

LAYING THE FOUNDATION FOR INTERNATIONAL STANDARDIZATION OF SMART CITIES

- D1.1 스마트시티 거버넌스 체계 및 프레임워크 개발 - 한국건설기술연구원, 가천대학교
- D1.2 통합플랫폼 등 공공서비스 표준화 및 표준 인덱스 개발 연구 - 스마트도시협회, 성균관대학교
- D1.3 스마트시티 포트폴리오-프로그램-프로젝트 프로세스 표준화 - 한양대학교
- D1.4 스마트시티 데이터 상호운용성 확보 및 가이드라인 정립 - 한국정보통신기술협회
- D2.1 스마트교통-에너지 연계분야 표준지침개발 - 한국지능형교통체계협회, 한국교통대학교, 홍익대학교, 한국전기연구원
- D2.2 스마트에너지-빌딩 연계분야 표준지침개발 - 스마트에너지협회, 한국기계전기전자시험연구원
- D2.3 스마트시티 공간정보 분야 표준지침개발 - 한국정보통신기술협회, 안양대학교
- D2.4 스마트 ICT분야 표준지침개발 - 한국전자기술연구원, 연세대학교
- D3.1 스마트시티 표준화 역량강화 프로그램 - 한국정보통신기술협회, 영국왕립표준협회(BSI)
- D3.2 스마트시티 품질 평가체계 구축 및 인증 - 한국정보통신기술협회, 영국왕립표준협회(BSI)
- D3.3 스마트시티 국제표준화를 위한 협력프로그램 - 한국정보통신기술협회, 영국왕립표준협회(BSI)

SMART CITY STANDARDS

For leading a smart cities technology and international standardization
Laying the foundation of standardization capability in smart cities related domains
Establishment of Smart cities standardization roadmap
Development of connecting standards between each smart city domain and technology

STRATEGIC PLANNING GOVERNANCE
스마트시티 표준화 전략 및 체계 구축

TECHNOLOGY STANDARDS
스마트시티 도메인별 기술표준 및 서비스 시험표준 개발

스마트시티 표준화역량개발 및 국제협력
CAPABILITY DEVELOPMENT

KAIA KOREA AGENCY FOR INFRASTRUCTURE TECHNOLOGY ADVANCEMENT



국토교통부

KAIA



Korea Agency for Infrastructure Technology Advancement



ISBN 979-11-963090-6-0 (세트)

DAMA-DMBOK DATA MANAGEMENT PRINCIPLES STRATEGY

07
ISSUE
SMART CITY
STANDARDS

DAMA
DATA MAN-
AGEMENT
FRAMEWORK
EVOLVED

DATA
MANAGE-
MENT

DATA

SMART CITY STANDARDS

DATA MANAGE-
MENT
INTRODUCTION
ESSENTIAL CON-
CEPTS
DATA

DATA AND IN-
FORMATION
DATA AS AN OR-
GANIZATIONAL
ASSET
DATA MANAGE-
MENT PRINCIPLES

DATA MANAGE-
MENT CHAL-
LENGES
DATA MANAGE-
MENT STRATEGY
DATA MANAGE-
MENT FRAME-
WORKS

STRATEGIC
ALIGNMENT
MODEL
THE AMSTER-
DAM INFORMA-
TION MODEL
THE DAMA-DM-
BOK FRAME-
WORK

데이터 매니지먼트 (Data Management)

제1절 서론	6
제2절 본질적 개념	10
1. 데이터	10
2. 데이터와 정보	11
3. 조직 자산으로서 데이터	12
4. 데이터 관리 원칙	12
5. 데이터 관리에 관한 과제	15
6. 데이터 관리 전략	23
제3절 데이터 관리 프레임워크	26
1. 전략적 정렬 모델(Strategic Alignment Model)	27
2. 암스테르담 정보 모형(The Amsterdam Information Model)	30
3. 데이터 관리 지식 체계 프레임워크(The DAMA-DMBOK Framework)	30
4. 데이터 관리 지식 체계 피라미드(DMBOK Pyramid; Aihen)	32
5. Dama 데이터 관리 프레임워크의 진화	35
제4절 데이터 지식 관리 체계(DAMA-DMBOK)	38
제5절 인용 및 참고 자료	40





DATA MANAGEMENT

데이터 매니지먼트

1 서론

많은 조직들은 그들의 데이터가 필수적인 기업 자산(Vital Enterprise Asset)이라는 것을 인식하고 있다. 데이터와 정보는 고객, 제품 및 서비스에 대한 인사이트(Insight) 제공한다. 이는 그들의 혁신과 전략적 목표 달성에 도움을 준다. 하지만 이러한 인식에도 불구하고 데이터를 지속적으로 가치를 창출할 수 있는 자산으로 적극적으로 관리하는 조직은 거의 없다(Evans and Price, 2012). 원래 형태로 저장된 데이터에서 우연하게 가치가 도출되지 않으며, 확고한 의지를 가지고 계획을 세워 조직화하여 전력을 다할 필요가 있다. 이에 매니지먼트와 리더십이 필요하다. 데이터 매니지먼트(Data Management)는 생애주기 전체에 걸쳐 데이터 및 정보 자산의 가치를 전달, 제어, 보호 및 향상시키는 계획, 정책, 프로그램 및 실행을 개발, 실행 및 감독하는 것이다. 데이터 매니지먼트 전문가는 조직의 전략적 목표를 달성하기 위해 데이터 매니지먼트의 모든 측면(전 생애주기에 걸쳐 데이터의 기술적 매니지먼트에서 데이터가 적절하게 활용되고 활용되도록 보장)에서 일하는 전문가를 말한다. 데이터 매니지먼트 전문가는 고도의 기술자(예: 데이터베이스 매니저, 네트워크 매니저, 프로그래머)

부터 전략적 비즈니스(예: 데이터 담당자, 데이터 전략가, 최고 데이터 책임자)에 이르기까지 수많은 역할을 수행한다. 데이터 매니지먼트 활동은 광범위하다. 여기에서는 데이터에서 전략적 가치를 얻는 방법에 대한 일관성 있는 의사결정 능력에서부터 데이터베이스의 기술적 배치 및 성능에 이르기까지 모든 것이 포함된다. 따라서 데이터 매니지먼트에는 기술적 및 비기술적(즉, '비즈니스') 기술이 모두 필요하다. 데이터 매니지먼트에 대한 책임은 비즈니스와 정보통신기술 역할 간에 공유되어야 하며, 조직의 전략적 요구에 부합하는 고품질 데이터를 보장하기 위해 두 분야의 사람들이 협력할 수 있어야 한다.

데이터와 정보는 조직이 미래 가치를 도출하기 위해 투자한다는 의미에서 단순한 자산이 아닙니다. 데이터와 정보는 또한 대부분의 조직의 일상적인 운영에 필수적입니다. 그것들은 'Currency(통화)', 'Life Blood(생명의 피)', 심지어 정보경제의 'New Oil(새로운 석유)'로도 불려왔다.¹ 조직은 분석에서 가치를 얻든 얻지 못하든 데이터 없이는 비즈니스를 처리할 수조차 없다. DAMA International (The Data Management Association)은 작업을 수행하는 데이터 매니지먼트 전문가를 지원하기 위해 The DAMA Guide to the Data Management Body of Knowledge (DMBOK2)를 도출했다. 여기에서는 데이터 매니지먼트를 위한 일련의 원칙을 개략적으로 설명한다. 이러한 원칙의 준수와 관련된 과제를 논의하고 이러한 과제를 해결하기 위한 접근법을 제안한다. 또한 여기에서는 다양한 데이터 매니지먼트 지식 영역 내에서 데이터 매니지먼트 전문가가 수행하는 작업에 대한 맥락을 제공하기 위해 데이터 매니지먼트 프레임워크를 설명한다.

1.1 비즈니스 동인(Business Driver)

정보와 지식은 경쟁력의 우위를 점하기 위해 필수적이다. 고객, 제품, 서비스 및 운영에 대해 신뢰할 수 있는 고품질 데이터를 보유한 조직은 데이터가 없거나 신뢰할 수 없는 데이터가 있는 조직보다 더 나은 의사결정을 내릴 수 있다. 데이터 매니지먼트 실패는 예산 매니지먼트 실패와 비슷하다. 이것은 낭비를 초래하고 기회를 잃게 만든다. 데이터 매니지먼트의 주요 동인은 금융 및 물리적 자산의 효과적인 매니지먼트를 통해 정부가 데이터 자산으로부터 가치를 얻을 수 있도록 하는 것과 마찬가지로 정부가 데이터 자산으로부터 가치를 얻을 수 있도록 하는 것이다.

1.2 목표

한 조직 내에서, 데이터 매니지먼트 목표는 아래의 사항을 포함한다:

- 고객, 직원 및 비즈니스 파트너를 포함한 기업 및 이해관계자의 정보 요구 이해 및 지원
- 데이터 자산의 무결성 습득, 저장, 보호 및 보장
- 데이터 및 정보의 품질 보장
- 이해관계자 관련 데이터의 개인 정보 보호 및 기밀성 보장
- 데이터와 정보의 무단 또는 부적절한 액세스, 조작 또는 사용 방지
- 데이터를 효과적으로 사용하여 기업에 가치를 더할 수 있도록 보장

¹ Google 'data as currency', 'data as life blood', and 'the new oil', for numerous references.





2 본질적 개념

2.1 데이터

데이터에 대한 오랜 정의는 세상의 사실들을 표현하는데 있어 데이터의 역할을 강조한다.² 정보기술과 관련하여 데이터는 디지털 형태로 저장된 정보로도 이해된다(데이터는 디지털화된 정보에 한정되지 않으며 데이터 매니지먼트 원칙은 데이터베이스뿐만 아니라 문서에서 포착된 데이터에도 적용된다). 하지만 오늘날 우리는 전자적으로 너무 많은 정보를 접할 수 있기 때문에, 이름, 주소, 생년월일, 토요일 저녁에 무엇을 먹었는지, 가장 최근에 구입한 책과 같이, 이전에는 '데이터'라고 불리지 않았을 많은 것들까지도 '데이터'라고 부른다.

개인에 대한 일련의 사실들은 종합되고, 분석되고, 이윤을 창출하고, 건강을 증진시키거나, 공공 정책에 영향을 미치기 위해 사용될 수 있다. 게다가 광범위한 사건 및 활동(빅뱅의 파장으로부터 우리 자신의 심장박동에 이르기까지)을 측정하고, 이전에는 데이터로 생각하지 않았던(비디오, 사진, 녹음파일, 문서) 것들의 전자화된 버전을 수집, 저장, 분석하는 우리의 기술 역량은 데이터를 사용 가능한 정보로 합성하는 능력을 능가하는데 근접했다.³ 양과 속도에 압도되지 않고 다양한 데이터를 활용하려면 신뢰할 수 있고 확장 가능한 데이터 매니지먼트 실행이 필요하다.

대부분의 사람들은 데이터가 사실을 나타내기 때문에, 그것은 세계에 대한 진실의 한 형태이며, 그 사실이 서로 들어맞을 것이라고 추측한다. 그러나 '사실'이 항상 간단하거나 직설적인 것은 아니다. 데이터는 표현의 수단이다. 이것은 그 자체 이외의 것을 나타낸다(Chisholm, 2010). 데이터는 데이터가 나타내는 객체와 해석해야 하는 객체에 대한 해석이다(Sebastian-Coleman, 2013). 이는 데이터가 의미 있게 쓰이기 위해서는 맥락이 필요하다는 또 다른 표현이다. 문맥은 데이터의 대표 체계라고 생각할 수 있다. 이러한 시스템은 공통 어휘와 요소들 간의 일련의 관계를 포함한다. 만약 우리가 이러한 시스템의 관계를 안다면, 우리는 그 안에 있는 데이터를 해석할 수 있을 것이다.⁴ 이러한 규약은 흔히 메타데이터라고 하는 특정 종류의 데이터에 문서화된다.

² The New Oxford American Dictionary defines data as 'facts and statistics collected together for analysis.' The American Society for Quality (ASQ) defines data as 'A set of collected facts' and describes two kinds of numerical data: measured or variable and counted or attributed. The International Standards Organization (ISO) defines data as 're-interpretable representation of information in a formalized manner suitable for communication, interpretation, or processing; (ISO 11179). The definition emphasizes the electronic nature of data and assumes, correctly, that data requires standards because it is managed through information technology systems. That said, it does not speak to the challenges of formalizing data in a consistent way, across disparate systems. Nor does it account well for the concept of unstructured data.

³ <http://ubm.io/2c4yPOJ> (Accessed 2016-12-04). <http://bit.ly/1rOQkt1> (Accessed 2016-12-04).

⁴ For additional information on the constructed-ness of data see: Ket, Data and Reality (2012) and Devlin

그러나, 사람들은 종종 개념을 표현함에 있어서 다른 선택을 하기 때문에, 그들은 같은 개념이라도 다른 표현방식을 만든다. 이러한 선택으로 말미암아 데이터는 다른 형태를 취한다. 합의된 정의가 존재하는 개념인 달력 날짜를 표현하는 방법의 범위를 생각해 보자. 이제 더 복잡한 개념(예: 시민 또는 서비스)을 고려해 보자. 여기서 표현해야 할 세부사항의 세부사항과 수준이 항상 자명하지는 않으며, 시간이 지남에 따라 해당 정보를 매니지먼트하는 과정이 더 복잡해진다.

단일 조직 내에서도 같은 생각을 표현하는 방법은 여러 가지가 존재하는 경우가 많다. 따라서, 데이터 아키텍처, 모델링, 거버넌스, 매니지먼트, 메타데이터 및 데이터 품질 매니지먼트의 필요성, 그리고 이 모든 것이 사람들이 데이터를 이해하고 사용하는 데 도움이 됩니다. 조직 전체에서 다중성 문제는 다중성을 배로 증가시킨다. 따라서 데이터에 대한 일관성을 높일 수 있는 업계 수준의 데이터 표준이 요구된다. 조직은 항상 데이터를 매니지먼트해야 했지만, 기술의 변화로 데이터가 무엇인지에 대한 사람들의 이해가 바뀌면서 이러한 매니지먼트 요구의 범위가 확대된다. 이러한 변화로 조직은 데이터를 새로운 방식으로 사용하여 제품을 만들고, 정보를 공유하고, 지식을 창출하고, 조직의 성공을 개선할 수 있게 되었다. 그러나 기술의 급속한 성장과 더불어 의미 있는 데이터를 생산, 캡처 및 채굴할 수 있는 인적 능력의 증가는 데이터를 효과적으로 매니지먼트해야 할 필요성을 더 심화시켰다.

2.2 데이터와 정보

데이터와 정보의 관계에 관한 많은 글들이 존재한다. 데이터는 ‘정보의 원재료’라고 불리고 정보는 ‘문맥 속의 데이터(Data in Context)’라고 불리기도 한다. 흔히 데이터(가장 아래), 정보, 지식, 지혜(맨 위)의 관계를 설명하기 위해 층을 이룬 피라미드를 사용합니다. 피라미드는 데이터를 잘 매니지먼트해야 하는 이유를 설명하는 데 도움이 될 수 있지만, 이 표현은 데이터 매니지먼트에 대한 몇 가지 문제가 제기된다.

- 데이터는 단순히 존재한다는 가정에 근거한다. 그러나 데이터는 단순히 존재하는 것이 아니다. 데이터를 생성되어야 하는 것이다.
- 데이터에서 정보, 지식, 지혜를 통한 선형적 순서를 기술함으로써, 애초에 데이터를 생성하기 위해서는 지식이 필요하다는 것을 인식하지 못한다.
- 데이터와 정보는 별개의 것이라고 암시하는데, 실제로는 두 개념이 서로 얽혀 있고 서로 의존하고 있습니다. 데이터는 정보의 한 형태이며 정보는 데이터의 한 형태이다.

조직 내에서, 서로 다른 이해당사자들의 서로 다른 사용에 대한 요구사항과 기대치에 대해 명확한 의사소통을 위해 정보와 데이터 사이에 선을 긋는 것이 도움이 될 수 있다. ‘여기 지난 분기의 판매 보고서가 있다[정보]. 이것은 데이터 웨어하우스의 데이터를 기반으로 합니다[데이터]. 다음 분기에 이러한 결과[데이터]는 분기별 성과 측정[정보]를 생성하는 데 사용될 것입니다.’) 데이터와 정보를 다양한 목적에 맞게 준비해야 한다는 인식은 데이터와 정보는 모두 적절히 관리되어야 한다는 데이터 매니지먼트의 중심 원칙을 형성한다. 두 가지 모두 사용되는 모습 및 고객 요구사항을 염두에 두고 함께 매니지먼트된다면 더 높은 품질을 얻을 수 있을 것이다. 데이터 매니지먼트 지식 체계 전체에서 이 용어는 서로 교환하여 사용될 것이다.

2.3 조직 자산으로써 데이터

자산은 소유하거나 통제할 수 있고 가치를 지니거나 산출하는 경제적 자원이다. 자산은 돈으로 환산할 수 있다. 데이터를 자산으로 매니지먼트하는 것이 무엇을 의미하는지 이해하는 방법은 여전히 진화하고 있지만 데이터는 기업의 자산으로 널리 인식되고 있다. 1990년대 초, 일부 단체들은 영업권의 가치에 금전적 가치를 부여해야 하는지에 대해 의문을 제기하였다. 이제 ‘영업권의 가치’는 흔히 손익계산서(P&L)에 항목으로 나타난다. 마찬가지로, 보편적으로 채택되지는 않지만, 데이터의 수익화는 점점 더 보편화되고 있다. 머지않아 손익계산서의 항목으로도 볼 수 있을 것이다.

오늘날의 조직은 보다 효과적인 의사결정과 보다 효율적인 운영을 위해 데이터 자산에 의존한다. 기업은 데이터를 활용하여 고객을 이해하고, 새로운 제품과 서비스를 창출하며, 비용을 절감하고 리스크를 통제함으로써 운영 효율성을 향상시킨다. 정부기관, 교육기관, 비영리단체도 운영, 전술, 전략적 활동을 지도할 수 있는 양질의 데이터가 필요하다. 조직이 점점 더 데이터에 의존함에 따라 데이터 자산의 가치는 더욱 명확하게 정립될 수 있다.

많은 조직이 스스로를 ‘데이터 중심적’이라고 인식합니다. 경쟁력 유지를 목표로 하는 기업은 직감이나 본능에 근거한 의사결정을 중단하고, 대신 이벤트 트리거(Trigger)를 사용하고 분석 및 적용하여 실행 가능한 통찰력을 얻어야 한다. 데이터 중심화에는 비즈니스 리더십과 기술 전문지식을 두루 익혀 데이터를 효율적으로 매니지먼트해야 한다는 인식이 포함된다.

더욱이 오늘날 비즈니스의 속도는 변화가 더 이상 선택사항이 아니라는 것을 의미한다. 디지털 붕괴가 일반적이다. 이에 대응하기 위해 기업은 LOB(Line-of-Business)와 함께 일하는 기술 데이터 전문가와 함께 정보 솔루션을 공동 개발해야 한다. 그들은 사업 전략을 지원하는 데 필요한 데이터를 얻고 매니지먼트하는 방법을 계획해야 한다. 또한 새로운 방식으로 데이터를 활용할 수 있는 기회를 활용할 수 있도록 스스로를 포지셔닝해야 한다.

2.4 데이터 매니지먼트 원칙

데이터 매니지먼트는 [그림]과 같이 다른 형태의 자산 매니지먼트와 특성을 공유한다. 여기에는 어떤 데이터와 조직이 어떤 데이터를 보유하고 있으며 이를 통해 무엇을 달성할 수 있는지 파악한 다음, 조직 목표에 도달하기 위해 데이터 자산을 가장 잘 사용하는 방법을 결정하는 작업이 함께 포함된다. 다른 매니지먼트 프로세스와 마찬가지로 전략적인 요구와 운영상의 요구의 균형을 맞춰야 한다. 이러한 균형을 맞추기 위해서 데이터 매니지먼트의 중요한 특징을 인식하고 데이터 매니지먼트 실행을 안내하는 일련의 원칙을 따라야 한다.

- 데이터는 고유한 속성을 가진 자산임: 데이터는 자산이지만, 데이터가 어떻게 관리되는지에 영향을 미치는 중요한 방식으로 다른 자산과 다르다. 이러한 속성 중 가장 명백한 것은 데이터가 사용될 때 소비되지 않는다는 것이며, 데이터는 금융자산과 마찬가지로 물리적 자산이다.
- 데이터의 가치는 경제적 용어로 표현될 수 있고 표현되어야 함: 데이터를 자산으로 부르는 것은 데이터가 가치가 있다는 것을 의미한다. 데이터의 질적, 양적 가치를 측정하는 기법은 있지만, 아직 표준은 존재하지 않는다. 데이터에 대해 더 나은 결정을 내리고자 하는 조직은



[그림1] 데이터 매니지먼트 원칙

해당 가치를 계량화할 수 있는 일관된 방법을 개발해야 한다. 또한 저품질 데이터의 비용과 고품질 데이터의 이점을 모두 측정해야 한다.

- 데이터 매니지먼트란 데이터의 품질을 매니지먼트하는 것을 의미함: 데이터가 목적에 맞는지 확인하는 것이 데이터 매니지먼트의 주요 목표이다. 품질을 매니지먼트하기 위해 조직은 품질에 대한 이해관계자의 요구사항을 이해하고 이러한 요구사항에 대한 데이터를 측정해야 한다.
- 데이터를 매니지먼트하는 데 메타데이터가 필요함: 자산을 매니지먼트하려면 해당 자산에 대한 데이터(직원 수, 회계 코드 등)가 필요하다. 데이터를 매니지먼트하고 사용하는 데 사용되는 데이터를 메타데이터라고 한다. 데이터를 보유하거나 만질 수 없기 때문에, 데이터가 무엇이고 어떻게 사용하는지를 이해하기 위해서는 메타데이터 형태의 정의와 지식이 필요하다. 메타데이터는 아키텍처, 모델링, 스투어드십, 거버넌스, 데이터 품질 매니지먼트, 시스템 개발, IT 및 비즈니스 운영, 분석을 포함한 데이터 생성, 처리 및 사용과 관련된 다양한 프로세스에서 비롯된다.
- 데이터 매니지먼트에는 계획이 필요함: 소규모 조직도 복잡한 기술 및 비즈니스 프로세스 환경을 가질 수 있다. 데이터는 여러 곳에서 생성되며, 사용하기 위해 여러 장소를 거쳐간다. 최종 결과(End Result)와 업무 과정을 일치시키기 위해서는 구조 및 프로세스 관점에서 계획을 수립해야 한다.

- 데이터 매니지먼트는 상호 기능적이며, 다양한 기술과 전문 지식이 필요함: 단일 팀이 조직의 모든 데이터를 매니지먼트할 수는 없다. 데이터 매니지먼트에는 기술 및 비기술적 지식과 협업 능력이 모두 필요하다.
- 데이터 매니지먼트에는 기업의 관점이 필요함: 데이터 매니지먼트는 부서별 특징을 가지고 있지만, 효과를 가능한 한 더 높이기 위해서는 기업 전체에 걸쳐 적용될 수 있는 전사적 관점이 필요하다. 이는 데이터 매니지먼트와 데이터 거버넌스가 얹혀 있는 한 가지 이유이기도 하다.
- 데이터 매니지먼트는 다양한 관점을 고려해야 함: 데이터는 유동적이다. 데이터 매니지먼트는 데이터가 생성되고 사용되는 방식과 그것을 사용하는 데이터 소비자의 트렌드를 따라잡기 위해 끊임없이 진화해야 한다.
- 데이터 매니지먼트는 수명주기 매니지먼트임: 데이터에는 수명주기가 있고 데이터를 매니지먼트하려면 수명주기를 매니지먼트해야 한다. 데이터로 말미암아 더 많은 데이터가 생성되기에 데이터 수명주기는 그 자체로 매우 복잡할 수 있다. 데이터 매니지먼트 원칙에는 데이터 수명주기를 고려하는 것도 포함되어야 한다.
- 데이터 유형마다 수명주기 특성이 상이함: 그리고 이 때문에 데이터 매니지먼트 원칙도 다르다. 데이터 매니지먼트 원칙은 이러한 차이를 인식하고 다양한 종류의 데이터 수명주기 요구사항을 충족할 수 있을 만큼 충분히 유연해야 한다.
- 데이터 매니지먼트에는 데이터 관련 리스크 매니지먼트도 포함됨: 자산이 되는 것 외에도 데이터는 조직에 대한 위험도 나타난다. 데이터는 분실, 도난 또는 오용될 수 있다. 조직은 데이터 사용의 윤리적 영향을 고려해야 한다. 데이터 관련 위험은 데이터 수명주기의 일부로 매니지먼트해야 한다.
- 데이터 매니지먼트 요건은 정보 기술 결정을 이끌어야 함: 데이터와 데이터 매니지먼트는 정보 기술 및 정보 기술 매니지먼트와 깊이 관련되어 있다. 데이터를 매니지먼트를 하려면 기술이 조직의 전략적 데이터 요구를 충족시키는 것이 아니라 이를 보장하는 접근 방식이 필요하다.
- 효과적인 데이터 매니지먼트에는 리더십이 필요함: 데이터 매니지먼트에는 효과적인 조정, 협업 및 헌신이 필요한 복잡한 프로세스 등이 포함된다. 이들을 달성하기 위해서는 경영 기술뿐만 아니라, 헌신적인 리더십에서 오는 비전과 목적이 필요하다.

2.5 데이터 매니지먼트에 관한 과제

데이터 매니지먼트는 데이터 자체의 속성에서 파생된 뚜렷한 특성을 가지고 있으며, 또한 이러한 원칙을 따르는 데 있어서 과제를 제시한다. 이에 대한 자세한 내용은 섹션 2.5.1에서 2.5.13까지에 설명되어 있다. 이러한 난제들 중 많은 것들이 하나 이상의 원칙을 언급하고 있다.

2.5.1 데이터는 다른 자산과 상이함⁶

물리적 자산은 가리킬 수 있으며 만지고 움직일 수 있다. 그리고 그들은 한 번에 한 장소에만 있을 수 있다. 금융자산은 대차대조표상 회계처리 되어야 한다. 그러나 데이터는 다르다. 데이터는 실체가 없다. 그러나 데이터는 영속성이 있다; 오래될수록 데이터의 가치는 변하지만 데이터가 사라지는 것은 아니며, 데이터는 복사하고 전송하기 쉽다. 그러나 잃어버리거나 파괴되면 재생산이 쉽지 않다. 왜냐하면 데이터는 사용할 때 소비하지 않으면, 없어지지 않고도 도난당할 수도 있기 때문이다. 데이터는 동적이며 여러 용도로 사용할 수 있다. 동일한 데이터를 여러 사람이 동시에 사용할 수도 있는데, 이는 물리적 또는 금융 자산으로는 불가능한 것이다. 많은 데이터 사용은 더 많은 데이터를 확보할 수 있게 해준다. 대부분의 조직은 증가하는 데이터의 양과 데이터셋(Data Sets) 간의 관계를 잘 관리해야 한다.

이러한 차이는 데이터에 금전적 가치를 부여하는 것을 어렵게 만든다. 이러한 금전적 가치 없이는 데이터가 조직의 성공에 어떻게 기여하는지를 측정하기 어렵다. 이러한 차이는 또한 데이터 소유권 정의, 데이터 사용량 재고 매니지먼트, 데이터 오용 방지, 데이터 중복과 관련된 위험 매니지먼트, 데이터 품질 표준 정의 및 시행과 같은 데이터 매니지먼트에 영향을 미치는 다른 문제를 야기한다. 데이터의 가치를 측정해야 하는 어려움에도 불구하고, 대부분의 사람들은 데이터가 실제로 가치를 가지고 있다는 것을 인식하고 있다. 조직의 데이터는 그 자체로 특별하다. 조직적 내 고유한 데이터(예: 고객 목록, 제품 재고 또는 청구 기록)가 손실되거나 파괴되는 경우, 이를 대체하는 것은 불가능하거나 비용이 매우 많이 들 것이다. 데이터는 또한 조직이 스스로를 알 수 있는 수단이다(그것은 다른 자산을 설명하는 메타자산이다). 그만큼 데이터는 조직의 통찰력을 위한 토대를 마련한다.

조직 내부와 조직 간에는 업무 수행에 데이터와 정보가 필수적이다. 대부분의 운영 업무 거래는 정보의 교환을 포함한다. 대부분의 정보는 전자적으로 교환되어 데이터 흔적을 만든다. 이 데이터 추적은 발생한 교환을 표시하는 것 외에 다른 목적으로 사용될 수도 있다. 조직이 어떻게 기능하는지에 대한 정보를 제공할 수 있는 것은 이의 한 예시이다. 어떤 조직에서든 데이터가 수행하는 중요한 역할 때문에 데이터를 주의 깊게 관리할 필요가 있다.

⁶ 이 장은 Redman, Thomas.의 Data Quality for the Information Age (1996) pp. 41-42, 232-36;와 Data Driven (2008), Chapter One. "The Wondrous and Properties of Data and Information."을 참고했다.

2.5.2 데이터 가치평가

가치는 물건의 원가와 그것으로부터 파생되는 이익의 차이이다. 주식과 같은 일부 자산의 경우 가치평가가 쉽다. 이는 매입 당시 주식 원가와 매각 대금의 차이가 있다. 그러나 데이터의 경우 이러한 계산은 더 복잡하다. 왜냐하면 데이터의 비용도 이익도 표준화되지 않기 때문이다. 각 조직의 데이터는 그 자체에만 고유하기 때문에, 데이터 가치평가에 대한 접근법은 조직 내에서 일관성 있게 적용될 수 있는 일반적 비용 및 이익 범주를 명확히 하는 것으로 시작할 필요가 있다. 샘플 범주에는 다음이 포함된다.

- 데이터 수집 및 저장 비용
- 데이터 손실 시 교체 비용
- 데이터가 누락된 경우 조직에 미치는 영향
- 데이터 관련 리스크 완화 비용 및 잠재적 리스크 비용
- 데이터 개선 비용
- 고품질 데이터의 이점
- 경쟁사가 데이터 비용으로 지불하는 액수
- 데이터를 판매할 수 있는 대상
- 혁신적인 데이터 사용으로 예상되는 수익

데이터 자산 가치평가에 대한 일차적인 어려움은 데이터의 가치가 문맥상(한 조직에 가치 있는 것은 다른 조직에 가치가 없을 수 있음)이며 종종 일시적(어제 가치 있었던 것은 오늘날 가치 있는 것이 아닐 수 있음)이라는 것이다. 즉, 조직 내에서 특정 유형의 데이터는 시간이 지남에 따라 일관되게 가치가 있을 가능성이 높다. 예를 들어 신뢰할 수 있는 고객 정보를 참조하라. 고객 활동과 관련된 데이터가 더 많이 축적됨에 따라 고객 정보는 시간이 지남에 따라 더욱 가치가 높아질 것이다.

데이터 매니지먼트와 관련하여, 조직은 일관된 의사결정을 하기 위해서는 재무적 측면에서 자산을 이해해야 하기 때문에 재무적 가치를 데이터와 연관시키는 방법을 확립하는 것이 중요하다. 데이터에 가치를 부여하는 것은 데이터 매니지먼트 활동에 가치를 부여하는 기초가 된다. 데이터 평가 과정도 변화 매니지먼트 수단으로 활용될 수 있다. 데이터 매니지먼트 전문가와 자신이 지원하는 이해관계자에게 작업의 재정적 의미를 이해하도록 요청하는 것은 조직이 자체 데이터에 대한 이해와 이를 통해 데이터 매니지먼트에 대한 접근 방식을 혁신하는 데 도움이 될 수 있다.

2.5.3 데이터 품질

데이터가 고품질인지 확인하는 것은 데이터 매니지먼트의 핵심이다. 조직은 데이터를 사용하고자 하기 때문에 데이터를 관리한다. 만약 그들이 비즈니스 요구를 충족시키기 위해 데이터에 의존할 수 없다면, 그것에 대한 수집, 저장, 보안, 접근을 가능하게 하는 노력은 낭비된다고 볼 수 있다. 데이터가 비즈니스 요구를 충족할 수 있도록 데이터 소비자와 협력하여 높은 품질의 데이터를 만드는 요구를 정의해야 한다.

주로 데이터는 정보통신기술과 매우 밀접하게 연관되어 있기 때문에, 데이터 품질을 매니지먼트하는 것은 역사적으로 사후고려 대상으로 취급되어 왔다. IT팀은 종종 자신이 만든 시스템이 저장해야 하는 데이터를 무시하곤 했다. 그는 아마 쓰레기통에 쓰레기가 차면 쓰레기통이 비워지듯 데이터가

‘들어오고 나가는 것’을 처음 관찰한 프로그래머였을 것이고 그는 이를 그냥 내버려둬야 한다는 의심 또한 가지고 있었을 것이다. 그러나 데이터를 사용하려는 사람들은 품질을 무시해서는 안 된다. 이들은 일반적으로 데이터가 신뢰할 수 있고 신뢰할 수 있다고 가정하며, 이를 의심할 이유가 있을 때까지 데이터를 신뢰한다. 한번 신뢰를 잃으면 이를 되찾는 것은 더욱 어렵기 때문이다.

대부분의 데이터 사용에는 그러한 학습을 적용하고 가치를 창출하기 위해 데이터로부터 배우는 것이 포함된다. 제품이나 서비스를 개선하기 위해 고객의 습관을 이해하고 조직의 성과를 평가하는 것 또는 더 나은 사업 전략 등을 개발하기 위한 시장 동향을 파악하는 등을 예로 들 수 있다. 질 낮은 데이터는 이러한 결정에 부정적인 영향을 미칠 것이다.

중요한 것은, 낮은 품질의 데이터는 어떤 조직에도 그저 많은 비용으로 작용할 뿐이라는 것이다. 추정치는 다르지만, 전문가들은 조직이 수익의 10-30%를 데이터 품질 문제를 처리하는 데 사용한다고 생각할 수 있다. IBM은 2016년 미국의 불량 품질 데이터 비용을 3조 1천억 달러로 추산했다. 품질 불량 데이터의 많은 비용은 숨겨져 있고, 간접적이며, 그러므로 측정하기 어렵다. 벌금과 같은 다른 비용이 직접적이며 계산하기 쉬운 것과 대조된다.

비용은 다음에서 발생한다:

- 스크랩 및 재작업
- 해결 방법 및 숨겨진 수정 프로세스
- 조직의 비효율성 또는 낮은 생산성
- 조직 갈등
- 낮은 직무 만족도
- 고객 불만족
- 혁신 불가능을 포함한 기회 비용
- 컴플라이언스 비용 또는 벌금
- 평판 비용

고품질 데이터의 이점은 다음과 같다:

- 고객 경험개선
- 생산성 향상
- 리스크 감소
- 기회에 민첩하게 대응할 수 있는 능력
- 매출 증대
- 고객, 제품, 프로세스 및 기회에 대한 인사이트를 통해 얻은 경쟁력

이러한 비용과 편익이 시사하듯이 데이터 품질 매니지먼트는 일회성이 아니다. 고품질의 데이터를 생산하려면 프로세스와 시스템에 품질을 구축하는 계획, 헌신, 사고방식이 필요하다. 모든 데이터 매니지먼트 기능은 좋은 나쁜 데이터 품질에 영향을 미칠 수 있으므로, 모든 데이터 매니지먼트 기능은 작업을 실행할 때 이를 설명해야 한다.

2.5.4 더 좋은 데이터를 위한 계획 수립

서론에서 언급된 바와 같이, 데이터로부터 가치를 도출하는 것은 우연히 일어나지 않는다. 여러 가지

형태로 계획을 세워야 한다. 계획을 세우는 것은 조직이 어떻게 데이터를 얻고 생성하는지를 통제할 수 있다는 인식에서 출발한다. 만약 그들이 데이터를 자신이 만들 수 있는 제품으로 간주한다면, 그들은 데이터에 관해 수명 주기 전반에 걸쳐 더 나은 의사결정을 내릴 것이다. 이러한 결정에는 다음과 같은 사항이 수반되기 때문에 시스템 사고가 필요하다.

- 데이터가 서로 별개로 보일 수 있는 비즈니스 프로세스를 연결하는 방법
- 비즈니스 프로세스와 이를 지원하는 기술 간의 관계
- 시스템 및 시스템의 생성 및 저장 데이터의 설계 및 아키텍처
- 데이터가 조직 전략을 발전시키는 데 사용되는 방법

더 나은 데이터를 계획하려면 아키텍처, 모델링 및 기타 설계 기능에 대한 전략적 접근이 필요하다. 또한 데이터 계획은 비즈니스와 IT 리더 간의 전략적 협업에 달려 있습니다. 그리고 물론 개별 프로젝트 및 단계에서 효과적으로 실행할 수 있는 개개인의 능력 또한 중요하다. 문제는 시간과 비용에 관한 지속적인 압력을 비롯한 조직의 압력이 더 나은 계획의 수립을 방해한다는 것이다. 조직은 전략을 실행할 때 장단기 목표의 균형을 맞춰야 한다. 트레이드오프를 명확하게 하는 것은 더 나은 결정을 가능하게 한다.

2.5.5 메타데이터와 데이터 매니지먼트

조직은 데이터를 자산으로 매니지먼트하기 위해 신뢰할 수 있는 메타데이터를 필요로 한다. 이런 의미에서 메타데이터는 종합적인 관점에서 이해되어야 한다. 여기에는 비즈니스, 기술 및 운영 메타데이터뿐만 아니라 데이터 아키텍처, 데이터 모델, 데이터 보안 요구사항, 데이터 통합 표준 및 데이터 운영 프로세스에 포함된 메타데이터 또한 포함된다.

메타데이터는 조직이 어떤 데이터를 보유하고 있는지, 무엇을 나타내고 있는지, 어떻게 분류되는지, 어디서 왔는지, 조직 내에서 어떻게 움직이는지, 사용을 통해 어떻게 진화하는지, 누가 사용할 수 있는지, 누가 사용할 수 있는지, 그리고 고품질인지 등을 기술한다. 데이터는 추상적이다. 문맥에 대한 정의와 다른 설명은 그것을 이해할 수 있게 도와준다. 이를 통해 데이터, 데이터 수명주기, 그리고 데이터를 포함하는 복잡한 시스템을 이해할 수 있다.

문제는 메타데이터 또한 데이터의 한 형태이기에 데이터와 동일하게 매니지먼트할 필요가 있다는 점이다. 데이터를 잘 매니지먼트하지 못하는 조직은 일반적으로 메타데이터를 전혀 매니지먼트하지 않는다. 메타데이터 매니지먼트는 데이터 매니지먼트 전반의 개선을 위한 시작이 되는 경우가 많다.

2.5.6 데이터 매니지먼트는 기능 횡단적(Cross-functional)이다

데이터 매니지먼트는 복잡한 과정을 수반한다. 데이터는 데이터 수명주기의 다른 단계에 대한 책임이 있는 팀에 의해 조직 내의 다른 장소에서 관리된다. 데이터 매니지먼트는 시스템을 기획할 수 있는 디자인 역량, 하드웨어 매니지먼트 및 소프트웨어 구축을 위한 고도의 기술 역량, 문제를 이해하는 데이터 분석 역량, 데이터를 해석하는 분석 역량, 각 개념의 정의와 분석 모델에 대한 합의를 이끌어내는 언어적 역량과 더불어 고객에게 서비스를 제공할 기회를 탐색하고 목표를 달성하게끔 이끄는 전략적 사고 역량을 필요로 한다. 문제는 이렇게 다양한 범위의 기술과 관점을 가진 사람들이 공통의 목표를 향해 일할 때, 그들이 서로 잘 협업할 수 있도록 서로가 어떻게 조화를 이루는지를 인식하게 하는 것이다.

2.5.7 전사적 관점 수립

데이터를 매니지먼트하려면 조직 내에서 데이터의 영역과 범위를 이해해야 한다. 데이터는 영업, 마케팅, 운영 등 여러 분야를 넘나들거나 넘나들어야 하는 조직의 '수평적 요소' 중 하나이다. 데이터는 조직에만 있는 것이 아니라 부서나 조직의 다른 하위 부분에 대해서도 고유하게 존재할 수 있다. 데이터는 단순히 운영 프로세스의 부산물(예: 판매 거래 기록이 판매 프로세스의 부산물)로 간주하는 경우가 많기 때문에, 데이터가 항상 즉각적인 필요에 따라 계획되는 것은 아니다.

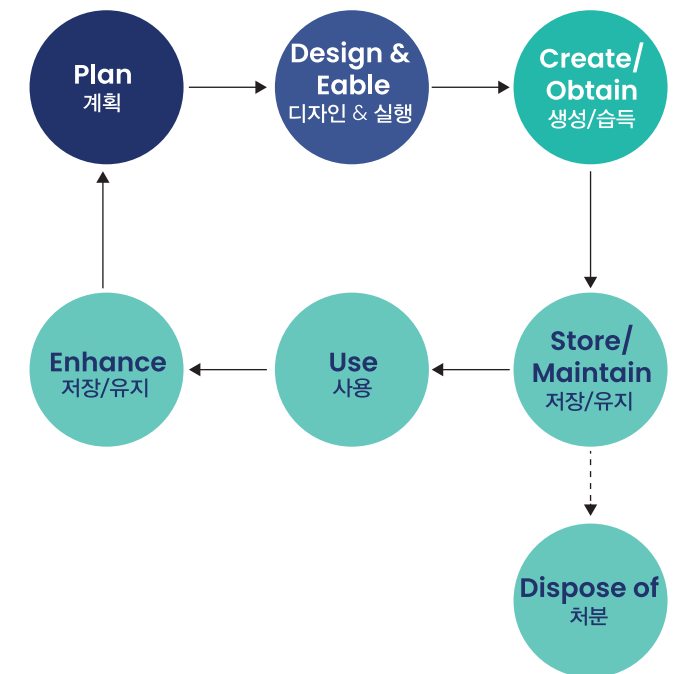
조직 내에서조차 데이터는 서로 다른 형태로 존재할 수 있다. 데이터는 조직 내의 여러 곳에서 생성된다. 부서마다 동일한 개념(예: 고객, 제품, 공급업체)이라도 표현하는 방법이 다를 수 있다. 데이터 통합 또는 마스터 데이터 매니지먼트 프로젝트에 관련된 모든 사람이 증언할 수 있듯이, 대표 용어를 선택하는 과정에서의 미묘한(또는 노골적인) 차이는 조직 전체에 걸쳐 데이터를 매니지먼트하는 데 있어 난제를 야기한다. 이와 동시에, 이해관계자는 조직의 데이터가 일관성이 있어야 한다고 가정하고 있으며, 데이터 매니지먼트의 목표 또한 일관성이 존재하고 유지되는 데이터를 통해 다양한 데이터 소비자들이 실제로 데이터를 활용할 수 있도록 하는 것이다. 데이터 거버넌스가 점점 더 중요해지고 있는 한 가지 이유는 조직이 수직적 데이터로 데이터에 대한 결정을 내릴 수 있도록 돕기 위함이다.

2.5.8 다른 관점에 관한 설명

오늘날의 조직에서는 내부적으로 생성한 데이터와 외부 소스로부터 획득한 데이터를 사용한다. 그들은 국가 및 산업 라인에 걸쳐 서로 다른 법적 및 규정 준수 요건을 설명해야 한다. 데이터를 만드는 사람들은 종종 다른 사람이 나중에 그 데이터를 사용할 것이라는 것을 망각한다. 데이터의 잠재적 사용 가능성에 대한 인지는 데이터 수명주기에 대한 더 나은 계획과 더 나은 품질의 데이터를 위한 더 나은 계획을 가능하게 한다. 또한 데이터도 오용될 수 있다. 이 위험 역시 미리 설명하면 오용의 가능성이 줄어든다.

2.5.9 데이터 수명주기

다른 자산과 마찬가지로 데이터에도 수명주기가 존재한다. 데이터 자산을 효과적으로 매니지먼트하기 위해 조직은 데이터 수명주기를 이해하고 계획해야 한다. 잘 관리되는 데이터는 조직이 데이터를 어떻게 사용할 것인지에 대한 비전을 가지고 전략적으로 매니지먼트되는 데이터이다. 전략적 조직은 데이터 콘텐츠 요구사항뿐만 아니라 데이터 매니지먼트 요구사항도 규정한다. 여기에는 사용, 품질, 제어 및 보안에 대한 정책과 기대, 아키텍처와 설계에 대한 전사적 접근법, 인프라와 소프트웨어 개발 모두에 대한 지속 가능한 접근법이 포함된다. 데이터 수명주기는 제품 수명주기에 기초한다. 이는 시스템 개발 라이프사이클과 혼동되면 안된다. 개념적으로 데이터 수명주기는 설명하기 쉽습니다(그림2 참조). 데이터를 생성하거나 얻는 프로세스, 데이터를 이동, 변환, 저장하여 유지·공유할 수 있는 프로세스, 데이터를 사용하거나 적용하는 프로세스, 폐기하는 프로세스 등이 이에 포함된다. 수명주기 전반에 걸쳐 데이터는 정리, 변환, 병합, 강화 또는 통합될 수 있다. 데이터가 사용되거나 고도화하면서 새로운 데이터가 생성되는 경우가 많기에 수명주기 내에는 다이어그램에는 표시되지 않는 내부적인 반복 단계도 존재한다. 데이터는 정적인 경우가 드물다. 이에 데이터 매니지먼트에도 또한 데이터 수명주기에 맞춰 상호 정렬되고 연결된 일련의 과정이 포함된다. 데이터에는 수명 주기가



[그림2] 데이터 라이프사이클 주요활동

있을 뿐만 아니라 계통(즉, 데이터 체인(Data Chain)이라고도 함)도 있기 때문에 특정 조직 내에서 데이터 수명주기의 구체적인 내용은 상당히 복잡할 수 있다. 데이터 계통을 이해하려면 데이터셋의 출처를 문서화하는 것은 물론, 데이터셋이 사용될 때 시스템을 통한 데이터셋의 이동과 변환이 필요하다. 수명주기와 데이터 계통이 교차하며 서로에 대해 이해할 수 있게 된다. 조직이 데이터의 수명주기와 계통을 더 잘 이해할수록 데이터를 더 잘 관리할 수 있게 된다.

데이터 수명주기에 대한 데이터 매니지먼트의 초점은 다음과 같은 몇 가지 중요한 의미를 가진다.

- 생성과 사용은 데이터 수명주기에서 가장 중요한 사항임: 데이터 매니지먼트는 데이터 생성 및 획득 방법과 데이터 사용 방법을 이해하여 실행되어야 한다. 데이터 생성은 비용을 수반한다. 데이터는 소비되거나 적용될 때만 가치가 있다.
- 데이터 품질은 데이터 수명주기 전체에 걸쳐 매니지먼트되어야 함: 데이터 품질 매니지먼트는 데이터 매니지먼트의 중심이다. 낮은 품질의 데이터는 가치보다는 비용과 위험을 내포한다. 조직들은 종종 데이터의 품질을 매니지먼트하는 데 어려움을 겪는데, 그 이유는 앞서 서술한 바와 같이 데이터는 종종 비즈니스 운영 과정의 부산물로 생겨나는 경우가 많고, 조직 또한 품질에 관한 명시적인 표준을 정하지 않는 경우가 많기 때문이다. 품질은 수명주기 이벤트의 범위에 의해 영향을 받을 수 있으므로 품질은 데이터 수명주기의 일부로 계획되어야 한다.
- 데이터 수명주기를 통해 메타데이터 품질을 매니지먼트해야 함: 메타데이터는 데이터의 한 형태고, 조직이 다른 데이터를 매니지먼트하기 위해 메타데이터에 의존하기 때문에 메타데이터 품질은 다른 데이터의 품질과 동일한 방식으로 매니지먼트되어야 한다.
- 데이터 보안은 데이터 수명주기 전체에 걸쳐 매니지먼트되어야 함: 데이터 매니지먼트에는 데이터의 보안과 데이터 관련 리스크의 완화가 포함된다. 보호가 필요한 데이터는 생성부터 폐기까지 수명주기 전체에 걸쳐 보호되어야 한다.

· 데이터 매니지먼트 작업은 가장 중요한 데이터에 집중해야 함: 조직은 많은 데이터를 생산함에도 실제로는 그 중 많은 부분은 사용하지 않는다. 모든 데이터를 매니지먼트하려고 하는 것은 불가능하다. 수명주기 매니지먼트를 위해서는 조직의 가장 중요한 데이터에 초점을 맞추고 데이터 ROT(중복(Redundant)되고 구식(Obsolete)이며 사소한(Trivial) 데이터)를 최소화해야 한다(Aiken, 2014).

2.5.10 다양한 종류의 데이터

수명주기 매니지먼트 요구사항이 각기 다른 다양한 데이터 유형이 있다는 점에서 데이터 매니지먼트는 더욱 복잡해진다. 모든 매니지먼트 시스템은 관리되는 객체를 분류할 필요가 있다. 데이터는 데이터 유형(예: 트랜잭션 데이터, 참조 데이터, 마스터 데이터, 메타데이터, 범주형 데이터, 리소스 데이터, 이벤트 데이터, 상세 트랜잭션 데이터) 또는 콘텐츠(예: 데이터 도메인, 주제 영역) 또는 데이터에 필요한 보호 수준별로 분류할 수 있다. 또한 데이터는 저장 방법과 장소에 따라 분류될 수 있다. 데이터 유형마다 요구사항이 다르고, 리스크가 서로 다르며, 조직 내에서 서로 다른 역할을 하기 때문에, 데이터 매니지먼트의 많은 톨은 분류와 통제 측면에 초점을 맞추고 있다(Bryce, 2005). 예를 들어, 마스터 데이터는 서로 다른 용도를 가지고 있으며, 결과적으로 트랜잭션 데이터와 다른 매니지먼트 사항을 요구한다.

2.5.11 데이터와 리스크

데이터는 가치를 나타낼 뿐만 아니라 위험을 나타낸다. 저품질 데이터(정확하지 않음, 불완전함 또는 구식)는 분명히 그 정보가 정확하지 않기 때문에 위험을 나타낸다. 그러나 데이터는 오해와 오용될 수 있기 때문에 위험하기도 하다. 조직은 최고 품질 데이터(사용 가능, 목적적합, 완전, 정확, 일관성, 시기적절, 사용 가능, 의미 및 이해)로부터 최대의 가치를 얻는다. 하지만, 많은 중요한 결정들을 위해, 우리는 정보격차 즉, 우리가 알고 있는 것과 효과적인 결정을 내리기 위해 알아야 할 것의 차이를 가지고 있다. 정보 격차는 운영 효과와 수익성에 잠재적으로 심대한 영향을 미칠 수 있는 기업 부채를 나타낸다. 고품질 데이터의 가치를 인식하는 조직은 규제 및 윤리적 문화 프레임워크 내에서 데이터와 정보의 품질과 사용성을 개선하기 위해 구체적인 사전 예방적인 조치를 취할 수 있다. 모든 분야에 걸친 조직 자산으로서의 정보의 역할 증가는 규제 당국과 입법자들에 의해 정보의 잠재적인 사용과 남용에 관한 이목을 집중시켰다. Sarbanes-Oxley(거래에서 대차대조표까지의 금융거래 데이터의 정확성과 타당성에 대한 통제에 초점을 맞춘 정책)부터 Solbanes-Oxley(보험 부문의 위험 모델과 자본적정성을 뒷받침하는 데이터 계통 및 품질에 초점을 맞춘 정책)에 이르기까지, 지난 10년 동안 급속도로 성장한 데이터 프라이버시 규정(광범위한 산업과 관할권에 걸친 사람에 대한 데이터 처리 포함)에 따르면, 대차대조표에 정보를 자산으로 표시하는 회계처리를 우리는 여전히 기다리고 있지만, 규제 환경은 적절한 완화 및 통제가 적용되는 위험 항목에서 이를 볼 수 있게 될 것으로 점점 더 기대하고 있다. 마찬가지로, 소비자들은 자신의 데이터가 어떻게 사용되는지를 더 잘 알게 되면서 프로세스의 보다 원활하고 효율적인 운영뿐만 아니라 정보 보호와 사생활에 대한 존중도 기대할 수 있게 되었다.

이는 데이터 매니지먼트 전문가로서 전략적 이해당사자가 누구인지에 대한 범위가 전통적으로 그랬을 수 있는 범위보다 넓을 수 있다는 것을 의미한다. 정보 매니지먼트의 대차대조표 영향은 불행하게도 이러한 리스크가 매니지먼트되지 않고 주주들이 자신의 주식 포트폴리오로 투표하고 규제당국이 벌금이나 영업제한을 부과하고 고객이 지갑을 가지고 투표할 때 너무 자주 발생한다.

2.5.12 데이터 매니지먼트 및 IT

서론 등에서 지적인 바와 같이, 데이터 매니지먼트 활동은 광범위하며, 기술력과 비즈니스 능력을 모두 필요로 한다. 오늘날 거의 모든 데이터가 전자적으로 저장되기 때문에 데이터 매니지먼트 전문은 기술의 영향을 강하게 받는다. 초기부터 데이터 매니지먼트 개념은 기술 매니지먼트와 깊이 얽혀 있었다. 이 같은 전통은 계속될 전망이다. 많은 조직에서, 새로운 기술을 구축하려는 추진력과 보다 신뢰할 수 있는 데이터를 확보하려는 열망 사이에 긴장이 지속되고 있다. 마치 이 둘이 서로에게 필요한 것이 아니라 서로 대립했던 것처럼 말이다. 성공적인 데이터 매니지먼트를 위해서는 기술에 대한 현명한 결정이 필요하지만, 기술 매니지먼트는 데이터 매니지먼트와는 같지 않다. 조직은 데이터에 대한 기술적 유혹이 데이터에 미치는 영향을 이해해야 한다. 대신, 비즈니스 전략과 일치하는 데이터 요구사항은 기술에 대한 결정을 촉진해야 한다.

2.5.13 효과적인 데이터 매니지먼트는 리더십과 커밋먼트를 필요로 함

리더스 데이터 매니페스토(The Leader's Data Manifesto, 2017)는 “조직의 유기적 성장을 위한 최선의 기회는 데이터에 있다(Organization's best opportunities for organic growth lie in data)”고 밝혔다. 대부분의 조직이 데이터를 자산으로 인식하고 있지만 데이터 중심적인 것과는 거리가 먼 것이 현실이다. 많은 이들은 어떤 데이터를 보유하고 있는지, 어떤 데이터가 비즈니스에 가장 중요한지 잘 알지 못한다. 그들은 데이터와 정보 기술을 혼동하고 둘 다 잘못 관리한다. 그들은 전략적으로 데이터에 접근하지 않는다. 그리고 그들은 데이터 매니지먼트와 관련된 일을 과소평가한다. 이러한 조건들은 데이터 매니지먼트와 조직의 성공 가능성에 중요한 요소, 즉 헌신적인 리더십과 조직의 모든 수준의 모든 사람의 개입의 어려움을 가중시킨다. 이 과제에 관해 다음과 같은 점은 확실하게 해 두어야 한다. 데이터 매니지먼트는 쉽지도 않고 단순하지도 않다. 하지만 이를 잘 하는 조직은 거의 없기 때문에, 그것은 대부분 개발되지 않은 기회의 원천이라는 사실이다. 이를 더 잘하기 위해서는 비전, 계획, 변화의 의지가 필요하다. 최고 데이터 책임자(CDO)의 역할을 옹호하는 것은 데이터 매니지먼트가 고유한 과제를 제시하며 성공적인 데이터 매니지먼트는 IT가 아닌 비즈니스 중심이어야 한다는 인식에서 비롯된다. CDO는 데이터 매니지먼트 이니셔티브를 주도하고 조직이 데이터 자산을 활용하고 이들로부터 경쟁 우위를 확보할 수 있도록 할 수 있다. 그러나 CDO의 역할은 이니셔티브를 주도하는 것에서 그치지 않다. 그는 또한 조직이 데이터에 대한 보다 전략적인 접근을 할 수 있도록 하는 문화적 변화를 이끌어야 한다.

2.6 데이터 매니지먼트 전략

전략은 높은 수준의 목표를 달성하기 위한 높은 수준의 행동 과정을 함께 도표화하는 일련의 선택과 결정을 말한다. 체스 게임에서의 전략은 체크메이트가 이기거나 교착상태에서 살아남기 위한 일련의 움직임이다. 전략 계획은 높은 수준의 목표를 달성하기 위한 고도의 행동 과정이다. 데이터 전략에는 정보를 경쟁 우위에 사용하고 기업 목표를 지원하기 위한 사업 계획이 포함되어야 한다.

데이터 전략은 비즈니스 전략에 내재된 데이터 요구사항으로부터 파생되어야 한다. 이 요구사항에는 조직이 필요한 데이터가 무엇인지, 어떻게 조직이 데이터를 확보할 것인지, 데이터를 어떻게 매니지먼트하고 시간이 지남에 따라 그 신뢰성을 보장할 것인지, 이를 어떻게 사용할 것인지 등이 포함된다. 일반적으로 데이터 전략에는 데이터 품질, 데이터 무결성, 접근 및 보안의 유지 및 개선을 위한 계획인 데이터 매니지먼트 프로그램 전략에 관한 지원이 필요하다. 또한 이 전략은 데이터 매니지먼트와 관련된 알려진 문제를 해결해야 한다.

많은 조직에서 데이터 매니지먼트 전략은 CDO가 소유하고 유지하며 데이터 거버넌스 위원회에서 지원하는 데이터 거버넌스 팀을 통해 제정된다. 종종 CDO는 데이터 매니지먼트 책임과 거버넌스 확립에 대한 고위 경영진의 의지를 얻기 위해 데이터 거버넌스 위원회가 구성되기 전부터 초기 데이터 전략과 데이터 매니지먼트 전략을 입안한다.

데이터 매니지먼트 전략의 구성요소에는 다음이 포함되어야 한다.

- 데이터 매니지먼트에 대한 강력한 비전
- 데이터 매니지먼트를 위한 요약 비즈니스 사례(선택한 예 포함)
- 원칙, 가치 및 매니지먼트 관점 안내
- 데이터 매니지먼트의 미션 및 장기적인 방향 목표
- 성공적인 데이터 매니지먼트를 위해 제안된 조치
- SMART(구체적인(Specific), 측정가능한(Measurable), 실행가능한(Actionable), 현실적인(Realistic), 시간이 정해져 있는(Time-bound))한 단기(12-24개월) 데이터 매니지먼트·프로그램 목표 설정
- 데이터 매니지먼트 역할 및 조직에 대한 설명과 책임 및 의사결정 권한 요약
- 데이터 매니지먼트 프로그램 구성 요소 및 이니셔티브에 대한 설명
- 업무 제한을 가진 우선순위 작업 프로그램
- 프로젝트 및 실행 항목이 포함된 로드맵 초안

데이터 매니지먼트를 위한 전략적 계획 수립의 결과물은 다음과 같다.

- 데이터 매니지먼트 차터(Charter): 전체적인 비전, 비즈니스 사례, 목표, 지침 원칙, 성공의 척도, 중요한 성공 요인, 인식된 리스크, 운영 모델 등
- 데이터 매니지먼트 범위 설명: 개별 계획(보통 3년)의 목표 및 이러한 목표를 달성하는 데 책임이 있는 개별 리더와 조직 및 구성원의 역할
- 데이터 매니지먼트 구현 로드맵: 특정 프로그램, 프로젝트, 작업 할당 및 전송 마일스톤 식별

데이터 매니지먼트 전략은 조직과 관련된 모든 데이터 매니지먼트 프레임워크 지식 영역(Data Management Framework Knowledge Areas)을 다루어야 한다.





3 데이터 매니지먼트 프레임워크 (Data Management Framework)

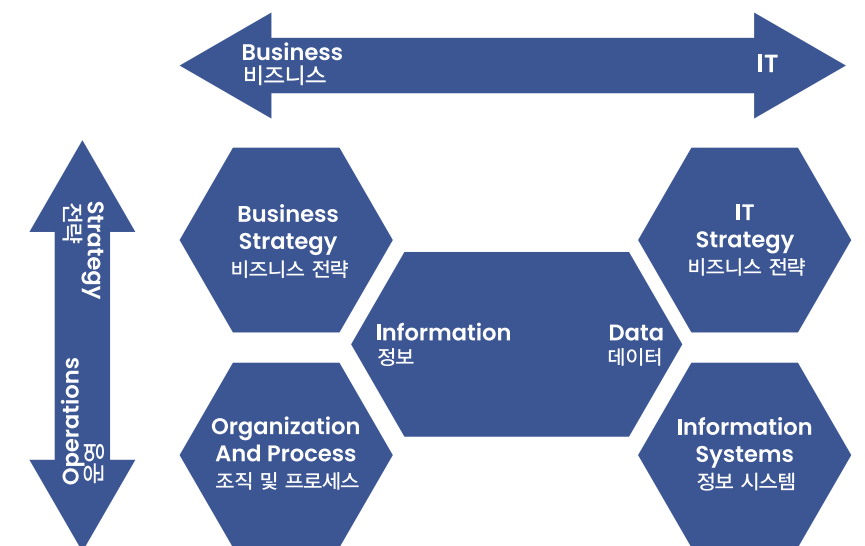
데이터 매니지먼트는 각각의 목표, 활동 및 책임을 갖는 일련의 상호의존적 기능을 포함한다. 데이터 매니지먼트 전문가는 전략적 및 운영 목표, 특정 비즈니스 및 기술 요구사항, 리스크 및 컴플라이언스 요구사항, 데이터가 무엇을 대표하는지, 데이터가 높은 자격을 갖는지 여부에 대한 상충되는 이해의 균형을 유지하면서 추상적 기업 자산에서 가치를 도출하는 데 내재된 과제를 고려해야 한다.

추적해야 할 부분이 많기에 데이터 매니지먼트를 종합적으로 이해하고 구성요소 간 관계를 파악할 수 있는 프레임워크를 갖추는 것이 이 때 도움이 된다. 기능은 서로 의존하고 조정되어야 하기 때문에, 어떤 조직에서든 데이터 매니지먼트의 다른 측면을 담당하는 사람들은 조직이 데이터로부터 가치를 이끌어내려면 협업이 필요하다.

다양한 추상화 수준에서 개발된 프레임워크는 데이터 매니지먼트에 접근하는 방법에 대한 다양한 관점을 제공한다. 이러한 관점은 전략을 명확히 하고 로드맵을 개발하며 팀을 구성하고 기능을 조정하는 데 사용할 수 있는 통찰력을 제공한다.

이 보고서에 제시된 아이디어와 개념은 기관마다 다르게 적용될 것이다. 데이터 매니지먼트에 대한 조직의 접근방식은 산업, 데이터 사용 범위, 문화, 성숙도 수준, 전략, 비전, 그리고 해결 중인 특정 과제와 같은 주요 요소에 따라 달라지기 때문이다. 여기에서 설명하는 프레임워크는 데이터 매니지먼트를 보고 제시된 개념을 적용하는 데 필요한 시야를 제공한다.

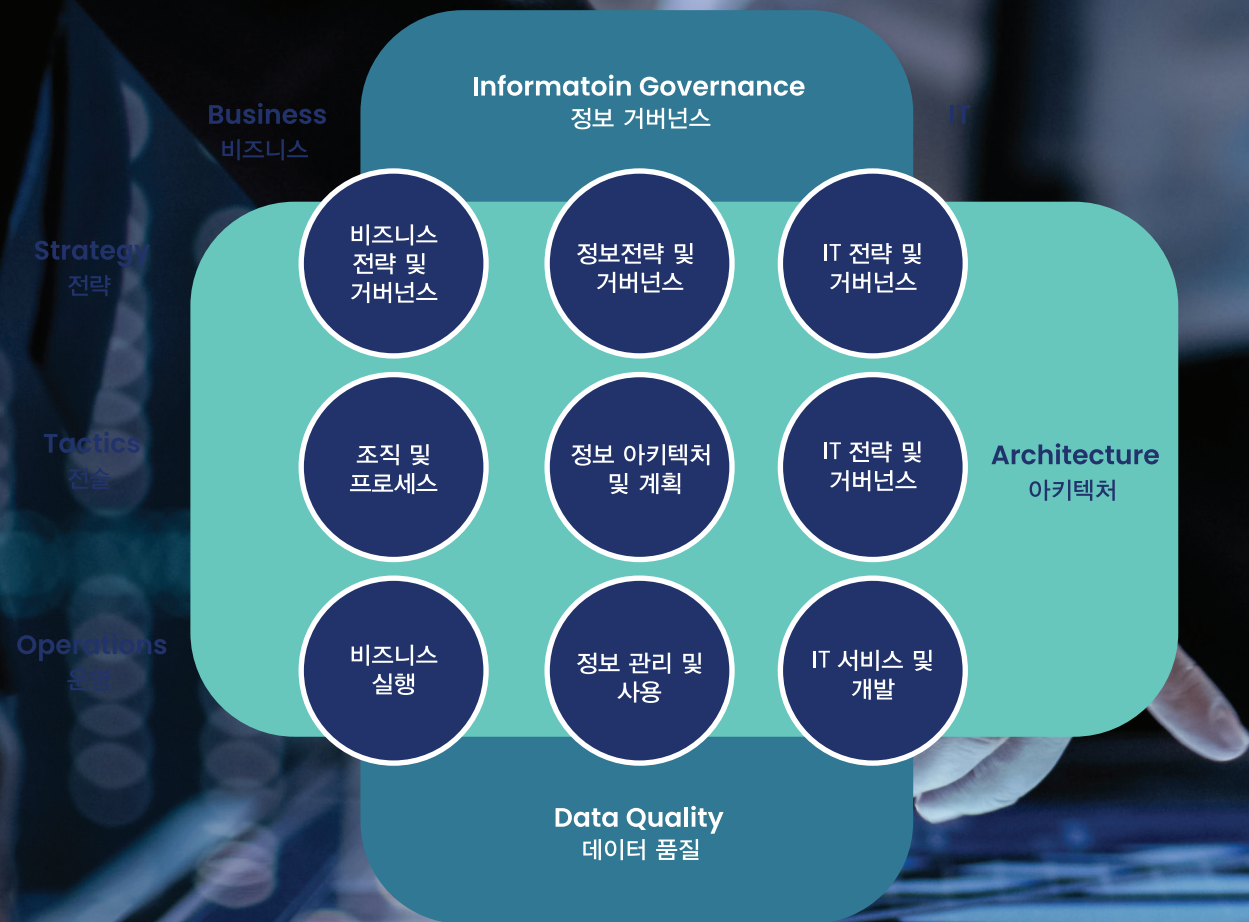
- 첫 번째 두 가지인 전략적 정렬 모델(Strategic Alignment Model)과 암스테르담 정보 모델(Amsterdam Information Model)은 조직이 데이터를 매니지먼트하는 방식에 영향을 미치는 높은 수준의 관계들을 보여준다.
- 데이터 매니지먼트 지식 체계 프레임워크(DAMA 휠, 육각형 및 문맥 다이어그램)는 DAMA 에서 정의한 데이터 매니지먼트 지식 영역을 설명하고 시각적 표현 방법을 보여준다.
- 마지막 두 가지는 DAMA 휠을 출발점으로 삼고, 각 부분 간의 관계를 더 잘 이해하고 설명하기 위해 각 부분들을 재배열한다.



[그림3] 전략적 연계 모델(Strategic Alignment Model, Adapted by Henderson and Venkatraman)

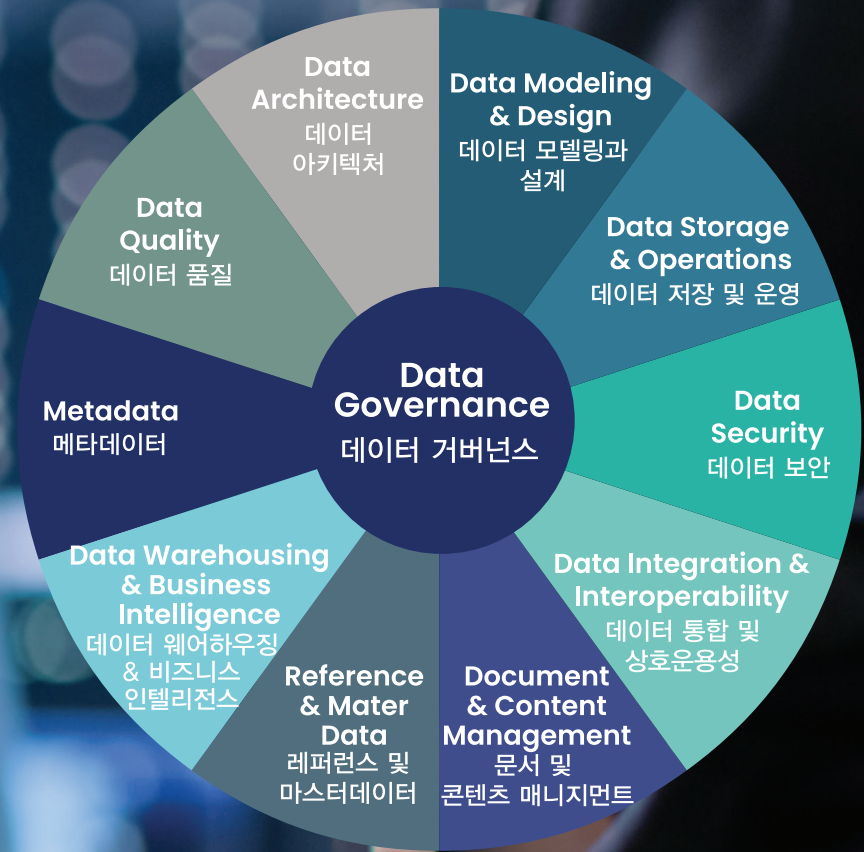
3.1 전략적 연계 모델(Strategic Alignment Model)

전략적 연계 모델(Henderson and Venkatraman, 1999)은 데이터 매니지먼트에 대한 모든 접근방식의 근본적인 동인을 추상화한다. 그 중심에는 데이터와 정보의 관계가 있다. 정보는 사업 전략 및 데이터의 운영적 사용과 가장 밀접하고 자주 관련되어 있다. 데이터는 데이터 접근을 가능하게 하는 시스템의 물리적 매니지먼트를 지원하는 정보 기술 및 프로세스와 관련된다. 이 개념을 둘러싸고 전략 선택의 4대 기본 영역인 비즈니스 전략, 정보기술 전략, 조직 인프라 및 프로세스, 정보기술 인프라 및 프로세스 등이 있다. 완전하게 연결된 전략적 정렬 모델은 그림 3에 나타난 것보다 더 복잡하다. 각 코너의 육각형에는 고유의 차원이 있습니다. 예를 들어, 비즈니스 전략과 IT 전략 모두에서 범위, 역량, 거버넌스를 고려할 필요가 있다. 운영은 인프라, 프로세스 및 기술을 고려해야 한다. 조각들 간의 관계는 조직이 다른 요소들의 전략적 적합성과 조각들의 기능적 통합을 이해하는 데 도움이 된다. 모델에 대한 높은 수준의 묘사는 데이터와 데이터 매니지먼트에 대한 의사결정에 영향을 미치는 조직 요소를 이해하는 데 유용하다.

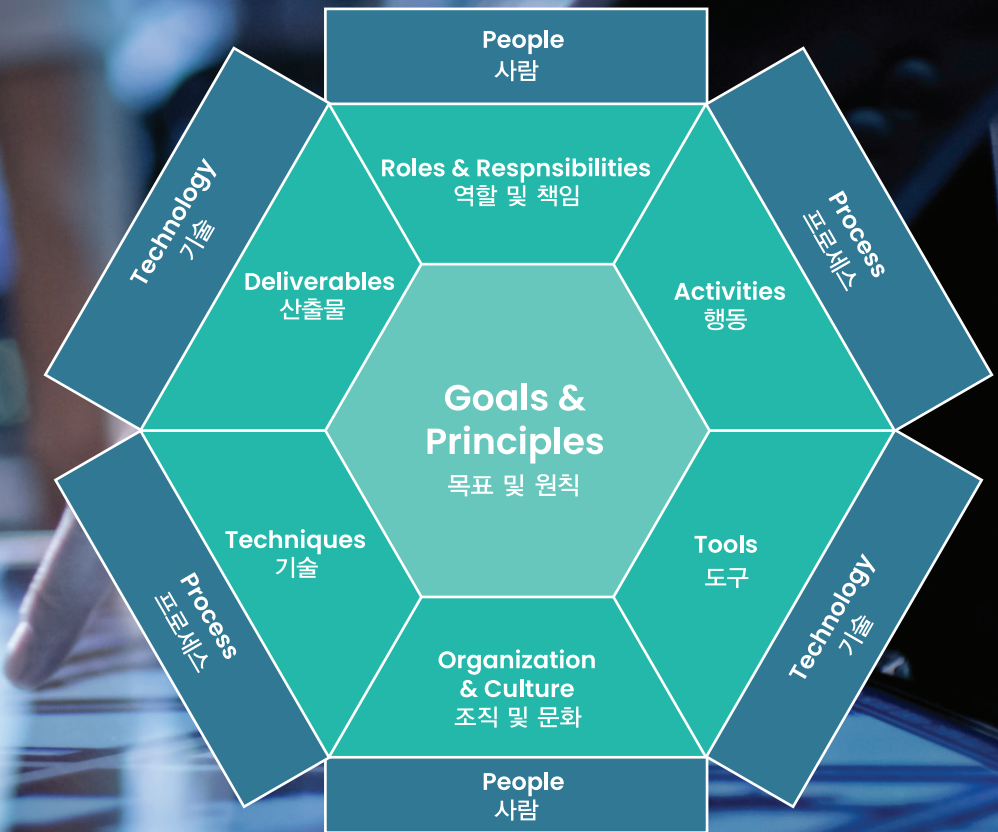


[그림4] 데이터 라이프사이클 주요활동

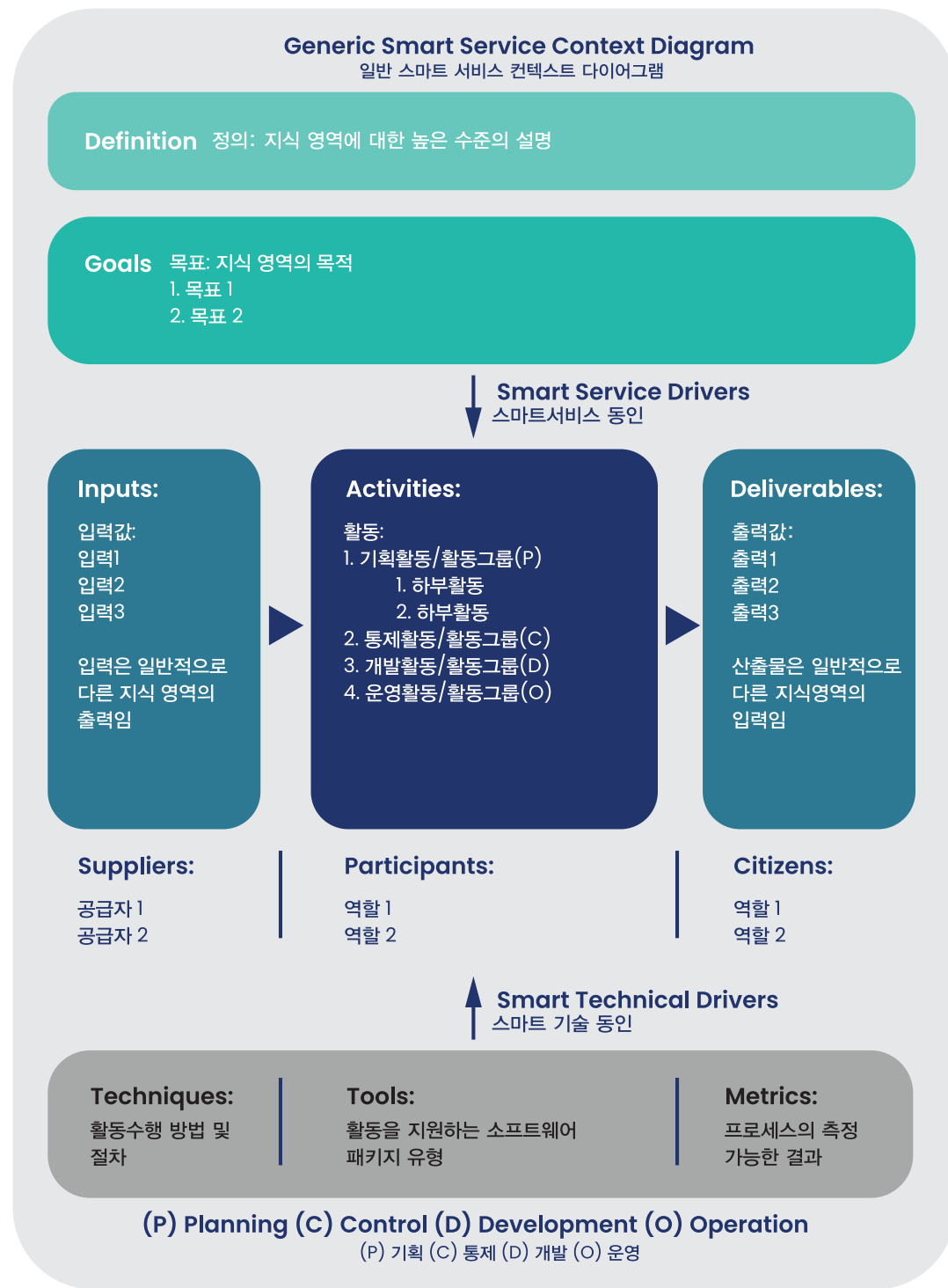
Technical Standards for Smart Cities
스마트시티 국제표준화 기반조성(R&D)



[그림5] 데이터 매니지먼트 프레임워크(DAMA Wheel)



[그림6] 데이터 매니지먼트 프레임워크(DAMA Wheel)



[그림7] 지식 영역 컨텍스트 다이어그램

3.2 암스테르담 정보 모형(The Amsterdam Infomation Model)

암스테르담 정보 모형은 전략적 제후 모형과 마찬가지로 비즈니스 및 IT 제후에 대한 전략적 관점을 취하고 있으며(Abuwer, Maes, Truijens, 1997), 9-cell로 알려져 계획과 아키텍처를 포함한 구조와 전술에 초점을 맞춘 중간층을 인정한다. 더욱이 이 모형은 정보 통신의 필요성(그림 4의 정보 거버넌스 및 데이터 품질 축으로 표현)을 인식하고 있다. SAM과 AIM프레임워크의 작성자는 수평적(비즈니스/IT전략)과 수직적(비즈니스 전략/비즈니스 운영) 관점에서 구성요소 간의 관계를 상세히 설명한다.

3.3 데이터 매니지먼트 지식 체계 프레임워크 (DAMA-DMBOK Framework)

데이터 매니지먼트 지식 체계 프레임워크는 데이터 매니지먼트의 전체적인 범위를 구성하는 지식 영역에 대해 더 심층적으로 고찰한다. 데이터 매니지먼트 프레임워크에 대한 세 가지 시각적 묘사는 아래와 같다:

- DAMA 휠(The DAMA Wheel, 그림5)
- 환경 요인 육각형(The Environmental Factors Hexagon, 그림6)
- 지식 영역 컨텍스트 다이어그램(The Knowledge Area Context Diagram, 그림7)

DAMA 휠은 데이터 매니지먼트 지식 영역을 정의한다. 이 모형은 기능 간의 일관성과 균형을 위해 거버넌스가 필요하기 때문에 데이터 매니지먼트 활동의 중심에 데이터 거버넌스를 둔다. 다른 지식 영역(데이터 아키텍처, 데이터 모델링 등)은 휠을 중심으로 균형을 이룬다. 이 모두는 성숙한 데이터 매니지먼트를 위해 필요한 부분이지만, 조직의 요건에 따라 다른 시기에 구현될 수도 있다(그림5 참조).

환경 요인 육각형은 사람, 프로세스, 기술 사이의 관계를 보여주고 지식 영역 컨텍스트 다이어그램을 해석할 수 있는 열쇠를 제공한다. 이 모형은 목표와 원칙을 중심에 두고 있는데, 이는 사람들이 어떻게 활동을 실행하고 성공적인 데이터 매니지먼트에 필요한 도구를 효과적으로 사용해야 하는지에 대한 지침을 제공하기 때문이다(그림6 참조).

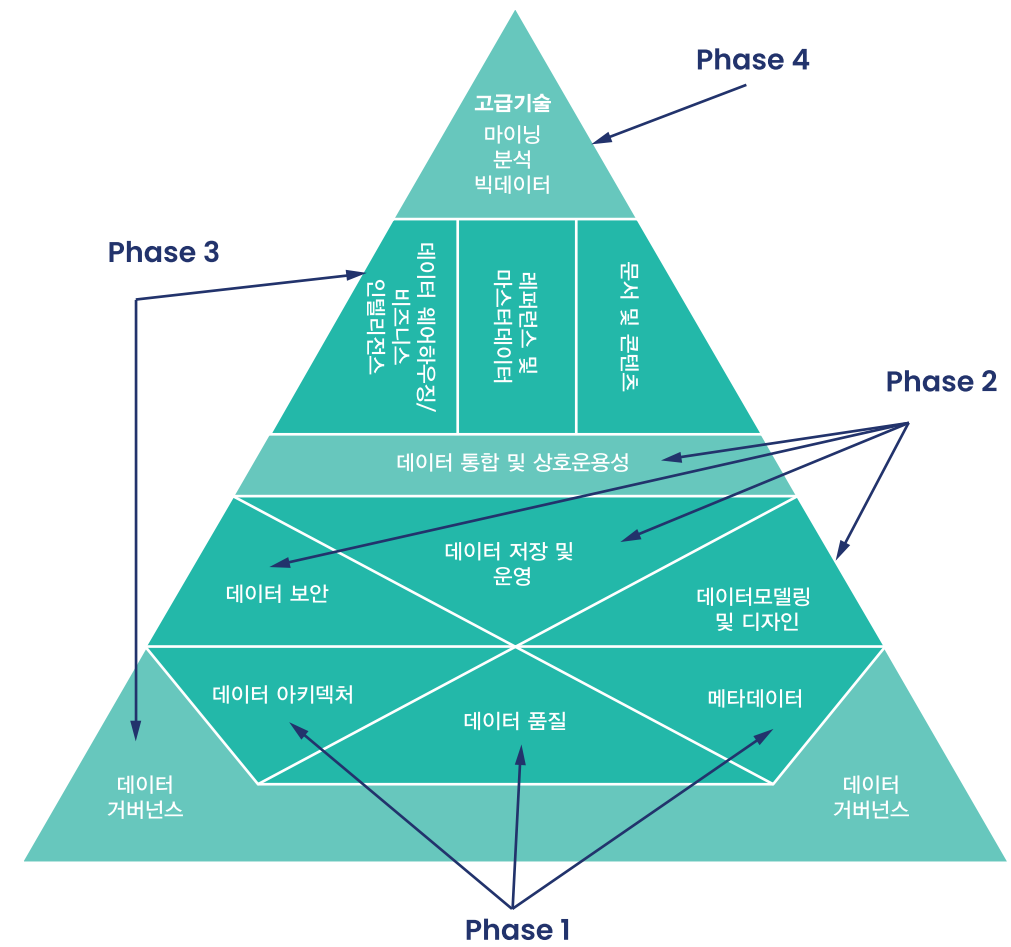
지식 영역 컨텍스트 다이어그램(그림7 참조)은 사람, 프로세스 및 기술과 관련된 세부사항을 포함하여 지식 영역의 세부사항을 설명한다. 서비스 매니지먼트에 사용되는 SIPOC 다이어그램(공급자(Suppliers), 입력값(Inputs), 프로세스(Processes), 출력값(Outputs) 및 소비자(Consumers)의 개념을 기반으로 한다. 컨텍스트 다이어그램은 이해관계자의 요구사항을 충족하는 결과물을 생산하기 때문에 활동을 중심에 둔다.

각 컨텍스트 다이어그램은 지식 영역의 정의와 목표로 시작한다. 목표(중심)를 추진하는 활동은 계획(Plan), 개발(Develop), 운영(Operate), 제어(Control)의 4단계로 분류된다. 왼쪽에는 (활동으로 유입되는) 입력값 및 공급자가 있다. 우측(활동에서 흘러나오는 흐름)에는 출력값과 소비자가 있다. 참가자는 활동 아래에 열거되어 있다. 아래에는 지식 영역의 측면에 영향을 미치는 도구, 기술 및 평가지표가 있다.

컨텍스트 다이어그램의 목록은 전체적이지 않고 예시적이다. 각 항목은 조직마다 다르게 적용된다. 높은 수준의 역할 목록에는 가장 중요한 역할만 포함된다. 각 조직은 자신의 필요에 따라 이러한 패턴을 조정할 수 있다.

컨텍스트 다이어그램의 구성 요소에는 다음이 포함된다.

1. 정의(Definition): 이 절에서는 지식 영역을 간결하게 정의한다.
2. 목표(Goals)는 지식 영역의 목적과 각 지식 영역 내에서 활동의 수행을 안내하는 기본 원칙을 설명한다.
3. 활동(Activities)은 지식 영역의 목표를 달성하기 위해 필요한 행동과 과제이다. 일부 활동은 하위 활동, 과제 및 단계에 대해 설명한다. 활동은 계획(Plan), 개발(Develop), 운영(Operate), 통제(Control)의 4가지 범주로 분류된다.
 - a. (P) 계획 활동(Planning Activities)은 데이터 매니지먼트 목표를 달성하기 위한 전략 및 기술적 과정을 설정한다. 기획활동은 반복적으로 일어난다.
 - b. (D) 시스템 개발 수명주기(System Development Lifecycle; SDLC)를 중심으로 개발 활동(Development Activities)을 조직한다. (분석, 설계, 구축, 테스트, 준비 및 배치).
 - c. (C) 제어 활동(Control Activities)은 데이터의 지속적인 품질과 데이터에 액세스하고 사용하는 시스템의 무결성, 신뢰성 및 보안을 보장한다.
 - d. (O) 운영 활동(Operational Activities)은 데이터에 액세스하고 사용하는 시스템과 프로세스의 사용, 유지보수 및 개선을 지원한다.
4. 입력값(Inputs)이란 각 지식영역의 활동을 개시하기 위해 필요한 가시적(Tangible)인 것이다. 많은 활동은 동일한 입력을 요구한다. 예를 들어, 많은 사람들은 비즈니스 전략에 대한 지식을 입력으로 요구한다.
5. 출력값은 지식영역 내 활동의 산출물(Deliverables)로서, 각 기능이 생산을 담당하는 가시적인 것이다. 출력값은 그 자체로 종료되거나 다른 활동에 투입될 수 있다. 몇 가지 주요 출력값들은 여러 기능에 의해 생성된다.
6. 역할과 책임(Role and Respsnsibilities)은 개인과 팀이 지식 영역 내의 활동에 어떻게 기여하는지를 설명한다. 역할은 대부분의 조직에서 요구되는 역할 그룹에 초점을 두고 개념적으로 설명된다. 개인의 역할은 기술과 자격 요건에 따라 정의된다. 정보 시대를 위한 기술 프레임워크(SFIA)는 역할 제목을 일치시키기 위해 사용된다. 많은 역할들이 교차 기능할 것이다.
7. 공급자(Suppliers)는 활동에 대한 투입물에 대한 접근을 제공하거나 가능하게 할 책임이 있는 사람들이다.
8. 주요 산출물의 직접적인 이익을 얻는 소비자(Consumers)는 데이터 매니지먼트 활동에 의해 생성된다.
9. 참가자(Participants)는 지식 영역의 활동을 수행, 매니지먼트 또는 승인하는 사람이다.
10. 도구(Tools)는 지식 영역의 목표를 가능하게 하는 응용 프로그램 및 기타 기술이다.
11. 기법(Techniques)은 지식 영역 내에서 활동을 수행하고 산출물을 생산하기 위해 사용되는 방법과 절차이다. 기법에는 공통 규약, 모범 사례 권고사항, 표준 및 프로토콜, 해당되는 경우 새롭게 등장하는 대안적 접근법 등이 포함된다.
12. 평가지표(Metrics)는 성능, 진행률, 품질, 효율성 또는 기타 효과의 측정이나 평가를 위한 기준이다. 평가지표 섹션은 각 지식 영역 내에서 수행되는 작업의 측정 가능한 측면을 식별한다. 개선이나 가치와 같은 추상적인 특성도 평가지표가 될 수 있다.

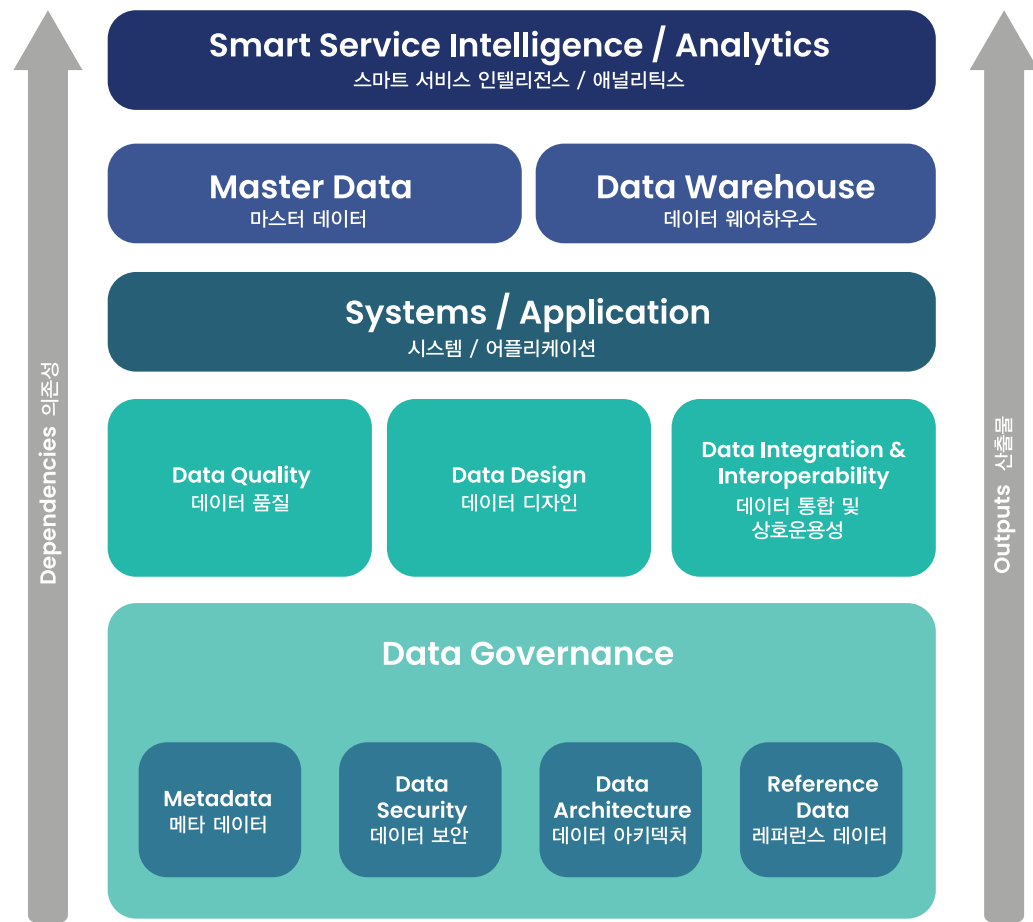


[그림8] 구축한 데이터베이스 수용력
(Golden Pyramid figure copyright Data Blueprint, used with permission)

DAMA 휠은 지식 영역 세트를 높은 수준으로 제시하는 반면, 육각형은 지식 영역 구조의 구성요소를 인식하며, 상황도는 각 지식 영역 내에 세부 사항을 제시한다. 기존의 데이터 매니지먼트 프레임워크의 어느 부분도 서로 다른 지식 영역 간의 관계를 기술하지 않는다. 그 문제를 해결하기 위한 노력은 다음 두 섹션에 설명되어 있는 DAMA 프레임워크의 개혁을 이끌었다.

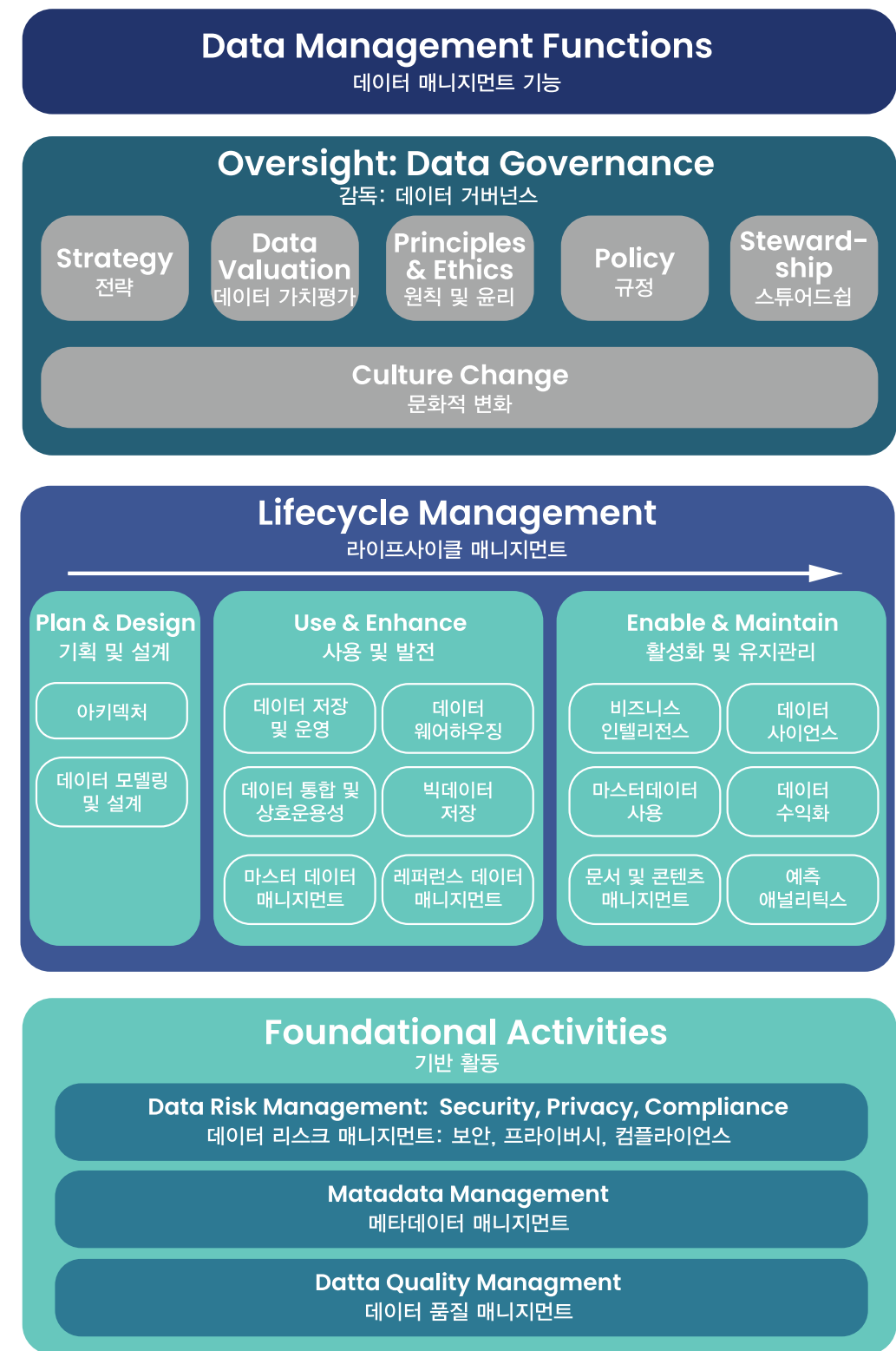
3.4 데이터 매니지먼트 지식 체계 피라미드 (DMBOK Pyramid; Aihen)

만약 질문을 받는다면, 많은 조직들은 그들의 데이터를 최대한 활용하기를 원한다고 말할 것이다 – 그들은 고급 기술(데이터 마이닝, 분석 등)들을 활용해 황금 피라미드를 쌓기 위해 노력하고 있다. 하지만 이 피라미드는 사실 더 큰 구조의 꼭대기에 불과하며 기초 위의 작은 첨탑에 불과하다. 대부분의 조직은 데이터 매니지먼트를 시작하기 전에 데이터 매니지먼트 전략을 정의하려고 하지 않는다. 대신에, 그들은 대부분 최적 조건보다 낮은 조건에서 그러한 능력을 지향한다. 피터 아이켄(Peter Aiken)의 프레임워크는 데이터 매니지먼트 지식 체계 내 기능적 영역을 사용하여 많은 조직들이 자신이 놓여있는 현실을 탐색할 수 있는 경우를 묘사한다. 조직은 이를 사용하여 전략적 비즈니스 목표를 지원할 수 있는 신뢰할 수 있는 데이터와 프로세스를 보유한 상태로 나아갈 수 있는 방법을 정의할 수 있다. 이러한 목표에 도달하기 위해 많은 조직들이 유사한 논리적 단계를 거친다 (그림8 참조).



[그림9] 기능 영역 의존성(DAMA Functional Area Dependencies)

- 1단계: 조직은 데이터베이스 기능을 포함하는 애플리케이션을 구입한다. 이는 조직이 데이터 모델링/설계, 데이터 스토리지 및 데이터 보안(예: 일부 사람을 들여보내고 다른 사람은 들어오지 못하게 하는 것)을 위한 출발점을 가지고 있음을 의미한다. 시스템이 그들의 환경 내에서 그리고 그들의 데이터와 함께 작동하도록 하기 위해서는 통합과 상호운용성에 대한 작업이 필요하다.
- 2단계: 일단 애플리케이션을 사용하기 시작하면 데이터의 품질에 대한 문제를 발견할 것이다. 그러나 보다 높은 품질의 데이터를 얻는 것은 신뢰할 수 있는 메타데이터와 일관된 데이터 아키텍처에 달려 있다. 이것들은 서로 다른 시스템의 데이터들이 어떻게 함께 작동하는지 명확하게 해준다.
- 3단계: 데이터 품질, 메타데이터 및 아키텍처 매니지먼트를 위한 규율화된 실행에는 데이터 매니지먼트 활동을 위한 구조적 지원을 제공하는 데이터 거버넌스가 필요하다. 또한 데이터 거버넌스는 문서 및 콘텐츠 매니지먼트, 참조 데이터 매니지먼트, 마스터 데이터 관리, 데이터 웨어하우스, 비즈니스 인텔리전스와 같은 전략적 이니셔티브의 실행을 가능하게 하여 황금 피라미드 내에서 고급 기술의 구현을 완벽하게 가능하게 한다.
- 4단계: 조직은 잘 관리되는 데이터의 이점을 활용하고 향상된 분석 기능을 누리게 된다.



[그림10] 데이터 매니지먼트 기능 프레임워크
(DAMA Data Management Function Framework)

Aiken의 피라미드는 DAMA 휠에서 파생되며 지식 영역 간의 관계를 시각적으로 알려준다. 이 관계는 모두 서로 교환할 수 있는 것은 아니며 다양한 종류의 상호의존성을 가지고 있기도 하다. 피라미드 프레임워크에는 두 개의 동인이 있다. 첫째는, 기초 위에 설계한다는 아이디어인데 이는 서로를 지원할 수 있는 각 부분들이 적절하게 구성되어 있는 것을 말한다. 둘째는 이와는 모순되는 생각인데 각 부분들이 임의의 순서로 배치되는 부분도 있다는 것이다.

3.5 데이터 매니지먼트 프레임워크의 확장

Aiken의 피라미드는 조직이 어떻게 더 나은 데이터 매니지먼트 실행을 위해 발전하는지를 설명한다. DAMA 지식 영역을 살펴보는 또 다른 방법은 그들 사이의 의존성을 탐구하는 것이다. Sue Geuens 가 개발한 그림 9의 프레임워크는 비즈니스 인텔리전스와 분석 기능이 다른 모든 데이터 매니지먼트 기능에 종속되어 있음을 인식한다. 이들은 마스터 데이터 및 데이터 웨어하우스 솔루션에 직접 의존하며 궁극적으로 공급 시스템과 응용 프로그램에 의존한다. 신뢰할 수 있는 데이터 품질, 데이터 설계 및 데이터 상호운용성 실행은 신뢰할 수 있는 시스템과 애플리케이션의 토대가 된다. 또한, 이 모델 내에서 메타데이터 매니지먼트, 데이터 보안, 데이터 아키텍처 및 참조 데이터 매니지먼트를 포함하는 데이터 거버넌스는 다른 모든 기능이 의존하는 기반을 제공한다.

DAMA 휠에 대한 세 번째 대안은 그림 10에 나와 있다. 이것은 또한 DAMA 지식 영역 사이의 일련의 관계를 제안하기 위해 아키텍처 개념을 끌어낸다. 이러한 관계를 명확히 하기 위해 일부 지식 영역의 내용에 대한 추가 세부사항을 제공한다. 프레임워크는 데이터 매니지먼트의 지침을 제공하려는 목적에서 시작한다. 조직이 다른 자산에서와 마찬가지로 데이터 자산에서 가치를 창출할 수 있도록 지원 가치를 도출하기 위해서는 라이프사이클 매니지먼트가 필요하므로, 데이터 라이프사이클과 관련된 데이터 매니지먼트 기능이 다이어그램의 중앙에 설명되어 있다. 여기에는 신뢰할 수 있는 고품질 데이터를 위한 계획 및 설계, 데이터를 사용할 수 있고 유지 매니지먼트할 수 있는 프로세스 및 기능 확립, 마지막으로 다양한 유형의 분석 및 프로세스를 통해 데이터를 사용하여 그 가치를 향상시키는 것이 포함된다.

라이프사이클 매니지먼트 섹션에서는 데이터의 전통적인 사용(비즈니스 인텔리전스, 문서, 콘텐츠 매니지먼트)을 지원하는 데 필요한 데이터 매니지먼트 설계 및 운영 기능(모델링, 아키텍처, 저장 및 운영 등)을 설명한다. 새롭게 등장하는 데이터 활용(데이터 사이언스, 예측 분석 등)을 지원하는 신흥 데이터 매니지먼트 기능(빅데이터 스토리지)도 포함한다. 데이터가 자산으로 진정으로 매니지먼트되는 경우, 조직은 데이터를 다른 조직에 판매함으로써(데이터 수익화) 자신의 데이터로부터 직접적인 가치를 얻을 수 있을 것이다.

직접적인 라이프사이클 기능에만 집중하는 조직은 기초 및 감독 활동을 통해 데이터 라이프사이클을 지원하는 조직만큼의 가치를 얻지 못할 것이다. 데이터 리스크 매니지먼트, 메타데이터 및 데이터 품질 관리와 같은 기초적인 활동은 데이터 라이프사이클에 걸쳐 있다. 이들은 설계 결정을 개선하고 데이터를 더 쉽게 사용할 수 있도록 한다. 이러한 것들이 잘 실행되면 유지 비용이 적게 들고, 데이터 소비자들은 이에 대한 신뢰도가 높아지며, 이를 사용할 기회가 확대될 것이다.

데이터 생산과 사용을 성공적으로 지원하고 기초적인 활동이 규율적으로 수행되도록 하기 위해 많은 조직들은 데이터 거버넌스의 형태로 감독을 확립한다. 데이터 거버넌스 프로그램은 조직이 데이터로부터 가치를 얻을 수 있는 기회를 인식하고 행동하도록 보장하는 전략과 지원 원칙, 정책

및 스텐어드십 실행을 시행함으로써 조직이 데이터 중심으로 변할 수 있도록 한다. 또한 데이터 거버넌스 프로그램은 조직을 교육하고 데이터의 전략적 사용을 가능하게 하는 행동을 장려하기 위한 조직 변화 매니지먼트 활동에 참여해야 한다. 따라서, 특히 조직이 데이터 매니지먼트 실행을 성숙시킬 때 문화 변화의 필요성은 데이터 거버넌스 책임의 광범위한 범위에 걸쳐 있다. 데이터 매니지먼트 프레임워크는 또한 거버넌스의 엄격성 내에서 라이프사이클과 사용 활동에 둘러싸인 핵심 활동을 DAMA 휠의 발전으로 묘사할 수 있다(그림11 참조).

이 그림에는 메타데이터 매니지먼트, 데이터 품질 매니지먼트, 데이터 구조 정의(아키텍처) 등 핵심 활동이 프레임워크의 중심에 있다.

라이프사이클 매니지먼트 활동은 계획 관점(위험 매니지먼트, 모델링, 데이터 설계, 참조 데이터 매니지먼트)과 지원 관점(마스터 데이터 매니지먼트, 데이터 기술 개발, 데이터 통합 및 상호 운용성, 데이터 웨어하우징, 데이터 스토리지 및 운영)에서 정의될 수 있다.

라이프사이클 매니지먼트 활동 전반에서 일어나는 활동인 마스터 데이터 사용, 문서 및 콘텐츠 매니지먼트, 비즈니스 인텔리전스, 데이터 과학, 예측 분석, 데이터 시각화 중 다수는 기존 데이터에 대한 통찰력을 강화하거나 개발함으로써 더 많은 데이터를 생성한다. 데이터 수익화의 기회는 데이터를 사용함으로써 식별될 수 있다.

데이터 거버넌스 활동은 전략, 원칙, 정책 및 매니지먼트 책임을 통해 감독 및 봉쇄를 제공한다. 이들은 데이터 분류와 데이터 평가를 통해 일관성을 유지하는 것을 가능하게 한다.

데이터 매니지먼트 프레임워크의 다른 시각적 묘사를 제시하고자 하는 의도는 데이터 지식 매니지먼트 체계에서 제시된 개념을 어떻게 적용할 것인가에 대한 추가적인 관점과 개방적인 논의를 제공하기 위함이다. 데이터 매니지먼트의 중요성이 증가함에 따라, 그러한 프레임워크는 데이터 매니지먼트 커뮤니티 내에서 그리고 데이터 매니지먼트 커뮤니티와 우리의 이해관계자들 사이의 유용한 커뮤니케이션 도구로 기능하다.

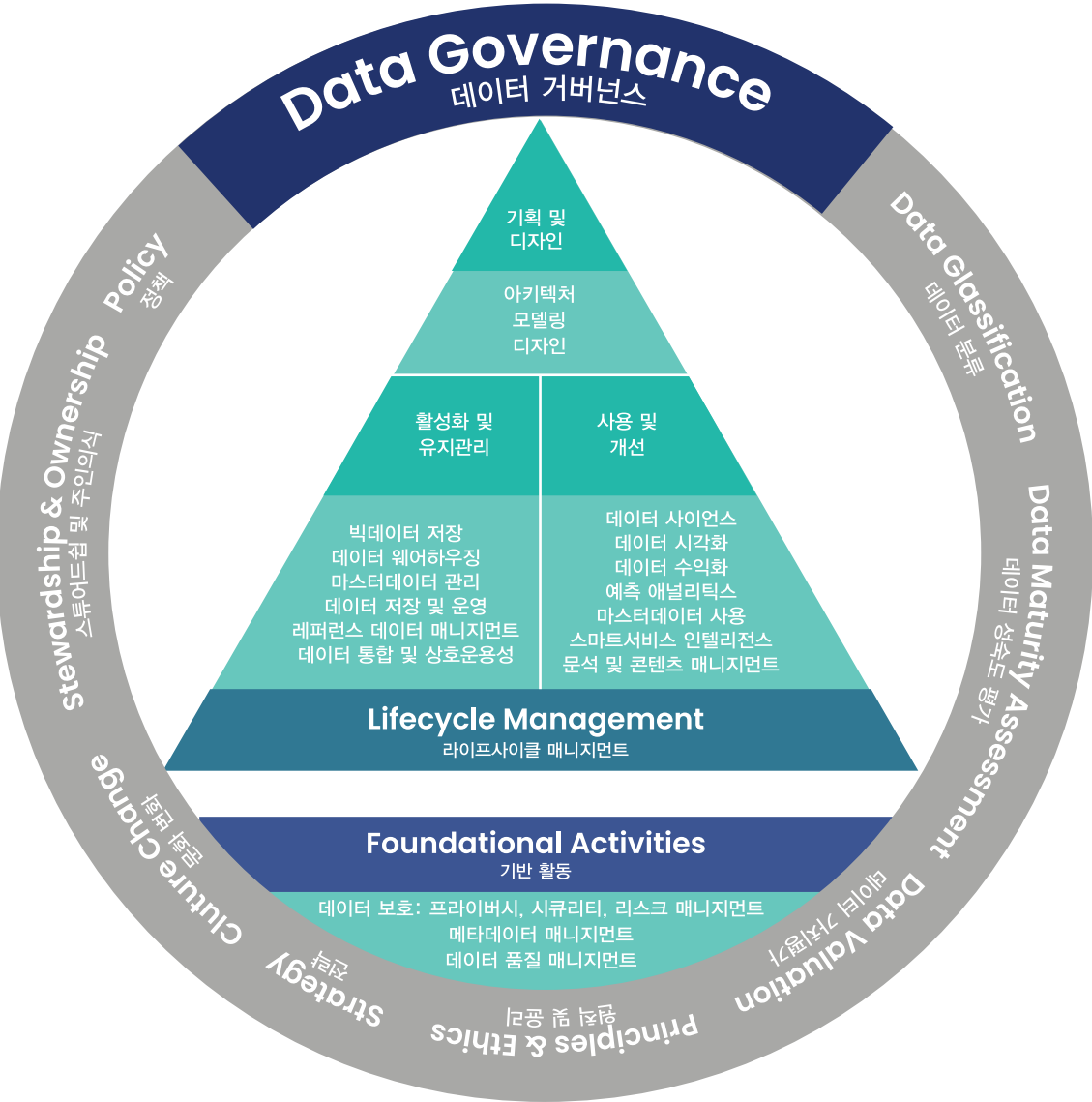


Technical
Standards
for Smart Cities
스마트시티 국제표준화
기반조성(R&D)

4 데이터 지식 매니지먼트 체계 (DAMA and the DMBOK)

데이터 매니지먼트는 많은 문제를 내포하고 있지만, 그 중 새롭게 등장한 것은 거의 없다. 적어도 1980년대 이후, 조직은 데이터 매니지먼트가 성공의 핵심이라는 것을 인식해왔다. 또한 데이터를 생성하고 활용하려는 능력과 욕구가 증가함에 따라 신뢰할 수 있는 데이터 매니지먼트 실행이 필요하게 되었다. 데이터 매니지먼트 전문가들을 위한 접근 가능하고 이 보고서는 다음과 같은 임무를 지원한다.

- 지침 원칙, 널리 채택된 실행, 방법 및 기법, 기능, 역할, 결과물 및 측정 기준을 포함한 기업의 전사적 데이터 매니지먼트 업무 구현을 위한 기능 프레임워크 제공
- 데이터 매니지먼트 개념에 대한 공통 어휘를 확립하고 데이터 매니지먼트 전문가의 모범 사례 제공



[그림11] 데이터 매니지먼트 기반 데이터 거버넌스

- CDMP(공인 데이터 매니지먼트 전문가; Certified Data Management Professional) 및 기타 인증 시험의 기본 참조 가이드 역할 수행

데이터 매니지먼트 프레임워크의 11가지 지식 영역(DAMA 휠, 그림5 참조)을 중심으로 구성되며, 다음과 같은 공통 구조를 따른다.

1. 소개
비즈니스 동인, 목표와 원칙, 필수 개념
2. 활동
3. 도구
4. 기법
5. 실행 지침
6. 데이터 거버넌스와의 관계
7. 평가지표

지식 영역은 일련의 데이터 매니지먼트 활동의 범위와 맥락을 설명한다. 지식 영역에는 데이터 매니지먼트의 기본 목표와 원칙이 포함되어 있다. 데이터는 조직 내에서 수평적으로 이동하기 때문에 지식 영역 활동은 서로 교차하고 다른 조직 기능과도 교차한다.

1. 데이터 거버넌스(Data Governance)는 기업의 요구를 감안한 데이터에 대한 의사결정권 체계를 확립함으로써 데이터 매니지먼트에 대한 방향과 감독을 제공한다.
2. 데이터 아키텍처(Data Architecture)는 이러한 요구사항에 맞는 전략적 데이터 요구사항 및 설계를 수립하기 위한 조직 전략에 맞추어 데이터 자산을 매니지먼트하기 위한 청사진을 정의한다.
3. 데이터 모델링 및 설계(Data Modeling and Design)은 데이터 모델이라 불리는 정확한 형태로 데이터 요구사항을 발견, 분석, 표현, 전달하는 과정이다.
4. 데이터 저장 및 운영(Data Storage and Operations)은 그 가치를 극대화하기 위해 저장된 데이터의 설계, 구현, 지원을 포함한다. 운영은 데이터 계획 수립부터 폐기까지 데이터 라이프사이클 전반에 걸쳐 지원을 제공한다.
5. 데이터 보안(Data Security)은 데이터 프라이버시와 기밀성이 유지되고, 데이터가 침해되지 않으며, 데이터에 적절하게 접근하도록 보장한다.
6. 데이터 통합 및 상호운용성(Data Integration and Interoperability)은 데이터 저장소, 애플리케이션 및 조직 간의 데이터 이동 및 통합과 관련된 프로세스를 포함한다.
7. 문서 및 콘텐츠 매니지먼트(Document and Content Management)는 특히 법률 및 규정 준수 요건을 지원하는 데 필요한 문서 등 다양한 비정형 미디어에서 발견된 데이터 및 정보의 라이프사이클 매니지먼트에 사용되는 계획, 구현, 통제 활동을 포함한다.
8. 레퍼런스 및 마스터 데이터(Reference and Master Data)에는 핵심 중요 공유 데이터의 지속적인 조정 및 유지보수가 포함되어 있어 필수 비즈니스 실체에 대한 가장 정확하고 시기적절하며 관련성이 높은 데이터를 전 시스템 전체에 걸쳐 일관성 있게 사용할 수 있게 도와준다.

9. 데이터 웨어하우징 및 서비스 인텔리전스(Data Warehousing and Service Intelligence)에는 의사결정 지원 데이터를 매니지먼트하고 지식 근로자가 분석 및 보고를 통해 데이터로부터 가치를 얻을 수 있도록 하기 위한 계획 수립, 구현, 제어 프로세스가 포함된다.

10. 메타데이터(Metadata)는 데이터 및 데이터 생성, 유지, 액세스에 중요한 기타 정보를 포함하여 고품질의 통합 메타데이터에 대한 액세스를 가능하게 하는 계획, 구현 및 제어 활동을 포함한다.

11. 데이터 품질(Data Quality)은 조직 내에서 사용할 데이터의 적합성을 측정, 평가 및 개선하기 위한 품질 매니지먼트 기법의 계획 및 구현을 포함한다.

특정 조직이 데이터를 관리하는 방법은 데이터의 목표, 크기, 리소스 및 복잡성과 더불어 데이터가 전체 전략을 지원하는 방법에 대한 인식에 따라 달라진다. 대부분의 기업은 각 지식 영역에 설명된 모든 활동을 수행하지는 않는다. 그러나, 데이터 매니지먼트의 더 넓은 맥락을 이해하면 조직은 이러한 관련 기능 내부 및 전체에서 실행을 개선하기 위해 작업할 때 어디에 초점을 두어야 할지에 대한 더 나은 결정을 내릴 수 있을 것이다.

5 인용 및 참고자료

Abcouwer, A. W., Maes R, Truijens J.: “Contouren van een generiek Model voor Informatienmanagement.” Primavera

Adelman, Sid, Larissa Moss Data Strategy Addison-Wesley Professional, 2005. Print.

Aiken, Peter and Billings ‘ Juanita · Monetizing Data Management Technics Publishing, LLC, 2014. Print.

Aiken, Peter and Harbour, Todd. Data Strategy and the Enterprise Data Executive. Technics Publishing, LLC. 2017. Print.

APRA (Australian Prudential Regulation Authority) . Prudential Practice Guide CPG 234, Management of Security Risk in Information and Information Technology May 2013

APRA (Australian Prudential Regulation Authority) Prudential Practice Guide CPG 235, Managing Data Risk. September 2013

Borek, Alexander et al. Total information Risk Management : Maximizing the value of Data and Information Assets. Morgan Kaufmann 2013, Print.

Brackett, Michael. Data Resource Design: Reality Beyond Illusion. Technics Publishing, LLC. 2014. Print.

Bryce, Tim. Benefits of a Data Taxonomy. Blog 2005-07-11. <http://bit.ly/2sTeUIU>.

Chisholm, Malcolm and Robyn-Lee, Diane. Definition in Data Management: A Guide to Fundamentals Semantic Metadata. Design Media, 2008. Print.

Devlin, Barry. Business Unintelligence. Technics Publishing, LLC. 2013 . print.

English, Larry. Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs And Increasing Profits, John Wiley and Sons, 1999. Print

Evans, Nina and Price, James. "Barriers to the Effective Deployment of Information Assets: An Executive Management Perspective." *Interdisciplinary Journal of Information, Knowledge, and Management* Volume 7, 2012. Accessed from <http://bit.ly/2sVwvG4>.

Fisher, Tony. *The Data Asset: How Smart Companies Govern Their Data for Business Success*. Wiley, 2009. Print. Wiley and SAS Business Ser.

Henderson, J.C, H Venkatraman H. "Leveraging Information technology for transforming Organizations" *IBM System Journal* Volume 38, Issue 2.3 1999. [1993 Reprint]

Kent, William, *Data and Reality: A Timeless Perspective on Perceiving and Managing Information in Our Imprecise World*. 3ded. Technics Publications, LLC, 2012, Print.

Kring, Kenneth L. *Business Strategy Mapping – The Power of Knowing How it All Fits Together*. Langdon Street Press (a division of Hillcrest Publishing Group, Inc.) 2009. Print.

Loh, Steve. *Data-ism: The Revolution Transforming Decision Making, Consumer Behavior, and Almost Everything Else*. HarperBusiness, 2015 Print.

"Loshin, David. *Enterprise Knowledge Management: The Data Quality Approach*. "Morgan Kafuman, 2001 Print.

Maes, R.: "A Generic Framework for Information Management." *Prima Vera Working Paper* 99-02, 1999."

McGilvray, Danette. *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information*. Morgan Kaufmann, 2008, Print.

McKnight, William. *Information Management Strategies for Gaining a Competitive Advantage with Data*. Morgan Kaufmann, 2013. Print. *The Savvy Manager's Guides*.

Moody, Daniel and Walsh, Peter. "Measuring The Value Of Information: n Asset Valuation Approach." *European Conference on Information Systems (ECIS)*, 1999. <http://bit.ly/29JucL0>

Olson, Jack E. *Data Quality: The Accuracy Dimension*. Morgan Kaufmann, 2003. Print.

Redman, Thomas. "Bad Data Costs U.S. \$3 Trillion per Year." *Harvard Business Review*. 22 September 2016. Web.

Redman, Thomas. *Data Driven: Profiting from Your Most Important Business Asset*. Harvard Business Review Press. 2008 Print.

Redman Thomas. *Data Quality: The Field Guide*. Digital Press, 2001. Print

Reid, Roger, Gareth Fraser-King and W. David Schwaderer. *Data Lifecycles: Managing Data for Strategic Advantage*. Wiley, 2007. Print.

Rockley, Ann and Charles Cooper. *Managing Enterprise Content: A Unified Content strategy* 2nd ed. New Riders 2012. Print *Voices That Matter*.

Sebastian-Colemann, an, Laure. *Measuring Data Quality for Ongoing Improvement: A Data Quality Assesment Framework*. Morgan Kaufmann, 2013. Print. *The Morgan Kafumann Series in Business Intelligence*

Simsion, Graeme. *Data Modeling: Theory and Practice*. Technics Publications, LLC, 2007. Print

