

# 웨어하우스 흥망성쇠 가름하는 열쇠

메타데이터의 저장소를 구축하는데 전적인 동의를 구하는 것은 의사결정 지원시스템(DSS)에 접근하고 그 시스템을 유지하는데 있어서 매우 중요한 사항이다. 메타데이터 없이 구축된 웨어하우스는 마치 자동차가 안 다니는 고속도로와 같은 형국이다. 우리들 모두가 이 개념을 이해하는데 왜 아무도 이것을 하지 않는 것일까?

이 용자들이 그들의 의사결정 지원시스템에 있는 보고서, 질의 및 데이터에 접근하기 위해 도입한 고성능의 메타데이터 저장소를 성공적으로 수행한 기업은 거의 없는 실정이다. 실제로 대부분 기업들의 메타데이터 저장소는 두 말할 것도 없이 실패작이다. 회사가 정말로 필요로 하는 실제적인 메타데이터에서 작고 부정확한 부분집합을 가지는 비구조적이며 접근성도 떨어지는 저장소가 대반이다.

이 기사를 통해 메타데이터 저장소의 정의, 요건, 제작, 제어 및 접근에 관해 전반적인 사항을 완벽히 언급하고자 한다.

메타데이터의 개념은 새로운 것이 아니다. 그러나 의사결정 지원시스템에서의 메타데이터의 역할과 중요성은 새로운 것임에 틀림없다. 우선, 비즈니스와 기술적 메타데이터의 차이점 및 용도에 대해 논의되어야 할 것이다. 둘째로 데이터 웨어하우스와 관련된 모든 메타데이터를 관리하는 일반적인 메타데이터 통합 전략을 얘기할 것이다.

이러한 전략은 다수의 플랫폼을 수용하는 메타데이터의 확산을 처리하며 저장소로 메타데이터를 옮기는 것에 가깝다고 정의할 수 있다. 끝으로, 메타데이터에 접근하는 것과 다름없음이 증명될 것이다.

## 기술 대 비즈니스 메타데이터

저장소에 있는 메타데이터에는 기술과 비즈니스의 두 유형이 있다. 기술적 메타데이터는 의사결정

이용자 보고와 질의 접근 유형, 빈도, 실행시간  
감사제어와 균형 정보  
데이터 웨어하우스를 만족시키는 레코드의 시스템  
데이터 웨어하우스에 대한 레코드의 시스템으로부터 매핑과 변환  
물리적, 논리적 데이터 모델  
데이터웨어하우스 테이블명, 요소, 색인  
데이터 웨어하우스 테이블 구조, 테이블 속성  
업무 의존도  
프로그래밍과 종류  
보안  
제거 기준

〈그림 1〉 기술적 메타데이터의 예

비즈니스 분석가에게 알려진 데이터 구조  
웨어하우스/시장에서 데이터 공통접근루틴  
테이블명과 비즈니스적 정의  
속성명과 비즈니스적 정의  
데이터 웨어하우스 필드 매핑, 변환, 요약  
데이터 웨어하우스 갱신일

〈그림 2〉 비즈니스 메타데이터의 예

지원시스템의 개발자들과 기술적 이용자들에게 데이터 웨어하우스의 데이터가 정확하다는 확신을 준다.〈그림 1〉 게다가, 기술적 메타데이터는 웨어하우스의 계속적 유지와 성장에 있어서 매우 중요하다. 기술적 메타데이터 없이 의사결정 지원시스템을 분석하고 변화를 시행하려면 일이 더욱 어려워지고 시간만 소모하게 될 것이다.

비즈니스 메타데이터는 데이터 웨어하우스와 비

즈니스 이용자들간의 연결점이다. 비즈니스 메타데이터는 이용자들에게 데이터 웨어하우스와 데이터 마트에 있는 데이터에 접근하기 위한 지도를 제공한다. 비즈니스 이용자들은 경영간부이거나 경영 분석가들인 경우가 대부분이며 기술적인 능력이 조금은 떨어질 수도 있다.

따라서 그들에게는 비즈니스 용어로 정의된 의사결정 지원시스템이 요구된다. 비즈니스 메타데이터는 데이터 웨어하우스 안에 있는 보고서, 질의, 데이터의 내용, 데이터의 위치, 데이터의 신뢰도, 데이터의 내용, 어떤 데이터 변환 규칙이 적용되었는지, 어떤 전송 시스템으로부터 데이터가 나왔는지를 비즈니스 용어로 보여준다.<그림 2>

### 메타데이터 고객

메타데이터 이용자들은 비즈니스 이용자부터 기술적 이용자에 이르기까지 광범위하다.<그림 3>

두 이용자 그룹 모두 데이터 웨어하우스 메타데이터의 다양한 이용자를 포함한다. 그들은 모두 의사결정 지원시스템에서 정보를 확인하고 효과적으로 사용하기 위해 메타데이터를 필요로 한다.

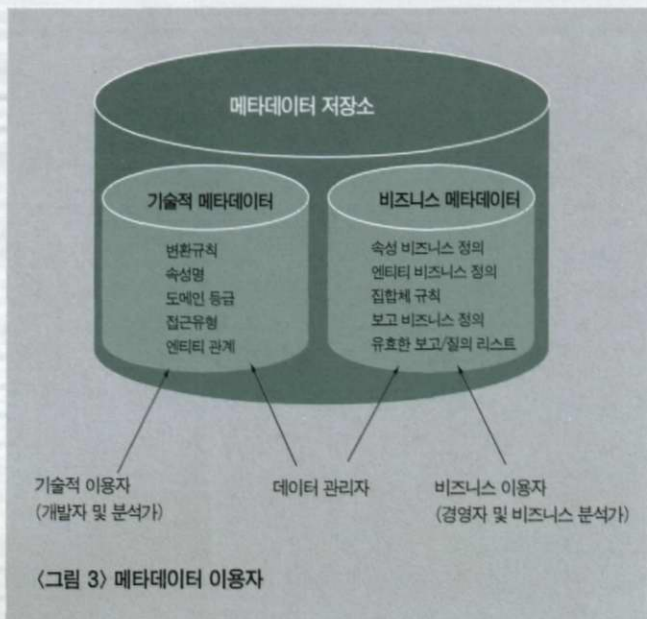
### 비즈니스 이용자

비즈니스 이용자들의 대부분은 그다지 기술적이지 못하다. 이러한 이용자들은 의사결정 지원시스템의 미리 정의된 질의와 보고서들로부터 정보를 얻는다. 이용자들이 요구하는 메타데이터는 주로 비즈니스명과 웨어하우스 테이블 및 칼럼의 정의, 기존 보고서의 기술과 관련이 있다.

이것은 비즈니스 이용자들이 데이터 웨어하우스에서 그들이 필요로 하는 정보를 확인하고 위치를 알아낼 수 있도록 한다. 또한 이러한 이용자들은 질의와 보고서 실행시간의 평가를 받는데 관심을 가지고 있다.

### 기술적 이용자

이 분야는 데이터 수집 개발자, 데이터 접근 개발자, 데이터 설계자, 상급 기술자 및 데이터 관리 직원 등 매우 다양한 이용자들을 가진다.



데이터 웨어하우스를 구축, 유지, 관리하는 시스템 지원인력은 비즈니스 이용자와 다른 경우의 메타데이터 필요성을 가진다. 그들은 프로그램이 어떤 방법으로 데이터를 데이터 웨어하우스와 데이터마트 작업에 확장, 변환, 로드시키는지 이해하기를 원한다.

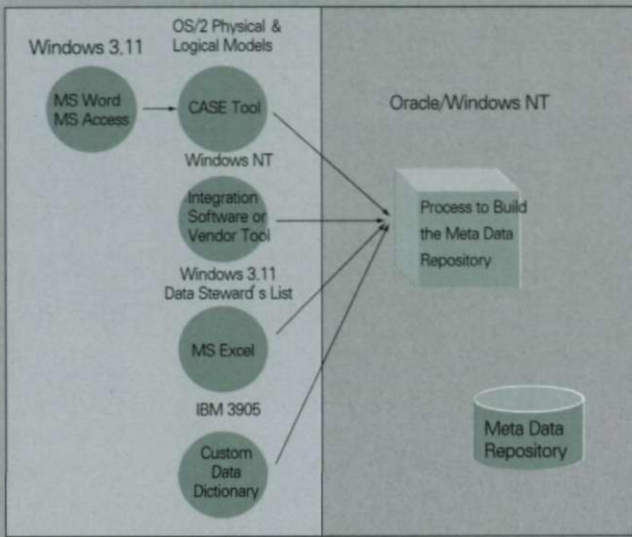
그들은 웨어하우스 변화가 어떤 프로그램, 파일, 테이블에 영향을 줄 것인지를 결정할 필요가 있다. 기술적인 메타데이터는 데이터 웨어하우스 담당직원이 의사결정 지원시스템의 부가적 개발 공개에 대해 보다 효과적이면서 정확하게 계획할 수 있도록 해준다. 데이터 관리자는 데이터 웨어하우스 이용자의 접근, 빈도 및 보고서/질의 실행시간 등의 유형을 알 필요가 있다.

이 메타데이터는 정보 습득을 위해 적용되는 제공자 접근 툴(예: Brio, Business Objects, MicroStrategy, Cognod, Seaate)에 의해 가장 손쉽게 변환된다. 메타데이터 저장소는 데이터 웨어하우스 담당직원이 의사결정 지원시스템을 효과적으로 유지하고 성장시킬 수 있도록 해주는 매우 중요하고도 생생한 정보를 제공한다.

### 메타데이터 통합 구조

오늘날 당신 회사의 모든 메타데이터 자원들을

메타데이터 통합구조



<그림 4> 표본 메타데이터 구조

하나의 통합 구조화된 저장소에 통째로 보관할 수 있는 틀이 있다면 대단한 일일 것이다. 그러나 그런 이상적인 경우는 존재하지 않는다. 오늘날의 메타데이터 통합구조는 통합된 제품을 의미하기보다는 껌, 줄과 종이 클립과 함께 서로 첨부되는 독립적인 제공자 제품 시리즈라고 할 수 있다.

지금 당신이 희망을 버리기 전에 Microsoft의 새로운 저장소 확장이 현 상황을 개선할 수 있는 희망을 준다. Microsoft 저장소는 다른 제공자들로부터 데이터 웨어하우징 제품들이 정보공유를 할 수 있도록 할 광범위한 저장소 확장 지원 장치이다.

확장은 다양한 데이터 웨어하우징 산업 지도자들의 광대한 합작의 산물이다. 이 새로운 저장소 확장은 모든 데이터 웨어하우스 틀 제공자들을 가로지르는 개방적이고 공통적인 하부구조를 제공함으로써 데이터 웨어하우징 시장을 확장할 수 있도록 도움을 줄 수 있어야 한다. 메타데이터 통합의 이러한 문제는 다수 조직들의 성공적인 데이터 웨어하우스와 데이터마트 수행을 방해한 중요한 요인들 가운데 하나이다.

제공자들이 이러한 표준을 받아들이고 이행할 때까지, 복잡한 통합 전략은 이렇게 다양한 원천을 가

지고 있는 메타데이터를 모으기 위해 필요할 것이다.

### 메타데이터 소스의 유형

메타데이터 소스에는 정형과 비정형의 두 가지 유형이 있다. 이러한 소스들은 조직에서 비즈니스와 기술적 메타데이터를 포함한다.

정형 메타데이터 소스는 기업의 의사결정자에 의해 논의되고 문서화되고 동의를 얻은 메타데이터의 소스들이다. 정형 메타데이터는 일반적으로 조직을 통해서 유지되고 배포되며 인가되는 틀이나 문서에 저장된다. 이러한 정형 데이터 소스는 기술적 메타데이터와 비즈니스 메타데이터 모두를 포함한다. (그림 5)

비정형 메타데이터는 공동의 지식, 정책 및 표준 형이 아닌 가이드라인으로 구성되어 있다. 이것이 사람들이 "그냥 아는" 정도의 정보이다. 이러한 정보 유형은 "회사의 지각도"에 있거나 중요한 종업원의 책상 위에 기록되어 있을 수 있다. 이것은 정식으로 문서화되거나 동의를 얻은 것은 아니지만 이러한 지식은 이전의 메타데이터 소스에서처럼 아주 가치가 있는 데이터들로 이루어졌다.

종종 비정형 메타데이터는 비즈니스와 관련이 있기 때문에 상당히 가치있는 정보를 제공하기도 한다. 비즈니스 메타데이터의 대부분이 비정형이라는 사실은 주의를 끝만큼 중요하다. 결과적으로 이 메타데이터가 변환되고, 문서화되며, 정형화되고 데이터 웨어하우스에 반영되는 것은 매우 중요하다.

이러한 과정을 통해서 당신은 메타데이터의 비정형 소스를 받아서 그것을 정형화된 소스로 변환한다. 각각의 기관이 다르기 때문에, 당신의 메타데이터 소스가 어디에 있는지를 확실히 얘기하는 것은 어려운 일이다. 그러나 다음은 비정형 메타데이터의 가장 일반적인 유형들의 목록이다.

- 데이터의 직위
- 비즈니스 역할
- 비즈니스 정의
- 경쟁사 제품 리스트
- 변형 및 요약

<그림 4>에서 데이터 관리자의 스프레드시트는

이 회사의 안에 존재하지 않는다. 그러나 대부분의 고용자들은 누가 유산 시스템 데이터에 대한 책임을 지는지를 알았다.

따라서 우리는 유형화된 이 리스트와 장소에서 중요한 관계를 얻었다. 이것은 이전의 정보원인 메타데이터가 정규의 정보원이 되는 전형적인 예이다.

## 대중을 위한 메타데이터

OLAP을 실행시키거나 데스크탑에 툴을 기록하고 그것을 수천 명의 이용자들에게 배포하는 것이 증가하는 최종 사용자 집단에 대한 비즈니스 정보의 잇점을 반드시 전달해 주는 것은 아니다. 새로운 경향은 인터넷이나 또는 더욱 일반적으로 인터넷 구조로 OLAP 기능을 웹으로 기록과 함께 이동시키는 것이다. <그림 6>

이것은 "클라이언트 중심"에서 "서버 중심"으로 구조를 변환시키며 비즈니스 이용자들이 웨어하우스와 데이터마트를 통해 메타데이터에 접근하기 위해 그들에게 친숙한 웹 브라우저를 사용할 수 있도록 해준다.

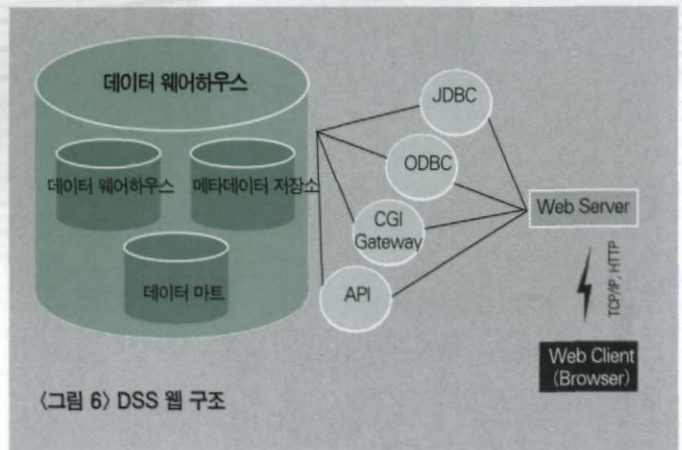
이 n-층 구조는 더 적은 수의 클라이언트를 위해 다수의 통신 서버들 가운데 분포된 논리와 함께 제공한다. 웹 클라이언트(브라우저)는 CGI 게이트웨이나 중간 층으로 작용하는 다른 인터넷 서비스를 통해 웹 서버에 접근할 수 있다. 웹에서 가능한 OLAP 툴은 데스크탑용 툴처럼 완벽하지 못하다. 그러나 대부분의 제공자들은 그들의 OLAP 툴을 가능한 빠른 속도로 맞추어서 작업하고 있다.

웹 액세스에는 두 가지의 넓은 기미가 있다. 정적 웹 액세스는 HTML의 정적 페이지를 만들기 위해서 표준 웹 브라우저(넷스케이프 또는 마이크로소프트 익스플로러)를 사용한다. 이러한 접근 방법은 매우 신뢰성을 가지며 비용을 절감시키고 실행하기도 쉽다. 역동적인 웹 액세스는 OLAP과 비즈니스 사용자 접근의 임시 질의 형태를 참고한다. 이러한 액세스 유형은 더욱 긴 응답 시간과 시장에서는 여전히 성숙해가는 더욱 복잡한 툴세트를 요구한다.

데이터 웨어하우스와 데이터마트 프로젝트는 그

정규 메타데이터 유형	통상적 기억장소
추출/변환 툴	프리즘 웨어하우스 관리자, VMARK, Datastage, D2K Tapestry, ETI · Extract, Apertus Carleton
논리적 물리적 데이터모델	케이스 툴(Cayenne, PowderDesigner, ERwin) and Vendor Applications(SAP, Peoplesoft)
비즈니스 정책	MS Word, MS Excel
데이터 사전	케이스CASE Tool, MS Access, MS Excel

<그림 5> 정규 메타데이터 정보원의 예



<그림 6> DSS 웹 구조

들의 프로젝트 개시의 핵심 개체의 일부로서 메타데이터 저장소를 소유할 필요가 있다. 이러한 저장소는 비즈니스 이용자들을 염두에 두고 구축되고 기술적으로 견고한 구조상에서 생성되어야 할 필요가 있다.

부가적으로 메타데이터는 기술을 갖추지 못한 비즈니스 이용자들이 저장소에 접근하여 돌아볼 수 있도록 하는 전달 체계를 갖추어야 할 필요가 있다. 심지어 미완의 상태인 메타데이터 시장에서도, 저장소를 구축하지 않는 대안은 비즈니스 이용자들이나 의사결정 지원시스템을 유지할 필요가 있는 데이터 웨어하우스 담당직원들의 요구를 만족시켜 주지 못할 것이다.

메타데이터 저장소는 웨어하우스나 데이터마트에서 보다 시각적이고 이해하기 쉽고 대중에 접근하기 쉬운 정보를 만드는데서 의사결정 지원시스템에 많은 도움을 줄 것이다.

요컨대, 저장소는 데이터 웨어하우스를 만들 수도 파괴할 수도 있다. 