발간등록번호
11-1543000-004314-01

농림축산식품부
Ministry of Agriculture, Food and Rural Affairs

스마트 농업 데이터 거래 환경 구축 및 운영 모델 연구

2023. 03. 30.

SUNCHON NATIONAL UNIVERSITY 순천대학교

목 차

제 1장 스마트팜 환경 데이터 분류방안 및 가치평가 방법 조사 및 검토 ····· 1
가. 스마트팜 환경 데이터 분류방안 ····· 1
나. 데이터 가치평가 방법 조사 ····· 20

제 2장 국내 데이터 거래 관련 플랫폼 및 가격산정 사례 조사 ····· 26
가. 국내 데이터 거래 관련 플랫폼 ····· 26
나. 국내 데이터 거래 관련 가격산정 사례 조사 ····· 33

제 3장 표준확산사업 연계 방안 및 관련 표준화 준수사항 도출 ····· 42
가. 표준확산사업 연계 방안 ····· 42
나. 데이터 가치평가 시 표준화 준수사항 ····· 43

제 4장 데이터 거래를 위한 예상 사용자 요구사항 조사 ····· 64
가. 데이터 거래 계약사항 ····· 64
나. 데이터 거래 관련 문제점 및 개선 방안 조사 ····· 67
다. 잠재 공급자 조사 결과 ····· 68
라. 잠재 수요자 및 농업 전문가 조사 결과 ····· 82
마. 생산자 측면에서의 소유권 문제 해결 방안 조사 ····· 94
바. 타 산업 데이터 거래소와의 차별성 도출 ····· 99

제 5장 국내외 오픈 데이터 활용 사례 ····· 102
가. 국외 오픈 데이터 현황 ····· 102
나. 국제기구의 오픈 데이터 포털 ····· 104
다. 국내 주요 오픈 데이터 플랫폼과 활용 사례 ····· 105
라. 국외 민간 및 학계의 데이터 개방/공유 현황 ····· 107
마. 주요 국내 민간 주도 데이터 공개와 활용 사례 ····· 109

제 6장 데이터 가격 평가 방법 ····· 113
가. 데이터 가격 결정 평가 ····· 113
나. 판매 및 가격 전략 수집 ····· 118
다. 데이터 가격 책정 방법 ····· 122

제 7장 스마트팜 데이터 거래환경(시스템) 구축 ····· 176
가. 시스템 운영 모델 ····· 176
나. 운영모델 가동을 위한 요구사항 ····· 181
다. 데이터 거래의 사례 ····· 185
라. 예상 문제점 및 개선방안 ····· 191
마. 데이터거래를 위한 스마트농업육성법 수정 제안 ····· 192

제 8장 결론 및 시사점 ····· 196
가. 연구용역 사업 보고서의 결론 ····· 196
나. 시사점 ····· 201

제 9장 참고문헌 ····· 206

1. 스마트팜 환경 데이터 분류방안 및 가치평가 방법조사 및 검토

가. 스마트팜 환경 데이터 분류방안

(1) 기본 데이터 유형 분류 체계

<데이터 유형 분류 체계>

○ 빅데이터는 생산 주체에 따라 프로세스에 의해서 생성되는 프로세스 생성, 각종 기기에 의하여 생산되는 기계 생산 및 사람에 의하여 생산되는 사람 생성으로 구분

(2) 스마트팜 데이터 유형 분류

데이터 유형 분류	설명
센서 데이터	• 기상 데이터 수집 • 토양 데이터 수집 • 생육 데이터 수집 등
생산 데이터	• 시설 및 농기계 현장 데이터 • 품질 데이터 • 농업 경영체 데이터베이스 및 소득자료
유통 및 소비 데이터	• 농산물 출하 정보 • 소비자 패널 자료 • SNS 소비 데이터 등
영상 장비 데이터	• CCTV • 적외선 카메라 • DVR 등
기타 데이터	• 농업 복지 데이터 • 농촌 인구 데이터 • 농업 일자리 데이터 등

(3) 스마트팜 데이터 분석 방법

데이터 유형	데이터 정의	환경생육데이터 정렬 및 결합	특이값 확인 및 결측치 보정
환경 데이터	• 온도 및 습도 • 일사량 • PH • EC 등	• 단위 : 분-시-주야-일-주차-월 • 시간 단위 환경/생육 정보 결합	• 시계열 그래프를 통한 특이값 확인 • 결측치가 연속으로 두 개 이상 발생 시 이동평균법을 통해 결측치 보정 • 3개 이상 연속으로 발생 시 패턴 자료 이용법을 통해 결측치 보정
구동기 데이터	• 유형 • 면적 • 천창 및 측창 • 유동팬 • 환풍기 등		
작물 생육 데이터	• 품목 • 정식일시 • 초장 • 엽수 등		
경영 데이터	• 수입 • 수량 • 품질 • 인건비 등		

(4) 스마트팜 데이터 분석 세부항목

○ 환경데이터 조사항목

조사항목	표준단어	표준단어(영문)	측정단위
외부감우	감우	RAIN_DETECTOR	O/X
외부상대습도	외부상대습도	EXTRL_PARTN_HD	%
외부야간평균온도	외부온도(야간평균)	EXTRL_NIGHT_AVRG_TP	℃
외부주간평균온도	외부온도(주간평균)	EXTRL_WIK_AVRG_TP	℃
외부최고온도	외부온도(최고)	EXTRL_TOP_TP	℃
외부최저온도	외부온도(최저)	EXTRL_LWET_TP	℃
외부평균온도	외부온도(평균)	EXTRL_AVRG_TP	℃
외부최고일사량	일사량(최고)	MXMM_SOLRAD_QY	W/㎡
외부누적일사량	일사량(누적)	ACCMLT_SOLRAD_QY	J/cm2
내부야간평균상대습도	내부상대습도(야간평균)	INNER_NIGHT_AVRG_PARTN_HD	%
내부주간평균상대습도	내부상대습도(주간평균)	INNER_WIK_AVRG_PARTN_HD	%
내부최고상대습도	내부상대습도(최고)	INNER_TOP_PARTN_HD	%
내부최저상대습도	내부상대습도(최저)	INNER_LWET_PARTN_HD	%
내부평균상대습도	내부상대습도(평균)	INNER_AVRG_PARTN_HD	%
내부야간평균온도	내부온도(야간평균)	INNER_NIGHT_AVRG_TP	℃
내부주간평균온도	내부온도(주간평균)	INNER_WIK_AVRG_TP	℃
내부최고온도	내부온도(최고)	INNER_TOP_TP	℃
내부최저온도	내부온도(최저)	INNER_LWET_TP	℃
내부평균온도	내부온도(평균)	INNER_AVRG_TP	℃

배양실내부CO2	배양실(CO_2)	CLTVT_CO2	%
배양실내부습도	배양실(습도)	CLTVT_PARTN_HD	%
배양실내부온도	배양실(온도)	CLTVT_LWET_TP	℃
생육실내부CO2	생육실(CO_2)	GRWH_CO2	%
생육실내부습도	생육실(습도)	GRWH_PARTN_HD	%
생육실내부온도	생육실(온도)	GRWH_LWET_TP	℃
내부CO2농도	CO_2	CO2	%
내부배지EC	배액EC	EXLQ_EC	dS/m
내부조도	조도	IMIT	Lx
내부차광율	차광율	SHADE_PT	%
내부드리퍼당공급량	드리퍼당 공급량	SUPLY_QY_PER_DRIP	l
내부지하부(배지)온도	배지온도	MDIM_TP	℃
내부양액공급량	양액 공급량	NUTRIENT_SLUTN_SUPLY_QY	cc/m2
내부CO2설정값	CO_2 설정값	CO2_SETUP_VALUE	%
내부공급횟수	양액공급횟수	NUTRIENT_SUPLY_CUNT	회차
내부관수배수율	관수 배수율	IRRN_DAIL_DRNG_PT	%
내부관수일배액량	관수 일배액량	IRRN_DAIL_EXLQ_QY	cc/m2
내부일회공급량	관수 일회공급량	IRRN_OTM_SUPLY_QY	l

○ 제어데이터 조사항목

조사항목	표준단어	표준단어(영문)	측정단위
3Way밸브	3Way밸브	THREE_WAY_VALVE	On/Off
3Way밸브공급수온도	공급수온도	SUPLY_WATER_TP	℃
CO_2 발생기	CO_2 발생기	OCPRD_CO2	On/Off
CO_2 발생기+난방기	CO_2 발생기+난방기	CO2_OCCRRNC_MHRLS	On/Off
CO_2 설정값	CO_2 설정값	CO2_SETUP_VALUE	ppm
관수/관비시스템	관수/관비시스템	IRRN_FRTLZR_SYS	On/Off
관수작동상태(수경)	관수작동상태(수경)	IRRN_WRFN_STTUS_WATER	On/Off
관수작동상태(토경)	관수작동상태(토경)	IRRN_WRFN_STTUS_SOIL	On/Off
난방기	난방기	HEAT_MHRLS	On/Off
난방조절	난방설정	HEAT_SETUP	℃
냉난방기	냉난방기	ARCLNG_HEAT_MHRLS	On/Off
냉방기	냉방기	ARCLNG_MHRLS	On/Off
두상살수장치	두상살수장치	OVERHEAD_IRRN_MHRLS	On/Off
배기팬	배기팬	EXHAUST_FAN	On/Off
보광등	보광등	SPPL_LMP	On/Off
보온덮개개폐기	보온덮개 개폐기	HEAT_COVER_SWITCHGEAR	On/Off
보온스크린(천장)	천장보온커튼	CELLARS_HEAT_CURTAIN	On/Off
보온스크린(측면)	측면보온커튼	WALL_HEAT_CURTAIN	On/Off
보일러	보일러	BLR	On/Off
분무(살수장치)	분무(살수장치)	SPRAYER	On/Off
수막시스템	수막시스템	WATER_CURTAIN_SYS	On/Off

양액시스템	양액시스템	NUTRIENT_SLUTN_SYS	On/Off
유동팬	유동팬	DYNMC_FAN	On/Off
유량계	유량계	FLWMTR	l
이중창	이중창	DPLC_WNDW	On/Off
차광커튼(스크린)	차광커튼	SHADE_CURTAIN	On/Off
환기구	환기구	VENTL_FAN	On/Off
천창	천창	CEILING_WNDW	On/Off
측창	측창	SIDE_WALL_WNDW	On/Off
환기조절	환기조절	VENTL_ADJST	On/Off
훈증기	훈증기	FMG_MHRLS	On/Off

○ 생육데이터 조사항목

조사항목	표준단어	표준단어(영문)	측정단위	비고
엽장, 엽길이,잎의길이	잎길이	LEF_LT	cm	딸기, 방울토마토, 오이, 참외, 토마토
엽폭, 잎의너비	잎너비	LEF_BT	cm	딸기, 방울토마토, 오이, 참외, 토마토
잎수, 엽수, 잎의갯수	잎개수	LEF_CUNT	개	딸기, 방울토마토, 오이, 토마토
초장	초장	TOT_LT	cm	딸기, 방울토마토, 참외, 토마토, 파프리카
생장길이	생장길이	GRWT_LT	cm	딸기, 방울토마토, 참외, 토마토, 파프리카
줄기굵기, 줄기직경	줄기굵기	STEM_THCK	mm	방울토마토, 오이, 토마토, 파프리카
착과화방	착과화방	FRTST_FCLU	화방	방울토마토, 토마토, 파프리카
화방당착과수	착과수	FRTST_CUNT	개	딸기, 방울토마토, 토마토, 파프리카
생장점과화방거리	화방높이	FCLU_HG	cm	방울토마토, 토마토, 파프리카
수확화방	수확화방	HVST_FCLU	화방	방울토마토, 토마토, 파프리카
수확수	수확수	HVST_CUNT	개	방울토마토, 토마토, 파프리카
개화화방	개화화방	FRMHS_FCLU	화방	방울토마토, 토마토, 파프리카
개화기	개화기	BLPR	년월일	딸기
화방당소화수, 개화수	개화수	FLAN_CUNT	개	딸기, 방울토마토, 토마토, 파프리카
과장, 과고	과장	FRUIT_LT	cm	딸기, 오이, 참외, 파프리카
과중	과중	FLSH_WT	g	딸기, 참외
과폭	과폭	FLSH_ARA	cm	딸기, 오이, 참외, 파프리카
관부직경,경경, 줄기굵기	관부직경	CROWN_DMT	mm	딸기, 참외, 파프리카
당도	당도	BRX	Brix	딸기, 참외
산도	산도	ACIDITY	pH	딸기
액아발생시기	액아 발생기	AXBD_OCPRD	년월일	딸기
엽병장	잎자루길이	PTL_LT	cm	딸기
화방출뢰기	화방출뢰기	FCLU_BUDDING_DT	년월일	딸기
마디수	마디수	JNT_CUNT	마디	오이
암꽃수(마디별)	암꽃수(마디별)	FEMALE_FLWR_CUNT_PER_JNT	개	오이
절간장	마디길이	INTERNODE_LT	cm	오이
주당과중	주당 과중	PLNTT_FRUIT_WT	g	오이
주당열매수	주당 열매수	PLNTT_FRUIT_CUNT	개	오이
착과습성	줄기유인방법	STEM_ENTICE_MTHD	-	오이
과육두께	과육두께	FLSH_THICK	cm	참외

○ 경영데이터 조사항목

조사항목	표준단어	표준단어(영문)	측정단위
CO2비	CO_2 비	CO2_COST	원
감가상각비	감가상각비	DPEX	원
인건비	인건비	LBCST	원
기타비용	기타 비용	ETC_COST	원
난방비(유류,전기)	난방비	HEAT_COST	원
농약비	농약비	AGCHM_COST	원
소농구비	소농구비	SMALL_FMHPS_COST	원
수리비	수리비	REPAIR_COST	원
수선비	수선비	MENDING_COST	원
양액(비료)비	양액비	NUTRIENT_SLUTN_COST	원
위탁영농비	위탁영농비	CNSGN_FARM_COST	원
임차료	임차료	HIRE_CHARGE	원
재료비	재료비	MATRL_COST	원
조수입	총수입	TOT_INCME	원
종묘비	종묘비	PLNSDL_COST	원
출하량(판매량)	출하량	SHIPMNT_QY	kg
품질상	품질(상)	QLITY_TOPFR	등급
품질중	품질(중)	QLITY_MDDL	등급
품질하	품질(하)	QLITY_BTTO	등급

(4) 농림축산식품부 공공데이터 주제별 현황

순번	주제	개방시기	데이터 건수
1	농산	2013년	36
2	축산	2013년	47
3	산림	2013년	97
4	유통	2013년	163
5	농지용수	2014년	46
총 합계	-		389

○ 총 다섯 가지 주제로, 농산(36건), 축산(47건), 산림(97건), 유통(163건), 농지용수(46건)로 이루어졌으며, 해당 주제별 데이터 총 개수로 389개가 있음.

<농림축산식품부 공공데이터 제공기관별 현황>

순번	제공기관	개방시기	데이터 건수
1	농림축산식품부	2013년	328
2	국립농산물품질관리원	2013년	53
3	국립종자원	2014년	44
4	농림축산검역본부	2013년	35
5	한국농수산대학	2014년	8
6	농식품공무원교육원	2013년	16
7	한국농어촌공사	2012년	84
8	한국농수산식품유통공사	2014년	41
9	축산물품질평가원	2013년	19
10	농림수산식품교육문화정보원	2012년	113
11	농림수산식품기술기획평가원	2014년	101
12	한국마사회	2012년	82
13	가축위생방역지원본부	2016년	2
14	농업정책보험금융원	2016년	1
15	국제식물검역인증원	2016년	4
16	한식진흥원	-	0
17	농촌진흥청	2013년	67
18	산림청	2013년	120
19	농촌경제연구원	2014년	23
총 합계	-		1,141

○ 총 19개의 제공기관이 있으며, 총 공공데이터 건수는 1,141건으로 제공되고 있음. 공공데이터의 건수가 많은 순으로 농림축산식품부(328건), 산림청(120건), 농림수산식품교육문화정보원(113건)이 있으며, 적은 순으로 한식진흥원(0건), 농업정책보험금융원(1건), 가축위생방역지원본부(2건)이 있음.

<농림축산식품부 공공데이터 서비스 유형별 공공데이터 현황>

서비스유형 / 분리체계	Grid	Link	File	Open API	원시 데이터	Map	Chart
농촌복지	40	27	31	43	14	4	3
말산업	11	19	10	71	4	1	0
교육	10	20	24	10	3	2	3
식품산업	31	16	37	31	22	1	8
축산	26	10	21	29	22	0	21
농산	14	8	23	15	10	0	8
귀농귀촌	4	2	1	4	3	1	3
시설장비	6	2	13	6	6	6	6
품질관리	22	14	22	24	7	6	1
농지용수	14	1	32	17	12	0	10
산림	9	48	27	30	6	0	5
농업정책	11	3	15	12	9	0	0
검역	9	4	10	9	6	2	4
식품안전	17	7	15	20	5	1	11
유통	94	38	64	102	49	3	17
농축산기술	71	118	154	93	49	2	12
합계	389	392	511	519	231	29	112

○ 데이터 제공은 주로 파일데이터(511건)와 OpenAPI(519건)로 제공되고 있으며, 그 밖에도 Grid(389건)와 Link(392건), Chart(112건), Map(29건), 원시데이터(231건) 서비스유형이 제공되고 있음.

<농림축산식품분야 개방 공공데이터 목록(2019년, 10월 기준)>

대분류	소분류	공공데이터명	기관명	개방서비스			
				그리드	Open API	파일	링크
농산	식량작물	고품질브랜드쌀현황	농림축산식품부	○	○	○	
		국영무역품목 시장접근물량	한국농수산식품유통공사			○	○
		국영무역품목 양허세율	한국농수산식품유통공사			○	○
		비축시설 현황	한국농수산식품유통공사			○	○
		작물영향 진단 서비스	농촌진흥청				○
	원예작물	계약재배(노지채소) 현황	농림축산식품부	○	○		
		과실출하약정 사업실적 현황 (요약)	농림축산식품부	○	○		
		과실출하약정 상세현황	농림축산식품부	○	○		
		시설채소 생산실적	농림축산식품부	○	○	○	
		시설채소 온실현황	농림축산식품부	○	○	○	
		채소류 생산실적	농림축산식품부	○	○	○	
		폐업지원 지역별 신청규모정보	농림축산식품부	○	○		
		화훼류 시도별 재배시설 현황	농림축산식품부	○	○	○	
		화훼류 시도별 재배현황	농림축산식품부	○	○	○	
		화훼류 시도별 재배현황 (재배농가)	농림축산식품부	○	○	○	
		화훼류 시도별 재배현황 (재배인력)	농림축산식품부	○	○	○	
		화훼류 시도별 재배현황 (품목별 재배현황)	농림축산식품부	○	○	○	
	특용작물	약용식물 서비스	산림청		○		
		특용작물 생산통계	농림축산식품부	○	○	○	
축산	곤충	누에사육 규모별 농가수	농림축산식품부	○	○	○	
		누에씨 생산실적 및 계획	농림축산식품부	○	○	○	
		뽕나무 묘목 생산 공급현황	농림축산식품부	○	○	○	
		뽕밭 규모별 농가수	농림축산식품부	○	○	○	
		양잠 농가현황 (기능성 양잠 경영주 연령 분포)	농림축산식품부	○	○	○	
		양잠 농가현황 (기능성 양잠업 종사경력)	농림축산식품부	○	○	○	
		양잠 농가현황 (양잠농가 소득현황)	농림축산식품부	○	○	○	
		양잠산물 생산현황	농림축산식품부	○	○	○	
		양잠현황 (총괄)	농림축산식품부	○	○	○	
	동물	가축 분뇨 살포 현황	농림축산식품부	○	○		
		가축 분뇨 작물별 액비살포현황	농림축산식품부	○	○		
		기타 가축통계	농림축산식품부			○	
		기타 가축통계(개/마필)	농림축산식품부	○	○		
		기타 가축통계(거위/사슴/칠면조)	농림축산식품부	○	○		
		기타 가축통계(꿀벌)	농림축산식품부	○	○		
		기타 가축통계(메추리)	농림축산식품부	○	○		
		기타 가축통계(면양)	농림축산식품부	○	○		
		기타 가축통계(산양)	농림축산식품부	○	○		
		기타 가축통계(토끼)	농림축산식품부	○	○		
		동물보호관리시스템 유기동물 조회 서비스	농림축산검역본부				○
		동물실험실윤리위원회 운영 및 동물실험 실태	농림축산검역본부			○	
		반려동물 등록대행업체 조회	농림축산검역본부	○	○		○

	사료	배합 사료 생산 실적	농림축산식품부	○	○	○	
		표준 사료 검색 서비스	농촌진흥청		○		
	유제품	시도별 우유 생산량	농림축산식품부	○	○	○	
		유제품별 생산 및 소비실적	농림축산식품부	○	○	○	
	축산농장	가축 거래 상인 등록현황	농림축산식품부	○	○		
		가축사육 통계제공 서비스	농림축산식품부		○		
		가축사육현황	가축위생방역지원본부	○	○	○	
		농가현황	가축위생방역지원본부	○	○	○	
		축산농장 통계제공 서비스	농림축산식품부		○		
		축산업허가등록지역별통계현황	농림축산식품부	○	○		
산림	산림자원	경제림 육성 단지 구역도	산림청				○
		기본정보(목재제품 DB백과)	산림청				○
		기본정보(임산물 DB백과)	산림청				○
		기증자료 서비스	산림청		○		
		단기임산물재배적지도	산림청				○
		둘레길 정보조회	산림청		○		
		맞춤형 조림지도 서귀포시	산림청			○	
		맞춤형 조림지도-제주시	산림청			○	
		백두대간 보호 구역도	산림청				○
		백두대간 서비스	산림청		○		
		보호림	산림청		○		
		산림 사업법인 정보	산림청				○
		산림 유전자원 서비스	산림청		○		
		산림사업용 묘목가격	산림청	○	○	○	
		산악기상정보	산림청				○
		산양삼 생산신고 통계	산림청	○	○	○	
		산촌생태마을 위치정보	산림청			○	
		산코드	산림청			○	
		상세정보(목재제품 DB백과)	산림청				○
		상세정보(임산물 DB백과)	산림청				○
		숲길 정보(지리산 둘레길 코스)	산림청			○	
		숲에 사는 식물 정보	산림청				○
		시군구별 관리소별 임산물 생산량(액)	산림청	○	○	○	○
		시도별 임산물 생산량(액)	산림청	○	○	○	○
		우리나라 주요산 등산로 정보	산림청			○	
		우리나라 주요산 등산로 포인트 정보	산림청			○	
		임도망도-서귀포시	산림청			○	
		임도망도-제주시	산림청			○	
		임상도-서귀포시	산림청			○	
		임상도-제주시	산림청			○	
		종류(임산물 DB백과)	산림청				○
		주요품목 주산지별 생산현황	산림청	○	○	○	○
		지리산 둘레길 안내	산림청			○	
		지방청별 임산물 생산량(액)	산림청	○	○	○	○
		품목(목재제품 DB백과)	산림청				○
		품목(임산물 DB백과)	산림청				○
		품목별 임산물 생산량(액)	산림청	○	○		○
	산림재해	산불 발생 위치도 정보	산림청				○
		산불 위험 예보	산림청				○
	산림통계	국립산림품종관리센터 생산수입판매신고	산림청	○	○	○	
		산불통계 서비스	산림청		○		

	연구정보	산림 ICT 서비스	산림청				○
		임업경제동향	산림청				○
	휴양문화	명산 등산로 정보	산림청				○
		산 정보	산림청				○
		산림치유정보(에코힐링)	산림청				○
		전통 마을 숲 위치도	산림청				○
		휴양림 고시 구역도	산림청				○
유통	관측전망	농업 관측 통계 정보	농촌경제연구원				○
		첨부파일 정보	농촌경제연구원	○	○		
		품목 정의	농촌경제연구원	○	○		
		품목관측 - 관측속보	농촌경제연구원	○	○		
		품목관측정보-월보	농촌경제연구원	○	○		
	소비동향	생산품목별원가구성	농림축산식품부	○	○		
		소비 관측	농촌경제연구원				○
		식품 산업별 원료 소비 실태	농림축산식품부			○	
		식품 소비 형태	농촌경제연구원				○
		원료별 국산 원료 사용 비중	농림축산식품부	○	○		
	유통가격	GPC 품목 매핑 목록 조회	농림수산식품교육문화정보원	○	○		
		국제 곡물 일별 선물 가격	농촌경제연구원	○	○		
		농수축산 유통정보 조사가격 (농수축산물 표준코드 변환)	농림수산식품교육문화정보원	○	○		
		농수축산물 도매가격	한국농수산식품유통공사	○	○	○	
		농수축산물 도매시장 코드	한국농수산식품유통공사	○	○	○	
		농수축산물 도매시장법인 코드	한국농수산식품유통공사	○	○	○	
		농수축산물 소매가격	한국농수산식품유통공사	○	○	○	
		농수축산물 조사가격 시장코드	한국농수산식품유통공사	○	○	○	
		농수축산물 조사가격 품목코드	한국농수산식품유통공사	○	○	○	
		농수축산물 조사가격정보	농림수산식품교육문화정보원	○	○	○	○
		농수축산물 친환경 가격	한국농수산식품유통공사	○	○	○	
		농축수산물 통합유통 정보제공 서비스	농림수산식품교육문화정보원		○		
		농수축산물 통합유통통계 정보제공 서비스	농림수산식품교육문화정보원		○		
		농수축산물 표준단위 코드	한국농수산식품유통공사	○	○	○	
		농수축산물 표준등급 코드	한국농수산식품유통공사	○	○	○	
		농수축산물 표준품목 코드	한국농수산식품유통공사	○	○	○	
		농협 공판장 경락가격	농림수산식품교육문화정보원	○	○	○	○
		농협 종합유통센터 매입가격	농림수산식품교육문화정보원	○	○		○
		도매시장별 주요 양채류 거래현황 (2013년01월~2016년05월)	농림수산식품교육문화정보원			○	
		돼지 검수(소비) 확인정보	축산물품질평가원	○	○		○
		서울 가락도매시장 조사가격	한국농수산식품유통공사	○	○		○
		서울 강서도매시장 조사가격	농림수산식품교육문화정보원	○	○		○
		소 검수(소비) 확인정보	축산물품질평가원	○	○		○
		수입농산물가격조회	농림수산식품교육문화정보원		○		
		수협 산지위판장 경락가격	농림수산식품교육문화정보원	○	○	○	○

		신구 표준단위 목록 조회	농림수산식품교육문화정보원	○	○		
		신구 표준등급 목록 조회	농림수산식품교육문화정보원	○	○		
		신구 표준산지 목록 조회	농림수산식품교육문화정보원	○	○		
		신구 표준시장 목록 조회	농림수산식품교육문화정보원	○	○		
		신구 표준크기 목록 조회	농림수산식품교육문화정보원	○	○		
		신구 표준포장 목록 조회	농림수산식품교육문화정보원	○	○		
		신구 표준품목 목록 조회	농림수산식품교육문화정보원	○	○		
		실시간 경매 속보 서비스	농림수산식품교육문화정보원		○		
		원천 산지 공판장 가격 목록 조회	농림수산식품교육문화정보원	○	○		
		원천 산지 위판장 가격 목록 조회	농림수산식품교육문화정보원	○	○		
		원천 실시간 경락가격 도매시장 목록 조회	농림수산식품교육문화정보원	○	○		
		원천 실시간 경락가격 도매시장 요약목록 조회	농림수산식품교육문화정보원	○	○		
		원천 실시간 경락가격 원시데이터	농림수산식품교육문화정보원				
		원천 실시간 경락가격 품목 목록 조회	농림수산식품교육문화정보원	○	○		
		원천 정산 경락가격 도매시장 목록 조회	농림수산식품교육문화정보원	○	○		
		원천 정산 경락가격 원시데이터	농림수산식품교육문화정보원				
		원천 정산경락가격 품목 목록 조회	농림수산식품교육문화정보원	○	○		
		원천 조사가격 목록 조회	농림수산식품교육문화정보원	○	○		
		임산물 조사가격	농림수산식품교육문화정보원	○	○	○	○
		전국 농협 하나로 종합유통센터 매입 가격 요약 정보	농림수산식품교육문화정보원	○	○		
		전국 농협산지공판장 경락가격 요약 정보 (산지공판장별 산지별)	농림수산식품교육문화정보원	○	○		
		전국 도매시장 실시간 경락가격 요약 정보 (도매시장별 산지별)	농림수산식품교육문화정보원	○	○		
		전국 도매시장 일별 경락가격	농림수산식품교육문화정보원	○	○	○	○
		전국 도매시장 일별 실시간 경락가격	한국농수산식품유통공사	○	○		○
		전국 도매시장 일별 실시간 경락가격 상세정보	농림수산식품교육문화정보원	○	○		
		전국 도매시장 일별 정산 경락가격 상세정보	농림수산식품교육문화정보원	○	○		
		전국 도매시장 정산 경락가격 요약 정보 (도매시장별 품목별 산지별)	농림수산식품교육문화정보원	○	○		
		중국 농산물 가격정보	농림수산식품교육문화정보원	○	○		
		지자체 농수축산물 물가조사가격 정보 제공	농림수산식품교육문화정보원		○		
		직거래 사업장 정보	한국농수산식품유통공사	○	○	○	
		축산물 경락가격	농림수산식품교육문화정보원	○	○	○	○
		축산물 부분육 경락가격	한국농수산식품유통공사	○	○	○	
		코드조회 서비스	농림수산식품교육문화정보원		○		
		표준단위매핑목록 조회	농림수산식품교육문화정보원	○	○		
		표준시장매핑목록 조회	농림수산식품교육문화정보원	○	○		
		표준코드국제코드 매핑 정보 서비스	농림수산식품교육문화정보원		○		
		표준품목매핑목록 조회	농림수산식품교육문화정보원	○	○		
		화훼류 시세현황	한국농수산식품유통공사	○	○	○	

	유통관리	APC 현황별 산지유통 관리시설 현황	농림축산식품부	○	○		
		가공용쌀 공급업체 단속현황	국립농산물품질관리원	○	○	○	
		농산물 도매시장 현황	농림축산식품부			○	
		농산물 브랜드 현황	농림축산식품부	○	○	○	
		농축산물 원산지 표시 품목별 적발현황	국립농산물품질관리원	○	○		
		농축산물 유통조사 현황	국립농산물품질관리원	○	○		
		로컬푸드 직매장 현황	농림축산식품부	○	○	○	
		비축 농산물 정보	한국농수산식품유통공사			○	
		산지유통 종합평가 결과	농림축산식품부	○	○		
		시도별 산지유통 관리시설 현황	농림축산식품부	○	○		
		시장 법인별 거래 실적	농림축산식품부	○	○		
		시장 연도별 거래 실적	농림축산식품부	○	○		
		운영주체별 산지유통 관리시설 현황	농림축산식품부	○	○		
		원산지표시 우수업체 현황	국립농산물품질관리원	○	○		
		정부지원 미곡종합처리장(RPC)현황	농림축산식품부	○	○	○	
		표준규격품 사후관리 실적통계	국립농산물품질관리원	○	○	○	
		학교급식전자조달시스템 계약정보	한국농수산식품유통공사	○	○		
	재해	년도별 농작물 보험가입 통계정보	농림축산식품부	○	○		
		년도별 농작물 보험금지급 통계정보	농림축산식품부	○	○		
		년도별 농작물보험료 통계정보	농림축산식품부	○	○		
		농업 재해 보험 연감	농림축산식품부			○	
		농작물 보험금지급현황	농업정책보험금융원				○
		시도별 농작물 보험가입 통계정보	농림축산식품부	○	○		
		시도별 농작물 보험금지급 통계정보	농림축산식품부	○	○		
		시도별 농작물보험료 통계정보	농림축산식품부	○	○		
		시도별 사고원인 가축보험금지급 통계정보	농림축산식품부	○	○		
		축종별 가축보험가입 통계정보	농림축산식품부	○	○		
		축종별 가축보험금지급 통계정보	농림축산식품부	○	○		
		품목별 농작물 보험가입 통계정보	농림축산식품부	○	○		
		품목별 농작물 보험금지급 통계정보	농림축산식품부	○	○		
		품목별 농작물보험료 통계정보	농림축산식품부	○	○		
농지 용수	농업용수	농업용 저수지 수위자료	한국농어촌공사		○		
		농업진흥지역도	한국농어촌공사			○	
		농지보전부담금 통계자료	한국농어촌공사	○	○		
		농지전용 허가(협의-신고)현황	한국농어촌공사	○	○	○	
		발전소 발전량 데이터	한국농어촌공사	○	○		
		새만금방조제 수위계측정보	한국농어촌공사	○	○	○	
		영농여건불리농지도	한국농어촌공사			○	
		조성토지관리 통계정보	한국농어촌공사	○	○		
		토양검정 화학성 분석 자료 서비스	농촌진흥청		○		
	농지	경영이양직접지불제 통계자료	한국농어촌공사	○	○		
		경영회생 지역별 추진실적	한국농어촌공사	○	○		
		과원규모화 사업추진실적	한국농어촌공사	○	○		
		농지 실거래가 동향분석결과	농림축산식품부	○	○	○	
		농지규모화 사업추진 실적	한국농어촌공사	○	○		
		담보농지 분석	한국농어촌공사	○	○		
		매입비축 도별 임대현황	한국농어촌공사	○	○		
		임대수탁 농지분석	한국농어촌공사	○	○		
		지급방식별 가입분석	한국농어촌공사	○	○		
		토양개량제 지원현황	농림축산식품부	○	○	○	
		환지 정보	한국농어촌공사			○	

시설 장비	시설장비	농기계안전정보 서비스	농촌진흥청				○
		농업기반시설 시설제원-관정	한국농어촌공사	○	○		
		농업기반시설 시설제원-방조제	한국농어촌공사	○	○		
		농업기반시설 시설제원-양배수장	한국농어촌공사	○	○		
		농업기반시설 시설제원-저수지	한국농어촌공사	○	○		
		농업기반시설 시설제원-집수암거	한국농어촌공사	○	○		
		농업기반시설 시설제원-취입보	한국농어촌공사	○	○		
		농작물시설표준 설계도	농촌진흥청			○	
교육	교육	(강의명)갈등과 협상	농식품공무원교육원				○
		(강의명)농식품 브랜드 마케팅 성공전략	농식품공무원교육원				○
		(강의명)미래 농식품트랜드 변화와 농식품산업의 미래	농식품공무원교육원				○
		(강의명)미래지향적 농업분야	농식품공무원교육원				○
		(강의명)스토리텔링과 스킬	농식품공무원교육원				○
		(강의명)지역활성화와 농촌관광 동영상 강의	농식품공무원교육원				○
		교육 과정 분류	농림수산식품교육문화정보원	○	○	○	○
		교육 실적 통계	농림수산식품교육문화정보원	○	○		○
		농식품공무원교육원 교육정보	농식품공무원교육원	○	○	○	
		농식품공무원교육원 시설대여 신청정보	농식품공무원교육원	○	○	○	
		농식품공무원교육원 안전관리매뉴얼	농식품공무원교육원			○	
		농식품유통교육원 연간교육일정	한국농수산식품유통공사	○	○	○	
		농업마이스터대학 현황	농림수산식품교육문화정보원	○	○	○	
		농촌교육농장30선 (2015년 품질인증농장)	농촌진흥청			○	
		만화로 보는 농업정책 햇살이네	농림축산식품부			○	
		산림교육_프로그램_운영_정보	산림청		○		
		식생활 교육기관 현황	농림축산식품부			○	
		약초의 효능	농식품공무원교육원				○
		오프라인 교육 과정	농림수산식품교육문화정보원	○	○	○	
		온라인 교육 과정	농림수산식품교육문화정보원	○	○	○	○
		졸업생 영농종사 현황	한국농수산대학			○	
		직무교육과정 현장학습 정보	농식품공무원교육원	○	○	○	
농촌 복지	농촌개발	경관보전직불 사업지구	농림축산식품부	○	○	○	
		과수거점 산지유통센터 건립지원사업 상세정보	농림축산식품부	○	○		
		과실전문 생산단지 기반조성사업 결과	농림축산식품부	○	○		
		과실전문 생산단지 기반조성사업 단지기반 정보	농림축산식품부	○	○		
		농산어촌지역 개발공간정보	농림수산식품교육문화정보원		○		
		농촌 마을 현황 정보	농림수산식품교육문화정보원	○	○		
		농촌마을별 기본정보	농림수산식품교육문화정보원	○	○		
		농촌마을별 연계자원 정보	농림수산식품교육문화정보원	○	○		
		농촌마을별 인원정보	농림수산식품교육문화정보원	○	○		
		농촌마을별 자원 및 조직정보	농림수산식품교육문화정보원	○	○		
		에너지효율화(목재펠릿, 에너지절감)	농림축산식품부	○	○		
		일반농산어촌지역 개발사업 정보	농림수산식품교육문화정보원	○	○	○	○
		일반농산어촌지역 개발사업 통계	농림수산식품교육문화정보원	○	○	○	○
		전원마을 분양 공고정보	농림수산식품교육문화정보원	○	○		
		지역별 과수시설 현대화 지원실적 현황	농림축산식품부	○	○		
		행복마을 만들기 콘테스트 실적	농림수산식품교육문화정보원	○	○		
		행복마을 만들기 선정마을 정보	농림수산식품교육문화정보원	○	○		

	농촌관광	객실 정보	한국농어촌공사	○	○	○	
		낙농체험 목장 정보	농림축산식품부	○	○	○	○
		농가 맛집 정보	한국농어촌공사	○	○		
		농어촌 체험 마을	한국농어촌공사	○	○		
		농촌관광사업 등급정보 현황	농림축산식품부	○	○	○	
		농촌체험휴양마을 으뜸촌 현황	농림축산식품부	○	○	○	
		동네맛집 정보	한국농어촌공사	○	○	○	
		마을 특산물 정보	한국농어촌공사	○	○		
		마을동영상	한국농어촌공사	○	○		
		마을축제 정보	한국농어촌공사	○	○	○	
		모두 투어 연계 상품	한국농어촌공사	○	○		
		식생활교육 우수농어촌 체험공간 지정 현황	농림축산식품부	○	○	○	
		여행작가 추천 여행	한국농어촌공사	○	○		
		이달의 추천여행	한국농어촌공사	○	○		
		주요시설 정보	한국농어촌공사	○	○	○	
		지역축제	한국농어촌공사	○	○	○	
		체험 마을 숙박 정보	한국농어촌공사	○	○		
		체험 마을 주변 숙박	한국농어촌공사	○	○	○	
		체험 마을 탐방	한국농어촌공사	○	○		
		체험 프로그램	한국농어촌공사	○	○		
		체험식단 정보	한국농어촌공사	○	○	○	
		캠핑정보	한국농어촌공사	○	○	○	
		포토 및 후기정보	한국농어촌공사	○	○		
		학교 장터 연계 상품	한국농어촌공사	○	○		
	도시농업	실내정원용 식물 목록 및 상세 정보	농촌진흥청		○		
		웰빙도시농업 정보	농림수산식품교육문화정보원			○	
		칼로리 정보	농림수산식품교육문화정보원			○	
말산업	말산업	경주거리별 최고기록	한국마사회		○		
		경주계획	한국마사회		○		
		경주마 등록 내역	한국마사회	○	○		○
		경주마 현황정보	한국마사회		○		
		경주마등급변동현황	한국마사회		○		
		경주마명 변경이력	한국마사회		○		
		경주마성적정보	한국마사회		○		
		경주마정보	한국마사회	○	○		
		경주마정보 (경주결과)	한국마사회	○	○		
		경주마정보 (주행심사결과)	한국마사회	○	○		
		경주마정보 (출전정보)	한국마사회	○	○		
		경주마진료현황	한국마사회		○		
		경주마출전취소현황	한국마사회		○		
		경주별 마권매출 순위정보	한국마사회		○		
		경주요약성적표	한국마사회		○		
		경주정보(경주개요, 상금, 승식별 매출액)	한국마사회		○		
		경주정보(복승식확정배당률)	한국마사회		○		
		경주정보(상세구간통과기록, 펄롱타임)	한국마사회		○		
		경주정보(심판사항)	한국마사회		○		
		경주정보(출전마 정보, 경주기록)	한국마사회		○		
		기수 정보	한국마사회	○	○		
		기수변경현황	한국마사회		○		
		기수성적정보	한국마사회		○		
		기수정보	한국마사회		○		
		대상경주 연간계획	한국마사회		○		
		렛츠런 CCC 문화센터 강좌정보	한국마사회		○		

		마카드 게좌고객 조건별 적중율 순위 Top 20	한국마사회		○		
		마카드 게좌고객 조건별 환수율 순위 Top 20	한국마사회		○		
		마주정보	한국마사회		○		
		마필, 기수, 조교사통산 경주기록	한국마사회		○		
		마필구간별 최고/최저/평균기록	한국마사회		○		
		마필종합정보(서울, 부산)	한국마사회		○		
		마필질병정보	한국마사회		○		
		말 방역 현황	한국마사회	○	○		
		사업장별 입장인원	한국마사회		○		
		승마장 정보	한국마사회	○	○	○	
		승식별 최고배당률정보	한국마사회		○		
		승용마 및 제주산 등록 내역	한국마사회	○	○		○
		승용마정보	한국마사회		○		
		예상배당률정보(단승식, 연승식)	한국마사회		○		
		예상배당률정보(복승식, 쌍승식, 복연승식)	한국마사회		○		
		예상배당률정보(삼복승식, 삼쌍승식)	한국마사회		○		
		일별훈련현황	한국마사회		○		
		전국승마장 정보	한국마사회		○		
		조교사 정보	한국마사회	○	○		
		조교사 정보	한국마사회		○		
		주행심사결과	한국마사회		○		
		출발심사결과	한국마사회		○		
		출발훈련결과	한국마사회		○		
		출전등록경주마 현황	한국마사회		○		
		출전마장구 사용현황 및 폐출혈 이력정보	한국마사회		○		
		출전마체중정보	한국마사회		○		
		출전표	한국마사회		○		
		훈련 정보	한국마사회	○	○		
농축산 기술	GIS	과수재배변동정보 조회 서비스	농촌진흥청		○		
	IoT	스마트팜 데이터 조회 서비스	농림수산식품교육문화정보원		○		
	R&D	IBT융합기술 R&D 과제정보	농림수산식품기술기획평가원	○	○	○	
		IBT융합기술 R&D 논문정보	농림수산식품기술기획평가원	○	○	○	
		IBT융합기술 R&D 특허정보	농림수산식품기술기획평가원	○	○	○	
		농림축산검역검사기술개발 연구성과 정보	농림축산검역본부	○	○	○	
		문화 R&D 과제정보	농림수산식품기술기획평가원	○	○	○	
		문화 R&D 논문정보	농림수산식품기술기획평가원	○	○	○	
		문화 R&D 특허정보	농림수산식품기술기획평가원	○	○	○	
		바이오 R&D 과제정보	농림수산식품기술기획평가원	○	○	○	
		바이오 R&D 논문정보	농림수산식품기술기획평가원	○	○	○	
		바이오 R&D 특허정보	농림수산식품기술기획평가원	○	○	○	
		산림 과학 연구정보	산림청				○
		생산·가공 R&D 과제정보	농림수산식품기술기획평가원	○	○	○	
		생산·가공 R&D 논문정보	농림수산식품기술기획평가원	○	○	○	
		생산·가공 R&D 특허정보	농림수산식품기술기획평가원	○	○	○	
		생산시스템 R&D 과제정보	농림수산식품기술기획평가원	○	○	○	
		생산시스템 R&D 논문정보	농림수산식품기술기획평가원	○	○	○	
		생산시스템 R&D 특허정보	농림수산식품기술기획평가원	○	○	○	
		식물검역연구장비 보유현황	농림축산검역본부	○	○	○	
		식품·유통 R&D 과제정보	농림수산식품기술기획평가원	○	○	○	
		식품·유통 R&D 논문정보	농림수산식품기술기획평가원	○	○	○	
		식품·유통 R&D 특허정보	농림수산식품기술기획평가원	○	○	○	
		자원·환경·생태기반 R&D 과제정보	농림수산식품기술기획평가원	○	○	○	

		자원·환경·생태기반 R&D 논문정보	농림수산식품기술기획평가원	○	○	○	
		자원·환경·생태기반 R&D 특허정보	농림수산식품기술기획평가원	○	○	○	
		작목별 농업기술 동영상 정보	농촌진흥청		○		
	가축질병	가축 질병 발생 정보	농림축산검역본부	○	○		
		동물 질병 정보	농림축산검역본부	○	○		
		소 브루셀라 검사결과 현황	농림축산식품부	○	○		
		소 브루셀라 이동제한 정보	농림축산식품부	○	○		
	경영	경영주 경영체 수 현황	농림축산식품부	○	○		
		농업 경영 컨설팅지원사업	농림축산식품부			○	
		농업 경제 동향	농촌경제연구원				○
		농업법인 경영체 수 현황	농림축산식품부	○	○		
		농축산물 소득상세정보 서비스	농촌진흥청		○		
		농축산물 소득정보	농촌진흥청		○		
		사단법인	농림축산식품부	○	○	○	
		재단법인	농림축산식품부	○	○	○	
		전체 노지작물 수확(재배) 경영체 수 현황	농림축산식품부	○	○		
		전체 축종별 경영체 수 현황	농림축산식품부	○	○		
		컨설팅 분야별 지원현황	농림축산식품부	○	○	○	
		품목별농산물소득정보	농촌진흥청				○
	농자재	녹비작물 종자 지역별 공급현황	농림축산식품부	○	○		
		농기계별 면세유 공급현황	농림축산식품부	○	○	○	
		농업기계 보유 현황	농림축산식품부	○	○	○	
		식물목록별 참고문헌 정보서비스	농촌진흥청		○		
		위반 정보 (면세유 사후관리)	농림축산식품부	○	○	○	
		유기농업자재 현황	농림축산식품부			○	
		재배품목별 면세유 공급현황	농림축산식품부	○	○	○	
		지역별 면세유 공급현황	농림축산식품부	○	○	○	
	병해충	국가농작물 품목별 병해충 도감정보 서비스	농촌진흥청		○		
		국가재난병 감염황룡병 방제 약제 등록 현황	농림축산검역본부	○	○	○	
		농림축산검역본부 식물병해충 소장 표본 현황	농림축산검역본부	○	○	○	
		병해충 예측 정보	농촌진흥청				○
		수출입식물 품목별 농약 등록 현황	농림축산검역본부	○	○	○	
		식물검역 병해충 정보	농림축산검역본부	○	○		
		작물 병해충 검색 서비스	농촌진흥청		○		
	생명자원	가축인공 수정사 면허취득 현황	농림축산식품부	○	○	○	
		곤충자원 정보	농림수산식품교육문화정보원	○	○		○
		국가 표준 식물 목록_정명	산림청	○	○	○	○
		국가 표준식물 목록_이명	산림청	○	○	○	○
		국가 표준 곤충목록 서비스	산림청		○		
		국립수목원 곤충표본	산림청				○
		국립수목원 귀화식물	산림청				○
		국립수목원 균류 도감	산림청				○
		국립수목원 식물도감	산림청				○
		누에자원 정보	농림수산식품교육문화정보원	○	○		○
		동물자원 정보	농림수산식품교육문화정보원	○	○		○
		미생물자원 정보	농림수산식품교육문화정보원	○	○		○
		민속식물 정보	산림청				○
		수산자원 정보	농림수산식품교육문화정보원	○	○		○
		수의유전자원 정보	농림축산검역본부	○	○		
		식물 세밀화 정보	산림청				○
		식물 표본 정보	산림청				○

		식물자원 정보	농림수산식품교육문화정보원	○	○		○
		유전체자원 정보	농림수산식품교육문화정보원	○	○		○
		인공수정소 현황	농림축산식품부	○	○	○	
		종자 정보	산림청				○
		지의류 자원 서비스	산림청		○		
		지의류 표본	산림청				○
		품목별 유전자원 정보	농촌진흥청		○		
	영농	귀농귀촌 지능형 상담	농림수산식품교육문화정보원		○		
		년도별 품목별 영농활용 기술	농촌진흥청				○
		농식품 뉴스	농림수산식품교육문화정보원				○
		농식품백과사전	농림수산식품교육문화정보원			○	
		농업 수입지출 통계정보(총괄통계)	농림수산식품교육문화정보원	○	○	○	
		농업농촌 종합정보(인테러뱅) 서비스	농촌진흥청			○	○
		농업용어 검색 및 설명 정보	농촌진흥청		○		
		농작업 일정 정보	농촌진흥청			○	
		분류별 정보 목록	농촌진흥청				○
		여성 농업담당 비중	농림축산식품부	○	○		
		여성 농업인 실태 조사	농림축산식품부			○	
		여성농업 외 소득 활동	농림축산식품부	○	○		
		여성시급해결과제	농림축산식품부	○	○		
		우리말 농업용어	농림축산식품부	○	○	○	
		작목 및 품목 정보	농촌진흥청				○
		작목별 기술 목록	농촌진흥청				○
		작목별 기술상세정보	농촌진흥청				○
		작목별 기술정보 분류	농촌진흥청				○
		작목별 영농기술 및 농사기술 정보	농촌진흥청				○
		작물 및 품목 분류 목록	농촌진흥청				○
		주간 농업정보 및 품목별 농사정보	농촌진흥청		○		
		주간농사정보 목록	농촌진흥청				○
		주간농사정보 상세정보	농촌진흥청				○
		지능형 검색	농림수산식품교육문화정보원		○		
		품목별 관리메뉴얼 상세정보	농촌진흥청				○
		품목별 관리메뉴얼 목록	농촌진흥청				○
	종자	거베라 품종보호 화보집	국립종자원			○	
		국립산림품종관리센터 품종보호 출원현황	산림청	○	○	○	
		국화 품종보호 화보집	국립종자원			○	
		농산물 품종 정보	농촌진흥청				○
		대한민국 우수품종상 통계	국립종자원			○	
		민간 지자체 생산종자 검사결과	국립종자원	○	○	○	
		분쟁종자시험분석데이터	국립종자원	○	○	○	
		소관작물 현황	국립종자원	○	○		○
		시도별 종자업 등록현황	국립종자원	○	○		○
		식물 신품종 품종보호권 소멸공고 자료	국립종자원	○	○		○
		연도별 종자 수출입 물량 및 국내외 채소종자 생산량 현황	국립종자원			○	
		일본의 품종보호 등록 품종현황	국립종자원			○	○
		전국 일반 종자관리사 및 버섯 종자관리사 등록 현황	국립종자원				○
		정부 보급종 생산공급 및 보증검사 통계	국립종자원			○	
		정부 보급종 품종 안내	국립종자원			○	

		정부 보급종 생산 통계	국립종자원	○	○		
		정부 보급종 생산공급 통계	국립종자원	○	○	○	
		정부생산 보급종자 검사결과	국립종자원	○	○	○	
		종자 유통조사 통계	국립종자원			○	
		종자 재배시험 및 종자검정 통계	국립종자원			○	
		종자 품종보호 분야 통계	국립종자원			○	
		종자수입생산판매 신고 정보	국립종자원	○	○	○	
		종자수출입현황	국립 종자원	○	○		
		종자품종 특성정보	국립종자원	○	○		
		채소종자 국내외 생산량 현황	국립종자원	○	○		
		채소종자 수입적응성 시험 목록	국립종자원			○	○
		품종보호권 출원/등록 정보	국립종자원	○	○	○	
		품종정보 목록	농촌진흥청				○
	허가정보	동물의약품 허가결과정보	농림축산검역본부	○	○	○	
검역	검사검수	선박 AGM 검사 통계	국제식물검역인증원	○	○	○	
		선박 AGM 검사 현황	국제식물검역인증원	○	○	○	
		수입동물 검역 불합격 현황	농림축산검역본부	○	○	○	
		수입동물 검역 합격현황	농림축산검역본부	○	○	○	
		수입식물 검역 정보	농림축산검역본부	○	○		○
		수입재 식용식물 검역장소 관리 통계 현황	국제식물검역인증원	○	○	○	
		수입재 식용식물 검역장소 관리 현황	국제식물검역인증원	○	○	○	
		수입축산물 검역 불합격 현황	농림축산검역본부	○	○	○	
		수입축산물 검역 합격 현황	농림축산검역본부	○	○	○	○
		수출동물 검역현황	농림축산검역본부	○	○	○	
		수출식물 검역 정보	농림축산검역본부	○	○		○
		수출축산물 검역 현황	농림축산검역본부	○	○	○	
	예찰방역	구제역 예찰 및 진단	농림축산검역본부			○	
		수출입식물방제업체운영정보	농림축산검역본부	○	○	○	
		열처리 업체 운영정보	농림축산검역본부	○	○	○	
식품 산업	6차산업	6차산업 인증 사업자 현황	농림축산식품부	○	○	○	
	가공식품	과실류 가공내역	농림축산식품부	○	○	○	
		식품 명인 현황	농림축산식품부	○	○	○	
		연도별 가공업체 현황	농림축산식품부	○	○	○	
		전통식품 산업실태 조사	농림축산식품부			○	
	수출입	농수축산물 수출입통계 조회	한국농수산식품유통공사				○
		수출국가별 식품첨가물 및 유해물질 정보	한국농수산식품유통공사	○	○		
	식품기업	네덜란드 푸드밸리 현황	농림축산식품부			○	
		달 평균 음용 횟수	농림축산식품부	○	○		
		도별 생산 주종	농림축산식품부	○	○		
		식품산업 일자리(채용) 정보	농림수산식품교육문화정보원	○	○		
		우수외식업지구 식당현황	농림축산식품부	○	○	○	
		주류 시장 현황	농림축산식품부	○	○		
		주류산업 정보 실태조사	농림축산식품부			○	
	외식산업	글로벌 외식 및 한식산업 목록정보	한식재단	○	○	○	○
		레시피 과정정보	농림수산식품교육문화정보원	○	○		
		레시피 기본정보	농림수산식품교육문화정보원	○	○		
		레시피 재료정보	농림수산식품교육문화정보원	○	○		
		외식산업 지식재산권 최종보고서	농림축산식품부			○	
		전국식재료 산지정보	한국농수산식품유통공사	○	○	○	
	통계	국내 외식기업 해외진출 현황	농림축산식품부	○	○	○	
		농림업 생산지수	농림축산식품부	○	○	○	

		동행 응답자 특성	농림축산식품부	○	○		
		외식업 경기지수	농림축산식품부	○	○		
		외식업 경기지수 규모별	농림축산식품부	○	○		
		외식업 경기지수 지역별 정보	농림축산식품부	○	○		
		외식업 트렌드 조사	농림축산식품부			○	
		음식점 인지경로	농림축산식품부	○	○		
		인삼 연도별 생산량	농림축산식품부	○	○		
		인삼 통계 자료	농림축산식품부			○	
		재배 면적별 농가수	농림축산식품부	○	○		
		한국 외식업 경기지수	농림축산식품부			○	
		한식 좌석별 외식업 경기지수	농림축산식품부	○	○		
		한식분기별 외식업 경기지수	농림축산식품부	○	○		
		한식지역별 외식업 경기지수	농림축산식품부	○	○		
식품 안전	검사	농산물 독소류 분석결과	국립농산물품질관리원	○	○	○	
		농산물 방사능 검사현황	국립농산물품질관리원	○	○	○	
		농산물 잔류농약 분석결과	국립농산물품질관리원	○	○	○	
		농산물 중금속 분석결과	국립농산물품질관리원	○	○	○	
		닭 품종별 도축실적	농림축산검역본부	○	○		
		도축장별 축종별 도축실적	농림축산검역본부	○	○		
		소 품종별 암수별 도축실적	농림축산검역본부	○	○		
		수입농산물 검사 통계정보	국립농산물품질관리원	○	○	○	
		수입농산물검사결과	국립농산물품질관리원	○	○	○	
		시도별 도축실적	농림축산검역본부	○	○		
		연간 정곡검사 통계	국립농산물품질관리원	○	○	○	
		축종별 생체중량	농림축산검역본부	○	○		
		축종별 암수별 도축실적	농림축산검역본부	○	○		
		친환경 인증 품목별 현황	국립농산물품질관리원	○	○		
		친환경 인증정보	국립농산물품질관리원				○
		하추곡 매입검사 실적	국립농산물품질관리원	○	○	○	
	이력추적	농산물 이력추적 등록정보	국립농산물품질관리원	○	○	○	
		돼지 묶음번호 구성내역 정보	축산물품질평가원	○	○		
		돼지 이력 정보	축산물품질평가원	○	○		
		돼지 포장처리 정보	축산물품질평가원	○	○		
		소 및 쇠고기 개체정보	축산물품질평가원	○	○		○
		소 이동정보 조회	축산물품질평가원		○		
		원산지 정보, 유통이력 정보, 위해 쇠고기 여부 조회	농림축산검역본부	○	○		○
		축산물DNA동일성검사 업소정보	축산물품질평가원		○		
		축산물이력정보조회	축산물품질평가원		○		
품질 관리	등급	공통 코드표	축산물품질평가원	○	○	○	
		돼지 등급판정 부위별 고기 및 부산물 정보	축산물품질평가원	○	○		○
		돼지 등급판정결과(확인서) 정보	축산물품질평가원	○	○		○
		돼지 박피, 탕박 권역별 경락가격 정보	축산물품질평가원	○	○		○
		소 등급판정(확인서) 부위별 고기 및 부산물 정보	축산물품질평가원	○	○		○
		소 등급판정결과(확인서) 정보	축산물품질평가원	○	○		○
		소 한우 육우 권역별 경락가격 정보	축산물품질평가원	○	○		○
		축산물 도매시장 소도체 등급별 경락가격 정보	축산물품질평가원	○	○		○
		축산물 돼지도체 경락가격 상세정보	축산물품질평가원	○	○		○
		축산물 돼지도체 등급별 경락가격 정보	축산물품질평가원	○	○		○
		축산물 소도체 경락가격 상세정보	축산물품질평가원	○	○		○
		축산물 소도체 육질등급별 경락가격 정보	축산물품질평가원	○	○		○
		축산물 실시간 경매시황 중계정보	축산물품질평가원	○	○		○

		축산물 실시간 도매시장 정보	축산물품질평가원	○	○		○
		축산물 실시간 대지도체 경매 세부현황 정보	축산물품질평가원	○	○		○
		축산물 실시간 대지도체 등급별 경매현황 정보	축산물품질평가원	○	○		○
		축산물 실시간 소도체 경매세부현황 정보	축산물품질평가원	○	○		○
		축산물 실시간 소도체 등급별 경매현황 정보	축산물품질평가원	○	○		○
		축산물등급판정확인서 발급번호 정보	축산물품질평가원		○		
	인증	GAP 인증농가 현황	국립농산물품질관리원	○	○	○	
		농림축산분야 녹색인증현황	농림축산식품부			○	
		농산물 안전성 검정기관 현황	농림축산식품부			○	
		농산물검정 현황	국립농산물품질관리원	○	○		
		술 품질 인증 업체 정보	국립농산물품질관리원	○	○	○	
		우수 식품 인증 정보	국립농산물품질관리원	○	○	○	
		우수농산물(GAP) 관리시설 정보	국립농산물품질관리원	○	○	○	
		우수농산물(GAP) 인증기관 정보	국립농산물품질관리원	○	○	○	
		우수농산물(GAP) 인증정보	국립농산물품질관리원	○	○	○	
		유기가공식품 인증현황	농림축산식품부	○	○	○	
		지리적 표시 관리 정보	국립농산물품질관리원	○	○	○	
		축산물 HACCP 컨설팅 등록업체 현황	농림축산식품부			○	
		친환경 농업단지 현황	농림축산식품부			○	
		친환경농산물 인증기관 현황	농림축산식품부			○	
		친환경농산물 인증현황	농림축산식품부			○	
농업 정책	국내정책	밭 농업 직불제	농림축산식품부	○	○	○	
		쌀소득등보전직불제 지급현황_고정	농림축산식품부	○	○	○	
		쌀소득등보전직불제 지급현황_변동	농림축산식품부	○	○	○	
		조건불리지역 직불 지급현황	농림축산식품부	○	○	○	
		조건불리지역 직접지불제 지원현황	농림축산식품부	○	○	○	
		친환경농업 직불제(인증종류별)	농림축산식품부	○	○	○	
		친환경농업직불제(지역별)	농림축산식품부	○	○	○	
		한우, 송아지 피해보전직불 지원현황	농림축산식품부	○	○	○	
		후계농업경영인 선정현황	농림축산식품부	○	○		
	국외정책	년도별 FTA협정세율 현황	농림축산식품부	○	○		
		농축산물 FTA협정세율 현황	농림축산식품부			○	
		농축산물 관세율 현황	농림축산식품부	○	○	○	
		농축산물 시장접근물량 현황	농림축산식품부	○	○	○	
		농축산물 특별긴급관세(SSG) 품목현황	농림축산식품부	○	○	○	
귀농 귀촌	귀농귀촌	귀농귀촌인 통계	농림축산식품부			○	
		귀농귀촌 성공의 지름길	농식품공무원교육원				○
		귀농귀촌 지원정책	농림수산식품교육문화정보원				○
		귀농주가구원성별연령별	농림축산식품부	○	○		
		농촌 빈집 정보	농림수산식품교육문화정보원	○	○		
		작물재배 가구 농지 임차	농림축산식품부	○	○		
		재배 작물별 귀농 가구수	농림축산식품부	○	○		

○ 기준데이터 20개 이상으로 데이터를 제공한 기관은 농림축산식품부, 국립농산물품질관리원, 국립종자원, 농림축산검역본부, 한국농어촌공사, 한국농수산식품유통공사, 농림수산식품교육문화정보원, 농림수산식품기술기획평가원, 한국마사회, 농촌진흥청, 산림청, 농촌경제연구원임을 확인할 수 있음

나. 데이터 가치평가 방법 조사

(1) 가치평가 방법

○ 일반적으로 자산의 가치평가 방법은 시장 접근법, 원가 접근법, 소득 접근법이 있음.
○ 유가증권, 부동산 등의 유형자산뿐만 아니라 소프트웨어나 데이터베이스 등의 무형 자산의 가치를 평가하는데 활용되어 왔음.

<시장, 원가, 소득 접근법>

(2) 시장 기반 접근법

○ 시장 기반 접근법은 일반적으로 시장에서 동일 또는 유사한 자산이 얼마의 가격에 거래되었는지 확인하는 것. 해당 자산이 활발하게 거래되는 거래소가 있다면 그 거래소의 가격을 기준으로 하고, 그렇지 않다면 유사한 거래 사례를 찾는 방법을 사용
○ 미국의 경우 여러 출처로 얻은 데이터를 수집·취합하여 판매하는 이른바 '데이터 브로커'가 매우 활성화되어 있음
○ 데이터 브로커들은 소비자가 온라인에서 게시한 정보, 온·오프라인 쇼핑 정보, 웹사이트 등록정보 등 다양한 출처로부터 데이터를 수집·가공하여 판매하는 사업을 수행

(3) 원가 기반 접근법

○ 원가 기반 접근법은 데이터를 생성하기 위해 소요된 비용을 기준으로 데이터의 가치를 산정하는 방법
○ 비용 접근법은 크게 두 가지로 나뉘며 평가 대상 자산을 자체적으로 다시 생산 또는 재현해 내는 현재 비용을 기준을 하는 방법(재생산 비용)과 평가 대상 자산을 그대로 재생산 또는 재현할 수 없는 경우, 동등한 효용을 갖는 유사한 자산을 생산하는 비용을 기준으로 하는 방법(대체비용)으로 구분

○ 해당 데이터베이스를 직접 구축하기 위해 현재 소요되는 비용을 기준으로 그 데이터의 가치를 평가하는 방식, 하지만 원가 기반 접근법은 해당 데이터를 통해 장래에 얼마나 되는 경제적 가치를 얻을 수 있을 것인지를 고려하지 못한다는 한계가 존재.

(4) 소득 기반 접근법

○ 소득 기반 접근법은 데이터를 활용함으로써 얻을 수 있는 미래의 소득이 얼마나 될 것인지 추정한 다음 이를 현재가치로 할인한 금액을 데이터의 가치로 평가하는 방법.
○ 소득 기반 접근법은 평가 대상 데이터의 유용성이 유지되는 수명 기간 동안의 미래의 매출 증가 혹은 비용 절감 효과를 추정하고, 데이터 활용에 드는 비용을 차감함으로써 데이터 활용에 따라 얻어질 수 있는 현금 흐름을 추정. 여기에 적절한 할인율을 적용하여 현재가치로 할인한 다음 이를 합산한 것이 해당 데이터의 가치
○ 소득 기반 접근법은 이론적으로는 정교한 방법이지만 미래의 현금흐름을 예측하기 위해 많은 추정을 해야 한다는 어려움이 있음.
○ 또한 평가자의 주관적 요소가 작용할 여지가 크며 현재가치는 할인율에 따라 민감하게 변화하므로, 다른 할인율을 적용하면 데이터의 현재가치 액수가 상당히 달라질 수 있다는 문제점이 존재.

(5) 데이터 가치평가

데이터 가치평가	주요 내용
컨텐츠 적합성	• 데이터가 내포한 정보, 데이터의 컨텐츠에 대한 가치 • 동일한 주제라도 좀 더 다양하고 깊이 있는 정보를 제공하는 데이터의 활용가치가 높음
품질・공급 신뢰성	• 데이터의 물리적인 품질과 데이터 공급에 대한 가치 • 데이터의 오류가 적고 향후에도 지속적으로 공급되는 데이터의 가치가 높음
기술적 사용성	• 구매한 데이터 이용 시 예상되는 기술적 제약사항에 대한 가치 • 이용이 편리하고 특정 소프트웨어에 의존적이지 않은 데이터의 가치를 높게 평가
경제성	• 구매한 데이터 활용의 제도적 제약사항과 대안 가능성에 대한 가치 • 활용의 제약이 없고 대안이 적을수록 가치를 높게 평가

(6) 데이터 가치평가지표

데이터 가치평가	가치평가요소	가치평가지표
컨텐츠 적합성	데이터 정보의 다양성	의미 컬럼수
		정보 범주수
	타 데이터와의 연계성	표준연계키 개수
	데이터 명세의 명확성	절대척도컬럼 비중
		연속형컬럼명의 표기명확성
		범주형컬럼값의 기준명확성
	데이터의 적시성	데이터의생성일의 최신성
		데이터 제공주기의 규칙성
		데이터 집계주기의 규칙성
품질·공급 신뢰성	데이터의 물리적 품질	데이터의 품질(적합성율)
	데이터 건수의 충분성	유효데이터건수
	데이터 공급의 지속성	공급기간
		공급건수
		공급최신성
	데이터 공급자의 신뢰도	데이터 수령방식
		확장자 개방성 단계
기술적 사용성	데이터 수령방식의 편리성	데이터 수령방식
	확장자의 개방성	확장자 개방성 단계
경제성	제약성	라이선스 유형
	대체가능성	동종 카테고리 상품 비중

(7) 데이터 가치의 평가 체계

○ 윤리기준

- 윤리기준의 목적은 평가자가 전문가로서 평가 업무를 수행할 때 준수하여야 할 기준으로 평가의 공정성 및 사회적 신뢰의 제고에 기여하고자 함.
- 평가자는 평가의 공공성을 충분히 이해하고 전문가로서의 사회적 역할과 책임을 인식하는 채무를 갖음.
- 평가자는 평가 업무를 수행함에 있어 다음의 윤리원칙을 준수해야 함.
  - 평가자는 객관성과 성실성을 유지하여야 하며, 고의로 사실을 왜곡하거나 정당한 사유가 없는 한 다른 사람에게 평가의 대리를 의뢰해서는 안됨.
  - 평가자는 평가 업무와 관련하여 일체의 외부 간섭으로부터 독립되어야 함.
  - 평가자는 의뢰자의 사전허락 없이는 제3자에게 평가 관련 자료를 제공해서는 안됨.

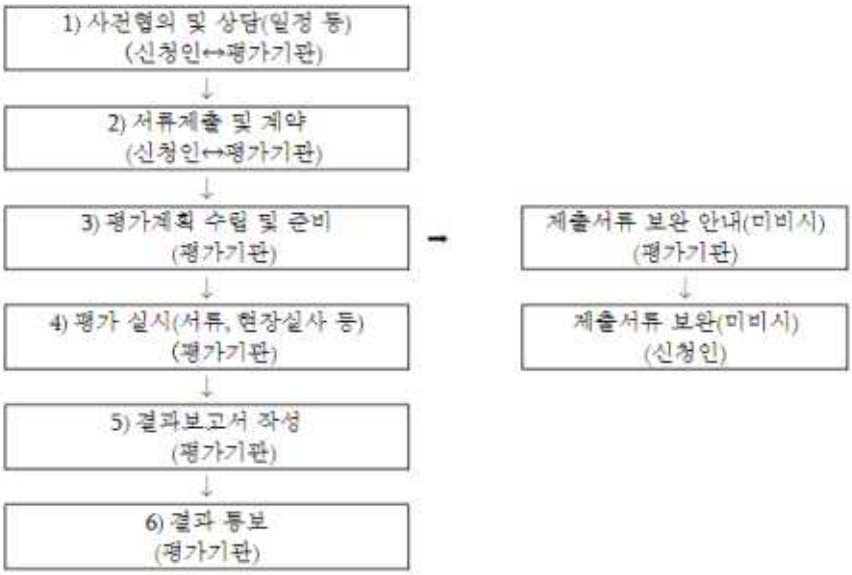
○ 평가 원칙

- 평가 업무를 수행함에 있어 다음의 평가 원칙 및 기준을 지켜야 한함.
  - 평가자는 평가의 수행함에 있어 객관성, 전문성 및 신뢰성을 확보해야 함.
  - 평가의 기준시점은 평가보고서에 명시하여야 함.
  - 평가의 결과 표시는 등급, 점수, 가액 등으로 표시할 수 있음.
- 기본사항의 특징
  - 의뢰자, 대상 데이터 평가 목적, 기준시점, 평가조건, 평가 결과 유형 및 활용 등을 특정해야 함.
- 목적과 용도의 명시
  - 데이터 가치평가는 목적과 용도에 따라 평가 관점이나 평가요인의 고려 요소가 달라지므로 평가의 목적과 용도를 명시하여야 함.
- 조건의 설정, 평가 기법 선정 및 적용
  - 데이터 가치평가는 여러 가지 대안 중에서 가장 발생 가능성이 높은 조건을 설정하여 효율적이고 효과적인 사용 원칙을 적용해야 함.
- 평가의 가정 및 한계
  - 평가기관은 평가의 결과를 신청인에게 통보할 때, 평가과정에서 사용된 가정과 제한적인 조건, 평가의 목적과 용도를 명시하고 평가 결과는 명시된 목적에서만 유효하며 상황 변화에 따라 평가 결과가 변동될 수 있음을 명시하여야 함.

○ 평가 대상

- 법 제14조 제1항에 따른 데이터 가치평가는 일정 수준의 품질이 담보되고 경제적 효익 측정이 가능한 데이터를 대상으로 함.

○ 평가 절차



(데이터 가치 평가 절차)

- 사전협의 및 상담
  · 신청인이 평가를 의뢰하려는 경우 평가기관은 신청인의 평가목적과 평가받으려는 데이터 등 상담에 필요한 사항을 확인하고 평가에 적용되는 절차, 기법(예시), 활용 범위, 가정 및 한계 등의 사항을 안내하고 일정, 수수료 등에 대해 협의
- 서류제출 및 계약
  · 신청인은 평가기관에 별지 제3호 서식의 신청서 및 평가 대상 데이터에 대한 현황 등이 담긴 서류 제출하고, 평가기관은 소요기간, 수수료, 비밀유지 등의 내용을 적시하여 신청인과 평가 계약을 체결
- 평가계획 수립 및 준비
  · 평가기관은 평가조직, 평가일정과 평가방법 등의 평가계획을 수립하여 평가 준비
- 평가 실시
  · 평가기관은 신청인이 제출한 서류를 검토하고 평가에 필요한 자료 수집과 현장 실사를 통해 평가 실시
- 결과보고서 작성
  · 평가기관은 조사·분석 내용을 종합하여 보고서를 작성하고, 수정 또는 보완이 필요한 경우 추가 조사 및 분석을 실시하여 수정 보고서를 작성
- 결과 통보
  · 평가기관은 신청인에게 결과보고서 송부

※ 다만, 합리적이고 능률적인 평가를 위하여 필요한 때에는 순서를 조정할 수 있음

(8) 해외 사례를 기반으로 한 데이터 가치평가 방법

○ 기업 가치 분석 (Enterprise Valuation)

- 기업의 가치를 분석하는 과정에서 데이터 자산의 가치를 평가하는 방법임. 이 방법은 기업의 재무제표, 수익성, 시장 점유율 등을 분석하여 데이터 자산이 기업의 가치에 미치는 영향을 계산함.

○ 데이터 수익 모델링 (Data Revenue Modeling)

- 데이터가 생성하는 수익을 모델링하여 데이터 자산의 가치를 평가하는 방법임. 이 방법은 데이터가

수익을 생성하는 방법과 수익의 크기를 분석하여 데이터 자산의 가치를 계산함.

○ 경쟁 분석 (Competitive Analysis)
- 비슷한 시장에서 경쟁사가 어떻게 데이터를 활용하고 있는지를 분석하여 데이터 자산의 가치를 평가하는 방법임. 이 방법은 기존 시장에서 데이터가 어떤 역할을 하는지를 파악하고 경쟁사의 데이터 활용 전략을 분석하여 데이터 자산의 가치를 계산함.

○ 미래 가치 모델링 (Future Value Modeling)
- 데이터 자산이 미래에 가질 수 있는 가치를 모델링하여 데이터 자산의 가치를 평가하는 방법임. 이 방법은 예측 분석, 데이터 과학, 인공지능 등을 활용하여 데이터 자산이 미래에 어떤 가치를 생성할 수 있는지 예측함.

(9) 해외 사례를 기반으로 한 데이터 자산의 특성 방법

○ 데이터 자산의 특성에 따라 적절한 방법을 선택하여 데이터 가치평가를 수행할 필요가 있음.

○ 데이터의 유형
- 데이터 자산은 다양한 유형의 데이터를 포함하고 있음. 이러한 데이터의 유형을 파악하여 데이터의 활용 가치를 평가할 필요가 있음. 예를 들어, 구조화된 데이터는 기존의 비즈니스 프로세스를 개선하고 최적화하는 데 매우 유용하며, 비구조화된 데이터는 예측 분석 및 인공지능 기술을 적용하여 새로운 비즈니스 모델을 창출하는 데 활용됨.

○ 데이터의 질
- 데이터 자산의 가치는 데이터의 질과 진위성에 따라 달라짐. 데이터의 질은 데이터의 정확성, 완전성, 일관성 및 유효성 등으로 측정됩니다. 이러한 질적 측면을 평가하여 데이터 자산의 가치를 결정할 필요가 있음.

○ 데이터의 양
- 데이터 자산의 양은 데이터의 수와 다양성을 나타냄. 데이터 양의 측면에서는, 데이터의 수는 데이터 자산의 가치와 직접적인 상관관계가 있음. 또한, 다양한 유형의 데이터를 포함하고 있는 경우 데이터 자산의 가치가 증가할 가능성이 높음.

○ 데이터의 사용 범위
- 데이터 자산의 가치는 데이터의 사용 범위에 따라 달라짐. 예를 들어, 데이터가 내부적인 용도로만 사용되는 경우보다는 외부적인 용도로도 활용될 수 있는 경우에 데이터 자산의 가치가 더 높아짐.

○ 데이터의 소유권
- 데이터 자산의 소유자와 이용자는 데이터의 가치를 다르게 인식할 수 있음. 데이터의 소유자는 데이터를 자사의 경쟁 우위를 확보하거나 기업가치를 높이는 데 활용하고자 할 수 있으며, 이용자는 데이터를 적극적으로 활용하여 새로운 가치를 창출하고자 할 수 있음. 이러한 소유권과 이용자의 요구사항을 파악하여 데이터 자산의 가치를 결정할 필요가 있음.

2. 국내 데이터 거래 관련 플랫폼 및 가격산정 사례 조사

가. 국내 데이터 거래 관련 플랫폼

(1) 데이터 플랫폼 현황

○ 다양한 분야에 걸쳐 공공(중앙・지방정부・공공기관) 주도 180개, 민간 주도 21개, 총 129개 데이터 플랫폼이 구축되어 운영

○ 개방, 분석, 거래 중 한 가지 기능만을 제공하는 플랫폼이 다수를 차지하며, 21개 플랫폼 만전 주기 지원

(2) 농식품 공공데이터 이용 현황 및 문제점

○ 농식품 공공데이터 포털을 통해 590종 데이터 개방, 누적 1.5억건 활용(‘21년’ 72백만건)

< 농식품 공공데이터 Top10 이용 비율 >

순위	데이터셋 명	설명	비율
1	경락가격 정보서비스	농수축산물 경락가격정보를 조회하기 위한 서비스로서 농산물, 축산물, 수산물, 화훼류에 대한 도매시장, 산지공판장, 종합유통센터 별로 농수축산물 경락가격 제공	74.1%
2	농수축산물 안심 레시피 정보	우리 농산물을 이용해서 만드는 요리정보로 요리기본정보, 레시피, 주요 재료 정보로 구성되어 있어 다양한 요리레시피 정보를 체계적으로 제공	5.8%
3	가축질병 발생정보	질병발생일자, 질병명(AI, 결핵, 브루셀라 등), 농장명, 소재지, 축종, 발생두수, 진단기관	3.7%
4	안심식당 정보	3가지 방역수칙을 준수하여 안심식당으로 선정된 식당 리스트 ①덜어먹기가능한도구비치・제공②위생적인수저관리③종사자마스크착용여부	3.0%
5	화훼류 시세현황	화훼류(절화, 난, 관엽, 춘란 등) 경락정보	2.8%
6	농업경영체등록관리	농업경영체 등록 현황 정보	2.7%
7	기관조사가격 서비스	품목의 가격정보를 조회하기 위한 서비스로서 가락 도매시장과 강서 도매시장의 가격정보와 농수축산물의 품목, 지역, 유통센터별 가격정보, 임산물의 품목, 지역별 가격정보를 제공	2.7%
8	농식품 안전성 정보	농식품 안전성에 대한 분석결과 및 안정성 검사기관 지정현황	2.6%
9	동물질병정보	동물질병정보제공	1.3%
10	농수축산물 도소매가격	농수축산물 도소매가격을 정보수요자에게 제공하여 시장출하 및 매매에 관한 의사결정, 원활한 수급조절 유도 및 가격안정대책 추진을 위한 기초자료	1.2%

○ Top 10 데이터 중 이용 대부분이 도매시장경락가격으로, 다양한 농식품 데이터의 개방・이용이 미흡

(3) 국내 플랫폼 산업 현황

○ 국가 경제의 경쟁력을 뒷받침하는 데이터의 산업 규모가 연평균 10% 이상 성장하며 혁신성장을 견인
○ 한국 IDC 연구보고서에 따르면 빅데이터 및 분석 시장은 2021년 전년대비 5.5%성장하여 2조 296억원의 매출 규모를 형성할 전망이며 향후 5년간 연평균 성장률 6.9%를 기록하며 2025년까지 2조 8,353억원 규모에 이를 전망

<국내 빅데이터 및 분석 시장 전망>

(4) 국내 데이터 플랫폼 구축・운영 현황

< 농식품 관련 빅데이터 플랫폼 구축・운영 현황 >

순번	플랫폼 명	구축 목적	보유 데이터
1	농정 빅데이터 플랫폼	농업 보조금 부정수급 차단・검증 및 농업인 제출서류 간소화 등	64종
2	농식품 빅데이터 플랫폼(KADX)	민간의 유통・소비 데이터를 중심으로 수집・융합・가공하여 데이터 유통・거래 생태계 구축 * 한국판 뉴딜 사업	320종
3	농산물유통 종합정보 플랫폼 (nongnet.or.kr)	농산물 수급 안정을 위한 농산물 생산・유통・소비 빅데이터 수집 및 분석 정보 제공	78종
4	스마트팜 빅데이터 플랫폼 (smartfarmkorea.kr)	스마트팜 데이터를 수집・활용하여 빅데이터 기반의 스마트 농업 전환 촉진	61종
5	축산업통합정보 플랫폼	기관별로 별도 운영중인 축산정보를 공유・활용할 수 있도록 통합 관리	26종
6	스마트팜 R&D 빅데이터 플랫폼	스마트팜 구축・확산을 통해 성장동력과 경쟁력 확보를 위한 R&D사업 * 농식품부, 농진청, 과기부 3개 부처 공동사업	시범 운영
7	축산물 경매시장 온라인 경매플랫폼	온라인으로 축산물 영상, 등급판정 등 정보를 제공하고 구매자가 경매에 참여	'22. 1월 시범 운영
8	공공급식 통합 플랫폼	공공 급식 식재료 현황 파악 및 원산지 이력관리로 안전한 식재료 공급	'22. 9월 구축 예정
9	귀농귀촌 통합 플랫폼	귀농·귀촌 희망수요를 적극 발굴하고, 맞춤 정보제공 및 정책 효율성 제고	'22. 12월 구축 예정

○ 개별 데이터의 생산 및 관리가 사업단위로 이루어지고 있어 체계적 관리 어려움
○ 민간 데이터의 경우 데이터 품질지표(Data Quality Index)가 있어 품질수준을 평가하고 있으나 농식품 공공데이터에는 평가기준이 없음

(5) 해외 데이터 플랫폼 현황

국가명	주요 내용
EU	• 독일, 프랑스 주도로 유럽 500여 개 기업·기관이 참여하는 클라우드 기반의 데이터 인프라 프로젝트 가이아-X 추진('19~) • 'EU 데이터전략('20.2')의 핵심으로서 주요 산업별로 개방형 데이터 인프라를 조성하여 다양한 주체의 참여를 유도하고 혁신을 촉진
미국	• 연방정부는 기업, 연구자 등이 활용할 수 있도록 오픈데이터 포털을 통해 공공 데이터를 개방('09~') • 각 연방기관이 준수할 가이드(표준화, 품질, 보안 등)를 개발하고 포털에 관련 정보(정책, 도구, 사례 등) 공개
중국	• 중국 각 지역에 민·관 합작(구이양 등 8개) 또는 민간 주도 8개로 오픈마켓 형태의 16개 데이터 거래 플랫폼 운영('14~) • 데이터거래플랫폼, 데이터서비스 업체 등을 통해 데이터 상품의 권리 확정을 진행하고, 데이터가격산정체계는 초기 단계

(6) KDX한국데이터거래소

○ KDX의 상품 수는 지속적으로 증가하고 있으며 통신/인구 데이터가 가장 많은 비율을 차지
○ 주로 업무 활용 목적으로 데이터의 판매가 이루어지고 있음
○ 데이터 공급자와 수요자의 데이터 거래 중개 진행
○ 맞춤형 데이터 분석 및 플랫폼 서비스 제공
○ AI 학습용 데이터 생산 및 유통 제공
○ 맞춤형 데이터 컨설팅 및 데이터 바우처 사업 제공
○ 데이터 전문 온라인 뉴스레터 '데이터 루' 발행

(7) 분야별 판매 데이터

○ 유통/마케팅

기업명	판매 데이터
GS리테일	• 지역별 GS25 점포 매출 분포 데이터 • 지역별-상품분류별 매출 구성비 데이터
엠코퍼레이션	• 온라인 배달 음식 카테고리 데이터 분석 • 온라인 상품 구매 데이터
MarketLink	• Retail Pos Index 데이터 • 대형마트/체인슈퍼 카테고리별 판매 데이터 • 치킨 프렌차이즈 판매 점유율 데이터
휴모스온	• 이커머스 상품별 구매 및 관심사 데이터

○ 소비/상권

기업명	판매 데이터
삼성카드	• 지역별/일별 소비 행태 데이터 • 반려동물 보유자의 소비 행태 데이터 • 고객유형별 소비 행태 데이터
신한카드	• 시군구별 가맹점 데이터
코리아크레딧뷰로	• 전국 시군구/행정동 단위 SOHO 데이터 • 전국 행정동 단위 소비 정보 데이터
식신	• 전국 핫플레이스 권역 정보(좌표) 데이터 • 전국 맛집 정보(메뉴, 가격) 데이터

○ 통신/인구

기업명	판매 데이터
SK텔레콤	• 지역별 서비스 인구 데이터 • 지역별 유입 인구 데이터
SK플레닛	• Syrup App 사용자 정보 데이터
LG유플러스	• 유동 인구 데이터 • 모바일TV/IPTV시청 통계 데이터
로플랫	• Wi-Fi 신호 기반 유동 인구 데이터 • 오프라인 매장 방문 통계 데이터

○ 유통/물류

기업명	판매 데이터
CJ올리브네트웍스	• 상품별 택배 송장 표준 인덱스 데이터 • 상품별 동 단위 물동량 데이터 • 고객 분류별 상품 매출 점유율 데이터

○ 경제/산업

기업명	판매 데이터
나잇스디앤비	• 산업 생태계 분석 데이터 • 기업 정보(재무, 국민연금, 등기) 데이터 • 중소기업 특화 ESG 평가항목 데이터

○ 인공지능

기업명	판매 데이터
매일경제미디어그룹	• AI 학습용 텍스트 원천/가공 데이터 • AI 학습용 영상 원천/가공 데이터 • AI 학습용 음성 원천/가공 데이터
에이아이스페라	• 사이버 공격 탐지 전 세계 위험 VPN IP 데이터
어반유니온	• AI 학습용 패션 데이터

○ 보건/의료

기업명	판매 데이터
온누리에이치엔씨	• 약국 처방(ETC) 데이터 • 약국 고객 판매 POS 데이터
이블루	• 의약품 판매 데이터

○ 금융/증권

기업명	판매 데이터
웰컴에프앤디	• 대부업 대출 신용등급 분포 정보 데이터 • 저축은행 대출 분포 정보 데이터
나이스디앤비	• 채무불이행 데이터 • 단기 연체 정보 데이터
코리아크레딧뷰로	• 시군구 단위 대출 정보 데이터 • 전국 행정동 단위 소득/대출/연체 정보 데이터
쿠콘	• 금융/보험 API 데이터

○ 부동산

기업명	판매 데이터
매일경제미디어그룹	• 부동산 분양 정보 데이터 • 부동산 매물/시세 정보 데이터 • 부동산 단지 정보 데이터
코리아크레딧뷰로	• 자산평가 데이터
비씨티원	• 공간 정보 기반 전국 아파트 환경 정보 데이터
경동도시가스	• 전입/전출 정보 데이터

○ 결합데이터

기업명	판매 데이터
신한카드XSK텔레콤	• 카드소비-이동동선 결합 분석 데이터
신한카드XGS리테일	• 카드소비-편의점 품목 결합 분석 데이터
코리아크레딧X신한카드XSK텔레콤	• 아파트 타겟 마케팅용 융복합 정보 데이터

○ 기타

기업명	판매 데이터
나이스디앤알	• 신차/중고차 등록 데이터 • 카테고리별 앱 사용성 평가(앱마인더) 데이터
KB캐피탈	• KB차차차 인기 검색 차종 데이터
바이브컴퍼니	• 소셜 미디어 분석 데이터
경동도시가스	• 도시가스 데이터

(8) 농식품 빅데이터 거래소(KADX)

◯ 농식품 빅데이터 거래소는 농식품 데이터를 온라인으로 판매하거나 구매할 수 있는 플랫폼
◯ 서비스 개발, 데이터 분석 등 사업에 필요한 데이터를 데이터 거래소에서 쉽게 사고 팔 수 있음.

(9) 사업영역

< KADX 주요 사업 영역 >

(10) 판매 데이터

◯ 농업 및 유통

기업명	판매 데이터
한국농수산식품유통공사	• 전국 공영도매시장 실시간 경매자료 데이터 • 농수축산물 일자별 도소매 가격 데이터 • 농가 및 영농정보 데이터 • 친환경 농산물 가격정보 데이터 • 유통실태조사 데이터 등
장보고식자재마트	• 거래방식 및 거래빈도 소비 데이터 • 유통사별 상품 데이터 • 기간별 거래 데이터 • 농산물 매입 카테고리 데이터 • 거점별 거래지역 및 소비데이터 등
유로지스넷(주)	• 월별/품목별 판매량 데이터 • 전국단위 품목별 생산량 추이 데이터 • 지역 간 품목별 유통량 추이 데이터 등

○ 축산

기업명	판매 데이터
WISELAKE	• 양계(산란계) 사양 정보 데이터 • 한우 개체 급이 정보 데이터 • 경영정보 모돈(월별) 데이터 • 경영정보 후보돈(산차별) 데이터 • 이유(생육) 정보(지역별) 데이터 등
한국마사회	• 지능형 마방 말 생체 데이터 • 지능형 마방 외부 센서 데이터 • 지능형 마방 내부 센서 데이터

○ 소비

기업명	판매 데이터
만개의 레시피	• 여러 가지 음식의 레시피 데이터
케이플러스	• 외식 소비패턴 데이터 • 맛집지수 및 관광지수 데이터 • POS가맹점 월별 매출 정보 데이터 등
나이스지니데이타	• 지역별 소비자유형별 농식품 소비데이터 • 소비자 성/연령대별 농식품 소비데이터 • 업종별 소비지역별 농식품 소비데이터 등

○ 국제

기업명	판매 데이터
플랜인피닛	• 농식품 해외 바이어 정보 데이터 • 국내외 전시회 데이터베이스 데이터 • 수출시 활용 인증 데이터베이스 데이터 • 국내외 B2C 플랫폼 데이터베이스 데이터 • 국내외 B2B 플랫폼 데이터베이스 데이터
한국무역통계진흥원	• 농수축산물의 중소기업 맞춤형 수출 통계 데이터 • 기업유형별 수출지표 통계 데이터 • 품목별 수출입 통계 데이터 • 수출입단가 및 교역조건지수 데이터 등

○ 농식품 공공데이터

기업명	판매 데이터
농림축산식품 공공데이터 포털 농림축산식품부	• 계약재배(노지채소) 현황 데이터 • 시설재배 계약재배 현황 데이터 • 농약안전사용기준 데이터 • 지능형 마방 말 생체 데이터 • 지역별 젖소 산지 가격 등

나. 국내 데이터 거래 관련 가격산정 사례 조사

(1) 데이터 판매가격 산정 방법론

접근법	의의	장점	단점
시장접근법	유사 데이터를 판매하는 경쟁사의 가격을 고려해 자사의 데이터 가격을 결정하는 방법	경쟁 업체 가격 경쟁력을 확보할 수 있음	가격 변동 시 그 이유를 추적하지 어려움 (시장 환경 변화, 동종업체간 비용 담합 등)
가치접근법	구매자가 인지하는 평가·수요·가치를 바탕으로 데이터 가격을 결정하는 방법	고객이 중심이 되어 가격을 책정하므로 거래 성사율을 높일 수 있음	가치 판단은 주관적 영역이므로 가격 책정 일관성을 확보하기 어려움
수익(소득) 접근법	데이터를 활용함으로써 얻을 수 있는 미래의 소득이 얼마나 될 것인지 추정한 다음 현재가치로 할인한 금액을 데이터 가치로 평가하는 방법	거래가격을 정교하게 산정할 수 있음	수입측정에 사용되는 변수들이 많고 모두 예측에 의하므로 평가자의 주관적 요소가 적용될 여지가 큼
원가접근법	원가(데이터 구축 운영비 + 최소이윤)를 충당하고 목표이익을 낼 수 있는 수준에서 가격을 결정하는 방법	비교적 공정하고 간단하게 가격을 설정할 수 있음	수요나 경쟁을 고려하지 않고 데이터가 가진 가치를 제대로 반영하지 못함

(2) 공공기관 A기관

○ 데이터 상품 및 제공방식

- A기관이 보유한 데이터 상품은 출원인이 출원할 때부터 권리소멸까지 생산 축적되는 데이터이며, API와 데이터 세트 2가지 형태로 제공하고자 함

○ 데이터 가격책정의 목표

- 대량의 데이터가 상업적 목적이 있는 이용자에게 보급될 경우 사적재의 특성을 가지므로 일정 수준의 수수료를 부과하는 것이 필요함. 다만, 공공 데이터는 공공활동을 위해 생산되었고 정부의 예산이 투입되었으므로, DB생산·구축을 위한 비용은 원가 책정에서 제외됨. 데이터 검색/제공 서비스 운영과 유지, 가공 비용을 회수하는 것을 목적으로 함

○ 가격 산정의 방법

- 데이터 제공 서비스의 운영을 위한 투입 원가를 기준으로 가격을 산정함
- 단위 상품에 대해 소요되는 개별 원가가 아닌 전체 상품에 소요되는 인건비, 경비 등을 기초로 총괄원가를 산정한 후 전체 상품 중 가중치에 따라 이용료를 배분
- 공공 기업 데이터 서비스원가 = (인건비, 경비) + 일반관리비 + 경우에 따른 최소 이윤 데이터 상품별 원가 = $\sum${(항목별 원가) X (상품별 가중치)}

○ 인건비 책정

- 데이터 구축 작업에 종사하는 종업원, 노무자가 제공하는 노동력의 대가
- 인건비 = 데이터 구축 작업 수행비용 + 상여금 + 퇴직급여 충당금

○ 직접비의 책정

- 데이터 구축 작업에 소요되는 직접 경비
- 직접비 = 경비 = 컴퓨터 시스템/SW 사용료 + 전문가 비용 + 여비 + 특수자료비(특허 등) + 인쇄 및 청사진비 + 자료조사비 + 기자재 시험비 + 위탁비 및 현장운영비 + 프로토타입 제작비 등

○ 재경비

- 인건비와 직접경비에 포함되지 않는 비용으로 간접비를 의미
- 제경비 = 간접비 = 프로젝트 관리자, 임원, 경리직원 등의 급여 + 사무실 임차료 + 사무용품비 + 비품비 + 기계수선 및 상각비 + 통신비 + 회의비 + 공과금 + 기타 운영비 등

○ 해당 데이터 상품 소요비용 계산 방법

- 기존에 누적된 데이터를 상품화하기 위해서 투입되는 업무(데이터 이관, 구축, 관리, 보급, 개발, 응대 등)들을 분류하고, 업무별 소요 시간과 비용을 산출합니다. 이 경우는 공통 업무와 데이터 세트 업무, API 업무 등으로 구분한 후, 상품별로 나눠 총 원가를 산정하고, 판매 가격을 결정한 사례

○ 데이터 상품 개발을 위한 업무 분류 및 소요시간 산출

구분	내용	투입 공수(시간)		
		공통 업무	데이터 세트 업무	API 업무
데이터 이관	국내외 DB 파일 입수, 데이터 백업 등	1,200		
데이터 구축	국내외 DB 탑재, 누락 데이터 재구축 등	1,642		
데이터 관리	서버 용량 관리, DB 성능, 품질, 모델 관리 등	466		
데이터 보급	FTP 서버 관리, 데이터 재구축 등		750	
개발	데이터 분석・API설계・구현, 테스트, 배포 등		70	30
민원 응대	전화・메일 등 응대, 서비스 유지보수 등	250		
웹서버 관리	웹서버・도메인・서비스 및 콘텐츠 운영관리, 유지보수 등	250		
시스템관리	전산자원관리, 백업 관리, 전산자원조사 등	130		
보안관리	보안유지보수, 보안관련조사	86		
문서작업・회의	문서작성 및 회의운영	350		

○ 업무별/상품별 소요시간과 참여인력 비율 책정

- 전체 업무 중 상품 개발에 따른 소요시간과 참여인력의 비율을 산정

투입 공수(시간)		
공통 업무	데이터 세트 업무	API 업무
150,000,000원/년	25,000,000원/년	17,500,000원/년

○ 업무별/상품별 직접비와 제경비 산출

비목	배부율(%)	데이터 세트 업무		API 업무	
		전체 금액(원)	금액(원)	전체 금액(원)	금액(원)
전력비	0.6	150,000,000	900,000	24,000,000	1,440,000
여비(교통비)	2.9		4,350,000		6,960,000
지급임차료	0.3		450,000		720,000
복리 생비	1.0		1,500,000		2,400,000
(생략)					
계	8.0		65,000,000		12,000,000

○ 일반관리비 비율 산정: 일반관리비는 원가 대비 일반관리비 비율을 계산

○ 각 상품별 총괄원가 산정

- 총괄원가는 인건비와 경비, 일반관리비를 더하고 부가가치세를 더해 반영

구분		데이터 세트 업무(원)	API 업무(원)
용역 원가	1. 재료비	-	-
	2. 인건비	25,000,000	17,500,000
	3. 경비	65,000,000	12,000,000
4. 일반관리비		6,500,000	1,300,000
5. 이윤(제외)		-	-
6. 원가(년간)		200,000,000	40,000,000
7. 부가가치세		20,000,000	4,000,000
8. 총괄 원가		220,000,000	44,000,000
9.계		264,000,000	

○ 판매 전략 및 할인 정책

- 최종 투입 원가를 파악한 후, 최종적으로 몇 년안에 총 원가를 회수하겠다는 계획 등 판매 전략과 시나리오를 설정
- 주요 사용자층에 따라 다양한 할인 정책을 제공할 수 있음

구분			할인 정책
데이터 세트	기관 할인		개인, 중소기업, 공공기관, 대학 50%
	데이터 특성	1	전문 혹은 이미지 택일 60%
		2	서지 데이터만 50%
	데이터 시기	과거 5년 미만	30%
		과거 10년 미만	60%
		과거 11년 이전	무료
API	기관 할인		개인, 중소기업, 공공기관, 대학 50%
	묶음 할인		2개이상 50%

(3) 민관기관 B사

○ 데이터 상품 및 제공방식

- 전국 연간/월별/일별 부동산 시세, 입주 정보, 분양정보 등을 제공하고 있으며, 데이터 판매를 통해 서비스 구축비와 운영비를 회수하고 이윤을 창출하고자 함. 이 경우는 해당 기업이 맞춤형 가공 데이터 세트를 제공할 때 가격을 산정한 사례

○ 가격 산정의 방법

- B기업은 동종업계의 경쟁상품 대비 방대한 표본수를 제공함으로써 상품을 차별화함. 따라서 경쟁사 가격을 고려하기보다 실제 투입된 원가를 바탕으로 가격을 제시하고 일부 비율만큼 이용료를 추가해 산정하고 있음

○ 인건비

- 기획재정부 고시 「2018년 학술연구용역인건비 기준 단가」를 바탕으로 투입비율과 시간을 고려 하여 산정

구분	항목	단가(원/월)	투입비율	투입기간	금액
인건비	팀장급 연구원(경력 5년이상)	3,169,323	30%	1.4개월	1,302,592
	대리급 연구원(경력 3년~5년)	2,430,194	90%	1.4개월	2,996,429
	사원급 연구원(경력 3년 미만)	1,624,503	80%	1.4개월	1,780,455

○ 경비: 데이터 조사 및 처리 비용, 자료수집비, 전산처리비 등의 비용을 계상

구분	항목	단가(원/월)	투입비율	투입기간	금액
데이터	데이터 조사 및 처리	1,000,000	1.5	1.0개월	1,000,000
부대비	자료 수집비	800,000	1,5	1.0개월	800,000
전산 처리비	전산비 및 전산운영비	300,000	1.0	1.0개월	300,000
	사무용품 및 소모품비	800,00	1.0	1.0개월	800,00

○ 일반관리비: 인건비와 경비의 4% 수준으로 책정하여 산정

○ 이윤: 인건비와 경비, 일반관리비의 4%로 책정하여 산정

(4) 민간기관 C사

○ 데이터 상품 및 제공 방식

- 연립・다세대・아파트 등 부동산의 매물 및 시세 정보, 주변 정보에 대한 데이터를 별도 계약을 통해 제공하고 있으며 데이터 구축을 위한 소요비용이 전체 지출의 많은 부분을 차지하고 있음. 따라서 데이터 판매를 통해 최소한의 데이터 수집・분석 운영비를 확보하고 기업을 영위 하는 것을 목표

○ 가격 산정의 방법

- 데이터 판매시 가격은 고객의 특성에 따라 원가기준 가격과 구매자 가치기준 가격을 혼합하여 제시
- 원가 기준 가격을 산정하는 경우 먼저 회수 목표와 기간을 고려하여 시나리오를 설정함. 예를 들어 시장을 전체 30%로 볼 때 소요되는 투입 원가가 5억 원이고 이윤을 10%로 남긴다면, 3년 안에 1억 원, 1.5억 원, 3억 원 회수한다고 설정하고 계약 가격을 제시
- 구매자 가치 기준 가격은 고객이 경험적으로 알고 있는 가격이 있다면 동일한 수준으로 제시하거나, 구매자가 대체할만한 데이터를 구축하는 비용의 1/3 가격을 제시하는 경우가 많음

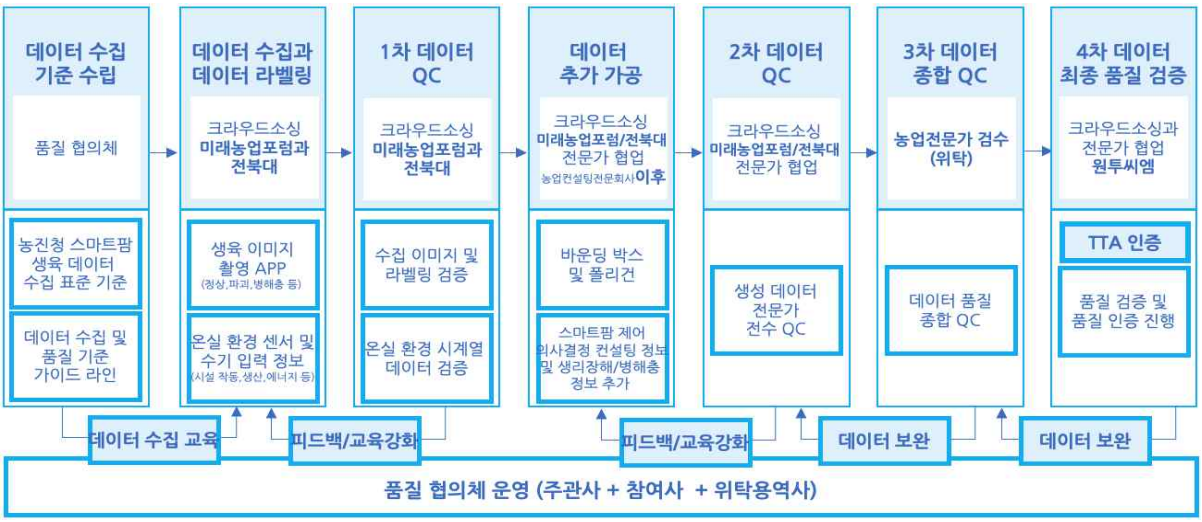
(5) 스마트농업 데이터의 거래 가능한 상품화 유형 조사

○ 토마토(베타티니)

- 현재 토마토 데이터의 수집 및 가공환경의 구축은 농업관련 전문기관의 스마트팜 작물 생육데이터 취득기준을 참고하여 상세기준을 마련하고, 기존 스마트팜 농자 현장에 추가로 센서를 설치하는 것으로 시작
- 기보유 스마트팜의 시계열 빅데이터 관리체계를 활용할 수 있는 기반을 마련
- 구축된 센서 데이터수집 환경과 더불어 필요 수기데이터의 수집환경도 함께 제공하며 어노테이션 플랫폼을 활용하여 크라우드 워커들에 의하여 데이터수집 활동 진행

시설ID	시도	시군구	품목명	품종	재배면적	시설유형	재배형태	단동/연동	수집일	외부온도	외부습도	외부 일사	외부풍향	외부풍속	외부강우	강우감지	내부온도	내부습도	내부CO2
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 0:00	-1	84	16	294	0	1		13	95	578
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 1:00	-2	84	26	295	0	1		13	95	573
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 2:00	-1	85	25	69	0	1		13	95	576
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 3:00	-1	85	23	108	0	1		13	95	578
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 4:00	-1	86	23	117	0	1		14	95	580
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 5:00	-2	87	23	68	0	1		13	95	584
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 6:00	-2	88	21	60	1	1		14	94	582
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 7:00	-2	86	32	60	1	1		13	94	586
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 8:00	-1	82	102	65	0	1		13	96	592
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 9:00	2	75	242	61	1	1		17	87	588
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 10:00	6	65	322	62	1	1		21	64	566
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 11:00	9	50	387	82	1	1		26	60	550
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 12:00	12	48	548	97	2	1		31	34	551
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 13:00	13	48	561	97	2	1		32	25	564
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 14:00	13	52	401	97	3	1		34	32	565
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 15:00	12	55	216	119	2	1		28	53	536
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 16:00	11	56	109	107	2	1		24	78	521
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 17:00	11	58	54	104	1	1		20	91	542
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 18:00	9	64	23	96	0	1		18	95	574
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 19:00	8	71	25	75	1	1		17	96	573
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 20:00	6	75	21	64	0	1		16	97	570
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 21:00	7	77	21	84	0	0		16	97	577
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 22:00	7	82	25	88	0	0		16	97	590
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-01-31 23:00	7	84	25	104	0	0		16	97	600
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 0:00	7	85	27	101	0	0		16	97	605
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 1:00	7	87	24	66	0	0		16	98	610
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 2:00	8	87	22	76	0	0		16	98	615
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 3:00	8	86	19	128	1	0		16	98	618
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 4:00	9	86	23	113	0	0		16	98	619
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 5:00	10	86	22	125	0	0		15	98	628
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 6:00	9	85	20	153	0	1		15	98	629
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 7:00	9	84	18	184	0	1		15	98	632
PF_0020070	전라남도	화순군	토마토	베타티니	2,100 비닐		수경	단동	2021-02-01 8:00	9	84	45	149	0	1		15	98	631

<토마토 원천데이터>



<토마토 데이터 공정 단계>

○ 오이

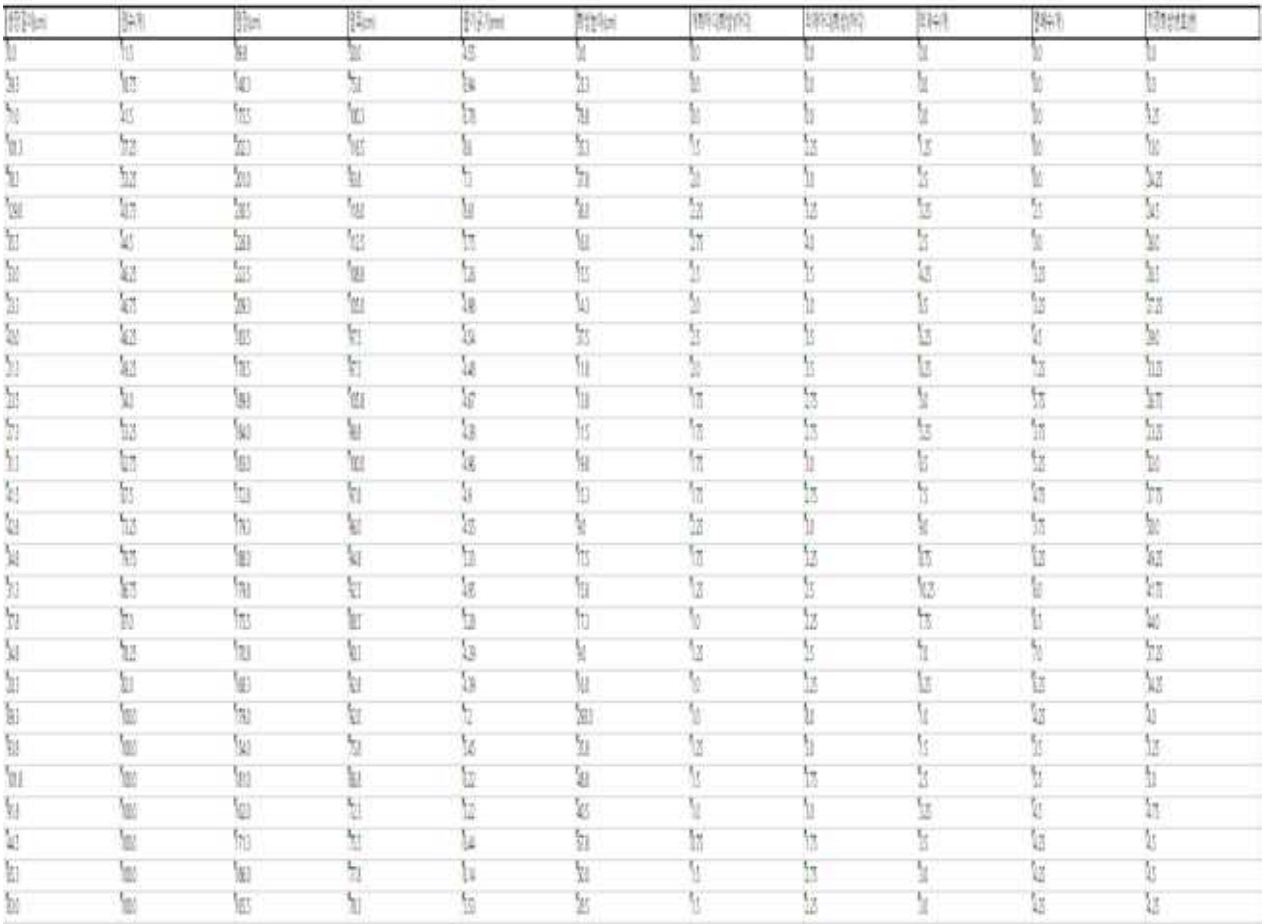
- 오이 스마트팜 외·내부 환경 정보 수집 기준으로는 내부 환경 (온도, 습도, 일사량, CO_2), 외부 환경 (온도, 풍속)으로 구성
- 오이의 생육 정보 수집은 1시간 단위로 실시하였으며 초기생육은 초장, 엽장, 엽폭, 엽수 등 4개의 항목을 수집
- 조사일지는 연월일로 표시하며, 생육 정보를 수집하는 식물체는 특정 번호를 부여하여 관리

수집일시	내부CO2(ppm)	이슬점온도(도)	토양EC(dS/m)	내부습도(%)	강우감지	일사량(W/m-2-s)	외부온도(도)	내부온도(도)	외부풍속(m/s)
2021-12-25 17:00:00	474.68	16.0	0.5	98.27	0.0	5.5	-6.0	17.35	3.27
2021-12-25 18:00:00	488.86	15.0	0.5	98.08	0.0	0.0	-6.45	16.03	2.82
2021-12-25 19:00:00	500.14	14.09	0.5	98.0	0.0	0.0	-7.0	15.15	3.0
2021-12-25 20:00:00	496.18	13.82	0.5	98.24	0.0	0.0	-7.05	14.3	2.73
2021-12-25 21:00:00	490.14	13.0	0.5	97.48	0.0	0.05	-8.0	13.64	1.23
2021-12-25 22:00:00	487.95	12.95	0.5	97.94	0.0	0.0	-8.0	13.03	1.09
2021-12-25 23:00:00	487.86	12.05	0.5	97.73	0.0	0.0	-8.0	12.98	1.73
2021-12-26 00:00:00	485.64	12.0	0.5	97.68	0.0	0.0	-9.0	12.61	1.64
2021-12-26 01:00:00	483.77	12.0	0.5	97.86	0.0	0.0	-9.0	11.98	1.77
2021-12-26 02:00:00	481.14	11.68	0.5	97.64	0.0	0.0	-9.09	11.91	1.41
2021-12-26 03:00:00	480.82	11.0	0.5	97.33	0.0	0.0	-10.5	11.2	0.09
2021-12-26 04:00:00	480.14	10.45	0.5	97.18	0.0	0.0	-11.0	11.0	0.18
2021-12-26 05:00:00	481.5	10.0	0.5	97.58	0.0	0.0	-11.5	10.26	0.14
2021-12-26 06:00:00	484.73	10.0	0.5	97.3	0.0	0.0	-11.95	9.94	0.0
2021-12-26 07:00:00	491.5	10.0	0.5	97.32	0.0	0.5	-12.0	9.88	0.0
2021-12-26 08:00:00	498.36	9.27	0.5	97.95	0.0	40.0	-11.09	9.35	0.0
2021-12-26 09:00:00	495.82	9.36	0.5	97.02	0.0	192.91	-8.0	9.58	0.0
2021-12-26 10:00:00	498.32	12.36	0.5	97.76	0.0	276.59	-7.55	11.41	1.82
2021-12-26 11:00:00	504.32	15.18	0.5	96.79	0.0	425.36	-6.05	14.3	2.59
2021-12-26 12:00:00	501.68	20.91	0.5	98.41	0.0	469.41	-4.82	19.7	2.41
2021-12-26 13:00:00	497.23	24.18	0.5	98.55	0.0	444.32	-3.95	23.05	2.68
2021-12-26 14:00:00	491.09	23.0	0.5	98.98	0.0	372.73	-3.55	23.05	3.36
2021-12-26 15:00:00	490.32	20.64	0.5	99.03	0.0	249.18	-3.23	20.82	3.36
2021-12-26 16:00:00	496.5	15.64	0.5	98.52	0.0	82.5	-3.91	17.3	2.82
2021-12-26 17:00:00	495.5	15.59	0.5	98.5	0.0	5.95	-5.05	16.88	2.82
2021-12-26 18:00:00	499.14	14.41	0.5	98.55	0.0	0.0	-6.0	15.5	1.05
2021-12-26 19:00:00	512.14	13.45	0.5	97.89	0.0	0.0	-7.23	14.33	0.05
2021-12-26 20:00:00	516.77	12.77	0.5	97.52	0.0	0.0	-8.73	13.52	0.05
2021-12-26 21:00:00	515.14	12.0	0.5	97.8	0.0	0.0	-9.32	12.77	0.0
2021-12-26 22:00:00	513.86	11.95	0.5	97.8	0.0	0.0	-10.0	12.33	0.18
2021-12-26 23:00:00	513.23	11.0	0.5	97.52	0.0	0.0	-10.05	11.76	0.0

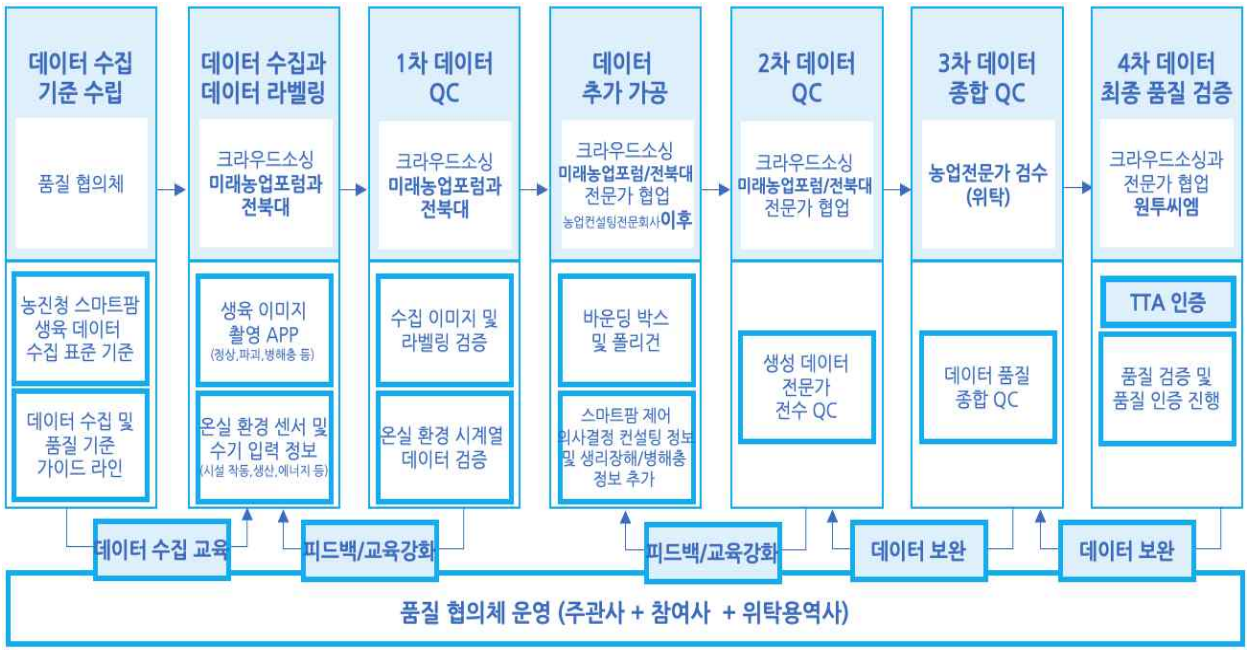
<오이 원천데이터>

○ 파프리카

- 파프리카의 효율적인 작물 재배를 위해 AI 기반의 신속·정확한 실시간 조기 작물 생육 진단 시스템 도입
- 농작물 품질관리 및 생산력 향상을 도모와 재배환경에 따른 시설 작물의 생육 및 생산량 변화 예측이 가능한 빅데이터 기반 농업 생산성 향상을 위한 AI 학습용 데이터 구축
- 4단계의 품질 검증 및 피드백 단계를 가진 전체 7단계의 공정으로 학습용 데이터 구축
- 데이터는 파프리카 농가 6개 이미지 및 시계열 데이터를 자체 수집
- 구축된 이미지 데이터는 생장, 화방, 줄기, 엽장, 꽃, 열매, 만개 꽃으로 총 207,000건의 데이터가 분포됨



<파프리카 원천데이터>



<파프리카 데이터 공정 단계>

○ 방울토마토

- 방울토마토의 스마트팜 ICT 데이터를 일별, 시간별로 수집
- 스마트팜 내·외부 환경 정보 수집
- 지역별, 온실 규모, 온실 형태, 재배형태, 온·습도, 일사량 등과 같은 데이터 수집

시도	시군구	품목명	품종	재배면적	시설유형	재배형태	단동/연동	수집일	외부온도	외부습도	외부 일사량	외부풍향	외부풍속	외부강우	강우감지	내부온도	내부습도	내부CO2
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	19		0		0		0	21	87	324
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		0		0		0	20	88	340
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		0		0		0	21	88	353
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		0		0		1	21	88	347
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		0		0		1	21	88	348
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		0		0		1	21	87	350
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		0		0		1	20	88	356
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		22		0		0	20	88	362
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		81		0		0	21	86	344
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	19		127		0		0	22	77	315
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	20		327		0		0	25	68	309
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	20		209		1		0	24	66	303
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	21		177		1		0	24	68	298
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	21		176		1		0	24	64	290
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	20		150		1		0	23	63	275
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	21		229		1		0	24	60	275
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	21		146		1		0	23	58	276
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	20		29		0		0	21	66	273
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	19		1		0		0	20	73	280
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	18		0		1		0	19	72	274
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	16		0		2		0	18	66	271
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	15		0		1		0	18	71	283
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	15		0		2		0	18	77	294
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	14		0		1		0	17	80	304
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	13		0		0		0	16	80	313
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	12		0		0		0	15	79	319
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	10		0		0		0	14	77	323
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	9		0		0		0	13	75	330
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	8		0		0		0	12	75	334
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	7		0		0		0	12	75	339
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	7		2		0		0	11	74	343
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	9		80		0		0	13	70	350
충청남도	부여군	방울토마토		3,306 비닐		수경	연동	********	11		238		0		0	20	62	357

<방울토마토 시간별 원천데이터>

○ 버섯

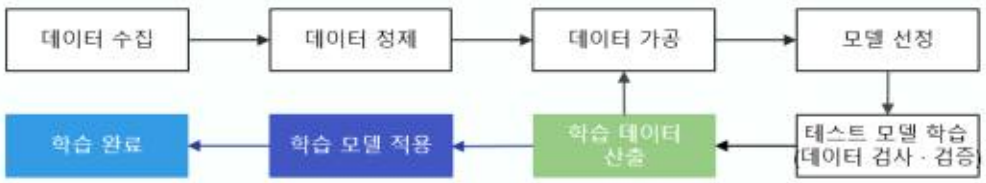
- AI Hub의 지능형 스마트팜 통합 데이터(버섯)
- 2021년도에 데이터가 구축되었으며 총 데이터양은 350,621건임.
- 각 데이터는 양송이 70,900건, 느타리 86,460건, 큰느타리 62,419건, 팽이 77,050건, 표고 53,792건으로 이루어져 있으며 각 품종마다 배양단계, 생육단계, 병해충 단계로 나뉘어져 있음.

품종	수량
양송이	70,900
느타리	86,460
큰느타리	62,419
팽이	77,050
표고	53,792
합계	350,621

주기	수량
배양	13,521
생육	244,700
병해	92,400
합계	350,621

<데이터 분포표>

- 센서를 통해 풍속, 온도, 습도, 조도, 이산화탄소 데이터를 측정 후 수집하고 카메라를 통해 병해 이미지 및 생육 이미지 수집
- 구축된 데이터를 기반으로 수확기 버섯 상태 구분, 버섯 품종 인식, 병충해 판단 모델 개발
- 데이터는 학습 데이터 80%, 검증 데이터 10%, 평가 데이터 10%로 구분하여 관리하고 학습 데이터를 적용하고 성능평가를 통해 품종 인식, 생육단계 판단, 병충해 판단 데이터 학습 모델 구축



<모델링 과정 도식도>

○ 딸기

- 농림축산식품부 공공데이터 포털의 지역별 딸기 스마트팜(시설원예) ICT 생육 융복합 정보
- 지역, 품종, 재배면적, 시설유형, 재배형태, 단동/연동을 기준으로 딸기의 생육데이터
- 생육데이터는 엽장, 엽폭, 엽병장, 엽수, 화방수 등 8개의 항목을 수집하고, 환경 데이터는 외부온도, 외부습도, 외부일사량, 외부풍속, 외부강우, 강우감지, 내부온도, 내부습도, 내부 CO2, 내부 일사량, 토양 EC, 토양 pH, 지온, 지습 총 14개의 항목을 수집하고 양액 데이터는, 총 배액, 총 급액, 배액 pH 등 총 8개의 항목별로 수집

시군구	품목명	품종	재배면적	시설유형	재배형태	단동/연동	수집일	외부온도	외부습도	외부 일사량	외부풍향	외부풍속	외부강우	강우감지	내부온도	내부습도	내부CO2
보령시	딸기	설향	1,488	비닐		연동	*******	19		0		0	0	0	21	92	420
보령시	딸기	설향	1,488	비닐		연동	*******	19		0		0	0	0	20	92	416
보령시	딸기	설향	1,488	비닐		연동	*******	19		0		0	0	0	20	92	398
보령시	딸기	설향	1,488	비닐		연동	*******	19		0		0	0	0	20	92	408
보령시	딸기	설향	1,488	비닐		연동	*******	18		0		0	0	0	19	92	418
보령시	딸기	설향	1,488	비닐		연동	*******	18		0		0	0	0	19	92	420
보령시	딸기	설향	1,488	비닐		연동	*******	18		6		0	0	0	19	93	419
보령시	딸기	설향	1,488	비닐		연동	*******	19		63		0	0	0	21	91	412
보령시	딸기	설향	1,488	비닐		연동	*******	21		99		0	0	0	23	89	378
보령시	딸기	설향	1,488	비닐		연동	*******	22		159		0	0	0	24	87	333
보령시	딸기	설향	1,488	비닐		연동	*******	25		299		0	0	0	28	71	294
보령시	딸기	설향	1,488	비닐		연동	*******	25		194		0	0	0	28	70	286
보령시	딸기	설향	1,488	비닐		연동	*******	25		211		0	0	0	27	76	291
보령시	딸기	설향	1,488	비닐		연동	*******	23		145		0	0	0	26	81	304
보령시	딸기	설향	1,488	비닐		연동	*******	23		109		0	0	1	26	83	307
보령시	딸기	설향	1,488	비닐		연동	*******	23		118		0	0	1	26	84	311
보령시	딸기	설향	1,488	비닐		연동	*******	23		45		0	0	0	24	88	312
보령시	딸기	설향	1,488	비닐		연동	*******	23		44		0	0	1	24	89	319
보령시	딸기	설향	1,488	비닐		연동	*******	22		8		0	0	1	22	84	305

< 수집 데이터 예시>

○ 스마트팜 데이터 마트

- 현재 스마트팜 데이터 마트에는 딸기의 생육데이터를 품종, 연도, 지역, 온실 형태, 온실 규모로 분류하여 데이터를 제공 중
- 딸기의 경우 품종은 금실, 설향, 죽향의 데이터가 있고, 2015~2021년도의 데이터가 있으며 온실의 형태로는 단동/연동이 있고 온실 규모는 소규모 1,000평 이하, 중규모 1,000평~3,000평, 대규모 3,000평 이상으로 분류된 데이터가 있음.

생육정보_일별(딸기)

조사일	주차	초장(cm)	엽장(cm)	엽폭(cm)	엽병장(cm)
2021-12-17	1	300.0	130.0	110.0	150.0
2021-12-23	2	290.0	129.5	120.5	162.0
2021-12-30	3	257.5	135.0	116.75	119.0
2022-01-07	4	256.25	126.5	110.0	127.75
2022-01-13	5	252.0	124.75	109.5	127.25
2022-01-21	6	261.25	108.25	90.25	153.0
2022-01-27	7	266.25	110.5	92.75	155.75
2022-02-03	8	271.0	123.25	105.0	147.75
2022-02-10	9	285.0	104.5	91.25	180.5
2022-02-17	10	243.25	98.75	85.75	144.5
2022-02-24	11	248.0	104.25	90.25	143.75
2022-03-03	12	262.25	109.0	93.5	153.25
2022-03-10	13	271.5	110.75	94.5	159.75

<스마트팜 데이터마트 수집 데이터>

3. 표준확산사업 연계 방안 및 관련 표준화 준수사항 도출

가. 표준확산사업 연계 방안

(1) 데이터 거래의 문제점

○ 국가통계포털(KOSIS)에 의하면 쓸만한 양질의 데이터가 부족, 불합리한 데이터 가격 선정, 데이터 유통 채널 부족, 데이터 유통 채널 부족, 데이터 품질 문제 등의 문제점이 발견

애로사항별(1)	2021	
	응답기업수 (개)	비중 (%)
전체	146	-
쓸만한 양질의 데이터 부족	76	52.1
구매 데이터 불합리한 데이터 가격 책정	54	37.0
데이터 유통 채널 부족	53	36.3
데이터 품질 문제	10	6.8
데이터 소재파악 및 검색의 어려움	46	31.5
개인정보 포함 데이터 유통/활용시 법적문제·법률지원	11	7.5
개인정보 처리 기술력 및 예산 부족	9	6.2
데이터 가공 문제	2	1.4
데이터 거래 절차 및 방법 미숙	3	2.1

<데이터 거래의 애로사항>

(2) 데이터 거래가 원활하지 않는 이유

○ 데이터를 무료로 사용하려는 이용자 인식, 데이터 상품 검색의 어려움, 개인정보 및 법·제도 문제, 거래방법 정보 부재

○ 공급자과 수요자를 연결하는 데이터 거래의 장을 마련하고, 원하는 데이터를 찾고 비교할 수 있는 검색체계를 갖춰야 함.

문제점	원인
• 데이터의 중복 및 조직, 업무, 시스템별 데이터 불일치 • 데이터에 파악 지연으로 정보 제공의 적시성 결여 • 데이터 통합의 어려움 • 정보시스템 변경 및 유지보수 곤란	• 동시 다발적인 정보시스템 개발 • 전사 데이터 관리 마인드 미형성 • 전사 데이터 관리 인력 부재 • 전사 데이터 표준 관리 도구 부재
개선방안	
• 데이터 표준화, 규격화를 위한 기본 방침 설정 • 전사적인 정보 공유를 위해 유지되어야 할 공통 데이터 요소의 도출 • 전사적인 데이터 요소 등록 및 관리 체계 구축	
기대효과	
• 명칭의 통일로 인한 명확한 의사소통의 증대 • 필요한 데이터의 소재 파악에 소요되는 시간 및 노력감소 • 일관된 데이터 형식 및 규칙의 적용으로 인한 데이터 품질 향상 • 정보시스템 간 데이터 인터페이스 시 데이터 변환, 정제 비용 감소	

나. 데이터 가치평가 시 표준화 준수사항

(1) 데이터 거래 표준화 필요성

○ (유통・거래) 데이터 공급기업은 수요의 불확실성과 플랫폼을 통한 유통 채널이 부족하다고 인식하고 있어 데이터 거래 활성화 제약으로 작용

- 다양한 데이터 플랫폼 운영에 따른 데이터 탐색비용 증가와 표준화・품질・가격산정 등 정보비대칭 발생 때문에 데이터 거래 애로 발생
- 데이터 내부사용과 외부제공에 별도 시스템 구축이 필요하나 투자를 주저, 수집 주체마다 목적이 달라 표준화와 품질 차이 발생, 데이터의 객관적 가치 평가 어려움
- 이용자가 데이터 플랫폼을 통해 데이터에 쉽게 접근・검색(편의성)하고 믿고 거래(신뢰성)할 수 있도록 유통·거래 역량 강화
- 대부분 플랫폼은 데이터 개방・공유에 중점을 두어 데이터 활용・분석인프라 제공 미흡 및 수요자 요구사항 반영, 활용성과 확산 부족
- 분야별 축적된 데이터 활용성을 고려하여, 데이터 플랫폼을 통한 수요자 맞춤형 분석・활용 지원을 통해 가시적 성과 조기 창출

○ (연계・결합) 수집・축적되는 데이터의 표준화 → 용이한 융합・활용의 핵심

- 다양한 분야의 데이터 간 연계・융합이 되지 않는다면 데이터 활용이 어렵고 데이터 가치도 떨어질 우려
- 데이터 간 연계・융합을 촉진하고 양질의 데이터 확보를 위해 데이터 수집・축적・가공 단계에서 공통기준(표준화・품질) 적용 중요
- 데이터 전 주기(표준화–수집–축적–가공–활용)의 선순환 생태계를 만들어 데이터 활용성과를 향상하기 위해 데이터 통합관리 필요성 증대

○ (데이터 확충) 공공・민간에서 수요와 활용가치가 높은 데이터의 지속적인 발굴 및 개방 촉진 필요

○ (표준화・연계) 공공・민간 데이터의 공통 표준 규격을 적용하여 AI 가공 표준 제시 및 국가 전반의 데이터 통합 검색 기반 마련

○ (거버넌스) 개별적으로 산재되는 데이터 플랫폼 중복성 방지 및 데이터 통합 관리를 위한 데이터 거버넌스 마련

○ (기계가독성) 기계(인공지능)가 데이터를 읽을 수 있는 형태로 가공하여 AI 알고리즘 개발에 적용하고 이를 활용해야 함

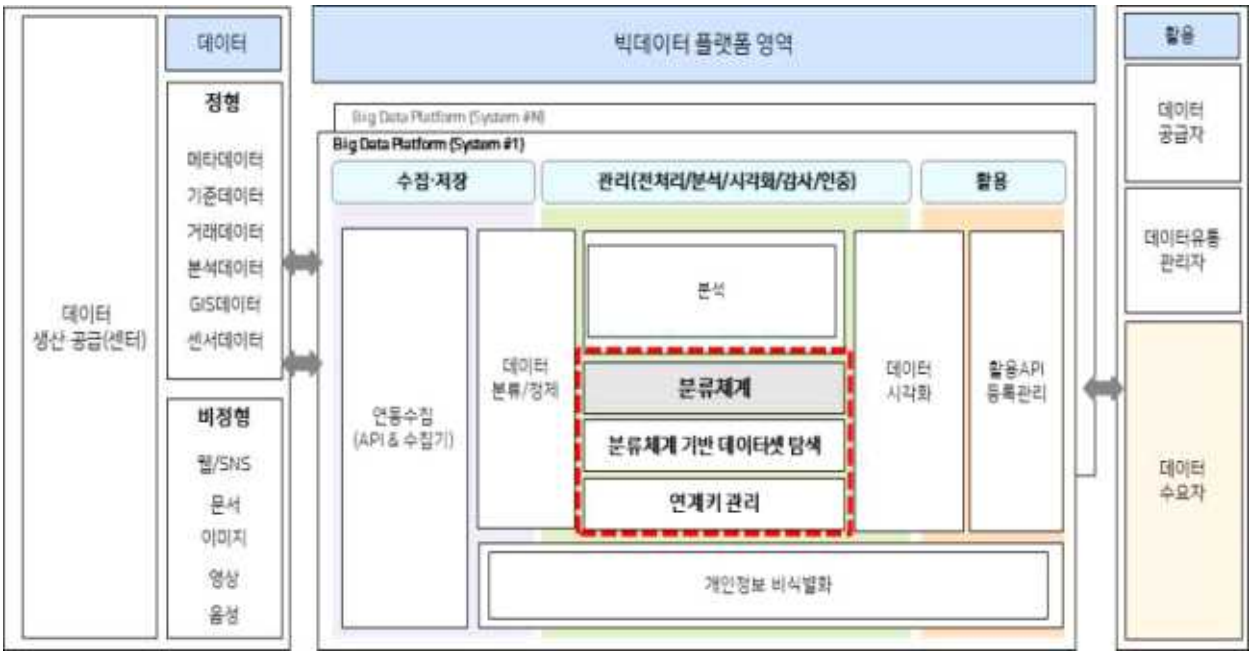
구분	원칙
관리 측면	• 인공지능 학습용 데이터의 전 생애주기의 품질을 보장해야 함 • 상시적이고 지속적인 품질개선이 가능해야 함 • 품질관리를 위해서는 데이터 품질관리를 위한 조직을 구성하고, 정해진 역할과 책임에 따라 수행해야 함 • 품질관리를 위해서는 조직의 품질관리 역량을 확보하도록 품질관리 교육 등 지원체계를 확보해야 함

데이터 측면	• 인공지능이 학습하는 데 필요한 요구사항을 충족해야 함 • 법률적 제약 없이 누구나 활용이 가능해야 함 • 학습 목적에 부합되도록 수집되어야 하며, 수집한 데이터는 중복 없이 원하는 목적에 따라 정제되어야 함 • 인공지능 학습용 데이터에 부여하는 참값(Ground Truth) 등의 라벨링 정보의 정확성이 확보되어야 함 • 인공지능 학습모델을 통해 목표로 하는 유효성이 확보되어야 함

(2) 데이터 관련 표준 분석 및 준수사항 도출

① 빅데이터 유통 플랫폼을 위한 프레임워크 - 제 1부 : 분류 체계

○ 빅데이터 분류 체계는 빅데이터 유통 플랫폼의 데이터셋을 ‘주제영역’, ‘데이터 유형’, ‘처리 유형’, ‘품질수준 유형’, ‘가격정책 유형’, ‘공간범위 유형‘, ‘발생주기 유형’ 등의 7가지 분류 유형으로 나눠 기술함



○ 분류 유형을 바탕으로 데이터 수요자, 공급자 등은 데이터셋을 검색할 수 있으며, 분류 유형을 조합하여 데이터셋을 다양하게 검색할 수 있게 함

- 데이터 유형 분류는 생산된 다양한 데이터에 대하여 생산 주체, 자료출처 및 내용의 특성별로 세분화하여 데이터의 유형으로 분류하는 것을 목적으로 설계됨.
- 이를 통해 데이터의 유형을 파악할 수 있으며, 아래의 표는 데이터 유형 분류의 하위 항목을 보여줌

<데이터 유형 분류 예시 >

구분	대분류(9)	중분류(17)	설명
정형 데이터	메타 데이터	기술 메타 데이터 (technical meta)	관리를 위해서 파악해야 하거나 통제해야 할 대상이나 항목으로 데이터 구조, 데이터 표준, 데이터 흐름, 데이터 권한 등에 대한 정보
		비즈니스 메타 데이터 (business meta)	데이터를 활용하기 위하여 파악해야 할 대상이나 항목으로 데이터를 설명하는 정보로 정보성, 주제영역, 품질수준, 다른 데이터와 연관성 등에 대한 정보
	기준 데이터	기준(master) 데이터	업무 프로세스의 중심이 되는 기준 정보 및 참조 정보(데이터 값이 참조하는 코드 정보)
	거래 데이터	거래(transaction) 데이터	기업 또는 기관의 고유한 업무 및 서비스 활동을 처리하는 정보시스템에 의해 생성, 관리되는 트랜잭션 정보
	분석 데이터	집계(summery)/통계 (statistics) 데이터	데이터를 단순 집계 정보 및 기초통계량을 생성한 정보
		고급분석(advanced analysis) 데이터	고급 분석기법을 통하여 분석한 결과 정보
	GIS 데이터	벡터/속성 (vector/attribute) 데이터	지형지물에 대한 공간적 정보를 GIS에 의하여 생성 관리되는 정보
		래스터(rester) 데이터	위성영상, 항공영상, 드론영상 등 지형지물에 대한 영상 정보
	센서 데이터	센서(sensor) 데이터	사물인터넷(IoT), 추적장치(Tracking Device), 공장자동화기기 등 각종 센서를 통하여 생성되는 정보
비정형 데이터	활동 데이터	웹(Web)/SNS 활동 데이터	의견정보(Opinion data), 웹 검색(Web search)정보 등을 포함한 온라인상에서 생성된 것으로 분석을 위하여 전처리(정재, 자연어처리 등)를 수행한 정보
	문서 데이터	간행물(publication)	연속간행물, 단행본, 보고서(논문, 연구보고서, 결과보고서) 등 문서
		매뉴얼/가이드 (Manual/Guide)	사용설명서, 운영매뉴얼, 가이드 등 문서
		법령/특허 (Law/Patent)	법령(법, 지침, 규칙 등) 및 특허관련 문서
		기타	기타(통계집류, 사료, 연혁집, 기관지, 교육자료, 회의자료, 목록류, 연설, 강연집, 전시・도감・화보집 등)
	미디어 데이터	이미지	사진, 엑스레이(x-ray), 초음파, CT, MRI 등 정보
		음성	자연의 소리, 상담내용 등의 음성 정보
		영상	CCTV, 영화, 재해상황, 뉴스, 스포츠, 다큐멘터리 등 영상 기록 정보

○ 주제영역 분류

- 주제영역 분류는 데이터에 대한 의미를 반영하고, 다수의 빅데이터 유통 플랫폼에 쉽게적용할 수 있도록 간결하고 유연한 주제영역 분류를 제시함
- 데이터 자체의 의미를 반영하는 주제영역 분류는 수요자들이 데이터의 의미를 유추하여 직관적으로 데이터를 탐색할 수 있는 체계적인 주제 분류를 구현할 수 있음
- 또한 사용자 경험을 고려하여 2단계로 계층화하여 데이터셋 검색의 사용성 제고를 도모할 수 있음
- 주제영역 분류의 개념은 인간 활동을 통해 생산되는 데이터를 기반으로 의사결정을 한다는 점에 주목한다면 주제 분류는 인간 삶과 활동을 기반으로 개념적으로 구현될 수 있으며, 이를 인간 삶의 기반, 삶의 활동, 삶의 향유라는 3가지 영역으로 구분하여 빅데이터 유통 플랫폼의 데이터를 개념적으로 분류할 수 있음
- 기존 다양한 주제영역을 검토 이후, 현재 운용 중인 다수 빅데이터 유통 플랫폼의 데이터셋을 기반으로 귀납적 분석을 통하여 빅데이터 유통 플랫폼에 적용될 수 있는 주제영역 분류를 도출함.
- 주제영역 분류는 앞서 개념화한 인간 삶의 기반에 속하는 생존영역과 제도영역, 삶의 활동에 속하는 산업과 경제영역, 그리고 삶의 향유에 속하는 문화영역 등 5가지 영역 11가지 대분류 35가지 중분류로 구분됨.

< 주제영역 분류의 중분류에 대한 설명 >

대분류(11)	중분류(35)	설명
문화	문화	각 분야별 축제, 이벤트, 공연 등의 데이터
	관광	다양한 여행 활동 데이터
	체육	다양한 체육 활동 데이터
교육과학	교육	교육활동 관련 데이터
	연구	연구 활동 관련 데이터
	과학기술	과학기술 개발 등에 관한 데이터
교육통신	교통	교통시설, 교통이용 등에 관한 데이터
	물류	물류의 이동 등 물류활동에 관련된 데이터
	통신	통신사의 기지국 정보를 기반한 위치 관련 데이터
경제금융	기업	기업의 경영활동으로 생산되는 데이터
	금융	개인 및 기업 신용관련 정보로 금융자산 등 데이터
	부동산	부동산의 위치, 시세 등 부동산에 관한 데이터
	무역	산업별 무역 활동 등에 관한 데이터
건설에너지	에너지	에너지의 생산, 소비 등에 관한 데이터
	토목건설	국토 개발에 관한 데이터
농수산식품	농식품	농산물 및 식품에 관한 데이터
	산림	산림에 관한 데이터
	해양수산	해양수산에 관한 데이터
제조소비	소비재	소비재 제품의 생산 활동 등에 관한 데이터
	생산재	생산재 제품의 생산 활동 등에 관한 데이터
	유통소비	개인의 소비활동에 관한 데이터
건강의료	건강	개인의 건강관링 관한 데이터
	의료	질병 및 의료 활동 등에 관한 데이터
재난안전	재난	자연재해, 안전사고 등에 관한 데이터
	안보	국가에 중대한 영향을 주는 재난 데이터
	치안	국민 생활 치안에 관한 데이터

환경자원	기상	기상관측정보 등 기상에 관한 데이터
	동식물	자연 생태계의 동식물에 관한 데이터
	국토	농지, 초지, 산지 및 해양 등 국토에 관한 데이터
	수자원	수자원에 관한 데이터
	광물자원	광물자원에 관한 데이터
공공/행정	행정	정부 정책 등 행정에 관한 데이터
	사법	법률에 관한 데이터
	정치외교	정치 및 외교에 관한 데이터
	재세정	국가 및 지자체의 제정 및 세제 등의 관한 데이터

○ 데이터 유형 분류

- 데이터 유형 분류는 생산된 다양한 데이터에 대하여 생산 주체, 자료출처 및 내용의 특성별로 세분화하여 데이터의 유형으로 분류하는 것을 목적으로 설계됨.

○ 가공 유형 분류

- ‘가공 유형’ 분류는 생산된 RAW 데이터의 전처리 수행에 따라 데이터의 가공 유형으로 분류하는 것을 목적으로 설계됨.

< 가공 유형 분류 예시 >

가공 유형	설명
무처리	플랫폼에서 최초로 등록한 상태의 데이터
정제 처리	등록된 데이터를 정제하여 등록한 데이터
융복합처리	등록된 데이터를 융・복합하여 등록한 데이터
태깅처리	음성, 이미지, 영상, 문서 등을 분석하여 세부 주제에 대하여 태깅 처리한 데이터

○ 품질수준 유형 분류

- ‘품질수준 유형’ 분류는 데이터의 품질진단 결과에 따른 품질수준 유형으로 분류하는 것을 목적으로 설계됨

< 품질수준 유형 분류 예시 >

품질수준	설명
1등급	데이터 품질 정합률 99.0% 이상
2등급	데이터 품질 정합률 98.0% 이상
3등급	데이터 품질 정합률 97.0% 이상
4등급	데이터 품질 정합률 96.0% 이상
5등급	데이터 품질 정합률 95.0% 이상

○ 가격 유형 분류

- ‘가격 유형’ 분류는 데이터에 대한 가격정책으로 분류하는 것을 목적으로 설계됨

< 가격 유형 분류 예시 >

가격유형	설명
무상	무상으로 제공되는 데이터
유상	유상으로 제공되는 데이터

○ 공간범위 유형 분류

- ‘공간범위 유형’ 분류는 데이터가 가지고 있는 공간적 범위를 지리적 공간 범위 유형으로 분류하는 것을 목적으로 설계됨.

< 공간범위 유형 분류 예시 >

공간 범위 유형	설명
전국	데이터셋의 데이터가 전국적인 데이터
시도	데이터셋의 데이터가 특정 시도의 데이터
시군구	데이터셋의 데이터가 특정 시군구의 데이터
읍면동	데이터셋의 데이터가 특정 읍면동의 데이터
특정 공간	데이터셋의 데이터가 특정 공간적 범위의 데이터

○ 발생주기 유형 분류

- ‘발생주기 유형’ 분류는 데이터셋의 데이터가 생성되는 시간적 주기를 발생주기 유형으로 분류하는 것을 목적으로 설계됨

< 발생주기 유형 분류 예시 >

발생 주기 유형	설명
년/분기	데이터의 생성주기가 년 또는 분기, 반기 주기로 발생
월	데이터의 생성주기가 월 주기로 발생
일	데이터의 생성주기가 일 주기로 발생
실시간	데이터의 생성주기가 실시간 주기로 발생
상시	데이터의 생성주기가 상시(준 실시간) 주기로 발생

② 빅데이터 유통 플랫폼을 위한 프레임워크 - 제 2부 : 분류 체계 기반 데이터셋 탐색 기능

○ 서로 다른 분류 체계를 쓰는 빅데이터 유통 플랫폼 환경에서는 각 플랫폼 간 데이터셋활용이 어렵다는 문제점이 있음.

○ 이 표준에서는 이를 해결하기 위해 빅데이터 분류 체계기반 데이터셋 탐색 방안과 기능들을 정의함.

- 빅데이터 분류 체계를 쓰는 플랫폼에서 활용하고 있는 데이터셋으로부터 의미태그를 생성함
- 의미태그는 해당 빅데이터 분류 체계의 중분류가 가진 의미를 기준으로 생성되며, 다른 분류 체계를 쓰는 이기종의 플랫폼은 보유한 데이터셋에서 자신의 분류 체계 중분류의 의미를 기준으로 의미태그를 추출함.
- 이기종 플랫폼의 의미태그와 빅데이터 분류 체계의 플랫폼의 의미 태그간의 거리를 측정하고, 의미의 유사성 계산을 통하여 유사한 의미태그 데이터셋을 탐색할 수 있음.
- 빅데이터 분류 체계 기반의 데이터셋 탐색 기능은 수요자에게 자기 플랫폼에 존재하는 데이터셋뿐만 아니라 이종의 플랫폼이 보유하고 있는 데이터셋의 정보를 빅데이터 분류 체계의 중분류 수준에서 탐색할 수 있게 함

○ 유사도 측정(Similarity Measure)

- 유사도(Similarity)란 두 데이터가 얼마나 같은지 나타내주는 척도로 다음의 표와 같은 유사도 측정법을 정의하고, 유사도는 데이터 과학에서 클러스터링(clustering), 분류(classification)의 기반이 되는 것이며, 이를 통해서 더 복잡한 것들을 측정할 수 있도록 함.
- 어떤 데이터가 n차원 상의 벡터로 표현된다면, 두 데이터의 유사도는 n차원상에서 두 벡터 사이의 거리(Distance)라고도 볼 수 있으며, 아래의 표에서 제시한 유사도 측정법을 바탕으로 측정된 유사도 결과를 의미 태그 생성 과정에 적용하도록 함

< 유사도 측정법 >

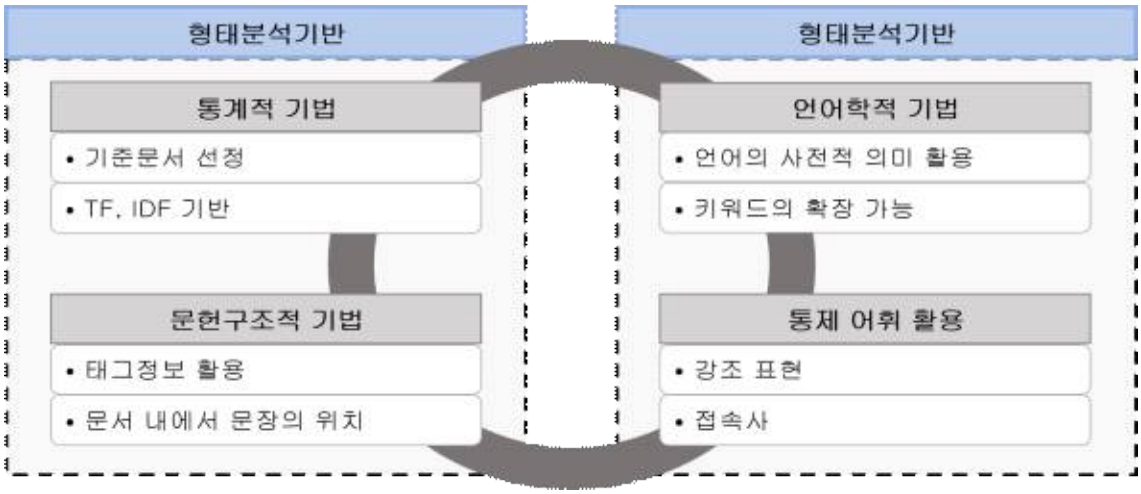
측정법	설명
Euclidean Distance	실제 거리라는 의미에 가장 부합하는 측정법으로, 두 벡터 간의 직선거리를 측정함.
Manhattan Distance	맨해튼과 같이 정사각형 나눈 곳에서 두 점 사이의 거리를 측정하는 방법. 두 벡터의 Cartesian coordinate의 차의 절닷값의 합을 측정함.
Minkowski Distance	민코프스키 거리는 유클리드 거리와 맨해튼 거리를 일반화하는 것임.
Cosine Similarity	코사인 유사도는 두 벡터가 이루는 각도를 통해 유사도를 측정하는 방식. 두 벡터가 이루는 각이 작을수록 유사도가 높은 것이고, 각이 클수록 유사도가 작다고 생각하는 것임.

○ 키워드 추출 프로세스

- 빅데이터 유통 플랫폼 환경에서는 다양한 문서의 구조화 및 저장이 필요하며, 문서 및 데이터에 담겨 있는 정보를 추출하기 위해서 가장 우선해야 하는 것이 키워드 추출로, 키워드 추출 프로세스는 다음과 같이 4가지 단계로 구분함.
  a) 첫 번째 단계는 키워드를 추출하기에 앞서 문서를 처리하는 선처리 단계. 문서를 지정된 포맷으로 준구조화하거나, 또는 형태소 분석기를 통하여 구성성분을 분해하는 작업을 수행함
  b) 두 번째 단계는 후보군을 선정하는 단계로 앞선 단계에서 구분된 각 단어들 중 키워드가 될 가능성이 있는 단어들만 추려내는 작업으로, 키워드 추출 알고리즘에서 가장 많이 사용되는 TF(Term Frequency), IDF(Inverted Document Frequency) 등도 활용함.
  c) 세 번째 단계는 선정된 후보군에 대하여 키워드 지표를 측정하는 알고리즘을 적용하는 단계.
  d) 네 번째 단계는 신경망 모델을 활용하여 키워드를 추출함.

○ 의미분석 기반의 키워드 추출

- 의미분석 기반의 키워드를 추출하기 위해서는 키워드일 가능성이 높은 단어의 특징을 정리하는 것이 선행되어야 함.
- 의미분석 기반의 키워드 추출 방법은 크게 통계적 기법, 문헌 구조적 기법, 언어학적 기법, 통제 어휘 활용의 4가지 기법을 방법론 측면에서 다음의 그림과 같이 정의함. 이와 같은 4가지 키워드 추출 방법을 본 표준의 키워드 추출에 적용하도록 함.



(의미분석 기반 키워드 추출 방법론)

○ 통계적 기법

- 통계적 기법은 첫 번째 단계로 선처리 단계임. 다양한 분야와 주제를 포괄할 수 있는 기준문서를 선정하고, 문서에서 숫자 및 특수문자를 제거함.
- 두 번째 단계는 알고리즘을 적용하는 단계. 단어의 사용 빈도가 높을수록 그리고 특정 문서에서만 사용될수록 키워드일 가능성이 높다는 특성을 이용함.
- 세 번째 단계는 보정하는 단계임. 문서의 길이가 길거나, 문장의 길이가 길면 사용하는 단어의 수가 많아지며, 문서의 길이에 따른 보정값을 적용할 필요가 있음.

○ 문헌구조적 기법

- 문헌구조적 기법에서는 기본적으로 문서의 태그 정보를 활용함. 문서의 형태에 따라 다르지만, 문서가 제목, 소제목, 본문 등으로 구성되어 있다면 본분에서 등장하는 단어보다 제목, 소제목에서 사용된 단어가 키워드일 가능성이 높음.
- 따라서 태그 정보를 활용하여 해당 단어가 제목에서 쓰였는지, 소제목에서 쓰였는지를 판단하면 키워드로 결정하기 쉽게 됨. 또한 문단 및 문장 내에서의 단어의 위치정보도 활용함.

○ 언어학적 기법

- 형태분석 기반의 방법을 보완할 수 있는 방법으로, 문서의 의미를 반영함. 예로 한 대상에 대하여 설명하는 문서가 있다고 가정해보면 대상에 대한 언급보다는 대상의 의미를 나타내는 단어, 대상과 관련이 있는 단어, 대상의 하위 개념을 나타내는 단어들이 주를 이루게 되는데 이 경우 대상에 대한 설명임에도 분구하고 대상을 나타내는 단어는 키워드로 선정되지 않을 수 있음.
- 이런 경우는 문서의 의미를 파악하지 않고, 형태에만 국한하여 판단했기 때문임. 단어의 사전적 의미를 반영하는 방법과 시소러스(Thesaurus)를 활용하는 방법을 고려할 수 있음.

○ 통제 어휘 활용

- 특정 단어 뒤에 오는 단어들은 그 중요성이 다른 단어보다 크다고 할 수 있음. 그 중요성을 결정지어줄 수 있는 단어를 통제 어휘라 칭함.
- 통제 어휘에는 해당 문장에 중요도를 부여하는 강조 표현, 접속사, 부정 표현 등이 포함됨.
- 강조 표현에는 다양한 형태가 있음. 특정 단어만을 강조하는 표현이 있고, 문장 전체를 강조하는 표현이 있으며, ‘무척’, ‘매우’ 등과 같이 극대의 의미를 갖는 표현이 있고, 겨우’, ‘고작’과 같이 극소의 의미를 갖는 강조 표현이 있음.

○ 태그 추출 기법

- 빅데이터 유통 플랫폼의 수요자, 공급자 등은 데이터를 분류하고 레이블링(Labeling)하여 만들어진 데이터 분류 시스템을 만들고 있음. 이때 자유롭게 키워드를 선택하고 모든 미디어, 텍스트 등의 항목에 태그를 붙일 수 있음.
- 태깅(Tagging)은 단어나 문구와 같이 자유로운 형태의 텍스트를 아이템에 추가하는 과정을 의미함

○ 태그 활용성

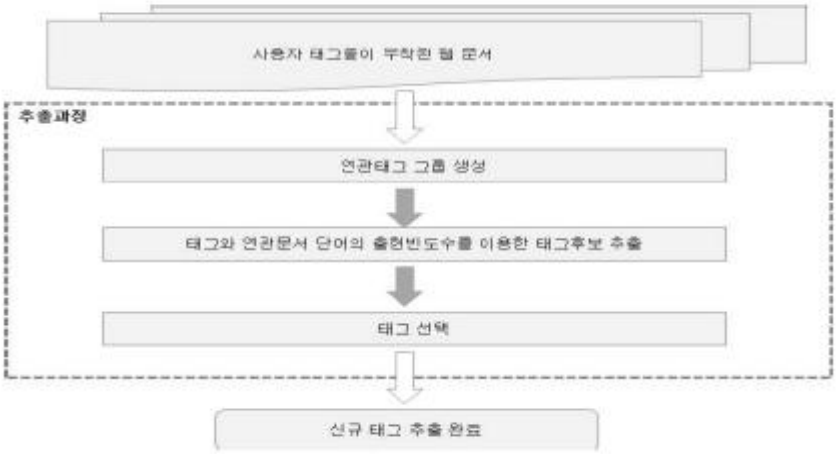
- 태그란 블로그, 사진, 동영상, 문서, 웹 페이지 URL 같은 웹 콘텐츠를 잘 설명할 수 있는 간단한 키워드 혹은 단어를 의미함.

○ 태그 추출 기법 순서도

- 태그 추출을 위한 기법 순서도는 아래의 그림처럼 태그들이 부착된 웹 문서의 집합을 사용하고, 추출 과정은 연관 태그 그룹 생성, 태그와 연관문서 단어의 동시 출현 빈도수를 이용한 태그 후보 추출, 태그 선택 후 신규 태그 추출 완료 등을 거쳐 진행함.

  a) 연관태그 그룹 생성 : 웹 문서에 태깅(Tagging)을 할 때 문서에 대한 사람의 생각이 연결된 태그 단위로 나열됨. 나열된 태그들은 의미적으로 연관적인 관계를 갖게 되는데 태그-문서 행렬을 이용하여 태그들의 연관성을 분석함.

  b) 태그 선택 방법 : 연관 태그 그룹 생성과 출현 빈도수에 의한 가중치 기법을 통해 가중치가 높은 단어를 추출하고, 신규 태그를 선택함.



(태그추출 기법 순서도)

○ 메타정보를 이용한 의미 태그의 생성

- 이종의 빅데이터 유통 플랫폼에 저장된 데이터셋의 명칭, 설명정보 등을 이용하여 의미 태그를 생성하도록 다음의 기능을 제공함.

○ 키워드 생성 기능

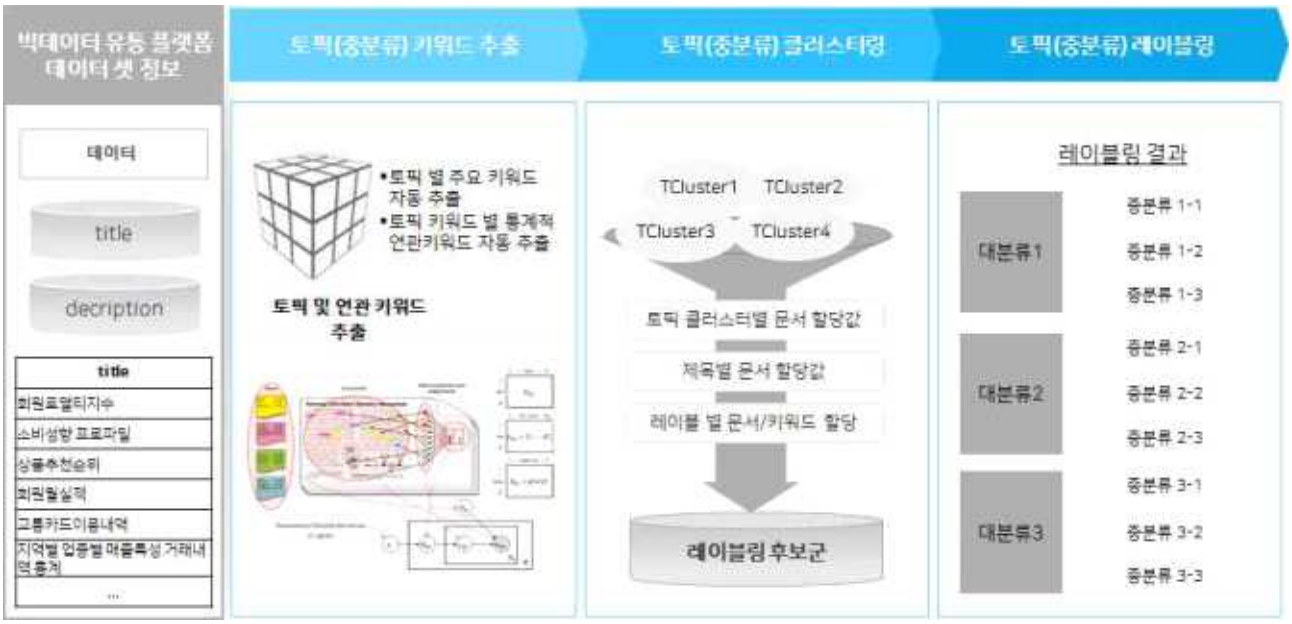
- 빅데이터 유통 플랫폼에 저장된 데이터셋의 명칭, 설명정보 등을 수집하여, 자연어 처리 기술의 형태소분석을 통하여 키워드를 추출하는 기능을 제공함.

○ 키워드의 군집화 기능

- 추출한 키워드를 군집분석을 통하여 주제 분류 후보군을 생성하는 기능을 제공함.

○ 레이블링 기능

- 생성된 주제 분류 후보군과 빅데이터 분류 체계의 주제와의 매핑과정을 통하여 빅데이터 분류 체계에 의미 태그를 부여하는 레이블링 기능을 제공함.



(분류 체계별 의미 태그 생성)

- 위의 그림의 분류 체계별 의미 태그 생성과정은 각 빅데이터 유통 플랫폼의 데이터셋의 명칭과 설명정보 등을 이용하여 연관 키워드를 추출한 후, 군집화 및 레이블링의 각 단계를 거치면서 분류 체계 중분류의 의미 태그를 생성하게 됨.
- 의미 태그 생성 결과는 아래의 표와 같이 주제영역 중분류의 의미 태그에서 예시로서 볼 수 있음.

< 분류 체계별 의미 태그 예시 >

대분류	중분류	의미 태그(키워드)
재난안전	재난	안전 붕괴 교육 조사 손해 사망 검사 구조 지휘 경호 훈련 조사 공사 추락 수사 표시 평가 지도 신고 부담 손실 침수 중독 소란 탈취
문화체육	체육	여행 놀이 개발 관광 유치 교류 게임 축제 질서
경제금융	금융	조회 저축 투자 파산 공개 미납 거래 보증 출자 가입 감독 개혁 조정 정리 인수 제재 매각 징수 재정 감시 부담 결제 송금 소멸 매매 고시 평가 장려 예측 공모 양도 청약
경제금융	기업	노동 고용 촉진 차별 평등 허가 폐쇄 감독 단축 파견 검사 안전 운동 보호 교육 협력 탄압 해산 교섭 은퇴 해고 연수 출퇴근 건강관리
행정법률	행정	세액 징계 보고 감사 청구 처분 공개 고시 규제 정책 기록 재정 과세 교류 검사 퇴출 신고 징수 체납 개발 행정 지원 협력 납세 감정 공제 등기 규율 양도 시찰 인증 산정 결산 우대 검역 집행 감시 제한 소환 등급 표시 해제 법인 허가 감리 기간

○ 빅데이터 분류 체계에 데이터셋 매핑 기능

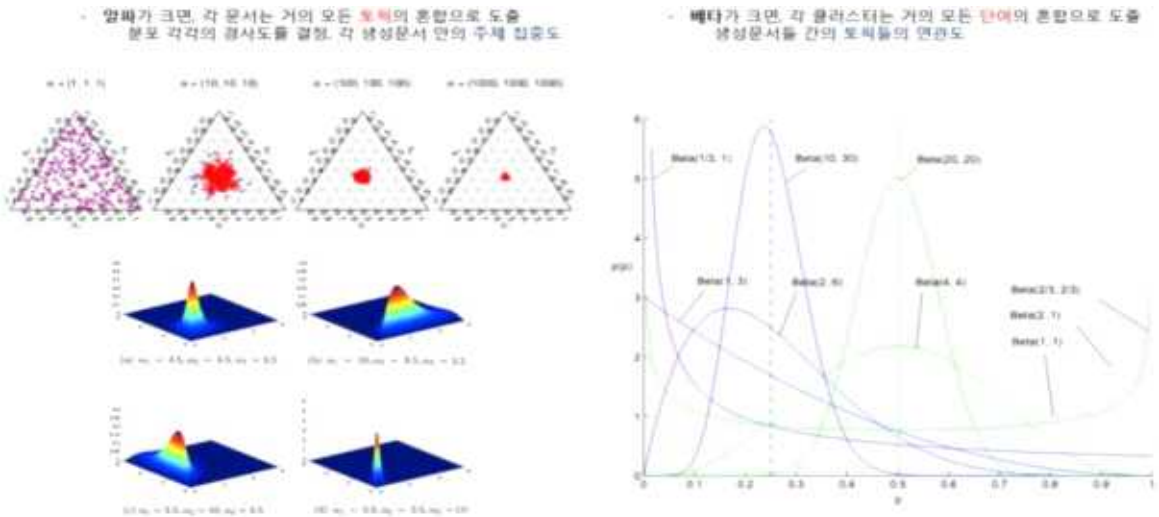
- 빅데이터 분류 체계에 부여된 의미 태그와 다른 분류 체계를 사용하는 플랫폼에 게시된 데이터셋의 의미 태그와의 매핑을 하도록 다음의 기능을 제공함.

○ 의미 태그 간의 거리 산출 기능

- 빅데이터 분류 체계에 부여된 의미 태그와 다른 분류 체계를 사용하는 플랫폼에 게시된 데이터셋의 의미 태그 간에서, 의미 거리를 산출하는 기능을 제공함.

○ 거리 기반의 매핑 기능

- 빅데이터 분류 체계에 부여된 의미 태그와 다른 분류 체계를 사용하는 플랫폼에 게시된 데이터셋의 의미 태그 간에서 산출된 의미 거리를 근거로 매핑 기능을 제공함.
- 이를 위해 구체적으로 두 가지 의미 태그들의 개별 거리를 합산 평균하는 방법을 적용할 수 있음.
- 아래의 그림에서 제시한 매핑 방안은 의미 태그(데이터셋 제목 및 설명) 분석을 통해 특정 주제에 대한 설명 여부(알파↑), 메타데이터 간 유사도(베타↑) 등에 해당하는 알고리즘을 반복적으로 수행하여 진행됨.
- 본 매핑 방안은 아래의 표의 분류 체계별 의미 태그에 적용하여, 실제 데이터셋의 의미 태그를 의미 거리에 근거하여 분류 체계의 의미 태그와 매핑하도록 함.
- 예를 들어, 앞서 제시한 의미 태그 생성과정을 거쳐 예금, 금리 등의 의미 태그가 추출되었다면, 매핑 과정을 통해 아래의 표의 대분류 ‘경제금융’, 중분류 ‘금융’의 분류 체계에 따른 의미 태그와 매핑되는 결과를 찾아낼 수 있다.



(의미태그 기반 매핑 방안)

○ 의미 태그와 분류군 간의 거리 산출 기능

- 빅데이터 분류 체계에 부여된 의미 태그와 이종 빅데이터 유통 플랫폼에 게시된 데이터 셋의 의미 태그 간의 거리를 산출하는 기능을 제공함.

○ 의미 태그와 매핑 기능

- 구체적으로 두 가지 분류군들의 키워드와 그 키워드들의 의미 태그 군들의 거리를 합산 평균하는 방법을 적용할 수 있음.
- 이 방법은 아래의 표의 분류군과 개별 데이터 매핑 방안과 같이 데이터 단위 의미 태그(데이터 제목 및 설명) 분석과 특정 주제가 속한 분류군 간의 연관성의 정도를 반복된 알고리즘 수행을 통하여 생성됨.

○ 크로스 플랫폼(Cross Platform) 개요

- 빅데이터 분야에서는 다중 데이터를 처리할 수 있는 이종 플랫폼 간의 분석이 가능한 기술들이 필요하고, 각 플랫폼이 다양한 데이터의 효율적인 연계 요구가 증대되고 있음. 이러한 요구에 크로스 플랫폼상에서 데이터가 처리될 필요가 있음.
- 크로스 플랫폼(Cross Platform)은 데이터를 처리하는 개별 플랫폼들을 유기적으로 통합하는 과정을 거쳐 여러 형태의 데이터들에 대한 분석 작업을 수행할 수 있도록 하는 하나의 플랫폼으로, 데이터셋

검색방안에 적합한 플랫폼 모델임.

○ 크로스 플랫폼 구조

- 데이터를 제공하는 플랫폼 단계, 제공받는 데이터를 연결하는 미디에이터(mediator) 단계, 데이터를 요청하는 사용자가 이용하는 사용자 인터페이스(interface) 단계로 구성함.

○ 데이터를 제공하는 플랫폼 단계

- 데이터를 제공하는 플랫폼은 크로스 플랫폼에 등록되어 데이터 제공에 동의를 한 기관을 의미함
- 기관은 본인들의 플랫폼을 공유하고, 플랫폼 내 DB의 구조적 업데이트를 실시간으로 알림.
- 개인의 사생활이나 노출될 여부가 존재하거나 법적으로 보호되어야 할 민감정보를 가진 데이터는 개인 식별정보, 인증정보, 신용정보와 같은 중요한 데이터는 비식별화를 통해서만 해당 단계를 진행함.

○ 미디에이터 단계

- 플랫폼을 등록한 기관이 신뢰하는 미들웨어로 사용자의 요청으로 확인하고자 하는 기관의 데이터를 받게 되며, 추가로 데이터의 연결을 수행할 때 서로 다른 기관의 데이터를 관계데이터(relation data)로 생성하고, 관계가 맺어진 데이터를 제공함.

○ 사용자 인터페이스 단계

- 사용자 인터페이스는 사용자 그룹의 요청을 입력받거나, 요청을 통해 연결된 정보의 View를 확인하거나 미디에이터로부터 전달받은 정보를 분석하여 시각화를 진행하는 소프트웨어를 의미함.

○ 시각화

- 크로스플랫폼의 시각화 기능은 데이터셋 탐색 결과를 데이터 수요자가 확인할 수 있는 기능으로 이 표준에서 정의한 데이터셋 탐색 기능을 구체화함.
- 시각화는 사용자 인터페이스에서 구현된 기능으로 크게 아래의 표와 같이 9가지 기능으로 나눌 수 있음.
- 이에는 지역 검색 기능, 정보를 요청하고 연결하는 정보 검색 기능, 연결한 정보와 연관된 데이터를 추가로 검색하는 기능, 연결한 데이터 범위내에서 검색 가능한 영상 정보 검색 기능, 검색된 내용을 기억하는 북 마크 기능, Map을 이용한 시각화 기능, 검색 가능한 영상을 재생 가능한 영상 시각화 기능, 실제 연결된 정보를 확인하는 View 기능, 검색한 정보를 초기화하는 닫기 기능 등이 있음.

< 사용자 인터페이스에서 구현된 기능 >

기능명	설명
지역 검색 기능	분석하고자 하는 지역을 찾고자 할 때 도로명, 지번 주소, 특정 키워드를 통해 검색할 수 있는 기능
정보 검생 기능	1차적으로 데이터를 연결함을 의미하며, Map 시각화 기능과 View 시각화 기능에 필요한 정보를 검색하고 연결하는 기능
연관 정보 검색 기능	2차적으로 데이터를 연결함을 의미하며, 1차로 연결된 데이터와 연관 있는 데이터를 불러와 연결하는 기능
영상 검색 기능	정보 검색 수행 시 변경 내 확인 가능한 영상 정보를 파악하고 시각화 하는 기능
북마크 기능	정보와 연관 정보를 검색하였을 때 이전의 검색 결과를 기억하고 불러올 수 있는 기능
MAP 시각화 가능	주소 및 거리 연결 방법을 이용하여 정보를 묶는 결과를 표현하는 기능
Wiew 시각화 기능	실제 연결된 데이터 및 연관이 있는 데이터 간에 공통으로 포함이 되는 정보의 자료를 확인할 수 있는 기능
Video 시각화 기능	사용자가 정보를 검색했을 때 설정한 범위 안의 영상정보를 재생하여 표현할 수 있는 기능
닫기 기능	데이터 처리 과정을 끝내고 나서, 데이터 내 민감정보나 개인정보를 보존 및 수집할 수 없도록 악용 가능성을 방지하는 기능

○ 데이터셋 탐색 신뢰 구축

- 데이터셋 탐색 결과 해당 데이터셋에 대한 신뢰도를 확인할 필요가 있음. 이에 데이터셋 탐색 신뢰 구축단계가 필요함.
- 이를 위해 먼저 적절한 데이터셋을 찾은 후, 해당 데이터를 신뢰할 수 있는지를 묻게 됨.
- 드물게는 깨끗하고 신뢰할 수 있는 정형화된 데이터셋을 찾기도 하지만, 대부분의 경우 데이터의 신뢰 여부를 판단해야 함.
- 신뢰 구축단계는 데이터 품질, 이력(또는 유래), 관리의 3가지를 기반으로 진행됨.
  a) 데이터 품질은 데이터셋이 얼마나 완전하고 깨끗한가?
  b) 이력(또는 유래)는 데이터가 어디에서 유래되었는가?
  c) 관리는 데이터셋을 누가 만들었고 왜 만들었는가?
- 이상과 같은 3가지에 대한 신뢰를 확보해야 함.

○ 데이터 품질

- 데이터 품질은 넓고 복잡한 문제이기 때문에, 데이터의 관련 규칙 준수(예를 들면, 고객명 필드는 비어있으면 안 된다)에서부터 복잡한 규칙까지 고려해야 함.
- 가장 대표적인 데이터 품질 규칙은 아래의 표와 같이 완전성, 데이터 유형, 범위, 포맷, 집합의 크기, 선택성, 참조무결성 등이 있음.

< 데이터 품질 규칙 >

데이터 품질 규칙 요소	규칙 정의
완전성	필드는 비어 있지 않음.
데이터 유형	필드가 올바른 유형이다. 예를 들면, 나이는 숫자여야 함.
범위	필드의 값은 특정 범위에 속해야 한다. 예를 들면, 나이는 0과 125 사이여야 함.
포맷	필드가 특정 포맷(format)으로 돼 있어야 한다. 예를 들면, 한국 우편번호는 5자리여야 함.
집합의 크기	필드에 포함된 고윳값의 개수가 특정한 수여야 함.
선택성	필드의 값은 고유한 값이어야 한다. 예를 들면, 고객 리스트의 고객 ID는 고유한 값이어야 함.
참조무결성	필드 값이 참조 값 리스트에 있음. 예를 들면, 모든 고객 상태 코드는 허용된 값 중에 있고, 주문 리스트의 각 고객 ID는 고객 리스트에도 있는 값임.

- 데이터 품질을 확인하는 가장 일반적인 방법은 데이터 프로파일링임.
- 각 필드의 데이터를 읽어 비어 있는 필드 수(완전성), 고유한 값 개수(집합 크기), 고유한 값 비율(선택성)과 같은 측정 지표뿐만 아니라 데이터 유형, 범위 등을 확인하고 참조무결성 검사까지 진행하는 것이 데이터 프로파일링임.

○ 이력(또는 유래)

- 데이터 품질검사는 데이터가 얼마나 좋은지 확인하고, 이력은 데이터가 어디에서 왔는지 알려줌.
- 데이터셋은 다양한 프로그램, 프로그래밍 언어, 스크립트, 도구 등으로 구성된 여러 단계를 거쳐 생성될 수 있기 때문에, 기술적 이력을 표현하는 것이 어려울 수 있음.
- 기술 이력에 관해서는 밀도, 변환 등 2가지 측면을 고려해야 함.

< 밀도와 변화 표현 >

밀도		변환	
데이터셋 수준밀도	필드 수준밀도	정규화된 표현	원래대로 표현
여러 데이터셋 간의 이력관계는 일반적으로 방향성 있는 그래프 형태로 포착하고 표현	각 필드의 이력은 방향성 있는 그래프의 형태로 포착하고 표현하는데, 이때 각 변환 작업은 별도의 노드로 변환	모든 변환을 공통 형태로 표현	모든 변환을 원래 언어나 스크립트로 표현하는 방법

③ 빅데이터 유통 플랫폼을 위한 프레임워크 - 제 3부 : 연계키 관리 기능

○ 이 표준에서는 다양한 분야에서 공통으로 활용할 수 있는 연계키 중, 활용도가 높은 ‘사람’, ‘기업’, ‘공간’ 및 ‘시간’ 등의 영역으로 구분하여 발굴된 연계키 항목을 제시함.

- 등록된 표준연계키는 표준연계키사전에서 관리한다. 한편 표준연계키의 효율적인 이용을 위해 표준연계키 신청, 등록, 검색 등의 기능과 표준연계인덱스의 생성, 검색 등의 기능을 정의함.

○ 표준연계키 필요성

- 빅데이터 유통 플랫폼 환경에서는 각 플랫폼이 자체 용어사전을 적용하여 다른 플랫폼과 상이한

속성(칼럼) 논리명을 사용함.

- 서로 다른 속성 논리명을 사용하는 환경에 서는 서로의 데이터셋 정보를 파악하기 쉽지 않아 데이터 연계가 어려움.
- 빅데이터 유통 플랫폼의 다양한 분야에서 공통으로 활용할 표준용어사전 기반 연계키를 발굴하여, 이를 기반으로 표준연계키를 추출하고 표준연계키사전, 표준연계인덱스 등을 구축함.
- 그 결과 서로 다른 플랫폼별 용어사전의 호환성을 확보할 수 있고, 플랫폼 간 데이터를 연계할 수 있음.

(표준 연계키, 표준연계키사전, 표준연계인덱스)

- 위의 그림은 표준연계키인 ‘전화번호’를 기반으로 데이터를 연계하는 예시임.
- 전화번호인 표준연계키를 통해 각 데이터셋인 데이터A와 데이터B의 전화번호를 서로 연계하여 사용할 수 있음.

(표준 연계키를 이용한 데이터 연계 예시)

○ 표준연계키 발굴 범위 사례

- 다양한 분야에서 공통으로 활용할 연계키는 먼저 ‘사람’, ‘기업’, ‘공간’ 및 ‘시간’ 영역에서 발굴함. 이들 영역은 그동안 활용도가 높았던 분야로, 이들을 중심으로 우선 표준연계키가 발굴되어 활용하도록 함. 예를 들어 사람영역에서는 표준연계키를 성, 연령, 연령대, 이메일, 거주지 주소, 우편번호 등으로 볼 수 있음. 이와 같이 각 영역에서 활용도를 기초로 표준연계키를 발굴하도록 함. 아래의 표는 표준연계키의 발굴 범위에 대한 예시임.

< 표준연계키 발굴 범위 사례 >

영역	표준 연계키
사람	성, 연령, 연령대, 이메일, 거주지 주소, 우편번호
기업	사업체 명칭, 사업체 전화번호, 사업체 주소, 법인번호, 업종
시간	시간, 시각, 날짜, 연도
공간	주소, 좌표, 행정구역, 국가기초구역, 건물 (집합건물, 빌딩)

○ 표준연계키 적용 영역

- 표준연계키에 적용할 영역은 그동안 활용도가 높았던 ‘사람’, ‘기업’. ‘공간’, ‘시간’ 등의 범위에서 정의함.

a) ‘사람’ 영역에 해당하는 표준연계키는 사람의 특성을 식별할 수 있는 성별코드, 연령, 연령대 코드 등의 항목을 이용한다. 이에 대한 표준연계키 예시는 부록에서 확인할 수 있음.

b) ‘기업’ 영역에 해당하는 표준연계키는 기업을 식별할 수 있는 사업체 명칭, 전화번호 등의 항목을 이용함.

c) ‘시간’ 영역에 해당하는 표준 연계키는 시간을 구분하는 연도, 연/월/일 등의 항목을 이용함.

d) ‘공간’ 영역에 해당하는 표준연계키는 지역을 식별할 수 있는 주소, 좌표, 행정구역 등의 항목을 이용함.

이상과 같은 ‘사람’, ‘기업’, ‘시간’, ‘공간’ 등 영역의 표준연계키 적용 예시는 부록 I을 통해 확인할 수 있음.

○ 표준연계키 관리

○ 표준연계키 관리 개요

- 빅데이터 유통 생태계 구성원인 빅데이터 유통 플랫폼 제공자, 데이터 수요자, 데이터 공급자 등이 데이터를 연계•활용하기 위해서는 표준연계키 관리가 필요함. 먼저 각 플랫폼으로부터 연계키를 추출하고 이들을 표준연계키로 등록한다. 표준연계키는 표준연계키 사전에 등록하여 표준연계키사전을 구축함. 그리고 각 플랫폼의 데이터 연계•활용을 위해 데이터셋과 표준연계키 사전과의 매핑 결과인 표준연계인덱스를 구축함. 또한 각 플랫폼은 표준연계키를 관리 또는 사용하기 위해 표준연계키 신청, 등록, 생성등의 기능과 표준연계 인덱스의 생성, 검색 등의 기능을 수행함.

○ 표준연계키사전 항목

- 표준연계키 등록 과정을 거친 뒤 구축된 표준연계키 사전은 아래 그림과 같은 항목들로 구성됨. 표준연계키항목, 데이터 상세설명, 영역 등이 하나의 레코드로 저장된 형태임.

(표준연계키사전 항목 예시)

표준연계키 명	표준연계키 영문명	표준연계키 항목명	표준연계키 항목영문명	데이터 타입	데이터 길이	영역ID(FK)
회원성별코드	MMBR_SEXDSTN_CD	성별코드	SEXDSTB_CODE	CHAR	1	C01(사람)
신청인성별코드	APPCN_SEXDSTN_CD	성별코드	SEXDSTB_CODE	CHAR	1	C01(사람)
기업명	ENTRPRS_NM	기업명	ENTRPRS_NM	CHAR	500	C02(기업)
업체명	ENTP_NM	기업명	ENTRPRS_NM	CHAR	500	C02(기업)
년월	YN	년월	TM	CHAR	6	C03(시간)
수집년월	COLCT_DT	년월	TM	CHAR	6	C03(시간)
기준년월	STRD_YM	년월	TM	CHAR	6	C03(시간)
법정동코드	LEGALDONG_CD	법정동코드	LEGALDONG_CD	VARCHAR	10	C05(공간)
지적법정동코드	LNRGS_LEGALDONG_CD	법정동코드	LEGALDONG_CD	VARCHAR	10	C05(공간)
행사장소법정동코드	EVEN_PLAC_LEGALDONG_CD	법정동코드	LEGALDONG_CD	VARCHAR	10	C05(공간)

○ 표준연계키사전 구축절차

- 연계키 선정을 통한 표준연계키의 구축 절차는 다음과 같음.

a) 연계키 영역을 기반으로 각 플랫폼의 데이터셋 속성에서 연계키를 발굴함.

b) 발굴된 연계키 중, 각 플랫폼 간의 연계•활용도 검증작업을 거친 연계키를 표준연계키로 등록함.

d) 검증과 등록과정을 거쳐 발굴된 표준연계키를 표준연계키 사전에 등록함.

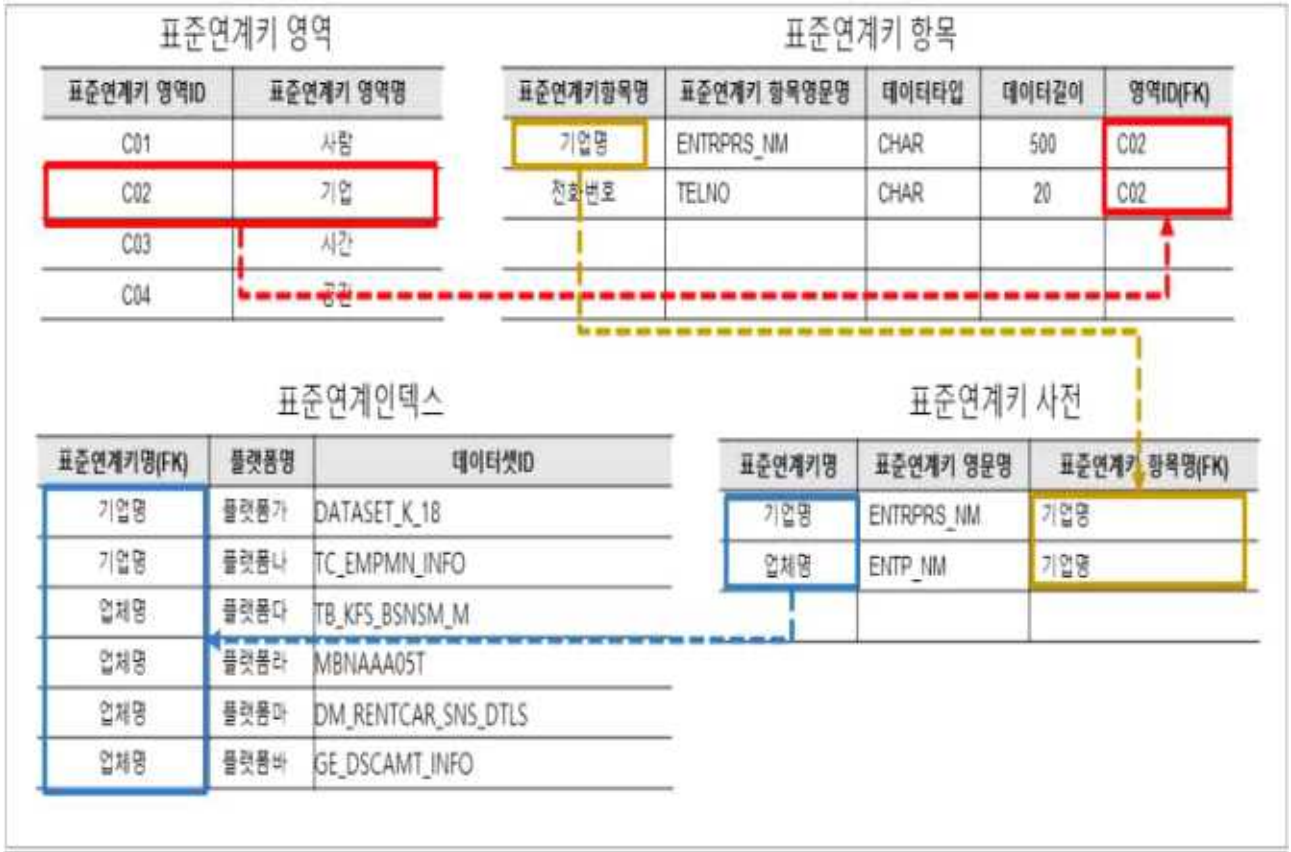
e) d)의 과정을 반복하여 표준연계키사전을 구축함.

○ 표준연계인덱스 정의

- 표준연계키사전이 구축되면, 각 플랫폼은 자신의 데이터셋과 표준연계키사전을 매핑시킴. 그 결과 플랫폼명, 데이터셋명, 속성명 등이 하나의 레코드로 구성된 표준연계 인덱스를 구축할 수 있음. 표준연계 인덱스는 플랫폼 간의 데이터를 연계활용할 수 있는 환경을 제시하여, 플랫폼 간 데이터 상호운용성을 확보할 수 있게 됨.

○ 표준연계인덱스 항목

- 표준연계인덱스 구축과정을 거쳐 플랫폼명, 데이터셋명, 속성명 등의 항목으로 구성된 표준연계 인덱스가 생성됨. 아래 그림에서 보이는 것과 같이 표준연계키 사전과의 매핑 과정을 거쳐 생성됨.



[SQL문 예시]

1. 표준연계키항목명(데이터셋의 속성명)이 '기업명'인 플랫폼의 데이터셋의 조회하고자 함.

```
select B.플랫폼명, B.데이셋ID
 from 표준연계키사전 A, 표준연계인덱스 B
where A.표준연계키항목명 = "기업명" and A.표준연계키명 = B.표준연계키명
```

2. 표준연계키항목의 '기업명'이 'SSSS'인 업체의 정보를 가지고 있는 (플랫폼가 및 플랫폼나)플랫폼의 모든 속성정보를 조회하고자 함

```
Select * from 플랫폼가.DATASET_K_18 A, 플랫폼나.TC_EMPMN_INFO B
        where A.기업명 = B.기업명 and A.기업명 = "SSSS"
```

(표준연계인덱스 항목과 활용 예시)

○ 표준연계 인덱스 구축 절차

- 표준연계 인덱스 구축을 통한 데이터 연계 활용은 다음의 절차를 거쳐 이뤄짐.

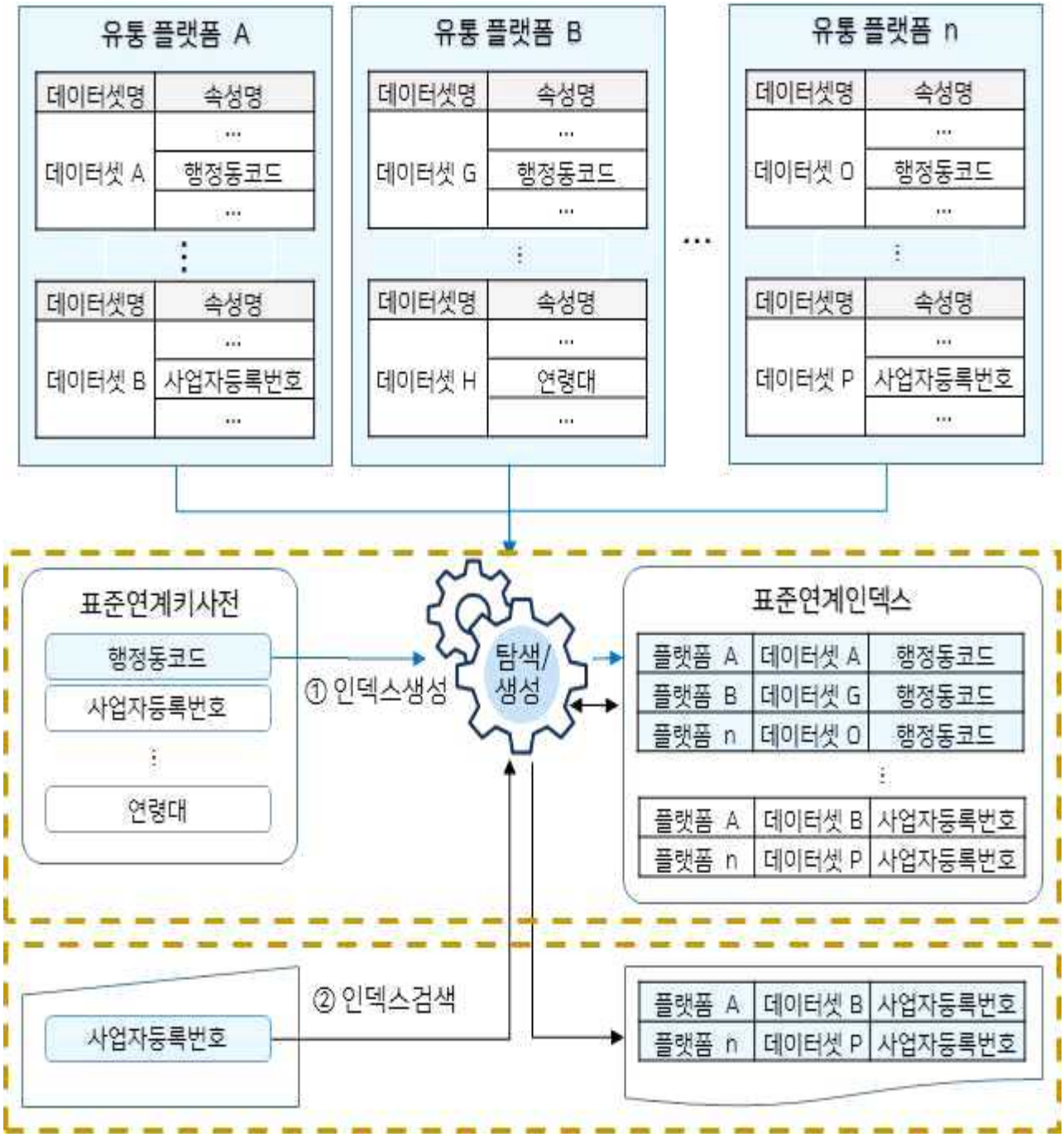
a) 표준연계키 사전을 바탕으로 각 플랫폼의 데이터셋과의 매핑 과정을 거침.

b) 표전연계키 사전에 매핑된 각 데이터셋들의 플랫폼명, 데이터셋명, 속성명 등이 하나의 레코드로 저장됨.

c) b)의 과정을 거쳐 표준연계인덱스가 생성됨.

d) 표준연계 인덱스 검색 기능을 수행하여 각 플랫폼은 서로의 데이터셋을 연계하여 활용 할 수 있음.

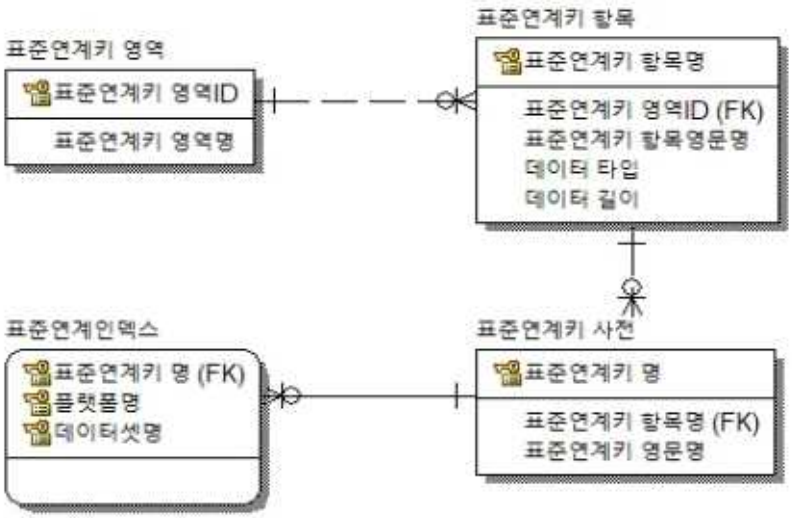
- 위 절차를 거쳐, 아래 그림에 제시된 것과 같이 표준연계 인덱스 생성, 표준연계 인덱스 검색 등을 수행할 수 있음.



(표준연계인덱스 생성 및 검색 설명)

○ 표준연계키 관리 구조

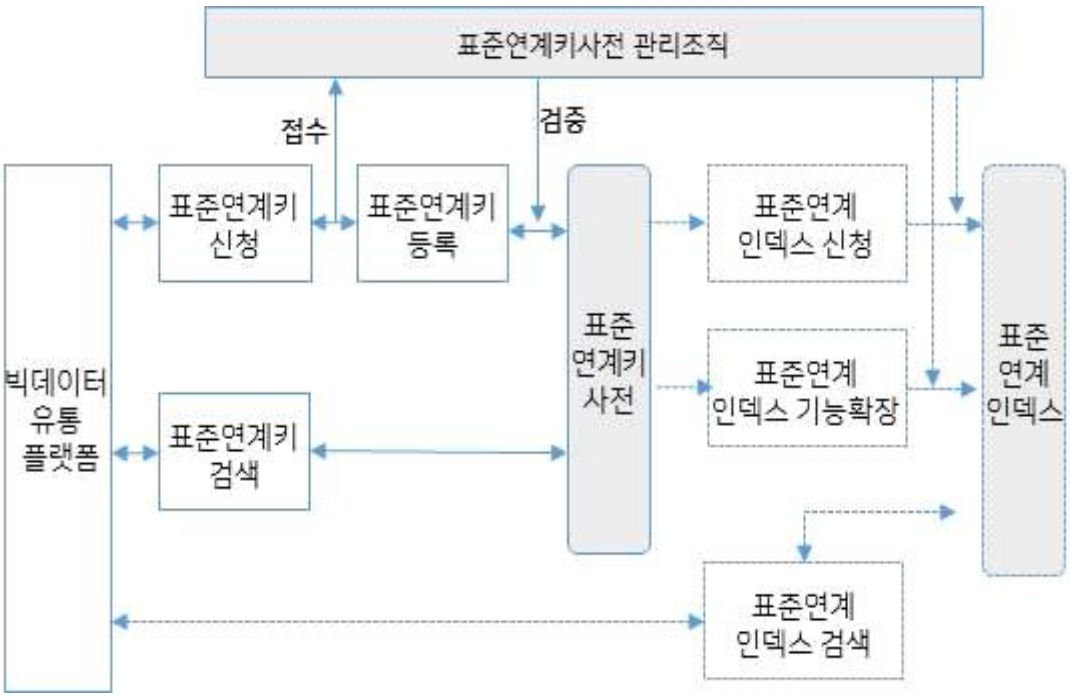
- 표준연계키 관리 구조에서는 표준연계키, 표준연계키사전, 표준연계인덱스의 각 요소를 정의하여 설명하고 표준연계키 관리를 위한 개략적인 흐름을 보여줌. 아래 그림에는 표준연계키 관리 구조를 다이어그램으로 표현한 것.



(표준연계키 관리 다이어그램(IE 표기법))

○ 표준연계키 관리 기능

- 표준연계키 관리 기능은 앞서 제시한 표준연계키 등록, 표준연계키사전 구축, 표준연계인 덱스 구축 등의 절차를 구현하기 위한 기능임. 이에는 표준연계키 신청, 등록, 검색 등의 기능과 표준연계인덱스의 생성, 검색 등의 기능이 있음. 아래 그림과 같이 표준연계키 관리 기능이 수행됨.



(표준연계키 관리 기능)

○ 표준연계키 신청

- 각 플랫폼은 서로 다른 데이터 환경 속에 있기 때문에, 데이터의 상호연계를 위한 표준 연계절차를 따라야 함. 표준연계절차는 플랫폼들 간의 데이터 연계를 위해 표준으로 정의된 절차임. 각 플랫폼은 표준연계를 위해 속성을 발굴하여 표준연계키로 신청해야함. 이 기능을 위해서 표준연계키 사전 관리조직이 필요하며, 이 관리조직은 각 플랫폼에서 발굴된 표준연계키를 신청을 받음.

○ 표준연계키 등록

- 표준연계키 사전의 관리조직은 각 플랫폼으로부터 표준연계키 등록 요청받으면, 이를 표준연계키로 사용할 것인지 검토해야 함. 검토 결과, 표준연계키 등록이 인정되면 관리조직은 이를 표준연계키사전에 등록할 수 있음. 이를 위해 표준연계키 등록 기능에서는표준연계키사전 등록기능도 함께 제공되어야 함.

○ 표준연계키 검색

- 각 플랫폼이 등록된 표준연계키를 표준연계키사전에서 검색할 수 있는 기능임. 각 플랫폼은 이 기능을 통해 자신이 찾고자 하는 표준연계키를 검색할 수 있음.

○ 표준연계인덱스 생성

- 각 플랫폼은 표준연계키를 기반으로 각 플랫폼의 데이터셋 속성정보를 탐색할 수 있어야 함. 이를 위해 표준연계인덱스 생성 기능이 제공됨. 표준연계인덱스는 표준연계키 사전과 매핑되는 데이터셋(테이블) 및 속성(컬럼) 정보를 바탕으로 구축된 것임. 따라서 각 플랫폼은 생성된 표준연계인덱스를 통해 각 플랫폼이 찾는 데이터셋의 속성정보를 탐색할 수 있음.

○ 표준연계인덱스 검색

- 빅데이터의 융•복합 운용을 위해서 먼저 표준연계키사전을 검색하여 데이터 연계에 활용 하려는 표준연계키와 관련된 데이터셋의 정보를 확인할 수 있어야 함. 이는 표준연계 인덱스 검색 기능을 통해 확인할 수 있음. 이 기능은 표준연계인덱스에서 해당 표준연계 키와 매핑되는 데이터셋 정보를 검색하여 빅데이터 유통 생태계내 각 구성원이 이용할 수 있게 함.

○ 표준연계인덱스 기능확장

- 표준연계키 사전 검색으로 제공받은 연계키와 관련된 데이터셋 정보를 확장할 수 있는 표준연계인덱스 기능확장 기능이 필요함. 표준연계키사전 관리조직은 표준연계인덱스에 해당된 연계키와 매핑되는 데이터셋의 정보를 확장하는 기능을 제공함.

4. 데이터 거래를 위한 예상 사용자 요구사항 조사

가. 데이터 거래 계약사항

○ 어떠한 거래를 하더라도, 공급자(판매자)와 수요자(구매자)가 계약과 합의를 통해 이용 대상과 범위, 제공방법, 권리관계 등을 명확히 해야 함.
○ 특히 데이터 거래를 할 때에는 더욱 주의가 필요하며 데이터를 상업적인 목적으로 사용하려면 온라인 등 공개된 데이터라 하더라도 권리자가 무료로 자유로운 이용을 명시적으로 허락하지 않은 경우 데이터 사용에 대한 별도의 확인이나 계약 절차가 필요함.
○ 데이터 거래 계약에서 핵심적으로 고려할 사항은 다음과 같음.

<데이터 거래 계약 순서>

(1) 계약대상 범위 설정

○ 계약 대상 데이터 특성, 데이터 및 데이터 소재 권리관계 확인, 계약 당사자 확인에 대한 사항임. 권리관계 확인, 계약 당사자 확인은 필수이며 데이터 거래를 할 때에는, 특히 개인정보 문제에 주의해야함.
○ 개인정보가 포함되어 있는지, 포함되어 있다면 사전 동의를 얻은 것인지 검토해야함.

방법	내용
계약대상 명확화	계약 대상 데이터가 무엇인지 정확히 식별할 수 있도록 특정
데이터 권리관계 확인	계약 당사자가 계약 대상 데이터를 거래할 수 있는 권리가 있는지 확인
계약 당사자 확인	계약 당사자 정보(성명, 법인명, 주소 등)를 확인
개인정보 포함확인	계약 대상 데이터에 개인정보 포함 여부와 그 개인정보를 이용하기 위한 사전 동의 획득
신용정보, 위치정보 등 관계법령 준수 확인	계약 대상 데이터에 신용정보, 위치정보 등 관계법령에서요구하는 보호조치를 이행

(2) 데이터 이용목적 및 방법

○ 구매자 이용목적 및 서비스 확인, 데이터 가공여부 확인, 계약기간 설정에 대한 사항임. 같은 데이터이더라도 여러 가지 목적으로 사용될 수 있음. 구매자가 어떤 목적으로 데이터를 이용할 것인지, 어느 기간동안 이용할 것인지 반드시 확인하셔야 예상치 못한 문제를 방지할 수 있음.

방법	내용
데이터 이용목적 및 서비스 확인	구매자가 계약 당사자 대상 DB를 이용하려는 목적은 무엇인지, 어떤 플랫폼에서 어떤 방식으로 이용할지를 확인
데이터 가공여부 확인	계약 대상 데이터의 변형, 변경 등의 가공이 필요한지 확인하고 이를 허용할 것인지 협의해야함
계약 기간 결정	계약 대상 데이터를 얼마동안 이용할지 계약기간을 협의해야함

(3) 데이터 이용허락 권리범위

○ 이용 허락 권리 설정, 재이용허락, 독점・비독점, 권리 양도 협의에 대한 사항임. 구매자가 데이터 사용에 있어 어느정도의 권리를 가지고 있는지 명확히 해야함.

방법	내용
이용허락 권리 설정	계약 대상 데이터를 이용할 수 있는 데이터의 권리범위(복제, 배포, 방송, 전송 등)를 명확히 설정해야함
재이용허락 허용여부	계약 대상 데이터의 이용 목적상 제 3자에게 재이용 허락이 필요한지, 필요하다면 허용할 것인지 협의해야함
독점/비독점 이용허락 여부	계약 대상 데이터를 독점적으로 이용하도록 할 것인지, 비독점적으로 이용하도록 할 것인지 협의해야함
권리의 양도	계약 당사자 간 상호간의 권리, 의무를 상호 동의없이는 제 3자에게 양도하지 못한다는 것을 상호 숙지해야함

(4) 데이터 제공방법

○ 데이터 파일 포맷 및 전달방식, 제공 데이터의 품질 수준, 제공주기, 보증 등에 대한 사항임. 데이터는 다양한 파일 형식을 가질 수 있음.
○ 파일 상품을 예를 들면 XLS, XLSX, CSV, TXT 중 어떤 포맷으로 제공할지, 다운로드, FTP, E-mail 중 어떤 방식으로 전달할지 확실히 해야함.

방법	내용
포맷 및 전달방법	계약 대상 데이터를 제공할 파일 포맷과 전달 방식 등을 협의해야함
제공 데이터의 품질 수준	계약 대상 데이터의 최신성, 정확성, 업데이트 주기 등의 품질 수준에 대하여 협의해야함

(5) 데이터 이용료 및 정산

○ 최종 가격이 결정되고, 구매자와 판매자 모두가 동의했는지 확인하여야 하며 지불 방법과 시기를 구체적으로 결정해야 예민한 문제를 방지하실 수 있음.

방법	내용
이용료 산정	계약 대상 데이터의 이용 대가는 충분히 협의하여 적정 이용료를 정해야함
이용료 정산 방법	계약 대상 데이터를 어떻게 지급할 것인지 지급 방법과 지급 시기에 대해 충분히 협의해야함

(6) 데이터 권리 침해예방 설정

○ 제3자 복제・배포・전송 등 금지, 출처 표시, 기술적 보호조치, 이용조건 준수 확인 협의에 대한 사항이며 판매자의 권리를 보호하는 사항이므로 반드시 확인해야함.

방법	내용
제3자 복제/배포/전송 등 금지	계약 대상 데이터를 제 3자가 무단 복제 등 데이터 권리를 침해하는 방법으로 사용하지 못하도록 협의해야함
데이터 출처 표시	계약 대상 데이터의 권리자가 누구인지 구매자 서비스에 표시 여부와 표시한다면 그 방법 등을 협의해야함
기술적 보호조치 등	계약 대상 데이터를 보호하기 위한 기술적 보호조치 등의 필요 여부와 필요하다면 조치 방법 등을 협의해야함
이용 조건 준수 확인	계약 대상 데이터를 이용 조건 내에서 이용하도록 하며, 이를 점검할 수 있음을 협의해야함

(7) 계약의 해지

○ 계약 해지 사유 및 계약 종료효과, 잔존 권리 의무 등 명시에 대한 사항이며 시작도 중요하지만, 끝도 중요하므로 계약 해지에 대해 확실히 해야함.

방법	내용
계약 해지 방법의 결정	계약 해지 사유와 행사 방법을 협의해야함
계약 종료 효과	계약 대상 데이터의 이용기간 만료 또는 계약이 해지된 경우 데이터의 반환, 폐기, 삭제방법 및 시점에대해 협의해야함

나. 데이터 거래 관련 문제점 및 개선 방안 조사

○ 목적
- ‘(가칭) 스마트농업 데이터 거래소’ 운영을 위한 주요 이슈에 대한 이해당사자 인식을 조사하고, 운영 모델 및 제도 개선 방향 도출에 활용

○ 대상 : 주요 이해 당사자
- 잠재 공급자(114명) : 농가, 스마트농업 데이터 생산 기업, 데이터 거래 기업, 연구기관 등
- 잠재 수요자 및 전문가(11명) : 농업 전문가 및 유통업체 실무자
- 가능한 최대한의 표본을 확보하여 조사

○ 조사 방법 및 기간
- 방법 : 면담 및 이메일 조사
- 기간 : 2023.3.13.~17(5일)
- 잠재 공급자는 (사)한국스마트팜산업협회 도움으로 조사

○ 주요 조사내용

1) 잠재 공급자
- 스마트농업 데이터 공급 현황 및 데이터 거래에 대한 인식
  · 생산/수집/분석 데이터, 데이터 공급 역량 및 제공 데이터 활용도, 데이터 판매 경험 및 판매 형태, 데이터 판매 의향 및 판매 시 소유권 이전 의향과 판매하지 않고자 하는 이유 등
- (데이터 판매 경험자) 판매 데이터 종류, 판매처, 판매 방식(양도 계약, 라이선스 계약, 기타), 판매 가격 책정 시 참고 정보, 데이터 거래소를 통해 판매한 경험 등
- (거래소를 통한 데이터 판매 경험자) 직거래 대비 장점 및 단점, 거래소 미이용 이유
- 스마트농업 데이터의 가치에 대한 인식
  · 데이터 판매가격 책정 시 중요하게 고려하는 속성, 데이터 가치에 영향을 미치는 주요 속성의 수준 평가, 원가 근거 판매가격 책정 시 적정 회수 기간 및 이윤율, 수요자에게 가치를 제공하는 최대 기간 및 적정 데이터 보존 기간, 데이터 신뢰성 확보를 위해 제공 의향이 있는 정보, 구매 데이터로 인한 손실 보상에 대한 인식 등

- 스마트농업 데이터 거래소 설립에 대한 인식
  · 설립 필요성, 거래소 설립이 필요하지 않거나 잘 모르겠다고 답한 이유, 데이터 거래 설립 관련 사항(운영 주체, 거래방식, 정부 및 공공기관 데이터 거래 필요성, 적합한 거래 상품), 거래소 설립 및 운영 시 고려해야 할 중요 문제점
- 스마트농업 데이터 거래 활성화를 위한 과제
- 일반현황 : 성별, 연령대, 분야(농가, 기업, 컨설팅, 대학, 연구기관, 정책부서, 기타), 업무

2) 잠재 수요자 및 농업 전문가
- 농업 데이터 이용 현황 및 데이터 거래에 대한 인식
  · 필요로 하는 농업 데이터 종류 및 유형, 농업 데이터 확보 방법, 농업 데이터 구매 경험, 향후 구매 의향 및 소유권 이전 의향 등

- (데이터 구매 경험자) 구매 데이터 종류, 판매처, 판매 방식(양도 계약, 라이선스 계약, 기타), 판매가격 책정 시 참고 정보, 데이터 거래소를 통해 판매한 경험 등
- (거래소를 통한 데이터 판매 경험자) 직거래 대비 장점 및 단점, 거래소 미이용 이유
- 농업 데이터의 가치에 대한 인식
  - 데이터 판매가격 책정 시 중요 속성, 데이터 가치에 영향을 미치는 주요 속성의 수준 평가, 원가 근거 판매가격 책정 시 적정 회수 기간 및 이윤율, 수요자에게 가치를 제공하는 최대 기간 및 적정 데이터 보존 기간, 데이터 신뢰성 확보를 위해 제공 의향이 있는 정보, 구매 데이터로 인한 손실 보상에 대한 인식 등
- 스마트농업 데이터 거래소 설립에 대한 인식
  - 설립 필요성, 거래소 설립이 필요하지 않거나 잘 모르겠다고 답한 이유, 데이터 거래 설립 관련 사항(운영 주체, 거래방식, 정부 및 공공기관 데이터 거래 필요성, 적합한 거래 상품), 거래소 설립 및 운영 시 고려해야 할 중요 문제점
- 스마트농업 데이터 거래 활성화를 위한 과제
- 일반현황 : 성별, 연령대, 분야(농가, 기업, 컨설팅, 대학, 연구기관, 정책부서, 기타), 업무

다. 잠재 공급자 조사 결과

(1) 응답자 현황

○ 응답자 현황은 표와 같음

- 남성(88.7%)이 여성(11.3%)보다 많고, 연령대는 50대가 59.6%로 가장 많음
- 종사 분야로 기업이 80.9%로 가장 많고, 다음이 농가(13.5%) 등의 순임
- 분야경력은 6~10년이 21.7%로 가장 많고, 다음이 16~20년(19.3%), 20~25년(16.9%) 등의 순임

< 잠재 공급자 대상 조사 응답자 일반현황 >

구분		빈도(명)	비율(%)
성별	남	125	88.7
	여	16	11.3
	합계	141	100.0
연령대	30대	7	5.0
	40대	27	19.1
	50대	84	59.6
	60대 이상	23	16.3
	합계	141	100.0
종사 분야	기업	114	80.9
	농가	19	13.5
	대학	1	0.7
	법인	1	0.7
	연구기관	4	2.8
	컨설팅	2	1.4
	합계	141	100.0
분야경력 (38명 미응답)	5년 이하	10	12.0
	6~10년	18	21.7
	11~15년	13	15.7
	16~20년	16	19.3
	20~25년	14	16.9
	25~30년	8	9.6
	31년 이상	4	4.8
	합계	83	100.0

(2) 스마트농업 데이터 공급 현황 및 데이터 거래에 대한 인식

○ 생산, 수집, 분석 중인 스마트농업 데이터 종류로 농작물 생육환경이 16.8%로 가장 많고, 다음이 생산량 등 생산 현황(13.6%), 농작물 생육(12.3%) 등의 순임

< 생산, 수집, 분석 중인 데이터 종류(복수 응답 가능) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
농작물 생육환경	54	6	0	0	2	2	64	16.8
생산량 등 생산 현황	43	5	0	0	2	2	52	13.6
농작물 생육	37	6	0	0	2	2	47	12.3
농업기상	33	5	1	0	2	1	42	11.0
생산비/경영 현황	31	5	0	0	1	2	39	10.2
요소 투입량	27	4	0	0	3	1	35	9.2
병해충	25	4	0	0	2	1	32	8.4
유통/판매 현황	24	5	0	0	0	1	30	7.9
기타	35	5	0	1	0	0	41	10.7
합계	309	45	1	1	14	12	382	100.0

○ 생산, 수집, 분석 중인 데이터 유형으로 정형 데이터(30.0%)가 가장 많고, 다음이 비정형 이미지(23.0%), 비정형 텍스트(15.6%), 비정형 동영상(14.4%) 등의 순임

< 생산, 수집, 분석 중인 데이터 유형(복수 응답 가능) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
정형 데이터(수치 등)	68	9	1	0	1	2	81	30.0
비정형 이미지	48	10	0	0	3	1	62	23.0
비정형 텍스트	33	6	0	0	1	2	42	15.6
비정형 동영상	26	9	0	0	3	1	39	14.4
비정형 오디오	22	3	0	0	2	1	28	10.4
기타	17	0	0	1	0	0	18	6.7
합계	214	37	1	1	10	7	270	100.0

○ 응답자의 스마트농업 데이터 공급 역량 및 제공 데이터 활용도에 대한 인식을 보면, 전반적 제공역량이 5점 만점에 3.52점(매우 낮음 10.6%, 약간 낮음 6.4%, 보통 26.2%, 약간 우수 34.0%, 매우 우수 22.7%)으로 보통 수준의 역량을 보유한다고 인식

- 4개 세부 역량 중 지속적인 데이터 제공역량(3.63점)이 가장 우수하고, 다음이 데이터 생산 또는 수집 역량(3.57점), 데이터 분석 및 활용 서비스 제공역량(3.50점), 데이터 활용 솔루션 제공 및 컨설팅 역량(3.48점)의 순임
- 제공되는 데이터가 데이터 활용자의 생산성, 부가가치 향상 등에 공헌하는 제공 데이터의 활용도에 대한 인식은 3.55점(매우 낮음 8.5%, 약간 낮음 8.5%, 보통 29.1%, 약간 우수 27.0%, 매우 우수 27.0%)으로 보통 수준임

< 데이터 제공역량 및 활용도 인식 >

구분		기업	농가	대학	법인	연구기관	컨설팅	전체	
								빈도(명)	비율(%)
데이터 생산 또는 수집 역량 (3.57점)	매우 낮음	11	2	0	0	1	0	14	9.9
	약간 낮음	11	0	0	0	0	0	11	7.8
	보통	27	2	1	1	0	1	32	22.7
	약간 우수	40	5	0	0	3	0	48	34.0
	매우 우수	25	10	0	0	0	1	36	25.5
	합계	114	19	1	1	4	2	141	100.0
데이터 분석 및 활용 서비스 제공역량 (3.50점)	매우 낮음	10	1	0	0	0	0	11	7.8
	약간 낮음	9	0	0	0	0	0	9	6.4
	보통	40	2	1	1	2	0	46	32.6
	약간 우수	40	6	0	0	2	1	49	34.8
	매우 우수	15	10	0	0	0	1	26	18.4
	합계	114	19	1	1	4	2	141	100.0
데이터 활용 솔루션 제공 및 컨설팅 역량 (3.48점)	매우 낮음	11	2					13	9.2
	약간 낮음	12						12	8.5
	보통	36	2	1	1	1	1	42	29.8
	약간 우수	35	4			3		42	29.8
	매우 우수	20	11				1	32	22.7
	합계	114	19	1	1	4	2	141	100.0
지속적인 데이터 제공역량 (3.63점)	매우 낮음	10	1					11	7.8
	약간 낮음	12				1		13	9.2
	보통	27	2	1	1		1	32	22.7
	약간 우수	41	3			2		46	32.6
	매우 우수	24	13			1	1	39	27.7
	합계	114	19	1	1	4	2	141	100.0
전반적 데이터 제공역량 (3.52점)	매우 낮음	14	1					15	10.6
	약간 낮음	8				1		9	6.4
	보통	31	3	1	1		1	37	26.2
	약간 우수	41	6			1		48	34.0
	매우 우수	20	9			2	1	32	22.7
	합계	114	19	1	1	4	2	141	100.0
제공 데이터의 활용도 (생산성, 부가가치 향상 등) (3.55점)	매우 낮음	11	1					12	8.5
	약간 낮음	11				1		12	8.5
	보통	36	2	1	1		1	41	29.1
	약간 우수	32	4			2		38	27.0
	매우 우수	24	12			1	1	38	27.0
	합계	114	19	1	1	4	2	141	100.0

주. ( ) 안의 값은 '매우 낮음 = 1점, 약간 낮음 = 2점, 보통 = 3점, 약간 우수 = 4점, 매우 우수 = 5점'을 적용한 산출한 값임

○ 응답자 중 데이터 판매 경험자는 전체의 12.1%(17명)로 미미

- 판매 데이터는 생산량 등 생산 현황(20.6%)이 가장 많고, 다음이 농작물 생육환경(17.6%), 농작물 생육(14.7%), 요소 투입량(13.2%) 등의 순임

< 스마트농업 데이터 판매 경험 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
판매 경험 있음	13	3				1	17	12.1
판매 경험 없음	101	16	1	1	4	1	124	87.9
합계	114	19	1	1	4	2	141	100.0

< 스마트농업 데이터 판매 경험자(17명)의 판매 데이터 종류(복수 응답 가능) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
생산량 등 생산 현황	9	3	0	0	0	2	14	20.6
농작물 생육환경	9	2	0	0	0	1	12	17.6
농작물 생육	7	2	0	0	0	1	10	14.7
요소 투입량	6	2	0	0	0	1	9	13.2
병해충	5	1	0	0	0	1	7	10.3
농업기상	4	1	0	0	0	1	6	8.8
유통/판매 현황	3	1	0	0	0	1	5	7.4
생산비/경영 현황	2	0	0	0	0	0	2	2.9
기타	3	0	0	0	0	0	3	4.4
합계	48	12	0	0	0	8	68	100.0

- 데이터 판매처로 농기업(23.3%)이 가장 많고, 다음이 연구기관(20.9%), 농가(18.6%), 대학(16.3%) 등의 순임

< 스마트농업 데이터 판매 경험자(17명)의 데이터 판매처(복수 응답 가능) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
농기업	8	1	0	0	0	1	10	23.3
연구기관	6	2	0	0	0	1	9	20.9
농가	5	2	0	0	0	1	8	18.6
대학	5	1	0	0	0	1	7	16.3
데이터 거래소 또는 플랫폼	3	2	0	0	0	1	6	14.0
기타	3	0	0	0	0	0	3	7.0
합계	30	8	0	0	0	5	43	100.0

- 데이터 판매 방식으로 이용만 허락하는 라이선스 계약이 64.7%로 양도 계약(35.3%)의 약 2배에 가까움

< 스마트농업 데이터 판매 경험자(17명)의 데이터 판매 방식 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
라이선스 계약(이용 허락)	8	3	0	0	0	0	11	64.7
양도 계약(소유권 이전)	5	0	0	0	0	1	6	35.3
합계	13	3	0	0	0	1	17	100.0

- 데이터 판매가격 책정 시 참고 정보는 업종 내 유사 데이터 판매사례(31.4%)가 가장 많았으나, 데이터 거래소 또는 플랫폼의 판매가격 정보(22.9%), 자사의 과거 판매가격(22.9%), 데이터 생산 및 수집 비용(원가)(20.0%)의 비중도 작지 않음

< 스마트농업 데이터 판매 경험자(17명)의 판매가격 책정 시 참고 정보(복수 응답 가능) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
업종 내 유사 데이터 판매사례	9	1	0	0	0	1	11	31.4
데이터 거래소 또는 플랫폼의 판매가격 정보	5	2	0	0	0	1	8	22.9
자사의 과거 판매가격	6	1	0	0	0	1	8	22.9
데이터 생신/수집 비용(원가)	4	2	0	0	0	1	7	20.0
기타	1	0	0	0	0	0	1	2.9
합계	25	6	0	0	0	4	35	100.0

○ 데이터 판매 경험자(17명) 중 데이터 거래소를 통해 판매한 경험이 있는 응답자는 절반 수준인 47.1%(8명)임

< 스마트농업 데이터 판매 경험자(17명)의 데이터 거래소를 통한 판매 경험 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
거래소 판매 경험 있음	6	1	0	0	0	1	8	47.1
거래소 판매 경험 없음	7	2	0	0	0	0	9	52.9
합계	13	3	0	0	0	1	17	100.0

- 거래소를 통해 데이터를 판매한 경험이 있는 응답자(8명)가 인식하는 직거래에 대비한 거래소를 통한 거래의 장점은 판매 조건, 가격 등의 협상이 불필요(24.2%)가 가장 많고, 다음이 계약체결 후 거래이행 확인 불필요(21.2%), 적절한 거래 상대 찾기 위한 노력 불필요(21.2%), 거래 관련 계약체결 등 거래 절차 용이(18.2%) 등의 순임

< 거래소를 통한 스마트농업 데이터 판매 경험자(8명)의 데이터 거래소를 통한 판매의 장점 >

(복수 응답 가능)

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
판매 조건, 가격 등의 협상 불필요	6	1	0	0	0	1	8	24.2
계약체결 후 거래이행 확인 불필요	5	1	0	0	0	1	7	21.2
적절한 거래 상대 찾기 위한 노력 불필요	5	1	0	0	0	1	7	21.2
거래 관련 계약체결 등 거래 절차 용이	4	1	0	0	0	1	6	18.2
판매대금을 확실히 정산 받을 수 있음	4	0	0	0	0	1	5	15.2
기타	0	0	0	0	0	0	0	0.0
합계	24	4	0	0	0	5	33	100.0

- 거래소를 통해 데이터를 판매한 경험이 있는 응답자(8명)가 인식하는 직거래에 대비한 거래소를 통한 거래의 단점은 판매 조건, 가격 등을 지나치게 제약(26.9%)가 가장 많고, 다음이 제공서비스 대비 비싼 수수료 요구(23.1%), 거래소의 불필요한 절차 또는 서류 요구(19.2%) 등의 순임

< 거래소를 통한 스마트농업 데이터 판매 경험자(8명)의 데이터 거래소를 통한 판매의 단점 >

(복수 응답 가능)

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
판매 조건, 가격 등을 지나치게 제약	5	1	0	0	0	1	7	26.9
제공서비스 대비 비싼 수수료 요구	4	1	0	0	0	1	6	23.1
거래소의 불필요한 절차 또는 서류 요구	4	0	0	0	0	1	5	19.2
거래소의 데이터에 대한 불필요한 간섭	3	0	0	0	0	1	4	15.4
거래소의 판매기업 또는 개인정보 요구	3	0	0	0	0	1	4	15.4
기타	0	0	0	0	0	0	0	0.0
합계	19	2	0	0	0	5	26	100.0

- 데이터 판매 경험자(17명) 중 데이터 거래소를 통해 판매한 경험이 없는 응답자(9명)가 거래소를 이용하지 않는 이유로 신뢰할 만한 거래소를 찾지 못함(33.3%)이 가장 높고, 다음이 데이터 거래소가 있는지를 몰랐음(20.0%), 거래소의 데이터에 대한 불필요한 간섭(13.3%), 거래소의 불필요한 절차 또는 서류 요구(13.3%) 등의 순임

< 거래소를 통한 스마트농업 데이터 판매 미경험자(9명)가 거래소를 이용하지 않는 이유 >

(복수 응답 가능)

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
신뢰할 만한 거래소를 찾지 못함	4	1	0	0	0	0	5	33.3
데이터 거래소가 있는지를 몰랐음	2	1	0	0	0	0	3	20.0
거래소의 데이터에 대한 불필요한 간섭	2	0	0	0	0	0	2	13.3
거래소의 불필요한 절차 또는 서류 요구	2	0	0	0	0	0	2	13.3
판매 조건, 가격 등을 지나치게 제약	1	0	0	0	0	0	1	6.7
거래소의 판매기업 또는 개인정보 요구	1	0	0	0	0	0	1	6.7
기타	0	0	0	0	0	0	0	0.0
합계	13	2	0	0	0	0	15	100.0

○ 응답자의 절반에 가까운 46.8%(66명)가 향후 데이터를 판매할 의향이 있다고 답함

- 데이터 판매 의향자(66명)의 59.1%가 소유권 이전 의향이 있다고 답함

< 향후 데이터 판매 의향 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
판매 의향 있음	47	15	0	0	3	1	66	46.8
판매 의향 없음	22	3	0	0	1	1	27	19.1
잘 모르겠음	45	1	1	1	0	0	48	34.0
합계	114	19	1	1	4	2	141	100.0

< 향후 데이터 판매 의향자(66명)의 데이터 소유권 이전 의향 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
소유권 이전 의향 있음	25	12	0	0	2	0	39	59.1
소유권 이전 의향 없음	15	1	0	0	0	1	17	25.8
잘 모르겠음	7	2	0	0	1	0	10	15.2
합계	47	15	0	0	3	1	66	100.0

- 향후 데이터를 판매할 의향이 없거나 잘 모르겠다고 답한 응답자(75명)는 그 이유로 판매가격 책정 어려움(19.4%)이 가장 중요하고, 다음이 판매 후 관리 어려움(18.2%), 판매에 부적합(18.2%), 적절한 판매 채널 확보 어려움(17.1%) 등의 순임

< 향후 데이터를 판매할 의향이 없거나 잘 모르겠다고 답한 이유(75명; 복수 응답 가능) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
판매가격 책정 어려움	30	1	0	0	1	1	33	19.4
판매 후 관리 어려움	28	1	0	0	1	1	31	18.2
판매에 부적합	27	1	1	1	0	1	31	18.2
적절한 판매 채널 확보 어려움	26	2	0	0	0	1	29	17.1
판매 후 불법 유통에 따른 손실 우려	23	1	0	0	1	1	26	15.3
개인정보 등 판매에 따른 법적 문제 발생 우려	18	1	0	0	0	1	20	11.8
기타	0	0	0	0	0	0	0	0.0
합계	152	7	1	1	3	6	170	100.0

(3) 스마트농업 데이터의 가치 인식

○ 데이터 판매가격 책정 시 가장 중요한 속성은 데이터 신뢰성(16.2%)이고, 다음이 정확성(15.8%), 생산 및 수집 비용(12.3%), 제공기관의 신뢰성(10.7%) 등의 순임

< 데이터 판매가격 책정 시 중요하다고 생각하는 속성(3개 이내) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
데이터 신뢰성	63	4	0	0	2	2	71	16.2
데이터 정확성	55	9	0	1	3	1	69	15.8
데이터 생산/수집 비용	45	5	0	0	3	1	54	12.3
데이터 제공기관의 신뢰성	38	7	0	0	1	1	47	10.7
제공정보 종류(기상, 경영 등)	35	7	1	0	1	2	46	10.5
데이터 품질	36	4	0	1	2	1	44	10.0
유형(텍스트, 이미지 등)	33	6	0	0	0	1	40	9.1
데이터 적시성	27	5	0	0	2	2	36	8.2
데이터 차별성/유일성	21	6	0	1	1	1	30	6.8
기타	0	1	0	0	0	0	1	0.2
합계	353	54	1	3	15	12	438	100.0

○ 데이터 가치에 영향을 미치는 주요 속성의 수준과 관련하여 전반적 데이터 수준은 5점 만점에 3.489점(매우 낮음 4.3%, 약간 낮음 12.1%, 보통 35.5%, 약간 우수 27.0%, 매우 우수 21.3%)으로 평가됨

- 세부 속성별 수준을 보면, 데이터 제공기관 신뢰성과 데이터 적시성이 3.674점으로 가장 우수하고, 다음이 데이터 정확성(3.603점), 데이터 품질(3.574점), 데이터 유용성(3.567점) 등의 순임

< 스마트농업 데이터의 주요 속성별 수준 >

구분		기업	농가	대학	법인	연구기관	컨설팅	전체	
								빈도(명)	비율(%)
데이터 유용성 (3.567점)	매우 낮음	6	0	0	0	0	0	6	4.3
	약간 낮음	17	0	0	0	1	1	19	13.5
	보통	29	2	1	1	1	0	34	24.1
	약간 우수	46	7	0	0	0	0	53	37.6
	매우 우수	16	10	0	0	2	1	29	20.6
	합계	114	19	1	1	4	2	141	100.0
데이터 신뢰성 (3.468점)	매우 낮음	7	0	0	0	1	0	8	5.7
	약간 낮음	16	0	0	0	0	0	16	11.3
	보통	40	3	1	1	0	1	46	32.6
	약간 우수	35	8	0	0	1	0	44	31.2
	매우 우수	16	8	0	0	2	1	27	19.1
	합계	114	19	1	1	4	2	141	100.0
데이터 제공기관 신뢰성 (3.674점)	매우 낮음	4	0	0	0	0	0	4	2.8
	약간 낮음	12	1	0	0	1	0	14	9.9
	보통	39	2	1	1	0	0	43	30.5
	약간 우수	35	6	0	0	1	1	43	30.5
	매우 우수	24	10	0	0	2	1	37	26.2
	합계	114	19	1	1	4	2	141	100.0
데이터 정확성 (3.603점)	매우 낮음	6	0	0	0	0	0	6	4.3
	약간 낮음	17	1	0	0	0	0	18	12.8
	보통	33	2	1	1	2	1	40	28.4
	약간 우수	34	5	0	0	0	0	39	27.7
	매우 우수	24	11	0	0	2	1	38	27.0
	합계	114	19	1	1	4	2	141	100.0
데이터 적시성 (3.674점)	매우 낮음	5	0	0	0	0	1	6	4.3
	약간 낮음	19	0	0	0	1	0	20	14.2
	보통	25	2	1	1	0	0	29	20.6
	약간 우수	41	4	0	0	0	0	45	31.9
	매우 우수	24	13	0	0	3	1	41	29.1
	합계	114	19	1	1	4	2	141	100.0
데이터 품질 (3.574점)	매우 낮음	5	0	0	0	0	0	5	3.5
	약간 낮음	19	0	0	0	1	0	20	14.2
	보통	34	3	1	1	0	1	40	28.4
	약간 우수	35	6	0	0	0	0	41	29.1
	매우 우수	21	10	0	0	3	1	35	24.8
	합계	114	19	1	1	4	2	141	100.0
데이터 차별성 또는 유일성 (3.489점)	매우 낮음	5	0	0	0	0	0	5	3.5
	약간 낮음	17	0	0	0	1	0	18	12.8
	보통	38	4	1	1	0	1	45	31.9
	약간 우수	40	8	0	0	1	0	49	34.8
	매우 우수	14	7	0	0	2	1	24	17.0
	합계	114	19	1	1	4	2	141	100.0
전반적 데이터 수준 (3.489점)	매우 낮음	5	0	0	0	1	0	6	4.3
	약간 낮음	16	0	0	0	0	1	17	12.1
	보통	44	4	1	1	0	0	50	35.5
	약간 우수	30	7	0	0	1	0	38	27.0
	매우 우수	19	8	0	0	2	1	30	21.3
	합계	114	19	1	1	4	2	141	100.0

주. ( ) 안의 값은 '매우 낮음 = 1점, 약간 낮음 = 2점, 보통 = 3점, 약간 우수 = 4점, 매우 우수 = 5점'을 적용하여 산출한 값임

○ 데이터 생산 및 수집에 드는 원가를 근거로 판매가격을 책정하는 경우 적정한 원가 회수 기간은 평균 4.28년(최대 20년, 최소 1년), 적정한 최소 이윤율은 평균 39.75%(최대 100%, 최소 1%)로 답함

< 원가 기반 판매가격 책정 시 적정 원가 회수 시간 및 최소 이윤율 >

구분		기업	농가	대학	법인	연구기관	컨설팅	전체
원가 기반 판매가격 책정 시 적정한 원가 회수 기간	빈도(명)	114	19	1	1	4	2	141
	평균(년)	4.26	4.58	5.00	5.00	4.00	2.00	4.28
	최대(년)	12	20	5	5	7	3	20
	최소(년)	1	1	5	5	1	1	1
	표준편차(년)	2.87	4.18	-	-	2.58	1.41	3.02
원가 기반 판매가격 책정 시 적정한 최소 이윤율	빈도(명)	114	19	1	1	4	2	141
	평균(%)	38.55	45.26	30.00	60.00	37.50	55.00	39.75
	최대(%)	100	77	30	60	70	60	100
	최소(%)	1	5	30	60	10	50	1
	표준편차(%)	23.89	20.43	-	-	27.54	7.07	23.31

○ 데이터 수명주기(life cycle)와 관련하여 데이터가 수요자에 가치를 제공하는 기간은 평균 6.62년(최대 50년, 최소 1년), 법적 문제 등에 대비한 적정한 데이터 보존 기간은 평균 7.83년(최대 50년, 최소 1년)으로 답함

< 데이터 수요자에 가치를 제공하는 최대 기간 및 적정 데이터 보존 기간 >

구분		기업	농가	대학	법인	연구기관	컨설팅	전체
데이터 수요자에 가치를 제공하는 최대 기간	빈도(명)	114	19	1	1	4	2	141
	평균(년)	6.92	4.68	5.00	5.00	8.00	6.50	6.62
	최대(년)	50	10	5	5	20	10	50
	최소(년)	1	1	5	5	2	3	1
	표준편차(년)	6.79	2.98	-	-	8.29	4.95	6.38
법적 문제 등에 대비한 적정 데이터 보존 기간	빈도(명)	114	19	1	1	4	2	141
	평균(년)	8.21	5.47	5.00	5.00	9.75	7.50	7.83
	최대(년)	50	20	5	5	20	10	50
	최소(년)	1	2	5	5	2	5	1
	표준편차(년)	7.46	4.38	-	-	7.59	3.54	7.06

○ 데이터 구매자가 판매 데이터의 신뢰성 확보를 위해 요청하는 경우 제공 의향이 있는 정보 등을 조사한 결과 데이터 생산 및 수집 방법(19.5%)이 가장 많고, 다음이 데이터 원천정보(16.1%), 데이터 품질 보증서(13.6%), 분야경력 및 실적(13.6%), 데이터 샘플 및 활용사례(10.8%) 등의 순으로 답함

< 구매자가 판매 데이터의 신뢰성 확보를 위해 요청하는 경우 제공 의향이 있는 정보 >

(복수 응답 가능)

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
데이터 생산/수집 방법	78	7	0	1	4	2	92	19.5
데이터 원천정보	67	4	0	1	2	2	76	16.1
데이터 품질 보증서	54	8	0	0	1	1	64	13.6
분야경력 및 실적	52	8	0	0	2	2	64	13.6
데이터 샘플 및 활용사례	47	2	0	0	1	1	51	10.8
전문가 평가서	41	3	0	0	2	1	47	10.0
회계정보(수입 등)	32	7	0	0	1	1	41	8.7
신용정보	31	3	1	0	0	1	36	7.6
합계	402	42	1	2	13	11	471	100.0

○ 판매 데이터의 오류 등으로 데이터 구매자에 손실이 발생하는 경우 이를 보상하는 방안에 대해 53.2%는 데이터 보상이 부적절하다(데이터 활용은 구매자 자율로 손실 보상 부적절 24.8%, 손실 원인 확인이 불가해 보상 부적절 27.7%)고 답하고, 46.8%는 조건부 또는 조건과 상관없이 보상이 필요하다(귀책 사유가 명확한 경우 보상 필요 27.7%, 손실 발생 원인 확인과 상관없이 보상 필요 19.1%)고 답함

< 데이터 오류 등으로 인한 구매자 손실 보상에 대한 인식 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
데이터 활용은 구매자 자율로 손실 보상 부적절	29	4	0	0	1	1	35	24.8
손실 원인 확인이 불가해 보상 부적절	28	10	0	1	1	0	40	28.4
귀책 사유가 명확한 경우 보상 필요	35	3	0	0	0	1	39	27.7
손실 발생 원인 확인과 상관없이 보상 필요	22	2	1	0	2	0	27	19.1
기타	0	0	0	0	0	0	0	0.0
합계	114	19	1	1	4	2	141	100.0

(4) ‘(가칭) 스마트농업데이터거래소’ 설립에 대한 인식

○ 스마트농업 데이터 거래소 설립에 대해 절반에 가까운 48.9%(69명)가 거래소 설립이 필요하다고 답했으나, 51.1%(필요하지 않다 14.9%, 잘 모르겠다 36.2%)는 필요하지 않거나 잘 모르겠다고 답함

< (가칭) 스마트농업데이터거래소 설립 필요성 인식 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
필요하다	53	13	0	0	2	1	69	48.9
필요하지 않다	17	2	0	0	1	1	21	14.9
잘 모르겠다	44	4	1	1	1	0	51	36.2
합계	114	19	1	1	4	2	141	100.0

- 데이터 거래소 설립이 필요하지 않거나 잘 모르겠다고 답한 가장 중요한 이유는 적정한 데이터 가격책정 어려움(17.8%)이고, 다음이 지적 재산권 등 데이터 거래 관련 제도 미비(17.3%), 데이터 거래에 대한 인식 미흡(16.2%), 데이터 거래시장 공급자와 수요자 부족(14.1%) 등의 순임

< (가칭) 스마트농업데이터거래소 설립이 필요하지 않거나 잘 모르겠다고 답한 이유 >

(복수 응답 가능)

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
적정한 데이터 가격책정 어려움	30	2	0	0	1	1	34	17.8
지적 재산권 등 데이터 거래 관련 제도 미비	28	3	0	0	1	1	33	17.3
데이터 거래에 대한 인식 미흡	26	3	0	0	1	1	31	16.2
데이터 거래시장 공급자와 수요자 부족	22	2	0	1	1	1	27	14.1
이미 설립된 데이터 거래소로 충분	13	2	0	1	1	1	18	9.4
개인정보 유출 등 거래에 따른 부작용	15	1	0	0	0	1	17	8.9
양질의 신뢰성 있는 데이터 상품 미흡	13	2	1	0	0	1	17	8.9
판매 데이터의 불법 유통으로 인한 혼란	9	1	0	1	1	1	13	6.8
기타	1	0	0	0	0	0	1	0.5
합계	157	16	1	3	6	8	191	100.0

- 데이터 거래소를 설립한다고 가정할 때 적정 운영 주체로 공공기관이나 공사가 43.3%로 가장 많고, 다음이 공공기관과 민간 공동(39.0%), 민간기업(17.0%) 등의 순임

< 스마트농업 데이터 거래소 설립 시 적정 운영 주체 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
공공기관/공사	48	11	1	0	0	1	61	43.3
공공기관-민간 공동	46	4	0	1	4	0	55	39.0
민간기업	19	4	0	0	0	1	24	17.0
기타	1	0	0	0	0	0	1	0.7
합계	114	19	1	1	4	2	141	100.0

- 데이터 거래소의 적합한 데이터 거래방식으로 수요자 요구 데이터를 중계하는 온디맨드 방식 중계가 50.4%로 가장 많고, 다음이 단순 중계(공급자 공급 데이터 중계), 데이터 매입 판매(22.7%) 등의 순임

< 스마트농업 데이터 거래소 설립 시 적정 거래방식 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
온디맨드 방식 중계(수요자 요구 데이터 중계)	61	6	0	0	3	1	71	50.4
단순 중계(공급자 공급 데이터 중계)	29	6	1	1	0	0	37	26.2
데이터 매입 판매	23	7	0	0	1	1	32	22.7
기타(단순중계 와 온디맨드 방식 혼합 1명 등)	1	0	0	0	0	0	1	0.7
합계	114	19	1	1	4	2	141	100.0

- 데이터 거래소 설립 시 정부나 공공기관 데이터 거래가 필요한지에 대해 절반 이상인 55.3%가 필요하다고 답함

< 스마트농업 데이터 거래소 설립 시 정부나 공공기관 데이터 거래의 필요성 >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
필요하다	60	12	1	0	4	1	78	55.3
필요하지 않다	21	3	0	0	0	1	25	17.7
잘 모르겠다	33	4	0	1	0	0	38	27.0
합계	114	19	1	1	4	2	141	100.0

- 데이터 거래소 설립 시 거래가 적합한 상품으로 원본 데이터를 분석한 분석보고서가 22.6%로 가장 많고, 다음이 거래소 정제 · 가공 원본 데이터(21.2%), 기존 원본 데이터를 거래소가 융합 · 가공한 원본 데이터(20.1%), 수요자 맞춤형 분석 서비스(19.8%), 원본 데이터(16.2%) 등의 순으로 조사됨
  · 이는 단순히 원본 데이터를 중계하기보다 수요자에게 의미 있는 정보를 제공하는 분석보고서나 서비스를 제공하고, 수요자의 데이터 활용도를 높일 수 있는 형태로의 데이터 전환이 필요함을 의미

< 스마트농업 데이터 거래소에서 거래가 적합한 상품(복수 응답 가능) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
원본 데이터 분석보고서	67	9	0	1	3	1	81	22.6
거래소 정제· 가공 원본 데이터	60	11	0	1	2	2	76	21.2
융합·가공 원본 데이터	64	3	0	0	3	2	72	20.1
수요자 맞춤형 분석 서비스	60	7	0	1	2	1	71	19.8
원본 데이터	49	6	1	0	1	1	58	16.2
기타	0	0	0	0	0	0	0	0.0
합계	300	36	1	3	11	7	358	100.0

- 데이터 거래소를 설립, 운영하는 경우 발생할 수 있는 문제점으로 양질의 신뢰성 있는 데이터 부족(16.7%)을 답한 응답자가 가장 많고, 다음이 공급자 신뢰 미흡(12.0%), 거래가격 불합리성(11.7%), 소유권이 불명확한 데이터 거래로 인한 법적 분쟁(10.0%) 등의 순임

< 스마트농업 데이터 거래소 설립, 운영 시 고려해야 할 중요 문제점(3개 이내) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
양질의 신뢰성 있는 데이터 부족	59	5	0	0	1	2	67	16.7
공급자 신뢰 미흡	36	8	0	1	2	1	48	12.0
거래가격 불합리성	39	5	0	0	2	1	47	11.7
소유권이 불명확한 데이터 거래로 인한 법적 분쟁	30	7	0	0	2	1	40	10.0
수요자 부족	34	2	1	0	1	1	39	9.7
판매 데이터의 불법유통	29	4	0	1	3	1	38	9.5
개인정보 유출로 인한 법적 분쟁	26	6	0	1	1	1	35	8.7
불필요한 데이터 저장/보관에 따른 비용 발생	24	2	0	1	2	2	31	7.7
공급자 부족	24	2	0	0	1	2	29	7.2
데이터 오류 등으로 인한 수요자 손실	23	3	0	0	0	1	27	6.7
기타	0	0	0	0	0	0	0	0.0
합계	324	44	1	4	15	13	401	100.0

(5) 스마트농업 데이터 거래 활성화를 위한 과제

○ 스마트농업 데이터 거래 활성화를 위한 과제로 양질의 신뢰성 있는 데이터의 생산, 수집 및 보급(13.9%)이 가장 중요하다고 답하고, 다음이 우수한 데이터 공급자 육성(13.6%), 데이터 품질개선 및 표준화(12.3%), 데이터 거래에 대한 인식 개선(10.0%), 데이터 관련 개인정보 보호 제도 정비(10.0%) 등의 순으로 답함

< 스마트농업 데이터 거래 활성화를 위한 과제(3개 이내) >

구분	기업	농가	대학	법인	연구기관	컨설팅	전체	
							빈도(명)	비율(%)
양질의 신뢰성 있는 데이터 생산/수집 및 보급	46	4	0	0	2	2	54	13.9
우수한 데이터 공급자 육성	42	6	0	1	2	2	53	13.6
데이터 품질개선 및 표준화	43	2	0	0	2	1	48	12.3
데이터 가치평가 방법 고도화	38	6	0	0	2	1	47	12.1
데이터 거래에 대한 인식 개선	30	4	0	0	3	2	39	10.0
데이터 관련 개인정보 보호 제도 정비	31	5	0	1	1	1	39	10.0
데이터 관련 지적 소유권 제도 정비	29	5	1	1	0	1	37	9.5
데이터 거래 및 활용 관련 서비스 개발/보급	21	4	0	1	3	1	30	7.7
데이터 활용 수요자 지원	23	5	0	0	0	1	29	7.5
데이터 거래 효율화를 위한 기술개발	11	1	0	0	0	1	13	3.3
기타	0	0	0	0	0	0	0	0.0
합계	314	42	1	4	15	13	389	100.0

○ 스마트농업 데이터 거래 활성화 또는 거래소 설립, 운영에 관한 의견을 주관식으로 질문한 결과는 다음과 같음

< 스마트농업 데이터 거래 활성화 거래소 설립, 운영에 관한 의견(주관식) >

- 거래소 설립보다는 각 회사의 데이터 관련 전문 종사자도 없고, 외부 인력으로 대체 해야 합니다. 간단하게라도 교육을 받을 수 있는 프로그램 운영이 있으면 좋겠습니다.
- 거래소 설립 이전에 주기가 긴 농업 데이터의 신뢰성 확보가 중요하며. 이를 위하여 지속적인 신뢰성 데이터를 수집하는 지원정책이 선행되어야 할 것으로 판단됨.
- 공급자 수요자 정확한 매칭과 적절한 단가 설정이 중요함.
- 공급자가 가진 데이터와 수요자가 원하는 데이터에 대해 먼저 확인할 필요가 있습니다. 그리고 부분적인 데이터는 오히려 시장에 혼란을 가져올 수 있습니다. 국내 ○○○ 작물 생산량, 출하량 등도 현재는 국내 시장 전체를 확인할 수 있는 데이터가 아닙니다.
- 공적 신뢰성 있는 기관설립 및 데이터 수집 공급자 보상과 수요개발 중요.
- 관련 업계에 대한 홍보 중요.
- 기존에 있는 데이터 형태라면 설문을 통해 알아내고 공급할 수 있는 곳을 찾으면 되지만, 새로운 형태의 데이터는 수요자가 요구하는 데이터가 정확히 어떤 것인지 알기가 쉽지 않을 것입니다. 해당 부분에 대한 데이터 생성 업체의 R&D 및 실증 연구에 대한 지원이 유효한 데이터를 얻는데 필요할 것으로 생각됩니다.
- 데이터 거래소 설립 시 민간 의견이 필요할 것으로 예상됨.
- 데이터 거래소 설립 후 데이터 판매전략 등에 대한 가이드가 필요할 것 같다.

· 데이터 거래소 자체 UI의 가독성 및 명확성 확보를 통한 데이터 거래 활성화가 될 수 있기를 바란다.
· 데이터 거래소가 창립된다고 해서 처음부터 다수의 데이터 공급자와 수요자가 넘쳐나지는 않을 것입니다. 따라서 거래소 개설 초기에는 유용한 거래 상품(데이터 상품)의 기획, 발굴 즉, 상품화에 많은 투자가 되어야 할 것입니다.
· 수요자(주로 기업, 대학, 연구소)들이 어떤 데이터 상품을 원하는지를 잘 기획하고, 이를 선도적으로 발굴하고 매입하여 거래소 쇼핑몰을 매력도 있도록 만드는 것이 중요할 것입니다.
· 데이터 거래에 집중된 기능으로 운영. 민간이 담당할 복합분석 지양
· 데이터 기반 실수요자 지원사업도 검토해 주세요.
· 데이터 카테고리의 세분화가 이뤄지면 좋을 것 같음.
· 데이터에 따른 생산 가치가 수치적으로 표현이 가능한 상황이 되었으면 합니다.
· 명확한 데이터 거래방식 확립이 필요함.
· 법적근거 확보 필요.
· 비슷한 데이터 유형 간 데이터 경쟁 유도가 필요할 것 같다.
· 신뢰성 있는 데이터 필요.
· 신뢰성 있는 데이터 확보가 중요함.
· 양질 데이터 확보가 시급.
· 양질의 신뢰성 데이터 및 꾸준한 관리 감독이 필요.
· 자사의 사례를 들면 운영 측면의 경우 데이터 분석/이해를 통하여 기계 사용의 효율을 향상시키고, 처리량을 높일 수 있음.
· 기계적인 측면의 경우는 기계 고장의 신호를 실시간으로 확인하여 바로 대처함으로써 A/S 비용을 줄일 수 있으나, 이에 대한 수요자(ex.농장)는 이러한 데이터에 대한 가치와 비용에 대하여 지불을 꺼림.
· 회사(데이터 개발/공급자)의 경우 데이터 생성을 위한 프로그램 개발과 부품비(ex.센서, 제어기 등), 유지보수비(ex.모니터링을 위한 인건비, 프로그램 개발/업그레이드 비용)등과 같이 계속적으로 비용이 발생하나, 데이터 제공으로 인한 매출을 창출하기가 매우 어려워 이러한 비용을 감당하기 어려움. 이렇게 되면 데이터 제공 자체도 어려울 뿐 아니라 양질의 신뢰성 있는 데이터를 제공하기 위한 투자를 하기가 어렵게 됨.
· 정보의 보안이 제일 중요할 것 같습니다.
· 지역 관련 한정된 데이터로 다른 지역에서 사용 시 참조 위주로 되기 쉬우므로 데이터를 활용한 1차 참조가 우선적이고, 이를 이용한 생육 관리에 대한 책임은 구매자가 부담해야 함.
· 처음에는 신뢰성 등 여러 요인이 있겠지만 반드시 필요하다고 판단.
· 충분한 검토 후 진행되기를 바람.
· 투명한 데이터 거래 과정 공개를 통한 신뢰성 확보가 필요함.
· 필요한 시스템이라고 판단된다.
· 홍보가 잘되어야 하겠습니다.
· 활성화 이후 홍보가 필요할 것으로 예상.

라. 잠재 수요자 및 농업 전문가 조사 결과

(1) 응답자 현황

○ 응답자 현황은 표와 같음

- 성별은 모두 남성이고, 연령대는 50세 이상이 54.5%(6명)로 가장 많음
- 종사 분야는 유통기업(롯데마트, SSG, 로컬푸드) 36.4%(4명), 대학 36.4%(4명), 연구기관 18.2%(2명), 언론사 9.1%(1명)의 순이고, 분야경력은 최소 12년으로 16~20년 36.4%(4명), 21~25년 36.4%(4명), 25년 초과(최대 27년) 18.2%(2명), 10~15년 9.1%(1명)의 순임

< 잠재 수요자 및 농업 전문가 대상 조사 응답자 일반현황 >

구분		빈도(명)	비율(%)	비고
성별	남성	11	100.0	
	합계	11	100.0	
연령대	40~44세	3	27.3	
	45~49세	2	18.2	
	50세 이상	6	54.5	
	합계	11	100.0	
종사분야	유통기업	4	36.4	MD(롯데마트, SSG), 본부장(로컬푸드)
	대학	4	36.4	교수
	연구기관	2	18.2	농진청(연구사), 농협경제연구소(부연구위원)
	언론	1	9.1	부국장
	합계	11	100.0	
분야경력	10~15년	1	9.1	
	16~20년	4	36.4	
	21~25년	4	36.4	
	25년 초과	2	18.2	
	합계	11	100.0	

(2) 농업 데이터 이용 현황 및 데이터 거래에 대한 인식

○ 응답자 또는 소속기관이 필요로 하는 데이터에 대해 11명 모두 생산량 등 생산 현황, 생산비/경영 현황, 유통/판매 현황 등 3개 종류의 데이터가 필요하다고 답했고, 다음이 농업기상(54.5%, 6명), 농작물 생육환경(45.5%, 5명) 등의 순임

< 필요로 하는 농업 데이터 종류(복수 응답 가능) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
생산량 등 생산 현황	11	19.6	100.0
생산비/경영 현황	11	19.6	100.0
유통/판매 현황	11	19.6	100.0
농업기상	6	10.7	54.5
농작물 생육환경	5	8.9	45.5
요소 투입량	5	8.9	45.5
농작물 생육	4	7.1	36.4
병해충	2	3.6	18.2
기타(농가경제조사, 농업경영체조사 등)	1	1.8	9.1
합계	56	100.0	-

- 필요로 하는 농업 데이터 유형에 대해 11명 모두 정형 데이터(수치 등)가 필요하다고 답했고, 다음이 비정형 텍스트(72.7%, 8명), 비정형 이미지(45.5%, 5명) 등의 순임

< 필요로 하는 농업 데이터 유형(복수 응답 가능) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
정형 데이터(수치 등)	11	42.3	100.0
비정형 텍스트	8	30.8	72.7
비정형 이미지	5	19.2	45.5
비정형 오디오	1	3.8	9.1
비정형 동영상	1	3.8	9.1
기타	0	0.0	0.0
합계	26	100.0	-

- 농업 데이터 확보 방법으로 11명 모두 통계청 MDIS와 같이 데이터를 제공하는 공공사이트에서 확보한다고 답했고, 다음이 직접 조사 및 수집(81.8%, 9명), 데이터 생산 및 수집 기관이나 기업에 직접 요청(81.8%, 9명), 데이터 전문기관이나 기업에 의뢰(18.2%, 2명)의 순임

< 필요로 하는 농업 데이터 확보 방법(복수 응답 가능) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
데이터 제공 공공사이트(통계청 MDIS 등) 이용	11	35.5	100.0
직접 조사/수집	9	29.0	81.8
데이터 생산/수집 기관이나 기업에 직접 요청(구매 포함)	9	29.0	81.8
데이터 전문기관이나 기업에 의뢰	2	6.5	18.2
데이터 거래소 이용(구매 포함)	0	0.0	0.0
기타	0	0.0	0.0
합계	31	100.0	-

○ 농업 데이터를 구매한 경험에 대해 27.3%(3명)가 구매 경험이 있다고 답함

< 농업 데이터 구매 경험 >

구분	빈도(명)	비율(%)
구매 경험 있음	3	27.3
구매 경험 없음	8	72.7
합계	11	100.0

- 데이터 구매 경험자(3명)가 구매한 데이터의 종류는 유통 및 판매 현황(3명), 생산량 등 생산 현황(2명), 생산비 및 경영 현황(2명), 농작물 생육(1명), 요소 투입량(1명)임

< 구매 농업 데이터 종류(복수 응답 가능; 구매 경험자 3명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
유통/판매 현황	3	33.3	100.0
생산량 등 생산 현황	2	22.2	66.7
생산비/경영 현황	2	22.2	66.7
농작물 생육	1	11.1	33.3
요소 투입량	1	11.1	33.3
농업기상	0	0.0	0.0
농작물 생육환경	0	0.0	0.0
병해충	0	0.0	0.0
기타	0	0.0	0.0
합계	9	100.0	-

- 농업 데이터 구매처는 공공기관(3명), 농가(1명), 농기업(1명)으로 조사됨

< 농업 데이터 구매처(복수 응답 가능; 구매 경험자 3명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
공공기관	3	60.0	100.0
농가	1	20.0	33.3
농기업	1	20.0	33.3
데이터 전문기업	0	0.0	0.0
데이터 거래소 또는 플랫폼	0	0.0	0.0
기타	0	0.0	0.0
합계	5	100.0	-

- 농업 데이터 구매 형태는 양도 계약(소유권 이전) 2명, 라이선스 계약(이용 허락) 1명임

< 농업 데이터 구매 형태(구매 경험자 3명 대상) >

구분	빈도(명)	비율(%)
양도 계약(소유권 이전)	2	66.7
라이선스 계약(이용 허락)	1	33.3
합계	3	100.0

- 구매 데이터를 가격책정 시 참고 정보로 데이터 거래소 또는 플랫폼의 판매가격 정보(2명), 업종 내 유사 데이터 구매사례(2명), 데이터 생산 및 수집 비용(1명)이었음

< 농업 데이터 가격책정 시 참고 정보(2개 이내; 구매 경험자 3명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
데이터 거래소 또는 플랫폼의 판매가격 정보	2	40.0	66.7
업종 내 유사 데이터 구매사례	2	40.0	66.7
데이터 생산/수집 비용(원가)	1	20.0	33.3
자사의 과거 구매가격	0	0.0	0.0
기타	0	0.0	0.0
합계	5	100.0	-

○ 농업 데이터 구매 경험자(3명) 1명만이 데이터 거래소를 이용한 경험이 있었음

- 이용한 거래소는 통계청 거래 사이트(MDIS)로 데이터 직거래에 대비한 거래소의 장점으로 적절한 거래 상대를 찾기 위한 노력 불필요, 거래 관련 계약체결 등 거래 절차 용이, 구매대금을 투명하게 정산할 수 있음 등 3가지를 답했으며, 직거래 대비 단점으로는 판매 조건이나 가격 등을 지나치게 제약한다는 점이라고 답함

< 데이터 직거래 대비 거래소 거래의 장점(복수 응답 가능; 거래소 경유 구매 경험자 1명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
적절한 거래 상대 찾기 위한 노력 불필요	1	33.3	100.0
거래 관련 계약체결 등 거래 절차 용이	1	33.3	100.0
구매대금을 투명하게 정산할 수 있음	1	33.3	100.0
판매 조건, 가격 등의 협상 불필요	0	0.0	0.0
계약체결 후 거래이행 확인 불필요	0	0.0	0.0
기타	0	0.0	0.0
합계	3	100.0	-

< 데이터 직거래 대비 거래소 거래의 단점(복수 응답 가능; 거래소 경유 구매 경험자 1명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
판매 조건, 가격 등을 지나치게 제약	1	100.0	100.0
거래소의 데이터에 대한 불필요한 간섭	0	0.0	0.0
거래소의 불필요한 절차 또는 서류 요구	0	0.0	0.0
제공서비스 대비 비싼 수수료 요구	0	0.0	0.0
거래소의 구매기업 또는 개인정보 요구	0	0.0	0.0
기타	0	0.0	0.0
합계	1	100.0	-

- 데이터 거래소 이용 경험이 없는 구매 경험자(2명)의 경우 거래소를 이용하지 않은 이유로 데이터 거래소가 있는지를 몰랐기 때문이라고 답함

< 데이터 구매 시 거래소 미이용 이유(복수 응답 가능; 거래소 미경유 구매 경험자 2명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
데이터 거래소가 있는지를 몰랐음	2	100.0	100.0
신뢰할 만한 거래소를 찾지 못함	0	0.0	0.0
거래소의 데이터에 대한 불필요한 간섭	0	0.0	0.0
판매 조건, 가격 등을 지나치게 제약	0	0.0	0.0
거래소의 불필요한 절차 또는 서류 요구	0	0.0	0.0
제공서비스 대비 비싼 수수료 요구	0	0.0	0.0
거래소의 구매기업 또는 개인정보 요구	0	0.0	0.0
합계	0	100.0	-

○ 향후 데이터 거래소에서 농업 데이터를 구매할 의향에 대해 81.8%(9명)가 구매 의향이 있다고 답하고, 18.2%(2명)는 잘 모르겠다고 답함

- 데이터 거래소에서 농업 데이터를 구매할 의향이 있다고 답한 9명에게 소유권 이전 의향을 질문한 결과 55.6%(5명)가 소유권 이전 의향이 있다고 답하고, 44.4%는 구매 의향이 없거나 잘 모르겠

다고 답함

< 향후 농업 데이터 거래소에서 데이터를 구매할 의향 >

구분	빈도(명)	비율(%)
구매 의향 있음	9	81.8
구매 의향 없음	0	0.0
잘 모르겠음	2	18.2
합계	11	100.0

< 거래소에서 농업 데이터 구매할 때 소유권을 이전받을 의향(구매 의향자 9명 대상) >

구분	빈도(명)	비율(%)
소유권 이전받을 의향 있음	5	55.6
소유권 이전받을 의향 없음	3	33.3
잘 모르겠음	1	11.1
합계	9	100.0

- 거래소에서 농업 데이터를 구매할 의향에 대해 잘 모르겠다고 답한 이유(3명)는 적절한 구매 채널 확보의 어려움(2명)과 구매가격 책정의 어려움(1명)이었음

< 향후 농업 데이터 거래소에서 데이터를 구매할 의향이 없거나 잘 모르겠다고 답한 이유

(복수 응답 가능; 3명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
적절한 구매 채널 확보 어려움	2	66.7	100.0
구매가격 책정 어려움	1	33.3	50.0
데이터의 품질이 낮아 구매에 부적합	0	0.0	0.0
구매 후 사후 서비스 미비	0	0.0	0.0
구매 데이터의 불법 유통에 따른 손실 우려	0	0.0	0.0
개인정보 등 구매에 따른 법적 문제 발생 우려	0	0.0	0.0
기타	0	0.0	0.0
합계	3	100.0	-

(3) 농업 데이터의 가치에 대한 인식

○ 농업 데이터의 판매가격 책정 시 중요하다고 생각하는 속성으로 데이터 신뢰성과 데이터 정확성을 가장 중요하게 인식하고(72.7%, 8명), 다음이 제공정보 종류(45.5%, 5명), 데이터 품질(45.5%, 5명) 등의 순임

< 농업 데이터 판매가격 책정 시 중요하게 생각하는 속성(3개 이내) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
데이터 신뢰성	8	22.2	72.7
데이터 정확성	8	22.2	72.7
제공정보 종류(기상, 경영 등)	5	13.9	45.5
데이터 품질	5	13.9	45.5
데이터 생산/수집 비용	3	8.3	27.3
데이터 적시성	3	8.3	27.3
데이터 차별성/유일성	3	8.3	27.3
데이터 제공기관의 신뢰성	1	2.8	9.1
유형(텍스트, 이미지 등)	0	0.0	0.0
기타	0	0.0	0.0
합계	36	100.0	-

○ 농업 데이터의 가치에 영향을 미치는 주요 속성의 전반적 수준에 대해 5점 만점에 3.636점으로 답하고, 세부 속성 중에서 데이터 정확성(4.000점)의 수준을 가장 높게 인식하고, 다음이 데이터 신뢰성(3.818점), 데이터 품질(3.636점), 데이터 유용성(3.636점), 제공기관의 신뢰성(3.545점), 데이터 차별성 또는 유일성(3.364점), 데이터 적시성(3.364점)의 순임

< 농업 데이터 가치에 영향을 미치는 주요 속성의 수준 인식 >

구분	데이터 정확성		데이터 신뢰성		데이터 품질		데이터 유용성	
	빈도(명)	비율(%)	빈도(명)	비율(%)	빈도(명)	비율(%)	빈도(명)	비율(%)
매우 낮음	1	9.1	1	9.1	1	9.1	0	0.0
약간 낮음	1	9.1	2	18.2	2	18.2	2	18.2
보통	2	18.2	1	9.1	2	18.2	4	36.4
약간 우수	0	0.0	1	9.1	1	9.1	1	9.1
매우 우수	7	63.6	6	54.5	5	45.5	4	36.4
합계	11	100.0	11	100.0	11	100.0	11	100.0
평균	4.000		3.818		3.636		3.636	
구분	제공기관의 신뢰성		데이터 차별성/유일성		데이터 적시성		전반적 수준	
	빈도(명)	비율(%)	빈도(명)	비율(%)	빈도(명)	비율(%)	빈도(명)	비율(%)
매우 낮음	0	0.0	0	0.0	1	9.1	0	0.0
약간 낮음	2	18.2	4	36.4	2	18.2	3	27.3
보통	2	18.2	1	9.1	3	27.3	1	9.1
약간 우수	6	54.5	4	36.4	2	18.2	4	36.4
매우 우수	1	9.1	2	18.2	3	27.3	3	27.3
합계	11	100.0	11	100.0	11	100.0	11	100.0
평균	3.545		3.364		3.364		3.636	

주. 평균은 ‘매우 낮음 = 1점, 약간 낮음 = 2점, 보통 = 3점, 약간 우수 = 4점, 매우 우수 = 5점’을 적용하여 산출한 값임

○ 데이터 생산 및 수집에 지출되는 원가를 기반으로 가격을 책정하는 경우 적정 원가 회수 기간은 평균 3.00년(최대 10년, 최소 1년) 이내로 답하고, 최소 이윤율은 평균 21.27%(최대 60%, 최소 2%)로 답함

- 잠재 공급자 인식과 비교할 때 원가 회수 기간(잠재 공급자 평균 4.28년)은 짧고, 최소 이윤율(잠재 공급자 평균 39.75%)은 낮음

< 원가 기반 가격책정 시 적정 원가 회수 기간 및 최소 이윤율 인식 >

구분	빈도(명)	평균	최대	최소	표준편차
적정 원가 회수 기간(년)	11	3.00	10	1	2.61
적정 최소 이윤율(%)	11	21.27	60	2	19.31

○ 농업 데이터의 수명주기와 관련하여 데이터가 수요자에 가치를 제공하는 기간은 평균 6.55년(최대 30년, 최소 1년), 법적 문제 등에 대비한 적정 데이터 보존 기간은 평균 8.91년(최대 30년, 최소 1년)은 답함

- 잠재 공급자와 비교하면, 수요자에 가치를 제공하는 기간(잠재 공급자 평균 6.62년)은 길고, 보존 기간(잠재 공급자 평균 7.83년)은 짧게 인식

< 농업 데이터의 가치 제공 및 보관 기간에 대한 인식 >

구분	빈도(명)	평균	최대	최소	표준편차
수요자에게 가치를 제공하는 최대 기간(년)	11	6.55	30	1	8.13
법적 문제 등에 대비한 적정 데이터 보존 기간(년)	11	8.91	30	1	8.53

○ 판매 데이터의 신뢰성 확보를 위해 제공이 필요한 정보로 11명 모두 데이터 생산 및 수집 방법을 가장 필요하다고 답하고, 다음이 데이터 원천정보(72.7%, 8명), 데이터 샘플 및 활용 사례(72.7%, 8명), 분야경력 및 실적(54.5%, 6명), 데이터 품질 보증서(45.5%, 5명) 등의 순임

- 1순위와 2순위는 잠재 공급자가 제공 의향이 있는 정보(1순위 데이터 생산 및 수집 방법, 2순위 데이터 원천정보)와 같음

< 판매 데이터의 신뢰성 확보를 위한 제공이 필요한 정보(복수 응답 가능) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
데이터 생산/수집 방법	11	26.2	100.0
데이터 원천정보	8	19.0	72.7
데이터 샘플 및 활용 사례	8	19.0	72.7
분야경력 및 실적	6	14.3	54.5
데이터 품질 보증서	5	11.9	45.5
전문가 평가서	3	7.1	27.3
신용정보	1	2.4	9.1
회계정보(수입 등)	0	0.0	0.0
합계	42	100.0	-

○ 데이터 오류 등으로 인한 구매자 손실 보상에 대해 응답자 기준 72.7%(8명)는 조건부 또는 조건과 상관없이 보상이 필요하다고 답하고(귀책 사유가 명확한 경우 보상 필요 63.6%, 손실 발생 원인 확인과 상관없이 보상 필요 9.1%), 36.4%(4명)는 데이터 보상이 부적절하다(데이터 활용은 구매자 자율로 손실 보상 부적절 18.2%, 손실 원인 확인이 불가해 보상 부적절 18.2%)고 답함

- 이는 53.2%가 데이터 보상이 부적절하다는 잠재 공급자의 인식과 차이가 있음

< 데이터 오류 등으로 인한 구매자 손실 보상에 대한 인식 >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
데이터 활용은 구매자 자율로 손실 보상 부적절	2	16.7	18.2
손실 원인 확인이 불가해 보상 부적절	2	16.7	18.2
손실 발생 원인 확인과 상관없이 보상 필요	1	8.3	9.1
귀책 사유가 명확한 경우 보상 필요	7	58.3	63.6
기타	0	0.0	0.0
합계	12	100.0	-

주. 1명은 데이터 활용은 구매자 자율로 손실 보상 부적절과 귀책 사유가 명확한 경우 보상 필요 등 2개 응답 항에 복수로 답함

(4) ‘(가칭) 스마트농업데이터거래소’ 설립에 대한 인식

○ 스마트농업 데이터 거래소 설립에 대해 72.7%(8명)가 필요하다고 답하고, 27.3%는 잘 모르겠다고 답하여 잠재 공급자의 필요성 인식(필요하다 48.9%, 필요하지 않다 14.9%, 잘 모르겠다 36.2%)과 차이가 있음

< (가칭) 스마트농업데이터거래소 설립 필요성 인식 >

구분	빈도(명)	비율(%)
필요하다	8	72.7
필요하지 않다	0	0.0
잘 모르겠다	3	27.3
합계	11	100.0

- 스마트농업 데이터 거래소 설립의 필요성에 대해 잘 모르겠다고 답한 응답자(3명)의 경우 양질의 신뢰성 있는 데이터 상품 미흡, 지적 재산권 등 데이터 관련 거래 제도 미비, 데이터 거래에 대한 인식 미흡 등을 그 이유로 답함

< (가칭) 스마트농업데이터거래소 설립 필요성에 대해 잘 모르겠다고 답한 이유(3명 대상) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
양질의 신뢰성 있는 데이터 상품 미흡	3	27.3	100.0
지적 재산권 등 데이터 거래 관련 제도 미비	3	27.3	100.0
데이터 거래에 대한 인식 미흡	2	18.2	66.7
데이터 거래시장 공급자와 수요자 부족	1	9.1	33.3
적정한 데이터 가격책정 어려움	1	9.1	33.3
판매 데이터의 불법 유통으로 인한 혼란	1	9.1	33.3
이미 설립된 데이터 거래소로 충분	0	0.0	0.0
개인정보 유출 등 거래에 따른 부작용	0	0.0	0.0
기타	0	0.0	0.0
합계	11	100.0	-

- 스마트농업 데이터 거래소 설립 시 적정 운영 주체로 공공기관이나 공사가 54.5%(6명)로 가장 많고, 다음이 공공기관과 민간 공동(36.4%, 4명), 민간기업(9.1%, 1명) 등의 순으로 잠재 공급자 인식과 순위가 같음

< 스마트농업 데이터 거래소 설립 시 적정 운영 주체 >

구분	빈도(명)	비율(%)
공공기관/공사	6	54.5
민간기업	1	9.1
공공기관-민간 공동	4	36.4
기타	0	0.0
합계	11	100.0

- 데이터의 거래소의 적정 거래방식에 대해 11명 모두 수요자 요구 데이터를 중계하는 온디맨드 방식의 중계가 적정하다고 답하여 잠재 공급자의 인식(온디맨드 방식 중계 50.4%, 단순 중계 26.2%, 데이터 매입 판매 22.7%, 단순 중계와 온디맨드 방식 혼합 0.7%)과 차이가 있음

< 스마트농업 데이터 거래소 설립 시 적정 거래방식 >

구분	빈도(명)	비율(%)
단순 중계(공급자 공급 데이터 중계)	0	0.0
온디맨드 방식 중계(수요자 요구 데이터 중계)	11	100.0
데이터 매입 판매	0	0.0
기타	0	0.0
합계	11	100.0

- 데이터 거래소 설립 시 정부나 공공기관 데이터 거래가 필요한지에 대해 81.8%(9명)가 필요하다고 답하여 잠재 공급자 인식(필요하다 55.3%, 필요하지 않다 17.7%, 잘 모르겠다 27.0%)과 차이가 있음

< 스마트농업 데이터 거래소 설립 시 정부나 공공기관 데이터 거래의 필요성 >

구분	빈도(명)	비율(%)
필요하다	9	81.8
필요하지 않다	2	18.2
잘 모르겠다	0	0.0
합계	11	100.0

- 데이터 거래소 설립 시 거래가 적합한 상품으로 원본 데이터 분석보고서(72.7%, 8명)와 수요자 맞춤형 분석 서비스(72.7%, 8명)가 가장 많고, 다음이 원본 데이터(63.6%, 7명), 거래소 정제 · 가공 원본 데이터(54.5%, 6명), 융합 · 가공 원본 데이터(54.5%, 6명)의 순이며, 수요자 맞춤형 분석 서비스(잠재 공급자 4순위)와 원본 데이터(잠재 공급자 5순위)의 순위가 잠재 공급자보다 높음

< 스마트농업 데이터 거래소에서 거래가 적합한 상품(복수 응답 가능) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
원본 데이터 분석보고서	8	22.9	72.7
수요자 맞춤형 분석 서비스	8	22.9	72.7
원본 데이터	7	20.0	63.6
거래소 정제·가공 원본 데이터	6	17.1	54.5
융합·가공 원본 데이터	6	17.1	54.5
기타	0	0.0	0.0
합계	35	100.0	-

- 데이터 거래소를 설립, 운영하는 경우 발생할 수 있는 문제점으로 양질의 신뢰성 있는 데이터 부족(81.8%, 9명)을 가장 중요하다고 답하고, 다음이 수요자 부족(54.5%, 6명), 공급자 부족(36.4%, 4명), 데이터 오류 등으로 인한 수요자 손실(36.4%, 4명) 등의 순이며, 양질의 신뢰성 있는 데이터 부족을 가장 중요하게 인식한다는 점은 잠재 공급자와 같으나, 공급자 신뢰 미흡(잠재 공급자 2순위), 거래가격 불합리성(잠재 공급자 3순위), 소유권이 불명확한 데이터 거래로 인한 법적 분쟁(잠재 공급자 4순위)의 중요도는 상대적으로 낮게 인식

< 스마트농업 데이터 거래소 설립, 운영 시 고려해야 할 중요 문제점(3개 이내) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
양질의 신뢰성 있는 데이터 부족	9	27.3	81.8
수요자 부족	6	18.2	54.5
공급자 부족	4	12.1	36.4
데이터 오류 등으로 인한 수요자 손실	4	12.1	36.4
거래가격 불합리성	3	9.1	27.3
판매 데이터의 불법 유통	3	9.1	27.3
공급자 신뢰 미흡	2	6.1	18.2
소유권이 불명확한 데이터 거래로 인한 법적 분쟁	2	6.1	18.2
개인정보 유출로 인한 법적 분쟁	0	0.0	0.0
불필요한 데이터 저장, 보관에 따른 비용 발생	0	0.0	0.0
기타	0	0.0	0.0
합계	33	100.0	-

(5) 스마트농업 데이터 거래 활성화를 위한 과제

○ 스마트농업 데이터 거래 활성화를 위한 과제로 잠재 공급자와 마찬가지로 양질의 데이터 생산, 수집 및 보급(81.8%, 9명)을 가장 중요하게 인식하고, 다음이 데이터 품질개선 및 표준화(63.6%, 7명), 데이터 거래 및 활용 관련 서비스 개발 및 보급(45.5%, 5명), 데이터 관련 지적 소유권 제도 정비(27.3%, 3명) 등의 순임

- 잠재 공급자가 중요하게 인식한 우수한 데이터 공급자 육성(잠재 공급자 2순위), 데이터 가치평가 방법 고도화(잠재 공급자 3순위)에 대한 중요도 인식은 상대적으로 낮음

< 스마트농업 데이터 거래 활성화를 위한 과제(3개 이내) >

구분	빈도(명)	비율(%)	
		합계 기준	응답자 기준
양질의 신뢰성 있는 데이터 생산/수집 및 보급	9	26.5	81.8
데이터 품질개선 및 표준화	7	20.6	63.6
데이터 거래 및 활용 관련 서비스 개발, 보급	5	14.7	45.5
데이터 관련 지적 소유권 제도 정비	3	8.8	27.3
데이터 거래에 대한 인식 개선	2	5.9	18.2
우수한 데이터 공급자 육성	2	5.9	18.2
데이터 가치평가 방법 고도화	2	5.9	18.2
데이터 관련 개인정보 보호 제도 정비	1	2.9	9.1
데이터 활용 수요자 지원	1	2.9	9.1
데이터 거래 효율화를 위한 기술개발	1	2.9	9.1
기타(데이터 거래 홍보 및 마케팅)	1	2.9	9.1
합계	34	100.0	-

○ 스마트농업 데이터 거래 활성화나 거래소 설립, 운영에 관한 추가 의견은 다음과 같음

< 스마트농업 데이터 거래 활성화 거래소 설립, 운영에 관한 의견(주관식) >

(전문가 1, 연구기관)
- 특히 농식품의 유통/판매 분야 데이터는 생산하여 거래할 수 있는 데이터가 많음
- 하지만 공공부문은 이유 없이 공개하지 않거나, 유용한 데이터 생성에 대한 전문성과 관심이 부족한 실정이며, 민간 부문도 공유가 가능함에도 공개를 거부하는 분야가 많은 상황 이에 따라 농업 전반의 연구와 이에 따른 발전이 굉장히 더디게 진행되었다고 생각함
- 스마트농업 데이터 거래 및 데이터 거래소 설립은 이러한 관행들을 개선하기 위해서 반드시 필요하다고 생각
- 데이터 거래소 적정 운영 주체의 경우 공공기관과 민간 모두 장단점이 명확하여 공동으로 운영하는 것이 효과적인 결과를 도출할 것으로 판단되고, 또한 정부나 공공기관에서 생산/수집한 데이터도 거래가 된다면 일종의 성과지표로 활용되어 양질의 데이터로 전환될 수 있음
- 데이터 수요자가 관심을 가질 수 있는 유용하고 정확한 데이터가 공급된다면 자연스레 수요가 늘어나 거래시장이 발달하고 농업 전 분야로 확산될 수 있을 것으로 기대

(전문가 2, 연구기관)
- 데이터 소유권과 관련하여 사용권이 아닌 소유권이 이전되면, 이전매매가 반복될 시, 가격이 지속 상승할 우려
- 원가 기반 가격책정 시 데이터가 3년을 초과할 경우 'out of date'가 되어 데이터의 유용성이 사라지기 때문에 3년 이내 회수 필요

(전문가 3, 유통기업)
- 데이터의 신뢰성이 최우선이 되어야 거래 가치가 발생하리 생각하고, 국가나, 공공기관의 신뢰성에 대한 보장(인증) 등이 사전에 준비되어야 함
- 수요자가 필요에 따라, 가공할 수 있는 형태의 가변성이 필요하다고 생각(피벗 기능 등)
- 데이터 활용 사례 및 우수 사례 등의 성공모델이 있다면 더욱 활성화 되리라 생각
- 데이터 보안성이 사전에 확보되어야 하는데, 한번 구매 후, 불법 유통되거나, 남용되어 진다며, 데이터의 가치가 감소할 것임
- 데이터의 안정성(바이러스, 악성코드, 오류)이 보장되어야 하고, 무분별한 데이터의 유통은 또 다른 부작용이 발생할 여지가 큼
- 데이터 생산자의 자율성과, 표준 규격, 검증 시스템 등이 적정하게 준비되어야 거래가 활성화 될 수 있음

(전문가 4, 유통기업)
- 유통기업의 경우 주로 공급자 미팅을 통해 데이터 확보
- 데이터의 정형과 비정형 시스템을 구축하고, 수요자 입장에서 쉽게 가공할 수 있도록 비정형의 데이터 공급이 중요할 것으로 판단됨

(전문가 5, 유통기업)
- 사업이 활성화되고, 지속적인 투자가 이루어지기 위해 정부 및 기관 차원의 지속적인 지원과 투자가 뒷받침되는 모델로 접근하였으면 함

- 데이터 거래가 활성화되려면 데이터를 이용한 부가가치 창출이 가능해야 하고, 데이터 연관 사업모델 필요하나 국내는 미숙하고, 특히 이러한 성과 창출을 위해서는 유료 거래 활성화가 필요하나 데이터를 돈 주고 산다는 데 거부감 있음
- 그린랩스가 투자금을 재원으로 정보를 바탕으로 팜모닝 사이트 등을 활용하여 유통사업 하려 했고, 많은 유통 인력 영입했으나 일년 반 정도만에 사업을 정리하는 실정
- 한국 현실에서 정부가 농업 분야 데이터를 이용한 부가가치 창출 사업의 마중물 역할이 필요

(전문가 6, 대학)
- 거래소 설립 시 공공데이터 거래가 필요하나, 이때는 '거래'가 아닌 공개 필요
- 거래소 설립, 운영 시 발생할 수 있는 문제점과 관련하여 '⑦ 소유권이 불명확한 데이터 거래로 인한 법적 분쟁'과 '⑨ 불필요한 데이터 저장, 보관에 따른 비용 발생' 문제는 계약서 작성 시 관련 조항을 마련함으로써 예방할 수 있으리라 판단

(전문가 7, 대학)
- 데이터 활용 분야 발굴 및 사례 제시
- 이용자 친화 형태의 데이터 포맷 형성

(전문가 8, 대학)
- 스마트농업 데이터수집 및 활용은 필요할 것으로 생각되며, 데이터의 경우 많은 사람들이 활용할 수 있어야 가치가 있고, 그에 맞는 데이터 거래소 설립이 필요함
- 원가 기준 판매가격 책정 시 데이터 조사 및 산출 시기별로 차이가 있고, 매년 산출되는 데이터이니 격년으로 산출되는 데이터 인지 여부가 중요

(전문가 9, 대학)
- 데이터 내용의 충분성과 성숙한 거래환경이 조성될 필요가 있음
- 데이터를 이용한 활용 효과가 명시적으로 발생해야 거래 활성화 가능
- 스마트농업을 포함하여 농업 전반의 데이터를 거래할 수 있는 자격을 부여하는 방안도 고려될 필요가 있는데, 스마트농업 or 농업 데이터거래사 등의 도입도 필요

(전문가 10, 언론사)
- 데이터의 품질 확보와 수요자가 신뢰할 수 있는 데이터에 대한 객관적인 인증제도 도입

마. 생산자 측면에서의 소유권 문제 해결 방안 조사

(1) 데이터 거래시 예상되는 소유권 문제 조사

○ 법률상 보호 체계에 따른 데이터 유형

- 데이터의 가공 단계에 따른 한 위와 같은 분류는 데이터 거래 관련 사업 모델을 파악하는 데에는 유리하나, 데이터 거래에 수반되는 법적 및 제도적 이슈를 파악하기 위해서는 개인정보에 해당하는 데이터와 비개인정보에 해당하는 데이터의 유형에 따른 분류가 더욱 적절함.
- 개인정보에 해당하는 데이터는 '개인정보 보호법」에 의하여 법적으로 정의되는데, 법 제2조 제1호에서는 "성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보, 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보를 가맹 처리함으로써 원래의 상태로 복원하기 위한 추가 정보의 사용, 결합 없이 이는 특정 개인을 알아볼 수 없는 정보(가명 정보)"로 개인정보를 정의하고 있음. 개인정보는 그 자체로는 유통이나 거래의 대상이 된다고 보기 어려우며, 가명처리나 익명처리 등의 비식별화 과정을 거쳐야만 이전이 가능함.
- 반면 비개인정보에 해당하는 데이터는 주로 기계 생성 데이터(machine-generated data)와 기업의 제조•영업 과정에서 발생하는 데이터, 정부에서 생산하는 데이터가 있으며, 특별한 경우로 개인정보 등 익명화 처리를 한 데이터가 포함됨. 비개인데이터의 경우 규제를 위한 일반화된 법제가 없기 때문에 데이터의 사회적 가치를 극대화하기 위한 자유로운 이동과 활용이 필요하며, 데이터 거래에서도 비개인정보에 해당하는 데이터가 핵심이 됨.

< 데이터의 특성과 파생되는 데이터 거래의 특징 >

데이터 특성	설명	데이터 거래의 특징
비가역성	가공이 이루어질수록 범용성이 저하	원시데이터에 대한 수요 증가
유연성	분할이나 결합, 가공 등 변형이 용이	협상에 의한 맞춤형 거래 발생가격 산정이 어려움
결합성	다른 데이터와 결합하여 가치가 변화	표현 형식의 일치(표준화) 필요
경험성	유용성 판단을 위해서는 분석 필요	품질 및 가치 평가 어려움

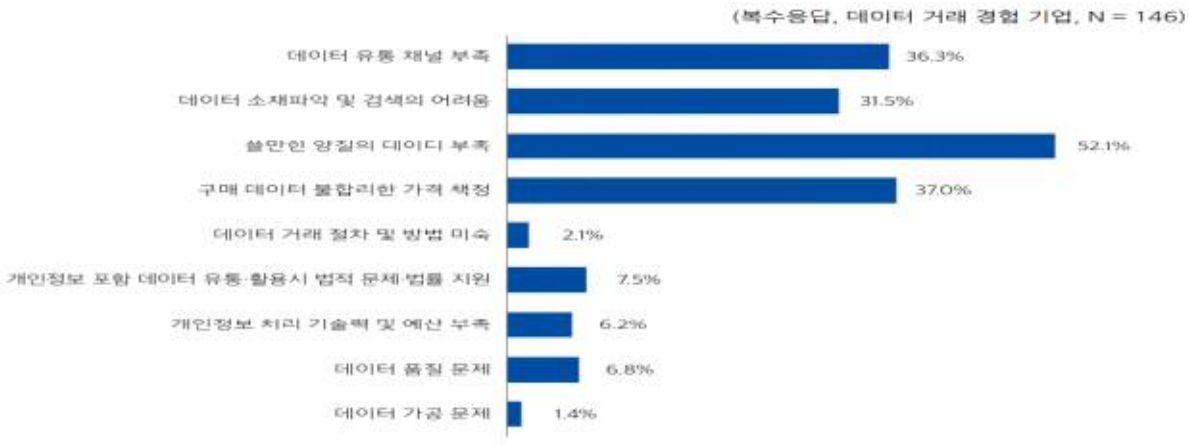
○ 데이터 산업 현황 조사 결과에 따른 문제점 도출

< 데이터 수집 방법별 비중 >

구분	비중
데이터 거래를 통해 데이터 확보	5.9%
공공기관에서 제공되는 데이터(공공데이터)	10.9%
수립 솔루션 (tool) 등에 의한 웹데이터 수집	6.5%
온라인 회원 및 고객이 이용 의한 데이터	24.9%
소설(SNS) • 인터넷 등을 통한 데이터 수집	18.6%
마케팅 등 고객 커뮤니케이션을 통해 수집	12.4%
내방객 등 오프라인을 통한 데이터 수집	3.5%
각종 센서를 동해 자동 데이터 수집	4.0%
자사의 데이터 가공(데이터화)에 의한 데이터 확보	13.3%
기타	0.0%

※ 자료: 과학기술정보통신부 한국데이터산업진흥원(2022)

- 2021년도 데이터 산업 현황 조사의 조사 결과에 따르면, 전체 478개 기업 중 데이터 거래 경험이 있는 기업은 242개 기업으로 50.6%에 불과하였음.
- 또한 데이터의 수집 방법에 대하여는 외부로부터 데이터를 수급받는 방법에 비하여 내부데이터의 이용 비중이 월등히 높게 나타남.
- ‘온라인 회원과 고객이 이용 동의한 데이터’가 24.9%, ‘소셜(SNS) 인터넷 등을 통한 데이터 수집이 18.6%.’, ‘자사의 데이터 가공(데이터화)에 의한 데이터 확보’가 13.3%. ‘마케팅 등 고객 커뮤니케이션을 통해 수집’한 경우가 12.4%로 내부데이터의 활용과 관련된 방법이 상위의 순위를 모두 차지하고 있는 반면에 데이터 거래를 통해 데이터를 확보하는 경우는 5.9%에 그치고 있음.



(데이터 거래 시 애로사항)

※ 자료: 과학기술정보통신부 한국데이터산업진흥원(2022)

< 데이터 거래 구분 별 고려사항 >

구분	유의사항
데이터 품질	• 데이터의 품질은 고려해야 할 중요한 요소이며, 데이터는 정확하고 요구 사항과 관련이 있어야 함 • 품질이 좋지 않은 데이터는 잘못된 정보에 기반한 잘못된 결론과 결정으로 이어질 수 있음.
데이터 소스	• 데이터 소스는 또 다른 필수 요소입니다. 데이터는 정확성과 무결성에 대한 평판이 있는 신뢰할 수 있고 신뢰할 수 있는 소스에서 가져와야 함.
데이터 볼륨	• 데이터 볼륨도 중요한 고려 사항입니다. 필요에 따라 대량 또는 소량의 데이터가 필요할 수 있음. • 데이터 볼륨이 요구 사항을 충족하기에 충분한지 확인하는 것이 중요함.
데이터 유형	• 데이터 유형도 중요합니다. 필요에 따라 정형 또는 비정형 데이터와 같은 특정 형식의 데이터가 필요할 수 있음. • 데이터 유형이 의도한 용도에 적합한지 확인하는 것이 중요함.
데이터 프라이버시 및 보안	• 데이터 프라이버시 및 보안은 데이터를 구매할 때 고려해야 할 중요한 요소임 • 귀하는 데이터를 합법적으로 획득하고 데이터를 보호하기 위해 필요한 모든 개인 정보 보호 및 보안 조치가 마련되어 있는지 확인해야 함.
비용	• 데이터 비용은 또 다른 중요한 요소임. • 데이터 비용과 제공된 데이터의 품질 및 양 측면에서 비용 대비 가치를 제공하는지 여부를 고려해야 함.
데이터 표준	• 데이터 표준을 설정하면 데이터의 일관성과 고품질을 보장하는 데 도움이 될 수 있음. • 표준은 데이터 형식, 메타데이터 및 데이터 사전을 지정하는 데 사용할 수 있으므로 비즈니스가 서로 데이터를 더 쉽게 공유할 수 있음.
데이터 공유 계약	• 데이터 공유 계약은 기업 간의 데이터 공유를 관리하는 데 사용할 수 있음. • 이러한 계약은 데이터 사용, 보안 및 소유권을 포함하여 데이터 공유 약관을 명시할 수 있음 • 블록체인 기술: 블록체인 기술을 사용하여 안전하고 투명한 데이터 거래를 생성할 수 있음. • 블록체인은 트랜잭션을 안전하고 투명하게 기록하는 탈중앙화 원장으로, 데이터가 변경되거나 조작되기 어려움.
데이터 분석 플랫폼	• 데이터 분석 플랫폼을 사용하여 데이터를 분석하고 정보에 입각한 비즈니스 결정을 내리는 데 사용할 수 있는 통찰력을 제공할 수 있음. • 이러한 플랫폼은 기업이 비즈니스 운영 및 고객 경험을 개선하는 데 사용할 수 있는 데이터의 패턴과 추세를 식별하는 데 도움이 될 수 있음.

- 데이터 거래 시 애로사항 즉, 데이터 거래에 소극적인 이유에 대하여는 '쓸만한 양질의 데이터 부족'이 52.1%로 가장 높게 나타났으며, 다음으로는 '구매 데이터 불합리한 가격책정'이 37.0%, '데이터 유통채널 부족'이 36.3%, '데이터 소재 파악 및 검색의 어려움이 31.5%, '개인정보 포함 데이터 유통·활용 시 법적 문제·법률 지원'이 7.5%, '데이터 품질 문제'가 6.8% 등의 순서로 나타남.
- 해당 조사가 민간 시장에서의 데이터 거래 시 애로사항을 조사한 것임을 감안할 때 정부 차원에서 정책적으로 지원할 수 있는 사항으로는 투명하고 합리적 기준에 따른 가격책정 정책 마련, 데이터 거래 확산을 위한 유통채널 확보, 데이터 소재 및 검색을 지원하기 위한 통합 포털서비스 혹은 데이터 수요자와 공급자의 연결 지원(수급 불일치 해소 정책), 데이터 거래 유통에 관한 법적 문제 개선, 데이터 품질 제고 방안 등을 고려할 필요가 있다는 것을 시사함.
- 위에서 살펴본 바에 의하면 데이터 거래의 활성화에 있어서 '쓸만한 양질의 데이터 구축에 대한 요구가 상당히 강하게 드러남.
- 다만 양질의 데이터가 부족하다는 것은 공급 차원에서 축적된 데이터 자체의 질적 수준이 낮다는 문제로 해석할 수도 있지만, 데이터가 판매·유통되기 위한 적절한 수급처('데이터 유통채널 부족', '데이터 소재 파악 및 검색의 어려움')를 확보하지 못한 점이 큰 문제로 지적되고 있는 것으로 보아 데이터를 효과적으로인 거래 유통시키기 위한 제반의 인프라 마련, 그중에서도 데이터거래소, 데이터 거래 플랫폼 등과 같은 유통채널을 마련하는 것이 중요함을 시사한다고 할 수 있음.
- 다만 이러한 데이터 거래 플랫폼의 활성화 쟁점은 민간에서 사적 계약 등을 통해 자율적으로 이루어지는

영역이므로 정부 주체가 과연 어떠한 형태로 어느 수준의 정책적 지원을 강구할 것인지의 문제는 면밀한 검토가 필요한 부분이라 할 것임.

○ 데이터 거래 관련 예상 소유권 문제점

- 데이터의 소유권은 데이터 제공자가 가지고 있다고 판단하며 사용자는 계약한 분석, 연구 목적에 따라 한시적인 이용 권한을 부여받는 것으로 허용하고 있으며, 실무상 권한 여부, 재사용에 대해서는 크게 제재를 하고 있지는 않음
- 데이터 판매자의 입장에서 거래 계약 시 목적에 맞춰 이용하도록 제한 요건을 계약서 내 부기, 그러나 데이터 재판매, 공공/대민 등 외부의 자유로운 접근이 가능한 환경에서의 데이터 오픈의 경우 금지하거나 협의를 통해 굉장히 제한적으로 허용하고 있음
- 수요기업의 데이터 가공은 공급자 입장에서 매우 민감한 문제임, 최소한의 안전장치로 계약/협약서에 데이터 재가공, 재판매 금지 조항을 명시하고 있으나, 실제로 데이터가 목적 외로 사용되고 있는지 확인은 어려움. 데이터 가공 및 재판매는 불허하나, 사용목적에 따라 일부 합의는 가능함.
- 데이터 수요기업은 데이터 가공을 통한 새로운 데이터 상품 개발, 내부 상품개발 등 세밀한 인사이트 확보를 원하기 때문에 데이터 가공에 대해서는 상호 협의를 통해 합의점을 찾을 필요가 있음
- 데이터의 재가공으로 인해 원천데이터의 훼손 등에 대한 이의 제기가 발생하여 원천데이터의 재사용에 있어서 신뢰성이 떨어지는 일이 발생.
- 판매사 입장에서 제한을 두고 싶으나, 계약 문구 협의 중 고객사의 법률검도 후 삭제되는 경우가 많음. 결국 판매사 입장에서는 고객사의 데이터 활용 후 발생하는 모든 문제는 고객사가 책임지도록 방어함
- 데이터의 판매자는 소유권, 지식재산권 및 기타 모든 권리를 가지며, 수요자는 비독점적이고 제한적인 사용 권한을 갖는 것이 일반적임
- 데이터 가공 후에 부가적인 가치가 생기는 경우가 많기 때문에, 파생데이터에 대한 소유권, 원천데이터 생산자와의 수익배분 등 법적으로 논의될 부분이 많음. 계약사항으로 동일한 데이터 재판매는 법적으로 문제가 되지만, 가공 등의 업무 처리를 진행한 뒤 판매. 재사용에 대한 부분은 기업간 논란의 여지가 있는 상황으로 이를 해결해 달라는 민원이 간혹 발생하고 있음
- 플랫폼 사업자를 통해 판매된 데이터 상품에 문제가 있는 경우 이는 플랫폼 사업자에게는 면책을, 데이터 상품 생산자에게는 책임을 부여해야 함.
- 데이터를 제공하는 기업이 법률 적용을 잘못하는 부분이 생겨 문제가 되면, 플랫폼도 일정부분 책임을 져야 하는 문제가 생기므로 이에 대한 해결이 필요해 보임
- 다만 해당 법령에 의해 데이터 자산 보호를 지나치게 강조하거나 데이터 거래만으로 외형적인 부가가치를 높이기 위해 데이디 가격을 높게 책정하는 방향으로 운영되는 경우 오히려 거래와 유동이 제한될 가능성이 높음
- 데이터의 거래 시 제공자의 법적 책임은 정량적으로 밖에 전근할 수 없기 때문에 플랫폼 사업자의 경우에도 객관적 지표들이 모두 충족되는지 여부를 확인하였다면 해당 데이터의 하자에 대해서는 플랫폼 사업자에게 법적 책임을 물을 수 없음
- 플랫폼 사업자가 거래시스템만을 제공하고, 거래자 간 분쟁에 개입하지 않는 경우라면, 분쟁의 결과로 인한 책임을 플랫폼 사업자에게 부담시키는 것은 부당

(2) 법제 연구 분석을 통한 소유권 문제 해결방안

- 공동 책임자라는 인식과 제도가 필요함, 하지만 판매자의 명백한 과실로 인해 발생한 부분에 대해서는 면책에 대한 부분도 일부 고려해야 할 것임
- 따라서 중개 플랫폼 사업자에게는 거래 데이터에 대한 검증 작업 등을 할 수 있는 권한이 부여되는 것이 필요하다고 보임.
- 그렇지 않을 경우 중개 플랫폼 사업자는 거래에 대해 직접적인 책임을지지 않는다는 점을 강조하면서 법적 책임 요소를 회피할 수 밖에 없을 것임
- 데이터 직접 제공 및 정보제공자와 연계 구간의 정보 유출사고 시 법적 책임, 장애발생으로 인해 수요기관 서비스 지연에 대한 사항은 중개 플랫폼 사업자의 책임이 수반되나, 제공된 개인 정보의

법령상 분리보관, 파기 등은 수요기관에 책임을 부과하는 것이 정당.
- 자사는 판매회원이 등록한 상품 및 용역 등에 관한 정보 또는 구매회원과의 거래에 관하여 분쟁이 발생한 경우 그 분쟁에 개입하지 않으며 그 분쟁의 결과로 인한 모든 책임은 판매회원 도는 구매회원이 부담하도록 함. 또한, 이와 관련하여 제3자에게 손해를 배상하거나 기타 비용을 제출한 경우 귀책사유 있는 자(판매회원 또는 구매회원)에게 구상권을 행사할 수 있도록 약관으로 규정해야함
- 자사는 「전자상거래 등에서의 소비자보호에 관한 법률」 제20조 제2항에 의거하여 판매회원의 정보를 열람할 수 있는 방법을 구매회원에게 제공할 수 있으며, 판매회원은 당해 정보를 기재하지 아니하거나, 허위로 기재함으로써 발생하는 모든 책임을 부담한다는 내용을 약관에 명시하고 있음. 데이터 공급/수요 공급자간의 분쟁 발생 시 중개 플랫폼의 법적인 면책 필요. 저작권, 파생데이터, 데이터 유출・노출 등 데이터의 소유권 및 원천 데이터의 신뢰성 등의 사항을 면책사항으로 지정
- 정보주체의 인센티브 역시 제도적으로는 개념을 형성할 수 있지만, 실무적으로는 정보주체가 어떤 데이터에 얼마나 기여를 했는지 추적할 수 있는 기술이 개발되어야 비로소 법에서 규정하는 취지가 달성될 수 있다고 생각됨
- 만일 정보주체에게 인센티브를 줄 수 있는 제도적, 기술적 환경이 갖춰진다면 데이터에 대한 저작권 같은 개념이 될 것이라고 생각됨, 이에 따라 정보주체들은 데이터 생성 및 이에 대한 유통에 있어 적극적인 모습을 보일것이고 데이터 생산량은 폭증할 것임
- 정당한 대가를 지불하고 데이터를 활용하며 그 대가가 유통채널에 고루 분배되는 구조는 시장의 투명성을 항상시켜 데이터 거래에 큰 기여를 할 것임
- 데이터의 흐름을 관리할 수 있는 Map이 필요하며. 데이터의 적절한 판매, 사용에 대한 Flow 관리를 통하여 생성자에 대한 보호가 필요하다고 보임. 한편 데이터산업법 제23조는 '데이터거래사 양성 지원'에 대하여 정하면서, 데이터거래 시 등록제도를 규정하고 있음.
- 정보주체의 인센티브 보장 방안, 데이터 자의 권한 여부, 플랫폼 기업의 책임 범위 등과 같이 그간 법리적인 쟁점으로서 중요하게 다루어진 주제에서 기술적인 조치와 현실적인 고려 등이 함께 수반되어야만 해당 제도의 진정한 취지를 달성할 수 있을 것이라는 점도 확인되었다. 따라서 데이터의 거래 및 유통의 활성화에 있어서 데이터 거래 실무의 사례와 특징들을 고려하여 적절한 대응방향을 설계하여야 할 것임.
- 하자 담보 책임보다는 분야 책임을 결정할 수 있는 기준이 마련되어야 할 것이며, 이를 위해 데이터의 표준화가 무엇보다 선행되어야 함

○ 데이터의 원천성 보장
- 데이터의 경우 무한대로 복제가 가능하다는 점을 고려하여 데이터 원천성, 즉 해당 데이터가 데이터 제공자에 의해 생성된 것이며 제 3사의 데이터가 아님을 담보할 수 있는 규정이 필요
- 계약의 해제 혹은 해지의 경우, 데이터의 반환 방식에 대한 약관 명시 방안 검도 필요
- 데이터의 유용성. 결합성, 경험성 등의 특징은 판매자가 아닌 구매자의 주관적 요소에 의존하므로 이를 계약에서 명확하게 규정하고 판매자에게 법적 책임을 지을 경우 이를 고려해야 함
- 플랫폼이 데이터의 단순 중개만을 할 경우 전자상거래 소비자보호법 등의 intermediary에 대한 법적 책임 규정(통신판매중개업자)을 적용할 수 있으며, 원칙적으로 거래당사자로서 직접 책임은 부정되어야 하지만 플랫폼이 동신판매업자의 지위도 겸할 경우 관련 규정이 적용되어야 하고 데이터 양도인으로서 민법상 담보적임도 직접 적용되어야 함. 데이터 플랫폼 중개사업자가 거래에서 차지하는 역할과 책임이 중요하기 때문에 거래의 안전과 활성화들 고려해 데이터 플랫폼이 일정 수준 데이터를 확인할 수 있게 하고, 그에 대한 책임을 지게 해야 함
- 플랫폼 사업자는 저작권자 또는 개인정보 주체 등 적법한 권리자의 요구가 있는 경우에는 해당 데이터 상품에 관한 정보를 삭제하거나 수정할 수 있도록 해야 함
- 데이터 거래 플랫폼은 거래를 중개하는 자에 불과하므로 공급자에 의해 제공되는 데이터가 개인정보 및 지식자산권을 침해하여 법적 분쟁-이 발생하거나 거래소가 제공하는 품질 평가 및 가치 명가로 인해 분이 발생하는 경우 일정한 면책 기준 마련 필요(저작권법(제102조 제103조), 정보통신망법(제44조의2) 참조)

바. 타 산업 데이터 거래소와의 차별성 도출

(1) 타 산업 데이터 거래소의 특징 및 문제점

○ KDX(한국데이터거래소)

- KDX는 데이터 보안 및 개인정보 보호에 중점을 둔 플랫폼으로 데이터 암호화, 액세스 제어 및 모니터링과 같은 다양한 보안 조치를 사용하여 데이터가 무단 액세스 및 오용으로부터 보호되도록 하는 장점이 있음
- 개인정보 보호법과 같은 엄격한 데이터 보호 규정을 준수하여 사용자의 개인 데이터가 법적 요구사항에 따라 처리되도록 함
- 다만, KDX는 통신판매중개자이며, 통신판매 당사자가 아님. 따라서 KDX는 상품·거래정보 및 거래에 대하여 별도의 책임을 지지 않음
- 스마트농업 데이터 거래소가 KDX처럼 민간화로 운용된다면 데이터 제공자에게 데이터에 의한 재배 결과까지 책임지게 할 수 없기 때문에 데이터 거래 사항의 분쟁이 다수 발생할 수 있을 것임

○ IBM Watson Health

- 의료 분야에서 사용되는 빅 데이터 플랫폼 중 하나는 인공 지능 및 기계 학습 알고리즘을 사용하여 전자 건강 기록, 의료 영상 및 유전체학에서 대량의 환자 데이터를 분석하는 IBM Watson Health이 있음
- 이 플랫폼은 연구자와 의료 제공자가 새로운 치료 옵션과 개인화된 의료 솔루션을 식별하는 데 도움을 제공함
- 의료 분야 빅데이터 거래소의 잠재적인 약점 중 하나는 분석 중인 데이터의 정확성과 신뢰성이 치명적으로 작용함
- 거래된 데이터는 불완전하거나 부정확한 환자 데이터에 대한 우려가 있을 수 있으며, 이는 진단 또는 치료 오류로 이어질 수 있음
- 또한 환자 정보의 개인 정보 보호 및 보안에 대한 우려가 있을 수 있으며, 이는 위반 또는 무단 액세스로 이어질 수 있음.

○ 금융 데이터 거래소

- 한국 금융 데이터 거래소의 주요 장점 중 하나는 금융 거래의 효율성과 속도를 향상시키는 데 기여한다는 점이 있음
- 금융기관이 실시간으로 데이터를 공유할 수 있도록 함으로써 거래 시간을 단축하고 금융 데이터의 정확성을 높일 수 있음
- 그러나 한국의 금융 데이터 거래소는 시스템이 사이버 공격이나 데이터 침해로부터 보호할 만큼 충분히 안전하지 않을 수 있음.
- 특히 금융기관 직원이 민감한 금융정보를 오용하거나 도용할 수 있는 내부자 공격의 위험이 우려됨.
- 한국의 금융 데이터 교환이 기술적 결함이나 오류에 취약할 수 있다는 것이 있고, 시스템에 다운타임이나 기타 기술적 문제가 발생하면 금융 거래가 지연되거나 오류가 발생하여 기업과 소비자에게 심각한 결과를 초래할 수 있음.
- 마지막으로, 한국의 금융 데이터 교환이 자금 세탁이나 테러 자금 조달과 같은 불법 활동에 악용될 가능성에 대한 우려가 있음.
- 이러한 유형의 활동을 방지하기 위해 금융 기관은 거래를 모니터링하고 관련 규정 및 법률을 준수하는지 확인하는 데 주의를 기울여야 함.

○ 국가 교통 데이터 오픈마켓

- 국가교통데이터 오픈마켓은 운송회사, 정부기관, 연구원 등 다양한 이해관계자들 간에 교통데이터를 공유하고 교환할 수 있는 시스템임.
- 이 시스템은 교통 상황, 차량 위치 및 기타 중요한 요소에 대한 실시간 데이터에 대한 액세스를 제공하여 운송 네트워크의 효율성과 안전성을 향상하도록 설계되었음.
- 국가 교통데이터 오픈마켓의 주요 특징 중 하나는 오픈데이터를 강조한다는 점이며, 이는 연구원, 스타트업 및 기타 조직을 포함한 광범위한 사용자가 교통 데이터를 사용할 수 있음을 의미함.
- 이는 교통 네트워크를 개선할 수 있는 새로운 기술 및 서비스의 혁신과 개발을 촉진하는 데 도움이 될 수 있음.
- 국가교통데이터 오픈마켓의 또 다른 특징은 서로 다른 이해관계자들 간의 협업에 초점을 맞추고 있다는 점이며, 운송 회사, 정부 기관 및 연구원이 협력하도록 장려함으로써 시스템은 운송 네트워크의 조정 및 효율성을 개선하는 데 도움이 될 수 있음.
- 하지만 국가교통데이터 오픈마켓에의 우려 사항은 개인 정보 침해 또는 민감한 교통 데이터에 대한 무단 액세스의 위험이 존재함.
- 또 다른 우려는 데이터 과부하의 가능성이며, 사용 가능한 교통 데이터가 너무 많기 때문에 적시에 효과적인 방식으로 모든 정보를 처리하고 분석하는 것이 어려울 수 있음.
- 이로 인해 운송 네트워크의 효율성과 안전에 영향을 미칠 수 있는 지연, 오류 또는 기타 문제가 발생할 수 있음.
- 마지막으로 국가교통데이터 오픈마켓을 통해 공유되고 있는 데이터의 정확성과 신뢰성에 대한 우려가 있을 수 있으며, 데이터가 불완전하거나 정확하지 않으면 운송 계획 및 의사결정에 오류가 발생할 수 있음.

○ 전체적인 데이터 거래소의 주요 문제점 도출

- 데이터 거래소가 어떠한 기능을 수행하고 어떠한 의무와 안전조치를 취해야 하는지가 불확실함
- 데이터산업법에 데이터 거래소의 요건·의무 등이 규정되어 있지 않기 때문에 온라인 쇼핑몰에 적용되는 일반규정인 「전자상거래 등에서의 소비자보호에 관한 법률」상 통신판매중개자의 의무와 책임, 「전기통신사업법」상 부가통신사업자 의무만 적용되므로, 데이터 거래소 이용자의 보호, 거래되는 데이터의 보호에 어려움이 발생할 수 있음.
- 정부가 데이터 거래소 운영에 많은 예산을 투입하고 있는데, 이는 시장의 건전한 성장을 왜곡시킬 우려도 큼.
- 정부가 예산 사업으로 기존에 없었던 새로운 데이터 거래 시장을 형성하는 것은 의미가 있지만, 거래소가 자칫 시장의 수요보다는 정부의 지원자격을 충족하는 방향으로 운영될 우려가 큼.
- 거래소가 열려 있고 진열된 상품도 많지만 쓸만한 양질의 데이터가 부족한 외화내빈을 피하기 어려움.
- 구매자 입장에서는 데이터 거래소가 많고 상품이 다양해서 필요한 데이터의 소재를 파악하고 찾아내기 어려운 문제도 있음.
- 정부는 통합 데이터 지도 서비스를 제공하고 있지만, 16개 빅데이터 플랫폼만 대상으로 하고, 다른 공공・민간 거래소는 포괄하지 못해서 완성된 지도로 보기는 어려움.
- 통합지도 안에서 상품 목록을 일괄 검색할 수 있지만 데이터 상품에 대한 설명이나 활용 방식, 연계가능한 데이터와 같은 필요한 정보가 제시되어 있지 않아서 현황판 이상의 활용을 기대하기 어려움.
- 데이터 거래에 관한 정책 거버넌스의 분절적 운영도 중요한 문제이며, 16대 빅데이터 플랫폼과 데이터 스토어는 데이터산업법 제6조의 국가데이터정책위원회 심의 사항임.
- 공공데이터포털은 「공공데이터의 제공 및 이용 활성화에 관한 법률」 제5조의 공공데이터전략위원회 소관임.
- 데이터의 생성 원천별로 소관 기관이 분리되어 있지만, 이를 종합적으로 조정하는 장치가 없기 때문에 소관부처의 경계를 넘나드는 데이터 거래와 융합을 효과적으로 관리하기 어려운 문제가 발생함.

(2) 스마트 농업 데이터 거래소의 차별성 도출

○ 설문조사 진행 후 도출된 예상 공급자 의견

- 2단계 스마트팜(지능화, 생육모델, 생체계측 등)에서 생산성 향상을 위한 데이터 활용은 필수사항으로 사료됨.
- 반면, ‘스마트 농업분야’의 데이터 거래는 실수요자의 적극적인 니즈가 부족하고, 또 일반 산업 분야처럼 원천 데이터의 품질 표준화가 취약하여 데이터 활용의 전문성이 부족한 농가에서 적용하기가 쉽지 않을 것임
- 자사에서 개발한 시스템이 탑재된 농장에서 수집된 데이터와 특정 재배환경에서의 데이터 수집에 대해서는 정확도가 높을 수 있지만, 원천데이터를 타사 시스템에서 그대로 적용하기에는 어려울 수 있음
- AWS(Amazon Web Service)에 쌓인 데이터를 타 농가나 기업체에서 활용하기에는 후속처리가 불가피할 것으로 사료됨
- 현재까지 데이터 시장은 데이터 사용자 요구사항이 전혀 반영되지 못한 공급자 중심의 일방적인 서비스 제공으로 의미 없는 시장이 형성되어 있으며, 그런 시장에서 생산된 데이터는 확산 측면에서 명확한 한계가 존재할 것임
- 실효성 수익모델 발굴을 위해서는 서비스 대상 및 영역을 명확히 한 후 그에 적합하게 가공된 데이터나 데이터를 적용하여 도출된 AI 학습모델을 유통하는 방식으로 개선할 필요가 있음

○ 조사 결과에 따른 데이터 거래소 차별성 도출

- 스마트 농업 데이터 거래소는 관계기관과 민간기업의 협업을 위하여 데이터에 대한 신뢰성을 강조하고 공급자입장에서 데이터 품질검증을 담당하고, 공급자에게 재배결과에 대한 책임부담을 없앨 수 있는 데이터 거래소의 약관 규정을 제공해야함
- 스마트 농업 데이터 거래소 또한 데이터 거래 시 발생되는 문제 중 하나로 작물 재배 결과물에 대한 책임을 배상할 수 없는 문제가 존재하나, 데이터의 정확성과 신뢰성에 대해서는 크게 지장이 가지 않을것으로 사료됨
- 따라서 스마트 농업 데이터 거래소에서는 재배 결과물에 대한 책임 배상 문제에 대해 약관으로 규정하고, 데이터의 정확성과 신뢰성은 데이터 공급자 측에서 책임을 맡는 것으로 운영한다면 이러한 문제는 해소될 수 있을것으로 판단됨
- 스마트 농업 데이터 거래소는 공공기관 주관하에 운영될 예정이며, 사이버 공격이나 데이터 침해로부터 방지하기 위해 공유된 데이터의 동기화 여부를 수시로 점검하고, 제공되는 데이터의 파일 형태, 전송방식, 보안 등에 대하여 거래소의 정책을 마련할 계획임
- 또한 개인정보 또는 영업비밀이 포함된 데이터 및 민감한 데이터로 분류되는 항목은 현행법 체계를 준수하여 전문가 의견을 통해 개방 여부를 결정할 계획임
- 스마트 농업 데이터 거래소 운영 시 발생될 수 있는 예상 문제점은 데이터 과부하문제가 발생할 수 있으나, 시스템 부하를 최소화 할 수 있는 메타데이터를 구축하여 문제를 해결할 수 있음
- 생장, 생육정보 분석을 통해 생산품종 결정, 출하시기 결정 등 농민 의사결정지원용 데이터셋 판매 및 연계 서비스 판매 추진
- 온실 내 환경제어 및 토양(배지)관리 서비스 등 작물 생산에 최적화된 대기환경(온습도, 광량 등) 조절, 관개 시스템 및 양액 제어모델, 에너지 비용 데이터셋 판매 및 연계 서비스 판매 추진 예정
- 매출 개선 및 비용 절감 가능 관련 데이터를 공공기관 클라우드의 인공지능 모델에 학습시켜 도출된 스마트팜 인공지능 모델을 스마트 농업 데이터 거래소에 등록 판매하여 플랫폼 자체 수익을 창출
- 인공지능 농업 상담 챗봇 서비스를 개발 및 운영하고, 부분 유료화를 통해 추가 수익모델을 발굴할 계획
- 실질적으로 다양한 현장의 애로사항을 해결하기 위한 서비스모델 개발 및 확산·보급에 많은 부분을 할당할 예정

5. 국내외 오픈 데이터 활용 사례

가. 국외 오픈 데이터 현황

○ 해외에서는 미국과 영국이 선두로 나서 OGD 포털을 구축함.
○ 양국은 각각 OGD 포털을 개설해 운영함, 이후 전 세계적으로 잇따라 연방 또는 지방 정부 차원의 오픈 데이터 이니셔티브 및 OGD 포털구축 움직임이 활발히 전개되어 옴.

(1) 미국

○ 미국은 2009년 5월 ‘오픈 거버넌스 이니셔티브(Open Governance Initiative)’에 의거해 정부가 보유하고 있는 대량의 데이터를 제공하는 오픈 데이터 포털 ‘DATA.GOV’를 개설함.
○ ‘DATA.GOV’는 각 정부 기관의 데이터를 카탈로그처럼 보기 쉽게 분류해 제공하고 있으며, 각종 필터링 옵션을 통해 원하는 데이터를 신속하게 검색 및 다운로드할 수 있도록 구성함.
○ 공개된 데이터 세트 수는 2012년 4만 2,000여 개에서 꾸준히 늘어나 2020년 22만 5,000여 개까지 증가함.

(2) 영국

○ 영국은 2010년 1월 공공 데이터 개방 포털 ‘data.gov.uk’를 개설함.
○ 영국의 공공데이터 포털은 2017년 ‘정부전환전략: 더 나은 데이터 활용(Government)
○ Transformation Strategy: better use of data)’의 일환으로 검색 편의성 개선을 위한 개편을 추진함.
○ 이후 2018년 3월에 ‘오픈 데이터 찾기(Find OpenData)’라는 명칭으로 개편되었는데, 그 결과 이전보다 쉽고 빠르게 원하는 데이터를 검색 및 다운로드할 수 있게 되었으며, 관심 있는 데이터에 대한 업데이트 알림 기능도 추가 됨.

(3) 호주

○ 호주는 2013년 중앙정부 차원의 오픈 데이터 포털 ‘data,gov.au’를 개설하였음.
○ 본래 ‘data.gov.au’는 호주 총리내각부에서 운영을 전담하고 있었으나 현재는 디지털 혁신국(Digital Transformation Agency)으로 이전되어 운영 중임.
○ 데이터 활용을 위한 오픈소스 플랫폼 ‘MAGDA(Making Australian GovernmentData Available)23)’를 제공해 이용자들의 자유로운 애플리케이션 개발을 지원함.
○ 연방정부, 주정부, 정부 산하기관 등에서 생성된 공공 데이터 이외에 공공 연구 데이터나 민간 조직에서 생성된 공공 분야 관련 데이터도 포함하고 있으며, 약 3만 개 이상의 데이터 세트를 제공하고 있음. 특히 위도나 경도와 같은 공간 정보를 포함한 데이터세트를 지도상에 표시하는 시각화 도구 ‘내셔널맵(NationalMap)’ 기능을 자체적으로 제공하고 있는 것이 특징임.

(4) 일본

○ 일본은 2014년 10월 1일 오픈 데이터 카탈로그 사이트인 ‘data.go.jp’의 운영을 개시함.
○ 내각관방 정보통신기술 통합전략실의 제안으로 개설되어 총무성에서 운영 중임.
○ 일본의 오픈 데이터 포털에서는 공공 데이터를 분야별로 검색 가능할 뿐만 아니라 각 부처별로도 검색할 수 있으며, 데이터 제공 관련 기관의 링크 정보와 데이터 활용 사례도 체계적으로 제공하는 특징을 가짐.

(5) 캐나다

○ 캐나다 정부는 2011년 3월 1세대 오픈 데이터 포털인 ‘data.gc.ca’를 시범 운영한 후 2013년에 현재의 오픈 데이터 포털인 ‘open.canada.ca’를 개설함.
○ ‘open.canada.ca’ 사이트에서는 크게 오픈 데이터 검색 캐나다 정부에서 제공하는 정책, 법률, 재무, 규제 등 모든 종류의 정보를 검색하는 ‘오픈 정보(OpenInformation)’ 정부 정책 및 프로그램 개발에 참여할 수 있는 모든 방법을 한눈에 확인할 수 있는 ‘오픈 다이얼로그(Open Dialogue)’ 등의 기능을 제공함.
○ 캐나다 정부 산하 기관들은 2014년 10월에 발효된 오픈 데이터 지침(Directiveon Open Government)에 따라 모든 공공데이터에 대한 목록(Inventory)을 작성해야 하며, ‘open.canada.ca’에서는 해당 목록을 확인할 수 있을 뿐만 아니라 목록을 통한 데이터 접근이 가능함.
○ ‘API 스토어(API Store)’를 마련하고 25가지의 API를 제공함으로써 공공데이터의 종류에 따라 필요한 API를 선택해 활용할 수 있도록 지원하고 있음.

< OECE 회원국 및 파트너국의 OGD 포털 >

국가	포털	국가	포털
호주	data.gov.au	오스트리아	data.gv.at
벨기에	data.gov.be	캐나다	open.canada.ca
칠레	datos.gob.cl	체코	data.gov.cz
덴마크	www.digitaliser.dk/catalogues	에스토니아	opendata.rilk.ee
핀란드	avoindata.fi/fi	프랑스	data.gouv.fr
독일	govdata.de	그리스	data.gov.gr
아일랜드	data.gov.ie	이스라엘	data.gov.il
이탈리아	dati.gov.it	일본	data.go.jp
한국	data.go.kr	라트비아	data.gov.lv
멕시코	datos.gob.mx	네덜란드	data.overheid.nl
뉴질랜드	data.govt.nz	노르웨이	data.norge.no
폴란드	danepubliczne.gov.pl	포르투칼	dados.gov.pt
슬로바키아	data.gov.sk	슬로베니아	data.gov.si
스페인	datos.gob.es	스웬덴	oppandata.se
스위스	opendata.swiss	영국	data.gov.uk
미국	data.gov	콜롬비아	data.gov.co
리투아니아	opendata.gov.it	페루	datosabiertos.gov.pe

나. 국제기구의 오픈 데이터 포털

(1) OECD 데이터(OECD Data)

○ OECD 데이터(data.oecd.org)에서는 경제협력개발기구(OECD: Organizationfor Economic Cooperation and Development)에서 보유하고 있는 모든 통계데이터와 분석 자료 및 보고서를 제공함.
○ 7,200여 개 이상의 데이터 세트를 데이터 종류별, 국가별, 분야별로 분류해 놓고 있으며, 통계데이터의 경우에는 Excel, CSV, XML 등의 다양한 형식으로 저장이 가능함.
○ 주요 데이터는 바로 사용 가능한 테이블 형태뿐만 아니라 그래프, 매핑 등을 통한 시각화된 자료로도 제공함.

(2) EU 오픈 데이터 포털

○ EU(European Union)의 오픈 데이터 포털은 EU 기관 및 기타 기관의 오픈 데이터를 통합적으로 검색하고 접근할 수 있는 단일 액세스 포인트임.
○ 이 포털은 2011년 12월에 발표된 '위원회 문서 재사용에 관한 유럽위원회의 결정(European Commission Decision 2011/833/EU on the reuse of Commission documents)' 에 따라 2012년에 개설됨.
○ 유럽 내 연합체와 정부 기관 데이터에 무료로 쉽게 접근함으로써 데이터를 혁신적으로 사용하고 경제적 잠재력을 발휘할 수 있도록 지원하고 나아가 해당 기관들의 책임성, 개방성. 투명성 도모를 목표로 함.
○ 포털에서는 지리적/지정학적/재무 데이터, 통계, 선거결과, 법적 행위, 범죄/건강/환경/운송/과학연구에 대한 데이터 등을 제공하며, 다양한 디지털 형식을 활용해 상업적 또는 비상업적 목적으로 자유롭게 사용 및 재사용할 수 있음.
○ 그리고 EU 오픈 데이터 검색 및 접근을 위한 표준화된 카탈로그, 데이터 재사용 앱 및 웹 도구 사례, 개발자들의 데이터 활용을 지원하기 위한 API 지원 등을 제공함.

(3) 유럽 데이터 포털

○ 유럽 데이터 포털은 2015년, 유럽위원회(EC)가 프랑스의 기술 컨설팅 기업인 캡케미니 인벤트(Capgemini Invent)가 이끄는 컨소시엄의 지원을 받아 개발되었는데, 이 포털은 28개 EU 회원국 및 4개 EFTA(European Free TradeAssociation) 국가의 공공데이터 포털에 있는 공공부문 정보를 수집해 제공하기 위한 목적으로 개설된 것임. 유럽의 모든 정부 기관이 공개한 데이터에 접속할 수 있는 단일 포인트를 제공하며, 24개 EU 공식 언어로 접속할 수 있음.
○ 주제/키워드/파일형식/데이터 라이선스 주체 등에 따른 데이터 검색, SPARQL 쿼리를 통한 검색, 카탈로그 검색 등 다양한 데이터 검색 및 접근 방식을 제공함.
○ 나아가 오픈 데이터 관련 분석 보고서를 비롯해 오픈 데이터 재사용 사례, 오픈 데이터 뉴스, 오픈 데이터 교육 자료, 오픈 데이터 뉴스 등 다양한 콘텐츠를 자체적으로 제작해 제공하고 있음

(4) UN 데이터 카탈로그

○ UN 데이터 카탈로그는 UN과 산하 조직에서 보유하고 있는 오픈 데이터를 검색하고 접근할 수 있도록 지원하는 플랫폼임.
○ UN 오픈 데이터에 대한 접근성과 가용성을 향상시키기 위한 목적으로 2016년 3월에 구축됨. UN 산하 조직, 지역, 태그, 시기 등의 데이터세트 검색 옵션을 제공하며, 데이터에 있는 링크와 API 및 시각화 도구 통해 해당 데이터에 바로 접근할 수 있음.
○ UN 산하 국제기구 및 조직의 오픈 데이터 포털로 직접 접속할 수 있는 링크 리스트를 제공하는데, 해당 리스트들을 50여 개 주제별로 검색할 수 있는 필터 기능도 지원함.

다. 국내 주요 오픈 데이터 플랫폼과 활용 사례

(1) 공공데이터포털

○ 공공데이터포털은 행정안전부와 한국정보화진흥원(이하 NIA)에서 운영하고 있음.
○ 공공기관이 생성하거나 보유하고 있는 공공데이터를 통합적으로 공유, 개방하기 위한 공공데이터 플랫폼임.
○ 행정안전부장관은 「공공데이터의 제공 및 이용활성화에 관한 법률」(이하 「공공데이터법」) 제21조에 따라 공공데이터의 효율적 제공을 위하여 통합제공시스템을 구축하여야 하는데 이 시스템이 공공데이터포털임.
○ 「공공데이터법」은 공공기관이 보유·관리하는 데이터의 국민 이용권을 보장하고 공공데이터의 민간 활용을 활성화하기 위한 법률로서 「공공데이터법」에서의 공공데이터 공개는 국가기관, 지방자치단체 및 공공기관이 보유·관리하는 데이터(DB, 전자화된 파일)를 민간에서 활용할 수 있도록 다양한 방식으로 제공함.
○ 공공데이터는 “데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리된 자료 또는 정보”를 의미함.(「공공데이터의 제공 및 이용 활성화에관한 법률(이하 공공데이터법)」 제2조)
○ 공공데이터포털의 주요 공공데이터 종류는 다음과 같음.

- 공공행정: 도로명주소, 공공시설물, 공용주차장, 도시공원, 공중화장실, 등산로 등
- 국토교통: 건축정보, 부동산정보, 공간융합정보, 자동차종합정보, 교통사고정보 등
- 산업고용, 사회복지, 식품건강, 환경/기상, 보건의료, 문화관광, 농축수산 등
- 제외: 단순행정업무, 업무추진비 등

○ 공공데이터포털에서는 공공데이터를 공개하는 기능 이외에도 필요한 공공데이터를 신청할 수 있으며, 주요 데이터를 시각화하거나 공공데이터를 활용한 사례도 공유하고 있음.

< 공공데이터 주요 활용 사례 >

서비스	활용데이터	서비스 내용
직방-(주)직방	실거래가격지수 통계정보(한국감정원), 아파트 매매 실거래 및 전월세 자료, 공공주택관리비정보, 건축물 대장정보(국토부) 등	원/투룸 · 오피스텔 · 아파트 매물 및 시세정보, 평형별 아파트 실거래가 및 변동추이정보 등 온라인 부동산 정보 서비스
김기사-카카오	도로이정표 정보(한국도로공사), 새주소정보(서울시), 가격정보(한국석유공사)	실시간 교통정보, 블랙박스, 길찾기 등을 종합적으로 제공
오픈라이더-(주)쿠핏	자전거 사고다발지역정보(도로교통공단), 생활기상지수, 동네예보정보조회(기상청), 대기오염정보(한국환경공단) 등	자전거 길찾기 내비게이션, 심박계, IoT 기기 연동 라이더 상태 체크 등
미세미세-(주)MiseMse	대기오염정보(한국환경공단)	실시간(초)미세먼지에 대한 지역별 현황, 일별/주별 예측서비스 등
스타포트키오스크-(주)아이오로라	국문 관광정보(한국관광공사), 공 연전시정보조회(한국문화정보원)	외국 관광객이 국내 여행 시 관광지정보를 포토티켓 속에 제공
스마트팔레트-(주)구니스	전통미술작품(국립중앙박물관)	아동 미술 교육용 디바이스, 전통 미술 작품 도안 채색, 앱 ‘키즈페인터’ 연동
후즈굿-지속가능발전(주)	화학물질 배출 및 이동량(국립환경과학원), 환경정보(환경부) 등	기업의 비재무 정보 분석을 통해 투자 시 판단기준을 제시할 수 있는 분석보고서 제공
크레딧잡-(주크레딧데이터)	사업장정보조회서비스(국민연금공단)	전국 사업장 고용정보 41만 개 기업, 980만 명에 대한 평균 월급 및 고용 정보 제공

(2) AI 허브

○ 한국정보화진흥원에서는 인공지능 기술 및 제품·서비스 개발에 필요한 AI 인프라(AI 데이터, AI SW API, 컴퓨팅 자원)를 지원하기 위한 AI 통합 플랫폼인 ‘AI 허브’를 운영하고 있음.

○ AI 허브는 2017년 인공지능 학습용 데이터의 구축·확산하는 사업의 일환으로 시작되었으며 AI 기술 및 서비스 개발에 필수적인 AI 데이터, 소프트웨어, 컴퓨팅자원, 소재정보 등을 원스톱으로 제공하는 것을 목적으로 하며 인공지능 역량 강화를 위해 데이터를 중심으로 필요한 자원들을 직접 지원하는 플랫폼임.

○ AI 허브에는 AI를 개발 및 활용하기 위한 인프라 서비스 4종(AI 데이터, AI 소프트웨어, AI 컴퓨팅, AI 이지빌더)과 AI 활성화를 위한 서비스 4종(AI 혁신체험, AI 리더보드, AI 커뮤니티, AI 경진대회) 등 총 8가지 서비스를 제공하고 있음.

○ AI 허브에서는 범용성이 높은 자연어 데이터 8종, 이미지 데이터 16종, 영상 데이터 11종과 함께 전문 분야의 지식베이스 6종을 공개하고 있음.

○ AI 허브에서 공개하고 있는 인공지능 학습용 데이터는 음성인식과 같은 분야에서는 실무적으로도 많이 이용하는 것으로 알려져 있음.

○ 네이버는 AI 허브의 한국어 대화 · 음성 데이터를 활용하여 인공지능 전화응대 서비스인 ‘AI Call’을 고도화했음.

○ 삼성전자도 한국어 음성 데이터를 활용하여 ‘빅스비’의 음성 인식률을 향상시키는 등 범용적 인공지능 학습데이터로서의 역할을 하고 있음.

라. 국외 민간 및 학계의 데이터 개방/공유 현황

○ 최근 몇 년 사이 Uber, AirBnB 등 일부 데이터 중심 기업이 자사의 데이터 일부를 오픈 데이터로 공표하고 있음.
○ 기업이 데이터를 공유하는 데는 데이터 공유의 홍보 가치에서부터 크라우드소싱을 통해 자사 데이터를 개선할 수 있는 기회 탐색 등 여러 이유가 있음.
○ 대부분의 민간 기업은 경쟁 기업들이 자사의 데이터를 공유하지 않는 한 그들의 데이터를 공유하는 것을 꺼려왔으나, 상대적으로 새로운 협력적 접근방식은 서로 다른 기업이 서로 다른 데이터로부터 이익을 얻으면서 동시에 공익에도 기여하기 위한 협력을 더 용이하게 하는 측면이 있음.
○ 데이터 협력자는 다양한 분야의 행위자들이 데이터를 교환하고 분석적 전문지식을 취합하여 복잡한 공공 문제를 해결하는 데 도움을 주는 민관 협력의 새로운 모델을 제시함.

(1) 우버의 무브먼트 오픈데이터 포털(Uber Movement Open Data Portal)

○ 우버는 2018년 ‘우버 무브먼트(Uber Movement)’라는 프로젝트를 시작으로 도시 교통체증 해결을 위한 방법과 관련된 데이터를 공개하였음.
○ 우버의 데이터서버에 저장되는 기하급수적인 정보는 개인의 위치정보, 신용카드 정보, 이동 정보 등이 포함되어 있으며, 도시의 교통에 관한 통찰력을 갖춘 정보를 제공함.
○ 우버의 데이터는 연구 집단이나 행정당국 등에 우선 제공되고 최종적으로는 일반인에게도 공개되는데, 이 데이터는 사기업의 데이터가 공공 정책에 쓰일 수 있음.
○ 데이터는 요일이나 시간을 지정해 특정 지점 혹은 특정 범위의 데이터를 불러내 다운로드할 수 있도록 했으며, 시계열 차트나 원천 데이터(raw data) 형태 등으로도 다운로드 가능함.
○ 우버의 데이터는 도시의 교통 인프라 발전에 쓰임으로써 장기적으로 우버 사용자의 만족도 향상에 도움이 되는 순환적 관계를 가짐.
○ 우버는 개인정보 이슈에 대해서는 비식별화를 통한 익명처리로 해결하고 있음.
○ 우버는 ‘우버 무브먼트’라는 프로젝트를 개시한 2018년에 ‘셰어드 스트리트(Sharedstreets)’라는 프로젝트를 시작하고 비영리단체 및 정부기관과 함께 모빌리티 데이터 기술 전반에 대해 연구 중임.
○ ‘우버 무브먼트’ 이외에 모빌리티와 관련된 민간 기업의 데이터 공개 사례로, 구글의 자율주행 프로젝트인 웨이모(Waymo)의 오픈 데이터, 동남아 모빌리티 시장에서 강세를 보이는 그랩(Grab)의 오픈 데이터, 중국의 우버라 불리는 디디추싱(Didi Chuxing, 滴滴出行)의 오픈 데이터 등이 존재함.

(2) 구글 리서치 데이터 세트

○ ‘구글 리서치 데이터 세트(Google Research Dataset)’는 광범위한 연구 커뮤니티를 지원하기 위해 구글의 연구조직인 ‘구글 리서치’에서 자체적으로 운영하는 일종의 오픈 데이터 포털임.
○ 주로 컴퓨터 과학 분야의 연구자들이 관심을 가지는 데이터가 주기적으로 공개됨.
○ 이미지, 동영상, 음성, 텍스트 주석(text annotation), 로봇공학(robotics), 기타 등으로 구분되는 다양한 유형을 제공하고 있음.

○ 원하는 데이터 세트를 선택하면 해당 데이터 제공자가 마련한 웹페이지나 깃허브(GitHub)로 연결되어 다운로드할 수 있으며, 데이터 세트를 활용한 챌린지(Challenge)를 개최해 개발자들의 참여를 독려하고 있음.

(3) 캐글

○ 구글의 자회사인 ‘캐글’은 데이터 전문가들로 구성된 온라인 커뮤니티이자 예측 모델 및 분석 대회 플랫폼임.

○ 기업 및 단체에서 데이터와 해결과제를 등록하면, 데이터 과학자들이 이를 해결하는 모델을 개발하고 경쟁함.

○ 2010년 머신러닝 관련 대회 플랫폼으로 시작했으나, 2017년 3월에 구글에 인수된 이후로 데이터 플랫폼, 클라우드 기반 워크벤치(workbench), 인공지능 교육 등의 서비스가 추가되어 개발자들을 위한 종합 플랫폼으로 발전하였음

○ 캐글의 홈페이지는 ‘컴피티션(Competitions)’, ‘데이터 세트(Datasets)’, ‘노트북(Notebooks)’ 등의 메뉴로 구성되어 있다. ‘컴피티션’은 알고리즘 경쟁을 위한 메뉴로서, 기업이나 특정 사용자가 과제를 게시하고 사용자들이 과제에 맞는 최적의 알고리즘이나 코드를 구축하기 위해 경쟁에 참여함.

○ ‘데이터 세트’ 메뉴는 오픈 데이터 세트 플랫폼으로서 커뮤니티 구성원들이 공유한 다양한 종류와 유형의 데이터 세트 검색, 데이터세트 제작 및 공유, 데이터세트 협업, 데이터 프로젝트 진행 등이 가능함.

○ ‘노트북’은 가상환경에서 코드를 작성하고 실행시킬 수 있는 클라우드 기반의 워크벤치로서, 스크립트(Script), R마크다운 스크립트(RMarkdown Scripts), 주피터 노트북(Jupyter Notebook) 등의 유형으로 제공되며, 파이썬(Python)과 R언어를 사용할 수 있음.

○ 현재 캐글에는 약 300만 명 이상이 등록되어 있으며, 약 5만 개의 데이터 세트와 약 40만 개의 노트북 결과물을 제공하고 있음.

(4) 아마존의 RODA

○ 아마존(Amazon)의 RODA(Registry of Open Data on AWS)는 AWS에서 분석을 위해 공개해 놓은 데이터를 검색하고 공유할 수 있는 레지스트리임.

○ RODA는 AWS에서 사용 가능한 유용한 데이터에 대한 접속을 확장하는 것이 목표이며, 연구 또는 교육 목적으로 사용 가능하고, AWS 도구를 사용한 분석에 최적화된 데이터세트 등록을 선호함.

○ 2020년 9월 13일(한국시간) 기준 총 180개의 데이터세트를 제공 중이며, 키워드나 태그로 데이터세트 검색이 가능하고, ‘깃허브’를 통한 데이터 추가도 가능함.

○ 구체적인 명시가 되어 있지 않는 한, 모든데이터 세트는 AWS가 아닌 라이선스에 따라 정부기관, 연구자, 기업 및 개인 등의 제3자가 소유 및 유지 관리하지만, AWS 오픈 데이터팀의 판단에 따라 언제든지 제거할 수 있음.

○ AWS에서 데이터가 공유되면 아마존 E2C(Amazon E2C), 아마존 아테나(Amazon Athena), AWS 람다(AWS Lambda), 아마존 EMR(AmazonEMR) 등 아마존의 컴퓨팅 및 데이터 분석 제품을 이용해 데이터를 분석하고 그 위에 서비스 구축이 가능함.

프로젝트 이름	관리자	데이터 세트 설명
암 게놈 아틀라스	Center for TranslationalData Science at TheUniversity of Chicago	암의 주요 유형 및 하위 유형의 주요 게놈 변화에 대한 포괄적인 다차원 지도를 생성 하기 위한 목적의 데이터세트
TARGET	Center for TranslationalData Science at The University of Chicago	유년기 암의 발병과 진행을 유발하는 유전적 변화 형태를 결정하기 위해 포괄적인 분자 특성을 제공하는 GDC (Genomic DataCommons)의 정량화 데이터 및 GDC 레거시 아카이브의 개방형 데이터
Landsat8	Planet	Landsat 8 위성이 지속적으로 수집한 지구상의 모든 육상 이미지
Common Crawl	Common Crawl	250억 개 이상의 웹 페이지로 구성된 웹 크롤링 데이터 모음
NEXRAD on AWS	US National Oceanic and Atmospheric Administration(NOAA)	차세대 기상 레이더(NEXRAD) 네트워크의 실시간 및 보관 데이터
CBERS on AWS	AMS Kepler	위성 미션이 시작된 이후에 획득한 모든 CBERS-4 MUX, AWFI, PAN5M 및 PAN10M 장면이 포함되어 있으며 매일 새로운 장면으로 업데이트
IRS 990 Filings	-	2013년부터 현재까지 IRS에 제출된 특정 전자 990 양식의 기계 판독 가능 데이터
Terrain Tiles		S3에서 쉽게 사용할 수 있도록 바둑판 모양 지형 높이를 제공하는 글로벌 데이터세트
NOAA Water-Column Sonar Data Archive	NOAA NCEI	NOAA 국립환경정보센터에 보관된 물기둥 (Water-Column) 음파탐지 데이터

마. 주요 국내 민간 주도 데이터 공개와 활용 사례

(1) 네이버 데이터랩

○ 네이버는 2016년부터 자사가 보유한 데이터를 공개하는 ‘데이터랩(datalab.naver.com/)’을 운영하고 있음. 데이터랩의 보유 데이터는 크게 급상승검색어 검색어트렌드, 쇼핑인사이트, 지역통계, 댓글통계로 구분되며, 통합 검색어 통계,쇼핑 트렌드 검색어 추이, 키워드 관심도, 카드 사용 통계 등을 제공함. 한편, 네이버는 이러한 서비스들을 API의 형태로 Naver Developer 사이트(developers.naver.com)를 통해 제공하여 개발자들이 이용할 수 있도록 하고 있음.

○ API는 무료로 운영되나, 필요한 경우 이용가능 시간 및 횟수를 제한할 수 있음을 밝히고 있음.

○ 오픈 API와는 별도로 네이버는 자회사 네이버 클라우드를 통해 Clova나 Papago같은 인공지능 어플리케이션 서비스, 텐서플로우 서버 등의 AI 관련 서비스를 제공하고 있음.

○ AI 기반 서비스는 음성인식, 음성합성, 자동번역 등이 있으며, 입력된 글자 수 또는 녹취록의 길이에 따라 차등 과금되는 가격체계를 유지하고 있음.

(2) SKT 빅데이터 허브

○ SKT는 자사가 보유하고 있는 데이터를 공개하기 위해서 2013년 10월부터 ‘빅데이터 허브(www.bigdatahub.co.kr)’를 운영하고 있음.
○ 이는 민간이 보유한 데이터를 일반에 개방한 첫 국내 사례로 평가됨.
○ 빅데이터 허브에는 휴대폰 통화량을 기반으로 한 업종별 이용현황, T map 사용 정보를 기반으로 한 시군구별 목적지 정보, 기지국 신호정보를 기반으로 한 서울시 유동인구 정보, 로밍 서비스 정보를 기반으로 한 출국 국가별 방문자 특징(성별/연령별) 정보를 월 단위로 제공하고 있음. 또한 업종별 통화량, T 멤버십 이용비율과 같은 정보도 제공하고 있음. 빅데이터 허브에서 제공하고 있는 데이터는 총 1,446건이며, 공개된 데이터는 무료로 이용할 수 있음.

(3) LG CNS

○ LG CNS는 일반인도 쉽게 데이터를 검색하고 분석할 수 있는 오픈 데이터 플랫폼인 ‘오디피아(ODPia, www.odpia.org)’를 운영하고 있음.
○ 오디피아는 SNS,블로그, 커뮤니티와 같은 소셜 미디어에서 정보를 수집, 분석하여 소셜 맵과 소셜인덱스로 제공하고 있는데 2016년 출시 당시에는 오픈 데이터 플랫폼을 지향하면서 소셜 분석, 데이터 통합 등 다양한 서비스 제공을 목표로 하고 있음.

(4) 타사 데이터의 활용 사례

○ 민간 주도로 국내에서 개방된 데이터의 실제 활용사례가 있더라도 찾기는 힘든 실정이임.
○ 다만 정부 보조를 통해 민간 보유 데이터를 개방, 유통하는 국내 ‘빅데이터 플랫폼’ 사업의 사례를 통해 개방된 데이터의 활용 사례를 유추해볼 수 있음.
○ 과기정통부는 2019년부터 공공과 민간이 협업하여 데이터의 생산, 수집, 분석, 유통의 전 과정을 지원하기 위해서 ‘빅데이터 플랫폼 및 센터 구축 사업’을 추진하고 있음.
○ 빅데이터 플랫폼은 빅데이터 플랫폼 및 빅데이터 센터를 중심으로 기업 및 공공기관의 데이터 구축 및 협업을 활성화하는 데 목적을 두고 있으며 금융, 환경, 문화, 교통, 헬스케어, 유통, 통신, 중소기업, 지역경제, 산림 등 총 10개 분야로 구성되어 있음.
○ 빅데이터 플랫폼은 분야별로 수행 기관이 지정되어 분야별 데이터 거래소 운영을 총괄하고 있으며, 참여기관들은 분야별 데이터를 생성하여 플랫폼에 데이터를 등록하며 등록된 데이터는 유무료로 데이터 수요자에게 유통되고 있음,
○ 빅데이터 플랫폼은 분야별 플랫폼 및 센터의 전반적인 정보를 제공하는 ‘빅데이터 플랫폼 통합 데이터 지도’라는 홈페이지(www.bigdata-map.kr)가 구성되어 있고, 각 분야별 플랫폼은 개별적인 홈페이지를 운영하고 있으며 개별 홈페이지가 데이터 유통의 플랫폼 역할을 수행함.
○ 이와 같이 개방, 유통되는 데이터는 다른 영역의 데이터와 결합되어 새로운 활용 사례를 등장시키고 있음.

분야	도메인 주소	수행 기관	센터 참여기관	플랫폼 주요내용
금융	금융 빅데이터 플랫폼 (www.bigdata-finance.kr)	비씨카드	10개(노타, 날슨코리아, 다음소프트, SBCN, 망고플레이트, 해빗팩토리, KT, 한국감정평가사협회, 기웅정보통신, 한국금융솔루션)	금융(대출, 보험, 증권), 비금융(통신, 소셜, 유통, 미디어, 상권) 데이터를 융합하여 소상공인 창업 지도 서비스 및 국민 금융생활 플래너 서비스 등 제공
환경	환경 빅데이터 플랫폼 (www.bigdata-environment.kr)	한국수자원공사	10개(기상산업기술원, 국립생태원, 환경정책평가연구원, 지질자원연구원, GDS컨설팅그룹, 그린에코스, 아이렉스넷, 한국과학기술원, 순천향대학교, 노비코스)	물, 기상기후, 미세먼지, 지질·재해, 생태·자원, 화학물질, 환경SNS 등의 데이터를 융합하여 맞춤형 수질정보 서비스 및 대기질 야외활동 추천 서비스 등 제공
문화	문화 빅데이터 플랫폼 (www.bigdata-culture.kr)	한국문화정보원	11개(국립중앙도서관, 국민체육진흥공단, 문화예술위원회, 야놀자, 청소년활동진흥원, 부산정보산업진흥원, 레드테이블, 레드타이, 원투씨엠, 데이터마케팅코리아, 타엔젬에스)	문화, 숙박, 레저, 음식, 상권, 도서·출판 등의 데이터를 융합하여 문화여가 종합정보 서비스 및 한류 비즈니스 매칭서비스 등 제공
교통	교통 데이터 거래소 (www.bigdata-transportation.kr)	한국교통연구원	10개(울산정보산업진흥원, 포항테크노파크, 아이나비시스템즈, SKT, KT, KST플레이스, 코리아크레딧뷰로, 성남시청, 진주시청, 대전광역시청)	실시간 교통량, 대중교통, 열차, 고속도로, 내비게이션, 블랙박스, 유동인구, 주차 등의 데이터를 융합하여 도로 및 대중교통 개선 서비스 및 스마트 시티 지원 서비스 등 제공
헬스케어	암 빅데이터 플랫폼 (www.bigdata-cancer.kr)	국립 암센터	10개(삼성서울병원, 세브란스병원, 건양대병원, 전북대병원, 대구가톨릭대병원, 서울대학교, 분당서울대병원, 화순전남대병원, 아주대학교, 길의료재단)	10대 암종별 임상데이터를 융합하여 암 진단·치료 의사 결정 및 항암 치료제 연구개발 등 활용
유통	KDX한국데이터 거래소 (www.bigdata-market.kr)	매일방송	10개(나이스디앤알, 다음소프트, 데이블, 로플랫, 빌트온, 식신, 온누리에이치엔씨, 지인플러스, 코리아크레딧뷰로, 우편사업진흥원)	유통상품, 카드결제, 택배송장, 통신, 부동산, 상권, 물류, 맛집, 중고차시세, SNS 등의 데이터를 융합하여 라이프스타일별 선호 외식업종 서비스 및 지역별 온라인 상품 구매정보 서비스 등 제공
통신	KT통신 빅데이터 플랫폼 (www.bigdata-telecom.kr)	KT	14개(비씨카드, 인터넷진흥원, 인천테크노파크, 소상공인연 합회, 경기대학교, 어메이징 푸드 솔루션, 두잉랩, 코난테크놀로지, 오픈메이트, 넥스 트	유통인구, 상권, 카드 소비, 관광, 교통카드 정보, SNS 등의 데이터를 융합하여 상권분석 서비스 및 생활인구 분석 서비스 등 제공

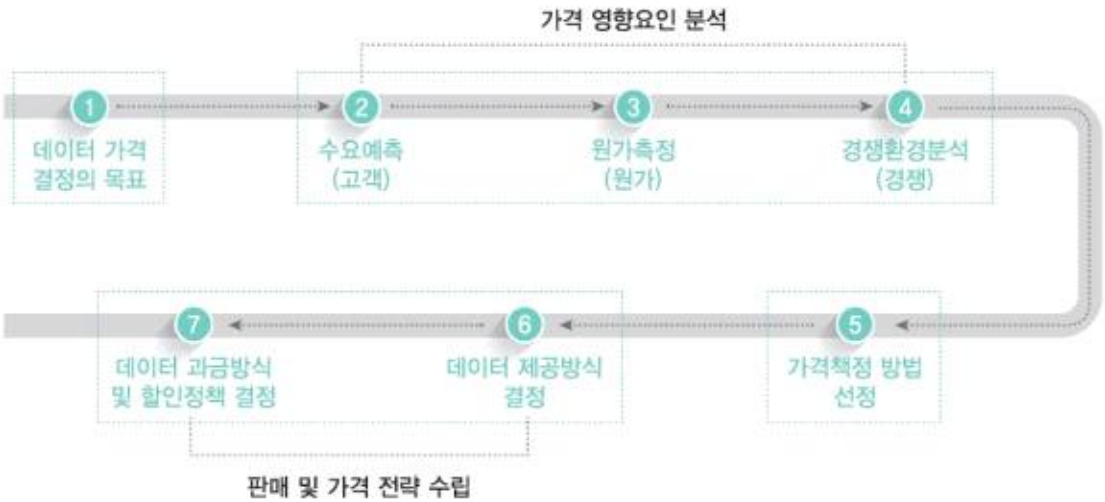
			이지, 네스, 한국스마트그리드사업단, 제로투원파트너스, 더큰나눔엠티엔)	
중소기업	데이터유통포털 (www.bigdata-sme.kr)	더존 비즈온	10개(빅밸류, 한화손해보험, 한국생산성본부, 한국무역정보통신, 와이즈넛, NICE평가정보, 산업기술진흥협회, 인크루트, 녹색기술센터, 선도소프트)	중소기업 회계정보, 부동산, 보험계약, 기업고용·복리후생, SNS등이 데이터를 융합하여 기업 경영정보 분석서비스 및 일자리 수요 예측 서비스 등 제공
지역경제	경기지역 경제포털 (www.bigdata-region.kr)	경기 도청	8개(경기콘텐츠진흥원, 경기일자리재단, 더아이엠씨, 경기신용보증재단, 한국기업데이터, 국토연구원, 한양대학교, 한국생산기술연구원)	지역화폐 결제정보, 기업정보, 일자리, 신용평가, 카드사정보, 경기도 인구·주거·환경 등의 데이터를 융합하여 지역 소비패턴 분석 서비스 및 맞춤형 일자리 매칭 서비스 등 제공
산림	www.bigdata-forest.kr	한국임업진흥원	10개(비글, 아로정보기술, 인포보스, 복지진흥원, 사산아이티, 우림엔알, 삼아항업, 우림인포텍, 마켓링크, 한국한의학연구원)	임업, 등산로·숲길·자전거, 대중교통, 산악기상, 산림재해, 항공영상 등의 데이터를 융합하여 트래킹 서비스 및 산림재해 예측 서비스 등 제공

6. 데이터 가격 평가 방법

가. 데이터 가격 결정 평가

(1) 데이터 가격 결정 절차

○ 데이터의 가치(valuation)는 데이터를 활용해 미래에 기대하는 편익을 현재의 값으로 환산 한 것이며, 가격(price)은 이를 투입된 비용에 의해 화폐 단위, 가액(價額)으로 환산한 것을 말함.
○ 다만, 데이터 가치와 가격은 반드시 비례하는 것이 아니고 가격은 원가, 경쟁, 고객, 환경 등 여러 가지 요인에 따라 지속적으로 변동될 수 있음.
○ 데이터와 데이터 거래에 대한 다양한 정보가 공유되고 여러 판매자가 공정하게 경쟁하는 시장에서 장기적으로 데이터가격은 데이터의 가치를 반영할 것으로 기대할 수 있음
○ 데이터 거래 시장에서 데이터 역시 주식이나 부동산처럼 미래 가치가 높고 가격이 낮을수록 좋은 데이터라고 볼 수 있고, 데이터 기업은 가치가 높은 데이터의 판매 가격을 조정해 기업 이익을 극대화할 수 있음.



(데이터 가격 결정 절차)

○ 위 그림의 절차에 따라 판매 기업은 데이터 가격 결정의 목표에 따라 고객, 원가, 경쟁 등 가격 영향 요인을 분석한 후 가격책정 방법을 수립하고, 판매 및 가격 전략들을 반영하여 최종 데이터 가격을 결정하게 됨.

(2) 데이터 가격 결정의 목표

○ 조직 내 데이터 거래가 필요하다고 판단한 경우, 데이터 상품을 얼마에 판매해서 어떠한 결과를 얻을 것인지 고민해야함.

○ 판매 기업이 데이터 가격을 결정하는 목표는 크게 매출중심적 가격 목표(Sales-Based Objectives), 이윤중심적 가격 목표(Profit-Based Objectives), 현상유지적 가격 목표(Status-Quo Objective) 등으로 구분함

○ 매출 중심적 가격 목표
- 매출 증대나 시장 점유율 확대를 위한 방법이며, 매출 확대를 목표로 가격을 설정하는 경우 대개 저가로 판매량을 늘리는 시장침투가격(penetration pricing)을 선택함.

○ 이윤 중심적 가격 목표
- 기간 내 목표 투자 수익률 달성 등 일정기간 동안의 이윤 극대화를 위한 방법이며, 이를 위해서 초기에 높은 마진을 남기는 초기고가전략(skimming pricing)을 채택함.

○ 현상 유지적 가격 목표
- 이미 시장에서 좋은 위치를 차지하고 있어 더 이상의 변화를 원치 않는 기업이 설정하는 방법으로, 기존 매출을 유지하는 수준으로 가격을 선택합니다.

(3) 가격 영향 요인 분석

○ 원가
- 데이터 상품에서 원가란, 데이터 확보・가공・판매를 위한 시스템 구축비, 운영비 등의 투입 비용을 의미함.
- 다른 제품이나 서비스와 마찬가지로 데이터의 가격을 결정할 때 원가는 중요한 요소로 작용하며, 가격은 원가와 최소이윤의 합을 의미하기 때문에 원가는 판매 가격의 최하한선을 결정하는 기준이 될 수 있음.
- 품질이 좋은 데이터 상품을 낮은 원가에 구축하면 저가 책정이 가능하고, 이에 따라 판매와 이윤이 증가해 경쟁우위를 점할 수 있음.

○ 고객
- 구매자는 외부로부터 데이터 상품을 구입하면 얻게 되는 편익이 크고 빠르다고 판단할 때 데이터 구매를 고려할 것임.
- 데이터에 대한 구매자의 인식과 수요, 구매자의 상황, 구입한 데이터를 통해 구매자가 얻는 편익과 가치, 구매자의 필요성, 구매자의 신뢰성, 데이터의 매력 등이 구매 요인이 될 수 있음.

○ 경쟁
- 데이터 시장과 수요
  ・시장의 유형이 다수의 공급자에 의한 순수 경쟁시장인지, 하나의 공급자가 주도하는 독점시장인지, 소수 기업이 차지하는 과점 시장인지에 따라 판매가격이 영향을 받음.
  ・쉽게 말해 독점 시장에서는 가격을 높이 책정해도 판매가 되는 반면, 경쟁 시장에서는 다수 경쟁사보다 낮게 책정해야 판매가 가능함.
  ・과점시장의 경우에는, 경쟁 기업의 가격 정책이 자사의 가격에 영향을 많이 미침
  ・또한, 수요에 따른 가격탄력성도 가격 책정의 중요 요인이 될 수 있음.
  ・수요의 가격탄력성이 크다는 의미는 데이터 판매가격 변화에 따라 수요량의 변동도 큰 것을 뜻하며, 탄력성이 작은 경우는 데이터 판매가격을 높이거나 낮추어도 구매자 수의 변동이 크지 않은 것을 의미함.

- 대체가능한 상품이 많거나, 구매자에게 해당 데이터가 필수적이지 않은 경우는 가격을 높이면 구매자 수가 줄어들 수 있으므로 가격 탄력성이 크다고 할 수 있음.
- 가격 탄력성이 큰 경우에는, 한번 정해진 가격을 바꾸는 것이 어렵기 때문에 처음부터 이를 고려해 가격을 책정함. 만일 가격 탄력성이 작으면 판매가격을 높여 수익을 극대화하는 방법이 적합함.

- 경쟁자 요인
- 경쟁사 비용, 가격, 반응 등도 데이터 가격 결정의 영향 요인이 됨.
- 만일, 거의 유사한 데이터를 제공하는 경쟁사가 있다면 경쟁사 데이터의 구성이나 구축・운영비용, 판매 가격, 구매자들이 경쟁사 데이터에 갖는 신뢰성과 반응, 경쟁사의 차별화된 서비스 등을 파악해 가격 결정에 반영할 수 있음.
- 경쟁사가 유사한 데이터를 가진 경우, 경쟁사 대비 저렴한 가격을 설정하는 것 외에도 방법이 있음.
- 데이터는 범위와 구성에 따라 활용범위가 천차만별이기 때문에 서비스와 정보를 차별화하고 별도의 가격 정책을 펼침으로써 경쟁우위를 확보할 수도 있음.
- 또한 정교한 수준이나 품질 완성도, 업데이트 주기 등에 따라 데이터 가격을 다르게 책정할 수 있음.

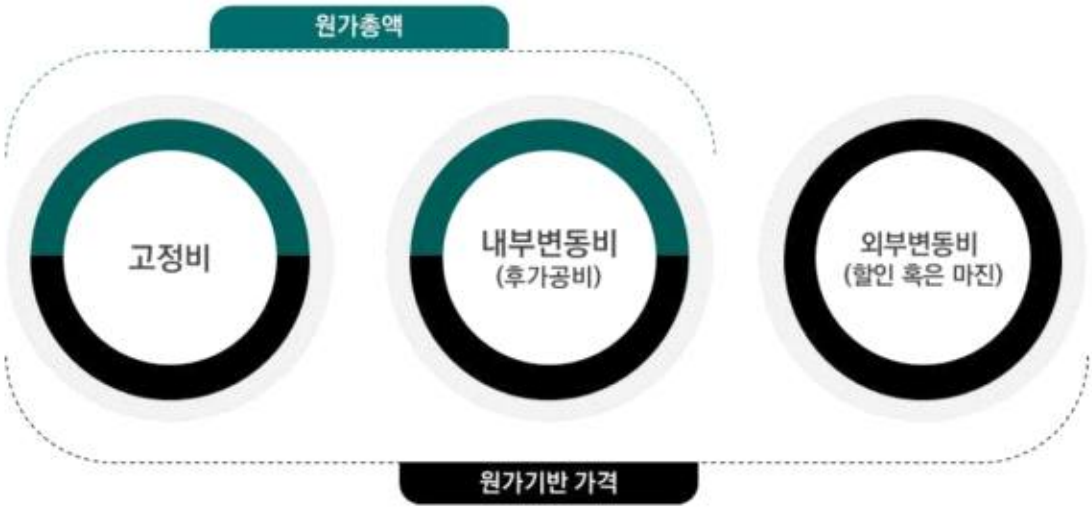
○ 환경
- 그 밖에 이자율, 인플레이션 등을 포함한 국내외 경제 수준, 데이터의 유통 경로, 데이터 거래 환경과 사회적 인식의 수준, 정부의 정책과 제도 등이 가격 결정에 영향을 미칠 수 있음.
- 데이터 활용 규제가 엄격하고 사회적 민감도가 높으면 민간 기업들은 데이터 거래 내용을 외부에 공개하기 꺼려하고 유통채널 없이 판매자와 구매자가 직접 거래하려는 경향이 나타남.
- 즉, 이러한 환경은 데이터 시장에서 가격 경쟁이 줄어들고 담합이 발생하거나 불합리한 조건이 추가되는 등 데이터 가격형성에 영향을 미칠 수 있음.

(4) 가격책정방법 선정

○ 원가 기준 가격 책정(Cost-Based Pricing)
- 원가 기준 가격 책정 방법은 비용(고정비, 변동비)을 충당하고 목표이익을 낼 수 있는 수준에서 가격을 결정하는 방법이며 이러한 가격책정방식을 비용회수형 방식이라고도 함.
- 이때 원가는 데이터의 구축・운영에 필요한 최소 비용과 자본을 유지할 수 있는 최소이윤을 포함한 총원가를 의미하기도 함.
- 원가 기준 책정방식은 비교적 공정하고 간단하게 가격을 설정할 수 있는 반면 수요나 경쟁을 고려하지 않아 데이터가 가진 가치를 제대로 반영하지 못한다는 단점이 있음.
- 원가만을 반영해 가격을 정하면 민간 기업 데이터는 고객이 지불하고자 하는 가격보다 다소 낮고, 정부를 비롯한 공공기관의 경우 고객이 인식한 가격보다 높게 책정될 가능성이 있음.
- 구입한 데이터의 실제 사용을 위해서는 반드시 가공이 수반됨.
- 데이터 판매자가 이종 데이터 간 결합을 위한 가공까지 제공할 경우인건비 등 원가가 매우 높게 책정될 수 있기 때문에 판매자가 어디까지 서비스할 것인지에 대해 제공 범위와 방법을 결정해야 함.
- 공공 부분 데이터는 공공기관의 고유 업무수행의 결과물로 생산되었고 공공부문의 예산(비용)이 소요되었기 때문에, 데이터 생산・구축을 위한 비용은 원가 책정에서 제외함. 반면, 데이터 보급이나 활용과정에서 사용자 편의성 및 부가가치 창출을 위해 투자가 이루어졌다면 이 비용은 원가에 포함될 수 있음.
- (예: 계측 데이터를 실시간으로 제공하기 위한 시스템 개발비용) 민간부문에서 데이터 판매는 주로 원가보다 경쟁 기준이나 고객 기준 단가를 반영함.

· 공공부문 데이터 서비스 원가는 물가안정에 관한 법, 국가정보화기본법, 공공데이터 제공 및 이용 활성화 법 시행령, 공공요금 산정 기준, 국가계약법의 예정가격작성 기준(기재부 계약예규 380호)을 준용하며, 일반적으로 수수료의 원가계산에 적용되는 용역원가 계산방식을 적용함.

- 원가 기준 가격책정에서 고려해야 할 사항은 고정비(인건비, 경비 등) 이외에도 다른 변동비가 포함될지를 결정해야함.
- 데이터의 특성상 수요기업이 원하는 형태로 상품을 가공해 판매할 수 있으므로 이때 발생하는 후가공비 등 내부변동비를 원가에 포함할 수 있고, 시장 환경에 따른 할인이나 마진 등 외부변동비 역시 가격 결정에 고려할 수 있음.



(원가 기준 가격책정)

< 원가 기반 제공방식별 과금체계 방식 >

제공방식	과금체계	요금제
데이터 셋	종량제	기본 종량제
		기본료 구간을 감안한 다운로드 횟수 기반 종량제
		기본료 구간을 감안한 사용량 기반 종량제
	정액제	기본 정액제
		사용일수(시간) 정액제
		용량기반 정액제
Open API	종량제	Open API 종량제
		기본료 구간을 감안한 call 수 기반 종량제
		기본료 구간을 감안한 사용량 기반 종량제
	정액제	Open API 정액제

○ 경쟁사기준 가격 책정(Competition-Based Pricing)

- 경쟁사기준 가격 책정 방법은 유사 데이터를 판매하는 경쟁사의 가격을 고려해 자사의 데이터 가격을 결정하는 방법임.
- 주요 경쟁자들이 정해놓은 가격을 감안하여 유사하게 책정하는 방법과 다수의 판매 기업이 경쟁 입찰에 따라 가격을 결정하는 방법 등이 있음.

- 이 방법은 유사한 데이터 상품들 간의 치열한 경쟁이 반복되는 환경에서 효율적으로 대처할 수 있음.
- 또한 경쟁기업의 데이터 서비스 대비 차별점이 크거나 이미지가 좋을 때에는 상대적으로 높은 가격을 책정하는 편이 좋고, 차별화가 어렵다면 시장의 수준과 유사하거나 경쟁기업 데이터 가격보다 낮게 책정하는 것이 유리함.
- 주로 기존 경쟁시장에 신규 진입하는 경우에 활용하지만, 시장 변화에 따라 경쟁사 가격이 자주 변경되는 상황이 발생할 수 있고 독과점 성격의 시장에서는 경쟁사의 판매 가격을 반영하는데 어려움이 있음.

○ 가치기준 가격 책정(Value-Based Pricing)

- 가치기준 가격 책정은 비용이나 시세를 기준으로 가격을 결정하는 전통적인 방식이 아니라 구매자가 인지하는 평가・수요・가치를 바탕으로 데이터 가격을 결정하는 방법임.
- 판매자는 비가격변수를 활용해 구매자의 지불의사와 가치를 확인하고 반영하거나 데이터를 활용할 때 예상되는 퍼포먼스 결과를 추적하여 제시함으로써 구매자의 가치기준 가격을 결정할 수 있음.
- 이 방법은 구매자가 기꺼이 지불하고자 하는 금액(willing to pay)을 기준으로 가격을 책정하기 때문에 구매자 입장에서는 합리적이지만 구매자가 인지하는 가치가 주관적이어서 일관되게 적용하기 어렵다는 단점이 있음.
- 데이터는 한번 적용하면 대체하기가 어려워, 지속적인 라이선스 비용을 받을 수 있다는 특성이 있음.
- 거래처 확보가 중요한 초기 데이터 사업자의 경우 구매자가 원하는 수준으로 가격을 정하고 꾸준한 거래를 하고자 함.
- 일부 회사에서는 데이터 상품을 무료로 제공함에도 간접 트래픽이 발생하여 전체 매출이 증가함.
- 다만, 이러한 매출의 비중이 높을 경우 타 포털, 타 유통플랫폼에 종속될 수 있음.
- 영세한 데이터 판매 중소기업들은 대형 포털과 거대 통신사들이 유통망을 장악하고 있는 상황에서 자체 유통망으로 수익을 내기 위해서는 독창적인 데이터 서비스와 수익 모델을 갖춰야 함.

< 가격 책정 방법별 장단점 >

가격 책정 방법	장점	단점
원가 기준 가격 책정	원가 중심으로 가격을 책정하기 때문에 마진폭을 쉽게 결정할 수 있음	무형가치를 반영하지 않아 큰 마진을 얻을 기회를 놓칠 가능성이 많음
경쟁사 기준 가격 책정	경쟁 업체에 대한 가격 경쟁력을 확보할 수 있음	지속적인 경쟁력을 확보하기 어렵고 가격변동이 심할 가능성이 매우 높음
가치 기준 가격 책정	고객 중심으로 가격을 책정하여 가장 잘 팔릴 수 있는 가격으로 접근 가능	가치가 주관적이어서, 일관된 가격을 적용하는 것이 매우 어려움

나. 판매 및 가격 전략 수집

(1) 데이터 제공 방식의 결정

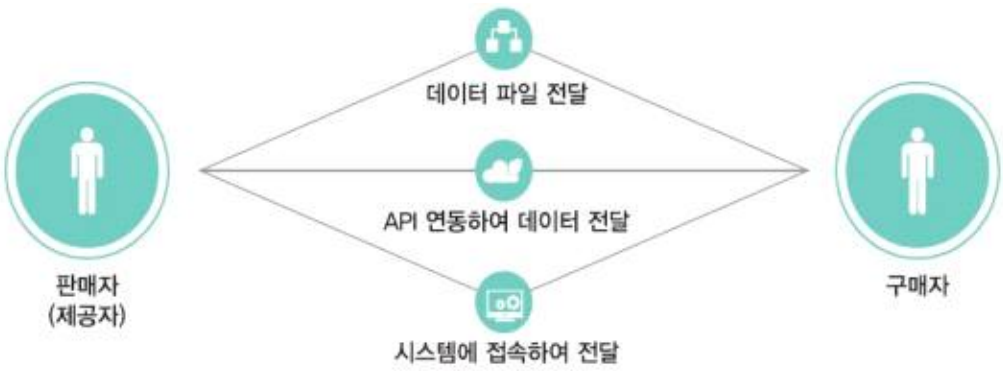
○ 데이터 제공 방식은 데이터 자체를 제공하는 방법과 서비스를 통해 데이터를 제공하는 방법으로 크게 2가지로 나뉨.

○ 데이터 자체를 제공하는 경우는 데이터 파일을 FTP・E-mail 등으로 전달하는 방법, Open API로 제공하는 방법, 구매자의 시스템에 직접 접속해서 제공하는 방법이 있음.

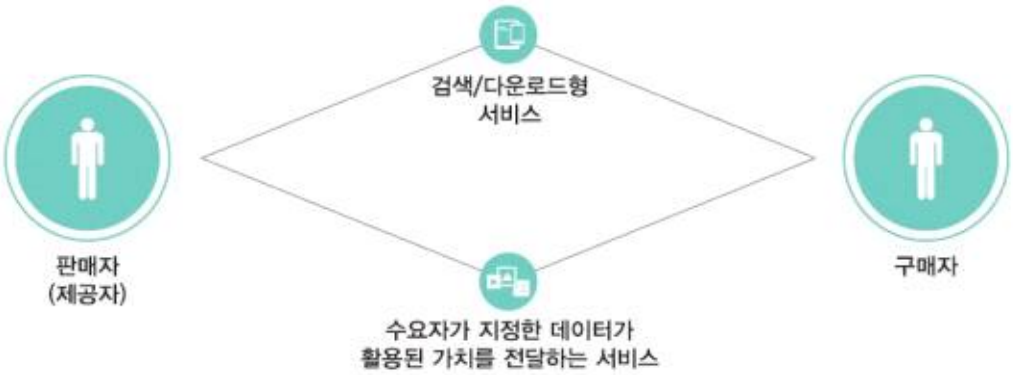
○ 위 3가지 방식은 기술적인 부분과 서비스 경험 관점에서도 매우 다르며, 이용료 과금 방법에도 차이가 있음.

○ 서비스를 통해 제공하는 방법의 경우, 조건에 맞는 데이터를 바로 검색・다운로드하는 솔루션 서비스와 수요자가 필요로 하는 조건의 데이터를 선택하면 그것을 활용해서 수요자에게 타겟 광고 집행 등 원하는 가치를 제공해주는 서비스(ex. DMP 서비스 등)가 있음.

○ 후자의 경우, 데이터를 직접 전달하지 않고 데이터가 활용된 부가가치를 제공하여 그에 따른 가격을 매기기 때문에, 데이터의 재사용 등으로 인한 권리 분쟁의 소지를 최소화할 수 있어 국내에서는 아직 미약하나 해외에서는 데이터 비즈니스의 대표적인 형태로 자리 잡아가고 있음.



(데이터를 직접 제공하는 경우)



(서비스를 통해 데이터를 전달하는 경우)

○ 데이터 이용 과금의 방식 결정

- 데이터 가격이 책정되면 이를 기준으로 이용료를 산정하여 부과하게 됨.
- 데이터 이용 계약을 할 때는 이용 허락에 대한 대가적 의미로 구매자가 판매자에게 이용료를 지급함.
- 일반적으로 데이터 이용의 과금은 ⑴개별 계약별로 권리 관계와 조건을 달리하여 과금하는 방식과 ⑵일반화된 하나의 서비스(API, 검색 솔루션 등)를 통해 동일한 조건과 권리를 제공한 후(이용약관), 예외 사항은 별도 협의하는 방식을 통해 이루어짐.

○ 개별 계약을 통한 데이터 이용료 과금

- 무형 자산인 특허권・상표권・저작권 등을 사용하고 지불하는 값을 로열티라고 함.
- 데이터의 거래 시 저작권, 사용권의 양도 등을 통해 이익을 창출하는 대가로 로열티를 지급할 수 있으며 조건에 따라 로열티를 다르게 책정할 수 있음.
- 데이터 이용에 대한 로열티는 확정된 금액을 일시/분할 지급하는 고정 이용료(Fixed Royalty)와 판매 수익이나 별도 정해진 산출 방법으로 이용료를 연동하여 지급하는 경상 이용료(Running Royalty)가 있으며 경우에 따라 두 가지를 혼합해 사용하기도 함.

< 개별 계약을 통한 데이터 이용료 과금 >

유형		내용
고정 이용료		데이터 이용에 대한 대가를 고정 금액으로 지급
선급 이용료		실시계약 체결시 데이터 이전을 위한 준비비용(또는 계약금) 형태로 지급되는 선불금
경상 이용료		생산량이나 판매량 등 데이터가 활용되는 정도에 비례해 대가를 지급하는 방식
혼합방식 (고정+경상)	최저 이용료	지급해야하는 최저 하한액을 설정하는 방식
	최대 이용료	지급 가능한 최고 상한액을 설정하는 방식

○ 고정 이용료 (Fixed Royalty)

- 데이터를 구매자가 어떻게 사용하는지 실적과 관계없이 일정한 금액을 부과하는 방법임.
- 이용료 지급 방식은 1회 이용료를 지급하는 '일시불 지급' 과 정해진 금액을 몇 차례 걸쳐 나누어 지급하는 '분할 지급' 방식이 있을 수 있음. 데이터의 제공 후 더 이상 판매자의 용역이 불필요한 경우 일시불로 지급함.
- 데이터 결합 등 추가 용역이 필요한 경우 분할 지급함으로써 위험부담을 낮추는 것이 일반적임.

○ 선급 이용료 (Initial Payment)
- 데이터 거래 계약이 성사된 직후, 계약금 형태로 일부 지급하는 이용료로서, 선불금 또는 착수금이라고 함.
- 데이터 판매자 입장에서 데이터 유출에 대한 일부 보장을 받는 것이고 데이터 구매자 입장에서는 상품화에 앞서 이용료를 지급하는 것이기에 위험부담이 될 수 있음.
- 선불금은 데이터 이전에 필요한 출장비, 교육비, 제작비 등을 포함한 수준으로 계상하는 것이 일반적임.

○ 경상 이용료 (Running Royalty)
- 경상 이용료는 구매한 데이터를 적용하여 새로운 서비스를 제조하거나 판매하는 경우, 판매량이나 매출액 등 데이터의 사용 실적에 비례하여 이용료를 산정하는 방식임.
- 데이터 서비스의 매출액 등의 비율(%)로 규정하는 요율 방식과 판매 횟수에 비례해 이용료를 책정하는 종량 방식이 있음.
- 경상 이용료는 매출액 기준이나 세금 공제 여부 등 명확한 기준을 상세히 계약서에 명시해야만 사후 분쟁을 예방할 수 있음. 초기에 용어와 방법에 대해 명백하게 정의하는 것이 좋음.

○ 최저 이용료 (Minimum Payment)
- 데이터 공급자(판매자) 입장에서 정하는 방법으로, 경상 이용료 등의 방식으로 산출된 이용료가 최저 이용료에 미치지 못하는 경우, 당초 약정한 최저 이용료를 지급하도록 하는 방식임.

○ 최대 이용료 (Maximum Payment)
- 주로 데이터 수요자(구매자)의 입장을 고려해 정해지며, 지불 총액이 일정 수준을 넘어서는 경우 초과분에 대해서는 지급을 면제한다는 개념임.
- 시장의 크기가 큰 경우 최대 이용료를 적용하면 이용료 금액을 낮출 수 있음.

○ 서비스를 통한 데이터 이용료 과금
- 데이터를 하나의 서비스로 구축하여 제공하는 경우, 데이터 이용대가가 곧 서비스의 요금을 의미하기도 함.
- 서비스 내 데이터의 과금 체계는 다운로드 수, 제공 건수(또는 호출 건수), 접속 시간, 접속 ID수 등에 따라 차등적인 요금을 부과하는 종량제와 데이터의 사용 정도에 관계없이 일정한 금액을 부과하는 정액제가 있음.
- 일반적으로는 정액제와 종량제를 혼합하여 사용하는 경우가 많음.
- 먼저 일정량 이하 데이터를 제공하는 경우에 고정 금액을 부과하고, 기준을 초과하면 이용량에 비례해 비용을 부과할 수 있으며 일정수준까지 이용량에 따라 비용을 부과하고 일정 수준을 넘을 때 고정 금액을 부과하기도 함.

< 서비스를 통한 데이터 이용료 과금 >

유형	내용
정액제	데이터의 사용정도와 관계없이 일정한 비용 부과
종량제	다운로드 수, 제공건수(호출수), 접속시간, 접속 ID수 등에 따라 차등적인 비용을 부과
혼합 (정액제+종량제)	일정량 이하 데이터를 사용하는 경우 고정금액을 부과하고, 기준을 초과하면 이용량에 비례해 비용을 부과
	일정수준까지 이용량에 따라 비용을 차등 부과하고 일정 수준을 넘으면 고정 비용을 부과

◯ 할인・가격 전략 및 정책

- 여타 상품과 같이 데이터 판매자는 이용자를 늘리기 위해 구매자를 대상으로 다양한 할인전략이나 차별화 전략을 적용할 수 있음.
- 예를 들어 학생이나 공공기관에 무료 이용권이나 할인권을 제공하기도 하고, 자주 구매하는 협력업체에게는 저렴하게 제공하기도 함.
- 데이터 가격 수준은 유사 데이터의 시장가를 기준으로 저가, 고가, 적정가로 나뉨. 만일 데이터 상품이 희소하고 기업 브랜딩이 필요한 상황이라면 초기 고가 전략을 선택할 수 있고, 후발 기업으로서 판로를 신속하게 확대하려는 경우에는 시장가보다 저가로 판매하는 시장침투 전략을 적용할 수 있음.
- 유사한 데이터 상품을 가진 기업이 많고 관습적으로 통용되는 가격대가 있는 경우 구매자의 심리를 반영해 유사 데이터 상품과 비슷한 가격으로 책정하기도 함.

◯ 데이터 가격 할인 및 차별화 방법

- 사용 기간 또는 사용량 할인
  ・사용 기간이나 사용량이 일정 수준 이상인 경우 이용료를 할인하는 방법으로, 주로 별도 협의를 통해 할인가를 책정함.
- 협력사 할인
  ・매자가 협력사인지 일반 고객인지에 따라 가격을 차별화하는 방법으로 고객 정보나 수익을 공유하는 협력사에게 가격을 할인하거나 무료 제공하기도 함.
- 이용 목적 할인
  ・학술・연구, 단순 열람, 공익적 목적 등 비상업적 목적이나 초기 테스트 목적으로 활용하는 고객에게 무료로 제공하거나 할인가격을 적용하기도 함.
- 전용망 할인
  ・특수 센터나 전용 서버를 통해 접속하는 고객에게 무료 제공 또는 할인 가격을 제공할 수 있음.
- 이용 시간 할인
  ・네트워크 과부하를 해결하기 위해, 서비스 이용 시간대에 따라 차등 가격을 적용하기도 함.
- 행사 할인
  ・고객 확보를 위해 신규고객에게 특별행사 가격을 제공하거나, 보완관계의 두 제품을 묶어 할인 가격에 제공하기도 함.

◯ 최종 이용과금의 결정

- 동일한 조건에서 생산된 데이터도 서로 다른 가격으로 판매될 수 있음.
- 이러한 가격 전략은 데이터 이용자를 늘리고 시장 점유율을 높이는 대안이 될 수 있을 것임.
- 가격 전략에 대한 고민 없이 고객이나 상황에 따라 제각각 다른 가격을 책정하는 것은 장기적 관점에서 고객에게 혼란을 줄 수 있음.
- 판매 기업이 추후 일괄적인 가격모델을 적용하는 데에도 장애로 작용할 수 있음.
- 따라서 데이터 판매 기업은 위와 같이 다양한 가격 전략을 펼침으로써 시장의 수요와 고객의 요구를 더 정확히 파악할 수도 있고, 더불어 수요에 맞는 데이터 상품을 개발・판매함으로써 이용자의 가치를 극대화할 수 있음.

다. 데이터 가격 책정 방법

(1) 데이터 및 데이터 거래 시장의 정의와 특성

○ 데이터 가치 평가나 가격책정을 위해서는 거래 상품으로서 데이터의 정의와 특성, 그리고 데이터 거래 시장의 정의와 범위 설정 중요

○ 윤경구 · 안명옥(2020)[1]은 데이터를 다음의 2가지로 정의

- 첫째, 분석을 위해 함께 수집된 사실 및 통계
- 둘째, 컴퓨터가 조작을 수행하고, 전기 신호의 형태로 저장 및 전송되며, 자기, 광학 또는 기계적 기록 매체에 기록되는 수량, 문자 또는 기호

○ 데이터 거래소의 거래 상품으로서 데이터의 정의와 관련하여 데이터와 정보의 관계 검토 필요

- 정보(information)는 가공된 데이터로서 누가, 무엇을, 어디서, 언제 등의 질문에 대한 답을 제공하는 유용성이 있음(Ackoff, 1989)[2]
- 실제 데이터 거래소에서는 데이터의 사전적 정의를 벗어나는 보고서 등의 정보가 거래되는 경우도 적지 않음 : 예를 들어, KADX에서는 '농식품 해외시장 동향 보고서', '농식품 해외 바이어 정보', '국제 곡물 시장 정보' 등 aT에서 생성한 보고서가 거래되고 있음

(KADX의 정보 거래 사례)

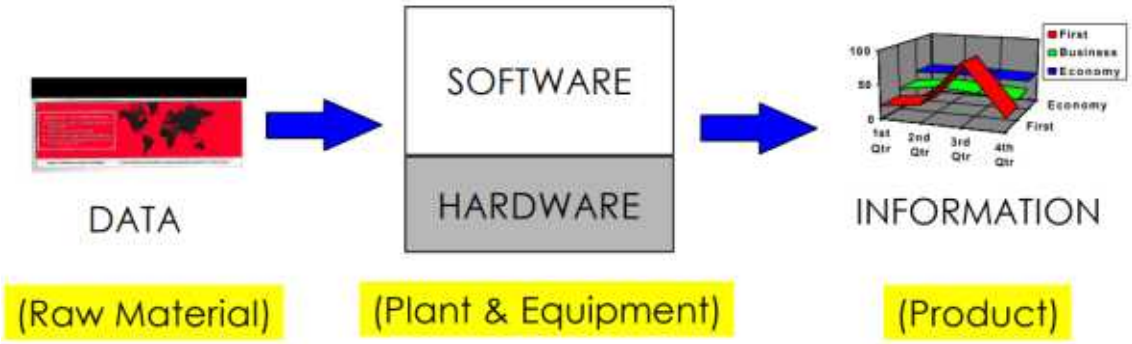
자료 : 농식품 빅데이터 거래소 홈페이지(kadx.co.kr)

○ 데이터와 정보의 관계와 관련하여 Moody and Walsh(1999), Ackoff(1989), Zeleny(1987) 등은 데이터와 정보 사이에 위계 구조가 존재하고, 데이터는 정보 생성에 필요한 투입 요소로 그 자체로는 가치가 없다고 주장

- Moody and Walsh(1999) : 정보 체계(information system)에서 데이터를 원재료(raw material), 소프트웨어와 하드웨어를 데이터를 이용해 정보라는 생산하는 플랜트 및 설비(plant and equipment), 정보를 고객에게 전달되는 최종 생산물(the end product that is delivered to the customer)로 정의
  · 특히, 데이터가 아닌 정보가 서비스 제공, 효과적 의사결정, 실적 향상, 비교 우위 확보 등의 역량을 제공하고, 최종 생산물로서 직접 판매될 수 있음(Information provides the capability to deliver

1) 윤경구, 안명옥, 2020, 데이터 가치 창출에 관한 사례 및 시사점, 우정정보, 2020 가을호: 91-114.
2) Ackoff(1989)는 데이터를 'symbols', 정보를 'Data that are processed to be useful; provides answers to who, what, where and when questions'로 정의(Rowley, 2007, p.167)

services, make better decisions, improve performance, achieve competitive advantage and can also be sold directly as a product in its own right.)



(정보 생산 과정)

자료 : Moody and Walsh(1999, p.2)

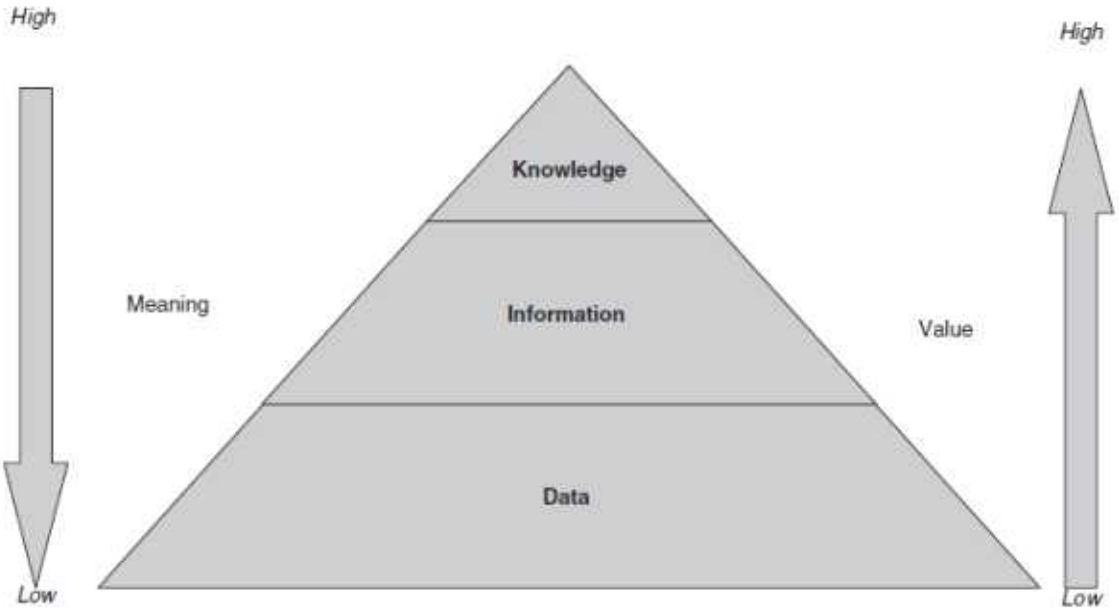
- Ackoff(1989), Zeleny(1987) : 데이터 자체는 아무런 의미가 없는 상징에 불과하고, 데이터를 가공하여 만들어진 정보가 지식 등의 형성에 유용하게 활용됨(Zeleny, 1987; Ackoff, 1989; Rowley(2007) 재인용)

< 데이터, 정보, 지식의 정의 비교 : Ackoff(1989) vs. Zeleny(1987) >

구분	Zeleny(1987)	Ackoff(1987)
Data (데이터)	· Know nothing	· Symbols
Information (정보)	· Know what	· Data that are processed to be useful; provides answers to who, what, where and when questions
Knowledge (지식)	· Know how	· Application of data and information; answers how questions

자료 : Rowley(2007, p.167) 수정 후 재인용

- Chaffey and Wood(2005) : 데이터, 정보, 지식 사이에 그림과 같은 위계 구조가 존재하고, 데이터에서 정보, 지식 단계로 올라갈수록 의미(meaning)와 가치(value)가 증가한다고 지적



(데이터-정보-지식의 위계 구조에서 의미와 가치의 수준)

자료 : Rowley(2007, p.167) 재인용

○ 본 연구는 가치 평가 및 가격책정의 대상이 되는 거래 상품으로서의 데이터에 정보를 포함하는 광의의 개념 적용

- 이는 첫째, 선행연구를 고려할 때 학술적으로 데이터와 정보의 개념이 다르고, 둘째, 데이터 구매자는 데이터 자체보다 데이터로부터 도출되는 정보에서 가치를 얻기 때문임
- 다만, 관행적으로 '데이터 구매 행위'에 '정보획득'의 의미가 포함되어 사용된다는 점을 고려할 때 광의의 데이터 정의 적용이 적절하다고 판단
- 광의와 협의의 데이터 구분이 필요한 경우 별도 설명 제시

○ 선행연구에서 제시된 거래 상품으로서 데이터의 특성

- 이규엽 외(2019) : 데이터는 생산요소의 특성 이외에도 비경합성, 중간 수준의 비배제성, 무형자산(정보 · 아이디어 · 지식), 프라이버시 침해 가능성 등 외부효과(externality) 발생 등의 특성을 가짐(이규엽 외, 2019, 수정 후 인용). 이규엽 외(2019)에서 제시된 각 특성은 다음과 같이 정리할 수 있음[3)]

① 생산요소 : 생산성 개선, 신상품 개발 등을 통한 부가가치 향상을 위한 정보제공

② 비경합성 : 경합성이란 한 사람이 소비하면, 다른 사람이 소비할 수 없는 속성으로 데이터의 경우 무제한 복제가 가능하므로 비경합성을 가짐

③ 부분적 비배제성 : 배제성이란 소유자가 타인의 소비나 사용을 제한할 수 있는 속성으로 데이터가 지적 재산권으로 보호되는 경우 배제성을 가짐. 그러나 복제가 쉽고 판매자와 구매자 사이에 정보 비대칭성(판매자가 구매자의 무단 복제 및 배포 행위를 쉽게 알 수 없음)으로 사용 제한에 한계

④ 무형자산 : (정보를 포함하는 광의의) 데이터의 자산으로서의 가치는 (협의의) 데이터로부터 도출되는 정보, 아이디어, 지식 등에 의해서임

⑤ 외부효과 : 외부효과는 '특정 주체의 경제활동 결과가 시장을 통하지 않고, 즉 가격 보상

3) 이규엽 외(2019)에서는 개별 특성만 제시되었으나, 연구자가 개별 특성의 세부 내용 추가

또는 비용 부과 없이 다른 주체의 후생에 영향을 미치는 현상’ 을 의미. 외부효과는 양의 외부효과(positive externality; 다른 주체의 후생을 증가시키지만, 가격 인상 등의 보상을 받지 못하는 상황)와 음의 외부효과(negative externality; 다른 주체의 후생을 감소시키지만, 비용을 부담하지 않는 상황)로 구분. 데이터 제공 시 개인정보 유출 등으로 프라이버시가 침해당하는 경우 음의 외부효과가 발생할 수 있고, (협의의) 데이터로부터 도출된 정보, 아이디어, 지식 등이 해당 데이터를 구매하지 않은 경제주체의 부가가치 향상에 영향을 주는 경우 양의 외부효과 발생. 영세 경영체가 많은 농업 분야의 경우 경영체 차원의 데이터 구축이나 구매가 어려울 뿐만 아니라 양의 외부효과로 인해 데이터를 구축하거나 구매하더라도 그에 상응하는 대가를 제대로 받지 못하므로 데이터 거래를 시장에만 맡기는 경우 사회적으로 필요한 데이터가 공급되지 못할 수 있어 정부 차원의 데이터 공급이나 지원 필요

< 데이터의 거래 상품으로서의 특성 >

구분(type)	데이터(data)	원유(oil)	산소(oxygen)
생산요소(input factor)	O	O	X
비경합성(non-rivalry)	O	X	O
비배제성(non-excludability)	△	X	O
무형자산(정보/아이디어/지식)	O	X	X
프라이버시 침해(privacy violation) 가능성 등 외부효과 발생	O	X	X

자료 : 이규엽 외(2019) 재인용

주1 : ‘O’는 그러함, ‘△’는 일부 그러함, ‘X’는 그렇지 않음을 의미

주2 : 이규엽 외(2019)에서는 표에 제시된 특성을 ‘경제학적 특성’이라고 하였으나, 본 연구에서는 이를 시장에서 거래되는 ‘거래 상품으로서의 특성’으로 변경

주3 : 이규엽 외(2019)의 ‘정보/아이디어/지식’에 ‘무형자산’, ‘프라이버시 침해 가능성’에 ‘외부효과’ 추가

- 윤경구 · 안명옥(2020) : 거래 상품으로서 데이터의 특성으로 비경합성, 외부효과, 부분적 배제성 지적

① 비경합성 : 비경합성으로 같은 데이터를 많은 사람이 이용할 수 있어 사회는 유용한 데이터로부터 큰 혜택을 받을 수 있음. 민간 데이터의 경우 지적 재산권 설정으로 데이터 접근 및 이용에 제한을 두는 경우가 적지 않은데, 이 경우 비경합성에 따른 혜택 확대에 한계. 반면, 윤경구 · 안명옥(2020)이 지적한 혜택 확대 한계라는 문제 해결을 위해 무차별적으로 데이터 접근 및 이용을 허용하는 경우 민간 데이터 공급 주체의 이익을 침해하여 데이터 공급을 위축시킬 우려도 있음

② 외부효과 : 개인 데이터의 수집, 공유 및 처리 과정에서 개인정보가 유출되어 데이터 거래에 참여하지 않은 개인에게 비용 또는 손실을 발생시킬 수 있어 개인정보 보호 중요. 윤경구 · 안명옥(2020)에서는 제시하지 않았으나, (협의의) 데이터로부터 도출된 정보, 아이디어, 지식 등이 해당 데이터를 구매하지 않은 경제주체의 부가가치 향상에 영향을 주는 경우 양의 외부효과 발생 가능

③ 부분적 배제성 : 부분적 배제성을 고려할 때 상호 접속된 시스템상에서 데이터를 보존

할 때, 사이버 공격 등에 의한 데이터 손실을 막기 위한 지속적 투자 필요. 윤경구 · 안명옥(2020)에서 언급하지 않았으나, 데이터 판매자와 구매자 사이의 정보 비대칭성이 존재하는 상황에서, 데이터 구매자의 무단 배포 등으로 데이터가 유출되는 경우 데이터 공급을 위축시킬 수 있음

○ 거래 상품으로서 데이터의 추가 특성

- 선행연구에서 제시한 생산요소, 비경합성, 부분적 비배제성, 무형자산, 외부효과 등의 특성 이외에 다음의 추가 특성 존재

① 제로에 가까운 한계 사용 비용(zero marginal use cost) : 데이터 제공을 위한 인프라가 안정적으로 구축되면, 데이터가 사용에 거의 비용이 소요되지 않음. 비경합성에 따른 사회적 혜택 확대는 제로 한계 사용 비용에 의해 배가될 수 있음

② 본원적 가치를 갖지 못하는 파생상품 : NFT(Non-Fungible Token, 대체불가토큰)와 같은 특별한 경우를 제외하고는 데이터의 가치는 ‘소유’가 아닌 ‘사용’으로부터 발현됨. 또한, 기초자산의 가치에 의해 가치가 결정되는 선물, 옵션 등의 파생상품과 마찬가지로 데이터 사용자가 속한 산업이나 경영체의 시장 규모나 소득에 의해 가치가 결정되는 파생상품. 예를 들어, 농촌진흥청의 농산물소득조사 데이터의 가치는 사용자가 농가인지, 농작물재해보험을 운영하는 보험사인지, 아니면 연구자인지에 따라 달라질 수 있음. 그러므로 데이터 가치 평가나 가격책정 시 수요자 특성 분석이 중요하고, 필요한 경우 차등적 가격책정 필요

③ 네트워크 외부효과(network externality) : 특정 재화에 대한 소비자의 수요가 그 재화를 소비하는 다른 소비자의 수요에 의해 영향을 받는 네트워크 외부효과로 데이터 사용자가 증가할수록 데이터 수요가 확대될 여지가 큼. 또한 데이터 사용이 증가할수록 사회적 가치도 증가

④ 정보 비대칭성(information asymmetry)에 따른 가치 손실 우려 : 데이터의 경우 사전적 정보 비대칭성에 따른 역선택(adverse selection; 의사결정이나 위험관리에 필요한 정보가 충분하지 않아서 저품질 상품이 많아지고, 고품질 상품은 시장에서 사라지는 문제), 사후적 정보 비대칭성에 따른 도덕적 해이(moral hazard; 본인이 대리인의 행동을 일일이 감시하지 못하는 상황을 이용해 대리인이 자신의 이익을 좇아 행동하는 문제)에 따른 가치 손실 우려가 있으나 관리가 쉽지 않음. 데이터 시장의 역선택 문제는 데이터 공급 주체가 정확하고 신뢰성 있는 데이터를 공급했는지를 확인하기 어려워 발생하는 문제로 담당 인력이 표본 확보의 한계로 확률표본추출(random sampling) 대신 편의표본추출(convenience sampling)을 사용하거나 조사비용 등을 이유로 왜곡된 데이터를 생성하는 경우 발생. 데이터 시장의 도덕적 해이는 공급자와 구매자 양자 모두에서 발생할 수 있는데, 공급자에 의한 도덕적 해이 문제는 공급자가 특정한 의도를 가지고 데이터를 왜곡하거나 데이터 관리 소홀로 공급 데이터 사용 기회가 상실 또는 감소하는 경우 발생. 구매자에 의한 도덕적 해이는 구매자가 공급자의 지적 재산권을 고려하지 않고, 데이터를 무단으로

배포하는 경우 발생. 데이터 시장의 역선택 및 도덕적 해이 문제는 거래소 설립 이후 지속적인 데이터 품질 관리와 지적 재산권 보호가 중요함을 의미

⑤ 공급 데이터 분할 판매 시 데이터 가치 보존 한계 : 농식품의 경우 분할 판매하는 경우 분할 이전의 평균 가치가 유지되거나 오히려 증가할 수 있으나, 데이터를 분할 판매하여 표본크기가 감소하거나 제공항목이 줄어드는 경우 분할 이전에 비해 평균 가치가 급감할 우려 있음. 예를 들어, 사과의 경우 낱개나 1/2개 등으로 분할 판매하는 평균 가치 역시 상승하여 판매가 인상이 가능하나, 농산물소득조사에서 총수입 관련 자료와 생산비 관련 자료를 분할 판매하는 경우 평균 가치 하락으로 판매가 인하 필요. 이러한 특성으로 공급원가를 기준으로 가격을 책정하는 데 어려움 있음. 예를 들어, 사과 50개를 생산에 10만 원을 지출하는 경우 개당 2,000원 이상으로 판매하여 생산비를 보존할 수 있으나, 완성된 형태의 사과 소득조사 데이터 50MB 구축에 10만 원을 지출하는 경우 1MB씩 분할 판매하면 MB당 2,000원으로 판매하기 어려움. 그러므로 데이터 생산비를 기준으로 가격책정을 하는 경우 완성된 형태의 데이터를 얼마나 많은 사용자에게 판매할 수 있는가에 대한 예측 필요. 생산비가 보존되는 손익분기점을 판매가격으로 설정하는 경우 사과는 식 (1), 사과 소득조사 데이터는 식 (2)와 같은 방식으로 가격책정 가능

(손익분기점 기준 사과가격 책정)

$$P_{\text{사과}} = \frac{TC_{\text{사과}}}{\text{총생산량}_{\text{사과}}} \quad (1)$$

(손익분기점 기준 사과 데이터가격 책정)

$$P_{\text{사과데이터}} = \frac{TC_{\text{사과데이터}}}{\text{사과데이터 예상 판매건수}} \quad (2)$$

⑥ 이종 데이터 간 결합을 통한 신규 가치 창출 : 이종 데이터 간 결합으로 신규 가치가 창출될 여지가 큼. 특히, 최근 민간 데이터 결합전문기관도 코드나 암호를 부여해 누구의 정보인지 알 수 없도록 처리한 가명 정보를 또 다른 가명 정보 데이터와 조합해 제3자에게 제공할 수 있도록 허용할 방침이라 이종 데이터 간 결합의 혜택이 증가할 전망

⑦ 사용에 따른 감가상각 제로 : 사용 증대에 따른 감가상각, 즉 가치의 손실이 발생하지 않으나, 최근 데이터센터 화재로 큰 파장을 일으켰던 카카오 사태에서 나타난 바와 같이 데이터 보존과 관리가 제대로 이루어지지 못하는 경우 큰 손실이 발생할 수 있음. 반면, 시간 경과에 따른 데이터 자체의 손실이 발생하지 않더라도, 다른 자산과 마찬가지로 가치 손실 발생. 그러나 시간이 경과한 데이터라 하더라도 최신 데이터와 결합하여 활용이 가능한 경우 가치가 상승할 수 있음

⑧ 강단조성(strong monotonicity) 미성립 : 신고전학파 경제학에서 상품 또는 재화는 강단조성, 즉 많을수록 좋다는 다다익선(the more the better)의 법칙이 성립하나, 데이터의 경우 일정 수준 이상을 보유하는 경우 의사결정에 혼란을 주어 데이터 사용에 따른 효용이 감소할 수 있음

⑨ 데이터의 정확성과 신뢰성이 중요한 경험재(experience goods) : 데이터의 경우 경험재로서 상품의 상세 정보만으로 품질 평가가 쉬운 탐색재(search goods)와 달리 실제로 사용한 이후에 품질과 유용성 평가가 가능. 특히, 품질과 관련하여 정확성과 신뢰성이 떨어지는 경우 가치가 급감할 수 있음

(3) 국내 데이터의 특성

○ 국내 농업 데이터의 라이프 싸이클

- 데이터 라이프 싸이클은 데이터의 현재 상태를 결정하는 중요한 정보로서 데이터 거래를 위해 정리 필요
- 데이터 라이프 싸이클 조사항목 : 데이터 명칭, 제공기관, 제공형태, 조사주기, 조사기간, 데이터설명, 조사항목, 조사방법 등

〈 국내 농업 공공 데이터의 라이프 싸이클 현황 〉

대분류	소분류	공공데이터명	기관명	제공형태	조사주기	조사기간	자료설명	조사항목	조사방법	데이터 관련 의견
농산	식량작물	국영무역품목 시장접근물량	aT	파일	연간	2010 ~ 2020	aT 관리 국영무역품목의 시장접근물량 수입 계획량	8개품목 수입량	수입실적 및 생산능력, 재배면적, 선호순 등 기준 적용 결정	
	식량작물	국영무역품목 양허세율	aT	파일	수시		국영무역품목의 양허세율	7개품목 저율관세, 고율종가, 종량	비공개	조사방법 Historical data, 일부품목 데이터 누락(땅콩, 팥(녹두))
	식량작물	작물영향 진단서비스	농진청	Open API						
	원예작물	시설채소 생산실적	농식품부	파일	연간		목적 : 농산물 생산 및 수급대책 수립 및 기초자료 제공	시도별, 작목별, 재배시설별 면적, 온실시설 생산량, 단수, 도매가격, 수입실적 등	표본조사(5종) 행정조사(38종)	연도별 데이터 제공 (연속 데이터는 통계청 이용)
	원예작물	채소류 생산실적	농식품부	파일	연간		목적 : 농산물 생산 및 수급대책 수립 및 기초자료 제공	시도별, 작목별 면적, 단수, 생산량 등	「농업통계조사 규칙」에 따라 조사	연도별 데이터 제공 (연속 데이터는 통계청 이용)
	원예작물	화훼류 시도별 재배현황	농식품부	파일	연간		전국 화훼 품목별 재배 및 판매 관련 정보를 제공	시도별, 재배시설별 농가수, 재배면적, 판매량, 판매액, 시설 면적 등		연도별 제공
	특용작물	약용식물 서비스	산림청	Open API	수시			약용식물, 식물명, 약재명		

대분류	소분류	공공데이터명	기관명	제공형태	조사주기	조사기간	자료설명	조사항목	조사방법	데이터 관련 의견
	특용작물	특용작물 생산통계	농식품부				특용작물(유지작물, 기호, 약용작물 등) 생산실적 통계자료			공공데이터 연결 안됨
축산	곤충	누에사육 규모별 농가수	농식품부	파일	연도별	2004-2016	전국의 모든 잠업가구의 양잠형태별 생산현황, 누에 생산 및 판매현황 조회 서비스	시도별 시군별 사육규모별 농가수		연속데이터는 시도별 제공
	곤충	양잠산물 생산현황	농식품부	파일	연도별	2008-2016	전국의 모든 잠업가구의 양잠형태별 생산현황, 누에 생산 및 판매현황 조회 서비스	시도별 연도별 종류별 농가수, 사육량, 생산량		시기별 조사 대상의 차이로 분할 데이터 형태 제공
	곤충	뽕나무 묘목 생산 공급현황	농식품부	파일	연도별	2011-2016	모든 잠업가구의 양잠형태별 생산현황, 누에 생산 및 판매현황 조회 서비스	시도별 연도별 종류별 생산농가 수, 재배면적, 생산주수		시기별 조사 대상의 차이로 분할 데이터 형태 제공
	동물	기타가축통계	농식품부	파일	연도별	2013-2014	시도별 기타가축(마필, 산양, 면양, 사슴, 토끼, 개, 오리, 칠면조 등) 사육농가 및 사육마릿수 자료 제공 목록 : 년도, 지역, 사육가구 수	시도별 연도별 기타가축(8종) 사육 농가수 규모별 분포		세부 축종별 자료 제공하지 않음 가구 규모에 따른 범위 형태로 자료 제공
	동물	가축분뇨실태현황	농식품부	Open API	연도별	2013-2015		읍면동별		신청서버 알수 없음(안내없음) 파일도 존재하나 변수명을 알수 없음
	사료	배합사료생산실적	농식품부	파일	연도별(수시)	2000-2015	2013년 분기별 배합사료생산실적 통계에 대한 데이터로 (대유용, 양축용, 어류용, 실험동물용) 항목제공	종류별 분기별 배합사료 생산량		단위 미표기 및 변수명 부정확(PRDCTN_QY)
	사료	표준사료 검색서비스	농진청	Open API	수시			사료별 조성분, 영양가, 무기물, 제조방법, 제조기술 제공		농진청 회원가입 후 API 신청
	유제품	시도별 우유 생산량 (우유 및 유제품 생산소비현황)	농식품부	파일	연도별	2011-2013	전국시도별 우유 생산 및 소비, 유제품 생산 및 소비 실적 통계자료 데이터 제공 제공 목록 : 년도, 유제품별 생산량 소비량		시도별 연도별 자료 제공 설명과는 다른 데이터 제공	목차명과 데이터 명의 차이가 존재 생산인지 소비인지 구분 안됨

대분류	소분류	공공데이터명	기관명	제공형태	조사주기	조사기간	자료설명	조사항목	조사방법	데이터 관련 의견
										단위 미표기
	축산농장	가축사육통계제공 서비스	농식품부	Open API	수시			읍면동(5개 지역) 가축 수		실질적 제공 데이터 부재
	축산농장	가축사육현황	가축위생방역지원본부	파일	연도	2017-2021	전국 가축사육농장 현황 시도별 집계 자료	전국 축산농장 총괄 축종별 호수 및 마리수		
	축산농장	축산농장 통계제공 서비스	농식품부	Open API						자료 확인 불가
	축산농장	화훼류 시세현황	aT	파일	일별		화훼유통정보시스템에 수집된전국 5대 화훼공판장 소속 도매법인을 통해 거래되는 품목, 품종들의 경락가격 등 시세자료	종류별 품목별 품종별 등급별 최고단가, 최저단가, 평균단가, 총수량, 총금액	화훼유통정보시스템 등록된 경락 가격 수집 및 정리	농식품 빅데이터 거래소 무료 제공 데이터 위치 재설정 필요
	유통관리	농산물 도매시장 현황	농식품부	파일	수시		공영도매시장, 일반법정, 민영 도매시장 현황 자료 농산물 도매시장 현황입니다 총 46개소(공영도매시장 32개소, 일반법정도매시장 11개소, 민영도매시장 3개소)	도매시장별 법인명, 대지 건물, 경매장, 주차장 면적 제공		한글파일 형태로 제공
	유통관리	로컬푸드 직매장 현황	농식품부	파일	수시		로컬푸드 직매장 현황(시군구, 사업장명, 농협명칭 등) 제공 목록 번호, 지역, 시군구, 사업장명, 농협명칭, 개장일, 전화번호, 주소	시군구별 농협명칭, 개장일, 전화번호, 주소		
	유통관리	시장 법인별 거래실적	aT	파일	연도	2017-2019	전국 공영 농수산물 도매시장의 정보(도매시장명, 건물규모, 부지규모) 및 법인별 거래금액(백만원) 및 거래물량(톤) 실적 자료	도매시장법인별 거래물량, 거래금액		
	재해	농작물 보험금 지	농업정책보험금	파일	연도	2001-2021	연도별로 농작물재해보험에 계약된	지급건수, 지급보험금		

대분류	소분류	공공데이터명	기관명	제공형태	조사주기	조사기간	자료설명	조사항목	조사방법	데이터 관련 의견
		급현황	융원				건수들을 지급건수(건), 지급보험금(백만원), 손해율(퍼센트) 등등 지급실적을 정리한 자료	(백만원), 손해율(퍼센트)		
농지/용수	농업용수	새만금방조제 수위계측정보	한국농어촌공사	파일	시간별		한국농어촌공사 새만금유역 가력도 수위계측정보에 대한 데이터로 해수, 담수의 수위정보를 일별 분단위 등 항목 제공	수위정보(일별)		
	농업용수	영농여건불리농지도	한국농어촌공사	파일			농업진흥지역도(시도별, 용도구역구분(농업진흥지역, 농업보호지역), 공간정보(Shape) 등 영농여건불리농지도(시도별 영농여건불리농지정도, 공간정보(Shape) 등)			dbf 파일 형태로 제공
	농지	경영이양직접지불제 통계자료	한국농어촌공사							확인불가
	농지	과원규모화 사업추진실적	한국농어촌공사	파일	연도별		과원규모화-매도, 과원규모화-임대 사업별 지역별 계약건수, 면적, 금액의 년도별 통계자료를 제공	지역, 계약건수, 면적, 금액		
	농지	농지 실거래가 동향 분석 결과	농림축산식품부	파일	분기별		신고된 농지(논, 밭, 과수원)별 실거래 가격현황(읍,면,동별 평균, 최저, 최고가)	농업진흥지역 내외 시도별 농지별 농지가격		한글파일 형태로 제공
시설장비	시설장비	농수축산물 조사가격 정보	농림수산식품교육문화정보원	파일	월별		전국에서 거래되는 농산물, 수산물, 축산물 중 국민들의 삶에 영향을 미치는 주요 품목들에 대하여 한국농수산식품유통공사, 농업협동조합중앙회, 수산업협동조합중앙회 및 축산물품질평가원에 소속된 조사원들이 조사한 가격 정보와 거래되는 상품들의 상세정보	농수축산물 원천조사가격정보 조사지역, 표준시장, 조사 단위, 조사 등급, 조사 가격		

대분류	소분류	공공데이터명	기관명	제공형태	조사주기	조사기간	자료설명	조사항목	조사방법	데이터 관련 의견
	시설장비	농협 공판장 경락가격	농림수산식품교육문화정보원	Open API	일별		농협 산지 공판장에서 거래되는 농산물들의 경락일자, 경락가격, 경매구분, 시장코드, 법인코드, 품목코드, 거래단량, 단위 등 경매에 대한 상세 정보 제공			
	도시농업	웰빙도시농업정보	농림수산식품교육문화정보원	파일	수시		농업관련 도시농업정보 및 웰빙정보 컨텐츠를 제공하며, 식품별 칼로리, 요리법, 공기정화 식물, 베란다 채소 등의 정보를 제공 함	메뉴별, 제목별 도시농업정보		
	도시농업	칼로리 정보	농림수산식품교육문화정보원	파일	수시		음식 및 식품 속에 있는 칼로리, 탄수화물, 단백질, 지방산 등 영양을 계산하여 정보 제공	음식별 칼로리, 탄수화물, 단백질, 지방, 콜레스트롤, 식이섬유, 나트륨		
말산업	말산업	경주마 등록 내역	한국마사회							확인불가
	말산업	말 방역 현황								확인불가
교육	교육	농식품공무원 교육훈련 통계자료	농림수산식품교육문화정보원	파일	수시		농식품공무원교육원에서 교육과정 운영 후 실시한 강의방식별, 과정별, 기관별 교육훈련 통계자료	교육훈련인원, 교육과목, 사이버교육 운영 현황, 직급별 현황, 연령별 현황		
	교육	식생활 교육기관 현황	농림수산식품교육문화정보원	파일	수시		국민건강증진 및 식생활개선을 목표로 식생활교육 인력 양성을 위해 지정된 전문 교육기관 현황	지역별, 교육기관별 연락처 및 주소		
농촌복지	농촌개발	경관보전직불 사업지구	농식품부	파일	연도별	2013-2021	지구별 경관작물 재배에 따른 경관보전직불 관련 자료로 사업년도, 지자체, 지구명, 작물명, 참여농가수, 지원면적 등의 항목 제공	읍면동별 대상지구별 농지별 작물별 참여자 수, 대상면적		2016-2020 자료 부재
	농촌개발	일반농산어촌지역개발사업	한국농어촌공사	Open API						

대분류	소분류	공공데이터명	기관명	제공형태	조사주기	조사기간	자료설명	조사항목	조사방법	데이터 관련 의견
	농촌개발	지역별 과수시설 현대화 지원 실적	농식품부	파일	연도별	2010-2016	생산시설 현대화를 통한 고품질 과수 생산 및 재해예방 등 과수산업 경쟁력 강화 등 고품질 생산을 위한 지역별 시설 지원 실적 현황	시군구별 작목별 시설별 지원금액		
	농촌관광	농촌관광사업 등급 정보 현황	한국농어촌공사	파일	연도별	2019-2021	농촌관광 관련 농촌체험휴양마을, 관광농원, 농어촌 민박을 대상으로 체험, 교육, 숙박, 음식에 대한 등급 결정 현황을 제공	시군구별 마을별 체험, 교육, 숙박, 음식별 등급 현황		
	생명자원	곤충자원 정보	농림수산식품 교육문화정보원	파일	수시		생물자원 곤충자원 소개, 특징, 사육방법 정보	곤충명, 학명, 사육방법 자원내용		
	생명자원	누에자원 정보	농식품부	파일	수시		누에사육 농가, 사육형태, 사육분포도, 생산량, 판매량 및 활용 분야 등의 정보			정보 식별 불가
	영농	여성 농업인 실태조사	농식품부	파일	연도별	2018	여성경영주, 노동시간, 농가수, 통계자료 등 여성농업인실태조사	연령, 가구형태, 가족수, 혼인, 월평균지출, 거주지, 영농규모, 주력품목, 경작면적, 농산물 판매액별 여성 농업인 비중		
	영농	작목별 기술상세정보	농진청	Open API	수시					
	종자	국립산림품종관리센터 품종보호 출원현황	산림청	파일	수시		국립산림품종관리센터에 접수된 품종보호출원내역	출원번호, 작물명, 품종명칭, 대조품종, 상태, 출원인, 출원일자		
	종자	연도별 종자 수출입 물량 및 국내외 채소 종자 생산량 현황	농식품부	파일	연도별		연도별 종자 수출량 및 수입량 현황, 채소종자의 생산량 및 해외채종 비율 등을 파악하기위해 만든 통계 데이터로 종자 수출입, 채소종자 생산량 등			접근오류 확인 불가
농축산	GIS	과수재배변동정보	농진청	OpenAPI			시도명, 시군명, 과수명, 품목 코드			

대분류	소분류	공공데이터명	기관명	제공형태	조사주기	조사기간	자료설명	조사항목	조사방법	데이터 관련 의견
기술		조회 서비스					등을 통해 시도별, 시군별, 년도별 품목별 과수 재배 현황(재배면적, 면적비율, 면적 순위) 정보를 제공하는 서비스			
	IoT	스마트팜 데이터 조회 서비스	농림수산식품 교육문화정보원	파일	연도별	2018-2021	농촌진흥청 내 "농업 빅데이터 수집 및 생산성향상모델개발" 사업에서 수집한 시설채소 농가 데이터	시군구별, 온실종류, 온실유형, 품목, 품종, 총면적, 식부면적, 재식밀도, 정식일, 출하일자, 출하량		
	R&D	생산시스템 R&D 특허정보	농림식품기술 기획평가원	파일	수시		농림식품 연구개발 과제를 통해 창출된 성과 정보를 공공데이터로 제공	특허명, 출원등록인, 출원등록년도, 출원등록국가		
	가축질병	가축질병발생정보	농림축산검역본부	Open API	수시					
	가축질병	동물질병정보	농림축산검역본부	파일	수시	2015	유전자원은행, 동물질병정보, 동물질병 등 동물질병정보	질병명, 학명, 등록자명, 등록일자, 대상 동물, 질병 분류		
	경영	농업경영체 지역별 법인 현황	농식품부	파일	수시		농림축산식품부 농업경영체 기준 지역별 법인 현황정보	지역별, 법인형태, 영농조합법인, 농업회사법인 수		
	경영	농축산물소득정보	농진청	Open API	연도별	-2013				
	농자재	농기계별 면세유 공급현황	농식품부	파일	연도별	2013-2014	면세유의 농기계별 공급현황 자료	지역본부, 시지부, 배정실적, 농기계 기종, 배정량		
	농자재	지역별 농기계보유 현황	농식품부	Open API						

○ 국내 농업 공공 데이터 라이프 싸이클 상의 문제점

① 현재 공공 데이터는 크게 4가지의 형태로 제공하고 있으나, 각 데이터의 접근성이 떨어짐
- 데이터의 형태는 그리드, OPEN-API, 파일, 링크 형태로 제공됨
- 파일 제공의 형태는 단절된 데이터의 형태를 제공하며, 일부 자료는 통계청을 이용하여 연속데이터 확인 가능
- OPEN-API 자료는 자료마다 제공 형태가 다르고, 공공 데이터 포털의 경우, OPEN-API 신청 후 승인까지 약 2시간 소요되고, 농식품부 및 농진청 OPEN-API의 경우, 회원 가입 및 신청 후 승인까지 2시간 소요됨
- 승인 후 데이터를 확인하더라도 원하는 형태의 데이터가 아닌 경우도 존재
- 데이터 거래를 위해서는 데이터의 형태나 구성 등 세부 항목에 대한 자세한 설명이 제공될 필요 있음

② 데이터 제공 시 변수명에 대한 설명 부족 및 단위 미표기로 사용자의 편의성을 고려하지 않음
- 변수명이나 단위에 대한 설명이 부족하여 데이터가 의미하는 바를 확인하는 추가적 시간투자 필요

③ 현재 분류된 공공 데이터 분류 체계는 복잡하고, 중복되며, 잘못된 분류가 포함되어 있어, 데이터 거래소 구축 시 데이터 분류 체계를 정리할 필요가 있음
- 현재 공공데이터 분류체계에서 축산 분류 내에 화훼나 농산물 등의 데이터 목록이 포함됨

④ 현재 시점 기준 제공(지속적인 업데이트)하는 자료는 제한적으로 존재하여 데이터로서의 가치는 불투명한 수준임
- 상당수의 제공 데이터가 과거에 사용하거나 조사된 자료이고, 시계열 측면에서 연속적 형태가 아니기 때문에 데이터로서의 가치가 낮음
- 그러므로 데이터 거래를 통한 사회적 가치 향상을 위해서는 연속적, 지속적인 데이터 제공이 필요하고, 최신 자료와 연계 필요

⑤ 개인정보 보호법 적용에도 불구하고 상호명, 전화번호, 주소 등 개인정보가 포함된 데이터도 일부 존재
- 데이터 거래 시 「개인정보 보호법」 위반 가능성 확인 필요
- 만약, 공공데이터의 상기 정보가 적법한 수준에서의 공개라면 현행 제공하는 데이터는 보다 세밀한 데이터로 제공 가능

⑥ 제공 기관이 동일하더라도 데이터 라이프 싸이클에 대한 정보는 상이한 형태로 제공됨
- 데이터 명칭, 제공기관, 제공형태, 조사주기, 조사기간, 데이터설명, 조사항목, 조사방법 등 제공기관에 따라 차이가 존재함
- 데이터 거래를 위해서는 라이프 싸이클의 형식을 정립하고, 데이터 판매자들은 데이터 거래를 위해 반드시 형식에 맞춰 작성할 것을 요구해야 함

⑦ 데이터 거래 적용 시 Open API 형태로 제공될 가능성이 크므로 이용자(구매자)의 데이터 접근의 어려움 존재
- 민간 및 공공기관의 데이터는 대부분 Open API 형태로 제공하고, 점차 확대하고 있음
- 데이터 구매자는 Open API 데이터를 확인할 수 있는 교육 과정 필요

(2) 시장의 정의와 데이터 거래 시장 이슈

○ 시장은 수요와 공급이 만나는 공간

- 시장은 가격 결정, 상품의 매매(상품수집, 거래 장소 제공, 소유권 이전, 상품배분 등), 정보의 수집 · 가공 · 분산 역할 수행

○ 시장의 정의

1) 지리적 시장(geographical market)

- 구매자와 판매자가 만나서 사회가 정한 일정한 조건에서 재화와 용역의 소유권을 이전하기 위한 교환 비율, 즉 가격을 결정하는 곳

2) 일물일가의 법칙(Law of one price)이 성립하는 공간

- 시장에는 매우 기민하고 다른 참여자들의 일을 잘 알고 있는 수많은 구매자와 판매자가 존재하므로 지리적으로 분리되었다 하더라도 같은 상품을 판매하는 경우 품질이나 거래비용 등의 차이를 제외한 가격이 항상 거의 같음
- 일물일가의 성립 여부는 공간적으로 분리되었으나 같은 상품을 판매하는 시장을 하나의 시장으로 규정할 수 있는지를 판단하는 기준으로 활용되고, 시장효율성 관점에서 시장이 효율적인지를 평가하는 기준으로 활용

3) 비즈니스 공간

- 기업 또는 경영체가 '무엇을 생산할 것인가? 얼마만큼 생산할 것인가? 어떻게 생산할 것인가? 생산품을 어떻게 분배할 것인가?' 등의 다양한 경제적 질문에 대응한 비즈니스 행위를 조직하고, 촉진하는 무대로 위치, 상품, 시점(시간), 소비자 그룹, 마케팅 시스템의 수준 등으로 정의됨
- 비즈니스 공간으로서의 시장에서 경영 성과를 높이기 위해서는 경쟁기업 또는 경쟁상품 확인과 자사의 경쟁력 수준 평가가 중요

4) 사회 후생에 영향을 미치는 경쟁 공간

- 기업 또는 상품의 경쟁 공간으로서 상품시장의 경계는 특정 상품과 그 상품의 대체재 사이의 합리적 교환가능성(reasonable interchangeability) 또는 수요의 교차탄력성(cross elasticity of demand)에 의해 결정
- 시장실패 방지와 사회 후생 증가를 위해서는 시장 주체, 특히 공급자 사이의 경쟁 촉진이 중요한데, 시장의 경쟁 수준 또는 공급자의 시장지배력 수준 평가를 위해 명확한 시장 정의 또는 범위 설정 중요

○ 시장 정의 또는 범위 설정 시 고려사항

- 수요 : 소비자의 최종 용도, 유통채널, 진열장소, 가격의 움직임 등
- 공급 : 원자재의 사용, 디자인, 포장, 진입과 퇴출 형태, 생산기술의 유사성, 가격 결정 방식, 판촉방법, 광고 등
- 지리적 거리 : 소비자의 쇼핑거리, 공급자의 유통거리

○ 시장의 정의 또는 범위 설정의 중요성

- 시장은 경제 분석의 기본 틀이지만, 명확한 범위를 두고 사용되지 않는 경우가 많음
- 시장의 정의 또는 범위 설정은 시장점유율 산정, 시장집중도의 추정, 경제 분석 모형의 범위를 결정하는 데 필수적 ex) 기업합병, 반독과점법(공정거래법), 판매촉진
- 시장의 범위는 시간, 장소, 형태에 따른 상호교환성의 정도에 따라 결정되고, 상호교환성은 주로

가격의 교차탄력성에 의해 측정
- 경제 분석에서 가장 기본적인 시장구조는 완전경쟁시장 : 현실적으로 많은 시장이 완전경쟁의 구조에서 벗어나 있지만, 완전경쟁시장에서 출발하여 완전경쟁시장의 가정과 조건을 완화함으로써 불완전경쟁시장을 보다 쉽게 분석할 수 있고, 불완전경쟁의 영향을 완전경쟁시장의 결과와 비교함으로써 분석할 수 있음

○ 데이터 거래 시장의 특성

○ 4차 산업의 확산, AI의 발전, 머신러닝, 딥러닝 등 빅데이터 활용의 다양한 수요가 증가하면서, 기업들은 해당 기술의 기본이 되는 데이터를 내부 데이터 뿐만 아니라 외부의 공공이나 일반 기업의 데이터를 보유하려 하고 있음(김재형 외, 2020)

○ 데이터 거래란 공급자(판매자)와 수요자(구매자) 사이에 온・오프라인 방식으로 데이터를 전송・사용・이전하는 행위를 의미(한국데이터산업진흥원, 2022)
- 데이터가 저작권법 등 법적 보호를 받는 경우, 구매자는 판매자로부터 적법하게 이용할 수 있도록 허락을 받아야 함(김재형 외, 2020)
- 데이터 판매자(데이터 권리자)란 데이터를 제작한 제작자, 제작자로부터 권리를 양수한 자, 데이터에 대한 권리를 가지고 있는자로부터 판매에 대한 대리권을 수여받은 자 등을 의미(김재형 외, 2020)

< 데이터 거래 유형 >

구분	거래유형
File	데이터 셋으로 구성된 TXT, CSV, XLS 등의 파일
API	Application Programming Interface의 약자로 JSON, RSS, ASP 방식의 API
IMAGE	지도, 사진, 예술 등 이미지 데이터를 디지털화한 상품

출처 : 김재형 외, 2020

○ 데이터의 거래는 크게 양도 계약과 라이선스 계약의 형태로 구분됨(김재형 외, 2020)
- 양도 계약이란 지적 재산권이 이전되어 권리 주체가 변경되는 방법으로 양도 계약 이후에는 판매자의 권리 행사 불가
- 라이선스 계약이란 데이터의 권리를 가진 자가 독점적 또는 비독점적으로 이용을 허락하고 이에 상응하는 이용료를 지급
- 실제 데이터 거래에서는 양도 계약보다 이용 권한을 제공하는 라이선스 계약이 더 많이 체결됨

○ 데이터 거래 계약 시 고려사항은 다음과 같음(김재형 외, 2020)
- 계약 대상의 범위 설정
- 데이터의 이용목적 및 방법
- 데이터의 이용허락 권리 범위
- 데이터의 제공방법
- 데이터 이용과금 및 정산
- 침해예방 설정
- 계약 해지

○ 데이터 거래 시장은 신규시장으로 관련 정책 중요

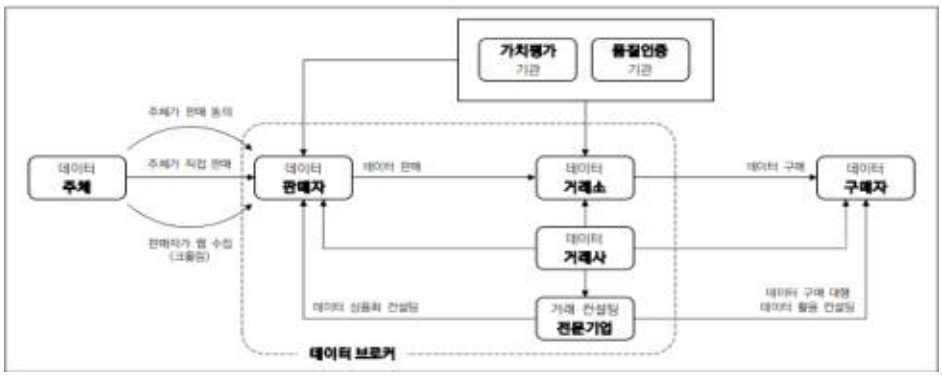
- 최근 과학기술정보통신부의 데이터 바우처 지원사업은 안전한 데이터 거래가 이루어질 수 있도록 바우처 형식의 데이터를 활용해 수요기업과 공급기업의 중간다리 역할을 하고 있음(한국데이터산업진흥원, 2022)
- 데이터 거래 관련 법률은 다음과 같음

< 데이터 거래 관련 법률 동향 >

▪ 2011년 제정·시행된 「개인정보 보호법」으로 엄격한 사전동의(opt-in) 방식을 적용하여 데이터 활용 위축 ▪ 2013년 「공공데이터의 제공 및 이용 활성화에 관한 법률」을 제정하여 공공데이터 개방의 틀을 마련 ▪ 2017년 인공지능 학습용 데이터를 대량으로 구축하는 '데이터 댐' 사업 추진 ▪ 2020년 소위 '데이터 3법(개인정보보호법, 정보통신망 이용촉진 및 정보보호 등에 관한 법률, 신용정보의 이용 및 보호에 관한 법률)'을 개정하여 개인정보와 연관된 데이터의 활용 가능성을 높임 ▪ 2022년 4월 22일 데이터 거래 활성화를 목적하는 「데이터 산업 진흥 및 이용 촉진에 관한 법률」이 시행되었음

○ 데이터 거래 체계는 주체 - 판매자 - 거래소(거래사) - 구매자로 구성됨

- 데이터 주체는 데이터를 최초로 만든 사람·기관·단체 등을 말하며 「개인정보 보호법」 제2조 제3호의 정보주체, 공공데이터를 생성하는 공공기관, SNS에 게시물을 올리는 개인, 사물데이터를 만들어 내는 사업자 등이 포함됨(정준화 외, 2022)
- 데이터 거래소의 형태는 하나의 판매자가 보유한 데이터만 판매하는 형태도 있고, 여러 판매자가 자유롭게 거래에 참여할 수 있도록 하는 형태도 존재함
- 데이터 브로커는 판매자·거래소·거래사가 결합되어 있는 형태로서 여러 기업·기관으로부터 데이터를 수입하여 분석·가공한 다음 구매자에게 판매하는 역할을 수행



(데이터 거래 체계)

자료 : 4차산업혁명위원회, 정준화 외(2022) 재인용

(3) 데이터 거래 규모

○ 전 세계의 데이터 총량은 2010년 기준 2ZB에서 2020년 64.2ZB로 32배 성장하였으며, 2024년에는 181ZB가 될 전망(스테티스타, 2022)

- 외국의 데이터 브로커 시장 규모는 2019년 기준 2,326억 달러에서 연평균 5.8%씩 성장하고 있으며, 세계 최대의 미국 시장 규모는 2018년 기준 약 1,832억 달러(약 220조원) 수준임(정준화 외, 2022)
- 중국도 정부 주도의 데이터 거래소를 운영하고 있으며, 구이저우성 구이양시에 빅데이터 거래소를 설립하여 데이터 산업의 빠른 성장을 촉진하고, 기술・안전・가격 등을 표준화하였음. 다만 정부 데이터가 80% 이상으로 민간의 참여는 높지 않음
- 한국의 경우, '데이터 판매 및 중개 서비스업'을 기준으로 2020년 기준 약 1조 6천억원 수준임
  - 데이터 산업의 전체 시장 규모는 약 20조 원 정도로 추정되지만, 이 중 데이터 거래가 차지하는 비중은 약 8% 수준임

구분	2019년		2020년		매출액 증감율
	매출액	비중	매출액	비중	
데이터산업 전체	168,582	100.0	200,024	100.0	18.7
데이터 처리 및 관리 솔루션 개발·공급업	20,805	12.3	25,133	12.6	20.8
데이터 구축 및 컨설팅 서비스업	65,412	38.8	76,999	38.5	17.7
데이터 판매 및 제공 서비스업	82,364	48.9	97,891	48.9	18.9
정보제공 서비스업[1]	71,033	42.1	81,838	40.9	15.2
데이터 판매 및 중개 서비스업[2]	**11,332**	**6.7**	**16,054**	**8.0**	**41.7**

(데이터 산업 시장 규모)

출처 : 과학기술정보통신부(2022), 「2021 데이터산업 현황조사」

○ 국내에서 운영중인 데이터 거래소는 크게 공공거래소, 민관합동 거래소, 민간거래소 등이 존재함

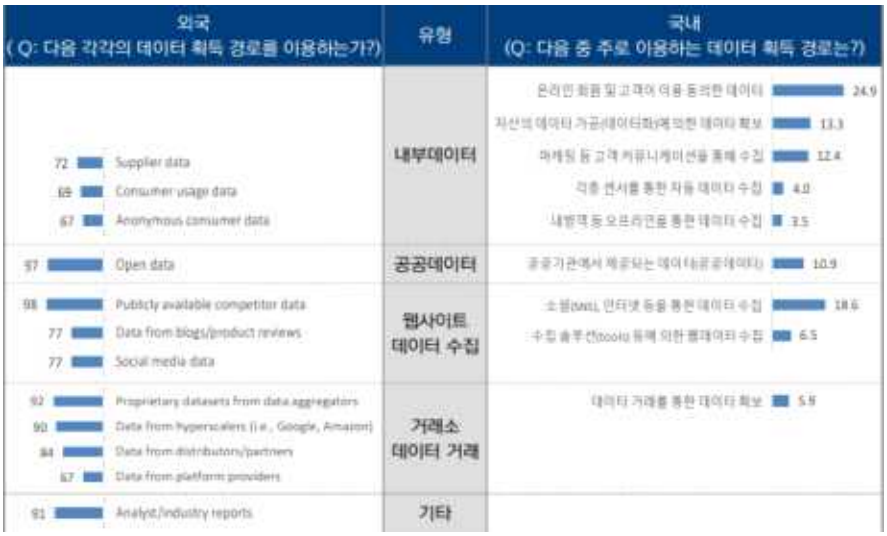
운영주체	데이터 거래소 사례
공공	• 공공데이터모델(행정안전부/한국지능정보사회진흥원 운영, data.go.kr) • AI허브(과학기술정보통신부/한국지능정보사회진흥원 운영, aihub.or.kr) • 데이터스토어(과학기술정보통신부/한국데이터산업진흥원 운영, datastore.or.kr) • 금융데이터거래소(금융위원회/금융보안원 운영, findatamall.or.kr) 등 다양함
민관합동 (정부지원 민간운영)	• 문화 빅데이터 플랫폼(주관: 한국문화정보원) • 통신 빅데이터 플랫폼(주관 : 케이티) • 유통 빅데이터 플랫폼(주관 : 매일방송) • 헬스케어 빅데이터 플랫폼(주관 : 국립암센터) • 교통 빅데이터 플랫폼(주관 : 한국교통연구원) • 환경 빅데이터 플랫폼(주관 : 한국수자원공사) • 금융 빅데이터 플랫폼(주관 : 비씨카드) • 중소기업 빅데이터 플랫폼(주관 : 더존비즈온) • 지역경제 빅데이터 플랫폼(주관 : 경기도청) • 산림 빅데이터 플랫폼(주관 : 한국임업진흥원) • 소방안전 빅데이터 플랫폼(주관 : 소방청) • 스마트치안 빅데이터 플랫폼(주관 : 경찰대학 치안정책연구소) • 해양수산 빅데이터 플랫폼(주관 : 한국해양수산개발원) • 농식품 빅데이터 플랫폼(주관 : 한국농수산식품유통공사) • 라이프로그 빅데이터 플랫폼(주관 : 원주연세의료원) • 디지털산업혁신 빅데이터 플랫폼(주관 : 한국산업기술시험원)
민간	• SK C&C 데이터 큐레이션 플랫폼 AccuRator(아큐레이터) (skdt.co.kr/accurator) • KT 빅사이트 (bigsight.kt.com) • 빅데이터 마트 (bigdatamart.co.kr) 등 다양함

(국내 주요 데이터 거래소 운영주체 및 사례)

출처 : 정준화 외(2022)

○ 데이터 거래 시장 규모

- 외국의 경우 내부 데이터만을 이용하는 경우보다는 닐슨 등 데이터 수집・분석 전문기업이나 구글, 아마존과 같은 IT 대기업을 통한 데이터 거래 이용률이 90% 이상으로 매우 높음(정준화 외, 2022)
- 반면, 국내의 경우, 기업이 보유한 내부데이터를 이용하는 비중이 매우 높으며, 데이터의 수집 경로는 '온라인 회원 및 고객이 이용 동의한 데이터(24.9%)', '자산의 데이터 가공에 의한 데이터 확보(13.3%)', '마케팅 등 고객 커뮤니케이션을 통해 수집(12.4%)' 등임

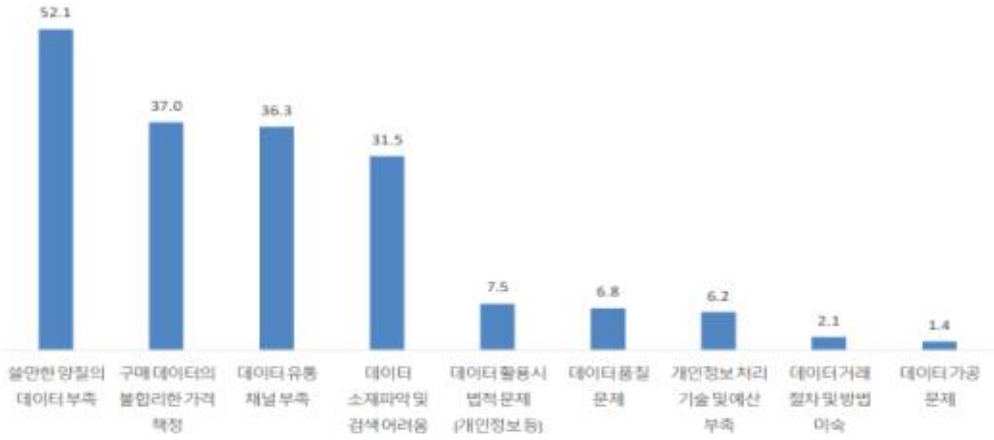


(국내외 데이터 획득 경로 비교)

출처 : 정준화 외(2022)

○ 국내 데이터 거래의 애로사항

- 데이터 내용의 불충분과 데이터 거래 환경의 미성숙으로 데이터 거래가 낮음(정준화 외, 2022) : 쓸만한 양질의 데이터 부족(52.1%) > 구매 데이터의 불합리한 가격 책정(37.0%) > 데이터 유통채널의 부족(36.3%) > 데이터 소재 파악 및 검색의 어려움(31.5%) 순임(과학기술정보통신부(2022)



(데이터 거래 시 애로사항)

자료 : 과학기술정보통신부(2022), 「2021 데이터산업 현황조사」

○ 데이터 거래에 관한 정책 거버넌스의 문제

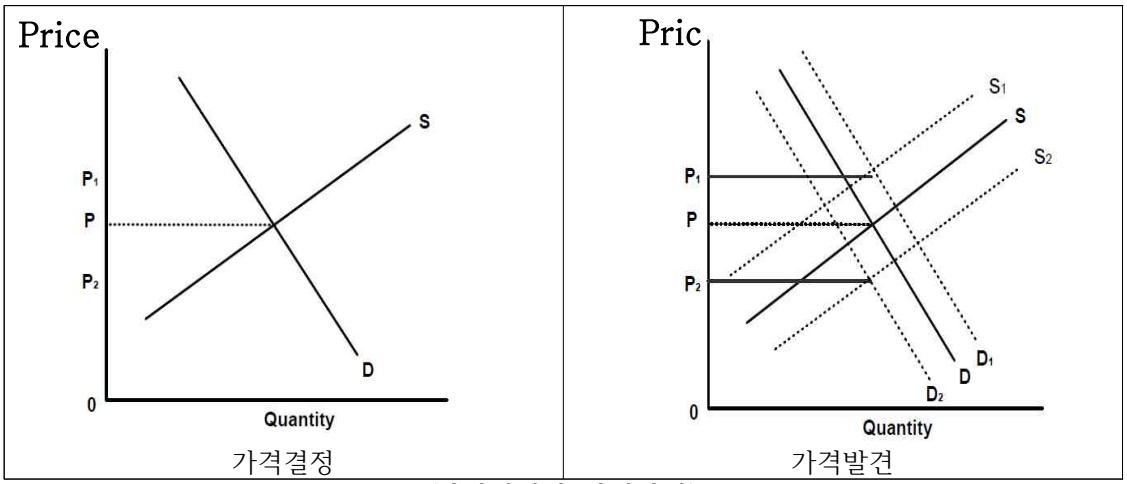
- 국내 데이터 거래에는 다양한 소관 부처가 연결되어 있으나, 이를 종합적으로 조정하는 장치가 없어 효과적인 관리가 어려움
- 16대 빅데이터 플랫폼과 데이터 스토어는 데이터 산업법 제6조의 국가데이터정책위원회 심의사항임
- 공공데이터포털은 「공공데이터의 제공 및 이용 활성화에 관한 법률」 제5조의 공공데이터 전략위원회 소관임
- 개인정보와 관련된 데이터의 경우, 「개인정보 보호법」 제7조의 개인정보보호위원회가 담당
- 금융데이터는 「신용정보의 이용 및 보호에 관한 법률」에 따라 금융위원회 소관 사항임

(4) 데이터 거래 시장 이슈

○ 이슈 1. 데이터 거래 시장 정의 또는 범위 설정
- 효율적 가격 설정을 통한 가격발견 기능 촉진, 거래 활성화, 독과점 방지 및 경쟁 촉진 등을 위해 시장 정의 또는 범위 설정 중요
- 현행은 농식품 데이터 거래 시장은 초기 시장으로 플랫폼 단위로 운영, 관리되고 있으나, 각 플랫폼에서 거래되는 정보의 상이성을 고려할 때 적절한 시장 정의 또는 범위 설정의 기준 수립 중요

○ 이슈 2. 데이터 거래 활성화
- 규제 완화와 홍보 등을 통한 구매자 및 판매자 유치 촉진
- 데이터 품질 관리와 등급표준화
- 시장정보 제공

○ 이슈 3. 데이터 거래 시장의 가격발견 기능 촉진
- 가격결정 : 시장 수요와 공급의 상호 작용을 통해 시장의 가격 수준이 형성되는 과정(Ward and Schroeder, 2002)
- 가격발견 : 구매자와 판매자가 특정한 품질의 상품을 특정 시간과 장소에서 사고 파는 거래가격에 도달하는 과정(Ward and Schroeder, 2002)



가격결정 가격발견

(가격결정과 가격발견)

자료 : Ward and Schroeder(2002) 재인용

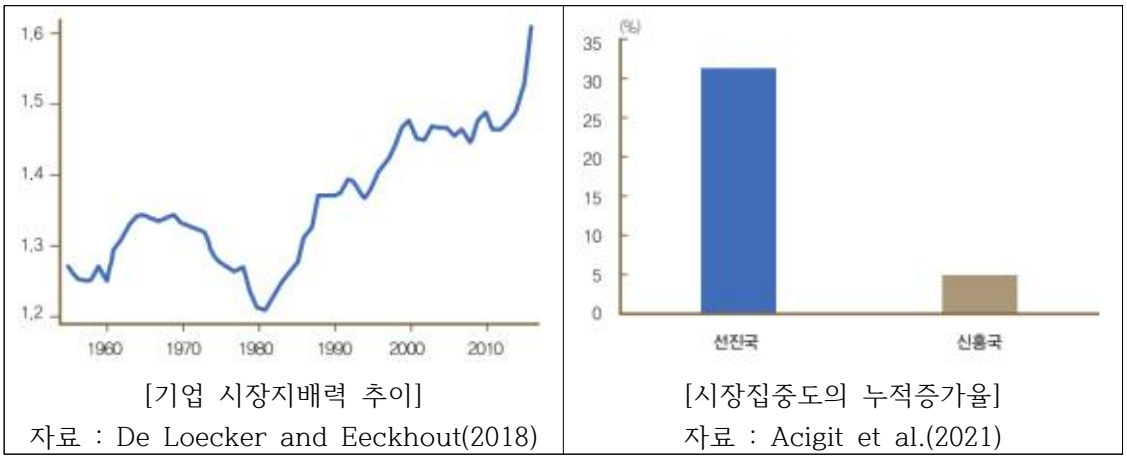
- 시장 균형가격에 대한 정보 불확실성(information uncertainty)으로 인해 구매자와 판매자가 불확실한 기대(uncertain expectation)를 바탕으로 거래가격을 발견하기 때문에 시장의 균형가격과 개별 거래자의 거래가격 사이에 격차 발생
- 가격발견에 영향을 미치는 요인으로 시장구조(참여자 수, 시장 크기와 규모, 위치, 경쟁 수준 등), 시장 행동(구매자와 판매자의 조달 및 가격책정 방식 등), 시장정보와 가격 보고(정보의 양, 전달 시점, 신뢰성 등), 시장 거래 주체의 가격위험 관리 수단 등이 있음

- 정보 불확실성이 증가하는 경우 개별 거래자의 가격발견 비용이 증가하는데, 데이터 거래 시장의 경우 시장 정의 또는 범위 설정이 미흡한 상황에서 거래 미흡, 시장정보와 가격 보고 미흡 등으로 가격발견 비용이 적지 않으리라 판단

○ 이슈 4. 데이터 거래 시장의 독과점 방지와 경쟁 촉진

- 국내 데이터 거래 시장이 초기라는 점을 고려할 때 독과점 논의가 이를 수 있으나, 최근 빅테크 중심의 새로운 독과점 시장구조에 대한 우려가 높아지고 있어(정선영, 2022) 대응 방안 필요
  · 국내에서 카카오모빌리티 같은 기업은 서비스 이용자와 플랫폼 노동자를 연결시켜 수익을 얻는 동시에 사업 과정에서 얻는 데이터를 회사 매출 증대에 활용하지만, 노동자나 소비자에겐 그 데이터에 대한 접근이 허용되지 않음(선담은 기자, 한겨레, 2021.10.23.)

- 1980년대 이후 미국 등 주요국에서 시장집중 현상이 심화되는 양상을 보이고, 신흥국보다 선진국에서 시장집중도 더 큰 폭으로 상승 (De Loecker and Eeckhout, 2018; Akcigit et al., 2021; 정선영, 2022 재인용)



[기업 시장지배력 추이]
자료 : De Loecker and Eeckhout(2018)

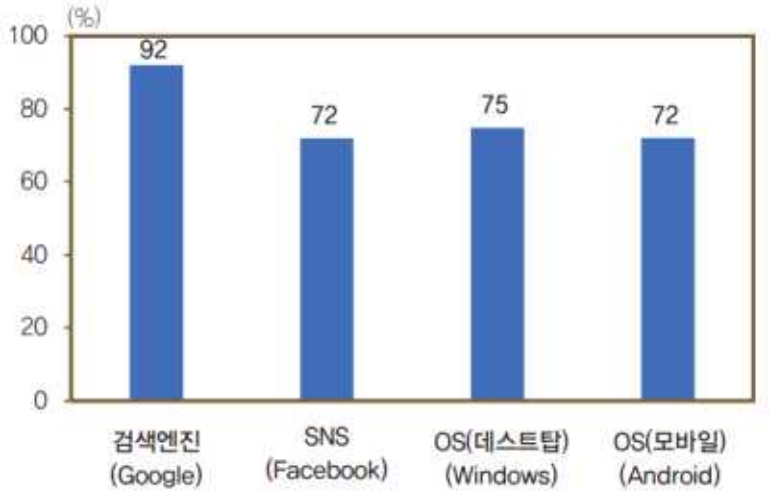
[시장집중도의 누적증가율]
자료 : Acigit et al.(2021)

주1 : 기업 시장지배력은 평균 마크업으로 추정
주2 : 시장집중도는 매출액 기준 상위 20개 기업 중 상위 4개 기업이 차지하는 비중인 CR4를 기준으로 시장집중도 정의
주3 : 시장집중도에서 선진국은 1980~2016년, 신흥국은 1992~2016년 기간 동안 시장집중도의 누적증가를 의미하고, 국가 내 산업평균과 국가 간 중앙값을 취함

그림 (기업 시장지배력 추이 및 시장집중도의 누적증가율)

자료 : 정선영(2022, p.2) 수정후 재인용

- 2010년대 이후부터는 거대 IT기업인 빅테크가 시장을 승자독식하는 등 과거와 다른 새로운 유형의 독과점 시장구조 형성(홍지연, 2021; 정선영, 2022, 재인용)



(주요 빅테크의 글로벌 시장점유율(2021년 기준))

자료 : Statcounter GlobalStats, 정선영(2022) 재인용

- 데이터 경제에서는 ① 승자독식의 시장, ② 집중화된 시장, ③ 게이트키퍼로서의 온라인 플랫폼 등으로 독과점적 시장구조가 발생할 우려가 있고, ① 혁신 및 기업가 정신의 약화, ② 프라이버시 및 데이터 보호의 저하, ③ 자유롭고 다양한 언론의 쇠퇴, ④ 독과점화된 플랫폼 기업의 경제적, 정치적 영향력 확대 등에 영향을 미칠 수 있음(유영상, 2020)
- 데이터 관련 경쟁 제한 우려 사항은 다음과 같음(Autorite' de la concurrence and Bundeskartellamt, 2016; 유영상, 2020, 재인용)

< 데이터 관련 경쟁 제한 우려 사항 >

구분	내용
진입장벽	진입장벽은 다음과 같은 요인에 의해 발생 가능 - 시장지배력의 원천으로서의 데이터, 네트워크 효과 - 전환비용, 규모 및 범위의 경제
인수와 합병	다음과 같은 경우의 인수와 합병은 경쟁 제한 가능성이 존재 - 합병 대상 신규사업자가 다른 시장의 대규모 데이터베이스 보유 - 데이터 조합을 통해 추출된 정보를 경쟁사가 복제 불가능 - 별도의 수직 시장에서 이미 강력한 시장 지위를 보유하고 있는 두 회사의 합병
배제적 행위	경쟁자의 데이터 접근을 박탈함으로써 경쟁자를 배제할 수 있는 행위 - 필수 설비에 해당하는 데이터에 대한 접근 거부, 차별적 접근 허용 - 배타적 계약, 데이터와의 결합 판매 및 교차 사용
가격 차별 행위	고객 데이터 수집을 통한 가격 차별이 다음에 해당하는 경우 - 시장지배적 지위의 남용 - 수직적 거래제한
개인 정보 보호 경쟁의 제한	지배적 사업자가 데이터 기반의 시장지배력을 이용하여 개인 정보 보호에 관한 경쟁을 제한하는 경우

자료 : Autorite' de la concurrence and Bundeskartellamt(2016), 유영상(2020) 수정 후 재인용

- 다양한 분야에서 독과점적 지위를 구축하고 있는 빅테크(Big 5기준)는 온라인 플랫폼 및 데이터 중심 비즈니스라는 공통점을 가지는데, 이들 기업은 전통 독과점 기업과 차이를 보이기 때문에 데이터 거래 시장에 부합하는 공정경쟁 유도 필요

< 기존 경제와 디지털 경제의 독과점 기업 특성 비교 >

구분	전통 독과점 기업	신생 독과점기업(빅테크)
산업	철도 등 기간산업	정보통신서비스업
시장구조	단면 시장을 가진 단방향 구조	양면 시장을 가진 플랫폼 구조
생산함수 형태	수확체감	수확체증
효율성 개선방법	공급 측 규모의 경제	수요 및 공급 측 규모의 경제
중요 경쟁력 지표	가동률, 생산단가, 시간상 생산성	MAU(월간 활동사용자), 체류시간, 재방문률, LTV(고객생애가치)
소비자에 대한 시각	소비자 규모(stock)에 초점	규모뿐 아니라 선호, 활동량 등 소비자 정보(flow)에 초점
단기 기업 목표	기업의 이익 극대화	기업의 점유율 극대화
경쟁 기업과의 관계	경쟁 관계	경쟁 및 종속의 이중적 관계
확장성	타 부문으로의 확장 제약	타 부문으로의 유영한 확장

자료 : 정선영(2022) 재인용

(5) 데이터 가치 및 가격의 정의와 영향요인

○ 가치(value)의 정의와 특성

- 가치 : 본인을 포함한 누군가에게 의미 있거나 만족감을 준다고 판단되는 성질로 만족감을 제공하는 수단은 물질적, 정신적 수단을 모두 포함
- 특성 1 : 주관적으로 평가
- 특성 2 : 관점에 따라 다양한 차원의 가치 존재하는데, 주체를 기준으로 국가, 지역사회 또는 공동체, 기업, 개인 등이 있고, 영역에 따라 정치, 군사, 경제, 문화, 사회, 과학, 기술, 종교 등으로 구분할 수 있으며, 가치의 충돌이 발생할 경우 어떠한 가치가 우월한지 판단이 어려움
- 특성 3 : 경제적 가치의 경우 시간, 장소, 상황 등에 따라 변화 ex) 사막과 오아시스의 물의 경제적 가치 다름
- 특성 4 : 경제적 가치의 경우 사회적, 역사적 산물 ex) 조선시대와 지금의 쌀의 가치 다름

○ 상품의 경제적 가치(economic value)

- 상품의 (경제적) 가치는 사용가치와 교환가치로 구분(Marx, 1867; 김수행, 2009, 재인용)
- 사용가치 : 특정 물품의 소비로 인해 발생하는 효용
- 교환가치 : 그 물품의 소유가 가져다주는 다른 재화에 대한 구매력
- 사용가치로서 상품은 질적으로 구분되나, 교환가치로서의 상품은 오직 양적 차이를 가지는데, 이러한 차이로 인해 사용가치와 교환가치 사이에 괴리 발생. 예를 들어, 물의 경우 생명 유지의 필수상품이기 때문에 사용가치가 크나, 교환가치는 낮고, 다이아몬드의 경우 공업용 등으로 활용되고 있으나 사용가치에 비해 교환가치가 월등히 큼

○ 경제적 가치와 가격

- 가격은 만족감 또는 효용을 위해 거래되는 상품의 교환가치(value placed on what is exchanged for satisfaction or utility)를 의미
- 일반적으로 화폐단위로 표시되는 값을 '가격'으로 인식하나, 경제적 가치의 관점에서는 상품 간의 교환 비율을 의미하는 교환가치가 중요하고, 여러 가지 한계에도 시장가격은 교환가치를 대표하는 지표로 볼 수 있음

- 교환가치로서 가격의 원천에 대해 고전학파를 중심으로 한 노동가치론에서는 노동만이 부를 창출하는 유일한 자원으로 그 물품의 생산에 사용된 노동량에 의해 결정된다고 주장했는데, 이는 교환가치 또는 가격이 공급 비용(공급)에 의해 결정된다는 의미
- 반면, 한계학파가 한계효용가치론을 제시한 이후 시장가격으로 대표되는 교환가치는 사용가치와 희소성을 의미하는 한계효용(수요)과 공급 비용(공급)의 균형에 의해 결정된다는 이론이 정립

○ 가격의 종류

- 대상 자산 기준 : 재화는 가격, 서비스는 수수료(commission or fee) 또는 요금(fare), 자본은 이자(interest), 노동은 임금(wage), 건물 또는 토지는 임대료(rent)로 지칭
- 가격 결정 방식 기준 : 가격 결정 방식을 기준으로 ① 시장경쟁가격(market competitive price), ② 개인이나 기업 간 교섭이나 거래로 형성되는 교섭가격(bargaining price; 사적 교섭가격과 단체 교섭가격으로 구분), ③ 경매를 통해 결정되는 경매가격(auction price), ④ 원가에 일정 수준의 마진 또는 이윤을 더하는 방식의 공식(formula)을 통해 정하는 공식에 의한 가격(formula price), ⑤ 기업이 마케팅 활동의 하나로 책정하는 전략가격(strategic price), ⑥ 정부가 정책 목적 달성을 위해 책정하는 정책가격(policy) 등으로 구분할 수 있음
- 가격의 적정성 기준 : 생산원가에 적정 이윤을 합한 가격은 정상 또는 자연가격(normal or natural price), 거래 쌍방이 수용할 수 있는 합리적 가격은 공정가격(fair price), 기준연도의 구매력과 같은 구매력을 유지하는 가격은 패리티 가격(parity price) 등으로 지칭
- 본 연구에서 목표로 하는 ‘합리적 가격’ 또는 ‘적정 가격’은 거래 쌍방이 수용할 수 있는 합리적 가격을 의미하는 공정가격으로 설정할 필요 있음.
  - 정상 또는 자연가격 설정을 목표로 할 수 있으나, 이 경우 적정 이윤 설정이 어려움
  - 마케팅적 관점에서 가격에 영향을 미칠 수 있는 시장지배력을 가진 기업의 경우 수요함수와 수리적 최적화 모형(mathematical optimization model)을 이용해 매출이나 이윤 극대화를 달성할 수 있는 ‘최적가격(optimal price)’ 도출이 가능하나, 초기 시장인 데이터 거래 시장에서 시장지배력을 가진 공급자가 존재하기 어렵고, 정부가 추진하는 데이터 거래소의 공공성을 고려할 때 최적가격을 목표로 하는 방식 부적절

○ 데이터 (거래) 가격의 역할과 중요성

- 첫째, 효율적 데이터 자원배분의 기준
- 둘째, 데이터 공급자의 수익의 원천
- 셋째, 데이터 가격의 수준이나 변동성을 평가하여 데이터 거래 시장의 적정성을 판단할 수 있고, 데이터 거래 시장이 Fama(1970)에 제시한 시장효율성 가설(market efficiency hypothesis)에 맞게 효율적으로 운용되는지를 판단하는 지표로 활용 가능
- 넷째, 가격에 대한 인센티브나 패널티 정책은 데이터 거래 시장의 활성화나 불공정거래를 규제하는 상대적으로 저렴하고 효과적인 정책수단으로 활용할 수 있는데, 이때 가격 변화에 대한 공급자나 구매자의 반응 중요
- 다섯째, 가격은 강력한 마케팅 도구로서 4P 전략 중 상품, 유통, 홍보 전략 등 다른 전략에 비해 저렴하고, 즉각적이며 가장 커다란 효과를 발생시킬 수 있음

(6) 데이터 가격에 영향을 미치는 요인

○ 데이터 가격 영향요인에 관한 선행연구

- 한국데이터산업진흥원(2020) : 데이터 가격의 영향요인을 원가, 고객, 경쟁, 환경 등 4가지로 구분

① 원가 : 데이터 상품에서 원가란, 데이터 확보·가공·판매를 위한 시스템 구축비, 운영비 등의 투입 비용을 의미함. 다른 제품이나 서비스와 마찬가지로 데이터 가격 결정 시 원가는 중요한 요소로 작용하고, 가격은 원가와 최소이윤의 합을 의미하기 때문에, 원가는 판매가격의 최하한선을 결정하는 하는 기준이 될 수 있음. 품질이 좋은 데이터 상품을 낮은 원가에 구축하면 저가 책정이 가능하고, 이에 따라 판매와 이윤이 증가해 경쟁우위를 점할 수 있음

(원가요소 기반 가격책정 범위)
자료 : 한국데이터산업진흥원, 2020.

② 고객(구매자 요인) : 데이터에 대한 구매자의 인식과 수요, 구매자의 상황, 구입한 데이터를 통해 구매자가 얻는 편익과 가치, 구매자의 긴급성, 구매자의 충성도 등이 구매 요인이 될 수 있음. 구매자 니즈가 높고, 구매자의 상황이 긴급하며, 예산도 충분한 경우에 판매자는 판매가격을 높일 수 있음. 또한 시장에 대체할만한 데이터 상품이 없거나, 해당 데이터 상품의 품질이 매우 좋고 모수(표본 수를 의미한다고 판단됨. 연구자 주)가 월등히 많아 활용성이 높은 경우에 판매자의 가격 협상력은 높아질 수 있음

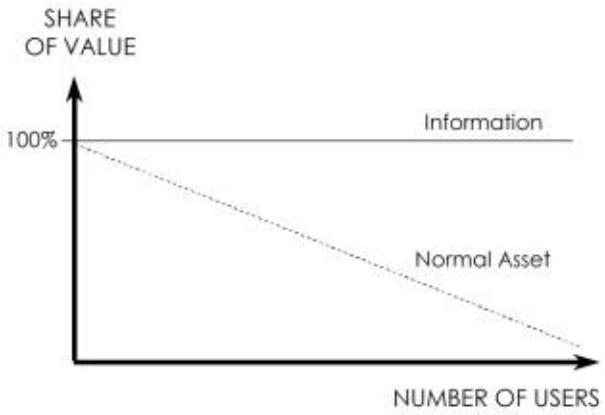
③ 경쟁 : 경쟁 요인은 크게 데이터 시장과 수요 요인과 경쟁자 요인으로 구분. 데이터 시장과 수요 요인 측면에서 시장의 유형(독점시장, 과점시장 등)이 판매가격에 영향을 미치고, 수요에 따른 가격탄력성도 가격책정의 중요 요인이 될 수 있음. 경쟁자 요인에서 경쟁사 비용, 가격, 반응 등도 데이터 가격 결정의 영향요인이 될 수 있고, 경쟁사가 유사한 데이터를 가진 경우, 경쟁사 대비 저렴한 가격을 설정하는 것 외에도 방법이 있는데, 데이터는 범위와 구성에 따라 활용범위가 천차만별이기 때문에 서비스와 정보를 차별화하고 별도의 가격 정책을 펼침으로써 경쟁우위를 확보할 수 있음

④ 환경 : 이자율, 인플레이션 등을 포함한 국내외 경제 수준, 데이터의 유통 경로, 데이터 거래환경과 사회적 인식의 수준, 정부의 정책과 제도 등이 가격 결정에 영향을 미칠 수 있음

- Moody and Walsh(1999) : 자산으로서 정보의 가치와 관련하여 7개 법칙을 제시했는데, 본 연구에서의 데이터는 정보를 포함하는 광의의 데이터로 설정한 상황에서 7개 법칙은 자산의 영향을 미치는 데이터의

특성으로 해석할 수 있음

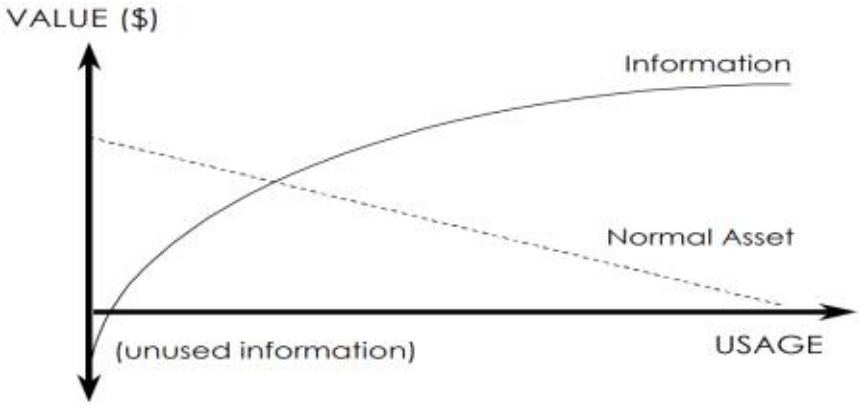
① 정보는 무제한 공유 가능(information is (infinitely) sharable) : 정보는 수많은 사람, 사업 지역, 그리고 조직이 거래하더라도 가치의 손실이 없고, 자산을 공유할 경우 공유자 수에 따라 가치가 감소하는 다른 자산과 달리 공유하더라도 가치의 감소가 없음. 또한 공유자가 늘어날수록 가치 증가



(정보 사용자의 수와 공유 가치)

자료 : Moody and Walsh(1999, p.4)

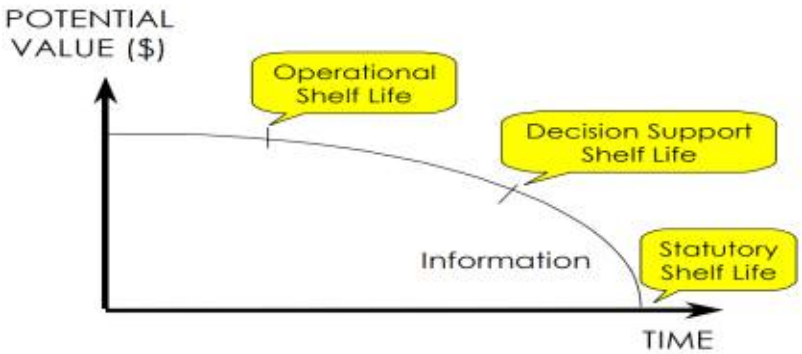
② 사용이 증가할수록 가치 증가(value of information increases with use) : 대부분 자산이 사용할수록 가치가 감소하나, 정보의 경우 대부분 비용이 수집(capture), 저장(storage), 유지(maintenace) 비용으로 한계사용비용 낮음. 관련하여 데이터는 사용하는 경우에만 가치가 발생함에도 많은 조직이 데이터를 수집, 저장한 이후에 전혀 사용하지 않는데, 데이터를 효과적으로 사용하기 위해서는 ① 데이터 존재 여부 알기(knowing it exists), ② 데이터가 어디에 존재하는지 알기(knowing where it is located), ③ 데이터 접근 수단 확보하기(having access to it), ④ 데이터 사용 방법 숙지(knowing how to use it) 등의 실천 필요. 또한, 정보가 의사결정 역량을 향상시키지만, 의사결정 역량 향상 수준은 제공되는 정보의 질과 의사결정자의 정보 활용 역량에 좌우되므로 정보 해석 능력(information literacy) 확보 중요



(데이터 사용량과 정보의 가치)

자료 : Moody and Walsh(1999, p.5)

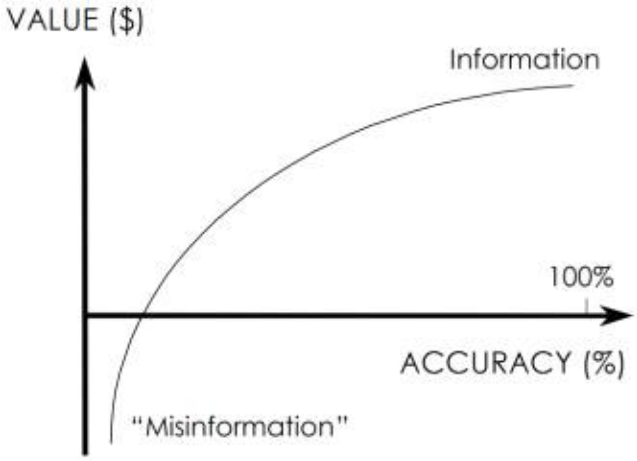
③ 시간 경과 시 정보의 가치 감소(information is perishable) : 다른 자산과 마찬가지로 정보의 가치도 시간이 흐를수록 감소하는데, 관련하여 정보는 운영 유통기한(operational shelf life), 의사 지원 유통기한(decision support shelf life), 법정 유통기한(statutory shelf life) 등 3가지 유형의 유통기한을 가짐. 운영 유통기한은 정보가 의사결정에 직접적 영향을 미칠 수 있는 기간, 의사 지원 유통기한은 의사결정 시 참고할 수 있는 기간, 법정 유통기한은 데이터의 특성상 개인정보 등 민감 정보가 포함되거나 조직의 내부 방침에 따라 의사결정에 도움을 주지 않음에도 법적으로 보유하는 기간으로 볼 수 있음



(시간과 정보의 잠재적 가치)

자료 : Moody and Walsh(1999, p.6)

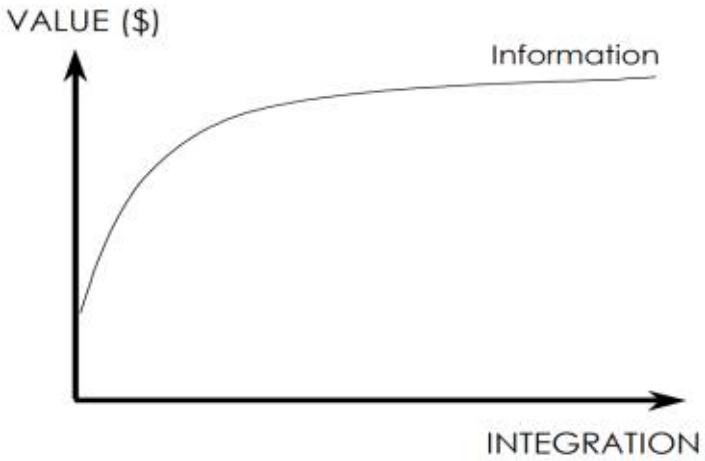
④ 정확도가 높을수록 가치 증가(value of information increases with accuracy) : 정보의 정확도가 높을수록 활용도가 커지므로 가치 증가하고, 정확도가 일정 수준 이상으로 떨어지는 경우 가치가 음(-)의 값이 될 수도 있음. 정보의 정확도 수준은 정보의 유형과 사용방식(또는 사용목적)에 따라 달라지는데, 예를 들어, 100% 정확도를 요구하는 데이터가 필요한 경우가 있으나, 사용 목적을 고려할 때 80% 정도의 정확도도 충분한 경우도 있음. 정보의 정확도 수준을 안다면, 의사결정 시 정확도 수준에 따른 손익을 고려할 수 있으나, 실제 데이터를 활용할 때 정보의 정확도는 의사결정자의 주관적 의견(subjective opinion)과 경험 증거(anecdotal evidence)에 의존하는 경우가 많음



(정보의 정확도와 가치)

자료 : Moody and Walsh(1999, p.7)

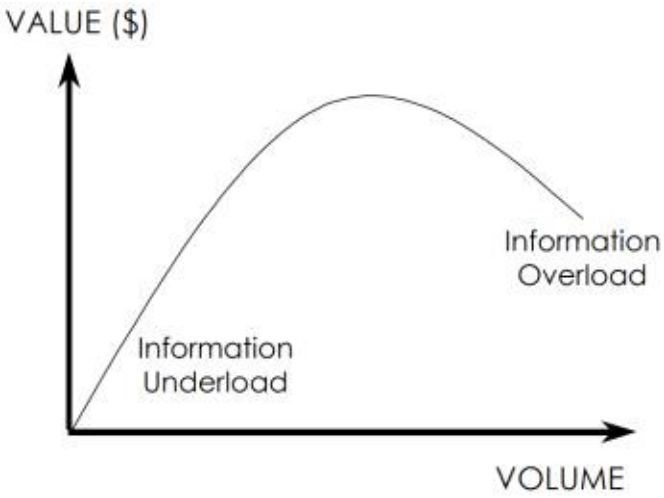
⑤ 다른 정보와 결합하는 경우 정보의 가치 증가(value of information increases when combined with other information) : 고객 정보와 판매정보를 결합하여 마케팅에 활용하는 사례처럼 정보를 결합하여 사용하는 경우 정보의 가치 증가. 정보 결합에 따른 혜택의 대부분은 전체 데이터의 일부 또는 작은 비율만을 표준화해도 확보할 수 있으므로 굳이 100% 정보 결합을 목표로 할 필요가 없고, 결합 비용과 편익을 고려한 효율적 결합 필요. 정보 결합 시 가장 중요한 데이터가 식별자(identifiers)와 코드체계(coding schemes)임



(정보 결합과 가치)

자료 : Moody and Walsh(1999, p.8)

⑥ 데이터양이 증가한다고 반드시 가치가 증가하지 않음(more is not necessarily better) : 대부분 자산은 자산 크기가 커질수록 편익이 커지나, 정보의 경우 일정 수준 이상으로 증가하는 경우 오히려 가치 감소. 특히, 정보통신기술의 발달로 정보의 부족(lack of information)보다 정보의 과잉(overabundance of information)이 더 문제가 되는 실정임



(정보의 양과 가치)

자료 : Moody and Walsh(1999, p.9)

⑦ 정보는 고갈되지 않음(information is not depletable) : 대부분 자산은 사용할수록 줄어드는데, 오히려 정보는 사용량이 늘어날수록 오히려 더 많은 혜택을 가지게 됨(information is self-generating- the more you use it, the more you have). 이는 새로운 또는 파생된 정보(new or derived information)가 다른 정보를 함께 활용하여 요약, 분석, 결합한 정보이기 때문임. 이 경우 기존 원본 정보(original information)는 그대로 유지되면서 파생정보가 자산으로 추가될 수 있음

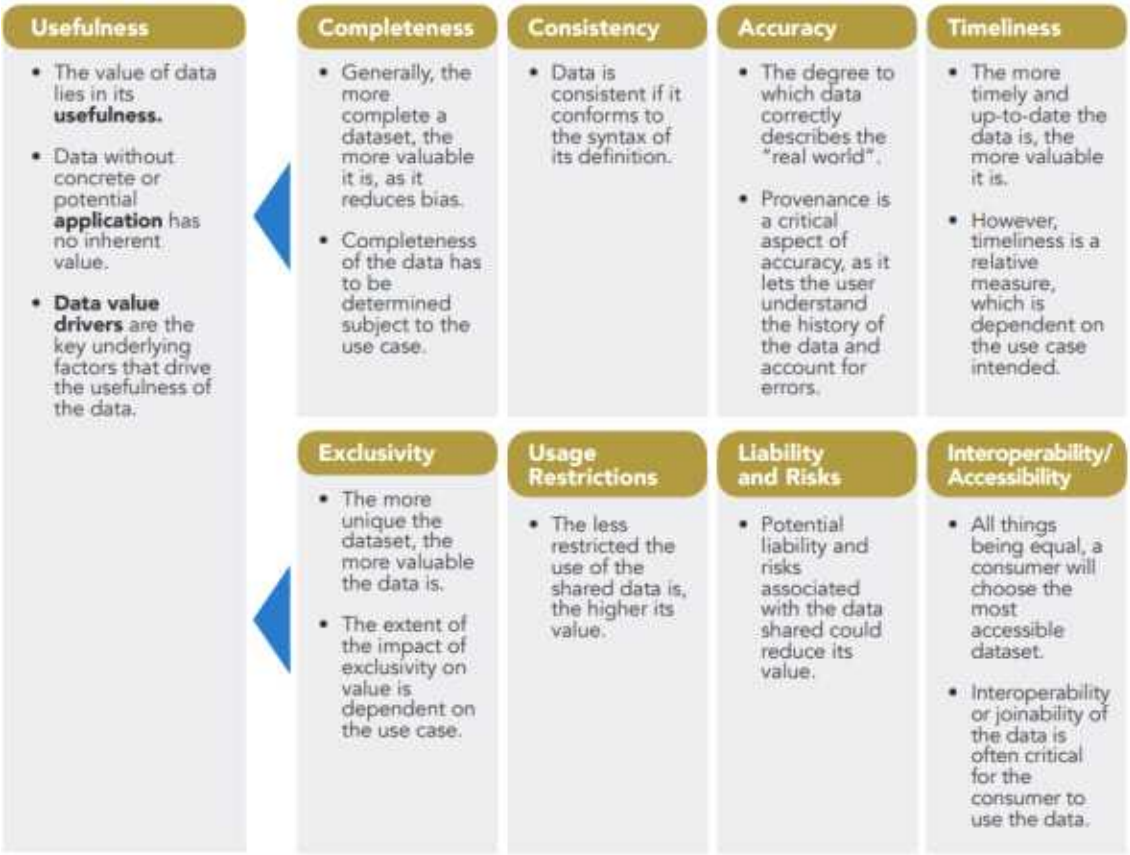
- 김옥기 외(2021) : 데이터 생명주기, 대체불가능성, 다른 데이터와의 지속적 융합 가능성, 위험관리(개인정보 보호, 보안 등), 데이터가 유의미하게 활용될 수 있는 기간, 데이터 품질 등을 데이터 가치 결정 요인으로 제시. 특히, 아무리 중요한 항목의 데이터를 많이 가지고 있더라도 품질이 좋지 않으면 활용 가치가 없어지기 때문에 데이터 품질이 가장 중요하다고 지적. 데이터 품질과 관련하여 Gartner(2015)의 조사결과에 따르면 비즈니스 프로젝트의 40%가 데이터 품질 때문에 실패
- Azocitia et al.(2021) : B2B 데이터 마켓플레이스(DMs)의 판매가격 데이터를 이용하여 데이터 가격에 연계된 데이터 품질 속성(features) 제시

< 데이터 가격에 영향을 미치는 데이터 품질 속성 >

Questions	Group	Definition	Number of features	Example of features
What?	Category	Labels attached to the product that define the type of data it contains	11	Weather, Gaming, Financial
	Descrip-tion	Stem-like features obtained from data product descriptions	up to 2000	wordmarket, wordientifi, wordlist
	Identifi-ability	Tells whether the product allows the buyer to recognize the activity of individuals or to identify specific companies	2	Idsession, Idcompanies
How much?	Volume	Normalized n^0 units covered broken down by the nature of such units	14	units, people, entities
	Update date	Defines the frequency between data updates as announced by the seller	11	real time, monthly, hourly
How?	Delivery method	Defines how the buyers can have access to data	8	S3Bucket, Download, FeedAPI
	Format	Defines the way in which data is arranged	17	txt, shapefile, 틴
	Add-ons	Tells whether the product attached any add-on or has any limitations	2	ProfServices, Limitations
When?	History	Time scope included	1	History
Where?	Geo scope	Metrics about countries included in the data product	up to 249	Ncountries, USA, Canada

자료 : Azcoitia et al.(2021) 재인용

- IMDA and PDPC(2019) : 데이터 가치에 영향을 미치는 요인(drivers)으로 유용성(usefulness), 완전성(completeness), 일관성(consistency), 정확성(accuracy), 적시성(timeliness), 독특성(exclusivity), 사용 제약(usage restrictions), 책임과 위험(liability and risks), 상호운영성/접근성(interoperability/accessibility) 등 9개 요인을 제시. 9개 요인 중 유용성은 가치에 직접적 영향을 미치고, 다른 8개 요인은 유용성에 영향을 미침으로써 데이터 가치에 영향을 미침



(데이터 가치에 영향을 미치는 요인)

자료 : IMDA and PDPC(2019)

- 김수진 외(2021) : 해외 데이터 상품의 가치평가요소 활용사례를 조사하고, 정형 데이터 세트 상품의 가치평가요소 제안

< 해외 데이터거래소별 가치평가요소 사례 >

거래소	평가요소					평가항목수	평가방식	평가목적
	데이터	업데이트	품질	제공방식	공급자			
다웍스 (프랑스)	대체가능성 업무중요성 지속성 상세설명 데이터볼륨	규칙성	사실성 완전성 정형화		평판	10개 영역	체크리스트 형식	잠재수요자에게 데이터 상품에 대한 정보 제공
구이양 (중국)	다양성 범위 상세성	경과도 적시성	무결성			6개 영역	가격산정모델 적용	데이터 공급자에게 거래가격 결정을 위한 정보 제공
프라운호퍼 (독일, 연구기관)	객관성 관련성 상세성 연계성 상세설명 데이터세트	적시성	명료성 무결성 타당성	플랫폼구조 접근방법 접근속도 라이선스		3개 영역 18개 지표	평점표 및 가중평균모델 적용	시장 내 데이터 상품의 가치 비교를 통한 구매 결정

거래소	평가요소					평가항목 수	평가방식	평가목적
	데이터	업데이트	품질	제공방식	공급자			
	수							
데이터 스트림 엑스 (싱가폴)	유용성 독창성 대체가능성 상세설명 활용효과 데이터볼륨	적시성 규칙성	정확성	편의성	브랜드	9개 영역	평점표 형식	데이터 공급자의 데이터 상품 공급가격 책정

자료 : 김수진 외(2021) 수정 후 인용

< 정형 데이터 세트 상품의 가치평가요소 >

평가영역		평가요소	측정지표	
영역	설명		지표명	산출방법
콘텐츠 적합성	데이터가 내포한 정보 즉, 데이터세트의 콘텐츠 평가	데이터정보 다양성	의미컬럼수	전체 컬럼수 - 컬럼 기준 결측비율 30% 이상 컬럼수
			정보범주수	데이터 컬럼이 표현하는 정보를 의미적 유사성에 따라 분류한 정보범주 개수
		타데이터와 연계성	표준연계키 개수	다른 데이터상품과의 결합이 가능하게 하는 표준연계키로 분류된 컬럼의 개수
		데이터 명확성	절대척도컬럼 비중	연속형컬럼의 절대값 컬럼수 / 전체 연속형컬럼수
			연속형컬럼명의 표기명확성	1 - (미표기컬럼수 / 전체 연속형컬럼수)
			범주형컬럼값의 기준명확성	기준제시컬럼수 / 전체 범주형컬럼수
		데이터 적시성	데이터 생성일의 최신성	현재일자 - max(데이트세트 내 기준일자 컬럼값)
			데이터 제공주기의 규칙성	공급자가 데이터를 업데이트하는 주기
			데이터 집계주기의 규칙성	데이터가 생성/집계되는 주기
품질/ 공급 신뢰성	데이터세트의 물리적 품질과 데이터상품 제공서비스 평가	데이터 물리적 품질	데이터품질 (정합성율)	데이터세트 품질 측면에서 도출된 품질지표의 지수화된 값
		데이터건수 충분성	유효데이터 건수	전체 데이터건수 - 컬럼 내 결측값이 있는 행의 수
		데이터 공급 지속성	공급기간	현재일자 - min(데이터스토어 상품등록일자)
			공급건수	현재일자 기준 최근 3년간 데이터스토어 상품등록건수
			공급최신성	현재일자 - max(데이터스토어 상품등록일자)
		데이터 공급자 신뢰도	공급자 기업신용등급	한국기업평가 기준 기업신용등급(회사채 기준)
			공급자 데이터품질 인증등급	한국데이터산업진흥원에서 수행하는 데이터품질인증(DQC-V) 등급

평가영역		평가요소	측정지표	
영역	설명		지표명	산출방법
기술적 사용성	데이터세트 이용 시 예상되는 기술적 제약	데이터 수령방식 편리성	데이터 수령방식	데이터제공방식
			확장자의 개방성	확장자 개방성 단계
경제성	데이터세트의 경제적 측면 평가	제약성	라이선스 유형	CCL(Creative Commons License) 유형
		대체 가능성	동종 카테고리 상품 비중	동일 카테고리 데이터상품수 / 전체 데이터상품수

자료 : 김수진 외(2021) 수정 후 인용

- 한국데이터산업진흥원(2020) : 정형 및 비정형 데이터의 품질 평가지표 제시

< 정형 및 비정형 데이터의 품질평가지표 >

데이터 유형	평가 지표	세부지표	
		지표	설명
정형 데이터	완전성	항목 완전성	필수 항목은 반드시 값이 입력되어야 함
		레코드 완전성	반드시 있어야 하는 레코드는 반드시 존재해야 함
		상품 완전성	상품 파일, API, URL은 반드시 존재해야 함
		메타 완전성	메타데이터의 필수 값은 반드시 존재해야 함
	유효성	범위 유효성	데이터 값은 유효 범위 내 값이어야 함
		형식 유효성	데이터 값은 유효 형식의 포맷을 준수해야 함
		목록 유효성	데이터 값은 유효 값의 목록에 존재하는 값이어야 함
		응답 유효성	API, 다운로드, 서비스상품 호출 시 응답이 되어야 함
	정확성	메타 정확성	메타데이터에 기재된 값과 실제 등록된 값은 일치해야 함
		의미 정확성	데이터 값은 업무적으로 의미 있는 값이어야 함
		계산/집계 정확성	항목 간 계산에 의해 산출되었을 경우 계산결과는 정확해야 함
		선후관계 정확성	선후관계를 갖고 있을 경우 선행과 후행이 정확해야 함
	적시성	데이터 제공 적시성	데이터는 갱신주기에 따라 정상적으로 갱신되어야 함
		데이터 최신성	데이터는 최근 데이터를 포함하고 있어야 함
		응답 적시성	API/서비스 호출 시 송수신 속도가 활용 가능한 수준이어야 함
	일관성	참조 무결성	하나 이상의 상품이 상속관계를 갖고 있을 경우 참조 상품은 기준상품에 반드시 정의된 값이어야 함
		항목 형식일관성	하나 이상의 상품의 동일한 데이터 항목은 동일한 데이터 형식으로 관리되어야 함
		항목 값 일관성	하나 이상의 상품의 동일한 항목은 동일한 데이터 값으로 관리되어야 함
		항목 관계 일관성	항목간에 연관관계가 있는 경우 서로 일관되게 관리되어야 함
	유일성	항목 유일성	식별자 단일 항목은 유일해야 함

데이터 유형	평가 지표	세부지표	
		지표	설명
		레코드 유일성	식별자 항목을 기준으로 레코드는 유일해야 함
	활용성	개인정보 익명성	개인정보 항목은 비식별화되어야 함
비정형 데이터	완전성	메타 완전성	메타데이터의 유무 및 필수 항목 내 Null 값 진단
	유효성	응답 유효성	데이터 파일 자체의 오류 여부 평가
		기능성	객체 데이터가 가지고 있는 자체적인 품질에 대해 평가
	정확성	메타 정확성	관리 메타 데이터에 기재된 값과 실제 객체 메타 데이터의 값이 정확한지 평가
		의미 정확성	객체 데이터의 내용의 의미가 정확하지 않은지 평가
		주제 정확성	데이터의 주제와 실제 객체 데이터의 내용이 일치하는지 평가
	적시성	최신성	데이터가 만들어진 시점과 현 시점을 기준으로 최신의 자료인지를 평가
	일관성	메타 일관성	객체 메타 데이터에 대하여 일관되게 관리하고 있는가를 평가
		메타 유사성	객체 메타 데이터에 대하여 균질한 규격을 가지고 있는지 평가
	유일성	객체 유일성	객체 데이터가 다른 파일명을 가지고 중복되어 있는지 평가
	활용성	친밀성	활용의 접근이 용이한 보편적인 확장자를 갖는 파일인가를 평가
		효율성	상품으로서 데이터를 활용하는 것에 특성을 평가
		개인정보 익명성	개인정보가 비식별화 정책에 따라 비식별화 되어 익명성이 보장되는지 평가
평가 지표 설명	(완전성) 상품의 활용을 위해서 필요한 항목과 레코드가 누락 없이 포괄적으로 정보를 제공하여 실물 세계를 완전하게 반영하는지 평가 (유효성) 데이터의 형식이나 범위, 목록 등이 데이터 성격(도메인)의 정의를 잘 반영하는지 평가 (유일성) 업무적으로 유일해야 하는 데이터가 1건만이 존재하는지 평가 (정확성) 실세계에 존재하는 상품 객체(사건, 사물, 개념 등)의 값이 오류 없이 저장되었는지 평가 (일관성) 상품 내의 동일한 데이터 간에 불일치가 발생하는지 평가 (적시성) 갱신주기에 대한 대응 및 응답시간과 같은 비기능적 요구사항 및 데이터의 최신성 유지와 같은 품질 요건에 대응하는지 평가 (활용성) 상품을 활용하는데 있어 사용자(이용자)가 요구하는 데이터의 범위와 상세화 정도를 충족시키는지 평가		

자료 : 한국데이터산업진흥원(2020) 재인용

○ 데이터 가격 영향요인에 대한 선행연구 검토 결과

- 한국데이터산업진흥원(2020)은 영향요인을 원가, 고객, 경쟁, 환경 등 4가지로 구분했는데, 원가는 상태변수, 고객은 행위자 변수, 경쟁은 활동 변수, 환경은 영역변수로서 동일한 위계에서 비교하기 어려움

  ⇒ 동일한 위계에서 비교할 수 있는 영역으로 구분할 필요가 있는데, 공급과 수요에 의해 가격이 결정되고, 시장 경쟁 수준을 의미하는 시장구조(완전경쟁, 독점 등)에 따라 가격 결정 행태가 다르다는 경제학의 가격이론을 고려하여 가치평가 요인을 수요, 공급, 시장구조 등 3개 영역으로 구분

  ⇒ 공급(supply)은 ‘주어진 시간과 공간에서 공급자의 이윤 극대화로부터 도출되는 가격 수준에 따라 가용한 또는 시장에 공급될 물량의 계획표’ 또는 ‘주어진 물량에서 한 단위 더 생산하기 위해 공급자가 요구하는 최소한의 가격(최소수용의사, WTA)’으로 정의되는데, 공급은 공급자의 생산비용이나 최소수용의사(WTA)에 영향을 미치는 생산기술, 요소 가격, 세금 또는 보조금, 가격위험, 미래에 대한 기대 등에 영향을 받음

  ⇒ 수요(demand)는 ‘효용 극대화로부터 도출되는 소비자가 주어진 가격에 구매하고자 하는 물량의 계획표’ 또는 ‘주어진 물량을 소비하기 위해 기꺼이 지불할 수 있는 최대 금액(최대지불의사, WTP)’으로 정의되는데, 수요는 효용이나 최대지불의사(WTP)에 영향을 미치는 소비자 선호, 가처분소득, 대체재 가격, 인구, 종교, 계절적 요인, 제도, 신상품 출시, 기대 등에 영향을 받음

  ⇒ 시장구조(market structure) 또는 시장 형태(market form)는 시장의 경쟁 수준에 의해 결정되는데, 시장구조를 기준으로 시장은 완전경쟁시장, 독점시장, 과점시장, 독점적 경쟁시장으로 구분. 시장구조를 결정짓는 요인으로는 시장참여자의 특징(가격수용자, 가격설정자) 또는 시장참여자의 수, 상품의 동질성 수준, 진입과 퇴출의 자유도, 지식과 정보의 완전성 여부 등이 있음

- 김수진(2021), 한국데이터산업진흥원(2020) 등에서 제시된 데이터 평가지표의 경우 데이터의 형식적 품질을 나타내는 요인으로 산업 또는 경영체의 성과에서 가치가 파생되는 데이터의 가치를 대표하는 데 한계 있음

  ⇒ 데이터의 형식적 품질과 질적 품질은 수요(한계효용)와 공급(한계비용) 모두에 영향을 미치므로 수요와 공급 영역의 요인으로 분류하고, 질적 품질을 나타내는 평가지표 개발 필요

  ⇒ 데이터 비즈니스 프로젝트의 40%가 데이터 품질 때문에 실패한다는 Gartner(2015)의 지적을 고려할 때 데이터 거래소의 브랜드 가치 향상을 위해 데이터 정확성, 완전성, 유효성, 적시성 등의 형식 품질은 거래소의 거래 데이터 선정기준으로 적용하는 방안 검토 필요

○ 본 연구의 데이터 가격 영향요인 분류

- 데이터 가격에 영향을 미치는 시장의 수요와 공급, 시장구조 요인은 다음과 같이 분류할 수 있음
- 특정 데이터 상품의 가격에 영향을 미치는 요인을 실증적으로 분석하기 위해서는 해당 상품의 수요와 공급 식별을 위한 시장의 정의와 범위설정 중요

< 영역별 데이터 가격 영향요인 >

구분	요인
수요	· 소비자 선호 : 대상 농식품, 대상 식품공급사슬 단계(생산, 유통, 소비 등), 데이터 상품에 대한 인식이나 선호, 충성도, 데이터 품질 - 데이터의 형식적, 질적 품질 모두 선호에 영향을 미칠 수 있음 · 가처분소득 : 구매자의 소득수준, 구매자가 속한 산업의 시장규모 · 대체재 가격 : 경쟁 데이터 가격 · 인구 특성 : 데이터 구매자 수, 구매자 긴급성, 데이터 활용 역량 · 계절적 요인 · 제도 · 신상품 출시 · 기대 : 기대소득 및 가치 · 데이터 생명주기(라이프사이클)
공급	· 생산기술 : 데이터 수집, 저장, 유지기술 · 요소가격 ; 데이터 수집, 저장, 유지 등에 투입되는 요소 가격 · 세금 또는 보조금 : 데이터 공급자에 대한 세제 혜택 · 가격위험 : 데이터 판매 및 요소가격 위험 · 미래에 대한 기대 : 데이터 산업에 대한 기대나 전망 · 데이터 형식적 품질 - 농산물과 같은 상품은 품질이 다른 경우 서로 다른 시장으로 구분되기 때문에 품질을 공급 요인으로 보기 어려우나, 데이터의 경우 형식적 품질이 달라도 내용적 품질이 같은 경우 동일한 상품이 될 수 있어 공급 요인으로 포함될 수 있음 - 데이터의 형식적 품질 향상 시 총비용이 증가하여 공급 감소
시장구조	· 시장참여자의 특징 : 가격수용자 vs. 가격설정자 · 시장참여자의 수 · 상품의 동질성 수준 · 진입과 퇴출의 자유도 · 지식과 정보의 완전성 여부

(7) 데이터 가치평가 및 가격책정 방안

- 주관적 또는 상황에 따라 달라지는 가치를 특정 시점을 기준으로 객관적으로 평가하는 과정
  - 가치 : 교환되는 자산에 지불될 가장 가능한 가격 또는 자산을 소유하는 경제적 이득에 관한 의견 (국제가치평가기준위원회)
  - 시장가치 : 불특정 다수 사이의 공정한 거래에서 반드시 거래되는 추정치를 지칭하는 개념으로 적절한 마케팅 기간이 주어지고, 평가일 현재 이해관계가 없는 판매자와 구매자 간의 추정 교환가격. 이 때 구매자와 판매자는 해당 상품에 대해 관련 지식이 있고, 사려 깊으며, 강제 없이 자유의지로 행동하는 사람을 지칭
  - 투자가치 : 개별 투자자산이나 비즈니스에 대한 소유자나 소유 예상자의 가치. 시장가치와 달리 특정 주체에 특화된 가치
  - 공정가치 : 공개되고 제약 없는 시장에서 비관계자 간의 거래로 가상적인 능력 있는 구매자와 판매자 사이에 거래가 이루어질 것으로 예상되는, 현금등가물로 표현된 가격. 이 때 시장참여자는 거래의 강제성이 없고, 합리적인 지식을 가지고 있음. '적절한 마케팅 기간'에 대한 언급이 없는 점을 제외하면 '시장가치' 정의와 거의 유사

- 가치평가 유형
  - 시장가치 평가 : 시장에서 교환되는 가장 가능한 가격을 추정하는 과정
  - 투자가치 평가 : 시장참가자 일반이 아닌 특정 개인이나 기업 실체에 특화된 가치를 측정하는 과정
  - 공정가치 평가 : 시장 일반이 아닌 거래 당사자 사이의 합의에 기반한 가치를 도출하는 과정

○ 객관적 또는 합리적 가치평가의 요건

① 가치평가 용어 표준화 : 현실에서 사용하는 다양한 용어를 정의하고, 가치평가 관행 표준화

② 개념 정립 : 가치평가와 관련된 이론적 개념 정립

③ 가치평가 원칙과 방법 표준화

④ 가치평가기준 제정, 운영 : 가치 평가자의 전문성, 가치평가에서의 성실성, 사회적 기준 준수, 법규 준수 등 가치 평가자가 지켜야 할 윤리 규정

○ 가치평가 목적

① 거래 : 특정 자산을 사거나 팔기 위한 평가

- 거래용, 합병용, 분할용(자산 분할 용도), 파산(파산기업 자산의 이해당사자에 대한 배분) 등에 활용

② 투자 : 투자나 금융거래를 위한 평가

- 투자용, 합작투자용, 기업공개(IPO), 담보 보증 등에 활용

③ 전략 : 전략기획을 위한 가치평가
- 사업타당성 사전평가, 전략기획 등에 활용

④ 회계 : 회계장부에 기재된 각종 자산을 현재 시가로 표현하기 위한 평가
- 시가평가, 손상평가(damage assessment), 구매가격 할당(purchase price allocation; 사업이나 기업을 구매할 때 각 자산에 대해 적정가격을 부여하는 작업) 등에 활용

⑤ 조세 : 조세 관련 업무 대응
- 상속 증여세, 이전가격(다국적 기업이 국가 간 거래 시 특정 국가의 조세 회피나 이익을 목적으로 거래가격을 조작할 수 있어 적절한 관세를 부여하기 위한 정상가격) 등의 산출에 활용

⑥ 소송 : 지적재산권 침해 등 법적 분쟁 대응을 위한 평가
- 법적 요구사항을 반영하거나 상거래 일반의 원칙에 입각한 또는 가치평가에서의 공정가치 원칙에 의거한 공정성 유지 중요

○ 가치평가원칙

① 시장가치평가 원칙
- 시장가치를 가치평가의 기본개념으로 적용한다는 원칙 : 다른 가치개념을 이용한 가치평가가 가능하나, 이 경우 해당 가치의 개념을 명확히 정의해야 함

- 시장가치 평가 시 추가 원칙
  - 거래비용 제외 원칙 : 판매비용, 운송비용, 거래세 등의 거래비용 제외
  - 최고 최선 사용(highest and best use)의 원칙 : 물리적으로 구현되고, 법적으로 허용되며, 재무적으로 가능한 상태에서의 가치 측정
  - 주체별 특수요인 배제원칙 : 특정 자산 소유자 또는 소유 예상자의 특수한 요인에 의해 발생하는 특수가치를 고려하지 않고, 시장 일반에서 가장 가능성이 높다고 생각하는 거래가격을 계산

② 실제 정보 우선원칙
- 실제 거래에서 발생한 실제 정보(과거 거래 사례, 비교 거래 사례, 유사 거래 사례 등)와 가상적인 상황에서의 가상 정보(미래소득 예측치, 미래소득 발생 확률, 가상 위험 등) 중 실제 정보를 우선함

③ 복수 방법론 사용원칙
- 여러 가지 가치평가 접근법을 사용하도록 권장 : 특정 접근법을 사용할 수 없는 경우 해당 이유를 명확히 제시하도록 권고
- 최근에는 복수주의 접근법 원칙을 완화하여 가장 적절한 가치를 도출하는 접근법을 사용하도록 하고 있음

④ 상관행법 사용원칙
- 경제원칙이나 이론에 근거하지 않고, 특정 분야에서 관행적으로 수행되는 가치평가 방법을 공식적인 가치평가 방법으로 채택해서는 안 됨
- 단, 특정 분야의 관행적 평가 방법을 공식적인 방법에서 도출된 가치지표의 적합성 판단을 위한 방법으로는 사용 가능

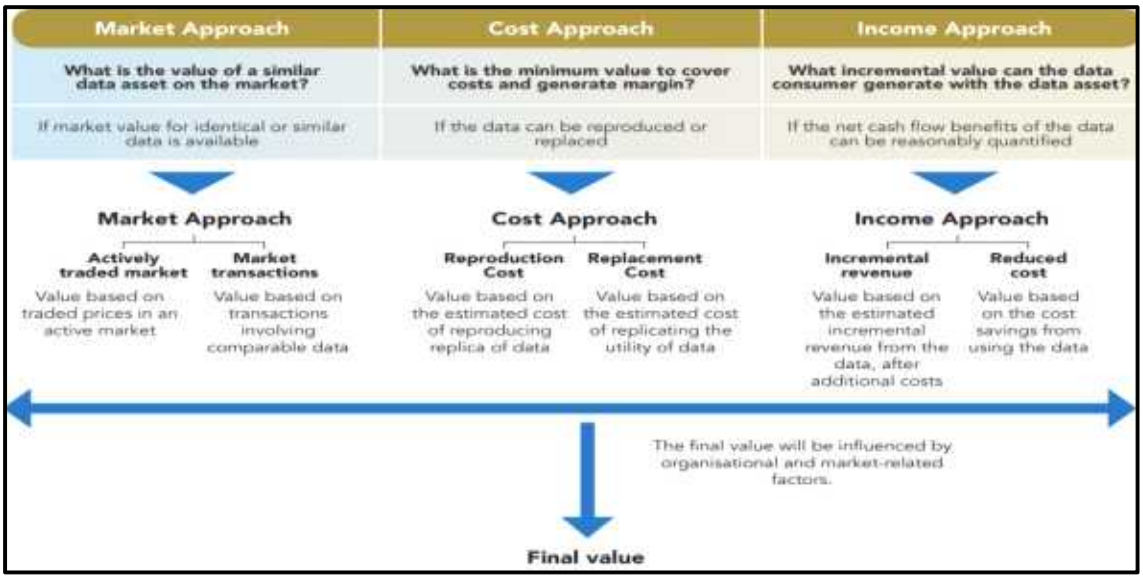
(8) 데이터 가치평가 방법론

○ 자산으로서 데이터의 가치는 시장 접근법(market approach), 비용 접근법(cost approach), 소득 접근법(income) 등의 회계적 가치평가 방법론을 이용하여 평가

① 시장 접근법 : 시장에서 이루어진 동일하거나 유사한 데이터 자산의 거래에 관한 정보를 바탕으로 가치를 평가하는 방법

② 비용 접근법 : 평가대상 데이터 자산과 유사한 효용을 제공하는 자산을 획득하는 데 소요되는 대체비용(replacement cost) 또는 동일한 자산을 생산 또는 재생산하는 데 소용되는 재생산비용(reproduction cost)을 이용하여 가치를 평가하는 방법

③ 소득 접근법 : 데이터 자산의 경제적 수명 안에서 발생이 예상되는 금전적 매출(이익)에 할인율을 적용하여 현재가치를 평가하는 방법



(데이터 가치평가 방법론 개요)

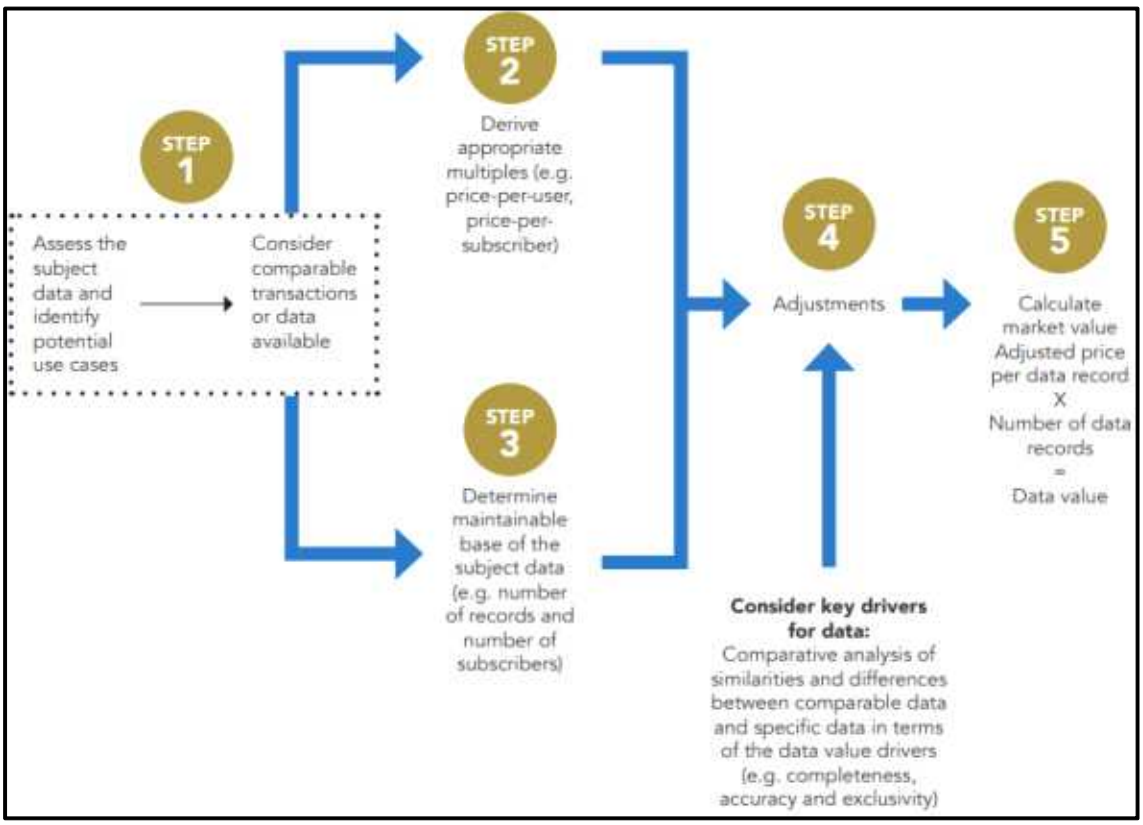
자료 : IMDA and PDPC(2019) 재인용

○ 가치평가 절차

① 시장 접근법

- 1단계 : 평가대상 데이터를 분석하고, 잠재적 사용 사례를 규명한 후 비교 가능한 거래 또는 확보 가능한 데이터 검토
- 2단계 : 데이터 사용자당 가격(price per user), 구독자당 가격(price per subscriber) 등 적절한 배수(appropriate multiples) 도출
- 3단계 : 데이터 레코드 수, 구독자 수 등 평가대상 데이터를 유지관리할 수 있는 기준을 결정
- 4단계 : 데이터 가치에 영향을 미치는 요인을 고려하여 조정
- 5단계 : 데이터 가치 추정

데이터 가치 = 데이터 레코드당 조정된 가격 × 데이터 레코드 수 (3-1)



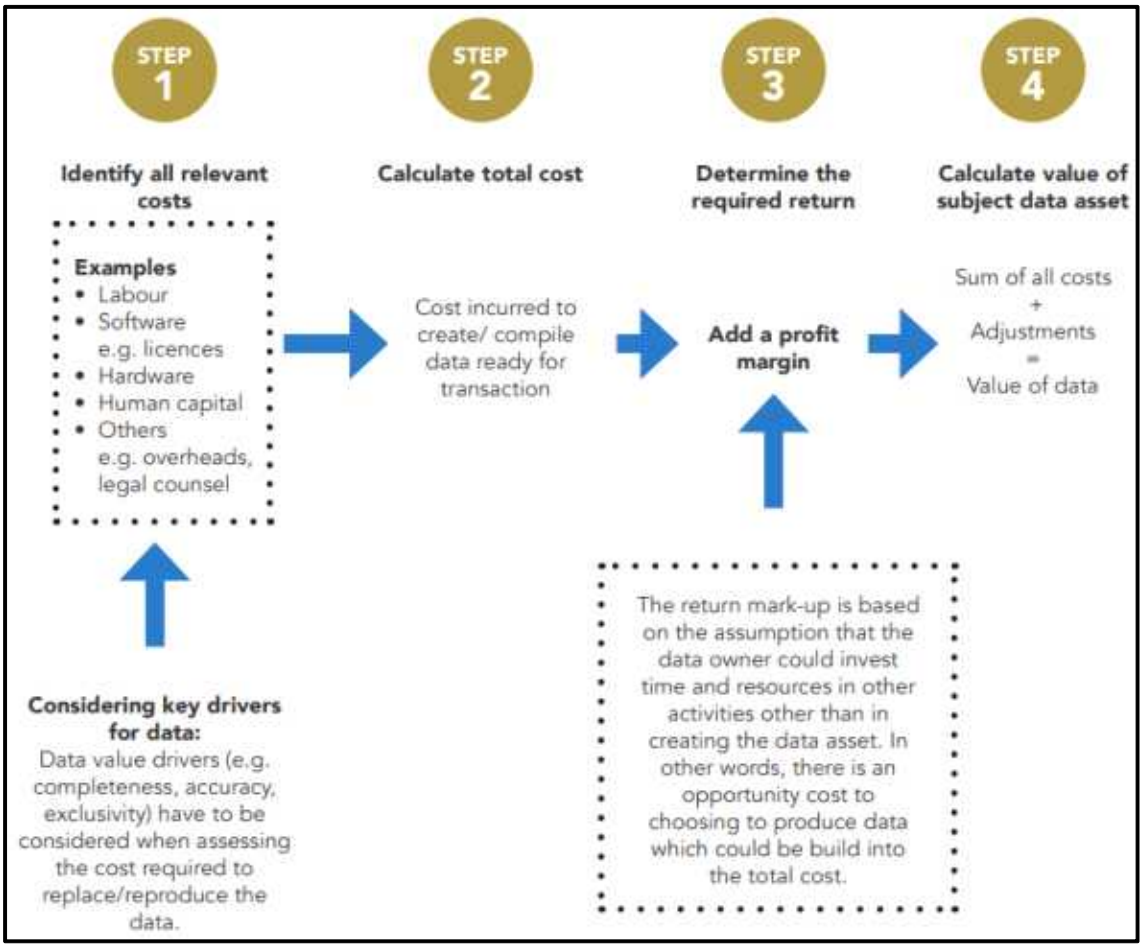
(시장 접근법을 이용한 데이터 가치평가 절차)

자료 : IMDA and PDPC(2019) 재인용

② 비용 접근법

- 1단계 : 평가대상 자산의 대체비용 또는 재생산비용 관련 비용 항목 규명
- 2단계 : 총비용 계산
- 3단계 : 데이터 대체 또는 재생산을 위해 투입한 데이터 소유자가 투입한 암묵적 기회비용(자원, 시간 등)을 '필요 수익(required return)'으로 추가
- 4단계 : 가치 추정

데이터 가치 = 총비용 + 필요 수익 등의 조정 가치 (3-2)



(비용 접근법을 이용한 데이터 가치평가 절차)

자료 : IMDA and PDPC(2019) 재인용

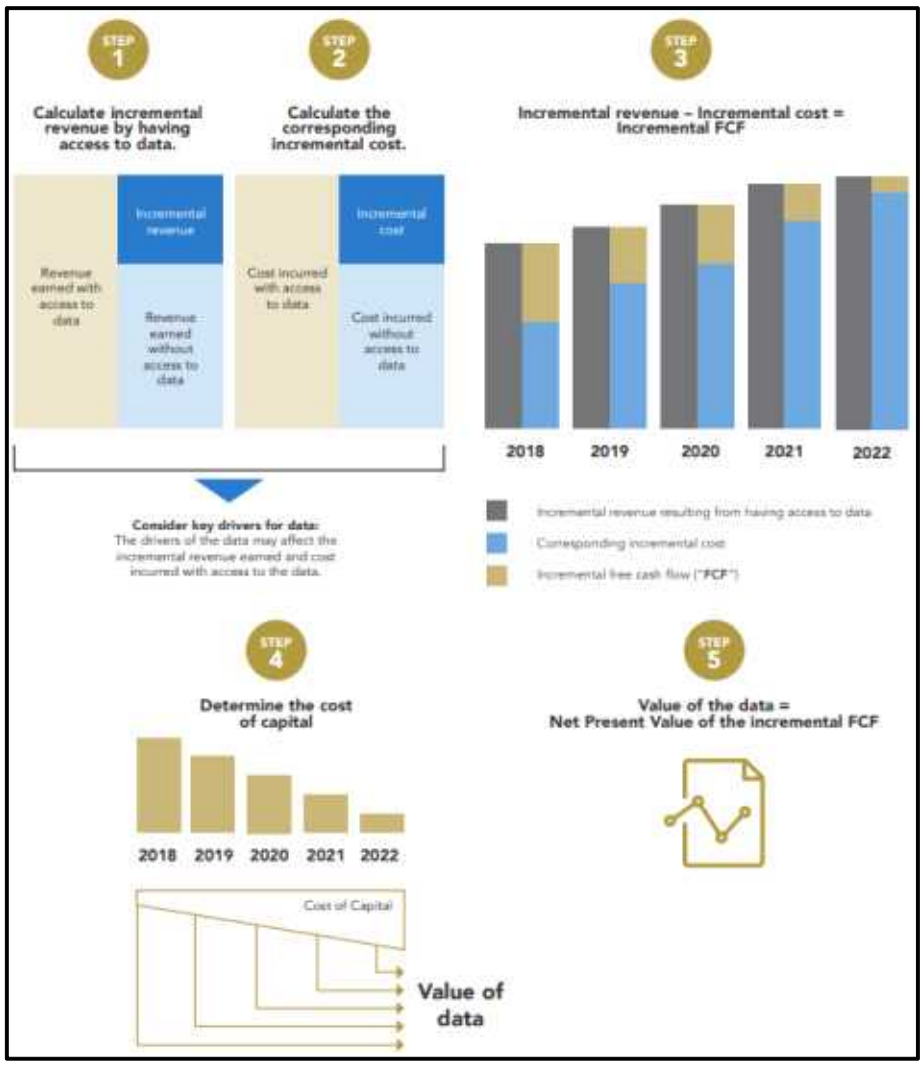
③ 소득 접근법

- 1단계 : 데이터 접근을 통한 시점별 (예상) 수입 계산
- 2단계 : 1단계의 (예상) 수입 실현에 필요한 시점별 (예상) 비용 계산
- 3단계 : 시점별 (예상) 잉여 현금 흐름(free cash flow) 계산

시점별 잉여 현금 흐름 = 시점별 수입 - 시점별 비용 (3-3)

- 4단계 : 현재가치 추정을 위한 자본비용(capital cost) 결정
- 5단계 : 데이터 가치 추정

데이터 가치 = 미래 잉여 현금 흐름의 현재가치의 합 (3-4)



(소득 접근법을 이용한 데이터 가치평가 절차)

자료 : IMDA and PDPC(2019) 재인용

(9) 데이터 가격책정 방법론

○ 가격책정(전략적 가격책정, pricing)의 과제
- 데이터 공급자가 경영목표를 달성할 수 있는 최적의 가격(optimal price) 부과

○ 전략적 가격책정의 중요성
- 가격은 매출이나 시장점유율에 결정적 영향을 미침
- 다른 마케팅 믹스 변수보다 신속하게 실행할 수 있고, 효과가 빠르게 나타남
- 여타 마케팅 변수에 비해 저렴

○ 가격책정 목표
- 이익 실현 : 매출 극대화, 이윤 극대화
- 투자 수익 확보 : 투자의 경제성 확보
- 최소한의 생존 : 생산비 이상의 이익 실현
- 일정 수준의 시장점유율 확보
- 현금 유동성 확보
- 현 상태 유지(status quo) : 안정성 추구
- 법, 제도적 이슈 ex) 가격 하한제 또는 상한제

○ 가격책정에 영향을 미치는 요인
- 구매자 또는 고객의 지불의사수준
- 조직적, 마케팅적 목적
- 비용 또는 원가 : 가격의 하한선 결정
- 기타 마케팅 믹스 요인 : 제품, 유통, 촉진
- 유통경로 구성원의 기대
- 시장의 경쟁수준 : 경쟁 제품의 가격과 품질
- 법, 제도

○ 가격책정 방법

① 원가가산 가격책정(cost-plus or cost-markup pricing)
- 쉽고 편리하나, 최적 가격책정이 어려움 : 가격 = 원가 + 목표 이윤
- 기업의 원가구조가 비슷한 과점시장에서는 담합의 수단으로 사용되기도 함
- 일반 상품과 달리 데이터의 경우 예상 판매 횟수를 고려해 책정

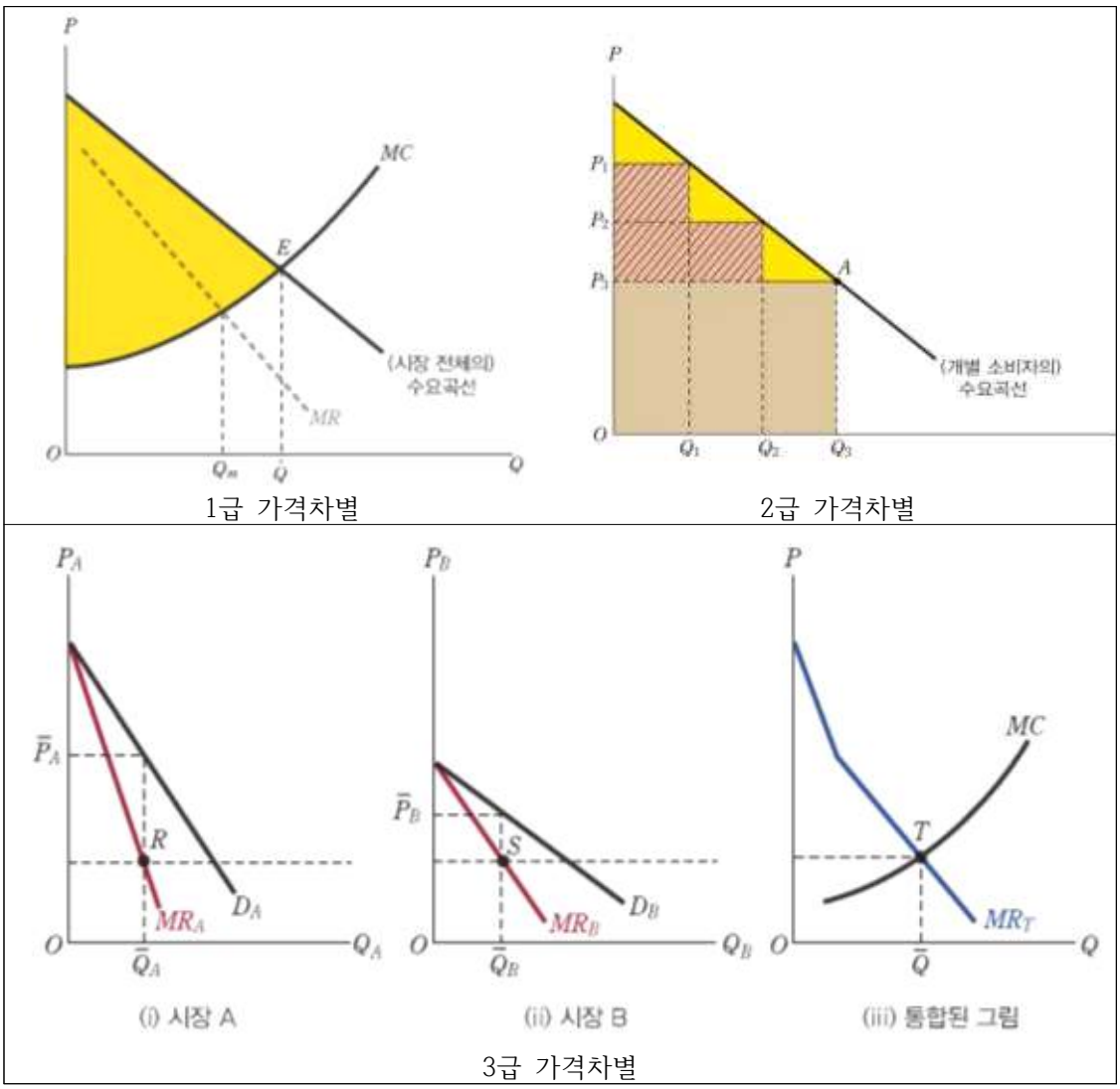
$$\text{(일반상품) 가격} = \frac{(\text{원가} + \text{목표 이윤})}{\text{생산량}} \quad \text{vs.} \quad \text{(데이터) 가격} = \frac{(\text{원가} + \text{목표 이윤})}{\text{예상 판매 횟수}} \qquad (4)$$

② 경쟁적 가격책정(competitive pricing)

- 저가 : EDLP(every day low price), 가격 할인 전쟁으로 치닫기도 함
- 경쟁사와 유사 가격 : 가격이 비가격 경쟁에 비해 중요하지 않은 경우 적용하고, 경쟁업체의 가격에 지속적인 주의를 기울여야 함
- 고가 : 고품질, 독점적 공급, 우수한 서비스, 높은 브랜드 충성도, 명품 이미지 등 비가격 경쟁에서 경쟁우위가 있는 경우 적용

③ 변동 가격책정(variable pricing, price differentiation)

- 가격 차별의 유형
  - 1급 가격차별 : 하나의 시장에서 개별 소비자에게 서로 다른 가격 적용
  - 2급 가격차별 : 이용 실적에 따라 차등 가격 부과, 구간가격 설정
  - 3급 가격차별 : 두 개 이상의 분리된 시장에 대해 다른 가격책정



(변동 가격책정)

- 가격 차별 조건
  · 시장 구분이 가능하고, 시장 구분 비용이 이익보다 적으며, 합법적이어야 함
  · 시장 구분에 따른 소비자 거부감이 없고, 목표시장의 경합 수준이 낮아야 함

- 가격 차별의 토대 : 구매자의 가격탄력성 차이
  · 구매자 특성 : 소득, 연령, 성별, 상태(신규 고객, 대규모 고객)
  · 상품의 품질
  · 소비 시점

④ 가격 고정(카르텔)
- 동질적 상품에 대한 공식적이고 공공연한 가격 합의
- 시장점유율, 생산능력, 협의 등을 통한 수익 배분
- 교란요인 : 무임승차, 신상품, 신규 진입, 시장지배력 변화
- 일반적으로 불법

⑤ 암묵적 담합(tacit collusion) : 가격 리더쉽이 존재하는 경우
- 가격 발표를 통한 공공연하지 않고, 비공식적인 모의
- 규모, 시장점유율, 활동기간, 비용효율성, 예측 능력 등을 기반으로 한 리더쉽 작용

⑥ 묶음 가격(pricing bundling)
- 반반음식, 햄버거 세트 메뉴, 통신요금 등에 적용
- 데이터의 경우 결합 데이터의 형태로 적용 가능
- 순수 묶음 : 묶음으로만 구매 가능
- 혼합 묶음 : 개발 상품으로도 구매 가능

⑦ 비선형 가격책정(non-linear pricing)
- 이부가격(이중요율, two part tariff) : 기본료와 이용 수준별 추가 요금으로 가격을 책정하는데, 소비자잉여를 기본료로 책정 ex) 통신요금, 놀이공원

- 단계별 요율(block tariff) : 이용량에 따라 차등 가격을 적용
  ex) 300MB 구매 시 100MB까지는 P1, 100~200MB는 P2, 200~300MB까지는 P3에 판매

- 전량 수량 할인(all unit quantity discount) : 전체 물량에 대해 이용 수준별 차등 가격을 적용
  ex) 100MB 이하 구매 P1, 100~200MB 구매 P2, 200MB 초과 구매 P3

- 비선형 가격책정 방식은 다양한 형태로 혼합하여 적용할 수 있음
- 한국데이터산업진흥원(2020)에서 제시한 이용료 책정방식 중 종량제와 혼합제는 비선형 가격책정 방식을 의미

< 개별 계약을 통한 데이터 이용료 책정방식 >

유형	내용
정액제	· 데이터의 사용 정도와 관계없이 일정한 가격을 적용
종량제	· 다운로드 수, 제공 건수(호출 수), 접속 시간, 접속 ID 수 등에 따라 차등적 가격을 적용(단계별 요율 or 전량 수량 할인)
혼합제 (정액제+종량제)	· 일정량 이하 데이터를 사용하는 경우 고정금액을 부과하고, 기준을 초과하면 이용량에 비례해 차등적 가격을 적용(이부 가격책정)
	· 일정 수준까지 이용량에 따라 비용을 차등 부과하고, 일정 수준을 넘으면 고정 비용을 부과(단계별 요율 + 이부 가격책정)

자료 : 한국데이터산업진흥원(2020) 수정 후 인용

⑧ 홀수 가격책정(odd pricing) 및 짝수 가격책정(round pricing)

- 홀수 가격책정
  - 소비자가 비싸다고 인식하는 문턱 가격을 피함
  - 왼쪽 자리 효과(left digit effect) : 10,000원 vs. 9,999원
  - 국가나 문화에 따라 가격책정 방식 다름 : 미국 소매업계는 끝자리가 9로 끝나는 경우가 많으나, 한국이나 일본의 경우 8-9-5-0 규칙 적용

- 짝수 가격책정 : 고급 이미지 제공

○ 가격경쟁전략

① 가격할인 : 수요 창출 효과 + 수요 전환 효과

- 수량 할인 : 구매자에게 규모의 경제에 따른 이익 제공
- 현금 할인 : 현금 구매에 대해 가격 인하
- 정기 할인 : 비성수기 할인
- 시간 할인 : 마감 시간이 임박할수록 가격 인하
- 쿠폰, 공제 등

⇒ 한국데이터산업진흥원(2020)에서는 ① 사용기간 또는 사용량 할인, ② 협력사 할인, ③ 이용목적 할인, ④ 전용망 할인, ⑤ 행사 할인 등의 가격 할인 및 차별화 방법을 제시했는데, 상시적 가격책정 방식인 비선형 가격책정과 마케팅 목적으로 일시적으로 가격을 변경하는 할인은 구분해야 함

- 제시 방법 중 행사 할인을 제외한 나머지 가격책정 방식은 비선형 가격책정을 의미

② 지정가(limit pricing)

- 장기 이익을 위해 단기 이익을 희생하여 가격 인하
- 주로 독(과)점 시장에서 신규 진입을 제한하기 위해 사용

③ 약탈적 가격(로스리더 또는 덤핑)

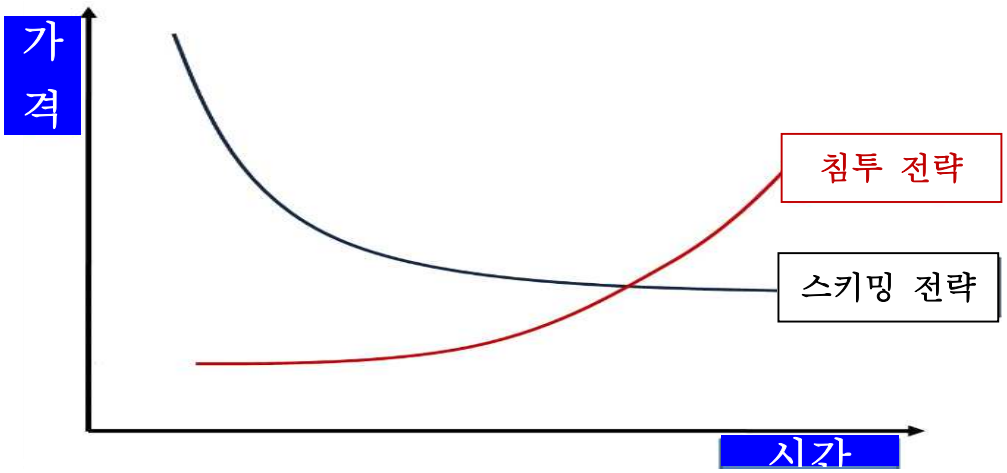
- 가격을 평균비용 미만으로 책정
- 주로 자본력이 풍부한 대규모 공급자가 경쟁자를 추출하기 위해 사용
- 현금 흐름이나 유동성 확보를 위해 사용하기도 함

- 대부분 불법이거나 가격 전쟁 유발

○ 상품 사이클에 기반한 동태적 가격책정

① 시장 스키밍 가격책정(market skimming pricing)
- 초기에 높은 가격을 책정, 경쟁자가 증가할 경우 가격 인하
- 수요가 많고, 초기 경합수준이 낮은 경우 사용
- 고소득의 비탄력적 수요를 가진 시장 대상

② 시장 침투 가격책정(market penetration pricing)
- 초기에 낮은 가격을 책정하여 경쟁자를 추출하고, 소비자 충성을 확보한 후 가격 인상
- 수요가 적고 소비자가 가격에 민감하며 잠재적 경쟁자가 많은 경우 사용
- 가격할인이 중요한 마케팅 수단으로 사용됨



(스키밍 가격책정 vs. 침투 가격책정)

(9) 데이터 속성가치평가 방안

○ 속성가치평가는 구매자의 선호를 반영한 데이터 품질 지표 구축과 이를 토대로 한 가격 책정에 활용할 수 있음
- 한국데이터산업진흥원(2020), 김수진 외(2021) 등의 선행연구에서는 전문가 조사등을 통해 도출된 가중치를 활용한 평가지수를 이용하여 데이터 품질을 평가하는 방안을 제시했으나, 해당 방식은 특정 품질지표의 유효성 평가에는 적합하나, 해당 지표가 가치 또는 가격에 미치는 영향 분석에 한계

○ 본 연구에서는 대표적 속성가치평가 방법인 가상가치평가법(CVM), 헤도닉 가격 분석(hedonic price analysis), 컨조인트 분석(conjoint analysis) 방법 제시

1) 가상가치평가법

○ 가상가치평가법(contingent valuation method)은 소비자가 이용하고 있는 비시장재화 이용을 위해 또는 비시장재화의 개선을 위해 어느 정도의 금액을 지불할 의사가 있는지를 소비자에게 직접 질문하여 비시장재화의 가치를 평가하는 방법임(오세익 외, 2001)
- 국내에서는 가상가치평가법을 조건부가치평가법이라 칭하기도 함(권오상, 2013)

○ 비시장재에 대한 가상적 시장(hypothetical market)을 설정하고, 지불의사를 직접 표현하도록 하여 소비자들의 선호를 측정함
- 가상가치평가법에서 평가된 경제적 가치는 보상잉여나 동등잉여 등의 Hicks의 후생개념이 응답자의 지불의사금액(WTP; willingness to pay)이나 보상수용액(WTA; willingness to accept)을 통해 표현되어 논리적인 이론적 근거가 있음
- 지불의사(WTP)나 수용의사(WTA)를 직접 도출할 수 있어 효용함수에 대한 가정이나 수요함수 유도 등의 복잡한 중간과정을 생략하고 Hicks의 후생변화 척도를 유도할 수 있음

○ 가상가치평가법은 적용 범위가 넓어 사용가치(use value)뿐만 아니라 비사용가치(non-use value)를 평가할 수 있는 장점이 있음
○ 그러나, 실제 상황이 아닌 가상상황을 근거로 개인의 의사를 묻기 때문에 오류를 범할 가능성 있음(오세익 외, 2001)

○ 그러므로 가상가치평가법을 이용하여 가치평가를 할 경우, 설문지 작성, 설문시행, 사후분석 과정에 세심한 주의를 기울여야 함(권오상, 2013)

○ 가상가치평가법의 설문은 크게 다음과 같이 세 부분으로 구성

① 평가대상이 되는 비시장재화에 대한 설명과 응답자가 처한 가상적 상황 설명

② 응답자의 지불의사를 유도하는 질문

③ 응답자의 사회경제적 상황 및 선호도에 대한 질문

○ 가상가치평가법은 질문 형태에 따라 개방형, 폐쇄형, 다지선다형으로 구분

① 개방형
- 응답자가 직접 지불의사를 직접 진술하는 방법으로 직접질문법과 지불카드법으로 구분
- 직접질문법(direct question) : 응답자에게 최대지불의사금액을 직접 질문. 일반적으로 응답자는 기대되는 편익을 화폐단위로 평가하는데 어려움을 느끼므로 지나치게 큰 금액이나 적은 금액을 말하거나 응답을 회피하는 경향 있음
- 지불카드법(payment card format) : 응답자가 속한 소득계층에서 소비를 위해 연간 부담하는 금액이 적혀 있는 지불카드를 보조자료로 제시. 응답자는 카드를 참조하여 자신의 최대 지불 의사 금액을 답함

② 폐쇄형

- 제시된 지불의사에 대해 ‘예’ 또는 아니오‘로 응답하는 방법으로 경매법과 이산선택형으로 구분
- 경매법(bidding game) : 경매방식을 사용해서 응답자에게 일정액수를 제시하고 ‘예 ' 또는 ‘아니오’의 응답을 유도. 가치평가를 위해 먼저 최초입찰가격을 제시하고 응찰여부를 확인. 응찰자의 지불의사금액이 제시된 금액보다 적으면 최초 입찰가격에 불응할 것이고, 이 경우 조사원은 금액을 낮추어 다시 질문. 최초 시작 금액과 입찰가격의 증가분(감소분)에 따라 지불의사가 달라지는 출발점 편의(starting point bias) 문제와 입찰가격의 증가분 편의문제가 있음
- 양분선택형(dichotomous choice) : 가상적 시장을 설정하고 각각의 응답자에게 무작위로 명시된 금액을 지불할 용의가 있는지를 물어보고 응답자는 ‘예’ 또는 ‘아니오’로 한 번만 대답함. 응답자는 제시된 금액이 본인의 지불의사금액보다 같거나 작으면 ‘예’ 라고 응답하고 높으면 ‘아니오’ 라고 응답. 이 기법은 응답자가 주어진 제시가격보다 본인의 지불의사금액이 큰지 작은지 여부만 판단하면 되기 때문에 응답하기가 용이. 실제 시장거래에서의 의사결정 상황과 매우 흡사하여 전략적 편의가 최소화 될 수 있고, 출발점 편의가 존재하지 않는 장점이 있음(Freeman, 1993). 최근에는 첫 번째 질문에서 ‘예’ 또는 ‘아니오’라고 응답한 경우 각각의 상황에 맞는 지불의사를 추가로 제시하여 ‘예’ 또는 ‘아니오’로 응답하게 하는 이중양분선택형(double dichotomous choice) 방식이 자주 활용됨

③ 다지선다식(multi-alternative question method)

- 다지선다형으로 제시된 지불의사 중 하나를 선택하는 방식
- 다지선다식 방식은 질문방식이 단순하면서도 응답의 범위를 한정함으로써 개방형에서 발생할 수 있는 무응답이나 극빈치의 발생을 방지할 수 있음
- 양분선택형(dichotomous choice) 방식의 경우 응답이 쉽고, 극빈치의 발생확률이 작다는 장점이 있으나, 지불의사 추정이 어렵다는 단점이 있고(곽승준 외, 2007), 지불 여부를 답하는 기준이 자의적으로 제시된다는 한계가 있는데, 다지선다식은 이러한 한계를 보완

추가지불의사	① 0%(지불하지 않음)			② 1%	③ 2%	④ 3%	⑤ 4%	⑥ 5%	⑦ 6%
	⑧ 7%	⑨ 8%	⑩ 9%	⑪ 10%	⑫ 11%	⑬ 12%	⑭ 13%	⑮ 14%	⑯ 15%
	⑰ 16%	⑱ 17%	⑲ 18%	⑳ 19%	㉑ 20%	㉒ 21%	㉓ 22%	㉔ 23%	㉕ 24%
	㉖ 25%	㉗ 26%	㉘ 27%	㉙ 28%	㉚ 20%	㉛ 30%	㉜ 기타( )%		

(다지선다식 가상가치평가법 질문 예시)

○ 가상가치평가법을 이용하여 지불의사 추정 시 지불수단 중요

- 다수의 연구에서는 재산세, 소득세, 목적세, 기부금, 입장료 등을 지불수단으로 사용하고, 특히 소득세 방식을 많이 사용
- 반면 이춘수 · 양승룡(2013)에서는 간접세율을 지불수단으로 사용
- 소득세의 경우 소득세에 대한 거부감으로 지불의사를 과소 응답할 가능성이 있고(홍종호 · 엄영숙, 2011), 실제로 부과된 사례가 없어 현실성 낮음
- 이와 달리 간접세 형태의 목적세는 이미 유류세에 교통 · 에너지 · 환경세, 교육세, 주행세, 농어촌 특별세 등의 목적세가 부과되고 있기 때문에 소득세에 비해 현실성이 높음
- 소비자들에게 익숙한 세율 방식을 적용함으로써 답변의 현실성을 높일 수 있고, 세금 부과 대상(또는 산업)을 명확히 할 경우 확률효용함수 또는 확률지불함수 등을 이용한 별도의 추정 과정을 거치지 않고 지불의사를 직접 추정할 수 있음

○ 가상순위결정법(Contingent Ranking Method: CRM)
- 가상순위결정법은 CVM의 설문 형식 중 하나인 투표모형(referendum model)을 확장한 분석 모형
- 응답자들이 비용과 함께 제시된 몇 가지 대안을 선호 순서대로 나열하게 하는 방법임(송운강, 2009)
- 연구대상의 개별 특성 각각에 대한 한계가치를 평가할 수 있다는 장점 존재(권오상, 2000)
- 국내에서는 권오상(2000), 오세익(2001), 권오상 · 윤태연(2004), 송운강(2009) 등에서 가상순위결정법을 이용하여 농업의 다원적 가치를 평가한 바 있음

2) 헤도닉 가격 분석

○ 헤도닉 가격 분석(hedonic price analysis)은 Waugh(1928)에 의해 제안된 방법으로 쾌적한 환경의 질에 대한 수요를 주택, 토지와 같은 재화의 속성 및 시장 가치(거래가격)로부터 추정하는 방법임

○ 주택 i의 가격(P_i)은 주택 특성(S_i), 주변 환경(N_i) 등에 의해 결정되고, 주변 환경 변수 중 환경질과 관련된 변수(Q_i)를 고려하여 주택 관련 자료를 구한 후 주택별 헤도닉 가격 방정식을 다음의 식과 같이 추정 가능
- 식에서 각 속성, 예를 들어 주택의 환경질 개선에 따른 주택가격 변화($\frac{\partial P_i}{\partial Q_i}$)를 잠재한계가격(implicit marginal price)이라 하고, 이를 해당 속성의 가치로 평가

$$P_i = f(S_i, N_i, Q_i) \quad (5)$$

○ 헤도닉 가격 분석 모형 모형으로 선형모형(linear model), 반대수모형(semi-log model), 양대수모형(log-log model) 등 다양한 함수 형태를 고려할 수 있음
- 다수 연구에서 추정된 계수값을 속성 가치로 평가할 수 있는 선형모형 활용

○ 헤도닉 가격 분석의 한계

① 개인별 한계지불의사 도출
- 개인별 한계지불의사함수를 추정하기 위해서는 거래된 주택을 구입한 사람의 소득을 포함한 소비자 특성 정보를 추가 파악해야 하므로 자료 수집 비용이 많이 듦
- 각 개인이 자신이 소비할 환경재의 양과 잠재한계가격을 동시에 선택하기 때문에 내생성과 같은 계량경제학적 문제가 발생 가능
- 실제 관측되는 데이터는 개인별 한계지불의사의 어느 한 점만을 모아서 연결한 선이기 때문에 추정된 개인별 한계지불의사가 사실상 순수한 개인의 환경질에 대한 한계지불의사를 반영하지 못하고, 개인 특성의 차이로 인해 발생하는 한계지불의사의 차이를 나타낼 가능성이 큼
- 개인별 한계지불의사를 정확하게 도출하기 위해서 취할 수 있는 한가지 방법은 주택 구매자가 다른 환경질을 지닌 지역에서도 주택을 구매했을 경우 얼마만큼 지불하였는지를 추정하는 것임
- 예를 들어 서울에서 주택을 구매한 구매자(j)의 개인특성(αj)을 추정된 부산의 잠재가격함수에 대입할 경우 이 주택자가 부산에서 주택을 구매했다면 얼마의 가격을 지불하였는지를 추정할 수 있음

- 그러나 현실적으로 자료의 한계 등으로 이러한 한계 보완 불가

② 함수 형태의 선택
- 헤도닉 함수를 선형으로 추정할 경우 잠재한계가격이 모든 주택이나 소비자에게 동일해지는 문제 발생

③ 공간 종속성과 이질성
- 공간 종속성은 자료들이 서로 독립되게 관측되었으나 공간적인 상관관계가 존재한다는 의미로 공간 종속성이 존재하는 상황에서 회귀모형을 추정할 경우 편의 발생 우려 있음(Irwin and Bockstael, 2001)
- 공간 이질성은 변수들이 상호 관련을 맺는 형태가 공간이 변함에 따라 달라짐을 의미

④ 편익의 계산
- 주변 환경의 변화로 사람들의 주택선택 자체가 달라지고 주택시장의 균형이 달라지는 경우 적분을 활용한 편익 계산이 다원적 기능의 가치를 나타낸다고 보기 힘듦
- 개인별 잠재한계가격함수가 효용이 아닌 소득의 함수이므로 특정 속성에 대한 한계지불의사를 나타낸다고 보기 어려움

3) 컨조인트 분석

○ 컨조인트 분석(conjoint analysis)은 확률효용모형(random utility model)에 근거를 둔 방법으로 여러 속성들에 대한 수준 조합으로 구성된 선택 대안을 응답자가 선택하도록 하여 해당 속성의 가치를 평가하는 방법임
- 컨조인트 분석을 통해 각 속성 수준에 대한 부분가치(part-worth)를 추정할 수 있고, 가격 속성 사이의 상충관계를 이용하여 개별 속성들의 한계지불의사금액을 추정할 수 있는 장점이 있음

○ 컨조인트 분석법의 응답 방법으로는 여러 속성으로 구성된 다양한 선택 대안에 대한 선호수준이나 순서를 표시하거나 제시된 대안 중 하나를 선택하는 방법이 있음
- 최근 연구에서는 응답자의 응답 용이성 등을 고려하여 제시된 대안 중 하나를 선택하도록 하는 선택형 실험법이 자주 활용됨

○ 선택형 실험법(choice experiment)은 여러 속성과 가격으로 구성된 2개 이상의 대안을 응답자에게 제시하면, 응답자가 제시된 대안 중 가장 선호하는 것을 선택하게 하여 속성들의 수준 변화에 대한 지불의사를 추정하는 방법임
- 응답 결과로부터 획득한 자료와 다항로짓모형(multinomial logit model)이 프로빗 모형(probit model) 등을 사용하여 속성 가치 추정

○ 속성들의 수준 변화에 대한 지불의사금액을 추정하기 위한 선택형 실험법은 확률효용모형을 이용하여 다음과 같이 정형화할 수 있음
- 응답자 i 가 선택대안집합 C 내의 한 선택 대안 j 로부터 얻는 간접효용함수는 식 (2)과 같음
- 이때 U_{ij} 는 응답자 i 가 선택대안 j 를 소비할 때 얻을 수 있는 효용수준이며, 이는 관측 가능한 확정적인(deterministic) 부분 V_{ij} 와 관측 불가능한 확률적인(stochastic) 부분 ϵ_{ij} 로 구분할 수 있고,

식 (6)에서 Z_{ij} 는 선택대안 j 와 관련된 속성벡터임

$$U_{ij} = V_{ij}(Z_{ij}) + \epsilon_{ij} \quad (6)$$

- 응답자 i 는 선택대안집합 C_i 내의 선택 대안들에 대해 $U_{ij} > U_{ik}$ $(k \in C_i, k \neq j)$ 일 때, 선택대안 j 를 선택하는데, 응답자 i 가 선택대안 j 를 선택할 확률은 식 (7)과 같음

$$P_{ij} = Prob(V_{ij} + \epsilon_{ij} \geq V_{ik} + \epsilon_{ik}) = Prob(V_{ij} - V_{ik} \geq \epsilon_{ik} - \epsilon_{ij}) \quad (7)$$
$$\text{for all } k \neq j$$

- 만약 오차항의 분포가 독립적이고 동일하게 분포(independently and identically distribution)하고, 제 I 형태 극치 분포(Type I extreme value distribution)을 따른다고 가정할 경우, 응답자 i 가 K 개의 선택 대안 중에서 대안 j 를 선택할 확률 P_{ij} 는 식 (4)과 같이 나타낼 수 있음(McFadden 1974)
- 식 (8)에서 μ는 오차항의 분산과 역의 관계를 갖는 비례 모수(scale)를 의미하는데, 일반적으로 불변오차분산(constant error variance)을 뜻하는 1과 같다고 가정

$$P_{ij} = \frac{\exp(\mu V_{ij})}{\sum_{k=1}^{K} \exp(\mu V_{ik})} \quad (8)$$

- 식 (6)에 나타낸 바와 같이 간접효용함수의 관측 가능한 부분인 V_{ij}는 식 (9)와 같이 속성벡터의 선형함수로 가정하고, 이때 Z_{ijm}은 m번째 속성을 의미하며, β는 응답자의 효용에 영향을 미치는 개별 속성들에 대한 추정 계수임

$$V_{ij} = \beta_0 + \beta_1 Z_{ij1} + \beta_2 Z_{ij2} + \cdots + \beta_m Z_{ijm} \quad (9)$$

- 식 (9)를 통해 추정된 β 값을 통해 속성별 부분가치(part-worth)와 한계지불의사금액을 구할 수 있음(식 10)
- 식 (9)의 간접효용함수에 가격 변수를 더미(dummy) 처리하여 투입한 후 다항로짓모형을 적용하면 각 속성 수준에 대한 부분 가치와 속성별 상대적 중요도를 도출할 수 있고(김근배 2006; 이훈영 2007), 실제 제시한 가격 자체를 변수로 투입한 후 전미분하면 다음과 같이 속성별 한계지불의사금액을 도출할 수가 있음
- 식 (10)에서 Z_m는 m번째 속성 변수를, Z_p 는 가격 속성 변수를 의미

$$MWTP_{Z_m} = \frac{dV/dZ_m}{dV/dZ_p} = -\frac{\beta_m}{\beta_p} \quad (10)$$

○ 선택형 실험법을 적용한 컨조인트 분석 사례

- 본 사례는 컨조인트 분석을 이용하여 막걸리의 속성 가치를 분석한 양성범 · 양승룡(2011)의 사례를 선택형 실험법에 맞게 수정한 것임
- 사전 소비자 설문조사를 통해 컨조인트 분석에 활용할 속성을 선택하고, 가격을 포함한 5가지의 속성 선택
- 발효에 사용되는 주원료를 쌀과 밀을 혼합하여 사용하는 경우와 쌀만을 사용하는 2가지로 구분하고, 쌀의 원산지는 ‘수입산’, ‘국산’의 2가지로 구분

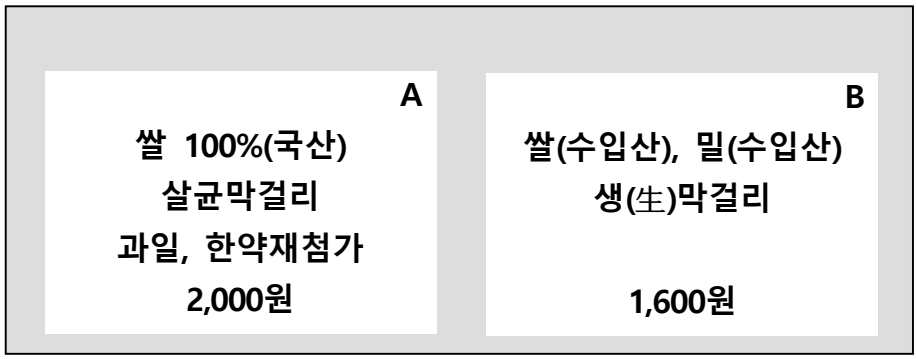
- 살균 방식은 ‘살균 막걸리’, ‘생(生) 막걸리’의 2개로 구분하고, 차별화 원료를 사용하지 않는 막걸리와 과일이나 한약재와 같은 차별화 원료를 첨가하는 2가지로 구분
- 여기에 각 속성의 경제적 가치를 추정하기 위해 가격을 속성으로 포함하였는데, 대형마트에서 실제 판매되고 있는 막걸리 가격을 참고하여 가격의 수준을 ‘800원’, ‘1,200원’, ‘1,600원’, ‘2,000원’의 4단계로 구분

< 컨조인트 분석 사례 : 막걸리의 속성 및 속성 수준 >

속성	속성별 수준
사용원료	‘쌀+밀’, ‘쌀 100%’
쌀의 원산지	‘수입산’, ‘국산’
살균 방식	‘살균 막걸리’, ‘생(生) 막걸리’
차별화 원료	‘무첨가’, ‘과일, 한약재 첨가’
가격	‘800원’, ‘1,200원’, ‘1,600원’, ‘2,000원’

자료 : 양성범 · 양승룡(2011)

- 막걸리의 속성과 속성 수준을 상기와 같이 설계할 경우 선택 대안을 구성하는 개별 속성들의 수준을 결합하면 (2×2×2×2×4)=64개의 선택 대안이 존재하며, 이 중에서 2가지 선택 대안을 뽑아 선택 대안 집합을 구성하면 2,016개로 매우 많은 대안이 존재
- 이 모든 대안을 질문하는 것은 현실적으로 어려우므로 직교설계(orthogonal design)로부터 16개의 최소 선택대안집합 도출(Kuhfeld, 2005)
- 직교설계로 도출된 다섯 가지 속성을 가진 상품을 2개씩 짝지어 선택대안 집합 8세트를 제시하고, 각 세트 내에서 선호하는 상품을 선택하도록 하고, 연구에 따라 ‘둘 다 선택하지 않음’을 선택대안으로 제시하기도 함



(컨조인트 분석 사례 : 선택형 실험법 상품 대안 예시)

자료 : 양성범 · 양승룡(2011)

○ 컨조인트 분석의 한계

① 상품을 구성하는 속성의 수가 증가할수록 대안의 현실성이 증가하는 반면 직교설계에도 불구하고 응답자가 선택해야 하는 대안의 수가 각 속성의 속성 수준의 배수만큼 증가하여 설문에 어려움 있음

② 속성 가치가 가격 속성의 수와 범위에 영향을 받음

7. 스마트팜 데이터 거래환경(시스템) 구축

가. 시스템 운영 모델

((가칭)스마트농업 데이터 거래소(안))

구분	유형
개방용 상품	• 플랫폼 포털을 통해 무상 개방되는 상품 • 원본 데이터를 사용 가능 한 수준으로 품질 개선하여 수요기업에서 바로 활용할 수 있는 형태로 가공 • 주로 정제가공 기법이 적용
양산형 판매 상품	• 플랫폼 포털에 전시되면서 상시 판매되는 상품으로서 비용 결제가 완료되면 곧바로 다운로드 받을 수 있는 상품 • 단일 데이터의 경우 개방용 상품보다 한단계 낮은 단계로 제공하거나, 데이터 제공 주기가 최신인 상품 또한 주제별로 연관된 데이터들을 융합하여 새롭게 생산된 데이터 상품 • 주로 정제가공 기법과 융합가공 기법 적용(맞춤형 판매 상품) • 수요기업 요구사항에 맞게 맞춤형으로 가공하여 제공 • 수요기업의 활용목적이 서비스일 경우 해당 요건에 맞게 가공되어 곧바로 활용될 수 있도록 가공하여 제공 • 활용목적이 분석일 경우 분석에서 설계된 내역에 맞게 가공되어 제공

○ 공유/개방형 상품 : 원본 데이터 품질개선을 통해 플랫폼 포털을 통한 무상개방
- 기존 개방된 데이터와의 차별성을 통해 플랫폼의 참여센터와의 가공・융합 재생산을 통한 공유/개방형 데이터 상품 생산
- 양산형 판매상품 : 플랫폼 포털을 통한 상시 판매로 주체별 연관 데이터의 융합상품

○ 데이터 검색, 결재, 다운로드/API, 연계, 통계 공유 등 거래기능을 통한 유통모델
- 직거래 유통모델 : 플랫폼의 유통 서비스 기능 활용을 통해 비용 결재 후 다운로드
- 위탁 유통모델 : 수요기업 접점에 있는 업체가 판매대행으로 대행업체에 영업 수수료 지급
- 패키지 유통모델 : 분석 전문업체가 데이터를 구매하여 분석 서비스와 패키지 상품으로 판매

○ 빅데이터 플랫폼 데이터 가격 및 판매정책
- 플랫폼 판매 데이터 : 양산형 판매 상품과 맞춤형 판매 상품에 대한 원가를 반영한 가격정책 수립을 통해 자체 수익 발생
- 원가고려 : HW/SW 등의 구매비용, 가공/운영을 위한 인건비, 마케팅 비용
- 가치고려 : 융복합 분석을 통해 생성된 맞춤형 데이터,
- 맞춤가공 : 원가를 고려한 유사 경쟁 상품과의 가격고려, 차별점에 대한 고려
  → 데이터 융・복합을 통한 플랫폼 데이터 판매를 통한 수익발생

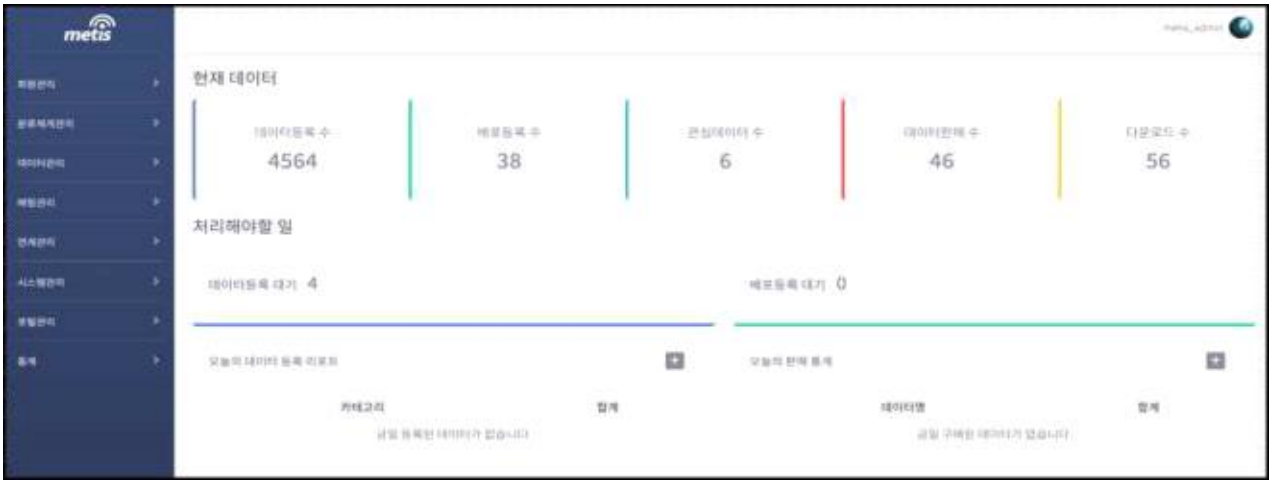
○ 혁신 모델발굴을 통한 서비스 제공으로 수익발생 방안 수립
- 센터 및 타 산업분야 데이터 융복합을 통한 혁신서비스 수익발생

○ 온라인상에서 데이터 마켓 모든 서비스 구현
- 데이터 공급, 판매, 구매, 유통을 위한 다운로드, 결재 등 플랫폼내에서 업무 프로세스가 처리될 수 있도록 빅데이터 플랫폼 구축 운영

○ 기존 개방데이터와의 차별화 방안
- 기존 데이터 개방 : 농정원, 농진청, 농협 보유데이터만을 활용한 데이터 개방
- 플랫폼 데이터 개방 : 농정원, 농진청, 농협 보유데이터 + 센터 + 연계기관 + 타 산업 융복합을 통한 데이터 생산 개방

○ 유통・거래 표준약관 수립 방안
- 관련 법규와 기존 유사 사이트들의 약관을 벤치마킹하여 유통/거래 표준약관을 마련
- 공공데이터의 제공 및 이용 활성화에 관한 법률, 공공기관의 정보공개에 관한 법률, 전자정부법, 국가정보화기본법, 행정조사기본법, 개인정보 보호법, 저작권법 등
- 벤치마킹 대상 유사 서비스 : K-ICT빅데이터센터, 공공데이터포털 등

○ 데이터 유통 시스템 운영



(데이터 유통시스템 예시 - 대시보드)

(데이터 유통시스템 예시 - 데이터셋 추가)

(데이터 유통시스템 예시 - 검색 및 시각화)

- 표준 데이터 유통 플랫폼 소스를 활용하여 데이터 연계 기능을 개발하고, 개방 데이터 등록 및 사용자 데이터 검색 지원을 위한 태그 설정 작업 등을 수행
- 데이터는 본 플랫폼에만 제공되는 게 아니며, 다수의 개방 플랫폼들과 연동하기 위해 표준화된 데이터 카탈로그 파일을 구축 및 관리

○ 스마트 농업 데이터 거래소 구축 및 운영 방안

(시스템 기능 구성도 예시)

- 인프라 부문
  a) 인프라(IaaS 및 SaaS)는 대기업 혹은 공공기관의 cloud를 이용하여 빅데이터 플랫폼 구성
  b) IaaS영역에서 Obect Storage, Elasticsearch, TensorFlow Server등의 다양한 빅데이터 인프라 서비스를 사용
  c) SaaS영역에서는 시각화 도구와 실시간 데이터 전송을 위한 NiFi, 비식별화 솔루션 등을 이용

- 데이터 부문
  a) 연계기관, 센터, 활용기업 등에서 스마트팜 데이터를 생산 및 공급
  b) 센터 및 플랫폼에서 가공 생산한 정형, 비정형, 반정형의 다량의 데이터를 빅데이터 기술을 활용 저장, 관리 및 서비스

- 빅데이터 플랫폼 부문
  a) 데이터의 유형별로 수집하는 방식과 저장하는 방식이 달라지며, 언제든지 수집된 데이터를 활용 가능하도록 관리
  b) 표준화/품질관리 : 분석, 보관, 공유를 위한 품질 기준을 정의하고, 규칙 기반으로 데이터 품질을 진단하여 품질 향상
  c) 데이터 융합·가공 : 수요기관이 활용 목적에 맞게 바로 사용할 수 있는 형태로 가공 및 융합하여 활용성 높은 데이터로 제공
  d) 데이터 유통/거래 : 개방형, 판매용, 주문생산형 데이터들로 정제 가공하여 유통
  e) 생태계 구축 : 인력양성 및 기술·연구·창업 지원·공공/민간 협업 거버넌스 구성 및 운영
  f) 거버넌스 참여 : 빅데이터 플랫폼 협의체 운영 및 의사 결정

○ 포털 서비스 부문 구축 및 운영 방안

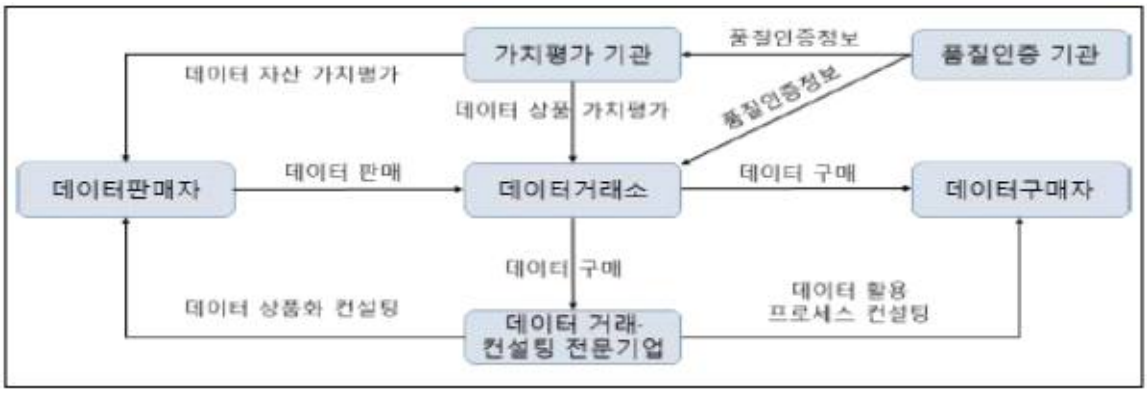
플랫폼 내 영역	구축 및 운영 방안
스마트팜 빅데이터 포털 서비스 포털	• 스마트팜 분야 빅데이터 플랫폼으로써의 브랜딩, PR, 콘텐츠영역의 서비스 제공 • 센터와 플랫폼에서 생산 및 유통하는 연관 데이터 검색 영역, 컨텐츠 응용서비스, 데이터 조회 순위, 맞춤형 추천서비스 제공
유통거래 포털	• 대국민 혁신서비스 소개 맞춤형데이터 요청, 창업지원 등의 기능 제공

- 스마트팜 빅데이터 구매, 이종 데이터 융합 및 활용 분석, 스마트팜 데이터 활용 사례 및 비즈니스 아이디어 정보를 찾는 스타트업 등 스마트팜 빅데이터에 대한 수요가 있는 이용자들에게 편의를 제공할 수 있는 서비스 포털사이트 구축이 필요
- 스마트팜 빅데이터 플랫폼에 참여하는 컨소시엄 및 센터의 사업 안내를 위한 홍보채널 제공
- 스마트팜분야의 데이터 수요자들을 위한 데이터 마켓, 분석 콘텐츠 제공, 교육 및 지원 사업 안내 등 수요자들의 기업 활동, 스마트팜 빅데이터와 관련한 지식정보 추구에 필요한 정보를 제공하는 정보전달 채널 제공
- 데이터 생산, 공급, 제공
  a) 센터 : 스마트팜 핵심데이터를 보유한 센터 및 활용기업, 연계기관에서 수요기반 활용가치가 높은 데이터 생산/구축/개방
- (빅데이터마켓) : 스마트팜 빅데이터플랫폼 및 각 센터들에서 가공, 정제 및 생산된 데이터의 유상 구매, 라이선싱 또는 무상 다운로드와 API 키발급을 통한 인증 및 테스트 진행 등 필요한 데이터의 검색과 사용까지 원스탑 활용 제공
- (데이터분석 콘텐츠) : 각 센터들에서 생산된 데이터와 플랫폼에 데이터를 공급하는 참여기관들의 융합된 데이터를 기반으로 수요자들이 궁금하고 흥미를 유발할 수 있는 데이터 시각화 분석 콘텐츠 제공
- (혁신서비스) : 스타트업 등 기업, 산학연구소 연구자 및 정책연구자를 포함한 농업인이 스마트팜 빅데이터의 활용성을 체감할 수 있는 혁신 서비스를 운영
- (커뮤니티) : 농/축업 산업분야 종사자, 빅데이터 기술전문인 등이 플랫폼을 통하여 데이터를 융합하여 사용할 수 있는 정보교류, 아이디어 교환, 토론공간 등을 제공
  a) 빅데이터 플랫폼 및 센터 협의체 커뮤니티
  b) 빅데이터 인력양성을 위한 교육 참가자 및 멘티 커뮤니티
  c) 빅데이터 수요기업 커뮤니티 및 자연발생적 빅데이터 기술 커뮤니티 등

○ 스마트 농업 데이터 거래 관련 단계별 역할

< 데이터 거래 관련 분류별 역할 >

분류	역할
데이터 판매자	직접 생산하거나 수집·가공한 데이터를 다양한 목적에 따라 판매하는 자
데이터 구매자	비즈니스 강화 및 확장을 위해 외부에서 데이터를 공급받고자 하는 자
데이터 거래소	데이터 공급자와 구매자를 이어주는 중개서비스를 제공하는자 (예시 : 데이터거래 전문기업, 유료데이터를 유통하는 빅데이터 플랫폼 등)
데이터 거래·컨설팅 전문기업	데이터 활용 프로세스 컨설팅, 외부데이터 조사·추천 컨설팅 및 데이터 상품화, 거래 등에 필요한 컨설팅을 수행하는 자
가치평가 기관	데이터산업법에 따라 데이터에 대한 객관적인 가치평가를 전문적·효율적으로 수행하는 자
품질인증 기관	데이터산업법에 따라 데이터의 내용, 구조, 기술적·절차적 관리체계 등 데이터 품질인증을 수행하는 자



(데이터 거래 관련 분류 체계)

○ 사전준비 단계

- 판매자와 데이터 거래소는 시장에서 필요한 상품 기획, 데이터가공 및 가격을 산정하고, 구매자는 업무에 필요한 외부 데이터 사용요건과 데이터 조달계획을 기획하고 데이터 상품을 검색
- 판매자, 구매자는 내부 데이터 전문인력 또는 데이터 컨설팅 기업을 활용하여 (판매자) 데이터 상품 기획, (구매자) 데이터 활용 요건 검토, 상품 검색 등 추진
- 판매자 또는 데이터 전문기업은 데이터 상품화 단계에서 가치평가와 품질인증 등을 통해 가격산정의 기초자료로 활용

○ 거래・계약 단계

- 데이터 거래소는 등록된 데이터의 가격과 품질의 적정성을 검토해야 하며, 계약시 필요한 표준계약서, 대금결제 등의 제반사항 지원

○ 사후관리 단계

- 판매자는 계약서에 따라 데이터 갱신시 무상 또는 별도의 수수료를 받고 추가 데이터를 제공할 수 있음
- 구매자는 구매한 일부 데이터(가명정보 등)에 따라, 계약서에 명시된 기간 동안 구매한 데이터의 보호조치 활동과 구매한 데이터를 이용 하여 서비스 개발・재가공・제3자 제공 등의 이용현황을 관리해야 함

나. 운영모델 가동을 위한 요구사항

○ (데이터 활용 서비스 개발 지원) 분야별 신산업 창출 및 서비스 효율화를 위해 플랫폼 데이터에 지능정보기술을 접목한 서비스 개발 지원

- (예시) 통신, 부동산 등의 비금융 데이터를 활용한 대안신용평가 모델 개발
- 데이터 분석 기술이 부재한 기업의 수요를 확보하여 데이터 가공 기업 연계를 통한 기업 수요에 맞는 데이터와 서비스 개발 지원
- 스타트업 및 중소기업과 기 구축 플랫폼이 컨소시엄을 구성하여 신규 비즈니스 개발 및 사업화 수행

○ (데이터 안심구역 확대) 플랫폼 내 미개방・가명정보를 다양하게 가공, 분석, 개발, 활용할 수 있도록 데이터 안심구역 확대

- 중소기업, 스타트업 등이 데이터를 자유롭게 분석하고 결과만을 반출할 수 있는 오프라인 분석환경(공간) 구축 확대

- 금융, 헬스케어 플랫폼 → 통신, 중소기업 플랫폼 → 문화, 유통 플랫폼
- 개별 데이터 안심구역의 미개방 데이터를 상호 연계·분석할 수 있는 데이터 안심구역 허브(Hub) 구축 추진
- K-DATA의 데이터 안심구역을 중심으로 다양한 데이터 안심구역 연계 추진

○ (기반) 농식품 데이터 축적-가공-유통 생태계 조성을 위해 가장 필요한 농식품 빅데이터 활용 지원 조직, 빅데이터 플랫폼, 빅데이터 생성-활용 관련 주체 간 유기적인 네트워크 등의 기반 구축이 우선되어야 함

- 다양한 주체의 활용 수요에 맞춰 데이터 구축 · 개방 · 활용의 체계적 지원을 위하여 정부와 공공기관 간 역할 분담이 필요하며, 기관별로도 역할에 맞는 전문 인력으로 구성된 조직 구성 및 기능 확대가 필요함
- 본격적인 생산 · 유통 · 소비 분야와 후방산업이 연계되어 데이터가 자동적으로 수집되어야 하며, 데이터 활용의 제약사항 개선 및 표준 관리 체계 정립이 이루어져야 함.
- 분야별 데이터 축적 · 가공 · 유통 표준 마련, 품질 관리 및 분산된 데이터 통합 관리, 이들 데이터 간 융합 분석을 위한 다양한 기능을 제공하는 “농식품 빅데이터 플랫폼” 구축을 우선적으로 추진해 나가야 할 것임

○ (활용) 부처, 각 기관별로 수집 · 생성되는 공공(행정) 데이터의 종류와 양은 지속적으로 증가하고 있으나, 대국민 대상 정책 서비스 개선 및 발굴 관점에서의 데이터 활용 노력은 부족한 상황이며, 단발적인 사업 형태에 그치고 있음.

- 농식품 분야 정보시스템(EA기준) 194개, 개방 공공데이터 375종을 보유하고 있고 농업 분야별 빅데이터 기반 융 · 복합 분석은 대표적으로 채소류 수급예측, 동물질병 확산 대응체계 구축 분야에서 이루어지고 있으나, 아직은 활용이 저조한 상황임.
- 각 기관 또는 업무 중심의 단순한 빅데이터 활용도 필요하지만, 현안 대응 등 정책적 활용과 대국민 서비스 분야에 대해서는 실현 가능성과 개선 중심으로 “농식품 빅데이터 활용 로드맵” 수립이 되어야 하고, 민관 및 학계가 함께 중점 과제 발굴 및 단계적 추진 방안 수립이 필요함.
- 지금까지는 공급자인 정책자 및 데이터 관리자 관점으로 과제가 발굴 · 추진되어 왔으나, 앞으로는 고객 서비스 및 민간 비즈니스 모델 관점에서의 데이터 수요 및 서비스 모델 발굴을 추진해나가야 할 것임
- 이런 수요 분석 과정을 통해 민간 비즈니스 영역과 중복되지 않는 공공 기능 및 서비스 제공 차원에서 지속적인 모델 발굴및 대국민 서비스 개선이 이루어져야 할 것임.
- 현재 구축 중인 행정데이터, 통계 등 정형데이터의 활용과 함께 영농현장 · 환경 변화, 농업인의 정책 수요 등 관심 사항을 파악할 수 있는 농식품 분야 비정형 데이터 수집 기반 구축과 정책 서비스 차원에서 활용 방안 마련이 필요함.

< 현장 요구사항에 기반한 분야별 데이터 거래 및 활용 문제점 >

현장의 목소리	공공데이터	민간데이터	융합데이터
①쓸만한 데이터 부족	개방건수는 증가하고 있으나, 기업의 활용에는 역부족	ROI부담으로 (특히 중소기업에서) 데이터축적과 활용에 소극적	인공지능 학습 및 서비스 개발에 필요한 데이터의 태부족
②표준화 미흡	데이터 구조 및 포맷이 상이하거나 품질이 고르지 않음	활용가치가 높은 민간 (통신사, 카드사 등) 데이터의 개방 및 유통여건 열악	공공-민간의 데이터 연계 · 공유 기반 미흡
③개인정보 활용 제약	(규제) 엄격한 규제로 개인정보의 연계와 활용 제약 (인식) 개인정보 유출에 대한 부정적 인식으로 이용활성화에 보수적		

○ (통합 데이터지도 중심 연계) 누구나 쉽게 플랫폼 데이터를 찾고 활용할 수 있도록 다양한 데이터 플랫폼을 ‘통합 데이터지도’와 연계
- 필요한 데이터를 찾는 이용자 또는 데이터 공급을 원하는 기업을 플랫폼과 연결해주는 ‘데이터114’ 서비스 제공

○ (통합 거래체계 구축) 통합 데이터지도를 중심으로 거래 기능이 있는 공공주도 플랫폼을 우선 연계하고 민간 플랫폼의 자발적 참여 유도
- 데이터 거래 관련 정책지원 창구 역할을 수행할 거래지원센터 운영, 데이터 가격산정 모델 고도화·적용 및 적용 확산

○ (수요자 중심 환경 조성) 일정요건을 갖춘 데이터 플랫폼에서 수요자가 자유롭게 구매·가공을 지원받도록「바우처 One Service」도입 추진
- 바우처 집행의 효율성·투명성 확보를 위해 관계부처 협의체에서 구체적 도입 방안 등 논의
- 데이터 바우처, 플랫폼 데이터 활용 분야 등을 분석하여 데이터 공급기업에게 데이터 수요를 미리 알려주는 수요 예보제 추진

○ (민간데이터의 공공활용 촉진) 데이터 플랫폼의 융합서비스 및 데이터API가 디지털서비스 전문계약제도를 통해 활용되도록 지원
- 각 플랫폼이 시장수요가 높은 데이터 상품·서비스를 디지털서비스 심사위원회를 거쳐 이용지원시스템에 등록하도록 유도

○ (플랫폼 관리기반 마련) 데이터 플랫폼에서 개방·유통되는 데이터가 상호 연계·활용 되도록 데이터 입력 항목의 단어·용어 등을 표준화
- 메타데이터 연계규격을 바탕으로 플랫폼이 데이터를 연계·융합· 활용하도록 연계기준과 규격 등 표준화 추진

○ (신뢰성 확보 관리체계 도입) 플랫폼의 데이터 중 오류추정 데이터를 분석하여 품질관리에 활용하고 품질관리 가이드를 지속 개정
- 플랫폼 구축 이후 데이터 현행화, 편의성 제고 등을 위한 플랫폼 운영 대가 산정기준 마련 추진(’ 2) 및 개별 데이터 식별 체계 도입 검토

○ (데이터 거래) 현재 거래시장은 브로커형 거래가 상당 부분 차지하나, 데이터 플랫폼을 중심으로 오픈마켓 형태의 유통채널이 형성
- 직접 수집 구매 데이터를 가공 및 활용하여 데이터 상품이나 분석 서비스를 판매하는 유형으로 국내 데이터 거래 매출 상위 30개 기업 중 67% 차지(‘20년 기준)
- 15개 빅데이터 플랫폼과 금융데이터 거래소 위주로 형성 중
- 데이터 거래는 전통적으로 금융·기업 정보나 마케팅·광고·소매 등데이터의 활용이 금전적 가치로 바로 이어지는 특정 분야에 집중
- 최근 맞춤형으로 데이터를 수집·가공해 제공하거나, 수요기업이원하는 데이터 분석 모형을 개발 및 판매하는 등 사업모델 다양화

○ (유통 기반 마련) 공공·민간 협업 기반의 빅데이터 플랫폼 구축과 거래 체계 지원을 통해 데이터 플랫폼 중심의 거래 활성화 유도
- 빅데이터 플랫폼의 데이터를 한 곳에서 검색 가능한 ‘통합 데이터 지도’를 개시하고 다른 플랫폼과도

연계를 확대
- AI허브, 데이터스토어, 무역빅데이터 및 한국관광 데이터랩과 검색 연계
- 데이터 판매자와 구매자가 거래 시 참고할 수 있는 가격·품질·법률 분야의 구체적 정보를 제시하는 가이드라인(3종) 개발

○ (표준화·품질 제고) 공공데이터포털과 빅데이터 플랫폼을 중심으로 데이터의 원활한 연계 및 활용을 위한 표준화 및 품질관리 추진
- 빅데이터 플랫폼은 데이터 구축 단계부터 표준화 및 품질관리 가이드를 제작하고 플랫폼별 데이터 구축·관리에 적용
- 공공데이터 공통표준용어를 토대로 빅데이터 플랫폼에 필요한 각 산업의 공공·민간 용어 등을 정리하여 연계표준사전 마련

○ (바우처 활용처 확대) 일정요건을 갖춘 데이터 플랫폼에서 기업 등 수요자가 자유롭게 구매 또는 가공을 지원받을 수 있는「바우처 One Service」도입 추진
- 자격요건을 미리 승인* 받은 플랫폼에 대해 데이터 바우처 이용을 허용하며, 관련 관계협의체를 구성하여 도입 논의
- 한국데이터산업진흥원(K-DATA)을 통한 승인을 전제로 바우처 집행의 효율성과 투명성 확보를 위해 관계부처 협의체에서 구체적 도입 방법 및 절차 등 논의

○ 바우처카드(신용, 체크)를 통해 편리하게 이용할 수 있는 방안 검토
- (예시) 보육료, 유아학비, 임신·출산 등 교육부·보건복지부의 바우처의 경우 통합카드 활용

(선택형 별 바우처 카드 이용 예시)

○ (데이터 수요 예보제) 데이터 공급기업의 수요불확실성을 완화하기 위해 시장의 현재 데이터 수요를 미리 알려주는 '수요 예보제' 추진
- 데이터 바우처 수요, 데이터 플랫폼별 데이터 활용 분야, 공공부문의 수요 등을 분석하여 차년도 바우처 사업 시행 전 수요현황 공개

○ (클라우드 기반 통합 유통) 클라우드 기반 데이터 거래 서비스 활성화를 통해 공급·수요자 모두의 편의성 향상 및 거래시장 활성화
- 클라우드 플랫폼에 축적된 다양한 대용량 데이터를 활용한 거래· 분석 통합 서비스(Data-as-a-Service) 제공 환경 구축
- 아마존, 알리바바, 바이두, 스노우플레이크(Snowflake) 등 해외 주요 클라우드 기업은 클라우드

플랫폼 기반 데이터 거래·분석서비스 제공 중

○ (분석서비스 바우처) 데이터 플랫폼을 중심으로 분석서비스 바우처를 제공하여 창업, 서비스 개발 및 사업 확대 등에 데이터 활용을 지원

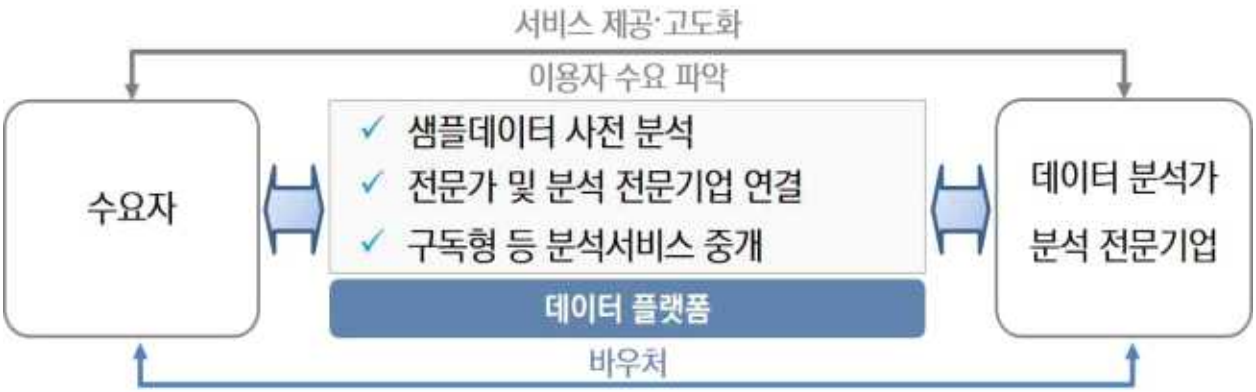
- 분석역량이 부족한 스타트업, 중소기업 등의 수요에 따라 플랫폼 (데이터)과 전문기업(분석기술)이 함께 분석서비스 및 컨설팅 등 제공

○ (분석 편의성 제고) 플랫폼의 분석기능과 유료데이터의 샘플 등록을 강화, 이용자가 원하는 데이터의 분석방향을 미리 쉽게 파악 지원

- 플랫폼 내에서 데이터 분석가·분석기업을 이용자와 연결하여 수요 맞춤형 데이터 상품 및 분석서비스 공급을 활성화
- 데이터 플랫폼은 신규·인기 데이터 현황, 데이터 분석사례, 데이터 상품 추천 등 유용한 정보를 이용자에게 지속 제공

○ (구독형 분석서비스) 시장 마케팅 분석 등 수요가 있는 주제별 정기 구독형 분석리포트를 제작하여 분석서비스 저변을 확대

- 산업 트렌드 및 품목별 매출 정보 등 중소기업, 스타트업, 소상공인 등에게 필요한 데이터 분석보고서를 저렴하게 제공
- (통신 빅데이터 플랫폼) 소상공인에게 필요한 상권정보(업종별 매출액·임대료·개폐업수 등) 분석보고서 및 분석전문가 지원 서비스 추진(‘21년~)



(바우처 서비스 예시)

다. 데이터 거래의 사례

○ 스마트팜 빅데이터 플랫폼

- 네이버클라우드의 노하우 및 인프라를 적극 활용하여 농업수요에 맞는 데이터 생산·유통과 혁신적인 서비스 모델 제공으로 스마트팜 생태계 조성
- 4차산업 혁명을 견인하는 스마트팜 빅데이터 플랫폼 활용을 통한 디지털 농업과 미래 농업 전환 기반 마련
- N 스마트팜 데이터 거래서에서 데이터 분석 결과를 쉽게 이해할 수 있도록 데이터 시각화에 따른 인사이트 제공

- 수요자는 기업, 일반인, 농민, 공공기관 등으로 구성되며 주요 서비스는 환경, 생육, 생산량, 판매, 질병 예찰/방역, 에너지소비/탄소저감 등으로 구성

< 스마트팜 빅데이터 플랫폼 거래 사례 >

상품 목록	유형	이용료
장미 병해충 방제방역 데이터	질병예측	11,000,000원
딸기 생육이미지 데이터	생육정보	10,000,000원
인삼 외부환경 데이터	생육정보	1,000,000원
대상 작물별 스마트-팜 환경제어관리 농작업 데이터	생육정보	1,000,000원
파프리카 내부환경 데이터에 따른 생산량	생산량정보	100,000원
송아지 체온, 우사 내・외부 온도 융합 데이터	생육정보	300,000원
방역 취약 농가 전담제 농가별 통계	질병예측	무료
임상예찰 통계	질병예측	300,000원
예방접종 점검 통계	질병예측	300,000원
딸기 혁신밸리 온실_유효전력량 데이터	에너지관련	무료

○ 중국 정부의 데이터 거래소(구이저우성 구이양시)

- 데이터 산업의 빠른 성장을 촉진, 기술・안전・가격 등을 표준화 함
- 거래소의 유통되는 데이터의 80%가 정부 데이터이고 민간의 참여는 높지 않다는 한계점

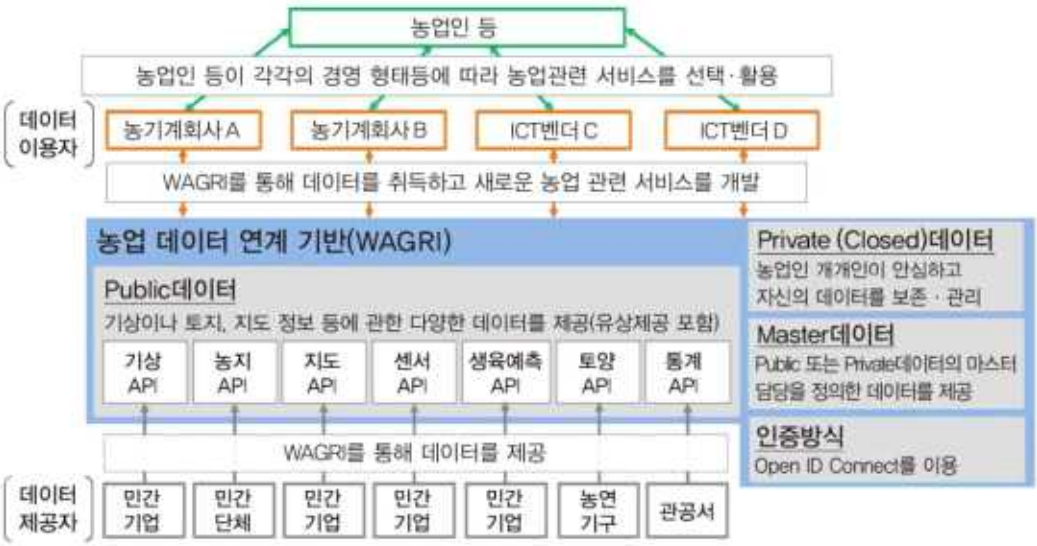
○ 수요조사 반영 서비스 카테고리 종류

- 환경정보, 생육정보 제공 서비스
  - 규격화된 데이터 인증 서비스 모델(친환경 인증, HACCP 등)
  - 환경측정/시설제어/생육 관리 모델
  - 작물 생산 최적화 제어 모델(대기, 관개, 양액)
  - 농산물 파동 예측 및 방지 모델
- 품질정보 제공 서비스
  - 생산품 품질(제품 등급) 상태 관리 모델
  - 환경제어 및 토양(배지) 관리 모델
- 생산량(수확량) 정보 제공 서비스
  - 생육 상태에 따른 생산량 예측 모델
  - 작물 개체 인식을 통한 생산량 예측 모델
  - 양돈 축사 생산성 향상 지원 모델

  · 농가 사료 사용량 및 사료 구매 현황 분석 모델
- 판매정보(가격, 소비량) 제공 서비스
  · 전국 기간별 소비량 및 판매 단가 검색 모델
  · 사료 가격 예측 모델
- 질병 예찰/방역 서비스
  · 질병 현황 및 이력 분석을 통한 예찰 및 방역 서비스 모델
  · 동물 영상 기반 AI분석 질병 예찰 모델
  · 동물 음성 기반 AI분석 질병 예찰 모델
  · 식물 영상 기반 AI분석 병해충 예찰 모델
  · 병해충 지역 발생 정보 공유 모델
  · 잡초 식별 모델
- 에너지/탄소 중립 서비스
  · 축사 냄새 민원 대응 서비스 모델
  · 돈사 악취 저감 관리 서비스 모델
  · 주변 농장 악취 정보 제공 모델
  · 에너지 저감 및 비용 최적화 모델
  · 가스 분석 정보 제공 모델(악취 저감 솔루션 기업)

○ 수요자 측면에서 요구되는 카테고리를 연계하여 제공
- 데이터 제공자에게 수집된 데이터를 수요자 측면의 카테고리로 분류하여 새로운 농업 서비스를 구축, 수요자 각각의 경영 형태 등에 따라 농업 관련 서비스를 선택·활용

○ 일본의 WARGI 데이터 플랫폼
- 데이터 기반 분석을 통한 전략적인 경영 수행 가능
- 카테고리별 데이터를 이용, 각 원인에 맞는 해결책을 구상하여 상품의 수확량 및 품질 향상
- 분류된 전문적인 경험과 감각을 데이터 기반으로 농업인들에게 노하우 제공



(WAGRI 데이터 거래 카테고리 제공 예시)

○ 일본의 WARGI 데이터 플랫폼 활용 사례

- Farm chat

(Farm chat 예시)

- WAGRI의 병충해 이미지 판정 프로그램, 청과물 시황 데이터를 FarmChat 연계
- 대화기능과 다양한 농업정보제공기능, 업무효율화기능을 갖춘 농업인용 스마트폰 어플
- 스마트폰 촬영 작물의 잎 이미지로 병해충 식별,적절한 방제 가능
- 시황 확인, 산지중계, 단경기 모니터링 가능. API 연계로 ICT벤더는 작업비용과 실수 경감

- 영농지원 어플

(영농지원 어플 예시)

- WAGRI의 청과물 시황데이터를 미도리클라우드 영농지원어플 연계
- 기상/시황정보 확인, 농약 검색, 생산/수확계획의 등록이 가능한 농업인 대상으로 한 무료 어플
- 스마트폰 등으로 청과물의 일일 가격변동과 거래수량 확인 가능, 농업인은 이 데이터를 기초로 출하량을 적절하게 판단

- ArcGIS Online

(ArcGIS Online 예시)

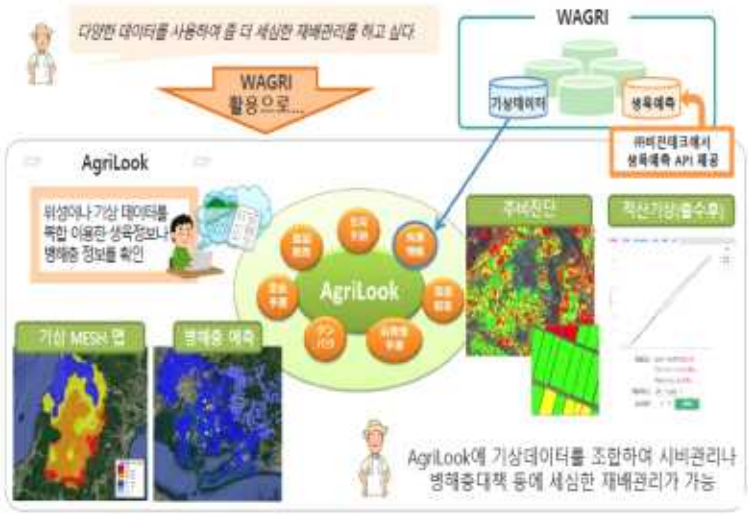
- WAGRI의 지적도, 농지핀, 토양지도, 기상데이터를 ArcGIS Online 연계, WAGRI 인증시스템과의 연계로 WAGRI 계정을 사용하여 ArcGIS Online 로그인 가능
- 다양한 지리공간정보를 다양한 디바이스에서 이용 가능한 환경을 제공하는 클라우드방식 GIS 서비스
- 많은 공공기관, 기업 등에서 이용되는 ArcGIS가 WAGRI와 연계, ArcGIS 콘텐츠나 자신의 GIS 데이터와 WAGRI를 용이하게 중첩하여 사용가능

- NEC 영농지도지원시스템

(NEC 영농지도지원시스템 예시)

- WAGRI의 지적도와 농약정보, 1km MESH 추정 기상정보를 활용
- 영농데이터 수집, 지도상에서 작업, 생육상황 파악 등에 따라 생육목표에 근거한 지도나 정보 공유를 가능하게 하는 시스템

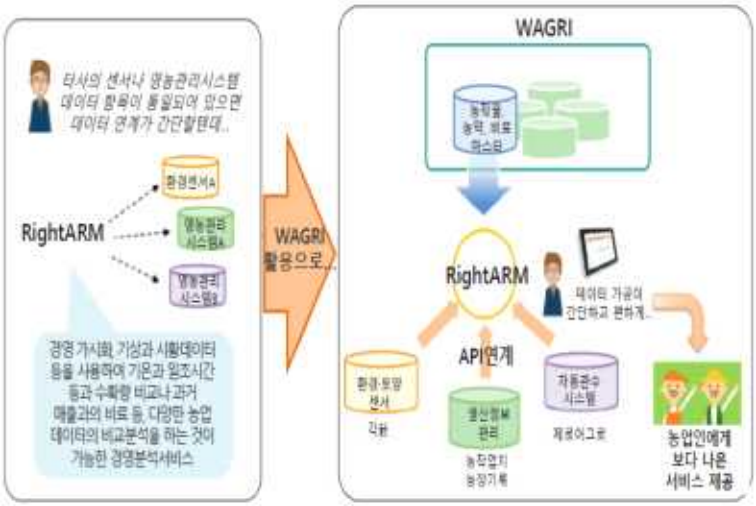
- AgriLook



(AgriLook 예시)

· WAGRI의 기상 데이터와 자사의 위성화상 및 생육예측 모델을 연계, 생육 단계에 맞는 시비관리나 병충해대책 등에 세심한 재배관리가 가능하도록 AgriLook 제공

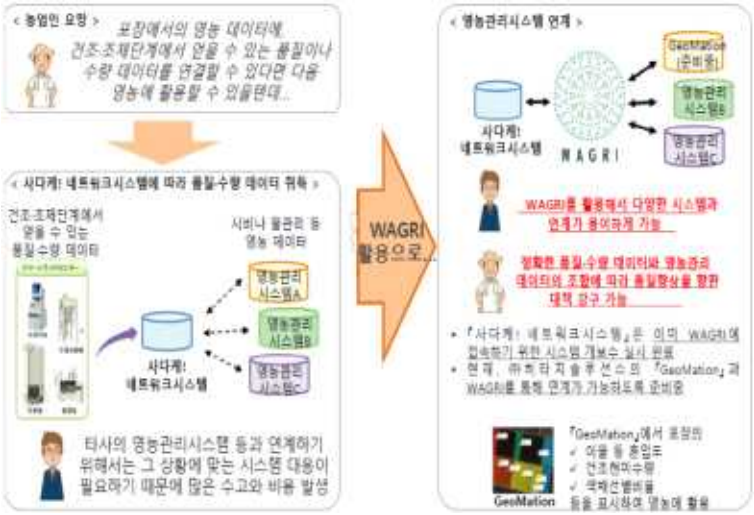
- RightARM



(RightARM 예시)

· 경영분석 서비스 RightARM에서는 다른 기업과의 시스템 연계를 원활하게 하고자, WAGRI상의 마스터 데이터(농작업,농약,비료 등)을 표준으로 취급

- 사타게! 네트워크시스템



(사타게! 네트워크시스템 예시)

- 곡물 건조, 조제단계에서 얻을 수 있는 정확한 품질, 수량 데이터를 영농관리 데이터에 연결하여 품질향상대책을 강구할 수 있도록 데이터를 관리하는 클라우드시스템 사다케! 네트워크시스템과 다양한 영농관리시스템이 연계되도록 WAGRI 활용

라. 예상 문제점 및 개선방안

○ 현황 및 문제점

- K-data에서 운영중인 데이터 품질인증(DQC)은 공공·민간에서 구축·활용중인 데이터베이스를 대상으로 도메인, 업무규칙을 기준으로 품질영향요소 전반을 심사하여 인증하는 방식으로 진행 중
- 데이터 관리수준(도입-정형화-통합화-최적화)을 심사 인증하는 데이터 관리인증 및 데이터 보안관리수준(접근제어, 암호화, 작업결제, 취약점분석)을 인증하는 보안인증이 운영 중임
- 안전하고 신뢰할 수 있는 데이터 거래시장 환경이 조성되기 위해서는 데이터 거래 시 권리보호(예: 지식재산권)와 부정사용 등에 따른 법적인 리스크 해소 방안 마련 필요
- 구매 데이터를 활용한 경제적 효익 창출, 보유한 데이터 기반으로 거래 및 대출 등을 위해 데이터를 회계상 자산으로 처리 필요성 증가
- 국내 기업은 회계기준의 모호함으로 데이터를 재무제표에 반영하지 못하거나, 데이터 거래를 비용으로만 처리함에 따라 데이터의 가치평가나 회계처리에 한계 존재
- 데이터산업법에서 경제적 가치를 가지는 데이터를 '데이터 자산'으로 정의하고 있으며, 부정경쟁방지법에서는 데이터자산의 부정사용 금지 행위를 규정

○ 개선 방안

- 정부에서 거래되는 데이터 품질평가를 위한 지표, 가이드라인 등을 마련하고, 데이터산업법에 따른 품질인증기관은 데이터거래소에서 판매자가 등록한 데이터에 대한 품질인증 실시
- 데이터 품질인증 평가 요소는 데이터 유형에 따라 달라질 수 있지만, 데이터 표준화, 완전성, 정확성 등에 대해 공통적으로 측정 필요
- 데이터에 대한 법적인 권리보호와 공정한 데이터 유통환경 조성을 위해 표준계약서 기준 마련, 법률상담을 통한 분쟁 예방 등을 주된 기능으로 하는 '데이터거래 법률지원센터' 설치·운영
- 데이터 거래 경험이 부족한 중소기업, 스타트업 등이 데이터 거래 계약 시 고려해야 할 표준계약서(안),

체크리스트 제공 등 컨설팅을 지원하여, 관련 법적 이슈 발생 최소화
- 계약조건별 옵션화로 선택 사용 가능한 표준계약서(안) 마련, 계약내용에 대한 검토 등 자문 기능 지원
- 데이터 자산에 대한 지식재산권(IP) 등록 시 정부의 적극 지원을 통해 데이터 자산에 대한 권리보호 강화
- 계약 및 거래과정, 거래 이후의 분쟁조정에 관한 사항은 데이터산업법에 따른 '분쟁조정위원회'에서 담당
- 데이터 판매자, 거래소 등은 품질평가에 따른 인증 결과를 등급 형태로 개별 판매 데이터에 표시하여 구매 시 부가정보로 제공
- 가치평가 시 품질인증 결과를 포함하여 추진할 경우 품질인증 활성화 견인
- 품질인증 기관이 민간·공공 데이터거래소의 데이터 품질관리 수준(도입-정형화-통합화-정량화-최적화), 보안관리 수준 등을 인증, 고시하는 품질인증 체계 도입
- 데이터 구매자가 데이터거래소 선택 시 전문분야, 경력, 가격 외에 기관의 품질관리 수준 활용 가능
- 기업은 필요에 따라 데이터를 (무형)자산으로 처리하고, 중장기적으로 일반기업회계기준에 데이터를 자산으로 처리하는 근거 마련 필요
- 데이터산업법에 따른 경제적 가치를 가지는 데이터는 일반기업 회계기준 상 무형자산 식별요건*과 인식요건을 모두 충족
- 기업은 우선적으로 지식재산권 등록, 회계기준원 질의 등의 방법을 통해 무형자산으로 처리 가능
- 보유한 데이터 기반으로 M&A, IPO, 담보대출 등에 필요한 경우, 산업은행의 데이터 담보대출 사례와 같이 플랫폼·앱·데이터를 포함하여 지식재산권으로 등록하고 이를 무형자산으로 처리
- 구매한 데이터의 회계 처리 등을 위한 경우, 기업마다 처한 상황, 데이터의 성격, 비용처리 시점 등 상이하므로 회계 기준원의 질의·회신을 통해 무형자산 등록 여부 판단
- 한국회계기준원은 질의회신 사례가 축적되면 향후 일반기업회계 기준서 개정시 일정 요건을 충족하는 데이터의 경우 무형자산으로 처리 가능하도록 기준서 개정 검토 필요

마. 데이터거래를 위한 스마트농업육성법 수정 제안

○ 스마트농업육성법은 농업 및 연관 산업과 첨단 정보기술 등의 융합을 통해 농업의 자동화·정밀화·무인화 등을 촉진함으로써 농업인과 농업·농촌의 성장·발전에 기여하는 것을 목적으로 한 법률임

○ 스마트농업육성법에서 제정한 '스마트 농업 데이터'란 스마트 농업에 활용되거나 스마트 농업과정에서 생산되는 생육환경 및 상황 등을 나타내는 기계적 활용이 가능한 수치·문자·영상 등 정보를 말함

○ 스마트 농업 육성에 적용되는 지원 및 특례 등에 관하여 다른 법률에 특별한 규정이 있는 경우를 제외하고는 이 법에서 정하는 바에 따르며, 본 보고서에서는 스마트 농업 데이터 거래를 위한 스마트 농업 육성법 수정을 제안하는 바임

(1) 제 2장 스마트 농업 육성 및 지원 체계

< 조문 대비표 >

조항	현행	개정안
제 5조 (육성계획의 수립)	② 육성계획에는 다음 각 호의 사항이 포함되어야 한다. 7. 스마트 농업 데이터 수집·분석 및 활용 및 거래에 관한 사항	② 육성계획에는 다음 각 호의 사항이 포함되어야 한다. 7. 스마트 농업 데이터 활용 및 거래에 관한 사항
제 7조 (시도계획의 수립 등)	① 시·도지사는 관할 지역의 스마트 농업화를 촉진하기 위하여 육성계획을 반영하여 다음 각 호의 사항이 포함된 5년 단위의 시·도 스마트 농업 육성계획 (이하 “시·도계획”이라 한다)을 수립하여야 한다. 6. (신설)	① 시·도지사는 관할 지역의 스마트 농업화를 촉진하기 위하여 육성계획을 반영하여 다음 각 호의 사항이 포함된 5년 단위의 시·도 스마트 농업 육성계획 (이하 “시·도계획”이라 한다)을 수립하여야 한다. 6. 스마트 농업 데이터 활용 및 거래에 관한 사항

(2) 제 3장 스마트 농업 육성 및 지원

< 조문 대비표 >

조항	현행	개정안
제 21조 (표준화의 추진)	① 정부는 스마트 농업의 확산 및 스마트 농업 기자재·데이터 산업 등의 육성을 위하여 다음 각 호와 같은 표준화 사업을 추진 할 수 있다. 2. 스마트 농업 데이터 수집 방식 등에 대한 표준 개발 5. 스마트 농업 단체표준 제정 지원	① 정부는 스마트 농업의 확산 및 스마트 농업 기자재·데이터 산업 등의 육성을 위하여 다음 각 호와 같은 표준화 사업을 추진 할 수 있다. 2. 스마트 농업 데이터 수집 방식, 무결성, 신뢰성 등에 대한 표준 개발 5. 스마트 농업 관련 표준 제정 지원
제 24조 (스마트 농업 데이터 기반 구축)	① 농림축산식품부장관, 시·도지사 또는 시장·군수·구청장은 스마트 농업의 고도화를 뒷받침하기 위하여 데이터 수집·제공·처리 등에 필요한 물적 기반 구축에 노력하여야 한다. ② 제1항에 따른 스마트 농업 데이터 기반은 경영체, 품목, 지역 및 국가 등 수준별로 구분하여 추진할 수 있다. ③ 농림축산식품부장관은 시·도지사 또는 시장·군수·구청장이 제1항에 따른 데이터 수집·제공·처리 등에 필요한 물적 기반 구축을 추진하는 경우 필요한 기술적·재정적 지원을 할 수 있다.	① 농림축산식품부장관, 시·도지사 또는 시장·군수·구청장은 스마트 농업의 고도화를 뒷받침하기 위하여 데이터 활용 및 거래 등에 필요한 물적 기반 구축에 노력하여야 한다. ② 제1항에 따른 스마트 농업 데이터 기반은 경영체, 품목, 지역 및 국가 등 수준별로 구분하여 추진할 수 있으며, 스마트 농업 데이터 활용은 수집, 연계, 저장, 보유, 가공, 분석, 이용, 제공, 공개 및 그 밖에 이와 유사한 행위를 말한다. ③ 농림축산식품부장관은 시·도지사 또는 시장·군수·구청장이 제1항에 따른 데이터 활용 및 거래 등에

		필요한 물적 기반 구축을 추진하는 경우 필요한 기술적·재정적 지원을 할 수 있다.
제 25조 (데이터 생산·수집)	② 농림축산식품부장관은 제27조의 데이터 플랫폼을 활용하여 농업인, 기업, 공공기관, 민간단체, 법인 등으로부터 스마트 농업 데이터를 수집할 수 있다. ③ 농림축산식품부장관은 제2항에 따라 농업인, 기업, 민간단체, 법인 등으로부터 스마트 농업 데이터를 수집하는 경우 데이터를 제공하는 자에게 대가를 지불 할 수 있다. ④ 농림축산식품부장관은 스마트 농업 데이터를 수집하는 농가 등을 대상으로 제21조 제2호에 따라 개발된 스마트 농업 데이터 수집 방식 등에 대한 표준 준수 여부를 확인하고, 확인을 받은 농가 등에게 필요한 지원을 할 수 있다. 이 경우 확인 절차 등 필요한 사항은 농림축산식품부령으로 정한다.	② 농림축산식품부장관은 제27조의 데이터 플랫폼을 활용하여 농업인, 기업, 공공기관, 민간단체, 법인 등이 스마트농업 데이터를 활용하여 사용·수익할 권리를 가진다. ③ 농림축산식품부장관은 제2항에 따라 농업인, 기업, 민간단체, 법인 등으로부터 스마트 농업 데이터를 활용하는 경우 데이터를 제공하는 자에게 대가를 지불 할 수 있다. ④ 농림축산식품부장관은 스마트 농업 데이터를 수집하는 농가 등을 대상으로 제21조 제2호에 따라 개발된 스마트 농업 데이터 수집 방식, 무결성, 신뢰성 등에 대한 표준 준수 여부를 확인하고, 확인을 받은 농가 등에게 필요한 지원을 할 수 있다. 이 경우 확인 절차 등 필요한 사항은 농림축산식품부령으로 정한다.
제 26조 (데이터 유통 ·거래 지원)	제목 변경 : 데이터 제공·활용·유통·거래 지원 ① 농림축산식품부장관은 제25조에 따라 수집된 데이터를 제27조의 데이터 플랫폼을 활용하여 제공할 수 있다. 이 경우 데이터 제공 방식·범위 등은 대통령령으로 정하는 바에 따른다. ③ 농림축산식품부장관은 데이터의 생산자와 수요자간 거래가 활성화 될 수 있도록 데이터 중개 등을 지원할 수 있다.	제목 변경 : 데이터 유통·거래 지원 ① 농림축산식품부장관은 제25조에 따라 수집된 데이터를 제27조의 데이터 플랫폼을 활용하여 거래할 수 있다. 이 경우 데이터 제공 방식·범위 등은 대통령령으로 정하는 바에 따른다. ③ 농림축산식품부장관은 데이터의 생산자와 수요자간 거래가 활성화될 수 있도록 데이터 중개 등을 지원할 수 있다. 다만, 당사자 간의 약정이 있는 경우에는 그에 따른다.
제27조 (데이터 플랫폼 구축 및 운영)	② 농림축산식품부장관은 관계 중앙행정기관의 장, 시·도지사, 시장·군수·구청장, 「공공기관의 운영에 관한 법률」 제4조에 따른 공공기관의 장, 그 밖에 대통령령으로 정하는 기관(이하 “중앙행정기관등” 이라 한다)의 장에게 데이터의	② 농림축산식품부장관은 관계 중앙행정기관의 장, 시·도지사, 시장·군수·구청장, 「공공기관의 운영에 관한 법률」 제4조에 따른 공공기관의 장, 그 밖에 대통령령으로 정하는 기관(이하 “중앙행정기관등” 이라 한다)의 장과 협의를

	공유 등 데이터 플랫폼의 구축 및 운영에 관한 협조를 요청할 수 있다. 이 경우 중앙행정기관등의 장은 정당한 사유가 없는 한 협조하여야 한다. ④ 농림축산식품부장관은 제3항에 따라 데이터 플랫폼을 위탁 운영하는 경우에 데이터 플랫폼의 운영 및 유지·보수에 필요한 비용 등을 지원할 수 있다.	거쳐 데이터의 공유 등 데이터 플랫폼의 구축 및 운영에 관한 지침을 마련할 수 있다. ④ 농림축산식품부장관은 제3항에 따라 데이터 플랫폼을 위탁 운영하는 경우에 데이터 플랫폼의 운영 및 유지·보수, 데이터 분실, 도난, 유출, 위조, 변조 또는 훼손되지 아니하도록 하며, 이에 필요한 비용 등을 지원할 수 있다.
제28조 (데이터 센터의 설치)	① 농림축산식품부장관, 시·도지사 또는 시장·군수·구청장 스마트 농업 데이터의 생산·수집·제공 등에 필요한 장비 및 시설 구축을 위하여 데이터 센터를 설치할 수 있다. ③ (신설)	① 농림축산식품부장관, 시·도지사 또는 시장·군수·구청장은 스마트 농업 데이터의 생산·수집·제공 등에 필요한 장비 및 시설 구축을 위하여 데이터 센터를 설치할 수 있다. ③ 농림축산식품부장관, 시·도지사 또는 시장·군수·구청장은 스마트 농업 데이터의 적절한 품질관리 및 거래 등을 촉진하는 데 필요한 시책을 마련할 수 있다.

8. 결론 및 시사점

가. 연구용역 사업 보고서의 결론

(1) 본 연구사업의 주요 목표

○ 본 연구과제는 스마트 농업 데이터의 연계·이용 활성화 및 거래 환경 조성을 위해 농업분야 데이터 가치평가 모델 및 거래 가이드라인 등을 마련하는 것을 목표로 수행되었음

○ 따라서 국내외 스마트 농업 데이터 거래·유통 시장 동향을 파악하고 타 분야 데이터 거래 가이드라인 개발 사례, 거래 가능한 데이터 상품화 유형 조사 등을 수행함으로써 궁극적으로 농업분야 데이터 가치평가 모델을 개발하는 연구가 수행되었음

○ 농업 분야 데이터 가치평가 모델 개발은 데이터 생성 원가, 구매수요, 거래유형 등 데이터 가치 평가시 주요 고려사항 도출 등을 통해 모델이 마련되었으며, 이를 통해 스마트 농업 데이터 거래환경 및 운영 모델이 설계 되었음

○ 데이터 거래 시 국내외 법률적 쟁점사항(거래방식, 대금지급, 개인정보 보호 및 분쟁 발생 시 해결방법, 데이터 거래 플랫폼 이용시 예상되는 문제점 해결사항 등)의 검토를 통해 데이터 유형별 거래 가이드라인 및 표준 계약서 마련을 위해 데이터 관련법률 및 표준문서의 검토가 이루어졌음.

○ 스마트 농업 데이터 거래환경 설계 및 운영 로드맵을 제시함으로써 향후 공공기간, 민간, 단체 등 데이터 거래·유통 운영체계(안)을 마련하여 스마트 농업 분야 데이터 거래 및 확산에 기여하고자 함

(2) 제 1장 스마트팜 환경 데이터 분류방안 및 가치평가 방법 조사 및 검토

○ 스마트팜 환경 데이터 분류 방안의 경우 기본 데이터 유형과 분류 체계를 제시하고 환경 데이터, 작물 생육 데이터 등의 데이터 분석 방법과 데이터 분석을 위한 세부항목등의 표준단어, 측정단위를 정의하여 조사를 위한 지표로 활용하고자 함

○ 현재 농림축산 식품부의 공공데이터 주제별 현황은 총 다섯 가지 주제가 있으며, 제공기관은 총 19개 기관, 총 공공데이터 건수는 1,141 건으로 제공되고 있음.

○ 기준데이터 20개 이상으로 데이터를 제공한 기관은 농림축산식품부, 국립농산물품질관리원, 국립종자원, 농림축산검역본부 순으로 조사되었음

○ 가치평가 방법 조사의 경우 시장 접근법, 원가 접근법, 소득 접근법을 중점으로 데이터베이스 등의 무형 자산의 가치를 활용하는 방법에 대해 구체적으로 제시하였음

○ 또한 데이터 가치평가를 위한 4가지의 요소(컨텐츠 적합성, 품질·공급 신뢰성, 기술적 사용성, 경제성)에 대한 상세 내용과 평가 지표를 제 1장에서 제시하였음

(3) 제 2장 국내 데이터 거래 관련 플랫폼 및 가격산정 사례 조사

○ 국내 데이터 거래 관련 플랫폼 현황은 공공 주도 180개, 민간 주도 21개, 총 129개 데이터 플랫폼이 구축되어 운영중에 있음

○ 국내 플랫폼 산업 현황 조사를 위한 참고문헌으로 한국 IDC 연구보고서를 분석하였으며,

빅데이터 분석 시장은 2021년 대비 5.5% 성장하여 2조 296억원의 매출 규모를 형성할 전망이며, 향후 5년간 연평균 성장률 6.9%를 기록하여 2025년까지 2조 8,353억원 규모에 이를 전망으로 분석되었음

○ 국내 데이터 플랫폼 구축 및 운영 현황으로, 현재 운영중인 플랫폼으로는 농정 빅데이터 플랫폼, 농식품 빅데이터플랫폼(KADX), 농산물 유통 종합정보 플랫폼, 스마트팜 빅데이터 플랫폼, 축산업 통합정보 플랫폼, 스마트팜 R&D 빅데이터 플랫폼으로 총 549종의 데이터를 보유하고 있는 상황으로 조사되었음

○ 다만 개별데이터의 생산 및 관리가 사업단위로 이루어지고 있어 체계적 관리가 어렵다는 문제가 도출되었으며, 민간 데이터의 경우 데이터 품질지표가 있어 품질수준을 평가하고 있으나, 농식품 공공데이터에는 평가기준이 없다는 문제가 조사되었음

○ 해외 데이터 플랫폼 현황으로는 주요 데이터 국가인 EU, 미국, 중국을 대상으로 조사하였으며, EU의 경우 독일, 프랑스 주도로 약 500여개 기업・기관이 참여하는 클라우드 기반의 데이터 프로젝트 가이아-X를 추진중임

○ 미국의 경우 연방정부 주도로 기업, 연구자 등이 활용할 수 있는 오픈데이터 포털을 통해 공공데이터를 개방하였으며, 중국의 경우 각 지역에 민・관 합작 또는 민간 주도로 오픈마켓 형태의 16개 데이터 거래 플랫폼을 운영중에 있는 것으로 조사되었음

○ 국내 데이터 거래 관련 가격산정 사례 조사로는 데이터상품 및 제공 방식/데이터 가격책정의 목표/가격 산정의 방법/인건비 책정 등으로 각각 공공기관, 민간기관에 대해 조사를 진행하였음

○ 스마트 농업 데이터의 거래 가능한 상품화 유형 조사로는 토마토, 오이, 파프리카, 방울토마토, 버섯, 딸기의 원천데이터와 공정단계를 제시하였으며, 스마트팜 데이터 마트 수집 데이터를 토대로 거래 가능한 상품화 유형에 대해 정리하였음

(4) 제 3장 표준확산사업 연계 방안 및 관련 표준화 준수사항 도출

○ 표준 확산사업 연계방안을 제시하기 위해 먼저 데이터 거래의 문제점과 원활하지 않는 이유에 대해 분석을 진행하였음.

○ 국가통계표털(KOSIS)의 조사 결과에 의하면 2021년 기준 응답기업의 52.1%가 쓸만한 양질의 데이터 부족, 37.0%가 불합리한 데이터 가격 책정, 36.3%가 데이터 유통 채널 부족 순으로 문제점이 발견 되었음

○ 동시 다발적인 정보시스템 개발, 전산 데이터 관리 마인드 미형성 등의 원인으로 인해 발생되는 문제로는 데이터의 중복 및 조직, 업무, 시스템별 데이터 불일치, 데이터 통합의 어려움, 정보시스템 변경 및 유지보수 곤란 등이 조사되었음

○ 이를위해 개선할 사항으로는 데이터 표준화, 규격화를 위한 기본 방침을 설정하고 전산적인 데이터 요소 등록 및 관리 체계를 구축하여 명칭의 통일로 인한 명확한 의사소통의 증대를 이끌어 내고 데이터 소재 파악에 소요되는 시간 및 노력을 감소하는 효과가 발생하도록 노력해야 함

○ 데이터 가치평가 시 표준화 준수사항으로 먼저 유통・거래 측면에서 수요의 불확실성과 플랫폼을 통한 유통 채널이 부족하다고 인식되고 있으므로 이용자가 데이터 플랫폼을 통해 데이터에 쉽게 접근・검색하고 믿고 거래할 수 있도록 편의성과 신뢰성을 보장하는 유통・거래 역량을 강화하고자 데이터 거래 표준화를 추진해야함

○ 또한 연계・결합 측면에서 수집 및 축적되는 데이터의 표준화를 통해 데이터 전주기

(표준화-수집-축적-가공-활용)의 선순환 생태계를 만들어 데이터 활용성과를 향상하기 위해 데이터 통합관리를 추진해야할 필요가 있음

○ 데이터 확충 측면에서 공공·민간 기준 수요와 활용가치가 높은 데이터의 지속적인 발굴 및 개방 촉진을 위해 데이터 표준화가 필요하며, 공통 표준 규격을 적용하여 AI 가공 표준 제시 및 국가 전반의 데이터 통합 검색 기반이 마련되어야 함

○ 거버넌스 및 기계가독성 측면에서는 개별적으로 산재되는 데이터 플랫폼 중복성 방지 및 데이터 통합 관리를 위한 데이터 거버넌스를 마련해야 하며, 이러한 데이터를 읽을 수 있는 형태로 가공하여 AI 알고리즘 개발에 적용하고 이를 활용할 수 있도록 해야함

(5) 제 4장 데이터 거래를 위한 정보 제공 방식

○ 데이터 거래를 위한 정보 제공 방식 중 먼저 데이터 거래 계약사항에 대한 조사를 진행하였으며, 데이터 거래 계약을 위해 핵심적으로 고려할 사항으로는 계약대상 범위설정 -> 데이터 이용목적 및 방법 -> 데이터 이용허락 권리범위 -> 데이터 제공방법 -> 데이터 이용과금 및 정산 -> 침해예방 설정 -> 계약의 해지 순으로 조사되었음

○ 계약대상 범위 설정을 위해 5가지의 방법(계약대상 명확화, 데이터 권리관계 확인, 계약당사자 확인, 개인정보 포함 확인, 신용정보/위치정보 등 관계법령 준수 확인)을 이용하여 검토를 진행해야함

○ 데이터 이용목적 및 방법을 위해서는 다음의 3가지 방법(데이터 이용목적 서비스 확인, 데이터 가공여부 확인, 계약 기간 결정)을 이용하여, 어떤 목적으로 데이터를 이용할 것인지, 어느 기간동안 이용할 것인지 파악함으로써 예상치 못한 문제를 방지할 수 있도록 해야 함

○ 데이터 이용허락 권리 범위를 위해서는 다음의 4가지 방법(이용허락 권리 설정, 재이용허락 허용여부, 독점/비독점 이용허락 여부/권리의 양도)을 이용하여 검토를 진행함으로써 구매자가 데이터 사용에 있어 어느정도의 권리를 가지고 있는지 명확히 정해야 함

○ 데이터 제공방법으로써 다음의 2가지 방법(포맷 및 전달방법, 제공 데이터의 품질 수준)을 이용하여 데이터의 품질 수준, 제공주기, 보증 등에 대한 사항을 파악하고 어떠한 포맷으로 제공할지, 어떤 방식으로 전달할지 확실히 정해야 함

○ 데이터 이용료 및 정산을 위해서는 다음의 2가지 방법(이용료 산정, 이용료 정산 방법)을 이용하여 최종 가격이 결정되고 구매자와 판매자 모두가 동의했는지 확인 후 지불 방법과 시기를 구체적으로 결정해야 예민한 문제를 방지할 수 있음

○ 데이터 권리 침해예방 설정 방법으로는 다음의 4가지 방법(제 3자 복제/배포/전송 등 금지, 데이터 출처 표시, 기술적 보호조치, 이용 조건 준수 확인)을 이용하여 판매자의 권리를 보호해야함

○ 계약의 해지 방법으로는 다음의 2가지 방법(계약 해지 방법의 결정, 계약 종료 효과)을 이용하여 데이터 거래의 끝을 명확히 함으로써 계약 종료효과, 잔존 권리 의무 등을 명시할 수 있어야함

(6) 제 5장 국내외 AI 산업의 오픈 데이터 활용 사례

○ 국내외 AI 산업의 오픈 데이터 활용 사례 조사를 위해 먼저, 국외 오픈 데이터 현황에 대한 조사가 진행되었음

○ 미국은 2009년 5월 ‘오픈 거버넌스 이니셔티브(Open Governance Initiative)’에 의거해

정부가 보유하고 있는 대량의 데이터를 제공하는 오픈 데이터 포털 'DATA.GOV'를 개설함.

○ 영국은 2018년 3월 '오픈 데이터 찾기(Find OpenData)'라는 명칭의 공공 데이터 개발 포털을 운영 중이며, 쉽고 빠르게 원하는 데이터를 검색 및 다운로드할 수 있으며, 관심 있는 데이터에 대한 업데이트 알림 기능을 제공함.

○ 호주는 2013년 중앙정부 차원의 오픈 데이터 포털 'data,gov.au'를 개설하여 연방정부, 주정부, 정부 산하기관 등에서 생성된 공공 데이터 이외에 공공 연구 데이터나 민간 조직에서 생성된 공공 분야 관련 데이터도 포함하고 있으며, 약 3만 개 이상의 데이터세트를 제공하고 있음

○ 일본은 2014년 10월 1일 오픈 데이터 카탈로그 사이트인 'data.go.jp'의 운영을 개시하여 공공 데이터를 분야별로 검색 가능할 뿐만 아니라 각 부처별로도 검색할 수 있으며, 데이터 제공 관련 기관의 링크 정보와 데이터 활용 사례도 체계적으로 제공하고 있음

○ 캐나다 정부는 2011년 3월 1세대 오픈 데이터 포털인 'data.gc.ca'를 시범 운영한 후 2013년에 현재의 오픈 데이터 포털인 'open.canada.ca'를 개설하여 오픈 데이터 검색 캐나다 정부에서 제공하는 정책, 법률, 재무, 규제 등 모든 종류의 정보를 검색하는 '오픈 정보(OpenInformation)' 정부 정책 및 프로그램 개발에 참여할 수 있는 모든 방법을 한눈에 확인할 수 있는 '오픈 다이얼로그(Open Dialogue)' 등의 기능을 제공하고 있음

○ 국제 기구의 오픈 데이터 포털로는 OECD 데이터, EU 오픈 데이터 포털, 유럽 데이터 포털, UN 데이터 카탈로그로 조사되었으며, 각 오픈 데이터 포털에서는 관련 분석 보고서를 비롯, 재사용 사례, 데이터 뉴스, 교육 자료, 데이터 활용 지원을 위한 API, 시각화 자료 등을 체계적으로 제공하고 있음

○ 국내 주요 오픈 데이터 플랫폼과 활용 사례로 먼저, 공공데이터 포털의 경우 행정안전부와 한국정보와 진흥원에서 운영하고 있으며, 공공행정, 국토교통, 산업고용, 농축수산 등의 공공데이터를 공유 및 데이터 시각화, 사례 공유 등을 제공하고 있음

○ 한국정보화진흥원에서 운영중인 AI 허브는 AI데이터의 구축/확산을 위한 사업의 일환으로 시작되어, AI 데이터, 소프트웨어, 컴퓨팅 자원 등을 제공하고 있음

○ 국외 민간 및 학계의 데이터 개방/공유 현황 조사 결과, 데이터 협력자는 다양한 분야의 행위자들이 데이터를 교환하고 분석적 전문지식을 취합하여 복잡한 공공 문제를 해결하는 데 도움을 주고 있었으며, 공익에도 기여하기 위한 협력을 용이하게 하는 측면이 있다고 도출 되었음

○ 대표적인 사례로 우버의 무브먼트 오픈데이터 포털, 구글 리서치 데이터 세트, 캐글, 아마존 RODA가 있음

○ 주요 국내 민간 주도 데이터 공개와 활용 사례 조사 결과, 네이버 데이터랩, SKT 빅 데이터 허브, LG CNS 등이 조사되었으며, 세곳 모두 공개된 데이터를 무료로 공유할 뿐 아니라 분석 기능도 제공한다는 특정이 도출되었음

(7) 제 6장 데이터 가격 결정 평가 및 절차

○ 데이터 가격 결정 평가를 위한 절차로는 데이터 가격 결정의 목표 -> 수요 고객 예측 -> 원가 측정 -> 경쟁환경 분석 -> 가격책정 방법 선정 -> 데이터 제공방식 결정 -> 데이터 과금방식 및 할인정책 결정 순으로 조사되었음

○ 먼저 데이터 가격 결정의 목표는 데이터 상품을 얼마에 판매해서 어떠한 결과를 얻을 것인지 고민해야 하며, 매출 중심적 가격 목표, 이윤 중심적 가격목표, 현상 유지적 가격

목표 등으로 구분됨

○ 다음으로 가격 영향 요인 분석으로는 다음과 같은 4가지 요인(원가, 고객, 경쟁, 환경)을 고려하여 가격 영향 요인에 대한 분석을 거침으로써 판매가격이 정해짐

○ 가격 책정방법 선정으로는 비용을 충당하고 목표이일을 낼 수 있는 수준에서 가격을 결정하는 방법인 원가 기준 가격 책정, 경쟁사의 가격을 고려해 자사의 데이터 가격을 결정하는 방법인 경쟁사 기준 가격 책정, 구매자가 인지하는 평가·수요·가치를 바탕으로 데이터 가격을 결정하는 방법인 가치 기준 가격 책정으로 구분됨

○ 판매 및 가격 전략 수집을 위한 방법으로 먼저 데이터 제공 방식의 결정을 위해 서비스를 통해 데이터를 전달하는 경우와 데이터를 직접 제공하는 경우로 나뉨.

○ 서비스를 통해 데이터를 전달하는 경우 데이터를 직접 전달하지 않고 데이터가 활용된 부가가치를 제공하여 그에 따른 가격을 매기므로, 국내에서는 미약하나 해외에서는 데이터 비즈니스의 대표적인 형태로 자리잡는 중임

○ 데이터를 직접 제공하는 경우 데이터 파일을 FTP 또는 E-mail 등으로 전달하는 방법, Open API로 제공하는 방법, 구매자의 시스템에 직접 접속해서 제공하는 방법이 있음

(8) 제 7장 스마트팜 데이터 거래환경(시스템) 구축

○ 본 연구용역의 목적인 데이터 거래환경 시스템 구축에 대한 시스템 운영 모델로써 다음과 같은 안건을 제시하였음

((가칭)스마트농업 데이터 거래소(안))

○ 해당 안건에서 고려해야할 사항은 공유/개방형 상품, 데이터 검색, 결재, 다운로드/API, 연계, 통계 공유 등 거래기능을 통한 유통 모델, 빅데이터 플랫폼 데이터 가격 및 판매 정책, 혁신 모델 발굴, 데이터 마켓 서비스, 기존 데이터와 차별화 방안, 유통 및 거래

표준약관 수립 방안이 도출되었음

○ 기존 개방된 데이터의 차별성을 통해 플랫폼의 가공 및 재생산을 통한 공유/개방형 데이터 상품을 생산할 수 있어야 하며, 데이터 공급, 판매, 구매 등 플랫폼 내에서 업무 프로세스가 처리될 수 있도록 빅데이터 플랫폼 구축 운영을 진행해야함

○ 또한 기존 데이터와 플랫폼 데이터를 개방함과 동시에 유통 및 거래 표준약관을 수립하여 공공 데이터의 제공에 따른 보호 조치를 수행할 필요가 있음

○ 운영 모델 가동을 위한 요구사항으로 먼저 데이터 활용 서비스 개발 지원사항이 있음.

○ 데이터 분석 기술이 부재한 기업의 수요를 확보하여 데이터 가공 기업 연계를 통한 기업 수요에 맞는 데이터와 서비스 개발 지원을 통해 신산업 창출 및 서비스 효율화를 이끌어 내야함

○ 다음으로 데이터 안심구역 확대를 위해 중소기업, 스타트업 등이 데이터를 자유롭게 분석하고 결과만을 반출 할 수 있는 오프라인 분석환경 구축을 진행해야 함

○ 기반구축을 위해 농식품 빅데이터 활용 지원 조직, 빅데이터 플랫폼, 빅데이터 생성-활용 관련 주체 간 유기적인 네트워크 등의 기반 구축이 우선되어야 하며, 본격적인 생산/유통/소비 분야와 후방산업이 연계되어 데이터가 자동적으로 수집될 수 있도록 해야함

○ 데이터의 활용을 위해 각 기관별로 현안 대응 등 정책적 활용과 대국민 서비스 분야에 대한 실현 가능성과 개선 중심으로 로드맵 수립이 되어야 하고, 민관 및 학계가 함께 중점 과제 발굴 및 단계적 추진 방안 수립이 필요함

○ 스마트 농업 거래환경 구축 시 예상되는 문제점 및 개선 방안으로, 회계 기준의 모호함으로 인해 발생되는 데이터의 가치평가나 회계처리에 한계가 존재할 수 있다는 점이 대표적으로 도출되었음

○ 이러한 문제를 개선할 방안으로 데이터에 대한 법적인 권리보호와 공정한 데이터 유통 환경 조성을 위해 표준계약서 기준 마련, 법류 상담을 통한 분쟁 예방 등을 주된 기능으로 하는 ‘데이터 거래 법률지원센터’ 설치 운영을 통해 관련 법적 이슈 발생을 최소화할 필요가 있다고 도출되었음

나. 시사점

(1) 민간 중심의 데이터 거래소 강화

○ 데이터 거래소에 대한 정의를 명확히 하고, 이용자와 데이터의 보호에 필요한 의무와 안전조치를 구체화할 필요가 있음

○ 이와 함께 데이터산업법 제19조의 ‘데이터 플랫폼’은 전용몰 형태의 민간 데이터 거래소도 포함될 수 있도록 관련 용어의 정비가 필요함.

○ 참고로, 「데이터 기본법안」(조승래의원 대표발의, 의안번호 2106182)과 「데이터 산업 진흥법안」(이영의원 대표 발의, 의안번호 219463)은 모두 데이터 플랫폼 대신 ‘데이터 거래소’라는 명칭을 사용하였다.

○ 민간 거래소가 태동하지 못한 상황에서 정부 주도의 데이터 공급은 불가피한 측면이 있었지만, 현재와 같은 정부 의존적인 데이터 생태계는 장기적으로 성장의 한계를 초래할 우려가 큼.

○ 따라서 데이터 거래소에 대한 정부개입과 민간자율의 균형을 어떻게 유지할 것인지에 대한 장기계획이 필요함.

○ 이를 바탕으로 데이터 거래소에 대한 정부의 지원이 시작되는 순간부터, 정부 지원이 종료되는 이후를 대비한 출구전략도 함께 마련 되어야 함.
○ 데이터 통합지도는 정부가 독점할 것이 아니라, 시장에서 데이터 검색·연계·추천이 하나의 서비스로 자리잡을 수 있도록 민간에 맡기는 것이 타당함.
○ 민간이 여러 데이터 거래소의 상품을 분석하고, 활용가치를 평가하고, 결합의 가능성 등에 대한 정보를 제공하는 것이 효과적이기 때문임.
○ 이를 통해 데이터 통합지도가 현재의 '현황판' 성격이 아니라 이종간 데이터 융합의 길을 찾을 수 있는 수단이 되도록 해야 함.

(2) 데이터 거래사 역할과 책임 정립

○ 데이터 거래사의 역할을 명확히 할 필요가 있으므로 데이터 거래사만이 할 수 있는 고유한 기능을 부여할 것인지, 데이터사업자 신고의 필수 조건으로 데이터 거래사 확보를 포함할 것인지 등을 면밀히 검토해야 함.
○ 데이터 거래사는 데이터 분석에 대한 경력·자격을 갖추고 교육을 이수한 인력이므로 데이터 분석·결합도 일정부분 담당할 수 있도록 관련 규정을 정비하는 것도 고려해 볼 수 있다. 데이터 거래사의 책임 확보 방안을 강화해야 함.
○ 데이터 거래사의 의무와 금지 사항을 명확히 밝히고, 현재의 등록취소 이외에 시정명령과 자격정지 등의 다양한 제재조치를 신설하여 안전한 업무수행을 담보해야 함.
○ 추가적으로 거래사 등록의 유효기간 및 갱신주기 등을 마련하여 거래사가 지속적으로 역량을 관리할 수 있도록 할 필요가 있음

(3) 크롤링 활성화 환경 조성

○ 경쟁업체의 정보를 그대로 이용하는 사례들이 부각되면서 크롤링 전체가 부정적으로 이해되는 경향이 있지만, 크롤링이 소비자 후생을 높이고 경쟁을 촉진하는 긍정적인 효과도 큼.
○ 크롤링을 통해 새로운 서비스를 제공하고 인공지능 학습에 필요한 데이터를 모을 수 있으며, 신규 사업자들이 시장에 진입하기 위한 데이터를 확보할 수 있음.
○ 크롤링은 데이터 보호뿐만 아니라 경쟁 촉진도 충분히 고려하여 균형있게 허용되어야 하고, 웹사이트에 공개된 자료에 대한 크롤링을 제한하는 것은 조심스럽게 접근하여야 함.
○ 「저작권법」과 부정경쟁방지법 등 타법에서 보호하고 있는 권리를 침해하지 않는 공개 자료에 대한 크롤링을 허용하고, 이를 통해 특정 주체가 데이터를 부당하게 독점하지 않도록 하는 방안을 적극적으로 논의할 필요가 있음.
○ 또한 기술적으로 크롤링이 가능한 데이터에 대해서 국내 법·제도가 크롤링을 제한한다면 외국 기업만 국내 데이터를 가져가게 되는 역차별 문제가 발생할 수도 있다는 점을 경계해야 함.
○ 이러한 맥락을 고려하여 「저작권법」과 부정경쟁방지법 등의 제한을 벗어나지 않는 범위 내에서 데이터산업법을 기반으로 안전하고 효과적인 크롤링을 지원하는 방안도 모색해 볼 수 있을 것임.

(4) 데이터 정책 거버넌스의 정비

○ 데이터 크롤링, 거래소, 거래사 관련 정책을 성공적으로 추진하기 위해서는 효과적인 정책 거버넌스를 갖추는 것이 중요하나, 데이터 거버넌스 정비는 중요하지만 어려운 과제임.
○ 따라서 데이터 거버넌스 문제를 해결하기 위해서는 형식적 옥상옥이 아닌 실질적으로

작동하는 대통령 소속의 조정기구를 두거나, 범정부 데이터 업무를 총괄하는 전담조직을 법률로 명확히 설정하는 것도 방법일 것임.

(5) 시설 보급 · 확산 부문

○ 스마트농업 육성 정책의 1차적인 목표는 혁신농업의 미래를 위해 스마트팜, 스마트축사를 농가에 보급하는 것인데, 2021년(잠정)까지 보급 실적을 살펴보면 시설 원예 스마트팜의 경우 6,485ha, 스마트축사의 경우 4,743호를 보급하여 「스마트팜 확산 방안(2018)」에서 목표로 한 2022년 7,000ha, 5,750호에 근접한 것으로 보임.

○ 그러나, 시설원예 부문 스마트팜은 기술수준이 상대적으로 낮은 1세대 소규모 중심으로 보급되고 있어, 향후 경영효율성 등 경쟁력을 강화하기 위해서는 고도화 및 규모화된 스마트팜 보급이 필요함.

○ 스마트팜의 경우 딸기 · 참외 · 토마토 · 파프리카 4개 품목에 80% 이상이 집중되어 있고 노지작물 · 과수의 보급 실적은 미미한 수준으로, 향후 시장성 있는 품목을 대상으로 실증 및 R&D 강화를 통해 추가 확산을 도모할 필요가 있음.

○ 스마트농업 육성 목표 중 하나는 스마트팜 청년 창업 생태계를 조성하는 것인데 스마트팜 경영주 중 청년층은 최근 3년간 10% 내외로 정체되고 있어, 임대형 스마트팜의 조속한 완공 · 운영, 청년층에 대한 창업자금 지원 확대 등을 통해 청년층 유입을 강화할 필요가 있음.

(6) ICT 데이터 수집 · 활용 부문

○ 스마트농업을 구현하기 위해서는 생산과 환경, 유통, 소비 전반에 이르는 데이터 수집 · 활용이 필수적인 요소인데, 현재 데이터 수집 · 분석 비중이 높지 못하고 대부분 생산 관련 데이터가 중심이 되고 있어 향후 적극적인 개선 노력이 필요함.

○ 그리고, 농가의 데이터 제공을 유도하고 농업 데이터 시장에서 농가 교섭력을 강화하기 위해 농업 데이터에 대한 권리 · 수집 · 접근 · 이용 등에 대한 기준을 마련할 필요가 있음.

○ 농업 데이터의 가용성 · 접근성 향상 및 체계적인 관리 등을 위해 현재기관별로 분산된 농업 데이터 플랫폼을 조속히 연계통합 운영할 필요가 있음.

(7) 연구개발(R&D) 부문

○ 스마트농업은 농업 밸류체인(생산과 유통, 소비) 전반에 첨단 ICT 기술이 접목되어 자동화와 지능화를 구현하는 것이며, 이를 위해서는 단계별로 다양한 연구 및 기술개발이 필요함.

○ 전체 농림식품 기술수준 대비 스마트농업 관련 농림식품 기계 · 시스템 및 융복합 기술수준이 상대적으로 낮은 것으로 나타나서, 향후 전문인력 양성 및 타 산업 전문가 그룹과 협력 강화, 연구비 투자 확대, 개발기술 시장적용을 위한 지속적인 산업화 지원 등을 통해 기술수준을 향상시켜 나갈 필요가 있음.

○ 「스마트팜 확산 방안(2018)」 수립 이후 전체 스마트농업 관련 예산 중에서 연구개발(R&D) 예산에 대한 비중이 감소 · 정체되고 있는 가운데, 관련 기술수준향상 등을 위하여 연구개발에 대한 투자 배분을 강화할 필요가 있음.

○ 연구개발 투자가 산업 경쟁력 강화의 핵심요소라는 점을 감안하여, 스마트농업 연구개발의 경우 국가 전체 및 농림식품 분야에 비해 상대적으로 저조한 민간 부문의 투자를 확대해 나갈 필요가 있음.

○ 그리고 스마트팜 연구 및 인력양성을 위해 만들어진 스마트팜 연구센터에 대한 평가가 다른 연구센터에 비해 상대적으로 낮고, 제품화 및 취업인력 등의 목표 대비 실적이 저조한 것으로 나타나서 향후 성과향상을 위한 노력이 필요함.

(8) 산업인프라 구축 부문

○ 전 세계적으로 스마트팜 시장규모가 지속적으로 성장하고 있는 가운데, 정부에서도 스마트팜의 국산화율을 높이고, 기자재 표준화, 청년인력 육성, 수출 등을 통해 전후방 산업 경쟁력을 강화한다는 계획임.

○ 낙농 · 양계 등 축산 부문과 중규모(2,000평) 이상 시설원예 부문의 스마트농업 시설 · 장비의 국산화율이 타 부문에 비해 상대적으로 낮고, 대부분의 스마트농업 관련 기업은 영세하여 국산화 및 기술고도화에 어려움을 겪을 수 있으므로, 제품 성능 테스트, 실증연구 등을 지원하여 스마트농업 기업의 성장을 지원해나갈 필요가 있음.

○ 스마트농업 ICT 기자재의 국가표준 및 국제표준 등에 있어서 지정 성과가 다소 미흡한 상황이므로, 향후 정부는 표준화 제정 대상 장비를 확대하고 국가표준 등이 제정된 기자재 등에 대해서는 제조기업이 표준을 적용할 수 있도록 지원을 강화해나갈 필요가 있음.

○ 스마트팜 플랜트 · 기자재의 수출 실적은 증가하고 있으나 목표 대비 미미한 실정이고 전략 국가와 실제 수출국가와의 괴리가 발생하고 있는 가운데, 판로개척 및 시장정보 지원 강화 등을 통해 수출성과를 개선해나갈 필요가 있음.

○ 스마트농업의 산업인프라 형성 및 관련 기술의 구현을 위해서 인력양성은 무엇보다 중요한 요소임.

○ 인력양성을 위해 시행하고 있는 스마트팜 청년창업 보육센터의 경우 수료생의 대부분이 관련 분야에 취 · 창업하는 등 성과가 나타나고 있으나 선발 후 중도에 포기하는 인원이 최소화될 수 있도록 노력할 필요가 있으며, 현재 스마트농업 보급 품목이 딸기 · 토마토 등에 집중되어 있는 측면을 감안해서 교육생의 재배품목도 다양화시켜 나갈 필요가 있음.

○ 그리고 농촌진흥청의 스마트농업전문가 양성의 경우, 40세 미만 청년층 선발 비율 및 취업률이 감소하고 있어 사업성과 향상을 위한 개선노력이 필요함.

○ 한편, 스마트농업 산업인프라 구축의 핵심거점인 스마트팜 혁신밸리의 경우 당초 계획 대비 사업이 지연되고 있어, 청년창업보육, 기업의 실증지원 등에 지장을 주고 있으므로 향후 조속한 완공을 통해 성과를 창출해나갈 필요가 있음.

(9) 종합

○ 스마트팜이 주로 생산 기능을 강화하기 위한 시설 · 장비(H/W)에 초점을 둔 협의적 개념이라고 하면, 스마트농업은 생산 뿐만 아니라 가공 · 유통 · 소비 등 농업 가치사슬(value chain)의 모든 단계에 걸쳐, 데이터 · 인공지능(S/W)에 기반을 두고 농업혁신을 창출하는 광의적 개념임

○ 그런데, 현재까지 우리나라 스마트농업 추진 현황을 종합해서 살펴보면, ①기술수준이 상대적으로 낮은 1세대 소규모 중심으로 자동화된 시설 · 장비(H/W)가 보급이 확대된 수준이고, 스마트 농업의 기반이 되는 데이터 수집 · 활용 비중도 높지 못하여 진정한 의미의 스마트농업이라고 보기 힘들고, ②딸기 · 토마토 · 파프리카 등 시설원예 부문의 일부 품목에 편중되어, 노지작물 중심으로 다품종 농업경영이 이루어지고 있는 우리나라 농업 전반에 정착되고 있다고 보기 곤란한 측면이 있으며, ③스마트농업 관련 기술수준도 타 부문에

비해 상대적으로 저조한 상황에서 연구개발(R&D) 투자도 최근 저하되고 있는 실정임.

○ ④산업인프라 구축을 위한 국내 기업 육성, ICT 기자재 표준화, 수출 확대, 인력육성 등의 성과도 아직 충분하지 못한 상황이며, 미국 · 네덜란드 · 일본 등 스마트농업을 선도하고 있는 국가들에서는 자국의 농업 구조 및 특성에 맞는 스마트농업을 추진하고 있는데, 미국의 경우 영농규모가 커서 노지 분야의 정밀농업과 데이터 비즈니스를 활성화하고 있음

○ 네덜란드에서는 좁은 면적을 효율적으로 활용하여 생산성을 극대화하기 위해 첨단유리온실을 통한 수출 중심의 농업을 시행하고 있음.

○ 그런데, 우리나라의 경우 시설원예, 노지작물, 축산 등 다방면으로 보급 확산을 도모하고 있으나 핵심 전략품목이 불분명하고, 스마트농업 생산물 및 관련 기자재의 목표시장도 수출 중심인지 국내시장인지 명확하지 못한 측면이 있음.

○ 향후 정부는 전략품목, 목표시장, 관련 기술 등에 있어서 한국형 스마트농업의 개념을 명확히 하고, 시설 보급을 넘어서 데이터에 기반을 둔 진정한 스마트농업으로 발전할 수 있도록 체계적으로 관리해나갈 필요가 있음.

9. 참고문헌

1. 농식품 공공데이터 Top10 이용 비율, 농식품 공공데이터 포털, 2021년도
2. 국내 빅데이터 및 분석 시장 전망, 한국 IDC 연구보고서, 2021년도
3. KADX 주요 사업 영역, KADX
4. 2018년 학술연구용역인건비 기준 단가, 기획재정부 고시
5. 데이터 거래의 애로사항, 국가통계포털(KOSIS), 2021년도
6. 데이터 수집 방법별 비중, 과학기술정보통신부 한국데이터산업진흥원(2022)
7. 데이터 거래 시 애로사항, 과학기술정보통신부 한국데이터산업진흥원(2022)
8. 데이터거래 중개 플랫폼의 이용과정 및 이용약관, 데이터스토어, 금융데이터 거래소 홈페이지 이용약관
9. 미국 오픈 데이터 현황 “오픈 거버넌스 이니셔티브“, 오픈 데이터 포털 ’DATA.GOV“, 2012년
10. 영국 오픈 데이터 현황 “정부전환전략 : 더 나은 데이터 활용“, 공공 데이터 개방 포털 “data.gov.uk“, 2017년
11. 호주 오픈 데이터 현황 “MAGDA(Making Australian GovernmentData Available)23)“, 오픈 데이터 포털 “data.gove.au“, 2013년
12. 일본 오픈 데이터 현황, 오픈 데이터 칼로그 사이트 “data.go.jp“, 2014년
13. 캐나다 오픈 데이터 현황, 오픈 데이터 포털 “oepn.canada.ca“, 2013년
14. OECE 회원국 및 파트너국의 OGD 포털, API 스토어
15. 공공데이터 주요 활용 사례, 공공데이터포털
16. 네이버 데이터랩, Naver Developer 사이트, 2016년
17. SKT 빅데이터 허브, www.bigdatahub.co.kr, 2013년
18. 오픈 데이터 플랫폼 LS CNS , ODPia, WWW.odpia.org, 2016년
19. 빅데이터 플랫폼 통합 데이터 지도, www.bigdata-map.kr,
20. 금융 영역의 데이터와 결합되어 새로운 활용 사례, 금융 빅데이터 플랫폼(www.bigdata-fi nance.kr)
21. 환경 영역의 데이터와 결합되어 새로운 활용 사례, 환경 빅데이터 플랫폼(www.big data-environme nt.kr)
22. 문화 영역의 데이터와 결합되어 새로운 활용 사례, 문화 빅데이터 플랫폼(www.bigdata-culture.kr)
23. 교통 영역의 데이터와 결합되어 새로운 활용 사례, 교통 데이터 거래소(www.big data - transportation.kr)
24. 헬스케어 영역의 데이터와 결합되어 새로운 활용 사례, 암 빅데이터 플랫폼 (www.bigdata-cancer.kr)
25. 유통 영역의 데이터와 결합되어 새로운 활용 사례, KDX한국데이터 거래소(www.bigdata-market.kr)
26. 통신 영역의 데이터와 결합되어 새로운 활용 사례,KT통신 빅데이터 플랫폼 (www.bigdata-telecom.kr)
27. 중소기업 영역의 데이터와 결합되어 새로운 활용 사례, 데이터유통포털(www.bigdata-sme.kr)
28. 지역경제 영역의 데이터와 결합되어 새로운 활용 사례, 경기지역 경제포털(www.bigdata- region.kr)
29. 산림 영역의 데이터와 결합되어 새로운 활용 사례, www.bigdata-forest.kr
30. KADX의 정보 거래 사례, 농식품 빅데이터 거래소 홈페이지 (kadx.co.kr), 2022년
31. 윤경구, 안명옥, 2020, 데이터 가치 창출에 관한 사례 및 시사점, 우정정보, 2020 가을호: 91-114.
32. Ackoff(1989)는 데이터를 ‘symbols’, 정보를 ‘Data that are processed to be useful; provides answers to who, what, where and when questions’로 정의(Rowley, 2007, p.167)
33. 데이터 정보 생성에 필요한 투입요소의 가치 주장, Moody and Walh, 1999, p.2
34. 데이터, 정보, 지식의 정보 비교 Ackoff(1989) vs. Zeleny(1987), Rowley(2007, p.167)

35. 정보의 양과 가치, Moody and Walsh, 1999
36. 정형 및 비정형 데이터의 품질 평가지표 제시, 한국데이터산업진흥원, 2020
37. 데이터 가치평가 방법론, MDA and PDPC, 2019
38. 데이터 관련 경쟁 제한 우려 사항, Autorite’ de la concurrence and Bundeskartellamt, 2016
39. 가격결정과 가격발견, Ward and Schroeder, 2002
40. 데이터 산업 시장 규모, 과학기술정보통신부, 2022
41. 데이터 거래 체계, 4차산업혁명위원회, 2022
42. 일본 WAGRI(농업데이터연계기반) 추진 동향, 지성훈, 2021
43. 일본 농업 빅데이터 현황, 유찬석, 2021
44. 스마트팜 빅데이터 플랫폼, https://www.n-farm.kr/intro
45. 데이터 거래 활성화를 위한 거래소・거래사・크롤링의 현황과 개선과제, 2022
46. 데이터 경제 활성화를 위한 이슈와 쟁점, 2020
47. 개인정보보호위원회,일본의 개인정보보호 법제・정책 분석에 관한 연구, 2017. 1
48. 고미리, 데이터를 활용한 소비자 주권 강화: 영국 Midata 프로그램, Data Issue Report 2016-24 제 98호, 한국 데이터진흥원, 2016.11
49. 김형건, 해외 mydata 정책의 전개 양상과 국내 도입방향, Data Issue Report 2016-25 제 99호, 한국 데이터진흥원, 2016.11
50. 방송통신위원회, 빅 데이터 환경의 개인정보에 대한 소비자 가치 추정 연구, 2014.1
51. 법제처, 개인정보 보호와 적정 활용의 조화를 위한 제도 도입 연구, 2009.9
52. 윤상오, 김기환, 빅데이터 시대의 한국과 영국 간 개인데이터 활용정책 비교 연구, 한국정책과학학회보 20(1) 29-56, 2016.3
53. 한국데이터진흥원, 국내 개인데이터 이용 환경 분석, 2017.
54. 한국데이터진흥원, MyData 시뮬레이션 및 시스템 설계 연구, 2017.11
55. 한국인터넷진흥원, 스마트 환경에서의 개인정보보호 중장기 발전방향 연구, 2014.
56. 한국인터넷진흥원, 신규 ICT 환경의 해외 개인정보 활용사례 및 법제동향 분석연구, 2016.10.