



2019년
비즈니스
인텔리전스
동향

목차

- 01 설명 가능한 AI의 부상**
조직이 인공 지능 및 기계 학습 모델에 점점 더 의존하는 이때, 그 기술이 신뢰할 만하다고 어떻게 보장할 수 있을까요?
- 02 데이터를 인간화하는 자연어**
NLP 시스템의 발전으로 모든 사람이 데이터와 자연스러운 대화를 나눌 수 있습니다.
- 03 실행 가능한 분석으로 상황에 맞는 데이터 제공**
BI 플랫폼은 사람들이 조치를 취하려는 상황에 적합한 데이터를 제공하도록 진화합니다.
- 04 사회 공헌을 확대하는 데이터 협업**
공공 및 민간 부문 조직의 집중적인 노력으로 'Data for Good' 운동이 강화됩니다.
- 05 데이터와 발맞추는 윤리 강령**
리더들이 GDPR과 같은 규정에 따라 윤리적 데이터 관행의 미래를 평가합니다.



- 06 **최신 BI 플랫폼으로 수렴되는 데이터 관리**
관리되는 데이터 큐레이션으로 데이터와 비즈니스 간의 격차가 해소됩니다.
- 07 **기업의 새로운 언어가 된 데이터 스토리텔링**
데이터 인사이트를 발견하고 전달하는 것은 이제 팀 스포츠입니다.
- 08 **더 스마트하게 분석을 채택하는 엔터프라이즈**
리더가 기술 채택보다 직원들의 참여에 더 주력하면 어떻게 될까요?
- 09 **데이터 과학자의 역량을 강화하는 데이터 민주주의**
데이터 과학자는 조직의 변화를 추진하는 소프트 스킬을 개발합니다.
- 10 **최신 BI 채택에 동력이 되는 가속화된 클라우드 데이터 마이그레이션**
데이터가 그 어느 때보다도 빠르게 클라우드로 이동하고 있으며, 이에 따라 조직은 데이터 전략을 재고하게 됩니다.

01

설명 가능한 AI의 부상

조직이 인공 지능 및 기계 학습 모델에 점점 더 의존하는 이때, 그 기술이 신뢰할 만한지 어떻게 보장할 수 있을까요?

기계가 의사 결정을 자동화함으로써 인간의 이해를 향상시키게 된다는 것이 인공 지능 (AI)에 대한 전망입니다. Tableau의 마켓 인텔리전스 책임자인 Josh Parenteau는 인공 지능과 기계 학습이 어떻게 “이전에는 발견하지 못했던 인사이트를 발견하도록 도와면서” 다른 관점으로 작용할지 설명했습니다. Gartner의 연구에 따르면, 2020년까지 “85%의 CIO가 구매, 구축 및 외주 인력의 방식을 결합하여 인공 지능 프로그램을 시험 운영하고 있을 것”이라고 전망합니다. 그런데 조직이 점점 더 기계 학습

모델에 의존하는 이때에, 인간은 인공 지능에 의한 권장 사항이 신뢰할 만한 것인지 어떻게 확신할 수 있을까요?

현재 대다수의 기계 학습 응용 프로그램에는 의사 결정 및 권장 사항 이면의 알고리즘 혹은 논리를 이해하는 '내막을 들여다보는' 방법이 없습니다. 따라서, AI 프로그램을 시험 운영하는 조직이 광범위한 채택에 대해 걱정하는 것은 무리가 아닙니다. 케임브리지 대학의 기계 학습 수석 연구원인 Adrian Weller가 요약한 대로, 기계 학습과

85%
의 CIO가 구매,
구축 및 외주
인력의 방식을
결합하여 인공지능
프로그램을 시험
운영하고 있을
것입니다.

같은 "지능형 시스템이 현실에 효율적으로 배포되는 데 있어 투명성은 종종 매우 중요한 요소로 간주"됩니다. 투명성이 필요한 이유는 여러 가지가 있지만, 모델이 설계된 대로 작동하는지 확인하거나, 사용자가 예측을 기반으로 자신 있게 의사 결정을 내릴 수 있도록 사용자와의 신뢰를 구축하기 위한 것이 주된 이유입니다.

투명성에 대한 요구는, 기계 학습 모델에 대한 투명성을 제공하여 그 이면에 대한 이해를 돕는 관행이 되는 설명 가능한 AI의 성장을 촉진하고 있습니다. 의사 결정권자는 리더가 중요한 결정을 할 때 인간 전문가에게 질문하는 방식과 매우 흡사하게, 기계 모델이 왜 그런 결론을 도출했는지, 이것이 얼마나 신뢰할 만한지, 그리고 입력이 다른 경우 결과가 어떻게 다를지에 대한 후속 질문을 할 수 있기를 기대합니다. Tableau의 AI 제품 관리자인 Richard Tibbetts는 "AI 및 기계 학습이 제공한 답을 설명할 수 없다면, 의사 결정권자는 여기에 회의적일 수밖에 없습니다. 분석과 AI는 인간의 전문 지식과 이해를 완전히 대체하는 것이 아니라 지원하는 도구가 되어야 합니다"라고 말합니다.

많은 조직, 특히 금융 서비스 및 제약 회사와

같은 위험 관련 업종의 사업 부문 리더는 데이터 과학 팀에게 좀 더 설명이 가능한 모델을 사용하고 모델 구축 방식에 대한 문서 또는 감사 추적 기능을 제공할 것을 요구하고 있습니다. 데이터 과학자는 이러한 모델에 대해 비즈니스 사용자에게 설명해야 하는 직무가 있으므로, 결론을 탐색하고 검증하기 위한 대화형 수단으로서 BI 플랫폼을 이용하고 있습니다.

결국, 기업들은 인공 지능 및 기계 학습의 가치를 수용할 수 있게 되었습니다. 그러나 조직에 혁신적인 영향을 주기 위해서는, AI의 신뢰성이 매우 중요합니다. AI는 제공하는 결론의 근거를 가능한 한 간결하고 이해할 수 있는 방식으로 보여줄 필요가 있으며, 후속 질문에 동적으로 응답할 수 있어야 합니다. 이를 통해, 인간은 자신의 데이터를 보다 잘 이해할 수 있게 됩니다.

“

**분석과 AI는 인간의 전문 지식과 이해를
완전히 대체하는 것이 아니라 지원하는 도구가
되어야 합니다.**

Richard Tibbetts, Tableau AI 부문 제품 관리자

02

데이터를 인간화하는 자연어

NLP 시스템의 발전으로 모든 사람이 데이터와 자연스러운 대화를 할 수 있습니다.

자연어 처리(NLP)는 컴퓨터 과학과 언어학을 결합하여 컴퓨터가 인간 언어 이면의 의미를 이해하도록 돕습니다. 오늘날, BI 공급업체는 사용자가 BI 도구에 대한 심층적 지식 없이도 질문을 하며 자연스럽게 데이터와 상호 작용할 수 있도록 비주얼리제이션에 대한 자연어 인터페이스를 제공하고 있습니다.

최신 BI라는 컨텍스트 내에서 자연어는 분석적 대화를 지원하기 위해 적용되고 있습니다. 분석적 대화란 인간이 시스템과 자신의 데이터에 대해 대화하는 것으로

정의됩니다. 이 시스템은 대화 내의 컨텍스트를 활용하여 질의에 담긴 사용자의 의도를 이해하고 대화를 발전시켜, 더욱 자연스러운 대화 환경을 만듭니다. 예를 들어, 어떤 사람이 자신의 데이터에 대해 후속 질문이 있을 때 더 자세히 파고 들거나 모호함을 없애기 위해 질문을 바꿔 말할 필요가 없습니다. BI 도구에 "캘리포니아 부근의 대규모 지진을 찾아보라"는 요청을 한 다음 두 번째에는 지진을 언급하지 않고 "텍사스 부근은 어떠한가?"라는 후속 질문을 할 수 있습니다. 기계 학습을 통해 시스템은

“

자연어는 모든 종류의 기술을 훨씬 광범위한 대상 사용자에게 도입하는 방법입니다. 자연어로 인해 기술의 장벽이 낮아지므로 사용자는 소프트웨어뿐 아니라 분석에 대해서조차 배울 필요가 없습니다. 비즈니스 컨텍스트만 알면 사용자는 적절한 질문을 할 수 있게 됩니다.

Stephanie Richardson, Tableau 제품 마케팅 담당 상임 이사

회사의 데이터 및 사용자가 하는 질문의 유형에 기반하여 시간이 지남에 따라 더 깊은 도메인 지식을 얻을 수 있습니다.

Tableau의 자연어 팀의 개발 관리자인 Vidya Setlur는 “분석적 대화의 핵심 특성 중 하나는 질문을 하고, 결과를 얻고, 원래 질문을 둘러싼 대화를 끊임없이 발전시키는 것입니다. 즉, 대화의 교착 상태를 피하는 것입니다”라고 설명합니다. “모든 사람들에게는 자신의 데이터에 대한 인사이트를 얻으려는 갈망이 있습니다. 그리고 자연어는 그 격차를 해소하는 중요한 방식 중 하나입니다.”

또한 자연어를 사용하면 사용자가 데이터 시각화를 기반으로 질문할 수 있게 됩니다. Tableau의 소프트웨어 엔지니어인 Ryan Atallah는 “BI 도구에서 질병 발발에 대한 질문을 하고, 그 결과로 어떤 비주얼리제이션을 얻는다고 가정해 봅시다. 그 비주얼리제이션을 보고 ‘그 주황색 스파이크는 무엇입니까?’라는 질문할 수 있습니다. 이것은 하나의 후속 질문이지만 자신의 데이터에 기반한 질문이 아닌, 비주얼리제이션의 인코딩에 기반한 것입니다”라고 말합니다. 만일 기존의

비주얼리제이션이 다음 질문의 맥락과 맞지 않다면 BI 도구가 대안을 제시합니다.

자연어는 사람들이 데이터에 대해 질문하는 방식에 대한 패러다임을 바꾸었습니다. 사람들이 마치 사람에게 하듯이 비주얼리제이션과 상호 작용할 수 있을 때, 이전에는 데이터 과학자와 고급 분석가들만 할 수 있었던 분석 파이프라인의 영역이 일반 사용자에게로 확대됩니다. 사용자의 역량이 더 이상 자신의 분석 기술에 의해 제한되기 보다는, 스스로가 가진 질문의 범위로 판별되는 것입니다. 또한 고급 사용자는 다양한 질문에 대한 답을 더 짧은 시간 내에 찾고, 다른 사람들에게 더 활용도 높은 대시보드 기능을 제공할 수 있게 됩니다. 자연어가 BI 산업에서 발전함에 따라, 조직 전반에서 분석 채택을 방해하던 장벽이 허물어지고 더 나아가 데이터가 직장 문화의 핵심에 자리잡게 됩니다.

자연어 생성 시장
규모는 2023년까지
8억 2530만 달러
규모로 성장할 것으로
예상됨 (Markets and Markets)

03

실행 가능한 분석으로 상황에 맞는 데이터 제공

BI 플랫폼은 사람들이 조치를 취하려는 상황에 적합한 데이터를 제공하도록 진화합니다.

데이터 작업자는 데이터 분석과 조치를 한 곳에서 수행할 수 있어야 합니다. 한 곳에서 분석을 수행하고 다른 곳에서 조치를 취하는 대신, 데이터로 작업하는 모든 사람이 비즈니스 프로세스 및 워크플로우의 컨텍스트 내에 머물 수 있어야 합니다. 비즈니스 인텔리전스 플랫폼은 모바일 분석, 내장된 분석 기능, 대시보드 확장 프로그램(추가 기능이라고도 함) 및 API와 같은 기능을 통해 핵심 비즈니스 운영, 워크플로우 및 프로세스와 병합함으로써 이러한 필요에 부응하고 있습니다. 그 결과, 실행 가능한

분석은 기술 및 비기술적 역할 모두에서 의사 결정 프로세스를 신속하게 처리합니다.

이러한 기능을 통해 데이터 작업자는 데이터를 분석하고 인사이트를 찾은 후에 조치를 취하는 것까지 모두 한 장소에서 처리할 수 있게 됩니다. 인사이트와 조치가 통합된 사례의 하나가 바로 내장된 분석 기능입니다. 내장된 분석 기능은 데이터와 인사이트를 이미 사용자가 작업하는 곳에 제공하여 또 다른 응용 프로그램이나 공유 서버를 찾아갈 필요를 없애줍니다. 분석 기능은 SharePoint와 같은

“

분석이 실행 가능하게 되려면 적절한 사람에게 적절한 메시지를 그들이 이해할 수 있는 방식으로 적시에 제공하고 있는지 확인해야 합니다.

Peter Benson, Automated Insights 전략적 제휴 책임자

사내 포털 또는 일반적으로 사용되는 응용 프로그램에 내장될 수 있습니다.

예를 들어, 조직은 분석을 Salesforce 등의 고객 관계 관리(CRM) 소프트웨어에 내장할 수 있습니다. 영업 담당자는 제품 선호 사항 또는 시간 경과에 따른 지출 등 상황에 맞는 중요한 고객 데이터를 볼 수 있으므로, 여기에 기반해 해당 고객과 더 원활한 대화를 할 수 있거나 다음 단계를 모색할 수 있습니다.

대시보드 확장 프로그램은 이러한 통합에 다른 측면으로 접근합니다. 분석 플랫폼에서 대부분의 시간을 보내는 사람들은 확장 프로그램을 통해 다른 시스템에 대한 액세스를 직접 대시보드로 가져올 수 있습니다. 이제 개인은 분석 워크플로우를 벗어나지 않고도 필요한 작업을 수행할 수 있게 됩니다.

예를 들어, 분석에 의존하여 티켓 대기열을 모니터링하는 IT 관리자는 티켓팅 시스템으로 전환하지 않고 대시보드 확장 프로그램을 사용하여 사례 정보를 편집하거나 대시보드 내에서 직접 작업을 수행할 수 있습니다. 이로써 자신의 작업 흐름에서 벗어나지 않고, 이미 작업을 하고 있던 곳에서 더 신속하게

조치를 취할 수 있습니다.

내장된 분석 기능과 대시보드 확장 프로그램이 작업과 인사이트를 플랫폼 및 도구 내로 통합한다면, 모바일 분석은 이러한 기능들을 사용자가 실제로 어디에 있든지 상관없이 활용할 수 있게 해줍니다. 모바일 분석을 통해 컨설턴트는 현장에서 고객 데이터를 활용할 수 있으며, 정비사는 사물 인터넷(IoT) 데이터를 현장에서 활용하여 기기를 수리할 수 있습니다.

상황에 맞게 제공되는 분석 환경은 특정 비즈니스 또는 산업 분야에 맞게 분석을 사용자 정의할 수 있으므로 그 효과가 더욱 큼니다. 분석과 조치의 통합으로 인사이트에서 의사 결정에 이르는 시간과 노력이 단축될 것입니다. 또한 비즈니스 워크플로우 내에서 사용할 수 있는 데이터의 범위가 더욱 확장되므로 더 많은 사람이 매일의 의사 결정에 데이터를 활용할 수 있게 됩니다.

2022년까지 50%
의 디지털 비즈니스
기술 플랫폼
프로젝트에서
이벤트와 비즈니스
성과가 연계될
것입니다. (Gartner)

04

사회 공헌을 확대하는 데이터 협업

공공 및 민간 부문 조직의 집중적인 노력으로 'Data for Good' 운동이 강화됩니다.

데이터는 비정부기구(NGO) 및 비영리 단체를 포함하여 조직의 운영 방식을 변화시켰습니다. 그리고 이제 "Data for Good" 운동이 민간 및 공공 부문 조직에서 폭발적으로 증가하고 있습니다. Gartner의 연구에 따르면 일반 대중이 어떻게 데이터가 사회에 긍정적인 영향을 미칠 수 있는지 더 많이 알게 됨에 따라 "소셜 미디어에서 'Data for Good'에 대한 언급이 작년에 68% 증가했다"고 합니다.

다국적 통신 회사인 Orange와 같은 사기업은 데이터 기반 인사이트를 사용하여 사회 복지 향상을 위한 프로젝트를 확립하고 있습니다. Orange는 자사의 OPAL 프로젝트를 통해 지방 자치 단체와 협력하여 데이터 수집, 익명화 및 보호에 관한 방식을 규제하는 거버넌스 위원회를 구성했습니다. 이를 통해 세부 통화 기록을 사회 공헌 조직과 안전하고 보안화된 방식으로 공유할 수 있습니다. 세네갈에서는 최근에 이 데이터가 문자 메시지 사용을

기반으로 문맹률을 평가하는 데 사용되어 사회 공헌 단체가 문맹 퇴치 프로그램에 대한 자원 배분 결정을 내리는 데 도움을 주었습니다.

한때, NGO 및 비영리 단체에는 정교한 데이터 인프라 또는 대규모 데이터 작업자 팀에 투자할 자원이 충분하지 않았습니다. 이제는 클라우드 컴퓨팅의 비용 효율성과 유연성으로 인해, 이러한 조직에서 대규모 온프레미스 투자 없이 정교한 데이터 환경을 개발할 수 있게 되어, 더 많은 데이터 기반의 사회 공헌 계획으로의 길을 마련할 수 있습니다.

실례로 공동의 목표를 달성하기 위해 조직 간 공유와 협력을 위한 플랫폼인 데이터 연방 (Data Commonwealth)의 출현과 성장을 들 수 있습니다. 예를 들어, Hutch Data Commonwealth는 "혁신적인 데이터 과학 도구, 인프라 기능 및 협업을 통해 Fred Hutchinson 암 연구 센터 연구원들의 보다 신속한 연구를 지원하는 것을 사명으로 하는 여러 분야가 융합된 팀"입니다. GeekWire에 실린 글에서 Hutch Data Commonwealth의 최고 정보 책임자이자 전무 이사인 Matthew Trunnell은 "클라우드를 우리가 회합하고 협업하며, 모두가

공동의 선을 위해 모이는 장소가 됩니다"라고 이야기하며, 클라우드가 어떻게 '과학 데이터의 연방의 기반' 역할을 하는지 설명합니다. Hutch Data Commonwealth는 그 사명의 중심에 데이터를 두고, 타 연구 기관 및 기술 제공 업체와의 파트너십을 기반으로 합니다.

공공 및 민간 부문 프로젝트 또는 연방을 통한 이러한 파트너십에는 신뢰할 수 있는 기반이 필요합니다. 조직은 데이터 공유와 관련된 법적 관련 사항 및 거버넌스 표준을 포함해 성공적인 파트너십의 중요 요소를 평가하고 있습니다.

“

조직은 데이터 연방을 통해 조직 및 전세계의 사용자와 안전하고 보안이 유지될 뿐 아니라 데이터가 수집되는 개인의 정보도 보호되는 방식으로 데이터를 공유할 수 있습니다.

Neil Myrick, Tableau 재단 글로벌 책임자

소셜 미디어에서 'Data for Good'에 대한 언급이 작년 68% 증가했습니다.

여기에는 개인 정보 보호 위험의 평가와 개인 식별 정보의 공유와 관련한 보호 의식을 고취하는 것도 포함됩니다.

NYU Tanden 공과 대학의 거버넌스 연구소 (Governance Lab)의 최근 보고서에서는 민간 및 공공 부문 조직¹ 간에 소셜 미디어 데이터의 공유를 둘러싼 문제점에 초점을 맞추었지만, 사실 이 원칙은 다양한 데이터 공유 파트너십에 적용할 수 있습니다. GovLab은 "데이터 요청에 응답하는 정당한 프로세스, 특정 유형의 정보를 필터링하거나 우선순위를 매기는 시스템, 공개되는 데이터가 대중의 필요와 요구에 부합하는지 확인하는 방법"을 보장한다는 개념 하에 데이터 협업을 추진하기 위해 더 많은 조직이

데이터 관리자를 임명할 것이라고 예측합니다.

Fred Hutch의 사례처럼, 적절한 제어 하에 방대한 양의 다양한 데이터 원본에 액세스할 수 있다면 혁신적인 효과를 기대할 수 있습니다. 이러한 대규모의 공동 프로젝트에 여전히 도전 과제는 남아있지만, 'Data for Good' 운동은 데이터 공유의 이타적인 가능성에 대한 증거입니다. 기술의 발전, 개선된 데이터 해석 능력, 그리고 협업에 초점을 맞추므로써, 세계가 직면한 가장 어려운 위기를 해결하는 데 적절한 환경을 조성합니다.

¹ <http://datacollaboratives.org/static/files/social-media-data.pdf>

05

데이터와 발맞추는 윤리 강령

리더들이 GDPR과 같은 규정에 따라 윤리적 데이터 관행의 미래를 평가합니다.

데이터 개인 정보 보호라는 주제가 두드러지게 대두되며 소비자는 개인 데이터 공유에 대해 그 어느 때보다 관심을 기울이고 있습니다. 이는 비즈니스가 데이터 수익화, 데이터 수집 및 데이터 공유를 다루는 방식에 영향을 줍니다. GDPR 같은 새로운 규정의 등장으로, 회사는 일상적인 비즈니스 관행의 맥락에서 데이터 윤리 및 개인 정보 보호에 관한 중대한 논의를 진행하고 있습니다. 대표적인 사안은 다음과 같습니다.

윤리 강령: 법률, 의학 및 회계 분야의 직책을

비롯하여 많은 직업군이 현재 직업 윤리 강령을 준수하고 있습니다. 모든 비즈니스 영역에서 데이터가 지속적으로 확산됨에 따라, 회사는 이러한 윤리 원칙을 데이터 분석 업무에 적용할 수 있도록 고려하고 있습니다. Gartner가 말한 바와 같이, “디지털 비즈니스 시대는 기술과 비즈니스의 경계를 모호하게 합니다.” 이제 데이터는 전략적 퍼즐에 핵심적인 요소입니다. 더 많은 회사가 모든 부서와 역할 내에서 비즈니스 의사 결정을 내리는 데 데이터에 의존하고 있습니다. 이는 데이터가 사용되고 공유되는 방법에 관여된

윤리가 직무의
일부라고 대답한
CDO의 수는
10
% 포인트 증가
(Gartner)

사람이 점점 늘어난다는 뜻입니다.

이에 따라 리더, 특히 최고 데이터 책임자(CDO)는 디지털 전환 노력의 일환으로 전사적 데이터 관행에 대한 내부 지침을 마련하는 책임을 맡고 있습니다. 실제로, 2017년 Gartner의 최고 데이터 책임자(CDO) 설문 조사에서 “윤리가 직무의 일부라고 대답한 CDO의 수는 2016년에서 2017년 사이에 10% 포인트 증가”했음이 밝혀졌습니다. 이러한 윤리 강령은 미래의 인프라, 거버넌스 및 인사 결정에 대한 프레임워크로 작용할 것입니다.

비즈니스 프로세스에서의 변화: 회사는 데이터 수집에서 분석에 걸친 데이터의 전체 수명 주기에 대해 비판적으로 생각하고 있습니다. 이는 회사의 데이터 관리 전략을 전체적으로 평가하여 규정과 내부 윤리 강령의 준수 여부를 확인할 기회를 마련합니다. 이 검토 프로세스는 일회성으로 끝나는 사안이 아닙니다. Accenture사의 데이터 윤리의 보편적 원칙(Universal principles of data ethics) 보고서에서 언급한 바와 같이, “거버넌스 관행은 강력해야 하고, 모든 팀원이 이해하고 정기적으로 검토되어야 하며,” 회사의 성장과 변화에 대응할 수 있어야 합니다.

데이터 윤리는 데이터 수집이나 데이터 거버넌스에만 제한되지 않습니다. 데이터가 해석되고 활용되는 방법에도 적용됩니다. 최신 BI 플랫폼을 통해 더 많은 사용자가 데이터 분석을 사용할 수 있게 되었으며, 이에 따라 더 많은 직무에서 데이터 윤리 원칙을 준수할 책임이 발생하게 됩니다. Teknion Data Solutions의 수석 컨설턴트인 Bridget Winds Cogley는 데이터를 분석하거나 인사이트를 전달하는 사람이라면 누구나 “편차 및 사실이 분명하게 표현되는지 여부를 고려”하고 “데이터의 한계가 이해되고 질문에 맞는지” 여부를 고려해야 한다고 제안합니다. 점점 더 많은 사람이 데이터 작업자가 됨에 따라, 데이터 윤리는 데이터 해석 노력의 핵심 요소가 되어, 사람들이 사적 및 업무적 맥락에서 데이터에 접근하는 방식에 영향을 주게 될 것입니다.

“

윤리 관행은 실무자로 하여금 한 걸음 뒤로 물러나 윤리적 관점에서 상황을 바라볼 수 있게 해줍니다. 무엇보다도 데이터 윤리는 우리의 업무에서 과속 방지턱과 같은 역할을 하도록 고안되었습니다. 이를 통해 개인적으로나 직업적으로 맞닥뜨리게 되는 딜레마에 어떻게 대처해야 하는지 이해할 수 있습니다.

Bridget Winds Cogley, Teknion Data Solutions 수석 컨설턴트

06

최신 BI 플랫폼으로 수렴되는 데이터 관리

관리되는 데이터 큐레이션으로 데이터와 비즈니스 간의 격차가 해소됩니다.

데이터 원본이 더 복잡해지고, 다양해지고, 수가 늘어남에 따라, 데이터 관리는 이제 최신 BI 배포에서 한층 더 필수적인 요소가 되었습니다. 더 많은 인력이 의사 결정을 추진하는 데 데이터를 사용함에 따라, 조직은 데이터의 정확성을 보장하고 데이터가 분석에 사용될 수 있도록 해야 합니다.

조직은 이렇게 광범위한 데이터 액세스에 따른 데이터 관리 및 거버넌스 문제를

해결하기 위해 데이터 큐레이션을 수행해 왔습니다. 데이터 큐레이션은 조직이 상이한 데이터를 수집, 정리, 정의 및 조정하는 방식을 모두 아우릅니다. 이 과정을 통해 데이터와 실제 응용 간의 격차를 메울 수 있습니다.

조직은 팀과 조직 전체에 존재하는 모호성을 제거하기 위해, 데이터 분석을 지원하는 분석 도구와 데이터 정의를 통합하는 기술에 이미 수백만 달러를 지출하고 있습니다. 이에

“

데이터 큐레이션이란 비즈니스 사용자가
데이터와 상호 작용하고 이해하여 분석
결과를 생성하는 데 데이터를 사용할 수
있도록 필요한 데이터 원본을 식별하고, 그
데이터를 비즈니스 컨텍스트와 연계하는
프로세스입니다.

Mike Hetrick, Tableau 제품 마케팅 담당 선임 관리자

따라, 데이터 큐레이션 도구 및 프로세스(데이터 카탈로그 및 의미론적 거버넌스 등)는 데이터를 비즈니스 컨텍스트에 연결하는 BI 플랫폼과 융합되고 있습니다.

데이터 카탈로그는 데이터 원본 및 일반 데이터 정의의 엔터프라이즈 비즈니스 용어집으로 활용됩니다. 데이터 엔지니어 및 데이터 관리자와 같은 주제 전문가는 데이터 원본 및 필드에 설명과 정의를 추가하고, 더 원활한 검색을 위해 태그를 지정할 수 있으며, 유용한 데이터 품질 지표까지도 추가할 수 있습니다. 여기에는 신뢰할 수 있는 콘텐츠의 인증 또는 데이터 자산의 유지 관리 또는 사용 중단에 대한 알림 등이 포함됩니다.

일반 사용자는 데이터가 원본의 어디에 위치하는지 알 필요는 없지만, 그 데이터가 실제로 무엇을 의미하는지는 반드시 이해할 필요가 있습니다. 예를 들어, 콘텐츠 분석가 및 이용자는 종종 데이터 출처를 검증(계통 분석이라고도 함)해야 합니다. 또한 데이터 집합이 변경될 경우, 데이터 엔지니어 및 데이터 관리자는 자신들이 관리하는 테이블 또는 스키마에 연결되어 있는 자산에 미치는 다운스트림 영향을 분석해야 합니다. 데이터 카탈로그와 BI 플랫폼을 결합하면 이러한 모든 작업이 간소화되고, 가장 자주 사용되는 데이터

원본과 대시보드를 신속하게 식별해 주는 사용 메트릭을 얻을 수 있습니다.

데이터 카탈로그가 필수적인 만큼, 의미론적 거버넌스 영역에도 메타데이터 거버넌스를 뛰어넘는 더 큰 가치가 있습니다. 의미 체계는 데이터의 컨텍스트뿐만 아니라 '주문 규모'와 '수량' 같은 명령들을 이어주는 동의어 매핑과 같이 분석 작업의 의도에 맞게 연결하는 데 도움이 됩니다. 이를 통해, 모든 범위의 데이터 작업자가 데이터와 상호 작용하여 신속하게 새로운 인사이트에 도달할 수 있는 새로운 방법이 제공됩니다. 그 중 한 가지 방법은 자연어와의 상호 작용을 사용한 것으로, 이를 통해 BI 플랫폼이 '최고, 최저, 평균 하이라이트'와 같은 다중 쿼리가 포함된 계층을 이해하게 됩니다.

이러한 기술과 프로세스가 계속해서 통합되면서, 데이터 큐레이션 및 의미 체계는 더욱 강력한 분석 환경 기반을 제공할 것입니다. 이에 따라, 정리 및 다운스트림 분석과 같은 보다 실질적인 데이터 환경 구성 요소를 통합하고, 테이블, 조인 및 데이터 모델에 대한 더 강력한 시스템 생성 권장 사항을 제공할 수 있습니다. 궁극적으로 데이터 큐레이션의 발전에 따라 직원들의 역량이 강화되어 직원들은 분석 중에 자신의 데이터에 대한 질문을 넘어 자신의 비즈니스에 대해 질문할 수 있게 됩니다.

디지털 데이터는
2020년까지 연평균
성장률(CAGR)
42%
로 성장할 것입니다(IDG)

07

기업의 새로운 언어가 된 데이터 스토리텔링

데이터 인사이트를 발견하고 전달하는 것은 이제 팀 스포츠입니다.

아무리 많이 자동화되고, 큰 데이터 집합과, 뛰어난 계산을 사용하더라도, 발견한 결과를 다른 사람에게 전달하지 못한다면 그 분석은 무용지물입니다. 여기에 데이터 시각화의 힘을 사용할 수 있습니다. 데이터 시각화는 하나의 언어이며, 분석가들이 정보를 의사 결정권자에게 실행 가능하고 이해하기 쉬운 방식으로 전달하는 방식의 표준이 되고 있습니다. 이 기술은 분석가가 데이터에서 인사이트를 발견하기 위해 수행한 단계를

공유하는 능력과 결합해 종종 '데이터 스토리텔링'이라고 정의됩니다.

데이터 스토리텔링은 분석 프로세스에서 중요한 요소입니다. 분석이 그 어느 때보다 중요해지고 있는 직장 환경에서, 데이터 스토리텔링은 재정의되고 있습니다. 조직이 분석 문화를 조성함에 따라, 분석가의 데이터 스토리텔링 방식은 하나의 결론을 이끌어 내기 위한 논증 방식에서 데이터를

“

청중인 우리는 정보를 얻고자 하는 열망과 정보를 해석할 수 있는 능력을 갖추어야 하며, 어느 정도의 전문 지식을 가지고 있어야 합니다. 우리가 이러한 자질을 갖추지 않았다면, 설계자가 이 모든 것을 이해할 수 있는 방식으로 청중에게 제공할 책임을 지니게 됩니다.

Andy Kirk, VisualisingData.com 설립자

Dresner 2018년도
시장 조사에 따르면
응답자의
75%
가 데이터
스토리텔링이
비즈니스
인텔리전스
전략 추진에
결정적이거나
반드시 필요하다고
답했습니다.

둘러싼 대화를 촉진하는 방식으로 변화하게 됩니다. 이러한 분석 문화는 또한 사람들에게 데이터를 진정으로 이해하도록 가르치고, 발견의 순간부터 비즈니스 의사 결정에 이르는 분석 대화에 참여하는 것을 목표로 하는 데이터 해석 노력을 강화합니다.

데이터 시각화 전문가이자 VisualisingData.com의 설립자인 Andy Kirk는 데이터 시각화의 7가지 직분을 이야기합니다. 그 중 하나는 바로 '기본적으로 어떤 프로젝트에 연관된 모든 인간관계에 관심이 있는 사람 (위원, 주주, 청중)'인 전달자입니다. Andy는 어떻게 "최소한 커뮤니케이션적 측면에서는, 모든 비주얼리제이션 작업이 청중 중심이 되어야 하는지"를 설명합니다. 데이터 작업자는 비주얼리제이션으로부터 결론을 도출하는 청중의 프로세스를 이해할 필요가 있습니다. 동시에, 청중은 데이터를 해석하는데 필요한 전문 지식을 갖추고 있어야 하며, '정보를 얻고자 하는 열망이 있어야' 합니다.

이러한 데이터 스토리텔링의 전환은 또한 데이터 시각화 동향을 드러내기도 합니다. 스크롤이나 여러 페이지로 되어 있는 대시보드 등을 이용하는 긴 양식의 스토리텔링 서식이 더 보편화되면서 분석가는 결론에 이르는 단계별 접근 방식을

보여줄 수 있습니다. 이러한 방법을 통해 분석가는 분석의 진행 상황을 보여 주며 데이터에서 발견된 인사이트와 그 결과로 얻은 가설을 강조할 수 있습니다. 그 다음 단계는 이러한 인사이트에 관해 열린 대화를 마련하는 것입니다. 이로써 비즈니스 의사 결정을 내리기 전에 다양한 직책 또는 부서의 사람들이 추가적으로 비즈니스 컨텍스트를 제공할 수 있는 여지가 생기므로 보다 다양한 관점을 수용할 수 있게 됩니다.

더 많은 조직이 분석적 협업에 중점을 둔 작업 스트림 및 팀을 구성하므로, 데이터 스토리텔링은 직장 문화에 계속해서 스며들 것입니다. 이 접근 방식은 조직이 사용자를 참여시키고, 정보를 전달하고, 아이디어를 검증하는 데 데이터를 활용하는 방식을 형성해줍니다. 또한 더 많은 사람이 데이터를 해석하고 자신의 분석 프로세스를 설명할 수 있으므로, 비즈니스에 미치는 영향력이 향상됩니다.

08

더 스마트하게 분석을 채택하는 엔터프라이즈

리더가 기술 채택보다 직원들의 참여에 더 주력하면 어떻게 될까요?

비즈니스 인텔리전스 채택 프로젝트에는 많은 경우 명확한 개시일과 종료일이 지정되어 있으며, 사용자에게 배포된 후에는 “완료됨”으로 여겨지는 것이 일반적입니다. 그러나 단순히 비즈니스 인텔리전스 솔루션에 대한 액세스를 제공했다고 채택이 끝난 것은 아닙니다. 최고 데이터 책임자는 주로 채택한 BI의 역할을 현대화를 향한 전략적 변환 측면에서 재평가합니다. 진정한 가치는 배포한 솔루션으로 측정되는 것이 아니라, 회사 인력이 솔루션을 어떻게 활용하여 비즈니스에 영향을 미치는가로 측정되기 때문입니다.

단순히 모든 사람이 BI 플랫폼에 액세스할 수 있으므로 모두가 BI 플랫폼을 통해 가치를 얻고 있다는 가정은, 분석을 통해 실질적인 진보를 가져오는 데 방해가 될 수 있습니다. Tableau 마켓 인텔리전스 책임자인 Josh Parenteau 는 라이선스 수를 기준으로 투자 수익률을 계산하면 “학습, 성장과 더 많은 성공을 거둘 가능성 모두를 배제할 수 있습니다.”²라고 말합니다. BI 플랫폼의 채택을 단순히 액세스 제공의 관점에서 생각하던 태도에서 벗어나, 이제 리더는 BI 플랫폼을 통한 데이터와 분석이 어떻게 조직 전체의 의사 결정 방식을 변화시키고 있는지에 초점을 맞추고 있습니다. 예를 들어, 리더는 'BI 플랫폼이 없어진다면 직원들이 일상적인 업무 의사 결정을 내리는 데 어떤 문제가 생길까'와 같은 질문을 하게 됩니다.

휴대폰에 응용 프로그램을 다운로드했다고 해서 내가 그것을 꼭 사용하는 것은 아니듯, 한 달에 한 번 보고서를 열어 본다든가 여기서 어떤 조치나 영향이 발생하지는 않습니다. 리더는 사내 커뮤니티와 사용자 그룹 등 참여를 촉진하는 프로그램을 평가하고 있습니다. 이전에는 일반 사용자를 위한 프로그램으로 간주되던 이러한 노력은 이제 조직의 BI 전략의 기본적 요소로서 사용자가 더 빨리 역량을 강화하고,

셀프 서비스를 촉진하며, 질문에 신속하게 답을 얻을 수 있도록 지원할 것입니다. 리더가 투자를 증대하고 커뮤니티의 확장을 유도하면 채택은 결과적으로 이루어지게 됩니다.

IT가 주도하는 JPMorgan Chase(JPMC)의 전문 센터(COE) 팀은 이 모델을 활용하여 수천 명에 이르는 분석가의 참여를 독려하고, 사용자 커뮤니티를 육성했습니다. 전문 센터(COE) 팀은 전일 세션을 개최하여, 데이터 시각화 및 거버넌스 모범 사례를 공유하였습니다. 참고로 부사장이자 BI 혁신 리더인 Steven Hittle은 이 세션을 '데이터 요법 세션'이라고 부릅니다. 이러한 세션은 JPMC에서 많은 직무 및 부서 간의 참여와 대화를 촉발하기 위해 구성된 다양한 활동 중 하나로, 최신 BI 플랫폼을 30,000여 명에 이르는 사용자에게 성공적으로 확장시키는 역할을 했습니다.

이러한 사내 커뮤니티를 통해 더 많은 작업자가 BI 플랫폼을 사용할 수 있게 되므로, 조직은 분석 업무를 위임하고 그 결과 새로운 사용자 챔피언이 양산됩니다. 이는 궁극적으로 기존에는 IT 전담이었던 유지 관리 및 보고의 막대한 업무를 덜어줄 수 있습니다. 모범 사례를 공유하고, 데이터 정의에 맞춰 사람들을

조율하는 주제 전문가로서 행동하는 더 많은 사내 챔피언이 등장하기 시작할 것입니다. 이러한 모든 움직임은 결국 더 많은 사람이 BI 소프트웨어를 사용하고 여기에서 가치를 얻을 수 있도록 합니다. 그리고 가장 중요한 것은, 회사 직원의 효율성은 향상되고 조직의 경쟁력은 더욱 강화된다는 점입니다.

60%
의 CIO가 다음 12
개월 동안 분석
지출을 늘릴
계획입니다.

(IDG CIO 기술 여론 조사: 2018년
기술 우선 순위)

² <https://www.tableau.com/about/blog/2017/10/three-reasons-your-business-intelligence-adoption-has-stalled-77448>

“

**BI의 이점을 실제로 측정하는 방법을 재고해야
합니다. 단순히 액세스 권한을 누구에게
부여하느냐의 문제가 아니라 실제로 분석을
사용해 얻은 정보가 어떻게 의사 결정
프로세스에 도움이 되는지에 대해 고민해야
합니다. 이것이 바로 진정한 채택입니다.**

Josh Parenteau, Tableau 마켓 인텔리전스 책임자

09

데이터 과학자의 역량을 강화하는 데이터 민주주의

데이터 과학자는 조직의 변화를 추진하는 소프트 스킬을 개발합니다.

데이터 과학자에 대한 수요가 급증하고 있습니다. 2017년 미국 신생 직업 보고서에서, LinkedIn은 “데이터 과학자 역할은 2012년 이래로 650퍼센트 이상 성장”하였으며 업종을 막론하고 “수백 개의 회사에서 해당 직책을 고용하고 있다.”라고 인용했습니다. “기계 학습 엔지니어, 데이터 과학자 및 빅 데이터 엔지니어가 신생 직업의 상위를 차지하면서” 데이터 관련업 지원자 풀은 점점 더 심화되는 양상입니다.

그러나, 데이터를 사용하는 부서와 직책이

점점 더 증가함에 따라, 데이터 해석 능력의 전반적인 향상과 함께 많은 시민 데이터 과학자의 출현도 관찰되고 있습니다. Gartner가 정의한 시민 데이터 과학자란 “고급 진단 분석 또는 예측 및 규범적 기능을 사용하는 모델을 만들거나 생성할 수 있으나, 주 업무는 통계 및 분석 분야가 아닌 사람”입니다. 이들은 데이터 과학자를 대체하지는 않지만, 가설을 세우고 검증하는 데 핵심적인 파트너가 되고 있습니다.

이는 데이터 과학의 정의를 변화시키고,

“

통계 모델링과 기계 학습은 이제 데이터 과학자가 되기 위한 최소한의 필요 요건에 불과합니다. 이런 환경에서 차별화 요소는 결국 그 분야에 종사하는 사람들이 얼마나 간단하고 실행 가능한 방식으로 자신이 발견한 사항을 전달하느냐입니다.

Sonic Prabhudesai, Charles Schwab 통계 분석 관리자

데이터 과학자
역할이 2012년
이후
650%
이상 증가함(LinkedIn)

기존의 데이터 전문가와 비즈니스 영역 지식 간의 경계를 희미하게 합니다. Charles Schwab의 통계 분석 관리자인 Sonic Prabhudesai는 어떻게 “데이터 과학자가 비즈니스 내부 작업에 좀더 익숙해지는 반면에 다수의 비즈니스 작업자는 데이터로 작업하는 방법을 이해하는지” 설명합니다.

오늘날, 데이터 과학자에게는 고급 통계 및 기계 학습 지식뿐 아니라, 업종에 대한 깊은 지식을 포함한 비즈니스에 대한 전략적 마인드도 요구되고 있습니다. Infosys의 최고 글로벌 데이터 과학자인 N. R. Srinivasa Raghavan 박사는 “데이터 과학은 수치 분석을 넘어, 업종의 특정 문제를 해결하기 위해 다양한 기술을 적용하는 것입니다”라고 설명하면서 “데이터 과학자에게는 자신의 인사이트가 적용될 영역에 대한 철저한 이해가 필요합니다”라고 덧붙였습니다.

알고리즘 및 모델에 의해 생성된 결과물은 적절한 맥락에서 적절한 문제를 해결할 수 있을 때에만 효과적입니다. 이는 프로세스의 시작부터 문제 기술 및 가정을 식별하고 조율하기 위해 이해 관계자와 협력하여 작업하고, 전체 워크플로우에 걸쳐 지속적으로 참여해야 함을 뜻합니다.

또한, 워크플로우³ 종료 시, 관련되고 실행 가능한 방식으로 비즈니스 파트너에게 결과를 전달하는 방법을 알고 있어야 함을 의미합니다.

Sonic은 “통계 모델링 및 기계 학습은 이제 데이터 과학자가 되기 위한 최소한의 필요 요건에 불과합니다”라는 의견을 밝히면서 “이런 환경에서 차별화 요소는 결국 그 분야에 종사하는 사람들이 얼마나 간단하고 실행 가능한 방식으로 자신이 발견한 사항을 전달하느냐입니다”라고 덧붙였습니다. 데이터 과학자는 단순히 결과만을 제공하는 것이 아니라, 그 결과를 어떻게 비즈니스에 적용하는가에 있어 핵심적 역할을 하게 됩니다.

셀프 서비스 분석 도구를 통해, 데이터 과학자와 고급 사용자는 모두 자신의 데이터를 탐색하고 그에 대한 더 나은 이해를 얻을 수 있습니다. 이로써 분석의 방향을 결정할 수 있는 인사이트를 얻게 되고, 궁극적으로 이는 비즈니스에 영향을 주게 됩니다.

³ <https://medium.com/@sonicmsba/how-to-build-an-effective-business-context-for-data-analytical-problems-cb02906341cd>

10

최신 BI 채택에 동력이 되는, 가속화된 클라우드 데이터 마이그레이션

데이터가 그 어느 때보다도 빠르게 클라우드로 이동하고 있으며, 이에 따라 조직은 데이터 전략을 재고하게 됩니다.

데이터 전략의 현대화는 종종 데이터 저장 위치를 재고해야 함을 뜻합니다. 보다 많은 회사가, 낮은 총 소유 비용으로 유연성 및 확장성 증대와 같이 데이터를 클라우드로 이전하는 데서 오는 혜택을 누리고 있습니다. 실제 Gartner 조사에 따르면 “퍼블릭 클라우드 서비스 시장은 2018년에 21.4% 성장하여 총 1,864억 달러 규모에 이를 것으로 전망됩니다.”⁴ 클라우드는

회사가 다양한 유형의 데이터를 수집하고 통합하는 것을 용이하게 해줍니다. 이것은 모든 데이터의 저장 환경이 고도로 구조화된 온프레미스 웨어하우스에서, 완전한 클라우드나 하이브리드 솔루션과 같은 확장 가능하고 더욱 유연한 인프라로 옮겨 간다는 의미입니다.

여기서 우리는 데이터 중력, 즉 데이터가 있는

방향으로 서비스와 응용 프로그램이 끌려간다는 5 개념을 발견합니다⁴. 더 많은 조직에서 워크로드의 클라우드 이전을 가속화함에 따라, 이러한 데이터 중력 현상은 분석 프로세스 역시 클라우드로 끌어당기고 있습니다. Google Cloud 제품 관리 책임자인 Sudhir Hasbe는 "여러 회사들이 Google Cloud로 이동함에 따라, 리더가 전체 데이터 분석 전략을 재고하고, 클라우드를 통해 어떻게 비즈니스 및 수익을 향상시킬 수 있을지 생각하는 것을 돕니다."라고 설명합니다.

이러한 중력으로 인한 전환의 이면에 있는 핵심 요소는 지연 시간(작업을 수행하는 데 필요한 시간의 양)과 처리량(지정된 시간 단위당 작업을 수행하거나 결과를 얻을 수 있는 횟수)입니다. 데이터, 응용 프로그램 및 서비스가 긴밀히 연결되어 있을 때 지연 시간과 처리량은 감소하고, 결과적으로 효율성은 증대됩니다. 데이터가 클라우드에 존재하면, 자연스럽게 응용 프로그램 및 서비스도 클라우드로 따라가기 시작합니다.

조직은 더 광범위한 데이터 전략을 평가하므로, 기존 BI에서 최신 BI로 이전함에 따른 분석 모델 역시 재고하게 됩니다. McKinsey에 따르면,

클라우드의 가치는 기업이 클라우드 인프라 및 시스템에 "일회성 전술적 의사 결정이 아니라 디지털 변환을 추구하는 전체론적 전략의 일환으로" 접근할 때 나타난다고 합니다.⁶

기존 비즈니스 인텔리전스는 질문에 대한 답을 얻는 데 IT 부서에만 의존하였습니다. 그 결과, 인력 부족으로 인한 병목 현상과, 비즈니스 컨텍스트가 연계되지 않은 분석이 초래되었습니다. 마찬가지로 기존 BI 배포는 경직된 온프레미스 모델을 기반으로 구축되어, 이러한 기업 보고 형식을 지원해왔습니다.

반대로 클라우드 분석은 새로운 배포 모델을 고려할 수 있는 기회 및 다양한 이점을 제공하므로, 리더는 클라우드 분석을 통해 이러한 기회를 적극적으로 활용하고자 합니다.

여기에는 현장의 근무 직원에게 모바일 대시보드를 제공하여 방화벽을 제거하지 않고도 데이터에 액세스할 수 있도록 하는 것이 포함합니다. 클라우드는 또한 내부 프로세스를 뛰어넘는 신뢰할 만한 단일 원본을 작성하여, 파트너 또는 고객과의 안전한 대시보드 공유도 보장합니다.

모든 회사에서 모든 데이터를 클라우드로 이전할 준비가 된 것은 아니지만 많은 회사들이 다양한 데이터 원본을 활용하는 하이브리드 솔루션을 실험하고 있습니다. 결과적으로 회사는 향후 완전한 클라우드 분석으로의 전환을 지원할 수 있는지 여부를 전제로 최신 BI 플랫폼을 평가하고 있습니다.

퍼블릭 클라우드
서비스 시장은 2018
년에 21.4% 성장하여
총 규모가
1864억 달러
에 이를 것으로 예상됨
(Gartner)

⁴ <https://www.gartner.com/newsroom/id/3871416>

⁵ <https://www.techrepublic.com/article/how-data-gravity-both-hurts-and-helps-cloud-adoption/>

⁶ <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/cloud-adoption-to-accelerate-it-modernization>

“

리더가 전체 데이터 분석 전략을 재고하고,
클라우드를 통해 어떻게 비즈니스 및
수익을 향상시킬 수 있을지 생각하는 것을
봅니다.

Sudhir Hasbe, Google Cloud 제품 관리 책임자

 + a b | e a u[®]