

데이터 품질관리 성숙모형 (Ver 1.0)

Data Quality Management Maturity Model

조직의 데이터 품질관리 수준 측정 모형

데이터의 유효성과 활용성 향상을 위해 필요한
데이터 품질관리 개선 전략과 방법



한국데이터베이스진흥센터
Korea Database Promotion Center

머리말

정보시스템은 민간 기업이나 공공 기관의 구분 없이 모든 조직의 업무 효율성을 획기적으로 개선시키고 있다. 일상의 많은 업무가 정보시스템을 경유해 수행되며, 정보시스템에서 제공하는 데이터는 조직의 크고 작은 의사결정의 중요한 근거 자료로 활용된다.

그런데 최근 정보기술의 발달에 따라 정보시스템 간의 통합·연계 활용이 확대되고 처리해야 할 업무가 복잡·다양해지면서 그 동안 감춰졌던 데이터 품질 문제가 표면화되기 시작하고 있다. 데이터 품질의 문제는 정보시스템에 대한 불신을 야기해 활용도를 떨어뜨리고 업무의 생산성은 물론 기업의 경쟁력 저하로까지 이어질 수 있다.

이에 따라 많은 기업들이 데이터에 대한 체계적인 품질관리를 시도하고 있다. 데이터 품질 관리를 도입하기 위해서는 현재의 품질 상태를 정확하게 인식하고 오류 발생 원인과 문제점을 분석해 대응 방안을 마련해야 한다. 그러나 이와 같은 문제를 효율적으로 해결하도록 지원하는 방법론이나 참조 사례가 부족해 데이터 품질관리를 도입하려는 기업의 실무 담당자들이 많은 어려움을 토로하고 있다.

이에 한국데이터베이스진흥센터는 조직이 자체적으로 데이터 품질관리 수준을 진단해보고 체계적인 데이터 품질관리 전략을 수립할 수 있도록 지원하기 위해 『데이터 품질관리 성숙모형』을 개발했다.

『데이터 품질관리 성숙모형』은 데이터 품질관리 성숙수준을 ‘도입–정형화–통합화–정량화–최적화’의 5단계로 구분해 데이터의 정확성, 일관성, 유용성, 접근성, 적시성, 보안성의 데이터 품질 기준별 관리 성숙수준을 측정하도록 고안되어 있다. 따라서 측정 결과에 따라 특정 기준의 데이터의 품질을 향상시키기 위해서 어떤 프로세스를 어떻게 개선해야 하는지 바로 알 수 있다. 투입 비용 대비 가시적 성과 측정도 용이하다. 특히 조직의 데이터 품질 기준 우선 순위나 가용 자원을 고려한 단계적인 품질관리 개선 계획 수립도 가능하다.

또한 본 책자는 한국데이터베이스진흥센터의 데이터 품질관리 첫번째 시리즈인 「데이터 품질관리 지침」을 기반으로 개발되었다. 「데이터 품질관리 지침」은 데이터 품질관리 대상과 관리 절차 및 방법에 대한 가이드를 제공하고 있으며, 「데이터 품질관리 성숙모형」에서 측정 대상으로 하는 제반 프로세스는 지침의 내용을 준용하고 있다. 따라서 데이터 품질관리 수준을 점검하고 관리 체계를 수립하고자 하는 조직이 「데이터 품질관리 성숙모형」과 「데이터 품

질관리 지침」을 상호 참조해 활용하면 많은 도움이 될 것이다. 데이터 품질관리 시리즈는 데이터 품질관리 인증센터 홈페이지(www.dbq.or.kr)에서 내려 받을 수 있다.

아울러 「데이터 품질관리 성숙도형」은 버전 1.0으로 개발되었으며 공공 기관과 민간 기업의 실무 담당자와 관련 업계 전문가들의 의견을 폭넓게 수렴해 지속적으로 개선해 나갈 것이다.

2006. 9

한국데이터베이스진흥센터

목 차

제1장 데이터 품질관리 성숙모형 개요	1
1. 데이터 품질관리 성숙모형의 필요성	1
2. 데이터 품질관리 성숙모형의 특징	5
제2장 데이터 품질관리 성숙모형의 이해	9
1. 데이터 품질관리 성숙모형 구조	9
2. 데이터 품질 기준	11
2.1 유효성 관련 품질 기준	12
2.2 활용성 관련 품질 기준	14
3. 데이터 품질관리 프로세스	17
3.1 요구사항 관리	17
3.2 데이터 구조 관리	18
3.3 데이터 흐름 관리	20
3.4 데이터베이스 관리	20
3.5 데이터 활용 관리	21
3.6 데이터 표준 관리	22
3.7 데이터 오너십 관리	23
3.8 사용자 뷰 관리	23
3.9 데이터 품질 기준과 관리 프로세스 관계	23
4. 데이터 품질관리 성숙수준	25
4.1 데이터 품질관리 성숙수준 단계	25
4.2 데이터 품질관리 프로세스와 성숙수준 관계	27

제3장 기준별 데이터 품질관리 성숙수준	29
1. 유효성 관련 성숙수준.....	29
1.1 정확성 성숙수준	29
1.2 일관성 성숙수준	33
2. 활용성 관련 성숙수준.....	39
2.1 유용성 성숙수준	39
2.2 접근성 성숙수준	43
2.3 적시성 성숙수준	47
2.4 보안성 성숙수준	51
부록 : 성숙수준별 품질관리 점검표.....	55

그림 목차

제1장 데이터 품질관리 성숙모형 개요

〈그림 1-1〉 데이터 품질관리 성숙모형 구성 요소	7
------------------------------------	---

제2장 데이터 품질관리 성숙모형의 이해

〈그림 2-1〉 데이터 품질관리 성숙모형 구조	10
〈그림 2-2〉 데이터 품질 기준	11
〈그림 2-3〉 데이터 품질관리 프로세스	17
〈그림 2-4〉 데이터 품질관리 성숙수준 단계	25
〈그림 2-5〉 데이터 품질관리 성숙수준 측정 결과	26

제3장 기준별 데이터 품질관리 성숙수준

〈그림 3-1〉 정확성 관련 데이터 품질관리 프로세스	30
〈그림 3-2〉 일관성 관련 데이터 품질관리 프로세스	34
〈그림 3-3〉 유용성 관련 데이터 품질관리 프로세스	39
〈그림 3-4〉 접근성 관련 데이터 품질관리 프로세스	43
〈그림 3-5〉 적시성 관련 데이터 품질관리 프로세스	47
〈그림 3-6〉 보안성 관련 데이터 품질관리 프로세스	51

표 목차

제1장 데이터 품질관리 성숙모형 개요

제2장 데이터 품질관리 성숙모형의 이해

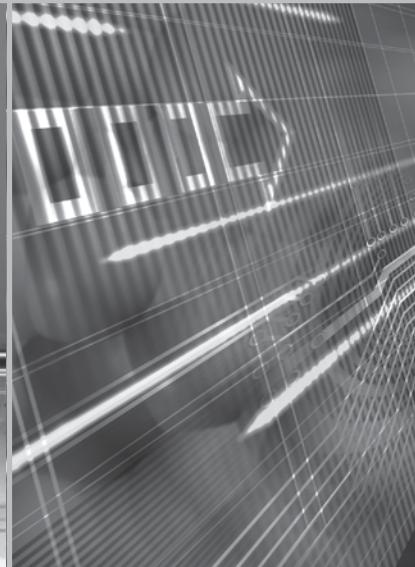
〈표 2-1〉 데이터 품질 기준과 관리 프로세스 연관도.....	24
〈표 2-2〉 품질관리 프로세스와 성숙수준 관계	27

제3장 기준별 데이터 품질관리 성숙수준

〈표 3-1〉 정확성 관련 데이터 품질관리 수준.....	32
〈표 3-2〉 일관성 관련 데이터 품질관리 수준.....	37
〈표 3-3〉 유용성 관련 데이터 품질관리 수준.....	42
〈표 3-4〉 접근성 관련 데이터 품질관리 수준.....	46
〈표 3-5〉 적시성 관련 데이터 품질관리 수준.....	50
〈표 3-6〉 보안성 관련 데이터 품질관리 수준.....	54

제1장 데이터 품질관리 성숙모형 개발

제1장에서는 『데이터품질관리성숙모형의 개발 배경과 필요성에 대해 설명하고, 제2장과 제3장에서 설명할 성숙 모형의 이론적 · 구조적 내용에 대한 이해를 돋기 위해 성숙모형의 특징에 대해 설명한다.



1. 데이터 품질관리 성숙모형의 필요성

업무의 효율성을 증가시키기 위하여 우리 주변의 많은 작업들에 대한 전산화가 광범위하게 이루어지고 있고, 이러한 추세는 향후에도 계속될 것으로 전망된다. 특히 전산 기술의 비약적인 발전은 업무 수행에 필요한 다양한 기능을 제공할 뿐만 아니라, 새로운 기술 접목을 통하여 기존의 업무 처리 방식을 보다 효과적인 방향으로 개선하고 있다. 전산의 적용은 단순 반복적인 업무뿐 아니라 핵심 업무에도 적용되고 있어, 일부 조직에서는 전산의 도움 없이 업무를 처리하는 것이 불가능한 경우도 발생하고 있다.

업무 수행에 있어서 전산의 의존도가 높아짐에 따라, 업무 처리 과정에서 발생하는 대부분의 데이터가 전산 시스템에 보관되며, 업무 처리에 필요한 데이터 또한 전산 시스템을 통하여

공급된다. 전산 시스템을 통한 데이터의 축적은 방대한 데이터를 여러 부문에서 필요에 따라 손쉽게 공유할 수 있을 뿐만 아니라, 원하는 형태로의 가공 작업도 용이하게 수행할 수 있다. 하지만, 업무의 복잡성 증가 및 처리량의 증가 등으로 인해 전산을 통하여 관리해야 하는 데이터의 종류 및 분량이 기하 급수적으로 증가하고 있어, 데이터 관리에 대한 체계적인 접근이 필요하다.

그러나 관리의 체계화를 수행할 수 있는 시간적 여유가 없을 정도로 전산에 대한 기대 사항들은 증가하고 있고, 단편적인 시스템 개발이 연속적으로 이루어짐에 따라 데이터 관리에 대한 어려움은 더욱 가중되고 있다. 또한 데이터 관리의 체계화를 수행하기 위한 재정적 지원도 미비하고 기술적인 한계도 존재하여 투자의 우선 순위에서 밀리는 실정이다. 그러나 보니 과거에는 전산을 통하여 제공되는 데이터는 항상 정확하다는 인식이 있었지만 근래에는 이와는 정반대의 현상이 발생하고 있다.

전산 시스템을 이용하는 많은 사용자들이 전산을 통해서 제공되는 데이터에 대하여 불만이나 불편을 제기하고 있고, 우리 주변에서도 잘못된 데이터로 인하여 문제가 발생하는 경우를 쉽게 찾아볼 수 있다. 사용자들의 데이터에 대한 부정적인 시각은 대략 다음과 같다.

- 정확하지 않아 신뢰할 수가 없다.
- 정작 필요한 정보는 누락되어 있다.
- 원하는 정보가 아니거나 원하는 형태가 아니다.
- 제공 시점이 너무 느리다.
- 어딘가 있을 것 같은데 요청하면 받아볼 수 없다.
- 정보가 훌러 나가거나 일순간에 없어질 것 같아 불안하다.

전산 데이터에 대한 강한 믿음이 깨지면서 기대가 실망으로 변하고 있으며, 전산 시스템에 대한 인식에도 부정적인 영향을 미치고 있다. 잘못된 데이터나 부족한 데이터는 업무 수행에 지장을 줄 뿐만 아니라, 경우에 따라서는 심각한 결과를 초래한다. 특히 금전적인 데이터에 대한 오류는 조직에 막대한 손해를 끼칠 수도 있다.

데이터 관리의 어려움은 관련 대상이 많고 복잡하다는 것이다. 시스템이 동시 다발적으로 개발되다 보니 데이터가 여러 시스템에 흩어지게 되고, 분산된 데이터의 일관성 유지에 어려움이 발생한다. 또한 데이터를 이용하는 사람이 많다 보니 잘못된 조작에 의한 데이터의 오류

가 발생하기 쉽다. 최신의 상태로 갱신되어야 하지만, 갱신 작업에서 누락되는 경우도 발생하고, 데이터의 진정한 의미를 모르고 잘못 사용하여 엉뚱한 결과를 얻기도 한다. 전산 애플리케이션은 한번 개발한 이후 요건의 변화가 발생하기 전까지는 고정적인 상태가 유지되지만, 데이터는 업무 처리에 따라 수시로 생성되고 변경되고 없어진다. 그러므로 세심한 관리가 필요한 대상이지만, 전산 분야 중에서는 상대적으로 관심이 적었던 부문이다.

데이터에 대한 무관심으로 인하여 전산 시스템의 활용에 심각한 제약이 발생하고 있고, 이에 대한 반작용으로 데이터 품질에 대한 관심이 높아지게 되었다. 데이터 품질관리를 통하여 데이터의 신뢰성이 확보되면, 자연스럽게 전산 시스템의 효율도 증가하게 될 것이다.

하지만 데이터 품질 확보는 일시적인 투자나 관심에 의하여 급속히 이루어지는 것은 아니다. 데이터 전반에 대한 체계적이고 장기적인 정비가 이루어져야 한다. 또한 한번 확보된 품질도 관리를 소홀히 하면 순식간에 나빠질 수 있다. 데이터 품질은 확보하기도 어렵고 유지하기도 어렵다. 하지만 데이터 품질의 확보 없이는 전산의 활용이 위축될 것이고 조직의 경쟁력도 저하될 것이므로, 가능한 모든 노력을 집중해야 할 과제라고 할 수 있다. 데이터 품질 확보의 어려움은 크게 두 가지 측면에서 생각할 수 있다.

첫째, 데이터 품질에 대한 인식은 개인적으로 차이가 존재하기 때문에, 모든 사람의 품질에 대한 기대를 충족시키는 것이 쉽지 않다. 완벽한 수준의 품질 확보를 요구하는 개인이 있는가하면 어느 정도의 오류는 인정하거나 특정 수준 이상을 유지하면 만족하는 경우도 있다. 한편 데이터 품질을 정의하는 데에도 많은 시각 차이가 있다. 데이터의 정확성을 품질로 생각하는 기본적인 수준에서부터, 데이터의 신속한 제공, 데이터 활용의 편의성 등을 품질로 생각하는 경우도 많다. 더 나아가 데이터의 안전성이나 보안성을 품질의 범주에 포함시키기도 한다. 데이터 품질을 확보하기 위한 출발점은 데이터 품질에 대한 정의를 공유하고, 현실적으로 달성 가능한 품질 수준을 설정하는 것이다.

데이터 품질 확보에 있어서의 두 번째 어려움은, 품질에 영향을 미치는 요인들이 다양하고 이를 관리하기 위한 기능이 복잡하게 연결되어 있다는 것이다. 그러다 보니 데이터 품질 확보를 위해서는 관련된 기능의 정비에 많은 노력을 들여야 한다. 예를 들어, 데이터 구조 관리, 데이터 흐름 관리, 데이터 표준 관리, 데이터 오너십 관리, 데이터 성능 관리 등이 데이터 품질과 직간접적으로 관련 있고, 이를 간에도 상호 복합적인 관계가 성립되어 있다. 그러므로 데이터 품질에 영향을 미치는 기능의 연관 관계를 명확하게 파악할 필요가 있다.

데이터 품질의 문제점을 해결하기 위해서는 현재의 품질 상태를 정확하게 인식해야 한다. 품질 수준이 파악 되어야만 이로 인한 문제점 및 원인을 정확하게 분석할 수 있고, 대응 방안도 마련할 수 있다.

일반적인 접근은 데이터 품질을 측정하고 품질 불량인 대상에 대한 개선을 수행하는 것이다. 이러한 방식은 일시적으로 품질 수준을 높일 수 있어 아주 효과적으로 보인다. 하지만, 이러한 접근에는 두 가지 문제점이 존재한다.

첫 번째는 품질 수준이 일정 수준 이상 밖에는 개선되지 않는다는 것이다. 그 원인은 데이터 품질에 영향을 미치는 요인이 단지 겉으로 드러나는 현상에만 있는 것이 아니고, 데이터 관리 기능의 미흡이라는 근본적인 원인이 존재하기 때문이다. 그러므로 이러한 근본적인 원인을 해결하지 않고 품질을 특정 수준 이상으로 끌어 올리기 힘들다.

두 번째는 품질 수준이 올라간 이후 일정 시간이 경과하면 품질이 다시 저하되는 경향이 많다는 것이다. 데이터 품질관리는 지속적으로 수행되어야 효과를 얻을 수 있으며 일시적인 조치를 통해서 확보된 품질은 오래 지속되기 어렵다. 그러므로 데이터 품질과 관련이 있는 핵심 프로세스를 정비하고 유지하는 것이 중요하다.

궁극적으로 데이터 품질 수준을 측정하는 것도 중요하지만, 데이터 품질을 관리하는 프로세스의 수준을 평가하고 이를 향상시키는 것이 근본적인 해결책이라고 할 수 있다.

2. 데이터 품질관리 성숙모형의 특징

데이터 품질관리의 중요성이 인식되면서, 품질관리에 대한 계획을 수립하거나, 구체적으로 품질관리 활동을 수행할 필요성이 높아지고 있다. 그러나 품질관리를 어디서부터 시작해야 하는지에 대한 의문이 많고, 지금 수행하고 있는 품질관리 활동을 제대로 하고 있는지에 대한 걱정도 많다. 더구나, 품질관리에 대한 장기적인 방향성을 가지고 단계적으로 품질 수준을 향상시키고 있는 사례는 많지 않다. 이러다 보니 데이터 품질에 대한 투자가 제대로 효과를 보지 못하거나 효과가 있다고 해도 비효율적인 과정을 통하여 진행되는 경우가 많다.

데이터 품질관리에 대한 체계적인 접근을 위해서는 객관적인 기준에 의하여 데이터 품질관리 수준을 측정하고 더 나은 수준을 확보하기 위해서는 무엇을 준비해야 하는지에 대한 가이드라인이 절실하다. 데이터 품질관리 성숙모형은 이러한 요건을 충족시키기 위한 목적에서 개발되었다.

데이터 품질관리 성숙모형은 세 가지 중심 축으로 구성되어 있다.

첫 번째는 데이터 품질에 대한 정의이다. 데이터 품질은 개인이나 조직의 시각에 따라 다양하게 정의될 수 있다. 그런데 품질에 대하여 상이한 인식을 바탕으로 품질관리를 수행한다면 대상이 명확하지 않아 문제가 발생한다. 그러므로 데이터 품질에 대한 정의가 가장 중요한 출발점이라고 할 수 있다. 본 모형에서는 데이터 품질에 대한 다양한 시각 중에서 보편성이 있는 것을 선정하고 이를 분류하여 품질의 다양한 측면을 반영할 수 있도록 구성하였다. 데이터 품질의 다양한 측면을 데이터 품질 기준으로 명명하고 6가지로 정리하였다. 정확성, 일관성은 데이터의 유효성에 대한 측면이고, 유용성, 접근성, 적시성, 보안성은 데이터의 활용적인 측면을 반영하고 있다. 데이터 품질을 총체적으로 말한다면 6개 기준을 모두 충족시켜야 하겠지만, 조직의 속한 환경이나 사용자의 관심에 따라 이중에서 특히 더 중요한 기준이 존재할 수 있다. 마찬가지로 데이터 품질에 대한 측정은 6개 기준을 모두 점검해 볼 수도 있지만, 이중 한 두 개의 비중 있는 기준만을 측정할 수도 있다. 데이터 품질의 정의는 조직의 품질에 대한 인식에 따라 6개 기준에 대한 비중을 고려하여 정해질 수 있다.

데이터 품질 성숙모형을 구성하는 둘째 축은 위에서 제시한 6개 기준의 품질을 향상시키기 위해서 필요한 프로세스 혹은 품질 기준에 영향을 주는 요소를 관리하는 프로세스를 식별하

는 것이다. 그런데, 이러한 프로세스는 대부분 데이터 관리에 필요한 프로세스로 구성되어 있다. 앞에서 언급했듯이 데이터 품질 자체에 관심을 두는 것보다는 이와 관련 있는 프로세스에 대한 총체적인 정비가 더 중요하다.

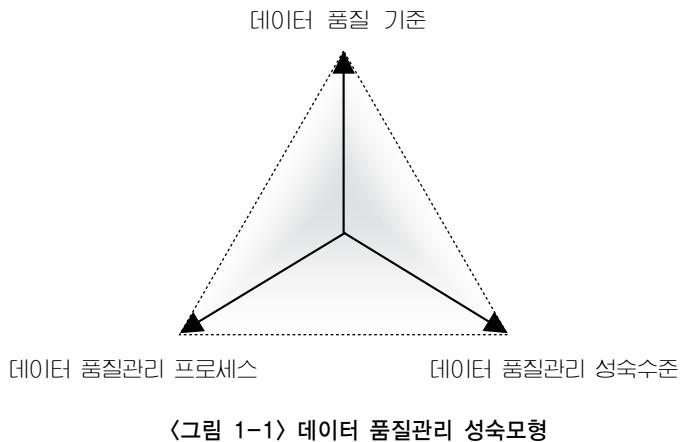
품질 기준과 데이터 관리 프로세스는 1:1로 대응되는 경우도 있지만, 품질 기준 여러 개와 관련이 있는 프로세스도 존재한다. 그런데 품질 기준과 관리 프로세스 간의 연관 관계는 데이터 품질관리 성숙모형을 어떤 모습으로 구성해야 하는지에 대한 중요한 방향을 제시한다. 만일 품질 기준과 관리 프로세스가 복잡하게 엮여있다면 특정 프로세스의 향상이 어떤 품질 기준에 영향을 주는지를 직접적으로 파악할 수 없다. 단지 프로세스의 향상을 통하여 총체적으로 품질에 도움을 주었다고 해석할 수 밖에 없다. 이러한 경우라면 관리 프로세스와 품질 기준 간의 관계 보다는, 관리 프로세스들 간의 선후 관계를 파악하여 단계적인 접근을 수행해야 한다. 즉 프로세스 간에 순서가 존재하여 앞의 프로세스를 개선시킨 후에야 뒤 프로세스를 향상시킬 수 있다. 이를 바탕으로 프로세스 간의 계층을 설정하고 어떤 계층까지 관리가 이루어지고 있는지를 바탕으로 관리 프로세스의 성숙수준을 측정한다. 이런 모형은 프로세스 개선이 품질 향상에 도움을 주는 것은 확실하지만 구체적으로 어떤 품질 기준에 영향을 주는지를 추적하기 어렵다.

반면, 품질 기준과 관리 프로세스 간에 맞대응이 이루어진다면, 특정 품질 기준을 향상시키기 위하여 어떤 프로세스를 개선해야 하는지를 알 수 있어, 실무적인 입장에서 더 많은 도움을 얻을 수 있고, 가시적인 성과를 바탕으로 품질 향상 활동을 전개할 수 있다.

품질 기준과 관리 프로세스와의 연관 관계를 분석해 본 결과 후자의 모습에 더 가깝고, 향후 성숙모형의 정밀도를 향상시키기 위해서도 후자의 형태로의 모형이 바람직하다고 판단된다. 그러므로 본 성숙모형은 품질 기준 별로 관련 있는 관리 프로세스를 식별할 수 있다는 기본적인 가정 위에서 만들어졌다.

데이터 품질관리 성숙모형을 구성하는 세 번째 축은 데이터 품질관리 성숙수준이다. 데이터 품질관리 성숙수준이 높을수록 체계적이고 정교한 관리가 수행되는 것이고 해당 품질 기준도 높은 품질을 유지한다. 데이터 품질관리 성숙수준은 품질 기준 별로 다르게 측정될 수 있다. 이는 각 품질 기준이 서로 다른 관리 프로세스에 의하여 영향을 받고 있기 때문이다. 데이터 품질관리 성숙수준은 1단계에서부터 5단계까지로 정의되어 있다.

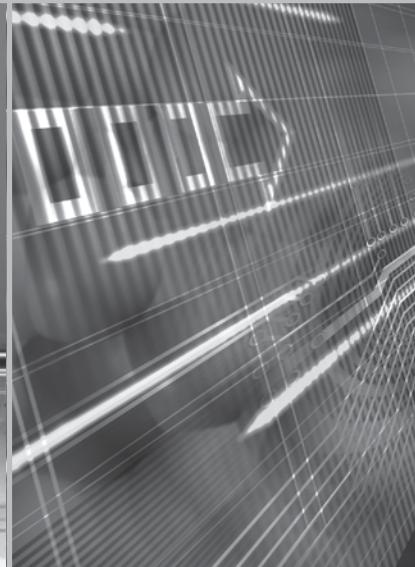
데이터 품질관리 성숙모형은 단지 현재의 품질 수준을 측정하기 위한 목적만을 가지고 있는 것이 아니고, 품질 향상을 위해서는 어떤 프로세스를 어떻게 개선해야 하는지에 대한 방향을 제시할 수 있도록 개발되었다. 또한 조직의 품질에 대한 인식에 따라 선별적으로 적용할 수 있는 구조를 갖추고 있다. 조직별 품질의 중요도 및 가용한 자원에 따라 단계적이고 지속적인 계획을 수립하여 품질 향상에 활용할 수 있기를 희망한다.



〈그림 1-1〉 데이터 품질관리 성숙모형

제2장 데이터 품질관리 성숙모형의 이해

『데이터 품질 관리 성숙모형』은 데이터 품질 기준, 데이터 품질 관리 프로세스, 데이터 품질 관리 성숙수준으로 구성된다. 제2장에서는 데이터 품질 관리 성숙모형 구조와 각 구성 요소별 세부 내용에 대해 설명한다.

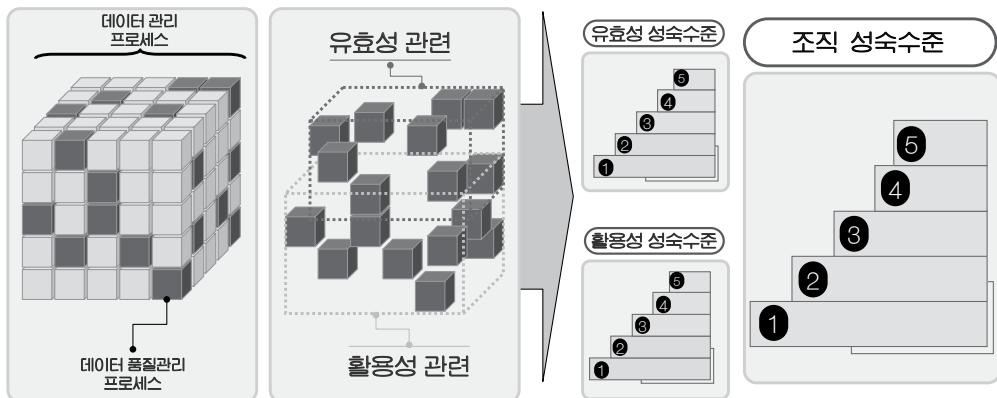


1. 데이터 품질관리 성숙모형 구조

『데이터 품질 관리 성숙모형』은 데이터 품질 기준, 데이터 품질 관리 프로세스 그리고 데이터 품질 관리 성숙수준을 기본 축으로 한다.

첫 번째 축인 데이터 품질 기준은 데이터 품질의 다양한 측면을 분류·정리해 정확성, 일관성, 유용성, 접근성, 적시성, 보안성의 6개로 정의하였다.

두 번째 축인 데이터 품질 관리 프로세스는 6개의 품질 기준을 향상시키기 위해서 필요하거나 품질 기준에 영향을 주는 프로세스를 식별해 8가지 프로세스를 도출하였다. 여기에는 요구사항 관리, 데이터 구조 관리, 데이터 흐름 관리, 데이터베이스 관리, 데이터 활용 관리, 데이터



〈그림 2-1〉 데이터 품질관리 성숙모형 구조

이터 표준 관리, 데이터 오너십 관리, 사용자 뷰 관리가 포함된다. 8개의 프로세스는 6개의 품질 기준과의 연관 관계에 따라 상호 매핑되어 있다.

세 번째 축인 데이터 품질관리 성숙수준은 조직의 데이터 품질관리 발전 단계를 고려해 ‘도입–정형화–통합화–정량화–최적화’의 5단계로 성숙 단계를 구분하고, 6개의 품질 기준별로 데이터 품질관리 성숙 단계를 정의하였다.

즉, 『데이터 품질관리 성숙모형』은 정확성, 일관성, 유용성, 접근성, 적시성, 보안성의 데이터 품질 기준별로 관련 프로세스의 품질관리 수준을 점검해, 도입–정형화–통합화–정량화–최적화의 5단계로 성숙수준을 측정하도록 고안되었다. 따라서 특정 품질 기준을 향상시키기 위해 어떤 프로세스를 개선해야 하는지 혹은 특정 프로세스를 도입하면 가시적으로 어떤 데이터 품질 개선 효과를 얻을 수 있는지 알 수 있다.

2. 데이터 품질 기준

데이터 품질은 바라보는 시각에 따라 다양하게 표현될 수 있지만 가장 중요한 판단 요소는 해당 데이터가 업무 처리에 도움이 되는지 여부이다. 데이터가 업무에 도움을 주기 위해서는 다음과 같은 두 가지 조건을 만족해야 한다. 첫째, 제공 데이터는 업무에 유효하며 신뢰할 수 있어야 한다. 둘째, 데이터를 필요한 시점에 손쉽게 활용할 수 있어야 한다. 따라서 아무리 유효한 데이터라 해도 사용하는데 어려움이나 제약사항이 있어 활용하기에 어려움이 있다면 데이터의 품질이 우수하다고 할 수 없다.

본 모형에서는 데이터 품질의 다양한 측면을 데이터 품질 기준으로 명명하고, 품질에 대한 다양한 시각 중에서 보편성이 있으며 업무 처리에 중요한 품질의 특성을 기준으로 도출해 정의하였다.

이에 따라 데이터 품질 기준을 크게 유효성과 활용성으로 구분하고, 각 기준별 품질의 특성에 따라 유효성은 정확성과 일관성으로 활용성은 유용성, 접근성, 적시성, 보안성으로 구분해 총 6가지를 데이터 품질 기준으로 정의한다.

다음절부터는 정확성, 일관성, 유용성, 접근성, 적시성, 보안성과 같은 6개 품질 기준의 개념에 대해 설명한다.

데이터 품질 기준																
유효성							활용성									
정확성				일관성			유용성				접근성	적시성	보안성			
사실성	적합성	필수성	연관성	정합성	일치성	무결성	충분성	유연성	사용성	추적성			접근성	적시성	보호성	책임성

〈그림 2-2〉 데이터 품질 기준

2.1 유효성 관련 품질 기준

2.1.1 정확성

정확성은 실 세계에 존재하는 객체(사건, 사물, 개념 등)의 값이 오류 없이 저장되어 있음을 의미한다. 정확성은 사실성, 적합성, 필수성, 연관성과 같은 품질 기준으로 구성된다.

- 사실성

사실성은 데이터가 실 세계의 사실과 동일한 값을 가지고 있음을 의미한다. 일반적으로 사실성은 데이터의 원천 오류, 입력 오류, 입력 프로세스 문제 등에서 비롯된다. 따라서 사실성 향상을 위해서는 데이터 오너십 관리와 입력 프로세스 개선 등이 필요하다.

- 적합성

적합성은 데이터 값이 정해진 데이터 유효 범위를 충족하고 있음을 의미한다. 예컨대 정합성 오류는 데이터가 표준 코드 값 또는 표준 도메인 값에 위배될 경우 발생한다. 따라서 적합성 향상을 위해서는 표준 도메인 및 표준 코드에 대한 개선이 필요하다.

- 필수성

필수성은 조직의 업무 지원을 위해 반드시 필요한 필수 항목에 데이터의 누락이 발생하지 않음을 의미한다. 필수성 오류는 데이터 수집 당시의 확인 부족이나 데이터베이스상의 Not Null 체크 조건의 누락 등이 원인이 되어 발생한다. 따라서 데이터 별로 중요도를 선정하여 관리할 필요가 있다.

- 연관성

연관성은 연관 관계를 갖는 데이터 항목간에 논리상의 오류가 없음을 의미한다. 예컨대 해지 일자가 가입 일자보다 앞서 있는 경우가 여기에 해당된다. 연관성 확보를 위해서는 업무 규칙상의 데이터 연관 관계를 파악해 시스템에 반영해야 한다.

2.1.2 일관성

일관성은 정보시스템 내의 동일한 데이터 간에 불일치가 발생하지 않음을 의미한다. 일관

성 오류는 데이터에 대한 정의가 정확히 이루어지지 않거나, 데이터 참조무결성(referential integrity)이 불분명한 경우, 개별 시스템 단위로 설계·관리돼 전사 관점의 접근이 부족한 경우 등의 원인에 의해 발생한다.

예를 들어 조직 내의 구매부와 판매부에서 사용하는 ‘고객’ 컬럼의 정의는 부서별 특성에 따라 달라질 수가 있는데, 구매부의 ‘고객’은 주로 기업 및 기관이 해당되며 판매부의 ‘고객’은 개인 또는 기업이 해당된다. 이에 따라 부서별로 요구하는 ‘고객’ 테이블의 속성이 달라질 수가 있으며 같은 ‘고객’ 이면서도 분명하게 다른 ‘고객’으로 관리된다.

따라서 일관성을 확보하기 위해서는 전사 관점의 데이터 용어 표준화, 데이터 구조의 전사적인 조율, 중복 데이터에 대한 관리 등을 통해 데이터 간의 통일성을 유지하는 것이 필요하다.

일관성은 정합성, 일치성, 무결성과 같은 품질 기준으로 구성된다.

• 정합성

정합성은 기능, 의미, 성격이 동일한 데이터가 상호 동일한 용어와 형태로 정의되어 있음을 의미한다.

• 일치성

일치성은 동일한 용어를 다르게 정의하여 사용하지 않도록 함을 의미한다. 한 예로 남녀의 성별을 구분하는 용어인 성별 구분이 단위 시스템별로 남녀구분, 성별구분, 성별코드, 남녀코드로 용어가 서로 다르게 사용되면 전사 관점에서 데이터 통합 시에 시스템 자원의 낭비를 가져오고 데이터 정확성의 확보에도 영향을 미친다.

• 무결성

무결성은 데이터 처리의 선후 관계가 명확하게 준수되고 있음을 의미한다. 예를 들어 선행 데이터가 존재해야 후속 처리가 가능한 경우가 있다. 고객의 주문 내역은 고객이 먼저 등록이 되어야만 생성 가능하며, 고객이 삭제되면 주문 내역도 함께 삭제되어야 하는 경우가 여기에 해당된다. 무결성은 참조무결성을 데이터 구조 설계에 반영하고 데이터베이스나 프로그램을 통해 이를 구현함으로써 확보할 수 있다.

2.2 활용성 관련 품질 기준

2.2.1 유용성

유용성은 조직이 요구하는 데이터의 범위와 상세화 정도를 충족시킬 수 있음을 의미한다. 유용성 확보를 위해서는 사용자의 요구에 대한 관리가 필요하다. 요구 관리는 사용자들의 추상적인 요구를 구체화하여 관리하는 제반 활동을 의미하며, 도출된 요구를 정해진 기준과 원칙에 따라 요구사항 수행 범위와 우선 순위를 결정하여 반영하고 일련의 결과를 문서화하여 관리해야 한다.

유용성은 충분성, 유연성, 사용성, 추적성과 같은 품질 기준으로 구성된다.

• 충분성

충분성은 제공 데이터가 사용자의 요구사항을 충분히 충족시킬 수 있음을 의미한다. 충분성의 확보를 위해서는 사용자의 현재 요구사항뿐만 아니라, 업계의 동향을 참조하여 향후 발생 가능한 요구사항에 대해서도 사전에 준비하는 노력이 필요하다.

• 유연성

유연성은 데이터가 사용자의 다양한 요구사항을 수용할 수 있는 유연한 구조를 가지고 있음을 의미한다. 즉, 데이터 구조는 다양한 요구사항과 업무를 수용할 수 있어야 하며 변경 사항 발생시에도 유연하게 대처할 수 있어야 한다.

• 사용성

사용성은 실제 공급되는 데이터가 현장에서 유용하게 사용되고 있어야 함을 의미한다. 사용성 향상을 위해서는 데이터 사용 현황 및 사용 빈도를 분석하여 활용 빈도가 높은 데이터에 시스템 자원을 우선 배분하는 등의 활동이 필요하다.

• 추적성

추적성은 데이터의 변경 내역이 관리되고 있음을 의미한다. 데이터에 대한 이력 관리를 수행하면 사용자의 요구사항 변경에 대해 유연하게 대처할 수 있다.

2.2.2 접근성

데이터에 대한 사용자 만족도를 총족시키기 위해서는 데이터 자체의 품질은 물론, 데이터를 효율적·효과적으로 제공하는데 필요한 인터페이스 등 제반 서비스 품질을 확보해야 한다. 접근성은 사용자가 원하는 데이터를 손쉽게 이용할 수 있음을 의미한다.

접근성은 사용의 용이성 관점과 검색의 용이성 관점으로 나누어 살펴 볼 수 있다. 사용의 용이성은 정보시스템에서 인터페이스, 도움말, 고객지원 등이 사용자가 데이터를 이용하는데 불편함이 없도록 제공되고 있음을 의미한다. 검색 용이성은 정보시스템에서 제공하는 검색 관련 제반 기능이 사용자가 원하는 데이터를 손쉽고 편리하게 추출하여 활용할 수 있도록 지원되고 있음을 의미한다.

2.2.3 적시성

적시성은 응답시간과 같은 비기능적 요구사항 그리고 데이터의 최신성 유지와 같은 품질 요건에 얼마나 잘 대처하고 있는지를 의미한다. 적시성은 특히 운영시스템의 데이터를 분석 데이터로 변환하여 조직 내부의 의사 결정 등에 활용할 경우 그 중요성이 더 커진다. 적시성이나 최신성이 확보되지 못한 데이터로 조직의 중요한 의사 결정을 내리게 될 경우 치명적인 결과를 가져올 수 있기 때문이다.

적시성을 향상시키기 위해서는 데이터베이스 튜닝을 통한 성능 개선, 최적의 데이터 구조 유지, 자원의 효율성 확보를 위한 정보의 생명주기 관리 등을 지속적으로 수행해야 한다.

2.2.4 보안성

보안성은 외부 및 내부 요인으로부터 데이터를 적절히 보호하고 있는지 여부를 의미한다. 여기에는 외부 침입이나 재난 등의 위험으로부터의 데이터 보호, 시스템의 에러나 장애 원인의 사전 차단, 에러나 장애 발생 시 데이터 사용의 중단 및 지역 최소화, 명확한 사용자 권한 정의 등이 포함된다.

보안성은 보호성, 책임성, 안정성과 같은 데이터 품질 기준으로 구성된다.

- **보호성**

보호성은 훼손, 변조, 유출 등의 다양한 형태의 위협으로부터 데이터를 안전하게 보호하고 있음을 의미한다.

- **책임성**

책임성은 사용자 접근 권한과 책임을 명확히 부여하고 있음을 의미한다.

데이터 보호 및 사용자 권한 접근 관리에서 중요한 것은 보호 대상에 대해 정확한 정의를 내리고 이에 대한 관리 체계를 구성하는 것이다. 보안 활동은 관리적 보안 활동과 기술 및 물리적 보안 활동으로 구분할 수가 있다. 관리적 보안 활동에서는 사람, 문서, 재해 복구 계획 등을 주요 관리 대상으로 선정해 이에 대한 조직 보안 규정 및 지침을 수립하여 시행해야 한다. 기술 및 물리적 보안 활동에 물리적 데이터베이스 저장소에 대한 출입 통제 관리와 외부에 의한 해킹 피해, 바이러스에 의한 피해로부터 보안 활동을 수행해야 한다.

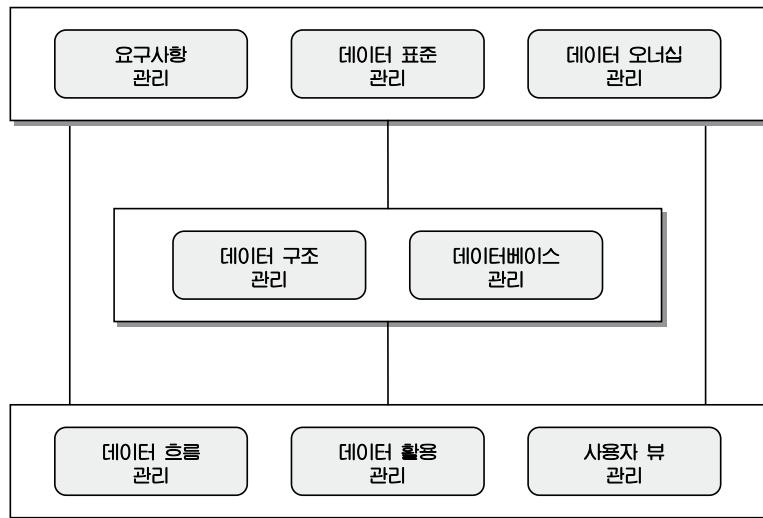
- **안전성**

안전성은 시스템의 여러나 장애를 사전에 차단하고 여러나 장애 발생 시 중단 및 자연을 최소화할 수 있는 체제로 운영함을 의미한다.

데이터 소실 사고는 다양한 요인으로 인해 언제 어디서나 일어날 수 있기 때문에, 오류 발생 시 데이터를 효율적으로 빠른 시간 안에 복구할 수 있는 절차가 필요하다.

3. 데이터 품질관리 프로세스

데이터 품질관리 프로세스는 요구사항 관리, 데이터 구조 관리, 데이터 흐름 관리, 데이터 베이스 관리, 데이터 활용 관리, 데이터 표준 관리, 데이터 오너십 관리, 사용자 뷰 관리의 8개 프로세스로 구분할 수 있다.



〈그림 2-3〉 데이터 품질관리 프로세스

3.1 요구사항 관리

요구사항 관리는 데이터에 대한 사용자의 요구를 도출·관리하는 프로세스로, 유용성 데이터 품질 기준과 상관 관계를 가진다.

사용자에게 유용한 데이터를 제공하기 위해서는 사용자가 원하는 것이 무엇인지 명확하게 파악해야 한다. 그런데 사용자의 요구는 대개의 경우 추상적이거나 개념적으로 표현된다. 따라서 사용자와의 지속적인 의견 교환을 통해 요구를 구체화하고 이를 효과적으로 지원할 수 있는 방안을 도출해야 한다.

요구사항 관리에서 데이터 품질과 관련이 있는 세부 프로세스에는 기능적 요구사항 관리와 비기능적 요구사항 관리가 있다.

• 기능적 요구사항 관리

기능적 요구사항 관리는 업무 처리 과정에서 기능적으로 필요한 데이터 요건을 도출·조정하여 관리함을 의미한다. 데이터에 대한 요건은 실무자, 중간관리자, 경영자 등 다양한 관리 계층에서 수렴되어야 한다.

• 비기능적 요구사항 관리

비기능적 요구사항 관리는 처리 시간, 응답 시간, 데이터 제공 주기 등과 같은 업무 내용적 이외의 요건을 도출·조정하여 관리함을 의미한다. 비기능적 요구사항에 적절히 대응하기 위해서는 풍부한 전산 자원의 지원이 필요하다. 그러나 투입 가능한 자원에는 한계가 있으므로 비기능적 요건과 투입 자원의 비용 대비 효과를 고려해 효율성 있는 관리를 수행해야 한다.

3.2 데이터 구조 관리

데이터 구조 관리는 사용자가 원하는 데이터를 제공할 수 있도록 데이터 모델 및 데이터베이스를 설계·관리하는 프로세스를 의미한다.

데이터 구조는 도서관의 책 분류에 비유할 수 있다. 만일 책의 분류 체계가 잘못되어 있다면 아무리 좋은 책을 많이 가지고 있다고 해도 책의 보관이나 활용이 원활하지 않게 된다. 데이터도 마찬가지로 적절하게 설계·관리되지 않으면 데이터의 제공에 많은 제약이 발생한다.

데이터 구조 관리에서 데이터 품질과 관련 있는 세부 프로세스에는 유연성 관리, 중복 관리, 참조무결성 관리가 있다.

• 유연성 관리

유연성 관리는 예측 가능한 미래의 변화를 고려해 종합적이고 장기적인 시각에서 데이터 구조를 도출·관리함을 의미한다.

조직의 업무 환경이 변화함에 따라 사용자의 데이터에 대한 요건도 수시로 변경된다. 그런데 요건이 바뀔 때마다 이에 대응하기 위해 데이터 구조를 변경해야 한다면 많은 혼선이 발생할 것이다. 따라서 데이터 구조는 추가적인 데이터 요건에 대해 변경을 최소화하면서 수용할 수 있는 체계를 갖추고 있어야 하며, 이를 위해 유연성을 고려한 데이터 구조 설계가 필요하다.

• 중복 관리

동일한 데이터는 한 곳에만 존재하는 것이 바람직하다. 그러나 업무적 · 시스템적인 이유로 동일 데이터가 여러 곳에 존재하는 경우가 많다. 데이터의 중복이 발생하게 되면 관리에 추가적인 노력도 필요하며, 중복된 데이터 간의 불일치가 발생하여 업무 혼선을 초래 할 수 있다. 그러므로 일차적으로는 데이터 중복을 최소화하는 것이 중요하며 불가피하게 데이터 중복이 발생하였다면 중복 데이터 간의 일치성을 확보해야 한다. 중복 관리는 중복에 대한 사전적인 예방 및 중복 발생 시의 사후적인 관리를 모두 포함한다.

• 참조무결성 관리

데이터의 생성, 변경, 삭제 과정에는 작업의 선후 관계가 존재하는 경우가 많다. 상품의 등록과 판매를 예로 들면 상품의 판매 데이터는 상품 데이터가 사전에 등록되어 있어야만 입력될 수 있다. 즉, 데이터 간의 선후 관계를 파악해 데이터 처리 과정에서 이를 준수 할 수 있도록 통제하는 참조무결성 관리는 데이터 품질 확보에 매우 중요한 활동이다.

• 통합 관리

통합 관리는 전사 시스템 관점에서 데이터 구조를 설계 · 관리함을 의미한다. 데이터 구조의 통합 관리는 개별 시스템 관점이 아니라 전사 조직 차원에서의 데이터 활용성 극대화를 목적으로 한다. 통합 관리는 데이터 구조 설계 이전에 고려되어야 하며, 시스템별로 개별적으로 관리하고 있다면 이를 통합하여 전체적인 관점에서 관리 · 제공할 수 있는 방안을 마련해야 한다.

3.3 데이터 흐름 관리

데이터 흐름 관리는 소스 데이터를 타깃에 생성하거나 추출, 변환, 적재하는 제반 과정을 관리하는 프로세스를 의미한다.

조직에 존재하는 시스템이 여러 개일수록 시스템간의 데이터 흐름도 많고 복잡해진다. 특히 데이터가 분산되어 있거나 중복이 많은 경우 데이터 오류 가능성이 높아지므로 품질 확보를 위해 데이터 흐름에 대한 철저한 관리가 필요하다. 데이터 흐름 관리를 위해서는 데이터 흐름에 대한 요건을 파악해 데이터 흐름을 통제하고 데이터 흐름의 처리 결과에 대한 정확성 확인이 병행되어야 한다.

데이터 흐름 관리에서 데이터 품질과 관련이 있는 세부 프로세스에는 흐름 주기 관리와 흐름 대사 관리가 있다.

• 흐름 주기 관리

흐름 주기 관리는 데이터 흐름의 주기를 파악해 필요한 시점에 필요한 흐름이 정상적으로 이루어질 수 있도록 관리하는 것을 의미한다. 데이터 흐름은 즉시, 지연, 일별, 주별, 월별, 수시 등과 같은 다양한 업무 요건을 반영하고, 각 흐름간의 선후 관계를 고려하여 관리해야 한다.

• 흐름 대사 관리

흐름 대사 관리는 정해진 요건에 따라 데이터 흐름이 정확히 이루어지고 있는지 확인 · 관리하는 것을 의미한다. 데이터 흐름에 대한 상세 내역이 관리되지 않으면 데이터 흐름 상에서 발생하는 오류를 발견하거나 추적하기 어렵다. 따라서 흐름 대사 관리를 통해 원천 데이터와 목표 데이터 간의 대응 관계를 파악해 상호 간의 데이터 흐름의 정확성을 확인하고 오류가 발생한 경우 원인을 추적하고 해결할 수 있어야 한다.

3.4 데이터베이스 관리

데이터베이스 관리는 데이터 값이 실제로 저장되는 데이터베이스를 운영 · 관리하는 프로세스를 의미한다. 데이터베이스는 데이터 값에 대한 지속적인 변경이 일어나는 장소이므로 유

동적이고 변화가 많다. 그러므로 다양한 상황에 대한 충분한 대응 방안이 마련되어야 한다.

데이터베이스 관리에서 데이터 품질과 관련이 있는 세부 프로세스는 성능 관리와 보안 관리가 있다.

• 성능 관리

성능 관리는 사용자가 원하는 데이터를 원하는 시간에 신속하게 제공할 수 있도록 수행하는 관리 활동을 의미한다. 데이터베이스의 처리 상황에 대한 지속적인 모니터링과 성능의 최적화를 위한 튜닝 활동, 그리고 적정 성능 유지를 위해 필요한 전산 자원을 예측하고 자원 조달에 대한 계획, 배분, 관리 활동 등이 포함된다.

• 보안 관리

보안 관리는 데이터베이스에 저장된 데이터가 오류 및 훼손 없이 안정적으로 서비스될 수 있도록 관리하는 활동을 의미한다. 이를 위해 데이터에 대한 권한을 통제하거나 백업 및 장애 대책 등을 통해 데이터베이스에 발생한 문제로부터 신속하게 복구할 수 있는 체계를 마련해야 한다.

3.5 데이터 활용 관리

데이터 활용 관리는 사용자의 데이터 활용 현황을 파악하고 데이터 활용을 극대화할 수 있는 환경을 구현하기 위해 수행하는 관리 활동을 의미한다.

데이터 활용 관리에서 데이터 품질과 관련이 있는 세부 프로세스는 업무 규칙 검증과 활용 모니터링이 있다.

• 업무 규칙 검증

업무 규칙 검증은 데이터 품질 점검 기준 및 방법 정의, 데이터 품질 점검, 데이터 정제를 수행하는 일련의 작업을 의미한다. 업무 규칙 검증을 통해 데이터의 비효율적 사용을 예방하고 데이터의 운용 중에 발생할 수 있는 데이터 품질 저하를 예방할 수 있다. 검증의 효과를 극대화하기 위해서는 데이터 품질 점검 기준과 방법이 정교하게 정의되어 있어야 하며 품질 점검 결과 발견된 오류를 체계적으로 추적·조치하는 절차가 강화되어야 한다.

- **활용 모니터링**

활용 모니터링은 사용자의 만족도를 제고하고 데이터의 활용도를 높이기 위해 주기적으로 데이터 사용 현황을 점검하는 작업을 의미한다.

3.6 데이터 표준 관리

데이터 표준 관리는 데이터에 대한 공통된 시각을 유지할 수 있도록 데이터 표준화 원칙 정의, 표준화 원칙에 따른 표준 정의, 표준 변경, 표준 적용 통제 등을 수행하는 관리 프로세스를 의미한다.

데이터 표준 관리와 관련성이 높은 데이터 품질 기준은 일관성 및 정확성이며, 데이터 관련자들 간의 공통된 용어와 기준을 사용하게 함으로써 데이터 품질을 향상시킬 수 있다.

데이터 표준 관리 프로세스 가운데 데이터 품질과 관련된 세부 프로세스에는 용어 표준 관리와 도메인 · 코드 표준 관리가 있다.

- **용어 표준 관리**

데이터를 표현하는데 사용되는 용어를 표준화하여 모든 사람이 동일한 인식 하에서 의사소통을 할 수 있는 기반을 마련하는 것이 중요하다. 이를 위해 데이터 관리에 수용되는 용어를 도출해 의미를 정의하고, 데이터를 경유한 의사소통이나 산출물의 작성 시에 표준화된 용어를 사용하게 함으로써 명확한 의사전달이 이루어질 수 있도록 통제해야 한다.

- **도메인 · 코드 표준 관리**

도메인 표준은 전사적으로 사용되고 있는 데이터 가운데에 논리적 · 물리적으로 유사한 유형의 데이터를 그룹화하여 해당 그룹에 속하는 데이터의 유형과 길이를 정의한 것을 말한다. 코드 표준은 다양하게 나타날 수 있는 데이터 값을 정형화하기 위해 정의된 기준에 따라 제한된 범위 내의 기호로 대치한 것을 의미한다. 표준 코드에는 각 산업별로 법적 · 제도적으로 부여하여 공통적으로 사용되는 코드뿐만 아니라 기관이나 기업 내부에서 정의하여 사용하는 코드가 대상이 된다.

3.7 데이터 오너십 관리

데이터 오너십 관리는 데이터에 책임과 권한을 가진 오너를 중심으로 데이터 품질관리 제반 절차나 과정을 정의하고 실행하는 모든 활동을 의미한다. 여기서 데이터 오너란 데이터의 생성, 변경, 활용, 삭제에 대한 권한을 가지고 있는 책임자를 의미한다.

데이터 오너는 데이터에 대한 품질 기준을 정의하고 데이터 처리 권한 및 프로세스를 통제하며 데이터 품질에 대한 사후 검증을 수행하여 데이터 오류에 대해 조치한다.

3.8 사용자 뷰 관리

사용자 뷰 관리는 사용자가 최종적으로 데이터를 이용하는 환경에 대한 관리를 의미한다. 사용자 뷰 관리의 목적은 사용자의 데이터에 대한 접근 장벽을 최소화하여 데이터 활용을 극대화하는 데 있다. 따라서 사용자 뷰는 사용자 관점에서 설계·구현되어야 한다. 또한 사용자 뷰를 구현하는 기술은 변화가 매우 빠른 만큼, 관련 기술 동향을 파악하여 시의 적절하게 적용하는 것도 중요하다.

3.9 데이터 품질 기준과 관리 프로세스 관계

앞에서 정의한 데이터 품질관리 프로세스는 정확성, 일관성, 유연성, 접근성, 적시성, 보안성과 같은 품질 기준에 대한 영향 관계에 따라 상호 매핑이 가능하다. 데이터 품질관리 프로세스와 데이터 품질 기준간의 관계를 정리하면 <표 2-1>과 같다.

예를 들어 요구사항 관리(기능적 관리) 프로세스는 유용성 데이터 품질에 영향을 주고, 요구사항 관리(비기능적 관리) 프로세스는 적시성 데이터 품질에 영향을 준다. 반대로 유용성 품질을 확보하기 위해서는 요구사항 관리(기능적 관리), 데이터 구조 관리(유연성 관리), 데이터 활용 관리(활용 모니터링) 프로세스를, 적시성 품질을 확보하기 위해서는 요구사항 관리(비기능적 관리), 데이터 흐름 관리(흐름 주기 관리), 데이터베이스 관리(성능 관리) 프로세스를 운용·개선해야 한다.

따라서 조직은 업무 및 관리 데이터의 특성을 고려해 데이터 품질 기준의 우선 순위를 정하고,

〈표 2-1〉 데이터 품질 기준과 관리 프로세스 연관도

프로세스	품질 기준	정확성	일관성	유용성	접근성	적시성	보안성
요구사항 관리	기능적 관리			●			
	비기능적 관리					●	
데이터 구조 관리	유연성 관리			●			
	종복 관리		●				
	참조 무결성 관리		●				
	통합 관리				●		
데이터 흐름 관리	흐름 주기 관리					●	
	흐름 대사 관리		●				
데이터베이스 관리	성능 관리					●	
	보안 관리						●
데이터 활용 관리	업무 규칙 검증	●					
	활용 모니터링			●			
데이터 표준 관리	용어 표준 관리		●				
	도메인. 코드 표준 관리	●	●				
데이터 오너십 관리	데이터 오너십 관리	●	●				
사용자 뷰 관리	사용자 뷰 관리				●		

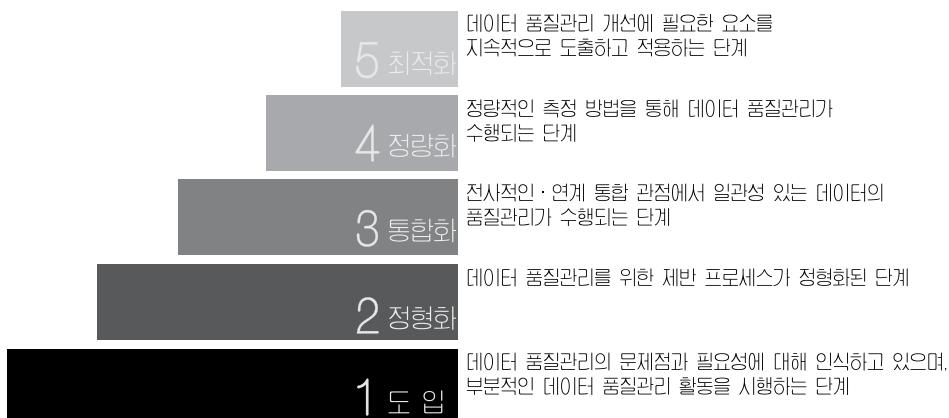
해당 품질 기준과 연계된 프로세스를 우선적으로 개선·관리할 수 있다. 또한 특정 관리 프로세스를 체계화함으로써 어떠한 데이터 품질 개선 효과를 얻을 수 있는지 알 수 있어 실무적인 입장에서 가시적인 성과를 바탕으로 품질관리 활동을 전개할 수 있다.

4. 데이터 품질관리 성숙수준

4.1 데이터 품질관리 성숙수준 단계

데이터 품질관리 성숙수준은 조직이 수행하고 있는 데이터 품질관리의 체계화 정도를 나타낸다. 즉, 성숙수준이 높을수록 체계적이며 정교한 데이터 품질관리가 수행되고 있음을 의미한다.

『데이터 품질관리 성숙도형』은 데이터 품질관리 성숙수준을 도입-정형화-통합화-정량화-최적화의 5단계로 구분한다.



〈그림 2-4〉 데이터 품질관리 성숙수준 단계

성숙 단계 1은 도입 단계이다. 도입 단계는 데이터 품질관리의 초기 단계로 데이터 품질관리에 대한 문제점과 필요성을 부분적으로는 인지하고는 있지만 표준화와 정형화된 프로세스가 미비하며, 데이터의 품질관리는 담당자 개개인의 능력에 의존해 수행된다. 기본적인 수준의 식별 및 인증, 백업 계획이 수립되고 있으나 전반적으로 데이터의 오류 발생 소지가 높다.

성숙 단계 2는 정형화 단계이다. 정형화 단계는 데이터 품질관리를 위한 정책 및 규정과 데이터 품질관리와 관련된 제반 프로세스가 정형화되어 있으며, 정의된 절차에 따라 기본적인

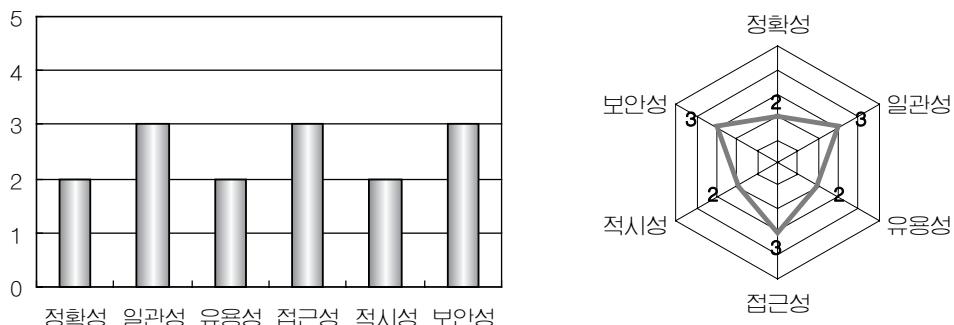
품질관리가 이루어진다. 데이터의 품질 상태는 누락 및 오류 발생에 대해 기본적인 대응이 가능하며 데이터 표준을 정의하여 단위 시스템 및 일부 조직에서 사용하여 기본적인 운영 활동을 수행하고 있다.

성숙 단계 3은 통합화 단계이다. 통합화 단계는 전사적인 연계 및 통합 관점에서 일관성 있는 데이터의 품질관리가 이루어지고 있으며 데이터가 누락이나 오류 없이 서비스되고 있다. 데이터 표준이 조직 전체에 반영되어 있으며, 메타데이터가 활용되어 데이터 요소 간의 연계성이 확립되어 있다. 또한 지속적인 성능, 보안, 재해 복구 등의 프로세스가 안정화돼 있으며 개선 작업이 수행된다. 데이터 품질은 정성적으로 관리되고 있으며 운영 성과가 정량화되어 있지 않다.

성숙 단계 4는 정량화 단계이다. 정량화 단계는 통계적 기법이나 정량적인 측정 방법을 통해 관리되는 단계로, 프로세스가 지속적·안정적으로 예측 가능하게 유지되고 있으며, 정량적인 품질의 목표를 설정하고 달성을 여부를 확인한다.

성숙 단계 5는 최적화 단계이다. 최적화 단계는 품질관리 프로세스의 개선 사항을 지속적으로 도출하고 실행하며 평가를 통해 사후 관리를 수행하는 단계로, 현재의 관점에서 최적화뿐만 아니라 지속적인 개선 노력을 통하여 미래의 환경 변화에 유연하게 대처할 수 있는 수준이다.

데이터 품질관리 성숙수준은 정확성, 일관성, 유용성, 접근성, 적시성, 보안성의 6가지 데이터 품질 기준별로 측정한다. 성숙수준의 측정 방법은 품질 기준에 따라 상이한데, 이것은 기준별로 품질에 영향을 주는 관련 프로세스가 서로 다르기 때문이다. 『데이터 품질관리 성숙모형』 및 성숙수준 적용 결과는 아래의 그림과 같이 도식화할 수 있다.



〈그림 2-5〉 데이터 품질관리 성숙수준 측정 결과

4.2 데이터 품질관리 프로세스와 성숙수준 관계

데이터 품질 기준별 관련 프로세스 및 성숙수준과의 관계를 정리하면 아래의 표와 같다.

〈표 2-2〉 품질관리 프로세스와 성숙수준 관계

품질 기준	품질관리 프로세스	성숙수준				
		1	2	3	4	5
정확성	데이터 활용 관리 (업무 규칙 검증)	●	●	●	●	●
	데이터 표준 관리 (도메인, 코드 표준 관리)		●			
	데이터 오너십 관리			●	●	
일관성	데이터 구조 관리 (중복 관리)	●	●	●	●	●
	데이터 구조 관리 (참조 무결성 관리)	●	●	●	●	●
	데이터 흐름 관리 (흐름 대사 관리)	●	●	●	●	●
	데이터 표준 관리 (용어 표준 관리)		●	●		
	데이터 표준 관리 (도메인, 코드 표준 관리)		●	●		
	데이터 오너십 관리			●	●	
유동성	요구사항 관리 (기능적 관리)	●	●	●	●	●
	데이터 구조 관리 (유연성 관리)			●		
	데이터 활용 관리 (활용 모니터링)		●			
접근성	사용자 뷰 관리	●	●	●	●	●
	구조 관리 (통합 관리)	●	●	●		
적시성	요구사항 관리 (비기능적 관리)	●	●	●	●	●
	데이터 흐름 관리 (흐름 주기 관리)	●	●	●	●	●
	데이터베이스 관리 (성능 관리)		●	●		
보안성	데이터베이스 관리 (보안 관리)	●	●	●	●	●

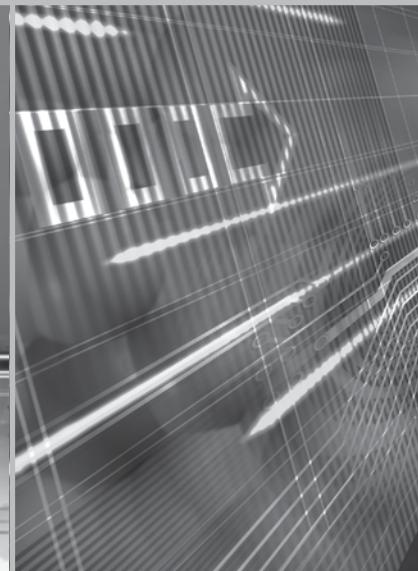
예를 들어 데이터 활용 관리(업무 규칙 검증), 데이터 표준 관리(도메인, 코드 표준 관리), 데이터 오너십 관리 프로세스는 정확성 품질 기준에 영향을 준다. 여기서 데이터 활용 관리(업무 규칙 검증)는 정확성 1단계에서 5단계의 모든 성숙수준과 관련이 있다. 반면 데이터 표준 관리(도메인, 코드 표준 관리)는 정확성 2단계 성숙수준과, 데이터 오너십 관리는 정확성 3단

계와 4단계 성숙수준과 관련이 있다.

따라서 데이터 품질관리 정확성 관련 성숙수준이 2단계인 조직이 3단계로 올라가기 위해서는 데이터 활용 관리(업무 규칙 검증)와 데이터 오너십 관리 프로세스에 대한 개선 작업을 수행해야 한다.

제3장 품질 기준별 데이터 품질관리 성숙수준

제3장에서는 데이터 품질 기준별로 운용 및 개선이 필요한 프로세스를 소개하고 각 기준별로 충족해야 할 품질관리 성숙수준에 대해 설명한다. 그리고 품질 기준별로 개선에 필요한 품질관리 활동에 대해서도 기술한다.



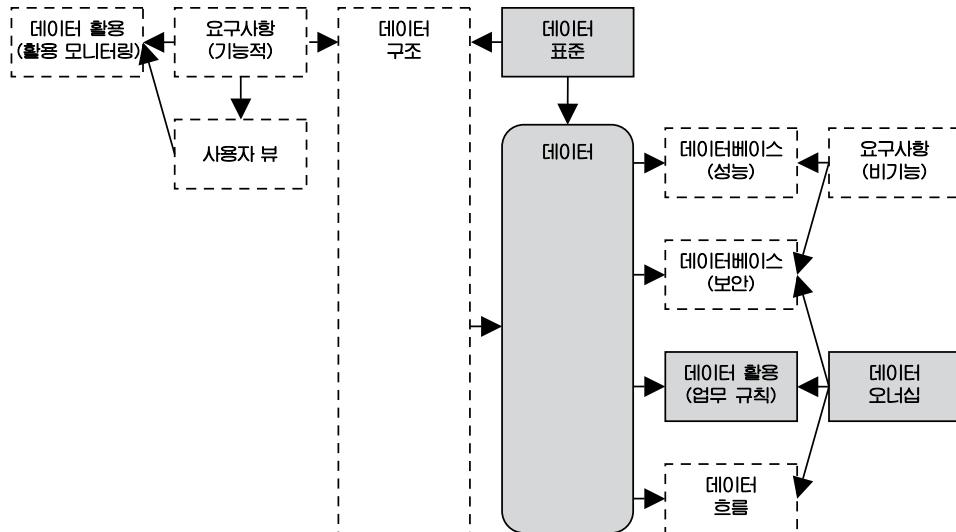
1. 유효성 관련 성숙수준

1.1 정확성 성숙수준

1.1.1 정확성 관련 프로세스

정확성은 실 세계에 존재하는 객체(사건, 사물, 개념 등)의 값이 오류 없이 정확하게 저장되어 있음을 의미한다. 데이터가 실 세계의 사실과 같은 값을 보유하고 있는지(사실성), 코드 및 구분 값이 유효한 범위 내에 있는지(적합성), 반드시 존재해야 하는 데이터가 존재하고 있는지(필수성), 연관 관계가 있는 항목간에 논리상 오류는 없는지(연관성) 등이 포함된다.

데이터의 정확성을 확보하기 위해 관리 및 개선되어야 할 프로세스에는 데이터 활용 관리(업무 규칙 검증), 데이터 표준 관리(도메인, 코드 표준 관리), 데이터 오너십 관리 등이 있다.



〈그림 3-1〉 정확성 관련 데이터 품질관리 프로세스

데이터 활용 관리(업무 규칙 검증)는 업무 규칙 정의 및 검증을 통해 데이터의 정확성을 검증하는 프로세스이며, 데이터 표준 관리(도메인, 코드 표준 관리)는 코드 및 도메인의 표준을 정의, 변경 및 운영하는 프로세스이다. 데이터 오너십 관리는 데이터에 책임과 권한을 가진 오너를 선정해 오너를 중심으로 데이터 품질관리 제반 절차나 과정을 정의하고 실행하는 프로세스를 의미한다.

1.1.2 정확성 관련 품질관리 성숙수준

데이터 활용 관리(업무 규칙 검증), 데이터 표준 관리(도메인, 코드 표준 관리), 데이터 오너십 관리에 기초하여, 도입–정형화–통합화–정량화–최적화의 데이터 품질관리 성숙 단계에 따라 정확성 관련 데이터 품질관리 성숙수준을 정의하면 다음과 같다.

- 단계 1: 도입

데이터 품질관리의 문제점에 대해 인지하고 있으며, 관리가 필요한 업무 규칙을 정의해

검증 작업을 수행하고 있다. 또한 제반 활동이 담당자 개인적 또는 부분적으로 수행된다. 즉, 문제점에 대해 인지하고 검증은 하지만 정형화되거나 전산화되어 있지 않아 체계적으로 관리되지 않고 오류도 많다.

• 단계 2: 정형화

관리가 필요한 업무 규칙을 명확히 식별해 표준화된 문서로 관리하며, 툴 등의 지원 도구를 활용하거나 체계적으로 검증·개선한다. 또한 도메인 및 코드 표준을 정의해 문서로 관리하며 데이터베이스 상의 준수 여부를 주기적으로 체크하고 개선한다. 즉, 데이터 품질 관리 검증 및 오류 개선과 관련된 제반 프로세스가 정형화되어 있다.

• 단계 3: 통합화

관리가 필요한 업무 규칙에 대해 전사적 차원에서 검증을 수행하여 발생 오류의 원인을 추적해 분석하고, 오류 개선 현황을 문서 또는 데이터베이스를 활용하여 체계적으로 관리한다. 즉, 전사적인 관점에서 데이터 품질 검증·추적·개선 관리가 안정적으로 운용되고, 정성적인 관리를 수행한다.

• 단계 4: 정량화

관리가 필요한 업무 규칙에 대한 검증 결과 또는 데이터 오류 개선 결과에 대해 정량적인 목표를 부여하고 측정 지표를 적용해 주기적으로 성과를 측정한다. 또한 데이터 오류 내역을 지식화하여 관리하며 필요에 따라 관련 프로세스 담당자의 책임과 역할을 조정한다. 즉, 전사 차원에서 정성적인 관리에서 정량적인 관리로 전이되고, 품질에 대한 성과가 명확히 측정되고 지식화되는 수준이다.

• 단계 5: 최적화

오류가 발생하지 않더라도 지식화되어 관리되고 있는 업무 규칙과 오류 패턴을 분석하여 데이터 품질관리 문제 발생의 근본 원인을 제거하고 지속적인 프로세스 개선 활동을 수행한다. 즉, 현재의 품질관리 최적화뿐만 아니라 향후의 환경에 유연하게 대처할 수 있다.

1.1.3 정확성 관련 데이터 품질관리

1.1.2에서 정의한 정확성 데이터 품질관리 성숙수준을 충족하기 위해 반드시 수행되어야 할

품질관리는 다음과 같다.

[데이터 활용 관리(업무 규칙 검증)]

- 관리 대상이 정의된다.
- 정확성에 대한 확인이 수행된다.
- 관리 대상이 체계적으로 정의된다.
- 정확성의 확인이 계획적으로 이루어진다.
- 체계적인 오류 추적이 가능하다.
- 관리 지표가 존재한다.
- 오류 해결에 대한 지식이 축적된다.
- 오류의 근본 원인 제거 작업이 수행된다.

[데이터 표준 관리(도메인, 코드 표준 관리)]

- 정확성을 위한 표준을 정의한다.

[데이터 오너십 관리]

- 역할과 책임이 명확하게 정의된다.
- 활동에 대한 성과 관리를 수행한다.

상기의 정확성 관련 데이터 품질관리를 성숙 단계별로 정의하여 상세화하면 다음과 같다.

〈표 3-1〉 정확성 관련 데이터 품질관리 수준

성숙수준	품질관리 수준
1	관리 대상이 정의된다. - 정확성에 대한 업무 규칙을 정의하며 이를 최신의 상태로 유지한다.
	정확성에 대한 확인이 수행된다. - 정확성을 확인하기 위한 방법이 존재하고 확인 결과를 문서로 기록한다.
2	정확성을 위한 표준을 정의한다. - 도메인, 코드 표준을 정의하고 유지하며 표준의 준수를 위한 절차를 정립하고 적용한다.
	관리 대상이 체계적으로 정의된다. - 컬럼의 필수 여부, 유효 범위, 컬럼간 관계 등의 정확성 기준을 식별하고 대상에 대한 기록 및 유지를 정형화 한다.
	정확성의 확인이 계획적으로 이루어진다. - 확인 방법이 정형화되거나 Tool을 활용하여 주기적으로 시행한다.

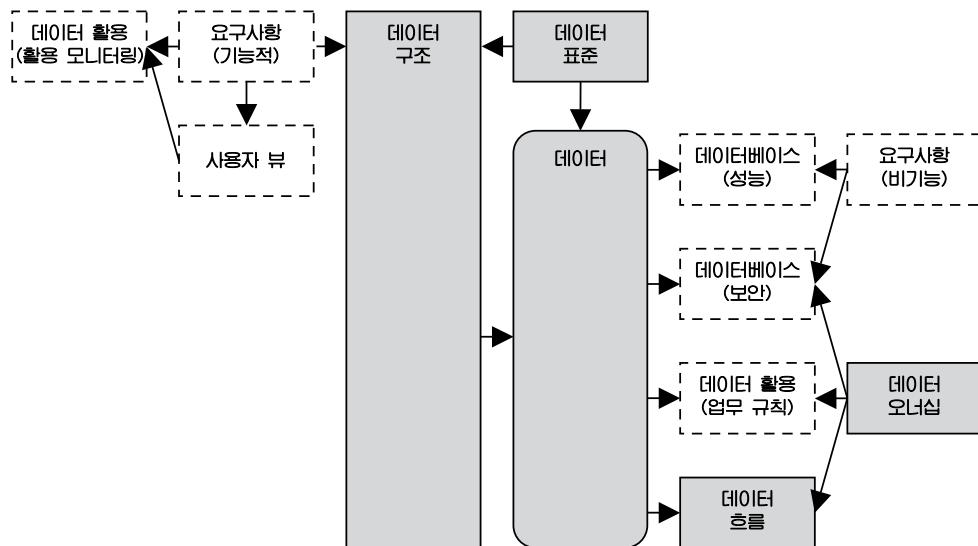
성숙수준	품질관리 수준
3	<p>역할과 책임이 명확하게 정의된다.</p> <ul style="list-style-type: none"> - 관리 대상에 대하여 오너를 부여하고 오너의 역할과 책임에 대하여 규정하여 관리 활동을 수행한다.
	<p>체계적인 오류 추적이 가능하다.</p> <ul style="list-style-type: none"> - 오류와 연관된 데이터베이스 또는 애플리케이션에 대한 원인 및 담당자 추적이 가능하고 오류에 대한 대응 활동을 추적하고 기록한다.
4	<p>관리 지표가 존재한다.</p> <ul style="list-style-type: none"> - 정확성 측정 지표를 정의하고 주기적으로 측정한다.
	<p>활동에 대한 성과관리를 수행한다.</p> <ul style="list-style-type: none"> - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
	<p>오류 해결에 대한 지식이 축적된다.</p> <ul style="list-style-type: none"> - 오류 해결 내역을 체계적으로 분류하여 지식을 축적하고 오류 대응 시 이를 적극적으로 활용한다.
5	<p>오류의 근본 원인 제거 작업이 수행된다.</p> <ul style="list-style-type: none"> - 오류의 근본 원인을 파악하여 이를 해결하기 위한 계획을 수립하고 관련 업무 프로세스를 개선한다.

1.2 일관성 성숙수준

1.2.1 일관성 관련 프로세스

일관성은 동일한 데이터 간에 통일성이 유지되고 있는지 여부를 판단하는 품질 기준이다. 여기에는 서로 다른 곳에 존재하는 동일한 데이터가 항상 일치하는지(정합성), 동일한 용어를 동일하게 정의하여 사용하는지(일치성), 정의한 규칙에 맞게 데이터가 존재하는지(무결성) 등이 포함된다.

데이터의 일관성을 확보하기 위해 관리 및 개선되어야 할 프로세스에는 데이터 구조 관리(중복 관리 / 참조무결성 관리), 데이터 흐름 관리(흐름 대사 관리), 데이터 표준 관리(용어 표준 관리 / 도메인, 코드 표준 관리), 데이터 오너십 관리 등이 있다.



〈그림 3-2〉 일관성 관련 데이터 품질관리 프로세스

데이터 구조관리(중복 관리 / 참조무결성 관리)는 데이터 모델의 정의 · 변경 및 테이블과 레코드에 대한 제반 관리 활동을 정해진 원칙과 표준에 근거해 관리하는 프로세스이다. 데이터 흐름 관리(흐름 대사 관리)는 소스에서 타깃까지의 데이터 흐름에 대한 검증을 통해 소스 와 타깃의 데이터를 대사하는 프로세스이며, 데이터 표준 관리(용어 표준 관리 / 도메인, 코드 표준 관리) 표준 원칙 정의 및 표준 정의, 표준 변경, 표준 적용 통제의 전 과정을 관리하는 프로세스이다. 마지막으로 데이터 오너십 관리는 보안, 업무 규칙, 데이터 흐름, 표준, 데이터베이스의 모든 데이터에 대해 책임과 역할을 지정하고 평가하는 프로세스이다.

1.2.2 일관성 관련 품질관리 성숙수준

데이터 구조 관리(중복 관리 / 참조무결성 관리), 데이터 흐름 관리(흐름 대사 관리), 데이터 표준 관리(용어 표준 관리 / 도메인, 코드 표준 관리), 데이터 오너십 관리에 기초하여, 도입-정형화-통합화-정량화-최적화의 데이터 품질관리 성숙수준 단계에 따라 일관성 관련 데이터 품질관리 성숙수준을 정의하면 다음과 같다.

• 단계 1: 도입

데이터 중복에 대한 정의, 데이터 간의 관계 정의가 포함된 테이블 및 컬럼 등에 대한 문서가 존재하고 중복 컬럼과 정의된 데이터 간의 관계를 주기적으로 검증한다. 데이터베이스 상의 프로그램과 매핑되는 구조 간의 데이터 흐름에 대해 정의한 문서가 존재하며, 매핑 로직을 별도의 프로그램을 활용하여 검증하고 수행 결과 에러 리스트를 관리한다.

• 단계 2: 정형화

전사 용어, 도메인, 코드를 정의한 문서가 있으며, 중복 데이터에 대한 정의와 데이터 간의 관계 정의가 명시되어 있다. 또한 테이블 및 컬럼의 관계를 명확히 파악할 수 있는 정형화된 방법(ERD 등)이 존재하며 표준을 이용하여 구조에 대한 표준 준수 확인 절차를 수립하여 주기적으로 수행한다. Trigger나 프로그램을 이용해 데이터의 일치성을 보장한다. 데이터 흐름은 Tool이나 정형화된 방법에 의해 일괄 진행하며 매핑 로직에 대한 체크, Layout 변경 체크, 흐름 진행 시 작업상태 체크 등을 수행하고 흐름 모니터링 시 발생한 에러 통제 방법 및 절차가 마련되어 있다.

• 단계 3: 통합화

표준에 정의된 테이블과 프로그램에 대한 확인이 가능하고, 표준 변경 시에는 영향도 분석을 수행하며 비표준 내역을 조회할 수 있다. 중복에 대한 정의와 데이터 간의 관계가 명시되어 있고 테이블 및 컬럼의 관계를 명확히 파악할 수 있는 정형화된 방법(ERD 등)이 전사적으로 일관되게 존재하고 유지된다. 전사적 차원에서 표준에 대한 검증 결과 각종 오류에 대한 담당자를 추적하여 원인 분석을 수행하고 관련 담당자의 오류 개선 현황을 문서 또는 데이터베이스를 활용하여 체계적으로 관리한다. 전사적인 관점에서 데이터 검증, 오류 추적, 처리 상태가 정성적으로 관리된다.

• 단계 4: 정량화

표준 준수도 일치성 및 적용 범위에서 개인에 대해 정량적인 목표를 부여해 주기적으로 성과를 측정하고 오류결과를 지식화하며 필요에 따라 역할과 책임을 조정한다. 전사 관점에서 볼 때 정성적인 관리에서 정량적인 관리로 품질관리 수준이 발전되고 품질에 대한 성과가 명확히 측정되며 지식화 된다.

• 단계 5: 최적화

오류가 발생하지 않더라도 지식화된 오류 패턴을 분석하여 근본 원인을 제거하고 프로세스 개선 활동을 수행하며 지속적으로 기록·관리한다. 즉, 현재의 품질관리 최적화 뿐만 아니라 향후의 환경에 유연하게 대처할 수 있다.

1.2.3 일관성 관련 데이터 품질관리

1.2.2에서 정의한 일관성 데이터 품질관리 성숙수준을 충족하기 위해 반드시 수행되어야 할 품질관리는 다음과 같다.

[구조관리(중복 관리 / 참조무결성 관리)]

- 데이터 목록이 파악된다.
- 데이터 간의 관계가 정의된다.
- 중복 데이터가 파악된다.
- 데이터 구조를 통하여 일관성을 유지한다.
- 중복에 대한 처리가 명확하게 정의된다.
- 전사적 파악 및 확인이 이루어진다.
- 체계적인 오류 추적이 가능하다.
- 관리 지표가 존재한다.
- 오류 해결에 대한 지식이 축적된다.
- 오류의 근본 원인 제거 작업이 수행된다.

[데이터 흐름관리(흐름 대사 관리)]

- 데이터 흐름이 파악된다.
- 데이터 흐름이 명확하게 정의된다.
- 전사적 파악 및 확인이 이루어진다.
- 체계적인 오류 추적이 가능하다.
- 관리 지표가 존재한다.
- 오류 해결에 대한 지식이 축적된다.
- 오류의 근본 원인 제거 작업이 수행된다.

[데이터 표준관리(용어 표준 관리)]

- 표준화를 통하여 일관성을 유지한다.
- 표준 적용 항목의 추적이 가능하다.

[데이터 오너십 관리]

- 역할과 책임이 명확하게 정의된다.
- 활동에 대한 성과 관리를 수행한다.

상기의 일관성 관련 데이터 품질관리를 성숙 단계별로 정의하여 상세화하면 다음과 같다.

〈표 3-2〉 일관성 관련 데이터 품질관리 수준

성숙수준	품질관리 수준
1	데이터 목록이 파악된다. - 테이블 및 컬럼 정의를 파악할 수 있는 문서가 존재해야 한다.
	데이터 간의 관계가 정의된다. - 테이블 간 또는 컬럼 간의 관계를 명시하고 이의 정합성을 유지한다.
	중복 데이터가 파악된다. - 중복 데이터의 목록이 파악하며 중복된 데이터의 일치성을 유지한다.
	데이터 흐름이 파악된다. - 소스, 타깃 데이터 간의 흐름을 정의하고 흐름 간의 일치성을 유지한다.
2	표준화를 통하여 일관성을 유지한다. - 업무 용어, 도메인, 코드에 대한 표준을 정의하고 이를 준수하기 위한 절차가 존재한다.
	데이터 구조를 통하여 일관성을 유지한다. - 데이터 구조를 ERD 등의 표준화된 방법으로 명확하게 정의하고 정의된 구조에 따라 물리적 DB를 구현한다.
	중복에 대한 처리가 명확하게 정의된다. - 중복 데이터의 오너, 사용자 관계가 정의되고 중복된 데이터간의 일치성을 보장하기 위한 처리 절차가 존재한다.
	데이터흐름이 명확하게 정의된다. - 데이터 흐름을 정형화하거나 Tool을 활용하고 주기적인 모니터링과 오류 사항에 대한 조치 등의 명확한 통제가 가능하다.

성숙수준	품질관리 수준
3	<p>전사적 파악 및 확인이 이루어진다.</p> <ul style="list-style-type: none"> - 전사적으로 데이터 구조를 파악하고 전사적인 데이터의 흐름 수행이 확인된다.
	<p>역할과 책임이 명확하게 정의된다.</p> <ul style="list-style-type: none"> - 관리 대상의 정의 및 확인에 대한 역할을 정의하고 주어진 역할에 따라 관리 활동을 수행한다.
	<p>체계적인 오류 추적이 가능하다.</p> <ul style="list-style-type: none"> - 오류와 연관된 DB 또는 애플리케이션에 대한 원인 및 담당자 추적이 가능하고 오류에 대한 대응 활동을 추적하고 기록한다.
4	<p>표준 적용 항목의 추적이 가능하다.</p> <ul style="list-style-type: none"> - 표준과 연계된 DB, 애플리케이션 등의 항목의 확인이 가능하고 표준의 추가, 변경, 삭제 등의 관리 절차가 존재한다.
	<p>관리 지표가 존재한다.</p> <ul style="list-style-type: none"> - 정확성 측정 지표를 정의하고 주기적으로 측정한다.
	<p>활동에 대한 성과 관리를 수행한다.</p> <ul style="list-style-type: none"> - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
5	<p>오류 해결에 대한 지식이 축적된다.</p> <ul style="list-style-type: none"> - 오류 해결 내역을 체계적으로 분류하여 지식을 축적하고 오류 대응 시 이를 적극적으로 활용한다.
	<p>오류의 근본 원인 제거 작업이 수행된다.</p> <ul style="list-style-type: none"> - 오류의 근본 원인을 파악하여 이를 해결하기 위한 계획을 수립하고 관련 업무 프로세스를 개선한다.

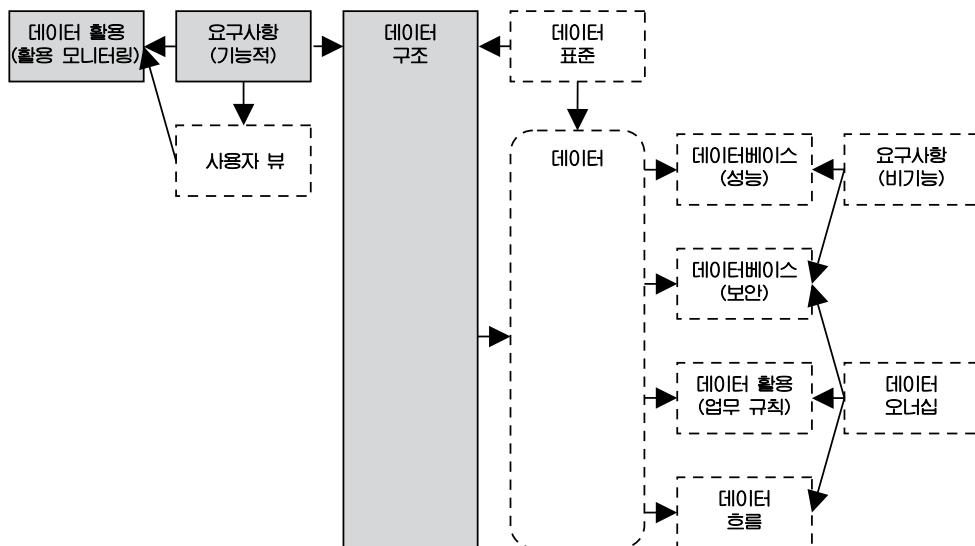
2. 활용성 관련 성숙수준

2.1 유용성 성숙수준

2.1.1 유용성 관련 프로세스

유용성은 제공 데이터가 조직 및 사용자의 다양한 요구사항을 충족시킬 수 있는지 여부를 의미하는 품질 기준이다. 즉, 요구사항에 적합한 충분한 데이터를 제공하는지, 데이터 구조가 유연하여 다양한 요구사항을 수용할 수 있는지, 사용 현황을 분석하여 관리하고 있는지 등이 포함된다.

데이터의 유용성을 확보하기 위해 관리 및 개선되어야 할 프로세스에는 요구사항 관리(기능적 관리), 데이터 구조 관리(유연성 관리), 데이터 활용 관리(활용 모니터링) 등이 있다.



〈그림 3-3〉 유용성 관련 데이터 품질관리 프로세스

요구사항 관리(기능적 관리)는 기능적 업무 요건, 업무 및 시스템 분석, 우선 순위 결정 등

을 관리하는 프로세스이며, 데이터 구조 관리(유연성 관리)는 신규 요건 및 요건 변경에 대응 할 수 있는 유연한 구조로 설계하기 위한 프로세스, 데이터 활용 관리(활용 모니터링)는 사용자의 활용 현황을 모니터링하여 자원 배분 및 사용자 만족도를 평가하기 위한 프로세스이다.

2.1.2 유용성 관련 품질관리 성숙수준

요구사항 관리(기능적 관리), 데이터 구조 관리(유연성 관리), 데이터 활용 관리(활용 모니터링)에 기초하여, 도입-정형화-통합화-정량화-최적화의 데이터 품질관리 성숙수준 단계에 따라 적시성 관련 데이터 품질관리 성숙수준을 정의하면 다음과 같다.

- 단계 1: 도입

요구사항이 정의되는 단계로 사용자 요구를 수집해 문서로 기록하고 요구에 대한 이해 당사자간의 확인이 이루어 진다. 또한, 요구의 대응 상황이 파악되므로 요구의 처리 현황이 기록되고, 요구 대응상의 문제점 및 한계를 파악하여 요구 당사자간의 협의가 이루어 진다.

- 단계 2: 정형화

요구의 대응 절차가 정형화되고 사용자의 반응이 반영되는 단계로 요구의 수집 및 확인, 요구에 대한 대응 절차가 수립되어 있으며 요구에 따른 변경 작업은 담당자의 승인을 얻어 실시한다. 또한 사용자를 위한 교육 및 가이드를 제공하며 데이터 활용 현황을 점검하여 사용자의 불만이나 개선 요구 사항을 파악하여 조치된다.

- 단계 3: 통합화

데이터 구조는 요구사항을 기반으로 설계되며 신규 요구사항의 발생이나 요구 변경에 따른 데이터 구조의 변경 영향도를 평가하고 있으며, 전사적인 데이터 구조 관리를 위한 조직이 존재한다. 전사적인 요구 협의를 위한 조직과 절차가 정의되어 있으며 요구 대응을 위한 자원 배분의 우선 순위가 정의 · 관리되고 요구사항과 프로젝트 계획 및 산출물 간의 양방향 추적 가능성이 유지된다. 즉 전사 차원의 데이터 구조 및 요구 관리가 수행되는 단계이다.

- 단계 4: 정량화

개인별 데이터 품질에 대한 정량적인 목표를 부여하고 측정 지표를 적용하여 주기적으로

성과를 측정한다. 측정 결과에 기초하여 기능적 요건에 대한 대응 지식을 체계적으로 분류하여 지식화하고 필요에 따라 관련 담당자의 역할과 책임을 조정하는 단계이다. 전사 관점에서 정성적인 관리에서 정량적인 관리로 전이되고 품질에 대한 성과가 명확하게 측정되고 지식화되는 단계이다.

- 단계 5: 최적화

업계의 Best Practice를 반영하여 지속적으로 유지·발전시키고 데이터 활용 기술 및 동향을 파악·적용하며 한정된 자원을 최적화하여 데이터 활용도를 극대화한다.

2.1.3 유용성 관련 데이터 품질관리

2.1.2에서 정의한 유연성 데이터 품질관리 성숙수준을 충족하기 위해 반드시 수행되어야 할 품질관리는 다음과 같다.

[요구사항 관리(기능적 관리)]

- 요건을 정의한다.
- 요건 대응 상황이 파악된다.
- 요건 대응 절차가 정형화된다.
- 요건이 전사적으로 조율 된다.
- 요건 대응을 위해 자원이 배분된다.
- 관리 지표가 존재한다.
- 활동에 대한 성과 관리를 수행한다.
- 요건 대응 지식이 축적된다.
- 요건의 내부 동향이 파악된다.
- 요건의 외부 동향이 파악된다.
- 데이터 활용이 최적화된다.

[데이터 구조 관리(유연성 관리)]

- 요건 기반의 데이터 구조 관리가 이루어진다.

[데이터 활용 관리(활용 모니터링)]

- 사용자의 반응이 반영된다.

상기의 유연성 관련 데이터 품질관리를 성숙 단계별로 정의하여 상세화하면 다음과 같다.

〈표 3-3〉 유연성 관련 데이터 품질관리 수준

성숙수준	품질관리 수준
1	<p>요건을 정의한다.</p> <ul style="list-style-type: none"> - 사용자의 요건을 수집하고 확인하여 문서로 정의한다.
	<p>요건 대응 상황이 파악된다.</p> <ul style="list-style-type: none"> - 요건의 처리 현황을 기록하고 요건 대응 시의 문제점을 파악한다.
2	<p>요건 대응 절차가 정형화된다.</p> <ul style="list-style-type: none"> - 요건의 수집 및 확인 절차가 규정되고 표준화된 구현 계획에 의하여 요건 대응이 수행되며 요건 변경에 대한 대응 절차가 수립된다.
	<p>사용자의 반응이 반영된다.</p> <ul style="list-style-type: none"> - 사용자를 위한 교육 및 가이드를 제공하고 사용자의 활용 현황을 주기적으로 점검하여 불만 및 개선 사항을 반영한다.
3	<p>요건 기반의 데이터 구조 관리가 이루어진다.</p> <ul style="list-style-type: none"> - 요건을 기반으로 데이터 구조를 설계하고 요건의 변경에 따른 영향도를 분석하여 데이터 구조 설계에 반영하는 절차가 존재한다.
	<p>요건이 전사적으로 조율된다.</p> <ul style="list-style-type: none"> - 전사적인 요건을 조율하는 공식화된 기능이 존재하고 요건을 통합하여 일관성 있게 처리한다.
	<p>요건 대응을 위해 자원이 배분된다.</p> <ul style="list-style-type: none"> - 요건에 대한 우선 순위를 관리하며 요건의 신규, 변경에 대한 영향도가 평가된다.
4	<p>관리 지표가 존재한다.</p> <ul style="list-style-type: none"> - 유연성 측정 지표를 정의하고 주기적으로 측정한다.
	<p>활동에 대한 성과 관리를 수행한다.</p> <ul style="list-style-type: none"> - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
	<p>요건 대응 지식이 축적된다.</p> <ul style="list-style-type: none"> - 요건 대응 내역을 체계적으로 분류하여 축적하여 추후 요건 대응에 필요한 기간 및 자원이 예측 가능하다.
	<p>요건의 내부 동향이 파악된다.</p> <ul style="list-style-type: none"> - 요건의 변경 내역이 분류되고 내부 요건의 변경 성향을 파악하여 요건 대응에 활용 한다.

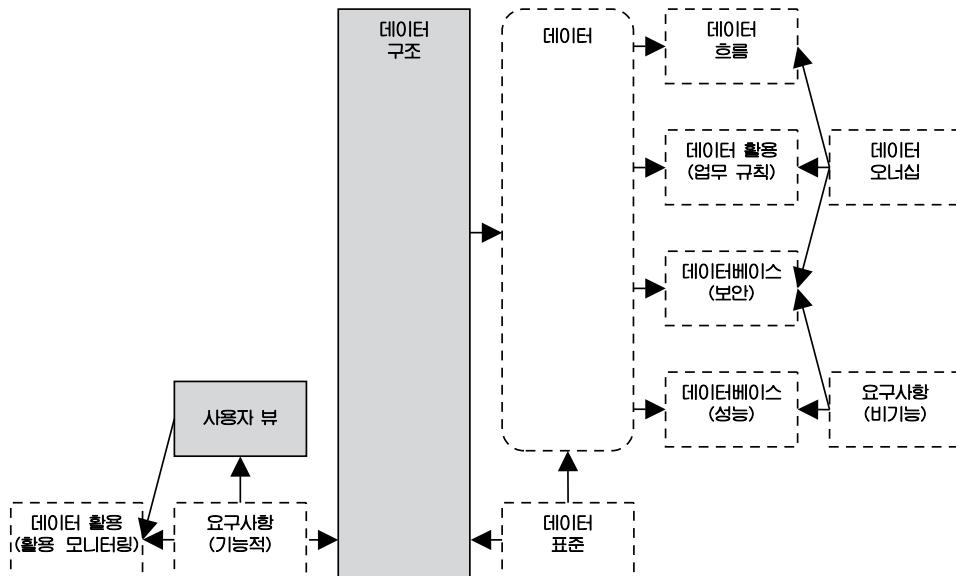
성숙수준	품질관리 수준
5	<p>요건의 외부 동향이 파악된다.</p> <ul style="list-style-type: none"> - 업계의 Best Practice를 반영하여 지속적으로 유지, 발전시키고 데이터 활용 기술 및 동향을 파악하여 적용한다.
	<p>데이터 활용이 최적화된다.</p> <ul style="list-style-type: none"> - 한정된 자원을 최적화하여 데이터 활용도를 극대화한다.

2.2 접근성 성숙수준

2.2.1 접근성 관련 프로세스

데이터에 대한 사용자 만족도를 충족시키기 위해서는 데이터 자체의 품질은 물론, 데이터를 효율적·효과적으로 제공하는데 필요한 인터페이스 등 제반 서비스 품질을 확보해야 한다. 접근성은 사용자가 원하는 데이터를 손쉽게 이용할 수 있음을 의미하는 품질 기준이다.

데이터의 접근성을 확보하기 위해 관리 및 개선되어야 할 프로세스에는 사용자 뷰 관리, 데이터 구조 관리(통합 관리) 등이 있다.



〈그림 3-4〉 접근성 관련 데이터 품질관리 프로세스

사용자 뷰 관리는 사용자와 데이터간의 인터페이스를 관리하는 프로세스이며, 데이터 구조 관리(통합 관리)는 여러 시스템에 분산된 데이터를 일관된 관점의 데이터 구조로 통합·관리하는 프로세스이다.

2.2.2 접근성 관련 품질관리 성숙수준

사용자 뷰 관리, 데이터 구조 관리(통합 관리)에 기초하여, 도입–정형화–통합화–정량화–최적화의 데이터 품질관리 성숙수준 단계에 따라 접근성 관련 데이터 품질관리 성숙수준을 정의하면 다음과 같다.

- 단계 1: 도입

메뉴 구조 및 사이트 맵이 제공되며 사용자 뷰에 대한 참조 문서가 존재는 등 기본적인 사용자 뷰가 제공된다. 또한 분산 데이터에 대한 조회가 가능하며, 분산 데이터의 통합 및 변환 기준이 마련되어 있고 이를 조회할 수 있는 사용자 뷰가 제공된다.

- 단계 2: 정형화

표준화된 사용자 뷰를 제공하며 사용자의 데이터 활용 능력 향상을 위해 주기적인 교육을 시행하고 사용자 매뉴얼은 최신의 상태로 유지된다. 분산된 데이터는 주제 지향적이고 통합된 형태의 데이터베이스로 제공되므로 사용자가 손쉽게 접근하여 직접 데이터를 활용할 수 있는 환경이 제공된다.

- 단계 3: 통합화

여러 시스템에 분산된 데이터가 통합된 사용자 뷰로 제공되고 업무에 필요한 데이터 접근 채널이 보다 다양해진다. 분산 데이터도 전사 통합된 관점에서 분석이 가능하며 전산 조직에서 지원되는 기본 데이터 외에 사용자의 고급 분석 요구에 대응할 수 있는 환경이 제공되고 데이터는 충분히 축적되어 부가 가치를 창출할 수 있을 정도로 성숙되어 있다. 기반이 되는 데이터 구조 또한 메타 정보가 사용자에게 제공되어 공유된다.

- 단계 4: 정량화

개인별 데이터 품질에 대한 정량적인 목표를 부여하고 측정 지표를 적용하여 주기적으로 성과를 측정한다. 측정 결과에 따라 접근성을 개선하고 사용자 활용 패턴을 분석하며 필

요에 따라 관련 담당자의 역할과 책임을 조정하는 단계이다. 전사 관점에서 정성적인 관리에서 정량적인 관리로 전이되고 품질에 대한 성과가 명확히 측정되고 지식화된다.

• 단계 5: 최적화

사용자 뷰를 최적화하기 위해 최신 기술과 활용 현황을 접목하여 지속적인 프로세스 개선 활동을 수행하는 단계로 현재 상태의 최적화뿐만 아니라 향후 환경에도 유연하게 대처할 수 있다.

2.2.3 접근성 관련 데이터 품질관리

2.2.2에서 정의한 접근성 데이터 품질관리 성숙수준을 충족하기 위해 반드시 수행되어야 할 품질관리는 다음과 같다.

[사용자 뷰 관리]

- 사용자 뷰 구조가 제공된다.
- 사용자 뷰를 표준화하여 제공한다.
- 사용자의 데이터 활용을 위한 매뉴얼을 지원한다.
- 사용자가 직접 데이터를 활용한다.
- 사용자 뷰가 통합된 환경에서 제공된다.
- 가치 창출을 위한 환경이 제공된다.
- 관리 지표가 존재한다.
- 활동에 대한 성과관리를 수행한다.
- 사용자 활용 패턴이 파악되고 지원된다.
- 사용자 뷰를 최적화한다.

[데이터 구조 관리(통합 관리)]

- 분산 데이터가 조회 가능하다.
- 분산 데이터를 통합한 구조가 존재한다.
- 전사 데이터 구조가 공유된다.
- 분산 데이터를 전사적으로 통합한다.

상기의 접근성 관련 데이터 품질관리를 성숙 단계별로 정의하여 상세화하면 다음과 같다.

〈표 3-4〉 접근성 관련 데이터 품질관리 수준

성숙수준	품질관리 수준
1	<p>사용자 뷰 구조가 제공된다.</p> <ul style="list-style-type: none"> - 메뉴 구조 및 사이트 맵이 제공되며 사용자 뷰에 대한 참조 문서가 제공된다.
	<p>분산 데이터가 조회 가능하다.</p> <ul style="list-style-type: none"> - 분산된 데이터를 조회할 수 있는 사용자 뷰가 제공되며 분산 데이터에 대한 통합, 변환 기준이 존재한다.
2	<p>사용자 뷰를 표준화하여 제공한다.</p> <ul style="list-style-type: none"> - 사용자 뷰의 구성 요소를 표준화하며 다양한 사용자 지원 기능을 제공한다.
	<p>분산 데이터를 통합한 구조가 존재한다.</p> <ul style="list-style-type: none"> - 분산 데이터가 업무 기능별로 통합되어 존재한다.
	<p>사용자의 데이터 활용을 위한 매뉴얼을 지원한다.</p> <ul style="list-style-type: none"> - 사용자 교육이 주기적으로 시행되며 사용자 매뉴얼은 최신의 상태로 유지된다.
	<p>사용자가 직접 데이터를 활용한다.</p> <ul style="list-style-type: none"> - 사용자의 직접적인 데이터 활용을 위하여 데이터 구조가 공유되고 활용 Tool 등이 제공된다.
3	<p>사용자 뷰가 통합된 환경에서 제공된다.</p> <ul style="list-style-type: none"> - 사용자 뷰를 다양한 매체를 통하여 제공하며 각각의 단위 시스템들은 통합된 사용자 뷰를 제공한다.
	<p>전사 데이터 구조가 공유된다.</p> <ul style="list-style-type: none"> - 데이터 구조가 전사적으로 관리되고 사용자에게 공유할 데이터 구조가 정의되어 제공된다.
	<p>분산 데이터를 전사적으로 통합한다.</p> <ul style="list-style-type: none"> - 전사적인 데이터 통합 구조가 존재하며 이를 활용한 사용자 뷰가 제공된다.
	<p>가치 창출을 위한 환경이 제공된다.</p> <ul style="list-style-type: none"> - 사용자의 가치 창출이 가능한 고급의 데이터 분석을 지원하기 위하여 데이터를 축적하고 환경을 제공한다.
4	<p>관리 지표가 존재한다.</p> <ul style="list-style-type: none"> - 접근성 측정 지표를 정의하고 주기적으로 측정한다.
	<p>활동에 대한 성과관리를 수행한다.</p> <ul style="list-style-type: none"> - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
	<p>사용자 활용 패턴이 파악되고 지원된다.</p> <ul style="list-style-type: none"> - 축적된 데이터를 바탕으로 사용자 활용 패턴이 파악되고 이를 근거로 사용자 뷰를 재구성하여 개인화된 뷰를 제공한다.

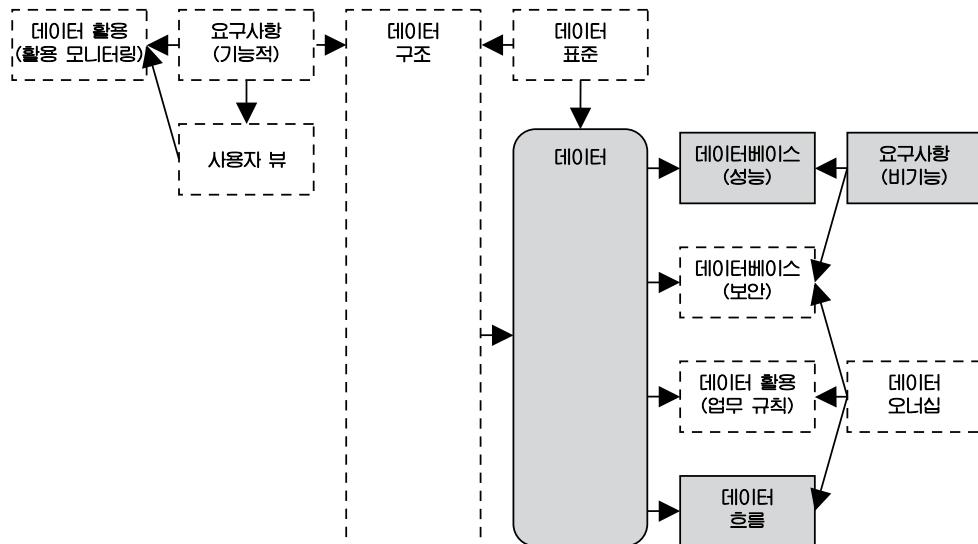
성숙수준	품질관리 수준
5	<p>사용자 뷰를 최적화한다.</p> <ul style="list-style-type: none"> - 사용자 뷰에 대한 최신 기술과 활용 현황을 접목하여 사용자 뷰를 지속적으로 최적화한다.

2.3 적시성 성숙수준

2.3.1 적시성 관련 프로세스

적시성은 비기능적 요구사항(응답 시간, 실행 속도 등)과 데이터의 최신성과 같은 요건에 얼마나 적절히 대처하고 있는지를 의미하는 품질 기준이다.

데이터의 적시성을 확보하기 위해 관리 및 개선되어야 할 프로세스에는 요구사항 관리(비기능적 관리), 데이터 흐름 관리(흐름 주기 관리), 데이터베이스 관리(성능 관리) 등의 프로세스가 있다.



〈그림 3-5〉 적시성 관련 데이터 품질관리 프로세스

요구사항 관리(비기능적 관리) 프로세스는 업무의 내용적 요건 외에 처리시간, 응답시간, 데이터 제공 주기 등과 같은 비기능적인 요구사항을 관리하는 프로세스이다. 데이터 흐름 관리(흐름 주기 관리) 프로세스는 소스에서 타깃까지의 데이터 흐름에 대해서 수행 시기, 수행 주기 등을 관리하는 프로세스이다. 데이터베이스 관리(성능 관리) 프로세스는 데이터 처리 작업을 모니터링하고 성능의 최적화를 위한 튜닝 등의 작업을 수행하는 프로세스이다.

2.3.2 적시성 관련 품질관리 성숙수준

요구사항 관리(비기능적 관리), 데이터 흐름 관리(흐름 주기 관리), 데이터베이스 관리(성능 관리) 데이터베이스 관리(보안 관리)에 기초하여, 도입–정형화–통합화–정량화–최적화의 데이터 품질관리 성숙수준 단계에 따라 적시성 관련 데이터 품질관리 성숙수준을 정의하면 다음과 같다.

- 단계 1: 도입

비기능적인 요구사항이 정의되는 단계로 사용자 요건이 수집되어 문서로 기록되며 요건에 대한 이해 당사자 간의 확인 과정이 이루어진다. 또한 요건의 대응 상황이 파악되므로 요건의 처리 현황이 기록되며, 요건 대응상의 문제점 및 한계를 파악하여 요건 당사자 간의 협의가 이루어진다. 기본적인 데이터 흐름에 대한 스케줄 관리가 수행되어 이행 결과가 주기적으로 파악되고 조치된다.

- 단계 2: 정형화

요건의 대응 절차가 정형화되고 데이터의 흐름 주기가 조율되며 적시성 관리를 위한 조직의 역할과 책임이 정의되어 있다. 요건 수집 및 확인 절차가 수립되어 있으며 일정 계획을 수립해서 요건에 대한 대응 작업을 수행하고, 주기적으로 운영 시스템의 튜닝 작업 등을 수행한다. 데이터 흐름의 주기는 명확히 정의되어 데이터의 분석 주기, 마감 기한, 요약 수준 등의 항목이 문서화되어 관리되며 흐름의 선후 작업이 조율 된다.

- 단계 3: 통합화

요건 대응을 위한 자원이 배분되고 우선 순위에 따라 자원 배분 계획이 수립되어 시행된다. 전사적으로 데이터 흐름이 관리되고 역할과 책임이 명확하게 정의되어 비기능적 요건에 대한 전문적인 대응이 이루어지는 단계이다.

• 단계 4: 정량화

개인별 데이터 품질에 대한 정량적인 목표를 부여하고 측정 지표를 적용해 주기적으로 성과를 측정한다. 측정 결과에 따라 비기능적 요건 대응 지식을 체계적으로 분류하여 지식화하고 필요에 따라 그 역할과 책임을 조정한다. 전사 관점에서 정성적인 관리에서 정량적인 관리로 전이되고 품질에 대한 성과가 명확히 측정되며 지식화되고 있다.

• 단계 5: 최적화

지속적으로 개선 요소를 발굴하고 튜닝 및 자원을 재분배하여 성능을 최적화하는 한편 데이터 흐름에 있어서는 불필요한 흐름의 원인을 제거하며 성능 향상을 위한 기술 동향을 파악하고 적용한다.

2.3.3 적시성 관련 데이터 품질관리

2.3.2에서 정의한 적시성 데이터 품질관리 성숙수준을 충족하기 위해 반드시 수행되어야 할 품질관리는 다음과 같다.

[요구사항 관리(비기능적 관리)]

- 요건을 정의한다.
- 요건 대응 절차가 정형화된다.
- 요건 대응을 위해 자원이 배분된다.
- 관리 지표가 존재한다.
- 활동에 대한 성과 관리를 수행한다.
- 요건 대응 지식이 축적된다.
- 성능을 최적화한다.

[데이터 흐름 관리(흐름 주기 관리)]

- 스케줄에 대한 관리가 수행된다.
- 데이터 흐름 주기를 명확하게 정의한다.
- 전사적인 데이터 흐름 주기를 정의한다.
- 관리 지표가 존재한다.
- 활동에 대한 성과 관리를 수행한다.
- 데이터 흐름을 최적화한다.

[데이터베이스 관리(성능 관리)]

- 요건 대응 절차가 정형화된다.
- 역할과 책임이 명확하게 정의된다.
- 요건 대응에 전문가를 활용한다.

상기의 적시성 관련 데이터 품질관리를 성숙 단계별로 정의하여 상세화하면 다음과 같다.

〈표 3-5〉 적시성 관련 데이터 품질관리 수준

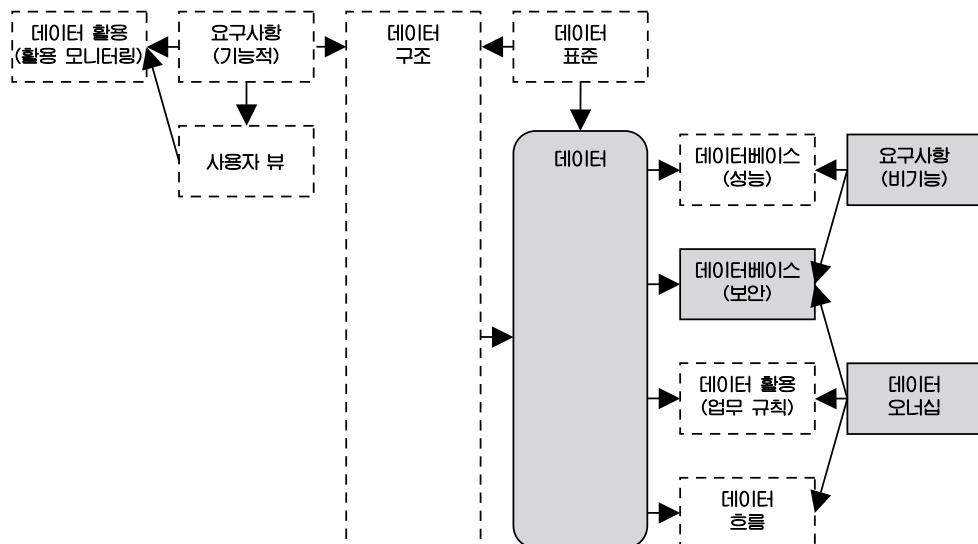
성숙수준	품질관리 수준
1	요건을 정의한다. - 적시성에 대한 사용자 요구사항을 파악되어 대응된다.
	스케줄에 대한 관리가 수행된다. - 데이터 이행에 대한 사용자 요구사항을 파악하고 스케줄 실행 결과가 주기적으로 파악된다.
2	요건 대응 절차가 정형화된다. - 응답 속도, 처리 속도, 데이터 생성 주기 등의 적시성의 기준을 수립하고 주기적으로 모니터링하고 튜닝한다.
	데이터 흐름 주기를 명확하게 정의한다. - 데이터의 생성 주기를 파악하고 데이터 흐름의 선후를 조율한다.
3	역할과 책임이 명확하게 정의된다. - 적시성 관리를 위한 역할을 정의하고 주어진 역할에 따라 관리 활동을 수행한다.
	전사적인 데이터 흐름 주기를 정의한다. - 전사적인 데이터 흐름의 선후 관계가 조율되고 우선 순위에 따라 지원이 배분된다.
	요건 대응에 전문가를 활용한다. - 전문가의 성능 진단을 필요 시 또는 주기적으로 수행하며 성능 진단의 결과가 성능 개선에 반영되어 활용된다.
	요건 대응을 위해 지원이 배분된다.
4	관리 지표가 존재한다. - 적시성 측정 지표를 정의하고 주기적으로 측정한다.
	활동에 대한 성과관리를 수행한다. - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
	요건 대응 지식이 축적된다. - 요건 대응 내역을 체계적으로 분류하고 축적하여 추후 요건 대응에 필요한 기간 및 지원이 예측 가능하다.

성숙수준	품질관리 수준
5	성능을 최적화한다. - 개선 요소를 발굴하고 지속적인 튜닝 및 지원을 재분배하여 성능을 향상을 위한 기술 동향을 파악하고 적용한다.
	데이터 흐름을 최적화한다. - 불필요한 흐름의 원인을 제거하고 최적화를 위한 최신 기술을 접목하여 활용한다.

2.4 보안성 성숙수준

2.4.1 보안성 관련 프로세스

보안성은 내·외부의 불법적이고 비인가된 위협으로부터 데이터의 기밀성, 무결성 및 가용성을 확보할 수 있음을 의미하는 품질 기준으로, 데이터베이스 및 적재된 데이터를 허가되지 않은 일련의 행위로부터 안전하게 보호하기 위한 인적, 정책적, 물리적, 논리적 대응체계가 포함된다.



〈그림 3-6〉 보안성 관련 데이터 품질관리 프로세스

데이터의 보안성을 확보하기 위해 관리 및 개선되어야 할 프로세스에는 데이터베이스 관리(보안 관리)가 있다. 데이터베이스 관리(보안 관리)는 데이터베이스에 저장된 데이터를 오류 및 훼손 없이 안정적으로 서비스 할 수 있도록 데이터에 대한 생성, 변경, 백업 관리를 지속적으로 수행하는 프로세스이다.

2.4.2 보안성 관련 품질관리 성숙수준

데이터베이스 관리(보안 관리)에 기초하여 도입-정형화-통합화-정량화-최적화의 데이터 품질관리 성숙수준 단계에 따라 보안성 관련 데이터 품질관리 성숙수준을 정의하면 다음과 같다.

- 단계 1: 도입

보안 대상 및 기본적인 보안 수준이 정의되는 단계로, 보안 대상에 대한 구분과 가치 평가가 수행된다. 시스템 단위의 접근 통제를 기반으로 기본적인 수준의 식별 및 인증이 수행되고 연속성을 보장하기 위한 백업 계획이 수립된다.

- 단계 2: 정형화

보안성을 충족시키기 위한 조직 구성과 업무 분담, 지침 및 절차가 수립되는 단계로 보안 담당 조직 및 지침에 기초한 정형화된 보안 관리가 이루어진다. 접근 통제, 백업, 복구에 대한 모든 절차가 수립되어 있으며, 보안 담당 조직원의 책임과 역할이 정의된다. 침해 사고 발생 시의 대응 절차 등 예상되는 모든 가능성에 대한 절차가 정의되어 있다.

- 단계 3: 통합화

전사적인 연계와 통합의 관점하에 일관성있는 데이터의 보안 관리가 이루어지며 모든 사용 내역에 대한 로깅 및 분석이 수행된다. 로깅 및 감사에 의해 위협 요소가 분석되고 판단되며, 보안 정책에 반영된다. 주기적으로 취약점에 대한 진단이 수행되고 위협 요소가 제거되며, 모든 행위에 대한 로깅 및 감사가 이루어 진다.

- 단계 4: 정량화

데이터 보안 관리가 심화되는 단계이다. 각 업무 담당자의 권한을 서로 견제할 수 있는 구조를 가지며, 인가된 관리자에 대한 적절한 통제가 수행된다. 인증서를 포함, 복합적인 식

별 및 인증이 수행되고 내부자에 의한 정보 유출 방지를 위한 방안이 수립되어 있다. 관리자의 권한 분담 및 중요 데이터에 대한 암호화가 적용되고 있다.

• 단계 5: 최적화

데이터 보안 관리가 완성되는 단계로 현재의 관점에서 최적화뿐만 아니라 미래의 환경 변화에 유연하게 대처할 수 있다. 수행되고 있는 보안 수준을 평가하여 조직의 데이터 보호 목적에 부합되는지 여부를 검증하여 보완하고, 계층적인 위협 관리 체계를 구축하여 예기치 못한 위협에 능동적으로 대처할 수 있는 방어 체계를 완성한다.

2.4.3 보안성 관련 데이터 품질관리

2.4.2에서 정의한 보안성 데이터 품질관리 성숙수준을 층족하기 위해 반드시 수행되어야 할 품질관리는 다음과 같다.

- 기본 수준의 식별 및 인증을 수행한다.
- 데이터에 대한 백업 계획을 수립한다.
- 보안 통제가 존재한다.
- 데이터 보안에 관련된 조직을 선정하고 업무 분장을 수행한다.
- 데이터 보안에 관련된 지침 및 절차를 수립한다.
- 침해 사고에 대한 복구 계획을 수립한다.
- 보안 감사를 수행한다.
- 취약점에 대한 분석 및 조치를 수행한다.
- 고급 수준의 식별 및 인증을 수행한다.
- 인가된 관리자에 의한 정보 유출을 방지한다.
- 계층적인 위협 관리 체계를 구축한다
- 보안 수준에 대한 평가를 수행한다.

상기의 보안성 관련 데이터 품질관리를 성숙 단계별로 정의하여 상세화하면 다음과 같다.

〈표 3-6〉 보안성 관련 데이터 품질관리 수준

성숙수준	품질관리 수준
1	<p>기본 수준의 식별 및 인증을 수행한다.</p> <ul style="list-style-type: none"> - ID, Password에 대한 설정 규칙, 애플리케이션이나 사용자에 대한 접근 권한, 인증의 내역 기록 등 기본적인 보안 관리를 수행한다.
	<p>데이터에 대한 백업계획을 수립한다.</p> <ul style="list-style-type: none"> - 백업 대상을 식별하고 보안 등급을 적용하며 백업 데이터의 이전, 보관에 대한 절차를 정의한다.
	<p>보안 통제가 존재한다.</p> <ul style="list-style-type: none"> - 보호되어야 할 자산의 파악 및 평가를 수행하고 접근 통제를 수행한다.
2	<p>데이터 보안에 관련된 조직을 선정하고 업무분장을 수행한다.</p> <ul style="list-style-type: none"> - 데이터 보안 관리자를 선정하고 업무 분장을 수행하며 주기적인 교육을 시행한다.
	<p>데이터 보안에 관련된 지침 및 절차를 수립한다.</p> <ul style="list-style-type: none"> - 정보 보호 지침이나 절차를 구비하며 ID별, 테이블별, 컬럼별 접근 권한에 대한 이력 및 백업 내역을 관리하고 변경 내역을 반영한다.
	<p>침해 사고에 대한 복구 계획을 수립한다.</p> <ul style="list-style-type: none"> - 데이터 복구 계획을 침해 사고의 유형에 따라 세부적으로 분류하여 수립한다.
3	<p>보안 감시를 수행한다.</p> <ul style="list-style-type: none"> - 보안 감시를 주기적으로 수행하여 보안 정책에 반영하고 감사 내역에 대한 백업 정책 및 접근 권한을 부여하여 관리한다.
	<p>취약점에 대한 분석 및 조치를 수행한다.</p> <ul style="list-style-type: none"> - 주기적으로 데이터 보안에 대한 취약점을 분석하고 보완하며 예측 가능한 취약점에 대하여 사전 조치한다.
4	<p>고급 수준의 식별 및 인증을 수행한다.</p> <ul style="list-style-type: none"> - 인증서 기반의 인증 및 최소 2가지 이상의 식별 요소를 부여하여 사용자 식별을 수행한다.
	<p>인가된 관리자에 의한 정보 유출을 방지한다.</p> <ul style="list-style-type: none"> - 시스템 관리자, 데이터베이스 관리자, 보안 관리자의 역할을 분담하고 이중화된 유출 차단 정책 및 중요 데이터를 선정하여 적절한 수준의 암호화를 수행한다.
5	<p>계층적인 위협 관리 체계를 구축한다.</p> <ul style="list-style-type: none"> - 데이터베이스 사용 내역을 주기적으로 분석하여 이상 징후에 대비하여, 직접적인 데이터베이스 접근 경로를 차단해 이상 징후 발견 시 자동으로 모든 접속을 차단할 수 있는 방안을 마련한다.
	<p>보안수준에 대한 평가를 수행한다.</p> <ul style="list-style-type: none"> - 주기적으로 데이터 보안에 관련된 지침 및 절차에 대해 평가를 수행하여 일련의 보안 행위들이 조직의 데이터 보호 목적에 부합하는지 검증한다.

부록 : 성숙수준별 품질관리 점검표

성숙 수준	품질 기준	품질관리 수준
1	정확성	관리 대상이 정의된다. - 정확성에 대한 업무 규칙을 정의하며 이를 최신의 상태로 유지한다.
		정확성에 대한 확인이 수행된다. - 정확성을 확인하기 위한 방법이 존재하고 확인 결과를 문서로 기록한다.
	일관성	데이터 목록이 파악된다 - 테이블 및 컬럼 정의를 파악할 수 있는 문서가 존재해야 한다.
		데이터 간의 관계가 정의된다. - 테이블 간 또는 컬럼 간의 관계를 명시하고 이의 정합성을 유지한다
		중복 데이터가 파악된다. - 중복 데이터의 목록이 파악하며 중복된 데이터의 일치성을 유지한다.
		데이터 흐름이 파악된다. - 소스. 타깃 데이터 간의 흐름을 정의하고 흐름 간의 일치성을 유지한다.
	유동성	요건을 정의한다. - 사용자의 요건을 수집하고 확인하여 문서로 정의한다.
		요건 대응 상황이 파악된다. - 요건의 처리 현황을 기록하고 요건 대응 시의 문제점을 파악한다.
	접근성	사용자 뷰 구조가 제공된다. - 메뉴 구조 및 사이트 맵이 제공되며 사용자 뷰에 대한 참조 문서가 제공된다.
		분산 데이터가 조회 가능하다. - 분산된 데이터를 조회할 수 있는 사용자 뷰가 제공되며 분산 데이터에 대한 통합, 변화 기준이 존재한다.
	적시성	요건을 정의한다. - 적시성에 대한 사용자 요구사항을 파악되어 대응된다.
		스케줄에 대한 관리가 수행된다. - 데이터 이행에 대한 사용자 요구사항을 파악하고 스케줄 실행 결과가 주기적으로 파악된다.
	보안성	기본 수준의 식별 및 인증을 수행한다. - ID, Password에 대한 설정 규칙, 애플리케이션이나 사용자에 대한 접근 권한, 인증의 내역 기록 등 기본적인 보안 관리를 수행한다.
		데이터에 대한 백업 계획을 수립한다. - 백업 대상을 식별하고 보안 등급을 적용하며 백업 데이터의 이전, 보관에 대한 절차를 정의한다.

성숙 수준	품질 기준	품질관리 수준
1	보안성	<p>보안 통제가 존재한다.</p> <ul style="list-style-type: none"> - 보호되어야 할 자산의 파악 및 평가를 수행하고 접근 통제를 수행한다.
정확성		<p>정확성을 위한 표준을 정의한다.</p> <ul style="list-style-type: none"> - 도메인, 코드 표준을 정의하고 유지하며 표준의 준수를 위한 절차를 정립하고 적용 한다.
		<p>관리 대상이 체계적으로 정의된다.</p> <ul style="list-style-type: none"> - 컬럼의 필수 여부, 유효 범위, 컬럼 간 관계 등의 정확성 기준을 식별하고 대상에 대한 기록 및 유지를 정형화 한다.
		<p>정확성의 확인이 계획적으로 이루어진다.</p> <ul style="list-style-type: none"> - 확인 방법이 정형화되거나 Tool을 활용하여 주기적으로 시행한다.
2	일관성	<p>표준화를 통하여 일관성을 유지한다.</p> <ul style="list-style-type: none"> - 업무 용어, 도메인, 코드에 대한 표준을 정의하고 이를 준수하기 위한 절차가 존재 한다.
		<p>데이터 구조를 통하여 일관성을 유지한다.</p> <ul style="list-style-type: none"> - 데이터 구조를 ERD 등의 표준화된 방법으로 명확하게 정의하고 정의된 구조에 따라 물리적 DB를 구현한다.
		<p>종복에 대한 처리가 명확하게 정의된다.</p> <ul style="list-style-type: none"> - 종복 데이터의 오너, 사용자 관계가 정의되고 종복된 데이터 간의 일치성을 보장하기 위한 처리 절차가 존재한다.
		<p>데이터 흐름이 명확하게 정의된다.</p> <ul style="list-style-type: none"> - 데이터 흐름을 정형화하거나 Tool을 활용하고 주기적인 모니터링과 오류 사항에 대한 조치 등의 명확한 통제가 가능하다.
유용성		<p>요건 대응 절차가 정형화된다.</p> <ul style="list-style-type: none"> - 요건의 수집 및 확인 절차가 규정되고 표준화된 구현 계획에 의하여 요건 대응이 수행되며 요건 변경에 대한 대응 절차가 수립된다.
		<p>사용자의 반응이 반영된다.</p> <ul style="list-style-type: none"> - 사용자를 위한 교육 및 가이드를 제공하고 사용자의 활용 현황을 주기적으로 점검하여 불만 및 개선 사항을 반영한다.
접근성		<p>사용자 뷰를 표준화하여 제공한다.</p> <ul style="list-style-type: none"> - 사용자 뷰의 구성 요소를 표준화하여 다양한 사용자 지원 기능을 제공한다.
		<p>분산 데이터를 통합한 구조가 존재한다.</p> <ul style="list-style-type: none"> - 분산 데이터가 업무 기능별로 통합되어 존재한다.
		<p>사용자의 데이터 활용을 위한 매뉴얼을 지원한다.</p> <ul style="list-style-type: none"> - 사용자 교육이 주기적으로 시행되며 사용자 매뉴얼은 최신의 상태로 유지된다.

성숙 수준	품질 기준	품질관리 수준
2	접근성	사용자가 직접 데이터를 활용한다. - 사용자의 직접적인 데이터 활용을 위하여 데이터 구조가 공유되고 활용 Tool 등이 제공된다.
		요건 대응 절차가 정형화된다. - 응답 속도, 처리 속도, 데이터 생성 주기 등의 적시성의 기준을 수립하고 주기적으로 모니터링하고 튜닝한다.
	적시성	데이터 흐름 주기를 명확하게 정의한다. - 데이터의 생성 주기를 파악하고 데이터 흐름의 선후를 조율한다.
		데이터 보안에 관련된 조직을 설정하고 업무 분장을 수행한다. - 데이터 보안 관리자를 설정하고 업무 분장을 수행하며 주기적인 교육을 시행한다.
	보안성	데이터 보안에 관련된 지침 및 절차를 수립한다. - 정보 보호 지침이나 절차를 구비하여 ID별, 테이블별, 컬럼별 접근 권한에 대한 이력 및 백업 내역을 관리하고 변경 내역을 반영한다.
		침해 사고에 대한 복구 계획을 수립한다. - 데이터 복구 계획을 침해 사고의 유형에 따라 세부적으로 분류하여 수립한다.
3	정확성	역할과 책임이 명확하게 정의된다. - 관리 대상에 대하여 오너를 부여하고 오너의 역할과 책임에 대하여 규정하며 관리 활동을 수행한다.
		체계적인 오류 추적이 가능하다. - 오류와 연관된 DB 또는 애플리케이션에 대한 원인 및 담당자 추적이 가능하고 오류에 대한 대응 활동을 추적하고 기록한다.
	일관성	전사적 파악 및 확인이 이루어진다. - 전사적으로 데이터 구조를 파악하고 전사적인 데이터의 흐름 수행이 확인된다.
		역할과 책임이 명확하게 정의된다. - 관리 대상의 정의 및 확인에 대한 역할을 정의하고 주어지 역할에 따라 관리 활동을 수행한다.
		체계적인 오류 추적이 가능하다. - 오류와 연관된 DB 또는 애플리케이션에 대한 원인 및 담당자 추적이 가능하고 오류에 대한 대응 활동을 추적하고 기록한다.
		표준 적용 항목의 추적이 가능하다. - 표준과 연계된 DB, 애플리케이션 등의 항목의 확인이 가능하고 표준의 추가, 변경, 삭제 등의 관리 절차가 존재한다.

성숙 수준	품질 기준	품질관리 수준
3	유용성	<p>요건 기반의 데이터 구조 관리가 이루어진다.</p> <ul style="list-style-type: none"> - 요건을 기반으로 데이터 구조를 설계하고 요건의 변경에 따른 영향도를 분석하여 데이터 구조 설계에 반영하는 절차가 존재한다.
		<p>요건이 전사적으로 조율된다.</p> <ul style="list-style-type: none"> - 전사적인 요건을 조율하는 공식화된 기능이 존재하고 요건을 통합하여 일관성 있게 처리한다.
		<p>요건 대응을 위해 자원이 배분된다.</p> <ul style="list-style-type: none"> - 요건에 대한 우선 순위를 관리하며 요건의 신규, 변경에 대한 영향도가 평가된다.
	접근성	<p>사용자 뷰가 통합된 환경에서 제공된다.</p> <ul style="list-style-type: none"> - 사용자 뷰를 다양한 매체를 통하여 제공하며 각각의 단위 시스템들은 통합된 사용자 뷰를 제공한다.
		<p>전사 데이터 구조가 공유된다.</p> <ul style="list-style-type: none"> - 데이터 구조가 전사적으로 관리되고 사용자에게 공유할 데이터 구조가 정의되어 제공된다.
		<p>분산 데이터를 전사적으로 통합한다.</p> <ul style="list-style-type: none"> - 전사적인 데이터 통합 구조가 존재하며 이를 활용한 사용자 뷰가 제공된다.
	적시성	<p>가치 창출을 위한 환경이 제공된다.</p> <ul style="list-style-type: none"> - 사용자의 가치 창출이 가능한 고급의 데이터 분석을 지원하기 위하여 데이터를 축적하고 환경을 제공한다.
		<p>역할과 책임이 명확하게 정의된다.</p> <ul style="list-style-type: none"> - 적시성 관리를 위한 역할을 정의하고 주어진 역할에 따라 관리 활동을 수행한다.
		<p>전사적인 데이터 흐름 주기를 정의한다.</p> <ul style="list-style-type: none"> - 전사적인 데이터 흐름의 선후 관계가 조율되고 우선 순위에 따라 자원이 배분된다.
	보안성	<p>요건 대응에 전문기를 활용한다.</p> <ul style="list-style-type: none"> - 전문가의 성능 진단을 필요 시 또는 주기적으로 수행하며 성능 진단의 결과가 성능 개선에 반영되어 활용된다.
		<p>요건 대응을 위해 자원이 배분된다.</p>

성숙 수준	품질 기준	품질관리 수준
4	정확성	관리 지표가 존재한다. - 정확성 측정 지표를 정의하고 주기적으로 측정한다.
		활동에 대한 성과관리를 수행한다. - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
		오류 해결에 대한 지식이 축적된다. - 오류 해결 내역을 체계적으로 분류하여 지식을 축적하고 오류 대응 시 이를 적극적으로 활용한다.
	일관성	관리 지표가 존재한다. - 정확성 측정 지표를 정의하고 주기적으로 측정한다.
		활동에 대한 성과관리를 수행한다. - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
		오류 해결에 대한 지식이 축적된다. - 오류 해결 내역을 체계적으로 분류하여 지식을 축적하고 오류 대응 시 이를 적극적으로 활용한다.
	유용성	관리 지표가 존재한다. - 유용성 측정 지표를 정의하고 주기적으로 측정한다.
		활동에 대한 성과 관리를 수행한다. - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
		요건 대응 지식이 축적된다. - 요건 대응 내역을 체계적으로 분류하여 축적하여 추후 요건 대응에 필요한 기간 및 자원이 예측 가능하다.
	접근성	요건의 내부 동향이 파악된다. - 요건의 변경 내역이 분류되고 내부 요건의 변경 성향을 파악하여 요건 대응에 활용 한다.
		관리 지표가 존재한다. - 접근성 측정 지표를 정의하고 주기적으로 측정한다.
		활동에 대한 성과관리를 수행한다. - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
	적시성	사용자 활용 패턴이 파악되고 지원된다. - 축적된 데이터를 바탕으로 사용자 활용 패턴이 파악되고 이를 근거로 사용자 뷰를 재구성하여 개인화된 뷰를 제공한다.
		관리 지표가 존재한다. - 적시성 측정 지표를 정의하고 주기적으로 측정한다.

부록 : 성숙수준별 품질관리 점검표

성숙 수준	품질 기준	품질관리 수준
4	적시성	<p>활동에 대한 성과관리를 수행한다.</p> <ul style="list-style-type: none"> - 목표 대비 성과를 측정하고 역할과 책임을 조정한다.
		<p>요건 대응 지식이 축적된다.</p> <ul style="list-style-type: none"> - 요건 대응 내역을 체계적으로 분류하여 축적하여 추후 요건 대응에 필요한 기간 및 자원이 예측 가능하다.
	보안성	<p>고급 수준의 식별 및 인증을 수행한다.</p> <ul style="list-style-type: none"> - 인증서 기반의 인증 및 최소 2가지 이상의 식별 요소를 부여하여 사용자 식별을 수행한다.
		<p>인가된 관리자에 의한 정보 유출을 방지한다.</p> <ul style="list-style-type: none"> - 시스템 관리자, 데이터베이스 관리자, 보안 관리자의 역할을 분담하고 이중화된 유출 차단 정책 및 중요 데이터를 선정하여 적절한 수준의 암호화를 수행한다.
5	정확성	<p>오류의 근본 원인 제거 작업이 수행된다.</p> <ul style="list-style-type: none"> - 오류의 근본 원인을 파악하여 이를 해결하기 위한 계획을 수립하고 관련 업무 프로세스를 개선한다.
		<p>오류의 근본 원인 제거 작업이 수행된다.</p> <ul style="list-style-type: none"> - 오류의 근본 원인을 파악하여 이를 해결하기 위한 계획을 수립하고 관련 업무 프로세스를 개선한다.
	유용성	<p>요건의 외부 동향이 파악된다.</p> <ul style="list-style-type: none"> - 업계의 Best Practice를 반영하여 지속적으로 유지, 발전시키고 데이터 활용 기술 및 동향을 파악하여 적용한다.
		<p>데이터 활용이 최적화된다.</p> <ul style="list-style-type: none"> - 한정된 자원을 최적화하여 데이터 활용도를 극대화한다.
	접근성	<p>사용자 뷰를 최적화한다.</p> <ul style="list-style-type: none"> - 사용자 뷰에 대한 최신 기술과 활용 현황을 접목하여 사용자 뷰를 지속적으로 최적화한다.
		<p>성능을 최적화한다.</p> <ul style="list-style-type: none"> - 개선 요소를 발굴하고 지속적인 튜닝 및 지원을 재분배하여 성능을 최적화하며 성능 향상을 위한 기술 동향을 파악하고 적용한다.
		<p>데이터 흐름을 최적화한다.</p> <ul style="list-style-type: none"> - 불필요한 흐름의 원인을 제거하고 최적화를 위한 최신 기술을 접목하여 활용한다.

성숙 수준	품질 기준	품질관리 수준
5	보안성	<p>계층적인 위협 관리 체계를 구축한다.</p> <ul style="list-style-type: none"> - 데이터베이스 사용 내역을 주기적으로 분석하여 이상 징후에 대비하며, 직접적인 데이터베이스 접근 경로를 차단해 이상 징후 발견 시 자동으로 모든 접속을 차단할 수 있는 방안을 마련한다. <p>보안 수준에 대한 평가를 수행한다.</p> <ul style="list-style-type: none"> - 주기적으로 데이터 보안에 관련된 지침 및 절차에 대해 평가를 수행하며 일련의 보안 행위들이 조직의 데이터 보호 목적에 부합하는지 검증한다.

참여인력

이진우	투이컨설팅 부사장
양승철	투이컨설팅 책임컨설턴트
이명주	투이컨설팅 책임컨설턴트
문화진	투이컨설팅 선임컨설턴트
박상원	투이컨설팅 선임컨설턴트
송동훈	투이컨설팅 선임컨설턴트
박민식	한국데이터베이스진흥센터 실장
김선영	한국데이터베이스진흥센터 팀장
신성수	한국데이터베이스진흥센터 선임연구원
김인창	한국데이터베이스진흥센터 선임연구원
이병준	한국데이터베이스진흥센터 연구원
배영준	한국데이터베이스진흥센터 연구원

데이터 품질관리 시리즈 2

데이터 품질관리 성숙모형 (Ver 1.0) Data Quality Management Maturity Model

1판 1쇄 인쇄 | 2006년 9월 15일

1판 1쇄 발행 | 2006년 9월 15일

발행인 | 원영희

펴낸곳 | 한국데이터베이스진흥센터

주 소 | 서울시 중구 다동 10 한국관광공사빌딩 9층

전 화 | 02-3708-5300

팩 스 | 02-318-5040

인 쇄 | 금영기획 · 인쇄

89-88474-06-6 93560

• 본 책자는 정보통신부 지원으로 한국데이터베이스진흥센터에서 출간하였습니다.

• 본 책자 내용의 무단 전재를 금하며, 인용할 경우 그 출처를 반드시 명기해 주시기 바랍니다.

데이터 품질관리 성숙모형 (Ver 1.0)

Data Quality Management Maturity Model

데이터 품질관리를 도입하기 위해서는 현재의 품질 상태를 정확하게 인식하고 오류 발생 원인과 문제점을 분석해 대응 방안을 마련해야 한다.

이에 한국데이터베이스진흥센터는 조직이 자체적으로 데이터 품질관리 수준을 측정하고 체계적인 데이터 품질관리 전략을 수립할 수 있도록 지원하기 위해 『데이터 품질관리 성숙모형』을 개발했다.

『데이터 품질관리 성숙모형』은 '도입-정형화-통합화-정량화-최적화'의 5단계 성숙수준에 따라 데이터의 품질 기준별로 관리 성숙수준을 측정하도록 고안되었다.

지침의 구성은 다음과 같다.

제1장 데이터 품질관리 성숙모형 개요

제2장 데이터 품질관리 성숙모형의 이해

제3장 품질 기준별 데이터 품질관리 성숙수준