

데이터 품질관리 프로세스 평가를 위한 프로세스 참조모델

The Process Reference Model for the Data Quality Management Process Assessment

김선호(Sunho Kim)*, 이창수(Changsoo Lee)**

초 록

데이터의 품질을 평가하기 위해서 데이터 자체의 품질을 측정하는 방법과 데이터 품질을 관리하는 프로세스를 측정하는 방법이 활용되고 있다. 최근에는 조직의 데이터 품질을 보장 및 인증하기 위해 데이터 품질관리 프로세스의 성숙도를 측정하는 방법을 활용하고 있다. 이러한 추세에 따라 본 논문에서는 데이터 품질관리의 프로세스 성숙도를 평가하는데 필요한 프로세스 참조모델을 제시한다. 우선 데이터 품질관리 프로세스 성숙도 평가 모델의 개요를 제시한다. 그리고, 프로세스 성숙도 평가에 기본이 되는 프로세스 참조모델을 제시한다. 여기서는 프로세스 도출 방안, 데이터 품질관리의 기본 원칙, SPICE 프로세스 참조 모델의 기본 개념을 기초로 하여 프로세스 참조모델의 구성과 세부 프로세스를 개발하였다. 그리고 본 모델의 특징 및 개선점을 ISO 8000-150의 프로세스와 비교하여 설명하였다.

ABSTRACT

There are two ways to assess data quality : measurement of data itself and assessment of data quality management process. Recently maturity assessment of data quality management process is used to ensure and certify the data quality level of an organization. Following this trend, the paper presents the process reference model which is needed to assess data quality management process maturity. First, the overview of assessment model for data quality management process maturity is presented. Second, the process reference model that can be used to assess process maturity is proposed. The structure of process reference model and its detail processes are developed based on the process derivation approach, basic principles of data quality management and the basic concept of process reference model in SPICE. Furthermore, characteristics of the proposed model are described compared with ISO 8000-150 processes.

키워드 : 데이터 품질관리, 프로세스 참조모델, 프로세스, 성숙도, 평가
Data Quality Management, Process Reference Model, Process, Maturity,
Assessment

이 논문은 미래창조과학부의 지원을 받는 정보통신표준화 및 인증지원사업의 연구결과로 수행되었음.

이 논문은 2012년도 강릉원주대학교 교수연구년연구 지원에 의하여 수행되었음.

이 논문은 2012년도 명지대학교 교수연구년연구 지원에 의하여 수행되었음.

* Department of Industrial and Management Engineering, Myongji University(shk@mju.ac.kr)

** Conesponding Author, Department of Industrial, Information and Management Engineering,

Gangneung-Wonju National University(cslee@gwnu.ac.kr)

2013년 07월 05일 접수, 2013년 09월 11일 심사완료 후 2013년 10월 10일 게재확정.

1. 개 요

데이터 품질을 개선하는 방식이 크게 데이터 중심의 품질관리와 프로세스 중심의 품질관리로 구별된다. 데이터 중심의 품질관리는 데이터 오류를 측정하고 발견된 오류를 수정하는 작업이다. 이 방법은 데이터 오류에 대해 신속히 개선하는 효과가 있으며 데이터 품질을 계량적으로 제시하는 장점이 있다. 그러나 시간이 지나면서 같은 종류의 오류가 반복해서 발생하며 데이터 품질의 수준이 하락하는 단점이 있다. 한편, 프로세스 중심의 품질관리는 데이터 품질관리 프로세스를 유지 및 개선하는 방법이다. 이 방법에서는 데이터 오류뿐만 아니라 프로세스 개선을 통해 데이터 오류의 근본적인 원인을 제거하여 동일한 데이터 오류가 재발되는 것을 방지하게 된다.

데이터 품질관리 프로세스를 간과한 사례로서, 2010년 1월 기상청 일기 오보 사례를 들 수 있다. 이 사례에서는 오보 원인이 일기 데이터의 품질관리 수준이 미흡했던 것으로 분석되었다. 관측된 일기 데이터 자체가 부정확하거나 관측점으로부터 일관되지 않고 일치하지 않는 데이터를 입수해 분석하면 아무리 좋은 예측 알고리즘과 빠른 성능의 슈퍼컴퓨터라도 정확한 기상 예측이 어렵다(전자신문 2010. 1. 18). 또한 공공 및 의료 DB 품질관리 미흡(전자신문 2009. 10. 12), 데이터 부실에 기인한 글로벌 금융위기(전자신문 2009. 10. 16) 등 데이터 품질관리 부재로 인한 사례를 볼 수 있다. 이러한 사례는 데이터 간 연계에서 데이터 일치성을 꾸준히 유지하는 활동으로 수집되는 데이터 값과 구조, 관리를 지속적으로 모니터링 하는 데이터 품질관리

프로세스를 체계적으로 수행하지 않은 데서 비롯된 것이다.

기존의 제품품질, 소프트웨어 품질을 다루었던 대부분의 표준이나 모델들은 제품의 품질을 측정 및 보완하는 방식이 아니라 프로세스를 고려하여 제품의 품질을 향상시키는 방식을 택하고 있다. 예를 들면 제품 품질을 다루는 ISO 9000은 품질관리 프로세스를 관리하는 표준이며, 소프트웨어 품질을 다루는 SPICE(ISO/IEC 15504-1)나 CMMI[3, 4, 5]도 소프트웨어 전반의 생명주기 프로세스를 평가하고 개선하여 소프트웨어의 품질을 향상시키는 표준과 모델이다. 마찬가지로 데이터 품질을 다루는 모델에서도 한국의 DQM3[8], 스페인의 IQM3[10], DAMA DMBOK Guide [24], ISO 8000-150[1, 2]등이 있는데 이 모델들은 프로세스를 통하여 또는 프로세스의 성숙도를 평가하고 프로세스를 개선함으로써 데이터의 품질을 향상시키는 모델들이다. 제품, 소프트웨어, 데이터에 대한 프로세스 중심의 품질관리 및 평가 표준 및 모델들이 <Table 1>에 요약되어 있다.

그러나, 기존의 프로세스 중심의 데이터 품질관리 모델들은 다음과 같은 내용에서 개발한 관점이 달라 프로세스들이 다양하며 일관되게 정의되어 있지 않다.

DQM3는 데이터베이스 중심의 데이터 품질을 관리하는 데이터베이스 통제자(database administrator), 데이터 구조 전문가(data architect) 등과 같은 데이터 관리 전문가(data management professional)의 관점에서 8개의 프로세스가 제시되었다. DQM3는 데이터 관리 경영진(data management executive 또는 CIO)에서 필요한 데이터 거버넌스의 측면이

〈Table 1〉 Examples of Process-Oriented Quality Management

Category	Standard/Model	Descriptions	Process/ Data
Product quality	ISO 9000 series	product quality management process standard	Process-oriented
Software quality	ISO 15504 SPICE	standard for software process improvement and capability assessment	
	ISO 15288 System life-cycle processes	standards for system lifecycle processes	
	CMMI	software process capability with five levels	
Data quality	DQM3	data quality management limited to database	
	IQM3	CMMI-based information quality management	
	DAMA DMBOK Guide	focused on data management rather than data quality management	
	Federal DAS-Data quality framework	PDCA processes for data quality improvement based on enterprise architecture	
	ISO 8000-150	master data quality management framework	
	ISO 25012 SQuaRE	the software quality model that includes both software and data properties and quality characteristics	
ISO 8000-100, 110, 130, 140, 311	exchange of characteristic data for master and product data		

나, 데이터 품질관리에 필요한 조직 및 자원의 지원이나 인력의 교육 등의 분야는 고려되지 않고 있다.

IQM3는 소프트웨어 품질을 다루는 CMMI 프로세스를 참조하여 14개의 정보품질관리 프로세스를 성숙도에 따라 5단계로 구분하여 제시하였다. 이 프로세스는 정보를 생성하는 기술적 프로세스와 정보 품질을 관리하는 프로세스를 포함하고 있다. IQM3에서는 데이터 자체보다는 데이터를 근거로 생성되는 정보(information)를 품질관리 대상으로 삼고 있어서 있어 데이터 자체에 대한 고려가 필요하다.

DAMA의 DMBOK Guide는 데이터 및 정보자산을 획득, 통제, 보호, 전달, 및 가치 향상시키기 위한 정책, 실무, 프로젝트를 계획하고 실행하는 비즈니스 프로세스를 정의하며 데이터 관리 프로세스를 크게 10개 분야로 분류하고 각각 프로세스별로 세부 활동(activity)을 정의하고 있다. 이들 프로세스는 데이터의 품질관리가 아닌 데이터 관리(data management)를 목적으로 제시되었다는 한계를 가진다.

Federal DAS Data Quality Framework는 데이터 품질을 개선하기 위해 Plan-Do-Check-Act의 관점에서 13개의 프로세스를 제시하고

있으며 전사적 아키텍처(enterprise architecture)에 기반을 두어 공공부문 적용 측면을 고려하였다. 그러나 프로세스간의 연관 관계와 프로세스를 수행하기 위한 계층이나 역할에 대한 고려가 필요하다.

ISO 8000-150는 데이터의 품질관리는 크게 데이터 활용, 데이터 품질 모니터링, 데이터 품질 개선의 3단계로 진행되며 각 단계는 조직의 역할에 따라 3개의 프로세스 레벨로 구분하여 모두 9개의 프로세스로 구성된 프레임워크를 제시하고 있다. 그러나 이 모델은 데이터 운용이나 조직 및 자원의 지원 등의 분야는 고려하지 않고 있다.

이와 같이 모델에 따라 프로세스나 활동들이 깊이(depth)나 범위(granularity)면에서 일관성이 없다. 따라서 어느 모델을 이용하여 프로세스 성숙도를 평가하느냐에 따라 결과가 달라질 수 있다. 이러한 문제점을 해결하기 위해, 본 연구에서는 데이터 품질관리 평가를 위한 표준화된 모델 구성을 제시하며 평가에 가장 기본적으로 필요한 프로세스 참조모델을 제시한다. 우선, 데이터 품질관련 국제 표준, 프로세스 기반의 데이터 품질관리 모델, 소프트웨어 프로세스 평가 모델, 데이터 품질관리 성숙 모델들에 대한 연구동향을 분석하였다. 둘째, 데이터 품질관리 프로세스 성숙도 평가 모델의 구성 개요를 제시한다. 셋째, 데이터 품질관리 프로세스 성숙도 평가를 위해 사용되는 프로세스 참조모델을 제시한다. 이 참조모델에서는 기존의 모델에서 제시되는 프로세스들을 고려하여 공통적이고 핵심적인 프로세스를 도출한다. 그리고, 도출에 필요한 데이터 품질관리 기본원칙, 프로세스 참조모델의 기본 구조 및 세부 구조를 제

시한다. 마지막으로, 기존의 모델들과의 비교를 통해 프로세스 참조모델의 특징 및 차이점을 제시한다.

2. 관련 연구 동향

2.1 데이터 품질 관련 국제 표준

데이터 품질에 대한 표준은 국제적으로 ISO 8000, ISO 25012, ISO 10303-59, ISO 191XX 등이 개발되고 있다. ISO 8000에서는 ISO 22745 개방형 기술사전(Open Technology Dictionary)과 같은 참조 데이터 사전을 기반으로 마스터 데이터(master data)를 교환하기 위한 표준을 개발하기 시작하였다. 처음에는, 표준화를 위한 데이터를 마스터 데이터, 트랜잭션 데이터(transaction data), 제품 데이터(product data)로 구분하여 개발을 진행하였다[11, 12]. 그러나, 2011년부터는 마스터 데이터와 트랜잭션 데이터를 구분하지 않게 되었으며 일반 데이터(general data)뿐만 아니라 정보(Information)의 품질로 개념이 확대되고 있다. 또한, ISO 8000-150 데이터 품질관리 프레임워크가 제시되면서 데이터 중심의 품질 개선 개념에서 프로세스 중심의 품질 개선 개념으로 확대되고 있다[13].

ISO/IEC 25000 소프트웨어 제품품질 요구 사항 및 평가(SQuaRE : Software Product Quality Requirements and Evaluation) 시리즈는 ISO JTC1 Information Technology 산하 SC7 Software and System Engineering 위원회에서 개발하고 있다. 이 시리즈 중에서

ISO 25012에서 데이터 품질을 다루고 있는데 데이터 품질 특성을 내부 데이터 품질, 외부 데이터 품질, 활용 데이터 품질 관점에서 20개의 특성으로 분류 및 정의하고 있다[20].

ISO 10303-59 제품 형상 데이터의 품질(PDQ-S : Quality of product shape data) 표준은 ISO TC184/SC4/WG12에서 개발하고 있다[16]. 이 표준은 SASIG(Strategic Automotive product data Standards Industry Group)와 일본의 주도로 개발된 것으로서 PDQ(Product data quality)를 정의하며 이 PDQ 정보의 활용법을 제시하고 있다.

ISO TC211 Geographic Information/Geomatics 기술위원회에서 ISO 19100 시리즈를 개발하고 있다[15]. 이 중에서 WG9 information Management 워킹그룹이 자료생산, 아이템 등록절차, 데이터 품질 측정, 크로스 도메인 용어에 대한 표준을 개발하고 있다. 19113에서 품질 원칙, 19114에서는 품질 평가 프로세스, 19138에서는 데이터 품질 측정을 규정하고 있으며 19157은 이 세 표준을 통합하여 데이터 품질 전반을 규정하고 있다. 19158에서는 19157을 기반으로 품질보증에 대한 표준을 규정하고 있다.

ISO 표준 이외에도 OASIS의 Customer Information Quality TC에서는 고객정보인 Name, Address, Party에 대한 XML 스키마를 명시하는 품질 표준을 개발하였다[25]. 국내에서는 ISO 8000 시리즈를 KS로 전환하는 작업을 진행하고 있으며 독자적인 데이터 품질 관련 표준은 아직 없는 실정이다.

김선호(2010)와 이창수(2010)는 ISO 8000-150(마스터 데이터 품질관리 프레임워크)을 기반으로하여 3가지 역할(관리자, 통제자, 실

행자)과 3가지 기능(데이터 활용, 품질 모니터링, 품질 개선)을 매트릭스로 구성하여 9개의 품질관리 프로세스를 제안하고 있다[1, 2].

소프트웨어 품질이나 제품품질에 대한 표준은 현업에서 널리 활용되어 사용하고 있으나 데이터 품질에 대한 표준은 최근 개념 정립이 되어 표준으로 만들어가고 있는 실정이다. 본 논문은 마스터 데이터 품질관리 프레임워크(ISO 8000-150)를 확장하여 일반 데이터에 대한 데이터 품질관리를 다루고 있다.

2.2 프로세스 기반 데이터 품질관리 모델

미국의 연방 데이터 아키텍처 소위원회(Data Architecture Subcommittee)에서는 데이터 공유 및 시스템 통합 운용을 위해서 연방 EA(Enterprise Architecture)기반의 데이터 품질 프레임워크를 제시하였다[7]. 이 프레임워크에서 데이터 품질은 관심 단계에서 데이터의 사용을 위해 필요한 품질 특성을 보장하는 일련의 원칙과 절차로 간주된다. 이것은 입증된 방법, 산업 표준, 과거 성과에 기반을 두어 사람, 프로세스, 기술을 위한 접근방법을 정의하고 있다. 여기서는 연방 EA를 5개의 참조모델로 구분하고 각 참조모델 별로 수행 프로세스를 정의하고 있다. 그리고 데이터 품질을 위한 개선 활동 프로세스로서 14개의 프로세스를 제시하고 있다.

유럽의 데이터정보품질협회(EIDIQ : European Institute for Data and Information Quality)는 데이터 및 정보의 효율적 관리를 지원하기 위하여 다양한 산업 협력 프로세스를 통해 정보 품질관리(IQM : Information Quality Management)에 대한 적절한 절차, 방법,

지침에 대한 표준 절차(SP : Standard Procedures)를 개발하였다[23]. 이 모델에서는 프로세스, 조직, 인력, IT 시스템, 데이터를 고려하여 품질관리 영역을 제시하고 있다. 그리고 각 품질관리 영역별로 정보 품질관리의 성숙수준을 3단계로 나누고 각 수준별로 수행 프로세스를 제시하고 있다.

DMBOK(Data Management Body of Knowledge) 가이드에서는 데이터 관리를 고수준의 비즈니스 프로세스로 정의하고 데이터에 대한 계획 및 실행, 정책, 프랙티스, 프로젝트, 획득, 통제, 보호, 전달, 가치 향상, 데이터 및 정보 자산의 관점에서 데이터 관리 프로세스를 정의하고 있다[24]. 여기서는 10개의 데이터 관리 기능으로 분류하여 각 기능에 따라 세부 프로세스로 정의하였다.

프로세스 중심의 데이터 품질관리는 데이터 중심의 데이터 품질관리의 내용을 내포하고 있다. 특히 데이터 거버넌스의 개념을 위해서는 프로세스 중심의 데이터 품질관리가 필수적이다. 본 논문에서는 데이터 자체뿐만 아니라 프로세스까지 확대한 관리 체계를 다룬다.

2.3 소프트웨어 프로세스 평가 모델

카네기멜론 대학의 SEI(Software Engineering Institute)는 미국 국방성(DoD)의 요청으로 소프트웨어 프로세스 능력을 평가하는CMMI(Capability Maturity Model Integration)를 개발하였다. CMM은 각 수준 별로 핵심적으로 수행되어야 하는 프로세스들을 정의하고 있고 업체에서 그 프로세스들을 모두 수행하고 있으면 해당 수준을 달성했다고 판단한다. 그리

고 프로세스 능력을 향상시킬 수 있는 방법을 명확히 제시하고 있다. CMM에서는 24개의 프로세스 영역을 고려하고 있으며, 프로세스의 능력(process capability)을 6단계로 평가한다. 조직의 성숙도(organizational maturity)는 여러 프로세스의 능력을 종합하여 5단계로 평가한다[3, 4, 5].

ISO 9000-3이 소프트웨어 분야의 특성을 고려하지 못함에 따라 ISO/IEC JTC1 SC7 WG 10에서는 소프트웨어 프로세스에 대한 능력 평가와 프로세스 개선을 목적으로 SEI의 CMM, Bell의 TRILLIUM, Esprit의 Boot Strap 등을 통합하여 ISO/IEC 15504 (SPICE : Software Process Improvement and Capability dEtermination)를 개발하였다[17]. SPICE는 고품질의 소프트웨어를 효율적으로 개발, 생산함으로써 소프트웨어 품질과 생산성을 향상시키는 역할을 한다. 또한, 최종 단계에서 산출물 중심의 평가를 실시하는 ISO 9000 시리즈의 취약점을 보완하는 모델로서 지속적인 프로세스의 개선에 역점을 두고 있다. SPICE의 심사모델은 프로세스 현 상태의 파악, 요구 사항에 대한 프로세스 수행능력, 적합성, 조직의 수행 능력을 심사하기 위해 사용된다. 이 평가 모델은 ISO 12207 소프트웨어 생명주기 프로세스[17]를 참조모델로 하여 프로세스를 평가한다. 소프트웨어 품질과 생산성을 향상시키는 표준 역할을 하는 CMMI와 SPICE는 본 논문의 모델이 포함되어 있는 제 3장 전체 모델의 기초 개념을 제공하고 있다.

시스템 및 소프트웨어 엔지니어링을 위한 시스템의 생명주기 프로세스에 대한 정의는 ISO 15288 : 2008[18]에서 정의되고 있다. 2009년부터는 ISO/IEC 15504 표준을 ISO/IEC

33000시리즈로 개편하기로 하고 ISO 33001~33099번호를 할당하여 새로운 표준화 작업을 진행하고 있다[21].

2.4 데이터 품질관리 성숙 모델

데이터 품질관리 성숙 모델은 데이터 품질 관리에 대한 체계적인 접근을 가능하게 하는 객관적인 기준에 의하여 데이터 품질관리 수준을 측정하고 더 나은 수준을 확보하기 위한 가이드라인을 제시하고 있다. 이것은 단지 현재의 품질 수준을 측정하기 위한 것뿐만 아니라 품질향상을 위해서 어떤 프로세스를 어떻게 개선해야 하는 지에 대한 방향을 제시하고 있다.

한국데이터베이스진흥원에서 제시한 DQM3 (Data Quality Management Maturity Model)은 크게 데이터 품질 기준, 데이터 품질 관리 프로세스, 데이터 품질관리 성숙수준의 요소로 구성되어 있다[8]. 데이터 품질 기준은 데이터 품질의 다양한 측면을 분류, 정리해 정확성, 일관성, 유용성, 접근성, 적시성, 보안성의 6개로 정의하였으며, 데이터 품질 관리 프로세스는 요구사항 관리, 데이터 구조 관리, 데이터 흐름 관리, 데이터베이스 관리, 데이터 활용 관리, 데이터 표준 관리, 데이터 오너십 관리, 사용자 뷰 관리가 포함된다. 세 번째 요소인 데이터 품질관리 성숙 수준은 조직의 데이터 품질관리 발전 단계를 고려해 ‘도입-정형화-통합화-정량화-최적화’의 5단계로 구분하고 6개의 데이터 품질 기준별로 데이터 품질관리 성숙 단계를 정의하였다.

한편, Ryu et al.[27]은 전사적 통합과 연계의 관점, 데이터의 구조적 품질 관리의 관점,

성숙 단계의 관점과 같은 3가지 관점을 반영한 성숙 모델을 제시하였다. 이 모델은 초기 데이터 관리, 정의데이터 관리, 관리데이터 관리, 최적화 데이터 관리의 4단계로 구성되어 있다.

IBM 데이터 거버넌스 협의회에서 제시한 성숙 모델은 조직의 소프트웨어 개발 프로세스를 개발하고, 정의하는 방법론을 이용하여 개발하였다[9]. 이 방법은 개발 프로세스의 성숙도에 따라 초기(initial), 관리(managed), 정의(defined), 정량적 관리(quantitatively managed), 최적화(optimizing)의 5단계로 구성된다. 성숙 모델은 조직 구조 및 인지도, 관리규정, 정책, 가치 창출, 데이터 위험 관리 및 순응, 정보 보안 및 프라이버시, 데이터 구조, 데이터 품질 관리, 분류 및 메타 데이터, 정보 생명주기 관리, 감사 정보, 로깅 및 보고와 같은 11개의 도메인으로 구성되어 있다. 각 도메인은 많은 서브 도메인으로 구성되어 있으며, 서브 도메인은 5단계의 성숙도로 분류되어 있다.

Pipino et al.[26]는 데이터 품질의 주관적 및 객관적 평가 요소를 나열하고 객관적 데이터 품질 평가 요소에 대한 3가지 기능적 형태(단순 비율, 최소 혹은 최대 운영, 가중 평균)를 제시하여 실제 평가에서 주관적 평가요소와 객관적 평가요소를 3단계로 적용하여 조직의 데이터 품질을 개선하는데 사용하고 있다[26].

Ismael Caballero et al.[10]은 정보품질관리 성숙모델(IQM3)과 정보품질관리의 평가 및 개선 방법론(MAIMIQ)을 제시하고 있다. IQM3 (Information Quality Management Maturity Model)는 정보 품질관리 성숙 모델로서 CMMI

의 성숙 모델을 참조하여 Initial, Defined, Integrated, Quantitatively Managed, Optimizing의 다섯 단계를 정의하고 있다. MAIMIQ (A Methodology for Assessing and Improving the Management of Information Quality)는 정보 품질관리의 평가 및 개선 방법론으로서 Plan-Do-Check-Act과정을 거쳐 정보 품질을 개선한다.

데이터 품질관리가 얼마나 잘 되어 있는가를 평가하여 개선하는 순환체계를 이루려면 데이터 품질 성숙도의 객관적 수준이 잘 정립되어 있어야 한다. 본 논문에서는 성숙도에 대한 심도 있는 논의는 하지 않지만 프로세스 참조모델이 궁극적으로 지향하는 목표는 고수준의 성숙도 수준을 달성하는 근거가 되는 것이다.

3. 데이터 품질관리 프로세스 평가 모델 개요

데이터 품질관리를 위한 프로세스의 성숙도를 평가하고 데이터 품질을 향상시키기 위한 모델은 <Figure 1>에서와 같이 프로세스 참조 모델, 프로세스 성숙 모델, 프로세스 측정 모델, 프로세스 개선 모델로 구성된다. 이 네 가지 모델은 ISO 15504 SPICE (소프트웨어 프로세스 개선 및 능력 평가 기준) 모델의 개념을 기반으로 제시하였다[19]. 데이터 품질관리 프로세스 평가 모델의 개괄적인 내용은 다음과 같다 :

1) 프로세스 참조모델(process reference model)

이 모델의 주요 내용은 본 논문의 제 4장에

서 다루고 있으며 정보 및 데이터품질관리를 위해 요구되는 프로세스를 정의한다. 이 프로세스는 성숙모델과 측정모델에서 참조된다. 이 모델은 다음과 같은 내용을 포함한다 :

- 프로세스 및 세부 프로세스 정의 및 분류
- 프로세스 별 목적, 결과물, 세부 활동

2) 프로세스 성숙모델(process maturity model)

이 모델은 정보 및 데이터 품질관리 프로세스의 성숙 레벨을 정의하고 각 레벨 별로 핵심 프로세스 영역(key process areas : KPA)을 정의한다. 이 모델은 다음과 같은 내용을 포함한다 :

- 성숙 레벨 정의
- 레벨별 KPA 속성 정의

3) 프로세스 측정 모델(process measurement model)

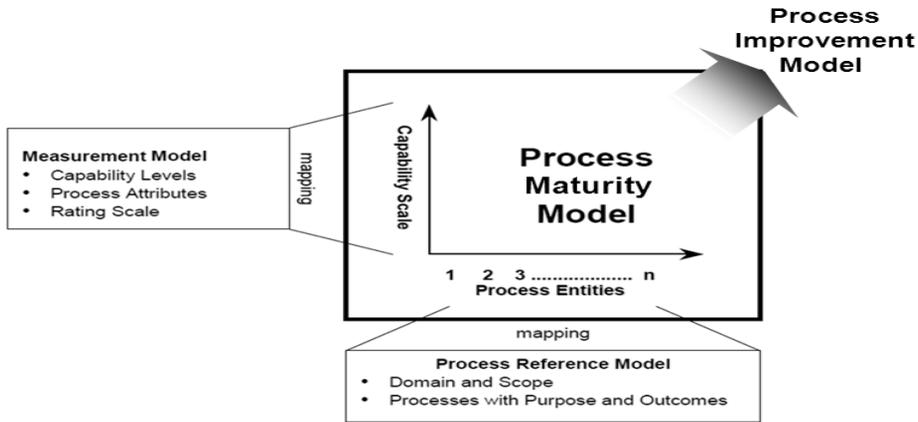
이 모델은 데이터 품질관리 프로세스 성숙 레벨을 측정하는 기준과 방법을 정의한다. 이 모델은 다음과 같은 내용을 포함한다 :

- 프로세스 성숙도 정의
- 프로세스 측정을 위한 측정항목 정의
- 각 측정항목별 측정 지표

4) 프로세스 개선 모델(process improvement model)

이 모델은 조직 내에서 정보 및 데이터 품질관리 프로세스 성숙도를 개선하여 데이터 품질을 향상시키기 위한 방안 또는 지침을 정의한다. 이 모델은 다음과 같은 내용을 포함한다 :

- 개선 프로세스 정의
- 프로세스 성숙도 개선을 위한 요구사항 및 절차



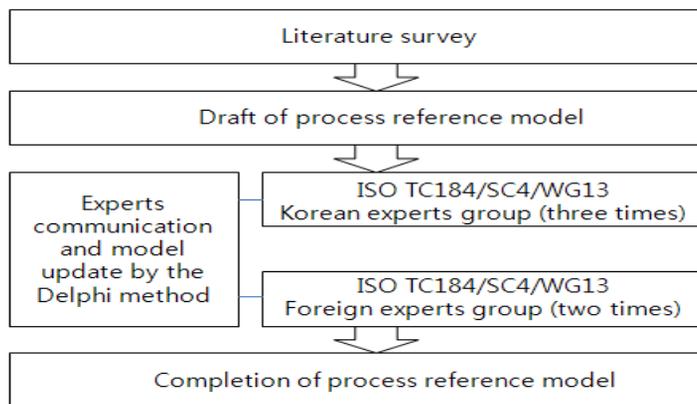
<Figure 1> The structure of Data Quality Management Process Assessment

4. 데이터 품질관리 프로세스 평가를 위한 프로세스 참조모델

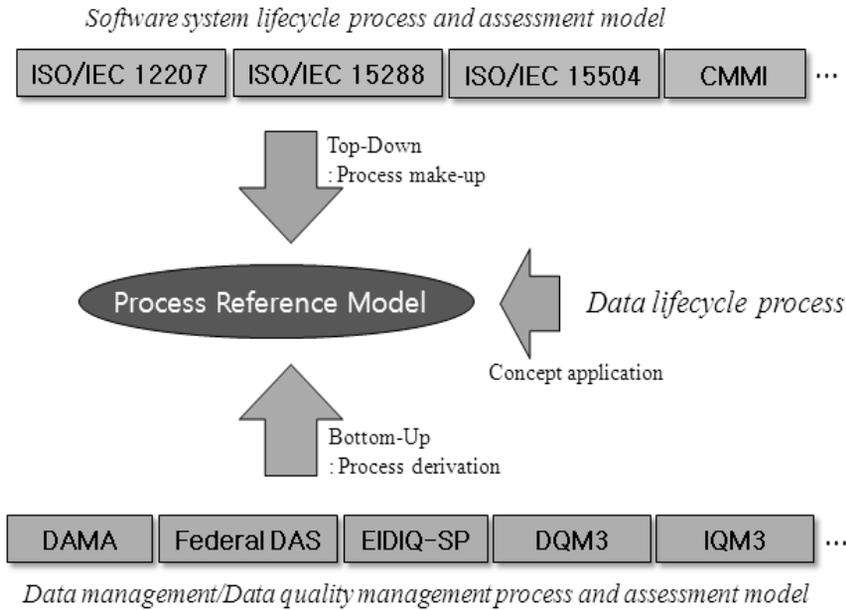
4.1 프로세스 참조모델 개발 과정

프로세스 참조모델 개발을 위해 신제품이나 신모델 도입 시 많이 활용되는 델파이 기법을 이용하여 <Figure 2>와 같이 ISO TC184/

SC4/WG13 국내외 전문가들을 대상으로 의견 조사를 실시하였다. 먼저, 본 논문의 저자들이 기존 문헌과 현장 평가 사례의 조사 연구를 통하여 초안을 작성한 뒤 데이터 품질 국내 전문가 그룹 4명의 의견을 3회 반복 수렴하여 초안을 수정하였다. 이 초안은 다시 영국, 스페인, 독일, 프랑스, 노르웨이 5개국 전문가 그룹 10명의 의견을 2회 반복 수렴하여 완성하였다



<Figure 2> The Development Procedure of the Process Reference Model



<Figure 3> The Derivation Procedure of the Process Reference Model

4.2 프로세스 참조모델의 프로세스 도출 과정

프로세스 참조모델의 프로세스는 <Figure 3>에서와 같이, 데이터 품질관리와 관련된 기존의 모델 중에서 정의된 데이터 관리 프로세스나 정보/데이터 품질관리 프로세스, 프로세스 성숙모델(ISO 8000-150, DQM3, IQM3, EIDIQ, DAMA DMBOK Guide, Federal DAS Data Quality Framework) 등에서 정의된 프로세스를 참고하여 기본적으로 필요한 프로세스를 도출하고, ISO 9000에서 사용된 PDCA(Plan-Do-Check-Act)의 개념에 따라 프로세스를 분류하게 된다. 추가로, 기존의 소프트웨어/시스템 생명주기 프로세스 및 성숙 모델(ISO 12207, 15504, 15388, CMMI 등)을 참조하여 프로세스를 보완한다. 이 프로세

스는 데이터 생명주기 프로세스 개념[6]도 반영하여 도출되었다.

4.3 프로세스 참조모델에서의 데이터 품질관리 기본 원칙

프로세스 참조모델의 프로세스를 정립할 때는 일정한 원칙이 필요하다. 이 프로세스는 데이터 품질관리를 위한 프로세스이며 본 절에서는 조직 차원에서 고려해야 할 데이터 품질관리의 기본 원칙을 설명한다.

ISO 9000 시리즈도 제품품질 개선을 위해 지속적 개선의 개념을 제안하고 있으며 프로세스 중심의 개념을 강조하고 있다[14]. 데이터 품질을 위해서는 데이터 값 중심과 데이터 프로세스 중심의 두 가지 방향이 있는데 지속적 개선을 위해서는 프로세스 방식의 접

근이 강조되어야 한다. 이는 ISO 8000-150 [13]에서도 지적한 바가 있으며 김선호(2010)와 이창수(2010)의 연구에서는 데이터 품질 개선에 수반된 다양한 계층의 이해관계자가 각자의 역할에 맞는 책임을 수행해야 한다는 것도 강조하고 있다[1, 2]. 이와 같은 기존의 연구 결과를 토대로 데이터 품질관리를 성공적으로 수행하기 위해서는 조직은 다음과 같은 세 가지 기본 원칙을 수립할 필요가 있다 :

- 1) 프로세스 방식의 접근 : 데이터 중심의 측정 및 수정은 조직 전체의 데이터 품질을 개선하기에 충분하지 않다. 바람직한 데이터 품질 수준은 데이터 품질과 관련된 활동과 지원 자원들이 프로세스로 관리될 때 더욱 효과적으로 달성될 수 있다.
- 2) 지속적인 개선 : 데이터 품질은 데이터 처리, 데이터 오류 측정 및 수정의 핵심 프로세스를 통해 지속적으로 개선된다. 그러나 이러한 프로세스만으로는 동일한 데이터 오류가 반복적으로 발생하는 것을 막을 수 없다. 동일한 데이터 오류 재발 방지는 이러한 프로세스들이 데이터 품질을 방해하는 근본원인을 분석, 추적, 제거하는 프로세스들과 함께 수행될 때 가능하다. 이러한 핵심 프로세스들이 더욱 효율적으로 수행되기 위해서는, IT 기반의 데이터 관리 프로세스들이 입력 데이터나 통제 정보를 핵심 프로세스에 제공해야 한다. 또한, 데이터 품질부서 및 자원 관리를 위한 프로세스들도 핵심 프로세스와 지원 프로세스에 자원을 제공해야 한다.
- 3) 인력의 관여: 데이터 품질관리를 위한

역할을 담당할 모든 레벨의 구성원들이 조직의 데이터 품질을 개선하는데 관여해야 한다. 하위 레벨에서 데이터 처리를 담당하는 사용자가 데이터 품질에 가장 직접적인 영향을 주지만, 중간 레벨의 데이터 관리자의 개입과 통계가 조직의 데이터 품질 개선을 위한 프로세스를 실행하고 정착시키는데 필요하다. 또한, 전사적인 데이터 품질에 책임을 지고 있는 상위 레벨의 경영층도 조직의 역할, 권한, 프로세스를 변경하고 최적화 하는데 반드시 참여하여야 한다.

4.4 프로세스 참조모델의 기본 구조

데이터 품질관리 프로세스 성숙도 평가를 위한 프로세스 참조모델의 기본 개념은 다음과 같다(<Figure 4> 참조).

- ① 프로세스 참조모델은 핵심 프로세스, 지원 프로세스, 조직 프로세스의 세 종류의 클래스로 구성된다.
- ② 지속적인 데이터 품질 개선을 이루기 위해, 핵심 프로세스는 Plan-Do-Check-Act(PDCA)의 개념에 따라 연속적으로 수행되어야 한다.
- ③ 데이터 품질 개선을 성공적으로 수행하기 위해서는 지원 프로세스는 데이터 관리에 관련되는 정보 및 IT 기술을 핵심 프로세스에 제공해야 한다. 지원 프로세스는 필요시 핵심 프로세스로부터 품질 개선 결과를 피드백 받아 보완될 수 있다.
- ④ 데이터 품질 개선을 효율적으로 수행하기 위해서는 조직 프로세스는 핵심 프

로세스와 지원 프로세스에 전사적 차원에서 조직, 예산, 인력 등과 같은 데이터 품질 자원과 훈련 서비스를 제공하여야 한다. 그리고 조직 프로세스는 필요시 핵심 프로세스로부터 개선 결과를 피드백 받아 보완될 수 있다

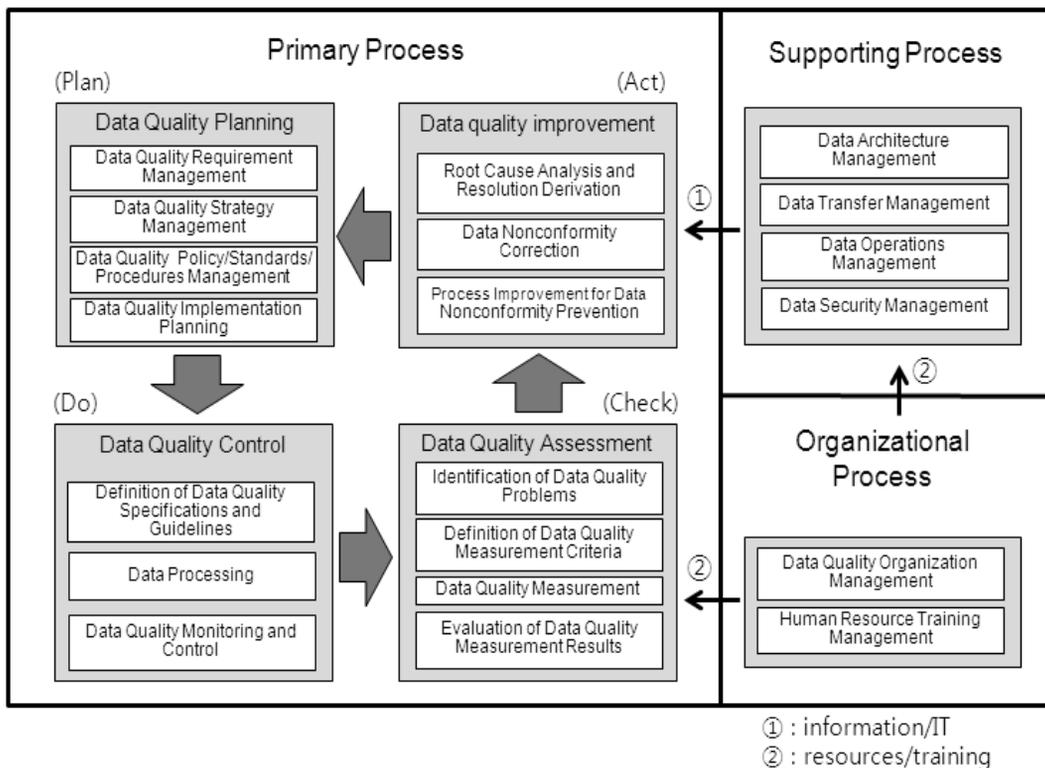
4.5 프로세스 참조모델의 세부 구조

<Figure 4>에서와 같이, 프로세스 참조 모델은 3개의 상위 프로세스 클래스로 구성되어 있으며 이중에서 핵심 프로세스 클래스는 데이터 품질 계획, 데이터 품질 통제, 데이터 품질 평가, 데이터 품질 개선의 4개 하위 프

로세스 클래스로 세분된다. 각 클래스는 다시 세부 프로세스로 구성되어 있다. 여기서 핵심 프로세스 클래스의 세부 프로세스들은 PDCA의 실행 순서에 따라 나열되어 있으며, 지원 프로세스 및 조직 프로세스 클래스의 세부 프로세스들은 임의의 순서로 나열되어 있다. 각 프로세스는 ISO 24774[22]에서 제시된 프로세스 정의 형식을 참조하여 프로세스의 목적(purpose), 프로세스의 수행으로 생성되는 결과물(outcomes), 프로세스를 구성하는 세부 활동(activity)을 명시하였다.

1) 핵심 프로세스

이 프로세스 클래스는 데이터 품질관리를



<Figure 4> The Structure of Data Quality Management Processes

위한 핵심 실행 프로세스들을 포함하고 있다. 이 클래스는 PDCA의 개념에 따라 다음과 같이 4개의 하위 클래스와 세부 프로세스로 구분된다.

① 데이터 품질 계획(Plan) : 이 프로세스 클래스는 전사적으로 수집된 데이터 품질 요구사항에 따라 품질 전략을 수립하고 이를 달성하기 위해 제반 여건을 확보하며 실행 계획을 수립 및 실행하고 실행결과를 평가하기 위한 프로세스들이다. 세부 프로세스는 다음과 같다 :

- 데이터 품질 요구사항 관리 : 이 프로세스는 데이터 사용자나 다른 이해당사자들의 정보를 수집하여 데이터 품질 요구사항을 수립하고, 이를 분석하여 그 결과를 데이터 품질 전략에 반영하는 것이다. 주요 결과물은 전사적 데이터 품질 요구사항 수집, 요구사항 정의, 식별 및 분류, 요구사항에 대한 기술, 비용, 인력, 일정의 타당성 분석, 요구사항의 우선순위 부여, 승인, 업데이트, 요구사항 반영 결과 평가, 요구사항관리 프로세스에 대한 전사적인 통합이 있으며 세부 활동으로는 데이터 품질 요구사항 수집, 우선순위 설정, 성과 평가가 있다.
- 데이터 품질 전략 관리 : 이 프로세스는 조직 내·외적인 데이터 품질에 대한 장기적인 목적(goal)을 설정하고 이를 달성할 수 있도록 단기적인 목표(objective)를 수립하는 것이다. 주요 결과물은 데이터 품질 개선을 위한 전사적 차원에서 최고 경영진의 헌신, 비전, 장기 목적, 구현 청사진이 정의

된 데이터 품질 전략, 정량적 결과물 관점에서 정의된 단기 목표, 데이터 품질 전략의 성과 결과 평가 및 이에 따른 전략 갱신, 전사적 차원에서 데이터 품질 전략 공유가 있으며 세부 활동으로는 데이터 품질 전략 수립, 성과 평가가 있다.

- 데이터 품질 정책/표준/절차 관리 : 이 프로세스는 데이터 품질 전략에 따라 조직에서 데이터 품질 관리를 위한 정책, 표준, 절차를 규정하고 관리하기 위한 것이다. 주요 결과물은 데이터 품질관리를 위한 기본적인 의도 및 원칙 정의, 요구사항 명세, 측정 방법, IT 시스템의 인프라, 데이터 품질관리를 위한 표준 절차 정의, 데이터 품질관리를 위한 정책, 표준, 절차의 전사적 공유가 있으며 세부 활동으로는 데이터 품질 정책 관리, 데이터 품질관리를 위한 표준관리, 데이터 품질관리를 위한 절차관리가 있다.
 - 데이터 품질 확보 계획 관리 : 이 프로세스는 데이터 품질의 범위와 대상을 정의하고 데이터 품질 목표 달성을 위해 상세 확보 계획을 수립하는 것이다. 주요 결과물은 데이터 품질 목표에 따른 데이터 품질 범위 및 대상 정의, 상세 실시 계획 수립, 실시 계획 수행에 필요한 자원(인력, 비용, 기술) 할당 및 관리가 있으며 세부 활동으로는 데이터 품질 실시 계획 수립, 자원 할당, 계획 실행, 성과 평가가 있다.
- ② 데이터 품질 통제(Do) : 이 프로세스 클래스는 데이터 품질 목표를 달성하기

위해 실시간 또는 일상적으로 수행하는 프로세스들로서 데이터 생명주기를 통하여 명세와 지침에 따라 데이터를 처리하고 통제하여 데이터의 품질을 유지하기 위한 것이다.

- 데이터 품질 명세 및 지침 정의 : 이 프로세스는 데이터 품질 통제를 수행하는 데이터 사용자와 데이터 품질 작업자를 위해 상세한 데이터 품질 관련 제반 명세 및 지침을 정의하는 것이다. 주요 결과물은 사용자를 위한 데이터 품질 명세, 데이터 처리 지침 및 데이터 오류 수정 지침 정의, 데이터 품질 모니터링 및 통제를 위한 명세, 운용 지침, 예방 활동 정의가 있으며 세부 활동으로는 데이터 사용자를 위한 명세 및 지침 정의, 데이터 품질 작업자를 위한 명세 지침 정의가 있다.
 - 데이터 처리 : 이 프로세스는 데이터 품질을 확보하기 위해 데이터 생애주기 동안 데이터 사용자나 애플리케이션이 데이터 품질 명세 및 지침을 준수하여 데이터를 처리하도록 한다. 주요 결과물은 데이터 수명주기내 각 단계별로 데이터 명세 및 지침, 데이터 사용자나 애플리케이션의 데이터 처리 기록 생성 및 저장이 있으며 세부 활동으로는 가이드라인에 따라 데이터 생성, 읽기, 갱신, 삭제, 변형, 이전을 수행하는 사용자 또는 애플리케이션의 데이터 처리 활동과 데이터 처리 기록을 저장 및 관리하는 시스템의 데이터 로깅 활동이 있다.
 - 데이터 품질 모니터링 및 통제 : 이 프로세스는 데이터 품질을 운용지침에 따라 정기적 또는 실시간으로 모니터링하고 발생한 데이터 오류를 교정하며 유사한 데이터 오류 재발 방지하기 위한 통제 활동을 수행하기 위한 것이다. 주요 결과물은 데이터 오류 식별 및 분류, 데이터 오류 수정 및 데이터 오류관련 측정 기록 저장이 있으며 세부 활동으로는 데이터 품질 모니터링 실시, 데이터 오류 교정, 데이터 오류 예방 활동이 있다.
- ③ 데이터 품질 심사(Check) : 이 프로세스 클래스는 정기적 또는 사후적으로 데이터 품질을 평가하기 위한 프로세스들로서 데이터 품질 통제 프로세스 클래스에서 해결되지 않은 데이터 오류를 측정 및 평가하기 위한 것이다.
- 데이터 품질 문제 파악 : 이 프로세스는 데이터 품질 통제 프로세스에서 해결되지 못한 문제를 식별하는 것이다. 주요 결과물은 조직 내외에서 제기된 품질 이슈를 수집 분석, 이슈의 데이터 품질 특성을 식별 분류가 있으며 세부 활동으로는 데이터 품질 문제 분석, 데이터 품질 특성 식별이 있다.
 - 데이터 품질 측정 기준 정의 : 이 프로세스는 데이터 품질 문제 식별 프로세스에서 식별된 데이터 품질 특성과 관련된 대상 데이터를 결정하고 데이터 품질을 측정하기 위한 측정기준과 방법을 결정하기 위한 것이다. 주요 결과물은 측정할 대상 데이터의 범위

정의, 데이터 품질 특성 관련 측정 지표 정의, 측정 지표 관련 측정 방법 정의가 있으며 세부 활동으로는 측정할 데이터 결정, 측정 지표 개발, 측정 방법 개발이 있다.

- 데이터 품질 측정 : 이 프로세스는 측정 요구사항에 적합한 자원을 구성하고 측정 기준에 따른 데이터 품질 수준을 측정하기 위한 것이다. 주요 결과물은 데이터 품질 측정에 대한 IT 환경의 적용 또는 확장, 측정 계획 수립, 측정 결과 도출이 있으며 세부 활동으로는 데이터 품질 측정 자원 설정, 데이터 품질 수준 측정이 있다.
 - 데이터 품질 측정 결과 평가 : 이 프로세스는 측정 결과를 바탕으로 데이터의 오류를 분석하여 데이터의 품질 수준과 데이터 품질 통제 프로세스의 성과를 평가하기 위한 것이다. 주요 결과물은 데이터 오류 식별 및 분류, 데이터 품질 특성에 의한 데이터 오류 분석, 분석 결과에 근거하여 데이터 오류의 심각성 또는 데이터 품질 수준 평가가 있으며 세부 활동으로는 데이터 품질 수준 평가, 데이터 품질 통제 프로세스 성과 평가가 있다.
- ④ 데이터 품질 개선(Act) : 이 프로세스 클래스는 데이터 품질평가 프로세스 그룹에서 나타난 데이터 품질평가 결과를 이용하여 데이터 오류의 근본원인 파악하고 동일 유형의 데이터 오류가 재발되지 않도록 관련 데이터를 수정하고 데이터 품질에 영향을 주는 프로세스를 개선하는 프로세스들이다.

- 근본원인 분석 및 해결방안 도출 : 이 프로세스는 데이터 오류의 원인을 분석하고 데이터 오류의 재발을 방지할 수 있는 해결 방안을 도출하기 위한 것이다. 주요 결과물은 데이터 품질 평가 프로세스에서 도출된 결과를 바탕으로 데이터 오류의 근본 원인과 그 영향을 분석, 근본 원인 재발 방지를 위해 프로세스 개선안 제시, 개선의 비용대비 효과 분석, 개선안의 우선순위 결정, 개선 실시 계획 수립이 있으며 세부 활동으로는 데이터 오류의 근본 원인 분석, 근본 원인 제거를 위한 개선안 도출이 있다.
- 데이터 오류 개선 : 이 프로세스는 도출된 프로세스 개선 방안에 따라 데이터 오류에 관련된 모든 데이터를 추적하고 수정하기 위한 것이다. 주요 결과물은 데이터 오류 재발 방지를 위해 상세 수정안 개발, 데이터 오류의 상세 수정안 개발, 이해당사자와 협의하에 데이터 수정 일정계획 개발 및 실시, 데이터에 행해진 수정 기록 유지, 데이터 오류 재발 방지를 위한 측정 개발이 있으며 세부 활동으로는 데이터 오류 및 관련 데이터의 수정, 데이터 오류 재발 방지가 있다.
- 데이터 품질관련 프로세스 개선 : 이 프로세스는 오류재발방지를 위한 상세 프로세스 개선 방안을 도출하고 이해당사자와의 합의를 통해 해당 프로세스의 활동 및 결과물을 개선하기 위한 것이다. 주요 결과물은 데이터 품질관리 프로세스에 대한 상세 개선

안 도출, 데이터 관리 프로세스에 대한 상세 개선안 도출, 비즈니스 프로세스에 대한 상세 개선안 도출, 이해 당사자와 합의된 프로세스 개선안에 대한 실시 일정 계획 및 시행이 있으며 세부 활동으로는 프로세스의 개선, 개선 프로세스의 검증이 있다.

2) 지원 프로세스

이 클래스는 핵심 프로세스 클래스의 지속적인 데이터 개선을 성공적으로 수행할 수 있도록 입력 데이터, 통제 정보 및 지원 정보를 제공한다.

- 데이터 아키텍처 관리 : 이 프로세스는 애플리케이션간에 공유된 데이터의 구조와 의미를 명시하고 이의 사용을 관리하는 것이다. 주요 결과물은 다른 애플리케이션과 조직의 서로 다른 부서가 공유하는 데이터를 정의하는 데이터 모델, 데이터 교환과 공유를 지원하는 각 애플리케이션에 대한 전송 메커니즘과 인터페이스 레이어, 데이터 요구사항 변경을 지원하는 데이터 아키텍처 개선이 있으며 세부 활동으로는 전사적인 공통 데이터의 교환 및 공유, 전사적인 데이터 표준 관리가 있다.
- 데이터 흐름 관리 : 이 프로세스는 애플리케이션 시스템이나 데이터베이스에 배포된 데이터 간의 관계와 흐름을 통제하여 데이터의 무결성을 보장하기 위한 것이다. 주요 결과물은 데이터 흐름의 이력 데이터 저장, 상세 명세에 따른 정기적 데이터 흐름의

모니터링 및 통제가 있으며 세부 활동으로는 데이터 흐름 이력, 데이터 흐름 모니터링 및 통제가 있다.

- 데이터 운용 관리 : 이 프로세스는 구조화된 데이터 자산의 무결성을 확보하여 보장하며 생애주기 동안 데이터의 가용성을 유지하고 데이터 품질 개선의 극대화를 위해 데이터베이스 트랜잭션의 성능을 최적화한다. 주요 결과물은 DBMS 소프트웨어 갱신 및 데이터베이스 연결성 관리를 포함하는 데이터베이스 환경 확보 및 통제, 교환 데이터는 표준화, 데이터 백업 및 복구가 있으며 세부 활동으로는 데이터베이스 지원, 데이터 기술 관리가 있다.
- 데이터 보안 관리 : 이 프로세스는 데이터 접근 권한을 통제하여 허가받은 사람이 데이터에 접근할 수 있게 하는 것이다. 주요 결과물은 데이터 보안을 위한 정책, 표준, 통제 및 절차 정의, 사용자 ID 및 암호 관리, 사용자 데이터 접근 모니터링 및 로그기록 유지, 데이터 보안 실시 성과 평가가 있으며 세부 활동으로는 데이터 보안 기준 설정, 데이터 접근 권한 관리, 데이터 보안 감사가 있다.

3) 조직 프로세스

이 클래스는 핵심 프로세스 및 지원 프로세스 클래스의 수행성과를 향상시키기 위해 전사적 차원에서 자원과 교육 서비스를 제공한다.

- 데이터 품질 조직 관리 : 이 프로세스는 데이터의 품질 향상을 위해 조직

내의 데이터 품질 담당 부서, 위원회, 팀을 운영하는 것이다. 주요 결과물은 데이터 품질관리 지원 부서운영, 데이터 품질관리 프로세스 수행, 데이터 품질관리 프로세스 활동에서 생성된 데이터 및 문서 유지관리가 있으며 세부 활동으로는 데이터 품질 조직 운영, 데이터 품질 관련 데이터 및 문서관리가 있다.

- **인력 교육 관리** : 이 프로세스는 데이터를 만들어 내는 사람과 사용자에게 데이터 품질의 중요성을 교육시킴으로써 데이터 품질에 대한 자신들의 책임성을 인식시키고 업무를 효율적으로 수행하여 데이터 품질관리에 기여할 수 있게 한다. 주요 결과물은 데이터 품질관리에 필요한 지식 및 역량 개발, 습득, 지원, 인력의 지식 및 역량 개발, 유지 또는 강화, 필수 데이터 품질관리 지식 및 역량을 지닌 인력을 조직에 제공, 전사적 차원에서 베스트 프랙티스 식별, 공유, 재사용, 강화가 있으며 세부 활동으로는 데이터 품질 역량 제공, 데이터 품질 역량 교육, 지식 관리가 있다.

5. 프로세스 참조모델의 특징 및 기존 프로세스 모델과의 비교

본 논문에서는 데이터 품질관리 프로세스 성숙도를 평가하기 위해 필요한 프로세스 참조모델에 대한 기본적인 개념과 구조를 제시하였고 각 프로세스를 정의하였다. 이 모델의

특징은 다음과 같다 :

첫째로, ISO 12207 : 2002 소프트웨어 생명주기 프로세스에서 제시된 핵심(primary), 지원(support), 조직(organizational) 프로세스 분류 방식과, ISO 9001에서 제시하고 있는 PDCA 개념을 활용하여 데이터 품질관리 프로세스에 적합하도록 새롭게 개발되었다는 점이다. 즉 핵심 프로세스는 데이터 품질관리를 지속적으로 실행하는 front-end 프로세스들을 PDCA 순서로 구성하였으며, 지원 프로세스는 핵심 프로세스가 원활히 실행될 수 있도록 데이터 품질관련 정보와 기술을 제공하는 back-end 프로세스들로 구성되어 있다. 조직 프로세스는 핵심 프로세스와 지원프로세스가 효율적으로 수행되도록 전사적 차원의 조직, 예산, 인력, 자료, 교육 등과 같은 자원 및 서비스를 제공하고 관리하는 back-end 프로세스로 구성되어 있다.

둘째로, 본 프로세스 참조모델은 기존의 데이터 품질관리 프로세스 모델[1, 2] 및 ISO 8000-150 데이터 품질관리 프레임워크[13]에서 정의된 프로세스가 데이터 품질관리 프로세스의 성숙도를 평가하기에 충분하지 못한 문제점을 보완하였다.

참고문헌[1, 2] 및 ISO 8000-150에서는 역할을 3개의 레벨로 나누고 프로세스를 역할별로 구분하여 3×3매트릭스 형태로 프레임워크를 구성하였다. 이 프레임워크는 매트릭스 형태로 프로세스를 단순화 하여 3개의 역할에 모두 공통적으로 필요한 부분들을 고려하지 못한 단점이 있다. 본 논문에서 제안한 프로세스 참조모델은 이 프레임워크에서 고려하지 못한 프로세스들을 추가로 포함하고 있다. 여기서는 프로세스 참조모델과 ISO 8000-

150의 프로세스를 비교하여 개선된 사항을 설명하였으며 그 내용이 <Table 2>에 요약되어 있다.

우선, 프로세스 참조모델에서는 핵심 프로세스들이 ISO 8000-150에 비해 더 확대되거나 세분화 되어 14개의 프로세스로 명시되었다. 데이터 품질계획 프로세스 클래스에는 데이터 품질 요구사항 관리 프로세스가 추가되었다. 데이터 품질 전략을 수립하기 위해서는 데이터 사용자의 정확한 요구사항을 수집 및 분석하고 그 결과를 전략에 반영하는 프로세스가 필요하다. 데이터 품질 전략 관리는 ISO 8000-150의 데이터 품질계획 내 통일된 목표 설정 및 관리 활동뿐만 아니라 데이터 품질 전략, 중장기 목표, 로드맵 등 중장기적으로 데이터 품질을 지속 관리하기 위한 전략과 목표를 포함하고 있다. 그리고 데이터 품질 확보계획을 수립하기 위해서는 전사적으로 사용되는 정책, 표준, 절차를 사전에 결정하고 관리하는 단계가 필요하다. 이 기준이 있어야 데이터 품질 통제 프로세스와 데이터 품질 개선 프로세스에서 혼란을 방지할 수 있다. ISO 8000-150의 데이터 권한 관리 부분은 프로세스 참조모델에서는 데이터 품질 확보계획 관리의 한 활동(activity)인 스텔어드쉽 배정으로 정의되었다. 스텔어드쉽 배정은 데이터 품질을 통제하거나 개선하는데 반드시 필요하므로 거버넌스 레벨에 해당되는 데이터 품질 계획 수립시 수행되어야 바람직하다.

데이터 품질 통제 프로세스 클래스에서는 사용자들이 데이터를 처리하고 모니터링 및 통제하기 위해 데이터 품질 명세와 데이터 처리지침을 제공하는 프로세스가 추가되었다. 이 기준을 제시함으로써 데이터 사용 및

처리시 혼란을 방지할 수 있다. 또한, 이 프로세스 클래스에서는 데이터 품질 모니터링 및 통제 프로세스가 추가되었는데 이것은 데이터 오류에 대한 사후 조치와 예방 활동을 모두 포함하고 있다. ISO 8000-150에서는 데이터 오류 원인 분석 프로세스 내 데이터 오류 예방 활동으로 정의되어 있으나 이 활동은 일상적으로 수행되는 조치 중의 하나이므로 데이터 품질 모니터링 및 통제 내에 포함되어야 바람직하다.

데이터 품질 심사 프로세스 클래스에서는 데이터 품질 통제 프로세스 클래스에서 해결되지 않은 문제점에 대한 정확한 정의를 내리고 측정할 품질 특성을 식별하는 프로세스가 추가되었다. 이 프로세스를 통하여 품질측정기준을 명확하게 결정할 수 있게 된다. 그리고 데이터 품질 측정결과 평가 프로세스가 추가되었는데 여기서는 데이터 품질 측정 후 데이터 품질 개선 조치를 취할지 여부를 판단하게 된다.

데이터 품질개선 프로세스 클래스에서는 데이터 품질관련 프로세스 개선 프로세스가 추가되었다. 이 프로세스는 ISO 8000-150에서 제시한 내용보다 더 포괄적인 내용을 제시하고 있다. 이것은 데이터 자체의 오류를 개선하는 것만으로는 데이터의 정합성을 유지하기 어려우므로 데이터 품질에 영향을 주는 모든 프로세스, 즉, 데이터 품질관리 프로세스, 데이터 관리 프로세스, 비즈니스 프로세스를 개선하는 것이다. 이 프로세스 개선에서는 프로세스의 결과물, 즉, 데이터 설계, 소프트웨어 어플리케이션 및 아키텍처, 비즈니스 아키텍처 등을 개선할 수 있다.

그리고, 지원 프로세스 클래스에서는 ISO

<Table 2> A Comparison of Processes in the Process Reference Model and ISO 8000-150

Processes in process reference model		Processes in ISO 8000-150	Comparison	Improvements	
Primary Process	DQ Planning	DQ Requirement Management		added	Possible to collect data user's requirements
		DQ Strategy Management	DQ planning : (A) Objectives setup and management	expanded	Possible to perform continuous quality management with the addition of the strategy, long/short-term objectives and the road map
		DQ Policy/Standards/Procedures Management		added	Criteria and procedures can be set up clearly to maintain DQ.
		DQ Implementation Planning	- DQ planning : (A) Identification of quality management items/action planning - Data stewardship management	expanded	The stewardship management is included in this process since it corresponds to one of DQ governance functions.
	DQ Control	Definition of DQ Specifications and Guidelines		added	Possible to prevent the confusion of DQ control by the provision of detailed DQ specifications and processing guidelines.
		Data Processing	Data Processing		
		DQ Monitoring and Control	Data error cause analysis : (A) Error recurrence prevention	expanded	Activities of data nonconformity correction and recurrence prevention are merged to this process.
	DQ Assessment	Identification of DQ Problems		added	Possible to determine the measurement criteria with the prior investigation of DQ problems.
		Definition of DQ Measurement Criteria	DQ criteria setup		
		DQ Measurement	DQ Measurement		
		Evaluation of DQ Measurement Results		added	Decide whether to perform DQ improvement after DQ measurement.
	DQ improvement	Root Cause Analysis and Resolution Derivation	Data error cause analysis		
		Data Nonconformity Correction	Data error correction		
		Process Improvement for Data Nonconformity Prevention	-Data error cause analysis: (A) Error recurrence prevention -Data design	expanded	Outcomes of the process can be improved by the improvement of DQ processes, data management processes, and business processes in detail.
Supporting Process	Data Architecture Management	Data Architecture Management			
	Data Transfer Management	Data stewardship(/flow) management			
	Data Operations Management		added	Possible to improve the efficiency of data operations.	
	Data Security Management		added	Possible to enforce DQ stability	
Organizational Process	DQ Organization Management		added	Possible to operate organizational units, committees, and teams for DQ, and administrate data and documents generated in relation with DQ.	
	Human Resource Training Management		added	Possible to improve the efficiency of data processing by educating data users about the importance of DQ and processing guidelines and methods.	

(A) : An activity within a process. DQ : data quality.

8000-150에는 고려되지 않은 데이터 운용관리, 데이터 보안관리가 추가되었다. 데이터 운용관리 프로세스는 데이터 교환, 백업 및 리커버리, 데이터베이스 효율화를 위한 데이터베이스 지원과, 데이터 품질관련 소프트웨어 및 툴을 제공하는 데이터 기술 관리 활동으로 구성되어 있다. 이 프로세스를 통하여 안정적인 데이터 제공, 빠른 데이터 검색, 신속한 데이터 교환 등의 효율성 높은 데이터 품질 서비스를 제공할 수 있다. 데이터 보안관리 프로세스는 데이터 접근 기준과 권한 통제, 데이터 보안 평가를 위한 감사 활동으로 구성되어 있다. 이 프로세스를 통하여 데이터 품질을 안전하게 유지할 수 있다.

조직 프로세스 클래스도 ISO 8000-150에는 고려되지 않은 것으로서 데이터 품질 조직 관리, 인력 교육 관리가 추가되었다. 데이터 품질 조직관리 프로세스를 통해서서는 담당 부서, 위원회 또는 팀을 운영하고 생성되는 데이터와 문서를 행정적으로 관리하게 된다. 인력 교육 관리 프로세스를 통해서서는 데이터를 다루는 모든 사용자와 데이터 품질관련 인력에게 데이터 품질의 중요성, 처리 지침 및 방법 등을 교육하여 전사적 차원에서 데이터 처리 효율을 높일 수 있다.

이 프로세스들은 데이터 품질관리를 위해 조직의 전사적 차원에서 자원을 지원하는 것으로서 핵심 프로세스 클래스 뿐만 아니라 지원 프로세스 클래스 지원하게 된다.

이 외에도, 한국에서 성숙도 평가 모델로 활용되는 DQM3[8]와 비교했을 때 이 프로세스 참조모델은 핵심프로세스 클래스가 더 포괄적으로 상세히 제시되었으며 조직 프로세

스 클래스의 데이터 품질 조직 관리, 인력 교육 관리가 추가되었다. IQM3[10]와 비교했을 때 데이터 흐름관리, 데이터 보안관리, 인력 교육관리가 추가된 특징이 있다. 본 모델을 개발하기 위해 참조하였던 다른 모델들은 프로세스의 생성관점, 깊이, 또는 범위가 달라서 여기서는 비교를 생략하였다.

6. 결 론

이 모델은 현재 개발 중인 ISO 8000(데이터 품질) Part 60시리즈(데이터 품질관리 프로세스 평가) 중에서 Part 61(데이터 품질관리 프로세스 참조모델)의 주 내용을 제시하고 있다. 본 모델은 향후 개발될 Part 62(데이터 품질관리 프로세스 성숙 모델), Part 63(데이터 품질관리 프로세스 측정 모델), Part 64(데이터 품질관리 프로세스 개선 모델)에 대한 프로세스 참조모델로 활용될 예정이다.

이 프로세스 참조모델에서는 데이터베이스에 저장된 정형화 데이터뿐만 아니라 CAD 파일, 전자문서 등 디지털화 된 비정형 데이터를 포함하는 일반 데이터의 품질을 관리할 수 있도록 프로세스를 포괄적으로 정의하였다. 이 모델은 데이터를 공유하고 교환해야 하는 복수의 응용 시스템을 가진 조직에서 전사적 차원에서 데이터 품질을 관리하려는 경우에 적용될 수 있다. 또한, 이 모델은 프로세스 성숙도를 측정하고 프로세스 개선을 통해 데이터 품질을 향상시키고자 하는 내부 부서에서만 아니라 인증기관과 같은 외부기관에서도 활용할 수 있다.

References

- [1] Kim, S. H., Lee, C. S., "A Master Data Quality Management Framework," *Entrue Journal of Information Technology*, Vol. 9, No. 2, pp. 109-121, 2010.
- [2] Lee, C. S., Kim, S. H., "A Data Quality Management Framework and its Business Scenario for Application," *The Journal of Society for e-Business Studies*, Vol. 15, No. 4, pp. 79-99, 2010.
- [3] CMMI for Acquisition, V.1.3, Carnegie Mellon, Software Engineering Institute, 2010.
- [4] CMMI for Development, V.1.3, Carnegie Mellon, Software Engineering Institute, 2010.
- [5] CMMI for Services, V.1.3, Carnegie Mellon, Software Engineering Institute, 2010.
- [6] Data Documentation Initiative(DDI) Technical Specification, Part I : Overview, Version 3.1, DDI Alliance, 2009.
- [7] Federal DAS-Data Quality Framework V1.0, US Federal Data Architecture Subcommittee(DAS), October 1, 2008.
- [8] Korea Database Agency, Data quality management maturity model(V1.0), 2006.
- [9] IBM, IBM Data Governance Council Maturity Model, 2007.
- [10] Caballero, I., Caro, A., Calero, C., Piattini, M., "IQMB : Information Quality, Management Maturity Model," *Journal of Universal Computer Science* Vol. 14, No. 22, pp. 3658-3685, 2008.
- [11] ISO 8000-1 Data quality -Part1 : Overview, ISO, 2009.
- [12] ISO 8000-110 Information quality -Part 110 : Master data quality : Syntax, semantic encoding, and conformance to customer requirements, 2008.
- [13] ISO/TS 8000-150 Master data : Quality management framework, ISO, 2011.
- [14] ISO 9000, Quality management systems -Fundamentals and vocabulary.
- [15] ISO TC211 Geographic Information/Geomatics, www.isotc211.org.
- [16] ISO/CD 10303-59 Product data representation and exchange : Integrated generic resource : Quality of product shape data, ISO TC184/SC4/WG12 N4866, 2007.
- [17] ISO/IEC 12207 Systems and software engineering -Software life cycle processes, ISO, 2002.
- [18] ISO/IEC 15288 System Engineering-System Life Cycle Processes, ISO, 2008.
- [19] ISO/IEC 15504-1 Information technology -Process assessment -Part 1 : Concepts and vocabulary, ISO, 2004.
- [20] ISO/IEC 25012 Software engineering -- Software product Quality Requirements and Evaluation(SQuaRE) -- Data quality model, ISO, 2008.
- [21] ISO/IEC 33000 series <http://www.spilab.co.za/iso-standards-watch/33-33000>.
- [22] ISO/IEC TR 24774, Systems and software engineering -Life cycle management -

- Guidelines for process description, 2010.
- [23] Mielke, M., Gebauer, M., Lüssem, Kutsche, R., Information Quality Management : Principles and Foundations, EIDIQ-SP-1-2008, EIDIQ, 2008.
- [24] Mosley, M., Brackett, M., Earley S., Henderson D., The DAMA Guide to the Data Management Body of Knowledge, DAMA International, 2009.
- [25] OASIS Customer Information Quality TC, Customer Information Quality Specifications Version 3.0 : Name(xNL), Address (xAL) and Party(xPIL), Public Review Draft, 2007.
- [26] Pipino, L. L., Lee, Y. W., Wang R. Y., "Data quality assessment", Communications of the ACM, Vol. 45, No. 4, pp. 211-218, 2002.
- [27] Ryu, K. S., Park, J. S., Park, J. H., "A data quality management maturity model," ETRI Journal, Vol. 28, No. 2, 2006.

저 자 소 개



김선호

(E-mail : shk@mju.ac.kr)

명지대학교 산업경영공학과 교수

Pennsylvania State University 산업공학과 (석·박사)

서울대학교 산업공학과 (학사)

관심분야

Data Quality, 전자거래 표준화, BPM, 분야

클라우드 컴퓨팅, 정책평가



이창수

(E-mail : cslee@gwnu.ac.kr)

강릉원주대학교 산업정보경영공학과 교수

서울대학교 산업공학과 (석·박사)

서울대학교 산업공학과 (학사)

관심분야

데이터 품질, 마케팅정보시스템, 검색엔진마케팅