

공공데이터 품질관리를 위한 조직 성숙도 평가 모델*

김 선 호**, 이 창 수***, 정 승 호****, 김 학 철*****, 이 창 수*****

요약 정부 3.0의 확산으로 공공 데이터의 활용요구가 증대되고 있으나, 정부가 보유하고 있는 공공데이터의 품질 및 관리체계는 아직 성숙화되어 있지 않아 데이터 개방 및 활용의 저해 요소로 부각되고 있다. 데이터 관리 및 연계·활용의 효과와 효율성을 향상시키기 위해서는 데이터 품질관리에 관한 표준 마련과 이를 평가할 수 있는 평가기준의 마련이 필요하며, 지속적인 품질관리가 이루어질 수 있도록 수준을 측정할 수 있는 체계의 마련이 필요하다. 본 연구는 공공데이터 품질관리 수준을 평가하고, 품질관리 수준을 지속적으로 판단할 수 있는 데이터 품질관리 성숙도 모델을 제시한다. 성숙도 모델은 프로세스 참조 모델과 측정 프레임워크로 구성하였다. 프로세스 참조 모델은 PDCA 기반의 15개 프로세스를 정립하였으며 이를 토대로 프로세스의 능력 수준과 데이터 품질관리 조직의 성숙도 수준을 평가하는 프레임워크를 마련하였다. 본 연구에서 제시한 성숙도 모델은 향후 공공기관이 보유하고 있는 데이터의 품질관리 현 수준을 진단하고 문제점을 개선하는 등 보유 공공데이터의 지속적인 품질향상을 위한 단계별 목표와 방향성을 수립하는 데 활용할 수 있으며, 궁극적으로는 공공데이터의 품질 신뢰도 향상을 통한 안정적인 공공데이터의 민간 개방을 촉진하고 이용을 활성화하는 데 적극 기여할 것으로 판단된다.

주제어: 공공 데이터 품질관리, 프로세스 참조모델, 프로세스 능력, 조직 성숙도, 측정 프레임워크

An Organizational Maturity Assessment Model for Public Data Quality Management

Sunho Kim, Changsoo Lee, Seungho Chung, Hakcheol Kim, Changsoo Lee

Abstract Although the demand for the use of public data increases in accordance with the expansion of Government 3.0, the poor level of data quality and its management currently implemented is becoming obstacles to opening data to the public. To improve the efficiency of management, linkage and usage for data, standardized processes for data quality management have to be prepared and appropriate data quality assessment criteria should be established. In this paper, we propose the organizational maturity model that can assess the public data quality management level. This model consists of the process reference model and the measurement framework. Fifteen processes grouped by the PDCA cycle are defined in the process reference model. The measurement framework measures the organizational maturity level based on process capability levels. The organizational maturity model can be used to establish objectives and directions for public data quality improvement by diagnosis of current level of public data quality management and problem solving. This model can also facilitate open to the private sector and activate usage of stable public data through reliability enhancement.

Keywords: public data quality management, process reference model, process capability, organizational maturity, measurement framework

2014년 11월 17일 접수, 2014년 11월 20일 심사, 2014년 12월 19일 게재확정

* 본 논문은 한국정보화진흥원이 지원한 "2014년 공공데이터활용지원센터 품질관리 기술지원용역"의 연구 결과로 수행되었음

** 명지대학교 산업경영공학과 교수(shk@mju.ac.kr)

*** 강릉원주대학교 산업정보경영공학과 교수(cslee@gwnu.ac.kr)

**** 한국정보화진흥원 수석연구원(jsh@nia.or.kr)

***** 투이건설링 수석연구원(danny68@gmail.com)

***** 명지대학교 책임연구원(calspdm@naver.com)

I. 개요

정부에서는 기관이 보유하고 있는 공공데이터의 개방정책을 수립하고, 이를 기반으로 기관이 보유하고 있는 정보와 데이터를 공개하는 정부 3.0을 추진하고 있다. 이러한 정부 정책의 효율적 추진을 위해 정부에서는 민간이 공공 데이터를 보다 편리하게 활용할 수 있도록 지원하는 단일창구(www.gov30.gov.kr)를 마련하여 운영 중에 있으며, 기관이 보유하고 있는 데이터의 품질을 일정수준 유지할 수 있도록 품질진단 및 개선사업을 진행 중에 있다.

2013년을 기준으로 전체 35개의 공공데이터에 대한 품질진단 및 개선 사업이 진행되고 있으며, 향후에도 지속적인 개선사업이 진행될 예정이다(www.data.go.kr).

공공기관이 보유하고 있는 데이터의 개방 및 활용은 국내뿐만 아니라 영국, 미국, 유럽위원회 등 여러 국가에서 이미 진행되고 있다.

먼저 영국에서는 국민에게 다양한 서비스를 제공하고 국민의 편익 향상을 위해 필요한 데이터의 접근성 개선과 유용성 향상을 위해 개방데이터 백서(HM Government)를 2012년에 마련하였으며, 정부 데이터 포털(data.gov.uk)을 마련하여 약 1만 5천여 개의 데이터 세트(2014년 10월 기준)를 민간이 활용할 수 있도록 공개하고 있다.

한편, 미국에서는 모든 부서가 보유하고 있는 데이터를 자산으로 관리할 것을 Open Government Plan(<http://www.state.gov/open/>)에 포함시켰으며, 여기에 포함된 개방 데이터 정책은 비용감소와 함께 운영의 효율성 증대, 서비스 개선, 정보에 대한 접근성 제고를 목표로 하고 있다. 이와 함께 미국무부 정보자원관리국(Bureau of Information Resource Management)은 개방 데이터 계획(US Department of State, 2014)을 통해 분기별로 개방할 데이터세트 건수와 목록을 공개하고 공개된 데이터세트는 오너를 지정하여 갱신과 유지관리에 대

한 책임제를 운영하는 등 보다 체계적인 데이터 품질 관리 정책을 추진하고 있는 상황이다.

이외에도 유럽위원회에서는 유럽 개방 데이터 전략(Open Data Strategy for Europe)을 통해 공공 부분이 보유하고 있는 정보를 재사용함으로써 가치를 창조할 것을 강조하였다. 또한 데이터 개방은 새롭고 혁신적인 솔루션을 제공할 수 있으며 데이터 공유를 통해 효율성 향상 및 정부의 투명성 제고 그리고 시민의 사회 참여가 가능하다고 제시하였다(<http://ec.europa.eu/digital-agenda/en/open-data-0>).

이와 같이 선진 각국에서는 공공 기관이 보유하고 있는 데이터를 가능한 많이 공개하기 위해 제도적 장치를 마련함과 동시에 공개된 데이터를 사용할 수 있는 기술적인 방법을 제공하여 공공데이터의 민간 활용 우수 사례를 지속적으로 발굴하고 있으나, 국내의 경우 데이터의 품질 수준 및 관리체계가 아직 미흡한 것이 현실이다.

'13년 1·2차 공공정보 품질관리 지원 사업(2013. 11)의 중간보고 결과에 따르면 국내 공공기관이 보유하고 있는 데이터 및 데이터베이스 품질 수준은 민간 기업에 비해 상대적으로 낮은 수준이며, 이는 다음과 같은 원인 때문인 것으로 나타났다.

첫째, 인력과 조직 측면에서는 전담인력의 수가 부족하고 전담인력의 전문성이 부족하며 외부 업체와 전문가에 대한 의존도가 지나치게 높고 둘째, 기관마다 데이터 연계 표준·품질이 상이하여 데이터의 신뢰도가 낮으며, 마지막으로 데이터 품질 표준·관리 체계 부재로 비효율적인 데이터 품질관리가 이루어지고 있기 때문인 것으로 분석되었다.

이는 효과적이고 신뢰도 높은 데이터의 확보를 위해서는 데이터 품질 표준 및 관리체계의 마련과 전담인력 등의 마련이 절실히 요구되며, 이를 지속적으로 모니터링할 수 있는 체계의 마련이 필요하다는 것을 의미한다고 할 수 있다.

Redman(1998)은 적정 수준의 데이터 품질을 보

유하고 있지 못할 경우 조직에 부정적 영향을 미칠 뿐만 아니라 비용의 손실도 유발될 수 있음을 제시하였고, Haug, et al.(2011)은 빈약한 데이터 품질로 유발되는 손실과 품질보장을 위해 투입되는 비용의 관계를 제시하고 이들 둘 사이의 균형이 필요함을 강조하였다.

데이터 품질을 유지하기 위해 투입되는 비용은 통상 내부인건비, 컨설팅 비용, 소프트웨어 비용 등이나, 낮은 수준의 데이터 품질로 야기되는 손실의 경우 구체적으로 측정이 어려운 것이 사실이다. 그렇기 때문에 데이터 품질에 투입되는 비용은 데이터의 품질 수준을 어느 정도로 유지할 것인지를 결정하는 것이 무엇보다 중요하고, 이를 위해서는 현재 데이터 품질 수준에 대한 진단과 목표수준을 선정하고 목표 수준의 달성을 위해 필요한 요소가 무엇인지 판단하는 것이 필요하다(Eppler, et al., 2004; Sadiq, 2013).

특히, 공공부문의 경우 현재 수준에 대한 진단 및 목표수준의 선정이 예산확보 등을 위한 주요한 지표로 활용될 수 있기 때문에 이를 체계적으로 진단하고 가시적인 목표수준을 선정할 수 있는 기준의 마련이 필요하다. 정승호 외(2013)는 이러한 공공부문의 데이터 품질문제를 고려하여 공공기관이 일반 사용자에게 공개하는 공공정보의 품질을 객관적으로 보장하는 것이 필요함을 제시하고, 이를 위해 공공정보 품질관리에 관한 성숙도를 평가할 수 있는 체계의 마련이 필요함을 제시하기도 하였다.

본 연구는 공공기관이 적정 수준의 품질을 보유한 공공데이터를 국민에게 제공하고, 공공데이터의 활용이 보다 활성화 될 수 있도록 지원하기 위해 기관이 보유하고 있는 데이터 품질을 진단하고 미래 목표를 설정할 수 있는 공공데이터 품질관리 성숙도 모델을 마련하였다.

이를 위해 II장에서 성숙도 모델에 관한 기존 연구 동향을 파악하였고, III장에서 성숙도 모델을 도출하기 위한 방법론 및 성숙도 모델 도출 과정을 제시하

였다. IV장에서는 공공데이터 품질관리를 위한 조직 성숙도 평가 모델을 구성하고 있는 참조 모델과 측정 프레임워크를 설명하였다. 먼저, 공공데이터 품질관리를 위한 기본 구조와 구성체계를 제시하고, 15개 공공데이터 품질관리 프로세스로 구성하는 참조 모델을 제시하였다. 참조 모델에서 제시한 15개 프로세스는 각 프로세스의 목표, 목표 달성에 필요한 활동, 활동의 결과로 나타나는 작업산출물 등을 포함하고 있다. 측정 프레임워크에서는 공공데이터 품질관리 프로세스의 능력 수준과 품질관리 조직의 성숙도 수준을 단계화하여 정의하고, 성숙도 수준별로 달성 기준을 제시하였다. V장에서는 본 연구에서 제시된 데이터 품질관리 조직 성숙도 모델의 검증을 위한 방법으로 전문가 의견을 수렴하고 프로세스 참조 모델에 대한 내용타당도 분석을 실시하였다. 마지막으로 VI장에서는 본 연구에서 개발된 성숙도 모델에 대한 의미와 향후 연구 방향을 결론으로 제시하였다

II. 성숙도 모델 관련 연구 동향

성숙도와 관련된 모델은 대체로 소프트웨어 품질에 대한 연구가 주류를 이루고 있다. 본 연구에서는 소프트웨어 성숙도 관련 모델뿐만 아니라 데이터 품질에 대한 평가 모델 등 다양한 관련 연구를 종합하여, 공공부문에 적합한 데이터 품질관리를 위한 성숙도 모델 수립의 근거로 활용하였다. 기존 연구 동향 분석은 공공부문의 성숙도 모델, 데이터, 제품, 소프트웨어 분야로 구분하여 품질, 프로세스, 성숙도, 국제 표준 관점에서 분석하였다.

1. 공공 부분의 성숙도 모델

EA(Enterprise Architecture) 성숙도 측정모델은 공공기관의 EA 활용 수준을 높이고 정책적 방향성 확립을 위해 개발된 성숙 모델로서 자원관리, 연계, 통합지원, 활용 성과 등 EA 관리 및 활동을 측정

영역으로 구분하고 각 측정 영역 별로 5단계의 성숙 수준을 정의하였고 세부적으로 산술 평균을 제시하고 있다(안전행정부, 2014).

한국정보화진흥원에서 국내 기업의 녹색정보화 수준을 진단하기 위하여 글로벌 IT 컨설팅 기업인 액센츄어(Accenture)의 녹색성숙도 모델을 기초로 국내 현실을 고려한 녹색정보화 수준 진단 지표 및 성숙도 모형을 개발하였다. 여기에는 액센츄어가 보유한 300여개 지표와 영국 내각부의 그린 ICT 스코어 카드 및 호주의 IT 컨설팅 기관인 커빅션 연구소의 그린 IT 준비지수의 지표를 분석하고 통합하여 66개 후보 지표를 제시하였다(박상현, 2011).

2. 데이터 분야

ISO 8000은 데이터 품질에 관한 국제 표준으로 ISO 8000-6x 시리즈에서 데이터 품질관리 프로세스의 평가와 개선 모델에 관한 표준을 다루고 있다. ISO 8000-60은 60시리즈의 전체 개념을 다루고 있으며, 시리즈 내의 ISO 8000-61은 프로세스 참조 모델을 정의하며 데이터 품질 관리에 필요한 프로세스를 정의하고 있다. 각각의 프로세스는 데이터 품질 제고를 위해 활용되며, 데이터 품질관리에 대한 프로세스 능력도나 조직 성숙도를 평가하는데 활용된다. 본 연구의 프로세스 참조 모델은 ISO 8000-61의 내용을 근간으로 하고 있다(ISO, 2014a; ISO, 2014b).

김선호 외(2010)와 이창수 외(2010)는 ISO 8000-150(마스터 데이터 품질 관리 프레임워크)을 기반으로 하여 품질조직에 관한 3가지 역할(관리자, 통제자, 실행자)과 3가지 기능(데이터 활용, 품질 모니터링, 품질 개선)을 매트릭스로 구성하여 9개의 품질관리 프로세스를 제안하였고 김선호 외(2013)는 ISO 8000-6x 시리즈의 참조 모델에 관한 프로세스를 제시하기도 하였다.

한국데이터베이스진흥원에서 개발한 데이터 품질

관리 성숙 모델(DQM3: Data Quality Management Maturity Model)은 데이터 품질기준, 데이터 품질관리 프로세스, 데이터 품질관리 성숙수준의 3개 요소로 구성되어 있다(한국데이터베이스진흥센터, 2006).

DQM3는 데이터 품질에 직접적으로 영향을 주는 프로세스를 측정 영역으로 고려하고 있으며, 조직의 데이터 품질 향상을 위한 지속적인 품질관리를 측정 항목으로 설정하고 있다. 이외에도 측정 영역에 대한 개별적인 성숙 수준뿐만 아니라 전사 차원에서 성숙 수준을 상호 비교하고, 목표설정이 가능하다.

DQM3는 데이터 품질의 다양한 측면을 분류·정리하여 데이터 품질 기준을 정확성, 일관성, 유용성, 접근성, 적시성, 보안성의 6개 기준으로 정의하고 있으며, 6개의 품질 기준 향상을 위해 필요하거나 품질 기준에 영향을 주는 8개의 프로세스를 마련하고 있다. 특히, 데이터 품질관리 성숙수준의 측정을 위해 조직의 데이터 품질관리 발전 단계를 고려해 '도입-정형화-통합화-정량화-최적화'의 5단계로 구성되며, 6개의 품질 기준별로 데이터 품질관리에 관한 성숙단계를 정의하고 있다는 특징을 가지고 있다.

국제데이터관리협회(DAMA: Data Management Association)는 데이터 관리의 전략적 목표를 제시하고 목표 달성을 위한 핵심 요인으로 거버넌스, 아키텍처 관리, 개발·운영 관리, 보안관리, 문서관리, 메타관리, 품질관리를 제시하고 있다. 데이터 관리는 계획 및 실행, 정책, 프랙티스, 프로젝트, 획득, 통제, 보호, 전달, 가치 향상, 데이터 및 정보 자산을 주요 내용으로 한다(DAMA International, 2009).

EIDIQ(European Institute for Data and Information Quality)는 데이터와 정보의 관리에 있어서 특정 목표를 실현하기 위한 조직들을 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 개발하여 보급한다

(EIDIQ, 2008).

연방 DAS(Data Architecture Subcommittee) DQF(Data Quality Framework)에서의 데이터 품질은 관심 단체에서 데이터가 사용을 위해 필요한 품질 특성에 직면하는 것을 보장하는 일련의 원칙과 절차로서, DAS DQF는 입증된 방법, 산업 표준과 과거 수행에 기초를 둔 사람, 프로세스와 기술을 위한 접근방법을 정의하고 있다. 각 기관에서 데이터를 활용하는데 있어서 데이터 품질관리가 필수적으로 선행되어야 한다고 보고 미 연방정부의 EA 참조 모델 관점에서 계획, 거버넌스, 관리 문화, 품질 측정 등 데이터 품질관리 개선 프로세스와 기준을 제시하였다(US Federal DAS, 2008).

Caballero, et al.,(2008)은 정보품질관리 성숙모델(IQM3)과 정보품질관리의 평가 및 개선 방법론(MAIMIQ)을 제시하였는데, 여기서 IQM3(Information Quality Management Maturity Model)은 정보품질관리 성숙모델로서 CMMI의 성숙모델을 참조하여 Initial, Defined, Integrated, Quantitatively Managed, Optimizing의 5단계로 정보품질관리 성숙도를 정의하고 있으며, MAIMIQ(A Methodology for Assessing and Improving the Management of Information Quality)는 정보품질관리의 평가 및 개선 방법론으로서 Plan-Do-Check-Act 과정을 거쳐 정보 품질을 개선한다.

IQM3는 CMMI를 기반으로 모델을 수립함으로써 모델 구조나 내용적인 측면의 구성 노력을 줄이고 적용 가능성을 높였다는 특징을 가지고 있으나, 데이터 품질 측면의 핵심 프로세스를 보다 세분화하여 측정 영역을 제시할 필요가 있다.

안전행정부는 공공데이터의 제공 및 이용 활성화에 관한 법률에 근거하여 공공데이터 관리지침(안전행정부, 2014)과 공공정보 품질관리 매뉴얼(한국정보화진흥원, 2014c)을 배포하였으며 여기에는 정보

생명주기를 고려하여 계획단계, 구축단계, 운영단계, 활용단계의 4가지 품질관리 활동을 제시하고 있으며 공공데이터 품질 진단 및 개선 절차로 진단대상 정의, 품질진단 실시, 진단결과 분석, 개선계획 수립, 개선수행 및 품질통제의 6단계를 제시하고 있다.

DQAT(Data Quality Assessment Tool)는 개선된 데이터 품질을 통한 비즈니스 성과 향상 프로젝트의 일부분으로서 영국의 국가 방위 산업 위원회(National Defence Industries Council)의 서브 그룹인 JIG(Joint Information Group)에서 개발하였다. 여기서는 조직이 얼마나 조직의 데이터 품질을 잘 관리하는지를 결정하는 메커니즘을 제공하고 조직이 데이터 품질을 개선하는데 도움을 주기 위해 데이터 품질 평가 도구를 제시하였다(LSC Group, 2012).

CMMI institute는 소프트웨어 프로세스 능력 성숙도 모델을 개발한 경험을 바탕으로 데이터 관리 성숙도(DMM: Data Management Maturity) 모델을 개발하였는데, 이 모델은 핵심 데이터 관리에 대한 20개의 프로세스 영역과 5개의 지원 프로세스 영역으로 구성되어 있으며 인프라 지원 활동을 필요로 한다. DMM은 5개 수준의 기능적 능력 및 성숙도 수준을 사용하며, 능력수준은 달성정도에 따라 비율로 측정되고 성숙도는 전체 프로세스 영역에 대한 능력도의 평균값으로 산정하고 있다(CMMI Institute, 2014).

3. 제품 분야

ISO 9001에서 조직은 고객 만족을 실현하기 위한 제품을 생산하기 위해 PDCA (Plan-Do-Check-Act) 사이클의 개념에 따라 경영책임, 자원관리, 제품구현, 측정·분석·개선의 4가지 프로세스 그룹으로 품질경영 시스템을 구성하고, 품질경영 시스템의 지속적인 개선을 통하여 조직의 수행능력 향상을 유도하고 있다(ISO 9001, 2008).

4. 소프트웨어 분야

ISO 9000-3이 소프트웨어 분야의 특성을 고려하지 못함에 따라 ISO/IEC JTC1 SC7 WG10에서는 소프트웨어 프로세스에 대한 능력 평가와 프로세스 개선을 목적으로 SEI의 CMM, Bell의 TRILLIUM, Esprit의 BootStrap 등을 통합하여 ISO/IEC 15504(SPICE: Software Process Improvement and Capability dEtermination)를 개발하였다(ISO, 2003). 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 심사하

트웨어 생명주기 프로세스(ISO, 2002)를 참조 모델로 하여 프로세스를 평가한다.

시스템 및 소프트웨어 엔지니어링을 위한 시스템의 생명주기 프로세스에 대한 정의는 ISO 15288:2008 (ISO, 2008)에서 정의되고 있다. 2009년부터는 ISO/IEC 15504 표준을 ISO/IEC 33000 시리즈로 개편하기로 하고 ISO 33001~33099 번호를 할당하여 새로운 표준화 작업을 진행하고 있다(ISO, 2014c).

한편, Peldzius, et al.(2010)은 CMMI의 성숙도 수준과 ISO/IEC 15504의 능력 프로파일을 매핑하는 연구에서 CMMI의 고유 프랙티스가 주로 ISO/IEC 15504의 성과(Outcome)와 일부 달성(Achievement) 부분에 대응되고, CMMI의 공통 프랙티스가 ISO/IEC 15504의 달성(Achievement) 부분에 대응된다는 것을 제시하여, CMMI와 ISO/IEC 15504와의 관계를 정의하기도 하였다.

〈표 1〉 연구동향 분석표

분야	분석 대상	관점			
		품질	프로세스	성숙도	국제 표준
공공부분 성숙도 모델	EA성숙도측정모델	○	○	○	
	녹색정보화성숙수준측정모델			○	
데이터	ISO 8000-60 시리즈	○	○		○
	데이터품질관리성숙모델(DQM3)	○	○	○	
	DAMA DMBOK		○		
	EIDIQ IQM-SP	○	○	○	
	Federal DAS DQF	○	○		
	IQM3/MAIMIQ	○	○	○	
	공공데이터 관리지침/공공정보 품질관리 매뉴얼	○	○		
	DQAT	○		○	
제품	DMM	○	○	○	
	ISO 9000 시리즈	○	○		○
소프트웨어	ISO 15504/33000 시리즈	○	○	○	○
	CMMI	○	○	○	○

Institute)에서 개발한 CMMI(Capability Maturity Model Integration)는 소프트웨어 프로세스 능력을 평가하는 도구로 각 수준 별로 핵심적으로 수행되어야 하는 프로세스들을 정의하고 있으며 단계별로 요구되는 모든 프로세스를 수행하고 있으면 해당 수준을 달성한 것으로 판단하고 있다. 그리고 프로세스 능력을 향상시킬 수 있는 방법을 명확히 제시하고 있다. CMMI에서는 22개의 프로세스 영역을 고려하고 있으며, 프로세스의 능력(Process Capability)을 4단계로 평가하고 있으며 조직의 성숙도(Organizational Maturity)는 여러 프로세스의 능력을 종합하여 5단계로 평가하고 있다(CMMI Product Team, 2010).

5. 연구 동향 비교 분석

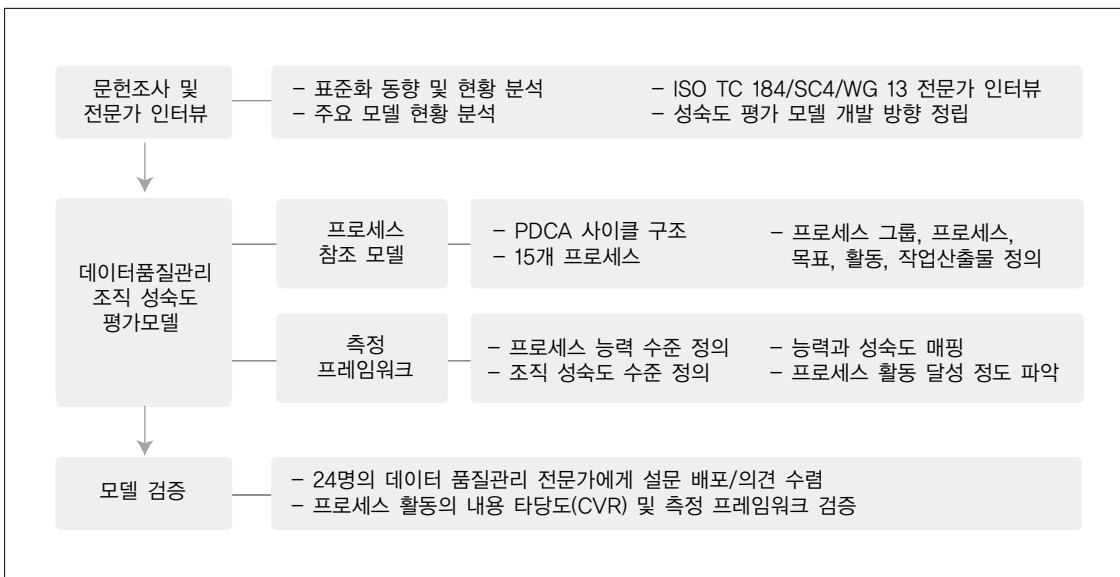
본 연구의 주제와 관련된 연구 동향은 품질, 프로세스 성숙도 등으로 분석 관점을 분류할 수 있다. <표 1>에 기존 연구 동향을 이러한 분석 관점에서 정리하였다.

Ⅲ. 공공데이터 품질관리를 위한 조직 성숙도 평가 모델 도출 방법론

공공데이터 품질관리를 위한 조직 성숙도 평가 모델의 도출은 <그림1>과 같이 문헌 조사 및 전문가 인터뷰, 성숙도 평가모델 정의, 모델 검증의 순으로 진행하였다.

먼저, 문헌 조사는 공공 부문의 유사모델, 데이터 분야, 제품 분야, 소프트웨어 분야로 구분하여 품질, 프로세스, 성숙도, 국제 표준의 관점에서 특징별로 분석 비교하였으며 ISO TC184/SC4/WG13 회원을 중심으로 전문가 인터뷰 및 의견 조사를 실시하였다.

성숙도 평가 모델은 크게 성숙도 평가의 대상이 되는 프로세스 참조 모델과 프로세스 참조 모델의 객관적 측정을 위한 측정 프레임워크로 구성된다. 프로세스 참조 모델은 공공데이터품질관리 매뉴얼과 ISO 8000-61에서 정의된 프로세스를 기반으로 공공기관의 조직 및 보유데이터 특성을 고려하여 새롭게 개발하였으며, 프로세스 기반의 소프트웨어 및 유사 데이터 품질관리 모델(CMMI, DMM, ISO 15504,



<그림 1> 공공데이터 품질관리 조직 성숙도 평가 모델 도출 절차

DQM3, IQM3 등)과 비교하여 프로세스 내용을 프로세스를 보완하였다.

측정 프레임워크는 CMMI의 능력수준과 성숙도 수준의 기본 원칙에 공공데이터 품질관리 프로세스 참조 모델의 각 프로세스를 적용하여 수준별로 그룹화한 뒤 프로세스의 목표와 활동의 달성 정도를 측정 가능하도록 정의하였다.

이를 통해 도출된 프로세스 참조 모델과 측정 프레임워크의 타당성을 판단하기 위하여 본 연구에서는 데이터 품질관리 전문가들을 대상으로 의견을 수렴하고, 제시된 의견을 중심으로 관련 내용을 수정/보완하였다.

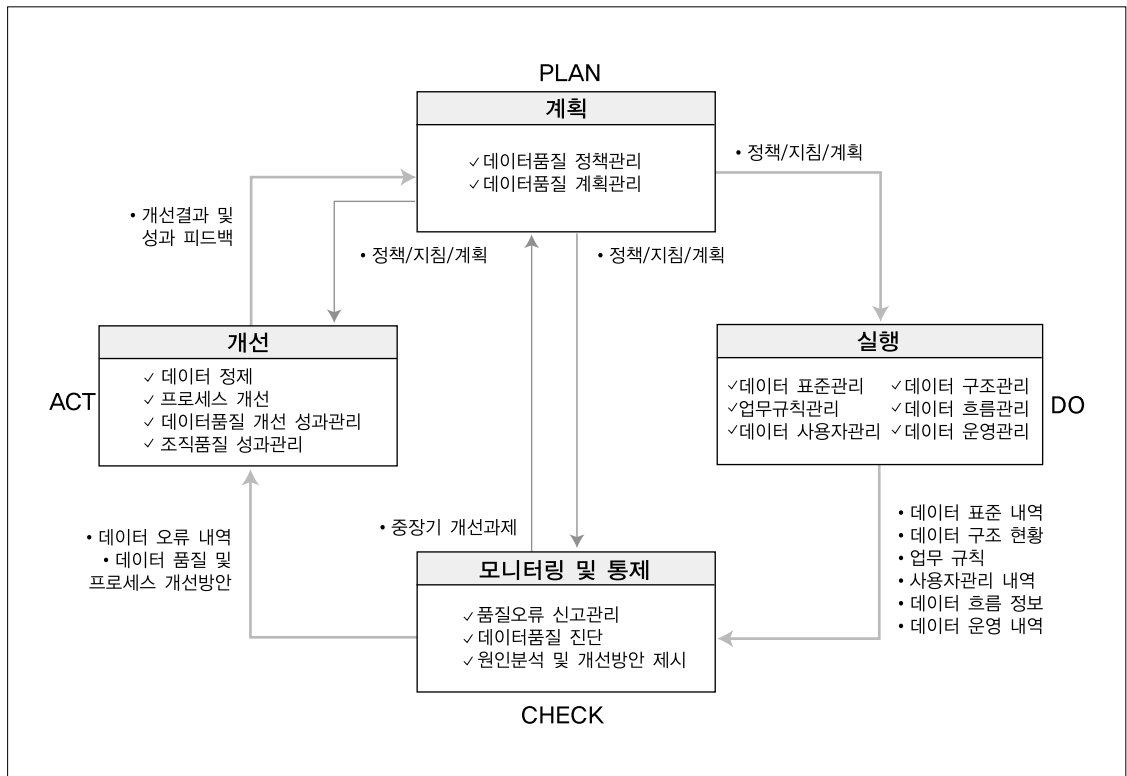
IV. 공공데이터 품질관리를 위한 조직 성숙도 평가 모델

1. 프로세스 참조모델

1) 프로세스 참조 모델의 기본 구조

공공데이터 품질관리 프로세스 참조모델은 Deming의 PDCA (Plan-Do-Check-Act) 사이클 (Deming, 1982; Moen, 2009)의 형태를 참조하여, 데이터 품질관리 계획(Plan), 데이터 품질관리 실행(Do), 데이터 품질관리 모니터링 및 통제(Check), 데이터 품질관리 개선(Act)의 프로세스 그룹으로 <그림 2>와 같이 구분하였다.

PDCA 사이클은 공공데이터 품질관리 프로세스 그룹에 속해 있는 각 프로세스들의 성과를 지속적으로 개선하는데도 적용이 가능하기 때문에 본 연구에서는 공공데이터 품질관리 프로세스는 프로세스 그룹들이 전체적인 순환 구조로 공공데이터의 품질을



<그림 2> 프로세스 참조모델의 PDCA 사이클

지속적으로 개선할 수 있도록 프로세스의 기본구조를 마련하였다.

2) 프로세스 참조 모델의 구성체계

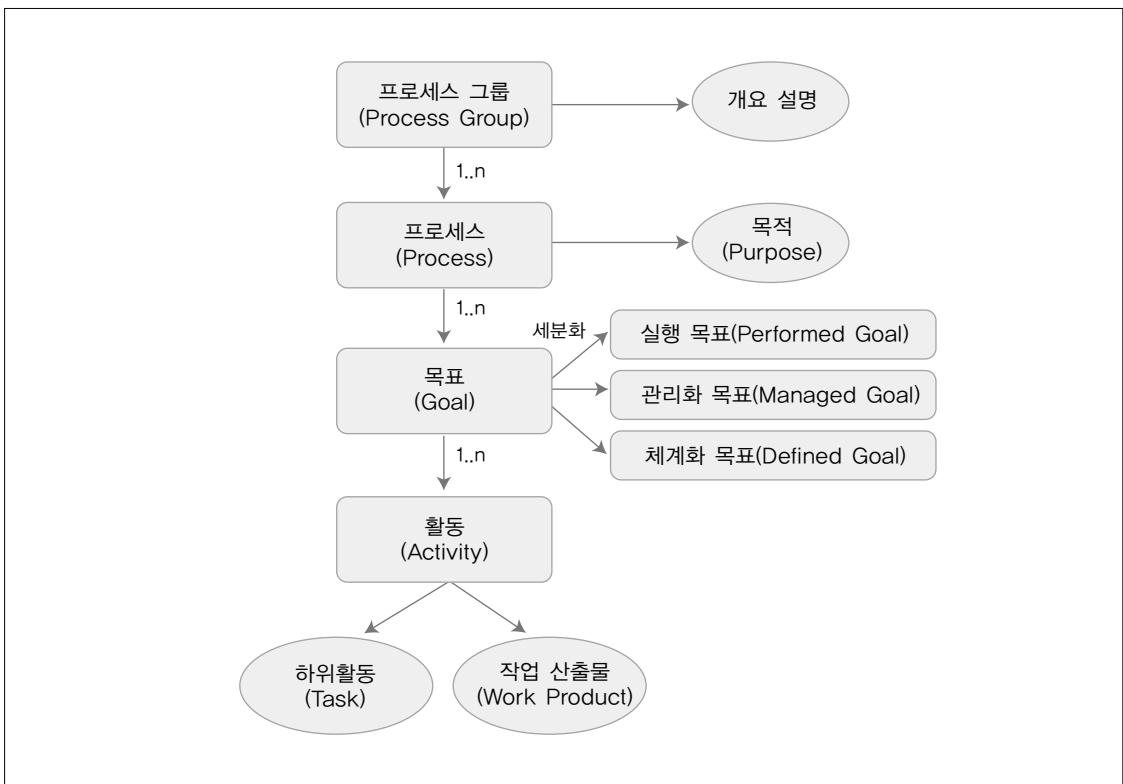
공공데이터 품질관리 프로세스 참조 모델의 각 프로세스는 <그림 3>과 같이 구성되어 있다.

(1) 프로세스 그룹(Process group) : 공공데이터 품질관리 프로세스를 PDCA 사이클에 따라 계획(Plan), 실행(Do), 모니터링 및 통제(Check), 개선(Act)의 4개 프로세스 그룹으로 구분한다.

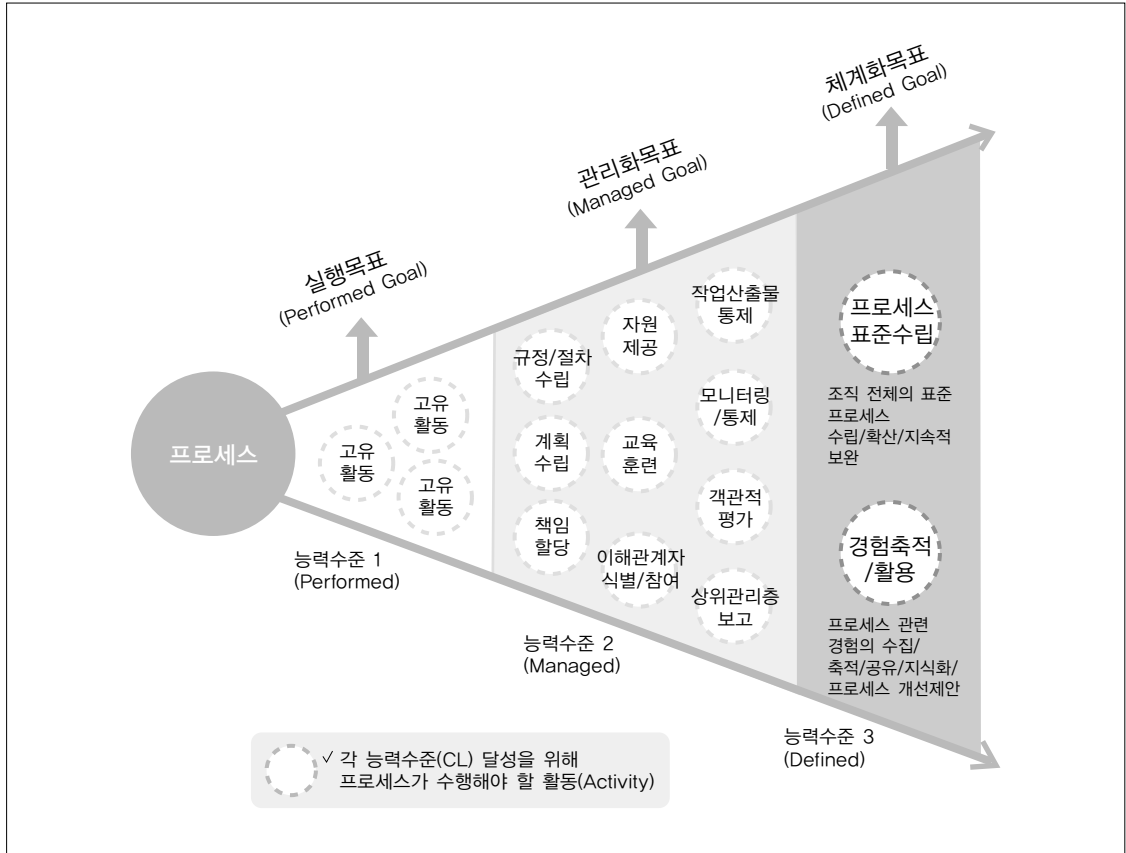
(2) 프로세스(Process) : 프로세스 그룹별로 데이터 품질관리에 필요한 프로세스들을 정의하고 있으며, 각 프로세스별로 목적(purpose)을 제시하고 있다.

(3) 목표(Goal) : 프로세스 수행을 통해 달성 가능한 특성으로서 본 모델에서는 프로세스의 능력 수준에 따라 3개 유형의 목표로 구분된다.

- 실행 목표(Performed Goal) : 각 프로세스의 기본적인 활동이 수행되었을 때 달성 가능한 목표로서 프로세스에서 정의된 고유 활동의 실행을 통해 해당 작업산출물을 생성하고 있는 상태를 의미한다.
- 관리화 목표(Managed Goal) : 각 프로세스가 규정 및 절차에 따라 계획 및 실행되고, 결과물을 생성하기 위해 적절한 자원과 함께 숙련된 인력을 할당하며, 관련 이해관계자가 참여하고, 모니터링 및 통제되며, 프로세스에 대한 검토 및 준수에 대한 평가를 통하여 프로세스를 관리하는 상태를 의미한다.



<그림 3> 프로세스의 구성



〈그림 4〉 프로세스의 목표별 활동

• 체계화 목표(Defined Goal) : 관리되는 프로세스가 표준화되고, 표준화된 프로세스들이 체계적으로 조직 전체로 활용되는 상태일 때 달성 가능한 목표로서, 표준, 프로세스 정의, 절차가 각 업무그룹이나 조직단위에 보완 적용이 가능하며 과거의 데이터 품질 관련 경험을 축적하고 활용하는 상태를 의미한다.

(4) 활동(Activity) : 목표를 달성하기 위한 각 프로세스 별 하위 작업들을 의미한다. 활동은 〈그림4〉와 같이 실행 목표를 달성하기 위한 고유 활동, 관리화 목표를 달성하기 위한 PDCA 관점의 10개의 공통 활동, 체계화 목표를 달성하기 위한 2개의 공통 활동으

로 구성된다.

① 실행 목표 달성을 위한 활동

프로세스의 실행 목표를 달성하기 위해 프로세스의 고유활동을 수행하여 작업 산출물을 생성한다. 예를 들면, 데이터 표준관리 프로세스의 실행목표를 달성하기 위해서는 데이터 표준 관리대상 및 기준 정의, 데이터 표준 정의, 데이터 표준 적용의 고유활동을 수행하게 된다.

② 관리화 목표 달성을 위한 활동

프로세스의 관리화 목표를 달성하기 위하여 규정 및 절차에 따라 프로세스 관리를 위한 10개의 활동

을 수행하여야 한다.

- 규정 및 절차 수립 : 프로세스 수행을 위한 규정 및 절차를 수립하고 유지한다.
- 프로세스 계획 수립 : 프로세스 수행을 위한 계획을 수립하고 유지한다.
- 자원 제공 : 프로세스를 수행하는데 필요한 적절한 자원을 제공한다.
- 책임 할당 : 프로세스를 수행하기 위한 책임과 권한을 할당한다.
- 교육훈련 : 프로세스를 수행 또는 지원하는 인력을 교육 훈련한다.
- 작업 산출물 통제 : 프로세스의 작업 산출물들을 적절한 수준으로 통제한다.
- 관련 이해관계자 식별 및 참여 : 프로세스의 관련 이해관계자들을 식별하고 참여시킨다.
- 프로세스 모니터링 및 통제 : 프로세스 수행 계획을 기준으로 프로세스를 모니터링 및 통제하고, 적절한 시정 조치를 취한다.
- 준수의 객관적 평가 : 프로세스 및 작업 산출물에 대한 기술서, 표준 및 절차를 준수하는지를 객관적으로 평가하고 부적합한 부분을 해소한다.
- 상위수준 관리성과 현 상태 검토 : 상위의 관리성과 함께 프로세스 활동, 현 상태 및 결과물을 검토하고 이슈를 해결한다.

③ 체계화 목표 달성을 위한 활동

프로세스를 표준화하고 조직 전체로 확산하기 위한 다음의 2개의 활동을 수행하여 프로세스의 체계화 목표를 달성한다.

- 표준화된 프로세스 수립 : 표준화된 프로세스 기술서를 수립하고 기관차원으로 확산한다.
- 프로세스 관련 경험 수집 : 프로세스와 프로세스 자산을 향후 데이터 품질 개선에 활용할 수 있도록 프로세스 관련 경험과 노하우를 수집하고 공유한다.

(5) 하위 활동(Task) : 활동에 대한 하위 작업으로서 활동을 수행하기 위한 요구사항, 추천사항, 허용행동을 의미한다. 이해를 돕기 위해 상세 설명이나 예시 등을 추가할 수 있다.

(6) 작업산출물(Work Product) : 활동의 수행에 따라 발생하는 결과물(Output)이다.

3) 프로세스 정의

프로세스의 내용의 많아 여기서는 프로세스의 의미를 파악할 수 있도록 구성체계 중 프로세스의 목적, 실행목표 달성을 위한 활동만을 <표 2>에 기술하였다.

4) 프로세스 비교 분석

<표 3>과 같이 프로세스 관점에서 한국정보화진흥원의 품질관리매뉴얼(v1.1), ISO 8000-61, 한국데이터베이스진흥원의 DQM3, Caballero, et al.(2008)의 IQM3 모델의 비교를 통하여 공공데이터 품질관리 프로세스를 검토하였다. 다음은 본 연구에서 제시한 프로세스 참조 모델과의 차이점을 설명한 것이다.

본 연구의 프로세스 참조 모델은 품질관리매뉴얼(v1.1)에서 다루지 않는 데이터 사용자관리가 추가되었고, 성과점검 등 부분적으로 다루던 데이터 운영관리를 재구조화 하였고, 성과관리도 데이터품질 개선 성과관리와 조직품질 성과관리로 구분하여 강화하였다.

ISO 8000-61은 본 연구의 프로세스 참조모델과 비교해 볼 때 프로세스에서 다루는 활동의 범주에서 차이가 존재하고, 프로세스의 활동이 구체화되어 있지 않은 점이 존재한다. 따라서 ISO 8000-61의 데이터 품질정책/표준/절차관리는 공공기관의 조직적 특성을 반영하여 본 연구의 참조모델에서는 데이터 품질 정책관리, 데이터 표준관리, 업무규칙관리 등으로 구체화하였다. 또한 ISO 8000-61의 하위활동에

〈표 2〉 공공 데이터 품질관리 프로세스의 목적과 활동 목록

구분	프로세스	목적	활동 목록
Plan	데이터품질 정책관리	조직의 데이터 품질관리를 위한 상위수준의 비전과 목표를 포함하는 데이터 품질관리 정책을 수립하고, 데이터 품질관리를 지속적이고 상시적으로 수행할 수 있도록 데이터 품질관리 규정 마련 및 준수	데이터 품질관리 정책 수립, 데이터 품질관리 규정 수립
	데이터품질 계획관리	조직의 데이터 품질관리 정책 및 정책 및 규정(지침)에 근거하여, 데이터 품질 목표 달성을 위한 조직의 데이터 품질관리 계획을 수립	데이터 품질 요구사항 정의, 데이터 품질관리 목표 정의, 데이터 품질관리 계획 수립
Do	데이터 표준관리	데이터 명칭, 정의, 형식, 규칙 등에 관한 표준을 수립하고, 데이터 구조에 일관되게 적용하며, 이를 체계적·지속적으로 관리함으로써, 데이터 품질 수준을 향상시키고 이해관계자(정보화부서, 업무부서, 유지보수사업자, 데이터 사용자 등) 사이의 의사소통을 원활히 함	데이터 표준관리 대상 및 기준 정의, 데이터 표준 정의, 데이터 표준 적용 대상 정의, 데이터 표준 적용
	업무규칙 관리	데이터 품질 측정 및 오류데이터 추출의 기준이 되는 업무규칙(Business Rule)을 도출하고, 체계적으로 관리	업무규칙 관리대상 정의, 업무규칙 정의
	데이터 사용자관리	데이터의 생성에 직접적으로 관여하는 자들을 대상으로 데이터의 오류입력을 최소화 하도록 데이터 처리 규칙을 수립하고, 교육하고, 지속적으로 모니터링	데이터 처리 이력관리 대상 정의, 데이터 처리 이력 기록
	데이터 구조관리	애플리케이션 간에 공유된 데이터의 구조와 의미를 명시하고 그 데이터의 사용을 관리	데이터 모델링 기준 정의, 데이터 모델 수립 및 유지, 중복데이터 목록 정의
	데이터 흐름관리	애플리케이션 시스템이나 데이터베이스에 배포된 데이터 간의 관계와 흐름을 통제하여 데이터의 무결성을 보장하고, 데이터 오류 발생에 따른 원인을 파악	흐름 관리 대상 데이터 식별, 데이터 소스-타겟 매핑정보 정의, 연계데이터 목록 정의
	데이터 운영관리	운영 중인 데이터 자산의 불법 유출 및 비정상적인 접속을 사전에 차단하고, 데이터의 가용성을 유지하며, 데이터 활용성 제고를 위한 데이터 성능을 최적화	데이터 접근 제한, 데이터 보안 및 통제, 데이터 백업, 데이터 복구, 주기적인 성능 점검
Check	품질오류 신고관리	기관에서 제공하는 데이터 및 정보에 대한 품질을 유지하기 위해 정보를 활용하는 국민 또는 타 기관 그리고 기관 내의 구성원들이 정보를 활용하는 과정에서 발견하는 다양한 품질 오류를 해당기관이 확인하고 처리	품질 오류 신고 접수 및 검토, 품질 오류 처리, 품질 오류 처리결과 통보
	데이터 품질진단	기관의 데이터 품질 수준을 진단하여 데이터 품질 오류를 식별	데이터 품질 측정대상 정의, 데이터 품질 측정기준 및 방법 정의, 데이터 품질 측정 실시, 데이터 품질 측정결과 분석, 데이터 품질 오류 식별
	원인분석 및 개선방안 제시	데이터 품질진단을 통해 도출된 데이터 품질오류에 대해 문제의 원인을 파악하고 데이터 품질문제들이 실제 업무에 미칠 수 있는 영향분석을 통해 품질개선의 방안을 제시	데이터 품질오류 원인 식별, 데이터 품질오류의 업무 영향 분석, 개선방안 제시, 개선계획 수립
	데이터 정제	데이터 품질 개선 계획에 따라 데이터 품질오류와 관련된 모든 데이터를 추적하여 수정	데이터 정제 대상 정의, 데이터 정제 실시
Act	프로세스 개선	데이터 품질 개선 계획에 따라 데이터 오류의 재발 방지를 위해 관련 프로세스를 개선	프로세스 개선 대상 정의, 프로세스 개선 실시
	데이터품질 개선 성과관리	데이터 품질 개선 전·후의 성과를 비교 및 평가하여 데이터 품질 개선을 검증	데이터 품질 개선 성과 측정기준 및 방법 정의, 데이터 품질 개선 성과 측정 및 분석, 데이터 품질 개선 성과를 위한 개선안 도출 및 반영
	조직품질 성과관리	조직의 데이터품질 정책관리에서 정의한 중장기 목표와 데이터품질 계획 관리에서 정의한 단기 목표에 따라 성과를 체계적으로 관리하여 그 목표를 달성하기 위해 개선안을 도출	조직품질성과 측정기준 및 방법 정의, 조직품질성과 측정 및 분석, 조직품질성과를 위한 개선안 도출 및 반영

〈표 3〉 공공데이터 품질관리 참조모델과 관련 품질관리 모델과의 비교

(○ : 전체 포함 △ : 일부 포함 × : 미포함)

프로세스 그룹	본 연구의 프로세스 참조모델	품질관리 매뉴얼(V1.1)	ISO 8000-61 프로세스	DQM3 프로세스	IQM3 프로세스
계획 (PLAN)	데이터 품질 정책관리	○	○	×	○
	데이터 품질 계획관리	○	○	△	○
실행 (DO)	데이터 표준 관리	○	△	○	○
	업무규칙 관리	○	△	○	×
	데이터 사용자 관리	×	○	○	×
	데이터 구조 관리	○	○	○	△
	데이터 흐름 관리	○	○	○	△
	데이터 운영 관리	△	○	△	×
모니터링 및 통제 (CHECK)	품질 오류 신고 관리	○	○ ¹⁾	○	○
	데이터 품질 진단	○	○	○	○
	원인분석 및 개선방안 제시	○	○	○	○
개선 (ACT)	데이터 정제	○	○	○	○
	프로세스 개선	○	○	○	○
	데이터 품질 개선 성과관리	△	○	△	×
	조직 품질성과 관리	△	○ ²⁾	△	×

1) 요구사항관리에 있음

2) 계획대비 성과평가로 되어 있음

서 다루는 요구사항관리, 계획대비 성과평가 등도 본 연구의 참조모델에서는 품질오류신고관리와 데이터 품질 개선 성과관리, 조직 품질성과관리 등으로 재구조화하였다.

국내의 DQM3나 해외의 IQM3는 주로 민간영역에 적용가능한 성숙도 모델을 제시하고 있는 한계점들이 존재한다. 예를 들어, DQM3에서는 공공기관의 특성인 데이터 정책관리가 프로세스에 누락되어 있고, 데이터 운영관리는 일부 보안관리만 관리하고 있으며, 데이터품질 개선 성과관리와 조직품질 성과관리도 일부만 관리되고 있다. 또한, IQM3에서는 공공데이터 품질관리에서 중요한 업무규칙관리, 데이터 사용자관리, 데이터 운영관리, 데이터품질 개선 성과관리, 조직품질 성과관리 프로세스가 제시되지 않거나 부분적으로만 다루어지고 있었다. 본 연구의

프로세스 참조 모델에서는 이러한 프로세스의 누락 및 일부 반영된 프로세스를 체계화하고 보완하여 구체화하였다.

2. 측정 프레임워크

측정 프레임워크는 프로세스의 능력을 측정하는 부분과 조직의 성숙도를 측정하는 부분으로 구성되며 능력 수준과 성숙도 수준을 결합하여 개별 프로세스의 능력과 해당 프로세스가 속한 성숙도 수준의 위치를 파악할 수 있다.

측정 프레임워크에서 '수준(Level)'은 데이터 품질관리를 수행하는 프로세스를 개선하고자 하는 조직에 대해 현재 상황을 객관적인 기준을 통해 표시하며 개선 목표 및 개선 방향을 기술하는데 사용된다.

측정 프레임워크에서 수준을 표현하는 방법은 두 가지이다. 하나는 조직이 선정한 개별 프로세스(또는 여러 프로세스 그룹)를 개선하는 방법이며 이는 능력 수준에 해당된다. 다른 하나는 프로세스를 몇 개씩 묶어 그룹으로 만들어서 그룹을 차례대로 같이 개선하는 방법이며 이는 성숙도 수준에 해당된다. 두 가지 방법 모두 프로세스를 개선해 사업 목적을 달성할 수 있는 방법을 제시한다.

1) 능력 수준과 성숙도 수준

능력 수준은 프로세스 참조모델에서 정의된 개별 프로세스에 대하여 해당 프로세스의 상태를 나타내기 위해 사용된다. 능력 수준은 개별 프로세스에서 조직의 프로세스 개선 달성에 적용된다. 능력 수준은 프로세스 목표에 맞추어 4단계로 정의한다. 능력 수준 0은 불완전한(Incomplete)로 정의하며 아직 수행되지 않았거나 부분적으로만 수행된 프로세스를 말한다. 능력 수준 1은 실행되는(Performed) 프로세스로 정의하며 작업 산출물을 생산하는데 필요한 작업을 실행한 프로세스이다. 능력 수준 2는 관리되는(Managed) 프로세스로 정의하며 정책에 따라 계획 및 실행되고, 통제된 결과물을 생산하기 위해 적절한 자원을 갖춘 숙련된 인력을 고용하며, 관련 이해관계자들을 참여시키고, 모니터링, 통제 및 검토 대상이 되며, 해당 프로세스 기술 준수 여부의 평가 대상이 되는 실행되는 프로세스를 말한다. 능력 수준 3은 체계화된(Defined) 프로세스로 정의하며 조직에서 유지 관리되는 프로세스 기술서가 있으며, 프로세스 관련 경험을 조직의 프로세스 자산에 축적하여 관리하는 프로세스를 말한다.

성숙도 수준은 전체 모델에 대비하여 조직 프로세스의 전반적인 상태를 나타내기 위해 사용된다. 성숙도 수준은 여러 프로세스에 걸쳐 조직의 프로세스 개선 달성에 적용된다. 성숙도 수준은 프로세스 그룹의 수준에 따라 6단계로 정의한다. 성숙도 수준 0은 도입전(Immature) 단계로 정의하며 정의된 프로세스

가 실행되지 않거나 부분적으로 실행되며, 작업산출물도 부분적으로만 생성되는 단계이다. 이 단계에서는 데이터품질문제가 지속적이고 반복적으로 발생하며, 품질 문제에 대한 해결 조치가 임기응변적으로 이루어진다.

성숙도 수준 1은 도입(Basic) 단계로 정의하며 PDCA 사이클에 기반하여 기본적인 데이터 품질관리 활동이 시작된다. 이 단계에서는 데이터 오류에 대한 적발 및 단순한 데이터 오류 정제 등 사후 조치가 가능하며 특정 DB나 단위조직의 단기적인 품질향상에 집중한다.

성숙도 수준 2는 관리화(Managed) 단계로 정의하며 프로세스 중심의 상시적인 품질관리 기반이 구축된다. 이 단계에서는 데이터 값에 영향을 주는 표준관리, 구조관리, 흐름관리, 사용자 관리 등의 통제 범위가 확대되며 데이터 오류에 대한 예방중심으로 활동이 강화된다. 성숙도 수준 1과 성숙도 수준 2의 차이는, 일정한 형태로 수행에 들어가면 수준 1이 시작되었다고 간주하며, 수행을 하되 정해진 절차를 따라서 수행하면 수준 2가 시작될 것으로 본다.

성숙도 수준 3은 체계화(Defined) 단계로 정의하며 단위조직 중심의 품질관리가 조직 전체차원으로 확대된다. 이 단계에서는 조직 전체의 품질관리를 효과적으로 지원하기 위해 데이터의 성능, 장애, 보안, 백업 등의 데이터 운영관리가 필요하며 조직차원의 반복적 품질 이슈에 관한 원인분석 및 개선방안 수립에 가능하다. 수준 2와 수준 3의 차이는, 수준 2는 이미 정형화되었다는 의미가 있으며 수행에 필요한 규정과 절차가 준비되어야 하며 이러한 절차나 규정 및 프로세스 등이 마련되어 있지만 실제 상황에서는 다소 지켜지지 않는다면 수준 3으로 간주한다.

성숙도 수준 4는 정량화(Measured) 단계로 정의하며 성숙도 수준 3의 원인분석 및 개선방안 제시에 따른 프로세스 개선이 수행된다. 이 단계에서는 프로세스 개선에 따른 성과측정 기준을 마련하여 개선 전후의 점검이 가능해져, 프로세스의 개선 효과가 검증

된다.

성숙도 수준 5는 최적화(Optimized) 단계로 정의하며 조직 전체의 지속적인 프로세스 개선이 가능하고, 이를 통해 조직의 품질관리에 대한 단계적 목표 달성이 가능해진다. 이 단계에서는 조직에 맞는 품질 성과관리 체계가 확보되고, 지속적으로 평가되어 목표수립이나 계획 수립에 반영되는 선순환 체계가 구축된다.

2) 수준 결정

(1) 성숙도 수준별 프로세스 배정

앞에서 정의한 성숙도 수준의 정의에 따라 각 프로세스의 특성을 매핑시키면 일정 수준에 속하는 프로세스 그룹을 형성할 수 있다. 성숙도 수준에 따라 배정된 프로세스가 <그림 5>에 나타나 있다. 예를 들어, 성숙도 수준 1의 특징으로는 데이터의 기본적인 활동 시작, 실행목표에서 정의한 활동 수행, PDCA 사이클 가동 등이 있으며 이러한 특징을 고려하여 업무 규칙 관리, 품질 오류 신고 관리, 데이터 품질 진단, 데이터 정제 프로세스를 성숙도 수준 1 그룹으로 묶어준다. 성숙도 수준 2는 데이터 값에 영향을 주는 품질관리의 통제 범위가 확대되며 근본적인 품질영향 요인의 관리가 필요하게 되어 품질관리 활동의 제도적 근거를 마련하는 단계이므로 데이터 품질 정책 관리, 데이터 품질 계획 관리, 데이터 사용자 관리, 데이터 표준 관리, 데이터 구조 관리 프로세스를 배정한다. 성숙도 수준 3은 단위조직 중심의 품질관리가 조직 전체차원으로 확대되는 단계로 표준 프로세스가 정의되고 문서화되며, 기관 전체의 이해관계자가 이를 준수하는 노력이 중요해지므로 데이터 흐름 관리, 데이터 운영 관리, 원인 분석 및 개선 방안 제시 프로세스를 배정한다. 성숙도 수준 4는 프로세스 개선에 따른 성과측정 기준을 마련하여 개선 전후의 점검이 가능해져 프로세스의 개선 효과가 검증되는 단계이므로 프로세스 개선, 데이터 품질 개선 성과

관리 프로세스를 배정한다. 성숙도 수준 5는 품질을 담당하고 있는 조직 전체의 지속적인 프로세스 개선이 가능하고, 이를 통해 조직의 품질관리에 대한 단계적 목표 달성이 가능해지는 단계이므로 조직 품질 성과 관리 프로세스를 배정한다.

(2) 능력 수준 결정

프로세스의 능력 등급을 매기기 위해서는 프로세스의 목표 달성 여부를 판별해야 한다. 프로세스 목표의 달성 여부는 활동이 얼마나 잘 이루어져 있는가에 달려있다. 각 활동의 결과로 작업 산출물을 생성하게 된다. 결국 프로세스의 능력 등급 산정은 최하위 단계인 작업 산출물의 점검을 통해 그에 해당하는 활동이 얼마나 잘 이루어 졌는지 판단하게 되며 활동이 잘 이루어지면 목표가 달성된 것으로 보며 목표가 달성된 프로세스는 적절한 능력 등급을 받게 된다.

즉, 각 프로세스에 속하는 활동의 결과로 해당 활동의 작업산출물이 생성되게 된다. 따라서 데이터 표준 관리 프로세스의 목표 달성 여부는 최종 작업 산출물의 이행 여부를 판단하여 이에 해당하는 활동 달성 여부→세부 목표 달성 여부→목표 달성 여부→프로세스 능력 수준 등급 부여의 절차를 거친다.

예를 들어, 데이터 품질관리 프로세스 참조모델에서, 데이터 표준 관리 프로세스의 목표는 실행 목표 2개(데이터 표준 수립, 데이터 표준 적용), 관리화 목표 1개(데이터 표준관리 프로세스 정형화), 체계화 목표 1개(데이터 표준관리 프로세스 체계화)로 구성되어 있다. 실행 목표 중 데이터 표준 수립 목표는 2개의 활동(데이터 표준관리 대상 및 기준 정의, 데이터 표준 정의)을 가지며 데이터 표준 적용 목표도 2개의 활동(데이터 표준 적용 대상 정의, 데이터 표준 정의)을 갖는다. 이어 데이터 표준관리 대상 및 기준 정의 활동의 결과로 두 가지 유형의 작업 산출물이 생성되며 데이터 표준 정의 활동의 결과로는 4가지 작업 산출물이 생성되어야 한다. 또한 데이터 표준 적용 목표는 각각 2개의 활동을 가진다. 관리화

목표인 데이터 표준관리 프로세스 정형화는 10개의 활동을 가지고 있다. 체계화 목표인 데이터 표준관리 프로세스 체계화 목표는 2개의 활동을 가진다. 데이터 표준 관리 프로세스는 성숙도 수준 2에 속하는 프로세스이므로 실행 목표 2개를 모두 만족하면 능력 수준 1, 관리화 목표를 만족하면 능력 수준 2, 더 나아가 체계화 목표를 달성하면 능력 수준 3으로 등급을 정한다.

(3) 조직 성숙도 수준 결정

조직의 성숙도 수준 1을 달성하려면 성숙도 수준 1에 할당된 프로세스가 능력 수준 1을 달성해야 한다. 조직의 성숙도 수준 2를 달성하려면 성숙도 수준 1과 2에 할당된 프로세스가 모두 능력 수준 1과 2를 달성해야 한다. 조직의 성숙도 수준 3을 달성하려면

성숙도 수준 1, 2, 3에 할당된 프로세스가 모두 능력 수준 3까지를 달성해야 한다. 조직의 성숙도 수준 4를 달성하려면 성숙도 수준 1, 2, 3, 4에 할당된 프로세스가 모두 능력 수준 3까지를 달성해야 한다. 조직의 성숙도 수준 5를 달성하려면 모든 프로세스가 능력 수준 3까지를 달성해야 한다. 만약 해당 수준에서 하나라도 만족하지 못하는 능력 수준이 있으면 그 조직 성숙도 수준에 도달하지 못한 것으로 판단한다. 이를 도식화하면 <그림 5>와 같다.

이 그림에서는 한 조직의 성숙도 수준의 예를 제시하였는데 능력 수준을 달성한 경우 “O”, 달성하지 못한 경우 “X”로 표시하였다. 이 예에서는 성숙도 수준 3에서 데이터 운영관리 프로세스와 원인분석 및 개선방안 제시 프로세스가 해당 능력 수준을 달성하지 못하였으므로 조직 성숙도 수준 2라고 판정한다.

범례 ○ 달성 × 미달성

성숙도 수준 (Maturity Level)	성숙도 수준 판정 기준	프로세스	프로세스 능력수준 (Capability Level)		
			CL1	CL2	CL3
ML1 도입(Basic)	• ML1에 속하는 프로세스의 능력수준(CL) 1 충족	업무규칙관리	○	○	○
		품질오류신고관리	○	○	○
		데이터품질진단	○	○	○
		데이터정제	○	○	○
ML2 관리화(Managed)	• ML1,2에 속하는 프로세스의 능력수준(CL) 1,2 충족	데이터품질정책관리	○	○	○
		데이터품질계획관리	○	○	○
		데이터표준관리	○	○	○
		데이터구조관리	○	○	○
		데이터사용자관리	○	○	○
ML3 체계화(Defined)	• ML1,2,3에 속하는 프로세스의 능력수준(CL) 1,2,3 충족	데이터흐름관리	○	○	○
		데이터운영관리	○	×	×
		원인분석 및 개선방안제시	○	○	×
ML4 정량화(Measured)	• ML1,2,3,4에 속하는 프로세스의 능력수준(CL) 1,2,3 충족	프로세스개선	Out of Scope (성숙도수준 판정 범위 제외)		
		데이터품질개선 성과관리			
ML5 최적화(Optimized)	• ML1,2,3,4,5에 속하는 프로세스의 능력수준(CL) 1,2,3 충족	조직품질 성과관리			

<그림 5> 조직 성숙도 수준 판정 예시

성숙도 수준 3을 달성하지 못함에 따라 성숙도 수준 4 및 5에 해당하는 프로세스는 평가할 필요가 없으므로 판정 범위에서 제외된다.

V. 검증

본 성숙도 모델의 개발 방향 정립을 위해서 사전에 전문가 협의회 2회 개최 및 전문가 의견 설문 조사를 실시하였다. 두 차례 전문가 협의회를 통해 본 성숙도 모델의 개발 방향에 대한 의견을 수렴하였다.

본 연구에서 제시된 데이터 품질관리 조직 성숙도 모델의 검증을 위해서 프로세스 참조 모델은 내용타당도 분석을 실시하였으며 측정 프레임워크는 참조 모델에 종속적이므로 별도의 정량적 검증을 수행하지 않고 전문가의 의견을 수렴하였다.

프로세스 참조 모델의 내용 타당도 분석을 위해서 24명의 데이터 품질 전문가에게 15개 품질관리 프로세스에 대한 활동들이 타당하고 적절하게 정의되어 있는지를 묻는 의견 조사서를 배포하여 그 의견을 수렴하였고 의견 결과의 내용타당도 검증을 수행하였다. 내용타당도(Content Validity Ratio; CVR) 분석은 전문가 패널을 대상으로 프로세스별 활동의 내용타당성을 검증하는 방법으로 해당 활동이 얼마나 타당하다고 생각하는지 질의하여 전체 패널의 수와 타당하다고 응답한 패널 수의 비율에 따라 적합성을 판단한다(Lawshe, 1975; 정운수 외, 2011; 박상현 외, 2011). 적합성 판단 방법은 다음과 같은 산식에 따라 측정된 각 지표별 CVR 값이 패널 수에 따른 최소 허용 값보다 높으면 채택된다.

$$CVR = (n_e - N/2) / (N/2)$$

n_e : 중요(Essential)하다고 응답(5점 척도에서 4점과 5점)한 전문가 패널의 수

N : 전체 전문가 패널의 수

Lawshe(1975)에 따르면 유의도 0.05 수준에서 24명의 전문가 패널 수에 따른 최소 허용값은 0.42이며, 응답결과 분석된 CVR 값이 0.42 이하인 활동

은 없으므로 본 모델에서 제시한 활동은 내용타당도가 있다고 판단되었다.

VI. 결론

본 연구는 공공데이터 품질관리를 위한 조직의 성숙도 수준을 측정하기 위한 모델을 제시한 것으로, 첫 번째는 공공부문에 적합한 품질관리 프로세스들로 구성된 프로세스 참조 모델을 제시하였고, 두 번째는 제시된 프로세스의 달성 수준을 측정하여 조직의 품질관리 성숙도 수준을 평가하는 기준을 제시하였다.

본 연구의 프로세스 참조 모델에서는 데이터 품질관리 조직(또는 해당부서)이 데이터 품질관리 관련 활동을 이행하는 여러 프로세스를 갖춰야 한다고 명시하고 있다. 이러한 프로세스들이 마련되어 있는지 여부를 판단하려면, 데이터 품질관리 조직(또는 해당부서)은 보유하고 있는 프로세스를 본 모델의 각 프로세스에 매핑해 보면 된다.

조직의 성숙도 수준은 데이터 품질 조직의 성과를 매길 수 있는 방법을 제공한다. 각 성숙도 수준은 조직 프로세스 중 사전에 정의된 프로세스 세트가 성숙하고, 다음 성숙도 수준으로 올라갈 수 있도록 준비하는 것이다. 조직 성숙 수준은 개인 차원에서 수행하는 것이 가장 낮은 수준이며, 수행의 단위가 부서 등으로 확대되면 수준이 높아지며, 수행 범위가 전사적으로 확대되면 수준은 점차 높아지게 된다. 또한 수행 범위가 전사적일 때 여기에 일반적인 목표를 부여하게 되면 한층 더 수준이 높아지게 된다.

특정 조직의 경우 데이터 품질 표준관리 프로세스를 다른 프로세스보다 중요도를 높게 줄 수도 있다. 이러한 경우 이 프로세스는 핵심 프로세스의 범주에 포함시킬 필요가 있다. 수준이 높ی 올라갈수록 흐름관리, 권한관리 등의 중요도가 높아지며 필요에 따라서는 사용자 관리의 중요도가 높아질 수 있다. 특정 조직의 관점에서는 구조관리나 표준관리에 비중을

높게 줄 수 있다. 따라서 프로세스의 중요도의 의미는 그 조직이나 데이터 형태에 따라서 상당히 달라질 여지가 있다.

본 연구는 공공데이터 품질관리에 데이터 품질에 대한 수준 진단을 넘어서 프로세스와 조직의 수준을 파악할 수 있는 근거를 제공한다. 데이터를 개방하고자 하는 조직은 본 모델을 적용함으로써 개방되는 데이터에 대한 신뢰성을 확보할 수 있으며 프로세스의 개선과 데이터 품질의 방향성을 설정할 수 있다.

■ 참고문헌

- 김선호·이창수 (2013). “데이터 품질관리 프로세스 평가를 위한 프로세스 참조모델.” 「한국전자거래학회지」, 18(4), 83-104.
- 김선호·이창수 (2010). “마스터 데이터 품질관리 프레임워크.” 「Entrue Journal of Information Technology」, 9(2), 109-121.
- 박상현·어재경·전효정 (2011). “녹색정보화 성숙도 진단 모델 개발 및 실증연구.” 「Information Systems Review」, 13(1), 115-141.
- 안전행정부 (2014). 「공공데이터 관리지침」. 안전행정부
- 이창수·김선호 (2010). “데이터 품질관리 프레임워크 및 사용 시나리오.” 「전자거래학회지」, 15(4), 79-99.
- 정승호·정덕훈 (2013). “공공기관의 데이터 품질에 영향을 미치는 요인에 관한 연구.” 「정보처리학회지」, 2(4), 251-266.
- 정윤수·김숙희·김태진·박상현 (2011). 「국가정보화 아키텍처 기반 공공-민간 협업시스템 구축방안 연구」. 정보통신정책연구원.
- 한국데이터베이스진흥센터 (2006). 「데이터 품질관리 성숙모형 (Ver 1.0)」. 한국데이터베이스진흥센터.
- 한국정보화진흥원, (2014a). “2014년 EA 실태조사 및 성숙도 측정 계획.” 안전행정부.
- 한국정보화진흥원 (2014b). “공공데이터 품질개선 성과 공유대회.” 안전행정부.
- 한국정보화진흥원 (2014c). 「공공정보 품질관리 매뉴얼 (v1.1)」. 한국정보화진흥원.
- Caballero, I. & Caro, A. & Calero, C. & Piattini, M. (2008). “IQM3: Information Quality Management Maturity Model,” *Journal of Universal Computer Science*, 14(22), 3658-3685.
- CMMI Product Team (2010). *CMMI for Development (CMMI-DEV), Version 1.3*, Technical Report, CMU/SEI-2010-TR-033, Software Engineering Institute, <http://www.sei.cmu.edu/cmmi/tools/cmmiv1-3/>.
- CMMI Institute (2014). *Data management Maturity(DMM) Model*. CMMI Institute.
- DAMA International (2009). *The DAMA Guide to The Data management Body of Knowledge (DAMA-DMBOK Guide), First Edition*, Technics Publications.
- Deming, W. E. (1982). *Out of the Crisis*, MIT Center for Advanced Engineering Study, Cambridge, MA.
- Eppler, M. & Helfert, M. (2004). A classification and analysis of data quality costs. In: 9th MIT international conferences on information quality, November 5-6, 2004, Boston, MA.
- Haug A. & Zachariassen F. & van Liempd, D. (2011). “The costs of poor data quality”, *J. of Industrial Engineering Management*, 4(2), 168-193.
- HM Government (2012). *Open data white paper, Unleashing the potential*, TSO.
- ISO 8000-60 (2014). ISO/CD TS 8000-60 Data quality-Part 60: Data quality management: Process assessment: Overview, ISO.
- ISO/CD 8000-61 Data quality-Part 61: Data quality management: Process assessment: Process reference model.
- ISO 9001:2008 Quality management systems-Requirements.
- ISO/IEC 12207:2002 Systems and software engineering-Software life cycle processes, ISO.
- ISO/IEC 15288:2008 System Engineering - System Life Cycle Processes, ISO.
- ISO 15504-2, ISO/IEC 15504-2:2003 Information technology-Process assessment-Part 2:

- Performing an assessment, ISO.
- ISO/IEC 33001:2014(E) Information Technology-
Process assessment-Concepts and
terminology.
- Lawshe, C. H. (1975). "A quantitative approach
to content validity", *Personnel Psychology*,
No. 28, 563-575.
- LSC Group (2012). *Data Quality Assessment Tool
Version 2.0, 2. Data Quality Maturity
Model*, Joint Information Group/MOD/LSC
Group.
- Moen R. & Norman C. (2009). "Evolution of
PDCA Cycle." *Asian Network for Quality
Congress*.
- Peldzius S. & Ragaisis S. & Simenas J. (2010).
"Mapping CMMI-DEV maturity levels to
ISO/IEC 15504 capability profiles."
Proceeding TELE-INFO'10 Proceedings of
the 9th WSEAS international conference
on Telecommunications and informatics.
- Redman, T. (1998). "The impact of poor data
quality on the typical enterprise." *Comm.
ACM* 41(2), 79-82.
- Sadiq, S. (2013). *Handbook of Data Quality:
Research and Practice*. Springer.
- United States. US Department of State (2014a).
Open data plan. [http://www.state.gov/
open/](http://www.state.gov/open/).
- United States. US Department of State (2014b).
Open government plan. [http://www.state.
gov/open/](http://www.state.gov/open/).
- US Federal Data Architecture Subcommittee
(DAS) (2008). *Federal DAS-Data Quality
Framework V1.0*. DAS.