

최 종 보 고 서			
관 리 번 호	2021-1	기 술 분 류	전 산
과 제 명	(한글) 과기정통부 메타데이터 및 품질관리 도입 방안 연구 (영문) A Study on the Introduction of Metadata and Data Quality Control Management of Ministry of Science and ICT		
주관연구기관 (협동연구기관)	기 관 명	소재지	대 표
	에버밸류컨설팅	서울특별시 영등포구 선유도 9 길 10 SK V1 센터 619,620호	정용균
주관연구책임자 (협동연구책임자)	성 명	소속 및 부서	전 공
	성기선	지능정보사업부	데이터베이스
총연구기간 (당해년도)	2021년 06월 01일 ~ 2021년 11월 28일(6개월)		
총연구비 (당해년도)	일금 오천팔백오십만원 정 (₩ 58,500,000)		
총참여연구원 (당해년도)	7 명(수석: 1명, 책임: 3명, 연구원: 3명)		
<p>2021년도 정책연구용역사업으로 수행한 연구과제의 최종보고서를 붙임과 같이 제출합니다.</p> <p>붙임 : 최종보고서 10부.</p> <p style="text-align: right;">2021년 12월 13일</p> <p style="text-align: right;">주관연구책임자 성기선 </p> <p style="text-align: right;">주관연구기관장 에버밸류컨설팅 </p> <p>과학기술정보통신부장관 귀 하</p>			

과기정통부 메타데이터 및 품질관리 도입 방안 연구에 관한 정책연구사업의

최종보고서를 별첨과 같이 제출합니다.

2021년 12월 13일

주관연구책임자 성기선(인) 

주관연구기관장 에버밸류컨설팅  직인

<p>정책연구 - 2021-1 데이터댐 구축을 위한 과기정통부 메타데이터 및 품질관리 도입 방안 연구</p> <p>과 학 기 술 정 보 통 신 부</p>	<p>정책연구 2021 - 1</p> <p>과기정통부 메타데이터 및 품질관리 도입 방안 연구</p> <p>(A Study on the Introduction of Metadata and Data Quality Control Management of Ministry of Science and ICT)</p> <p>과학기술정보통신부</p>
---	--

제 출 문

과학기술정보통신부장관 귀하

본 보고서를 “과기정통부 메타데이터 및 품질관리 도입 방안 연구“
최종보고서로 제출합니다.

2021년 12월 13일

- 주관연구기관명 : 에버밸류컨설팅
- 연구기간 : 6개월
- 주관연구책임자 : 성기선
- 참여연구원
 - 연구원 : 최정수
 - 연구원 : 김지현
 - 연구원 : 이혜련
 - 연구원 : 이경란
 - 연구원 : 김희준
 - 연구원 : 금재철

※ 주관연구기관 및 주관연구책임자, 연구원은 실제 연구에 참여한 기관 및 참여자의 명의로 함

목 차

요약문(국문)	I
요약문(영문)	II
본문	III
I. 연구의 필요성 및 목적	14
1. 연구의 목적	14
2. 연구의 배경	14
3. 연구의 필요성	15
4. 연구의 내용 및 추진방법	15
II. 공공 데이터 품질 정책 동향	23
1. 공공 데이터 품질관리	23
2. 공공 데이터 품질관리 단계 별 활동 및 산출물	24
3. 국가 차원의 데이터 품질관리	27
4. 공공데이터 품질관리 수준평가	28
III. 기관 데이터 품질 연구 진행 결과	31
1. '21년 공공데이터 품질관리 추진과제 점검 결과	31
2. '21년 공공데이터 품질관리 선정 DB 내역	33
3. 기관 별 표준화 진행결과	36
IV. 데이터 관리체계 품질관리 수준 평가 대응	43
1. '21년 품질관리 수준평가 대상 시스템 목록	43
2. '21년 품질관리 수준평가 결과	43
3. 평가대상 DB별 평가결과	44
V. 데이터 품질관리 도구 도입 방안	50
1. 기관 품질관리 현황	50
2. 데이터 품질진단 도구(DQMS) 도입방안	53
VI. 데이터 품질관리 역량 강화	63
1. 데이터 품질관리 체계 수립	63
2. 업무규칙 가이드 수립	97
3. 데이터 품질관리 역량 교육	112
VII. 최종 평가 결과 및 결론	114
1. 적합성 평가 품질 진단 최종 결과	114
2. 결론	115

표 목 차

【표1-1】 공공데이터 정의	17
【표1-2】 공공데이터 제공 관리 단계.....	18
【표1-3】 공통 표준 용어 관리체계도.....	19
【표1-4】 공통 표준 용어 관리 원칙.....	19
【표1-5】 개방 표준 관리체계도.....	20
【표1-6】 공공데이터 개방 표준 제·개정 절차.....	20
【표1-7】 행정표준용어 관리체계.....	21
【표1-8】 행정표준용어/기관표준용어 관리 절차.....	21
【표2-1】 공공데이터 품질관리 단계 별 산출물.....	26
【표2-2】 데이터 품질관리 주요 지침.....	27
【표2-3】 품질관리수준 5단계 설명.....	28
【표2-4】 품질관리수준평가 영역 별 지표 설명.....	29
【표2-5】 과학기술정보통신부 평가대상 DB 및 점검 범위.....	29
【표3-1】 표준화 대상 DB 표준화 진척율.....	32
【표3-1】 표준화 대상 DB 표준화 진척 변동 사항.....	33
【표4-1】 ‘21년 평가대상 시스템 목록.....	43
【표4-2】 ‘21년 품질관리 수준평가 결과.....	43
【표4-3】 과기부 대표홈페이지 수준평가 결과.....	44
【표4-4】 국립전파연구원 방송통신통합정보시스템 수준평가 결과.....	44
【표4-5】 국립전파연구원 우주전파환경홈페이지 수준평가 결과.....	45
【표4-6】 중앙전파연구소 비정지위성전파감시시스템 수준평가 결과.....	45
【표4-7】 중앙전파연구소 전파관리정보시스템 수준평가 결과.....	46
【표4-8】 국립중앙과학관 국가자연사연구종합정보시스템 수준평가 결과.....	46
【표4-9】 국립중앙과학관 홈페이지 수준평가 결과.....	47
【표4-10】 우정사업본부 우체국금융시스템 수준평가 결과.....	47
【표4-11】 우정사업본부 인터넷우체국 수준평가 결과.....	50
【표5-1】 ‘20~22년 중장기 품질관리 목표	50
【표5-2】 데이터 품질 진단 영역 별 측정방법 및 평가기준	56
【표5-3】 소프트웨어 기술성 평가기준 적용 가이드	58
【표5-4】 데이터품질 관련 솔루션 목록	59
【표5-5】 공공기관 도입사례	59
【표5-6】 [붙임1] 국산 솔루션별 기능 비교.....	61
【표6-1】 [데이터 품질관리 총괄 책임자 및 총괄부서* 업무 정의	66
【표6-2】 업무팀 데이터 품질 담당자 역할	67
【표6-3】 데이터 품질관리 프로세스	67
【표6-4】 계획단계 품질관리 주요 산출물	70

【표6-5】 증장기 데이터 품질관리 계획 수립 (예시)	71
【표6-6】 진단 대상 DB 선정 기준	72
【표6-7】 데이터 표준화 목적	74
【표6-8】 외부 교육 과정 사례	75
【표6-9】 구축단계 품질관리 주요 산출물	78
【표6-10】 계획단계 표준화 vs 구축단계 표준화	79
【표6-11】 과학기술정보통신부 표준 용어 사전 vs 실제 DB 테이블의 컬럼	80
【표6-12】 표준 용어 사전에 건별 신규·변경·삭제일자를 관리하고 있는 사례	81
【표6-13】 연계데이터 목록 정의서 (예시)	83
【표6-14】 연계 실제 데이터 오류 유형 (예시)	83
【표6-15】 운영단계 품질관리 주요 산출물	85
【표6-16】 데이터 품질 진단 및 개선 절차 설명	85
【표6-17】 데이터 품질 진단 방법	86
【표6-18】 Inside-Out 방식과 Outside-In 방식 비교	87
【표6-19】 데이터 품질진단 범정부 진단기준 정의(예시)	88
【표6-20】 데이터 프로파일링 분석 내용(예시)	89
【표6-21】 데이터 오류 유발 원인(예시)	91
【표6-22】 정량적 업무 영향 분석(예시)	91
【표6-23】 개선기회(예시)	91
【표6-24】 데이터 보정 방법(예시)	94
【표6-25】 활용단계 품질관리 주요 산출물	95
【표6-26】 업무규칙 근거구분	99
【표6-27】 업무규칙 품질지표	100
【표6-28】 업무규칙 품질레벨 정의서	101
【표6-29】 데이터 오너쉽 분류기준	102
【표6-30】 Inside-Out 업무규칙 도출 순서	103
【표6-31】 Outside-In 업무규칙 도출 순서 (업무 이슈 근거)	104
【표6-32】 Outside-In 업무규칙 도출 순서 (법령/규정 근거)	105
【표6-33】 조직 담당자 간 업무규칙 도출 절차 설명	106
【표6-34】 업무규칙 도출 주체 별 주요 업무	107
【표6-35】 업무규칙 도출 방법 교육 문서 주요 내용	107

그 립 목 차

【그림2-1】 공공데이터 품질관리 단계.....	23
【그림2-2】 데이터 품질관리 개념도.....	24
【그림2-3】 데이터 생명주기 별 평가지표 및 판정기준.....	28
【그림2-4】 21년 공공데이터 품질관리 수준평가 추진 계획.....	28
【그림3-1】 프로젝트 관리 및 현황분석 WBS 진행 결과.....	30
【그림3-2】 미래모델 수립 및 교육 수행 WBS 진행 결과.....	30
【그림3-3】 기관 별 품질/표준 개선 결과도.....	32
【그림5-1】 데이터 품질관리 활동도.....	51
【그림5-2】 ‘20~22년 중장기 품질관리 목표.....	54
【그림5-3】 데이터 관리지침.....	56
【그림5-4】 진단 프로세스.....	57
【그림5-5】 DQI 및 CTQ 정의 절차.....	57
【그림6-1】 단계별 데이터 품질관리 활동 데이터 품질관리 구성 요소	64
【그림6-2】 단계별 데이터 품질관리 활동	65
【그림6-3】 단계별 품질관리 활동	68
【그림6-4】 계획단계 품질관리 주요 활동 내역	69
【그림6-5】 중점 관리 DB 선정 절차	72
【그림6-6】 데이터 표준관리 기준	74
【그림6-7】 데이터 표준화 요소 간 관계	75
【그림6-8】 구축단계 품질관리 주요 활동 내역	78
【그림6-9】 구축단계 표준적용률 개념	80
【그림6-10】 연계 관리 대상	82
【그림6-11】 운영단계 품질관리 주요 활동 내역	84
【그림6-12】 데이터 품질 진단 및 개선 절차(6단계)	85
【그림6-13】 진단 대상 정의(Define) 단계 수행 내용	86
【그림6-14】 진단 대상 정의(Define) 단계 상세 수행 내용	87
【그림6-15】 품질진단 실시(Measure) 단계 수행 내용	89
【그림6-16】 품질진단 실시(Measure) 단계 상세 수행 내용	89
【그림6-17】 품질진단 분석(Analyze) 단계 상세 수행 내용	90
【그림6-18】 개선계획 수립(Improvement Plan) 단계 수행 내용	92
【그림6-19】 개선계획 수립(Improvement Plan) 단계 상세 수행 내용	92
【그림6-20】 개선계획 도출(예시)	92
【그림6-21】 개선 실시(Implement) 단계 수행 내용	93
【그림6-22】 개선 실시(Implement) 단계 상세 수행 내용	93

【그림6-23】 품질 통제(Control) 단계 수행 내용	94
【그림6-24】 품질 통제(Control) 단계 상세 수행 내용	94
【그림6-25】 활용단계 품질관리 주요 활동 내역	95
【그림6-26】 품질 오류 신고 처리 절차	96
【그림6-27】 품질오류 신고관리 담당자별 역할	96
【그림6-28】 업무규칙 후보군 도출 방법	99
【그림6-29】 업무규칙 관리 절차도	102
【그림6-30】 Inside-Out방법 품질이슈를 파악 도출방식	103
【그림6-31】 Outside-In방법 품질이슈를 파악 도출방식 (업무 이슈 근거)	104
【그림6-32】 Outside-In방법 품질이슈를 파악 도출방식 (법령/규정 근거)	105
【그림6-33】 조직 담당자 간 업무규칙 도출 절차	106
【그림6-34】 업무규칙정의서 양식	108
【그림6-35】 업무규칙 변경/폐기 절차	108
【그림6-36】 업무규칙 진단 화면 입력 절차	109
【그림6-37】 업무규칙 진단 결과 확인 화면	110
【그림6-38】 업무규칙 진단 결과 엑셀 보고서 출력	111
【그림6-39】 공공데이터 품질관리 역량 향상 교육 1차 교육자료.....	112
【그림6-40】 공공데이터 품질관리 역량 향상 교육 2차 교육자료.....	112
【그림7-1】 고유업무 표준/품질 측정 결과.....	114
【그림7-2】 지원업무 표준/품질 측정 결과.....	114

요 약 문 (국 문)

본 사업 연구의 명칭은 “과기정통부 메타데이터 및 품질관리 도입 방안 연구“로써 주관기관 및 소속기관이 보유한 공공데이터의 품질을 제고하여 표준과 값 측면에서 데이터 품질의 신뢰도를 향상하고 개방의 기틀을 마련하여 4차 산업혁명의 핵심 아이콘인 데이터댐 구축에 과학 기초 데이터를 제공하기 위함에 있다.

이를 위하여 과기정통부 및 과기정통부의 4개 소속기관으로부터 고유업무 6개, 지원업무 4개 등 총 10개 업무 데이터를 관리하는 DB를 2021년 공공데이터 품질 수준평가 대상으로 선정하였다. 각 선정된 DB는 공공데이터 품질관리 수준평가 추진계획에 의거 고유업무인 경우 2개 영역(데이터 관리체계, 데이터 값 관리), 11개 평가지표를, 지원업무인 경우 1개 영역(데이터 값 관리) 6개 평가지표를 기준으로 평가를 진행한다.

이번 연구 과제 수행자는 데이터 품질관리 체계 수립을 위한 3개 평가지표(‘데이터 표준 확산’ , ‘데이터 구조 안정화’ , ‘데이터 연계 관리’), 데이터 값 관리 수립을 위한 4개 평가지표(‘데이터 품질진단’ , ‘품질진단 결과 조치’ , ‘데이터 오류율’ , ‘오류신고 요구사항 분석 및 개선’) 에 대해 각각 평가에 대응할 수 있도록 수행 평가 컨설팅 기술을 제공하는 것이 본 연구의 핵심 과제 내용이다.

하여 연구의 진행내용은 선정된 기관으로부터 표준 및 품질에 관한 데이터 관리 산출물을 받아 현황을 분석하고 범정부 표준, 품질 가이드에 맞게 개선하고 이를 반영할 수 있도록 기관 담당자를 독려하고 이를 최종 취합하여 과기정통부 기관 차원의 데이터 품질 가이드를 제공하였다.

또한 서브 과제로서 단기적 평가 대응이 아닌 기관 자체의 데이터 품질 운용 능력 향상을 위한 데이터 품질관리 도입 방안을 제시하였다.

본 연구과제가 성공적으로 끝날 수 있도록 적극 협조해준 기관 부처의 담당자에게 감사의 뜻을 전하며, 마지막으로 본 과제 미래에 이후 대응하여야 할 요소는 보고서 맨 끝에 결언 부분을 참고하기 바란다.

S U M M A R Y

(영 문 요 약 문)

The name of this project is “Study on Metadata and Quality Management of the Ministry of Science and ICT,” which aims to improve the quality of public data held by organizers and direct agencies, improve data quality in terms of standards and values, and lay the foundation for openness.

To this end, DB, which manages a total of 10 task data, including 6 unique tasks and 4 support tasks, was selected as the target of public data quality evaluation in 2021 from 4 direct agencies of the Ministry of Science and ICT. Each selected DB is evaluated based on two areas (data management system, data value management), 11 evaluation indicators for its own work, and six evaluation indicators for one area (data value management) for support work.

The core task of this study is to provide evaluation consulting technology to respond to three evaluation indicators (data standard spread, data structure stabilization, data linkage management), and four evaluation indicators (data quality diagnosis, quality diagnosis result measures), data error rate, and improvement).

Therefore, the progress of the study was to receive data management outputs on standards and quality from selected institutions, analyze the current status, encourage institutional personnel to improve and reflect pan-government standards and quality guides, and finally collect them to provide data quality guides at the Ministry of Science and ICT.

In addition, as a sub-task, a plan to introduce data quality management was proposed to improve the institution’s own data quality management capability, not a short-term evaluation response.

I would like to express my gratitude to the person in charge of the institutional ministries for their active cooperation in the successful completion of this research project, and lastly, please refer to the resolution at the end of the report for the factors to respond to in the future.

분 문

CHAPTER I

연구의 필요성 및 목적

1. 연구의 목적
2. 연구의 배경
3. 연구의 필요성
4. 연구의 내용 및 추진방법

CHAPTER I 연구의 필요성 및 목적

1. 연구의 목적

- 범정부 데이터 댐 사업에 과학기술 데이터 제공
 - 과학기술 및 정보통신 관련 데이터에 대한 표준단어, 용어, 도메인 등 표준을 수립하여 개방 공공데이터 품질 향상 및 데이터간의 융합에 활용할 수 있도록 데이터 활성화를 위한 생태계 기반 마련
- 고품질 과학기술 데이터 운영
 - 과학기술정보통신부 및 소속기관 정보시스템 데이터 표준 및 관리 방안 수립
 - 상시 데이터 품질관리 체계 마련을 위한 방안 수립

2. 연구의 배경

- 공공기관 데이터 품질관리 기반 강화
 - 2014년 공공데이터법 시행, 2015년 공공데이터 품질관리 수준평가 시범 적용 등 공공 데이터 개방을 위한 데이터 품질 제고에 대한 필요성이 대두되었으며 2016년 ~2020년 기간에 수준평가의 범위는 DB단위에서 기관단위로 확대
 - 특히 2021년에는 기관의 운영DB에 대한 *수준평가 범위가 확대 및 평가지표의 강화로 품질 평가 등급 상향화에 더욱 신경을 써야함
 - (*수준평가 범위) 고유업무 DB를 중심으로 품질관리를 강화(40% 수준)하고, 지원업무 DB에 대해 점진적 확대(30% 수준) 적용
 - 이에따라 과학기술 및 정보통신 관련 데이터에 대한 표준단어, 용어, 도메인 등 표준을 기준에 맞춰 수립하고 개방 공공데이터 품질 향상 및 데이터간의 융합에 활용할 수 있도록 기반 마련 필요
 - 이 연구를 기반으로 데이터 표준, 데이터 품질, 데이터 거버넌스 등 주요 데이터 부문 별 품질 관리 활동에 대한 역량을 강화하고자 함
- 운영 데이터베이스의 데이터 표준 적용 방안 수립
 - 공공기관의 데이터베이스 표준화 지침에 의거 공공기관은 기관표준 수립 시 공통표준용어 등을 우선적으로 적용하여야 하며 이에 대한 평가(데이터 표준 확산)를 받아야 함
 - 이에따라 공통표준을 최대한 준수할 수 있도록 표준화 가이드 정비 및 표준 적용 연구를 통해 시스템 영향도를 최소화하면서도 평가 등급이 상향될 수 있도록 조치

2. 연구의 필요성

- 기관 차원에서 통일된 데이터 표준/품질 방안 마련 필요
 - 행안부가 제시하는 범정부 공동 표준용어 준수 필요
 - 공공데이터 공통표준용어 (행안부고시 제2020-66호, 2020.12.10.) 및 공통표준용어 관리 기준(2020.12) 준수 필요
 - 국가가 정한 상위 참조 표준의 준수 필요
 - 과학기술정보통신부 차원의 기관표준 확장 갱신
 - 소속기관이 개별적으로 운영하는 비표준 데이터의 표준화 방향 수립
 - 소속기관에서 많이 사용되는 용어를 기관표준으로 지정

3. 연구의 내용 및 추진방법

- 연구의 목표
 - 개정된 공통표준용어를 적용하여 기관표준을 개선하고 표준 가이드를 정비
 - 품질평가 결과 품질 개선과제를 정의하여 품질 적용을 위한 중장기 계획 수립
- 연구 내용
 - 표준적용 방안 검토
 - 과학기술정보통신부 기관 차원의 데이터베이스 조사 및 분석을 통해 표준 단어, 용어, 도메인, 코드 메타데이터 정의
 - 기존 데이터베이스의 데이터 표준 적용 방안 수립
 - 공공데이터 품질수준 분석
 - 본부 및 소속기관 데이터의 중요도 분석을 통한 중점 데이터 선정 및 표준화 적용 현황 등 품질 수준 분석 및 개선 방안 수립
 - 데이터품질진단 도구 도입 검토 및 운영·활용 방안 수립
 - 공공데이터 관리 거버넌스 체계 수립
 - 공공데이터 관리지침(행안부고시 제2019-77호, 2019.9.3.)을 고려한 표준 명명 규칙, 표준관리 조직 및 절차 등 데이터 표준 가이드 정비
 - 공공데이터 품질관리 실무담당자 협의체 운영 방안 도출
 - 품질관리 역량 교육
 - 본부 및 소속기관 데이터 품질관리 역량 향상 방안 및 교육
 - 상반기 교육, 하반기 교육 등 데이터 품질관리 교육 2회 실시

□ 연구 범위

○ 데이터 현황조사

- 과학기술정보통신부의 21년도 평가대상 데이터베이스의 데이터 품질, 표준, 연계 등 현황조사 및 데이터 거버넌스 운영체계 현황조사
- 평가대상 데이터베이스와 연결된 정보시스템의 영향도 분석
- 최근 품질평가 대상 공공기관 중 우수 평가 기관에 대한 선진사례 벤치마킹 분석
- 상시 품질관리를 위한 품질진단 솔루션 비교분석 및 도입방안 검토
- 데이터 품질 및 표준에 대한 모든 관리 산출물에 대한 현황분석
- 과학기술정보통신부의 진단 대상 DB의 정의
- 모의 품질 진단 실시 (사전 진단 or 시범 진단) 및 결과 분석
- 각각 현황조사 결과 고려사항 도출

○ 방향성 수립

- 고려사항을 취합하여 품질 개선에 대한 시사점 도출 및 종합
- 품질 개선에 대한 시사점을 종합하고 개선기회 정의
- 현행 표준, 기관표준 및 DB표준에 대한 표준화 방향성 정의

○ 목표 수립

- 21년 공공데이터 품질관리 수준평가 대응 데이터 품질 지침 수립
 - 품질관리 정책 수립 및 원칙 정의
 - 적용 대상, 조직 R&R, 절차 등 정의
- 21년 공공데이터 품질관리 수준평가 대응 데이터 표준 지침 수립
 - 표준관리 정책 수립 원칙 정의
 - 적용 대상, 조직 R&R, 절차 등 정의
- 21년 공공데이터 정부 평가 지원
 - 품질관리 수준평가 대응 지원
 - 품질평가 2개 영역(데이터 관리체계, 데이터 값 관리) 11개 평가지표에 대한 품질 진단 실시
- 데이터 품질 오류신고 관리체계 수립
 - 담당자 R&R, 프로세스 등 정의

○ 이행계획 수립

- 전체 표준/품질/거버넌스 개선 과제 정의
- 최우선 과제 선정 및 시범 개선
- 중장기 이행계획 수립(연간, 단기, 중장기 이행계획을 수립)

□ 연구 수행 과제 개념

○ 공공데이터 관리

- “공공데이터”의 개념 및 대상

[표 1-1] 공공데이터 정의

“공공데이터”란 데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리된 자료 또는 정보로서 다음 각 목의 어느 하나에 해당하는 것을 말한다.(법 제2조제2호)

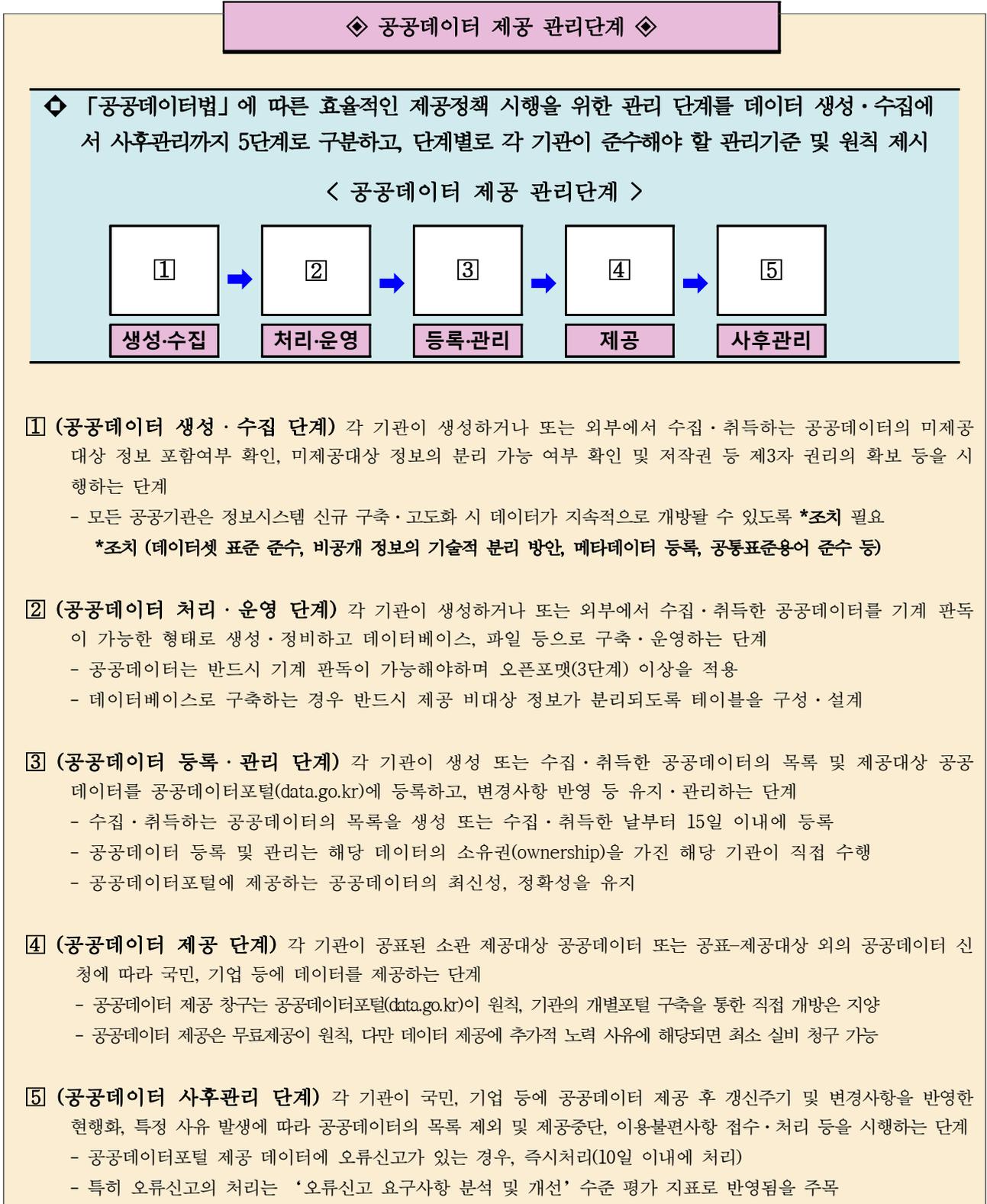


▶ 공공데이터의 정의 ◀	
「공공데이터법」 제2조제2호	내용
가. 「전자정부법」 제2조제6호에 따른 행정정보	• 행정기관등이 직무상 작성하거나 취득하여 관리하고 있는 자료로서 전자적 방식으로 처리되어 부호, 문자, 음성, 음향, 영상 등으로 표현된 것
나. 「국가정보화 기본법」 제3조제1호에 따른 정보 중 공공기관이 생산한 정보	• 특정 목적을 위하여 광(光) 또는 전자적 방식으로 처리되어 부호, 문자, 음성, 음향 및 영상 등으로 표현된 모든 종류의 자료 또는 지식 중 공공기관이 생산한 정보
다. 「공공기록물 관리에 관한 법률」 제20조제1항에 따른 전자기록물 중 대통령령으로 정하는 전자기록물	• 웹기록물 및 행정정보 데이터세트 등의 기록정보 자료
라. 그 밖에 대통령령으로 정하는 자료 또는 정보	

- 데이터는 개별 공공기관이 일상적 업무수행의 결과물로 생성 또는 수집·취득한 다양한 형태(텍스트, 수치, 이미지, 동영상, 오디오 등)의 모든 자료 또는 정보가 대상
- 공공공공데이터 목록(데이터셋) : 기관에서 관리하고 있는 원천데이터 내 유사데이터의 묶음 및 집합. 공공데이터포털에서는 데이터셋을 기준으로 관리하고 있으며, 이를 목록이라 함

- “공공데이터” 제공 절차

[표 1-2] 공공데이터 제공 관리 단계

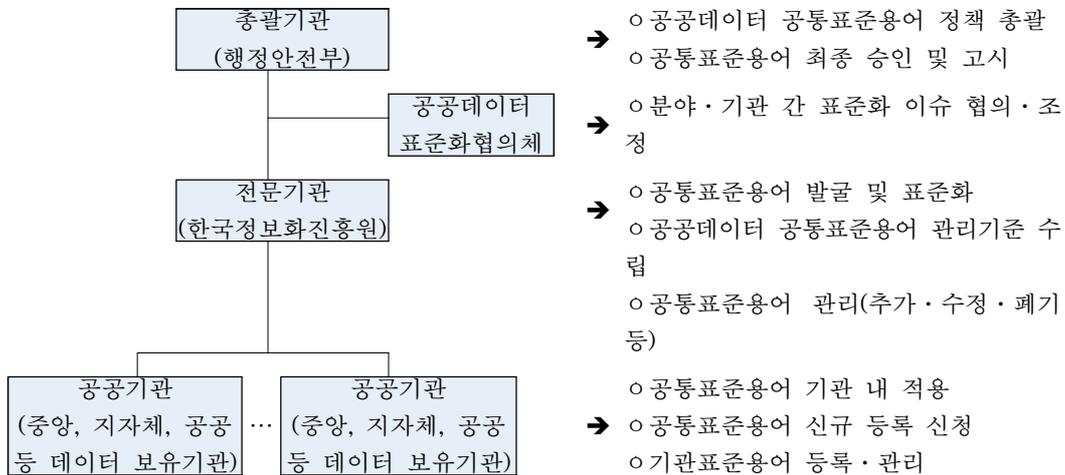


○ 공공데이터 공통표준

- 목적: 공공데이터법 제23조(공공데이터의 표준화), ‘공공기관의 데이터베이스 표준화 지침’에 따라, 공공데이터를 누구나 같은 의미로 이해하고 같은 방식으로 사용할 수 있도록 공통표준용어를 정의하여 기관 간 공동활용 효율성 극대화를 목적으로 함.

- 총괄기관-전문기관공통-주관기관 간 표준 용어 관리체계

[표 1-3] 공통 표준 용어 관리체계도



- 공통 표준 용어 관리 원칙

[표 1-4] 공통 표준 용어 관리 원칙

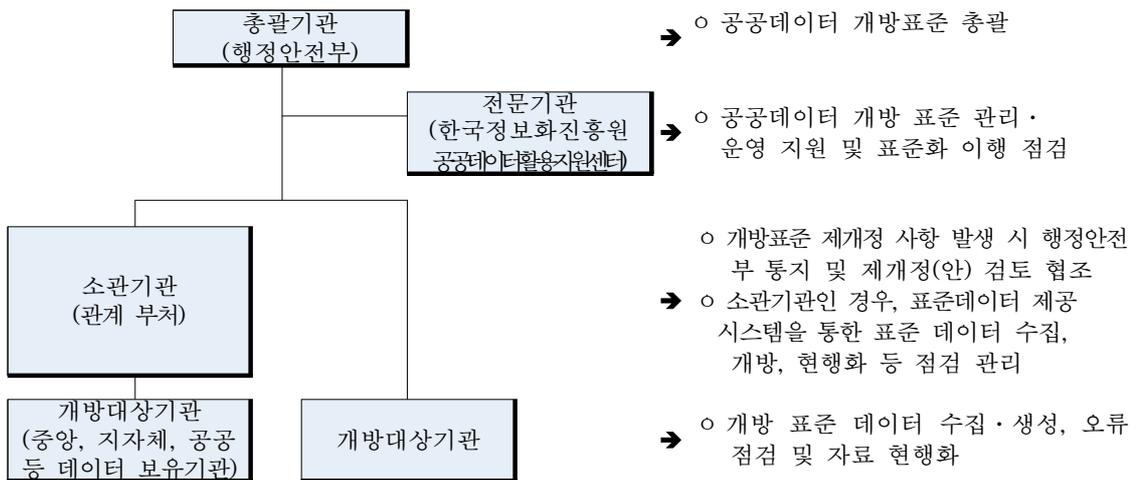
원칙1 (일관성)	□ 범정부 차원에서 정의한 공통표준용어는 공공기관이 구축하여 운영 중인 모든 공공데이터에 일관되게 적용하는 것을 원칙으로 한다.
원칙2 (체계성)	□ 공통표준용어는 범정부에서 표준으로 정한 단어와 도메인의 조합으로 이루어지며, 이를 체계적으로 관리하기 위해 공통표준단어와 공통표준도메인을 정의한다.
원칙3 (범용성)	□ 공공기관은 ‘공공데이터 공통표준용어 관리 기준’을 참조하여, 기관의 소관 공공데이터에 일관된 적용이 가능하도록 자체적인 데이터 표준 검토체계를 마련해야 한다.
원칙4 (표준성)	□ 공공기관은 기관의 표준용어를 정의할 경우, 범정부에서 제공하는 공통표준용어를 우선 적용해야 한다.
원칙5 (효율성)	□ 공공기관은 기관 메타데이터 관리시스템을 활용하여 기관표준용어를 마련, 등록 관리해야 하고, 공동 활용이 필요한 경우 공통표준용어로 등록될 수 있도록 상호 협력해야 한다.

○ 공공데이터 개방표준

- 목적: 공공데이터법 제23조 등에 따라 공공기관이 활용도가 높은 공공데이터를 제공할 때 적용할 표준을 정함으로써 데이터 가공·처리에 추가로 소요되는 비용을 절감하고 공공데이터의 민간 활용 제고를 목적으로 함

- 총괄기관-전문기관공통-주관기관 간 개방 표준 관리체계

[표 1-5] 개방 표준 관리체계도



- 표준데이터셋 생성 및 제공 시 준수사항

- 공통표준용어 준수
(공공데이터 개방 표준의 제공항목 및 항목 별 속성정보는 ‘공공데이터 공통표준용어’ 를 준수하여 정의)
- 표준표준용어 외 추가 정의
(공공데이터 공통표준용어’ 외 표준데이터셋에 공통적으로 적용이 가능한 항목을 추가로 정의)

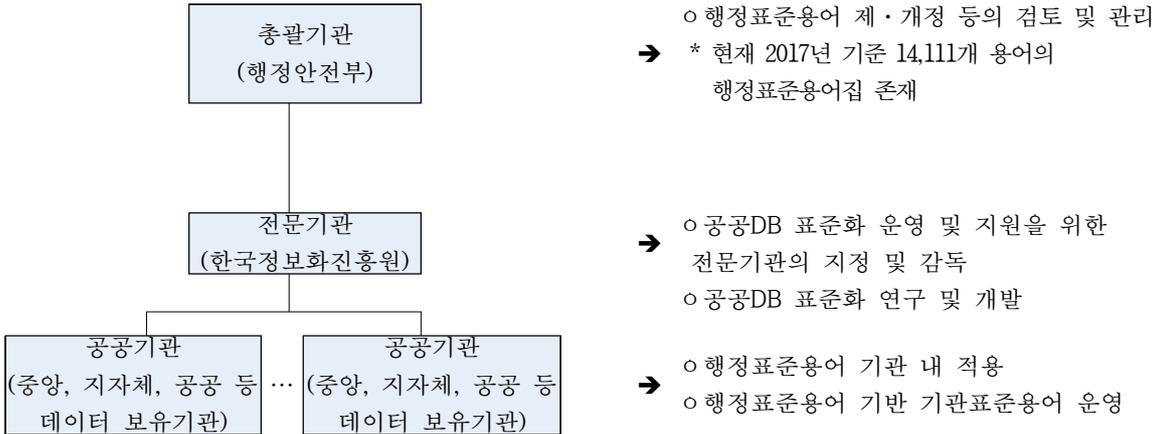
[표 1-6] 공공데이터 개방 표준 제·개정 절차

제·개정 절차	수행 주체
① 공공데이터 개방 표준 제개정 수요 제기	공공기관 (소관기관, 제공기관, 전문기관)
↓	
② 공공데이터 개방 표준 제개정 수요 검토	총괄기관(전문기관)
↓	
③ 공공데이터 개방 표준 제·개정(안) 마련 - 민간 및 공공의 데이터 활용·보유기관(이해관계자), 해당 분야 전문가의 의견을 수렴	총괄기관(전문기관), 소관기관, 제공기관, 수요제기기관 등
↓	
④ 공공데이터 개방 표준(안) 의견수렴 및 행정예고 - 제·개정(안)에 대한 의견 검토 및 반영	총괄기관
↓	
⑤ 공공데이터 개방 표준 확정 및 고시 - 관보 및 홈페이지, 공공데이터포털에 게재	총괄기관

○ 기관 표준 (행정표준 기반 기관 표준)

- 목적: 공공데이터법 제23조(공공데이터의 표준화), ‘공공기관의 데이터베이스 표준화 지침’에 따라, 공공기관이 생성 또는 취득하여 관리하는 데이터베이스의 표준화에 필요한 세부사항을 정하여 기관 표준용어 구성의 근간을 제공함을 목적으로 함.
- 행정표준용어 관리체계

[표 1-7] 행정표준용어 관리체계



- 기관 표준 운영 시 주의사항

- 행정표준용어 권고
(공공기관의 장은 기관표준을 위한 표준용어정의서 작성 시 행정표준용어사전에 등재된 용어를 우선적으로 사용하여야 하며, 행정표준용어사전에서 제공하지 않는 용어에 한하여 별도로 작성)
- 공통표준용어 준수
(공통표준용어 관리 원칙에 벗어나는 기관 표준 존재 시 공통표준에 맞게 수정 (표준 확산 정책))

[표 1-8] 행정표준용어/기관표준용어 관리 절차

절차	수행업무	수행주체
① 행정표준용어 신청	<ul style="list-style-type: none"> • 국가차원의 공동 활용에 필요한 행정표준용어의 발굴 및 효율적 관리를 위해 노력 • 국가차원에서 공동 활용이 필요하다고 판단되는 경우 행정표준용어 제·개정을 신청 	공공기관 (소관기관, 제공기관, 전문기관)
② 행정표준용어의 검토	<ul style="list-style-type: none"> • 행정표준용어의 제·개정 및 폐기 신청을 받은 경우 이를 접수하고 검토 	총괄기관
③ 행정표준용어의 제·개정 공지	<ul style="list-style-type: none"> • 제·개정된 행정표준용어를 정기적으로 년 1회 이상 공지 • 행정표준용어의 공지는 관보에 게시하거나, 행정안전부 홈페이지 또는 행정표준용어를 관리하는 시스템에 게시 	총괄기관

CHAPTER II

공공데이터 품질 정책 동향

1. 공공 데이터 품질관리
2. 공공 데이터 품질관리 단계 별 활동 및 산출물
3. 국가 차원의 데이터 품질관리
4. 공공데이터 품질관리 수준평가

CHAPTER II 데이터품질 정책 동향

1. 공공 데이터 품질관리

□ 수준평가 의미

- 공공데이터 품질관리 수준평가 제도 시행을 통해 기관이 적정 수준의 품질관리 활동의 수행 여부를 점검하여 개선사항 정비 등 품질개선 촉진

□ 공공데이터 품질관리 수준평가 제도

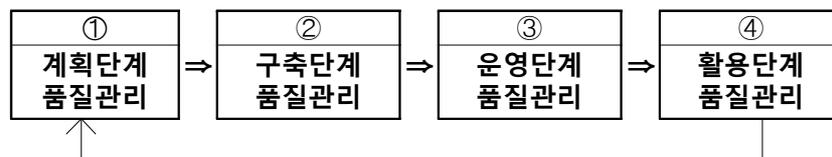
- 공공데이터의 적정한 품질수준 확보와 제공 촉진을 위하여 법 제22조에 따라 공공데이터 품질관리 수준평가를 실시
- 기관의 전사적 품질관리 활동 평가를 위해 기관 내 품질관리 업무의 명확한 역할과 책임을 정의하고 총괄부서와 DB 운영부서의 역할을 구분하여 평가

□ 공공데이터 품질관리 단계

○ 품질관리 단계 (정보 생명주기(Life Cycle) 기반)

- 공공기관이 생성·취득하여 관리하고 있는 공공데이터의 안정적 품질관리 및 적정한 품질수준을 확보하기 위한 관리단계를 계획-구축-운영-활용의 4단계로 구분하고, 각 단계별로 준수해야 할 관리기준과 원칙을 제시

[그림 2-1] 공공데이터 품질관리 단계



- (계획단계 품질관리) 각 기관의 품질관리를 위한 내부규정 마련, 품질관리 조직 구성, 공공데이터의 구축, 운영, 활용 단계별 품질관리 활동을 반영한 품질관리 계획을 수립
 - (20' 수준평가지표) 데이터 품질관리 기반, 데이터 품질관리 역량
 - (21' 수준평가지표) 품질관리 계획 수립, 예방적 품질진단

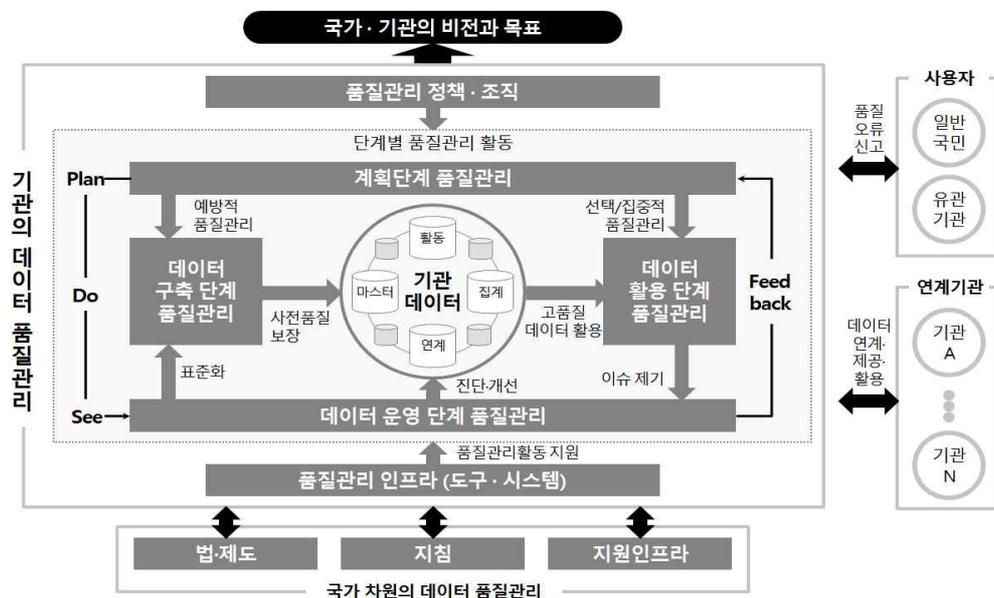
- (구축단계 품질관리) 각 기관이 데이터를 구축 시 준수해야 할 데이터 표준의 적용·점검, 데이터 구조의 일관성 확보, 데이터관리 산출물의 품질확보, 연계데이터 정합성 유지 등을 수행
 - (20' 수준평가지표) 데이터 표준 확산, 데이터 구조 안정화, 연계데이터 연계체계 정비
 - (21' 수준평가지표) 데이터 표준 확산, 데이터 구조 안정화, 데이터 연계 관리

- (운영단계 품질관리) 각 기관이 보유하고 있는 데이터를 운영하는 단계에서 준수해야 할 품질진단·개선, 주요 산출물의 현행화 점검 등을 수행
 - (20' 수준평가지표) 데이터 품질 진단 및 개선 관리, 데이터값 오류진단 기준
 - (21' 수준평가지표) 데이터 품질 진단, 품질진단 결과 조치, 데이터 오류율

- (활용단계 품질관리) 각 기관의 데이터를 활용하면서 발생하는 품질오류신고 및 차기 품질계획에 반영하기 위한 데이터 활용 사례, 개선사항 관리 등을 수행
 - (20' 수준평가지표) 데이터 활용성 제고
 - (21' 수준평가지표) 오류신고 요구사항 분석 및 개선, 개방데이터셋 오류 조치, 개방표준 데이터셋 현행화 관리

2. 공공 데이터 품질관리 단계 별 활동 및 산출물

[그림 2-2] 데이터 품질관리 개념도



□ 계획단계 품질관리 활동 점검사항

- 기관 데이터베이스 품질관리 조직 및 인력
- 기관 데이터베이스 품질목표 정의
- 기관 중점 데이터베이스 품질관리 대상 선정
- 기관 데이터베이스 품질 진단·개선 계획
- 기관 데이터베이스 표준화 방안
- 연계데이터 품질확보 방안
- 그 밖에 기관 데이터베이스 품질관리를 위해 필요한 사항

□ 구축단계 품질관리 활동 점검사항

- 데이터 표준화
 - 코드, 데이터사전, 도메인, 명명규칙 등의 표준화
 - 연계 데이터 표준화 (연계 대상 데이터 및 연계 절차, 규칙 등)
- 데이터 구조의 일관성 확보 및 오류 데이터의 입력 방지 등
- 연계 데이터 정합성 관리
- 데이터베이스 구축 관련 산출물 관리
 - 엔터티·속성 정의서, 테이블·컬럼 정의서, 논리·물리 ERD(Entity Relationship Diagram)
 - 연계 데이터 정의서, 업무규칙(BR; Business Rule) 정의서 등

□ 운영단계 품질관리 활동 점검사항

- 데이터 품질 진단 및 개선
- 데이터 품질관리 관련 산출물 점검
- 데이터 변경에 따른 문서 최신성 확보 및 이해관계자 통지

□ 활용단계 품질관리 활동 점검사항

- 데이터 품질 오류 신고 관리
 - 품질 오류 신고 접수 및 처리
 - 품질 오류 조치 결과 통보
 - 품질 오류 현황 및 보고 관리
- 데이터 활용 성과 평가

□ 공공데이터 품질관리 단계 별 산출물

[표 2-1] 공공데이터 품질관리 단계 별 산출물

단계	산출물 명
계획단계 품질관리	○ 데이터 품질관리 계획서
구축단계 품질관리	○ 데이터 사전 정의서 ○ 도메인 정의서 ○ 코드 정의서 ○ 엔터티 정의서 ○ 속성 정의서 ○ 논리데이터 모델 다이어그램 (논리ERD) ○ 테이블 정의서 ○ 컬럼 정의서 ○ 물리 데이터 모델 다이어그램 (물리ERD) ○ 데이터베이스(DB) 정의서 ○ 연계 데이터 목록 ○ 오너쉽 정의서 ○ 업무규칙 정의서 ○ 테이블 대 응용 프로그램 상관도
운영단계 품질관리	○ 품질 이슈 목록 ○ 요구사항 목록 ○ 품질 현황 분석서 ○ 품질 진단 계획서 ○ 업무규칙 진단 보고서 ○ 값·구조·표준화 진단 결과서 ○ 오류 유형 및 원인 분석 결과서 ○ 품질 개선 계획서 ○ 품질 개선 결과보고서
활용단계 품질관리	○ 데이터 품질 오류 신고서 ○ 데이터 오류 신고 및 처리 대장

- 데이터 산출물 점검 및 최신성 확보는 품질관리 계획 수립 시 포함된 산출물 점검 및 최신성 확보 계획에 의거 데이터 품질관리를 위한 관련 산출물 점검을 시행하고, 미흡한 산출물을 대상으로 최신성을 수행
- 매년 품질관리 계획 수립 단계에서 선정된 데이터베이스의 관련 산출물과 실제 구축 및 운영 현황 사이의 오류 또는 문제점을 식별하고, 이에 대한 개선 조치를 실행
- 각 기관은 매년 품질관리계획에 데이터 산출물 점검대상 DB를 품질관리계획에 반영하고, 운영단계에서 해당 DB의 산출물이 실제 데이터베이스의 내용과 일치하는지를 점검하여, 미흡한 산출물을 대상으로 현행화 작업을 수행
 - 산출물 점검은 해당 산출물이 존재하는지를 우선 확인하고, 주요 산출물에 대해서는 구축된 데이터베이스와 비교하여 최신성 여부를 점검
 - 각 기관은 산출물 점검에 따른 시정조치 사항을 관련부서(또는 담당자)에게 통보하여 조치토록 하며, 담당부서는 조치계획 및 조치결과를 통지

3. 국가 차원의 데이터 품질관리

□ 주요 규정 사항

- 국가 데이터 품질관리의 정책 방향을 규정하는 상위 수준의 “법·제도” 와 이를 뒷받침하는 데이터 품질 및 표준화 등과 관련된 세부 “지침” 그리고 품질관리 정책 실행을 지원하기 위한 “지원 인프라” 로 구성
- 품질관리 정책에 대한 법적 근거를 명확화하고 구체화하기 위해 공공기관의 정보화 추진 과정에서 생산·유통 또는 활용되는 데이터베이스 품질을 관리하기 위한 세부 사항을 규정하는 「공공데이터 관리지침」을 제정·고시
- 또한, 국가 차원에서는 데이터 품질관리 활동과 밀접하게 연관성을 갖는 데이터 표준화, 공동 활용 및 운영 등과 관련된 여러 지침을 국가 차원에서 규정

□ 데이터 품질관리 주요 지침

[표 2-2] 데이터 품질관리 주요 지침

지침	주요 내용
공공데이터 관리지침 (행정안전부 고시 제2019-71호, 2019. 9. 3. 개정)	<ul style="list-style-type: none"> ○ 제공정책의 효율적 시행을 위해 각 기관 및 업무담당자들이 준수해야 할 관리 원칙 및 기준 제시 ○ 추진근거 : 공공데이터법 ○ 공공데이터 생애주기별 관리기준, 공공데이터 개방 표준, 공공데이터 품질 관리 활동, 공공기관의 중복·유사 서비스 금지·정비 기준 및 절차를 제시
공공데이터 개방표준 (행정안전부고시 제 2020-54호)	<ul style="list-style-type: none"> ○ 공공데이터 제공 및 이용 활성화를 위하여 공공데이터 개방 시 적용하기 위한 공통 개방 기준과 데이터셋 분야별 개방 기준(제공항목, 속성정보, 제공형식 등) 및 기타 데이터 개방 표준을 정의 ○ 공공데이터법 제23조 ○ 공통 개방 기준, 행정표준용어 준수, 제공 방식·경로, 데이터셋 분야별 개방 기준 제시
공공기관의 데이터베이스 표준화 지침 (행정안전부고시 제2019-20호, 2019. 3. 20., 일부개정)	<ul style="list-style-type: none"> ○ 공공기관이 구축 운영하는 행정정보 데이터베이스의 설계·구축, 운영관리, 품질관리에 관한 표준을 규정 ○ 추진근거 : 전자정부법 제50조, 동법 시행령 제59조, 공공데이터법 제23조 ○ 공공DB표준화 관리체계 수립, 공공DB표준 수립, 정기점검및평가 기준 제시 ○ 주요 메타데이터 관리 서식 제공 <ul style="list-style-type: none"> - 공공데이터베이스 메타데이터 표준 관리 항목 - 표준용어정의서, 도메인정의서 - 논리데이터모델 다이어그램, 엔티티(개체)정의서, 애트리뷰트(속성)정의서 - 물리데이터모델 다이어그램, 테이블정의서, 컬럼정의서
행정기관 코드표준화 추진 지침 (행정안전부고시 제2021-6호, 2021.1.19., 일부개정)	<ul style="list-style-type: none"> ○ 행정기관 간 행정정보의 원활한 공동이용을 위해 각급 기관에서 행정표준 코드를 활용하고, 체계적인 관리를 하는데 필요한 세부 지침 ○ 추진근거 : 전자정부법 제50조, 동법 시행령 제59조 제1항, 제2항 ○ 행정기관에서 정보화를 추진할 경우 코드는 행정표준코드를 사용 ○ 행정표준코드의 최신자료는 행정표준코드관리시스템(www.code.go.kr) 이용

- 외 ‘행정기관 및 공공기관 정보시스템의 구축·운영 지침(행정안전부고시 제2021-3호, 2021)’ , ‘정보기술아키텍처 도입·운영 지침’ (행정안전부고시 제2018-2호, 2018. 1. 8) 참조



4. 공공데이터 품질관리 수준평가

□ 품질관리 수준평가 단계

- 데이터 생명주기를 고려한 [계획-구축-운영-활용] 영역으로 구성, 각 영역별 평가항목의 충족여부를 판단하여 5단계로 결과 도출

[그림 2-3] 데이터 생명주기 별 평가지표 및 판정기준



- 2020년 ‘3개 영역(기관, DB, 데이터 품질 관리) 9개 평가지표’ 에서 2021년 ‘2개 영역(데이터 관리체계, 데이터 값 관리) 11개 평가지표’ 로 평가 영역 통폐합 및 평가 지표 확대

[그림 2-4] 21년 공공데이터 품질관리 수준평가 추진 계획



- 평가등급은 ‘1레벨(도입前)’ → ‘2레벨(도입)’ → ‘3레벨(관리화)’ → ‘4레벨(체계화)’ → ‘5레벨(최적화)’ 의 5단계 등급 체계로 정의

[표 2-3] 품질관리수준 5단계 설명

등급	설명
5레벨	조직 전체의 데이터 품질관리 활동의 선순환 체계가 확립되고, 이를 통해 공공데이터의 안정적 품질 향상 및 유지가 보장되는 수준
4레벨	조직 전체의 데이터 품질관리 프로세스가 이행되고, 데이터 품질관리 활동 수행에 따른 성과측정이 가능한 수준
3레벨	데이터 품질관리를 위한 전반적인 활동들이 관리 및 통제되어, 이를 통해 데이터 품질 향상이 가능한 수준
2레벨	데이터 품질관리가 인식되고, 품질진단에 따른 개선 조치 등 기본적인 품질관리 활동들을 수행하는 수준
1레벨	데이터 품질관리가 인식이 미흡하여 기본적인 품질관리 활동의 수행이 불가능하거나 부분적인 품질관리 활동만 수행되는 수준

□ 품질관리 수준평가 범위

○ 2021년 공공데이터 품질관리 수준평가 지표 및 배점

[표 2-4] 품질관리수준평가 영역 별 지표 설명

				사업범위
영역	지표	세부지표	배점	
데이터 관리 체계 (32)	① 품질관리 계획 수립(2)	<ul style="list-style-type: none"> 기관 데이터 품질관리 계획 수립 및 이행 ※ 각 기관에서 제출한 '공공데이터의 제공 및 이용 활성화에 관한 시행계획'에 포함된 내역으로 평가 예정이며, 시행계획 제출 대상이 아닌 기관은 N/A 처리. '22년부터는 시행계획 작성 시 품질관리 계획 수립과 관련하여 별도 안내 예정 	2	
	② 예방적 품질관리 진단(10)	<ul style="list-style-type: none"> 예방적 품질관리 진단 예방적 품질관리 진단결과 반영 	3	
	③ 데이터 표준 확산(10)	<ul style="list-style-type: none"> 데이터 표준 정의 데이터 표준 적용율 	4	
	④ 데이터 구조 안정화(5)	<ul style="list-style-type: none"> 데이터 구조 정의 및 산출물 관리 데이터 구조 산출물 현행화 	2	
	⑤ 데이터 연계 관리(5)	<ul style="list-style-type: none"> 연계데이터 관리 정보 정의 연계데이터 송수신 내역 관리 	2	
데이터 값 관리 (68)	⑥ 데이터 품질진단(15)	<ul style="list-style-type: none"> 품질진단 기준 정의 및 진단 업무규정에 따른 품질진단 규칙(BR) 정의 및 진단 	5	
	⑦ 품질진단 결과 조치(23)	<ul style="list-style-type: none"> 품질진단에 따른 개선계획 수립 및 개선과제 정의 개선조치 이행률 	8	
	⑧ 데이터 오류율(15)	<ul style="list-style-type: none"> 데이터 오류율 	15	
	⑨ 오류신고 요구사항 분석 및 개선(4)	<ul style="list-style-type: none"> 데이터 품질 오류 분석 및 문제점 개선 	4	
	⑩ 개방 데이터셋 오류조치(4)	<ul style="list-style-type: none"> 개방 데이터셋 품질오류 조치결과 오픈API서비스 품질오류 조치결과 	2	
	⑪ 개방표준 데이터셋 현행화 관리(7)	<ul style="list-style-type: none"> 개방표준 데이터셋 현행화 관리 ※ 개방표준이 없는 기관은 [개방 데이터셋 오류조치] 지표의 평가 점수를 변경(4점→11점)하여 적용 	7	

○ 2021년도 선정된 과학기술정보통신부 평가대상 DB

[표 2-5] 과학기술정보통신부 평가대상 DB 및 점검 범위

	DB유형	기관명	시스템명	DB명	점검 범위	
1	고유업무	과학기술정보통신부 우정사업본부	우체국금융시스템	전자금융(FNIPD)	표준 점검 대상	품질 점검 대상
2		과학기술정보통신부 우정사업본부	인터넷우체국(epost)	통합DB1		
3		과학기술정보통신부 중앙전파관리소	비정지위성전파감시시스템	비정지위성전파감시 DB		
4		과학기술정보통신부 중앙전파관리소	전파관리정보시스템	전파관리 DB		
5		과학기술정보통신부 국립전파연구원	방송통신통합정보시스템	방송통신통합 DB(적합성평가)		
6		과학기술정보통신부 국립중앙과학관	국가자연사연구종합정보시스템	US_NTISAPP		
7	지원업무	과학기술정보통신부	대표홈페이지	대표홈페이지 DB		
8		과학기술정보통신부 국립전파연구원	우주전파환경홈페이지	우주전파환경홈페이지 DB		
9		과학기술정보통신부 국립중앙과학관	국립중앙과학관 홈페이지	NSM2018		



CHAPTER III

기관 데이터 품질 연구 진행 결과

1. '21년 공공데이터 품질관리 추진과제 점검 결과
2. '21년 공공데이터 품질관리 선정 DB 내역
3. 기관 별 표준화 진행결과

CHAPTER III 기관 데이터 품질 연구 진행 결과

1. '21년 공공데이터 품질관리 추진과제 점검 결과

□ WBS 진행결과

○ 프로젝트 관리 및 현황분석

[그림 3-1] 프로젝트 관리 및 현황분석 WBS 진행 결과

작업*	시작일*	완료일*	Calendar	총		계획		담당	산출물	계획	실적*
				작업량	작업량	기간	기간				
프로젝트 관리	2021-06-01	2021-11-26		86	77	126	55			100.00%	100.00%
자료수집	2021-06-01	2021-06-30		22	22	22	22	공동	요청자료 목록	100.00%	100.00%
데이터관리 준비	2021-06-01	2021-06-30		22	22	22	22	MIT	사업 관리/수행 맵플릿	100.00%	100.00%
사업환경 준비	2021-06-01	2021-06-30		22	22	22	22	에버블류		100.00%	100.00%
사업수행계획 확정	2021-06-01	2021-06-11		9	9	9	9	에버블류	사업수행계획서	100.00%	100.00%
착수보고	2021-06-18	2021-06-18		1	1	1	1	에버블류	착수보고서	100.00%	100.00%
중간보고	2021-08-16	2021-08-20		5	1	5	1	에버블류	중간보고서	100.00%	100.00%
완료보고	2021-11-22	2021-11-26		5	0	5	0	에버블류	완료보고서	100.00%	100.00%
고객 프로젝트 지원	2021-06-07	2021-09-24		127	81	77	51			100.00%	100.00%
현행산출물 제공	2021-06-07	2021-06-18		10	10	10	10	고객		100.00%	100.00%
진단환경 안내	2021-06-07	2021-06-18		10	10	10	10	고객		100.00%	100.00%
품질진단 대상 정의	2021-06-14	2021-07-02		15	15	15	15	고객		100.00%	100.00%
업무규칙 정의	2021-06-14	2021-07-02		15	15	15	15	고객		100.00%	100.00%
오류데이터 개선	2021-07-19	2021-09-03		35	21	35	21	고객		100.00%	100.00%
오류신고 요구사항 수립	2021-08-02	2021-08-13		10	10	10	10	고객		100.00%	100.00%
연계데이터 정의	2021-08-23	2021-09-03		10	0	10	0	고객		100.00%	100.00%
문제점 개선	2021-09-06	2021-09-17		10	0	10	0	고객		100.00%	100.00%
경방데이터셋 오류개선	2021-09-06	2021-09-24		12	0	12	0	고객		100.00%	100.00%
현황분석	2021-06-14	2021-08-13		112	112	45	45			100.00%	100.00%
표준분석	2021-06-21	2021-07-02		10	10	10	10			100.00%	100.00%
공동표준용어 분석	2021-06-21	2021-06-25		5	5	5	5	에버블류	현황분석서	100.00%	100.00%
공동표준 거점표준 비교 분석	2021-06-28	2021-07-02		5	5	5	5	에버블류	현황분석서	100.00%	100.00%
품질분석	2021-06-14	2021-07-23		70	70	30	30			100.00%	100.00%
중성관리 DB 선정	2021-06-14	2021-06-18		5	5	5	5	MIT	현황분석서	100.00%	100.00%
데이터 구조 산출물 분석	2021-06-14	2021-07-09		20	20	20	20	MIT	현황분석서	100.00%	100.00%
품질 진단 수행(프로파일링)	2021-06-14	2021-07-02		15	15	15	15	MIT	현황분석서	100.00%	100.00%
품질 진단 수행(업무규칙)	2021-06-21	2021-07-09		15	15	15	15	MIT	현황분석서	100.00%	100.00%
품질진단 결과 보고	2021-07-05	2021-07-23		15	15	15	15	MIT	현황분석서	100.00%	100.00%
거버넌스분석	2021-07-01	2021-08-13		32	32	32	32			100.00%	100.00%
품질 실무 협의제 분석	2021-07-01	2021-07-16		12	12	12	12	에버블류	현황분석서	100.00%	100.00%
표준 및 품질 가이드 분석	2021-07-15	2021-07-30		10	10	10	10	MIT	현황분석서	100.00%	100.00%
품질진단 솔루션 비교 분석	2021-08-02	2021-08-15		10	10	10	10	MIT	현황분석서	100.00%	100.00%

○ 미래모델 수립 및 교육 수행 결과

[그림 3-2] 미래모델 수립 및 교육 수행 WBS 진행 결과

작업*	시작일*	완료일*	Calendar	총		계획		담당	산출물	계획	실적*
				작업량	작업량	기간	기간				
미래모델	2021-07-05	2021-11-12		158	42	92	31			100.00%	100.00%
표준수립	2021-07-05	2021-09-17		55	31	55	31			100.00%	100.00%
DB데이터표준 정의	2021-07-05	2021-08-13		30	30	30	30	에버블류	미래모델수립서	100.00%	100.00%
거점데이터표준 정의	2021-08-16	2021-09-03		15	1	15	1	에버블류	미래모델수립서	100.00%	100.00%
데이터표준적용을 산정내역	2021-09-06	2021-09-17		10	0	10	0	MIT	미래모델수립서	100.00%	100.00%
품질수립	2021-07-26	2021-10-01		71	11	47	16			100.00%	100.00%
품질 진단 개선 계획 수립	2021-07-26	2021-08-06		10	10	10	10	MIT	미래모델수립서	100.00%	100.00%
데이터 구조 산출물 현행화	2021-08-23	2021-09-03		10	0	10	0	MIT	미래모델수립서	100.00%	100.00%
데이터 연계 산출물 현행화	2021-09-06	2021-09-24		12	0	12	0	MIT	미래모델수립서	100.00%	100.00%
데이터 오류율 측정	2021-09-06	2021-09-24		12	0	12	0	MIT	미래모델수립서	100.00%	100.00%
개선 조치 수행	2021-09-27	2021-10-01		5	0	5	0	MIT	미래모델수립서	100.00%	100.00%
개선 제시 품질 과제 수행	2021-08-16	2021-09-03		15	1	15	1	MIT	미래모델수립서	100.00%	100.00%
문제점 개선 보고	2021-09-20	2021-10-01		7	0	7	0	MIT	미래모델수립서	100.00%	100.00%
거버넌스수립	2021-08-30	2021-11-12		32	0	52	0			100.00%	100.00%
품질 협의제 운영 방안 도출	2021-10-04	2021-10-15		10	0	10	0	에버블류	미래모델수립서	100.00%	100.00%
표준 및 품질 가이드 정비	2021-10-16	2021-10-29		10	0	10	0	에버블류	미래모델수립서	100.00%	100.00%
품질진단 솔루션 고려사항 도출	2021-11-01	2021-11-12		10	0	10	0	에버블류	미래모델수립서	100.00%	100.00%
품질관리 역량 교육 2차	2021-08-30	2021-08-31		2	0	2	0	MIT	미래모델수립서	100.00%	100.00%
교육계획	2021-07-29	2021-08-31		4	2		13			100.00%	100.00%
1차교육	2021-06-29	2021-06-29		2	2	2	2			100.00%	100.00%
품질관리 역량 교육 1차	2021-06-29	2021-06-29		2	2	2	2	MIT	교육자료	100.00%	100.00%
2차교육	2021-08-18	2021-08-18		2	0	2	0			100.00%	100.00%
품질관리 역량 교육 2차	2021-08-18	2021-08-18		2	0	2	0	MIT	교육자료	100.00%	100.00%

□ 기관 별 데이터 품질/표준 개선 수행 결과

○ 프로젝트 관리 및 현황분석

[그림 3-3] 기관 별 품질/표준 개선 결과도

● 최종완료 ○ 1차진단완료 △미진행

평가 DB유형	평가지표	데이터 관리체계			데이터 값 관리		
		③표준	④구조	⑤연계	⑥품질진단	⑦진단결과 조치	⑧오류율
고유업무 (40%)	개별	●	●	●	●	●	●
	표준	● (소관부처)	● (소관부처)	● (소관부처)	● (중앙 및 기초)	● (중앙 및 기초)	● (중앙 및 기초)
지원업무 (30%)	개별	×	×	×	●	●	●
	표준	×	×	×	● (중앙 및 기초)	● (중앙 및 기초)	● (중앙 및 기초)

No	소속기관 명	시스템 명	데이터 관리체계			데이터 값 관리		
			데이터 표준 확산 (에버블류)	데이터 구조 안정화 (에버블류)	데이터 연계 관리 (과기정통부)	데이터 품질 진단 및 오류율 (MITSOFT)	품질 진단 결과 조치 (MITSOFT)	오류 신고 분석 및 개선 (과기정통부)
1	우정사업본부	우체국금융시스템	●	●	●	●	●	●
2	우정사업본부	인터넷우체국(epost)	●	●	●	●	●	●
3	중앙전파관리소	비정지위성전파감시시스템	●	●	●	●	●	●
4	중앙전파관리소	전파관리정보시스템	●	●	●	●	●	●
5	국립전파연구원	방송통신통합정보시스템	●	●	●	●	●	●
6	국립중앙과학관	국가자연사연구종합정보시스템	●	●	●	●	●	●
7	과학기술정보통신부	대표홈페이지	비대상 영역			●	●	●
8	국립전파연구원	우주전파환경홈페이지				●	●	●
9	국립중앙과학관	국립중앙과학관 홈페이지				●	●	●

○ 표준 평가 대상 DB 표준화 완료 보고

- DB 표준화 대상 6개의 DB표준화 진척율 (2021.11.20. 기준)

[표 3-1] 표준화 대상 DB 표준화 진척율

기관명	시스템명 DB명	완료여부 진척율	기관 별 담당자 및 완료일자	
				1
2	과학기술정보통신부 우정사업본부	인터넷우체국(epost) 통합DB1	완료 100%	담당자: 이효진 주무관(044-200-8182) / 한상인 주무관 (061-338-2206) ●테이블정의서 수정본 / 컬럼정의서 수정본 완료 (2021.8.30) ●표준정의서 중 표준용어, 표준도메인 작성 완료 (2021.8.30)
3	과학기술정보통신부 중앙전파관리소	비정지위성전파감시시스템 비정지위성전파감시DB	완료 100%	담당자: 허수진 주무관 (061-644-5961) ●테이블정의서 수정본 / 컬럼정의서 수정본 완료 (2021.9.10) ●표준정의서 중 표준 용어, 표준 도메인 작성 완료(2021.9.10)
4	과학기술정보통신부 중앙전파관리소	전파관리정보시스템 전파관리DB	완료 100%	담당자: 박홍선주무관 (02-3400-2445) ●테이블정의서 수정본 / 컬럼정의서 수정본 완료 (2021.9.20) ●표준정의서 중 표준 용어, 표준 도메인 작성 완료 (2021.9.20)
5	과학기술정보통신부 국립전파연구원	방송통신통합정보시스템 방송통신통합DB(적합성평가)	완료 100%	담당자: 전태식 주무관 (061-338-4934) / 신철희 차장(061-337-9417) ●테이블정의서 수정본 / 컬럼정의서 수정본 완료 (2021.9.30) ●표준정의서 중 표준 용어, 표준 도메인 작성 완료 (2021.9.30)
6	과학기술정보통신부 국립중앙과학관	국가자연사연구종합정보시스템 US_NTISAPP	완료 100%	담당자: 서정원 주무관 (042-601-8088) ●테이블정의서 수정본 / 컬럼정의서 수정본 완료 (2021.10.10) ●표준정의서 중 표준용어, 표준도메인 작성 완료 (2021.10.10)

- DB 표준화 대상 6개의 DB표준화 진척 추가 변동 사항

[표 3-2] 표준화 대상 DB 표준화 진척 변동 사항

기관명	DB명	테이블수	컬럼수	표준용어수 [신규추가]	표준도메인수 [신규추가]	표준코드수	인스턴스개수	특이사항
1	우정사업본부	122	1,883	32,450	4,615	5,663	51,821	기존표준사용
	전자금융(FNIPD) DB			[355]	[67]			
2	우정사업본부	140	2,417	[1,250]	[171]	140	2,278	신규표준제정
	통합 DB1							
3	중앙전파관리소	355	3,738	999	243	10	24	기존표준사용
	비정지위성전파감시 DB			[150]				
4	중앙전파관리소	128	1,186	647	105	13	89	기존표준사용
	전파관리 DB			[192]	[33]			
5	국립전파연구원	31	597	[451]	[72]	43	1,256	신규표준제정
	방송통신통합정보시스템 방송통신통합DB(적합성평가)							
6	국립중앙과학관	51	706	[417]	[74]	32	392	신규표준제정

2. '21년 공공데이터 품질관리 선정 DB 내역

□ 우체국금융시스템 전자금융(FNIPD) 데이터베이스 정의서

기관명	부서명	적용업무	관련법령	논리DB명	물리DB명
우정사업정보센터	예금정보과	전자금융이용자원장 관리 등	전자금융거래법 제22조(전자금융거래기록의 생성·보존 및 파기)	전자금융AP/DB	FNIPD
구축일자	DB설명		DBMS정보	운영체제정보	수집제외사유
2010-08-25	전자금융이용자원장 관리 등		ORACLE 10G	AIX 6.1	해당없음

□ 인터넷우체국(epost) 통합DB1 데이터베이스 정의서

기관명	부서명	적용업무	관련법령	논리DB명	물리DB명
과학기술정보통신부 우정사업본부 우정사업정보센터	우편정보과 우편기반팀	인터넷우체국 공통	우편법 전자정부법 소프트웨어산업진흥법	통합DB1	PPOSTDB1
구축일자	DB설명		DBMS정보	운영체제정보	수집제외사유
2013.07.25	인터넷우체국 공통정보(회원, 우편번호, 주소, OPENAPI, 알뜰폰)를 관리하는 DB		Oracle 11g	AIX 6.1	해당없음

□ 비정지위성전파감시DB 데이터베이스 정의서

기관명	부서명	적용업무	관련법령	논리DB명	물리DB명
과학기술정보통신부 중앙전파관리소	위성전 파감시 센터 위성관 리과	비정지위성 전파감시시스템정보 DB	전파법, 전파규칙(ITU-RR)	비정지위성 NGRMDB	NGRMDB
구축일자	DB설명		DBMS정보	운영체제정보	수집제외사유
2016.06.29	비정지위성전파감시시스템의 주 기능은 비정지궤도위성을 감시하여 그 전파의 특성을 측정하여 기록하고 보고하는 것이다. 사용자는 감시하고자 하는 위성망정보(SRS : Space Radiocommunication Stations, IFIC : International Frequency Information Circular, GIMS : Geographic Information Management System, NORAD/TLEs : North American Aerospace Defense Command / Two Line Elements, 전자 파일 등)를 비정지위성전파감시시스템 소프트웨어를 통해 간편하게 등록할 수 있으며 필요에 따라 등록된 정보를 수정 및 삭제할 수 있도록 한다.		ORACLE12C	Microsoft Windows Server 2012	해당없음

□ 전파관리DB 데이터베이스 정의서

기관명	부서명	적용업무	관련법령	논리DB명	물리DB명
중앙전파관리소	전파계 획과	전지소감시시스템 지 휘통제 및 사용자, 기초자료 관리	전파법	통합DB	CRMO DB
구축일자	DB설명		DBMS정보	운영체제정보	수집제외사유
2015	전파관리정보시스템 운영 및 관리		ORACLE 11G	Solaris 11(SunOS 5.11)	해당없음

□ 방송통신통합DB(적합성평가) 데이터베이스 정의서

기관명	부서명	적용업무	관련법령	논리DB명	물리DB명
국립전파연구원	국립전 파연구 원 정보 운영과	방송통신 통합시스템	방송통신기자재 등 적합성 평가에 관한 고시 제8조	ICRS	적합성평가
구축일자	DB설명		DBMS정보	운영체제정보	수집제외사유
2010.10.23	방송통신 기자재 등의 인증관리 및 수입통관		Oracle 11g	AIX	해당없음

□ US_NTISAPP DB 데이터베이스 정의서

기관명	부서명	적용업무	관련법령	논리DB명	물리DB명
과학기술정보통신부 국립중앙과학관	연구개발과	종정보(학명,국명,생태적,일반적특징), 표본·관찰·생태정보(지역,날짜,채집자,동정자,GPS), 이미지, 동영상, 음성, DNA 바코드	국립중앙과학관 자연사정보콘텐츠 제공및관리를위한규정	국가자연사 연구종합정보 시스템	NARIS
구축일자	DB설명		DBMS정보	운영체제정보	수집제외사유
2004.12	국립중앙과학관은 생물자원 주권실현을위해 기존 소장중인 자료 또는 신규로 채집하는 표본 및 생물자원정보를 메타데이터 구축 지침에 의거하여 표준화 및 디지털DB화하고, 국가자연사 연구종합정보시스템(http://www.naris.go.kr)을 통하여 자연사 및 생물종에대한 종설명정보등 기타 멀티미디어 정보를 제공		Oracle 10g Enterprise Edition	Solaris9	해당없음

3. 기관 별 표준화 진행결과

□ 6개 기관 별 표준화 결과는 최종보고서와 같이 제출되는 PPT 형태의 완료보고서 및 첨부되는 상세한 산출물 내용을 참고하면 된다.

□ 우체국금융시스템 전자금융(FNIPD) 표준화 결과

○ DB표준용어정의서 → 355개 신규표준용어 추가

- 신규 표준 용어 추가 예시

순서	표준용어명
32096	메시지함고유번호
32097	푸시고유번호
32098	앱아이디
32099	사용자아이디
32100	전송메시지
32101	확장데이터
...	...
32449	MMS맵
32450	등록년월일시

○ DB표준도메인정의서 → 67개 신규도메인 추가

- 신규 표준 도메인 추가 예시

순서	도메인분류명	도메인명
4549	번호	IDV255
4550	코드	보안카드종류
4551	번호	조작자번호
4552	코드	전자금융상태코드
4553	코드	보안카드사용용도
4554	코드	TOKEN아이디
...
4614	코드	VM모바일입금계좌서비스신청구분
4615	코드	인터넷입금계좌서비스관리구분

○ 코드도메인에 해당하는 표준용어는 ‘표준코드명’ 이 매핑되도록 보완

- 코드 도메인 표준용어 매핑 예시

표준영어명	영문약어명
11분별사용여부	ELDR_UTLZ_YN
120일한도유무	D120_LIMIT_EXNX
180일경과시입원여부	D180_PASS_HOUR_ADMS_YN
1개월적부거절여부	MON1_PPON_RJCT_YN
1개월적부여부	MON1_PPON_YN
1건거래구분	CNT1_TX_DVSN
1년적부거절여부	YY1_PPON_RJCT_YN
1년적부여부	YY1_PPON_YN
1단계결제유형	STG1_APVL_TYPE

□ 인터넷우체국(epost) 통합DB1 표준화 결과

○ DB표준용어정의서 → 1,250개 신규표준용어 추가

- 신규 표준 용어 추가 예시

순서	표준용어명
1242	우체국우편번호
1243	우편일련번호
1244	주소우편번호
1245	우편번호주소
1246	우편번호메일수신여부
1247	임시6자리우편번호
1248	우편번호변경순번
1249	우편번호일련번호
1250	우체국우편코드

○ DB표준도메인정의서 → 171개 신규도메인 추가

- 신규 표준 도메인 추가 예시

순서	도메인분류명	도메인명
163	내용	내용V1500
164	내용	내용V2000
165	내용	내용V3000
166	내용	내용V4000
167	URL	URL128
168	URL	URL256
169	URL	URL400
170	URL	URL1100
171	내용	내용CLOB

○ 코드명(한글)은 표준용어명, 코드명(영문)은 영문약어명을 반드시 기입되도록 개선

- 코드 도메인 표준용어 매핑 예시

코드명(한글)	데이터타입	데이터길이	코드값	코드의미
인터넷우체국서비스별이용방법	VARCHAR2	5	C01	우편서비스
인터넷우체국서비스별이용방법	VARCHAR2	5	C02	우체국소포
인터넷우체국서비스별이용방법	VARCHAR2	5	C03	EMS
인터넷우체국서비스별이용방법	VARCHAR2	5	C04	우표
인터넷우체국서비스별이용방법	VARCHAR2	5	C05	우체국쇼핑
인터넷우체국서비스별이용방법	VARCHAR2	5	C06	우체국장터
인터넷우체국서비스별이용방법	VARCHAR2	5	C07	우체국B2B
인터넷우체국서비스별이용방법	VARCHAR2	5	C08	전통시장
인터넷우체국서비스별이용방법	VARCHAR2	5	C09	이메일

□ 비정지위성전파감시DB 표준화 결과

○ DB표준용어정의서 → 150개 신규표준용어 추가

- 신규 표준 용어 추가 예시

순서	표준용어명
850	A6A3플래그
851	BR의견
852	BSTAR
853	CI값
854	CN총합수
855	EPCI최대수
...	...
998	허용오차
999	효율

○ DB표준도메인정의서 → 신규추가 없음(기존 도메인 변화 없음)

- 신규 표준 도메인 추가 예시

순서	도메인분류명	도메인명
235	코드	코드V3
236	코드	코드V30
237	코드	코드V4
238	코드	코드V5
239	코드	코드V6
240	코드	편파코드V2
241	수	숫자N10
242	방위각	방위각N12
243	시간	시간N20

○ 코드명(한글)은 표준용어명, 코드명(영문)은 영문약어명을 반드시 기입되도록 개선

- 코드도메인에 해당하는 표준용어는 ‘표준코드명’ 이 매핑되도록 보완한 예시

용어표준명	영문약어명	표준도메인명	표준코드명
ITU강우기후지역코드	DENSITY	구분코드N18	ITU강우기후지역코드
전력제한준수코드	F_OFF_AXIS	코드V1	전력제한준수코드
간섭발생코드	F_CAUSE	코드V1	간섭발생코드
간섭수신코드	F_REC	코드V1	간섭수신코드
경도방향	LONG_EW	방향코드V1	경도방향
구조비교결과	F_CMP_STR	결과코드V1	구조비교결과
궤도유지상태	F_STN_KEEP	상태V1	궤도유지상태
규제상태코드	REG_ST	상태코드V2	규제상태코드
기본수정코드	F_BASIC	코드V1	기본수정코드

□ 전파관리DB 표준화 결과

○ DB표준용어정의서 → 192개 신규표준용어 추가

- 신규 표준 용어 추가 예시

순서	표준용어명
456	오류코드
457	종료사용자ID
458	FMS장비명
459	장애주파수
460	서비스유형
461	RS코드
...	...
646	담당업무
647	직위

○ DB표준도메인정의서 → 33개 신규도메인 추가

- 신규 표준 도메인 추가 예시

순서	도메인분류명	도메인명
73	코드	코드N1
74	내용	명칭V100
75	내용	명칭V1000
76	식별자	번호V12
77	내용	명칭V14
78	내용	명칭V15
...
104	수량	숫자N11
105	좌표	위도좌표N22

○ 코드명(한글)은 표준용어명, 코드명(영문)은 영문약어명을 반드시 기입되도록 개선

- 코드도메인에 해당하는 표준용어는 ‘표준코드명’ 이 매핑되도록 보완한 예시

용어표준명	영문약어명	표준도메인명	표준코드명
게시판구분	BRD_GBN	코드C3	게시판구분
전파관리소	OFFC_CD	코드C2	전파관리소코드
행정동코드	H_CODE	코드V10	행정동코드
권한코드	AUTH_CD	코드C2	권한코드
법정동코드	B_CODE	코드V5	법정동코드
메뉴코드	MENU_CD	코드V30	메뉴코드
우편번호	ZIP_CD	코드C6	우편번호
상우회코드	CMPGRP_CD	코드C6	상우회코드
측정종류코드	MAS_TYPE_CD	코드C2	측정종류코드

□ 방송통신통합DB(적합성평가) 표준화 결과

○ DB표준용어정의서 → 451개 신규표준용어 추가

- 신규 표준 용어 추가 예시

순서	표준용어명
443	불확도내용
444	사용주파수
445	사용여부
446	사용저조직부서ID
447	적용시작일
448	적용종료일
449	중량
450	중량단위
451	위텍스등록일련번호

○ DB표준도메인정의서 → 72개 신규도메인 추가

- 신규 표준 도메인 추가 예시

순서	도메인분류명	도메인명
73	코드	코드N1
74	내용	명칭V100
75	내용	명칭V1000
76	식별자	번호V12
77	내용	명칭V14
78	내용	명칭V15
...
104	수량	숫자N11
105	좌표	위도좌표N22

○ 코드명(한글)은 표준용어명, 코드명(영문)은 영문약어명을 반드시 기입되도록 개선

- 코드도메인에 해당하는 표준용어는 ‘표준코드명’ 이 매핑되도록 보완한 예시

용어표준명	영문약어명	표준도메인명	표준코드명
추가기자재구분코드	ADD_MATL_DIV_CD	코드V2	추가기자재구분코드
지정시험기관코드	AEA_CD	코드V6	지정시험기관코드
지정시험기관소재지구분코드	AEA_LOC_DIV_CD	코드V2	지정시험기관소재지구분코드
사후관리계획구분코드	AFMG_PLAN_DIV_CD	코드V2	사후관리계획구분코드
지정시험기관변경구분코드	CHG_DIV_CD	코드V4	변경구분코드
교정구분1코드	CRE_DIV_1_CD	코드V5	교정구분1코드
교정구분2코드	CRE_DIV_2_CD	코드V5	교정구분2코드
인증서재발급사유코드	CTFA_RIS_RSN_CD	코드V2	인증서재발급사유코드
신청문서코드	DECL_DOC_CD	코드V10	신청문서코드

□ US_NTISAPP DB 표준화 결과

○ DB표준용어정의서 → 417개 신규표준용어 추가

- 신규 표준 용어 추가 예시

순서	표준용어명
409	상위분류군순번
410	사용여부
411	사용언어
412	사용목적
413	사용자명
414	사용자보안설정ID
415	사용자검색여부
416	사용자구분
417	단어사전ID

○ DB표준도메인정의서 → 72개 신규도메인 추가

- 신규 표준 도메인 추가 예시

순서	도메인분류명	도메인명
66	순번	순번N4
67	순번	순번N8
68	순번	순번N10
69	순번	순번N15
70	순번	순번V10
71	순번	순번V20
72	정렬순서	정렬순서N
73	정렬순서	정렬순서N4
74	정렬순서	정렬순서N8

○ 표준코드정의서 작업을 위해 코드명, 코드설명 등 표준코드 보완 작업 완료

- 코드명, 코드설명 등 표준코드 보완 작업 완료 예시

코드명(한글)	코드명(영문)	코드설명	코드값의미
이미지구분코드	IMAGE_TPCD	이미지를 구분하는 코드	일반이미지
이미지구분코드	IMAGE_TPCD	이미지를 구분하는 코드	인쇄이미지
이미지구분코드	IMAGE_TPCD	이미지를 구분하는 코드	대표이미지
연령코드	AGE	동물의 나이 연령대 구분 코드	확인불가
연령코드	AGE	동물의 나이 연령대 구분 코드	어린개체
연령코드	AGE	동물의 나이 연령대 구분 코드	중간성체
연령코드	AGE	동물의 나이 연령대 구분 코드	성체
성별코드	SXDST	동물의 성별 구분 코드	수컷
성별코드	SXDST	동물의 성별 구분 코드	암컷
성별코드	SXDST	동물의 성별 구분 코드	양성
성별코드	SXDST	동물의 성별 구분 코드	확인불가

CHAPTER IV

데이터 관리체계 품질관리 수준 평가 대응

1. '21년 품질관리 수준평가 대상 시스템 목록
2. '21년 품질관리 수준평가 결과
3. 평가대상 DB별 평가결과

CHAPTER IV 데이터 관리체계 품질관리 수준 평가 대응

1. '21년 품질관리 수준평가 대상 시스템 목록

□ 고유DB 6개, 지원DB 3개 총 9개 DB에 대한 품질진단 수행

[표 4-1] '21년 평가대상 시스템 목록

순번	시스템명	구분(고유/지원)
1	대표홈페이지	지원DB
2	방송통신통합정보시스템	고유DB
3	우주전과환경홈페이지	지원DB
4	비정지위성전과감시시스템	고유DB
5	전과관리정보시스템	고유DB
6	국가자연사연구종합정보시스템	고유DB
7	국립중앙과학관홈페이지	지원DB
8	우체국금융시스템	고유DB
9	인터넷우체국	고유DB

2. '21년 품질관리 수준평가 결과

□ 모든 지표에 대해 만점을 획득하였으며, 우수한 품질관리 수준을 유지하고 있음

[표 4-2] '21년 품질관리 수준평가 결과

영역	평가지표	구분	배점	취득점수
데이터 관리체계 (32)	품질관리 계획 수립	총괄	2	2
	예방적 품질관리 진단	총괄	10	10
	데이터 표준 확산	DB	10	10
	데이터 구조 안정화	DB	5	5
	데이터 연계 관리	DB	5	5
데이터 값관리 (68)	데이터 품질진단	DB	15	15
	품질진단 결과조치	DB	23	23
	데이터오류율	DB	15	15
	오류신고사항 분석 및 개선	총괄	4	4
	개방데이터셋 오류조치	총괄	11	11
합 계			100	100 (1등급)

3. 평가대상 DB별 평가결과

1) 대표홈페이지(지원DB)

[표 4-3] 과기부 대표홈페이지 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	-	N/A
		데이터 표준 적용율	6	-	N/A
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	-	N/A
		데이터 구조 산출물 현행화	3	-	N/A
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	-	N/A
		연계데이터 송수신 내역 관리	3	-	N/A
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	8개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.007%
합 계			53	53	

※ N/A 처리된 지표는 환산점수로 계산하여 만점 획득

2) 방송통신통합정보시스템(고유DB)

[표 4-4] 국립전파연구원 방송통신통합정보시스템 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	4	
		데이터 표준 적용율	6	6	100%
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	2	
		데이터 구조 산출물 현행화	3	3	100%
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	2	
		연계데이터 송수신 내역 관리	3	3	
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	13개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.0%
합 계			73	73	

3) 우주전파환경홈페이지(지원DB)

[표 4-5] 국립전파연구원 우주전파환경홈페이지 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	-	N/A
		데이터 표준 적용율	6	-	N/A
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	-	N/A
		데이터 구조 산출물 현행화	3	-	N/A
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	-	N/A
		연계데이터 송수신 내역 관리	3	-	N/A
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	5개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.0%
합 계			53	53	

※ N/A 처리된 지표는 환산점수로 계산하여 만점 획득

4) 비정지위성전파감시시스템(고유DB)

[표 4-6] 중앙전파연구소 비정지위성전파감시시스템 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	4	
		데이터 표준 적용율	6	6	93.7%
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	2	
		데이터 구조 산출물 현행화	3	3	99%
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	2	연계없음
		연계데이터 송수신 내역 관리	3	3	
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	9개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.0%
합 계			73	73	

5) 전파관리정보시스템(고유DB)

[표 4-7] 중앙전파연구소 전파관리정보시스템 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	4	
		데이터 표준 적용율	6	6	93.2%
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	2	
		데이터 구조 산출물 현행화	3	3	100%
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	2	
		연계데이터 송수신 내역 관리	3	3	
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	8개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.002%
합 계			73	73	

6) 국가자연사연구종합정보시스템(고유DB)

[표 4-8] 국립중앙과학관 국가자연사연구종합정보시스템 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	4	
		데이터 표준 적용율	6	6	97.9%
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	2	
		데이터 구조 산출물 현행화	3	3	100%
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	2	
		연계데이터 송수신 내역 관리	3	3	연계없음
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	10개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.005%
합 계			73	73	

7) 국립중앙과학관 홈페이지(지원DB)

[표 4-9] 국립중앙과학관 홈페이지 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	-	N/A
		데이터 표준 적용율	6	-	N/A
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	-	N/A
		데이터 구조 산출물 현행화	3	-	N/A
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	-	N/A
		연계데이터 송수신 내역 관리	3	-	N/A
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	8개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.003%
합 계			73	53	

※ N/A 처리된 지표는 환산점수로 계산하여 만점 획득

8) 우체국금융시스템(고유DB)

[표 4-10] 우정사업본부 우체국금융시스템 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	4	
		데이터 표준 적용율	6	6	91.2%
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	2	
		데이터 구조 산출물 현행화	3	3	99.4%
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	2	
		연계데이터 송수신 내역 관리	3	3	
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	9개정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.006%
합 계			73	73	

9) 인터넷우체국(고유DB)

[표 4-11] 우정사업본부 인터넷우체국 수준평가 결과

영역	평가지표	세부지표	배점	취득점수	비고
데이터 관리체계 (32)	데이터 표준 확산(10)	데이터 표준 정의	4	4	
		데이터 표준 적용율	6	6	94.2%
	데이터 구조 안정화(5)	데이터 구조 정의 및 산출물 관리	2	2	
		데이터 구조 산출물 현행화	3	3	99.4%
	데이터 연계 관리(5)	연계데이터 관리 정보 정의	2	2	
		연계데이터 송수신 내역 관리	3	3	
데이터 값관리 (68)	데이터 품질진단(15)	품질진단 기준 정의 및 진단	5	5	
		업무규정에 따른 품질진단 규칙(BR) 정의 및 진단	10	10	8개 정의
	품질진단 결과조치(23)	품질진단에 따른 개선계획 수립 및 개선과제 정의	8	8	
		개선조치 이행률	15	15	100%
	데이터오류율(15)	데이터 오류율	15	15	0.007%
합 계			73	73	

데이터 품질관리 도구 도입 방안

1. 기관 품질관리 현황
2. 데이터 품질진단 도구(DQMS) 도입방안
3. 데이터 품질관리 역량 교육

CHAPTER V. 데이터 품질관리 도구 도입 방안

1. 기관 품질관리 현황

1.1. 데이터 품질관리 계획 및 목표

□ 품질관리 계획

- 과학기술정보통신부 품질관리 관련지침*을 근거로 대상DB선정 및 기관·DB·데이터 관점의 추진과제(단기,중장기) 발굴을 통한 계획 수립

* 공공데이터 품질관리지침(과학기술정보통신부훈령 제72호. 2019.6.5. 제정) 및 공공기관의 데이터베이스 품질관리 지침(이하 품질관리 지침)

- 매년 실시되는 공공데이터 품질관리 수준평가(행안부) 영역별 평가 지표와 연계하여 품질관리계획 개선과제에 따른 이행과 그 성과를 평가

□ 품질관리 목표

- 과학기술정보통신부는 품질관리 체계기반 마련을 목표로 전략과 추진방향을 설정하여 중장기 품질관리 계획을 수립('20~22년) 함

[표 5-1] '20~22년 중장기 품질관리 목표

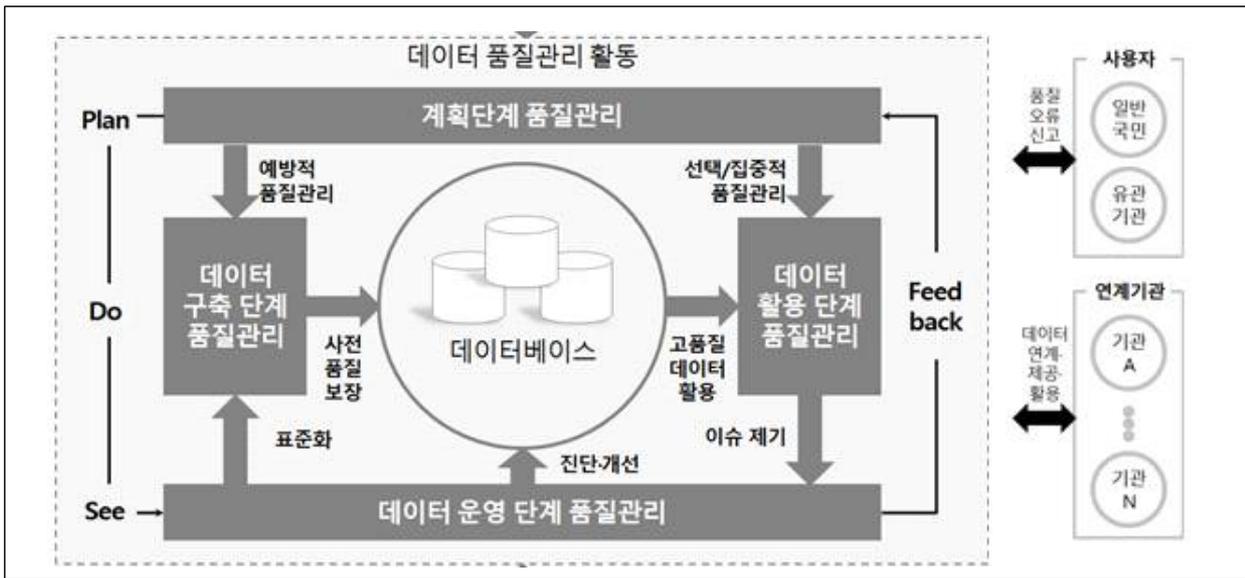
품질관리목표	데이터의 신뢰성 확보를 위한 “과학기술정보통신부 데이터 품질관리 기반 마련”
추진 전략	<ul style="list-style-type: none"> ■ 과학기술 및 정보통신 데이터 표준 확립 ■ 계획에 의한 지속적 품질관리 활동 강화 ■ 공공데이터 품질관리 수준 평가 향상
추진 방향	<ul style="list-style-type: none"> ■ 고품질의 데이터 관리를 위한 주기적 품질 진단 및 개선활동 ■ 데이터베이스 구조 안정화를 위한 산출물 현행화 및 연계 데이터 품질 확보 ■ 기관 차원의 데이터 표준 정립 및 확산 ■ 기관 공통의 품질관리 정책 및 조직 체계화 ■ 품질관리 시스템화를 위한 인프라 마련

1.2. 과학기술정보통신부 품질관리 프로세스

□ 품질관리 표준 프로세스

- 과학기술정보통신부 데이터 품질관리 프로세스는 행정안전부의 데이터 품질관리 단계(4단계)를 준용하여 적용하며 운영 중임

[그림 5-1] 데이터 품질관리 활동도



□ 품질관리 프로세스 단계별 수행 내용

- (계획단계 품질관리) 각 기관의 품질관리를 위한 내부규정 마련, 품질관리 조직 구성, 공공데이터의 구축, 운영, 활용 단계별 품질관리 활동을 반영한 품질관리 계획을 수립
- (구축단계 품질관리) 각 기관이 데이터를 구축 시 준수해야 할 데이터 표준의 적용·점검, 데이터 구조의 일관성 확보, 데이터관리 산출물의 품질 확보 등을 수행
- (운영단계 품질관리) 각 기관이 보유하고 있는 데이터를 운영하는 단계에서 준수해야 할 품질진단·개선, 연계데이터 정합성 유지, 주요 산출물의 현행화 점검 등을 수행
- (활용단계 품질관리) 각 기관이 데이터를 활용하면서 발생하는 품질오류 신고 및 차기 품질계획에 반영하기 위한 데이터 활용성과 등을 평가

1.3. 품질관리 관리 현황 및 관리도구 필요성

□ 과기부 품질진단 관리 현황 분석

- 과기정통부의 공공데이터제공책임관 및 실무담당자는 기관 전체의 데이터 품질관리를 총괄하여 수준평가 등 행안부 평가 업무를 대응하고 있으며, 각 시스템의 공공데이터 품질관리는 정보화담당관의 공공데이터 실무담당자가 총괄하며 DB 별 품질관리 담당자를 지정하여 수행 중임(정책 일관성 유지)
- 품질관리 수준평가의 평가대상DB 품질진단은 행정안전부가 제공해주는 진단도구(SDQ)를 활용하여 가이드에 따라 각 DB별 담당자가 1차 진단 후 개선 활동을 수행하고 있음(향후 도구변경 가능성)
- 매년 선정된 품질진단 대상DB의 1차 값진단 시 전반적으로 양호한 오류율을 보이고 있으나, 일부 DB의 경우 오류율이 높아 지속적이고 체계적인 진단관리 필요성 대두(DB 품질제고 필요)

※ '21년 품질진단 대상DB(9개) 1차 진단시 오류율 : 평균 0.0240% (2등급 수준)

- 품질진단 시 범정부 품질진단 기준(정합성, 완결성, 유효성 등 16개 유형)과 공간정보 품질기준에 의거하여 컬럼별 진단기준을 설정하고 수행하고 있음(DQI정의, 관리 복잡도 증가)
- 특히, 행정안전부는 품질진단 시 업무규칙에 의한 진단을 확대하는 상황이어서 이를 도출하고 관리해야 할 필요성이 증가함(업무규칙관리강화)
- 업무담당자 및 DB 담당자가 인사이동에 의해 변동 가능성이 많아 시스템을 통해 체계적으로 관리가 필요한 시점(담당자 변경대비 필요)
- 품질진단 관리도구 도입 필요성
- (품질관리대상정의) 과학기술정보통신부 차원의 품질관리방안을 수립하고 체계적인 품질관리를 위해 품질관리대상을 관리 필요
- (품질지표운용) 데이터 사용목적을 충족시키기 위해, 지속적으로 품질점검을 통해 관리되어야 할 진단 기준을 확정하고 관리
- (핵심 품질관리 대상 정보항목 관리) 데이터의 신뢰도가 고객, 프로세스 등 기관업무수행에 중요한 영향을 미치는 핵심 품질관리 정보항목 관리

- (업무규칙수집) 정부의 품질관련 평가를 대비하고 정량적인 목표치를 통해 품질을 관리하기 위해 업무규칙의 발굴 및 진단 프로세스 관리 필요
- (담당역량강화) 업무담당자 및 품질관리 담당자 교육 등을 통해 데이터 품질관리 역량 확보하고 업무 이동시에 인계인수 용이성 확보
- (품질진단관리도구) 부서의 품질관련 지표 및 핵심관리항목을 정의하고 이를 체계적으로 관리하여 품질을 진단·개선할 수 있는 체계 도입 필요

□ 품질진단 관리도구 도입 필요성

- (품질관리대상정의) 과학기술정보통신부 차원의 품질관리방안을 수립하고 체계적인 품질관리를 위해 품질관리대상을 관리 필요
- (품질지표운용) 데이터 사용목적에 충족시키기 위해, 지속적으로 품질점검을 통해 관리되어야 할 진단 기준을 확정하고 관리
- (핵심 품질관리 대상 정보항목 관리) 데이터의 신뢰도가 고객, 프로세스 등 기관업무수행에 중요한 영향을 미치는 핵심 품질관리 정보항목 관리
- (업무규칙수집) 정부의 품질관련 평가를 대비하고 정량적인 목표치를 통해 품질을 관리하기 위해 업무규칙의 발굴 및 진단 프로세스 관리 필요
- (담당역량강화) 업무담당자 및 품질관리 담당자 교육 등을 통해 데이터 품질관리 역량 확보하고 업무 이동시에 인계인수 용이성 확보
- (품질진단관리도구) 부서의 품질관련 지표 및 핵심관리항목을 정의하고 이를 체계적으로 관리하여 품질을 진단·개선할 수 있는 체계 도입 필요

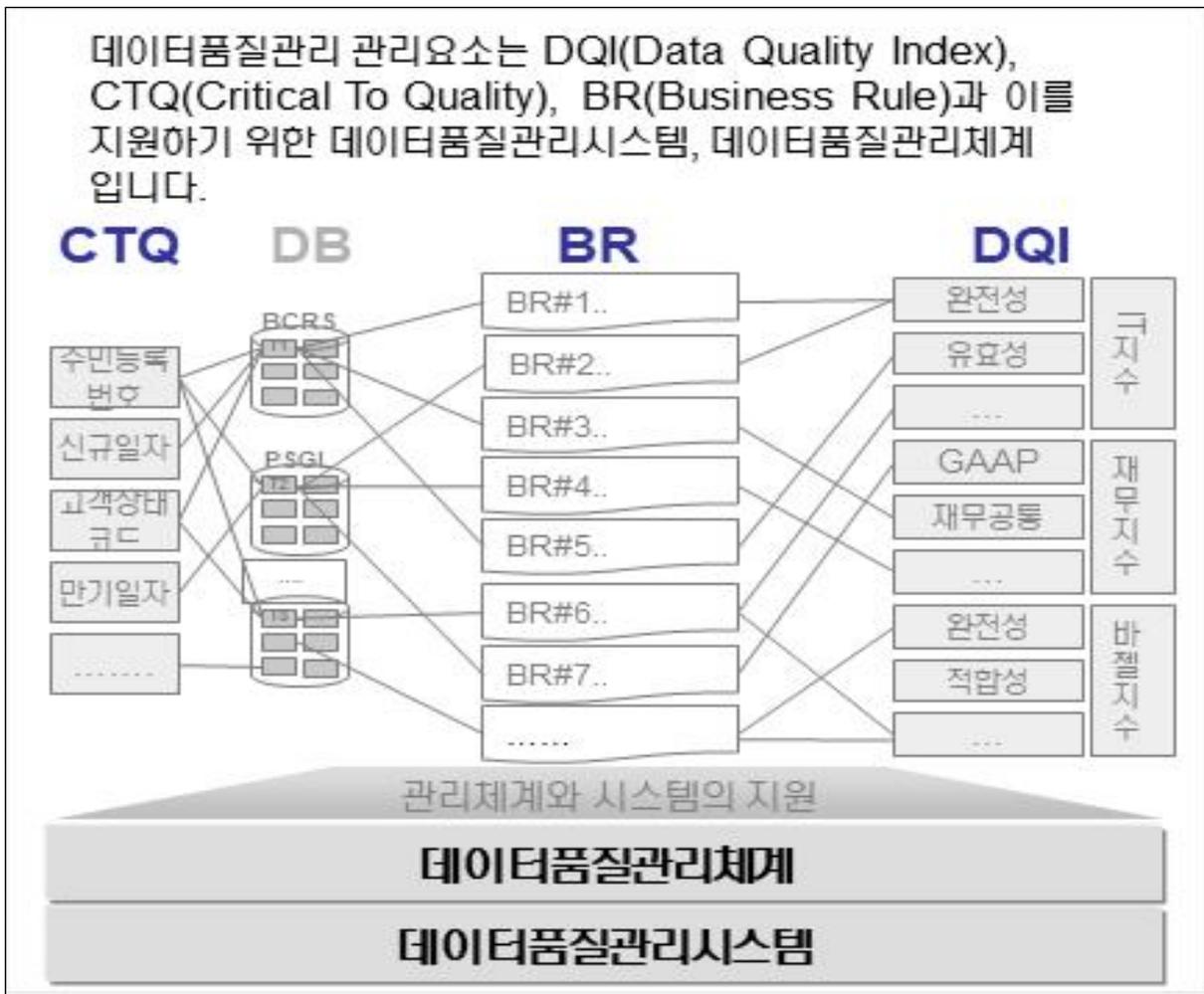
2. 데이터 품질진단 도구(DQMS) 도입방안

2.1. 품질진단 관리도구 개요

□ 도입 배경

- 품질진단 관리도구는 기관(기업)의 경영목표를 달성하기 위해 데이터 측면에서 데이터품질관리체계(데이터품질거버넌스)를 수립하고 데이터품질진단 관점에서 체계적으로 수행하기 위한 체계로서,
- 품질관리 대상(CTQ)과 품질측정 규칙(DQI, BR)이 정의하고 상시적인 품질 모니터링 및 품질개선 작업을 위해 도입이 필요

[그림 5-2] '20~22년 중장기 품질관리 목표



□ 품질진단 도구 주요 기능

1) 데이터품질 기준관리

- 데이터 품질관리의 기준이 되는 진단대상DB(시스템), 품질지표(DQI), 핵심관리항목(CTQ²⁾), 업무규칙(BR³⁾) 등의 기준정보 관리

- 사용자 정의 속성 기능으로 기관의 데이터품질 관리활동 특성에 따른 다양한 부가정보 관리 환경 제공

- 메타시스템과의 연계를 통한 도메인 자동 설정 기능 제공

2) 데이터품질 측정관리

- 프로파일링 분석을 통한 컬럼 Level의 데이터 오류 징후 도출 지원
- 업무규칙 분석을 통한 지속적인 데이터 품질 현황 모니터링 지원
- 다차원분석을 통한 이기종 DB간의 데이터 일관성 검증 지원
- JOB Workbench를 통한 정기/비정기 작업 수행 스케줄 관리 지원
- DataLoad(DB 간 데이터 이행), SQL/Procedure/JavaClass실행 등 다양한 전/후처리 기능을 통한 고 효과적인 데이터 품질 관리 활동 지원
- 데이터 원천구분 기능으로 데이터 원천별 오류 현황 파악과 그로 인한 원천적 원인분석 및 개선 활동 기반 제공

3) 데이터품질 통제관리

- 오류원인 분석 및 개선활동 내역 관리를 통한 데이터 품질 관리 활동의 통제력 강화 지원

4) 보고관리

- DocumentCenter 기능을 통한 Plexible한 보고서 생성 및 관리 기능 제공
- SMTP, SMS, FTP 연계 등 다양한 방식을 통한 데이터 품질 현황 및 품질 관리 활동 현황에 대한 전사적 차원의 공유 환경 지원

1) DQI(Data Quality Index: 데이터 품질지표): 데이터 사용목적을 충족시키기 위해, 지속적으로 품질점검을 통해 관리되어야 할 진단 기준
 2) CTQ(Critical To Quality: 핵심 품질관리 대상 정보항목): 데이터의 신뢰도가 고객, 프로세스 등 기업 경영에 중요한 영향을 미치는 핵심 품질관리 정보항목
 3) BR(Business Rule): 데이터 사용자가 요구하는 수준을 만족시키기 위하여 업무적으로 규정된 기준에 맞도록 데이터 값을 관리하기 위한 조건에 대한 일반적 표현



2.2. 품질관리 도구 도입 방안

□ 과학기술정보통신부 품질진단 기준

- 진단 대상은 해당 년도에 선정된 중점 품질관리대상 DB로 정의하고 이에 따른 영역별 진단 기준을 도출하여 수행

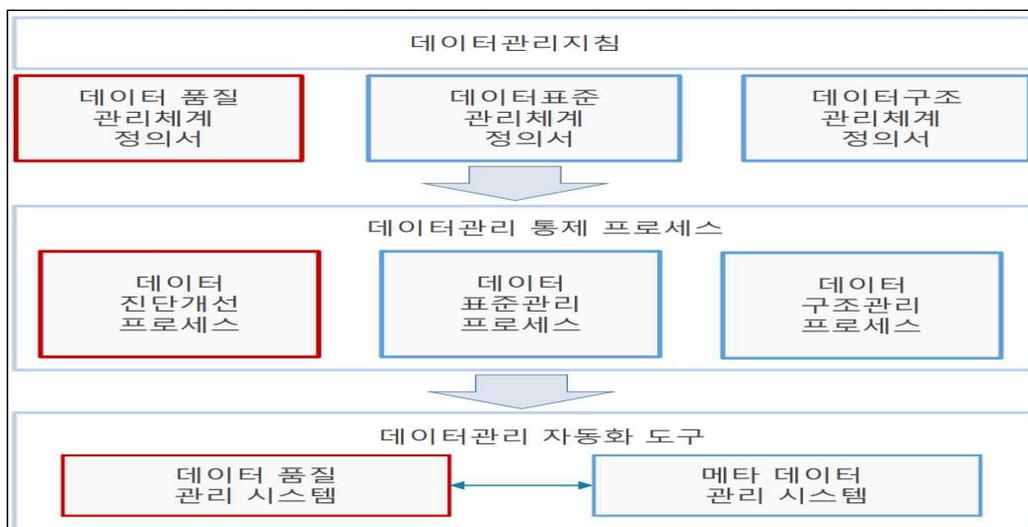
[표 5-2] 데이터 품질 진단 영역 별 측정방법 및 평가기준

영역	진단대상	측정방법	평가기준
값	<ul style="list-style-type: none"> 데이터값: 날짜, 번호, 여부, 코드, 금액, 수량, 율 업무규칙: 시간 순서, 선후, 계산식, 참조 	<ul style="list-style-type: none"> 데이터실사 프로파일링 업무규칙정의서 	<ul style="list-style-type: none"> 정량적 평가 값오류를 계산
구조	<ul style="list-style-type: none"> 산출물 현행화 	<ul style="list-style-type: none"> ERD 분석(정규화, 미사용테이블 등) 테이블정의서, 컬럼정의서와 물리 DB간 차이분석 	<ul style="list-style-type: none"> 정량적 평가 산출물과 물리DB간의 일치율 계산
표준화	<ul style="list-style-type: none"> DB표준과의 차이분석 	<ul style="list-style-type: none"> 테이블, 컬럼 분석 표준, DB간 차이분석 	<ul style="list-style-type: none"> 정량적 평가 DB표준과 물리적 DB간의 일치율
연계	<ul style="list-style-type: none"> 연계관리 활동 	<ul style="list-style-type: none"> 연계데이터 목록 정의 여부 로그기록관리, 품질 점검 여부 	<ul style="list-style-type: none"> 관리 유무

□ 품질진단 도구 필요 기능

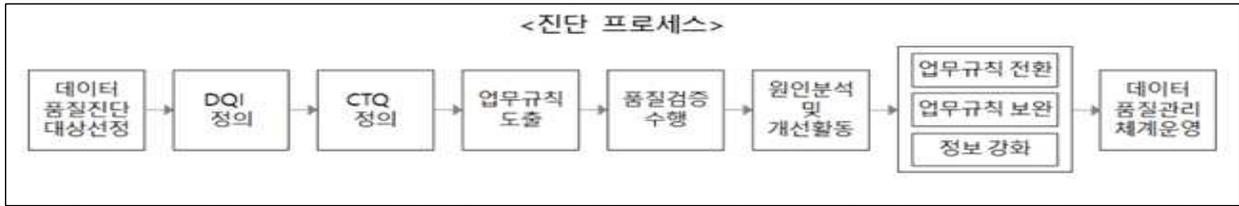
- 과학기술정보통신부 데이터 품질관리 프로세스의 생명주기 중 3단계인 “운영단계 품질관리”에 관련된 아래 품질진단 활동에 대한 기능을 충족하는 수준으로 검토함

[그림 5-3] 데이터 관리지침



- 품질진단 프로세스는 대상(DB)을 선정하고 이에 따라 DQI, CTQ, 업무규칙(BR)를 정의하고 수행하는 일반적인 절차로 수행함

[그림 5-4] 진단 프로세스



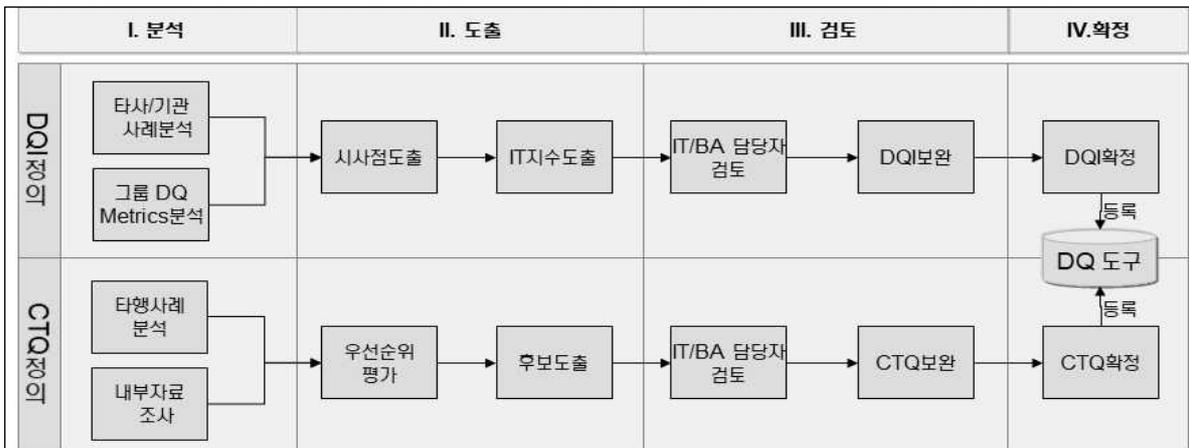
1) 데이터품질진단 대상선정

- 품질진단 관리대상 DB별로 진단대상 테이블을 선정함

2, 3) DQI 및 CTQ 정의

- 데이터품질대상인 DQI 및 CTQ의 정의는 분석, 도출, 검토, 확정의 절차로 수행하며, 지속적인 관리를 위하여 이를 데이터 품질관리 시스템에 등록하여야 함

[그림 5-5] DQI 및 CTQ 정의 절차



4) 업무규칙 도출

- 현업담당자와 공동으로 진단 규칙 정의하고 시스템에 등록

5) 품질검증(진단)

- 품질진단 수행

6) 원인분석 및 개선활동

- 진단결과에 따라 원인을 분석하여 단계별 개선 계획을 수립하고 개선수행

7) 업무규칙보완

- 오류율 개선을 위해 기존 업무규칙을 보완하거나 신규로 도출
- 8) 데이터품질 관리체계 운영
- 데이터품질 관리체계에 따라 관리체계 가동

2.3. 품질관리 도구 선정 방법

□ 품질진단 도구 선정 평가기준

- (선정기준)도입 필요 도구에 대해 기능성, 사용성, 효율성 등 7개 기준으로 평가하여 선정함

※ 정보통신산업진흥원(NIPA)의 “소프트웨어 기술성 평가기준 적용 가이드” 참조

[표 5-3] 소프트웨어 기술성 평가기준 적용 가이드

평가기준	설명	비고
기능성	<ul style="list-style-type: none"> ▪ 도입되는 상용 소프트웨어 제품이 제안요청서에서 요구하는 기능의 충족 정도 	솔루션
사용성	<ul style="list-style-type: none"> ▪ 목적을 달성하기 위해 사용될 때 사용 용이성, UI 편의성 평가 	솔루션
효율성	<ul style="list-style-type: none"> ▪ 상용SW 제품이 업무지원 및 자원 활용의 효율화 기여도 평가 	솔루션
이식성	<ul style="list-style-type: none"> ▪ 상용SW 제품이 다른 환경이나 기종으로 옮기는 것에 대한 용이성 	솔루션
유지관리성	<ul style="list-style-type: none"> ▪ 상용SW 제품의 유지관리가 얼마나 용이 한지 정도 	솔루션
신뢰성	<ul style="list-style-type: none"> ▪ SW 프로그램이 다운되지 않고 안정적으로 수행되는지에 대한 평가 	솔루션
공급업체지원	<ul style="list-style-type: none"> ▪ 제품의 원활한 활용을 위한 공급업체의 지원 수준에 대해 유지관리, 하자보수,교육훈련 등의 관점에서 평가 	업체평가

○ 제품 선정시 고려사항

- 효율성, 신뢰성, 공급업체 지원 등은 제품공급업체의 타 기관 적용 사례를 벤치마킹하여 의사결정에 반영 필요
- 제품평가 시 실제 사용 시연(BMT)을 통한 검증 과정이 필요
- 기준별 점수는 내부 협의를 거쳐 배점 조정 필요
- 진단솔루션을 포함하여 기관의 데이터품질 관리체계를 적용한 시스템 도입 부분도 포함하여 검토 필요

2.4. 국내 품질진단 솔루션 현황

□ 품질진단 관리도구 제품 현황

- 데이터품질 관련 솔루션 목록(붙임1참조)

[표 5-4] 데이터품질 관련 솔루션 목록

분류	글로벌 벤더	국산 벤더	역할
데이터 품질	<ul style="list-style-type: none"> TalenD Oracle 	<ul style="list-style-type: none"> DQube(MITSOFT) WISE DQ(위세아이텍) DQMiner(지티윈) 	데이터 품질 진단

- 공공기관 도입사례(제안요청서 기준)

[표 5-5] 공공기관 도입사례

기관명	범위	예산(백만원)	발주년도
환경부	<ul style="list-style-type: none"> 메타데이터 관리 솔루션 데이터 품질관리 솔루션 	77	2018
법무부 출입국	<ul style="list-style-type: none"> 메타데이터 관리 솔루션 데이터 품질관리 솔루션 	통합사업의 일부	2017
해양수산부	<ul style="list-style-type: none"> 메타데이터 관리 솔루션 데이터 품질관리 솔루션 	228	2016
한국교육학술정보원	<ul style="list-style-type: none"> 메타데이터 관리 솔루션 데이터 품질관리 솔루션 	통합사업의 일부	2015
한국산업은행	<ul style="list-style-type: none"> DQMS 엔진 도입 커스터마이징 메타시스템연계 	297	2014
산업자원부	<ul style="list-style-type: none"> 데이터 표준관리 데이터 모델관리 데이터 품질관리 	1,132	2008

2.5. 품질진단 관리도구 도입 효과

□ 전사 데이터자원 품질관리 기반 마련

- 데이터품질관리 원칙 정의
- 전사 데이터품질대상(DQI/CTQ) 정의
- 데이터품질규칙(BR) 정의

- 데이터품질관리시스템 구축을 기반 제공
- 데이터품질관리체계(프로세스/기능) 정의
- 기관 보유시스템의 데이터 품질진단 기능 제공
 - 수준평가 대상 데이터의 수시 프로파일링 실시
 - 전 영역 원천데이터의 표준화 기반 마련
 - 지속적 관리를 위한 BR 도출 관리
- 조직의 데이터품질관리 역량 강화
 - 데이터품질 진단도구 사용교육을 통한 활용지원
 - 데이터품질대상(DQI/CTQ) 도출 및 규칙(BR) 정의 방법에 대한 내재화
 - 데이터품질을 관리하기 위한 업무담당자의 BR 관리 용이
- 데이터품질 상시 모니터링 기반 마련
 - 데이터품질관리 프로세스와 조직 역할에 따른 점검 활동
 - 오류 결과 보고서의 사용자 관점 별 제공으로 종합적인 품질활동 수행
 - 오류 발생 시점별 이력 및 실 오류 데이터 제공으로 개선을 위한 분석 가능

[표 5-6] [붙임1] 국산 솔루션별 기능 비교

제품명	DQube	DQMiner	WiseDQ
벤더	엠아이티소프트	지티원	위세아이텍
기능	<ul style="list-style-type: none"> § 진단기준관리 (진단기준 자동설정) § 프로파일링 § 업무규칙진단 § 진단대상테이블관리 § 진단 JOB 관리 	<ul style="list-style-type: none"> § 진단기준관리 § 프로파일링 § 업무규칙진단 § 진단대상테이블관리 § 추이및다차원분석 § 진단 JOB 관리 	<ul style="list-style-type: none"> § 진단기준관리 § 프로파일링 § 업무규칙진단 § 진단대상테이블관리 § 추이및다차원분석 § 진단 JOB 관리
특징	<ul style="list-style-type: none"> § 공공데이터 품질관리 수준평가 전용 S/W § 품질진단 기준(물세팅) 자동설정기능으로 진단에 소요되는 시간과 노력을 대폭 감소 § 편의성과 효율성 향상을 위한 사용자 중심의 UI 설계 § 진단결과보고서는 수준평가 증적자료로 인정 	<ul style="list-style-type: none"> § 데이터 품질 관리 활동 과정을 시나리오 형태로 설계 및 Job 디자이너 제공 § Staging 영역의 데이터에 대해서 데이터 수정을 통한 시뮬레이션 기능 제공 § 서로다른DB의 값과 Cross SQL로 비교기능 § 원천구분코드 자동 분석 및 보고서 제공 	<ul style="list-style-type: none"> § 다수의 차세대 경험 § SI개발형태의 사이트버전 지원 § 80여회 이상 금융, 공공 분야 사업수행 § 다양한 플랫폼 지원 유지보수 용이 § 자사 메타데이터관리 솔루션통합시스템 구성 가능
운용환경	C/S 기반	WEB 기반	WEB 기반
가격 (솔루션+커스터마이징)	50백만원	345백만원	189백만원

CHAPTER VI

데이터 품질관리 역량 강화

1. 데이터 품질관리 체계 수립
2. 업무규칙 가이드 수립
3. 데이터 품질관리 역량 교육

CHAPTER VI 데이터 품질관리 역량 강화

1. 데이터 품질관리 체계 수립

□ 데이터 품질관리 개요

○ 목적

- 본 정의서는 과학기술정보통신부에서 생성·취득·연계하여 관리하는 데이터에 대한 품질관리체계를 수립함으로써 안정적·지속적·상시적으로 고품질의 데이터를 확보하기 위함이다.

□ 데이터 품질관리 주요 용어

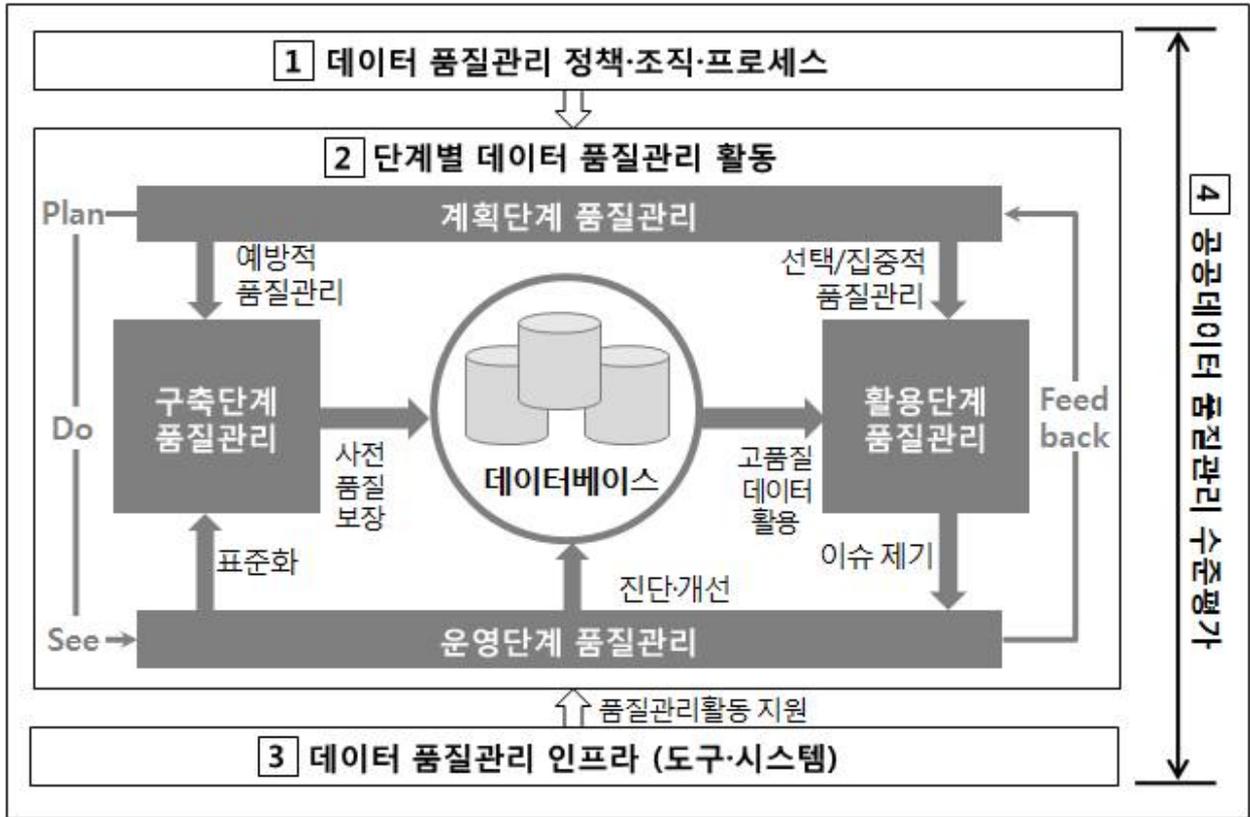
- 데이터 품질 : 데이터의 최신성, 정확성, 상호연계성 등을 확보하여 사용자에게 유용한 가치를 줄 수 있는 수준
- 데이터 품질관리 : 데이터 사용자에게 유용한 가치를 제공하도록 데이터의 품질을 확보하기 위한 품질 목표 설정, 품질 진단 및 개선 등 일련의 활동
- 공통표준용어 : 「공공기관의 데이터베이스 표준화 지침」 제2조제8호에 따라 행정안전부 장관이 정하여 공공기관들의 업무 및 데이터베이스에 공통 활용하는 표준용어
- 도메인 : 「공공기관의 데이터베이스 표준화 지침」 제2조제5호에 따라 데이터 값이 갖는 데이터 형식과 허용되는 값의 영역을 정의한 것
- 행정표준코드 : 「행정기관의 코드표준화 추진지침」에 따라 행정안전부 장관이 공공기관의 업무에 필요한 행정코드를 표준으로 정하여 제정 고시한 행정코드(www.code.go.kr에 고시)
- 데이터 표준화 : 「공공기관의 데이터베이스 표준화 지침」 제2조 제6호에 따른 코드, 용어, 도메인, 메타데이터 등의 표준을 수립하여 공공데이터베이스 등에 일관되게 적용하는 일련의 활동
- 메타데이터 : 「공공기관의 데이터베이스 표준화 지침」 제2조 제4호에 따른 공공데이터베이스내 데이터의 체계적인 관리와 편리한 검색 및 활용을 위하여 데이터의 구조, 속성, 특성, 이력 및 용어 등이 표현된 자료
- 기타 주요 용어는 「공공데이터 관리지침」, 「공공기관의 데이터베이스 표준화 지침」, 「공공데이터 품질관리 매뉴얼」 등 준용

□ 데이터 품질관리 범위

- 데이터 품질관리는 전반적인 품질관리 활동에 대한 정책적 방향성과 추진체계를 정의하는 정책·조직·프로세스, 계획·구축·운영·활용의 각 단계별 품

질관리 활동과 이를 지원하기 위한 품질관리 인프라로 구성됨

[그림 6-1] 데이터 품질관리 구성 요소



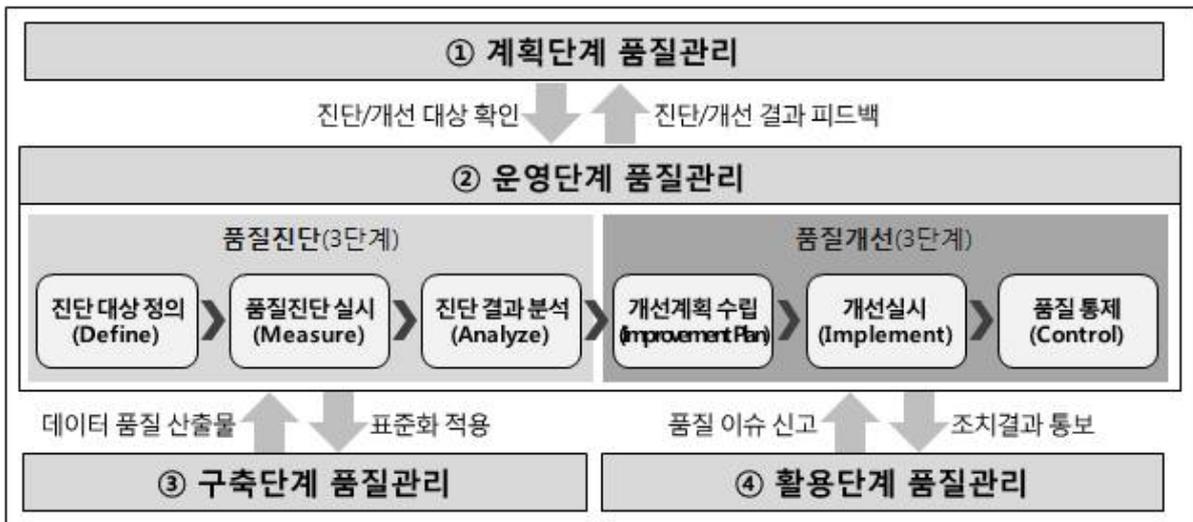
1 데이터 품질관리 정책·조직·프로세스 수립

- (정책) 데이터의 효과적 관리·운영·활용을 위해 과학기술정보통신부 차원에서 수립된 데이터 품질관리에 대한 원칙이 되는 데이터 품질관리 지침 수립
- (조직) 데이터 품질관리를 위해 수행되는 업무 정의 및 품질관리 수행 담당부서 또는 담당자 지정 등 품질관리 조직 구성
- (프로세스) 각 데이터 품질관리 업무의 흐름, 절차 등을 정의하는 품질관리 프로세스 수립

2 단계별 데이터 품질관리 활동

- 과학기술정보통신부에서 생성·취득하여 관리하고 있는 데이터의 품질 확보를 위하여 계획·구축·운영·활용의 4단계로 구분

[그림 6-2] 단계별 데이터 품질관리 활동



- ① (계획단계) 데이터의 구축, 운영, 활용 단계별 품질관리 활동을 반영한 품질관리 계획, 품질관리 교육계획 등 수립
- ② (구축단계) 데이터 구축 시 준수해야 할 데이터 표준의 적용·점검, 데이터 구조의 일관성 확보, 데이터관리 산출물의 품질확보, 연계데이터 정합성 유지 등 수행
- ③ (운영단계) 보유 데이터를 운영 시 준수해야 할 품질진단·개선, 주요 산출물의 현행화 점검 등을 수행
- ④ (활용단계) 데이터 활용 시 발생하는 품질오류신고 및 차기 품질 계획에 반영하기 위한 데이터 활용 사례, 개선사항 관리 등 수행

③ 데이터 품질관리 인프라 구축

- 데이터 품질관리 활동을 효과적으로 수행하기 위해 이를 지원하는 도구 및 관련 시스템 등 활용

④ 데이터 품질관리 수준평가

- 공공데이터 품질관리 수준평가 제도 시행을 통해 과학기술정보통신부가 적정 수준의 품질관리 활동의 수행 여부를 점검하여 개선사항 정비 등 품질개선 촉진
- 전사적 품질관리 활동 평가를 위해 과학기술정보통신부 내 품질관리 업무의 명확한 역할과 책임을 정의하고 총괄부서와 DB 운영부서의 역할을 구분하여 평가

□ 관련 근거

- 「공공데이터의 제공 및 이용 활성화에 관한 법률」(이하 ‘공공데이터법’) 제 22조(공공데이터의 품질관리)제1항, 제3항 및 동법 시행령 제17조(공공데이터의 품질 진단 및 개선)제1항제2호

※ 「공공데이터법」 제22조(공공데이터의 품질관리) 1항

“① 공공기관의 장은 해당 기관이 생성 또는 취득하여 관리하는 공공데이터의 안정적 품질 관리 및 **적정한 품질수준의 확보**를 위하여 필요한 조취를 취하여야 함”

“③ 행정안전부장관은 과학기술정보통신부장관과 협의하여 정기적으로 사회적·경제적 파급효과가 큰 제공 대상 공공데이터에 대한 **품질 진단·평가를 실시하고...**(중략)”

※ 「공공데이터법 시행령」 제17조(공공데이터의 품질 진단 및 개선) 제1항 제2호

“① 법 제22조제3항에 따른 공공데이터의 품질 진단평가의 기준 중 ‘**데이터 품질 관리체계**’가 포함됨”

□ 정책·조직·프로세스

○ 데이터 품질관리 지침

- 「과학기술정보통신부 공공데이터 품질관리 지침」(시행 2019. 6. 5) 제3조(적용범위)에 따라 과학기술정보통신부의 데이터 품질관리 지침을 준수하고, 과학기술정보통신부의 「데이터 표준화 지침」 등을 참고하여 수립

○ 데이터 품질관리 조직 체계

- 「과학기술정보통신부 전체의 데이터 품질관리를 총괄하는 부서를 지정하고, 데이터 품질관리 실무를 총괄하는 데이터 품질관리 총괄 책임자를 지정해야 함

[표 6-1] 데이터 품질관리 총괄 책임자 및 총괄부서* 업무 정의

업무명	업무 내용
품질관리 총괄	▪ 전사적 데이터 품질관리 업무 총괄
품질관리 지침 관리	▪ 지침 등 제·개정
계획·실적 관리	▪ 전사적 데이터 표준/품질 계획 수립
품질 진단·개선	▪ 데이터품질관리시스템 관리 ▪ 업무규칙 생성·변경 관리 ▪ 업무규칙 오너십 관리 ▪ 데이터 품질 모니터링 ▪ 품질 진단 결과 공유 ▪ 데이터 정제 관리 ▪ 과학기술정보통신부에 결과 보고
표준 관리	▪ 메타데이터관리시스템 관리 ▪ 단어·도메인·코드·용어 표준화, 현행화 관리 및 점검 (분기별 1 회 보고서 작성 및 보고) ▪ DB 모델 및 표준 관련 DB 디플로이 지원
구조 관리	▪ ERD 현행화 점검 및 관리
연계데이터 관리	▪ 연계데이터 목록, 송·수신 내역 관리
품질 교육	▪ 품질관리 교육 계획 수립 및 시행

업무명	업무 내용
평가·인증 대응	▪ 각종 평가 및 인증 대응
대내·외 협업	▪ 내·외부 기관 간 협력체계 구축
사업관리	▪ 데이터 품질 및 표준 관련 사업 관리

* 데이터 품질관리 총괄부서 : 정보화담당관('21년 7월 기준)

[표 6-2] 업무팀 데이터 품질 담당자 역할

업무명	업무 내용
대내·외 협의체 지원	▪ 데이터 품질관리 전담팀 및 과학기술정보통신부, 관련기관 등 데이터 품질 관련한 대내·외 협의체 지원활동 수행
업무규칙 도출 및 검증	▪ 신규 업무규칙 도출 및 기존 업무규칙 변경 관리 ▪ 데이터 품질관리 전담팀과 협의하여 업무규칙 오너십 관리 ▪ 데이터 품질관리 전담팀으로부터 통보받은 시스템 별 주요 품질오류 결과 검토 후 오류원인 분석 등을 수행한 후 데이터 품질관리 전담팀에 회신 ▪ 품질 오류 개선을 위한 응용프로그램 개선 및 데이터 정제 내역을 데이터 품질관리 전담팀으로 통보
연계데이터 관리	▪ 대내·외 연계데이터 목록 작성 및 관리 ▪ 연계데이터 정합성 모니터링 및 문제 발생시 해결방안 모색
데이터 품질교육 참석	▪ 과학기술정보통신부 내 데이터 품질 교육, 범정부 데이터 품질 교육 참석
평가 및 인증 대응	▪ 공공데이터 품질관리 수준평가 대응 ▪ 데이터 품질 인증 대응
데이터 품질 관련 사업 지원	▪ 데이터 품질 및 표준 관련 사업 지원

○ 데이터 품질관리 프로세스

- 과학기술정보통신부의 데이터 품질관리 업무를 중심으로 각 조직 간 역할 및 흐름을 정의한 9개 관리영역의 30개 프로세스는 다음과 같음

[표 6-3] 데이터 품질관리 프로세스

[관리영역] 프로세스 레벨1	프로세스 레벨2
지침 관리 프로세스	지침 제정
	지침 개정
계획·실적 관리 프로세스	연간 데이터 품질관리 계획 수립
	중·장기 데이터 품질관리 계획 수립
	데이터 품질관리 실적 평가·관리
	데이터 표준관리 계획 수립
데이터 품질 교육 관리 프로세스	데이터 표준관리 실적 평가·관리
	데이터 품질 교육 계획·실적 관리

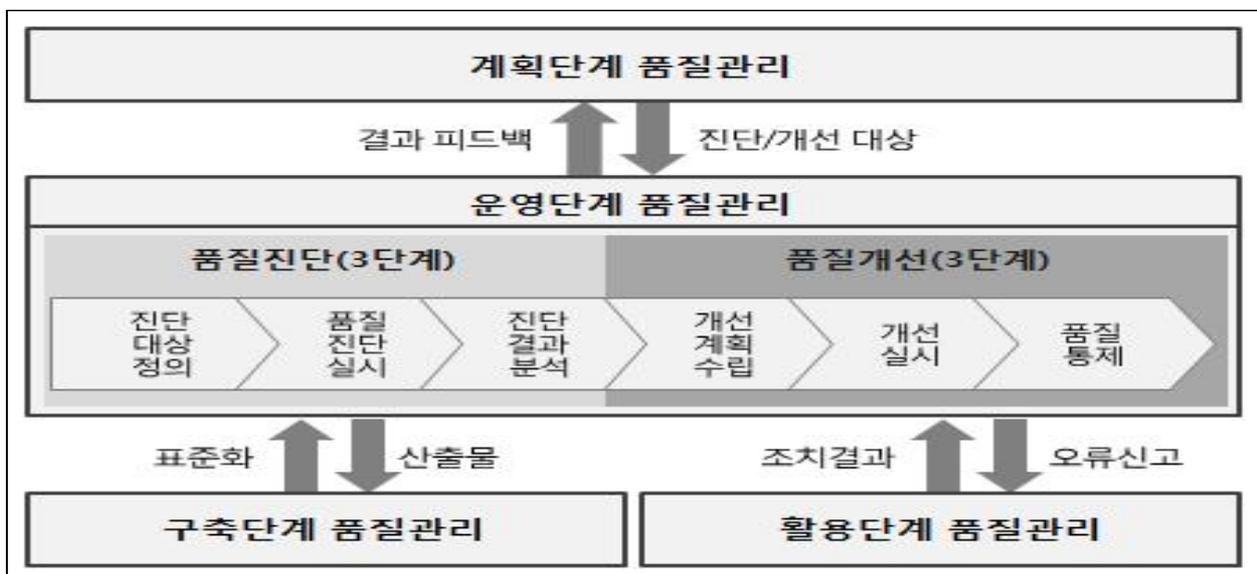


[관리영역] 프로세스 레벨1	프로세스 레벨2
데이터 표준 관리 프로세스	데이터 표준 단어 신청
	데이터 표준 용어 신청
	데이터 표준 도메인 및 인포타입 신청
	데이터 표준 코드 신청
	데이터 표준 점검 및 개선
	데이터 표준 변경
데이터 구조 관리 프로세스	데이터 모델 등록 및 변경
연계데이터 관리 프로세스	연계데이터 모니터링
	연계데이터 품질개선 활동 ✓ 과학기술정보통신부 산하기관(단기) ✓ 과학기술정보통신부 산하기관(장기) ✓ 타 부처
품질 진단·개선 프로세스	데이터 품질 진단 ✓ 진단 대상 정의 ✓ 품질 진단 실시 ✓ 진단 결과 분석
	데이터 품질 개선 ✓ 개선 계획 수립 ✓ 개선 수행 ✓ 품질 통제
데이터 활용 관리 프로세스	데이터 품질오류 신고 관리
	수요자 요구사항 관리
데이터 품질관리 협의회 운영 프로세스	데이터 품질관리 협의회 운영

□ 단계별 품질관리 방안

- 과학기술정보통신부의 데이터 품질관리 활동은 정보 생명주기(Life Cycle)에 기반 하여 ‘계획-구축-운영-활용’의 4단계 품질관리 활동으로 구성됨

[그림 6-3] 단계별 품질관리 활동

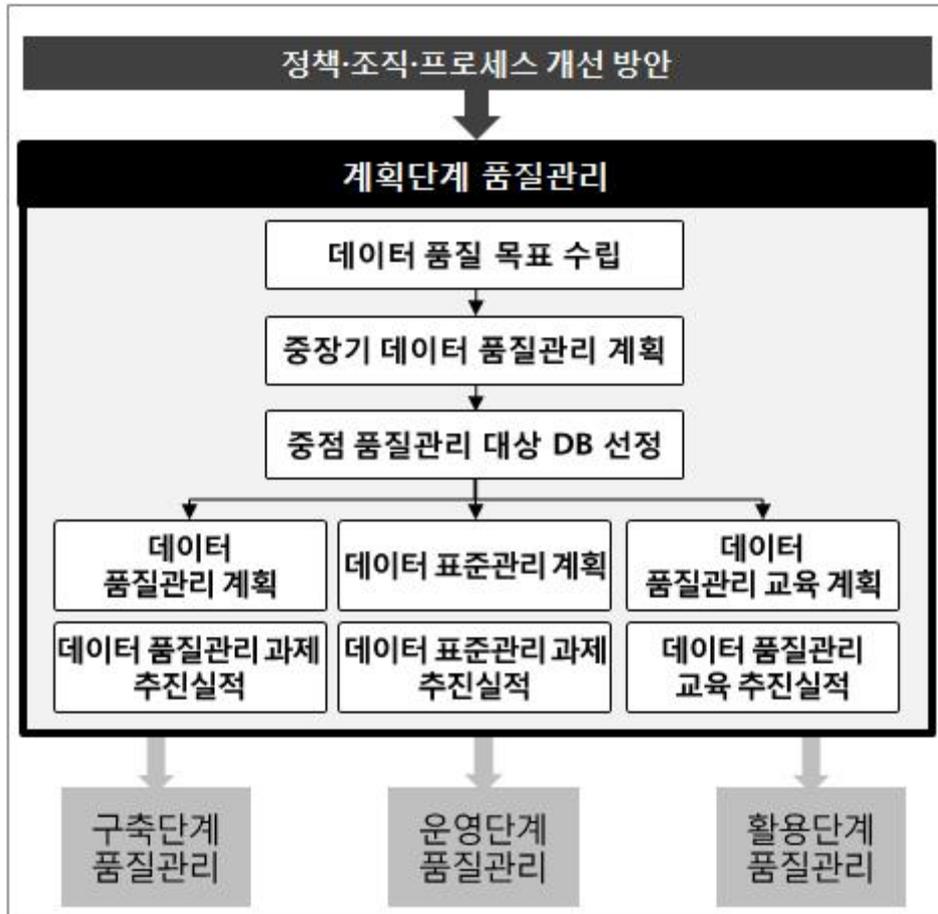


- 데이터의 최신성, 정확성, 상호연계성이 확보되도록 효과적으로 관리하여 고품질의 데이터를 제공토록 지속적이고 상시적인 데이터 품질관리체계를 확보해야 함

□ 계획단계 품질관리 방안

- 계획단계 품질관리는 과학기술정보통신부에서 매년 정보화 계획 등 수립 시 고려해야 할 데이터 품질 측면의 계획을 수립하는 활동을 의미함

[그림 6-4] 계획단계 품질관리 주요 활동 내역



- 데이터 품질 목표 수립
 - 데이터 품질관리를 위해 수립하는 각종 계획 수립 시 명확한 품질관리 목표 제시가 필요하며, 특히 「중장기 데이터 품질관리 계획」 및 「데이터 품질관리 계획」 수립 시 수립하는 것이 원칙
- 중장기 데이터 품질관리 계획
 - 보통 3년~5년 동안의 기간동안 데이터 품질관리체계 수립 혹은 고품질의 데이터 품질 확보를 위한 일련의 활동에 대한 계획
- 중점 품질관리 대상 DB 선정
 - 일정 기간(단기, 중기, 장기) 동안 집중적으로 관리할 대상 DB를 선정하는 것

- 데이터 품질관리 계획 및 추진실적 점검
 - 「중장기 데이터 품질관리 계획」에 따라 매년 「데이터 품질관리 계획」을 수립하고, 계획 대비 추진실적을 점검
- 데이터 표준관리 계획 및 추진실적 점검
 - 매년 「데이터 표준관리 계획」을 수립하고, 계획 대비 추진실적을 점검
- 데이터 품질관리 교육 계획 및 추진실적 점검
 - 매년 「데이터 품질관리 교육 계획」을 수립하고, 계획 대비 추진실적을 점검

○ 주요 산출물은 데이터 품질관리 계획서 및 추진실적 점검 결과서, 데이터 품질관리 교육 계획서 및 결과서, 데이터 표준관리 계획서 및 결과서, 표준 정의서 등이 있음

[표 6-4] 계획단계 품질관리 주요 산출물

산출물 명	기간 및 횟수
▪ 데이터 품질관리 계획서	1회/연
▪ 품질관리 대상 DB 목록	
▪ 데이터 품질관리 표준 프로세스	
▪ 데이터 품질관리 추진실적 점검 결과서	
▪ 중장기 데이터 품질관리 계획서	
▪ 중장기 (3~5년) 데이터 품질관리 추진실적 점검 결과서	3~5년
▪ 데이터 품질관리 교육 계획서	1회/연
▪ 데이터 품질관리 담당자 지정 내역	
▪ 데이터 품질관리 교육 자료	
▪ 데이터 품질관리 교육 결과서	1회/연, 2회/연 이상
▪ 데이터 품질관리 교육 참석 확인증 (또는 참석자 친필 싸인 목록)	
▪ 데이터 품질관리 의사소통 회의자료	2회/연 이상
▪ 데이터 표준관리 계획서	
▪ 데이터 표준관리 절차서	
▪ 데이터 표준 정의서 (용어 ,코드 ,도메인)	
▪ 데이터 표준관리 추진실적 점검 결과서	

중장기 데이터 품질관리 계획 수립

○ 과학기술정보통신부의 지속적인 데이터 품질관리체계 확립을 위해서는 중장기적 관점에서의 데이터 품질관리 목표, 단계별 실행 계획 및 로드맵 등 품질관리 계획 수립이 필요함

[표 6-5] 중장기 데이터 품질관리 계획 수립 (예시)

구분	도입기:~2019년	확산기:2020년~2022년	안정기:2023년~	
품질관리목표	데이터 품질관리 기반 마련	데이터 품질관리 확대 적용	데이터 품질관리 안정화	
중장기 품질 관리 주요 과제	품질관리 지침 수립	<ul style="list-style-type: none"> 과학기술정보통신부 품질관리 지침 제정 	<ul style="list-style-type: none"> 품질관리 지침 검토 및 필요 시 개정 	
	중점 관리 대상 DB	<ul style="list-style-type: none"> 과학기술정보통신부에서 보유한 전체 DB 중 중점 품질관리 대상 DB 선정 기준에 따라 선정 데이터 품질관리가 시급한 DB 부터 우선 선정하여 단계적으로 확산 관리 되도록 진행 		
	품질관리 교육	<ul style="list-style-type: none"> 데이터 품질 교재 제작 및 커리큘럼 수립 품질 부서 (데이터 표준팀, 중점 관리 DB 데이터 품질관리 담당자 대상 교육 실시 <ul style="list-style-type: none"> -기본교육 : 전체 품질 관련 담당자 대상 -전문교육 : 각 세부과목에 해당하는 DB 담당자 및 협력업체 대상 세부 교육 	<ul style="list-style-type: none"> 데이터 품질 교재 제작 및 커리큘럼 수정·보완 중점 관리 대상 DB 변경에 따른 담당자 교육 실시 점진적 데이터 품질 문화 확산을 위한 교육 대상 확대 	
	데이터 표준화	<ul style="list-style-type: none"> 과학기술정보통신부 전체의 데이터 표준화 기준 수립 <ul style="list-style-type: none"> -표준 사전 작성 (용어, 코드, 도메인) 중점 관리 대상 DB 표준화 	<ul style="list-style-type: none"> 데이터 표준화 대상 DB 확대 표준 적용률 (=표준적용 컬럼수 /전체 컬럼 수 비율)상향 조정 	
	품질 진단 및 개선	<ul style="list-style-type: none"> 중점 품질관리 대상 DB 품질 진단 및 개선 <ul style="list-style-type: none"> -오류율 측정 및 개선 과제 도출 -오류 개선 	<ul style="list-style-type: none"> 중점 품질관리 대상 DB 품질 진단 및 개선 <ul style="list-style-type: none"> -오류율 측정 및 개선 과제 도출 -오류 개선 	
	산출물 관리	<ul style="list-style-type: none"> 데이터 품질관리 산출물 작성 	<ul style="list-style-type: none"> 데이터 품질관리 산출물 현행화 	
	품질관리 수준평가	<ul style="list-style-type: none"> 공공데이터 품질관리 수준평가 대응 평가 대상 DB 선정 목표 등급 설정 	<ul style="list-style-type: none"> 평가 대상 DB 선정 전년도 목표 등급 대비 상향 목표 설정 	
	기타	<ul style="list-style-type: none"> 데이터 품질관리를 위해 필요한 과제 추가 발굴 		

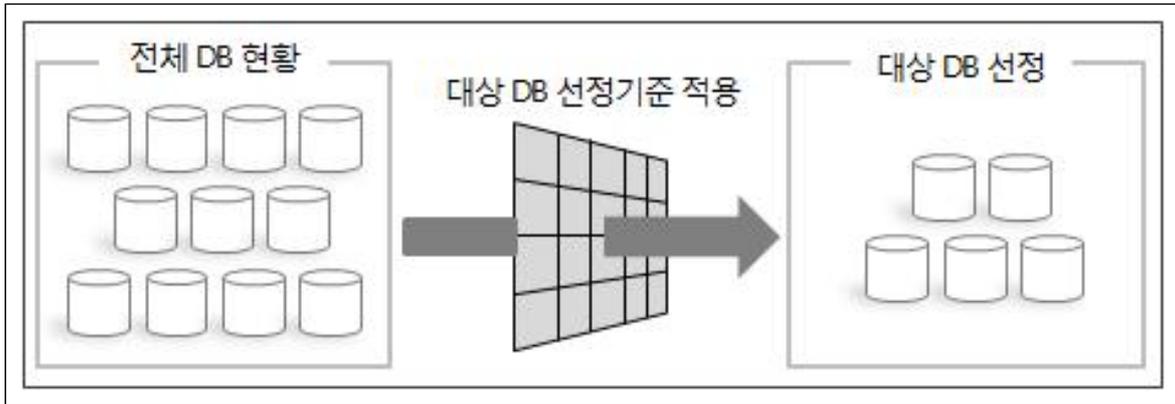
□ 중점 품질관리 대상 DB 선정

- 궁극적으로는 과학기술정보통신부에서 운영 중인 모든 데이터에 대한 품질관리를 수행하는 것이 목표이나, 단기 및 중장기적으로는 중점적으로 품질관리해야 할 대상 DB를 선정하여 선택적·집중적으로 관리하는 것이 합리적임



○ 중점 관리 DB 선정 기준

[그림 6-5] 중점 관리 DB 선정 절차



- 영향도, 이용자 활용도, 데이터 연계 규모 등을 기준으로 점수가 높은 상위 일정 DB를 대상으로 중점적으로 관리할 DB를 선정함
- 데이터 품질 진단 시 예상되는 적정 DB 규모를 고려하여 선정함
- 적정 규모는 과학기술정보통신부 전체 DB중에서 고유업무DB 40%, 지원업무DB 30% 수준으로 최소 5개 이상, 최대 25개 이하로 선정하는 공공데이터 품질관리 수준평가의 기준을 따르는 것을 권장함
- 선정 점수가 동점일 경우에는 영향도 > 연계시스템 수 > 이용자 수 > DB규모 (데이터건수) 순으로 순위를 선정함

[표 6-6] 진단 대상 DB 선정 기준

선정기준		점 수	분류 기준	비 고
영향도 (40점)		40점	업무 영향범위 1등급	
		30점	업무 영향범위 2등급	
		20점	업무 영향범위 3등급	
		10점	업무 영향범위 4·5등급	
이용자 활용도 (40점) ※3개 중 택1	대국민 사용자	40점	50만 이상	사용자수, 처리 건수는 일 평균치임
		20점	50만 미만에서 10만 이상	
		10점	10만 미만	
	기관 사용자	40점	50% 이상	
		20점	50% 미만에서 25% 이상	
		10점	25% 미만	
	처리 건수	40점	100만건 이상	
		20점	100만건 미만에서 50만건 이상	
		10점	50만건 미만	
연계규모 (20점)		20점	연계된 타 정보시스템 수가 10개 이상인 DB	평가대상 DB와 연계된 내· 외부 정보시스템 수
		10점	연계된 타 정보시스템 수가 5~9개인 DB	
		5점	연계된 타 정보시스템 수가 5개 미만인 DB	

□ 데이터 품질관리 계획 수립

○ 데이터 품질관리 계획 수립 원칙

- 기 수립된 「중장기 데이터 품질관리 계획」 중 당해 연도에 해당하는 「데이터 품질관리 계획」을 매년 초 수립함
- 「데이터 품질관리 계획」은 중장기 품질관리 계획에서 수립된 과제와 일관성을 유지해야 함
- 만일 「중장기 데이터 품질관리 계획」을 수립하지 않았다면, 별도의 「데이터 품질관리 계획」만을 수립할 수도 있음

○ 데이터 품질관리 계획 주요 내용 (예시)

- 품질관리 목표 : 과학기술정보통신부의 정보화 로드맵 등을 고려하여 당해 연도 (또는 중·장기) 데이터 품질관리 목표 기술
- 품질관리 현황 및 시사점 : 전년도의 품질관리 실적, 성과 등을 고려하여 과학기술정보통신부의 데이터 품질관리 전반에 대한 현황 및 시사점 기술
- 품질관리 추진 방향 : 데이터 품질 관리 목표, 현황 및 문제점에 따른 과학기술정보통신부의 당해 연도 품질관리 추진 방향 기술
- 중점 관리 대상 DB 선정 : 과학기술정보통신부의 특성을 고려하여 선정하고, 해당 DB의 주요 품질관리 대상 정보 등을 기재
- 대상 DB 일반 현황 : DBMS 종류·버전, 테이블수, 컬럼수, 용량 등
- 수행조직 : 품질관리를 주체적으로 수행할 조직 정의
- 데이터 품질관리 주요 과제 정의 : 당해 연도 수행해야 할 과제 정의
 - 「중장기 데이터 품질관리 계획」의 과제와 일관성을 유지해야 하며, 데이터 품질 진단 및 개선 계획, 표준화 적용 계획, 연계데이터 품질 확보 계획, 산출물 관리 계획 등을 기재
 - 「데이터 품질관리 계획서」에는 공공데이터 품질관리에 대한 전반적인 내용이 포함 될 수 있기 때문에 「데이터 품질관리 교육 계획서」 등 타 계획 관련 내용이 포함될 수 있음
- 기타 : 인프라 구축 및 운영계획 등 과학기술정보통신부의 데이터 품질관리 계획에 필요한 일반 사항을 기재

○ 데이터 품질 이행 점검

- “2021년 공공데이터 제공 및 이용활성화 시행계획”에 명시된 데이터 품질 및 표준화 관련 추진 내용 이행 점검
- 품질관리 규정 마련, 데이터 품질 오류 진단·개선, 데이터 표준 점검 및 표준화, 차세대 시스템 데이터표준 적용 등(32건) 등 추진 점검

□ 데이터 표준관리 계획 수립

○ 데이터 표준 관리 목적

- 데이터에 대한 공통된 시각을 유지할 수 있도록 표준 단어, 표준 용어, 표준 도메인, 표준 코드 등 과학기술정보통신부 전체적으로 관리 기준을 정의하기 위함

[표 6-7] 데이터 표준화 목적

구분	내용
명칭통일을 통한 의사소통 증대	동일한 데이터에 대해서는 동일한 명칭을 사용함으로써 개발자-현업, 운영자-현업, 운영자-운영자 등 다양한 계층간 명확하고 신속한 의사소통이 가능
데이터 소재 파악의 시간 및 노력 감소	정보 사용자는 데이터의 의미와 데이터의 위치 등을 신속하게 파악할 수 있어 적시에 정확한 정보를 활용
데이터 품질 향상	데이터 형식 및 규칙을 데이터 표준에 맞게 적용하여 데이터의 입력 오류를 예방함으로써 데이터의 품질을 향상
데이터 변환, 정제 비용 감소	데이터의 전송, 공유, 가공을 위해 정보시스템 간 또는 정보시스템 내에서 별도의 포맷 변환이나 정제 작업 불필요

- 데이터 표준관리는 데이터 항목들에 대해서 과학기술정보통신부 전체적으로 일관된 명명법과 정의를 부여하고 이를 정제, 개선하는 활동임

[그림 6-6] 데이터 표준관리 기준



- 단어, 용어, 도메인, 코드 등에 대한 데이터 표준화를 수립하며 주요 내용은 다음과 같음

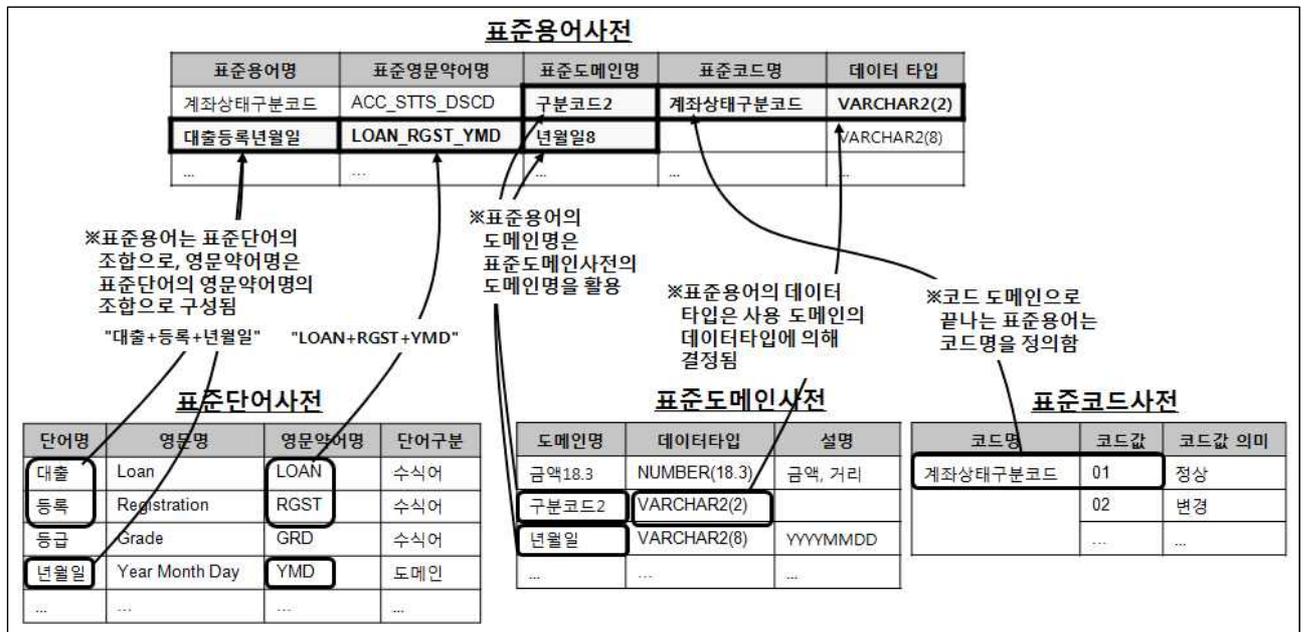
[표 6-7] 표준 데이터 종류

표준 종류	내용
표준 단어	<ul style="list-style-type: none"> 일반적으로 ‘단어’란 문법상 일정한 뜻과 구실을 가지는 일의 최소 단위를 의미함 ‘표준 단어’란 업무 수행 상 사용하며 일정한 의미를 갖고 있는 최소단위를 의미함 표준 용어를 구성하는데 사용됨
표준 용어	<ul style="list-style-type: none"> ‘표준 용어’란 업무 수행 시 사용되는 용어를 정보시스템에서 사용하는 기술적인 용어로 전환하여 이것을 일관되게 사용 할 수 있도록 정의한 것을 의미함 ‘표준 용어’는 ‘표준 단어’와 ‘표준 도메인’, ‘표준 코드’ 등을 조합하여 구성하며 최소한 한 개의 표준 도메인으로 구성됨 모델에서는 속성명으로 사용되며 과학기술정보통신부 전체적인 관점

표준 종류	내용
	에서 유일함
표준 도메인	<ul style="list-style-type: none"> ‘표준 도메인’ 이란 데이터의 성격을 분류한 것으로 동일한 형식을 가진 집합을 의미하며, 하나의 용어는 하나의 도메인만 지정됨 동일한 형식을 가진 데이터에 대해서 같은 도메인을 적용함으로써 속성의 의미 및 데이터의 범위를 명확히 할 수 있고 컬럼에 대한 일관적인 관리가 가능하게 됨 용어 조합의 마지막에 위치하는 도메인 단어(분류어)가 도메인의 후보가 됨
표준 코드	<ul style="list-style-type: none"> ‘코드’ 란 활용하고자 하는 데이터를 약어 혹은 기호로 함축하여 사용하는 데이터를 의미 ‘표준 코드’ 란 도메인의 한 유형으로써, 속성(칼럼)에 허용된 데이터 값을 제한된 범위 내에서 구체적으로 열거하여 정의한 것을 의미 이 데이터 값을 코드 값 또는 유효값이라 하며 각각의 코드값에는 의미를 부여하고, 이 의미를 ‘코드값명’ 또는 ‘코드유효값정의’ 라 함

○ 표준 용어 사전은 표준 단어 사전, 표준 도메인 사전, 표준 코드 사건의 내용을 토대로 그 조합으로 생성됨

[그림 6-7] 데이터 표준화 요소 간 관계



- '표준 용어명' 은 표준 단어 사전의 '단어명', '영문약어명' 각각의 조합으로 생성
- '표준 도메인명' 은 표준 도메인 사전의 '도메인명' 을 활용하고, '데이터 타입' 은 표준 도메인 사전의 '데이터타입' 에 의해 결정됨
- '표준코드명' 은 표준 코드 사전의 '코드명' 을 사용함

○ 데이터 표준관리 계획 수립 원칙

- 「데이터 품질관리 계획」 내 표준화 관련 과제가 존재할 경우 상세화 하여 당

해 연도에 해당하는 데이터 표준관리 계획을 매년 초 수립함

- 데이터 표준 관리 계획은 「데이터 품질관리 계획서」에 포함되어 작성이 가능함
- 「데이터 표준관리 계획」은 「데이터 품질관리 계획」에서 수립된 과제와 일관성을 유지해야 함
- 만일 「데이터 품질관리 계획」을 수립하지 않았다면, 별도의 연도별 「데이터 표준관리 계획」만을 수립할 수도 있음

○ 데이터 표준관리 계획 주요 내용 (예시)

- 표준화 목표 : 「데이터 품질관리 계획」 등을 고려하여 당해 연도 데이터 표준 관리 목표 기술
- 표준화 현황 및 시사점 : 전년도의 표준관리 실적, 성과 등을 고려하여 과학기술정보통신부의 데이터 표준관리 전반에 대한 현황 및 시사점 기술
- 표준화 추진방향 : 데이터 표준화 목표, 현황 및 문제점에 따른 과학기술정보통신부의 당해 연도 표준관리 추진 방향 기술
- 표준화 적용 대상 DB 선정 : 사전 예방적 품질관리 및 선택·집중적 표준관리 측면을 모두 고려하여 선정
 - 「데이터 품질관리 계획」 내 정의된 중점 관리 대상 DB로 선정 가능
 - 사전 예방적 품질관리 측면 : 당해 연도에 신규 또는 고도화 사업 등을 통해 구축되는 DB를 파악하여 표준화 적용 대상에 포함
 - 선택·집중적 품질관리 측면에 : 중점 품질관리 대상 DB 중에서 표준화 수준이나 개선의 긴급성 등을 고려하여 당해 연도 표준화 적용 대상을 선정
- 데이터 표준관리 주요 과제 정의 : 당해 연도 수행해야 할 과제 정의
 - 데이터 표준 관리 지침 마련, 표준용어·표준코드·표준도메인 등 정의 및 관리계획 수립, 표준관리 시스템 도입, 표준적용 확산방안, 표준준수 여부 점검 및 조치 계획 수립, 표준화 수행 일정 및 소요 예산 등

○ 데이터 표준 이행 점검

- “2021년 공공데이터 제공 및 이용활성화 시행계획”에 명시된 데이터 품질 및 표준화 관련 추진 내용 이행 점검
- 품질관리 규정 마련, 데이터 품질 오류 진단개선, 데이터 표준 점검 및 표준화, 차세대 시스템 데이터표준 적용 등(32건) 등 추진 점검

□ 데이터 품질관리 교육 계획 수립

○ 데이터 품질관리 교육 목적

- 품질 관련 담당자를 대상으로 데이터 품질관리 교육을 수행함으로써 데이터

품질관리 업무 수행을 위한 역량 강화를 위한

○ 교육 대상

- 데이터 품질관리 총괄책임자, 데이터 품질관리 실무담당자, 데이터 품질관리 업무담당자(중점 품질관리 대상, 품질관리 수준평가 대상 DB 등) 전원 및 기타 데이터 품질에 관심이 있는 구성원 전체
- 따라서, 교육 대상 선정을 위해서 품질관리 대상 DB 선정 및 데이터 품질관리 담당자 지정이 선행되어야 함

○ 교육 시간

- 중점 품질관리 및 품질관리 수준평가 대상 DB 업무담당자는 반드시 일정 시간 (연 2회 이상)의 교육을 이수해야 함

○ 교육 방법

- (내부 교육) 과학기술정보통신부 보유 및 운영 중인 데이터에 대한 품질관리를 위한 집합교육
- (외부 교육) 행정안전부·한국지능정보사회진흥원에서 실시하는 범정부 차원의 집합교육
- 내부 교육 대체 방안
 - 내부 교육이 여의치 않을 경우 행정안전부, 한국지능정보사회진흥원 등 범정부에서 실시하는 품질관리 교육에 참여 가능함

○ 교육 내용

- (내부 교육) 품질관리 지침, 데이터 품질관리체계, 공공데이터 품질관리 수준평가, 데이터 품질관리 시스템 및 메타데이터 관리시스템 교육 등
- (외부 교육) 품질관리 기본, 실무, 전문교육 등

[표 6-8] 외부 교육 과정 사례

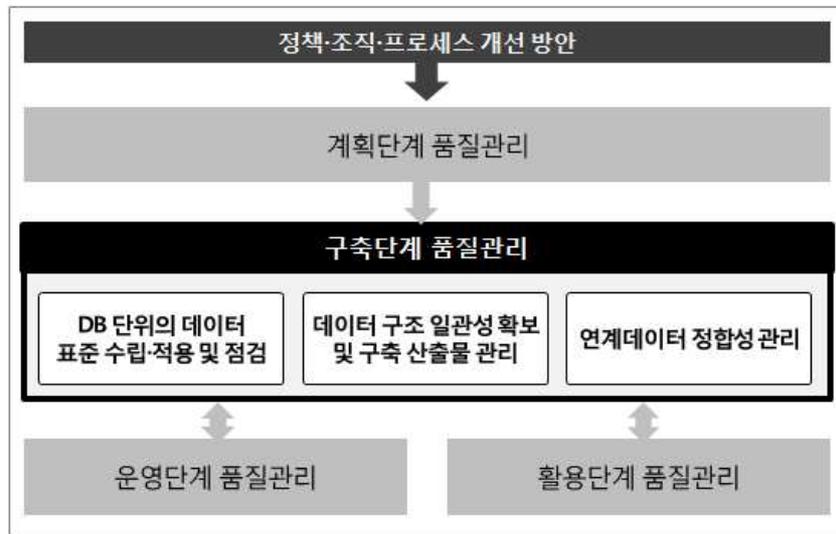
과정명		교육내용	시간	비고
기본 (초급)	품질관리 기본	▪ 품질관리 방법, 수준평가 지표 이해 등	3시간	이론
	품질 진단 기본	▪ 품질진단 이해, 진단도구 사용법 등	3시간	실습
실무 (고급)	품질관리 실무	▪ 품질관리 수준평가 지표 해설, 품질진단 유형 및 오류사례, 품질진단도구 활용 실습 등	4시간	이론 / 실습
전문 (고급)	품질관리 전문	▪ 품질관리 체계, 데이터 정규화, SQL 활용 프로파일링, 업무규칙 도출, 오류개선 등	18시간	이론 / 실습

□ 구축단계 품질관리 방안

- 구축단계 품질관리는 과학기술정보통신부에서 정보시스템 신규 도입 또는 고도화 등을 추진하기 위한 구축사업 단계에서 고려해야 할 데이터 품질관리 활동을 의미함



[그림 6-8] 구축단계 품질관리 주요 활동 내역



- DB 단위의 데이터 표준 수립·적용 및 점검 : 과학기술정보통신부의 표준을 반영하여 각 DB 별 데이터 표준을 정의하고, DB 단위의 데이터 표준 적용을 위한 계획 수립
- 데이터 구조 일관성 확보 및 구축 산출물 관리 : 정보시스템 구축 시 데이터 구조 정의 및 관련된 산출물을 지속적으로 관리하는 활동 수행
- 연계데이터 정합성 관리 : 연계데이터 관리 정보를 정의하고, 연계데이터의 정확성·신뢰성이 확보될 수 있도록 지속적으로 관리하는 활동 수행

○ 주요 산출물은 데이터 표준 적용 확산 계획서, DB별 데이터 표준 정의서, 테이블 정의서, 컬럼 정의서, 데이터 모델 다이어그램, 연계데이터 목록 정의서 등이 있음

[표 6-9] 구축단계 품질관리 주요 산출물

산출물 명	산출물 명
▪ 데이터 표준 적용 확산 계획서	▪ 논리데이터모델 다이어그램
▪ DB 별 데이터 표준 정의서(DB 표준용어/코드/도메인 정의서)	▪ 엔터티 정의서
▪ 데이터 표준 현황 분석서	▪ 속성정의서
▪ 데이터 표준 적용 컬럼 현황	▪ DB 산출물과 실제 DB 구조와의 현행화 (데이터 구조 산출물과 DBMS 카탈로그와의 비교표)
▪ 데이터 표준 정의서 변경 내역 (DB 표준 정의서 변경 이력 관리)	▪ 연계데이터 목록 정의서
▪ 데이터 표준 매핑테이블	▪ 연계 솔루션 캡처 화면
▪ 범정부 데이터 표준 반영 활용 내역 (범정부 표준 적용 검토 의견서 또는 범정부 표준 적용 결과 보고서 명칭의 문서 요구 대응 시 제출 가능)	▪ 연계데이터 품질점검 내역
	▪ 연계데이터 송수신 이력(혹은 로그를 기록한 시스템 내역)
	▪ 연계데이터 품질 개선 조치 결과서 (보고, 공유 내역 포함)
▪ 테이블 정의서	▪ 연계 관련 정기 회의체 운영 실적 (회의록 등)
▪ 컬럼 정의서	▪ 연계 (주기, 방법, 기술 등) 개선 내역
▪ 물리데이터모델 다이어그램	▪ 연계데이터 중장기 개선계획서 (3~5년)

□ DB 단위의 데이터 표준 수립 · 적용 및 점검

[표 6-10] 계획단계 표준화 vs 구축단계 표준화

구분	계획단계 표준화	구축단계 표준화
표준 수립 대상	▪ 기관 단위	▪ DB 단위
표준 수립 계획 내용	▪ 기관 단위 데이터 표준관리 계획 수립	▪ 데이터 표준 개발 및 적용 확산 계획 ※ 『데이터 표준관리 계획서』 내에 포함 되어 작성 될 수도 있음
표준 정의	▪ 기관 단위의 용어, 도메인, 코드 표준 수립	▪ DB 단위의 용어, 도메인, 코드 표준 수립
주요 측정 지표	▪ 계획대비 실적 달성률	▪ 데이터 표준 적용률

○ DB 단위로 표준화를 적용하지 못했거나, 혹은 부분적용 한 경우 과학기술정보통신부 차원의 차세대 시스템(시스템 고도화, 데이터 통합, 기관 간 데이터 연계 사업 등) 도입 추진 시 적용하는 것이 원칙이나, 운영 단계에서 지속적으로 표준화를 위한 노력을 수행해야 함

○ 데이터 표준 적용 확산 계획 수립

- 데이터 표준 적용 확산 계획 수립 원칙

- 「데이터 품질관리 계획」, 「데이터 표준관리 계획」 등에서 수립된 과학기술정보통신부 표준을 반영하여 DB 단위의 데이터 표준을 정의
- 「데이터 표준 적용 확산 계획」은 「데이터 표준관리 계획」에서 수립된 표준과 일관성을 유지해야 함
- DB 단위의 표준 용어, 표준 코드, 표준 도메인 등이 과학기술정보통신부의 표준을 포함하여 정의 되어야 함

- 데이터 표준 적용 확산 계획 주요 내용 (예시)

- 과학기술정보통신부 표준과 DB 표준 간 비교분석 등을 통한 표준 현황 파악, 과학기술정보통신부 표준과 DB 표준 간 용어 매핑, 실제 DB 등에 표준을 반영할 활동계획 등

○ ‘데이터 표준 적용률’ 점검

- 과학기술정보통신부 표준과 DB 표준간의 데이터 표준을 비교 분석, 중복 제거 후 적절히 적용 되었는지를 판단하는 도구로 ‘데이터 표준 적용률’을 활용

- ‘데이터 표준 적용률’ 개념

- 데이터 표준 적용률은 DB 내 전체 컬럼 수 중 표준 용어, 표준 도메인이 적용된 컬럼수의 비율을 의미

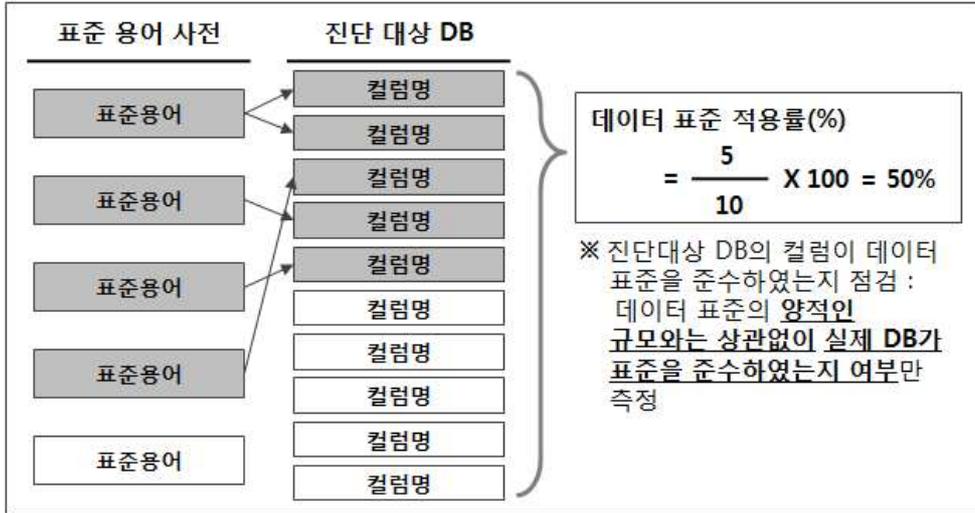
$$\text{데이터 표준 적용률(\%)} = \frac{\text{② 진단 대상 DB의 표준(용어+도메인)이 적용된 컬럼 수}}{\text{① 전체 컬럼 수}} \times 100$$

① 전체 컬럼 수 : 표준화 진단 대상 DB의 전체 컬럼 수로 중복 제거된 개수

② 대상 DB의 데이터 표준(용어+도메인)이 적용된 컬럼 수 : 표준화 진단 대상 DB의 컬럼에 데이터 표준(용어와 도메인 모두)이 적용된 컬럼의 개수로 실제 DB의 테이블 컬럼(용어+도메인 비교 컬럼) 기준 중복이 제거된 수

- ‘데이터 표준 적용률’ 산정 절차
 - 표준 적용 비교표 작성→전체 컬럼 수 계산→진단대상 DB의 표준(용어+도메인)이 적용된 컬럼 수 계산→데이터 표준적용률 계산
- ‘데이터 표준 적용률’ 산정 (예시)

[그림 6-9] 구축단계 표준적용률 개념



- 아래 [표 6-11]에서 과학기술정보통신부의 표준 용어 사전과 실제 DB 테이블의 컬럼을 비교해 보면, 대상 DB의 12개 컬럼 중 중복인 되는 테이블 ‘TB_TEST1’의 용어 비교 컬럼인 ‘1_AMT (NUMBER11.3)’, 테이블 ‘TB_TEST2’의 용어 비교 컬럼 ‘CODE(CHAR4)’ 등 2개 컬럼을 중복제거 하면 전체 컬럼 수는 12-2=10(개)이 됨
- 그리고, 위에서 구한 전체 컬럼 수(10개)에서 비표준 컬럼인 ‘DATE(DATE)’, ‘CODE(CHAR4)’ 등 6개 컬럼을 제외하면 실제 DB의 용어, 도메인이 적용된 컬럼 수는 10-6=4(개)가 됨

[표 6-11] 과학기술정보통신부 표준 용어 사전 vs 실제 DB 테이블의 컬럼

번호	표준 용어 사전 (전체)				실제 진단 대상 DB 테이블의 컬럼					표준 적용	중복 여부	
	한글명	영문명	타입	길이	용어비교 컬럼	테이블명	영문명	타입	길이			용어비교 컬럼
1	1_금액	1_AMT	NUMBER	11.3	1_AMT (NUMBER11.3)	TB_TEST1	1_AMT	NUMBER	11.3	1_AMT (NUMBER11.3)	Y	중복
2	2_일자	2_DATE	DATE		2_DATE (DATE)	TB_TEST1	2_DATE	DATE		2_DATE (DATE)	Y	
3	3_코드	3_CODE	VARCHAR2	5	3_CODE(VARCHAR25)	TB_TEST1	3_CODE	VARCHAR2	5	3_CODE (VARCHAR25)	Y	
4	4_번호	4_NO	NUMBER	6	4_NO(NUMBER6)	TB_TEST1	4_NO	NUMBER	6	4_NO (NUMBER6)	Y	
5						TB_TEST2	1_AMT	NUMBER	11.3	1_AMT (NUMBER11.3)	N	중복
6						TB_TEST2	DATE	DATE		DATE(DATE)	N	
7						TB_TEST2	CODE	CHAR	4	CODE(CHAR4)	N	중복
8						TB_TEST2	NO	CHAR	4	NO(CHAR4)	N	
9						TB_TEST3	CODE	CHAR	4	CODE(CHAR4)	N	중복
10						TB_TEST3	FIL2	CHAR	10	FIL2 (CHAR10)	N	
11						TB_TEST3	FIL3	CHAR	10	FIL3 (CHAR10)	N	
12						TB_TEST3	FIL4	CHAR	10	FIL4 (CHAR10)	N	
13	5_명	5_NAME	VARCHAR2	30	5_NAME(VARCHAR230)							
14	6_여부	6_YN	CHAR	1	6_YN(CHAR1)							
15	7_납부구분코드	7_MF_DIV_CODE	CHAR	2	7_MF_DIV_CODE(CHAR2)							
16	8_사원번호	8_MEMBER_NO	NUMBER	6	8_MEMBER_NO(NUMBER6)							

- 따라서, ‘데이터 표준 적용률(%)’ 은 $(4 \div 10) \times 100 = 40(\%)$ 로 산정됨

○ 표준 변경 이력 관리

- 데이터 표준 용어, 표준 코드, 표준 도메인 각각의 생성, 변경, 삭제 등 변경 이력을 기록하여 관리해야 하며, 이때 문서 단위의 이력이 아닌 건별 생성, 변경, 삭제 일자를 관리 해야 함

[표 6-12] 표준 용어 사전에 건별 신규·변경·삭제일자를 관리하고 있는 사례

변경 구분	변경일자	변경 요청자	용어명	영문약어명	도메인명	데이터 타입	코드명	용어설명
신규	2016.1.6	홍길동	가감구분 코드	ADSUB_DSTCD	구분코드 1	VARCHAR2(1)	가감구분 코드	+인지(가) -인지(감) 구분하는 코드
변경	2018.2.3	김정호	가감부호 구분코드	ADSUB_SIN_DSTCD	구분코드 1	VARCHAR2(1)	가감구분 코드	+, - 부호를 구분하는 코드
삭제	2018.6.2	허준	가감비율 개시년월일	ADSUB_RATIO_OPENING_YMD	년월일 8	VARCHAR2(8)		가감 비율을 적용하기 시작한 개시 년월일
...

□ 데이터 구조 일관성 확보 및 구축 산출물 관리

- 정보시스템 구축 시 데이터의 중복이 최소화되도록 데이터 모델을 설계 및 구축해야 하며, 불가피한 데이터 중복이 발생하는 경우 중복 데이터의 일관성이 확보되도록 관리하기 위해 산출물을 작성하여 운영단계의 입력물로 활용
- 정보시스템의 DB 설계 시 준수사항
 - 데이터 모델은 집을 지을 때의 설계도와 같아서, 데이터 모델의 설계가 잘못되면 데이터의 중복이 많아지고, 시스템의 구조가 복잡해지는 등 오류 데이터가 발생할 가능성이 높음
 - 데이터의 중복을 최소화하고, 불가피한 중복의 경우에는 데이터의 일관성 확보를 위해 검증규칙을 마련하여 지속적인 모니터링이 필요함
 - 정보시스템 구축사업 시 중요 데이터에 대해서는 구축사업자가 데이터의 일관성 확보 및 오류 데이터 검증을 위한 업무규칙 등을 작성·제출토록 사업요건에 반영 필요
- 정보시스템 구축 시 오류데이터 유입 방지를 위한 준수사항
 - 오류 데이터의 사후 개선은 원인분석에 장시간이 소요되는 등 오류 정제를 위한 비용과 노력이 많이 요구됨
 - 따라서, 오류데이터의 유입을 사전에 차단할 수 있도록 중요 데이터를 등록하는 입력 프로그램의 경우 입력 오류 검증 체계를 마련하여 정확한 데이터의 입력을 유도
- 산출물 작성 및 관리
 - DB 설계·구축 시 작성해야 할 표준 산출물 목록을 정의하고, 해당 산출물을 작성·관리해야 함

- 구축단계 데이터 품질관리 대상 산출물 목록에서 제시하는 필수 산출물은 DB 설계, 구축 시 반드시 작성하여 관리
- DB 설계·구축 각 단계별로 관련 산출물이 적절하게 작성되었는지, 실제 DB와 산출물이 일치하는지를 점검하여 불일치 사항을 즉시 보완토록 산출물 관리 절차와 담당자를 지정하여 운영 ※사업 발주 시 제안요청서 등에 산출물 명확히 제시 및 산출물 관리 담당자 지정, 검수 시 실제 DB와 일치 여부 점검 및 보완조치 등 실시

○ 구축단계 데이터 산출물 관리의 중요성

- 구축단계의 데이터 산출물 관리는 운영단계의 데이터 품질진단 및 개선 등의 품질관리 활동에 필요한 입력정보로 활용됨
- 또한, 업무 변화에 따른 데이터 모델 변경, 응용 프로그램 개선 등이 필요한 경우 데이터에 영향을 미치는 사항을 파악하는 데 기초 자료로 이용됨
- 따라서, 구축단계에서는 산출물 목록을 참조하여 실제 DB와 일치하도록 정확한 산출물을 확보하는 것이 매우 중요함

○ 데이터 구조에 대한 점검 및 개선을 산출물의 현행화(테이블 정의서, 컬럼 정의서 기준)된 정도를 기준으로 산정하므로, 진단 대상 DB의 구조 산출물을 실제 DB의 구조에 맞게 현행화하고 관리해야 함

- 구조 산출물 현행화율(%) = [테이블 정의서] 현행화율과 [컬럼 정의서] 현행화율의 산술평균

$$= \frac{\left[\frac{\text{현행화 테이블 수}}{\text{실제 DB 기준 전체 테이블 수}} \times 100 \right] + \left[\frac{\text{현행화 컬럼 수}}{\text{실제 DB 기준 전체 컬럼 수}} \times 100 \right]}{2}$$

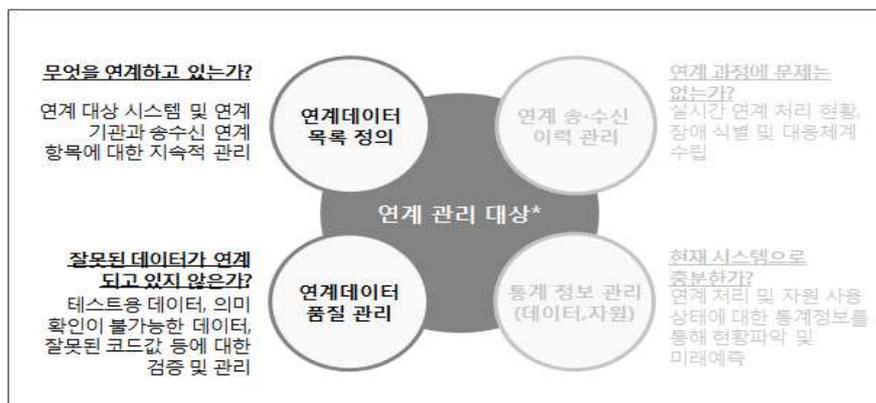
※비교 : 테이블(테이블명), 컬럼(영문명+타입+길이)

※진단 제외 테이블 유형 : 임시테이블, 미사용·중복 테이블, 백업테이블, 로그 기록테이블 등

□ 연계데이터 정합성 관리

- 연계데이터에 대한 현황을 체계적으로 관리하고, 상호 연계에 대한 표준과 절차를 명확히 하여 연계 전 품질 검증 및 연계 후 품질에 대한 수준을 검증하는 과정이 필요

[그림 6-10] 연계 관리 대상



- 내 외부 연계 대상 : 내부(시스템 간), 외부(기관 간)

DB to DB 대상 : 개방 포털 등을 통해 제공되는 XML, 오픈API 등은 연계 대상에서 제외

○ 연계데이터 목록 정의

- 과학기술정보통신부는 타 기관에 제공하거나 타 기관으로부터 제공받아 활용 중인 연계 데이터의 정합성 확보를 위해 연계데이터 목록을 작성하고 관리하여야 함
- 연계데이터 목록은 보유기관 입장에서 제공하는 정보와 활용기관 입장에서 이용하는 연계정보를 모두 관리

[표 6-13] 연계데이터 목록 정의서 (예시)

연계 데이터 관리항목		설명
연계 정보 정의	연계정보 구분	연계 정보 제공 시 “제공”, 활용 시 “활용” 으로 기재
	연계 정보명	연계 정보가 무엇을 의미하는지를 이해할 수 있는 수준에서 정의
	연계 주기	연계 정보의 연계 주기를 정의(실시간, 매주, 매월 등)
연계 항목 정의	연계항목명	연계 정보를 구성하는 세부 연계 항목들에 대한 명칭
	연계항목 설명	연계 정보를 구성하는 연계 항목들에 대한 정의
	연계항목 데이터타입	연계 항목의 데이터 타입(문자형, 숫자형 등)
	연계항목 데이터길이	정의된 연계 항목별 데이터의 길이(10 자리, 20 자리)
연계 항목 출처	연계 DB명	연계 항목을 제공 또는 활용하는 DB명
	연계 테이블명	연계 항목을 제공 또는 활용하는 테이블 이름
	연계 컬럼명	연계 항목을 제공 또는 활용하는 컬럼 이름
오너십 정의	제공기관	연계 정보를 제공하는 기관 정보를 관리(기관명, 담당자, 연락처를 관리하여 향후 연계 협의 시 담당자 식별 등에 활용)
	활용기관	연계 정보를 활용하는 기관 정보를 관리(다수의 활용 기관이 존재하는 경우 전체 기관에 대한 정보를 관리)
비고		연계 정보 관련 행정·기술적 사항을 기재

- 일반적인 연계데이터 목록 현행화가 미흡한 사유

- 설계 완료 후 구축단계에서 ‘연계데이터 목록 정의서’ 를 최신화하지 않고 직접 연계시스템에 적용하는 경우
- 시스템 구축 이후 운영담당자에게 인수인계가 부족한 경우
- 운영담당자가 추가·수정·삭제 시 현행화를 누락하는 경우

○ 연계데이터 품질 관리

- 연계가 정상적으로 수행되었다고 하더라도 실제 데이터의 값이 유효하지 않은 경우에는 별도의 품질 진단을 통해 지속적인 관리가 필요함

[표 6-14] 연계 실제 데이터 오류 유형 (예시)

기관명	테이블명칭	테이블아이디	항목명	실 데이터 값 (예시)	오류 유형
기관A	상세제원 _항만 _계류시설	TCM_KISTEC_DT L_HM	L형 블록식 유무	<u>공백</u> , O, Y, <u>NULL</u>	항목 의미에 부적합
기관A	경계구역	TCO_NEM_BOZO	offline 관리 번호	001, 1-1, <u>가_테스트</u>	항목 의미에 부적합, 테스트 데이터
기관C	점검진단이력	TCM_KISTEC_DIG	상태등급	<u>1, 2</u> , A, B, C, D, E	항목 의미에 부적합

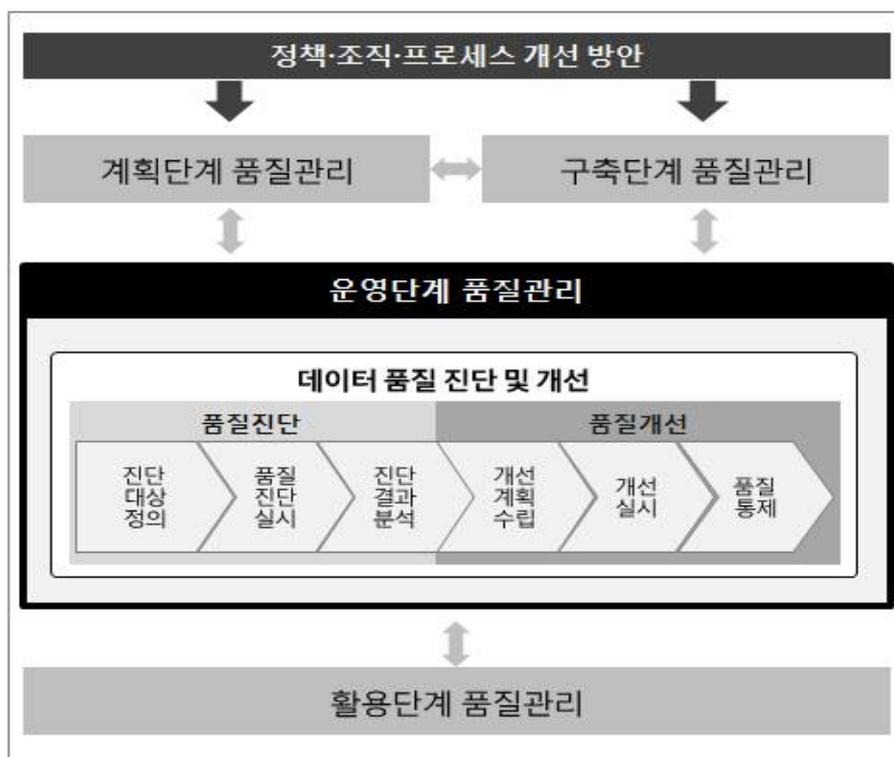


기관명	테이블명칭	테이블아이디	항목명	실 데이터 값 (예시)	오류 유형
		N_RESULT			
기관D	이재민발생및수용정보	TAC_NEM_OCMD	수용세대수	-1000, -25, 0, 148	항목 의미에 부적합
기관D	시장우량	TOB_KOW_DUBHRRFIF	시장우량	-27, 11.5, 44710, .2	항목 의미에 부적합
기관F	의료기관기준병상정보	TCO EMC_HOSP_RES_EINF	입원병상	0, 00, 704, -, 유	항목 의미에 부적합
기관F	구급차정보	TCO EMC_ERCAR_FIX	재해지원	9, N, Y, NULL	항목 의미에 부적합
기관G	돌발상황발생(유고)정보	TOB_RTSA_OCCURID	전송여부	N, O, Y	항목 의미에 부적합
기관H	RTU 정보	TOB_NPA_RTUINFO	전화번호	1,6721665,sorakaws/1	형식 부적합, 테스트 데이터
기관I	상세제원_건축물	TCM_KISTEC_DTL_AR	최고층고_해당층	-3~-2, 01월 25일, 1.9, 1.2.3	형식 부적합, 테스트 데이터

□ 운영단계 품질관리 방안

- 운영단계 품질관리는 과학기술정보통신부가 생성·보유·연계하여 활용하고 있는 데이터를 운영함에 있어 데이터의 품질수준을 향상시키기 위한 제반 활동을 의미함

[그림 6-11] 운영단계 품질관리 주요 활동 내역



- 운영단계의 데이터 품질관리는 데이터 품질 진단 및 개선 활동과 동일시 되는 경우가 많은데, 이는 데이터 품질의 문제는 운영 및 활용 단계에서 주로 진단을 통해 이슈화 되고, 이를 해결하기 위한 개선활동이 수행되기 때문임

- 주요 산출물은 데이터 품질진단 기준 정의서, 데이터 품질 개선 계획, 데이터 품질 개선 조치 결과서 등이 있음

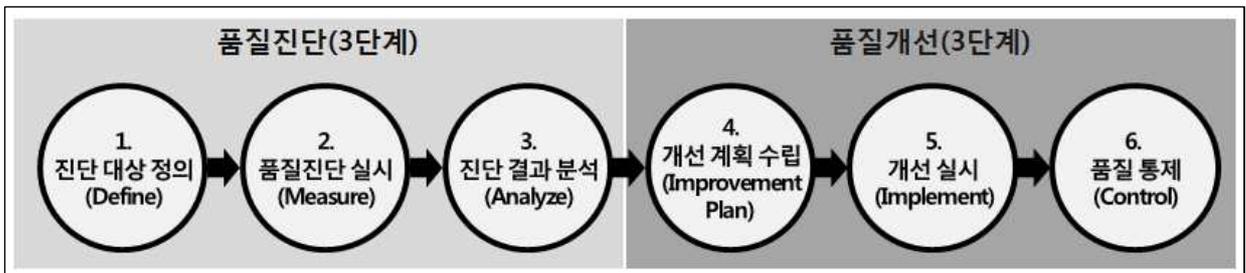
[표 6-15] 운영단계 품질관리 주요 산출물

산출물 명	비고
▪ 데이터 품질진단 기준 정의서	
▪ 데이터 품질진단 대상 정의서	
▪ 데이터 품질 진단 실시 내역	
▪ 데이터 품질 개선 계획 (단기 /중장기 개선 구분)	
▪ 데이터 품질 개선 조치 결과서	
▪ 업무규칙 정의서*	*별도 과제에서 작성중
▪ 데이터 품질진단 테이블 선정 내역	
▪ 데이터 품질진단 항목 진단기준 선정내역	
▪ 데이터 품질진단 결과 내역	

□ 데이터 품질 진단 및 개선

- ‘데이터 품질 진단 및 개선’은 과학기술정보통신부가 생성·보유·활용하고 있는 데이터에 대한 전반적인 진단 결과에 대해 원인분석 및 개선사항을 도출하고, 가능한 경우 실제 개선 활동을 수행하는 것을 의미함

[그림 6-12] 데이터 품질 진단 및 개선 절차(6단계)



[표 6-16] 데이터 품질 진단 및 개선 절차 설명

단계		정의
품질 진단	1.진단 대상 정의 (Define)	품질 이슈에 대한 수요 및 현황을 조사하여 진단 대상을 선정하고, 진단 방향성 정의
	2.품질진단 실시 (Measure)	품질 진단 대상에 대한 상세 수준의 품질 진단 수행 계획 수립 후 영역별 진단 실시
	3.진단 결과 분석 (Analyze)	오류원인 분석, 업무 영향도 분석을 통해 개선과제 정의 (단기과제, 중·장기과제 개선과제 등)
품질 개선	4.개선 계획 수립 (Improvement Plan)	품질 개선 과제별 개선 방향 정의 및 품질 개선 추진을 위한 추진 계획 수립
	5.개선 실시 (Implement)	상세 수준의 품질 개선 수행 계획 및 개선 영역별 품질 개선
	6.품질 통제 (Control)	목표 대비 결과 분석, 평가를 통한 품질관리 목표 재설정 및 지속적 품질 통제 수행

○ 데이터 품질 진단 방법

- 데이터 품질 진단 방법에는 값, 구조를 진단하는 프로파일링, 업무를 중심으로 도출된 품질지표를 활용하여 진단하는 업무규칙 진단 등이 있음

[표 6-17] 데이터 품질 진단 방법

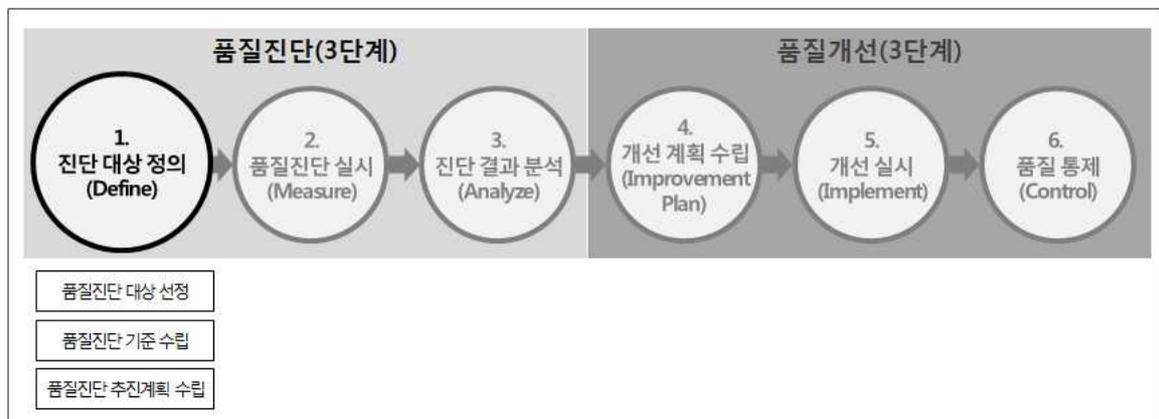
진단방법		내용
프로파일링	값 진단	<ul style="list-style-type: none"> 데이터 값의 유효성, 정확성 등 데이터 값 자체 오류를 분석하는 방법 - 컬럼분석, 날짜분석, 패턴분석, 코드분석 등을 통해 데이터 값의 정확성을 중심으로 진단
	구조 진단	<ul style="list-style-type: none"> 논리적 데이터 구조의 오류로 인한 일관성, 정합성 등을 확보하지 못하는 결함을 분석하고 진단하는 방법 - 표준화 수준(코드, 도메인 등), 테이블 구조, 정규화 수준, 컬럼 및 관계 정의 등 데이터의 구조적 결함 측정
업무규칙 진단		<ul style="list-style-type: none"> 업무기준에 근거하여 데이터가 관리되고 있는지를 진단하는 방법 - 업무규칙(BR, Business Rule)을 준수하고 있는지에 관한 측정 스크립트(SQL 등)를 실행하여 오류 값을 추출

- 값과 구조 진단의 경우, 일반적으로 데이터 프로파일링(Data Profiling) 기법을 활용하여 다양한 종류의 데이터 분석을 위한 패턴들이 정의되어 있어 매우 편리하게 사용할 수 있지만, 패턴이 정의되지 않은 데이터의 경우 프로파일링 도구만으로 모든 대상을 측정하기 힘들며, 이런 경우 업무규칙 진단 기법을 활용함
- 업무규칙은 범정부진단기준을 우선 적용하고 데이터관리문서, 법령, 지침, 업무매뉴얼 등에서 추가 도출할 수 있으며 다수 이해관계자와 협의하여 확정함
- 과학기술정보통신부에서는 데이터 품질관리시스템을 활용하여 값 진단 중심의 프로파일링과 업무규칙** 진단 방법을 주로 사용하고 있음
- **업무규칙 진단 방법에 대해서는 ‘업무규칙관리가이드’ 참고

□ 데이터 품질 진단

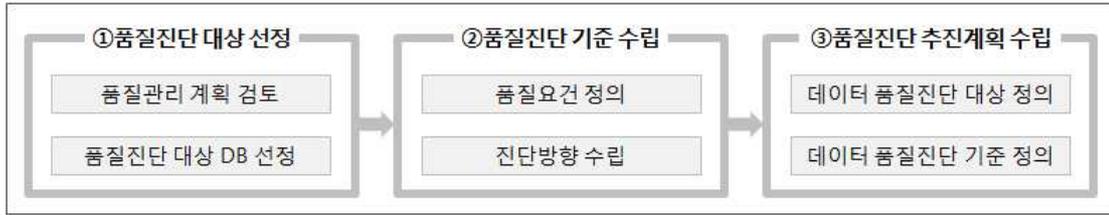
○ 진단 대상 정의

[그림 6-13] 진단 대상 정의(Define) 단계 수행 내용



- 진단대상 정의(Define) 단계는 품질진단 대상 DB를 선정하고, 해당 DB에 요구되는 품질요건과 진단방향 등을 파악하여 추진계획을 수립하는 단계임

[그림 6-14] 진단 대상 정의(Define) 단계 상세 수행 내용



- ① 품질진단 대상 선정
 (품질관리 계획 검토) 계획단계에서 작성한 「데이터 품질관리 계획서」를 검토하여 과학기술정보통신부의 중점 관리 DB, 품질 진단 과제 여부 및 내용을 확인함
 (품질진단 대상 DB 선정) 중점 관리 DB 중 ‘[표 6] 진단 대상 DB 선정 기준’을 활용하여 품질 진단 대상 DB를 선정하며, 진단 대상 DB는 매년 하반기에 실시되는 행정안전부 주관의 ‘공공데이터 품질관리 수준평가’의 대상 DB로도 활용함
- ② 품질진단 기준 수립
 (품질요건 정의) 진단 대상 DB의 관련 업무, DB 종류, DB 버전, 용량, 테이블 수, 컬럼 수 등 기본 현황을 토대로 인터뷰, 설문 등의 방법을 통해 품질 관리수준, 표준화 등의 관점에서 문제가 되는 주요 품질 이슈 목록을 작성함
 (진단 방향 수립) 품질 이슈를 해결할 수 있도록 진단 방향을 수립하며, 일반적으로 진단방향 수립에 사용할 수 있는 전략은 Inside-Out과 Outside-In 방식이 있음
 - Inside-Out 방식 : 과학기술정보통신부가 보유한 DB의 실제 내용, 구조 등을 분석하여 규칙(Rule)에 대한 위반 사례를 발견하여 해결하는 방식으로 전략 수립
 - 업무 관점의 이슈나 현업 담당자의 인터뷰 등을 통해 업무 수행에 핵심이 되는 요건으로부터 도출된 외부 비즈니스, 서비스 품질이슈 등을 해결하는 방식으로 전략 수립

[표 6-18] Inside-Out 방식과 Outside-In 방식 비교

구분	Inside-Out	Outside-In
장점	<ul style="list-style-type: none"> ▪ 도구 활용으로 한 사람의 분석가가 짧은 시간 내에 많은 데이터를 분석할 수 있음 ▪ 잠재적인 품질이슈 파악이 가능함 	<ul style="list-style-type: none"> ▪ 데이터 영역 뿐만 아니라 비즈니스 영역의 문제점까지 파악이 가능함 ▪ 데이터 값이 유효범위 내에 있지만, 부정확한 경우의 이슈를 찾아낼 수 있음
단점	<ul style="list-style-type: none"> ▪ 데이터 구조를 잘 아는 실력 있는 분석가가 필요함 ▪ 데이터 영역에 한정된 문제점 파악만 가능함 	<ul style="list-style-type: none"> ▪ 많은 시간이 소요되고 업무를 잘 아는 이해관계자의 도움이 필요함 ▪ 현재 나타나는 품질이슈만 파악 가능함
예시	<ul style="list-style-type: none"> ▪ 프로파일링과 같은 기술적 점검을 바탕으로 오류의 징후가 있는 데이터 영역으로 점차적으로 확대해 나가는 방법 	<ul style="list-style-type: none"> ▪ 내부 사용자의 오류 접수 혹은 외부로부터의 민원 등과 같이 전체 업무 영역이 아닌 특정 오류의 징후가 있는 영역으로 점차 줄여 나가는 방법

- ③ 품질진단 추진계획 수립
 (데이터 품질진단 대상 정의) 진단 대상 DB의 품질진단 수행에 필요한 대상 테이블 및 컬럼을 식별하고 진단 대상을 정의함
 - 진단 대상은 진단 기준에 해당하는 전체 테이블 및 컬럼을 의미
 - 데이터 품질진단 대상은 「데이터 품질관리 계획서」에 정의 되는 경우가 일반적이며 필요시 별도의 문서로 작성해도 됨

[표 6-16] 품질진단 대상 데이터베이스 선정(예시)

No	시스템 명	DB 명	주요 활용 정보
1	OO민원정보시스템	OO민원DB	(예) 공통코드정보, 고객 마스터 정보 등
2

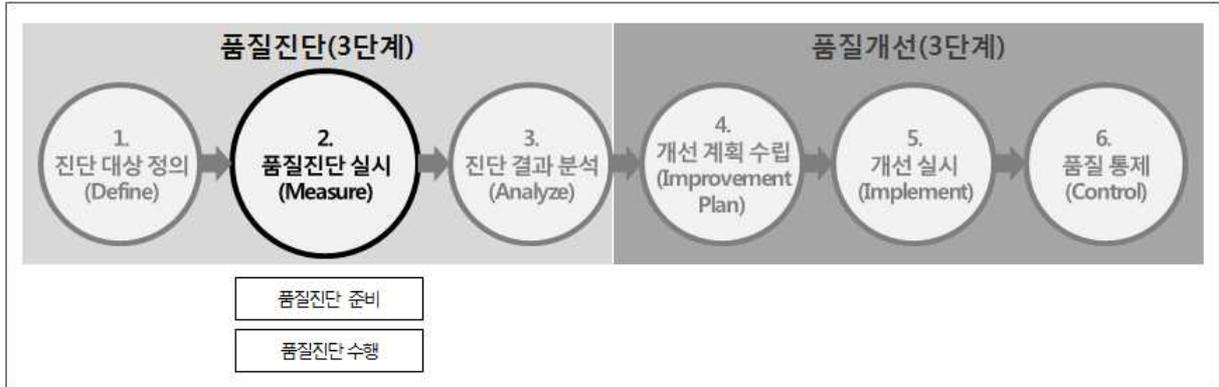
- (데이터 품질진단 기준 정의) DB 단위의 표준 및 구조 산출물 등과 일관성을 가질 수 있도록 하며, 범정부 데이터 품질 진단 기준 등을 참고하여 데이터 품질진단 기준을 정의함
 - 품질진단 기준은 정의된 품질진단 대상에 매핑 되어서 관리 될 수 있도록 함
 - 즉, 데이터 품질진단 대상 컬럼에 매핑된 품질진단 기준에 대한 식별이 가능해야 함
 - 업무규칙정의를 도출된 업무요건을 테이블에 매핑하고 분석영역을 범정부 기준에 부합하도록 분류함
 - 범정부 기준 외에 업무규정과 법령에 근거한 기준은 별도의 업무규칙으로 분류함

[표 6-19] 데이터 품질진단 범정부 진단기준 정의(예시)

진단 기준		진단 대상	
진단 항목	설명	DB	파일
여부	▪ OO여부, OO유무 등 분류 컬럼의 저장된 값이 유효값의 범위를 벗어나는 오류 측정	○	○
수량	▪ 숫자로 저장된 ‘수량’ 컬럼의 값이 유효한 범위를 벗어나는 오류 측정	○	○
금액	▪ 숫자로 저장된 ‘금액’ 컬럼의 값이 유효한 범위를 벗어나는 오류 측정	○	○
율	▪ 숫자로 저장된 ‘율’ 컬럼의 값이 유효한 범위를 벗어나는 오류 측정	○	○
날짜	▪ 날짜 컬럼 중 데이터 값이 유효한 범위를 벗어나는 오류 측정	○	○
코드	▪ 동일한 의미의 코드값으로 일관되게 적용하고 있지 못한 코드 컬럼의 데이터 오류 측정	○	○
번호	▪ 번호 생성 규칙이 존재하는 번호 컬럼의 데이터 값이 정확하게 저장되고 있는지 측정	○	○
선후관계 정확성	▪ 특정 컬럼 간 시차적인 발생 규칙이 존재하는 경우, 이를 위배한 데이터 오류 측정	○	○
시간순서 일관성	▪ 날짜 컬럼 중에서 날짜의 선후관계가 명확히 정의된 컬럼의 데이터 오류 측정	○	○
컬럼 간 논리관계	▪ 컬럼 간 관계에 따른 특정 컬럼의 논리적 일관성 오류 측정	X	X
계산 및 집계	▪ 원천데이터의 계산 등을 통해 저장되는 컬럼 값이 정확하게 관리되고 있는지 측정	○	○
참조값	▪ 참조하는 컬럼과 참조되는 컬럼 사이의 일관성을 위배하는 데이터 오류 측정	○	X

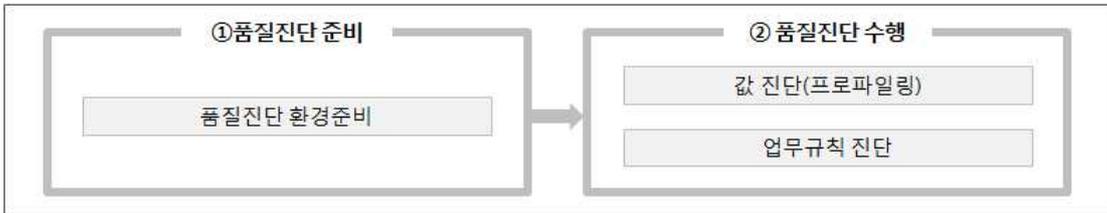
□ 품질 진단 실시

[그림 6-15] 품질진단 실시(Measure) 단계 수행 내용



- 품질진단 실시(Measure) 단계는 진단 대상 데이터베이스의 품질측정을 위한 진단 환경을 준비하고 진단 절차에 따라 데이터 값과 구조의 진단, 업무규칙 진단, 품질 관리 수준 진단을 수행하는 단계임

[그림 6-16] 품질진단 실시(Measure) 단계 상세 수행 내용



- ① 품질진단 준비
(품질진단 환경준비) 품질진단 기준에 따라 진단을 위한 전담 인력 지정, 데이터 품질 전담팀과 업무팀 간 협조체계 구성, DB 접근 권한, 진단 수행 장소 마련 등 데이터 품질진단의 효율적 수행을 위한 진단 환경을 준비함
- ② 품질진단 수행
(값 진단(프로파일링)) 값 진단을 위한 품질 수준의 측정은 프로파일링(Profiling) 기법을 통해서 수행함 - '프로파일링' 이란, 기본적으로 준수되어야 하는 데이터 값의 유효성, 정확성, 표준 적용 여부 등이 도메인별 속성에 위배되어 잠재적으로 업무에 영향을 줄 수 있는 오류데이터를 파악하기 위한 분석기법

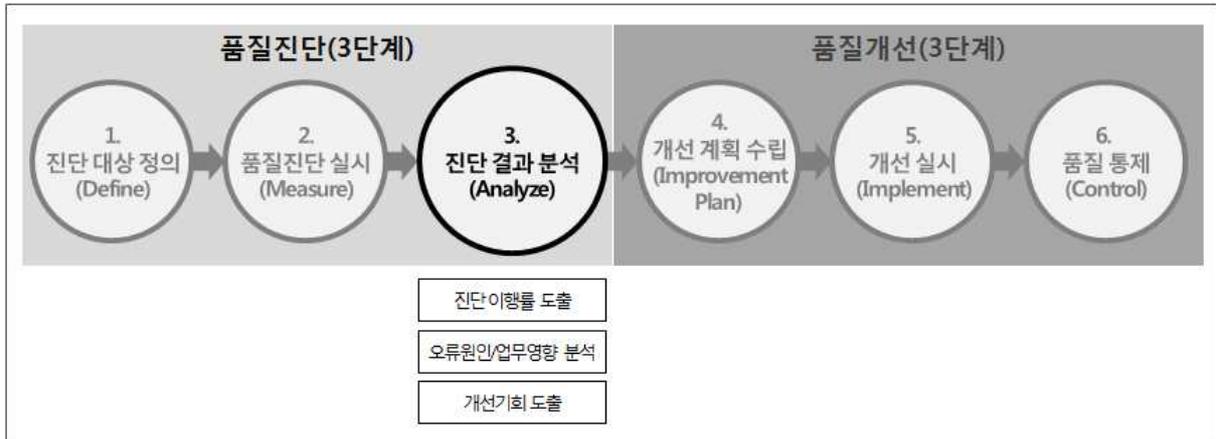
[표 6-20] 데이터 프로파일링 분석 내용(예시)

구분	분석 내용
컬럼 분석 (Column Analysis)	▪ 컬럼에 대한 속성에 대한 준수 여부 검증(Min/Max, Null/Space)
패턴 분석 (Pattern Analysis)	▪ 데이터를 구성하는 값에 대한 패턴을 분석
날짜 분석 (Date Analysis)	▪ 날짜 유형의 데이터로 구성된 값의 유효성을 검사
상관 분석 (Mutual Analysis)	▪ 상호 연결된 컬럼간의 데이터 정합성을 검증
코드 분석 (Code Analysis)	▪ 컬럼 내 코드 값이 정의된 표준에 따라 구성되었는지를 검증

- (업무규칙 진단) 업무규칙(Business Rule)은 기관에서 데이터 품질을 검증하기 위해 관리하는 규칙으로 중요도별, 시스템별, 업무별 등으로 구분된 진단 스크립트를 과학기술정보통신부의 데이터 품질관리시스템에 등록하여 진단을 수행함

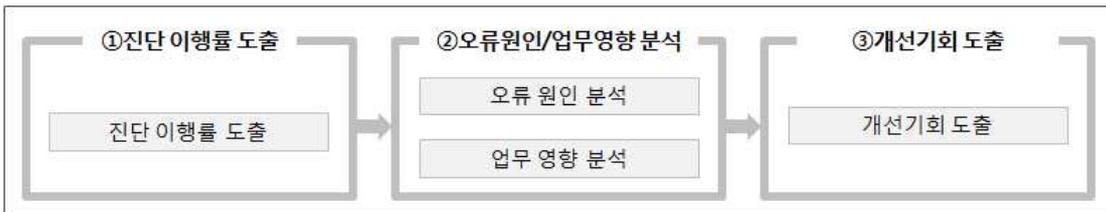
□ 진단 결과 분석

[그림 6-15] 품질진단 분석(Analyze) 단계 수행 내용



- 진단결과 분석(Analyze) 단계는 값 진단, 업무규칙 진단을 통해 도출된 품질 이슈에 대해 유의미한 데이터 품질 문제를 식별하고 문제의 근본 원인을 파악하며 품질 개선 기회를 도출하는 단계임

[그림 6-17] 품질진단 분석(Analyze) 단계 상세 수행 내용



- ① 진단 이행률 도출

(진단 이행률 도출) ‘진단 대상 정의’ 단계에서 정의된 품질진단 대상 및 기준의 대상 테이블 대비 실제 진단을 수행한 테이블 건수의 비율인 ‘진단 이행률’을 도출하여 계획대로 진단을 수행하였는지를 검증

$$\text{진단 이행률(\%)} = \frac{\text{당해 연도 진단수행 테이블 건수}}{\text{당해 연도 진단 대상DB 테이블 건수}} \times 100$$

- ② 오류원인/업무영향 분석

(오류 원인 분석) 오류에 대한 진위 여부를 확인하여 오류로 분류하고, 업무팀 담당자의 의견을 참조하여 오류 원인으로 판단되는 원인을 추적하여 분석

[표 6-21] 데이터 오류 유발 원인(예시)

구분	분석 내용
관리체계 측면의 오류 원인	<ul style="list-style-type: none"> 모델·DB의 관리 프로세스 통제(등록, 수정, 삭제) 미흡으로 인하여 표준, 구조, 응용프로그램 측면의 다양한 오류 데이터를 생성
표준 측면의 오류 원인	<ul style="list-style-type: none"> 기관이 정의한 표준 용어·도메인·코드의 미준수로 인하여 DB별로 표준이 서로 상이하게 정의어 기관의 데이터 일관성을 위배
구조 측면의 오류 원인	<ul style="list-style-type: none"> 모델 정규화 원칙 위배 및 주 식별자 정의 원칙 위배, 참조무결성 위배, 데이터 구조의 중복 정의로 인하여 기관이 목표로 하는 일관성을 위배
응용프로그램 측면의 오류 원인	<ul style="list-style-type: none"> DB 설계 사상이 충분히 응용프로그램DP 반영이 되지 않아 데이터 선·후행관계, 필수입력항목 위배로 인해 데이터의 정합성·유효성이 위배

- ② (업무 영향 분석) 품질 오류가 업무에 미치는 영향을 분석하는 것으로 품질문제로 인해 예상되는 업무적 파급효과를 추정
 - 정성적 업무 영향 분석 : 업무 담당자의 인터뷰 또는 설문을 통하여 품질오류로 인하여 발생이 예상되는 업무 피해 상황, 업무 수행 어려움 등을 분석
 - 정량적 업무 영향 분석 : 품질오류가 업무에 미치는 파급효과를 계량적으로 산출하여 분석하는 기법이며, 주로 경제적 측면이나 업무 효율적 측면으로 파급효과를 도출

[표 6-22] 정량적 업무 영향 분석(예시)

구분	데이터 전체건수	오류율	오류건수	횟수/규모	단가	손실액
고지서 발송 피해	고지서 발송 시행 시 데이터 오류로 인해 고지서가 잘못되어 발생되는 우편 요금					
: 우편 비용	10,715,184건	2.5%	267,880건	고지서 30종 * 30%	220원	약 5.3억
고지서 재발행 피해	시행 후 데이터 오류로 인해 고지서를 재발행할 시 소요되는 인쇄 비용					
: 인쇄 비용	794,123건	41.2%	327,179건	재발행 시 2회 인쇄	100원	약 0.7억
안내문 재발행 피해	시행 후 데이터 오류로 인해 안내문을 재발행할 시 소요되는 인쇄 비용					
: 인쇄 비용	10,715,184건	58.2%	6,236,237건	재발행 시 2회 인쇄	100원	약 12.5억

- ③ 개선기회 도출

[표 6-23] 개선기회(예시)

구분	분석 내용
관리체계 측면의 개선 기회	<ul style="list-style-type: none"> 모델·DB 변경·통제 프로세스 수립, 데이터 표준에 대한 변경·통제 프로세스 및 조직 등
표준 측면의 개선기회	<ul style="list-style-type: none"> 표준 용어 개정 필요, 표준 도메인 추가 제정, 표준 단어 개정, 표준 도메인의 중복 제거 등
구조 측면의 오류 원인	<ul style="list-style-type: none"> “개인” 테이블 정규화 필요, “발급” 테이블 식별자 항목 조정, “민원인·대리인·동반자” 테이블의 데이터 구조 통합 등
응용프로그램 측면의 오류 원인	<ul style="list-style-type: none"> “개인” 테이블의 “주민등록번호” 정제, “분실” 테이블의 “분실일”과 “회수일” 간의 선·후행 관계 정제 등

데이터 품질 개선

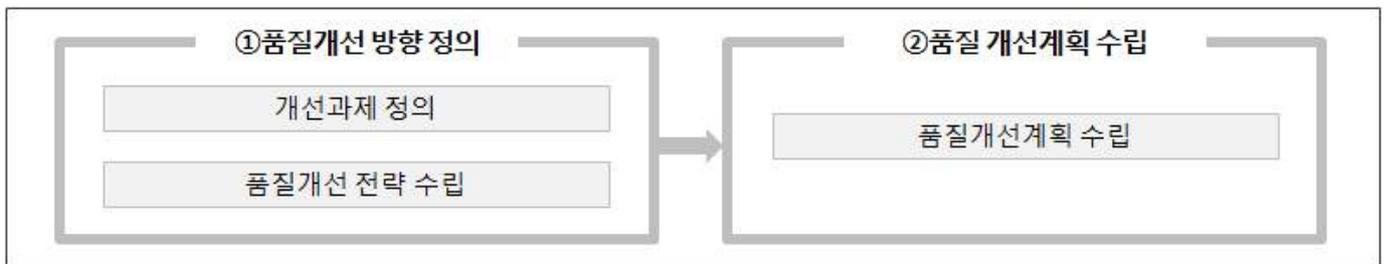
개선 계획 수립

[그림 6-18] 개선계획 수립(Improvement Plan) 단계 수행 내용



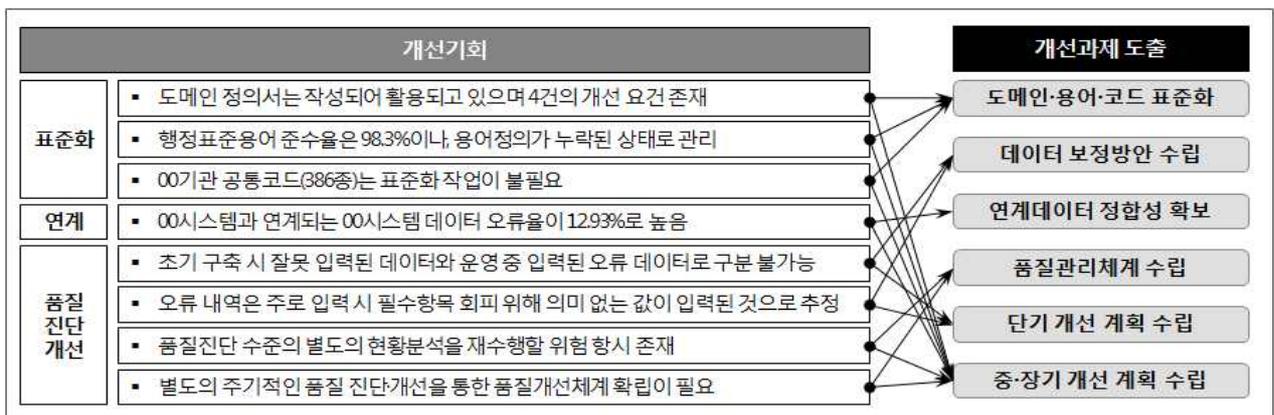
- 개선계획 수립(Improvement Plan) 단계는 ‘진단결과 분석’ 단계에서 도출된 개선기회를 종합하여 개선과제를 도출하고, 품질개선을 위한 개선 전략을 정의하여 상세 추진계획을 수립하는 단계임

[그림 6-19] 개선계획 수립(Improvement Plan) 단계 상세 수행 내용



- ① 품질개선 방향 정의
(개선과제 정의) ‘진단 결과 분석’ 단계에서 개선 기회들 간의 유사성을 고려하여 그룹핑하여 표준, 구조, 조직, 정책 등 품질관리 영역으로 재분류하고, 이로부터 품질개선과제를 정의함

[그림 6-20] 개선계획 도출(예시)



- (품질개선 전략 수립) 개선과제별로 구체적인 개선전략을 수립하며, 개선전략은 기술적 진단 결과를 중심으로 한 데이터 중심의 개선전략과 정보생명주기에 따라 단계별로 사전예방적 품질관리를 수행하는 프로세스 중심의 개선 전략이 있음
 - 데이터 중심의 개선(Reactive) : 데이터 값을 직접 수정하여 데이터 품질을 개선
 - 프로세스 중심의 개선(Proactive) : 저품질 데이터 생산의 원인 제거를 위해 데이터 생성 및 수정 프로세스 등을 재설계하여 데이터 품질을 개선

- ② 품질개선계획 수립

(품질개선계획 수립) 데이터 품질진단에 따른 오류의 원인을 분석하고, 오류 정제를 위해 개선 목표, 개선과제, 수행 조직 및 역할, 일정, 예산 등이 포함된 품질개선계획을 수립

- 개선계획 내 개선과제는 당해 연도 조치 가능한 개선과제(단기)와 차년도 이후 조치 가능한 개선과제(중장기)로 구분하여 정의해야 함
- 단기 개선과제에는 실제 오류데이터 정비를 통한 오류 개선율(%) 제고 활동이 포함되어야 함
- 개선계획에는 오류데이터에 대한 실제 오류개선을 제고하기 위한 노력(활동)이 반영되어야 하며, 반드시 범정부진단기준을 활용한 오류진단결과에 따른 개선활동을 수립하여야함

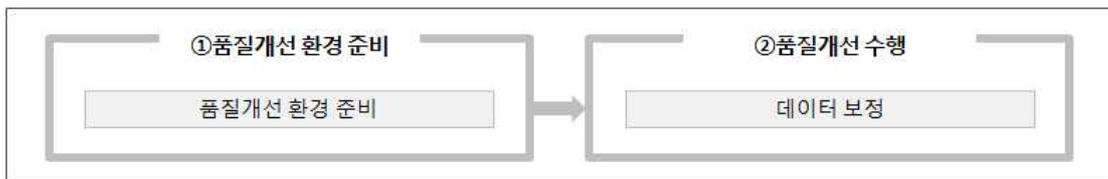
□ 개선 실시

[그림 6-21] 개선 실시(Implement) 단계 수행 내용



- 개선 실시(Implement) 단계는 수립된 품질개선 추진계획 중 단기 과제를 중심으로, 사업수행을 위한 환경을 준비하고, 실제 개선을 수행하는 단계임

[그림 6-22] 개선 실시(Implement) 단계 상세 수행 내용



- ① 품질개선 환경 준비

(품질개선 환경 준비) 품질개선을 효과적으로 수행하기 위해 과학기술정보통신부에서 사전에 준비해야 할 단계로 정의된 과업범위를 확정하고, 확정된 과업 범위를 고려하여 수행일정을 상세하게 수립하는 단계

- 사업수행을 위한 추진반 및 협의체 구성, 보정 대상 데이터 확보, 시스템 접근권한 및 개선 수행 장소 확보 등의 사전 환경 준비가 필요하며 ‘품질진단 환경준비’와 유사함

- ② 품질개선 수행

(데이터 보정) 오류 데이터의 값을 보정하는 것 외에도 오류 데이터를 발생 시키는 DB 구조의 보정, 오류 데이터 유입 방지를 위한 응용 프로그램의 수정 등을 포함

- 데이터 보정을 위해서는 원인과 오류 현상을 정확히 알고, 그에 맞는 보정 방식을 정하여야 하는데, 다양한 방법이 있음

[표 6-24] 데이터 보정 방법(예시)

구분	보정 내용
수작업 보정	<ul style="list-style-type: none"> 오타, 필수 데이터 누락 등에 대해 올바른 값을 아는 업무팀 혹은 현업 담당자가 직접 보정하는 것 대상이 많을 때는 외주 용역을 통해 보정하거나, 업무 담당자에게 보정 대상과 보정 방법을 안내하고 기간을 정하여 보정 할 수 있음
Script 방식 보정	<ul style="list-style-type: none"> 데이터 오류가 일정한 방식에 따라 동일하게 일어난 경우, 이를 자동으로 보정해줄 Script를 작성해서 보정하는 것 예를 들어, 날짜 표시가 YY-MM-DD 방식이어야 하는데, YYYY-MM-DD로 입력된 것이 있다든지, 전화번호에 '-' 표시가 들어간 것을 제거하는 등
원천 데이터 재입력 보정	<ul style="list-style-type: none"> 대량의 데이터를 일괄 입력 받을 때, 잘못된 데이터가 함께 유입되거나, 입력 과정에서 손상된 경우, 해당 부분(Block)만 다시 원천 데이터를 찾아 재입력하는 것

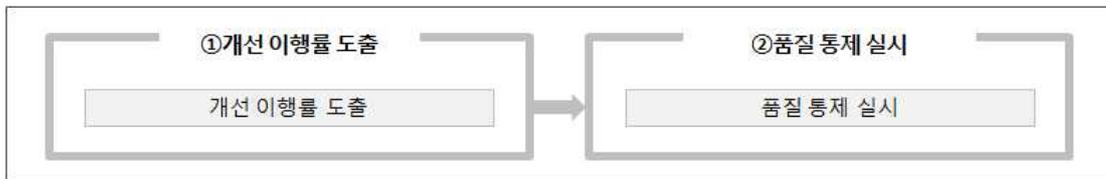
□ 품질 통제

[그림 6-23] 품질 통제(Control) 단계 수행 내용



- 품질 통제(Control) 단계는 개선 사업의 효과에 대한 결과평가를 하고, 지속적인 데이터 품질을 유지하고 더 높은 품질을 확보하기 위한 통제활동을 수행하는 단계임

[그림 6-24] 품질 통제(Control) 단계 상세 수행 내용



- ① 개선 이행률 도출

(개선 이행률 도출) 개선계획에서 수립된 당해 연도 과제 건수 대비 실제 개선을 수행한 과제 건수인 ‘개선 이행률’을 활용하여 수립된 개선계획에 따라 조치 가능한 개선 활동을 수행하고 있음을 검증

$$\text{개선 이행률(\%)} = \frac{\text{당해 연도 개선수행 과제 건수}}{\text{당해 연도 개선계획 과제 건수}} \times 100$$

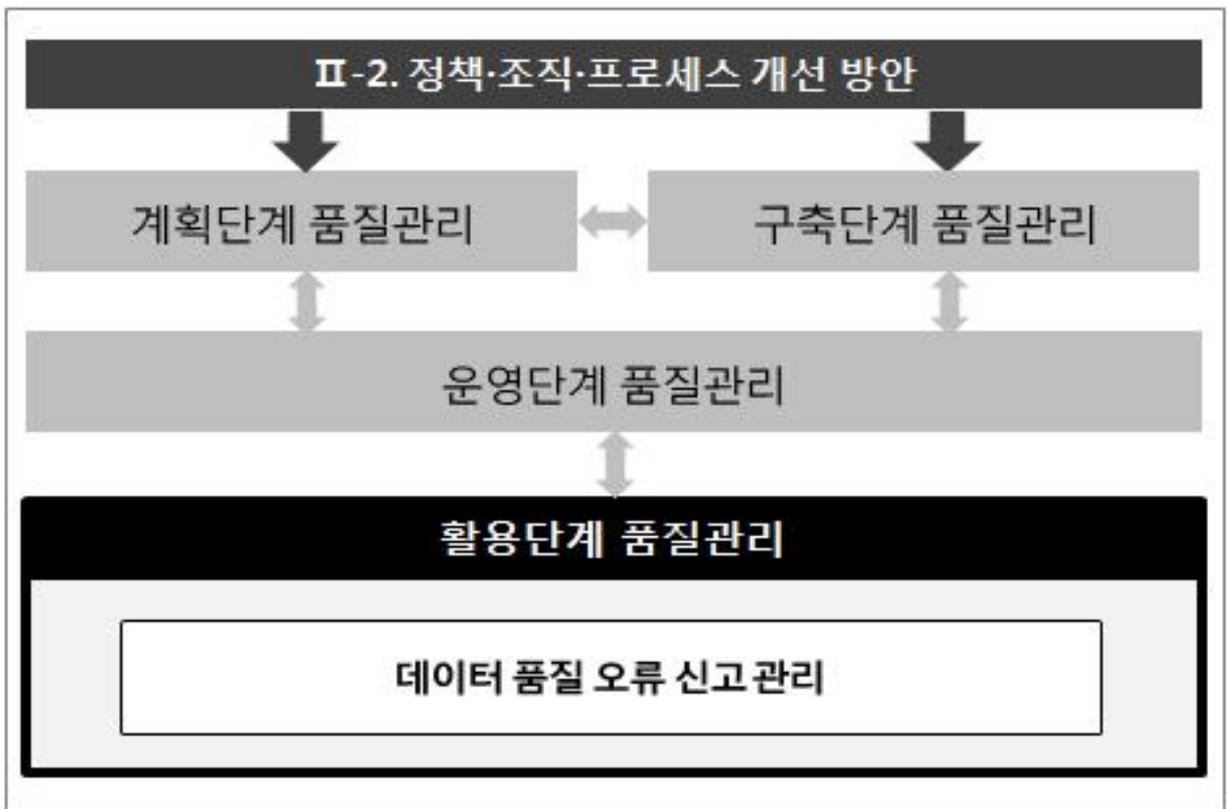
- ② 품질 통제 실시

(품질 통제 실시) 품질 개선이 완료된 대상에 대한 지속적인 모니터링 활동 뿐만 아니라 구성원들에 대한 품질 교육, 내·외부 사용자로부터의 품질오류 신고 접수 및 처리에 관한 활동 등을 포함

□ 활용단계 품질관리 방안

- 데이터를 활용함에 있어 발생하는 품질 이슈를 인지하여 개선 활동으로 연계하는 품질 오류 신고 관리와 수요자 시고 접수 내역 및 처리 결과를 분석하여 문제점을 발굴하고 개선하는 활동 등을 의미함

[그림 6-25] 활용단계 품질관리 주요 활동 내역



- 주요 산출물은 수요자 요구사항 기록 및 처리결과 내역, 이용실태 조사 결과서 등이 있음

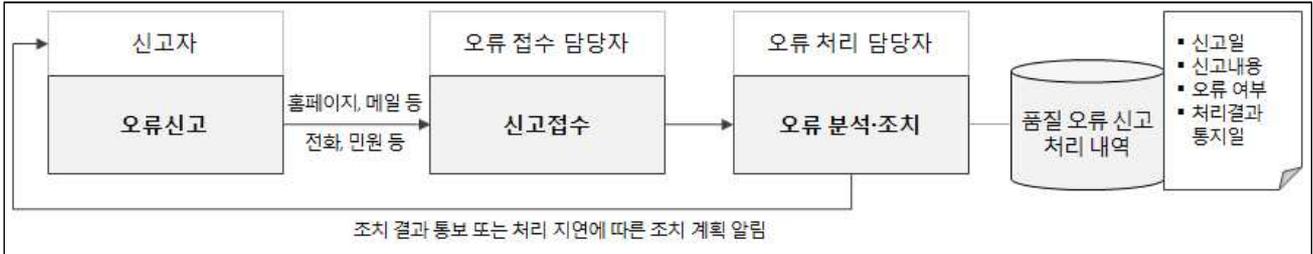
[표 6-25] 활용단계 품질관리 주요 산출물

산출물 명	비고
▪ 수요자 요구사항 처리 절차 정의	
▪ 수요자 요구사항 기록 및 처리결과 내역	
▪ 수요자 요구사항 분석 및 개선조치 활동 실적	
▪ 이용실태조사 결과서 (만족도, 활용성 높은 데이터 수요조사 등)	최근 2년간
▪ 수요자 요구사항 분석의 정책반영 내역서	1회/연 이상

○ 데이터 품질 오류 신고 절차

- 데이터 품질 오류 신고는 기본적으로 오류 발생 부분·시각·내역 등에 대해 이루어지며, 접수된 오류신고 내역에 대한 평가와 개선이 필요할 경우 개선 후 그 결과를 신고자에게 통보함

[그림 6-26] 품질 오류 신고 처리 절차

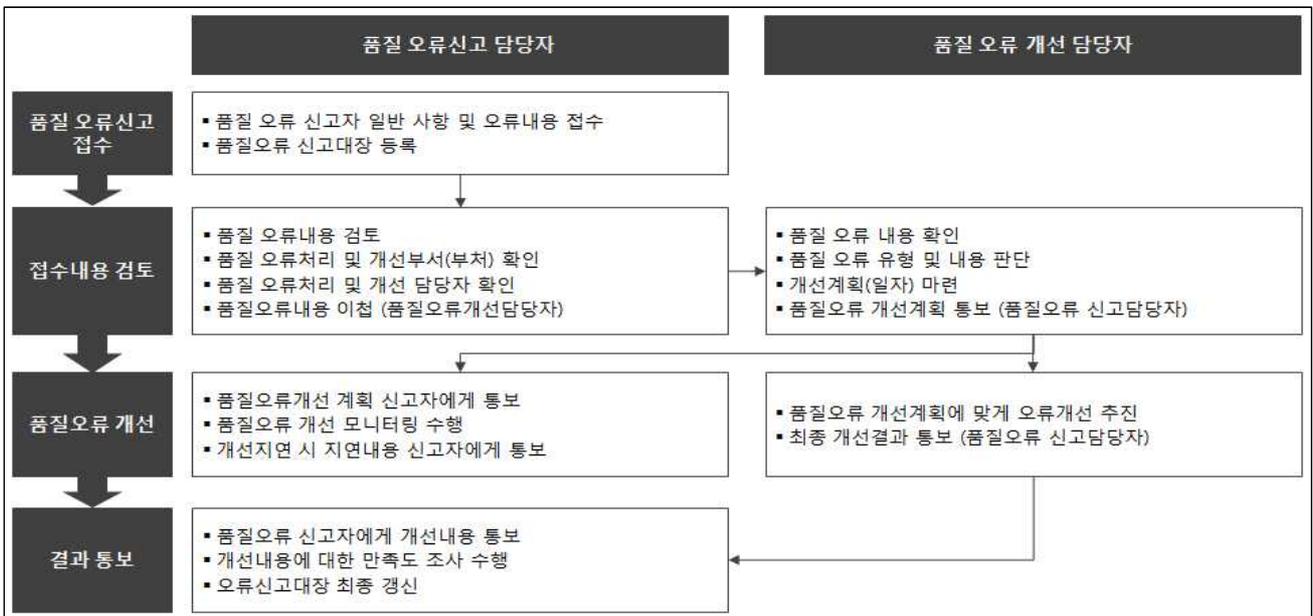


- ① 사용자가 오류를 발견하면, 발견된 오류에 대한 정보를 각종 채널을 통해 신고 및 접수
- ② 데이터 품질 오류 관리 담당자는 접수된 오류신고 내용을 파악하고, 해당 담당자에게 이관
- ③ 해당 담당자는 오류내역에 대해 상세 분석하고, 오류를 해결하기 위한 방안마련 및 오류개선 등 일련의 조치 활동 수행
- ④ 해당 담당자가 제공한 오류 개선 일정 및 처리 결과를 신고자에게 통지하고, 오류 처리 결과를 품질오류신고 처리대장에 기록
- ⑤ 만약, 단 기간 내에 오류를 처리할 수 없을 경우에는 오류처리 계획을 신고자에게 통지하고, 해당 기간 동안 오류가 처리되었는지 등에 대해 지속적으로 모니터링을 실시

○ 데이터 품질 오류 신고 관리 담당자별 역할

- 제공하는 다양한 정보에 대한 품질오류를 신고할 수 있는 체계를 마련하기 위해서는 기본적으로 품질오류에 대한 신고접수 및 처리를 전담할 수 있는 담당자 지정이 필요함

[그림 6-27] 품질오류 신고관리 담당자별 역할



2. 업무규칙 가이드 수립

□ 업무규칙 개요

○ 업무규칙 정의

- 업무규칙(Business Rule)은 기관에서 데이터 품질을 검증하기 위해 관리하는 규칙으로서 데이터의 값이 정확한지 여부를 판단하기 위한 조건에 대한 표현이다.
- 업무규칙은 법령·규칙 등에 따라 데이터의 정확성 확보를 위해 만족해야 하는 기준으로써 업무 데이터의 정합성을 검증하는 식이며, 이 방식은 데이터 품질 관리의 기본적인 방법론이다.

○ 업무규칙 작성 원칙

- (원칙 1) 업무규칙은 물리적인 진단이 가능한 테이블, 컬럼 단위로 작성한다.
 - 동일 항목 명, 항목ID 이라도 다른 테이블에 속해 있으면 업무규칙 내용이 다를 수 있으므로 반드시 테이블 별로 업무 기준의 차이점을 검토해보고 결정해야 한다.
 - 만약 타 테이블의 동일 항목에 대한 업무규칙이 도출되어 있고 그 내용이 동일하거나 비슷할 경우에는 해당 내용을 참조하여 작성한다.
- (원칙 2) 업무규칙은 가능하면 가장 작은 단위로 원자화하여 작성한다.
 - 업무규칙의 대상 컬럼에 여러 종류의 업무규칙이 존재할 경우, 향후 업무규칙분석 및 정비 등의 작업을 위하여 업무규칙을 원자화하여야 한다.
 - 예) ○○테이블의 개인정보공개여부는 Y 또는 N의 유효한 값을 가져야 한다.
 - ① ○○테이블의 개인정보공개여부는 반드시 존재해야 한다.
 - ② ○○테이블의 개인정보공개여부가 존재할 경우 값은 ‘Y’ 또는 ‘N’ 으로 구성되어야 한다.
- (원칙 3) 업무규칙은 가능하면 모든 품질지표를 대상으로 모든 유형의 업무규칙을 도출한다.
- (원칙 4) 업무규칙은 반드시 누구나 알기 쉬운 자연어와 (관련)참조 IT 용어로 작성하여야 하며 업무규칙을 분류할 수 있는 분류 및 명 체계를 관리하여 업무규칙을 쉽게 검색하거나 식별할 수 있도록 해야 한다.
 - 업무규칙을 (전문적) IT용어가 아닌 업무적 (자연어)서술형식으로 작성하여 현업부터 IT담당자까지 동일한 기준으로 의사소통이 가능해야 한다.
 - 전사차원의 동일한 업무규칙의 명명규칙에 따라 작성함으로 검색 및 식별이 용이 하여야 하며 해당 업무규칙 식별자로 담당자와의 의사소통이 가능해야 한다.
- (원칙 5) 업무규칙은 데이터 품질 점검을 위한, 즉 “Data Error Check“를 주 대상으로 한다.
 - 업무규칙은 부정(Negative)에서 출발한다. 즉, 품질의 관점에서는 ‘True’ 가 아니라 ‘Not True’ 에 관심을 가진다.

- 데이터의 발생 (통계)추이에 대한 업무규칙은 'Not True' 가 아닌 전체 데이터이므로 반드시 따로 분류하여 관리하여야 한다.
- (원칙 6) 컬럼이 코드화 되지 않았을 경우나 코드 정의 값이 명확하지 않은 경우 코드 값을 직접 풀어서 기술하고 추후 값을 코드테이블에 등록하도록 한다. 예) 사용계좌유형코드는 {'01','02','03','04'} 중에 하나여야 한다.
- (원칙 7) 2개 이상 테이블 간의 업무규칙 정의 시 필히 관련 테이블과 컬럼을 함께 기술한다.
 - 2개 이상의 테이블간의 업무규칙 작성 시 관련된 모든 테이블과 컬럼을 기술하여 업무규칙 설명이 명확하도록 해야한다.
 - 예) 기술기업신청[TB_ABC]에서 사업장규모코드[BUSI_SCALE_CD='0']가 5인 미만이면 사업장규모서브코드[BUSI_SCALE_SUB_CD] 값은 반드시 존재해야한다.
- (원칙 8) 진단 SQL은 품질점검 대상 시스템의 테이블을 기준으로 작성하여야 한다.
 - 품질 점검 대상 DB(Oracle, SQL Server 등)의 특징에 맞게 진단 스크립트를 작성한다.
 - 품질 점검 대상 데이터의 범위는 전체기간으로 하되 업무특성을 고려하여 필요한 경우(데이터 유효성, 복잡도, 진단속도, 데이터 보정 가능 여부 등) 범위를 조정하도록 한다.
 - 업무규칙 작성 시 변환, 적재에 대한 정보는 데이터담당자와 협의 또는 메타시스템의 관련 정보를 참조한다.
- (원칙 9) 금액차이에 대한 검증 업무규칙일 경우 반드시 Source 금액, Target 금액, 차이 금액을 모두 기술하고 금액의 기준이 되는 정보를 기술한다.
 - 금액, 건수차이를 검증하는 업무규칙의 경우 전체금액이 분모로 관리되며 차이금액이 오류 건의 분자로 관리된다.
- (원칙 10) 진단 대상의 총건수를 측정할 수 있는 총건수 SQL을 기술한다.
 - 업무규칙 작성 시 오류추출 진단 SQL 외에 대상이 되는 분모군에 대한 총 건수 SQL도 같이 기술한다.
 - 총 건수 SQL가 테이블 전체 건일 경우 생략이 가능하며 특정 조건을 만족하는 건만을 대상으로 할 경우 사용한다.
 - 예) 업무규칙 : 개인고객 테이블의 직장주소 칼럼의 데이터 값은, 종업원 칼럼의 데이터 값이 종업원('Y')이면, 반드시 존재해야 한다.

○ 업무규칙 도출근거 구분

- 업무규칙(Business Rule)은 도출 근거에 따라 (단순)기술적, (복잡)업무적 업무규칙으로 구분된다.
- 도출 근거가 중복된다면 상위 근거(규정/법령>일반>기술)에 해당하는 종류로 구분한다.

[표 6-26] 업무규칙 근거구분

구분		도출근거	예시
(단순)기술적		<ul style="list-style-type: none"> 품질표준(데이터모델링 상의 표준)을 근거로 도출한 BR 물리DB관점에서 컬럼 값의 단순 검증 시 적용 프로파일링에 의한 데이터 검증식이 적용됨 	수급자격인정의 수급자격신청일자 값은 'YYYYMMDD' 형식이어야 한다.
(복잡)업무적	지침 및 매뉴얼	<ul style="list-style-type: none"> 법령, 지침 또는 품질관리 매뉴얼을 근거로 도출한 BR 	과학기술정보통신부 데이터 품질관리 지침 및 데이터 품질관리체계정의서

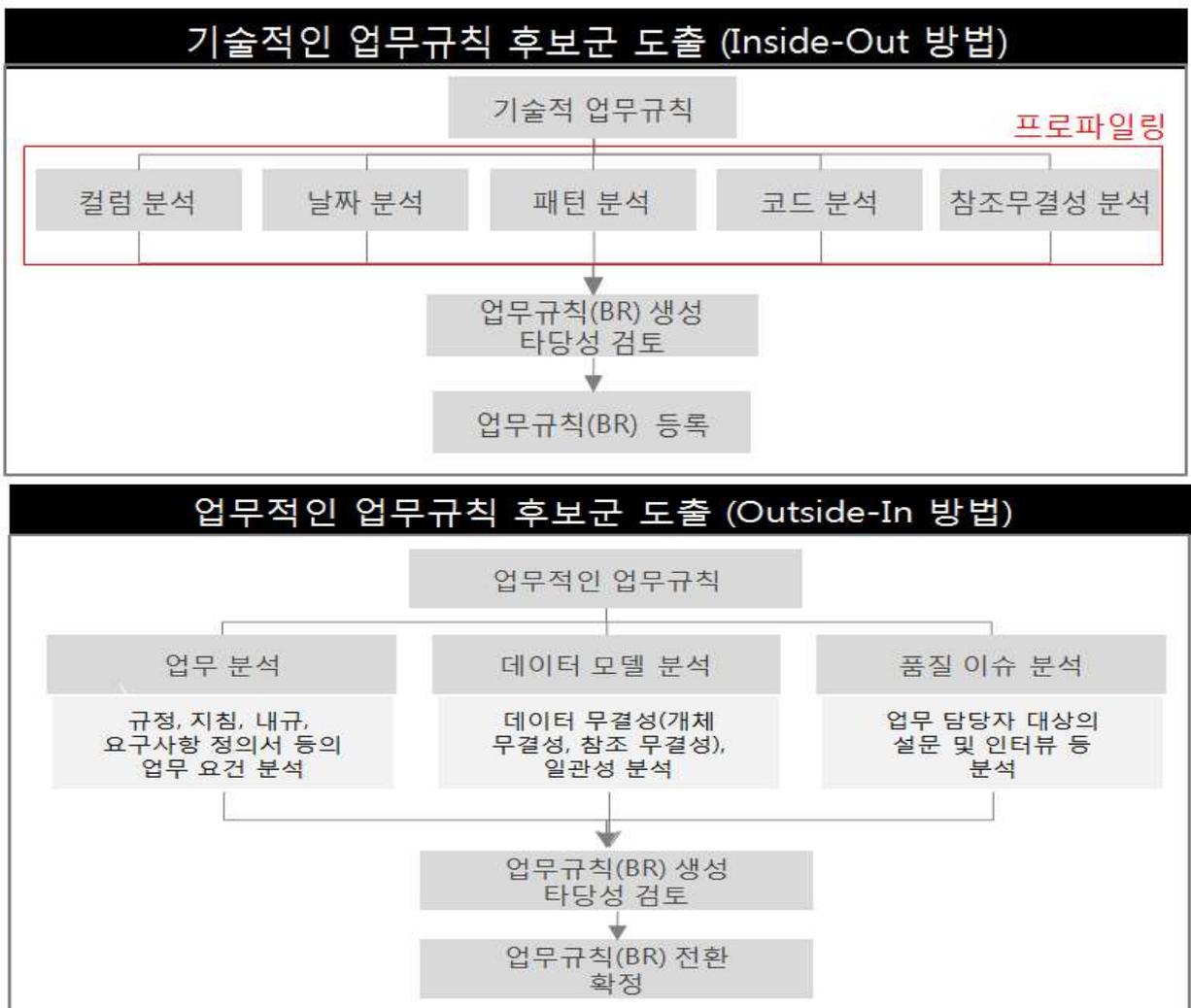
- 과학기술정보통신부의 업무규칙 근거규정

- 과학기술정보통신부 「데이터품질관리지침」 제3조 (적용범위) 및 「데이터품질관리체계정의서」 제3항 (데이터품질 진단 및 개선)에 근거하여 도출함

○ 업무규칙 도출방법 개요

- 업무규칙은 도출근거에 따라서 Inside-Out 방법론(프로파일링 분석) 및 Outside-In 방법론(업무규칙 분석)을 통해 도출된다.

[그림 6-28] 업무규칙 후보군 도출 방법



○ 업무규칙 품질지표

- 업무규칙 품질지표는 데이터의 품질수준을 측정하기 위한 기준이 되는 지표이다.

[표 6-27] 업무규칙 품질지표

구분	품질지표	코드	품질지표 설명
완전성	단독완전성 (SINGLE)	SG	값이 항상 채워져 있어야 함
	조건완전성 (CONDITION)	CN	값이 다른 데이터 항목의 조건에 따라 항상 채워져 있어야 함
유효성	범위유효성 (RANGE)	RG	Column Value List 내에 존재하여야 하거나, 정해진 범위내에 존재하여야 함
	날짜유효성 (DATE)	DT	Column 정의가 일시일 때 값은 유효한 일시의 값을 가져야 함
	포맷유효성 (FORMAT)	FM	정해진 형식(format)에 맞게 값이 존재하여야 함
	코드유효성 (CODE)	CD	Column Value가 개별코드 또는 통합코드 내에 정의된 값을 가져야 함
유일성	유일성 (UNIQUE)	UQ	PK, UK 또는 업무적으로 유일해야 하는 데이터는 1건만 존재하여야 함
무결성	참조무결성 (REFERENCE)	RF	참조되는 속성의 데이터는 해당 속성을 참조하는 속성의 데이터와 일치해야한다(공공데이터품질관리매뉴얼)
정합성	계산/집계정합성 (CALCULATION/CO UNT)	CC	한 컬럼의 값은 다수 컬럼의 계산된 값일 경우 계산 값이 정확해야 한다(데이터품질진단 절차 및 기법)
일관성	컬럼간일관성 (COLUMN)	CL	동일 테이블 또는 2개 이상의 테이블 간의 Column 값이 서로 모순되지 않아야 함
	ROWS간일관성 (ROWS)	RO	동일 테이블 또는 2개 이상의 테이블 간의 Row간 값이 서로 모순되지 않아야 함
	시간순서일관성 (TIME_SEQUENCE)	TS	복수의 컬럼값이 선후 관계에 있을 경우 이 규칙을 지켜야 한다(데이터품질진단 절차 및 기법)
최신성	최신성 (NEWEST)	NW	정보의 발생, 수집, 그리고 갱신 주기를 유지해야 한다.(데이터품질진단 절차 및 기법)

○ 업무규칙 레벨

- 업무규칙은 검증하고자 하는 진단 관점에 따라 레벨을 정의한다.
- 1, 2레벨은 (단순 기술적)IT관점의 정합성을, 3레벨 이상은 (복잡)업무 및 법령/규정의 관점에서 정합성을 점검하여 지속적으로 확보하고 중점적으로 관리해야 하는 대상이다.
- 주요지표를 기준으로 하되 전사품질관리자가 업무상 중요도를 판단하여 업무 규칙(전사공통코드 테이블을 참조하는 코드유효성 등)의 레벨을 조정할 수 있다.

[표 6-28] 업무규칙 품질레벨 정의서

구분	레벨	진단 관점	설명	주요지표	도출 시점	예시
(하위) (단순) 기술적	1	데이터 값 자체	데이터 값의 유효성, 정확성 등 값 자체 오류를 분석	날짜유효성 범위유효성 포맷유효성	시스템개발	수급자격인정의 수급자격신청일자의 값은 'YYYYMMDD' 형식이어야 한다.
	2	컬럼간 논리적 관계	테이블별 컬럼간 논리적인 참조무결성 관계를 확인(물리적 RI관계 제외)	참조무결성 코드유효성		훈련_과정회차별훈련생기 본의 훈련과정ID의 값은 훈련_과정공통기본의 훈련과정ID에 존재해야 한다.
(상위) (복잡) 업무적	3	업무적 관점	행정업무 중 민원이나 이슈가 발생하였거나 데이터 오류 발생 가능성이 높은 업무 등을 선별	유일성 단독완전성 조건완전성 시간순서일관성	시스템개발 시스템운영	구인자기본의 상시근로자수의 값은 반드시 존재해야 한다.
	4			계산/집계정합성 컬럼간일관성 ROWS간일관성 최신성 (기타 품질관리자 지정 BR)		구인신청의 접수마감일자 항목 입력 또는 채용시까지 구인여부 항목이 체크되어야한다.
(법률적)	5	법령/규정 관점	관련 법령/규정 등에 근거하여 선별			

○ 데이터 오너십

- 데이터 오너십이란 데이터 값에 대한 권한과 책임을 가지는 주체로서 데이터 오류를 정비하고 개선하는 데에 주도적인 역할을 수행할 수 있는 담당자 또는 담당 부서를 의미한다.

- 데이터 오너십 기본원칙

- (원칙 1) 데이터 오너십은 전사품질조직이 아닌 데이터를 취급하는 업무조직에 존재해야 한다.
- (원칙 2) 데이터 품질 이슈에 대한 의사결정권이 있는 부서에 우선으로 부여한다.
- (원칙 3) 세부적이고 구체적인 상세단위(시스템<엔티티<속성)의 오너십을 더 우선시 한다.
- (원칙 4) 정보 흐름에 일관성 있게 오너십을 부여한다.
- (원칙 5) 정보의 정합성 유지를 위해 데이터를 정의하는 부서는 유일해야 한다.

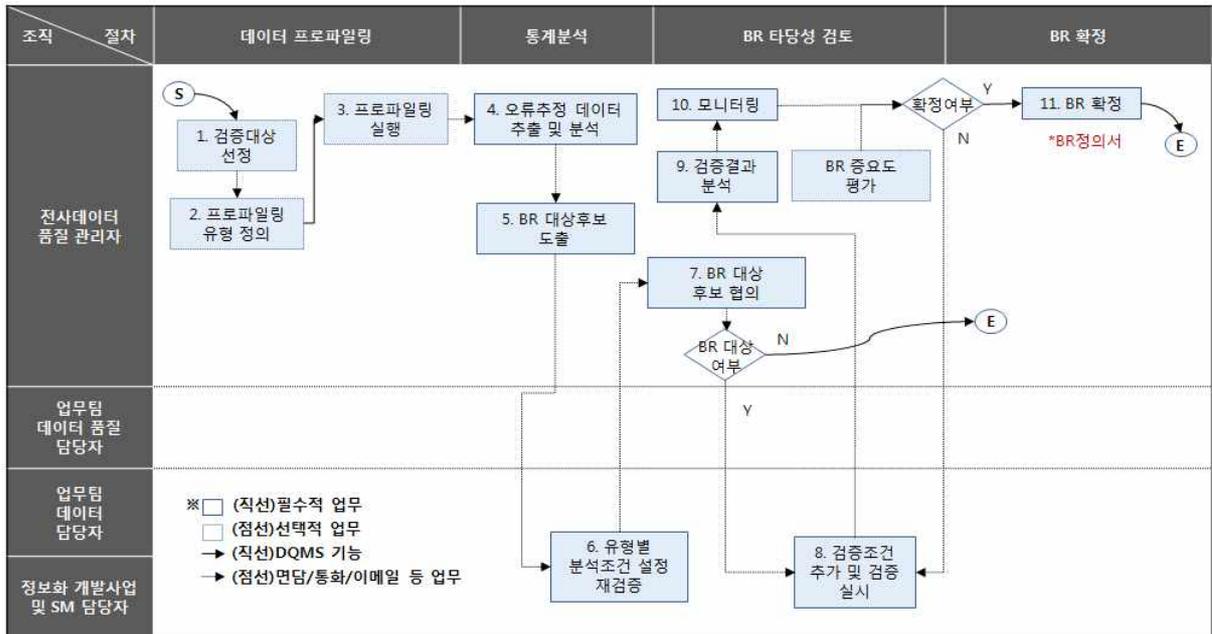
- 데이터 오너십 분류기준

- 품질진단시스템(DQMS)을 통해 관리되는 업무규칙 전체에 대해 데이터품질담당자 분석의견 및 DB모델상의 근거를 토대로 데이터 담당부서를 배정한다.



- 운영 중인 데이터베이스의 (선택)데이터 프로파일링을 실시한 후 본 가이드의 <업무규칙 대상 항목 선정>을 기준으로 오류 징후가 파악되거나 오류 위험이 높은 컬럼에 대하여 Inside-Out방법으로 품질이슈를 파악하기 위한 도출방식이다.

[그림 6-30] Inside-Out방법 품질이슈를 파악 도출방식



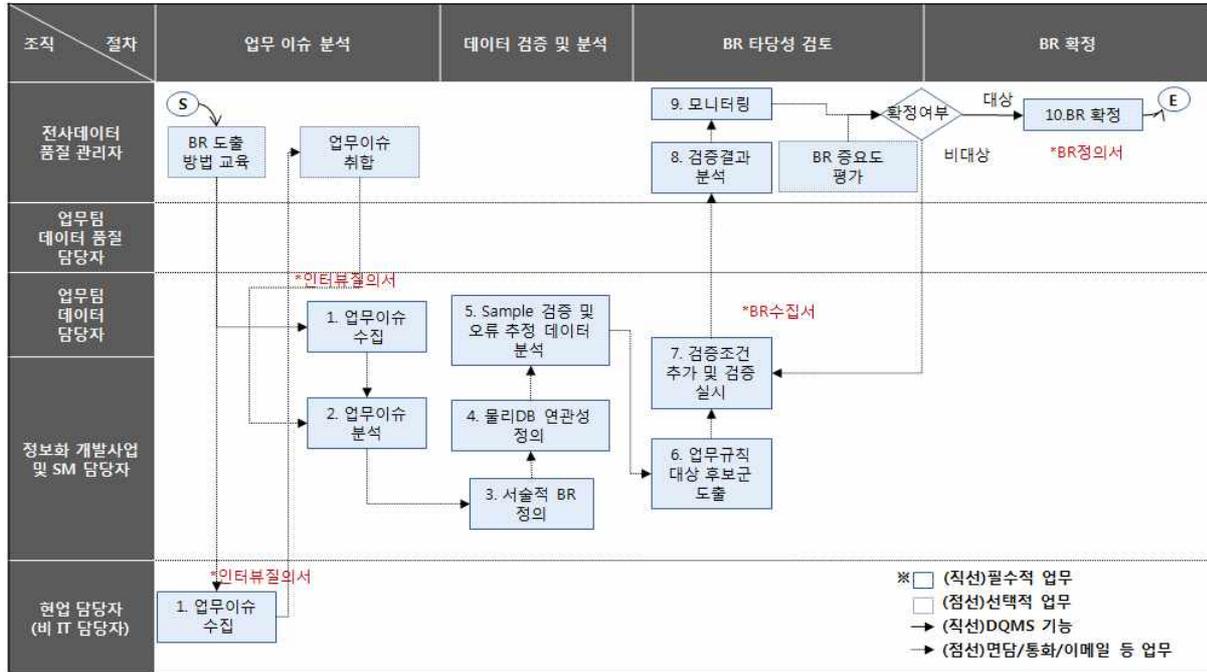
[표 6-30] Inside-Out 업무규칙 도출 순서

순서	절차	내용
1	검증대상 선정(선택)	<ul style="list-style-type: none"> 검증범위, 검증계획, 목표 기존 컬럼은 업무분류 체계에 따라 물리DB 검증대상 선정
2	프로파일링 유형 정의(선택)	<ul style="list-style-type: none"> 검증내용에 따라 컬럼, 패턴, 날짜유형, 코드, 참조무결성
3	프로파일링 실행(선택)	<ul style="list-style-type: none"> 유형별 계획 수립 프로파일링 실행 -> 프로파일링실행 작업명, 작업주기, 작업대상(시스템) 확정
4	오류추정 데이터 추출 및 분석	<ul style="list-style-type: none"> 오류추정 데이터 추출 통계적 기법을 이용하여 유형에 맞는 데이터 분석 재검증실시, 분석
5	BR대상 후보 도출	<ul style="list-style-type: none"> -> 기술적, 업무적 관점의 중요도 평가 -> 오류 데이터 및 오류 징후 데이터에 대해 BR대상 도출
6	유형별 분석조건 설정 및 재검증	<ul style="list-style-type: none"> 업무적, 법적 요건 및 요구사항 고려 - 추가분석조건 설정, 재검증 실시 -> 진단 SQL 작성
7	BR대상 후보 협의	<ul style="list-style-type: none"> 업무규칙 후보의 업무적 중요도를 협의. 도출된 업무규칙 후보의 업무간, 시스템간 연관관계 및 영향도를 파악
8	검증조건 추가 검증실시	<ul style="list-style-type: none"> -> 업무적 요건사항 고려하여 검증조건 추가 -> 진단 SQL에 대한 검증실시
9	검증결과 분석	<ul style="list-style-type: none"> 감시 필요성 있는 데이터를 BR대상으로 선정
10	모니터링	<ul style="list-style-type: none"> 선정 BR대상 감시 필요성 여부에 따라 주기 지정->모니터링 ->오류데이터 추이파악
11	BR 확정	<ul style="list-style-type: none"> 업무규칙정의서를 작성하고 DQMS에 등록한다.

- 업무적 접근에 의한 업무규칙 도출

- 업무이슈 분석을 위한 분석관점을 정의하고 현업의 요구(CSR), 업무 개발/운영 시 자체진단, 인터뷰 및 설문 등 이슈 사항을 중심으로 자료를 Outside-In방법으로 수집하는 방식이다.

[그림 6-31] Outside-In방법 품질이슈를 파악 도출방식 (업무 이슈 근거)



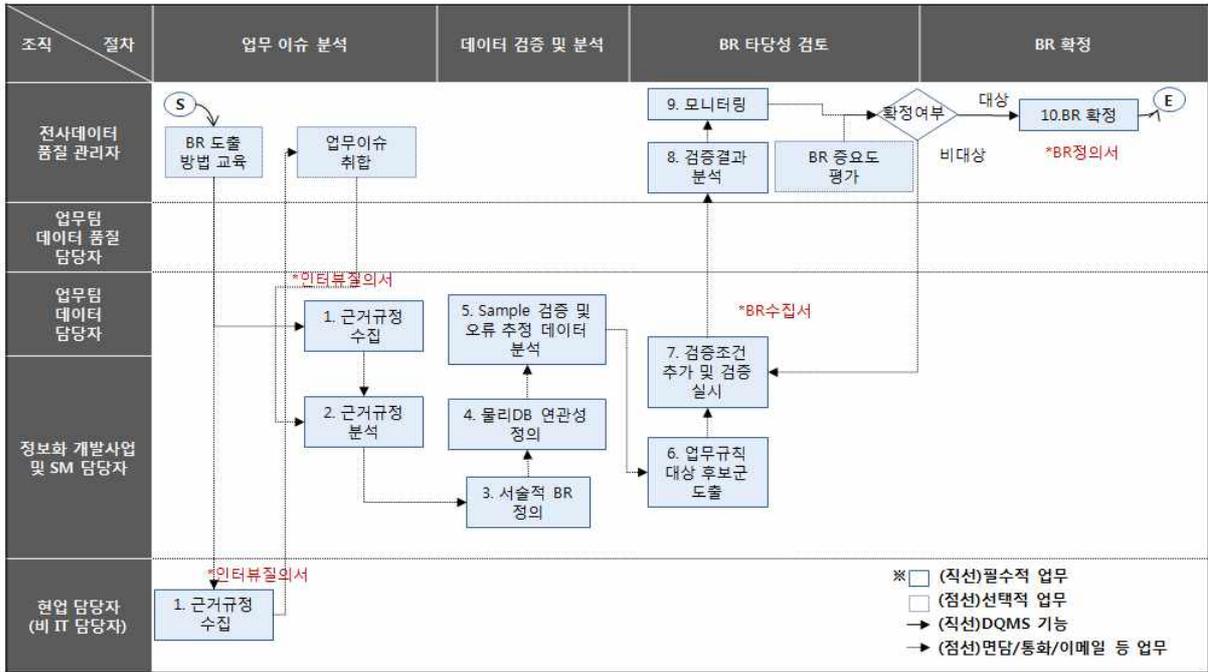
[표 6-31] Outside-In 업무규칙 도출 순서 (업무 이슈 근거)

순서	절차	내용
1	업무이슈 수집	• 현업의 요구(CSR), 업무 개발시 자체진단, 인터뷰 및 설문 등 이슈 사항을 중심으로 자료 수집
2	업무이슈 분석	• 업무분석, 현업분석(인터뷰, 설문), 업무 프로그램 및 산출물, 품질핵심항목(CTQ)등 분석 관점별 수집된 산출물 리스트 및 시스템 등 분석 수행
3	서술적 업무규칙 정의	• 업무규칙을 서술적 형태로 정의
4	물리DB연관 정의	• 검증에 필요한 물리DB와 연관관계를 정의위해 시스템분석->테이블/컬럼단위 매핑
5	Sample 검증 및 오류 추정 데이터 분석	• 진단 SQL 정의, 진단 실시->추출 오류 추정데이터에 대한 현상 및 원인 분석
6	업무규칙 대상 후보군 도출	• 분석조건을 설정하고 재검증 실시 및 분석 이후 기술적 및 업무적 관점의 중요도를 평가하여 오류 데이터 및 오류 징후 데이터에 대해 업무규칙 대상 후보를 도출한다.
7	검증조건 추가 및 검증 실시	• 도출된 업무후보에 대한 업무간, 시스템간 연관관계 및 영향도를 파악하고 업무적 요건 사항을 고려한 후 검증조건을 추가하여 검증을 실시한다. • 품질진단 대상은 현재 시점 기준 2년 이내로 하되, 업무특성(차세대 등)을 고려하여 시점을 적용한다.
8	검증결과 분석	• 추가 검증조건을 통해 업무규칙 대상 후보의 검증결과를 분석하여 감시 필요성이 있는 데이터는 업무규칙 대상으로 선정한다
9	모니터링	• 선정된 업무규칙대상에 대해 감시 필요성 여부에 따라 주기를 지정하고 모니터링을 수행하여 오류데이터에 대한 추이를 파악한다.
10	업무규칙 확정	• 업무규칙정의서를 작성하고 DQMS에 등록한다.

○ 법령/규정에 근거한 업무규칙 도출

- 법령, 지침, 매뉴얼 등 행정업무처리의 기준이 되는 자료를 통해체계적으로 관리되어야 할 데이터의 업무규칙을 도출하고 작성하는 일련의 절차이다.

[그림 6-32] Outside-In방법 품질이슈를 파악 도출방식 (법령/규정 근거)



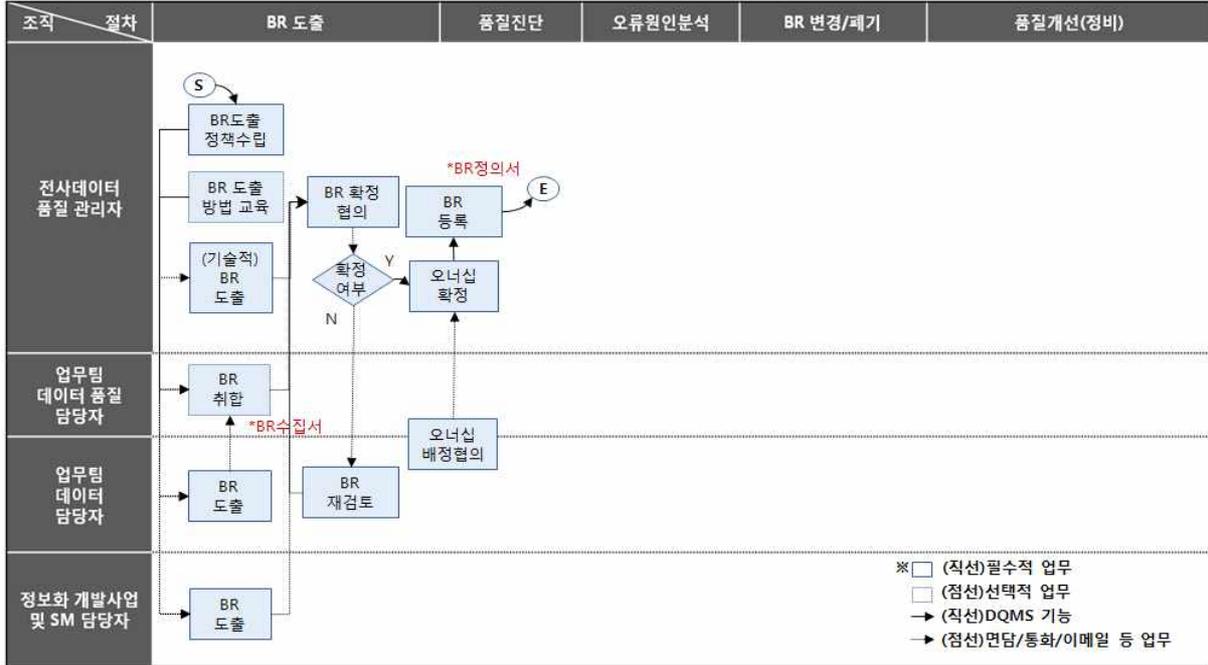
[표 6-32] Outside-In 업무규칙 도출 순서 (법령/규정 근거)

순서	절차	내용
1	근거규정 수집	• 현업의 요구(CSR), 업무 개발시 자체진단, 인터뷰 및 설문 등 이슈 사항을 중심으로 법령, 지침, 매뉴얼 등 자료 수집
2	근거규정 분석	• 수집된 근거규정에 대한 분석 수행
3	서술적 업무규칙 정의	• 업무규칙을 서술적 형태로 정의
4	물리DB연관 정의	• 검증에 필요한 물리DB와 연관관계를 정의위해 시스템분석->테이블/컬럼단위 매핑
5	Sample 검증 및 오류 추정 데이터 분석	• 진단 SQL 정의, 진단 실시->추출 오류 추정데이터에 대한 현상 및 원인 분석
6	업무규칙 대상 후보군 도출	• 분석조건을 설정하고 재검증 실시 및 분석 이후 기술적 및 업무적 관점의 중요도를 평가하여 오류 데이터 및 오류 징후 데이터에 대해 업무규칙 대상 후보를 도출한다.
7	검증조건 추가 및 검증 실시	• 도출된 업무후보에 대한 업무간, 시스템간 연관관계 및 영향도를 파악하고 업무적 요건 사항을 고려한 후 검증조건을 추가하여 검증을 실시한다. • 품질진단 대상은 현재 시점 기준 2년 이내로 하되, 업무특성(차세대 등)을 고려하여 시점을 적용한다.
8	검증결과 분석	• 추가 검증조건을 통해 업무규칙 대상 후보의 검증결과를 분석하여 감시 필요성이 있는 데이터는 업무규칙 대상으로 선정한다
9	모니터링	• 선정된 업무규칙대상에 대해 감시 필요성 여부에 따라 주기를 지정하고 모니터링을 수행하여 오류데이터에 대한 추이를 파악한다.
10	업무규칙 확정	• 업무규칙정의서를 작성하고 DQMS에 등록한다.

○ 업무규칙 도출 및 DQMS 등록 절차

- 데이터를 체계적으로 관리하기 위하여 관리되어야 할 데이터의 업무규칙을 도출하고 작성하는 일련의 절차이다.

[그림 6-33] 조직 담당자 간 업무규칙 도출 절차



[표 6-33] 조직 담당자 간 업무규칙 도출 절차 설명

업무	주요 내용
업무규칙 도출 정책 수립	<ul style="list-style-type: none"> • 업무규칙을 확보하기 위하여 업무규칙도출 주체를 정의하고, 도출 계획 및 방법을 수립한다.
업무규칙 도출 방법 교육	<ul style="list-style-type: none"> • 업무규칙 도출 주체들에 대해 업무규칙 가이드의 내용을 기준으로 역할 및 도출 방법을 교육한다.
업무규칙 확정협의	<ul style="list-style-type: none"> • 자체도출 및 접수된 업무규칙에 대해 타당성을 (필요시 작성자와 협의) 검토하고 적합한 건에 대해 업무규칙을 작성한다 • 신규 업무규칙에 대해 필요시 전사데이터품질관리자는 담당자에게 진단 SQL 변경 등의 재검토를 요청한다.
오너십 확정	<ul style="list-style-type: none"> • 등록 요청된 업무규칙에 대해 데이터품질담당자, 데이터담당자와 협의 후 오너십을 확정한다.
업무규칙 등록	<ul style="list-style-type: none"> • 검토 완료된 업무규칙에 대해 업무규칙 정의서를 기준으로 DQMS에 등록한다

○ 업무규칙 도출 정책 수립

- 전사 데이터품질관리자는 업무규칙이 필요한 항목을 점검하고 이에 필요한 업무규칙을 확보하기 위하여 업무규칙도출 주체를 정의하고 도출 계획 및 방법을 수립한다.

[표 6-34] 업무규칙 도출 주체 별 주요 업무

업무규칙 도출 주체	주요 업무
전사 데이터품질 관리자	<ul style="list-style-type: none"> 기술적 접근에 의한 업무규칙 도출 업무규칙 도출방법에 대한 교육
데이터품질 담당자	<ul style="list-style-type: none"> 오너십배정 협의 신규 업무규칙 취합 진단결과 분석요청 참조
데이터 담당자	<ul style="list-style-type: none"> 업무적 관점에 의한 업무규칙 도출 법률/규정에 근거한 업무규칙 도출
정보화 개발사업 담당자	<ul style="list-style-type: none"> 정보화 개발사업 수행시 발생한 업무규칙 도출
정보화 운영사업 담당자	<ul style="list-style-type: none"> 운영업무 수행시 필요한 업무규칙 도출

○ 업무규칙 도출 방법 교육

- 전사 데이터품질관리자는 업무규칙 도출 주체들에 대해 업무규칙 가이드의 내용을 기준으로 역할 및 도출 방법을 교육한다.

[표 6-35] 업무규칙 도출 방법 교육 문서 주요 내용

관련 문서	주요 내용
업무규칙 관리 가이드	<ul style="list-style-type: none"> 업무규칙의 개요 및 도출 방법
업무규칙 수집서	<ul style="list-style-type: none"> 업무규칙 수집 방법 및 작성 예시
업무규칙 품질지표 정의서	<ul style="list-style-type: none"> 업무규칙 품질지표 개요 업무규칙 품질지표별 주요 사례
데이터품질 업무 매뉴얼(DQMS)	<ul style="list-style-type: none"> DQMS 접속 및 사용방법

○ 신규 업무규칙 타당성 검토

- 전사 데이터품질관리자는 자체도출 및 접수된 업무규칙에 대해 타당성을 (필요시 작성자와 협의)검토하고 적합한 건에 대해 업무규칙을 작성한다. 타당성 검토항목은 다음과 같다.

신규 업무규칙 타당성 검토 항목
<ul style="list-style-type: none"> DQMS를 통한 기등록 BR과의 중복배제 및 유사 업무규칙 통합가능여부 검토 메타관리시스템을 통한 적합여부 검토 물리DB 적용 시 측정 적합여부 검토

○ 오너십 확정

- 전사 데이터품질관리자는 등록 요청된 업무규칙에 대해 데이터품질담당자, 데이터담당자와 협의 후 오너십을 확정한다.

○ 업무규칙 정의서

- 신규로 확정된 업무규칙을 DQMS에 등록하기 위해 전체내용을 확인할 수 있도록 각 항목을 다음과 같은 기준에 따라 정의서를 작성한다.

[그림 6-34] 업무규칙정의서 양식

업무규칙 ID	ROTA-공정-RNL-202	작성자	박OO	작성일자	2020.09.xx
시스템 명		DB 명	OOOO DB	담당자	공공데이터 담당관
업무규칙 명	DB구축기준				
업무규칙 내용	컬럼에 등록된 합산 금액은 (인건비+설치비+예비비) 합산금액이 동일해야함				
근거 규정	지능형 OOOO 표준노드링크구축관리지침				
테이블 명	LINK	컬럼 명		품질지표	
스크립트	SELECT COUNT(*) FROM (select nvl(M,0) + nvl(M,0) + nvl(P_M,0) tot, SM , (nvl(LM,0) + nvl(FM,0) + nvl(PPM,0) - M) as zero from TB_RPT where SM is not null) WHERE t_z > 0				
원인분석					
업무영향	필수				

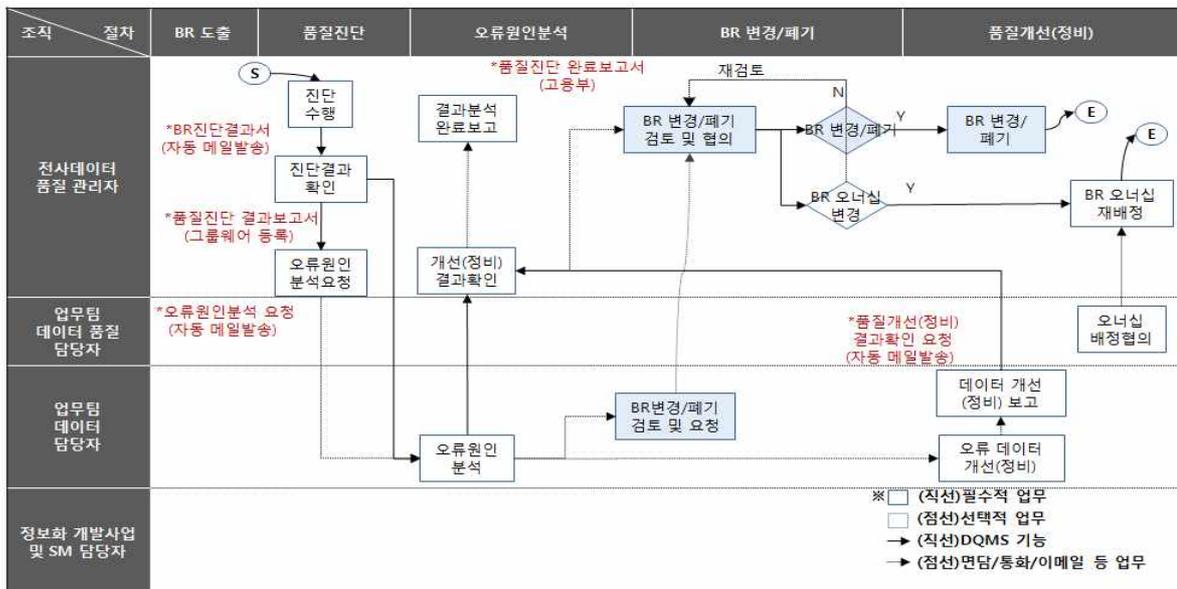
○ 업무규칙 등록

- 데이터품질관리자는 검토 완료된 업무규칙에 대해 업무규칙 정의서를 기준으로 DQMS에 등록한다. 업무규칙 등록에 대한 자세한 사항은 데이터품질 (DQMS) 업무 매뉴얼을 참고한다.

○ 업무규칙 변경/폐기

- 업무규칙의 변경 또는 폐기가 필요한 경우 다음과 같은 절차에 따라 업무를 수행한다.

[그림 6-35] 업무규칙 변경/폐기 절차



- 업무규칙 변경/폐기 검토 및 요청

- 데이터품질담당자는 업무규칙의 원인분석을 토대로 업무규칙에 대한 변경/폐기를 검토한다. 검토할 근거기준은 아래와 같다.

업무규칙 변경/폐기 근거 기준
<ul style="list-style-type: none"> • 법령/규칙의 변경 • 업무절차 변경(전산 프로세스 변경) 포함 • 데이터베이스 모델변경(테이블, 컬럼의 변경/통합/삭제) • 자체판단(1년 이상 오류 미 검출시 등) • 진단결과 갑작스런 변화(오류율 급증 등)가 생겼을 경우 업무규칙 변경/폐기가 발생했을 가능성 있음

- 업무규칙 변경/폐기 협의

- 데이터품질담당자는 업무규칙의 변경/폐기에 대한 검토결과를 협의한다.

- 업무규칙 폐기 확정

- 데이터품질담당자는 폐기사유가 합당할 경우 업무규칙폐기를 승인하고, 데이터담당자에게 결과를 통보한다.

□ 품질진단도구를 통한 업무규칙 진단

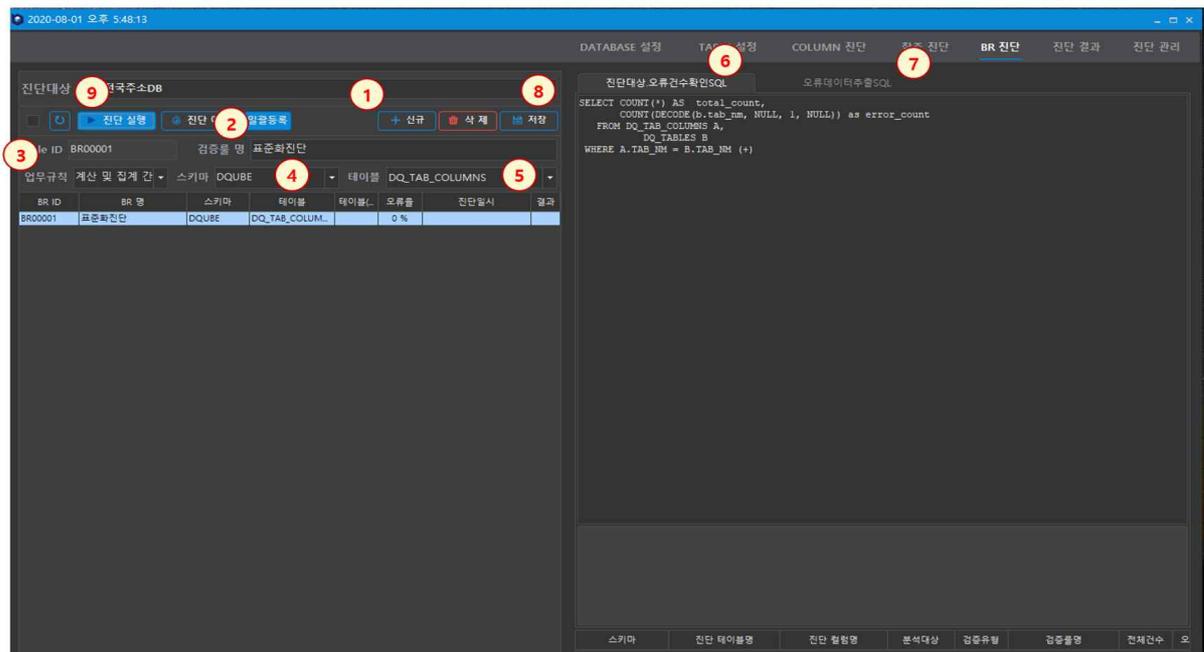
○ 진단대상 업무규칙

- 진단대상 업무규칙의 종류

- 시간순서 일관성
- 선후관계 정확성
- 계산 및 집계간 일관성
- 칼럼간 논리관계 일관성

○ 업무규칙 진단

[그림 6-36] 업무규칙 진단 화면 입력 절차



- ① 진단대상 DB를 선택한 후 [신규]버튼을 클릭합니다.
- ② 검증룰 명을 입력합니다.
- ③ 업무 규칙의 종류를 선택합니다.
- ④ 진단 대상 스키마(계정)를 선택합니다.
- ⑤ 진단 대상 테이블을 선택합니다.
- ⑥ 오류데이터 추출 SQL을 작성합니다. 오류건수 추출SQL은 내부적으로 자동으로 생성됩니다
- ⑦ [저장]버튼을 클릭합니다. 이 때 작성된 SQL들의 Syntax를 체크하여 Syntax에 문제가 있는 경우 알려줍니다.
- ⑧ [▶ 진단 실행]버튼을 클릭하여 업무규칙(BR) 진단을 수행합니다.
 - ※ 진단 결과는 [진단 결과]화면에서 확인하실 수 있습니다.

○ 업무규칙 진단결과

- [진단 결과] 화면에서 업무규칙 진단 결과 확인

[그림 6-37] 업무규칙 진단 결과 확인 화면

2020-08-01 오후 5:48:13

DATABASE 설정 TABLE 설정 COLUMN 진단 참조 진단 BR 진단 **진단 결과** 진단 관리

진단대상 DB 전국주소DB

전체건수 6,536,327
오류건수 6,449,563
오류율 98.672%

전체 테이블수: 512 진단대상 테이블수: 504 전체 컬럼수: 8,222
 도메인 (실형/실정): 15 / 61 업무규칙 (실형/실정): 0 / 1 참조 (실형/실정): 2 / 2

분석대상	검증유형	진단대상(오류/전체)	테이블	전체 데이터	오류 데이터	오류율
도메인	불차	3 / 3	3	54,086	54,086	100.00 %
도메인	반복	8 / 10	6	6,444,595	6,394,703	99.22 %
도메인	수평	0 / 1	1	3	0	0.00 %
도메인	코드	1 / 1	1	773	773	100.00 %
참조	참조	1 / 2	2	36,670	1	0.00 %

스키마	진단 테이블명	진단 컬럼명	검증유형	전체건수
DQUBE	DQ_GO_WORD	WORD_NM	계시판일련번호	200
DQUBE	DQ_TAB_COLUMNS	CUST_NO	공공표준화	36,670

○ 보고서 출력

- [진단결과 엑셀저장] 버튼을 클릭하여 보고서 출력

[그림 6-39] 업무규칙 진단 결과 엑셀 보고서 출력

값 진단 종합 현황						
						출력일 2020-08-01 18:14:53
진단 데이터베이스 기본 정보						
진단기간	2020-08-01 ~ 2020-08-01					
기관명	MITSOFT					
DB명	전국주소DB	DB 설명	전국 도로명주소 및 지번주소 데이터			
DBMS 종류	Oracle	DBMS 버전	11.2.0.4.0			
IP	*.0.7	PORT	1521			
진단규칙 설정 현황	전체 테이블수	진단대상 테이블수	전체 컬럼수	도메인 (실행/설정)	참조 (실행/설정)	업무규칙 (실행/설정)
	512	504	8222	15 / 61	2 / 2	1 / 1
분석영역	검증유형	진단대상	테이블	전체데이터	오류데이터	오류율(%)
도메인	수량	1	1	3	0	0 %
	날짜	3	3	54,086	54,086	100 %
	번호	10	6	6,444,595	6,394,703	99.225 %
데이터규칙	코드	1	1	773	773	100 %
	참조	2	2	36,870	1	0.002 %
업무규칙	계산 및 집계 간 일관성	1	1	38,999	0	0 %
전체				6,575,326	6,449,563	98.087 %

3. 데이터 품질관리 역량 향상 교육

○ 총 2회에 걸쳐 공공데이터 품질관리 역량 향상 교육을 진행하였음

[표 6-30] 품질관리 역량 향상 교육 현황

구분	교육일시	교육명	교육 내용
1차	2021.06.29	공공데이터품질관리 기본교육	<ul style="list-style-type: none"> ▪ 공공데이터 개요 ▪ 공공데이터 품질관리 ▪ 공공데이터 표준화 정책 ▪ 품질관리 수준평가 개요 ▪ 수준평가 산출물 작성 방법 ▪ 데이터기반행정 개요
2차	2021.07.27	공공데이터품질관리 수준평가 DB담당자 교육	<ul style="list-style-type: none"> ▪ 공공데이터 품질관리 수준평가 개요 ▪ 데이터 구조산출물 작성 가이드

[그림 6-40] 공공데이터 품질관리 역량 향상 교육 1차 교육자료



[그림 6-41] 공공데이터 품질관리 역량 향상 교육 2차 교육자료



CHAPTER VII

최종 평가 결과 및 결언

1. 적합성 평가 품질 진단 최종 결과
2. 결언

CHAPTER VII 최종 평가 결과 및 결론

1. 적합성 평가 품질 진단 최종 결과

□ 4개 고유업무 6개 시스템 표준/품질 평가 결과 만점 획득

[그림 7-1] 고유업무 표준/품질 측정 결과

DB유형	기관명	시스템명	DB명	점검 범위
고유업무	우정사업본부	우체국금융시스템	전자금융(FNIPD)	데이터 관리 체계 데이터 값 관리
		인터넷우체국(epost)	통합DB1	
	중앙전파관리소	비정지위성전파감시시스템	비정지위성전파감시 DB	
		전파관리정보시스템	전파관리 DB	
	국립전파연구원	방송통신통합정보시스템	방송통신통합 DB(적합성평가)	
	국립중앙과학관	국가자연사연구종합정보시스템	US_NTISAPP	
지원업무	과학기술정보통신부	대표홈페이지	대표홈페이지 DB	
	국립전파연구원	우주전파환경홈페이지	우주전파환경홈페이지 DB	
	국립중앙과학관	국립중앙과학관 홈페이지	NSM2018	

평가영역	기준점수	세부지표	내용	점수		
데이터 표준 확산	10	데이터 표준 정의(4)	데이터 표준용어 정의(1.5점) 데이터 표준도메인 정의(1.5점) 데이터 표준코드 정의(1점)	1.5 1.5 1		
		데이터 표준 적용률(6)	데이터 표준 적용률(6점)	6		
		데이터 구조 안정화	5	데이터 구조 정의 및 산출물 관리(2)	데이터베이스 정의서(0.5점) 테이블 정의서(0.5점) 컬럼 정의서(0.5점) 물리데이터모델 다이어그램(ERD) 또는 관계 정의서(0.5점)	0.5 0.5 0.5 0.5
				데이터 구조 산출물 현행화(3)	데이터 구조 산출물 현행화(3점)	3
데이터 연계 관리	5			연계데이터 관리 정보 정의(2) 연계데이터 송·수신 내역 관리(3)	2 3	
데이터 품질진단	15	품질진단 기준 정의 및 진단(5) 업무규정에 따른 품질진단 규칙(BR) 정의 및 진단(10)	5 10			
품질진단 결과 조치	23	품질진단에 따른 개선계획 수립 및 개선과제 정의(8) 개선조치 이행률(15)	8 15			
데이터 오류율	15	데이터 오류율(15)	데이터 오류율(15점)	15		
총 기준점수	73		총 평가 획득 점수	73		

□ 3개 지원업무 3개 시스템 표준/품질 평가 결과 만점 획득

[그림 7-2] 지원업무 품질 측정 결과

DB유형	기관명	시스템명	DB명	점검 범위
고유업무	우정사업본부	우체국금융시스템	전자금융(FNIPD)	데이터 관리 체계 데이터 값 관리
		인터넷우체국(epost)	통합DB1	
	중앙전파관리소	비정지위성전파감시시스템	비정지위성전파감시 DB	
		전파관리정보시스템	전파관리 DB	
	국립전파연구원	방송통신통합정보시스템	방송통신통합 DB(적합성평가)	
	국립중앙과학관	국가자연사연구종합정보시스템	US_NTISAPP	
지원업무	과학기술정보통신부	대표홈페이지	대표홈페이지 DB	
	국립전파연구원	우주전파환경홈페이지	우주전파환경홈페이지 DB	
	국립중앙과학관	국립중앙과학관 홈페이지	NSM2018	

평가영역	기준점수	세부지표	내용	점수
데이터 표준 확산	10	데이터 표준 정의(4)	데이터 표준용어 정의(1.5점) 데이터 표준도메인 정의(1.5점) 데이터 표준코드 정의(1점)	
		데이터 표준 적용률(6)	데이터 표준 적용률(6점)	
데이터 구조 안정화	5	데이터 구조 정의 및 산출물 관리(2)	데이터베이스 정의서(0.5점) 테이블 정의서(0.5점) 컬럼 정의서(0.5점) 물리데이터모델 다이어그램(ERD) 또는 관계 정의서(0.5점)	
		데이터 구조 산출물 현행화(3)	데이터 구조 산출물 현행화(3점)	
		데이터 연계 관리	5	연계데이터 관리 정보 정의(2) 연계데이터 송·수신 내역 관리(3)
데이터 품질진단	15	품질진단 기준 정의 및 진단(5) 업무규정에 따른 품질진단 규칙(BR) 정의 및 진단(10)	5 10	
품질진단 결과 조치	23	품질진단에 따른 개선계획 수립 및 개선과제 정의(8) 개선조치 이행률(15)	8 15	
데이터 오류율	15	데이터 오류율(15)	데이터 오류율(15점)	15
총 기준점수	53		총 평가 획득 점수	53

2. 결론

- 2021년 공공데이터 제공 및 이용 활성화 시행계획 및 향후 진행될 국가 시행계획에 따라 기관이 보유한 데이터가 활용도 높게 개방되도록 본 기관 및 산하기관의 데이터 품질을 장기적 관점에서 제고
- 데이터 품질관리 수준평가에 의한 단기적 대응도 필요하나 장기적 관점에서 본 기관(과학기술정보통신부)의 기관표준 운영 시 용어표준의 기반이 되는 단어표준 운용 및 단어 기반 표준용어 운용을 권고
- 기관 차원의 근본적 품질관리 및 상시 품질관리를 위해 평가용 품질관리도구 외 기관이 직접 품질관리도구를 보유하고, 상시 품질 관리가 되도록 품질 관리가 되도록 품질 관리 계획 및 활동 계획을 수립
- 공공데이터 제공 및 이용 활성화의 목적에 부합하도록 개방 보유 데이터를 사전 점검하여 민간 데이터 활용도가 높다고 판단되는 기관 데이터를 선정하고 개방될 수 있도록 기관/직속기관 담당자 독려 및 개방 추진
- 특히 과학기술정보통신부 보유 데이터는 국민 생활형 데이터 보다는 이에 대한 근간을 제공하는 기초형 데이터에 가까우며 특히 인공지능 학습 데이터로의 활용도가 크므로 이에맞게 정형/비정형 데이터가 개방되도록 데이터 품질기준을 수립할 것을 권고