 부서 사용자에게는 일상적인 업무 중에 정기적으로 액세스해야 하는 중요한 데이터 세트가 있습니다. 또한 데이터를 찾는 데 오랜 시간이 걸리기 때문에 현업 부서 사용자의 업무가 지연되거나 병목 현상이 발생하는 업무 영역이 많습니다. 규정 준수 의무의 영향을 받는 데이터는 더 자동화된 방식으로 관리해야 하는 다른 영역에 해당합니다. 모든 항목을 지능형 데이터 카탈로그로 한 번에 통합하려고 시도하는 대신 사용자에게 가장 중요한 영역을 식별하는 방법에서 시작해 보십시오. 또한 셀프서비스 툴을 통해 데이터를 사용하고 소비하는 현업 부서 사용자의 고충을 이해하는 것도 중요합니다. 예를 들어, 현업 부서 사용자는 데이터 자산에서 도출한 정보와 관련된 기술 용어 이면의 의미를 알지 못할 수 있습니다.



2. 전사적인 참여 촉진. 데이터 소스는 기업 전반에 걸쳐 수집되며 기업 외부에서도 가져오므로 데이터의 가용성을 확인해야 할 뿐만 아니라 가장 우수한 품질과 정확성을 확보해야 합니다. 이를 위해서는 데이터 제공자와 데이터 소비자가 긴밀하게 협력해야 합니다. 모든 직원이 기여자와 사용자 역할을 할 수 있는 기회를 갖는 협업이 관건입니다.



3. 강력한 분석 및 데이터 거버넌스 내에 카탈로그 배치. 기업이 메타데이터가 "우수한 타겟팅, 예측, 문제 해결, 고객 서비스 제공 등을 지원하는 방식을 결정해야 한다"고 운을 댄 Khodabandeh는 활용 사례를 개발할 때 고려해야 할 다른 요인에는 어떤 변수가 예측 가능한지 판단하는 것이 있다고 말합니다. 이 모든 것은 강력한 분석에 달려 있습니다. "패턴은 기존의 회귀 기반 기술을 활용하기 때문에 분명하지 않은 경우가 많지만 기계 학습을 활용할 경우 분명해집니다. 또한 카탈로그를 고객이 생각을 바꾸고 규제 기관이 새로운 규제를 생각해 낼 때 이 메타데이터를 체계적으로 추가 및 제거할 수 있는 데이터 거버넌스 프로세스의 지침에 따라 구축해야 합니다. 데이터는 모델을 축소하지 않고 제거 가능해야 합니다."라고 Khodabandeh는 덧붙였습니다.

자세한 내용은 다음을 확인해 주십시오.

www.informatica.com/products/big-data/enterprise-data-catalog.html.

Forbes insights

FORBES INSIGHTS 소개

Forbes Insights는 전 세계에서 매월 약 9천 4백만 명의 비즈니스 의사 결정자가 활용하는 플랫폼을 갖춘 글로벌 미디어, 브랜딩 및 기술 회사인 Forbes Media의 전략적인 리서치 및 생각의 리더십 실용서입니다. *Forbes* 커뮤니티에서 선임 임원진의 독점 데이터베이스를 활용하여 Forbes Insights는 생각의 리더로서 브랜드 입지를 다지고 이해 관계자의 참여를 촉진하기 위해 광범위한 주제에 대한 리서치를 수행합니다. 리서치 결과물은 다양한 디지털, 인쇄 및 실제적인 실행을 통해 제공되며 *Forbes*의 소셜 및 미디어 플랫폼 전반에서 자세히 설명됩니다.

FORBES INSIGHTS

Casey Zonfrilli

계정 관리 부문 이사

Tori Kreher

프로젝트 관리자

Todd Della Rocca

프로젝트 관리자

편집자

Erika Maguire

총괄 이사

Kasia Wandycz Moreno 이사

Hugo S. Moreno 이사

Joe McKendrick 보고서 작성자

Zehava Pasternak 디자이너

리서치

Ross Gagnon 이사

Scott McGrath 리서치 분석가

영업

North America

Brian McLeod 총괄 이사

bmcleod@forbes.com

Matthew Muszala 이사

mmuszala@forbes.com

William Thompson 관리자

wthompson@forbes.com

Kimberly Kurata 영업 총괄

kkurata@forbes.com

EMEA

Tibor Fuchsel 관리자

tfuchsel@forbes.com

APAC

Serene Lee 총괄 이사

slee@forbesasia.com.sg