

논리적 데이터 관리가 실현하는 새로운 AI 시대





2024년이 생성형 AI가 혁신적 기술로 인식되기 시작한 해였다면, 2025년부터는 비즈니스 및 재무적 가치가 실현되는 시대가 될 것입니다. 기업들은 여러 산업 분야에서 비용 절감, 프로세스 효율성, 수익 창출, 서비스 개선 등의 형태로 생성형 AI의 가치를 실현하고 결과를 수치로 확인하기 시작했습니다.

생성형 AI 애플리케이션의 경우, 기반을 이루고 있는 데이터의 정확도만큼만 신뢰할 수 있습니다. 맥킨지 조사에 따르면 응답기업의 72%가 데이터 관리를 AI 구현의 핵심 요건으로 꼽았으며, 보스턴 컨설팅 그룹은 “22%의 기업만이 생성형 AI의 개념 증명 단계를 넘어 어느 정도 가치를 창출하고 있고, 단지 4%만이 상당한 수준의 가치를 만들어냈다”고 밝혔습니다.

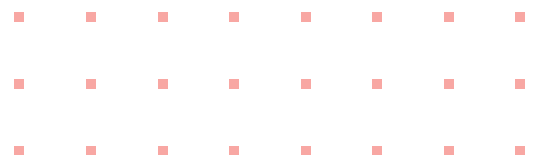
많은 기업은 계속적으로 증가하는 데이터 소스와 늘어나는 데이터 양에 맞춰 확장할 수 있는 데이터 기반을 구현하여 생성형 AI 애플리케이션과 데이터 소비자가 안전하고 개인정보를 준수하는 방식으로 필요한 데이터에 실시간으로 액세스할 수 있도록 하는 데 어려움을 겪고 있습니다. 이러한 데이터 관리의 어려움에도 불구하고 생성형 AI의 이점은 여전히 많습니다. 업계를 선도하는 데이터 관리 공급업체들은 데이터 준비와 소비를 가속화하고, 성능과 비용 효율성 개선을 위해 자동으로 최적화하며, 품질이나 보안을 유지하면서도 증가하는 비즈니스의 데이터 요구에 대응할 수 있도록 생성형 AI 기능을 플랫폼에 추가했습니다.

디노도 플랫폼은 모든 데이터 소비자에게 필요할 때, 필요한 형태로 데이터를 제공할 수 있습니다. 이를 위해 디노도 플랫폼은 데이터 관리 및 통합에 대한 논리적 접근 방식을 활용합니다. 이러한 방식은 생성형 AI 애플리케이션과 데이터 소비자를 위해 여러 데이터 시스템에 대한 액세스를 추상화하여 복잡한 시스템을 숨기고 사전 정의된 시맨틱 및 데이터 거버넌스 규칙에 따라 데이터를 배포합니다.

디노도는 데이터 민주화 실현에 있어 생성형 AI의 혁신적 잠재력을 이해하고 있습니다. 디노도 플랫폼은 데이터 관련 작업을 자동화할 뿐만 아니라 최종 사용자부터 개발자, 운영, 데이터 관리자에 이르기까지 다양한 역할의 사용자들에게 데이터 접근 권한을 부여하고 이를 관리합니다.

본 백서에서는 다음 사항을 자세히 다룹니다:

1. 생성형 AI가 기업에서 조기 도입되어 성공을 거두고 있는 활용 사례
2. 기업에서 생성형 AI를 사용할 때의 문제점
3. 디노도 플랫폼이 문제를 해결하는 방법
4. 생성형 AI가 디노도 플랫폼에서 활용되는 방식



기업에서 사용하는 생성형 AI 사례

가트너는 2026년까지 80% 이상의 기업에서 API 또는 생성형 AI 모델을 활용하거나 또는 운영 환경에서 생성형 AI 애플리케이션을 구현할 것으로 예측하며, 이는 2023년에 발표한 5% 미만에서 크게 증가한 수치입니다.

딜로이트 AI 연구소는 6개의 주요 산업에서 가장 주목할 만한 생성형 AI 활용 사례 60가지를 정리했으며, 가장 일반적인 사례는 다음과 같습니다:

고객 이해: 생성형 AI는 정보를 요약하고 고객 파악을 위한 인사이트를 생성하며 이 정보를 기반으로 타겟팅된 개인화 메시지를 생성할 수 있습니다.

마케팅 콘텐츠 제작: 마케팅 콘텐츠 관리는 다양한 언어로 된 여러 웹사이트를 다루는 기업에게 어려운 과제입니다. 제품 설명, 이미지, 동영상, 오디오 등 개인화된 콘텐츠 제작에 필요한 시간과 리소스를 할당하는 것은 쉽지 않습니다. 생성형 AI는 더 빠르고 일관된 솔루션을 제공하여 효율적인 콘텐츠 생성을 지원합니다.

제품 디자인: 제품 디자인은 많은 아이디어 중 단 하나의 아이디어만이 시장에 출시될 정도로 시간이 오래 걸립니다. 인간의 한계를 극복하면서 다양한 아이디어를 만들고 새로운 영감을 주면서 컨셉 테스트를 개선하기란 쉽지 않습니다. 이와 같은 과정은 생성형 AI를 CAD와 같은 소프트웨어 도구와 통합하여 프로토타입 설계, 창의적인 사고, 브레인스토밍을 촉진함으로써 개선할 수 있습니다.

개발자를 위한 코딩 지원: 개발자와 같은 숙련된 전문 인력의 수요는 높은 반면 인력 공급은 부족한 상황입니다. 이러한 인력 격차를 극복하기 위해 생성형 AI는 코드 생성 및 유지 관리의 일부를 자동화하여 개발자의 노력을 보완함으로써 개발자가 보다 복잡한 코드 작성에 집중할 수 있도록 지원합니다.



고객 지원: 생성형 AI는 자연어로 공감적이고 개인화된 대화를 생성함으로써 애프터서비스 지원 및 고객 불만 처리 분야에서 고객과의 소통을 혁신할 수 있는 잠재력을 가지고 있습니다.

장비 유지보수 계획: 제조 산업에서 장비의 유지보수 계획은 고장과 값비싼 수리를 예방하고 자산의 수명을 연장하는 데 매우 중요합니다. 생성형 AI는 유지보수 일정을 최적화하여 효율적이고 비용 효과적인 계획을 추천하는 동시에 장비 데이터를 분석하여 가동 중단 시간을 최소화하여 장비의 가용성을 극대화합니다.

엔지니어를 위한 가상 어시스턴트: 생성형 AI 기반 가상 어시스턴트는 광범위한 기술 정보에 신속하게 액세스할 수 있습니다. 가상 어시스턴트는 엔지니어가 요청한 세부 정보를 제공하고 엔지니어에게 적합한 리소스를 안내하는 것 외에도 특정 엔지니어링 개념과 원리, 계산과 관련된 문의를 처리하여 문제 해결을 지원할 수 있습니다.

물류 및 공급망 관리: 생성형 AI는 공급망의 잠재적인 중단과 위험을 식별하여 시뮬레이션해 보는 데 도움을 줄 수 있습니다. 생성형 AI는 항만 혼잡도, 배송 경로, 공급업체를 분석하여 잠재 위험과 그에 따른 운영 영향을 예측하고 이러한 위험을 완화하기 위한 조치를 추천할 수 있습니다.

디지털 공공 서비스: 정부 및 공공 서비스 데이터는 온프레미스와 클라우드에 여러 형식으로 분산되어 있는 경우가 많습니다. 생성형 AI 가상 비서가 정부와 일반 시민 사이에서 인터페이스 역할을 수행하며 자연어로 질의 응답이 가능합니다.

지금까지 살펴 본 내용은 생성형 AI가 비즈니스 운영 방식을 어떻게 변화시키고 있는지를 보여주는 몇 가지 예에 불과합니다. 향후 몇 년간은 인터넷에 버금가는 창의성과 혁신으로 생성형 AI가 우리의 생활과 업무 방식을 변화시킬 수 있는 잠재력이 더욱 커질 것으로 예상됩니다.



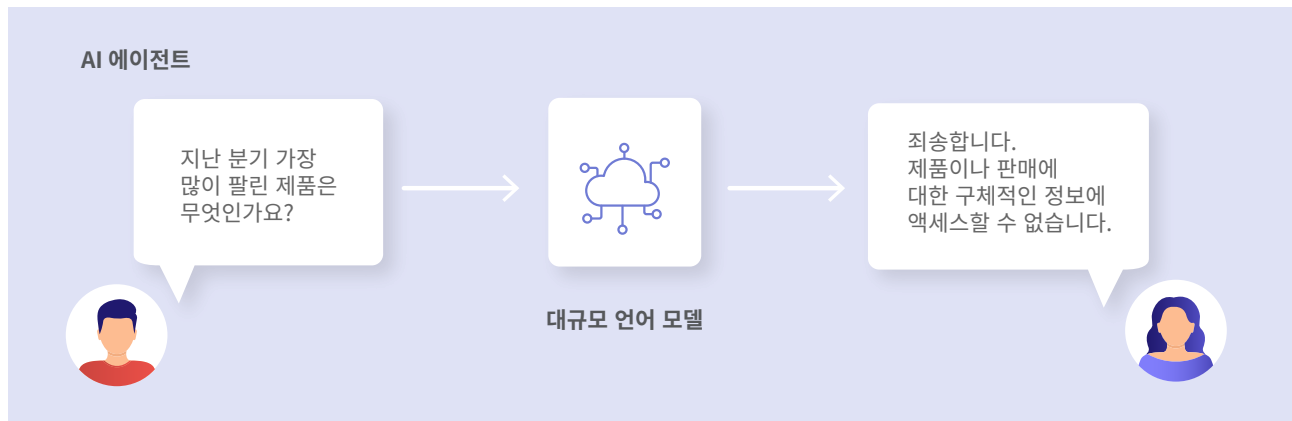
기업에서 생성형 AI를 사용하는 데 따르는 어려움

기업이 생성형 AI의 혁신 잠재력을 인식하면서 관련 기술과 역량에 투자하기 시작했고 비즈니스 적용을 시험해 보고 있습니다. 하지만, 데이터와 관련된 다양한 문제에 직면하게 됩니다. 맥킨지에 따르면, 72%의 기업에서 데이터 관리를 생성형 AI 활용 범위 확장에 있어 풀어야 할 핵심 과제로 인식하고 있습니다. 예를 들어, 생성형 AI를 통한 고객 맞춤 서비스가 안전하고 효과적이라면 기초가 되는 고객 데이터가 정확하고 최신 상태여야 하며 다양한 데이터 소스로부터 실시간에 가깝게 데이터에 액세스할 수 있어야 합니다.

“

데이터 품질 저하, 부적절한 위험 관리, 비용 상승, 불분명한 비즈니스 가치로 인해 2025년 말까지 개념 증명 이후 생성형 AI 프로젝트의 최소 30%가 중단될 것입니다.

— 가트너



특정 기업 고유의 질문에 답변하지 못하는 생성형 AI 에이전트

기존 LLM 모델을 추가 정보로 훈련시키고 미세 조정하여 기업 데이터와 정보를 인식하게 할 수는 있지만, 이 방법은 종종 더 복잡하고 어려운 문제로 이어집니다. LLM을 다시 교육하는 데 필요한 기술과 비용이 대부분의 기업에 부담스러울 뿐 아니라, LLM이 최신 데이터와 정보를 인식하도록 지속적으로 만들어야 하기 때문에 LLM을 재교육시키는 반복적인 과정은 비현실적입니다.

또한, 민감 정보를 갖고 있는 기업 내부 정보를 LLM 모델 자체에 포함시키기 위해 LLM을 훈련하는 것은 데이터 유출의 가능성이 항상 존재하기 때문에 위험이 따릅니다. AI에 정통한 사용자는 LLM이 정보의 민감도를 인식하도록 훈련된 경우에도 그러한 정보를 공개하도록 유도하는 프롬프트를 만들 수 있습니다. LLM 재교육은 완벽하지 않습니다.

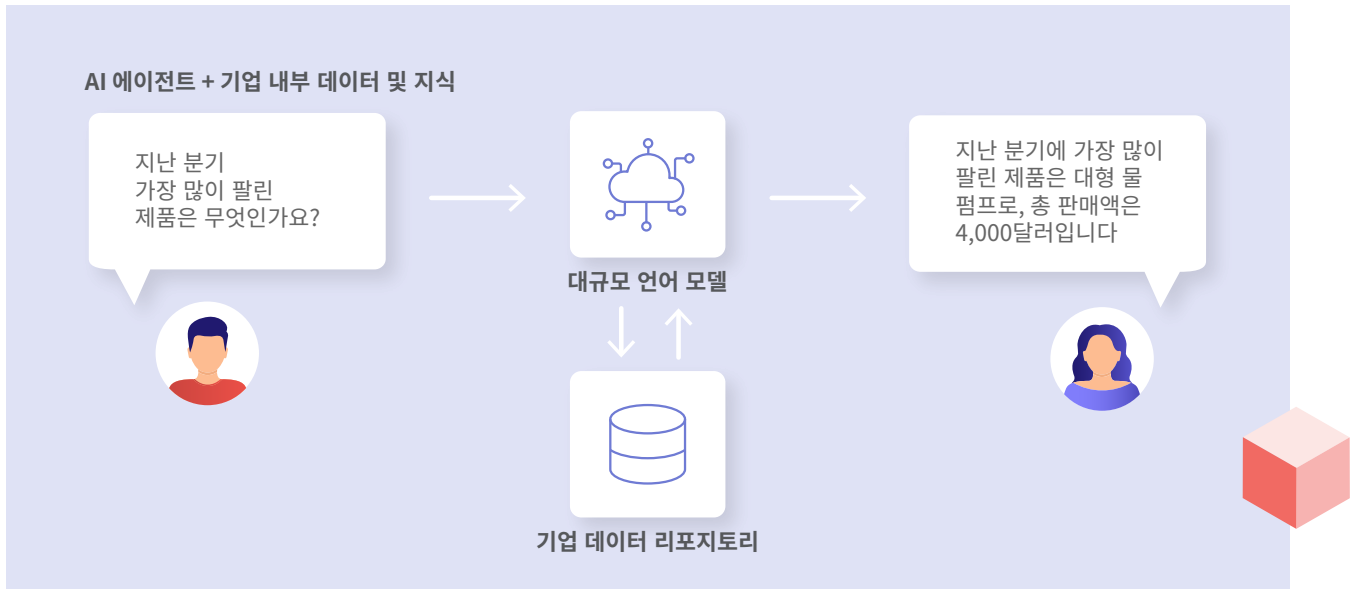
“

AI 전략에 있어 그다지 중요하지 않은 것은 대규모 언어 모델의 선택입니다. 좋은 옵션이 많이 있으며, 모두가 그것을 사용할 것입니다. 대신, 현명한 전략은 여러분을 차별화할 수 있는 요소, 즉 기업 내부의 지식과 데이터를 이용하는 AI 활용 방법을 찾아내는 것입니다.”

— 2025년 PwC인공지능 비즈니스 전망

다행히도, 새롭게 등장한 아키텍처 덕분에 LLM의 한계를 극복하고 기업 환경에서 필요한 지식을 안전하고 효과적인 방법으로 제공할 수 있을 것으로 기대됩니다. 검색 증강 생성(RAG) 아키텍처는 2020년 Meta의 논문에서 처음 논의되었고, 추가 데이터와 정보로 LLM을 증강하는 데 있어 비용 효율적이고 안전한 방법으로 빠르게 자리 잡고 있습니다.

RAG를 사용하면 실시간 데이터를 생성형 AI의 답변 결과에 반영할 수 있지만, 관련성이 높고, 품질이 우수하며, 개인정보 보호 규정을 준수하는 데이터를 실시간으로 제공하려면 포괄적인 데이터 관리 솔루션이 필요합니다.



기업 내부 데이터와 지식으로 증강된 생성형 AI 에이전트

기존의 머신러닝 활용 사례에서 데이터는 초기 학습 과정에서 중요한 역할을 했지만, 생성형 AI와 RAG는 이러한 패러다임을 바꾸고 있으며, 추론 과정에서 데이터로 하여금 지식 보강 계층의 역할을 하도록 요구합니다. 모델을 재학습시켜 LLM에 지식을 주입하는 대신, RAG 아키텍처는 프롬프트 과정에서 추가적인 컨텍스트 창을 통해 지식을 추가하는 방식을 사용합니다. 그런 다음, LLM은 추가적인 컨텍스트를 활용하여 LLM 자체에 지식이 내장되어 있지 않아도 필요한 응답을 생성할 수 있습니다.

RAG의 장점은 환각 현상 감소, 실시간 최신 정보, 제품 판매와 같은 특정 도메인 지식, 비용이 많이 드는 재교육 제거, 검색 데이터 출처에 대한 가시성 등이 있습니다.

RAG 접근 방식은 정형 및 비정형 정보에 적용할 수 있으며, LLM에 추가 지식과 정보를 제공하는 데 있어 보다 민첩하고 비용 효율적이며 안전한 방법입니다. LLM의 강력한 코드 생성 기능과 함께 RAG를 사용하는 차세대 AI 애플리케이션은 새로운 유형의 사용자 상호 작용과 기업에서 데이터의 가치를 극대화할 수 있는 새로운 방법을 열어줄 수 있습니다.

그러나, 기업 환경에서 RAG를 구현하려면 데이터 관리 기반이 필요합니다. 데스크톱 환경에서 단일 테이블에 대한 RAG의 성능을 보여주는 간단한 GitHub 프로젝트를 많이 찾을 수 있지만, 사일로화된 복잡한 기업 데이터 환경에서 실제 데이터에 대하여 RAG 아키텍처를 구현하는 것은 악몽과도 같은 일입니다. LLM과 기업 데이터 소스 간의 통합이 필요함에 따라, 데이터 사일로 극복, 다양한 데이터 소스 유형 처리, 길고 복잡한 데이터 전달 파이프라인 관리와 같은 데이터 관리 문제에 필연적으로 직면하게 될 것입니다.

기업은 생성형 AI를 도입할 때 다음과 같은 데이터 관리 문제를 고려해야 합니다.

규정: 새로운 규정, 지침, 프레임워크가 빠르게 등장하고 있습니다. 이러한 규정은 생성형 AI 시스템의 기능, 시스템 훈련에 사용되는 데이터, 편향성 및 공정성 문제, 지적 재산권 침해, 개인 정보 침해, 제3자 위험, 보안 문제 등 생성형 AI의 투명성 부족을 해결하기 위한 것입니다. 유럽연합의 AI 법은 투명성, 데이터 거버넌스, 관리와 감독, 위험 평가에 대한 구체적인 요구 사항과 함께 고위험 AI 시스템에 대한 포괄적인 규정 수립을 목표로 합니다. 미국 대통령이 발표한 행정명령은 AI 안전과 보안에 대한 새로운 기준을 수립하고, 미국인의 사생활을 보호하며, 형평성과 시민권을 향상시키고, 소비자 및 근로자를 보호하고, 혁신과 경쟁을 촉진하는 등 다양한 내용을 담고 있습니다. 중국은 2023년 세계 최초로 생성형 AI를 구체적으로 규제하는 법을 제정하여, 이러한 서비스를 제공하는 기업에 훈련 데이터 및 산출물에 관한 제한을 부과했습니다. 한국의 인공지능법은 고위험 인공지능 서비스에 대한 엄격한 고지 요건과 인공지능 인증을 통해 인공지능 시스템의 신뢰성을 보장함으로써 인공지능 서비스 사용자 보호를 목표로 합니다. 다른 국가의 관련 규정과 지침에서도 데이터 수집을 위한 개인정보 보호 강화 기술의 사용을 권장하거나 의무화하고 있습니다. 싱가포르의 생성형 AI를 위한 거버넌스 모델 프레임워크 (Model AI Governance Framework for Generative AI)이 좋은 예입니다.

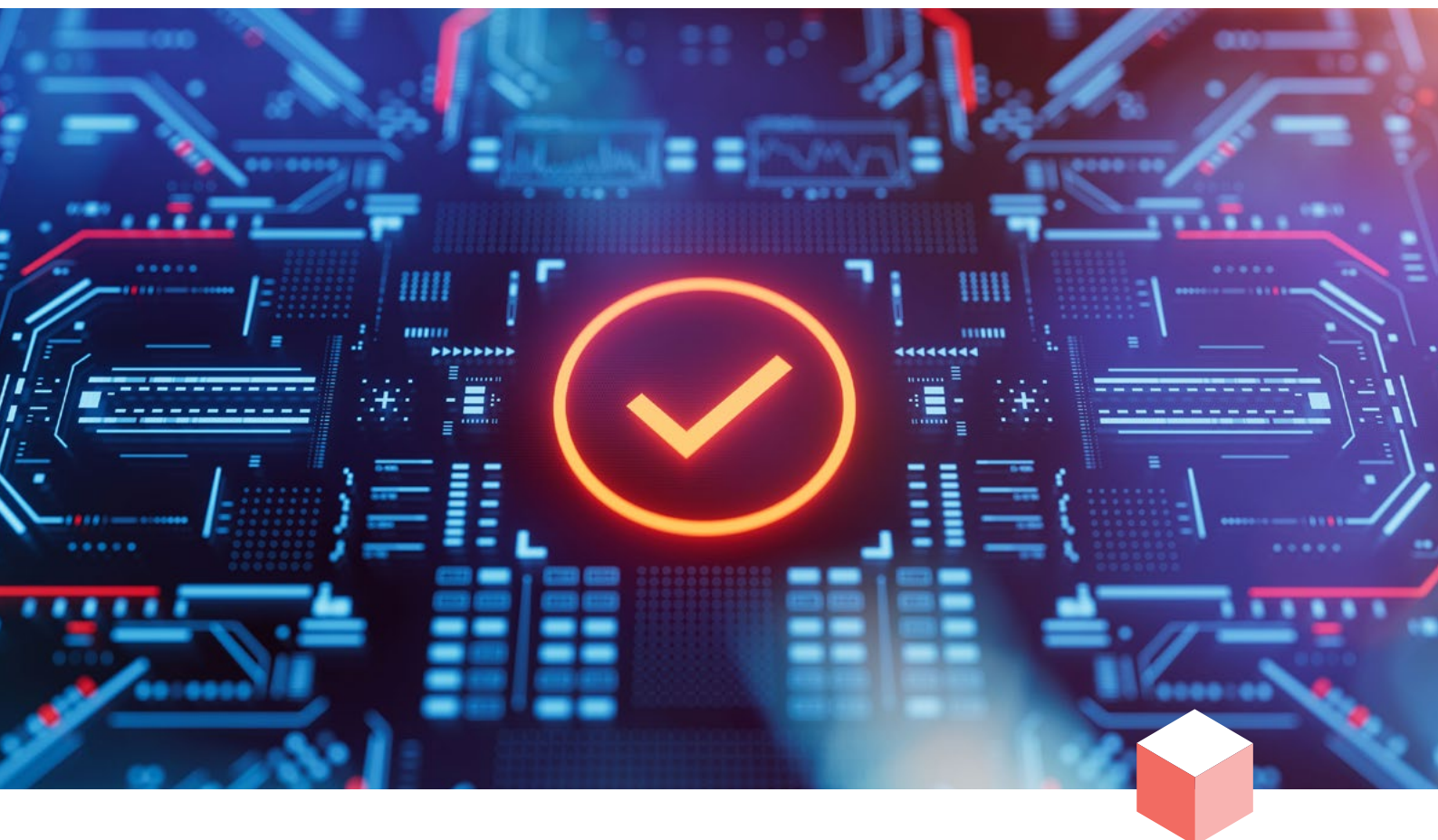
또한 이러한 규정은 생성형 AI가 만든 결과물에 대한 설명과 감사 가능성을 요구합니다. 예를 들어, 유럽 연합의 AI법은 생성형 AI가 편견이 없고 소비자에게 피해를 주지 않는다는 것을 감사할 수 있고 소비자와의 소통 등 특정 상호작용 중에 생성형 AI가 사용되었다는 사실을 소비자에게 투명하게 공개할 것을 요구합니다. 이와 같은 규정에 대응하려면 특정 결과물을 만들 때 생성형 AI가 어떤 데이터를 사용했는지 확인할 수 있어야 하며, 생성형 AI가 활용되는 분야에서 어떤 데이터를 언제 사용하는지 모니터링할 수 있는 기능을 갖추어야 합니다.

데이터 품질과 설명 가능성: 품질이 낮은 데이터는 잘못된 행동이나 일관성이 없는 행동, 즉 ‘쓰레기 정보 투입-쓰레기 정보 배출’을 초래할 수 있습니다. 이는 LLM의 훈련 과정과 RAG를 통한 실시간 기업 데이터 액세스 과정 모두에서 마찬가지입니다. 그 결과로 나타나는 환각은 최종 사용자의 불신으로 이어지고, 윤리적 AI 사용에 관한 규제를 받는 기업의 입장에서는 잠재적인 규정 위반으로 이어질 수 있습니다.

결과적으로, 엄격한 데이터 품질이 필요하지만, 그것만으로는 충분하지 않습니다. 생성형 AI는 또한 ‘데이터 설명 가능성’을 요구하는데, 이는 결과를 생성하기 위해 어떤 데이터에 액세스했는지 명시할 수 있는 기능입니다. 이 방법을 통해 생성형 AI 사용자는 사용된 데이터가 정확하고 올바른지 쉽게 확인할 수 있으며, 그렇지 않은 경우 쉽게 수정할 수 있습니다. 데이터 설명 가능성은 투명성을 제공하며, 이는 최종 사용자의 신뢰를 높이는 데 도움이 됩니다. 실제로 위에서 언급한 규정 중 일부, 특히 유럽연합의 AI 법은 이러한 수준의 투명성을 요구합니다.

개인 정보 보호 및 보안: 유럽의 GDPR과 같은 개인 정보 보호 규정이 여전히 적용되지만, 올바른 데이터 접근 통제 및 데이터 비식별 조치의 시행을 보장하는 것만으로는 충분하지 않습니다. 생성형 AI는 추가적인 위험을 야기합니다.

생성형 AI 모델이 어떻게 개인 정보 보호 위험을 초래할 수 있는지를 보여주는 예를 살펴보겠습니다. Jane Doe라는 사람의 개인 정보가 생성형 AI 모델 훈련에 사용된 경우, ‘제 이름은 Jane Doe이고 제 전화번호는...’와 같은 프롬프트가 개인 정보를 노출시킬 수 있습니다. 생성형 AI 모델이 훈련 데이터를 기억하므로 문제가 더욱 복잡해졌습니다. 훈련 데이터 추출과 같은 유형의 공격이 LLM으로부터 민감한 데이터를 유출하는 것으로 나타났습니다. 위에서 언급한 모든 AI 규정은 기업이 개인 정보 침해 위험을 완화하기 위한 통제 수단을 갖추도록 요구합니다.



디노도에서 문제를 해결하는 방법 - 논리적 데이터 관리

디노도는 데이터 관리를 새롭게 구상하고 발전시켜야 한다고 생각합니다. 그 어느 때보다 위험이 높기 때문입니다. 기업은 지금 바로 조치를 취하여, 빠르게 다가오고 있는 새로운 AI 기반의 경쟁 환경에 대비할 수 있도록 강력한 데이터 관리 기반을 구축해야 합니다.

차세대 AI 애플리케이션을 구동하는 데 필요한 데이터 관리 기능을 고려해 볼 때, 데이터 관리에 대한 논리적 접근 방식이 필요하다고 생각합니다.

데이터 가상화 기술을 활용하는 디노도 플랫폼은 데이터 복제나 이동이 필요 없는 논리적 데이터 관리 솔루션입니다. 디노도 플랫폼은 통합된 데이터에 대한 단일 게이트웨이를 AI 애플리케이션에 제공합니다. 디노도 플랫폼은 다음과 같은 주요 이점을 제공합니다.

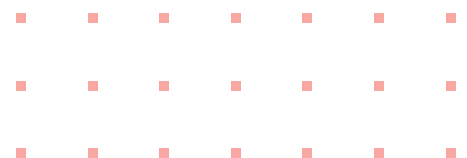
- ERP, 운영 데이터 마트, EDW, 애플리케이션 API 등 모든 기업 데이터와 상호 작용하고 쿼리할 수 있는 통합된 보안 액세스 포인트
- 비즈니스 시맨틱 레이어, LLM에 필요한 비즈니스 컨텍스트와 지식 제공
- LLM이 사용하기 어려운 기술적인 데이터 뷰로부터 추상화되고 분리된 논리 데이터 뷰의 빠른 배포
- 여러 데이터 세트를 물리적으로 결합할 필요 없이 LLM 친화적인 논리 테이블 뷰 전달
- 쿼리 최적화 기능이 내장되어 있어, LLM이 특정 데이터 소스의 제약 조건이나 조인 전략을 따르지 않아도 됨

이와 같은 장점을 제공하는 디노도 플랫폼은 RAG를 위한 완벽한 도구입니다. 시맨틱 레이어는 데이터 플랫폼에 저장된 데이터에 접근할 수 있도록 생성형 AI 애플리케이션을 위한 일관되고 안전한 인터페이스를 제공합니다. 또한, 디노도 플랫폼은 상황 정보가 포함된 필드 설명, 필드의 비즈니스 이름, 데이터 스키마 등 생성형 AI 애플리케이션에 정보를 제공하는 데 필요한 메타데이터를 갖추고 있습니다.

쿼리 RAG

RAG는 주로 비정형 데이터로 LLM을 보강하는 데 초점을 맞추고 있습니다. 그 결과, 웹사이트와 PDF에서 스크랩한 데이터를 사용하여 RAG를 구현하는 기술이 빠르게 발전했고, 그러한 기술의 견고성과 확장성이 입증되었습니다. 그러나, 정형 데이터로 LLM을 보강하는 것은 다음과 같은 일련의 문제를 야기합니다.

- 테이블 형식의 데이터를 가져와야 할 뿐만 아니라 올바른 SQL을 통해 액세스하고 쿼리해야 합니다.
- 데이터 검색에 필요한 액세스 메커니즘과 최적의 SQL은 공급업체마다, 플랫폼마다 다를 수 있습니다.
- 벡터 임베딩 전략은 메타데이터와 실제 데이터를 모두 포함해야 합니다.
- 정형 데이터 및 관련 메타데이터는 끊임없이 변화합니다.
- 기존의 데이터 보안 및 권한 관리 기능을 수용하고 통합해야 합니다.

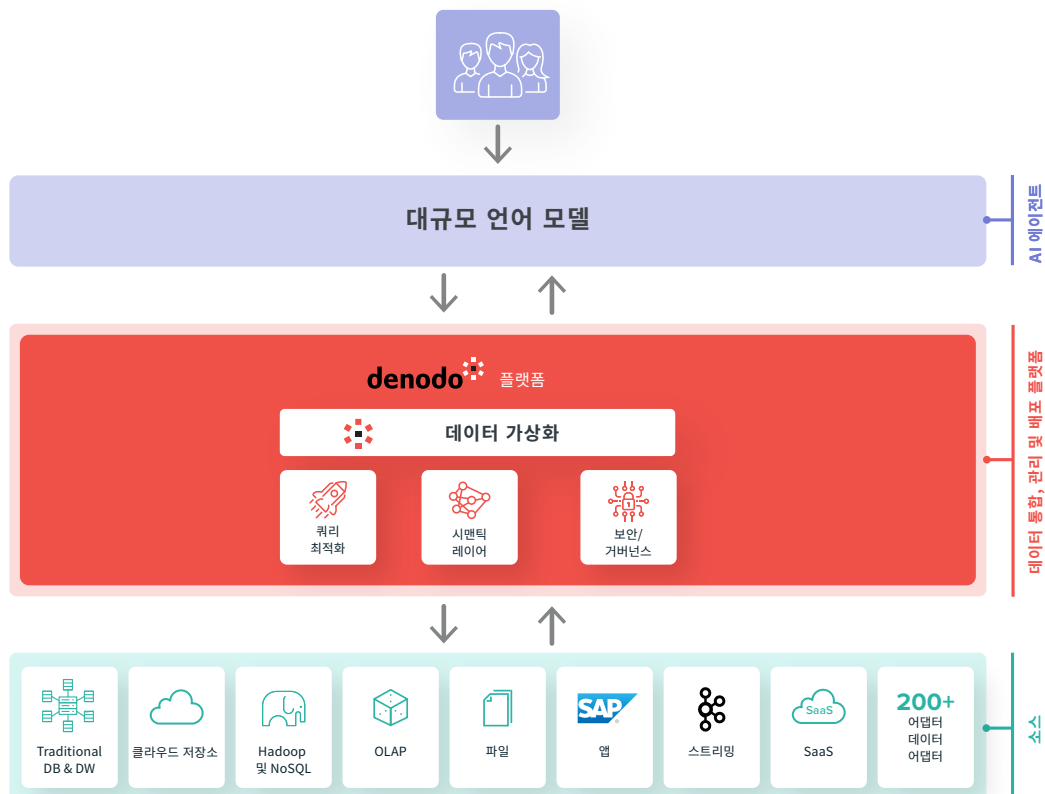


이러한 문제는 정형 데이터를 처리하는 데 있어 다른 접근 방식의 필요성을 강조합니다. 디노도 플랫폼은 이를 해결하기 위해 설계된, RAG의 변형인 쿼리 RAG의 구현을 지원합니다. 쿼리 RAG는 다음과 같은 이점을 제공합니다.

- 모든 메타데이터와 데이터 소스, 플랫폼에 걸쳐 벡터 임베딩을 생성하는 강력하고 확장 가능한 방법
- 데이터 소스와 시스템의 복잡성을 추상화하는 통합 SQL 액세스 엔진
- LLM이 생성한 SQL이 여러 데이터 아키텍처 패턴과 사용 사례에서 높은 성능을 발휘할 수 있도록 하는 강력한 쿼리 최적화 도구.
- 새로운 데이터 뷰를 쉽게 개발하고 메타데이터를 확장할 수 있는 유연하고 강력한 시맨틱 레이어.
- 모든 단계에서 사용자 권한과 데이터 보안을 확인하고 적용하는 간단한 프로세스.

쿼리 RAG의 독특한 특징 중 하나는 임베딩 검색의 목표가 프롬프트를 보강할 텍스트를 찾는 것이 아니라 쿼리할 올바른 데이터 뷰를 찾는 것이라는 점입니다. 올바른 데이터 뷰와 올바른 SQL이 확인되면, 테이블 형식의 데이터가 여러 데이터 소스에서 검색되어 LLM에 제공됩니다.

쿼리 RAG를 사용하면 데이터 저장소를 강력하고 확장 가능한 방식으로 활용할 수 있습니다. 쿼리 RAG는 실시간 데이터를 AI 애플리케이션에 내장하고 노출함으로써 비정형 데이터를 넘어 생성형 AI 기능을 확장하고, 실시간 통찰력을 제공하고, 새롭고 혁신적인 방식으로 의사 결정을 개선할 수 있도록 지원합니다.



AI 상호작용을 강화하고 지원하는 디노도 플랫폼

디노도 플랫폼은 생성형 AI 애플리케이션의 가장 중요한 구성 요소인 데이터에 대한 통합된 단일 공급자 역할을 할 수 있습니다. LLM과 디노도 플랫폼의 조합은 강력하고 효율적인 AI 에이전트의 개발을 훨씬 더 빠르게 할 수 있습니다.

디노도 플랫폼은 생성형 AI를 위한 데이터 기반을 제공하기 위해 다음과 같은 기능을 추가로 제공합니다:



데이터 품질 및 사용 모니터링

디노도 플랫폼은 변환, 필터, 매칭 기능을 포함한 다양한 라이브러리와 데이터의 유효성 검증, 정제, 보강, 표준화, 매칭, 병합을 위한 품질 관리 규칙을 제공합니다. 여기에는 조건부 처리, 분할, 중복 제거 및 정제를 위한 퍼지 매칭 알고리즘 등이 포함됩니다. 또한, 디노도 플랫폼은 누가, 언제 어떤 데이터에 액세스했는지 지속적으로 모니터링합니다. 디노도 플랫폼은 RAG를 통해 시작된 모든 생성형 AI 쿼리에 대해 쿼리뿐만 아니라 반환된 데이터도 모두 기록하여, 언제나 설명 가능하고 투명한 환경을 유지합니다.



데이터 보안 및 개인정보 보호 규정 준수

애플리케이션에 대한 일원화된 데이터 접점으로서 디노도 플랫폼은 데이터 필드 수준까지 세분화할 수 있는 기준 모델에 따라 데이터 접근 보안 규정을 정의하고 이를 시행하기 위한 일종의 중앙 집중식 레이어를 제공합니다. 디노도 플랫폼은 스키마 전체에 대한 권한과 데이터에 대한 개별적인 권한을 이용하여 사용자 및 역할 기반의 인증과 권한을 부여하는 방식을 취합니다. 디노도 플랫폼은 행과 열 기반의 보안을 모두 제공하며, 특정 필드를 가릴 수 있는 기능도 갖추고 있습니다. 예를 들어, 상급자의 ‘급여’ 열을 볼 수 없도록 지정할 수 있어, 해당 셀이 결과에서 가려진 채로 표시됩니다.

이러한 기능이 있기 때문에 디노도 플랫폼은 RAG가 적용된 AI 애플리케이션의 데이터 보안을 강화하고, 위에서 설명한 바와 같이 기업이 다양한 법적 요건을 이행하는 데 적합한 플랫폼입니다.

LLM 기반 AI 에이전트의 잠재력을 최대한 활용하기 위한 여정은 계속되고 있으며, 이를 위해서는 전체 업계가 지속적으로 기술 발전과 혁신을 이루어야 합니다. 디노도는 AI 시대의 요구에 부응하기 위해 꾸준히 제품을 개선하고 있습니다. 디노도의 숙련된 제품 개발팀은 데이터 관리를 위한 더 나은 기반을 제공하기 위해 AI 및 기타 최신 기술의 발전을 지속해서 모니터링하고 평가합니다. 디노도는 고객이 이러한 혁신적인 기술을 효과적으로 활용할 수 있도록 최선을 다하고 있습니다.



차세대 생성형 AI 에이전트 및 애플리케이션 실현

차세대 생성형 AI는 고객 소통, 비즈니스 운영 등 기업 부문의 여러 측면을 향상시키거나 자동화하는 특정 목적의 에이전트와 애플리케이션으로 구성될 것입니다. 이러한 에이전트와 애플리케이션은 LLM이 모바일 기기에 내장되거나 최신 애플리케이션 스택의 인메모리로 제공되면서 그 수가 급증할 것입니다. 그렇지만, 이런 스마트 앱 개발자들은 다음과 같은 데이터 관리 문제를 겪는 일이 많습니다.

- 사일로화된 복잡한 데이터 소스에 연결하고 그 데이터에 접근하는 일
- AI 앱이 소비할 데이터 프로젝트의 반복적 개발을 쉽게 하고 개발 속도를 높여서, 이 작업이 앱을 실제로 배치하는 데 방해가 되지 않도록 하는 일
- 애플리케이션 운영 환경을 고려하여 확장이 가능한 실시간 데이터 액세스

- 애플리케이션 성능과 정확도를 개선하기 위한 풍부하고 통일된 시맨틱 레이어 개발
- 다양한 데이터 저장소와 기술 전반에 적용되는 데이터 보안과 거버넌스 지원

디노도 AI SDK는 이러한 문제를 해결하여 AI 기반 에이전트와 애플리케이션 개발자들에게 탁월한 생산성과 민첩성을 제공합니다. 디노도 AI SDK는 오픈 소스 소프트웨어 패키지로, 사전에 통합된 컴포넌트와 재사용 가능한 최적화된 코드를 가지고 있습니다. 디노도 플랫폼을 사용하고 있는 고객들은 SDK를 다운로드하여 RAG 기반 아키텍처를 구현하고 AI 에이전트나 애플리케이션의 개발을 빠르게 진행할 수 있습니다. SDK는 환경 설정 파라미터를 통해 널리 사용되는 상용 LLM과의 통합을 쉽게 해 줍니다. 또한, 이 패키지에는 프론트 엔드 애플리케이션과 챗봇의 샘플이 포함되어 있어, 테스트와 자체 개발에 도움이 됩니다.

쿼리 RAG를 위해 특별히 설계된 API 세트를 제공함으로써, 디노도 플랫폼과 AI SDK는 어떤 데이터 저장소에서도 대화형 사용자 경험을 쉽게 구축할 수 있도록 지원합니다. 이는 데이터 분석 작업에 필요한 질문은 물론, 실시간으로 데이터를 이용해야 하는 운영 요청도 지원할 수 있다는 의미입니다. 이 모든 과정에서, 디노도 플랫폼은 추가적인 수동 개입 없이도 데이터 접근과 사용을 통제 가능한 범위 안에 두는 등 필요한 데이터 보안과 거버넌스 정책을 시행합니다.

SDK는 개발자에게 다음과 같은 이점을 제공합니다

- 여러 데이터 소스와의 통합을 쉽게 만들어 커스텀 커넥터와 수동 데이터 처리의 필요성을 줄여줍니다.
- 데이터 변환과 병합을 더 효율적으로 만들어 결합 분석과 LLM 입력을 위해 데이터를 자동으로 준비합니다.
- 임베딩, 벡터화, LLM 쿼리를 단일 흐름으로 처리하는 기본 제공 API를 통해 오케스트레이션의 복잡성을 줄여줍니다.
- SDK 내에서 액세스 제어 및 컴플라이언스를 관리하여 보안 및 거버넌스를 강화하고 위험을 최소화합니다.
- 반복적인 코딩 작업을 없애고 더 가치 있는 기능에 집중함으로써 개발 기간을 단축합니다.
- 데이터 일관성과 정확성을 향상시켜 생성형 AI 애플리케이션에서 상황에 맞는 적절한 응답을 가능하게 합니다.
- 오픈 소스 컴포넌트를 통해 확장성과 유연성을 지원하여 특정 용도에 맞는 수준의 커스터마이제이션을 제공합니다.



생성형 AI가 데이터 관리를 향상시키는 방법

생성형 AI는 또한 데이터 관리 개선에도 사용되었습니다. AI는 사람의 실수가 발생하기 쉬운 수작업 업무에 매우 가치 있는 자동화 기능을 선보입니다.

지금까지 노동 집약적이고 오류가 발생하기 쉬웠던 기초적인 데이터 관리 작업을 AI를 통해 보다 효율적이고 정확하게 처리할 수 있습니다. 가트너의 2025년 데이터 관리 계획 가이드에 따르면, 2026년까지 생성형 AI는 많은 수작업을 필요로 했던 데이터 관리 비용을 매년 최대 20%까지 줄이는 동시에 4배 더 많은 새로운 활용 방법을 가능하게 할 것입니다.

디노도 어시스턴트(Denodo Assistant)는 지능형 데이터 패브릭의 실현을 가능하게 함으로써 사용자의 생산성을 향상시킵니다. 디노도 어시스턴트는 디노도 플랫폼이 제공하는 시맨틱 레이어와 액티브 메타데이터, AI와 머신 러닝 기술을 활용하여 디노도 플랫폼의 모든 기능을 강화합니다.

디노도에서 개발한 알고리즘을 활용하고 업계 선두업체가 제공하는 LLM과 통합함으로써, 디노도 어시스턴트는 데이터 프로젝트 개발 전 과정에서 개발자와 사용자를 지원합니다. 질문에 지능적으로 답변하고, 제안을 제공하며, 해결책을 추천합니다.

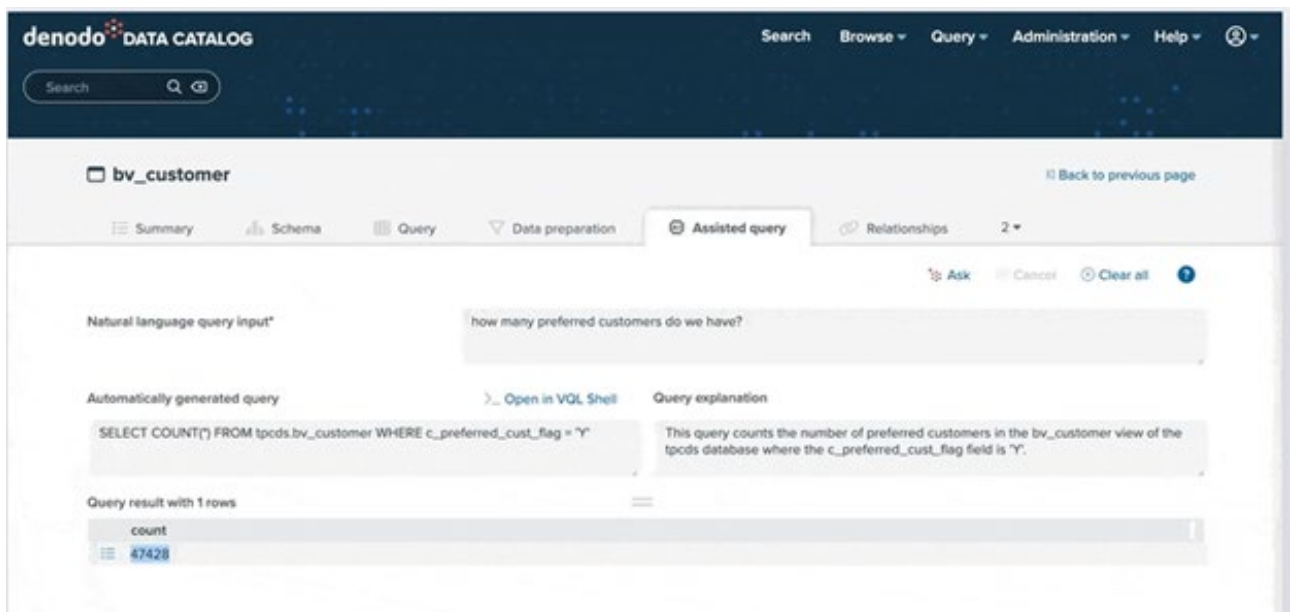
디노도 어시스턴트는 다음과 같은 기능으로 구성되어 있습니다:

자연어 쿼리. 누구나 자연어로 질문을 입력할 수 있으며, 입력된 질문은 SQL 코드로 변환되어 플랫폼에서 실행됩니다. 이것은 데이터 민주화의 큰 진전입니다. SQL이나 BI 도구에 대한 지식이 없는 비즈니스 전문가도 데이터에서 인사이트를 얻을 수 있게 해줍니다. 예를 들어, 영업 관리자는 ‘미국 내 모든 선호 고객의 이름, 성, 이메일 주소가 필요합니다’와 같은 일상적인 언어를 사용하여 데이터를 쿼리할 수 있습니다. 생성형 AI를 사용하는 데이터 플랫폼은 자연어 쿼리를 SQL 코드로 변환하고, 코드의 설명까지 생성할 수 있습니다. 그런 다음 데이터 플랫폼은 코드를 실행하고 결과를 가져옵니다.

“

디노도의 생성형 AI 기능은 누구나 자연어를 사용하여 데이터에 액세스하고 데이터와 상호 작용할 수 있도록 하는 획기적인 혁신입니다. 데이터베이스 이름이나 쿼리 언어와 같은 기술적 세부 사항을 알 필요가 없도록 해줍니다. 걸림돌을 제거함으로써 디노도의 생성형 AI 기능은 데이터 액세스를 단순화하고 가치를 극대화하며 데이터 민주화를 촉진할 수 있습니다.

— 마크 소로굿(Mark Thorogood), 미국 법률회사 퍼킨스 코이(Perkins Coie) 데이터 운영 책임 이사



이것은 데이터 분석가, 마케팅 담당자, 영업 사원, 의료 전문가 등 비즈니스 전문가에게 큰 잠재력을 가지고 있습니다. 이들은 IT 및 데이터 전문가의 도움 없이도 데이터에 독립적으로 액세스하고 통찰력을 얻을 수 있습니다. 디노도 데이터 카탈로그에는 데이터 검색 및 셀프 서비스를 지원하는 추천 데이터 뷰와 데이터 자산이 포함되어 있습니다. 이를 통해 비즈니스 의사 결정권자는 훨씬 더 민첩하게 의사 결정에 필요한 데이터를 얻을 수 있습니다.

“

2025년까지 자연어 기반의 데이터 관리 API가 데이터 생태계 전반에 걸쳐 데이터 소비를 100배 증가시키는 지배적인 인터페이스가 될 것입니다.

— 데이터 및 분석 분야의
주요 트렌드, 2024 가트너

쿼리 최적화. 디노도 플랫폼은 조인 순서, 연결, 집계 등 쿼리 최적화를 위해 오랫동안 AI를 사용해 왔습니다. 디노도 어시스턴트에는 머신러닝과 액티브 메타데이터를 사용하는 요약 캐시가 포함되어 있습니다. 이 기능은 비용을 줄이면서 쿼리 성능을 최적화하는 데 도움이 됩니다.

데이터 준비 및 확충: 디노도 플랫폼은 AI가 추천하는 데이터 변환을 비롯한 데이터 준비 기능도 갖추고 있습니다.

데이터 사용자는 특정 데이터 세트에 액세스할 때, 동일하거나 유사한 데이터 세트를 사용한 이전의 행동에 근거하여 데이터 변환을 추천하는 일련의 정보를 제공받습니다. 이 기능은 사용하는 횟수가 늘어날수록 향상됩니다. 디노도 플랫폼은 데이터에 대한 메타데이터를 학습하고 업데이트하기 때문입니다. 여기에는 데이터가 사용된 시기와 장소, 다른 데이터와의 관계 등이 포함됩니다.

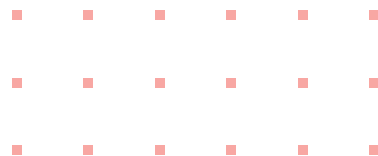
디노도 어시스턴트는 데이터 거버넌스 전문가의 부담을 줄여주기 위해, 뷰와 컬럼에 대한 설명을 제안합니다. 개발자의 생산성을 향상시키기 위해, 디노도 어시스턴트는 ‘SQL 및 쿼리 마법사 권장 사항에 대한 자동 완성’과 같은 기능을 갖추고 있습니다.

AI가 실현하는 미래를 위한 준비

인공지능(AI)은 모든 산업 분야에서 변화를 일으키는 원동력이 되었습니다.

LLM에 쉽게 접근할 수 있고 누구나 똑같은 파운데이션 모델을 사용한다면, 여러분에게 경쟁 우위를 가져다줄 수 있는 것은 바로 데이터입니다. 그런 관점에서 보면, 실제로 변한 것은 없습니다. 변한 것은 AI의 모든 잠재력을 발휘하기 위해 여러분이 데이터 관리에 대해 어떻게 생각해야 하는가입니다.

생성형 AI가 실현할 미래의 문턱에 서 있는 지금, 다음 질문을 던져 보는 것이 매우 중요합니다. 여러분은 생성형 AI를 효과적으로 활용하는 것은 물론 지속 가능한 비즈니스 경쟁 우위를 확보하기 위한 데이터 인프라와 데이터 관리 기반을 갖추고 있습니까?



디노도테크놀로지(이하 디노도)는 데이터 관리 솔루션 선도 기업이다. 다수의 수상 경력을 자랑하는 디노도 플랫폼(Denodo Platform)은 선도적인 논리적 데이터 관리 플랫폼으로, 데이터 셀프서비스, 인공지능(AI) 등 기업 전반의 데이터 관련 전략을 위해 데이터를 신뢰할 수 있는 인사이트와 자산으로 전환한다. 전 세계 모든 산업의 디노도 고객들은 데이터 레이크하우스 및 타 데이터 플랫폼 사용 시와 비교했을 때 3분의 1로 단축된 시간 안에 10배 더 우수한 성능으로 인공지능과 비즈니스에 적합한 신뢰할 수 있는 데이터를 제공했다. 더 자세한 내용은 디노도코리아 홈페이지(www.denodo.com/kr)에서 확인할 수 있다.