

# IBM의 마스터 데이터 통합 기술에 의한 마스터 데이터 관리

---

## 목차

---

**2 1장 : MDM의 정의 및 중요성**

**6 2장 : 마스터 데이터 관리의  
기능적 구성요소**

**16 3장 : 마스터 데이터의 관리 경로**

**25 4장 : 베스트 프랙티스,  
이점 및 성공 사례**

기업의 전사적 비즈니스 애플리케이션이 복잡한 정보 기술 아키텍처에서 구현되는 오늘날, 몇 해 전에 다뤘던 주제들이 다시 새로운 화제거리가 되어 등장하는 것을 종종 경험합니다. 기본적인 프론트 오피스/백 오피스 기능을 지원하던 본래의 코어 애플리케이션들은 그간 핵심적인 비즈니스 엔티티에 대한 참조 데이터를 자체적으로 관리해 왔습니다. 요즘과 같은 정보 집약적 환경에서 고객, 공급업체, 품목 및 제품에 관한 참조 데이터는 정확하고 여러 부서에 걸쳐 공유가 가능하며 다양한 이종 애플리케이션에서 ‘온 디맨드’ 방식으로 사용 가능해야 합니다. 앞 글자를 따 “MDM”으로 표기되는 “마스터 데이터 관리”는 공유되는 참조 데이터를 현업 부서 및 기술 부서 모두에서 관리할 수 있는 기회를 제공한다는 측면에서 많은 주목을 받고 있는 만큼, 면밀히 살펴볼 필요가 있습니다.

마스터 데이터 관리는 사실 복합적인 소프트웨어 솔루션입니다. 한편 MDM 개념은 이미 레거시 및 패키지 엔터프라이즈 애플리케이션 포트폴리오에 포함되어 있습니다. 다른 한편으로, 이는 근본적으로 핵심적인 비즈니스 엔티티에 대한 사실 정보를 다루는 애플리케이션들 간에 실시간으로 이용해야 하는 데이터 통합 기능입니다. 비용 감소, 법적 요건 준수, 파트너와의 ‘표준’ 데이터 교환, 새로운 애플리케이션의 구현 또는 이전 애플리케이션의 통합 등, 그 목적이 무엇이든 상관 없이 마스터 데이터는 일관된 방식으로 처리되어야 합니다. 베스트 프랙티스를 살펴보면, 데이터 통합 개발 및 관련 비즈니스 프로세스의 정렬(alignment)을 이끄는 방법론이 중요한 성공 요건임을 알 수 있습니다. 데이터 통합과 비즈니스 프로세스 방법론의 틀 안에서, 모든 MDM 솔루션은 기존의 레거시 및 비즈니스 애플리케이션 풋프린트를 활용해야 합니다. 마스터 데이터 통합 서비스는 전사적 범위에서 마스터 데이터 관리 애플리케이션을 다양한 비즈니스 애플리케이션과 연계하고 실시간으로 마스터 데이터를 획득, 변환하고 전달하는 데 필요한 통합 서비스를 제공합니다.

본 백서에서는 성공적인 MDM을 위해 필요한 정보 통합에 대해 자세히 기술합니다. 우선, MDM의 정확한 의미를 살펴보고 왜 기업 경영진이 MDM에 관심을 기울여야 하는지 알아보겠습니다. 다음으로, 마스터 데이터 통합 요건을 중심으로, 완전한 마스터 데이터 관리 솔루션의 기능적 구성요소에 대해 기술합니다. 이어서 도입 및 운영 아키텍처의 대안을 살펴 본 후 마지막으로, 방법론의 중요한 역할 및 이점, 그리고 IBM MDM 솔루션을 이용한 고객 성공 사례에 대해 알아보겠습니다.

## 1장: MDM의 정의 및 중요성

시작하기에 앞서, 실질적 정의를 내려보자면, 마스터 데이터 관리는 기업 내 각기 다른 애플리케이션을 망라해 핵심적인 비즈니스 엔티티에 대하여 정확하고 일관적인 뷰를 구축하기 위한 비즈니스 프로세스, 애플리케이션 및 기술 통합 아키텍처입니다. 즉 고객이 어떤 방식으로 거래하는지 상관 없이(예: 인터넷, 지점, 신용카드, 마케팅 부서 및 고객센터 이용) 금융 기관이 어떤 방식으로든 정확한 고객 정보를 확보하느냐에 관한 것입니다. 그러나 이는 지금까지 해오던 일이 아닙니까? 즉 중복되거나 잘못된 데이터 없이 정확한 사실 정보를 반영해 고객/협력업체/직원/제품/원료에 대한 단일 뷰를 구축하지 않았습니까?

불행히도, 기업의 마스터 데이터 관리와 관련된 문제의 심각성은 이전 어느 때보다도 높습니다. 고객, 제품, 장소 등 핵심적인 비즈니스 엔티티를 기술하는 데이터를 작성하고 전달하는 활동은 사내 외에서 기하급수적으로 증가하고 있습니다. 대기업 대부분은 마스터 데이터를 위한, 일관적이고 정확한 '레코드 시스템'을 갖추고 있지 않습니다. 현실적으로, 많은 애플리케이션 저장소들이 마스터 데이터의 단편적인 버전만을 수록하고 있으며, 하나라도 완전한 것이 없습니다. 동시에, 상호 운용성, 컴플라이언스와 관련된 법적 요건 때문에 데이터의 질적인 문제점들이 지속적으로 지적되고 있습니다. 설상가상으로, 브라우저 기반의 차세대 포탈형 애플리케이션 덕분에 더욱 분리된 "마스터" 데이터 세트가 실시간으로 전달되도록 하는 요건이 증가하고 있습니다.

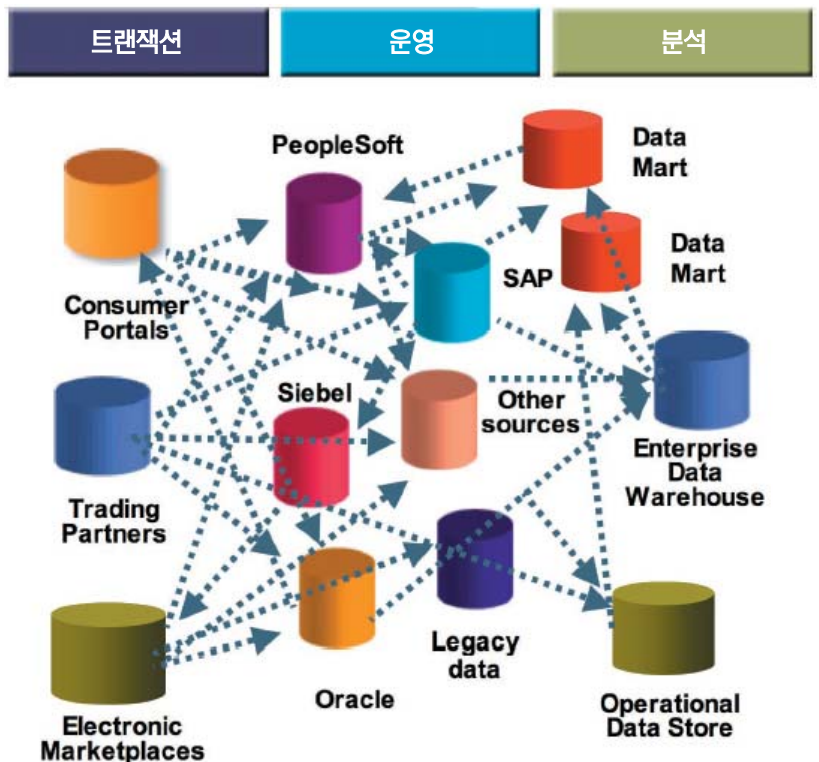


그림 1: 전사적 범위에서 마스터 데이터 관리의 복잡성

예를 들어, 첨단 기술 제조 분야에서는 어떤 신제품을 출시해 출고하기까지 다양한 프로비저닝 시스템, 현지 애플리케이션 및 파트너 시스템에 이 신제품을 추가하느라 최대 4주가 걸립니다. 새로운 절차 및 방법으로 제품 출시 일정을 며칠 수준으로 단축하기 위해 MDM에 투자하는 기업들이 많습니다.

**해석:** 핵심 비즈니스 엔티티에 대해 고품질의 정보를 즉시 입수해야 하는 그 순간, 그러한 정보가 수록된 데이터가 셀 수 없이 많은 데이터 트랜잭션 구조, 데이터베이스, 데이터마트 및 스프레드시트(그리고 미처 알지 못한 사내의 여러 장소) 곳곳에 흩어져 있습니다!

고객 식별 번호는 1,100만 개나 되지만, 실제 고객 수는 8백만 명 뿐이더라도 놀랄 일은 아닙니다. 핵심 비즈니스 엔티티에 대한 일관성 있는 뷰가 존재하지 않으면 비즈니스에 심각한 영향을 줍니다.

몇 가지 사례를 살펴보겠습니다.

## “기대에 미치지 못한 CRM”

그 까닭은 무엇입니까? 지금까지 고객들이 누구인지, 어떻게 연관되어 있는지, 어디에 위치하는지 또한 어떤 제품을 구입했는지에 대한 논리적으로 정확한 뷰가 어디에도 없었습니다.

- **수익성:** 66%의 응답자들은 전체적인 회사의 수익성이 부실한 정보로 인해 타격을 받았다고 밝혔습니다.
- **고객 서비스:** 75%의 응답자들은 부실한 고객 데이터가 고객 서비스, 품질 및 충성도에 지장을 주었다고 밝혔습니다.
- **부정확성:** 52%의 응답자들은 온갖 시스템이 통합된 까닭에 부정확한 정보가 흘러나온다고 지적했습니다.
- **잘못된 정보:** 한 정보통신 업체에서는 데이터 입력 오류로 인해 계정이 잘못 코드화되어 청구서가 발송되지 않았고, 그 때문에 한달 동안에 8백만 달러의 손실을 입었습니다.

## “동기화되지 않는 품목”

그 까닭은 무엇입니까? 100여 곳에서 구입, 제조 및 재고 사항을 기록하지만, 각 품목에 대한 설명이 부족한 이종 데이터 소스 전반에 걸쳐 동일한 원료, 부품 또는 제품과 링크되는 공통 키는 존재하지 않습니다. 이러한 작업에는 전문가가 필요합니다!

- **데이터의 오류** – 소매업체 시스템의 데이터 중 30% 가 정확하지 않습니다.
- **생산력의 손실** – 매년 각 SKU(Stock Keeping Unit) 당 25분의 수동 정제 작업이 필요합니다.
- **출시 일정 지연** – 새로운 제품을 출시하는 데 4주의 시간이 걸립니다.
- **송장 금액 누락** – 43%의 송장에서 금액이 누락됩니다.
- **검사 실패** – 매주 최대 70,000회(미국의 한 대형 소매업체)
- **매출 손실** – 매년 최고 3.5%

출처: A.T. Kearney, GMA

## “잘못된 답을 내놓는 BI”

그 까닭은 무엇입니까? 고객, 공급업체 및 제품 등의 일차적인 요소에 할당된 고유의 키는 현실 정보를 반영하지 않습니다. 품목과 금액 등의 기본적인 사항을 취합했을 때 그 답은 부정확하고 다른 출처에서 얻은 정보와 일치하지 않습니다.

- **제한적인 수용** – 2007년까지 50% 이상의 데이터 웨어하우스 프로젝트는 제한적으로 수용되는 데 따른 어려움을 겪을 것입니다. 완전한 실패는 아니더라도, 데이터의 질적인 문제에 대한 주의가 부족하여 이러한 결과가 나타날 것입니다.
- **생산력 손실** – BI 프로젝트에서 IT 관련 작업 중 약 80%가 인프라, 데이터 분석, 데이터 획득 및 데이터의 질적인 문제 해결에 집중될 것입니다.

그 밖에도 여러 가지 요인들이 존재합니다. 벤더를 통합할 수 없습니다. 회원, 가입자 및 공급사를 상호 비교할 수 없습니다. 부채 규모를 정확하게 기술할 수 없습니다. 게다가 규정과 표준 규격을 따르기도 어렵습니다.

다시 MDM에 관한 얘기로 돌아가자면, 마스터 데이터 관리는 다음과 같은 기회를 제공합니다.

- 기업 내 어느 곳에서도 핵심 비즈니스 엔티티에 관한 사실 정보를 입수할 수 있는 정보 통합 플랫폼을 구축합니다.
- 실제의 사실과 일치하는 비즈니스 규칙을 기반으로 논리적으로 정확한 단일 뷰를 자동으로 생성합니다.
- 마스터 데이터 시스템과 더불어 마스터 데이터 저장소를 능동적으로 관리합니다.
- “신상품 출시”, “신규 고객 설정”과 같은 프로세스를 자동화하고 표준화합니다.
- 기존 비즈니스 애플리케이션에 질적으로 우수한 마스터 데이터를 실시간으로 제공합니다.

물론 이러한 노력은 줄곧 이루어져 왔습니다. 그러나 차세대 IT 아키텍처에서 마스터 데이터를 활용하고 업무상 중요한 여러 이니셔티브에서 확실한 효과를 얻을 수 있다는 측면에서 이를 더 정확히 이해할 필요가 있습니다. 2장에서는 마스터 데이터 솔루션을 구성하는 기능적인 구성요소를 상세히 살펴보고, 3장에서는 현재의 애플리케이션 포트폴리오에서 전개 및 운영 통합을 위한 아키텍처적 대안을 살펴봅니다. 4장에서는 방법론의 중요한 역할에 대해 알아보고 그 이점에 대해 알아볼 것입니다. 마지막으로 IBM 고객의 MDM 성공 사례를 소개합니다.

## 2장: 마스터 데이터 관리의 기능적 구성요소

마스터 데이터 관리란 사실상 하이브리드 오퍼링입니다. 한편 MDM은 브랜드 매니저가 상품 품목 포트폴리오를 관리하고, 재무 계획 담당자가 계정 과목표를 관리하는 등 현업 부서 사용자의 업무에 의해 사용됩니다. 다른 한편으로 MDM은 기본적으로 전사적으로 고객, 협력업체, 공급업체와 같은 핵심 비즈니스 엔티티를 관리하고 접근할 수 있게 하는 통합 구조의 일부라고 할 수 있습니다. 성공적인 마스터 데이터 관리 솔루션에서 가장 중요한 이점은 하나의 마스터 데이터 엔티티의 "인스턴스"가 다른 인스턴스와 동일하지 판단하는 과정을 자동화할 수 있다는 것입니다. 예를 들어, 어떤 고객의 새 "인스턴스"(고객의 이름과 주소 등 그 고객을 설명할 수 있는 기본적인 사항으로 구성)와 함께 구매를 처리하려면, MDM 솔루션은 비즈니스 규칙에 따라 해당 고객이 이미 존재하는지 아니면 새로운 고객인지 결정하여야 합니다. 쉬운 일처럼 들리지만 결코 그렇지 않습니다.

## IBM 마스터 데이터 관리

### 종합적인 마스터 정보와 비즈니스 프로세스 관리 능력 확보

- 마스터 데이터 솔루션, 산업 및 콘텐츠 모델을 통해 신속한 전개 및 ROI를 실현합니다.
- 종합적인 마스터 정보와 비즈니스 프로세스를 관리합니다.
- 업계 최고 수준의 마스터 데이터 통합 능력을 활용합니다.
- 수백 건의 구축 성공 사례에서 얻은 비즈니스 변환, 베스트 프랙티스 및 방법론을 망라해 깊이 있는 전문성을 확보합니다.
- 강력한 파트너 에코시스템을 구축합니다.



그림 2: IBM 마스터 데이터 관리

## BM의 MDM 솔루션 공급 방식의 주요 구성 요소

**1. 마스터 데이터 관리 시스템** – 비즈니스 서비스 패키지를 통해 참조 데이터 저장소를 관리하는 트랜잭션 소프트웨어의 인프라입니다. 이와 같은 서비스에는 관리자/사용자 화면, 저장소 관리 기능, (데이터 계층 구조를 정의하고 관리하는) 워크플로우 및 이벤트 그리고 특정 데이터 요소/산업과 관련된 관계/속성 등이 포함됩니다. 예를 들면,

- 소매업체 및 CPG(Consumer Packaged Goods) 제조사 환경에서 IBM WebShpere® Product Center와 같은 MDM 시스템은 제품 번호, 명세, 가격, 이미지 및 패키지 세부 사항, 협력업체 및 공급업체 등을 관리합니다. 이와 같은 형태의 MDM 시스템은 보통 통합 제품정보관리 (PIM – product information management)로 불립니다.
- 금융 서비스, 소매금융, 통신 및 보험 업계에서 IBM WebShpere Customer Center와 같은 MDM 시스템은 재무 상태, 위치, 인구통계 데이터, 대금 청구 및 고객 정보 등을 관리합니다. 이와 같은 MDM 시스템은 개인 및 법인과 관련된 모든 역할을 포함하는 ‘관계자(party)’의 개념을 구현하는데, 이를 보통 고객 데이터 통합(CDI – customer data integration)이라고 합니다.



- 그 외에도 기업들은 제품 지향적 또는 고객 지향적인 정보 관리를 위해 맞춤형 MDM 애플리케이션을 구축하기도 합니다.

**2. 마스터 데이터 통합** – 모든 산업, 업무 기능 및 업무 데이터 범위를 망라하여 구현 문제와 지속적인 운영 데이터 문제를 해결할 수 있도록 하나의 통합 인프라를 제공합니다. 여기에는 마스터 데이터 비즈니스 트랜잭션의 관리 그리고 전사적 범위에서의 데이터 동기화 보장이 포함됩니다. 또한 연합된 미들웨어를 통해 MDM 시스템은 외부 데이터 소스에 동적으로 액세스하면서 관리 대상 엔티티와 관련된 콘텐츠(이미지, 문서 등)를 입수할 수 있습니다.

**3. 마스터 데이터 솔루션** – 수평적인 기술 차원을 넘어 MDM은 핵심 업무 영역을 대상으로 하는 MDM 시스템의 신속한 맞춤화를 지원하도록 산업별 데이터 모델, 워크플로우 및 업무 프로세스로 구성되어 있습니다. 여기에는 재사용 가능한 통합 템플릿도 포함되는데, 이 템플릿은 MDM 저장소로의 초기 로드 과정에서 ERP, CRM과 같은 공통 애플리케이션 소스로부터 더욱 신속하게 데이터를 옮기고 정제하는 데 도움이 됩니다.

## 마스터 데이터 근본 원인 분석

마스터 데이터 저장소 구축에 앞서 전사적 범위에서 엔티티, 즉 고객 또는 제품 등에 대해 논리적으로 정확한 뷰와 함께 구축되어야 합니다. 안타깝게도 전사적 범위에서 애플리케이션에 고객에 대한 사실 정보가 저장되는 방식에 대해 '의미론'적 합의가 이루어지지 않았습니다. 마스터 데이터 엔티티를 설명하는 사실 정보의 의미에 대해 합의를 도출하는 것은 마스터 데이터 저장소를 성공적으로 구축하고 지속적으로 통합하는 데 절대적으로 중요합니다.

예를 들어, 전사적으로 어떤 고객을 설명하는 데 사용되는 속성에 공통적인 정의가 존재하지 않습니다. 이름의 구성 요소(성, 이름, 칭호, 경칭) 및 위치(번지, 주소, 동 호수, 아파트 명, 우편번호, 시, 도, 사서함)의 정의 수준은 "40자 이내의 이름"과 다를 바 없습니다.

제품 관련 속성(제품명, 제품 브랜드, 크기, 색상, 포장, 무게)도 마찬가지로, “50자 이내의 제품 설명” 수준에 그칠 뿐입니다.

### 근본 원인 분석:

### 전사적 범위에서 소스 간에 일관성이 없는 마스터 데이터

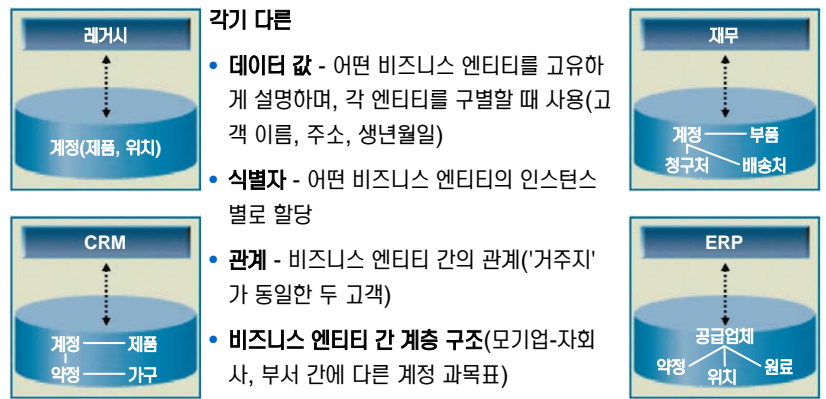


그림 3: 근본적인 원인 분석

고객, 위치, 공급업체, 부품, 제품, 원료와 같은 핵심 비즈니스 엔티티에 대한 데이터를 관리하는 각 애플리케이션은 각자의 목적에 따라 (고유 데이터 모델을 사용해) 기능을 수행하지만, 전사적 시스템을 망라해 일관성을 띠지 못합니다. ‘레거시’ 애플리케이션은 주로 계정 중심이므로, “이름 및 주소 열 1~5”로 표현되곤 하는 고객 마스터 데이터 엔티티에 대한 특별한 개념이 존재하지 않습니다. 금융 애플리케이션에서는 청구처 고객의 신용도에 문제가 없는지, 공급된 제품 및 서비스에 대한 비용을 지급할 것인가에 관심을 갖습니다. ERP는 특정 부품, 제품 또는 원료를 특정 장소로 발송하기 위해 누구와 접촉할 것인가를 다룹니다. CRM에서는 고객 문의처, 고객들 간의 상호 관계, 여러 접점을 망라해 문의처를 최적화하는 방법에 주목합니다.

원칙적으로 마스터 데이터를 필요로 하는 프로젝트는 여러 애플리케이션군의 집합체에서 데이터를 수집해야 합니다. 소스 데이터를 변환하고 이동하기에 앞서 그 소스 데이터를 제대로 이해하는 데 적절한 주의를 기울이지 않는 경우가 많습니다. 부실한 데이터는 프로젝트 실패를 야기합니다. 다음과 같은 데이터 품질에 대한 잘못된 상식을 믿는 과오를 범하지 마십시오.

- 우리 회사의 데이터는 우리가 잘 안다.
- 소스 문서가 존재할 뿐 아니라 정확하다.
- 우리 회사의 메타데이터는 데이터를 정확하게 설명하고 있다.
- 사용자들은 시스템을 "알맞게" 사용하고 있다.
- 시스템 구축 후 비즈니스 규칙은 바뀌지 않았다.
- 주제별 전문가가 있다.
- 데이터 품질을 평가하고 재조정하기에 충분한 시간이 프로젝트 범위에 포함되어 있다.

성공적인 마스터 데이터 관리는 그 시스템 내에 핵심 참조 데이터 저장소를 만들어 여러 애플리케이션 간의 마스터 데이터를 논리적으로 연결시킵니다. 상호 참조 키는 MDM 시스템에 존재하는 객체를 나머지 애플리케이션, 데이터베이스 또는 외부 참조 소스의 해당 객체와 연결시킵니다. 실제 현실을 반영하는 단일 버전의 정보로 업무를 처리할 수 있게 하는 것이 바로 이 상호 참조 키입니다.

이렇게 데이터 정보를 "재설계"하고, 핵심 비즈니스 엔티티 간의 관계를 규명하며, 논리적으로 정확한 키를 작성, 유지 관리하고, 상호 참조를 통해 이를 관리하며 공통의 데이터 통합 플랫폼을 활용하는 이 기본적인 과정이 성공에 관건이 됩니다.

## 마스터 데이터 통합 개요

초기 도입 단계에서 그리고 MDM 시스템의 데이터 무결성을 지속적으로 유지하기 위해 사내의 어떤 곳에 있는 사실 정보도 고품질의 마스터 데이터로 변환하는 데 필요한 기능적 구성 요소를 살펴보겠습니다.

## 마스터 데이터 관리의 기능 스택



그림 4: 마스터 데이터 관리의 기능 스택

### 프로파일링 & 규명

우선, 소스 시스템의 데이터를 이해하는 것이 무엇보다 중요합니다. 이는 현재 존재하는 상태에서 데이터 품질을 확인하는 데 도움이 됩니다. 이 과정을 자동화하는 데이터 프로파일링 툴은 기존의 데이터 분석 방식과 달리 번거로운 수작업 없이 소스 시스템 데이터를 심층적으로 분석합니다. IBM WebSphere ProfileStage™는 column, cross-column 및 cross-table 분석을 수행하는 자동 프로파일링 솔루션 중 하나입니다. 또한 데이터 분석가는 링크된 레코드를 확인하고 조합하는 다운스트림 프로세스에서 활용 가능한, 소스 시스템의 메타데이터 맵을 이 제품을 이용해 정의할 수 있습니다. 새로운 마스터 데이터 모델이 소스 시스템 스키마에서 만들어질 경우 그 모델의 기반으로 또는 단순히 소스 스키마를 새로운 대상 모델에 매핑하기 위해 메타데이터를 이용할 수 있습니다.

일단 소스 시스템의 내용과 구조가 파악되었다면, 이제 활발한 인터랙션이 소스 시스템에 영향을 주지 않는 작업 영역으로 데이터를 이동해야 합니다. 그러기 위해 일반적으로 임시 스테이징 데이터베이스를 만듭니다.

이 데이터베이스는 소스 시스템 데이터의 복사본을 제공하는데, 이 복사본은 데이터 정제, 매칭 및 상호 참조 생성 로직을 만드는 데 사용할 수 있습니다. 소스 데이터 프로파일링의 필요성, 방법 및 그 효과에 대한 자세한 내용은 IBM 백서 'Profiling: First Step to Assure Data Quality' 를 참조하십시오.

### 표준화

스테이징 데이터베이스가 마련되었다면, 전사적으로 소싱된 레코드에서 찾은 핵심 비즈니스 엔티티 관련 사실 정보를 표준화하는 규칙을 만들 수 있습니다. 표준화는 사실 정보 또는 속성에 대한 '고정 필드'를 만들고 고정 필드가 만들어진 각 사실 정보에 비즈니스 지향적인 의미론적 레이블을 부여하는 과정입니다. 마스터 데이터의 표준화는 공통 플랫폼에서 다양한 개인 이름, 법인 이름, 전세계 위치/ 언어, 제품, 부품 및 원료 데이터를 능히 처리할 수 있을 만큼 유연해야 합니다. 표준화가 이루어지면, 비즈니스 맥락에서 사실 정보를 이해하고 매칭 및 레코드 연결을 위한 토대가 마련됩니다.

### 글로벌 주소 확인 및 인증

지금과 같은 글로벌 경제 시대에 많은 기업들은 전세계적인 범위에서 핵심 비즈니스 엔티티에 대한 정보를 통합할 방안을 모색 중입니다. 따라서 전세계의 위치 정보가 현지 비즈니스 규칙을 적용하되, 일관성 있게 표준화되어야 합니다. 위치 정보의 구성 요소가 표준화되었다면, 이제 이러한 위치 정보가 알맞은 주소를 나타내는지 검증하는 것이 중요합니다. 글로벌 주소 확인에서는 각 위치 인스턴스에 대해 표준화된 사실 정보를 제3자가 취합한 참조 데이터베이스와 비교해 해당 지역의 주소가 유효한지 확인합니다. 미국 체신부 인증을 취득한

WebSphere QualityStage™ CASS와 같이 우수한 주소 정보 표준화 능력을 갖추었음을 "인증"받은 소프트웨어를 사용할 경우, 우편 서비스 기관으로부터 추가적인 우편 요금 할인 혜택을 받기도 합니다.

### 매칭 및 레코드 연결

매칭 및 레코드 연결은 기술적으로 까다로운 작업입니다. 어떤 마스터 데이터 엔티티(고객)의 인스턴스와 관련된 사실 정보 세트(로버트 리치, 브라이어힐 가 17번지)를 입수하고 이를 참조 저장소와 비교해 해당 인스턴스(로버트 리치, 브라이어힐 가 17번지)가 이미 알려져 있는지 여부를 판단합니다.

매칭이 필요한 이유는 전사적으로 사용 가능한 공통 키가 없기 때문입니다. 매칭 과정에서는 데이터의 변칙성과 충돌을 간파하여야 합니다. 구체적으로 설명하자면, 한 고객을 설명하는 입력 데이터 세트를 선택합니다. 0.5초 안에, 99.5%의 정확성으로, 해당 고객이 고객 마스터 파일에 존재하는지 자동적으로 결정합니다. 과거에는 이러한 작업이 몹시 어렵고 많은 리소스와 시간을 요했습니다. 따라서 일정한 주기를 두고 오프라인상에서 이러한 작업을 진행할 수 밖에 없었으며 서비스 부서가 모든 일을 떠맡기도 했습니다. 오늘날 마스터 데이터 관리가 전사적 차원의 전략으로 부상하는 까닭도 다양한 엔터프라이즈 애플리케이션을 망라해 핵심 비즈니스 엔티티에 대한 레코드를 실시간으로 매칭하여 연결시킬 필요성에 있습니다.

아키텍처의 관점에서 보면, 그와 같은 작업을 다양한 애플리케이션 및 비즈니스 목표에 따라 일관성 있는 매핑/연결 규칙의 집합을 배치 형식으로, 실시간에 가깝게 또는 실시간으로 이행하는 공통적인 "서비스"로 여기는 기업들이 많습니다. 매칭 기술에 대한 자세한 내용은 IBM 백서 "IBM WebSphere Quality Stage™: Superior Technology Produces Superior Results"를 참조하십시오.

## 마스터 데이터 보강

마스터 데이터 보강은 각 고객 인스턴스를 폭 넓고 깊이 있게 이해하기 위해 개인 및 기업 고객과 같은 핵심 비즈니스 엔티티에 관한 표준화된 사실 정보를 제3자의 참조 데이터와 연결하는 과정입니다. 예를 들어, Acxiom InfoBase 등에서 공급하는 인구 통계 및 사이코그래픽스 정보를 접목시켜 개별 고객에 대한 뷰를 확장하는 기업이 많은데, 수익, SIC 코드, 기관 간의 법적 관계와 같은 기업 속성 정보를 Dun & Bradstreet와 같은 벤더가 공급하기 때문입니다. 제3자의 데이터를 마스터 데이터 관리의 핵심적인 역량으로 활용함으로써, 그 데이터를 사용하는 애플리케이션에 상당한 부가 가치를 제공할 수 있습니다.

## 서바이버십

매칭 및 레코드 연결 과정의 산출물은 흔히 레코드의 집합, 즉 현실 세계에서 동일한 엔티티를 나타내는 것으로 매핑 기술에서 판단한 사실 정보의 집합으로 표현됩니다.

이러한 집합에 하나의 키를 할당함으로써 '논리적 키의 무결성'이 발생하며, 이는 '360도 전방위적 뷰' 및 '단일 버전의 진실'이 구현될 기반을 조성합니다. 서바이버십이란 특정 업무 목적에 어떠한 사실 정보의 조합이 '최선'인지 결정하는 것입니다. 아래에서 다루겠지만, 이 집합에 있는 레거시 키에서 그 집합에 할당된 "그룹 키"로의 상호 참조만 관리되는 경우도 있습니다. 이 집합의 레코드를 결합해 해당 고객, 제품 또는 부품에 대해 더 우수한 뷰를 구성하곤 합니다. 개별 사실 정보는 가장 빈번히 발생하는 값, 가장 최신 버전인 값 또는 특정 소스에서 얻어진 값을 기준으로 선택할 수 있습니다. 표준화 및 매칭 규칙과 마찬가지로, 서바이버십 규칙은 다양한 업무 목적에 맞게 유연하게 적용될 수 있으며, 매칭 프로세스의 산출물을 바탕으로 한번에 하나의 트랜잭션 또는 배치에 적용될 수 있습니다.

## 변환

MDM에서는 배치 과정으로 또는 트랜잭션별로 소스 시스템으로부터 많은 양의 데이터를 추출하고 이를 공통 포맷으로 변환시키는 경우가 많습니다. 이론상으로는, 이러한 규칙을 한번만 정의하고 두 시나리오 모두에서 일관성 있게 적용해야 합니다. IBM의 WebSphere DataStage®는 다수의 프로젝트에 걸쳐 전체 및 실시간 시나리오에서 사용될 수 있도록, 변환 및 전달 규칙을 정의합니다. 예를 들어, WebSphere DataStage는 매우 많은 양의 벌크 데이터 세트를 이동해야 하는 프로세싱 집약적인 요구 사항과 데이터 웨어하우스 또는 MDM 시스템과 같은 다른 전사적 대상 시스템에 로드하기 위해 데이터의 분석적 뷰를 생성해야 하는 복잡한 트랜잭션 요구 사항을 지원합니다.

## 동기화

동기화 과정은 마스터 데이터 스토어 및 모든 소스 시스템 전반에 걸쳐 일관성을 유지하는 작업입니다. 모든 소스 및 마스터 데이터 저장소 내의 변동 사항을 기록하여야 하므로, 이는 매우 복잡한 과정입니다. 변동이 발생하면, (보통 데이터 품질 솔루션에 의해 제어되는) 서바이버십 규칙에 따라 모든 관련 시스템에 이러한 변동 사항을 복제하여야 합니다. 동기화 과정이 이루어지는 동안에는 트랜잭션의 무결성 등과 같은 요소가 매우 중요합니다. 시스템을 업데이트할 경우, 모순이 발생하지 않도록 모든 시스템을 최대한 신속하게 업데이트해야 합니다. 업데이트가 실패하면 나머지 시스템은 롤백해야 합니다.

이러한 수준의 프로세스 동기화에서는 어떤 유형의 소스 시스템과 통합이 가능하고 트랜잭션의 무결성을 확보할 수 있는 IBM WebSphere DataStage TX 및 IBM WebSphere Business Integration과 같은 전문적 기술이 필요합니다. 동기화 과정은 메시지 버스 및 프로세스 이벤트 아키텍처와 긴밀하게 연결되어 있습니다. 따라서 MDI 계층에서 지원하는 MDM 시스템의 프로세스를 어느 지점에서 기존 애플리케이션 간 데이터 흐름으로 분할해야 하는지 면밀하게 검토할 필요가 있습니다. 실제로 MDM 시스템은 또 다른 데이터 게시자 및 사용자가 됩니다.

### 연합

MDM 시스템을 전개하면 다양한 후속 프로젝트로 이어지는 경우가 많습니다. 신뢰성 높은 참조 데이터 소스가 구축되면, MDM 전개는 다양한 새 애플리케이션을 서비스하는 기반이 될 수 있습니다. 한 예로, 제품의 경우 전자상거래 사이트 및 인쇄판 카탈로그 제작을 지원하게끔 MDM 솔루션의 가치를 확대하겠다고 결정할 수 있습니다. 중앙 집중식 제품 정보 저장소를 확보하는 것이 당연한 출발점이 되는 이유를 쉽게 알 수 있습니다.

그러나 어떤 경우에서든 이러한 솔루션은 제품 사진, 실시간 공급 정보, '고객이 구입한 다른 품목' 과 같은 친화성 데이터 등 그 밖의 다양한 온 디맨드 정보를 필요로 합니다. MDM 시스템 외부에서 이러한 정보에 액세스하는 권한은 연합(federation) 작업을 통해 관리하고 제공할 수 있습니다. 연합 과정에서는 애플리케이션, 메인프레임 데이터베이스(예: VSAM, IMSTM, DB2® Universal Database™), 콘텐츠 소스(예: IBM DB2 Content Manager, FileNet, Documentum)를 망라해 다양한 정보 소스를 동시 쿼리하여 연관된 정보를 반환하는 메커니즘을 제공합니다. 그 예로, 데이터 웨어하우스로부터 로버트 리치라는 사람의 계약 내역을 입수하고 지원 센터 전화 내역을 조회하며 계약직 HR 시스템에서 이 사람의 사진을 검색하는 작업을 들 수 있습니다.

이러한 모든 작업이 완전한 마스터 데이터 관리 솔루션 지원에 필요한 데이터 통합 기능에 속합니다.



마스터 데이터를 중심으로 데이터를 통합하는 것은 다양한 전사적 이니셔티브를 지원하기 위해 갖춰야 할 핵심적인 역량 중 하나입니다. 이 "마스터 데이터 통합 서비스"는 기본 인프라의 상위에 존재하면서 "마스터 데이터 관리 시스템"을 지원하는 데이터 통합 계층이라고 할 수 있습니다.

### 3장: 마스터 데이터 관리 경로

마스터 데이터 관리는 전사적 범위에서 수평적 통합을 이루는 진정한 전사적 차원의 이니셔티브입니다. MDM을 성공으로 이끄는 데 큰 어려움을 던져주는 요소 중 상당수는 기술이 아니라 프로세스 및 구성과 관련된 경우가 많습니다. 사실 MDM은 효율성, 유효성, ROI, 경쟁적 차별성과 관련된 핵심적인 목표를 추구하는 데 있어 프로세스, 애플리케이션 및 기반 기술을 통합하는 여정의 하나로 도입되고 있습니다. MDM "여정"에서는 방법론을 수용하고 전개 및 운영을 위한 아키텍처 차원의 대안을 검토해야 합니다.

### 데이터 통합 방법론의 중요한 역할

모든 데이터 통합 프로젝트, 특히 MDM 관련 프로젝트 성공의 가장 중요한 요소 중 하나는 바로 방법론입니다. 방법론은 다음과 같이 구성됩니다.

- **프로세스** – 데이터 통합 솔루션을 구축하고 구현하는 데 필요한 단계(및 그 단계 간의 관계)
- **기술** – 데이터를 다루는 각 프로세스 단계에 적용되는 기능적인 구성요소
- **리소스** – 뚜렷하게 정의된 역할 및 책임과 함께 기술을 적용하고 프로세스 단계를 이행하는 데 필요한 기술력
- **산출물** – 데이터 획득, 변환 및 공급을 자동화하는 일련의 프로세스 단계에서 생성되는 결과물

### IBM MDI 방법론 - 성공적인 MDM을 위한 필수 요건



그림 5: IBM 마스터 데이터 통합 방법론

방법론을 적용하면 마스터 데이터 통합을 성공적으로 실현할 수 있습니다! IBM의 Information Integration Solutions Iterations® 방법론은 데이터 통합 인프라 프로젝트를 위한 프레임워크로 검증받았으며, 500여 건의 프로젝트에서 IBM 컨설턴트와 파트너가 쌓은 지식을 바탕으로 합니다. 포괄적이고 단계별로 구성된 로드맵을 통해 초기 계획 및 전략을 시작으로 기술적 구현 및 운영 환경 전개까지 모든 과정을 안내합니다.

IBM Iterations® 방법론은 마스터 데이터 통합 서비스를 설계하고 관리하는 다음과 같은 상위 수준의 프로세스 단계로 구성됩니다.

- **역량 평가 & 청사진 단계**에서 프로젝트는 비즈니스/IT 인프라의 현 상태를 기반으로 하고, 가장 중요한 기회에 대한 이해를 반영하며 최우선적인 비전을 가진 로드맵의 맥락에서 구성됩니다.
- **데이터 평가 및 검증 단계**에서는 마스터 데이터의 현 상태를 기술 및 비즈니스 규칙 관점에서, 소스 단위로 파악합니다.
- **데이터 정렬 단계**에서는 초기 로드, 배치 변경, 트랜잭션 및 실시간 업데이트를 위해 소스별로 마스터 데이터 수집, 표준화, 검증 및 매핑에 필요한 처리 절차를 정의합니다.

- **데이터 조율 단계**에서는 초기 로드, 배치 변경, 트랜잭션 및 실시간 업데이트를 위해 여러 소스를 망라해 마스터 데이터에 대한 중복 제거, 참조 소스로부터의 보강, 레코드 연계, 통합, 서바이버십(survivorship), 변환, 취합 및 매핑 규칙을 정의합니다.
- **데이터 무결성 및 평가 단계**에서는 운영 설정 상태에서 마스터 데이터의 프로세스, 측정, 유지 관리 메커니즘 및 보고 방식을 정의하고, 획득한 이점 및 학습한 교훈을 점검해 프로젝트 루프를 종결합니다.

방법론은 비즈니스 프로세스 관점에서도 중요한 의미를 갖습니다. 즉 전체 솔루션이 운영 환경에서 효과를 발휘할 수 있도록 데이터 통합 노력과 병행해서 활용하는 비즈니스 프로세스를 최적화하기에 알맞은 기회가 되곤 합니다.

그림 6에서 보여주는 것처럼, 마스터 데이터 통합과 관련된 기능적 역량 모두를 이 방법론에서 사용합니다

### 단계별 IBM MDI 기능

역량 범주	요구사항		전개			운영	
	1단계 역량 평가	2단계 청사진	3단계 데이터 평가	4단계 데이터 정렬	5단계 조율	6단계 데이터무결성	7단계 평가
연합			✓		✓	✓	
동기화						✓	
데이터 품질 모니터						✓	✓
변환 & 보강				✓	✓	✓	
매칭&표준화				✓	✓	✓	
프로파일링&규명	✓	✓	✓	✓	✓	✓	
연계&투입		✓	✓	✓	✓	✓	
메타데이터 관리				✓	✓	✓	

그림 6: 단계별 마스터 데이터 통합 역량

IBM은 마스터 데이터 관리 솔루션을 구현하고 운영할 수 있도록 가장 폭 넓은 정보 통합 제품을 공급합니다.

### IBM의 마스터 데이터 통합 제품

역량 범주	요구사항		전개			운영	
	1단계 역량 평가	2단계 청사진	3단계 데이터 평가	4단계 데이터 정렬	5단계 조율	6단계 데이터무결성	7단계 평가
연합	WebSphere Information Integrator Standard and Content Editions						
동기화	WebSphere Business Integrator, DataStage TX, II Replication Edition						
데이터 품질 모니터	WebSphere AuditStage						
변환 & 보강	WebSphere DataStage and QualityStage						
매칭&표준화	WebSphere QualityStage						
프로파일링&규명	WebSphere ProfileStage						
연계&투입	WebSphere DataStage						
메타데이터 관리	WebSphere MetaStage						

그림 7: IBM의 마스터 데이터 통합 제품

완벽한 마스터 데이터 관리 솔루션은 각기 다른 수많은 구성 요소, 몇 가지 기본 논리적 설계 그리고 다양한 운영 아키텍처로 이루어져 있습니다. 어떤 경우에서든, 마스터 데이터의 본질을 고려한다면 데이터 통합 방법론을 사용해 데이터 통합 인프라를 설정하는 것이 중요합니다.

### 마스터 데이터 관리 "스타일"

적극적인 마스터 데이터 관리의 범위는 추출된 데이터를 서비스 부서에 보내거나 심지어 웨어하우스에서 "오프라인"으로 마스터 데이터를 유지 관리하는 것에 그치지 않습니다. 완벽한 MDM 솔루션은 운영적으로 통합되어 있고, 애플리케이션 및 업무 프로세스에 논리적으로 올바른 마스터 데이터 뷰를 실시간 제공합니다. 세 가지 기본적인 전개 스타일이 존재합니다.

- 상호 참조(Cross-Reference)
- 기본 골드 복제(Light Gold Copy)
- 확장 골드 복제(Extended Gold Copy)

## 상호 참조

상호 참조 방식에서는 신뢰성 있는 마스터 데이터 정보를 위해 경량급 내부 저장소를 제공합니다. Gartner에서는 이를 고객 데이터 통합을 위한 "레지스트리 스타일"이라고 부릅니다<sup>2</sup>. 상호 참조에서는 데이터 품질 기술을 활용해 시스템 내부 및 여러 시스템 간에 일치하는 고객 레코드를 고유하게 식별합니다. 상호 참조 데이터베이스에서는 각 고유 레코드에 기본 키를 할당하고, 모든 소스 시스템에서 연계된 모든 레코드의 식별자를 저장합니다. 뿐만 아니라 고객/위치/제품에 대한 식별 데이터를 충분히 저장해 놓기 때문에 불완전한 레코드를 수신하거나 기본 키가 알려지지 않은 경우 식별자를 매칭할 수 있습니다. 소스 시스템에 대해 업데이트가 이루어지므로, 매칭 서비스가 데이터 중복을 방지합니다. 인바운드 레코드가 일치하는 항목으로 식별되면, 연계 정보 및 서바이버십 규칙을 바탕으로 소스 시스템으로부터 기존 고객 데이터가 즉시 조합될 수 있습니다. 또한 상호 참조 데이터베이스에 영향을 미치는 모든 업데이트는 거의 실시간으로 포착해 실행할 수 있습니다.

이러한 상호 참조를 생성하는 것은 MDM에서 가장 까다로운 일 중 하나입니다. 소스 시스템에 대한 심도 있는 이해 그리고 비즈니스 중심적인 매칭/서바이버십 규칙 구현이 필요합니다. 데이터 로드에서 사용되는 엔진은 모든 소스 시스템을 망라해 많은 양의 데이터를 처리할 수 있어야 합니다. 어떤 소스의 컨텍스트도 잃거나 대체하는 일 없이 각기 다른 소스 시스템 내부에 존재하는 레코드들이 어떻게 상호 연관되는지 상세하게 규명합니다. 또한 소스와 데이터 처리 간의 메타데이터 연계성을 유지할 수 있어야 합니다.

## “기본” 골드 복제

"골드 복제"란 고객과 같은 마스터 데이터에 대해 인정된 "공식" 버전을 의미합니다. "기본" 버전의 골드 복제에서는 서로 다른 고객을 구별하는 것처럼 어떤 엔티티의 인스턴스를 차별화하는 데 필요한 모든 속성을 포함합니다. 어떤 개별 고객을 고유하게 정의하는 40개의 요소가 있다고 결정한 경우(이름 구성 요소, 위치, 전화 번호, 이메일, 주민등록번호 등), 골드 복제는 그러한 모든 요소의 "레코드 시스템"이 됩니다. 그렇지 않으면 상호 참조와 정확하게 동일한 범위를 갖습니다.

<sup>2</sup> Radcliffe, J. 2004. "Learn the Four Styles of Customer Data Integration," Gartner, 10월 6일

## 확장 골드 복제

이 스타일에서 MDM 스키마는 상호 참조/기본 골드 복제의 속성 및 관계 모두를 비롯해 해당 엔티티에 대해 알고 있는 모든 사항을 포함합니다. 고객이라면 고객 데이터 통합(CDI)이라고도 부르는데, 해당 고객에 대한 구매/판촉/서비스/지급 이력 전체, 고객 선호 사항 그리고 BI 애플리케이션에서 얻은 통찰력이 발휘된 단계별 오퍼링 등을 포함할 수 있습니다. IBM WebSphere Customer Center는 고객, 직원, 조직 등 임의의 "관계자"에 대한 확장 골드 복제의 예입니다. Gartner는 이를 "트랜잭션 허브"라고 부르는데, CDI를 가장 포괄적으로 구현한 것입니다.

이 방식은 개념적으로는 상호 참조 방식과 비슷하지만, 단순히 상호 참조를 저장하기 보다는 각 고객에 대해 완전한 고유 레코드를 저장합니다. 각기 다른 소스 시스템으로부터 최상의 데이터 복사본을 생성하고, 이 복사본을 운영 데이터 저장소에 보관합니다. 이렇게 보관된 데이터는 신뢰성 있는 고객 정보 소스로 활용될 수 있습니다. 상호 참조 방식과 유사하게, 시스템 내부 및 여러 시스템 간에 레코드를 고유 식별, 매칭하고 연계하는 데이터 품질 기술을 활용합니다. 뿐만 아니라 다양한 소스를 망라해 얻은 서바이버십 규칙을 기준으로 하나의 통합 레코드를 저장할 수도 있습니다. 운영 환경에서 효과를 거두기 위해서는 이 마스터 데이터 저장소를 소스 시스템과 계속 동기화시켜야 합니다. 이 동기화가 보장되려면 프로세스가 제 위치에 있어야 합니다. 이 프로세스는 개별 소스 시스템의 업데이트에 의해 트리거되며, 선택적으로는 마스터 데이터베이스 업데이트에서 트리거될 수도 있습니다.

확장 골드 복제 방식에서는 모든 비즈니스 요구 사항을 나타내기에 적합한, 완벽한 고객 데이터 모델이 마련되었음을 전제로 합니다. 이 데이터 모델은 소스 시스템 스키마에서 유도되어 이 스키마의 "수퍼 세트"가 되거나, 산업별 또는 솔루션별 모델이 될 수도 있습니다. 예를 들어, IBM WebSphere Customer Center 또는 WebSphere Product Center 솔루션을 구현할 경우, 이 모델이 솔루션의 일부로 제공되고 소스는 이 모델에 간단히 매핑됩니다. 데이터 웨어하우스와 같은 다른 분석 기능이 여전히 이 모델에 도움이 될 가능성도 있으므로, 그러한 데이터와의 외부 연계가 계속 존재할 수도 있습니다.

확장된 골드 복제는 마스터 고객 데이터 문제를 해결할 가장 완벽한 솔루션입니다. 실제로 Gartner는 2008년까지 이기종 IT 환경에서 마스터 고객 정보 데이터베이스를 구축함으로써 여러 채널 및 사업부를 망라해 고객에 대해 가장 정확하고 최신 버전이며 완전한 뷰를 마련할 수 있다고 예상합니다<sup>3</sup>. 그러나 이 방식을 완벽하게 이행하는 것은 매우 힘든 과제가 될 수 있습니다. 동기화 및 관리 프로세스를 마련하는 것은 가장 까다로운 작업 중 하나이기도 합니다.

어떤 벤더가 공급한 패키지 솔루션을 사용할 경우에도 구현에 앞서 여러 시스템에 걸쳐 마스터 고객 데이터 집합을 합리화하는 것이 중요합니다. META Group에 따르면, 이러한 패키지 솔루션의 기능은 마스터 고객 데이터 객체를 신디케이트하고 관리하는 것이므로 마스터 데이터 합리화의 필요성이 강조됩니다.

### MDM 전개 프레임워크

마스터 데이터 엔티티 및 운영 스타일과 관계 없이, 전개 과정은 항상 4가지 단계로 구성됩니다. 또한 이는 다양한 전사적 이니셔티브에서 사용되는 프로세스 및 데이터 통합 구성 요소이기도 합니다.

### MDM 전개

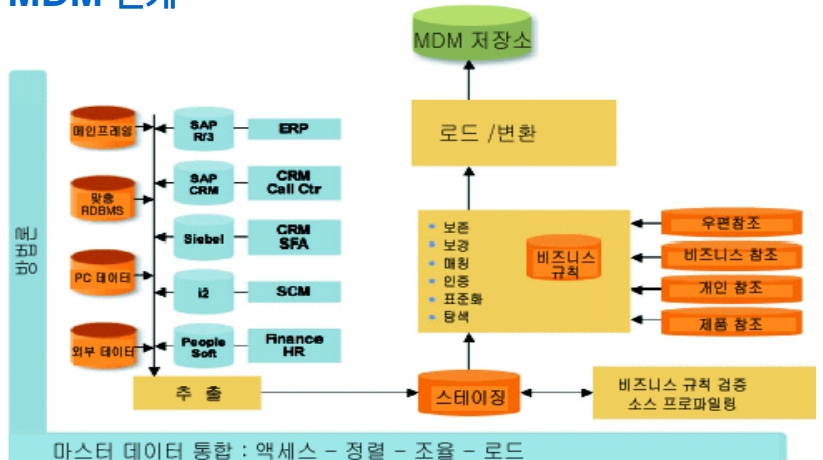


그림 8: MDM 전개

<sup>3</sup> Friedman, T., S. Nelson 및 J. Radcliffe. 2004. "CRM Demands Data Cleansing," Gartner, 12월 3일

<sup>4</sup> Hudson, Bruce. 2004. "ERP Consolidation and SAP's MDM Highlight the Need for Data Rationalization," META Group, 9월 21일

## 1. 스테이징 영역으로 데이터 추출

먼저 평가 및 후속 변환 작업을 위한 "안전한 항구"로 데이터를 추출해야 합니다.

## 2. 개별 소스 프로파일링 & 정렬

앞서 설명한 대로, 대상 및 변환 요구 사항을 확고하고(데이터 값 자체에 대한 면밀한 검사를 통해서만 드러나는)비즈니스 프랙티스를 규명할 수 있도록 먼저 프로파일링을 통해 소스 시스템에 대한 이해 기반을 마련하는 것이 중요합니다. 정렬은 각 소스 시스템을 한번에 하나씩 재조정하는 수단을 제공합니다. 애플리케이션 데이터의 품질을 높이도록 소스 시스템 애플리케이션, 프로세스 흐름 및 책임 사항의 변경 사항을 살펴볼 기회가 마련되곤 합니다.

## 3. 표준화, 검증, 매칭, 보강 및 보존

이제 데이터를 변환할 수 있습니다. 즉 데이터 값을 시작으로 리엔지니어링하여, 마스터 데이터 저장소를 채우는 데 필요한 고품질의 사실 정보, 논리적 키, 관계 및 계층 구조를 만들어냅니다. 프로젝트 개발 과정에서 규칙 세트를 확정하며, 이 규칙 세트는 마스터 데이터 솔루션에서 필요한 데이터 구조, 내용 및 품질을 생성합니다. 이러한 비즈니스 중심적인 비즈니스 규칙이 운영 환경에서 높은 수준의 데이터 품질을 유지할 수 있는 기반을 조성합니다.

## 4. 변환 및 MDM 저장소 로드

마지막으로 고품질의 데이터가 마스터 데이터 저장소로 이동합니다. 이 데이터는 변환을 거쳐 대상의 구체적인 요구 사항에 따라 형식 지정됩니다. 저장소에서 직접 관리하는 구조로 데이터가 로드되는 경우가 많습니다. 그렇지 않으면 변환 프로세스의 산출물이 XML 형식을 띠므로, 마스터 데이터 관리 솔루션의 비즈니스 서비스 계층에서 그 로드를 입력 트랜잭션으로 처리하고 직접 기본 스키마에 기록할 수 있습니다.

## MDM 운영 프레임워크

이 세 가지 전개 스타일에서 마스터 데이터를 만들고 그 운영에 필요한 정보 관리를 수행하려면 인라인 정제와 동기화라는 두 가지 추가 구성 요소가 필요합니다.



## MDM 운영 프레임워크

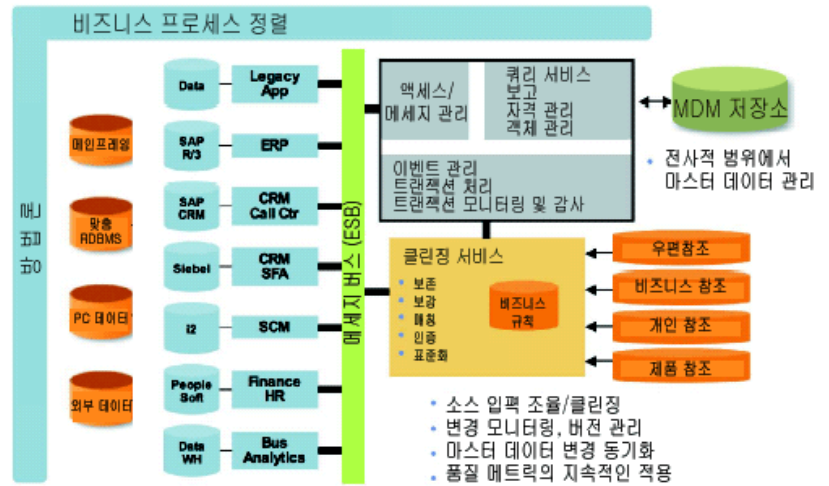


그림 9: MDM 운영 프레임워크

### 클린징 및 매칭 서비스

매칭 서비스에서는 데이터의 지속적인 무결성을 보장합니다. 이 서비스에서는 데이터 품질 기술에서 규정된 규칙을 도입해, 마스터 데이터를 생성하는 어떤 애플리케이션에서든 즉시 호출할 수 있게 만듭니다. 따라서 시스템에 입력되는 새 데이터는 중복 레코드를 생성하지 않으며, 기존 레코드와 적절히 연계됩니다. 느슨하게 결합된

SOA(service oriented architecture)는 동일한 클린징 서비스를 마스터 데이터 관리 시스템에서 또는(선택적으로)MDM 시스템과 관계 없는 다른 애플리케이션에서 직접 이용할 수 있다는 추가적인 이점이 있습니다. 예를 들어, CDI의 관계자 매칭 로직은 마케팅 애플리케이션에서도 사용할 수 있습니다. 인바운드 중복 레코드에 존재하는 새 정보를 마스터 레코드에서 "보존"시킬 수도 있습니다. 예를 들어, 기존 고객에 대해 새 주소 데이터를 수신할 수 있습니다. 중복 레코드를 만들지 않겠지만, 새 주소를 잃을 수도 없는 일입니다.

### 동기화 서비스

동기화는 마스터 데이터 저장소 및 모든 소스 시스템 간에 일관성을 유지하는 과정입니다. 동기화는 모든 소스 및 마스터 데이터 저장소의 변경 사항을 추적하므로 매우 복잡한 프로세스입니다.

변경 사항이 발생하면 서바이버십 규칙(주로 데이터 품질 솔루션에서 제어)에 따라 해당되는 모든 시스템에 복제되어야 합니다. 동기화 과정에서 트랜잭션 무결성이 매우 중요합니다. 어떤 시스템이 업데이트되면 가급적 빨리 모든 시스템에 그 업데이트가 적용됨으로써 불일치를 방지하는 것이 중요하며, 업데이트가 실패할 경우 나머지 시스템은 롤백되어야 합니다. 이러한 수준의 프로세스 동기화에서는 어떤 유형의 소스 시스템과도 통합할 수 있고 트랜잭션 무결성을 보장할 수 있는 전문화된 기술이 필요합니다.

운영적 측면에서 새 소스 시스템과 배치 업데이트는 전개 모델과 유사한 방식으로 마스터 데이터 저장소에도 지속적으로 적용됩니다. 변환 및 로드 모델은 역방향으로 가동될 때가 많으므로, MDM 참조 데이터를 BI, 사기 행위 탐지 시스템과 같은 다른 솔루션으로 전달할 수 있습니다.

## 4장: 베스트 프랙티스, 이점 및 성공 사례

### 베스트 프랙티스

IBM은 마스터 데이터 관리 프로젝트를 성공으로 이끄는 데 도움이 될 베스트 프랙티스를 제안합니다.

- **상세한 비즈니스 사례**를 구축하고, 프로젝트 착수에 앞서 가치 분석을 완료해야 합니다. 이러한 프로젝트에는 상당한 시간과 리소스 투자가 필요하며, 많은 수의 부서 및 그룹이 참여하므로 이 요건을 반드시 준수해야 합니다.
- **비즈니스 목표와 프로젝트 이니셔티브를 매칭**시키고, 이 목표를 임원진에서 적극 지원해야 합니다. 프로젝트가 회사 차원의 전략을 뒷받침하는지 확인하고, 임원진이 여러 관계 부서 및 기능 분야가 조율할 수 있는 인센티브를 마련하도록 장려합니다.
- **측정 가능한 ROI**로 신속하게 성과를 낼 수 있도록 프로젝트를 설계합니다. 가급적 프로토타입 및 PoC(proof of concept)를 활용해 긴 제품 주기를 방지하고 프로세스 전체에 걸쳐 현업 부서가 계속 참여하게 합니다.

- **"마스터 데이터" 인터랙션 생명 주기를 이해합니다.**  
마스터 데이터와 관련된 인력/프로세스 생명 주기 전체를 점검함으로써 비즈니스 요구 사항을 더 정확하게 파악하고, 올바른 데이터를 확보하도록 보장하며 프로젝트를 더 효과적으로 지원할 수 있습니다.
- **가능한 영역에서 데이터 관리(stewardship) 프로세스를 구현합니다.**  
이를 통해 데이터 경영 및 감독이 가능하고, 그룹 간의 충돌을 해결할 수 있으며, 데이터 소유권은 업무 책임 중 하나로 확정할 수 있습니다.
- **복잡한 데이터 아키텍처 및 통합 기술 프로젝트를 피합니다.**  
완전히 하향식으로 부분적인 범위를 제공하면서 각 릴리스에 중점을 둡니다. 그러면 각 프로젝트가 진행되면서 아키텍처가 발전하고, 그와 동시에 비즈니스 목표도 달성할 수 있습니다.
- **성공 및 측정 기준을 처음부터 정의해두어야 합니다.**  
성공 측정 기준에 대한 합의를 이끌고 측정 기준을 마련한다면 그 가치를 검증할 수 있습니다.
- **검증된 방식과 기술을 선택합니다.**  
이러한 프로젝트가 성공하기 위해서는 경험이 필수적입니다.

## 이점

Gartner에 따르면, 여러 채널과 현업 부서를 망라해 고객에 대해 정확하고 시기적절하며 강력한 단일 뷰를 구축하는 것은 고객 중심의 조직에서 비용을 줄이고 리스크를 관리하며 매출 및 이익을 높이는 데 핵심적인 역할을 합니다<sup>5</sup>. 마스터 데이터 이니셔티브를 이행하는 기업은 상당한 ROI를 확인했습니다.

고객 보존율이 5% 증가하면 개별 고객으로부터 얻는 평생 수익이 75% 증가하는 것으로 밝혀졌습니다<sup>6</sup>. 고객 보존율과 충성도는 이 방식에서 얻을 수 있는 여러 이점 중 일부에 불과합니다. 그 밖에도 다음과 같은 효과를 누릴 수 있습니다.

<sup>5</sup> Radcliffe, J. 2004. "Create a Single Customer View with Customer Data Integration," Gartner, 10월 7일

<sup>6</sup> Reichheld, Frederick F. and Thomas Teal. 1996. The Loyalty Effect. Boston: Harvard Business School Press.

- **출시 일정 단축** - MDM 환경에서는 하나의 시스템을 통해 제품 정보 생성 및 관리, 판촉, 온라인/기존 채널을 통한 강력하고 정확한 소비자와의 의사 소통이 이루어집니다. MDM을 도입한 한 대형 소매업체는 신상품 출시 일정을 4주에서 1일로 단축했습니다.
- **공급망 개선** - MDM에 의해 정확하고 합의된 제품/공급자 단일 정의가 이루어지므로 중복을 막고 구매력을 높이며 공급자 관계에 대한 통찰력을 높일 수 있습니다. 유럽의 한 대표적인 소매업체는 MDM과 WebSphere Portal을 결합해 3만 여 곳의 공급자들이 자체 마스터 정보를 관리할 수 있게 하는 중입니다.
- **수익 증가** - MDM은 고객 관계 개선에 기여합니다. 고객에 대해 360도 전방위적인 뷰를 마련함으로써 MDM은 영업/마케팅/서비스 조직이 고객의 요구 사항을 정확하게 예측하고 타겟화된 오퍼링을 내놓으며 고객 서비스를 개선할 수 있도록 뒷받칩니다.
- **통합 향상** - IT 부서에서는 MDM을 통해 전사적으로 생성되어 온 정보 사일로를 해소하고 통합 비용을 절약하며 협업을 증진시키고 비즈니스 생산성을 높일 수 있습니다. MDM을 도입한 유럽의 한 제조사는 200여 개의 SAP 및 기타 ERP 시스템 인스턴스를 망라해 정확하고 강력한 마스터 정보를 클린징하고 동기화하고 있습니다.
- **산업 규정 및 정부 법률 준수** - IBM은 미국의 사베인 옥슬리법, 애국법, Basel II, ACORD, HIPAA와 같은 정보 중심적인 규정을 비롯해 광범위한 업종에 걸쳐 글로벌 데이터 동기화 요건을 준수하는 데 도움이 될 솔루션과 수준 높은 전문 지식을 제공합니다.

뿐만 아니라 MDM은 모든 애플리케이션을 망라해 정보 중심적인 절차를 뒷받침할 수 있도록 공통적인 비즈니스 서비스를 제공합니다. 거시적/미시적 비즈니스 프로세스 기능을 통해 비즈니스 애플리케이션과 통합하거나 이를 구동합니다. 신제품 출시를 준비하거나 신입 직원을 모집하거나 공급자 수를 줄이는 등 어떤 경우에서든 마스터 정보 관리와 관련된 비즈니스 프로세스가 존재합니다. MDM 시스템은 이러한 프로세스에 참여하면서 프로세스 전체를 이끌거나(거시적) 다른 시스템으로부터 호출될 수 있습니다(미시적).

거시적 사례의 예를 소개합니다. MDM이 등장하기 전, 각기 다른 여러 사내 애플리케이션에 제품 정보를 입력하고 "새로운 제품 설정 프로세스"를 거쳐야 하는 기업들이 많았습니다(소매 분야의 경우, 이러한 애플리케이션은 30여 종이 넘으며, 제조사에서는 수백 여 종에 이르기 도 함). IBM의 MDM 방식에서는 고객이 단일 복합 애플리케이션을 개발함으로써 하나의 프로세스를 통해 모든 신상품 정보를 한 곳으로 수집하고, 그 단일 마스터 프로세스를 적용해 정보의 완전성과 유효성을 보장할 수 있습니다. 이 작업을 마치면 MDM 소프트웨어는 자동으로 데이터를 전송하고 이 30여 종의 애플리케이션 각각에서 새 항목을 설정하는 데 필요한 개별 업무 프로세스를 자동화할 수 있습니다.

미시적 사례는 더 간단합니다. 은행의 한 사업부에서 어떤 CRM 시스템이 새 고객을 설정하는 워크플로우를 갖습니다. 이 CRM 시스템은 고객 정보와 인터랙션하는 다른 시스템이 존재하는지 모르기 때문에 간단히 MDM으로 프로세스를 호출하며, 이 프로세스에서 고객 서비스 담당자가 해당 고객이 이 은행과 관계를 맺고 있는지 확인할 수 있습니다.

MDM이 제공하는 또 다른 이점은 정확한 BI가 이루어진다는 것입니다. 지금까지 데이터 웨어하우스 이니셔티브에서는 애플리케이션부터 다운스트림 방향으로 데이터 품질 문제를 해결하려고 했습니다. 데이터 웨어하우스에서 잘못된 마스터 데이터가 생성되는 업무 프로세스를 바로잡도록 시도하거나 애플리케이션의 마스터 데이터를 수정하려고 시도하지 않습니다. MDM 환경에서는 잘못된 데이터 및 잘못된 데이터를 양산하는 프로세스를 소스 애플리케이션에서 수정할 수 있습니다. 마스터 데이터 관리에서는 정확한 객체/구조가 데이터 웨어하우스 및 분석 애플리케이션과 자동으로 동기화됩니다.

## 고객 성공 프로필

마스터 데이터를 효과적으로 관리하는 기업은 고객, 벤더 및 파트너 관계에서 더 큰 성과를 이끌어낼 수 있습니다. 기업이 성공을 거두기 위해서는 마스터 데이터 문제를 인식하고, 측정 가능한 구체적인 비즈니스 목표와 해결책을 연결하며, 기존 기술의 구성 요소 및 기능을 극대화할 방식을 선택해야 합니다.

검증된 방식과 기술을 선택하고 베스트 프랙티스를 채택해 이행한다면 매우 신속하게 ROI를 실현하고 단기적/전략적 비즈니스 목표를 달성할 수 있습니다.

아래에서는 IBM WebSphere 정보 통합 포트폴리오를 활용해 마스터 데이터 통합을 중심으로 상당한 효과를 거둔 기업들을 간략하게 소개합니다.

고객	프로젝트	비즈니스 추진 요소
JP Morgan Chase	비즈니스 성과 관리	고객 수익성
Rockwell Automation	업무용 고객 마스터	폐쇄 루프형 마케팅
Aetna	가입자 레코드 시스템	마케팅 및 고객 서비스
Scott	공급망 최적화	재고 비용 감소
DHL Worldwide Express	CRM 서비스 개시	합병 중심의 통합
Celanese	SAP 통합	영업 비용 감소, 가치 창출 시간 단축
Bombadier	단일 제품 뷰	재고 관리 효율성 강화
Outo Kumpu	글로벌 부품 마스터	중복 해소, 재고 관리 효율성 강화
Colonial Electric	단일 고객 뷰	통합, 가치 창출 시간 단축, 구현 비용 절감
GMAC Mortgage	고객/상품 상호 참조	크로스셀링
영국 국방성	통합 재고 품목 & 공급자	재고, 가치 창출 시간, 구현 시간 단축
Tyson's	글로벌 ERP 단일 뷰	인수, 레거시 통합, 비용 절감

IBM의 정보 통합 포트폴리오를 활용한 마스터 데이터 통합이 마스터 데이터 관리를 실현하는 데 어떻게 기여할 수 있는지 더 자세한 내용은 IBM에 문의해 주십시오.

## 자세한 정보

마스터 데이터 관리 및 마스터 데이터 통합에 대한 자세한 내용은 아래 링크를 참조하십시오.

IBM 마스터 데이터 통합: [ibm.ascential.com/solutions/master\\_data\\_management.html](http://ibm.ascential.com/solutions/master_data_management.html)

IBM WebSphere Product Center: [www.ibm.com/software/integration/wpc/](http://www.ibm.com/software/integration/wpc/)

IBM WebSphere Customer Center: [www.dwl.com/products/customer/](http://www.dwl.com/products/customer/)

## IBM WebSphere 정보 통합

WebSphere 정보 통합 솔루션은 어떤 데이터와 콘텐츠로 통합하고 변환시켜 중요한 비즈니스 이니셔티브에서 믿고 활용할 수 있는 정보를 공급하도록 포괄적인 통합 플랫폼을 제시합니다. WebSphere 정보 통합 플랫폼은 획기적인 생산성, 유연성 및 성능을 바탕으로 귀사와 귀사의 고객 및 파트너가 업무 운영 및 확장에 필요로 하는 정확한 정보를 전달합니다. 정보 이해, 클린징 및 향상에 기여할 뿐 아니라 품질 관리를 통해 궁극적으로는 신뢰할 만한 정보 보급을 지향합니다. 확장된 기업 환경 전체에 걸쳐 통합되고 필요한 시점에 공급되는, 이 일관성 있고 시기 적절하며 완전한 정보는 업무 프로세스를 강화하고 컨텍스트에 대한 핵심적인 통찰력을 실현하며 자신 있는 비즈니스 의사 결정을 뒷받침합니다.

## 자세한 정보

IBM 정보 통합 솔루션의 근간이 되는 기술 및 제품에 대한 자세한 내용은 IBM 마케팅 담당자 또는 비즈니스 파트너에게 문의하거나 [ibm.com/software/data/integration](http://ibm.com/software/data/integration)을 참조하십시오.



© Copyright IBM Corporation 2005  
(135-270) 서울시 강남구 도곡동 467-12  
군인공제회관빌딩

한국아이비엠주식회사

한국에서 인쇄  
2005년 10월  
All Rights Reserved

IBM, IBM 로고, On Demand Business 로고,  
DataStage, IMS, ProfileStage, QualityStage 및  
WebSphere는 미국 및 기타 국가에서 International  
Business Machines Corporation의 상표입니다.

그 밖의 회사, 제품 및 서비스 이름은 타 회사의 상표 또는 서비  
스 상표입니다.

본 자료에서 IBM 제품이나 서비스가 언급되었더라도 IBM이  
영업 중인 모든 국가에서 해당 제품이나 서비스가 공급되는 것  
을 의미하지는 않습니다.

오퍼링은 예고 없이 변경, 확장되거나 취소될 수 있습니다.

IBM의 향후 방향 또는 의도에 대한 내용은 예고 없이 변경되거  
나 취소될 수 있으며, 목표 및 목적만을 나타낸 것입니다.

소비 후 재생 섬유 10%가 함유된 재생 용지를 사용하여 미국  
에서 인쇄

IBM 인터넷 홈 페이지는 [ibm.com](http://ibm.com)입니다.

GC-18-9747-00