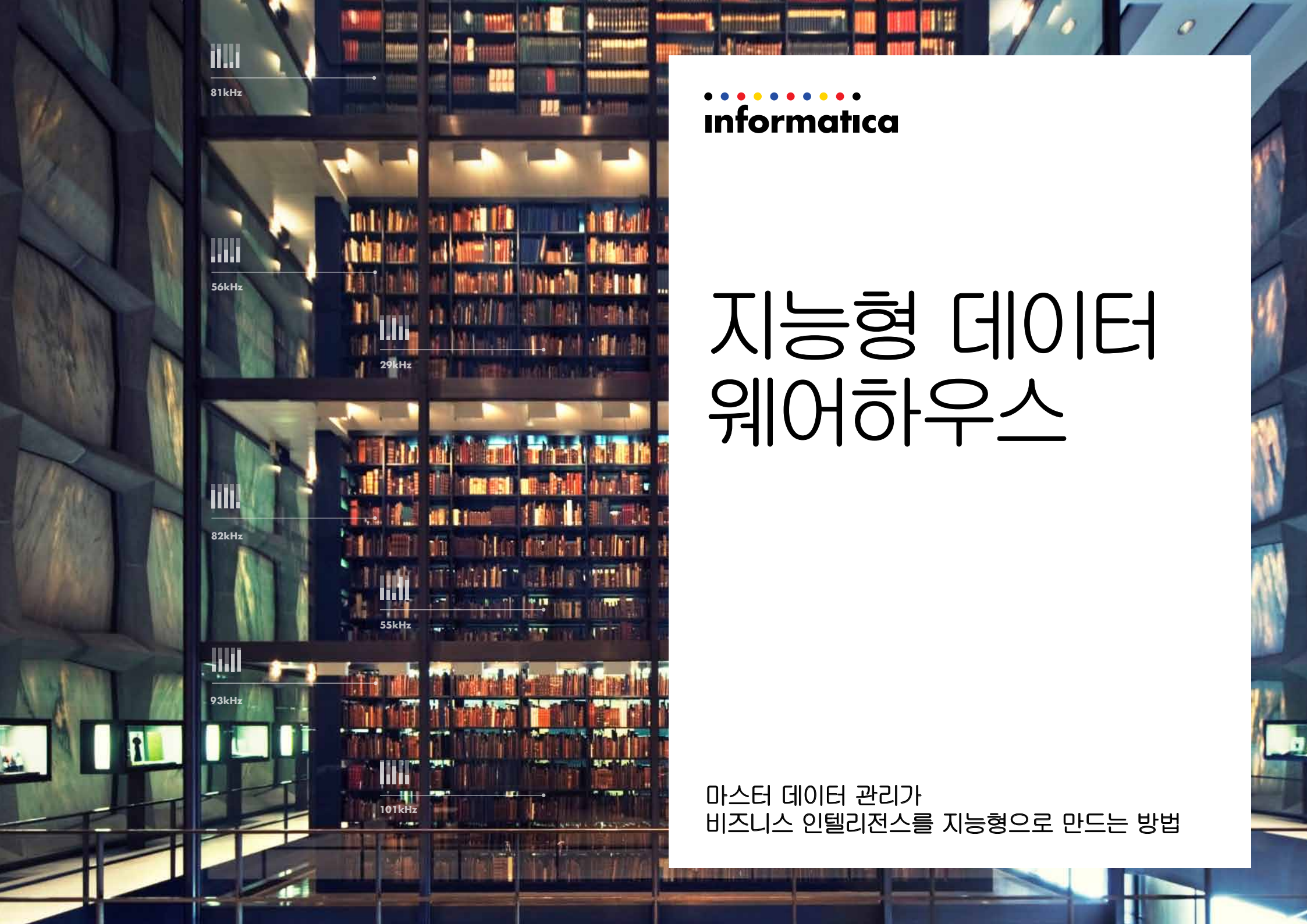


informatica

지능형 데이터 웨어하우스

마스터 데이터 관리가
비즈니스 인텔리전스를 지능형으로 만드는 방법

인텔리전스의 가능성

조직이 생성하는 데이터가 증가할수록 기업의 의사결정을 지원하기 위해 비즈니스 인텔리전스 시스템에 더 많이 의존합니다. 또한 데이터가 점점 더 유용하고 경쟁력 있는 자산이 되면서 기업에 비즈니스 인텔리전스가 그 어느 때보다 중요해졌습니다.

그러나 BI 소프트웨어, 데이터 웨어하우스 및 데이터 통합 프로젝트에 상당한 투자를 했음에도 분석팀은 정확한 보고서를 정시에 제출하는 데 여전히 어려움을 겪고 있습니다. 이유가 무엇일까요? 작업의 기반이 되는 데이터가 정리되지 않고 뒤엉켜 있으면 작업 속도가 느려지고 보고서의 정확성도 떨어집니다.

귀사의 비즈니스 인텔리전스는 결코 스마트해질 수 없습니다.



만약 귀사 역시 다른 기업 대부분과 마찬가지 상황이라면, 운영 시스템 내의 데이터가 부정확하고 일관성 없으며 많은 부분이 중복되어 모든 분석 결과의 신뢰성이 떨어지는 현상을 겪습니다. 이 때문에 귀사의 고객, 공급업체, 제품 및 직원에 대한 뷰가 단편화 되고 부정확해지는 결과가 나타납니다.

따라서 보고서와 그 보고서를 바탕으로 내리는 의사 결정 역시 부정확하고 신뢰성이 떨어질 수밖에 없습니다. 또한 그 보고서와 데이터를 기반으로 작업하는 모든 사람에게 악영향을 끼칩니다.

모두에게 영향을 주는 불량 데이터

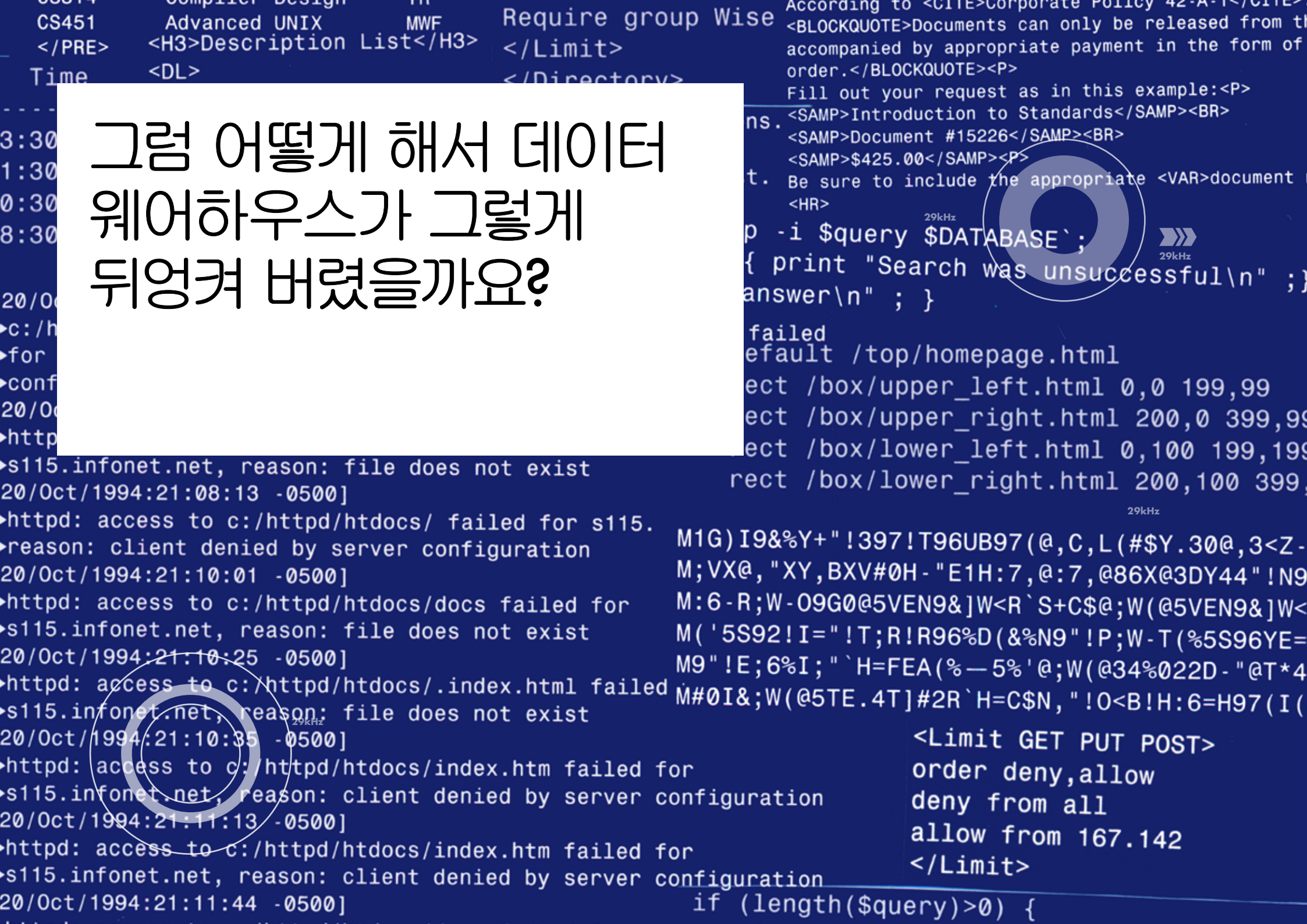
비즈니스 리더는 사용하는 보고서가 부정확하여 신속 정확성이 중요한 전략적 의사 결정에 차질이 생깁니다.

비즈니스 분석가는 데이터를 수동으로 조정하는 데 많은 시간을 낭비하여 판매 실적, 수익 배분 및 위험 완화에 대해 제대로 분석할 시간이 부족합니다.

IT 전문가와 데이터 아키텍트는 데이터의 정제, 교정 및 표준화를 오류가 발생하기 쉬운 대규모의 수동 작업으로 수행해야만 하기 때문에 변화 과정은 매우 괴로운 일이 됩니다. 또한, 인수 합병 같은 중요한 변경 작업도 시간 소모가 크고 지긋지긋한 고역이 됩니다.

규정 준수 담당자는 당국에 보고서 제출 시 자료의 정확성을 확신할 수 없습니다. 즉 완전히 예방 가능한 기본적인 행정 관리상 오류로 인해 회사 전체가 엄청난 처벌을 받을 위험에 노출된다는 의미입니다.

이 전자책은 비즈니스 인텔리전스가 본래의 가치를 제공할 수 있도록 귀사가 데이터를 준비하는 방법에 관해 설명합니다. 이를 실천하지 않으면 불량 데이터가 잘못된 인텔리전스를 형성하고 그 결과 잘못된 의사 결정을 내리게 되는 악순환으로 조직 전체가 어려움을 겪게 됩니다.


그럼 어떻게 해서 데이터
웨어하우스가 그렇게
뒤엉켜 버렸을까요?

현재 귀사의 비즈니스가 영위하는 데이터 환경은 귀사의 BI 시스템이 구축될 당시의 환경과 완전히 다릅니다. 따라서 귀사의 데이터 웨어하우스는 데이터가 야기하는 혼란에 대비되어 있지 않습니다. 특히 데이터 웨어하우스에는 다음과 같은 두 가지 중요한 문제가 있습니다.



1

다양한 소스로부터 입력되는 데이터

귀사의 데이터는 최소한 수십 개, 많으면 수백 개의 시스템에 저장되어 있습니다. 여러 부서의 다양한 사용자가 고객, 거래처, 제품에 대한 정보를 이러한 시스템 모두에 입력하기 때문에 정보는 그 형식과 정확도가 제각기 다른 채로 저장됩니다. 그 결과 데이터는 중복되고 일관성 없으며 부정확한 정보 덩어리가 되어 귀사의 데이터 웨어하우스에서 자체 정리할 수 없는 상태가 됩니다.

이런 문제는 한 국가 내에서만 발생해도 처리하기 어려운데, 글로벌 조직에서 발생하게 되면 그 혼란이 훨씬 증폭될 수밖에 없습니다. 귀사의 BI 도구가 이러한 복잡성을 해결할 수 없으므로 결국 그 공백을 메우는 것은 비즈니스 분석가의 몫이 됩니다.

더욱이 분석가가 이 방대한 작업을 해낸다고 해도 그 문제 해결은 일시적인 것으로 끝날 가능성이 높습니다. 즉 다음 보고서 작성 시에도 필요 이상의 작업을 수행하다가 결국 이전과 동일한 혼란에 맞닥뜨린다는 것입니다.

2

불일치 및 중복

귀사의 BI 시스템은 GE와 General Electric Company를 구별할 능력이 없습니다. 소스 시스템 내의 데이터 불일치는 BI 시스템을 통해 퍼져 나갑니다. 쓸모없는 데이터가 입력되면 쓸모없는 보고서가 작성되는 법입니다.

재무 보고가 정확성을 갖추려면 서로 다른 고객의 트랜잭션이 서로 다른 시스템에서 처리되더라도 고객의 정보를 종합해 내는 기능이 필수입니다. 데이터를 정확하게 요약하는 기능 없이는 분석 및 회계작업을 제대로 수행할 수가 없습니다.

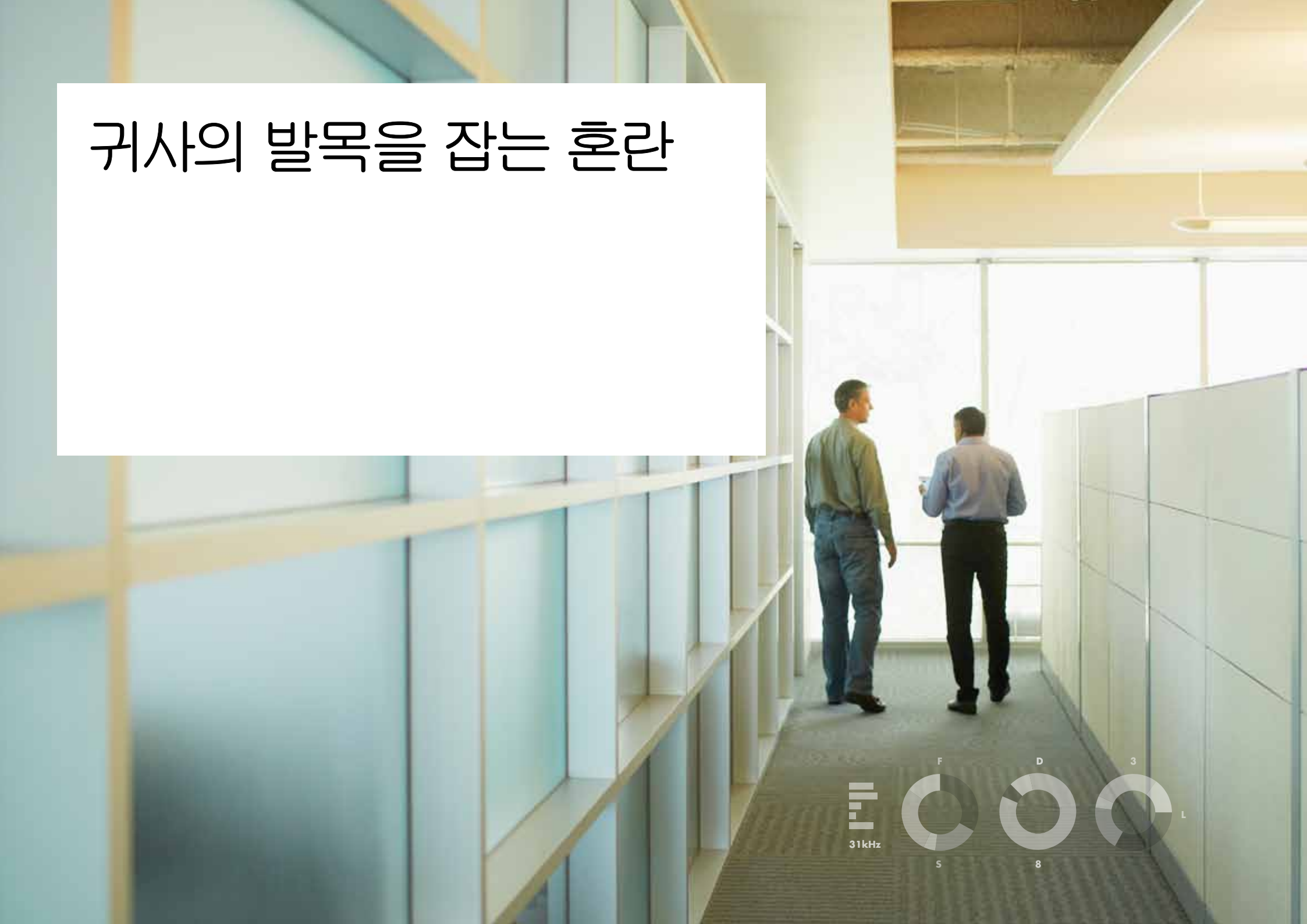
고객에 대한 단일 뷰가 없으면 영업과 마케팅에 해가 됩니다. 공급업체에 대한 단일 뷰가 없으면 공급망이 비효율적으로 구성됩니다. 또한 제품에 대한 단일 뷰가 없으면 제품 출시가 지연됩니다.

비즈니스 인텔리전스는 사용자가 신뢰할 수 있는 비즈니스의 권위 있는 뷰를 기반으로 합니다. 그러나 데이터에 신뢰성이 없다면 인텔리전스 역시 믿을 수 없습니다.





귀사의 발목을 잡는 혼란



신뢰할만한 정보의 단일 소스를 확보하지 않으면 인텔리전스가 필요한 부서는 거의 모두 어려움을 겪습니다.

영업

점점 경쟁이 치열해지는 시장에서 귀사의 영업팀이 고객을 계속 유지하기 위해서는 고객에 대해 최대한 많은 정보를 확보해야 합니다. 정보 확보가 부족하면 경쟁력이 현저하게 약화되는 것은 자명한 일입니다.

예를 들어, 브라질에 본사를 둔 화장품 제조 및 판매 회사인 Natura는 영업 인력이 고객과의 관계 및 상호 작용에 대해 정확한 뷰를 가지고 있지 않다는 사실을 알았습니다.

마케팅

마케팅 조직이 리소스를 전략적으로 할당하려면 고객층과 잠재 고객을 다양한 특성으로 세분화할 수 있어야 합니다. 더욱 중요한 것은 마케팅 캠페인 성공을 추적할 수 있는 능력입니다.

첨단 기술 업체인 Citrix는 채널 파트너, 고객 및 잠재 고객에 대한 고품질 데이터가 없었다면 Salesforce.com, Marketo 및 MyCitrix.com에서 현재와 같은 고차원의 전체 뷰를 확보하지 못했을 것입니다. 결과적으로 얼마나 많은 잠재 고객이 진정한 기회로 발전했는지 알 수 없었을 것입니다.

인사 관리

HP의 시스템에 고객, 파트너 및 영업사원에 대한 단편화되고 중복된 데이터가 존재함으로 인하여 인사 관리 부서가 영업 담당자에게 보너스를 잘못 지급하였습니다.

직원 데이터의 뷰가 상이할 경우 지나치게 높은 금액을 지급할 수 있다는 것입니다. 보너스가 영업 효과에 직접 연관된 경우에는 특히 그렇습니다.

재무

한 대형 생명공학 회사가 정부 기관 고객사와의 계약 요율을 정할 때, 이 회사는 그들의 활동과 요구 사항에 대해 보유하고 있던 뷰를 신뢰할 수 없다는 사실을 알았습니다. 관계에 대한 정확한 보고서에 액세스할 수가 없어 계약 요율 재협상 시 직접적인 장애 요소가 되었습니다.

더 넓은 관점에서 보면, 불일치와 중복 데이터를 완전히 조정하지 않으면 재무 보고의 무결성에 해가 됩니다. 또한 머지않아 이러한 신뢰 부족으로 인하여 항상 조직 재무 건전성에 대한 리더의 관점과 연관될 수밖에 없는 비즈니스 리더의 의사 결정이 질적으로 훼손됩니다.

규정 준수

정부 규정을 준수하려면 조직은 운영 현황을 추적하고 이를 보고할 수 있어야 합니다. 그러므로 벌금 및 징역형을 피하려면 전례 없이 높은 수준의 데이터 정확성을 기해야 합니다.

예를 들어, 제약회사를 대신해 임상시험을 행하는 임상시험수탁기관인 Quintiles는 FDA(Federal Drug Administration)의 정밀 조사 규정을 지속적으로 준수해야 합니다. 준수하지 않으면 FDA는 약물 승인을 오랫동안 지연시킬 수 있고 이에 대한 대가는 가혹할 수 있습니다.

일부 산업은 상대적으로 규정 준수에 더 엄격하기는 하지만 규정 준수 보고가 필요 없는 산업은 없습니다. 그리고 이 의무를 쉽게 수행할 수 있는 조직 역시 분명히 없습니다.

연구 개발

더블린에 본사를 둔 다른 임상시험수탁기관인 ICON은 임상시험 연구 진행 결과에 대해 고객사인 제약회사와 논의해야 합니다. 보고서는 제약회사에 임상시험의 효능을 알려주고 부작용에 대해 경고합니다.

연구자, 시험 현장, 위치 등 중요한 엔티티에 대한 뷰가 부정확하고 불완전하면 임상시험의 위험에 대한 모니터링과 연구자 성과 추적 및 정확한 보고서 작성이 불가능합니다.

그 외 다수

귀사의 BI 시스템이 올바른 정보를 올바른 방법으로 고속 처리하고 있지 않다면 판매 및 공급업체에 대한 위험부터 수익까지 모든 분석, 보고 또는 정확한 데이터에 의존하여 내리는 의사 결정에 어려움을 겪게 됩니다.

데이터 통합과 데이터 품질 관리가 좋은 방법이지만 충분하지 않은 이유



데이터 웨어하우스를 구축하려면 두 가지 중요한 도구, 즉 데이터 통합 및 데이터 품질에 의존해야 합니다. 그러나 안타깝게도 이 도구는 상당한 이점을 제공하기는 하지만 비즈니스 인텔리전스를 더욱 스마트하게 만들어주지는 않습니다. 그 이유는 다음과 같습니다.

1

데이터 통합

서로 다른 소스를 단일 시스템으로 합치는 일이 매우 중요하기는 하지만, 데이터 통합은 BI 시스템의 요청에 따라 데이터를 데이터 웨어하우스로 가져간 후 그대로 남겨두는 단방향 연결 방식으로 데이터를 조정하기 위해 구축되는 것이 아닙니다.

더 중요한 것은 데이터 통합이 코딩 작업으로 이루어진 경우입니다. 따라서 스키마가 변경되거나 차원이 업데이트될 때마다 ETL 개발자는 대량 수동 작업을 수행해야 합니다. 변경 사항을 적용하기 위해 수백만 개가 될 수도 있는 스크립트를 처리해야 하고, 그 후 해당 차원의 계보와 이력을 추적해야 합니다.

2

데이터 품질

정확성을 향상하고 데이터 자산 가치를 높이기 위해 데이터 품질 도구는 일반적으로 데이터를 데이터 웨어하우스에 보내기 전에 데이터를 정제하는 데 사용합니다.

그러나 데이터 웨어하우스에 보내기 전에 데이터를 정제한다는 것은 소스 시스템이 여전히 부정확하고 불완전한 데이터로 가득 차 있다는 뜻도 됩니다. 데이터 품질 소프트웨어는 정제된 좋은 데이터의 리포지토리를 관리하는 것이 아니기 때문에 이러한 소스 시스템의 사용자는 아직 정제되지 않은 데이터에 계속 의존합니다.

따라서 두 사람이 같은 데이터(예: 월간 판매 수익)를 두 개의 다른 소스 시스템에서 본다면 서로 다른 두 가지 버전의 사실을 보게 됩니다. 그리고 이 두 가지 버전이 나타내는 사실은 실제 판매량이 아니라 각 시스템의 업데이트 주기에 따라 좌우됩니다.

따라서 이러한 도구처럼 데이터 웨어하우스로 통합하는 일이 매우 중요하기는 하지만 BI라는 퍼즐을 완성하기 위해서는 이것이 전부가 아니라는 것입니다.

마스터 데이터 관리 및 지능형 데이터 웨어하우스



이제 데이터 웨어하우스에 방해되는 제한 사항을 열거하면서 개요를 정리해 보겠습니다.

데이터 소스가 통합되기는 하지만 데이터 자체는 여전히 중복과 불일치가 팽배한 상태입니다.

데이터가 정제되기는 하지만 정제 데이터는 실제로 어디에도 저장되어 있지 않습니다.

데이터가 변경될 수는 있지만 그러한 변경을 수행하고 기록하려면 오류가 발생하기 쉬운 수동 작업을 대규모로 수행해야 합니다.

따라서 이 세 가지 제한 사항을 모두 해결하는 일이 바로 인텔리전스 퍼즐에서 빠진 조각입니다.

데이터 전문가의 만병통치약과도 같은

마스터 데이터 관리(MDM)는 제어 프로세스로써 코딩이 없고 유지보수가 적게 소요되는 솔루션을 제공하며, BI 시스템, DI 도구, DQ 프로세스에서 수행할 수 없는 작업도 수행할 수 있습니다.

1

신뢰할 수 있고 조치할 수 있는 차원의 권위 있는 단일 뷰로 데이터를 통합합니다. 이는 일반적인 BI 시스템의 오류를 유발하는 모든 불일치 및 중복 문제(예: GE와 General Electric Company를 같은 항목으로 인식)를 해결합니다.

2

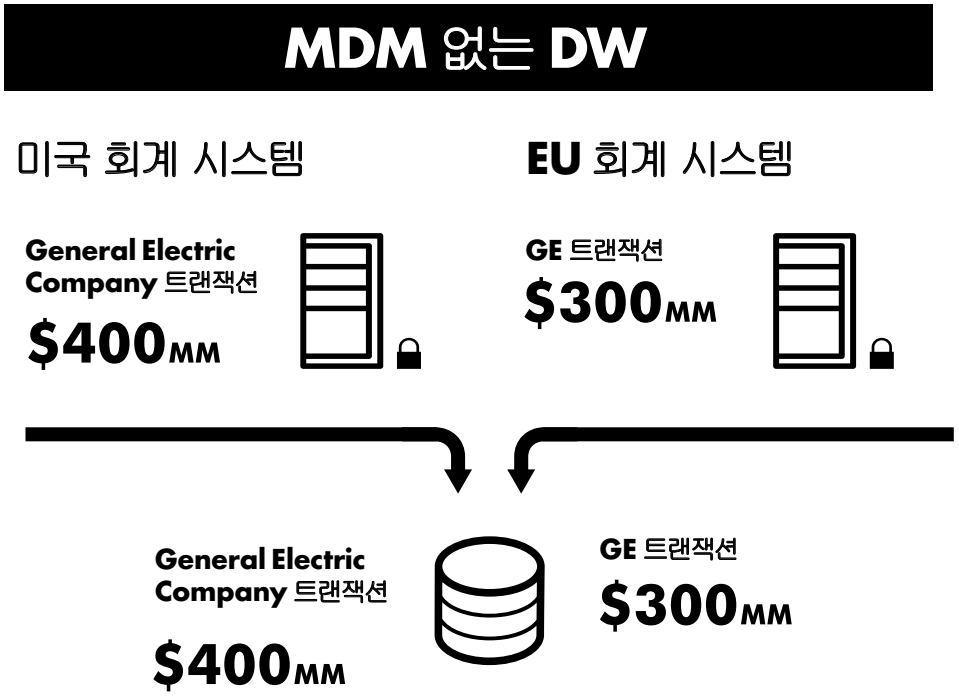
데이터를 정기적으로 자동 정제합니다. 그런 다음 정제된 고품질 데이터를 중앙 리포지토리에 모두 저장하고 연결된 모든 소스 시스템을 업데이트 합니다. 따라서 단말 시스템인 BI와는 달리 MDM은 양방향 데이터 흐름이 가능합니다.

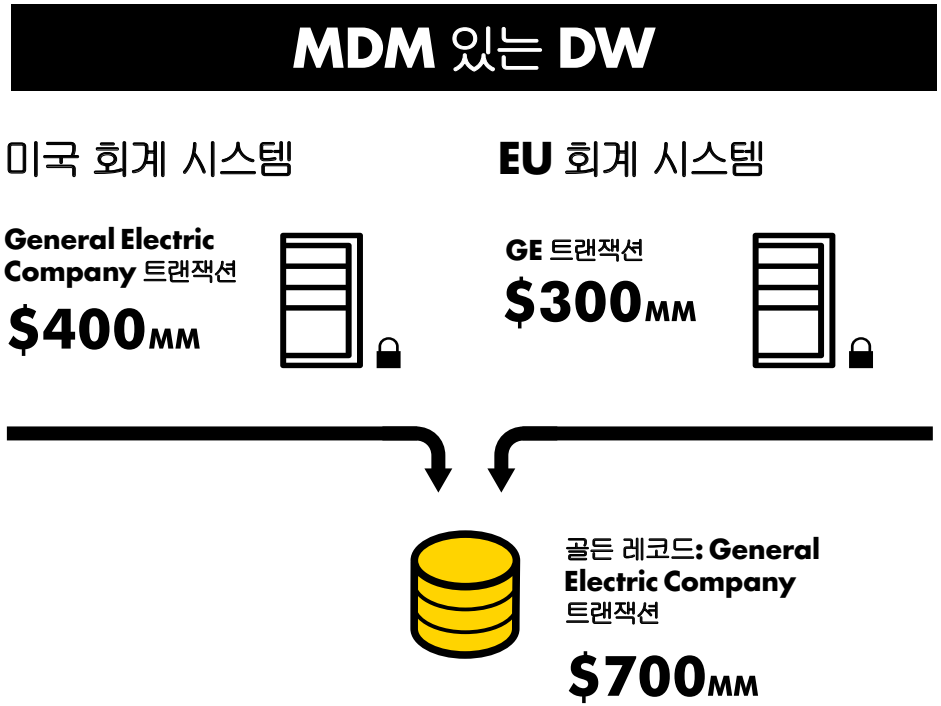
3

데이터가 변경되면 이상 현상을 즉시 알려주므로 변경 사항을 검사하고 저장하며 추적할 수 있습니다. 즉 데이터의 계보와 이력에 대한 신뢰할 수 있는 뷰를 가질 수 있습니다. 또한 ETL 개발자는 데이터의 사용을 준비하기 전에 모든 변경 사항을 수용하고자 과도한 작업을 할 필요가 없습니다.

이는 사실의 단일 버전을 보여주는 통합되고 일치되며 정제되고 자동으로 업데이트 되는 단일 리포지토리로 언제든지 이용할 수 있습니다.

MDM이 신뢰도 높은 데이터를 제공하는 방식





데이터 웨어하우징 시나리오 2가지: 좌측은 마스터 데이터를 데이터 웨어하우스로 보내기 전에 MDM 시스템을 사용하여 신뢰할 수 있는 마스터 데이터를 생성하지 않을 경우 보고가 부정확해질 수 있는 위험성을 보여줍니다.

우측은 MDM 시스템이 마스터 데이터를 데이터 웨어하우스에 보내기 전에 어떻게 데이터 문제를 해결하여 정확한 보고가 되도록 지원하는지를 보여줍니다.

"당사의 주목표는 영업 파이프라인 구축이며, **MDM**을 통하여 잠재 고객의 실제 고객 전환율을 **20%** 향상할 수 있었습니다. 데이터의 전체 뷰와 예측 분석 수행 기능 덕분에 우리는 거의 점쟁이 같은 능력을 갖추게 될 것입니다."

Dagmar Garcia, Senior Manager Marketing Operations

CITRIX®

MDM 기술의 장점

BI의 관점에서 보면 MDM의 가장 두드러진 장점은 다음 세 가지 기술 문제를 매우 효율적으로 처리하는 능력에서 기인합니다.

1 일치된 차원

MDM 시스템을 이용하면 변환에 따르는 수고, 즉 일반적으로 일치된 데이터 구조를 구축하는 데 따르는 수고를 대부분 덜 수 있습니다. 따라서 데이터 웨어하우스 인력은 데이터 통합이 아니라 비즈니스 문제에 대한 해답을 찾는 데 중점을 둘 수 있습니다.

2 느린 변경 차원 (Slowly changing dimensions)

MDM은 데이터 웨어하우스 개발자가 느린 변경 차원(SCD) 접근 방식을 채택할 때 일반적으로 발생하는 성능 문제를 해결합니다. MDM은 이력 추적 옵션을 제공하므로 참조 데이터 엔티티의 모든 변경사항을 기록할 수 있습니다. 따라서 데이터 웨어하우스는 느린 변경 차원(SCD)에서 데이터 세그멘테이션의 변경 내용 추적에 집중할 수 있습니다.

또한 사용자가 레코드에서 비점진적 변화 차원의 이력을 조회하고자 할 경우, MDM 시스템에서 히스토리 테이블을 찾아 볼 수 있습니다.

3 데이터 계보

MDM은 모든 레코드의 모든 필드에 자세한 데이터 계보를 가지고 있기 때문에(셀 수준까지), 소요 시간이 큰 수동 데이터 추적을 할 필요가 없습니다.

'한 레코드의 고객 이름 값이 John Smith인 이유'를 찾기 위해 모든 스테이징 테이블에 걸쳐 쿼리하는 대신에, 단순히 MDM 시스템에 액세스하여 그 값이 나온 소스 시스템을 식별할 수 있습니다. 또한 레코드 병합 이력까지 추적할 수 있습니다.



MDM 중심 데이터 웨어하우스의 역할

(그리고 그것이 대단한 이유)

데이터 웨어하우스로 가는 데이터가 지속적으로 업데이트 되는 정제 소스가 있다는 것은 사용해야 하는 유일한 종류의 데이터, 즉 오로지 사실만을 담은 단일 및 통합 버전의 신뢰성과 권위를 갖춘 정확한 데이터에서 분석 및 인텔리전스 시스템을 실행하고 있다는 뜻입니다.

이는 많은 시간과 비용을 투자한 **BI** 시스템에서 도출된 보고서와 데이터 사용자 모두에게 매우 반가운 소식입니다.

비즈니스 리더는 보고서가 전체 그림을 매우 분명하게 보여준다는 것을 확신해야 마음 놓고 더욱 좋은 의사 결정을 내릴 수 있습니다.

비즈니스 분석가는 오류와 불일치한 사항을 찾는 데 시간을 낭비할 필요가 없습니다. 회사에 고용된 본래 목적에 부합하는 업무, 즉 데이터 분석을 수행해야 합니다. 또한 더욱 쉽게 자신의 직감에 대한 유효성 검사를 수행하고 더욱 철저하게 가설을 시험할 수 있습니다.

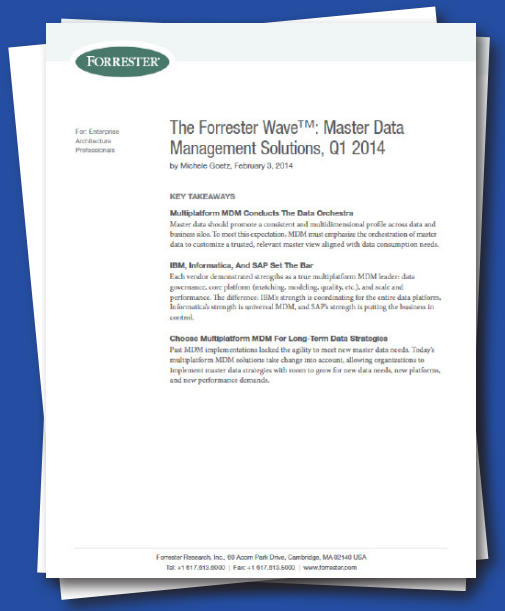
IT 전문가 및 데이터 아키텍트는 단순히 데이터를 전달하는 배관 역할로 끝나지 않고 데이터 자체의 품질을 향상하는 데 중점을 둘 수 있습니다. 또한 지속적인 개선 및 통합 준비가 된 시스템으로 작업하기 때문에 스키마가 변경될 때마다 하드 코딩을 할 필요가 없습니다.

규정 준수 담당자는 규정에서 요구하는 정확도 수준에 부합하는 보고서를 작성할 수 있습니다. 더 중요한 점은, 더 나아진 보고서가 그들의 본분인 기준 준수 수행 시 큰 힘이 된다는 사실입니다.

실제로 믿을만한 정확한 정보를 바탕으로 실행되는 비즈니스 인텔리전스 시스템에는 아주 많은 이점이 있습니다. 또한 MDM은 BI 시스템이 진정한 인텔리전스라는 궁극의 목표에 도달하는 데 주된 역할을 합니다. 비즈니스에서 더욱 지능적인 의사 결정을 내려야 한다면 비즈니스 인텔리전스는 더욱 좋은 데이터로 작업해야 합니다.

마스터 데이터 관리

당사의 MDM 제품과 멀티 도메인 기능은 규모와 업종에 상관없이 모든 조직에 파장을 일으켰습니다. 더 중요한 것은 MDM이 Forrester의 MDM Wave 보고서에서 선두 주자로 선정됐다는 사실입니다.



FORRESTER®

For: Enterprise Architecture Professionals

The Forrester Wave™: Master Data Management Solutions, Q1 2014

by Michele Goetz, February 3, 2014

KEY TAKEAWAYS

Multiplatform MDM Conducts The Data Orchestra

Master data should promote a consistent and multidimensional profile across data and business silos. To meet this expectation, MDM must emphasize the orchestration of master data to customize a trusted, relevant master view aligned with data consumption needs.

IBM, Informatica, And SAP Set The Bar

Each vendor demonstrated strengths as a true multiplatform MDM leader: data governance, core platform (matching, modeling, quality, etc.), and scale and performance. The difference: IBM's strength is coordinating for the entire data platform, Informatica's strength is universal MDM, and SAP's strength is putting the business in control.

Choose Multiplatform MDM For Long-Term Data Strategies

Past MDM implementations lacked the agility to meet new master data needs. Today's multiplatform MDM solutions take change into account, allowing organizations to implement master data strategies with room to grow for new data needs, new platforms, and new performance demands.

Forrester Research, Inc., 60 Acorn Park Drive, Cambridge, MA 02140 USA
Tel: +1 617.613.6000 | Fax: +1 617.613.5000 | www.forrester.com

'The Forrester Wave: Master Data Management Solutions, Q1 2014'(영문)에서 자세한 내용을 알아보십시오.

Informatica 정보

Informatica는 제품, 공급업체 및 고객에 관한 비즈니스에 중요한 데이터의 완전하고 정확한 뷰를 제공하여 조직에서 데이터를 우선 고려하도록 지원하고 있습니다.

Informatica에 문의하십시오.

IN18-0714-2667