KData 17-002 | 2017. 10.

CKAN을 활용한 민간데이터 연계 활성화 방안 연구 결과 보고서

주관기관 : 한국데이터진흥원

수행기관 : ㈜엔코아

한국데이터진흥원
Korea Data Agency

제 출 문

한국데이터진흥원 원장 귀하

본 보고서를 “CKAN을 활용한 민간데이터 연계 활성화 방안 연구”에 관한 결과 보고서로 제출합니다.

수 행 기 관 : ㈜엔코아
총괄 책임자 : 김 옥 기 ((주)엔코아 데이터서비스 센터장)
사업 참여자 : 서 학 준 ((주)엔코아 수석 연구원)
홍 성 진 ((주)엔코아 책임 연구원)
박 세 진 ((주)엔코아 선임 연구원)
최 준 영 ((주)엔코아 선임 연구원)
장 준 규 ((주)엔코아 연구원)
류 진 걸 ((주)엔코아 연구원)
김 용 연 ((주)엔코아 연구원)

< 목 차 >

제1장 서론 ····· 1

제1절 연구의 배경과 필요성 ····· 1
1. 연구의 배경 ····· 1
2. 연구의 필요성 ····· 5
제2절 연구의 목적과 범위 ····· 6
1. 연구의 목적 ····· 6
2. 연구의 범위 및 추진 절차 ····· 7
3. 기대효과 ····· 10

제2장 국내외 공공 · 민간데이터 개방 및 공유 현황 ····· 11

제1절 국내·외 공공·민간 데이터 포털 현황 ····· 11
1. 해외 데이터 포털 현황 ····· 11
2. 국내 데이터 포털 현황 ····· 24
3. 비교 및 시사점 ····· 26
제2절 국내·외 데이터 플랫폼 동향 ····· 27
1. 데이터 플랫폼 개요 및 특징 ····· 27
2. 국내·외 데이터 플랫폼 적용 현황 ····· 33
3. 비교 및 시사점 ····· 37
제3절 오픈 데이터 관련 기구 및 기관 활동 현황 ····· 37
1. 해외 오픈 데이터 관련 기구 및 기관 현황 ····· 37
2. 국내 관련 기구 현황 ····· 45
3. 비교 및 시사점 ····· 47

제3장 데이터 연계 도구 CKAN ····· 48

제1절 CKAN의 의의와 발전 과정 ····· 48
1. CKAN의 개요 ····· 48
2. CKAN의 주요 기능 ····· 48
3. CKAN의 특장점 ····· 50
4. CKAN의 기술 발전 동향 ····· 52
5. 시사점 ····· 54

제2절 CKAN의 기능 확장(CKAN Extension) ····· 55
제3절 국내외 CKAN 활용 현황 ····· 59
1. 국내외 데이터 포털의 CKAN 활용 현황 ····· 59
2. 기업의 CKAN 활용 사례 ····· 60

제4장 공공·민간 데이터 연계 현황 ····· 63

제1절 공공·민간 데이터 연계 사례 ····· 63
1. 데이터 플랫폼 사례 ····· 63
2. 데이터 연계 분석 사례 ····· 69
제2절 연계 가능한 민간데이터 ····· 71
1. 수요가 많은 민간데이터 주제 분야 ····· 71
2. 연계 가능한 민간 데이터 현황 ····· 75

제5장 공공·민간 데이터 연계 활성화 방안 ····· 77

제1절 데이터 연계 필요성 및 장애요인 ····· 77
1. 데이터 연계의 필요성 ····· 77
2. 민간 데이터 연계 장애요인 ····· 79
제2절 공공·민간 데이터 연계 활성화 제언 ····· 80
1. 민간 데이터 연계 활성화를 위한 개선 방안 ····· 80
2. CKAN 적용 확대 방안 ····· 87

제6장 결 론 ····· 90

참고문헌 ····· 95

< 표 목차 >

<표 2-1> 지역별 데이터 포털 현황 ····· 11
<표 2-2> 주요 국가별 데이터 포털 수 ····· 12
<표 2-3> 지역별 조사된 데이터포털 수 상위 5위 국가 ····· 12
<표 2-4> 선진국 국가별 주요 오픈 데이터 포털 일반사항 ····· 14
<표 2-5> 선진국 국가별 대표 오픈 데이터 포털 - 기능 비교 ····· 16
<표 2-6> 선진국 국가별 대표 오픈 데이터 포털 데이터 제공 건수 및 이용 권한 ····· 17
<표 2-7> 해외 주요 데이터 포털 주제별 데이터셋 집계 현황 ····· 23
<표 2-8> 국내 데이터 포털 및 데이터 마켓 현황 ····· 24
<표 2-9> 국내 데이터 포털 상세 현황 ····· 25
<표 2-10> 데이터 플랫폼 개발주체 및 오픈소스 활용 여부 ····· 28
<표 2-11> 데이터 플랫폼별 기능 비교 ····· 33
<표 2-12> 지역별 데이터 포털의 데이터 플랫폼 적용 현황 ····· 35
<표 2-13> 오픈 데이터 관련 기관 및 기구 ····· 41
<표 3-1> CKAN의 주요 기능 ····· 49
<표 3-2> 데이터 연계 관점에서의 CKAN의 주요 사용 기능 ····· 49
<표 3-3> CKAN 개발 활용 측면의 특장점 ····· 51
<표 3-4> 타 플랫폼과 비교를 통한 CKAN 적용의 타당성 ····· 51
<표 3-5> CKAN Extension 기능별 분류 ····· 56
<표 3-6> 데이터 연계와 관련된 CKAN Extension 패키지 ····· 57
<표 3-7> 기업의 CKAN 활용 사례 (CKAN 기반 상용 서비스) ····· 62
<표 4-1> 국내 주요 공공/민간 데이터 포털 및 마켓에서 제공하는 데이터셋의 주제 ···· 73
<표 4-2> 전체 및 상품 유형별 데이터 상품 등록 건수 주제별 순위 ····· 74
<표 4-3> 데이터스토어 상품 조회수 주제별 순위 ····· 74
<표 4-4> CKAN 도입 가능성 있는 기업들의 업종별 현황 ····· 76
<표 5-1> 데이터의 공공재적 성격 ····· 78

< 그림 목차 >

<그림 1-1> 2016년 ODB Top 10 국가 현황 ····· 2
<그림 1-2> 지역별 ODB 지표 Top 5 국가 현황 ····· 3
<그림 1-3> 최근 3년간 우리나라의 ODB 지표 변화 ····· 3
<그림 1-4> 우리나라의 ODB 개방도 상세항목 평가 결과 ····· 4
<그림 1-5> 오픈소스 데이터 연계 플랫폼 활용 데이터 공개 포털 예시 ····· 5
<그림 1-6> 연구의 목적 ····· 7
<그림 1-7> 연구 추진 절차 ····· 8
<그림 2-1> 미국 대표 국가 데이터 포털 DATA.GOV ····· 18
<그림 2-2> 미국 공공데이터 정보 시각화 사이트 DATA.USA ····· 18
<그림 2-3> 캐나다 대표 오픈 데이터 포털 ····· 19
<그림 2-4> 오픈 데이터 오타와 ····· 19
<그림 2-5> 인도 오픈 데이터 포털 ····· 20
<그림 2-6> 영국 공공데이터 포털 ····· 21
<그림 2-7> 오스트레일리아 공공데이터 포털 ····· 22
<그림 2-8> CKAN.org ····· 29
<그림 2-9> Socrata 시스템 아키텍처 ····· 30
<그림 2-10> Open Government Platform ····· 31
<그림 2-11> 데이터 플랫폼 사용 현황 ····· 34
<그림 2-12> 지역별 데이터 포털 비율과 CKAN 적용 비율 ····· 34
<그림 2-13> 데이터 플랫폼 사용 현황 ····· 36
<그림 2-14> 오픈 데이터 관련 기구들의 활동 ····· 43
<그림 2-15> 국내외 데이터 연계 가이드라인 사례 ····· 44
<그림 2-16> 오픈 데이터 포럼 운영 방향 ····· 46
<그림 3-1> CKAN 기술 발전 동향 ····· 52
<그림 3-2> 농업 분야 기업 몬산토 지리정보기반 데이터 카달로그 ····· 61
<그림 3-3> Link Digital을 통해 호스팅 하는 데이터 포털 ····· 62
<그림 4-1> 오픈 이미지 데이터셋 ····· 63
<그림 4-2> 오픈 이미지 데이터셋 ····· 64
<그림 4-3> YouTube-8M Dataset ····· 64
<그림 4-4> 구글북스 Ngrams ····· 65
<그림 4-5> 구글 트렌드 데이터스토어 ····· 66

<그림 4-6> 네이버 데이터랩 ····· 66
<그림 4-7> ADAM DATA HUB ····· 67
<그림 4-8> AWS Public Datasets ····· 68
<그림 4-9> Microsoft DataMarket ····· 68
<그림 4-10> ODPia ····· 69
<그림 4-11> 미국 내 데이터 서비스 분야 스타트업 투자현황 (2016년) ····· 71
<그림 4-12> 데이터스토어 상품 등록건수 및 카테고리 별 조회수 비교 ····· 75
<그림 4-13> 연계 대상 기업 도출 절차 ····· 75
<그림 5-1> 민간데이터 개방 시 예상되는 장애요인 ····· 80
<그림 5-2> 민간데이터 연계 활성을 위한 온/오프라인 마케팅 플랫폼 ····· 84
<그림 5-3> CKAN 민간 데이터 활성화 홍보 방안 ····· 85
<그림 5-4> 단계별 CKAN 적용 로드맵 ····· 88

제1장 서 론

제1절 연구의 배경과 필요성

1. 연구의 배경

○ 최근 데이터의 폭발적인 증가와 데이터 활용을 통한 새로운 가치창출의 중요성이 증대됨에 따라 데이터가 중심이 되는 전략이 요구됨

- 공공에서는 ‘정부 3.0 추진계획’을 통해 데이터 개방을 극대화하고 데이터 활용을 증진하기 위해 노력하고 있음
- 민간에서는 네이버, 다음 등 포털을 중심으로 오픈 API 제공을 통해 데이터 개방에 동참하고 있음
- 오픈 데이터 정책은 주로 공공데이터를 대상으로 하나, 민간데이터를 융·복합 하여 보다 정교한 통찰력을 확보하여야 실질적 시너지를 기대할 수 있음

○ 데이터가 ‘21세기 원유’로 인식되고, 4차 산업혁명이 국가 주요 현안으로 되면서 각 국가별로 공공데이터 개방 및 공공·민간 데이터 간 연계·활성화를 통해 경제 발전을 도모하고자 ‘오픈 데이터’정책을 수립 및 시행 중

○ 우리나라는 2013년 ’공공데이터의 제공 및 이용 활성화에 관한 법률’을 제정 후 2차에 걸쳐 ‘공공데이터의 제공 및 이용활성화 기본계획’을 수립하는 등 중앙정부 및 지방정부의 공공데이터 개방 및 이용활성화를 위한 법제도, 마스터 플랜 등의 정책 활동을 적극적으로 추진하고 있음

○ 기존의 개별적 빅데이터 활용 단계를 넘어 데이터 간 연계를 통한 새로운 가치 창출 단계로 진입

- 기업들이 비식별 정보를 활용하여 신상품 개발, 신약개발, 고객별 맞춤형 마케팅 등 다양한 목적의 빅데이터 활용 요구 증대
- 정부는 공공데이터포털, 빅데이터 공통기반, 빅데이터 전략센터 구축 등 공공데이터 활용에 필요한 기반 마련 및 확산
- 새로운 가치창출을 위해 데이터 간 연계·분석이 필요하다는 공감대 형성 및 이를 뒷받침하는 ‘개인정보 비식별화 가이드라인’ 공표(‘16.6.30)

○ 정부 3.0 구현을 위해 공공데이터의 개방 및 민간 빅데이터와 연계를 통한 민간 기업의 경제 활동 지원을 노력하고 있음

- 행정자치부 발표 '제2차(2017~2019)「국가중점개방 데이터」[1] 개방계획'[2]에 따르면, 국민이 선정한 국가중점 데이터 36개 분야 중 2016년까지 33개 분야의 공공데이터가 개방 완료, 3개 분야는 점진적 개방 예정
- 행정자치부에 따르면 공공데이터법이 제정된 2013년 대비 이용건수는 120배, 공공데이터를 활용한 서비스 개발 사례는 25배 증가
- 4차 산업혁명 대응을 위한 인공지능·IoT등의 지능형, 신사업 분야 데이터 개방수요를 반영한 '제2차「국가 중점개방 데이터」개방계획'을 수립
- 38개 분야의 국가중점 개방데이터 선정, 2017년부터 데이터 개방 예정

○ 이러한 노력으로, 2016년 ODB(Open Data Barometer)보고서[3]에 의하면 한국은 종합점수 81점으로 세계 5위를 차지, 2014년 이후 지속적으로 지표가 향상되고 있음

Country	Rank	Score OUT OF 100	Change	Score Trend OVER PAST EDITIONS	Readiness OUT OF 100	Implementation OUT OF 100	Emerging Impact OUT OF 100
United Kingdom See details	1	100	0 -		99	100	94
Canada See details	2	90	2 ▲		96	87	82
France See details	3	85	-1 ▼		100	71	88
United States of America See details	4	82	-2 ▼		96	71	80
Korea See details	5	81	3 ▲		95	59	100
Australia See details	5	81	5 ▲		85	78	78
New Zealand See details	7	79	-1 ▼		92	58	99
Japan See details	8	75	5 ▲		84	60	89
Netherlands See details	8	75	-1 ▼		94	64	68
Norway See details	10	74	7 ▲		77	71	73

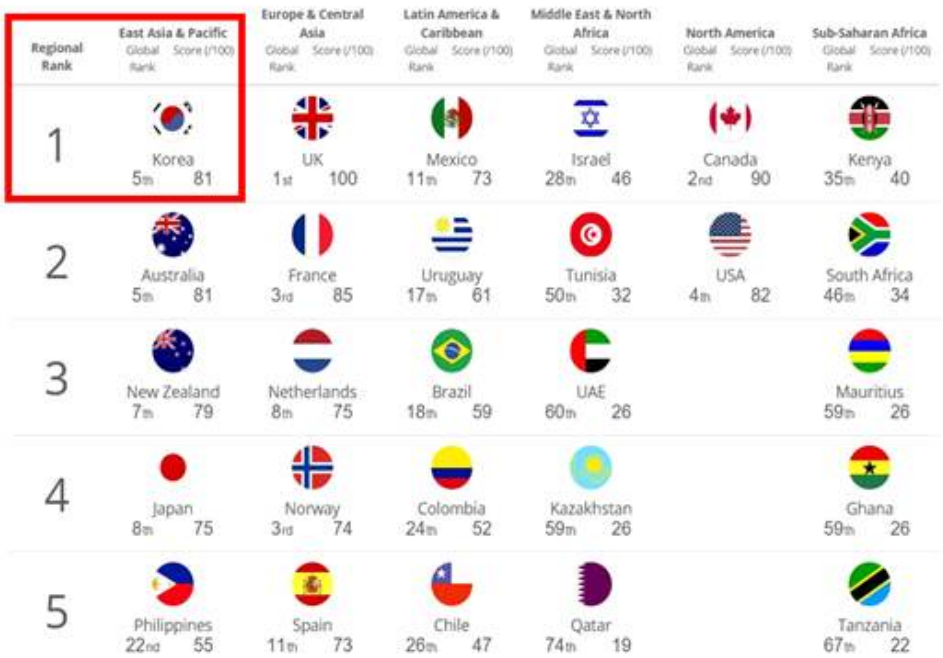
<그림 1-1> 2016년 ODB Top 10 국가 현황

※ 출처 : 오픈 데이터 바로미터(The Open Data Barometer),
http://opendatabarometer.org/?_year=2016&indicator=ODB 참조 (접속일: 2017.07.24.).

---

1) 국가중점데이터: 민관합동 TF를 구성하여 국민·기업·전문가 수요조사 등을 통해 사회 경제적 편익과 파급효과가 큰 대용량 중심의 개방 대상 공공데이터를 의미
2) 공공데이터전략위원회 (2016.12.15.)「제2차 국가중점개방 데이터 개방계획」, 공공 데이터전략위원회.
3) ODB 보고서는 WWW재단(월드와이드웹 컨소시엄, World Wide Web Consortium)에서 전 세계 77개국을 대상으로 오픈데이터 현황과 효과 분석을 목적으로 발표(2013.10.31)한 오픈데이터 현황 보고서이다. 이 보고서는 데이터 아젠 다별 분석, 국가 및 지역별 분석과 함께 각국의 준비도, 실행력, 효과에 대한 국가별 순위를 제공함

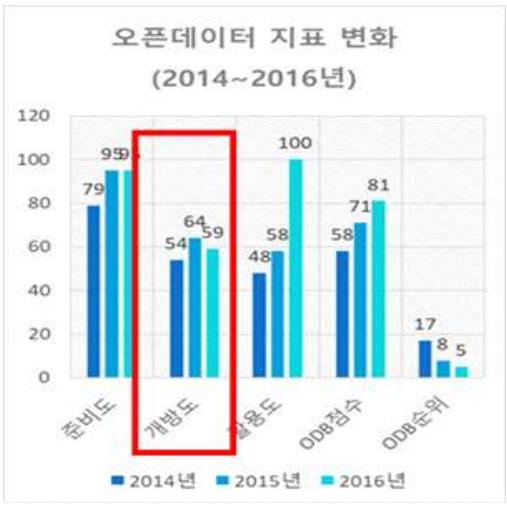
○ 아시아·태평양 지역 순위에서는 1위를 차지, 데이터 개방 및 활용 강국에 진입[4)]



<그림 1-2> 지역별 ODB 지표 Top 5 국가 현황

※ 출처: 오픈 데이터 바로미터(The Open Data Barometer), http://opendatabarometer.org/?_year=2016&indicator=ODB&region=:PA 참조 (접속일: 2017.07.24.).

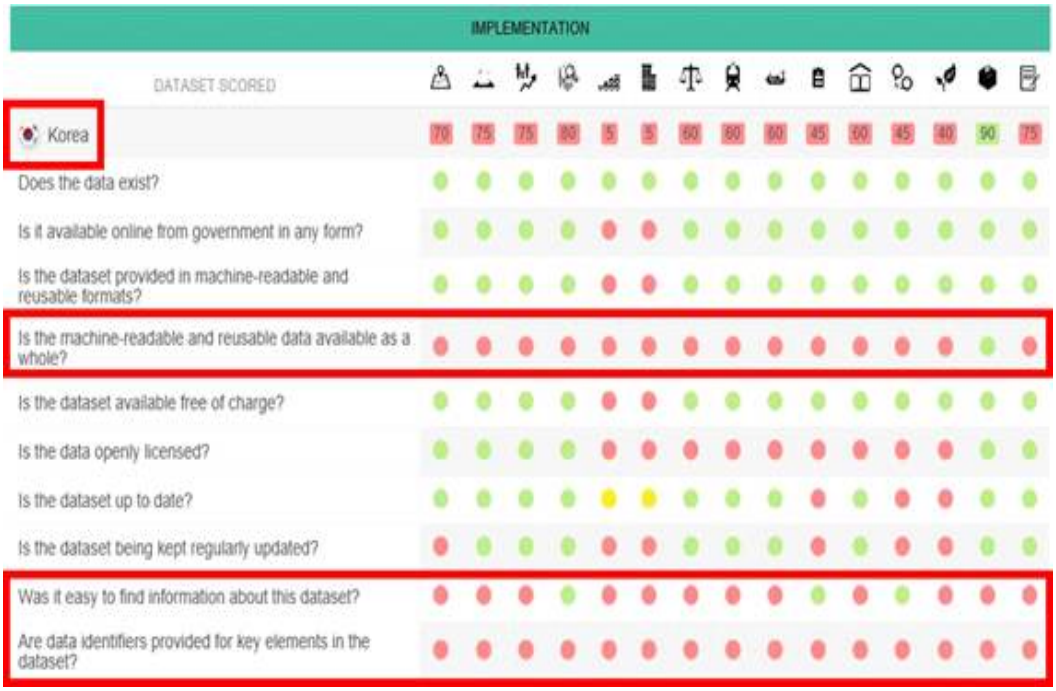
○ 그러나, 2015년 대비 준비도 지표는 95로 높은 수준이 유지되고 있고, 활용도는 100으로 비약적으로 상승한 반면 개방도는 59로 하락하며 상대적으로 취약



<그림 1-3> 최근 3년간 우리나라의 ODB 지표 변화

4) 행정자치부 보도자료 (2017.05.29.), “대한민국, 공공데이터 개방 세계 5대 강국 우뚝”.

- 최근 3년간 우리나라는 ODB 종합 점수 측면에서 꾸준히 지표가 좋아지고 있고, 특히 활용도는 공공데이터를 활용한 서비스 개발 등 경제적 효과창출 등에 대해서 높게 평가를 받아 100점을 달성
- 그러나, 오픈된 데이터 셋 평가지표인 개방도[5]는 59점으로 3년간 비교적 개선이 저조한 것으로 나타남
- 특히, 개방도 관련 세부 항목 중 데이터 연계와 관련된 '공개된 데이터 셋 전체의 기계판독 및 재사용 가능성' 과 '데이터 식별자가 데이터셋의 KEY요소로 제공', '관련 데이터셋의 검색 편리성' 등의 항목에 대해서는 거의 모든 항목에서 제공되지 않는 것으로 조사 됨
- 이는 수차례 공공데이터 포털 개편 중 기 오픈한 데이터 파일의 일부가 유실되어 추적이 되지 않은 이유라는 일부 전문가 의견이 있음
- 또한, 이번 ODB 지표가 높게 나온 것은 수차례에 걸친 공공데이터 포털의 개편시 ODB의 지표측정 방식에 맞게 준비를 했기 때문이라는 전문가 의견도 있음



<그림 1-4> 우리나라의 ODB 개방도 상세항목 평가 결과

※ 출처: 오픈 데이터 바로미터(The Open Data Barometer)-한국 , http://opendatabarometer.org/4thedition/detail-country/?_year=2016&indicator=ODB&detail=KOR 참조 (접속일: 2017.07.24.).

5) ODB 개방도(Implementation): 가용성을 통해 측정한 3가지 영역의 15가지 핵심 데이터의 개방과 이행, '오픈 정의' 및 열린 정보 데이터 원칙에서 규정하고 있는 일반적인 관례의 데이터 셋을 채택하고 있는가를 의미

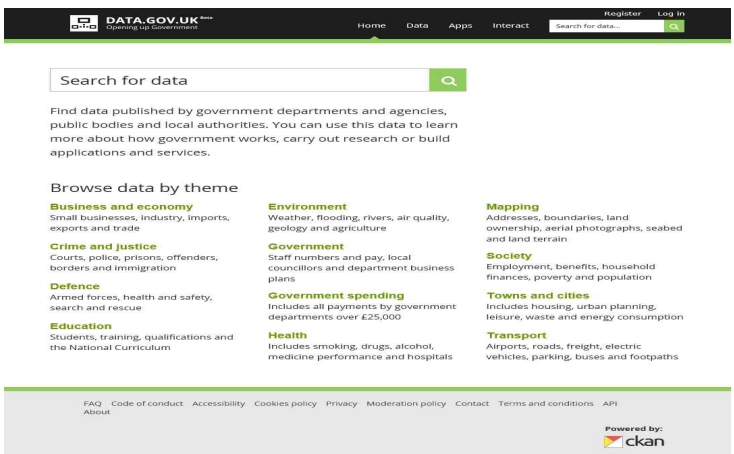
○ 국민 삶의 질 향상을 위해 데이터를 각 산업계에서 다양하게 활용하여 실질적 효과를 창출하기 위해서는 민간데이터와의 연계를 통한 데이터 융·복합 활성화 방안이 필요

- 빅데이터 환경에서 데이터의 연계·활용은 공공과 민간 부문의 데이터 산업을 활성화하고 데이터 자산의 국가 경쟁력을 좌우하는 중요한 요소임
- 데이터의 공개 및 연계를 하고 있는 우리나라의 공공 기관 및 민간 기관은 대부분 자체 개발 데이터 연계 플랫폼 활용, 호환성 및 확장성의 문제로 플랫폼 상호간 데이터 연계가 어려움

2. 연구의 필요성

○ 공공 및 민간 데이터의 개방 및 연계 활용 활성화 기반 연구를 위해 국내외 오픈 플랫폼 활용에 관한 사례 조사 필요

- 데이터 개방 및 연계 활용이 활발히 이루어지기 위해 공공기관 및 민간기관이 보유하고 있는 데이터를 편리하고 손쉽게 개방, 활용할 수 있는 데이터 연계를 위한 오픈 플랫폼의 필요성 증대
- 우리나라보다 앞서 공공 데이터 개방을 추진하고, 데이터의 공개 및 확산이 활발히 이루어지고 있는 영국, 미국 등 해외 주요 국가들 에서는 CKAN[6]과 같은 오픈소스 기반의 데이터 연계 플랫폼을 활용하여 공공기관 및 민간기관 간 플랫폼을 통한 자유로운 데이터 연계를 지원하고 있음



<그림 1-5> 오픈소스 데이터 연계 플랫폼 활용
데이터 공개 포털 예시(영국)

※ 출처: DATA.GOV.UK, https://data.gov.uk 참조 (접속일: 2017.08.01.).

6) CKAN(Comprehensive Knowledge Archive Network): Open Knowledge에서 개발, 관리 및 배포하고 있는 오픈소스 기반의 데이터 연계 플랫폼

- 4차 사업혁명의 핵심 원동력인 다양한 분야로부터 연계 및 가공된 융·복합 데이터의 중요성이 부각되고 있으나 우리나라는 이에 대응하기 위한 플랫폼 기반이 취약
- 특정 데이터 플랫폼에 종속되지 않고 타 시스템과의 확장성이 뛰어난 오픈 플랫폼의 적용을 위한 선진 활용 사례 연구 필요

○ 4차 산업혁명 대응 및 데이터 주도의 경제 전환을 위해 공공 및 민간 데이터 간 연계 활성화 방안 연구가 필요
- 지능정보사회로의 빠른 전환에 따른 공공·민간, 민간-민간 데이터 연계의 수요는 지속적으로 증가할 것으로 예상
- 現 공공 중심의 데이터 개방의 한계를 극복하고, 민간이 보유하고 있는 다양한 분야의 데이터와 융·복합을 통해 새로운 산업 및 경제 부가가치를 창출하기 위한 민간의 데이터 공개 및 연계 참여 활성화 방안의 연구 필요
- 국내 공공·민간 데이터 연계 활용 활성화 방안 연구를 위한 해외 선진국의 공공·민간 데이터 연계 및 활용 성공 사례 분석 필요

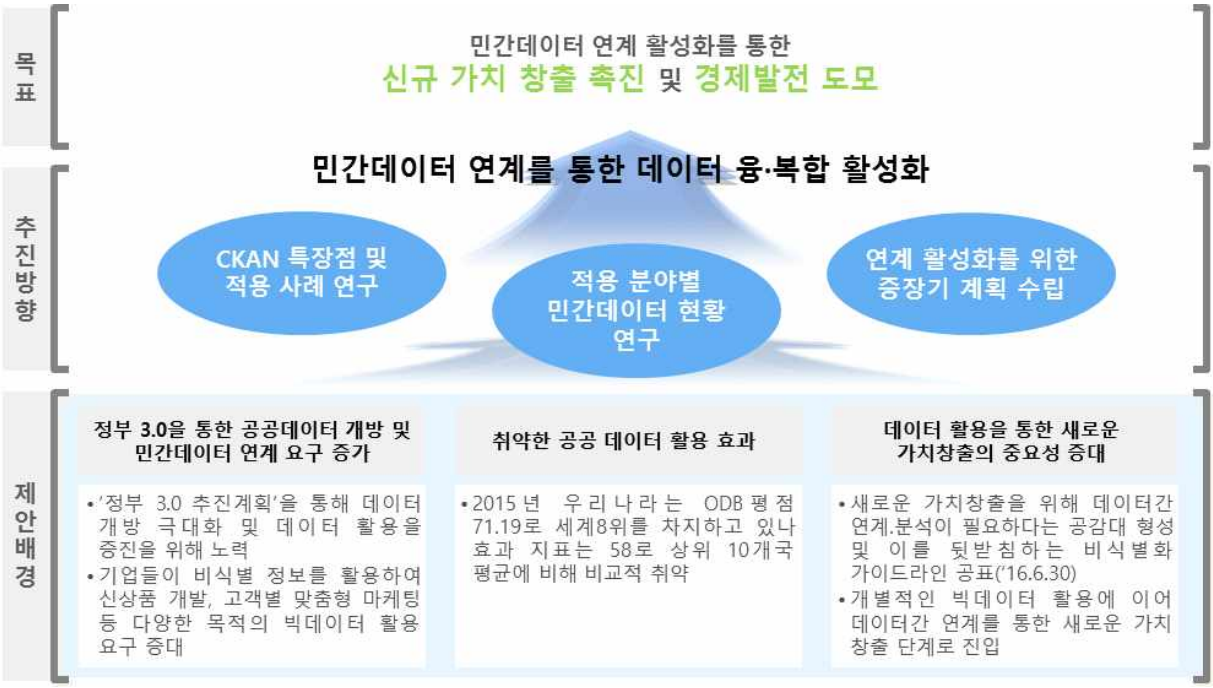
○ 이러한 필요성에 따라 오픈소스 데이터 연계 플랫폼인 CKAN을 활용한 공공·민간 데이터 연계 활성화를 위한 방안 연구를 수행 하고자 함

제2절 연구의 목적과 범위

1. 연구의 목적

○ 국내외 데이터 포털 및 플랫폼 활용 현황 분석, 오픈 연계 플랫폼으로서의 CKAN 장점 분석, 민간 데이터 보유 현황 분석을 통해 CKAN 활용 민간기관 보유 데이터의 연계 활성화 방안 제시
- 이를 통해 공공기관 및 민간기관 상호간 데이터 유통 활성화와 CKAN기반 데이터스토어(www.datastore.or.kr)[7]의 향후 발전 방향 수립에 기여

7) 데이터스토어(www.datastore.or.kr) : 한국데이터진흥원에서 운영·관리 중인 온라인 데이터 오픈 마켓



<그림 1-6> 연구의 목적

○ 민간데이터 연계 활용 활성화를 통해 새로운 가치창출 촉진을 위한 방안을 제시함
- CKAN 적용 사례 연구를 통하여 오픈소스 플랫폼 CKAN의 데이터 연계 관점에서의 장점을 살펴보고자 함
- 적용 분야별 민간 데이터 현황 연구를 통하여 CKAN 적용 대상을 찾아보고자 함
- 연계 활성화를 위한 중장기 계획을 통해 CKAN의 단계적 적용을 통한 민간 데이터 연계 활성화 방안을 마련하고자 함

2. 연구의 범위 및 추진 절차

○ 연구의 효율적인 추진을 위해 과제를 3단계의 절차로 나누어 단계적으로 진행
- 인터넷 검색을 통한 자료수집 방법인 탐색적 조사 연구방법을 활용하여 조사연구 수행
- 조사대상 웹사이트 게시 글, 공공 및 민간기관의 연구조사보고서, 기관 보도자료, 논문 등 인터넷 상에서 공개된 글이나 문서를 대상으로 자료를 수집

<그림 1-7> 연구 추진 절차

(1) (1단계) 국내·외 공공·민간 데이터 개방 및 공유 현황 조사

○ 데이터 포털 간 기능 및 제공 데이터 현황 비교를 위하여 국내외 공공 및 민간 기관에서 운영하고 있는 데이터포털 현황 조사
- 지역별 데이터 포털 현황을 집계하고, 주요 선진국별로 데이터포털 상세 현황을 분석
- 국내외 데이터 포털 구축 및 운용 현황에 대한 국내외 사례조사 실시함
- 주요 선진국의 데이터 포털 기능 및 제공 데이터 현황을 비교함

○ 주요 데이터 플랫폼(솔루션)간 기능 비교 및 CKAN 적용 현황 분석을 위하여 국내외에서 공급 및 활용되고 있는 주요 데이터 플랫폼의 상세 현황을 조사
- 데이터 플랫폼 솔루션 사용 현황에 대한 국내외 사례조사를 실시하고 CKAN 적용 비율을 집계함
- 조사 대상 플랫폼별 기능을 비교 분석하고 시사점을 도출함

○ 데이터 연계 플랫폼으로서의 CKAN의 타당성 분석을 위하여 CKAN의 특징과 국내의 기업에서의 CKAN 활용 및 적용 사례 조사
- CKAN의 개요, 주요기능, 특장점과 기술 발전 동향 등에 대하여 조사
- 국내외 CKAN을 활용한 데이터 포털 구축 및 운용 사례를 조사함
- CKAN 적용을 위한 국내외 기업들의 개발 사례를 조사함
- CKAN 확장 기능인 extension을 데이터 연계 관련 기능을 중심으로 살펴봄

○ 오픈데이터 관련 기구 및 기관의 현황 및 활동을 조사함
- 국내외 주요 국가의 오픈 데이터 관련 기구 및 기관 현황 조사하고, 각 기구 및

기관의 활동 내용과 산출물, 주요 참여 인물들을 조사
- CKAN과 관련 산출물 및 활동 현황을 조사
- 데이터 개방 및 연계 활성화를 위한 정책 제언을 위하여 국내외 주요 데이터 연계 가이드라인 사례를 조사

(2) (2단계) 주요 연계대상 민간데이터 현황 조사

○ 해외 데이터 서비스 기업에 대한 주제 분야별 투자 현황 조사하여 국내 및 해외 데이터 공개 현황과 비교하여 시사점 도출
- 주요 선진국(영국, 미국)의 데이터 주제 분류 현황을 조사 분석함
- 데이터 연계 및 분석 잠재력을 갖춘 주제 분야별 민간데이터 도출

○ 주요 주제별 민간데이터 현황 조사를 통해 주제 분야별 데이터 공개 현황 조사
- 주요 주제별 데이터 보유 기업 현황 조사하여 연계대상 민간데이터 보유 기업 현황을 도출하고 CKAN 적용 가능성 높은 기업 도출

(3) (3단계) 민간데이터 연계 활성화를 위한 제언

○ 민간데이터 연계 활성화를 위한 정책 및 CKAN 기술 제언을 수행
- 국내·외 데이터 오픈 가이드라인 현황 및 시사점을 살펴보고 CKAN 기술 발전 동향 및 기술을 제언함

○ 공공·민간, 민간-민간 데이터 연계 필요성 및 샘플 데이터 무료 공개 확대 필요성에 대한 인식 확산 방안 도출

○ 공공·민간 데이터 연계 방안
- 지재권, 가격, 연계상의 제약 등 공공데이터 개방과 달리, 민간데이터 연계 시 유의사항 등 차이를 감안한 연계 방안 도출

○ 데이터스토어의 CKAN 적용에 대한 당위성 도출함
- 데이터 연계 활성화를 위한 CKAN 도입 당위성 도출
- 데이터 플랫폼 중 CKAN 채택 이유 및 CKAN의 장점 조사

○ CKAN을 적용한 데이터스토어 중장기 계획을 수립하고 CKAN 기능의 개선 방향 설정

○ 주요 데이터 공개 가이드라인 및 CKAN의 기술 발전 동향을 조사하여 데이터 연계 활성화를 위한 CKAN의 적용 필요성 분석함

○ 민간데이터의 연계 필요성을 분석하고, 공공・민간 데이터 연계 사례 조사 및 민간 데이터 연계 활성화 방안 제시

○ CKAN 적용한 데이터스토어 중장기 발전 계획을 제시함
- CKAN 적용 기능개선 방향을 제시
- CKAN 적용 데이터스토어의 단계별 중장기 발전 계획을 제시

3. 기대효과

○ 오픈소스 플랫폼인 CKAN을 활용한 다양한 주제 분야의 공공 및 민간 데이터 간 상호 연계 활성화로 공공・민간 상생의 생태계 환경 조성

○ 오픈소스 데이터 연계 플랫폼인 CKAN을 활용한 새로운 융·복합 서비스 개발 촉진

○ 공공 및 민간 데이터의 융·복합을 통한 다양한 데이터 가치 창출로 데이터 브로커(Data Broker)와 같은 신규 산업 활성화 및 경제적 파급 효과 확산

○ 데이터를 중심으로 하는 새로운 산업·경제 생태계 활성화로 4차 산업혁명 및 지능정보시대 도래에 따른 국가적 대응체계 기반 마련

제2장 국내외 공공·민간 데이터 개방 및 공유 현황

국내외 공공, 민간 데이터 포털과 데이터 플랫폼 현황 조사를 통해 국내외 공공, 민간 데이터 개방 및 공유현황 파악, 오픈데이터 관련 기구를 대상으로 활동 현황과 산출물 조사

제1절 국내·외 공공·민간 데이터 포털 현황

국내외 공공, 민간 데이터 포털 현황을 조사하여 대륙 및 국가별로 데이터 포털 현황을 알아보고, 선진국과 국내 주요 데이터 포털에 대한 상세조사를 통해 데이터 플랫폼 사용현황과 데이터 포털 기능 및 데이터 제공 현황을 살펴봄

1. 해외 데이터 포털 현황

가. 국내외 데이터 포털 현황 조사

○ 전 세계 2883개[8]의 데이터 포털을 조사하여 <표 2-1>과 같이 지역별로 북아메리카(43.4%), 유럽(31.7%), 아시아(8.3%), 아프리카(7.0%), 남아메리카(5.3%) 순으로 많은 수의 데이터 포털이 집계됨

<표 2-1> 지역별 데이터 포털 현황 (개수, %)

북아메리카	유럽	아시아	아프리카	남아메리카	오세아니아	전세계	총합계
1251 (43.4%)	915 (31.7%)	239 (8.3%)	201 (7.0%)	152 (5.3%)	121 (4.2%)	4 (0.1%)	2883 (100%)

8) 웹사이트 자체 수집 및 CKAN, DKAN, OpenDataSoft 등 데이터플랫폼 운영 사이트에서 제공하는 데이터 포털 사이트 현황 자료를 바탕으로 수집
1. 자체 수집한 국내외 데이터 포털 정보
2. CKAN을 사용하는 데이터 포털 목록, https://CKAN.org/about/instances/ 참조.
3. OpenDataSoft (2015), "A COMPREHENSIVE LIST OF 2600+ OPEN DATA PORTALS AROUND THE WORLD", https://www.OpenDataSoft.com/a-comprehensive-list-of-all-open-data-portals-around-the-world/ 참조 (접속일: 2017.05.12.).
4. Open Knowledge International (2015), 전세계 오픈 데이터 포털 리스트(Data Portals), http://dataportals.org 참조 (접속일: 2017.5.14.).
5. NuCivic (2017), DKAN을 사용하는 데이터 포털 목록, https://github.com/NuCivic/DKAN-sites/ 참조 (접속일: 2017.5.12).

○ 주요 국가별로 데이터 포털 현황을 집계한 결과, <표 2-2>와 같이 미국(995개, 전체 중 34.5%), 캐나다(174개, 전체 중 6%), 프랑스(154개, 전체 중 5.3%), 영국(151개, 전체 중 5.2%)의 순으로 많은 데이터 포털이 집계됨

<표 2-2> 주요 국가별 데이터포털 수
(상위 16개국, 데이터 포털 수 기준으로 정렬)

순위	국가	지역	조사된 데이터포털 수 (개)	비율(%)
1	미국	북아메리카	995	34.5
2	캐나다	북아메리카	174	6.0
3	프랑스	유럽	154	5.3
4	영국	유럽	151	5.2
5	스페인	유럽	77	2.7
6	오스트레일리아	오세아니아	75	2.6
7	이탈리아	유럽	65	2.3
8	독일	유럽	60	2.1
9	네덜란드	유럽	55	1.9
10	콜롬비아	남아메리카	37	1.3
11	핀란드	유럽	33	1.1
12	브라질	남아메리카	28	1.0
13	스웨덴	유럽	28	1.0
14	한국	아시아	28	1.0
15	벨기에	유럽	27	0.9
16	일본	아시아	25	0.9

○ (지역별 상위 5위 국가) 지역별로 북아메리카와 유럽의 국가들에서 많은 수의 데이터 포털이 조사되었으며, 아시아권에서는 한국의 데이터 포털이 28개 조사됨
- 우리나라는 아시아권에서 가장 많은 데이터 포털이 조사 되었으나, 포털의 개수가 중요한 것은 아니며, 데이터 품질 개선에 힘써야 한다는 전문가 의견도 있음

<표 2-3> 지역별 조사된 데이터포털 수 상위 5위 국가 (데이터포털 개수)

순위	북아메리카	유럽	남아메리카	아시아
1	미국 (995)	프랑스 (154)	콜롬비아 (37)	한국 (28)
2	캐나다 (174)	영국 (151)	브라질 (28)	일본 (25)
3	멕시코 (19)	스페인 (77)	아르헨티나 (21)	중국 (18)
4	코스타리카 (14)	이탈리아 (65)	페루 (20)	타일랜드 (17)
5	엘살바도르 (6)	독일 (65)	칠레 (14)	타이완 (16)

나. 선진국 국가별 대표 오픈 데이터 포털 상세 조사

○ 선진국(27개) 국가별 대표 오픈 데이터 포털[9][10][11]에 대하여 포털 정보, 개시연도, 현재 사용 중인 데이터 플랫폼, 포털 제공 기능 조사에 대한 상세 조사를 실시함

○ 선진국의 국가별 대표 오픈 데이터 포털들의 개시년도는 2009년을 시작으로 2011년부터 본격적으로 많은 나라의 데이터 포털이 오픈되기 시작하여 2013-2016년도까지 꾸준히 새로운 오픈 데이터 포털이 개시되고 있음

- 앞에서 가장 많은 데이터 포털 수가 집계되었던 미국의 경우는 서비스 개시연도에서 보는 바와 같이 정부가 운영하는 공공데이터포털(Data.gov)의 역사가 선진국 27개 국가별 대표 오픈 데이터 포털 중 2009년도로 가장 오래되었으며, 200여개 이상의 지방 정부 및 도시, 기관들의 데이터 포털이 다양하게 운영되고 있음
- 2010년도에 영국과 이스라엘의 서비스가 개시되면서, 2011년도를 기점으로 한국을 포함, 프랑스, 스페인, 네덜란드, 브라질, 오스트리아, 싱가포르, 노르웨이, 핀란드의 공공 오픈 데이터 포털 서비스가 개시되는 등 세계 각국의 공공 데이터 포털이 본격적으로 활동하기 시작함

○ 국가별 대표 오픈 데이터 포털은 공공 데이터 포털을 중심으로 조사되었기 때문에, 포털 운영주체는 각 국가별 정부 및 정부 산하 기관에서 운영하는 것으로 나타남

- 미국과 같이 국가 차원에서 추진한 오픈 데이터 정책에 의해 시작된 경우와 영국과 같이 공공데이터 관련 법적 기반을 바탕으로 시행해오다가 공공데이터 개방원칙을 마련하면서 일원화된 공공데이터 포털 개방 및 정책 개선을 시행해온 사례도 있음[12]

---

9) Renata Máchová (2017), “Evaluating the Quality of Open Data Portals on the National Level”, *Journal of Theoretical and Applied Electric Commerce Research* [online], 12(1).

10) UNDP, GDP 등의 기준으로 선진국 선별

11) 해외 데이터 포털의 경우 공공 포털들의 경우 기존 문헌 등을 통해 비교적 쉽게 정보를 얻을 수 있으나, 민간 데이터 포털의 경우, 중소 포털 또는 데이터 마켓의 전체 규모 파악이 어렵고 개발 사이트의 상세 조사가 용이하지 않아 조사대상에서 제외함

12) 유재영 (2015), “해외 오픈데이터 정책 현황과 한국의 오픈데이터 정책 발전방안”, 한국정보화진흥원.

<표 2-4> 선진국 국가별 대표 오픈 데이터 포털 일반사항

(서비스 개시 연도 순으로 정렬, 개시연도가 같은 경우 <표2-6>의 제공하고 있는 데이터 건수로 정렬)

No	국가	주요 데이터 포털	구분	포털 운영주체	포털접근경로(URL)	서비스 개시 연도
1	미국	Us Government Open Data (Data.gov)	공공	미국정부	http://data.gov	2009
2	영국	data.gov.uk	공공	영국정부	https://data.gov.uk/	2010
3	이스라엘	Israel Government Portal Open Data	공공	이스라엘정부	https://data.gov.il	2010
4	프랑스	Data.gov.fr	공공	프랑스정부	http://www.data.gouv.fr	2011
5	한국	공공데이터포털	공공	행정자치부, NIA	https://www.data.go.kr/	2011
6	이탈리아	Italian Government Open Data	공공	정부 (디지털이태리)	http://www.dati.gov.it/	2011
7	스페인	Spain Government Data Portal	공공	에너지·관광·디지털	http://datos.gob.es	2011
8	네덜란드	Netherlands Public Data Catalogue	공공	네덜란드 정부	http://data.overheid.nl/	2011
9	브라질	dados.gov.br	공공	정부(정보기술 사무국)	http://dados.gov.br	2011
10	오스트리아	Open Government Data Portal Austria	공공	오스트리아정부	https://www.data.gv.at/	2011
11	싱가포르	Singapore's Open Data Portal	공공	싱가포르 정부	http://data.gov.sg	2011
12	노르웨이	Norwegian Government Open Data	공공	노르웨이 정부	http://data.norge.no/	2011
13	인도	Data Portal India	공공	인도정부	https://data.gov.in/	2012
14	캐나다	Open data in Canada	공공	캐나다정부	http://open.canada.ca/	2013
15	오스트레일리아	Australian Government Catalogue	공공	호주정부	http://data.gov.au/	2013
16	인도네시아	Indonesia Data Portal	공공	대통령직속 개발통제 관리부서	http://data.go.id	2013
17	핀란드	Finland Open Data	공공	핀란드정보	https://www.opendata.fi/	2013
18	멕시코	datos.gob.mx	공공	멕시코 정부	http://datos.gob.mx	2014
19	일본	Japanese Government Open Data	공공	정부	http://www.data.go.jp	2014
20	스위스	Opendata Swiss	공공	연방정부	https://opendata.swiss/en/	2014
21	덴마크	Denmark Opendata	공공	덴마크 수도 및 지방자치단체	http://portal.opendata.dk/	2014
22	러시아	Russie	공공	러시아정부	http://data.gov.ru/	2015

No	국가	주요 데이터 포털	구분	포털 운영주체	포털접근경로(URL)	서비스 개시 연도
23	벨기에	Belgium Government Open Data	공공	벨기에 정부	http://data.gov.be	2015
24	아일랜드	data.gov.ie	공공	정부 (공공집행 및 개혁부)	https://data.gov.ie/data	2015
25	독일	German Government Open Data	공공	독일정부	https://www.govdata.de/	2015
26	뉴질랜드	New Zealand Government Open Data	공공	정부	https://data.govt.nz/	2016
27	스웨덴	Sweden Open Data	공공	국립 문서 보관소	https://oppnadata.se	2016

○ 데이터 포털에서 사용하고 있는 데이터 플랫폼에 대하여 조사 결과 <표 2-5>와 같이, 현재 27개국 중 21개국에서 오픈소스 플랫폼인 CKAN을 사용하고 있으며, 그 외에도 OGPL(인도), OpenDataSoft(프랑스) 등 데이터플랫폼을 사용하거나, 자체 플랫폼을 사용하고 있음

- 한국의 경우 CKAN 카달로그 기능을 일부 사용하고 있으면서, CKAN 도입계획을 가지고 있음
- 노르웨이의 경우, 메타 데이터 연계 표준인 DCAT표준을 제공(노르웨이)하는 등, 데이터 연계를 위한 정책을 가지고 있는 경우가 있음

○ 조사대상 27개 포털에서 공통적으로 데이터셋에 대한 메타 데이터의 검색과 조회가 가능한 데이터 카달로그 기능을 제공하고 있음

○ 데이터 융합[13] 및 분석 기능의 경우, 일부 포털(미국의 Data.gov, 러시아의 공공데이터 포털)에서만 제공하고 대부분은 제공하고 있지 않음

- 가장 많은 공공 데이터 포털에서 사용하고 있었던 오픈소스 플랫폼인 CKAN을 사용하고 있는 경우 메타 데이터 통합 및 관리에 초점을 두고 있어 융합 및 분석 시각화의 경우 데이터 포털 자체적으로 제공하는 경우는 드물고 추가적인 tool 및 활용을 위한 웹사이트에서 제공하는 경우가 있었음

○ 시각화 기능은 대부분 플랫폼 기본 제공 요소인 리소스 미리보기 형태로 제공함

- 차트 및 그래프 시각화 기능은 일부 포털(미국, 멕시코, 벨기에, 뉴질랜드, 브라질, 스위스, 싱가포르, 노르웨이 등)에서 제공하고 있음

13) 데이터 융합 기능은 서로 다른 데이터셋을 통합하는 기능이 있는 것으로 봄

○ 결제 기능은 조사 대상이 공공 데이터 포털이어, 무료 이용가능 하고, 대량의 데이터를 이용하기 위한 소정의 수수료를 부가하는 경우가 있어 포털 자체적인 결제 기능은 가지고 있는 경우는 없었음

<표 2-5> 선진국 국가별 대표 오픈 데이터포털 - 기능 비교

No	국가	데이터포털 명	플랫폼정보	목록	데이터 융합	데이터 분석	시각화	결제
1	미국	Us Government Open Data	CKAN	o	x	o	o	x
2	캐나다	Open data in Canada	CKAN	o	x	x	x	x
3	인도	Data Portal India	OGPL	o	x	x	o	x
4	영국	data.gov.uk	CKAN	o	x	x	x	x
5	오스트레일리아	Australian Government Catalogue	CKAN	o	x	x	x	x
6	멕시코	datos.gob.mx	CKAN	o	x	x	o	x
7	프랑스	Data.gov.fr	OpenDataSoft	o	x	x	x	x
8	한국	공공데이터포털	자체플랫폼 (CKAN도입계획)	o	x	x	x	x
9	독일	German Government Open Data	CKAN	o	x	x	x	x
10	일본	Japanese Government Open Data	CKAN	o	x	x	x	x
11	이탈리아	Italian Government Open Data	CKAN	o	x	x	x	x
12	스페인	Spain Government Data Portal	CKAN	o	x	x	x	x
13	러시아	Russie	DKAN	o	x	o	x	x
14	네덜란드	Netherlands Public Data Catalogue	CKAN	o	x	x	x	x
15	벨기에	Belgium Government Open Data	Drupal 웹사이트	o	x	x	o	x
16	아일랜드	data.gov.ie	CKAN	o	x	x	x	x
17	뉴질랜드	New Zealand Government Open Data	CKAN	o	x	x	o	x
18	브라질	dados.gov.br	CKAN	o	o	x	o	x
19	오스트리아	Open Government Data Portal Austria	OGD	o	x	x	x	x
20	스위스	Opendata.Swiss	CKAN	o	o	x	o	x
21	인도네시아	Indonesia Data Portal	CKAN	o	x	x	x	x
22	싱가포르	Singapore's Open Data Portal	CKAN	o	x	x	o	x
23	핀란드	Finland Open Data	CKAN	o	x	x	x	x
24	노르웨이	Norwegian Government Open Data	자체플랫폼 (DCAT표준)	o	x	x	o	x
25	덴마크	Denmark Opendata	CKAN	o	x	x	x	x
26	스웨덴	Sweden Open Data	CKAN	o	x	x	x	x
27	이스라엘	Israel Government Portal Open Data	CKAN	o	x	x	x	x

○ 데이터 제공 현황은 미국, 캐나다, 인도의 데이터셋 제공 건수가 10만 건 이상으로 가장 많았으며, 대다수 공공 데이터 제공 비율 100%로 민간데이터를 제공하는 경우는 없고, 무료로 데이터를 제공

○ 데이터 이용 권한은 열람, 다운로드를 허용하는 경우가 가장 많고, 융합, 가공 가능한 경우가 일부 있음

- 해당 포털 데이터셋 전체에 해당하는 오픈 데이터 라이선스를 자체적으로 보유하는 사례(이탈리아, 싱가포르, 노르웨이, 덴마크 등)를 찾아볼 수 있음

〈표 2-6〉 선진국 국가별 대표 오픈 데이터포털 데이터 제공 건수 및 이용권한

(제공 데이터셋 건수 기준으로 정렬)

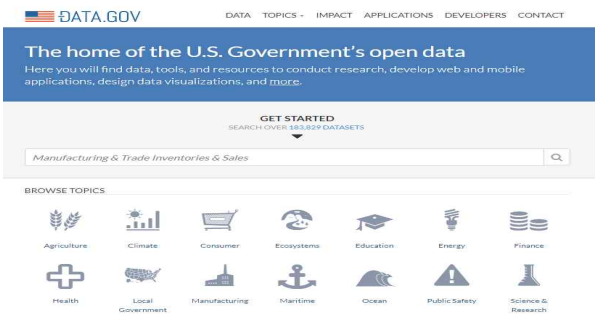
No	국가	데이터포털 명	건수	데이터이용권한
1	미국	Us Government Open Data	192,507	열람, 다운로드, 융합가공 가능 (데이터셋에 명시된 licence에 따름)
2	캐나다	Open data in Canada	118,265	열람, 다운로드
3	인도	Data Portal India	106,372	열람, 다운로드
4	영국	data.gov.uk	41,987	열람, 다운로드
5	오스트레일리아	Australian Government Catalogue	28,305	열람, 다운로드
6	멕시코	datos.gob.mx	26,156	열람, 다운로드
7	프랑스	Data.gov.fr	25,076	복사,배포,융합,가공 가능
8	한국	공공데이터포털	22,069	열람, 다운로드
9	독일	German Government Open Data	19,877	열람, 다운로드
10	일본	Japanese Government Open Data	19,368	열람, 다운로드
11	이탈리아	Italian Government Open Data	17,907	열람, 다운로드 이탈리아 오픈데이터 라이센스 (IODL v2.0) 적용
12	스페인	Spain Government Data Portal	15,323	열람, 다운로드
13	러시아	Russie	14,038	열람, 다운로드
14	네덜란드	Netherlands Public Data Catalogue	10,958	열람, 다운로드
15	벨기에	Belgium Government Open Data	9,816	열람, 다운로드
16	아일랜드	data.gov.ie	5,331	열람, 다운로드
17	뉴질랜드	New Zealand Government Open Data	5,069	열람, 다운로드
18	브라질	dados.gov.br	2,977	열람, 다운로드
19	오스트리아	Open Government Data Portal Austria	2,425	열람, 다운로드
20	스위스	Opendata.Swiss	2,346	열람, 다운로드
21	인도네시아	Indonesia Data Portal	2,223	열람, 다운로드
22	싱가포르	Singapore's Open Data Portal	1786	열람, 다운로드, 복사, 배포, 수정 (싱가포르 오픈 데이터 라이센스 1.0 적용)[14]
23	핀란드	Finland Open Data	1,720	열람, 다운로드
24	노르웨이	Norwegian Government Open Data	854	다운로드, 복사, 사용, 변경, 수정 및 결합, 상업적 활용 가능 / (노르웨이 오픈데이터 라이센스 (NLOD) 적용)
25	덴마크	Denmark Opendata	788	열람, 다운로드, 복사, 배포, 게시, 수정 및 결합, 상업적/비상업적 활용가능 (Open Data DK 라이센스 1.0 적용)
26	스웨덴	Sweden Open Data	493	열람, 다운로드
27	이스라엘	Israel Government Portal Open Data	406	복사, 배포, 공개, 변경, 파생물 작성, 상업적/비상업적 활용가능

다. 해외 주요국 데이터 포털의 특징

○ 해외 주요 국가에서 운영 중인 데이터 포털의 특징과 주제 분야별 데이터셋 현황을 살펴봄

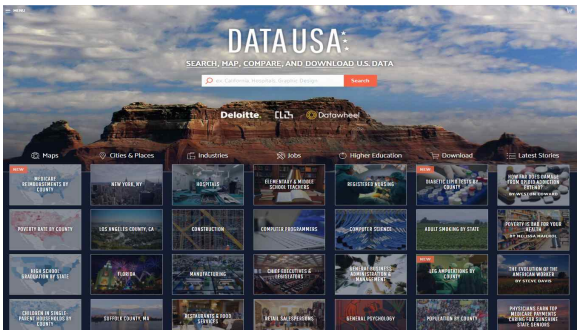
○ 미국 정부는 공공데이터 포털 Data.gov를 구축했던 플랫폼을 오픈 소스로 공개하였고, 이후 개방형 데이터 플랫폼 CKAN 기반으로 개편하였음

- 2009년 열린 정부 지침을 발표하면서 공공데이터포털 Data.gov를 발표하고, 데이터 구축 플랫폼인 'Open Government Platform(OGPL)'을 오픈소스로 공개하였음[15)]
- 2009년 47개 데이터셋으로 시작한 Data.gov는 2013년, 오픈소스 데이터 플랫폼 CKAN을 기반으로 한 데이터 카탈로그를 기반으로 한 Data.gov 2.0으로 다시 시작하였으며, [16)17)] 2017년 현재, 190,000여개의 검색 가능한 데이터셋을 제공함[18)]
- 미국 정부는 민간과 학계, 정부의 협업으로 운영되는 공공데이터의 정보 시각화를 강화한 DATA USA (https://datausa.io/)를 별도로 운영 중[19)]
- 데이터 포털에서 제공 중인 데이터셋 건 수 상위 5위를 살펴보면 건강(Health), 농업(Agriculture), 기후(Climate), 교육(Education), 에너지(Energy)의 순서로 나타났으며, 전체 197,575건을 제공하고 있음



<그림 2-1> 미국 국가데이터포털, DATA.GOV

※ 출처 : DATA.GOV http://data.gov 참조 (접속일: 2017.08.21.).



<그림 2-2> 미국 공공데이터 정보 시각화 사이트 DATA.USA

※ 출처 : DATA.USA http://data.usa 참조 (접속일: 2017.08.21.).

---

14) Data.gov.sg (2017), Singapore Open Data Licence, https://data.gov.sg/open-data-licence 참조 (접속일: 2017.05.17.).

15) Open Government Platform(OGPL), https://github.com/opengovplatform/opengovplatform-beta 참조 (접속일: 2017.05.12).

16) DATA.GOV (2013.01.28.), "CKAN: the Horizon for Data.gov 2.0", https://www.data.gov/blog/CKAN-horizon-datagov-20 참조 (접속일: 2017.05.12.).

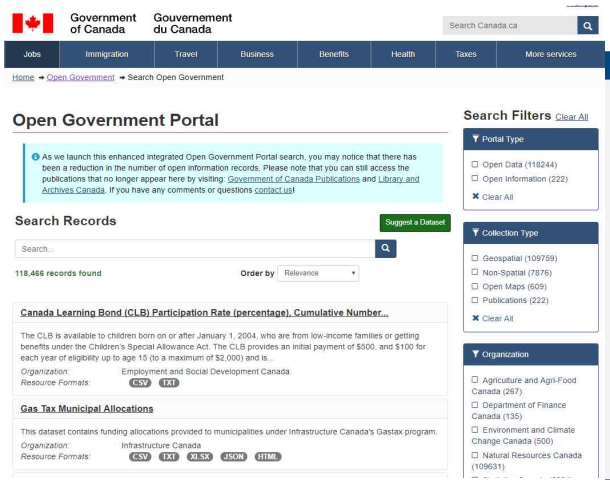
17) Mark Wainwright (2013), "US government's data.gov to use CKAN", https://CKAN.org/2013/02/04/us-data-gov-to-use-CKAN/ 참조 (접속일: 2017.5.19).

18) 정유철 (2016), "오픈 데이터 플랫폼 동향", 「정보처리학회지」, 23(5): 53-63.

19) 이재원 (2016), "공공데이터 포털 국내외 동향 분석", 「ICT Veiwer」, 한국정보화진흥원.

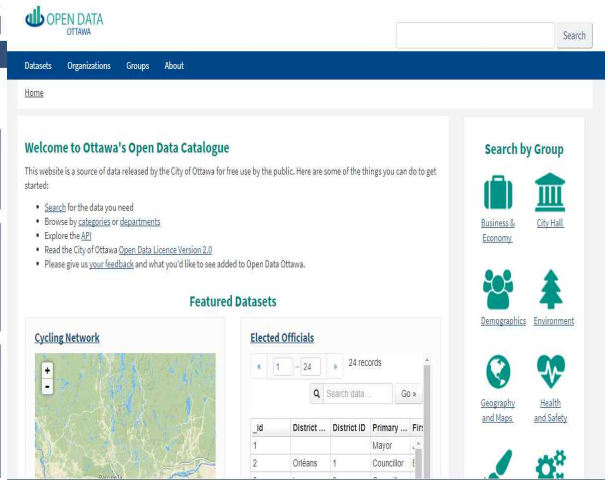
○ 캐나다 정부는 대표 오픈 데이터 포털인 Open Data Portal(data.gc.ca)을 2011년도에 오픈, 2013년도 CKAN 2.0을 기반으로 개편[20][21]함

- 캐나다 정부의 오픈 데이터 포털이 개편될 당시, API기능, 검색 기능 등을 개선하기 위해 CKAN으로 개편한 오타와 시의 데이터 카탈로그와 연계하여 중앙정부와 지방정부의 데이터 카달로그 연계가 가능했음
- 캐나다 공공데이터 포털의 경우 환경(Nature and Environment), 경제/산업(Economics and Industry), 정부/정치(Government and Politics), 과학/기술(Science and Technology), 개인(Persons)의 순으로 총 118,472건 제공하고 있음



<그림 2-3> 캐나다 대표 오픈 데이터 포털

※ 출처: http://open.canada.ca 참조 (접속일: 2017.08.21.).



<그림 2-4> 오픈 데이터 오타와

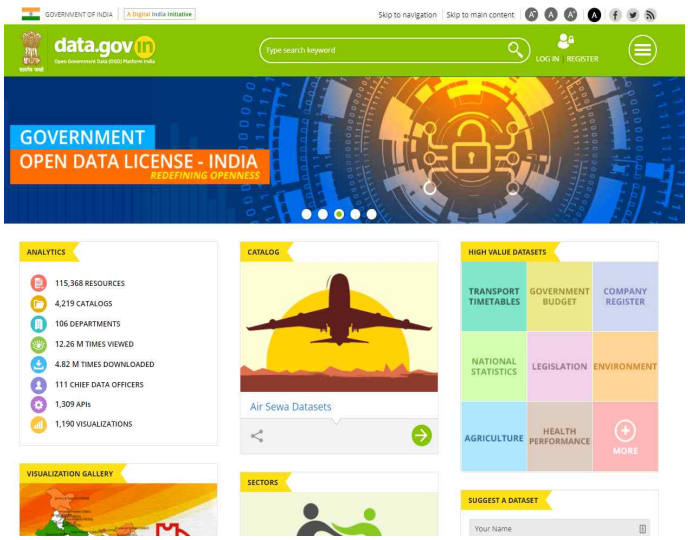
※ 출처: http://data.ottawa.ca 참조 (접속일: 2017.08.21.).

20) 조선일보 (2015.11.09.), "오바마의 '열린 정부'… 캐나다의 '서비스 캐나다'", http://biz.chosun.com/site/data/html_dir/2015/11/08/2015110802578.html 참조 (접속일: 2017.08.21.).

21) iCON 브리프 (2016. 03.09), "캐나다 오픈데이터 현황", http://icon.ndsl.kr/i_trend/icon_briefDetail.jsp?record_no=5261&trendType=O 참조 (접속일: 2017.08.21.).

○ 인도 정부는 2011년 National Data Sharing and Accessibility policy(Nasap)을 승인하면서, 오픈 데이터 포털(data.gov.in)을 발표[22]한 후, 2012년도 10월 사이트를 오픈[23]하여 미국, 인도 정부가 공동 개발한 오픈 데이터플랫폼 OGPL(Open Government Platform)을 바탕으로 운영되고 있음

- 미국 General Service Administration 팀에서 데이터 관리 시스템 모듈을 개발하고, 인도 National Informatics Center에서 웹사이트, 콘텐츠 관리 시스템(CMS)를 개발하여 인도와 미국이 공동으로 개발함[24][25]
- 인도 공공데이터 포털의 경우 통계(Statistics), 수자원(Water Resources), 농업(Agriculture), 보건/가족(Health and Family welfare), 인도의회(Parliament Of india)의 순으로 총 자원 106,372건 (카달로그 4,236건) 제공하고 있음



<그림 2-5> 인도 오픈데이터 포털

※ 출처 : http://data.gov.in 참조
(접속일:2017.08.21.)

22) Nikhil Pahwa (2011.06.06.), "Data.Gov.In To Launch; Use Cases: Powercuts.in; Challenges", MEDIANAMA, https://www.medianama.com/2011/06/223-data-gov-in-to-launch-use-cases-powercuts-in-challenges/ 참조 (접속일: 2017.08.21.).

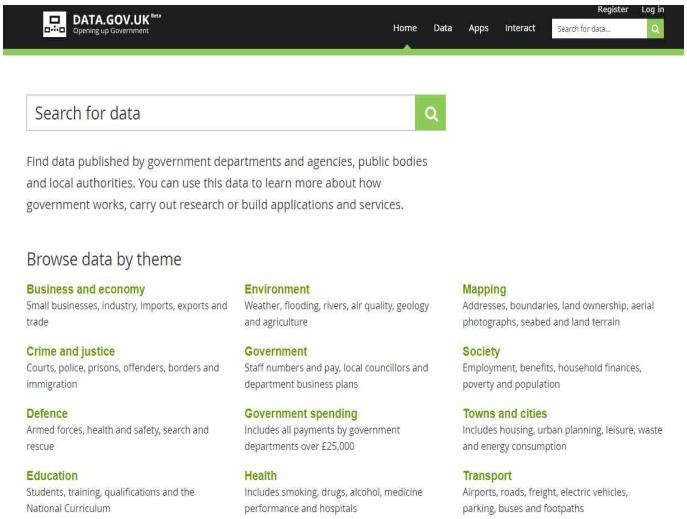
23) The Economic Times (2011.06.04.), "New Govt portal Data.Gov.In launch next month to ease information search", https://en.wikipedia.org/wiki/Data.gov.in#cite_note-2 참조 (접속일: 2017.08.21.).

24) Press Information Bureau, Government of India (2012.03.30.), "Shri Kapil Sibal Launches 'Open Government Platform' Developed by India and US Initiative will Spur Citizen Engagement and Interface with Government", http://pib.nic.in/newsite/erelease.aspx?relid=82025 참조 (접속일: 2017.08.21.).

25) Nikhil Pahwa (2012.03.30.), "US &India Governments Launch Open Government Platform", MEDIANAMA, https://www.medianama.com/2012/03/223-open-government-platform/ 참조 (접속일: 2017.08.21.).

○ 영국 정부는 공공데이터 포털 DATA.GOV.UK을 2009년 9월 비공개 베타버전으로 시작

- 2010년 1월 정식으로 공개 베타버전을 오픈하여 운영하고 있으며 Open Knowledge International에서 개발한 오픈소스 데이터 플랫폼 CKAN을 사용하고 있음[26)]
- 영국 공공데이터 포털의 경우 환경(Environment), 지역/도시(Towns & Cities), 정부(Government), 사회(Society), 지도(Mapping), 건강(Health)의 순으로 총 41,987건 제공하고 있음



<그림 2-6> 영국 공공데이터 포털

※ 출처 : http://data.gov.uk 참조
(접속일:2017.08.21.)

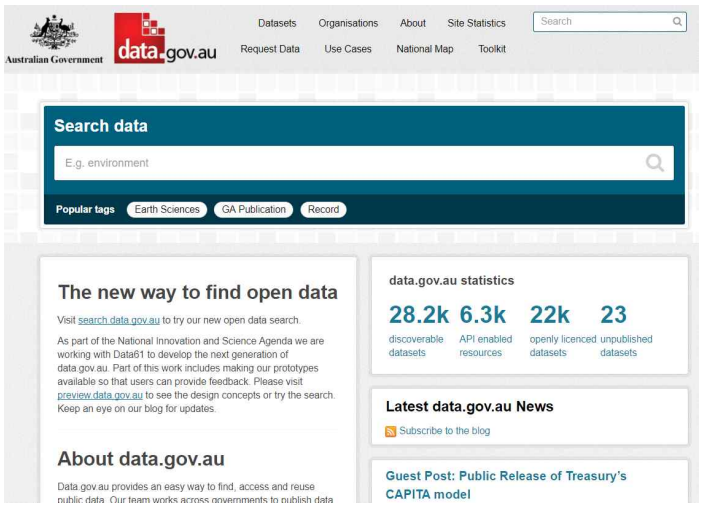
○ 오스트레일리아 정부는 2013년 공공데이터 포털인 data.gov.au를 공개하였으며 처음에는 Wordpress 기반으로 되어있던 공공 데이터 포털을 클라우드 기반의 CKAN 플랫폼으로 변경하여 사용하고 있음

- 당시 흔치 않았던 CKAN의 클라우드 기반 호스팅 사례[27)]로서, 내보내기 기능이 없던 기존 레거시 시스템에서 데이터를 이행하기 위해 수행사[28)]가 데이터 이행 도구를 개발함
- 기존 데이터에 대한 프로그래밍 방식의 접근이나 API에 대한 링크 및 데이터 스프레드 시트, 그래프/매핑과 같은 미리보기를 즉시 사용할 수 있도록 구성하였으며,

26) 한은영 (2014), “영국 오픈데이터 정책의 특징 및 시사점”, 「정보통신방송정책」, 26(23).
27) CKAN, “CKAN Portfolio - data.gov.au”, https://CKAN.org/portfolio/data-gov-au/ 참조 (접속일: 2017.09.09.)
28) Link Digital, “Federal Government Open Data Portal”, https://linkdigital.com.au/content/datagovau-0 참조 (접속일: 2017.05.25.)

아마존 웹서비스(AWS)에서 활용할 수 있는 참조 아키텍처를 개발함

- 데이터셋을 자동으로 업데이트 하는 CKAN의 기능 사용으로 매일 여러 데이터셋을 자동으로 게시하는 것이 가능해짐
- 오스트레일리아 공공데이터 포털의 경우 경영지원/교육(Business Support and Education), 지역사회 서비스(Community Services), 환경(Environment), 과학(Science), 시민사회기반(Civic Infrastructure)의 순으로 총28,305건 건 제공하고 있음



<그림 2-7> 오스트레일리아 공공데이터 포털

※ 출처: http://data.gov.au 참조
(접속일: 2017.10.11.)

○ 주요국 데이터 포털의 주제 분야별 분류체계가 상이하여 일괄적인 비교는 어려우나, 환경, 과학, 정부 관련 데이터셋이 상대적으로 많이 공개되고 있음

<표 2-7> 해외 주요 데이터 포털 주제별 데이터셋 집계 현황

데이터 포털 (건수)	US Government Open Data (건수 :197,575)	Open data in Canada (건수 : 118,472)	data.gov.uk (건수: 41,987)	Dataportal india (건수: 106,372) (카달로그4,236건)	Australian Government Catalogue (건수: 28,305)
구분	공공	공공	공공	공공	공공
국가	미국	캐나다	영국	인도	오스트레일리아
주제 (건수)	Local Government (16391) Health (1854) Agriculture (648) Climate (654) Education (353) Energy (263) Ecosystems (234) Ocean (179) Consumer (131) Finance (129) Public Safety (83) Maritime (48) Manufacturing (38) Science & Research (37)	Form Descriptors (106061) Nature and Environment (70736) Economics and Industry (65782) Government and Politics (63912) Science and Technology (6292) Persons (1748) Society and Culture (1722) Agriculture (1495) Health and Safety (720) Transport (613) Labour (553) Education and Training (318) Information and Communications (291) Law (167) History and Archaeology (97) Arts, Music, Literature (53) Military (25) Processes (21) Language and Linguistics (20)	Environment (14395) Towns & Cities (7257) Government (3156) Society (3035) Mapping (2218) Health (2087) Government Spending (1861) Education (1350) Transport (1204) Business & Economy (1132)	Statistics (625) Water Resources (562) Agriculture (435) Health and Family welfare (301) Parliament Of india (254) Finance (178) Education(174) Transport(166) Home Affairs and Enforcement (148) Power and Energy (139) Environment and Forest (106) Industries (88) Animal Husbandry (76) Rural (73) Labour and Employment (66) Science and Technology (53) Infrastructure (52) Social Development (48) Biotechnology (42) Information and Communications (40) Commerce (37) Governance and Administration (29) Defence (28) Urban (22) Mining (21) Water and Sanitation (21) Housing (17) Travel and Tourism (11) Information and Broadcasting (9) Economy (9) Food (4) Foreign Affairs (3) Art and Culture (2)	Business Support and Education (76) Community Services (66) Environment (60) Science (49) Civic Infrastructure (46) Finance Management (34) Statistical Services (32) Governance (32) Sport and Recreation (29) Health Care (20) Communications (19) Cultural Affairs (15) Employment (13) Transport (11) Trade (7) Education and Training (6) Primary Industries (4) Indigenous Affairs (4) Government (4) Natural Resources (3) Tourism (2) Planning (2) GovHack (2) Geography (2) Emergencies (2) Technology (1) Society (1) Safety (1) Property (1) News (1) Law (1) Justice Administration (1) Information-communications-technologies (1) Immigration (1)

2. 국내 데이터 포털 현황

○ 국내 데이터 포털 및 데이터 마켓 현황을 조사하고 운영주체에 따라 중앙정부・지방정부・공공기관・민간으로, 포털 운영방식에 따라 데이터 포털・데이터 마켓으로 구분함

<표 2-8> 국내 데이터 포털 및 데이터 마켓 현황

No	운영 주체 구분	데이터 포털 /데이터 마켓 구분	데이터포털	포털 접근경로(URL)
1	중앙 정부	데이터포털	공공데이터포털	http://data.go.kr/
2			국가공간정보포털	http://nsdi.go.kr
3			재난안전데이터포털	http://data.mpss.go.kr
4			농림축산식품 공공데이터 포털	http://data.mafra.go.kr
5			교통정보공개서비스	http://openapi.its.go.kr
6			기상자료개방포털	http://data.kma.go.kr/cmmn/main.do
7			공간정보 오픈플랫폼	http://www.vworld.kr
8			KOSIS 국가통계포털	http://kosis.kr
9	지방 정부	데이터포털	수원시 공공데이터	http://data.suwon.go.kr/
10			부산시 공공데이터	http://opendata.busan.go.kr/
11			강원 공공데이터	http://data.gwd.go.kr/
12			경기데이터드림	http://data.gg.go.kr/
13			인천광역시 데이터공유센터	http://open.incheon.go.kr/
14			제주특별자치구역	http://data.jeju.go.kr/
15			경상북도 공공데이터포털	http://data.gb.go.kr/
16			전라북도 오픈데이터포털	http://opendata.go.kr/
17			세종특별자치시 데이터실록	http://data.sejong.go.kr/
18			서울 열린 데이터 광장	http://data.seoul.go.kr
19			전라남도 공공데이터 커뮤니티 센터 (공공데이터 포털)	http://data.jeonnam.go.kr/
20			울산광역시 공공데이터 포털	http://data.ulsan.go.kr/
21	공공 기관	데이터마켓	데이터스토어	http://datastore.or.kr
22		데이터포털	한국언론진흥재단(빅카인즈)	http://www.kinds.or.kr/
23			보건의료빅데이터개방시스템	http://opendata.hira.or.kr
24	민간	데이터포털	Linked Data Hub	http://datahub.kr
25			한국에스리	http://data.esrikr-defense.opendata.arcgis.com
26		데이터마켓	빅데이터허브	http://bigdatahub.co.kr
27			API store	http://apistore.co.kr
28			오디피아	http://odpia.org

○ 국내 대표 데이터 포털 및 데이터 마켓 현황 상세 조사를 위해 공공 및 민간에서 운영하는 국내 데이터 포털 및 데이터 마켓 중 6개를 선정[29]하여, 서비스 개시연도, 데이터 플랫폼 정보와 제공하는 기능, 데이터 제공 현황 등 상세조사를 실시함

- 국내 데이터 포털의 경우 중앙정부, 지방정부, 공공기관, 민간에서 운영되고 있는 데이터 포털들을 찾아볼 수 있음

○ 공공과 민간에서 운영하는 데이터 포털과 마켓을 살펴본 결과, 대부분 자체 플랫폼을 사용하고 있음

- 공공에서 운영하며 자체 플랫폼을 활용하고 있는 경우에도 일부 CKAN을 활용하여 데이터 목록(카달로그)을 제공 하거나, 일부 데이터셋을 CKAN API로 제공(수원시)하고 있거나, 향후 오픈소스 데이터 플랫폼 CKAN을 기반으로 데이터를 연계하려는 움직임을 보이고 있으나 민간 대기업에서 운영하고 있는 데이터 마켓의 경우 자체 플랫폼을 통해 운영되고 있음

<표 2-9> 국내 데이터 포털 상세 현황

운영 구분	포털 마켓 구분	이름	포털 운영주체	서비스 개시 연도	플랫폼 정보	목록	데이터 융합	데이터 분석	시각화	결제	건수	유,무료 현황	데이터 이용권한 (단순열람, 다운로드, 융합 가공)
공공	포털	공공 데이터 포털	행정 자치부, NIA	2011	자체 플랫폼, CKAN	o (CKAN catalog 활용)	x	x	x	x	22069	무료	열람, 다운로드, 융합가공 가능
	마켓	데이터 스토어	한국 데이터 진흥원	2013	자체 플랫폼 → CKAN (구축 중)	o	x	x	o	o	7226	유료, 무료	(샘플데이터) 열람, (결제 후) 다운로드
	포털	Data for Suwon	수원시	2015	CKAN API제공	o	x	x	o	x	590	무료	열람, 다운로드, 융합가공 가능
민간	마켓	빅데이터 허브	SKT	2013	자체 플랫폼	o	o	o	o	x	958	유료, 무료	열람, 다운로드 가능
	마켓	오디피아	LG CNS	2016	자체 플랫폼	o	x	o	o	x	4513	무료 (유료 서비스 별도)	열람, 다운로드 (준비 중)
	마켓	API store	KT	2013	자체 플랫폼 (Waple)	o	x	x	x	o	2955	유료, 무료	결제 후 API 사용 가능 (개인사용자 사용불가)

29) 공공에서 운영하는 데이터 포털에 대하여 중앙정부 1곳, 지방 정부 1곳, 공공 기관에서 운영하는 데이터 마켓 1곳을 각각 지정하고, 현재 민간에서 운영되고 있는 대표적인 데이터마켓 3곳을 지정하여 상세조사함

3. 비교 및 시사점

○ 선진국의 대표 오픈 데이터 포털을 대상으로 상세조사를 한 결과, 대륙별로는 북아메리카와 유럽이, 국가별로는 미국, 캐나다, 프랑스, 영국이 조사된 데이터 포털 개수가 상위권에 속함

○ 선진국 국가별 대표 데이터포털인 공공데이터 포털의 서비스 개시연도를 보면 데이터 포털 개수가 가장 많이 조사되었던 미국이 2009년도에, 영국이 2010년의 뒤를 이어 2011년에 많은 국가들이 공공데이터 포털을 서비스하기 시작했고, 한국 역시 2011년 같은 시기에 공공데이터 포털이 서비스 개시되었음

○ 선진국 국가별 데이터 포털에서 제공하는 기능은 공통적으로 데이터셋에 대한 메타 데이터의 검색과 조회가 가능한 데이터 카달로그 기능제공이었고, 이는 국내 데이터 포털 및 데이터 마켓에서도 모두 지원하고 있음

○ 선진국 대표 오픈 데이터 포털들은 대다수 데이터 연계가 용이한 오픈소스 기반의 글로벌 데이터 플랫폼 기반으로 운영되고 있는 반면, 국내는 아직 자체 플랫폼 기반으로 서비스되고 있는 비중이 높으나, CKAN 카달로그 기능 활용 및 차후 CKAN 적용을 위한 구축계획을 가지고 있는데 데이터 연계를 향한 움직임을 보이고 있음

○ 해외 주요국가의 사례를 보면, 데이터 포털을 오픈하면서 자체 플랫폼이나 상용 데이터 플랫폼(Socrata 등)으로 시작하였으나, 세계적으로 타 데이터 포털과의 연계를 위하여 오픈소스 기반의 CKAN과 같은 데이터 플랫폼으로 전환한 사례(미국, 캐나다 등)를 찾아 볼 수 있고, 정부의 대표 오픈 데이터 포털을 운영하면서 쌓인 노하우를 데이터 플랫폼의 형태로 개방한 사례(미국, 영국 등)를 찾아볼 수 있음

○ 해외 주요 국가의 공공 데이터 포털에서 공급되고 있는 데이터셋에서 가장 높은 비율을 차지하는 데이터셋은 공통적으로 환경, 기후, 농업, 복지, 교육, 지역, 경제/산업에 대한 주제들로 연계 시 부가가치가 높은 주제 분야들임

○ 앞서 선진국 공공데이터 포털에서 결제기능이 없었던 것과는 달리, 국내 데이터 포털 중 데이터스토어(www.datastore.or.kr) 같은 경우, 다른 공공데이터 포털과는 다르게 유료 데이터를 찾아볼 수 있으며 결제기능을 포함하고 있는 이유는 국가 공공

데이터 이외에 민간 기업들이 판매하는 데이터가 다수 존재하는 데이터 마켓 형태를 띠고 있기 때문임

○ 데이터 마켓의 경우 공공에서 운영하는 데이터스토어(www.datastore.or.kr)와 민간 기업에서 운영하는 데이터 마켓 3개를 살펴본 결과 데이터 포털에 비하여 데이터의 가치를 탐색할 수 있는 시각화 기능이 더 많이 제공되고 있었으며, 특히 민간 데이터 마켓의 경우 데이터 분석 기능 및 융합이 특화된 것을 찾아볼 수 있음

○ 국내외 데이터 포털의 경우 시각화, 융합, 분석 기능이 미약한 것에 비해 민간 데이터 마켓에서는 연계, 융합을 포함한 데이터의 가치와 활용에 중점을 두고 있는 것을 확인할 수 있음

○ 국내외 데이터 포털 조사 결과, 정부를 중심으로 공공데이터의 개방에 중점을 두고 있고, 향후 데이터를 연계할 수 있는 방향으로 데이터 플랫폼이 변화하고 있으며, 점차 가치 있는 민간데이터와의 연계를 위한 노력이 확산되고 있음

제2절 국내외 데이터 플랫폼 동향

국내외 데이터 포털에 적용된 솔루션 중심으로 데이터 플랫폼 동향을 조사하여, 가장 많이 활용되고 있는 플랫폼과 그 특징을 살펴보고자 함. 또한 플랫폼 간 장단점을 파악하여 국내 적용 시 참고할 수 있도록 그 타당성을 알아보고자 함

1. 데이터 플랫폼 개요 및 특징

○ 국내외 데이터 포털에서 사용되고 있는 주요 데이터 플랫폼들을 개발주체에 따라 민간, 비영리로 나누고, 오픈소스 활용 여부에 대하여 조사함

- 비영리 단체인 Open Knowledge에서 개발한 CKAN이 있으며, 커뮤니티 기반으로 만들어지고 있는 오픈소스 DKAN의 경우 오픈소스 CKAN을 기반으로 확장성이 좋은 Drupal을 활용하여 만들어진 개방형 데이터 플랫폼으로 민간 기업의 지원을 받고 있음
- 비영리 단체인 Open Knowledge에서 개발된 CKAN과 미국 정부에서 개발된 OGPL을 제외한 데이터 플랫폼 이외에 다수의 데이터 플랫폼(Socrata, Junar 등)이 민간 기업에 의해 개발, 유지되고 있음

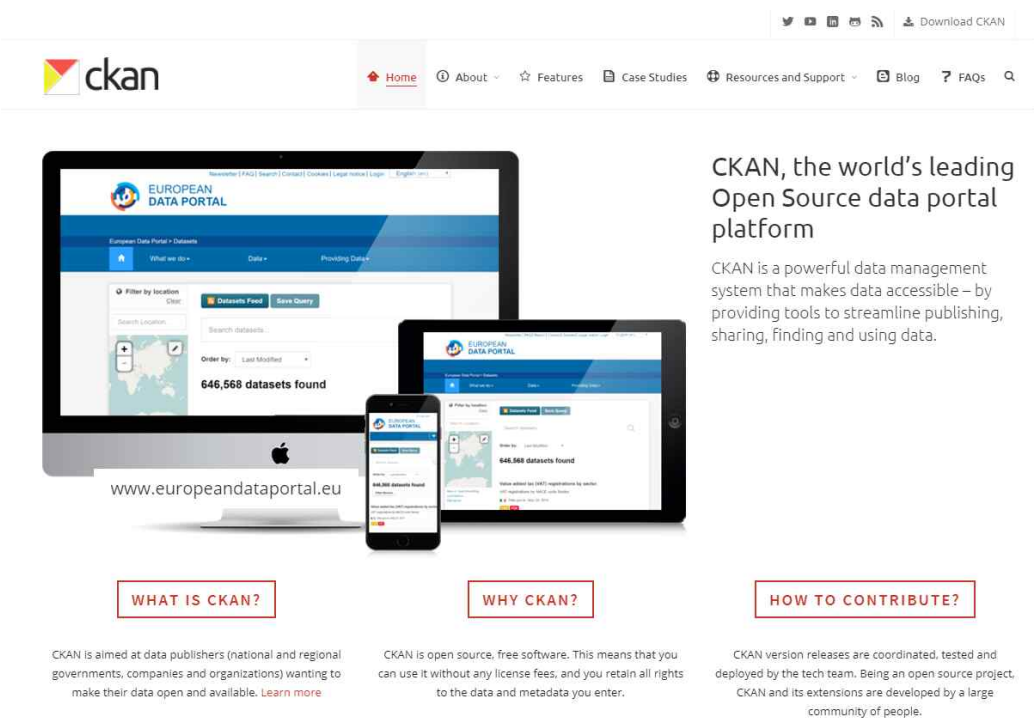
○ 국내외 데이터 포털에서 사용되고 있는 주요 데이터 플랫폼은 CKAN, Socrata, DKAN, Junar, OpenDataSoft이며, 그 외 소수로 Datatank, Enigma Public, Information Workbench, OGPL, Open Data Store 등이 있음

<표 2-10> 데이터 플랫폼 개발주체 및 오픈소스 활용 여부

플랫폼 이름	개발주체	구분	오픈소스 활용 여부
ArcGIS Open Data	esri	민간	x
CKAN	Open Knowledge International	비영리	o
Datatank	Datatank	민간	o
DKAN	커뮤니티 개발 Granicus[30] 지원	비영리, 민간	o (CKAN기반)
Enigma Public	Enigma Technologies, Inc	민간	x
esri GeoPortal	esri	민간	o
Information Workbench	fluid Operations AG	민간	x
Junar	Junar	민간	x
Open Government Platform (OGPL)	미국 정부, 인도 정부	공공	o
Open Data Store	Open Data Store Inc.	민간	o
OpenDataSoft(ODS)	OpenDataSoft	민간	x
Socrata	Socrata	민간	x

30) 민간기업 Granicus(Granicus와 GovDelivery가 합병됨, 구 NuCivic)에서 DKAN의 개발 및 유지보수를 지원해 왔으며, 2017년 11월부터는 CivicActions에서 DKAN 프로젝트를 지원하기로 함

○ CKAN은 Open Knowledge에서 개발한 공공데이터 오픈소스 플랫폼으로, 영국의 오픈데이터 포털 data.gov.uk 및 EU의 publicdata.eu, 브라질의 dados.gov.br 등의 국가 및 국제 정부 포털 및 다양한 데이터 포털에서 사용되고 있는 데이터 연계 플랫폼



<그림 2-8> CKAN.org

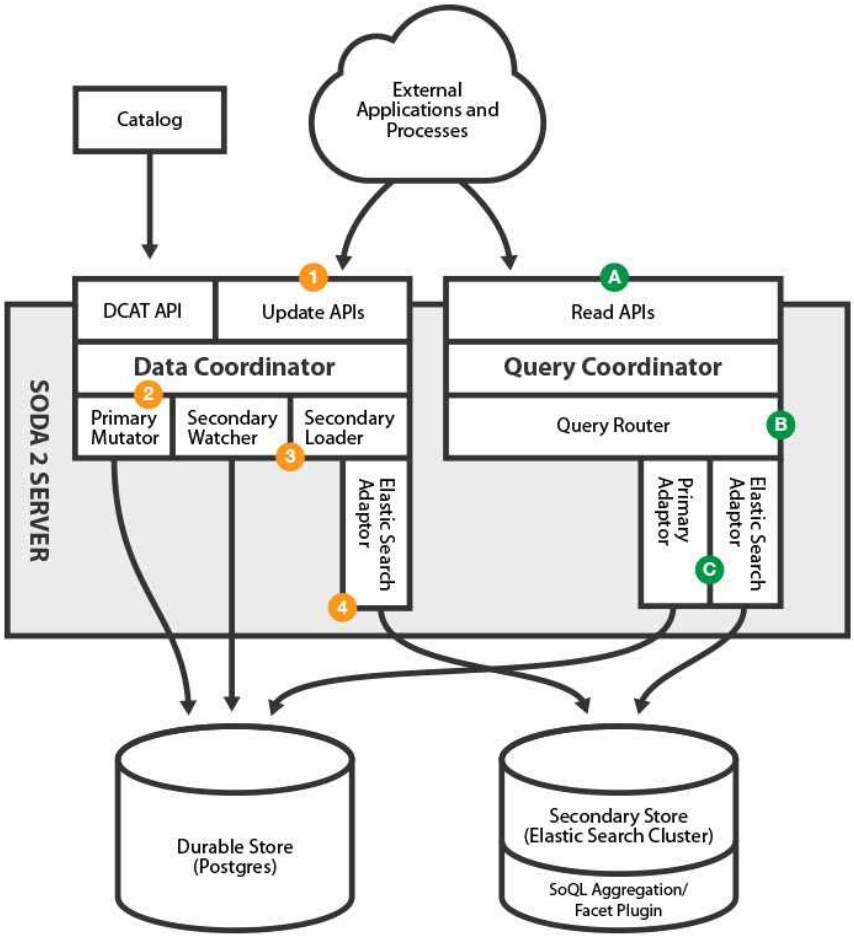
※ 출처: CKAN.org http://CKAN.org 참조 (접속일: 2017.08.21.).

○ DKAN은 Granicus[30]에서 개발, 지원하고 있는 CKAN과 Drupal 기반의 오픈소스 개방형 데이터 플랫폼임

- 카탈로그, publish, 시각화 전반에 걸친 기능을 가지고 있고, CKAN의 기능적 요소를 기반으로 하고 있기 때문에, DCAT, INSPIRE, CSV, JSON, XML, RDF를 포함하는 표준 파일 포맷과 API 구성을 지원
- CMS(콘텐츠 관리 프레임워크) 등 Drupal의 확장성 높은 기능 지원으로, 소셜미디어, 블로그 등 콘텐츠 관리 기능이 뛰어남

○ Socrata는 클라우드 기반의 상용 오픈 데이터 플랫폼 서비스로서 미국 정부 및 지방정부 기관에서 주로 사용

- 오픈소스 기반의 데이터 플랫폼에 비해 시각화, 분석 등에 강점이 있음

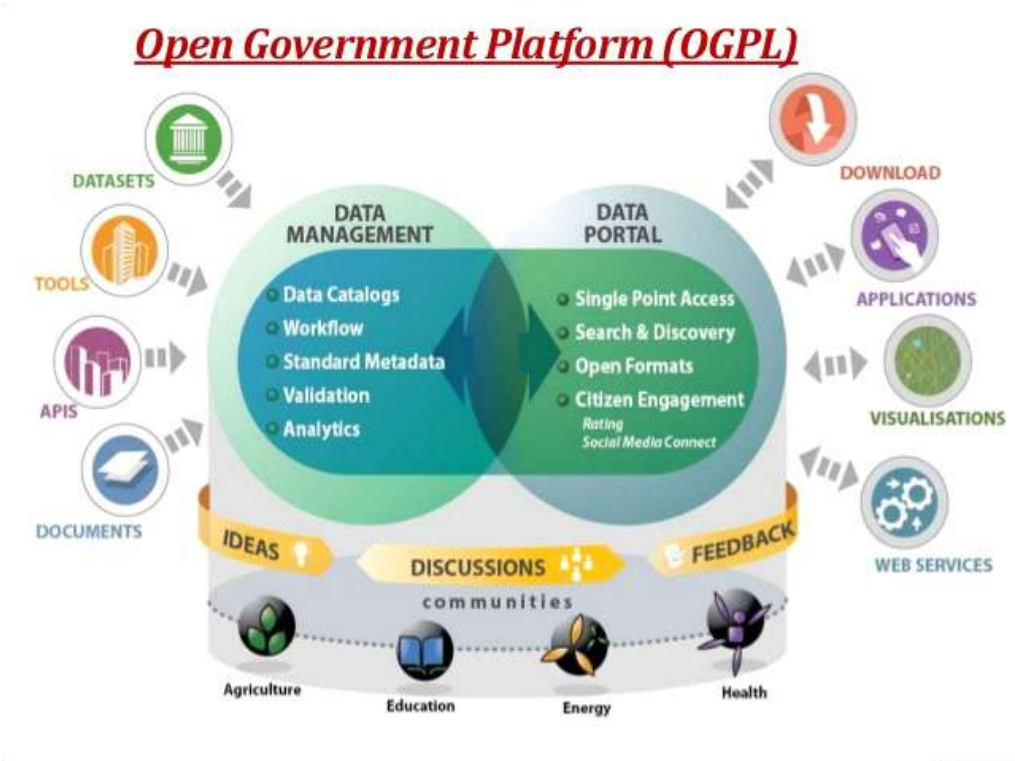


<그림 2-9> Socrata 시스템 아키텍처

※ 출처: Socrata System Architecture http://open-source.socrata.com/architecture 참조 (접속일: 2017.08.21.).

○ OGPL은 미국의 Data.gov 플랫폼을 오픈소스 버전으로 변환한 공개형 데이터 플랫폼[31]으로 현재 인도의 공공데이터 포털 Data Portal India(http://data.gov.in)는 OGPL 기반으로 구성되어 있음

- 2012년도 인도의 National Informatics Center와 미국의 Data.gov팀이 6개월간 협력하여 개발한 플랫폼
- 다양한 확장 기능의 활용이 가능한 Drupal 기반으로 개발된 플랫폼



<그림 2-10> Open Government Platform

※ 출처: Open Data Initiative of India
https://www.slideshare.net/DataPortalIndia/open-data-initiative-of-india 참조
(접속일: 2017.08.21.).

31) Nikhil Pahwa (2012.03.30.), "US &India Governments Launch Open Government Platform", MEDIANAMA, https://www.medianama.com/2012/03/223-open-government-platform/ 참조 (접속일: 2017.08.21.).

○ OpenDataSoft는 구조화된 데이터를 API 및 시각화 할 수 있는 상용(commercial) 오픈 데이터 플랫폼으로 클라우드 기반으로 데이터를 호스팅, 공유할 수 있도록 하며, turnkey 형식으로 판매하는 솔루션임

- 지도에 데이터 세트를 여러 레이어로 추가하여 레이어 맵을 만들고, 카탈로그에 제공하는 데이터 세트에 대하여 고급 시각화 차트 분석이 가능하며 시각화 기능이 뛰어남
- 포털의 Look&Feel(모양과 느낌)에 통일성을 주어 포털에 브랜드 통일성을 줄 수 있는 포털 브랜딩 기능을 가지고 있고, 반응형 포털 디자인을 제공
- 메타데이터 표준으로 DCAT의 하위집합 사용 및 INSPIRE 템플릿 사용 가능

○ 상용 소프트웨어 Socrata, OpenDataSoft, Junar의 경우 데이터 연계플랫폼으로서의 기능인 메타 데이터 관리, 검색, 소셜, 수집통합 카달로그, 확장성, 데이터 분석, 시각화, 사용자테마, 라이선스 기록에 대하여 모두 가지고 있음

○ 오픈소스 데이터 플랫폼인 CKAN, OGPL의 경우, 데이터 분석 기능 이외에 모든 기능을 가지고 있어 상용 데이터 플랫폼 못지않게 다양한 기능을 가지고 있음

<표 2-11> 데이터 플랫폼별 기능 비교

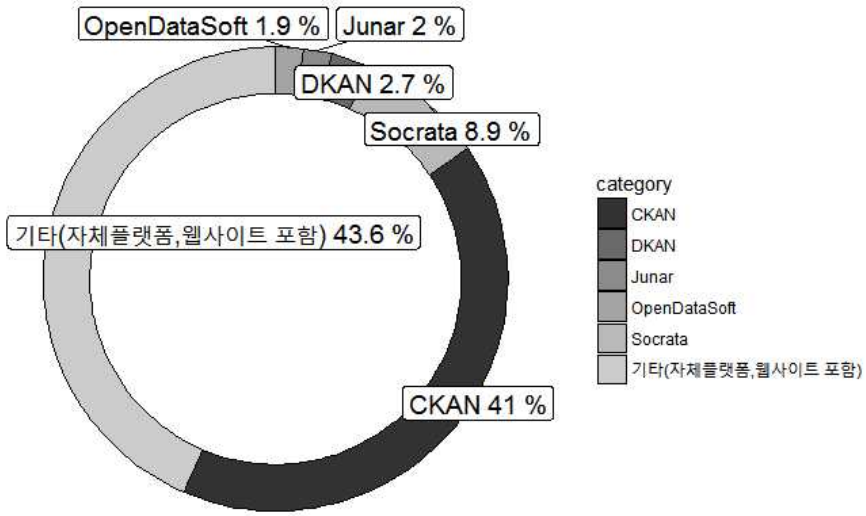
구분	플랫폼 이름	메타 데이터 관리	검색	소셜 기능	수집 통합 카달로그	확장성	데이터 분석	시각화	사용자 테마	라이센스 기록
오픈소스	CKAN	o	o	o	o	o	x	o	o	o
	DKAN	o	o	o	o	o	x	o	o	o
	OGPL	o	o	o	o	x	o	o	o	x
	esri GeoPortal	o	o	x	o	o	x	o	o	o
상용	ArcGIS Open Data	o	o	o	o	o	o	o	o	o
	Socrata	o	o	o	o	o	o	o	o	o
	OpenData Soft	o	o	o	o	o	o	o	o	o
	Junar	o	o	o	o	o	o	o	o	x
	Open Data Store	o	o	o	o	o	x	o	o	o
	Datatank	o	o	x	x	o	x	x	o	x
	Enigma Public	o	o	x	x	o	o	x	x	x

2. 국내외 데이터 플랫폼 적용 현황

○ 데이터 포털 현황 조사를 실시했던 2,883개의 국내외 데이터 포털 중, 지역적 분포 및 포털 규모를 고려해 1,673개를 선정, 이를 대상으로 데이터 플랫폼 적용 현황을 조사함

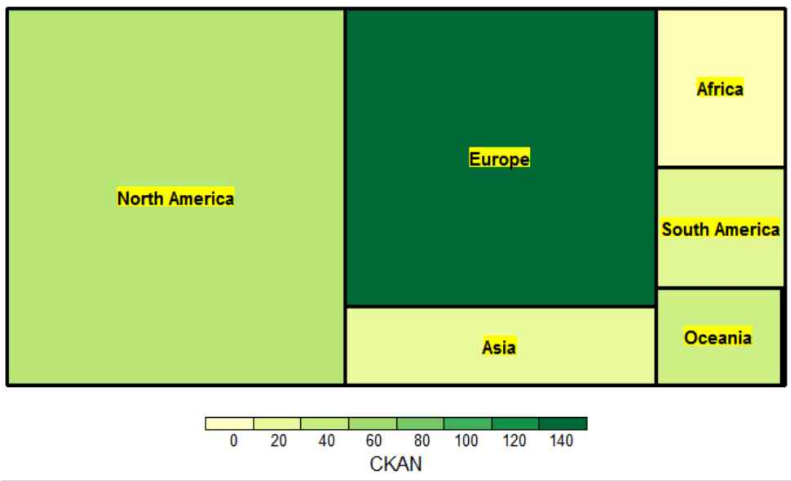
- 오픈소스 데이터 플랫폼인 CKAN, DKAN, OGPL과 민간에서 개발된 상용 데이터플랫폼인 Socrata, Junar, OpenDataSoft, 그리고 지리정보(GIS) 특수목적 플랫폼인 ArcGIS Open Data 등이 사용되고 있음
- 데이터 플랫폼으로는 CKAN(286건), 자체플랫폼(193건), Socrata(62건) 순으로 가장 많은 사용량을 보임
- 민간에서 개발된 상용 데이터 플랫폼이 DKAN(19건), Junar(14건), OpenDataSoft(13건)의 순으로 비슷한 비율(1%)로 사용되고 있음

- 그 외 기타 Open Government Data Initiative(OGDI), Open Government Platform (OGPL), knoema, Drupal 웹페이지, Stoplight, Trac, DataMarket, datashare, FLOSS 등의 데이터플랫폼이 소수 사용되고 있으며, 그 외에 지리정보 특수목적으로 사용된 플랫폼(975건)이 있음



<그림 2-11> 데이터 플랫폼 사용 현황

- 아래 트리맵 차트[32]에서 보이는 바와 같이 북아메리카 지역의 데이터 포털 수가 가장 많은 반면, 유럽 지역의 CKAN 적용비율이 가장 높음을 알 수 있음



<그림 2-12> 지역별 데이터 포털 비율(사각형 크기)과
CKAN 적용 비율(색깔)

32) 트리맵 차트 - 데이터 포털 수를 사각형 넓이로, 사각형 색깔의 진한 정도를 CKAN 사용비율로 나타냄

○ 지리정보 특수목적 플랫폼을 사용한 데이터 포털을 제외한, 데이터 포털 698개에 대한 데이터 플랫폼의 사용 현황[33]을 조사함

- 지역별로 보았을 때, 유럽의 경우 CKAN이 가장 높은 비율을 차지하고 있으며, 미국, 캐나다가 포함되는 북아메리카의 경우 CKAN과 Socrata의 비율이 비슷함
- 플랫폼별로 보았을 때 CKAN의 경우 유럽에서 강세를 보이고, Socrata의 경우 북아메리카에서 사용량이 많은 것으로 나타남
- 국가별로 보았을 때, 앞서 가장 많은 데이터 포털을 보유하고 있는 것으로 조사된 미국의 경우 Socrata(51개), 자체플랫폼(36개), CKAN(30개) 순으로 데이터플랫폼을 사용하고 있는 것으로 나타남

<표 2-12> 지역별 데이터 포털의 데이터 플랫폼 적용 현황 (데이터 포털 수)

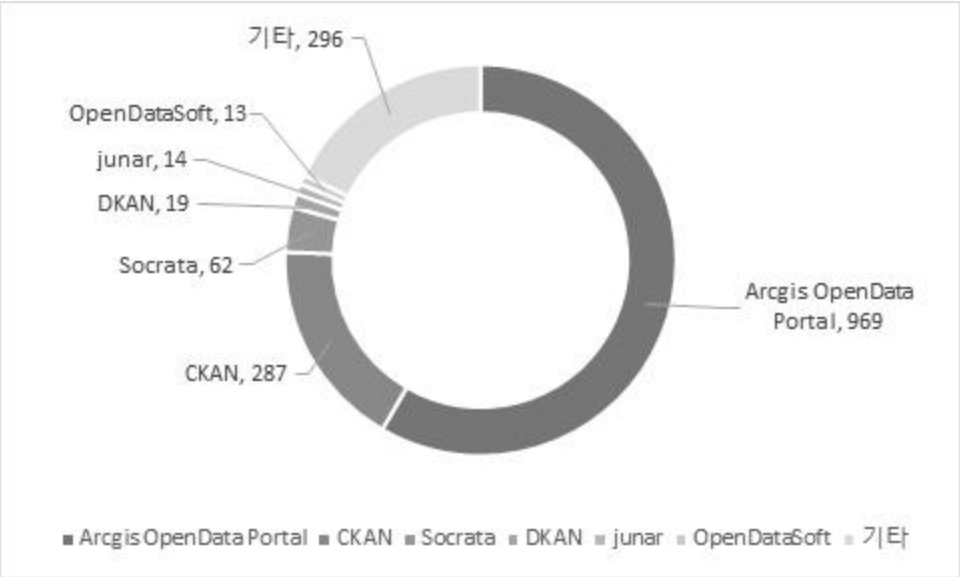
플랫폼 / 지역	CKAN	Socrata	DKAN	junar	OpenData Soft	기타 (자체플랫폼, 웹사이트 포함)	합계
유럽	140	4	11	-	10	115	280
북아메리카	50	53	7	9	2	99	220
아시아	21	-	-	-	-	49	70
오세아니아	39	3	-	-	-	8	50
남아메리카	27	1	-	5	1	8	42
아프리카	7	1	1	-	-	25	34
Worldwide	2	-	-	-	-	-	2
합계	286	62	19	14	13	304	698

33) 지리정보를 위한 플랫폼(ArcGIS 등)은 일반 데이터 플랫폼과는 구분되어 특수목적을 위해 사용되기 때문에 데이터 플랫폼 사용현황 집계 시 제외하고 <참고>로 별도 제시

<참고>

※ 지리정보 특수목적 플랫폼 ArcGIS

- 지리정보를 표시하기 위해 개발된 플랫폼이나 데이터 공유 목적의 툴인 ArcOpenData로 보편화되어 확장되었음
- ArcGIS Open Data가 적용된 데이터 공유 플랫폼이 포함할 경우 (967건)으로 제일 많았고, esri GeoPortal, GeoNetwork, GeoBlackLight, geobase 등이 소수 집계됨



<그림 2-13> 데이터 플랫폼 사용 현황 (특수목적 플랫폼 포함)

- 그러나 지리정보 표시를 위한 플랫폼으로 시작한 한계에 따라 전문가들이 일반적인 데이터 공유 플랫폼과의 차별성을 강조해야 한다는 의견에 따라 제외하고 분석

3. 비교 및 시사점

○ 국내외 데이터 포털에서 적용하고 있는 대표적인 데이터 플랫폼 솔루션들은 공통적으로 데이터 연계를 위해 메타 데이터의 검색과 활용이 가능한 데이터 카달로그를 제공하고 있음

○ 전세계적으로 지리정보 특수 목적 플랫폼을 사용한 데이터 포털을 제외하면, 비영리기관에서 오픈소스 기반으로 개발된 CKAN의 활용 사례가 가장 많음

○ 데이터 플랫폼 솔루션들은 개발 주체(민간/공공/비영리) 및 오픈 소스 활용 여부에 따라 나누어보면, 민간에서 개발한 경우 메타 데이터 관리, 검색, 소셜, 수집통합 카달로그, 확장성, 데이터 분석, 시각화, 사용자테마, 라이선스 기록 등의 기능을 모두 가지고 있는 경우가 많으며, 오픈 소스 기반의 솔루션의 경우 분석, 시각화 부분의 기능이 제한적으로 제공되거나 없는 경우가 있음

○ 선진국의 국가별 대표 오픈 데이터 포털에서는 데이터를 열람, 다운로드 할 수 있는 이용 권한을 대부분 제공함
- 데이터셋별로 다르게 라이센스를 적용하는 경우도 있음
- 국가 별로 오픈 데이터 라이센스 정책을 별도로 정하여 열람, 다운로드, 복사, 배포, 수정 등을 가능하도록 하는 등 데이터 연계와 활용을 위한 라이센스 정책을 도입하는 사례도 있음

제3절 오픈 데이터 관련 기구 및 기관 활동 현황

본 절에서는 국내외 오픈데이터 관련 기구 및 기관의 특징과 활동 현황을 살펴보고, 오픈 데이터 개방과 관련된 가이드라인 관련 활동을 살펴보고자 함

1. 해외 오픈 데이터 관련 기구 및 기관 현황

○ 지금까지 세계 각국의 공공 및 민간에서 데이터 플랫폼을 중심으로 살펴본 바에 따르면, 정부와 비영리 기관 그리고 민간에서 데이터 연계와 활용을 위한 활동이 활발하게 진행되고 있음

- 최근 활동이 두드러지고 있는 CKAN과 같은 오픈소스를 이용한 개방형 데이터 플랫폼을 활용하여 기업의 다양한 서비스와 연결하고자 하는 기업들의 노력을 볼 수 있음

○ 이러한 노력들은 정부나 기업 중심의 활동 뿐 아니라, 국제기구, 비영리재단, 협회 등이 주축이 된 오픈 데이터 관련 기구 및 기관의 통해서도 많은 활동이 이루어지고 산출물이 개발되어옴

- 오픈데이터 관련 기구 및 기관들은 다양한 활동을 통해 정부와 민간 부문의 파트너 역할, 정부기관 데이터를 사용한 비즈니스 창출, 지역사회의 공공데이터 개방 참여, 정부에 의한 오픈데이터의 활동의 홍보, 정부의 공공지출 정보 제공, 의회 활동 감시, 개방형 데이터 솔루션 및 도구 개발, 시민이 만드는 데이터에 대한 정보 공유의 역할을 해옴
- 이들의 주요 산출물로는 오픈 데이터 라이선스 개발, 가이드, 방법론 및 보고서 개발, 사례연구, 정책문서, 무료 공개 데이터 오픈, 데이터 솔루션 및 도구가 있음

가. 오픈데이터 관련 기구 및 기관

○ 미국의 대표적인 오픈데이터 관련 단체로는 USAspending.gov, U.S.OpenData, OpenData500 등이 있음

- USAspending.gov[34]은 2006년 연방 기금책임 및 투명성 법에 따라 공개적으로 접근 가능하고 검색 가능한 웹사이트 형태로 2007년 설립되었으며, 세금 공제 방법에 관한 정보를 미국 대중에게 제공함
- U.S.OpenData는 2013년 설립되어 정부의 공개데이터 생태계를 위해 정부와 민간 부문의 파트너 역할을 하며, 재현 가능한 코드를 개발하고 시민데이터 표준을 만드는 등의 활동을 해옴
- OpenData500은 GovLab에서 2013년 시작한 프로젝트로서 정부기관 데이터를 사용한 비즈니스 창출과 제품 서비스 개발을 위한 연구 및 데이터를 공개

○ 영국의 대표적인 오픈데이터 관련 단체로는 The public whip, Open Knowledge international, Open Rights Group, mySociety, Where Does My Money Go, Free our data, Open Data Institute 등이 있음

- The public whip은 2003년 설립되었으며, 영국 하원 의원들의 1997년부터 투표 결과를 공식 웹페이지에서 뽑아내어 구조화된 데이터로 공개하고 있음

34) USAspending.gov, http://www.usaspending.gov 참조 (접속일: 2017.08.17.).

- Open knowledge international[35]은 2004년 루퍼스 폴락을 중심으로 영국에서 출발한 Open Knowledge International은 비영리 조직으로서, CKAN 개발과 관련 활동을 하고 있음
- mySociety[36]는 영국을 기반으로 하지만, 국제적으로 파트너와 협력하는 비영리 사회적 기업으로 민주주의, 정보의 자유와 더 나은 도시를 만들기 위한 디지털 기술을 구축하고 공유하며 웹 도구와 앱을 제작하고 기술을 지원함
- Where Does My Money Go[37]는 2006년 설립된 비영리 단체로 공공 지출 정보 분석, 시각화로 투명성과 시민참여를 증진시킴
- Free our data[38]는 2006년에 설립되었으며 납세자 데이터를 활용할 수 있도록 함
- Open Data Institute(ODI)는 웹을 발명한 팀 버너스리와 인공지능 전문가 Sir Nigel Shabolt가 2012년도 공동 창립한 오픈 데이터 연구소로서 런던에 위치하며, 정부 정책을 지원하고, 신생 기업을 지원, 장려하며, 훈련, 연구, 혁신 등의 리더십 제공하고 전략 개발을 지원함

○ 캐나다의 대표적인 오픈데이터 관련 단체로는 OpenNorth, OpenDataBC 등이 있음
- OpenNorth는 2011년 설립된 비영리 기관으로서 정부 투명성 확대와 공공데이터 확대를 위한 활동을 하며, 다양한 데이터 개방 솔루션과 시민을 위한 도구를 개발
- OpenNorth의 설립자이자 캐나다 공공데이터 운동의 선두주자인 제임스 매키니[39]는 기업 정보의 투명성 확대를 위하여 전세계 기업의 정보를 모아 기업에 대한 접근성과 투명성을 높이기 위한 단체인 “Open Corporate”에서의 활동 계획을 가지고 있음
- OpenDataBC는 2012년 설립된 단체로 공공데이터를 오픈하는 활동 및 해커톤[40]을 주최함

35) Open Knowledge International, https://okfn.org/about/ 참조 (접속일: 2017.08.17.).
36) 박주성 (2011.11.14.), “Open Government 2.0 – Case Study: Australia & United Kingdom”, 발표자료.
37) Where Does My Money Go?, http://www.wheredoesmymoneygo.org/ 참조 (접속일: 2017.08.17.).
38) Free our data, http://www.freeourdata.org.uk/ 참조 (접속일: 2017.08.17.).
39) 한겨레 21 (2016.03.22.), “와글이 만난 몽상가들- 캐나다 ‘오픈노스’ 설립자 제임스 매키니”, http://h21.hani.co.kr/arti/society/society_general/41380.html 참조 (접속일: 2017.09.13.).
40) 해커톤(hackathon)은 해커와 마라톤의 합성어로서, 기획부터 프로그래밍까지 단기간에 거쳐 마라톤을 하듯 24~48시간 내외로 쉬지 않고 집중적으로 작업을 하는 이벤트

○ 유럽 및 전세계의 대표적인 오픈데이터 관련 단체로는 OpenData France, ROUTE-TO-PA Project, DATA SHIFT 등이 있음

- OpenData France는 2013년 프랑스에서 설립된 단체로, 공공데이터 개방에 참여하는 커뮤니티가 오픈데이터 홍보 목적으로 모든 단계를 장려하고 있음
- ROUTE-TO-PA Project(유럽)는 Open Knowledge International에서 후원하고 있는 프로젝트로 CKAN에 연계하여 활용할 수 있는 분석 및 시각화 TOOL인 TET 및 데이터 연계 솔루션을 개발함
- DATA SHIFT (Global 연합, 2015년 설립)[41]는 시민이 생성한 데이터를 생산하고 사용하여 시민 사회 조직의 역량과 신뢰를 구축하는 단체로 아르헨티나, 네팔, 케냐 및 탄자니아에서 시민이 생성한 데이터를 생산하고 사용하는 시민 조직 및 단체를 지원함

41) DATA SHIFT, http://civicus.org/thedatashift/ 참조 (접속일: 2017.08.27.).

<표 2-13> 오픈데이터 관련 기관 및 기구

No	기구, 기관명	구분	설립 연도	국가	주요 역할	산출물	주요멤버
1	Sunlight foundation	비영리	2006	미국	정부 정보의 투명성	법 개정 활동 및 보고서 국회 및 입법기관 모니터링 OpeningParliament.org	마이클 클라인 엘렌 S. 밀러,
2	DATA SHIFT	시민 단체	2015	전세계	시민이 창출한 데이터 이니셔티브 역량과 효율성을 높임	학습자료 제공 사례, 연구보고서	잭 콘 포스, 캐시아 모라에 니콜 비비
3	U.S.OpenData	비영리	2013	미국	정부의 공개데이터 생태계를 위해 정부와 민간 부문의 파트너 역할	재현 가능한 Code 개발, 시민데이터표준	Waldo Jaquith
4	OpenData500	비영리	2013	미국	정부기관 데이터를 사용한 비즈니스 창출과 제품 서비스 개발을 위한 연구 및 데이터 공개	무료 공개 데이터 오픈	베스 시몬 라벡, 슈테판 베를루스트
5	OpenData France	공공	2013	프랑스	지역사회의 공공데이터 개방에 참여, 오픈데이터 홍보를 위한 정부 활동 장려	공공 데이터 오픈	버트랜드 SERP, 프랑소아즈 브루네토, 장 다니엘 브루소, 등
6	OpenDataInstitute (ODI)	비영리	2012	영국	리더십 및 전략개발 지원, 연구 및 혁신, 언어개발 및 정책수립, 데이터훈련제공, 신생기업 지원, 글로벌 네트워크 만들기, 정부사업추진	가이드, 방법론 및 보고서, 사례연구, 정책문서, 연구논문, 토론, 백서 등	팀 버너스리, 나이젤 샤드볼트
7	OpenDataBC	비영리	2012	캐나다	British Columbia의 오픈 데이터 커뮤니티	해커톤, 공공데이터 오픈	
8	Open Contracting Partnership	비영리	2012	미국	국제 계약과 관련된 데이터 표준의 작성과 활동	공개 계약 데이터 표준	Gavin Hayman Kathrin Frauscher Lindsey Marchessault
9	WHERE DOES MY MONEY GO?	비영리	2007	영국	공공 지출에 대한 정보를 분석하고 시각화함으로써 투명성과 시민 참여를 증진	세금 납부 및 지출 정보	Jonathan Gray
10	USAspending.gov	비영리	2007	미국	세금 공제 방법에 관한 정보를 미국 국민에게 제공	세금 관련 정보 및 Tool	
11	Open Knowledge International	비영리	2004	영국	데이터 개방을 통해 데이터에 액세스하고 활용할 수 있도록 지원. 공개데이터의 가치를 알리고 도구와 기술을 제공함	CKAN, Open Data Commons, OpenSpending, Open knowledge labs, 오픈데이터핸드북, 데이터 저널리즘 핸드북	루퍼스 폴락, 마틴 키건, 조 월시
12	mySociety	비영리	2003	영국	영국을 기반으로 하지만 국제적으로 파트너와 협력하는 비영리 사회적 기업. 민주주의 , 정보의 자유 , 더 나은 도시 의 세 가지 영역에서 변화를 가져올 수 있는 힘을 주는 디지털 기술을 구축하고 공유함	- FixMyStreet -여행시간을 보여주는지도 플랫폼기반의 언론인과 캠페인를 위한 정보 도구 등	Mark Cridge, James Cronin, Jonathan Flowers
13	OpenNorth	비영리	2011	캐나다	공공의 참여와 정부 투명성을 높이는 개방적 민주주의를 위해 지역 및 국가 차원에서 다양한 개방형 데이터 솔루션과 시민 기술 도구를 제공	The Gazette, Gazette Docs, My City Hall 시민 예산	James McKinney Jean-Noé Landry Alyssa Beaton Rachel Bloom

나. 오픈 데이터 관련 기구들의 활동

○ 오픈 데이터 관련 기구 및 기관들의 활동을 통해, 개방형 데이터 솔루션 및 도구의 개발, 데이터 표준, 가이드, 방법론 개발 및 보고서 등의 산출물과 정부 공공정보 제공 및 감시, 정부와 민간의 파트너역할 및 비즈니스 창출 등의 효과를 얻고 있음

○ 개방형 데이터 솔루션 및 도구 개발과 관련된 활동을 통해서 CKAN 개발 및 연계 활용 도구 개발, extension 개발 등의 활동을 진행
- Open knowledge international[42] 은 CKAN 개발과 관련 활동을 하고 있음
- ROUTE-TO-PA Project를 통해 CKAN에 연계하여 활용할 수 있는 분석, 시각화 TOOL인 TET를 개발하고 있음
- OpenNorth는 CKAN-automator Extension 개발

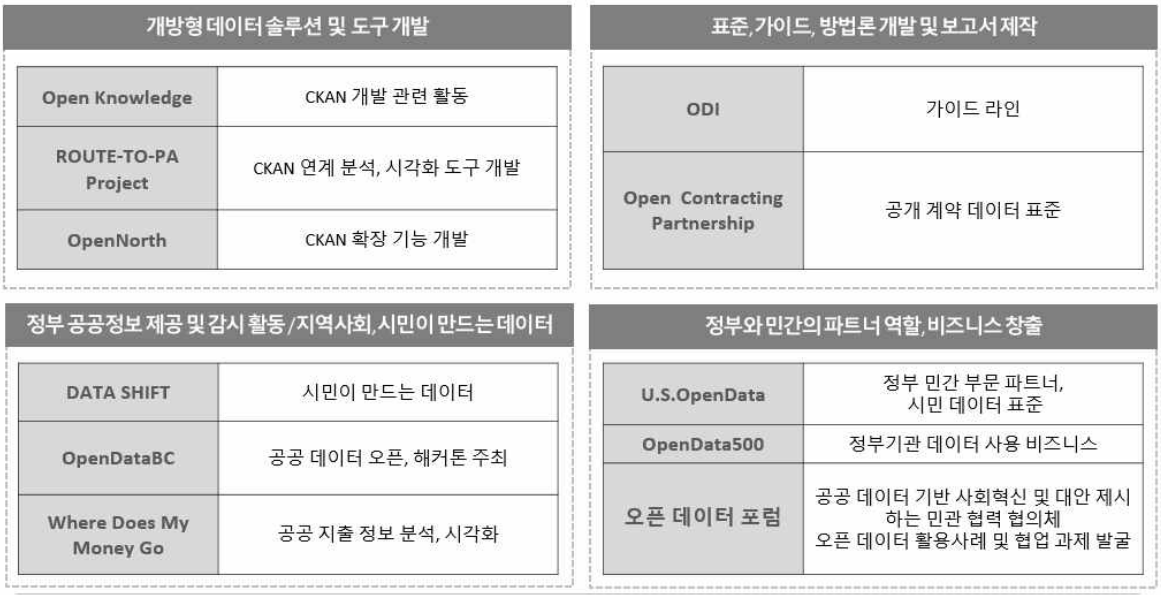
○ Open Data Institute(ODI), Open Contraction Partnership은 오픈 데이터 관련하여 가이드 라인의 제작 및 배포를 하고 있음

○ 정부 공공정보 제공 및 감시와 관련된 활동 부문에서는, 2006~2007년부터 정부의 공공정보를 공개하여 데이터로 활용하고자 하는 움직임이 활발해짐
- DATA SHIFT[43]는 시민이 만드는 데이터(Citizen Generated Data, CGD)를 생산하고 활용하는 활동을 함
- OpenDataBC는 공공데이터를 오픈하는 활동 및 해커톤을 주최
- Where Does My Money Go[44]는 공공 지출 정보 분석과 시각화로 투명성과 시민참여 증진함
- USAspending.gov[45]는 세금 공제 방법을 미국 국민에게 제공함
- Free our data[46]는 납세자 데이터를 활용함

○ 정부와 민간의 파트너 역할 및 비즈니스 창출하는 활동을 함
- U.S.OpenData는 정부와 민간 부문의 파트너 역할을 하며, 재현 가능한 코드를 개발하고 시민데이터 표준을 만드는 등의 활동 수행
- OpenData500는 정부기관 데이터를 사용한 비즈니스 창출 역할을 함

---

42) Open Knowledge International, https://okfn.org/about/ 참조 (접속일: 2017.08.17.).
43) DATA SHIFT, http://civicus.org/thedatashift/ 참조 (접속일: 2017.08.17.).
44) Where Does My Money Go?, http://www.wheredoesmymoneygo.org/ 참조 (접속일: 2017.08.17.).
45) USAspending.gov, http://www.usaspending.gov 참조 (접속일: 2017.08.17.).
46) Free our data, http://www.freeourdata.org.uk 참조 (접속일: 2017.08.17.).



<그림 2-14 오픈 데이터 관련 기구들의 활동>

다. 데이터 연계 가이드라인 사례

○ 데이터 연계 및 개방에 관련된 가이드라인으로서, 해외에서는 비영리에서 추진 한 데이터 소유자의 개방 가이드라인이 다수와, 여러 국가에서 참여한 공공에서 추진 한 공공데이터 개방을 위한 가이드라인이 존재

- 국내에는 공공 주도의 공공데이터 개방 표준이 있음. 그 이에도 개인 정보에 대한 법률 및 가이드라인이 있으며, 내년 시행 예정인 유럽 개인정보보호법(GDPR)의 영향이 주목받고 있음

○ Open Data Commons (Making Your Data Open: A Guide)[47]는 데이터를 개방하는 방법, 데이터에 적합한 라이선스를 부여하는 방법에 대해 가이드하고 있으며 공개 라이선스로 Public Domain(PDDL[48]), Share-Alike (Open Database License[49])를 제안

47) Open Data Commons, “Making Your Data Open: A Guide”, https://opendatacommons.org/guide/ 참조 (접속일: 2017.09.12.).
48) Open Data Commons, “Public Domain Dedication and License(PDDL)”, https://opendatacommons.org/licenses/pddl/ 참조 (접속일: 2017.09.12.).
49) Open Data Commons, “Open Database License”, https://opendatacommons.org/licenses/odbl/ 참조 (접속일: 2017.09.12.).

<그림 2-15> 국내외 데이터 연계 가이드라인 사례

○ OPEN DATA HANDBOOK에서는 데이터 소유자가 데이터를 공개하는 방법에 대한 가이드를 목적으로 하며 기초적인 부분에 대한 검토와 문제점을 살펴보고, 발생할 수 있는 이슈에 대하여 논의하고 있음

- 데이터를 공개하는데 있어서 세 가지 규칙(① 단순명료하게 공개 ② 신속하게 자주 공개 ③ 우려와 오해에 대하여 충분히 설명), 데이터를 공개하는 4단계(① 데이터셋 선택함, ② 오픈 라이선스 적용함, ③ 데이터를 이용 가능하도록 함, ④ 데이터를 탐색 가능하도록 함) 등 제시
- 데이터셋의 선택, 커뮤니티 활용, 비용 근거, 공개의 용이성에 대해서 가이드 중
- 오픈 라이선스에 대한 지침과 데이터를 활용, 검색이 가능하기 위한 기술적, 법적 지침을 소개

○ W3C guidelines for Open Data[50]은 비영리 조직인 W3C에서 정부가 데이터를 열어 공유할 수 있도록 개발한 지침

- 정부 데이터 공개를 위한 절차, 공개데이터의 유용성(식별자, 문서화, 연결성, 보존, 인터페이스, 객체에 대한 표준 식별자, URI), 게시할 데이터 선택, 데이터에 적합한 형식, 법적 및 규제상의 제한사항을 문서화함

50) W3C (2009), "W3C guidelines for Open Data", https://www.w3.org/TR/gov-data/ 참조 (접속일: 2017.08.03.).

○ 국제 데이터 오픈 헌장(Open Data Charter) [51]은 공공데이터 개방에 대한 정부와 공공기관이 지켜야할 여섯 가지 원칙(① 기본적으로 공개, ② 시의적절하고 포괄적, 접근가능하며 사용가능, ③ 비교할 수 있고 상호연동 가능, ⑤ 거버넌스와 시민의 참여증진, ⑥ 포괄적인 개발과 혁신을 위한 공개)에 대해서 정의하고 있음

○ 부패 방지를 위한 공개 데이터 사용을 위한 개방 가이드는 사용사례 및 방법론, 우선순위 데이터 세트 및 핵심속성, 데이터 표준을 위한 기술적 세부사항. 부패 방지 활동을 위한 표준이 포함됨

○ 영국 Digital Economy Act 2017[52]는 전자통신 인프라 및 서비스와 관련된 정책, 저작권 침해 처벌 강화, 연구목적의 데이터 공유 등의 내용을 담고 있는 영국 디지털 경제 법안 2017 (Digital Economy Act 2017)을 국왕의 승인을 받아 법제화함

- 디지털 서비스에 대한 접근, 디지털 인프라 구축, 온라인상의 음란물 및 불법 사이트로 부터 아동보호를 위한 조치, 온라인 저작권 강화, 디지털 정부 등 6개 파트로 구성
- 정부 부처가 보유한 개인 정보를 조사 목적의 익명화된 데이터셋으로 변환하는 방법을 제시

2. 국내 관련 기구 현황

가. 오픈데이터 관련 기구 및 기관

○ 국가 오픈데이터 포럼 (2017)은 정부(행정자치부), 공공(한국정보화진흥원), 시민, 기업, 학계가 협력하여 공공데이터 기반의 사회 혁신에 참여하고 대안을 제시하는 목표를 가진 민관 협력 데이터 활용 소통 협의체

- 오픈데이터 활용사례 발굴, 민간의 공공데이터 활용 수요 파악 및 협업 과제 발굴, 관련 지식 공유를 통한 공공 데이터 개방 정책의 실효성 제고
- 운영방향은 데이터 정책을 국민과 함께 공동 생산하는 열린 정부, 이용자 간 교류를 통한 공공데이터 활용 비즈니스 창출 및 확대, 공공데이터 기반 시빅 테크 활성화를 통한 사회혁신 기반 조성

---

51) Open Data Charter (2015), “Towards an International Open Data Charter”, http://opendatacharter.net/history/ 참조 (접속일: 2017.08.02.).

52) 한국정보화진흥원 (2017), “오픈데이터 해외 동향 : Digital Economy Act 2017”, 「Global Open Data, Now 오픈데이터 해외주요동향」, 2017(2):5-6, 대구: 한국정보화진흥원.

<그림 2-16> 오픈 데이터 포럼 운영 방향

※ 출처: 오픈 데이터 포럼, http://odkf.kr 참조
(접속일: 2017.08.21.).

나. 오픈데이터 관련 활동

○ 행정자치부와 한국정보화진흥원은 국내의 공공데이터 개방 표준[53]으로 공공데이터 제공 및 이용 활성화와 공공 기관에서 데이터 개방 시 공통으로 적용할 표준 기준을 마련하여 민간 활용을 촉진을 목적으로 하며 그 범위는 행정 기관과 공공기관으로 하고 있음

- 공공데이터 개방 시 적용할 공통 개방 표준으로 표준용어, 데이터 제공방식(다운로드 방식, 오픈 API), 제공 경로, 공통항목 개방기준과 공공데이터 개방 시 적용할 데이터셋 분야별 개방 기준(제공항목, 항목별 속성정보, 제공 형식 등) 및 기타 데이터 개방 표준을 정의하고 있음

○ Digital-5 헌장은 대한민국 행정자치부, 영국 내각사무처, 에스토니아 경제통신부, 뉴질랜드 내무부, 이스라엘 국무총리실이 전자정부 혁신과 국가 간 디지털 격차 해소를 위한 'Digital-5 부산선언문'을 공동으로 선언한 헌장으로 9대 원칙[54] 중 공개표준과 오픈소스 기반, 연결성에 대한 원칙을 이야기 하고 있음

53) 행정자치부 (2016), 「공공데이터 개방 표준」(행정자치부고시 제 2016-46호) 2014년도 10월 제정, 2016년도 12월 5차 개정, 행정자치부.
54) Digital-5 부산선언문 9대 원칙: 시민을 위한 공공 서비스 디자인, 공개 표준 정책 필요, 정책의 회원국 간 공유, 공정한 경쟁, 개방형 라이센스 활용, 온라인 이용자 활성화, 차세대 기술인재 양성, 시민의 디지털 서비스 접근 지원, 회원국 간 협력

3. 비교 및 시사점

○ 미국, 영국, 유럽을 포함한 국내외 관련 기구 및 기구들의 활동을 비교해 본 결과, 국내는 관 주도에서 민간 주도로 바뀌어 가는 과도기에 있으나 해외의 경우 설립 초반부터 민-관 협력 등에 의한 비영리법인이 설립하는 경향이 보임

○ 해외 오픈데이터 연계 가이드라인의 경우, 공개된 데이터의 이용, 활용 가능하게 하기 위한 기술적, 법적 고려사항과 활성화를 위한 다양한 솔루션(커뮤니티, 시민활동, 정부의 역할 등)을 제공

○ 국내 공공데이터 개방 표준의 경우 정부의 공공데이터를 개방하는데 필요한 용어 표준 및 항목에 대한 표준까지 구체적으로 다루고 있으며, 데이터를 개방하는 쪽에서의 표준에 비중을 두고 있음

제3장 데이터 연계 도구 CKAN

제1절 CKAN의 의의와 발전 과정

1. CKAN의 개요

○ CKAN은 Open Knowledge에서 개발한 공공데이터 오픈소스 플랫폼으로, 영국의 오픈데이터 포털 data.gov.uk 및 EU의 publicdata.eu, 브라질의 dados 등의 국가 및 국제 정부 포털 및 다양한 데이터 포털에서 사용되고 있는 데이터 연계 플랫폼

2. CKAN의 주요 기능

○ CKAN은 데이터의 배포, 저장, 시각화, 연계, 검색 및 활용, 기능 확장, 보안 및 인증, 이력 및 사용 통계, 커뮤니케이션에 대한 주요기능을 지원함
- 데이터셋을 쉽게 배포하고 데이터에 대해 설명하고 있는 메타 데이터의 저장과 업데이트가 용이함
- 데이터스토어(DataStore), 파일 스토어(FileStore), 링크드 데이터와 RDF, 카달로그를 저장할 수 있도록 지원
- CKAN 기본 기능인 미리보기(리소스뷰)와 시각화, 사용자 테마 등을 이용한 기본적인 시각화 기능을 가지고 있음
- API연계, 데이터 통합(federation), 메타데이터 연계를 통한 연계 기능을 제공
- 오픈 소스 검색 도구인 Solr를 이용한 메타데이터 검색 및 활용 기능 제공
- 파이썬(Python) 패키지 형태로 기능 확장(extension) 개발이 가능하여 기능의 추가 개발이 용이
- 확장 기능(Extension)을 통한 사용자 인증 기능 지원
- 사용자 활동 이력 및 데이터셋에 대한 사용 통계 정보 제공

○ 데이터를 올리거나, 웹 인터페이스를 통해 게시하는 기능이 타 플랫폼 대비 편리하게 수행이 가능

○ 다양한 데이터에 대해 메타데이터의 관리가 편리하다는 이점과 상호 연관되는 테이블, 그래프, 맵을 통해 사용자에게 익숙한 형태의 미리보기를 제공

○ CKAN의 경우 또 다른 CKAN 네트워크와 연합이 가능하여 사용자 참여가 용이한 기능과 커뮤니티를 통한 의견을 공유할 수 있는 기능을 제공

○ CKAN의 확장기능(extension)을 통해, 커스터마이징 및 확장성에 있어서도 CKAN은 뛰어난 기능을 제공하며, API 제공 또는 오픈소스 라이선스 정책을 통한 무료 다운로드 및 실행이 가능함

<표 3-1> CKAN 주요 기능

CKAN 주요기능	설 명
배포	데이터 셋을 쉽게 배포하고 메타데이터의 저장과 업데이트가 용이함
저장	데이터스토어 , 파일스토어 , 링크드 데이터와 RDF, 카달로그 저장
시각화	미리보기 (리소스뷰), 시각화 , 사용자 테마
연계	API 연계 , FEDERATION(통합), 메타데이터 연계
검색 및 활용	Solr를 이용한 검색
기능 확장	파이썬(Python) 패키지 형태로 기능 추가 개발 가능
보안 및 인증	데이터 보안 및 사용자 인증 지원
이력 및 사용 통계	사용자 활동 이력 및 데이터셋에 대한 사용 통계
커뮤니케이션	커뮤니티 기능 및 소셜 기능

○ 데이터 연계 관점에서 CKAN 기능을 보면, 데이터 공급자 측면의 조직 생성관리, 데이터셋과 리소스 추가, 삭제, 편집 등 관리 기능과 데이터 수요자 측면의 검색 및 탐색 기능을 제공하고 있음

<표 3-2> 데이터 연계 관점에서의 CKAN의 주요 사용 기능

데이터연계 관점	주요 사용 기능
데이터 공급자	조직 생성 및 관리 데이터셋 추가 삭제 편집 리소스 추가 석제 편집
데이터 수요자	검색 탐색

3. CKAN의 특·장점

○ CKAN과 타 데이터 연계 플랫폼과 비교를 하면 데이터 관리, 운영, 라이선싱 등 데이터 연계 유통에서 타 플랫폼 대비 우수

- 타 플랫폼 대비 분석 및 시각화에 대한 기능은 뚜렷한 장점을 보이지 않으나, 본연의 목적인 데이터 연계 활성화에 대한 강점을 가지고 있음
- 분석 및 시각화의 경우 Extension 및 타 플랫폼과의 연계를 통해 보강할 수 있음 (미국의 data.gov 와 datausa.io 의 사례)
- Drupal과 같은 타 오픈소스 플랫폼과 결합을 통한 시각화 기능 강화
- 다양한 목적에 따른 Plug-in 기능(Extension) 등을 통해 분석 기능 보완
- 대표 사례로 트래픽이나 데이터 셋 다운로드 분석을 위한 Google Analytics와의 Plug-in 형태 지원이 있음

○ CKAN은 특정 플랫폼에 종속되지 않으면서도 데이터를 효율적으로 연계・검색・활용이 가능하여 데이터 연계 활용 플랫폼으로서의 가장 중요한 장점을 보유

○ CKAN 개발 활용측면에서 아키텍처 관점과 메타 데이터 관점으로 나누어 정리할 수 있음

- 아키텍처 관점은 배포 방법, 저장소의 위치, 소스에 대한 비용, 콘텐츠의 버전관리로 나누어 확인 가능
- 손쉬운 패키지 인스톨 방식의 배포와 오픈소스 그리고 콘텐츠에 대한 버전 지원은 데이터 연계 활용을 위해 필수적인 요소임
- 커스터마이징 수준은 CKAN 기반의 데이터 연계 플랫폼 활성화를 고려할 때 필요한 요소
- 메타 데이터의 관점은 메타 데이터 제공 정보에 대한 항목과 메타 데이터의 스키마 측면으로 나누어 확인 가능
- 유연한 메타 항목과 스키마는 다양한 데이터 오픈 플랫폼과 연계하며, 데이터 소비자의 편의성을 높이기 위해 필요한 항목임
- CKAN 메타 기능은 데이터 연계에 유연성을 위해 DCAT 연계를 통해 확장 가능하며, CKAN의 DataStore 적용 기능에서 추가되어야 하는 부분임

〈표 3-3〉 CKAN 개발 활용 측면의 특・장점

분류	항목	내용
아키텍처	배포 방법	패키지 인스톨
	저장소	Local / 원격 모두 지원
	소스	오픈소스
	커스터마이징 수준	사용자 수준에 따라 커스터마이징 가능 (Fully Customizing 지원)
	버전 관리	콘텐츠 버전 지원
메타 데이터	메타 항목	확장을 통해 다양한 항목 지원
	스키마	유연한 스키마 지원

〈표 3-4〉 타 플랫폼과 비교를 통한 CKAN 적용의 타당성

구분	소스 공개	특징	사용
CKAN	오픈 소스	- 전세계 개발자들이 CKAN 프로그램 개발을 자발적으로 지원 - 가장 활성화된 데이터 연계 플랫폼 - 콘텐츠 관리, 시각화, API 추출 등의 특화된 기능을 Drupal과 같은 타 오픈 소스와 결합	영국, 미국, 캐나다 등 40개 이상 국가 활용 중
OGPL	오픈 소스	- 미국과 인도 정부에서 공동으로 개발하고 있는 데이터 연계 플랫폼 - 공공기관 투명성을 목적으로 개발	인도 정부 부처
Socrata	상용	- 시각화 및 분석 등에 있어서 오픈소스 플랫폼에 비해 장점이 존재	미국 연방정부 및 10여개 주정부의 데이터 포털에 사용
Junar	상용	- 클라우드 기반 데이터 연계 플랫폼 서비스	세크라멘토, 팔로알토 등 도시 중심으로 사용 중

4. CKAN의 기술 발전 동향

○ CKAN의 버전별 역사[55]를 살펴보면 CKAN은 2004년 설립된 비영리 기관인 Open Knowledge의 주최로 2006년에 개발을 시작해 2007년에 첫 공식 배포판을 선보임

- 개인보다는 기관을 대상으로 적용되었으며 세계의 중앙정부, 지방정부에서 인기를 끌기 시작, 2010년에 영국 정부에 의해 http://data.gov.uk에 도입되면서 CKAN 개발이 급성장
- 처음 40개의 데이터셋으로 시작했던 http://CKAN.net 은 1년 안에 200개, 2013년도에는 6400개의 데이터셋으로 증가하였으며, 2011년에는 Data Hub라는 새로운 브랜드로 태어남
- CKAN은 1.0이 출시된 지 3년 만에 2.0을 출시하였으며 2006년 이후 2.5개월에 한 번 씩 꾸준히 새로운 버전을 출시
- 처음에는 단순히 오픈 데이터를 찾는 카탈로그에서 출발했으나 현재는 live, active 데이터를 위한 저장소와 데이터스토어로서의 역할로 확장
- 1.7.1 버전에서는 관리(management)보다는 배포(publish)에 초점이 맞추어짐

<그림 3-1> CKAN 기술 발전 동향

※ 출처: CKAN history, https://delftswa.gitbooks.io/desosa2016/content/CKAN/chapter.html의 그림을 일부 인용하여 재구성함

○ CKAN은 기능별로 보안, 인증, 이력, 미리보기, 시각화, API, 링크드 데이터 지원, 메타데이터 지원, 파일스토어(FileStore) 및 데이터스토어(DataStore) 지원, 데이터베이스 및 검색의 방면에서 발전해옴

55) Winn, Joss (2013), "Open Data and the Academy: An Evaluation of CKAN for Research Data Management", IASSIST 2013, 5-7.의 내용을 인용하여 재구성

○ 연구 데이터 관리(RDM)를 위해 CKAN을 적용했던 사례 보고서[56] 에 따르면, 1.7 버전에서는 공공데이터의 publish에 초점이 맞추어져 있어, security model 에 있어서는 지나치게 단순한 구조로 되어있었으나, 2013년도 5월에 출시된 2.0에서부터는 다양한 특징들이 추가됨

- 인증(authorization) 부분은 개인으로 등록하던 시스템에서 기관(organization)의 도입으로 데이터셋에 대해 기관별로 권한(permission)을 지정 가능
- 이전에 기관에 대한 권한의 역할을 대신했던 group 기능은 데이터셋 수집에 대한 태그 정보로 활용할 수 있도록 CKAN에 계속 존재

○ 사용자, 데이터셋, 그룹, 기관의 활동 내역이 이력으로 표시됨. 다른 사용자 및 데이터셋, 그룹 기관의 활동 내역을 “follow”를 통해 지속적으로 뉴스피드를 통해 갱신

○ 확장 기능인 CKAN Extension을 통해, 데이터의 형태에 따라 table 형태, 그래프, 타임라인, 지도상에 미리보기가 가능하며, v2.0부터는 회사(ArcGIS 등) 및 서비스에서 만든 각종 Extension을 통해 다양한 파일형식에 대한 미리보기 기능 확장

○ CKAN 2.0부터 JSON기반의 API를 통해 프로그래밍이 가능한 RESTful API를 지원하며, DataStore와 FileStore 모두 CRUD(쓰기, 읽기, 수정, 삭제)가 가능한 인터페이스로 API를 제공하며 필요시 인증 요청이 가능

- 강력한 API 지원으로 다른 기관 시스템 및 사용자들의 데이터 저장, 쿼리 및 활용이 가능

○ 링크드데이터와 RDF는 1.7버전부터 지원되었으며, 그 이전에는 Extension으로 지원

- 링크드데이터의 경우 RDF에서 데이터셋을 불러오는 것에 특별한 설정이 불필요
- 링크드데이터는 API를 통해서 ‘application/rdf+xml’ 또는 ‘text/n3’로 불러오거나 데이터셋 URI 마지막에 mimetype을 추가하는 것만으로 불러올 수 있음

○ CKAN의 메타데이터 스키마는 더블린코어와 DCAT 스키마를 기본적으로 사용하여 매우 심플하며 DCAT-AP와 같이 확장 적용이 가능

- 더블린코어(Dublin Core)와 DCAT 스키마를 기본적으로 사용

56) Winn, Joss (2013), “Open Data and the Academy: An Evaluation of CKAN for Research Data Management“, IASSIST 2013, 14-17.의 내용을 인용하여 재구성

○ 처음 데이터 카탈로그로서의 역할로 출발했던 CKAN의 개발은 파일 스토리지(FileStore) 그리고 테이블 형태 데이터의 데이터스토어(DataStore)로 확장됨

- 스프레드시트(엑셀 파일)의 업로드와 저장과는 달리, DataStore는 직접 PostgreSQL 관계형 데이터베이스나 경우에 따라 적당할 파일로 변환하거나 저장 시 DataStore로 Import
- 1.8버전에서는 DataStore에 ElasticSearch를 사용했으나, 현재는 PostgreSQL을 사용하여 full SQL 인터페이스를 제공
- DataStore가 PostgreSQL을 사용하면서부터 다른 관계형 데이터베이스가 사용하는 것과 같은 방식으로 관리가 가능해짐
- CKAN FileStore는 CKAN이 다른 로컬 및 원격 저장소(아마존 S3)를 사용할 수 있게 해주고, 데이터를 업로드하고 메타데이터를 수정하기 위한 API가 존재

○ 2.0부터 웹UI가 완전하게 새 디자인으로 바뀌면서 사용성 측면이 크게 개선 됨

- 새로운 데이터셋과 자원을 추가하기 위한 간단한 WorkFlow를 선보이면서, 이전에 가능했던 기능들(예를 들면 metadata history 보기)이 없어진 경우도 있음

5. 시사점

○ 지식을 세계에 오픈하고 이를 사용하고 유용하도록 하는 것이 목적인 Open Knowledge에 의해 주도된 CKAN은 이미 어딘가에 게시(publish)된 데이터의 목록인 데이터 카탈로그로서 출발했고 데이터를 게시하는 데이터 소유자에 초점이 맞추어져왔음

○ 공공 서비스의 투명성을 개선하는 목적을 가졌던 공공기관과 정부에 의해 널리 후원을 받아오면서 CKAN의 개발이 진행되어왔으며, 기관의 데이터를 저장하고 유지하는 아카이브 시스템으로서 개발되어오지 않았기 때문에, 현존하는 프로세스와 시스템을 사용하여 수행되어왔을 것임

○ 상용 데이터 연계 플랫폼 솔루션 Socrata의 사례를 보면, 최근 기업 내부의 데이터와 공공 데이터를 연계하여 활용하고자 하는 니즈가 대두됨에 따라 정부 기관에서 주로 사용되어 오던 Socrata를 기업 내부에서도 활용하고자 하는 고객의 요구사항에 따라 기업의 내부 데이터 및 외부 데이터 연계용으로 사용하기 위한 기능이 확장되고 있음[57)]

○ 오픈소스인 CKAN을 사용을 사용 시 꾸준한 업데이트는 기능 개선 측면에서 장점이 있으나, 유지보수 측면에서는 부담이 됨

○ OPEN DATA HANDBOOK에 따르면, 정부 데이터 목록(catalog)을 만드는데 있어 처음부터 소프트웨어를 개발하는 것보다 세계 여러 나라 정부에서 도입하여 사용 중인 무료 및 오픈소스 기반 플랫폼(CKAN 등)을 활용 시 초기 인프라에 대한 투자를 줄일 수 있음

제2절 CKAN의 기능 확장 (CKAN Extension)

○ CKAN 확장 기능(Extension) 공식 레지스트리(http://extensions.CKAN.org)에 등록된 CKAN 확장 기능(Extension) 223개를 대상으로 데이터 연계 플랫폼과 관련하여 확장 가능한 기능에 대한 현황을 살펴보고, 데이터 연계에 특화된 기능 확장 및 기업에서 CKAN을 활용하고자 하는 노력에 대해서 살펴보고자 함

○ CKAN 확장 기능(Extension)은 CKAN 기본 플랫폼의 기능을 확장, 변경하는 방법으로, 2010년 가을에 처음 도입된 후 계속 확장되어 옴
- 초기에는 파이썬(Python) 패키지 형태로 Extension 목록을 WIKI에 수집, 관리하였으나, 2014년도 이후 레지스트리에 등록하여 관리하는 것으로 변경
- 시각화, 문서 미리보기 등의 기능추가, 커스텀 테마, 검색기능 추가, 타 서비스 및 저장소, 사이트 연계, 인증, 메타데이터, 데이터 관리 등의 확장 기능이 있음

○ CKAN 확장 기능(Extension) 중 CKAN에 기능을 추가하는 역할을 하고 있는 것이 51%에 해당하며, 미리보기 및 시각화, 사용자 테마, 검색, 연계, 인증, 데이터 관리 등의 기능이 해당됨
- 공간 시각화, 그래프 시각화, 대시보드, DB미리보기 등 시각화에 해당하는 부분과 검색 관련 기능 추가가 있음
- 데이터 연계 관련 기능으로는 타서비스 연계, 클라우드 등 저장소 연계, 웹사이트 연계가 있음
- Google 인증, LDAP 인증, ODI Data 인증, Drupal 인증, 오픈 ID 인증 등을 지원하도록 함
- 데이터 관리 측면에서 메타 데이터의 관리 및 연계 기능이 있음

57) 함유근 (2015), 「이것이 빅데이터 기업이다」, 삼성경제연구소, 193.

<표 3-5> CKAN Extension 기능별 분류

기능	분류	extension 종류	개수 (중복가능)
미리 보기 및 시각화, 사용자 테마	리소스뷰	공간, 그래프, 시각화, 대시보드, DB미리 보기 등	13개
	테마	커스텀 테마	16개
검색		공간정보 관련	9개
연계	서비스 연계	데이터 포털, 시각화 도구, 문서 도구	16개
	저장소 연계	DB 및 저장소 연계 (CartoDB, MongoDB, hadoop) 클라우드 저장소 연계 (Amazon S3 등)	6개
	웹사이트 연계	데이터 포털 연계 등	5개
인증		Google, LDAP, ODI Data, Drupal, 오픈 ID 인증 등	12개
데이터 관리		메타 데이터 관련 기능	56개
기타		속도개선, 데이터 사전 등	115개
합계			223개

○ 기업에서의 CKAN을 적용하고자 하는 노력으로 서비스 연계 및 저장소 연결, DB 및 검색 엔진과의 연계 등의 확장 기능(Extension) 개발이 있어왔음

- 기업 및 서비스 관련 Extension으로는 데이터 연계와 관련된 기능으로, Google, Amazon, Socrata, Junar, CartoDB, 기업 서비스와 연계 가능하도록 기업 주체로 개발된 40여개의 CKAN Extension이 개발되어 있음
- 로그 분석 서비스와의 연계로 Google Analytics 등이 있음
- 데이터 저장 서비스 및 DB와의 연계 사례로 Amazon S3, mongoDB 등이 있음
- 검색 엔진과의 연계 사례로 Amazon CloudSearch, 타 데이터 플랫폼과의 연계사례로 Socrata-harvester, Junar 등이 있음

<표 3-6> 데이터 연계와 관련된 CKAN Extension 패키지

No	기능	확장 기능 이름	관련 서비스	연계 대상 기업,서비스,언어	제작자	설명
1	타서비스와의 연계	Active Directory authentication for CKAN	Active Directory Federative	Microsoft	NHS England	Microsoft의 Active Directory Federated Service(ADFS) API
2	타서비스와의 연계	Array Express importer for CKAN	ArrayExpress	ArrayExpress	Jim McCusker	An ArrayExpress importer for CKAN
3	검색 서비스와의 연계	B2FIND Extension for CKAN	B2FIND	B2FIND	B2FIND	유럽 데이터 인프라 구축 프로젝트 EUDAT의 연구데이터 검색 프로그램 B2FIND와의 연계
4	GIS 서비스와의 연계	CartoDB to CKAN Extension	Jim McCusker	CartoDB	Chris Whong	CartoDB의 테이블에 대하여 CKAN 데이터셋 자동 생성
5	파일 업로드	CKAN upload Python tool	Python	Python	Steve Bennett	
6	데이터 연계	CKAN4j	Java	Java	SciamLab	CKAN JAVA 클라이언트 라이브러리
7	데이터 연계	R client for CKAN RESTful API	R	R	rOpenSci	R client for CKAN RESTFUL API
8	데이터 연계	Reegle Climate Tagger CKAN Extension	Climate Tagger	Climate Tagger	Reegle	Climate Tagger(기호관련 지식 커뮤니티에 연결하는 도구세트)에서 제공된 tag를 제공하는 API
9	타서비스 연계 – 클라우드 저장소	Store resources on cloud storage (AWS)	S3	AWS	Link Web Services Pty Ltd	아마존 AWS S3 스토리지에 CKAN 리소스 저장
10	타서비스 연계 – 클라우드 저장소	Libcloud support (AWS, Azure, Google, and 50+ cloud providers) for CKAN	Libcloud	Libcloud	open.canada.ca	클라우드 서비스 (아마존 S3, 마이크로소프트 Azure, 구글 등)에 연결가능한 아파치 libcloud의 CKAN Extension으로 CKAN 저장소를 클라우드로 지정
11	데이터 연계	Project Open Data Extension	Project Open Data Extension	Project Open Data Extension	data.gov -GSA	US project open data 의 데이터를 CKAN 카달로그로 가져오기 위한 datajson 플러그인
12	데이터 연계	CKAN+DCAT	DCAT연계	DCAT 연계	CKAN	CKAN을 다른 DCAT 사용하는 RDF 카달로그에서 사용하도록 하는 플러그인
13	데이터 연계	DCAT-AP-IT	DCAT연계	이탈리아 오픈 데이터 포털	GeoSolutions (https://github.com/geosolutions-it)	이탈리아 오픈 데이터 포털 확장팩(DCAT-AP-IT)을 사용
14	데이터 연계	HTTP PROXY	HTTP PROXY	HTTP PROXY	Open Knowledge	drupal, wordpress 같은 컨텐츠 관리 시스템(CMS) 연계
15	타서비스 연계	CKANext-drupal	drupal	drupal	Open Data Initiative	Open Data Initiative (ODI) 오픈데이터 drupal 모듈 Open data package의 drupal 모듈을 사용할 수 있게 해줌
16	타서비스 인증	CKAN Drupal7 Authentication	Drupal7	Drupal7	Open Knowledge	Drupal7 계정을 CKAN에서 사용할 수 있도록 인증체계연동
17	타서비스 연계	jQuery widget to embed CKAN datasets on external sites	외부 사이트	외부 사이트	Oleg Lavrovsky (https://github.com/datalets)	외부 사이트에 CKAN 데이터셋을 embed 시키는 Jquery 위젯
18	데이터 연계	CKANext-environmentcanada CKAN 2.0 Extension	캐나다 환경 데이터	data.gc.ca	thriuin	캐나다 환경 데이터를 data.gc.ca으로 import
19	데이터 연계	Feed Extension	RSS/ATOM	RSS/ATOM	Open Knowledge	RSS/ATOM 피드로부터 컨텐츠 페이지 유입
20	데이터 연계	CKAN Google Analytics Extension	Google Analytics	Google	CKAN	웹로그분석(Google Analytics, GA)데이터의 CKAN 통합 플러그인
21	데이터 연계	Google 인증 연동	Google 인증	Google	Yacme S.r.l (dev@yacme.com)	구글 인증 연계

No	기능	확장 기능 이름	관련 서비스	연계 대상 기업,서비스,언어	제작자	설명
22	데이터 연계	Google Docs 플러그인	Google Docs	Google	Consumer Data Research Center	구글 문서 도구 플러그인
23	타사이트 연계	GovData	GovData.de	GovData	Fraunhofer FOKUS	GovData.de 사이트 연계
24	데이터 수집 연계	iRODS Extension for CKAN	iRODS	iRODS	Kata team repository	iRODS 플러그인
25	데이터 연계	CKAN extention for junar	junar	junar	Ciudadano Inteligente	데이터플랫폼 junar 연계
26	인증	CKAN Extension to manage logins and logouts from Drupal	Drupal 인증	Drupal	weso	Drupal에서 로그인, 로그아웃 관리
23	인증	CKAN plugin to provide LDAP authentication	LDAP 인증 플러그인	LDAP	U.K. Natural History Museum	LDAP 인증을 제공
24	타서비스 연계	CKAN Extension for MongoDB	몽고DB연동	mongoDB	Alex Byrnes	MongoDB연계
25	연계-실시간 리소스생성	Create and visualize NGSI resources provided by a Context broker	Context broker	Context broker	Guillermo Zarzuelo	Context broker 에 의해 제공되는 실시간 리소스 생성 기능
26	타사이트 연계	CKAN Extension for data	data.nhm.ac.uk 데이터 스토어	Natural History Museum	Natural History Museum	data.nhm.ac.uk 데이터 스토어 업데이트 기능
27	타서비스 연계	OData endpoint; fork optimized for Tableau Public	Tableau 시각화	Tableau	OpenGov	Tableau 시각화
28	타서비스 연계	basic OData endpoint	OData endpoint	OData endpoint	whythawk	Odata 3.0 endpoint 제공
29	타서비스 연계	OpenRefine CKAN Storage Extension	OpenRefine	Google	OpenGov	구글 Refine에서 직접 CKAN으로 데이터를 로드하는 API
30	타서비스 연계	A CKAN Extension for logging in using Mozilla Persona	Persona	Mozilla	CKAN	모질라 Persona에 등록된 사용자의 CKAN 로그인 기능 추가
31	타서비스 연계	QGIS CKAN-Browser	QGIS	QGIS	BergWerk GIS	QGIS plugin to load and display Open Data from CKAN
32	타서비스 연계	Archive CKAN datastore to S3	S3	amazon	CKAN	amazon 클라우드 스토리지 S3 연계
33	타서비스 연계	CKAN Sync for Screendoor	screendoor	Better Technology	The Department of Better Technology	Screendoor에서 CKAN 데이터셋 자동으로 동기화
34	타사프레임워크로부터 통합	A plugin to integrate Silk Link Discovery Framework into CKAN	Silk Link Discovery Framework	Silk Link Discovery Framework	memaldi	Silk Link Discovery Framework로부터 CKAN으로 통합
35	타 플랫폼 연계 데이터 수집	Socrata harvester to allow CKAN to keep in sync with a Socrata store	Socrata store	Socrata	OpenDataMonitor	Socrata 스토어와 CKAN의 동기화를 유지시켜주기 위한 socrata harvestor
36	타플랫폼 연계	CKANext-stromnetzberlin	Stromnetz Berlin	Stromnetz Berlin	Fraunhofer FOKUS	Stromnetz Berlin(독일 전기회사)로 부터 데이터를 harvest 하기 위한 플러그인
37	메타데이터 연계	US Common Core Schema additional custom fields	US Common Core	US Common Core	data.gov -GSA	US Commn Core 스키마 설치와 관련하여 사용자 정의 필터 제공 (USmetadata)

제3절 국내외 CKAN 활용 현황

1. 국내외 데이터 포털의 CKAN 활용 현황

○ 중앙 정부 및 지방 정부 기관에서는 국내 공공데이터 개방 포털에서 CKAN 활용 중 또는 활용 예정임

- 2011년도 7월 오픈한 공공데이터 개방 포털(data.go.kr)은 자체 플랫폼으로 개발되었으며 CKAN 카탈로그를 활용 중임[58)59)]
- 서울시 정보 소통광장은 “열린 시정 2.0”(Nude Project)를 바탕으로 CKAN 및 drupal과 전자정부프레임워크(eGov Framework)를 활용하여 플랫폼을 구축[60)]됨
- 수원시 공공데이터 포털은 전체 603건의 데이터 중, 중소기업 현황 등 30건 데이터셋[61)]를 CKAN API의 유형으로 제공하고 있음
- 세종시 공공데이터 포털인 세종 데이터 실록은 전체 186건의 데이터 중, 115건의 공공데이터셋[62)]를 CKAN API의 유형으로 제공하고 있으며, 시민의 참여로 등록되는 시민 참여형 데이터를 가공하여 참여형 공공데이터로 함께 API로 제공함

○ 국외 CKAN 개발 및 도입사례에는 영국의 공공 데이터 포털 data.gov.uk, 프랑스의 data.gouv.fr 등이 있음

- 2010년에 오픈한 영국의 공공데이터 포털 data.gov.uk은 CKAN 기반으로 데이터 카탈로그를 운영하고 있으며 블로그, 포럼, 코멘트 등을 이용하기 위해 Drupal을 함께 사용
- 프랑스의 총리 task force 팀 Etalab이 주도하는 공공데이터 포털 data.gouv.fr은 CKAN을 기반으로 함[63)]

○ 미국, 오스트레일리아 등은 기존에 다른 플랫폼을 도입했었으나 CKAN으로 전환한 사례임

- 미국의 공공 데이터 포털인 Data.gov의 경우 2009년 Socrata 플랫폼으로 구축, 오

58) 백인수 (2013), “오픈데이터 플랫폼과 국가 데이터 전략방향”, 「정보처리학회지」, 16.
59) 박경현 외 (2017), “공공 빅데이터 개방과 표준화”, 「OSIA S&TR Journal」, 30(2): 21-25.
60) 서울시 전자정부 (2015.1.23.), https://seoulsolution.kr/ko/content/시민의-알권리-행정정보-공개-자동시스템화 참조.
61) 수원 공공데이터 포털, http://data.suwon.go.kr/pubr/data/viewDataSetSelfList.do?gubun=&mId=0101000000 참조 (접속일: 2017.8.31.).
62) 공공데이터 포털 세종 데이터 실록, https://data.sejong.go.kr/ 참조 (접속일: 2017.8.31.).
63) Cyrille CHAUSSON (2014.10.31.), “Data.gouv.fr: the French Open Platform for Open Data”, https://joinup.ec.europa.eu/community/epractice/case/datagouvfr-french-open-platform-open-data 참조 (접속일: 2017.09.29.).

픈하였으나, 2012년도 CKAN 카달로그 도입과 함께, 2013년 Data.gov 2.0의 새로운 이정표[64]를 세우면서 Open Knowledge과의 협력을 통해 영국의 오픈소스 데이터 플랫폼 CKAN으로 전환하면서 데이터 카탈로그, 데이터 관리, 데이터 사용성에서 큰 개선을 보였고, 특히 지리 정보와 일반 정보를 결합한 단일 카탈로그로 개방성과 호환성을 개선함

- 오스트레일리아의 정부기관 데이터 플랫폼인 data.gov.au는 CKAN을 이용한 개방형 데이터 플랫폼(Gov2.0)으로 전환함[65]

2. 기업의 CKAN 활용 사례

가. CKAN을 이용한 데이터 연계 및 플랫폼 구축 사례

○ 농업 빅데이터 기업 몬산토[66]는 오픈소스 구성 요소만을 사용하여 클라우드에서 확장 가능한 플랫폼 구축하고 기업 데이터, 고객 데이터, 공공 데이터를 연계하여 활용 가능하도록 함

- 사용된 기술 요소는 CKAN 카달로그(데이터 목록), PostGIS(데이터베이스), QGIS(지리정보시스템), AWS ECS(클라우드 호스팅서비스), amazon S3(클라우드 저장소) 등으로 기존의 시스템을 오픈소스 기술로 전환함
- 몬산토의 고객, 제품, 회사, Event, 농작물 관련 데이터(경지, 환경, 날씨)는 지리정보 서버에 있는 데이터와 함께 데이터 카달로그에 저장되어 검색과 동시에 해당 지역의 지리정보와 함께 데이터 연계 활용이 가능함
- 몬산토에서는 현재 자사의 오픈소스 기반의 지리정보 기반 데이터 플랫폼을 운영하기 위하여 AWS 클라우드, CKAN, QGIS의 운영이 가능한 플랫폼 엔지니어를 채용하고 있음[67]

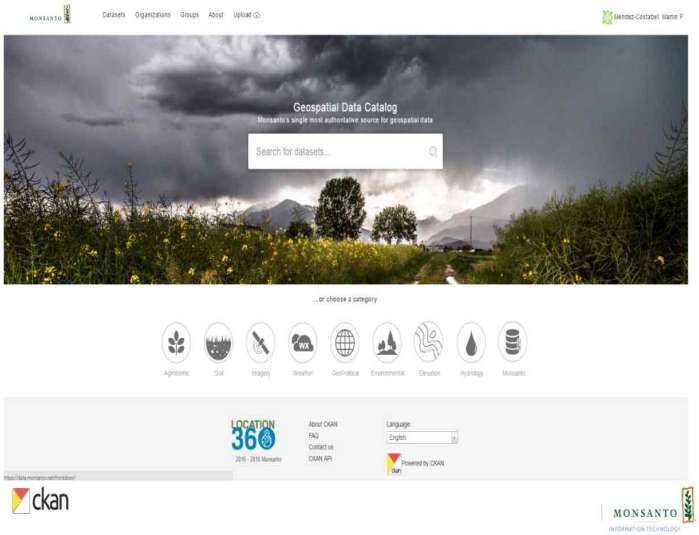
---

64) DATA.GOV (2013.01.28.), “CKAN: the Horizon for Data.gov 2.0”, https://www.data.gov/blog/CKAN-horizon-datagov-20 참조 (접속일: 2017.08.21.).

65) 과학기술정보통신부, 정보통신산업진흥원 (2013), 「2013 소프트웨어 산업 연간보고서」, 과학기술정보통신부, 정보통신산업진흥원.

66) Naghman Waheed, Martin Mendez-Costabel (2017.03.15.), “Enterprise Geospartial Platform - A perfect fusion of cloud and open source technologies”, MONSANTO, Strata + Hadoop World.

67) Monsanto Geospatial Platform Engineer Job offer (2017), https://jobs.monsanto.com/job/st-louis/geospatial-platform-engineer/769/5100033 참조 (접속일: 2017.09.25.).



<그림 3-2> 농업 분야 기업 몬산토 지리정보기반 데이터 카달로그

나. CKAN 기반 데이터 플랫폼 구축 및 기술 지원 사례

○ 오픈소스 기반의 데이터 플랫폼인 CKAN을 활용한 수익 모델을 가지고 있는 기업에는 Link Digital, Viderum, OpenGov 등이 있으며, 주요 사업유형은 주로 정부를 대상으로 하는 B2G사업이 공통적이며, Link Digital 같은 경우는 기존 사업영역인 웹디자인, 개발, 컨설팅 서비스와 함께 B2B, B2C 서비스도 함께 하고 있음

(1) 오픈 소스 CKAN의 민간 비즈니스 모델 사례 (Viderum사 사례)

○ CKAN을 개발하고 있는 주체인 비영리 기관인 Open Knowledge International의 상업적 자회사로서 클라우드 기반의 인프라, CKAN 기술지원, 개방형 데이터 웹사이트 및 CKAN 확장 개발과 관련된 기술 컨설팅을 제공

(2) 클라우드 서비스를 이용한 금융권 CKAN 이용사례 (Link Digital사 사례)

○ 아마존 웹 서비스에서 AWS Marketplace를 통해 CKAN을 이용한 오픈 데이터 플랫폼을 서비스

○ 금융 분야에서 은행, 보험회사들을 대상으로 데이터 가공, 분석 전반에 걸쳐서 CKAN에 있는 데이터를 활용하는 비즈니스 인텔리전스 플랫폼으로 활용하여 AWS CLOUD를 통해 제공

(3) 연방 및 지방정부 오픈데이터 포털 지원 및 호스팅 (Link Digital사 사례)

○ 연방 및 지방정부의 오픈 데이터 포탈을 지원하고 호스팅 (data.gov.au, data.nsw.gov.au, data.vic.gov.au, data.sa.gov.au, data.wa.gov.au) 하고 있음

<그림 3-3> Link Digital을 통해 호스팅하는 데이터 포털

<표 3-7> 기업의 CKAN 활용 사례 (CKAN 기반 상용 서비스)

No.	기업(관)명	홈페이지	설립 연도	CKAN활용 수익모델	사업 유형	주요 고객
1	Link Digital	http://linkdigital.com.au	2001	웹사이트 디자인, 호스팅, CMS구성, UX디자인, 프로젝트관리, 마케팅캠페인 및 플랫폼 구축/관리, 비즈니스 분석, 유지보수, 데이터자문서비스	B2C, B2G, B2B	학교, 정부, 공익단체
2	Viderum	http://www.viderum.com	2015	클라우드 기반 인프라, CKAN기술지원, 개방형데이터웹사이트 및 CKAN Extension 개발, 오픈데이터 기술 컨설팅, CKAN 호스팅	B2G	정부, 기관, 지방정부
3	OpenGov	http://OpenGov.com	2012	Smart Government Cloud	B2G	정부

제4장 공공·민간 데이터 연계 현황

선진국 주요 데이터 포털에서 CKAN 등의 데이터 플랫폼을 활용하여 데이터를 제공 및 연계하고 있으며, 국내외 민간기업도 데이터 연계를 위한 Extension 개발 등 CKAN을 활용하여 데이터를 연계하려는 다양한 시도가 이루어지고 있음. 본 장에서는 공공·민간 데이터 연계사례를 통해 데이터 연계의 필요성을 알아보고, 주요 주제별 민간데이터 현황을 통해 CKAN을 도입하여 민간 데이터를 연계할 가능성이 있는 기업들을 도출하고자 함

제1절 공공·민간 데이터 연계 사례

1. 데이터 플랫폼 사례[68]

○ 구글은 기계학습 스타트업을 위해 4종의 데이터셋인 오픈이미지, YouTube-8M 데이터셋, 구글북스 Ngrams, 구글 트렌드 데이터 스토어 자료를 공개함[69]

- 오픈 이미지 데이터셋(Open Image Data Set)[70] 9백만 건을 재사용 가능하게 공개하였으며 사람이 직접 눈으로 검증하여 6000개의 카테고리로 나누어서 라벨을 할당하여 활용도를 높임
- 오픈 이미지 데이터셋에서 사용된 이미지들은 CC BY 2.0 라이센스 하에 공개되어 있는 이미지를 사용하였으며, 각 사진의 특징을 나타내는 라벨을 붙이고, 이미지 속의 객체들에 사각형을 씌우고 수천 개의 클래스로 라벨을 할당된 주석은, Google Inc. 에 의해 CC BY 4.0 라이센스로 재사용 가능하도록 공개되고 있음

<그림 4-1> 오픈 이미지 데이터셋

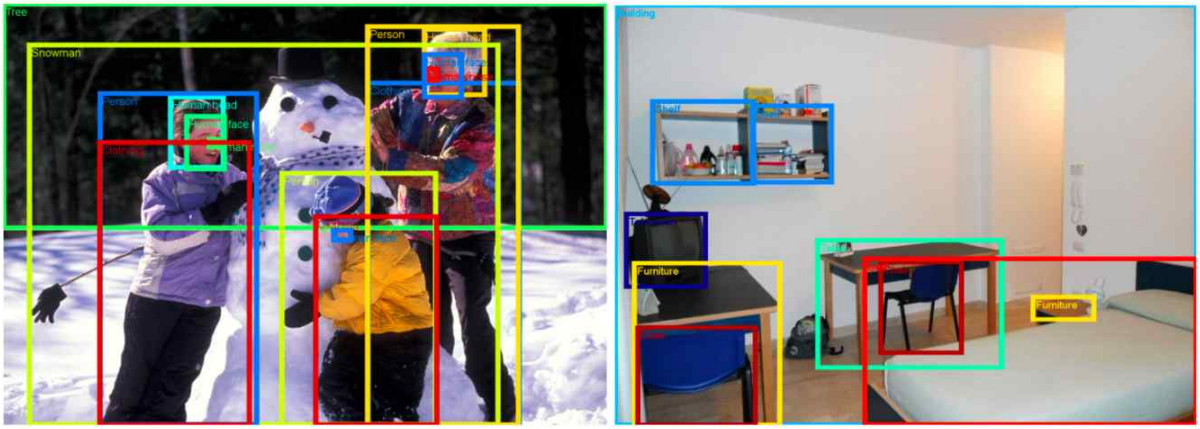
(좌: Kevin Krejci, Ghost Arches, 우: J B, Some Silverware)

※ 출처 : Introducing open image dataset
https://research.googleblog.com/2016/09/introducing-open-images-dataset.html 참조

68) 한국데이터진흥원 (2017), "민간데이터 연계와 CKAN의 활용", 「Data Issue Report」, 2017-02(106)의 민간데이터 연계 사례의 목록을 일부 참고하여 재구성함.

69) InfoWorld (2016.10.17.), "4 Google data sets to kickstart machine learning", https://www.infoworld.com/article/3131515/artificial-intelligence/4-google-data-sets-to-kickstart-machine-learning.html 참조 (접속일: 2017.08.24.)

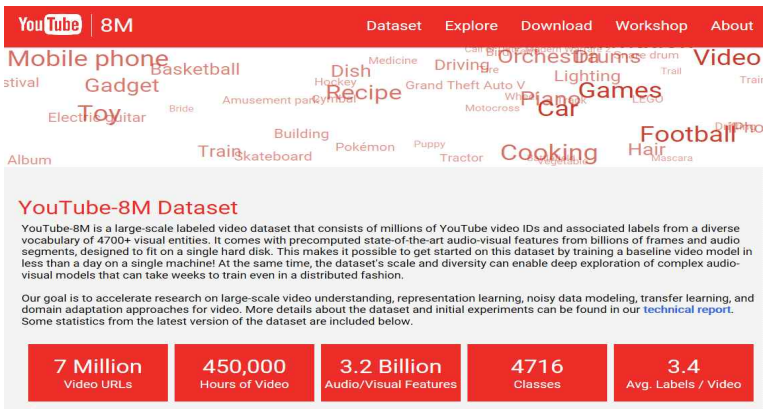
70) Open Images Dataset (2017.07.20.), Open images dataset, https://github.com/openimages/dataset 참조



<그림 4-2> 오픈 이미지 데이터셋

※ 출처: Krasin I. et al (2017), "Open Images: A public dataset for large-scale multi-label and multi-class image classification,", https://github.com/openimages 참조.

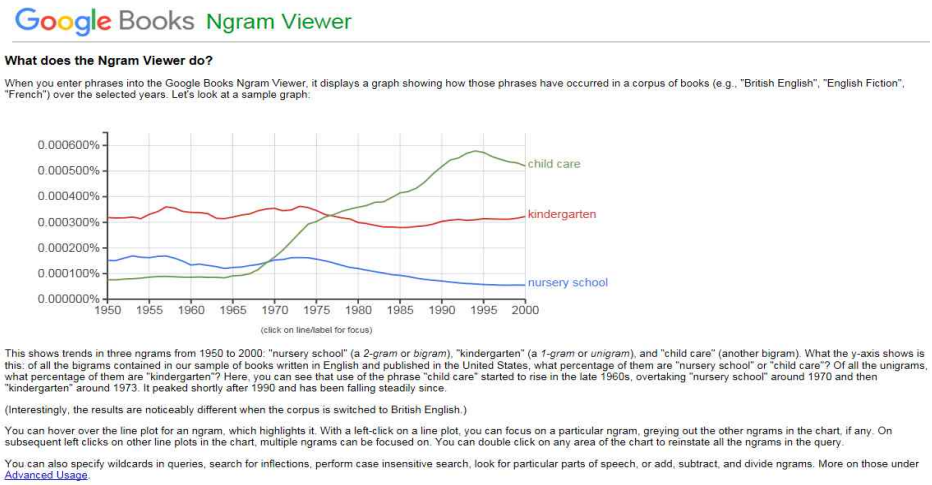
- YouTube-8M 데이터셋은 조회수 1,000회 이상, 최소 2분 이상의 YouTube 기본 카테고리로 사전 분류된 8백만 개의 동영상으로 구성되어있으며, YouTube-8M은 8백만 개의 YouTube 비디오와 4800개의 클래스를 가지고 있어 기존에 가장 큰 비디오 데이터 세트로 알려져 있던 Sports-1M이 약 1백만 개의 YouTube 비디오와 500개의 스포츠 관련 클래스를 가지고 있는 것에 비해 한 단계 발전함
- 비디오는 이미지 보다 수작업으로 주석을 달기위해 많은 시간과 처리 및 저장에 많은 비용이 드는 작업인 비디오 메타데이터 및 컨텐츠 분석 작업이 통합된 상태의 데이터가 제공되며, Tensorflow를 이용한 딥러닝 분석에 사용 가능한 형태로 제공되어 연구자와 학생들의 데이터 전처리에 드는 자원 병목현상을 극복 가능함



<그림 4-3> YouTube-8M Dataset

※ 출처: YouTube-8M Dataset https://research.google.com/youtube8m/ 참조

- 구글은 YouTube에 공개된 영상을 재가공하여 CC BY 4.0 라이센스[71]로 재사용 가능하도록 온라인 탐색과 오프라인 다운로드 형태로 공개하고 있음[72]
- 구글은 1500년부터 2012년까지 출간된 영어책을 디지털화 하는 작업을 통해 구축된 구글북스 Ngrams 자료를 기반으로 검색어에 대한 단어의 빈도를 웹 인터페이스를 통해 Ngram viewer를 통하여 대중에게 공개하고 있음
- 아마존 웹 서비스(AWS)에서는 공개된 구글북스 Ngrams 자료 2.2TB의 데이터셋을 CC 3.0 라이센스를 통해 하둡에서 사용하기 쉬운 형태로 아마존 S3[73] (클라우드 저장공간 서비스) 상에서 무료로 사용하도록 제공하고 있음

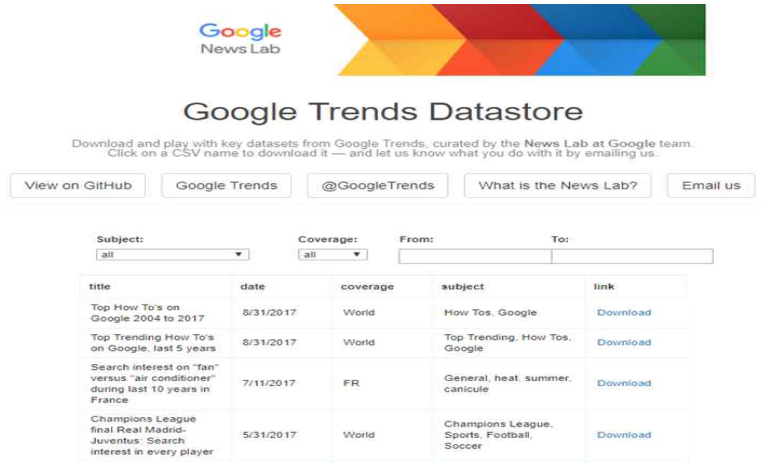


<그림 4-4> 구글북스 Ngrams

※ 출처: 구글북스 Ngrams, https://books.google.com/ngrams 참조

- 구글 트렌드 데이터 스토어에서는 검색을 통해 트렌드를 알 수 있는 구글 트렌드 자료를 저널리스트 등이 활용하기 쉽도록 구글 뉴스랩(News Lab at Google) 팀에서 큐레이팅한 뒤 무료로 제공하고 있음
- 구글 트렌드 데이터 스토어 github 저장소[74]에 100건의 데이터셋(총6MB)이 등록되어 있으며 용량이 크지 않은(수KB~수MB) 데이터셋들로 구성되어있음

71) CC BY 4.0 license, https://creativecommons.org/licenses/by/4.0/ 참조
72) Sudheendra Vijayanarasimhan and Paul Natsev (2016.09.28.), "Announcing YouTube-8M: A Large and Diverse Labeled Video Dataset for Video Understanding Research", Google Research Blog (접속일: 2017.08.30.)
73) 아마존 S3 : 클라우드 저장 공간 서비스
74) Google Trends Datastore, http://googletrends.github.io/data/ 참조 (접속일: 2017.8.31.)

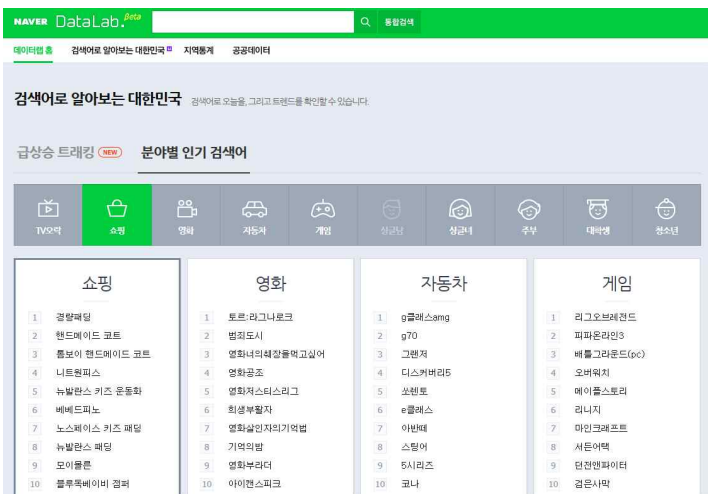


<그림 4-5> 구글 트렌드 데이터스토어

※ 출처: Google Trends Datastore, http://googletrends.github.io/data/ 참조

○ 네이버 데이터랩은 "데이터랩"을 통해 10년간 축적한 데이터를 무료로 공개하여 민간기업과 공공기관이 보유한 데이터를 활용할 수 있도록 함

- 검색어 데이터 기반의 쇼핑 빅데이터를 공개, 온라인 쇼핑몰을 이용하는 중소 사업자 및 일반 이용자들의 활용도를 높임[75]
- 지역별, 업종별로 이용자들의 관심도를 파악할 수 있는 지역별 관심도 기능을 추가하여 이용자가 실제로 클릭한 업체 정보를 제공[76]



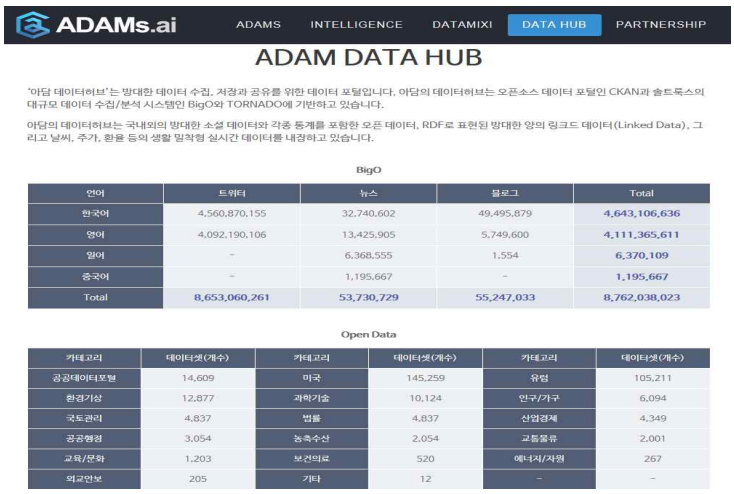
<그림 4-6> 네이버 데이터랩

※ 출처: NAVER DataLab http://datalab.naver.com/ 참조

75) 디지털데일리 (2016.07.20.), "네이버 데이터 랩, 쇼핑 빅데이터 공개" (접속일: 2017.08.17.)
76) 디지털데일리 (2016.04.22.), "강릉서 가장 인기 업종은? 네이버, 지역정보 빅데이터 공개" (접속일: 2017.08.17.)

○ 솔트룩스는 인공지능 플랫폼과 함께 소셜데이터, 오픈 데이터를 무료로 개방하여 인공지능 인지분석 서비스를 무료로 제공[77]하며 맞춤형 수집, 심층 분석, 레포트로 프리미엄 서비스를 유료로 제공

- 100억건의 소셜데이터를 활용하여 인공지능 인지분석 서비스 제공
- 100억건의 소셜데이터와 34만건의 오픈데이터를 무료 개방하여 학습분석에 활용



<그림 4-7> ADAM DATA HUB

※ 출처: ADAM DATA HUB http://datahub.adams.ai/ 참고

○ 아마존 웹 서비스(AWS)는 AWS Public Datasets[78] 서비스를 통해 생물학, 화학, 경제학, 백과사전 등 다양한 주제의 대용량 공개 데이터셋을 클라우드 상에 무료로 제공하고 있어 이전에 검색, 다운로드, 가공, 분석하는데 몇 시간, 또는 며칠이 걸리던 것을 AWS 클라우드 기반에서 다운로드나 저장 없이 활용이 가능하며 그 범위는 정부 공공데이터부터 아마존 물류센터 제품정보 등 자사 민간 데이터까지 다양함

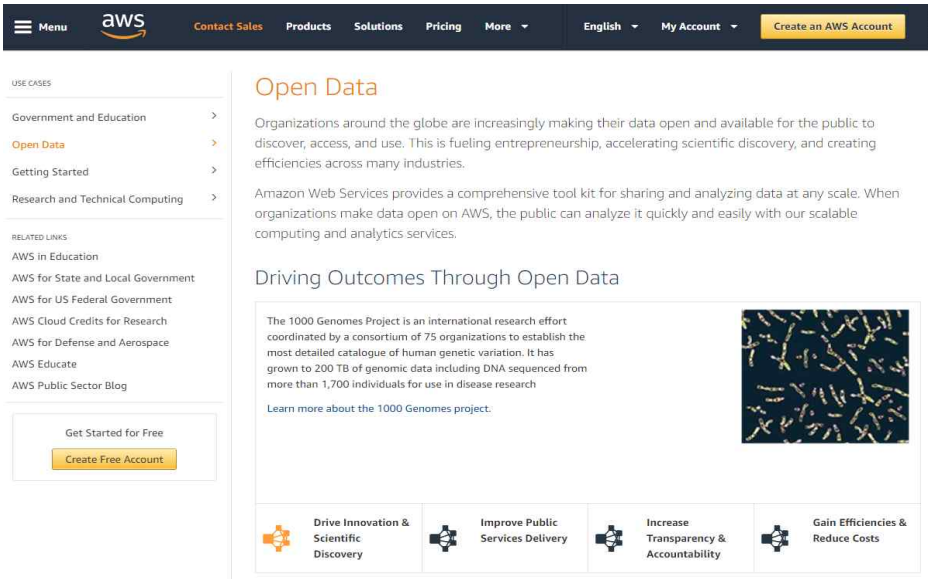
- Earth on AWS를 통해 위성에서 생성하는 지구상의 육지에 대한 위성 이미지, 날씨와 기상 및 우주 환경 데이터 모니터링, OpenStreetMap, NASA에서 관리하는 기후 및 지구 표면 위성 이미지 등을 포함한 지리정보 및 환경 데이터 세트를 제공
- 1000 Genomes Project(1000명의 인간 유전체 변이 지도), TCGA(암 유전체 공개 데이터셋), 3000개의 쌀 품종의 유전체 서열 등 유전체 및 생명 과학 관련 데이터셋 제공
- 50억개가 넘는 웹페이지로 구성된 웹 크롤링 말뭉치, 500,000개 넘는 아마존 물류

77) IT DAILY (2017.04.26.), “솔트룩스, AI 기반 데이터 클라우드 공개…데이터 100억건 무료 개방” (접속일: 2017.05.24.)

78) AWS Public Datasets, https://aws.amazon.com/ko/public-datasets/ 참조 (접속일: 2017.8.31.)

센터의 제품 설명 이미지 및 메타파일, 1억개 가량의 동영상, 이미지 모음, 구글 북스 Ngrams 등 인공지능 및 기계학습 데이터셋을 제공

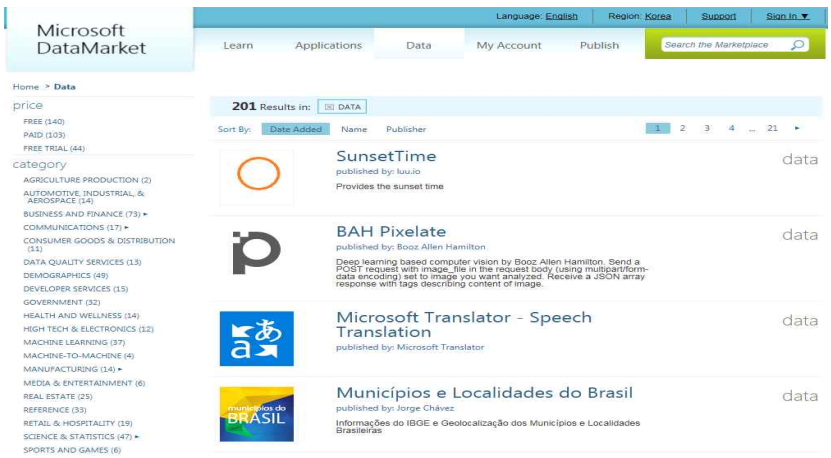
- 미국 통계국 ACS 데이터 및 USAspending.gov 등 미국 정부 공공데이터 제공



<그림 4-8> AWS Public Datasets

※ 출처: AWS Public Datasets https://aws.amazon.com/datasets/ 참고

○ Microsoft사는 Microsoft DataMarket을 통해 140건의 무료 데이터셋과 103건의 유료 데이터셋 그리고 44건의 무료 시범 (Free trial) 데이터셋을 서비스하였으며, 비즈니스/금융, 인구통계, 과학/통계, 기계학습, 교통, 소매 및 스포츠 등 22가지의 공공 및 민간 데이터를 포함한 주제로 서비스함 (2017년 3월 서비스 종료)

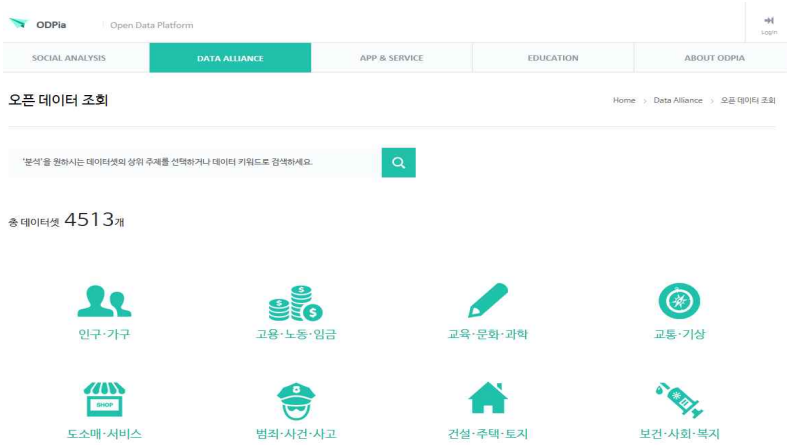


<그림 4-9> Microsoft DataMarket

※ 출처: Microsoft DataMarket https://datamarket.azure.com 참고

○ LG 오디피아는 빅데이터 공유 플랫폼 오디피아(ODPia)를 통해 공공 데이터와 소셜 데이터, 카드사 관련 데이터, 공간 데이터, 유통 데이터를 연계 및 수집 서비스함

- 공공데이터 4000여건을 인구/가구, 고용/노동/임금, 교육/문화/과학, 교통/기상, 도소매/서비스, 범죄/사건/사고, 건설/주택/토지, 보건/사회/복지 등으로 공개하고 있음
- 공공데이터 길잡이를 통해 수집한 데이터 이외에도 공공데이터포털, 부동산/국토, 일반행정, 의료/보건, 산업/경제, 인구/사회, 문화/관광, 교육, 자원/환경, 교통 등 분야별로 공공데이터를 공개하는 웹사이트 정보를 제공
- 소셜 분석을 통해 일상, 장소, 시간, 인물, 관심사, 감성, 행동, 제품 카테고리를 분류된 27만개의 키워드에 대한 소셜 맵과 함께 기업 소셜 인덱스 순위를 제공함



<그림 4-10> ODPia

※ 출처: ODPia https://www.odpia.org 참고

2. 데이터 연계 분석 사례

가. 공공・민간 연계

○ 아이슬란드 농업 자문센터[79] 사례[80][81] 에서는 아이슬란드 국립 레지스트리의 지리정보와 농장별 양고기 품질 정보를 연계 활용함

79) 아이슬란드 농업 자문 센터(Icelandic Agricultural Advisory Center) : 농민들을 위하여 축산, 사육, 수유, 관리, 기록, 부기 등과 같은 전문적인 컨설팅 제공, http://www.porthonnun.is/en/portfolio/icelandic-agricultural-advisory-centre 참조 (접속일: 2017.08.17.).

80) Linking Paths (2014.03.06.), “Icelandic Open Data, lambs and maps”, http://weblog.linkingpaths.com/post/78767700473/icelandic-open-data-lambs-and-maps 참조 (접속일: 2017.08.17.).

81) Europeandataportal e-learning module, https://www.europeandataportal.eu/elearning/en/module2/#/id/co-01 참조 (접속일: 2017.08.17.).

- 아이슬란드 농민들을 위해, 아이슬란드 국립 레지스트리의 지리정보와 농장별 양고기 품질 정보를 결합하여 양고기 품질 지도를 만들어 제공
- 최고 수준의 양고기를 생산하는 농장의 클러스터를 발견하여 고객 유치, 지역적인 요인과 계절적 요인에 대한 연구도 함께 가능

○ 통계청 공공데이터와 민간 신용정보기관의 신용 데이터 연계사례에서는 공공데이터-통계청 통계 조사 및 공공 행정자료, 민간 데이터 - 신용정보회사의 부채, 카드 사용액 등 자료를 활용함
- 통계청이 공공데이터를 비식별화 하여 데이터를 민간기관에 제공
- 통계 작성 목적에 대하여 예외를 인정하는 개인정보 보호법과 통계법을 근거로 데이터 연계 방식을 마련, 단계적 비식별화 조치를 취하여 식별자 암호화 후 속성정보와 함께 제공
- 연계 DB작성 후 연계 키는 삭제하고, 연계 DB의 외부 반출을 금지하여 분석 결과만 집계표현식 형태로 반출하여 활용
- 5만 신혼부부 가구 표본 분석하여 가구 단위의 부채 및 신용 DB구축. 개인별 신용데이터와 가구 데이터를 연계하여 가구 단위의 신용평가 모델 활용 가능

○ 서울시 교통정보, 상가정보와 비씨카드사 카드 소비데이터 연계사례에서는 서울시 교통정보, 상가정보와 비씨카드사 3000만 고객 개인정보를 제외한 카드 소비데이터를 활용하여 골목 상권 소비지출 동향을 분석

○ 신한카드의 카드 정보와 한국은행 통계 정보 연계 사례 (국내 사례)[82)83)]에서는 공공의 한국은행 제주 본부 통계 자료와 민간의 신한카드 카드결재 정보를 활용, 제주도청의 제주도 관광 통계 지표 수립하고 이를 기반으로 관광정책을 수립

○ 성남시는 건강보험심사평가원의 질병 데이터와 KCB 신용 정보를 연계 분석하여 행정동 단위 통계 데이터를 통해 공공 주치으 제도 우선 도입 여부 검토

82) 신한카드블로그 (2016.2.23.), “신한카드・제주도・한국은행, 공동 빅데이터 협업 추진”, http://www.shinhancardblog.com/202 (접속일 : 2017.2.21.).
83) 한국데이터진흥원 (2017), “민간데이터 연계와 CKAN의 활용”, 「Data Issue Report」, 2017-02(106).

제2절 연계 가능한 민간데이터

○ 앞 장에서 조사를 통해 살펴본바와 같이 국내외 주요 데이터 포털 및 데이터 마켓에서는 데이터셋을 주제별로 분류하여 제공하고 있음

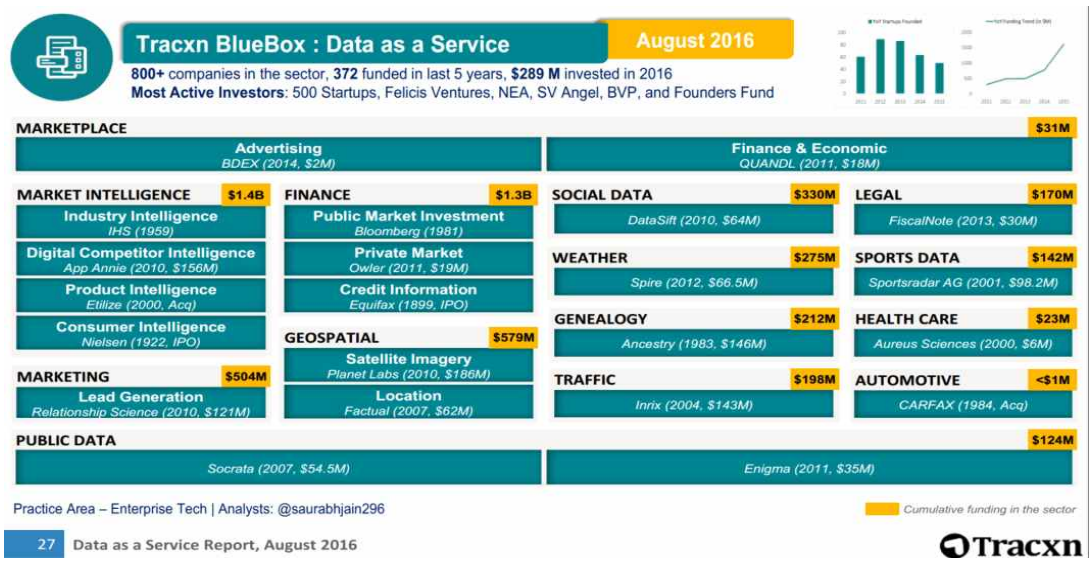
○ 본 절에서는 데이터 간 연계 가능성이 높은 주제 분야의 선정을 위해 국내외 주요 주제별 데이터 공개현황 및 민간데이터 수요를 파악하고, 사례 조사를 통해 공공·민간 데이터 연계의 시사점과 필요성을 도출하고자 함

1. 수요가 많은 민간데이터 주제 분야

○ 시장성과 활용도가 높은 주제를 탐색하기 위하여, 해외 데이터 관련 시장에서 주제 분야별 투자현황, 주요 공공데이터 포털 및 데이터스토어(www.datastore.or.kr)의 주제 분야를 살펴봄

○ 2016년에 발간된 데이터 서비스에 관련된 스타트업에 대한 투자보고서[84]에 따르면 2016년도 미국 데이터 서비스 분야 800개 회사가 289M달러의 투자를 유치

- 최근 5년간 주제 분야별 투자유치 현황을 보면 마켓 인텔리전스 1.4B달러, 금융 1.3B달러, 공간정보 579M달러, 마케팅 504M달러, 소셜데이터 330M달러, 기상데이터 275M달러, 교통 198M달러, 법률 170M달러, 스포츠 142M달러, 공공데이터 124M달러, 마켓플레이스 31M달러, 헬스케어 23M달러, 자동차 1M달러 미만 순으로 나타남



<그림 4-11> 미국 내 데이터 서비스 분야 스타트업 투자현황 (2016년)

84) Tracxn (2016.8), "Tracxn BlueBox : Data as a Service ", 「Data as a Service Report」, California: Tracxn.

○ 한편, 해외 주요 공공데이터 포털은 환경, 건강, 기후, 농업 등의 분야의 데이터셋이 높은 비율로 공개되고 있음

- 2016년 소개된 데이터 서비스 관련 스타트업 투자현황을 보면 마케팅, 지리정보, 소셜데이터, 기상, 교통, 법률, 스포츠, 헬스 등의 분야가 높은 비중을 차지
- 이 중 공공데이터에서 제공되고 있지 않은 마케팅, 소셜데이터, 스포츠, 헬스 등 민간 데이터에 해당하는 분야는 데이터 서비스 및 데이터 연계를 통한 신규 비즈니스 창출의 기회가 될 수 있음

○ 국내 주요 공공/민간 데이터 포털 및 데이터마켓에서 제공하는 데이터셋의 주제 분야를 비교함

- 데이터 마켓(데이터스토어, 빅데이터 허브 등)은 영화/공연/음악, 쇼핑/고객, 맛집/여행, 쇼핑, 채팅, 메시지, 금융, 산업동향 등 민간 기업 또는 개인이 생산하는 데이터 주제 분야가 많은 것을 확인할 수 있음

<표 4-1> 국내 주요 공공/민간 데이터 포털 및 마켓에서 제공하는 데이터셋의 주제

데이터 포털명	공공 데이터포털	수원시 공공데이터 포털	데이터 스토어	빅데이터 허브	오디피아	API store
공공 /민간	공공	지방정부	공공	민간	민간	민간
구분	데이터포털	데이터포털	데이터마켓	데이터마켓	데이터마켓	데이터마켓
주제	교육, 국토관리, 공공행정, 재정금융, 산업고용, 사회복지, 식품건강, 문화관광, 보건의료, 재난안전, 교통물류, 환경기상, 과학기술, 농축수산, 통일외교 안보, 법률	공공행정, 과학기술, 교육, 교통물류, 국토관리, 농축수산, 문화관광, 보건의료, 사회복지, 산업고용, 식품건강, 재난안전, 재정금융, 환경기상	맛집/여행, 날씨/기상, 교통, 지도/지리, 영화/공연/음악,쇼핑/고객, 기업일반, 건강/보건, 취업/창업, 금융/증권, 통계/산업동향, 부동산, 생활일반, 경매/입찰, 과학기술/자연, 논문/도서, 사진/미술, 스포츠/레저, 여성/육아, 오락/게임, 인문/사회, 자격증/시험, 전통문화, 특허/신안, 유학/연수, 물류/운송, 교육	금융, 산업동향, 유통, 날씨, 환경, 복지, 과학, 생활, 교육, 컴퓨터/정보, 게임, 미디어, 공연/예술, 관광, 건강/식품	인구/가구, 고용/노동, 임금, 교육/문화/과학, 교통/기상, 도소매/서비스, 범죄/사건/사고, 건설/주택/토지, 보건/사회/복지	메시지 올레맵, 지도/로케이션 대중교통 정보/검색 비즈니스 미디어 날씨 빅데이터/분석, 서울시 공공데이터, 쇼핑, SNS/Social, 경제/금융, 포탈/인터넷, 채팅/메신저, 사진, 음악, Blog, schedule, database, 기타

○ 데이터스토어(www.datastore.or.kr)에서 제공하는 주제별 데이터셋을 통해 수요가 높은 주제분야와 파일 유형별, 주제별로 상품 등록건수를 비교함

- 건당 조회 수가 가장 높은 카테고리는 취업/창업 분야이며, 공개된 데이터 건수가 많으면서 조회 수도 높은 분야는 지도/지리, 맛집/여행, 영화/공연/음악, 통계/산업동향, 생활일반, 날씨/기상, 과학기술/자연으로 나타남
- 특허/신안, 기업일반, 통계/산업동향, 지도/지리 등의 주제가 공통적으로 등록 건수가 높은 것으로 나타남
- 유형별로는 FILE로 제공되는 데이터의 경우 통계/산업동향, 과학기술/자연 등이 상위권에 있는 반면 API로 제공되는 데이터의 경우 지도/지리, 날씨/기상, 금융/증권 분야가 상위권으로 확인 됨

<표 4-2> 전체 및 상품유형별 데이터스토어 상품 등록 건수 주제별 순위

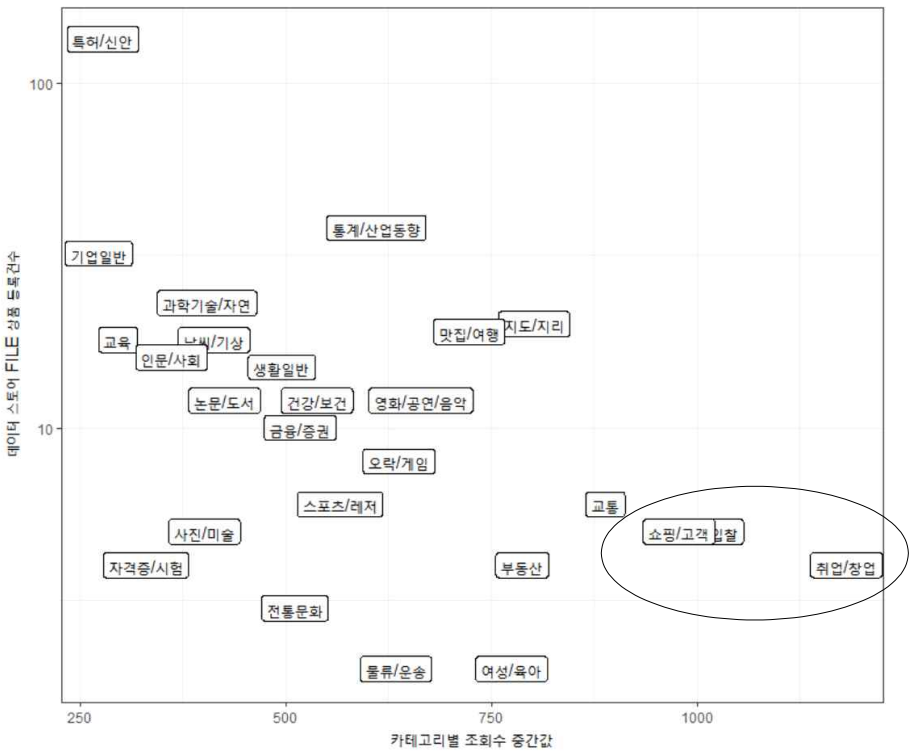
순위	전체	FILE	API
1	특허/신안	특허/신안	특허/신안
2	기업일반	통계/산업동향	지도/지리
3	통계/산업동향	기업일반	기업일반
4	지도/지리	과학기술/자연	생활일반
5	과학기술/자연	지도/지리	날씨/기상
6	날씨/기상	맛집/여행	금융/증권
7	생활일반	교육	통계/산업동향
8	맛집/여행	날씨/기상	과학기술/자연
9	교육	인문/사회	영화/공연/음악
10	인문/사회	생활일반	맛집/여행

○ 상품 유형별 조회 수가 높게 나타난 FILE 상품은 취업/창업, 경매/입찰, 쇼핑/고객으로 나타남

<표 4-3> 데이터스토어 상품
조회 수 주제별 순위

순위	주제
1	취업/창업
2	경매/입찰
3	쇼핑/고객
4	교통
5	지도/지리
6	부동산
7	여성/육아
8	맛집/여행
9	영화/공연/음악
10	오락/게임

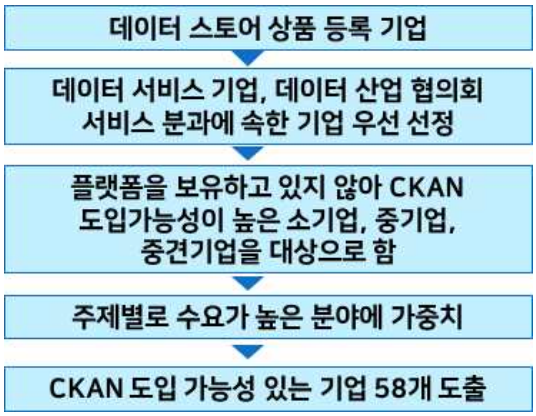
○ 데이터스토어(www.datastore.or.kr)의 주제별 상품 등록건수를 카테고리별 조회 수 중간값과 비교해보면 취업/창업, 경매/입찰, 쇼핑/고객과 같이 조회 수가 많은 카테고리의 데이터셋은 수요가 많은 반면 등록 건수가 많지 않은 것으로 파악됨



<그림 4-12> 데이터스토어 상품 등록건수 및 카테고리 별 조회수 비교

2. 연계 가능한 민간 데이터 현황

○ 연계 대상 기업 도출을 위해 데이터스토어에 등록된 민간데이터 549개에 대하여 데이터 상품을 등록한 기업 127개를 대상으로 58개 기업을 도출



<그림 4-13> 연계대상 기업 도출 절차

- 다양한 데이터 상품을 보유하고 있는 데이터스토어(www.datastore.or.kr) 상품 등록 건수 상위 목록 기업을 우선 선정하였고, 데이터 상품 개발에 대한 관심이 높은 기업군인 데이터 서비스 기업에 가중치 반영
- 자체 데이터 플랫폼을 보유하고 있지 않아, CKAN 도입이 필요할 가능성이 비교적 큰 소기업, 중기업, 중견기업을 대상으로 함
- 해외 데이터 서비스 산업별 투자현황과 해외 주요 데이터 포털 주제 분야 및 국내 데이터 포털 주제 분야를 참조, 수요가 높을 가능성이 있는 분야에 가중치 반영하고, 선별된 기업의 데이터스토어 상품 건 당 조회 수를 참조

<표 4-4> CKAN 도입이 필요한 업종별 기업 현황

업종	기업 수
출판, 영상, 방송통신 및 정보서비스업	37
전문, 과학 및 기술 서비스업	7
도매 및 소매업	4
공공행정, 국방 및 사회보장 행정	2
제조업	2
협회 및 단체, 수리 및 기타 개인 서비스업	2
교육 서비스업	1
사업시설관리 및 사업지원 서비스업	1
기타	2
합계	58

제5장 공공·민간 데이터 연계 활성화 방안

제1절 데이터 연계 필요성 및 장애요인

1. 데이터 연계의 필요성

○ 세계 40여개 이상의 국가에서 공공데이터를 통한 비즈니스 창출의 1차적 목표 달성을 위해 공공데이터를 국민들에게 공개하는 오픈데이터 정책을 시행[85]하고 오픈데이터 포털을 운영

- 오픈 데이터 정책은 주로 공공데이터를 대상으로 하나, 민간데이터를 융·복합하여 분석해야 보다 정교한 통찰력(insight)을 얻을 수 있음

○ 국가·공공기관 중심의 공공데이터 포털 개설 증가와 함께, 민간 기업들의 데이터 연계 노력도 점차 확대 중

- 데이터 포털들은 공공데이터 뿐만 아니라 민간데이터도 꾸준히 연계하고 있으며, 공공·민간데이터의 경계가 허물어지는 현상도 발생
- 뿐만 아니라, 최근 기업들은 오픈 이노베이션의 중요성을 인식, 오픈비즈니스 실천을 위해 오픈데이터를 적극 활용한 수익창출 비즈니스모델 설계를 위해 노력
- 기업들의 API를 통한 데이터 활용은 상당 부분 확산되고 있으며, 글로벌 플랫폼 기업들도 점차 데이터 포털을 개설하는 등 플랫폼을 확장시키려는 시도가 계속되고 있음

○ 오픈데이터를 공공 및 민간의 다양한 비즈니스 분야에 활용하기 위해서는 공공데이터와 민간데이터의 연계 분석이 필수

- 오픈데이터 분석에 활용되는 데이터 유형은 공공데이터, 기업데이터, 소셜데이터가 있음
- 연계 가능 민간데이터는 오픈데이터, 사업데이터, SNS등의 데이터, 기업의 사업과 관련된 비밀을 제외한 일반 사업 데이터 등[86]이 있음

---

85) 이재성, 박한우 (2015), “미국 기업의 오픈데이터 활용사례와 비즈니스 네트워크 분석 : 데이터와 기술 분야를 중심으로”, 「Journal of the Korean Data Analysis Society」, 17(5):2525-2536.

86) 한국데이터진흥원 (2017), “민간데이터 연계와 CKAN의 활용”, 「Data Issue Report」, 2017-02(106).

○ 데이터는 다른 데이터와의 연계와 결합을 통해 보다 가치 있는 통찰력을 도출할 수 있고, 정부의 공공데이터 개방 노력과 부합하여 4차 산업혁명 및 지능정보사회로의 진입으로 데이터가 의사결정의 중요 요소로 활용됨

- 이러한 환경으로 기업의 생존을 위한 새로운 비즈니스 가치 창출을 위해 공공 및 민간의 다양한 분야의 데이터 간 상호 연계 필요성 증대
- 데이터 중심 지능정보사회의 핵심 자원인 데이터 비즈니스의 순환형 생태계 조성을 위해서는 공공데이터와 민간데이터의 연계를 통한 새로운 데이터 자원의 생성 및 유통 필수
- 데이터는 한 사람이 소비를 해도 다른 사람이 소비를 못하게 되지 않는 '비경합성(non-rivalry)'으로 공공재적 성격이 강함

<표 5-1> 데이터의 공공재적 성격

배제성 \ 경합성	있음	없음
있음	Private Goods - 옷, 휴대폰, 자동차 식품	Club Goods(자연독점재) 수도, 전기, 철도, 이동통신 **민간데이터**
없음	Common Resources 바다 속 물고기, 환경	Public Goods 무료 공원, 국방, 치안 **공공데이터**

※ 출처 : 한국데이터진흥원, (2017),「Data Issue Report 2017-02 제106호」

○ 플랫폼 사례에서 보는 바와 같이, 기업의 이미지를 높이는 마케팅 수단으로서의 활용이 가능하며, 기업의 신규 인력 고용 창출이 가능함

- 데이터 연계 분석 사례에서 보는 바와 같이, 데이터 연계를 통하여 정부는 다양한 가치 높은 통계 작성으로 정책 수립에 도움이 되며, 민간은 고객 유치, 신규 사업 발굴 및 전략 수립에 도움이 될 수 있음

○ 민간데이터는 기업 또는 개인의 시간과 노력이 투입된 무형자산으로 무조건적인 공개를 강요할 수 없으나, 일정수준의 공개를 통해 활용의 폭을 넓혀야 시너지를 발휘할 수 있을 것임

2. 민간데이터 연계 장애요인

○ 공급자 중심의 데이터 유통 구조 등 데이터 유통 체계상의 다양한 문제점으로 인해 활성화에 어려움이 존재[87)]

- 공급자 중심의 데이터 생산 및 유통으로 공급과 수요 간 불일치가 발생하고 있으며, 사용자 중심 분류체계 미비로 국내 API 유통의 13%가 '기타' 범주로 지정되어 있는 상황
- API/데이터 유통 라이선스 체계 미비로 공급자와 소비자간 분쟁의 소지가 있으며, 데이터의 범주와 방식, 포맷 등의 표준화가 되어있지 않아 합리적 가격정책 수립 어려움
- 이러한 사유로 인한 제한된 API/데이터의 가격 및 이용 정책으로 데이터의 연계/유통을 원하는 다양한 유형의 소비자 요구가 충족되기 어려움

○ 데이터 연계 환경이 제대로 갖춰지지 않아 데이터 공급자와 수요자의 참여 비용이 높은 편

- 2장에서 살펴본 국내 데이터 포털 및 데이터 마켓 상세조사 현황에서와 같이 공공, 민간의 데이터 포털 및 마켓이 자체 플랫폼으로 구축되어 있는 비중이 높음
- 이로 인해 플랫폼 별로 기능과 사용 방법, 특징이 상이하여, 일반 비전문가를 위한 데이터 연계 및 유통 플랫폼 기반 활용 환경이 상대적으로 미비

○ 민간과 공공의 분류 간 매칭, 민간데이터와 공공데이터 간 표준화 문제의 해결

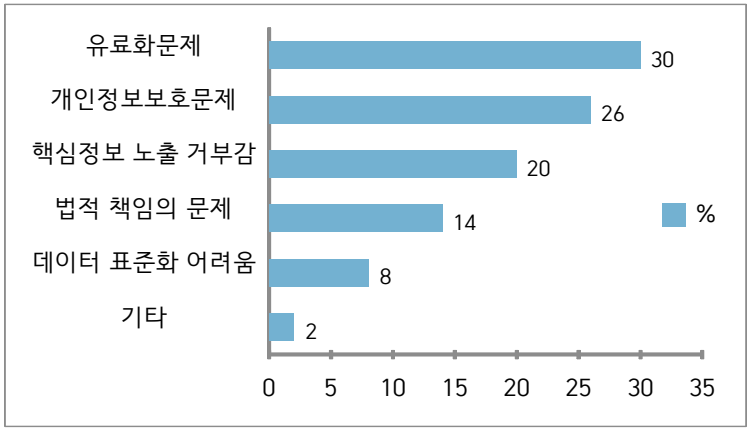
- 2장에서 살펴본 국내외 데이터 포털 현황, 국내 주요 공공/민간 데이터 포털 및 마켓에서 제공하는 데이터셋의 주제 분류에서 보듯, 공공과 민간의 데이터 주제 분류에 있어 많은 차이가 존재
- 전문가의 의견에 따르면 민간에서 활용하고자 하는 데이터를 찾는 데는 큰 어려움이 없으나, 민간과 공공에서 사용하는 분류의 차이로 데이터의 통합이 어려움
- 공공·민간 데이터의 효율적 통합(주제별, 지역별 등)을 위한 분류의 표준화가 필요

○ 데이터의 자산적 가치에 대한 인식 부족

- 민간데이터의 경우, 기업 또는 개인이 시간과 노력을 투입한 자산으로서 자산적 가치에 대한 대가 지불이 필요[88)]

87) 이규철 (2015), 「오픈 API기반 데이터 유통 활성화 방안」, 대전: 충남대학교.

- 이는 실제 조사 결과에서도 지적되고 있으며, 공공데이터 개방 수요조사 결과[89]에 포함되어 있는 “민간데이터 개방 시 예상되는 장애요인”으로 유료화 문제, 개인정보보호문제, 핵심정보노출에 대한 거부감, 법적 책임문제 순(順) 으로 조사됨



<그림 5-1> 민간 데이터 개방 시 예상되는 장애요인

※ 출처 : 월드리서치, (2015), 공공데이터 개방 수요조사 58page 차트 재구성

○ 데이터 개방과 연결의 중요성에 대한 사회적 인식이 부족하여 공공기관과 민간기업, 전 국민(시민)을 대상으로 한 사회적 공감대 형성이 필요

- 도시 최초로 CTO(최고 데이터 관리자)를 도입했던 미국 시카고 시 사례를 보면 도시의 공공데이터 오픈과 연계 활용에 앞서 사회적 인식의 개선을 위해 저널리즘을 활용하여 지속적으로 사회적 인식을 개선하여 시민들에게 정책에 대한 호응을 얻을 수 있었음[90]

제2절 공공·민간 데이터 연계 활성화 제언

1. 민간 데이터 연계 활성화를 위한 개선 방안

가. CKAN을 적용한 데이터 연계 환경 조성

○ CKAN 적용을 통해 공공·민간 데이터 연계에 필요한 환경 제공 및 사용자 편의성 증대

88) 한국데이터진흥원 (2017), “민간데이터 연계와 CKAN의 활용”, 「Data Issue Report」, 2017-02(106), 서울: 한국데이터진흥원.
89) 월드리서치 (2015), 「정부 3.0 공공데이터 개방 설문조사 연구용역 결과보고서」, 서울: 월드리서치
90) Brett Goldstein, Lauren Dyson (2013), 「Beyond Transparency: Open Data and the Future of Civic Innovation Open Data」, Code for America Press.

- 직접적인 데이터 게시 및 웹 인터페이스를 통해 게시하는 것을 사용자가 편리하게 수행 가능
- 이를 통해 데이터 보유 기업의 데이터 포털로의 접근이 쉬워져 데이터 연계가 더욱 활성화될 수 있는 기반 마련 가능
- 또한 CKAN 적용 시 현재의 데이터 정보 목록을 제공하는 것을 넘어 실제 데이터를 연계할 수 있는 기능 제공이 가능
- CKAN의 라이선스 정책 적용 기능 등을 통해 가격 산출 및 운영 정책에 대해 사용자 편의 기능 제공

○ CKAN의 데이터 카탈로그(DCAT[91]:Data Catalog Vocabulary)의 확장 및 메타데이터 표준 적용을 통해 플랫폼 간 상호 연계성을 강화
- DCAT Catalog는 RDF 어휘에 정의된 용어로 데이터셋에 대한 메타 데이터 모음
- DCAT은 데이터 카탈로그의 상호 운용성을 향상시켜 응용 프로그램이 여러 카탈로그의 메타데이터를 쉽게 사용할 수 있게 함
- DCAT 메타 정보를 통해 제목, 설명, 수집 방법, 저자 또는 발행자, 다뤄지는 지역 및 시간대, 라이선스, 발행일자, 주기 등의 정보를 획득
- 오픈 데이터 플랫폼에서 제공하는 데이터 셋 뿐만 아니라, 타 시스템에서 제공하는 오픈데이터에 대한 메타 정보 제공
- 이로 인해, 데이터 소비자는 카탈로그만 검색만으로 원하는 데이터의 기본 정보 및 위치를 쉽게 검색
- 유럽(DCAT-AP), 네덜란드(DCAT-NL), 이탈리아(DCAT-AP-IT)등의 해외 사례에서와 같이 우리나라 실정에 맞는 한국형 DCAT 확장 표준 개발 및 적용

나. 데이터 연계 표준화 및 거버넌스 체계 마련

○ 표준 분류 체계 제공을 위하여 빅데이터 거버넌스 체계 적용
- 데이터 수집 확보, 기술 요건 검토와는 별도로 데이터 공개 기준 및 교환 방식 표준을 마련하여 데이터 연계 활성화를 위한 선행 체계 수립이 필요함
- 다양한 리소스들과 데이터 유통을 위해서는 데이터 관리를 위한 체계적인 거버넌스 도입이 중요함
- 데이터 제공 기관 담당자가 공개한 정보를 관리할 수 있는 운영 관리 환경을 제공함

91) DCAT(Data Catalog Vocabulary) : 웹에 게시 된 데이터 카탈로그 간의 상호 운용성을 용이하게 하기 위해 설계된 RDF Vocabulary

○ 데이터의 연계의 효율성 재고를 위해 데이터 연계 방식에 대한 기술표준 개발 및 배포 필요

- 데이터의 공개 및 연계에 관한 기술적 표준 개발을 통해 공공 및 민간 기관 간 상호 연계 효율성 재고
- 이를 통해 민간 기업의 데이터 개방 및 연계 관련 기술적 장벽을 낮추어 참여 확대 유도

○ 데이터 개방에 따른 표준 가이드 뿐 아니라 민간 데이터 연계 활용을 위한 가이드라인 제공이 필요

- 해외 가이드라인은 공개된 데이터의 이용, 활용 및 활성화를 위한 다양한 솔루션 제공 중심
- 국내 공공데이터 개방 표준은 정부 기관이 공공데이터 개방 시 따라야 할 표준 가이드 제시 중심
- 데이터의 연계 활용이 활성화 되어있는 해외 선진사례와 같이 공공 데이터의 개방을 위한 표준 제시와 함께 이용, 활용 및 활성화를 위한 솔루션의 제공도 중요

다. 데이터 연계를 위한 법률 문제 해소 방안 마련

○ 데이터 개방 및 연계활용 시 발생할 수 있는 저작권·개인정보보호 등의 문제 해소를 위해 ‘클리어링 하우스’ 운영

- 데이터의 연계 활용 관련 불명확한 법리 해석 등의 문제로 공공 및 민간 기관의 소극적 참여
- 서로 다른 인증 및 연계 방식을 가진 공공·민간 기관 간 데이터의 연계를 위한 통합 서비스 개발 및 전담 운영기관을 설립
- 이를 통해 법적 문제 해소 및 연계를 위한 비용 최소화함

○ 앞서 언급된 통계청 사례[92]에서와 같이 정부가 보유하고 있는 개인 정보를 민간과 연계 시 익명화된 데이터셋으로 변환하는 데이터 연계방식을 제시

- 개인정보 비식별화 처리 등 기존 관련 법·제도 내용을 포함한 실제 공공·민간 데이터의 융복합을 위한 연계 절차, 데이터 제공 방법, 결과물 제공방법 등에 대한 가이드 개발

92) 통계청 보도자료 (2016.08.25), “공공·민간 빅데이터 연계 현황 및 활성화 방안 - 신혼부부 통계를 중심으로 “

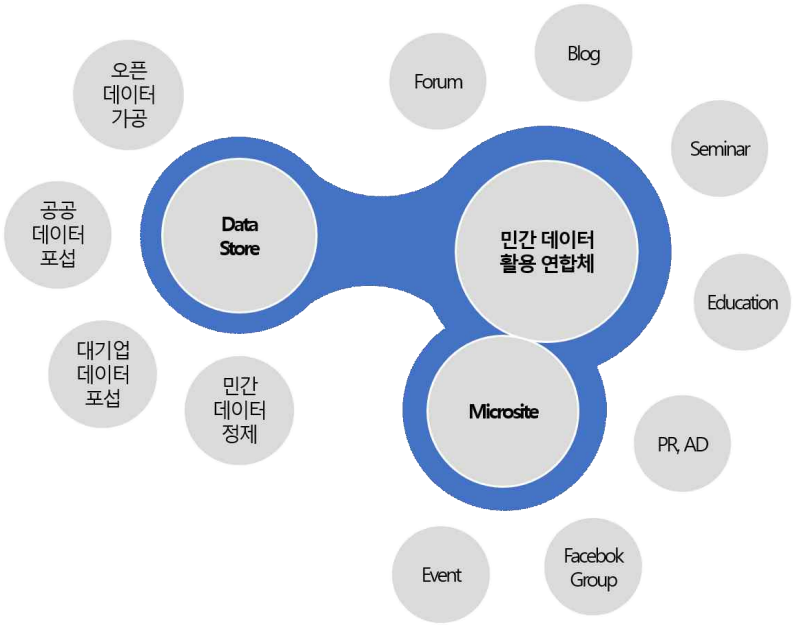
라. 데이터 연계 플랫폼 '데이터스토어' 역할 확대 및 홍보 강화

○ 공공·민간 데이터 연계 및 활성화를 위한 데이터스토어의 역할 확대
- 공공기관이 필요한 민간 데이터를 데이터스토어를 통해 구매 및 연계
- 민간 기업이 원하는 공공 데이터는 CKAN이 적용된 데이터스토어를 통해 타 공공기관에서 공개한 데이터를 연계 제공
- 데이터 유통에 걸림돌이 되는 저작권, 개인정보보호 등의 법·제도적 문제의 해소를 위해 '클리어링 하우스[93]' 역할 수행
- 이를 통해 공공기관과 민간 기업 사이의 데이터 가교 역할 수행

○ 4차 산업혁명은 데이터의 연결을 통해 완성되며, 데이터 개방과 연결이 새로운 비즈니스의 시작이라는 인식의 전환이 핵심 요소
- 데이터 개방과 연결은 국가 경쟁력의 상징이며, 데이터 융복합과 연결의 중요함을 홍보하고 데이터 개방의 대중화를 선언하여 잠재적 데이터스토어 및 데이터 마켓플레이스의 활성화를 도모
- "Connect DATA, Powered by DATA, Open & Success"를 핵심 키워드로 함

○ 데이터 융복합 홍보 마케팅 플랫폼 구축하여 민간 데이터 연계 활성화 및 홍보에 활용
- 연합체의 마이크로사이트(홍보용 웹페이지) 및 Facebook, youtube, blog 등을 활용한 SNS 플랫폼을 구축하여 홍보에 활용

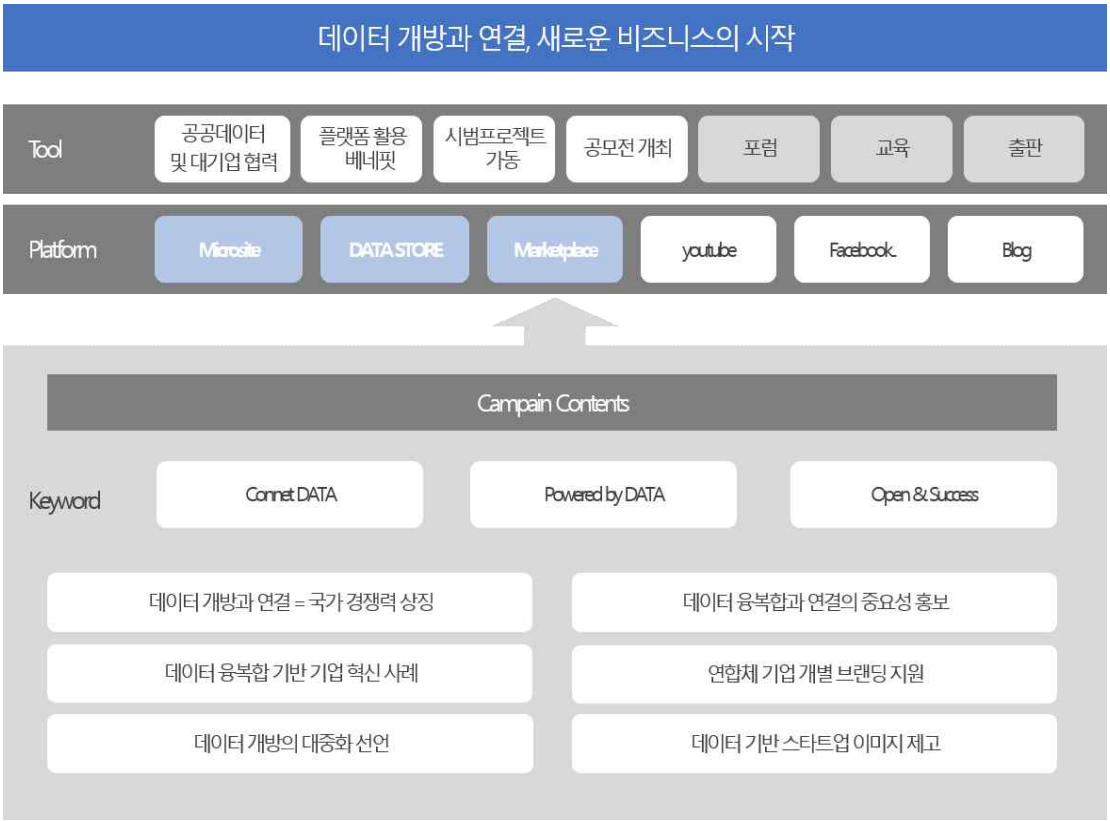
93) 클리어링 하우스 : 서로 다른 시스템이나 서비스를 연결하여 통합 서비스를 지원할 때 서로 다른 인증 방식이나 과금 방식을 처리하고 정산하는 서비스 혹은 정산소



<그림 5-2> 민간데이터 연계 활성을 위한
온/오프라인 마케팅 플랫폼

○ 데이터 개방과 연결 활성화를 위한 통합 마케팅 커뮤니케이션 전략을 진행

- 민간 데이터 연합체 중심의 지속적인 콘텐츠 생산하고 대기업 및 공공기관과의 협력을 통한 성공사례를 발굴
- 온/오프라인 마케팅 플랫폼, 채널, 홍보, 제작물, 프로모션의 통합 캠페인을 진행하고 마이크로사이트 중심으로 데이터스토어 및 마켓플레이스 활용
- 민간 데이터 개방과 융복합 사례 기반의 국내외 사례 공유 정기 세미나 및 데이터 융복합 공모전 개최
- 컨퍼런스 홍보 및 기술 교육 동영상 배포하고 지속적인 기획기사 및 보도자료 배포, TV, 다큐멘터리, SNS 활동을 통해 홍보 활동
- 홍보 동영상 제작, 기획기사 및 보도자료 배포, TV 다큐멘터리, 팸플릿 및 동영상 자료 제공 및 각 프로모션별로 온라인 이벤트를 개최하고 동일하게 연속성을 유지하여 연결의 중요성 강조
- 포럼(facebook) 운영을 통해, 잠재 수요 및 공급자 유치, 홍보 플랫폼으로 활용



<그림 5-3> CKAN 민간 데이터 활성화 홍보 방안

마. 공공·민간데이터 활용을 위한 소통 채널 마련 및 상설화

○ 데이터 활용이 필요한 스타트업을 대상으로 민간 데이터 활용 연합체를 모집하여 운영

- 스타트업의 데이터 연계 유도 및 적정 가치·인센티브 수혜 방안 마련
- 연합의 모집 및 발기, 간사 및 운영주체를 민간에서 자발적으로 이루어지도록 지원하며 데이터스토어와 마켓 플레이스를 제공
- 연합체 기업 개별 브랜딩(기업 이미지 메이킹)을 지원하여, 데이터 기반 스타트업 이미지를 제고
- 비즈니스 교류 및 데이터 활용 방안 교류를 통해, 공공 및 대기업 데이터 활용 포섭 단체의 역할을 수행

○ 민간 데이터 연합체 주요 활동을 통해 사회적 인식 개선을 위한 홍보 활동과 민간 데이터 활용을 위한 공공 및 대기업과의 협력 활동을 수행

- 민간 데이터 융복합 온/오프라인 캠페인 실시, 데이터 융복합 시범 프로젝트 및

공모전을 개최

- 공공 데이터 접근성 및 데이터스토어를 중심으로 한 플랫폼 및 마켓 플레이스 활용 프로모션을 통한 컨텐츠 확보
- 공공기관 및 대기업 협력 활동을 통한 오픈 데이터 가공, 민간 데이터 정제, 공공 데이터 활용률 증진을 위한 요구사항 발의
- 데이터 경제 활성화를 위한 블로그 및 포럼 운영, 컨퍼런스 및 세미나를 개최, 각종 보도자료 배포 및 기획 기사를 통한 홍보

○ 민간 데이터 활용 연합체 활동으로 데이터 융복합의 물리적 플랫폼 역할 및 민간 주도 데이터 융복합 시장 창출 효과와 회원사의 기존 비즈니스 분석 및 신규 모델 개발 효과의 기대가 가능

○ 공공기관과 민간 기업이 공동 참여하는 데이터 공개 및 연계 활성화를 위한 소통 채널 상시 운영

- 해당 단체를 통해 데이터의 공급자와 수요자 양방향 소통의 장치 조성을 통한 데이터 유통 및 활용 활성화 방안 제시 및 정책 수립 지원
- 다양한 홍보, 마케팅을 통한 데이터 개방, 연계 및 활용 관련 중요성 및 데이터의 자산 가치에 대한 인식 재고
- 데이터의 개방 및 유통 활성화 정도가 낮은 국내 환경을 고려 민간 유통 플랫폼 간 상호 연계 체계 구축

○ 수요자가 원하는 데이터를 수집·가공하여 제공하거나 데이터의 소재·활용방법 등을 추천하는 전문 중개 사업자인, 데이터 브로커(Data Broker) 육성

- 민간 및 개별 기업이 보유하고 있는 다양한 데이터 포털을 활성화 하고 상호 연계 추진
- 이를 통해 데이터 브로커, 데이터 포털, 데이터 마켓 간 데이터 연계 네트워크 형성 추진
- 민간 기관 대상 연계 플랫폼 또는 확장 모듈 무상 지원으로 데이터 연계 및 샘플 데이터 공개 참여 확대
- 국내 데이터 산업 환경 특성 상 공공이 주관이 되어 데이터스토어와 같은 공공 포털/마켓을 중심으로 데이터 연계 네트워크 활성화를 통한 데이터 유통 시장 활성화를 위해 주도적, 지속적 노력 필요

2. CKAN 적용 확대 방안

○ 오픈 데이터 실행력을 증진시키기 위해서는 체계적인 데이터 관리, 공유, 연계를 위해 표준화를 지원하고 데이터 공개시 기계 판독이 가능한 원천데이터를 개방하고 관리할 수 있는 오픈 플랫폼 필요

- CKAN은 영국에서 공공데이터 개방을 목적으로 개발하여 오픈소스로 공개한 바가 있음
- 영국, 미국 등 데이터 연계 활용이 활성화된 선진 국가의 다양한 기관에서 데이터 개방을 위해 오픈소스 또는 상용 CKAN 계열의 데이터 연계 플랫폼 사용을 통해 활용 우수성 검증

○ ODB 지표에서 보여주듯 개방도 측면이 취약한 우리나라 상황에서 CKAN의 도입은 오픈 데이터 실행력을 높일 수 있는 효과적인 접근법이며, 장기적인 관점에서 보면 CKAN 기반의 한국형 데이터 연계 플랫폼의 개발이 필요

○ 단계적 CKAN 적용 로드맵에 의한 CKAN 확장을 통해 민간 데이터 연계 기능 강화

- 1단계 CKAN 도입기에는 데이터 연계를 위한 필수적인 기능 위주 접근하여 CKAN을 통한 데이터 오픈 및 사용자 관리에 중점
- 2단계 CKAN 적용 활성화 시기에는 연계를 위해 중요한 메타 데이터 관리 및 DCAT 적용에 중점을 두어 민간 또는 타 데이터 오픈 플랫폼과 1차원적인 접근 및 연계를 지원
- 3단계 CKAN 적용 성숙기에는 자원 관리를 비롯해 보다 원활한 데이터 연계를 위한 데이터 맵 제공하며 CKAN 확장 부분과 모니터링 부분, 비즈니스 영역과 API 및 사용자 인증을 위한 게이트웨이 모듈 형태로 구성

<그림 5-4> 단계별 CKAN 적용 로드맵

○ 유럽 및 캐나다의 공공분야 데이터포털 사례와 같이 중앙정부, 지자체 및 산하 기관이 CKAN과 같은 연계 오픈플랫폼을 활용하여 데이터 사용자가 각 조직별 데이터를 하나의 포털에서 쉽게 접근하여 사용할 수 있도록 구성

○ 데이터의 활용 요구가 높은 민간기업 대상 데이터 연계를 위한 오픈 플랫폼(CKAN) 제공 및 적용 기술 지원

- 데이터 연계를 위한 확장 기능(Extension)개발 시 관련 기술 지원
- 표준화된 오픈 플랫폼 기반 수요자가 활용 가능한 데이터 분석환경, 개인 및 중소기업 등 수요자가 제품개발 등을 위해 활용할 수 있는 데이터 활용환경 제공
- CKAN기반의 데이터스토어와 민간기관의 데이터분석 및 시각화 플랫폼과의 연계를 통해 데이터 간 융복합, 분석 및 시각화 등의 다양한 기능 추가 제공
- 이러한 기능을 활용, 민간 기업의 내·외부 데이터 융복합 활동 활성화를 통한 샘플 데이터 공개 확대 유도

○ CKAN을 기반으로 한 데이터스토어 기능을 제공하는 지원 상품 활용으로 연계를 유도

- CKAN기반 데이터스토어 플랫폼 기능 패키징을 통해 데이터 활용 중심 중소기업(스타트업 포함) 지원 상품으로 활용

- 커스터마이징 및 확장성이 뛰어난 장점을 활용 민간 데이터 기업의 CKAN 적용 지원 및 데이터스토어와 연계 유도

○ CKAN의 확장성은 데이터 연계 활용 플랫폼으로서의 기능을 손쉽게 제공하는 환경을 제공함

- CKAN Extension 개발 및 Drupal과 같은 타 플랫폼과의 결합을 통해 데이터 융복합, 분석, 시각화 등 데이터 연계 및 활용 기능 강화
- 사용자가 데이터 연계를 통해 얻고자 하는 다양한 인사이트 도출에 대한 니즈를 만족시킬 수 있는 환경 구성

제6장 결 론

○ ‘정부 3.0’ 공공데이터 오픈 정책 등으로 정부는 공공데이터를 체계적으로 분류하고 공개하려는 노력을 지속적으로 하고 있으며, 민간 기업들도 4차 산업혁명으로 데이터를 기업 비즈니스의 중심에 두고, 오픈데이터와의 연계를 통해 새로운 가치 창출 기회를 얻고자 노력하고 있음

- ODB 지표에 의하면 우리나라는 데이터 개방에 대한 준비도와 활용도 측면은 상당히 우수하나 데이터 연계 및 이행과 관련된 개방도 측면은 상대적으로 취약함
- 이런 부분을 보완하기 위해 데이터 간 연계 및 활용을 위한 전략이 필요

○ 국내외 데이터 포털 활동 현황과 데이터 플랫폼 사용 현황을 조사한 결과 대륙별로는 북아메리카, 유럽이, 국가별로는 미국, 캐나다, 프랑스, 영국이 데이터 포털이 많이 운영되고 있음

- 국내외 선진국 대표 데이터 포털들이 제공하는 기능에서, 데이터를 설명하는 메타데이터를 등록, 관리하여 검색하고 활용할 수 있는 데이터 카탈로그(목록)는 대부분 제공
- 데이터 연계에 유리한 데이터 플랫폼을 사용하는 경우가 대부분이며 열람, 다운로드는 대부분 허용하고, 융합 가공되는 경우가 일부 있음
- 데이터 플랫폼으로는 오픈 소스 기반의 개방형 데이터 플랫폼 CKAN이 가장 많이 사용되고 있음

○ 국내 데이터 포털 및 데이터 마켓 현황에서 공공데이터 포털, 데이터스토어 등 중앙정부 및 기관 뿐 아니라, 서울시 ,수원시, 경기도청 등 지방정부에서도 데이터 연계에 유리한 CKAN을 데이터 포털에 적용하고자 하는 시도가 있어옴

- 국내외 선진국의 대표 데이터 포털들은 대상이 대부분 공공데이터 포털이었으므로, 데이터의 융합 및 분석 시각화 기능을 제공하는 경우는 거의 없었으나 국내 데이터 마켓의 경우는 대부분 이를 제공하고 있음
- 데이터 융복합을 통해 가치를 발견하고자 하는 데이터 마켓에서는 연계 뿐 아니라, 분석 및 시각화 기능을 필요로 하고 있음을 유추할 수 있음

○ 오픈 데이터 관련 기구들은 CKAN 등 개방형 데이터 솔루션 및 도구 개발, 표준 가이드, 방법론 개발 및 보고서 제작, 지역 시민의 데이터 생성 활동 지원, 정부 민간 연결 파트너의 역할을 해왔음

○ CKAN은 영국 정부가 공공데이터 개방을 위해 오픈 플랫폼으로 개발되어 Open Knowledge International을 통해 무료 배포 하고 있으며, 전 세계 국가에서 데이터 개방을 위한 플랫폼으로 CKAN을 활용하고 있음

- 특히 미국 정부와 같이 Socrata와 같은 자국 내 민간 플랫폼에서 데이터 연계 및 확장성이 우수한 CKAN으로 전환하는 사례가 발생하는 등 데이터 연계 플랫폼으로서의 CKAN의 우수성이 전 세계적으로 인정받고 있음
- CKAN 주요기능으로는 데이터셋 배포, 저장, 시각화, 연계, 검색 및 활용, 확장 기능(Extension), 인증, 이력 및 사용 통계, 커뮤니케이션의 기능이 있으며 데이터 공급자, 수요자 측면에서의 기능을 제공하고 있음
- CKAN은 DCAT 표준을 통한 데이터 연계 유연성을 갖는 메타 데이터 관리 체계를 가지고 있으며, 기술 발전 동향에서 연계를 위한 데이터 카달로그로 시작해서 앞으로 데이터 저장소로서의 기능을 포함한 데이터 연계 플랫폼의 방향으로 발전됨

○ 해외 농업 정보 활용사례, 국내 공공 기관와 금융 기관의 데이터 연계 분석 사례와 같이 정부의 공공데이터를 보완할 만한 민간 기업의 데이터를 연계, 분석하여 공공에게는 데이터 기반 정책 수립에 도움이 되며, 민간에게는 고객 유치, 신규 사업 아이템 발굴 및 전략 수립에 도움이 되고 있음

- 민간 기업에서도 농업분야의 빅데이터 기업 몬산토에서 CKAN을 기반으로 기업, 고객, 공공 데이터의 데이터 연계 및 플랫폼 구축 사례를 찾아볼 수 있음
- Link Digital, Viderum 같은 CKAN기반의 플랫폼 구축 및 기술 지원, 데이터 활용 컨설팅의 비즈니스 모델을 가지는 기업들의 등장함
- 기업 서비스와 연계가 가능하도록 기업 주체로 40여개의 CKAN 확장 기능(Extension)이 개발되고 있는 등 기업에서 CKAN을 적용하기 위한 노력이 진행되고 있음
- 구글, 네이버와 같은 민간 기업들이 데이터 플랫폼을 통해 자사 데이터, 공공 데이터, 고객 데이터를 연계하여 활용 가능하도록 하면서, 이를 유상 또는 무상으로 공개하고 있음

○ 해외 공공데이터 포털을 통해 환경, 건강, 기후, 농업 등의 분야의 데이터셋이 높은 비율로 공개 되고 있으며, 해외 데이터 서비스 관련 스타트업 투자현황을 보면 마케팅, 지리정보, 소셜 데이터, 기상, 교통, 법률, 스포츠, 헬스 등의 분야가 높은 비중을 차지하고 있으며, 이 중 공공데이터에서 제공되고 있지 않은 마케팅, 소셜데이

터, 스포츠, 헬스 등 민간 데이터에 해당하는 분야는 데이터 서비스 및 데이터 연계를 통한 비즈니스 창출을 통해 스타트업에 있어 기회가 될 수 있음

○ 국내 주요 공공/민간 데이터 포털 및 데이터 마켓에서 제공하는 데이터셋에 대하여 제공하는 주제 분야를 비교해 보았을 때, 데이터 마켓에서는 공공 데이터 포털과는 달리 영화/공연/음악, 쇼핑/고객, 맛집/여행, 쇼핑, 채팅, 메시지, 금융, 산업동향 등 민간 기업 또는 개인이 생산하는 민간데이터에 속하는 주제가 많은 것을 확인할 수 있음

○ CKAN의 적용 가능성이 높은 기업을 선별하고자 연계 대상 민간 데이터를 보유하고 있는 기업의 데이터 보유현황을 알아본 결과, 국내의 공공기관에서 운영하는 데이터스토어에서 제공하고 있는 데이터셋의 주제 분야를 살펴보면, 수요가 높은 주제 분야는, 건당 조회 수가 높은 취업/ 창업 분야와 공개된 데이터 건수가 많으면서 조회 수도 높은 분야인 지도/지리, 맛집/여행, 영화/공영/음악, 통계/산업동향, 생활이반, 날씨/기상, 과학기술/자연으로 나타남

○ 상품 등록 건수의 상위를 차지하고 있는 주제가 특허/신안, 기업일반, 통계/산업동향, 지도/지리, 과학기술/자연, 날씨/기상 등이 공통적이며, 파일 유형별로 파일로 제공 및 거래 되고 있는 데이터의 경우 통계/산업동향, 과학기술/자연의 주제가 상위권인데 비하여 API로 제공하는 데이터의 경우 지도/지리, 날씨/기상, 금융/증권 분야가 상위권에 위치하며, 취업/창업, 경매/입찰, 쇼핑/고객과 같은 조회 수가 많은 카테고리의 데이터셋은 수요가 많지만 등록 건수가 많지 않음을 알 수 있음

○ 데이터스토어에 등록된 민간 데이터 상품을 등록한 기업을 대상으로 데이터 서비스 및 연계와 관련 여부 및 플랫폼을 보유하고 있지 않아 CKAN 도입이 필요한 소기업, 중기업, 중견기업을 여부, 주제별로 수요가 높은 분야에 가중치를 두어 CKAN 도입이 필요한 기업을 도출함

○ 데이터는 다른 데이터와의 연계와 결합을 통해 보다 가치 있는 통찰력을 도출 할 수 있고 4차 산업혁명과 지능 정보 사회의 진입으로 데이터가 의사결정의 중요 요소로 활용되고 있으며, 새로운 비즈니스 창출을 위해 공공 및 민간분야의 다양한 데이터의 상호 연계 필요성이 증대되고 있음

○ 공공·민간 데이터 연계 활성화를 위하여, 오픈 소스 기반의 데이터 연계 플랫폼의 단계적 적용을 확대할 필요가 있으며 이를 위해, 연계 범위를 지방 정부 및 공공기관 대상을 확대하고, 민간 기업의 데이터 개방 참여를 확대 시킬 필요가 있음

- 지방 정부 및 공공기관 대상 확대를 위하여 중앙 정부, 지자체 및 산하기관이 연계 될 수 있는 CKAN과 같은 오픈 플랫폼을 활용하여 각 조직별 데이터를 쉽게 접근 할 수 있도록 함
- 공공·민간의 연계 및 활성화를 위하여 데이터스토어가 데이터 개방 및 연계 활용 시 발생할 수 있는 저작권, 개인정보 보호 등 법 제도적 문제를 해소하는 클리어링 하우스의 역할을 수행
- 민간 기업의 데이터 개방 참여 확대를 위해 수요자가 원하는 데이터를 수집 가공하여 제공하고 데이터의 소재 및 활용방법을 추천하는 전문 중개자 (데이터 브로커)를 육성하고 데이터 연계활용 서비스를 통하여 새로운 가치를 창출

○ 민간 데이터 연계에 대한 장애요인으로는 데이터 개방과 연결의 중요성에 대한 사회적 인식의 중요성과 민간과 공공의 분류 간 매칭 미 표준화 문제, 데이터 유통체계 및 구조의 문제, 그리고 표준화된 연계 및 유통 방식 지원 플랫폼의 미비를 들 수 있으며, 이를 해소하기 위하여 민간 데이터 연계 활성화를 위한 개선방안을 제시함

○ 사회적 공감대를 확산시키기 위하여, 4차 산업혁명은 데이터의 연결을 통해 완성되고, 데이터 개방과 연결이 새로운 비즈니스의 시작이라는 인식의 전환과 함께 스타트업 중심의 민간 데이터 활용 연합체를 구성하여 융복합의 물리적 플랫폼 역할 및 민간 주도 데이터 융복합 시장 창출 효과를 기대하고, 공공 기관 및 대기업 협력 활동을 통한 대기업 데이터 활용을 위한 포섭 단체로서 발의하고, 민간 데이터 융복합 온/오프라인 캠페인 및 CKAN 기술 교육과 사용자 교육을 실시

○ 표준화 문제 개선을 위하여 빅데이터 거버넌스 체계 적용을 통한 표준 분류 체계 제공 및 CKAN의 데이터 카달로그의 확장 및 메타 데이터 표준 적용을 통해 플랫폼 간 상호 연계성을 강화하고 데이터 연계의 효율성 제고를 위해 데이터 연계 방식에 대한 기술 표준 개발 및 배포

- 해외 가이드라인은 공개된 데이터의 활용 및 활성화를 위한 다양한 솔루션을 제공하고 있어, 공공 데이터 개방 시 따라야 할 표준 가이드 제시 중심의 국내 공공 데이터 개방 표준을 보았을 때 활용 및 활성화를 위한 솔루션의 제공도 중요

○ 공급자와 사용자가 함께 활용 가능한 유통 체계 및 구조를 개선하기 위하여 CKAN을 적용하여 데이터 게시를 위한 사용자 편의성을 증대하고 데이터 연계 활용 및 라이센스 정책을 통한 가격 산출 및 운영상 사용자 편의 기능 제공

○ 데이터 연계 및 유통 방식을 지원하는 플랫폼의 단계적 적용을 위하여, 1단계 CKAN 도입기를 통해 데이터 연계를 위한 필수적인 기능과 사용자 관리에 중점을 두고 2단계 CKAN 적용 활성화 단계에서는 연계를 위해 중요한 메타 데이터 관리 및 DCAT 적용에 중점을 두고 민간 또는 타 데이터 오픈 플랫폼과 1차원 적인 접근과 연계를 지원 3단계 CKAN 적용 성숙기에는 자원 관리를 비롯해 데이터 연계를 위한 데이터 맵을 제공하여 CKAN 적용 로드맵에 의한 CKAN 확장을 통해 민간 데이터 연계 기능을 강화함

참고문헌

과학기술정보통신부, 정보통신산업진흥원 (2013), 「2013 소프트웨어 산업 연간보고서」, 과학기술정보통신부, 정보통신산업진흥원.

박경현 외 (2017), “공공 빅데이터 개방과 표준화”, 「OSIA S&TR Journal」, 30(2): 21-25.

박주성 (2011.11.14.), “Open Government 2.0 - Case Study: Australia & United Kingdom”, 발표자료.

백인수 (2013), “오픈데이터 플랫폼과 국가 데이터 전략방향”, 「정보처리학회지」, 16.

월드리서치 (2015), 「정부 3.0 공공데이터 개방 설문조사 연구용역 결과보고서」

유재영 (2015), “해외 오픈데이터 정책 현황과 한국의 오픈데이터 정책 발전방안”, 한국정보화진흥원.

이규철 (2015), 「오픈 API기반 데이터 유통 활성화 방안」, 대전: 충남대학교.

이승환 (2015), 「지방자치단체 정책지원을 위한 데이터 분석 운영체계 연구」, 대구대학교 산학협력단, 행정안전부 수탁연구.

이재성, 박한우 (2015), “미국 기업의 오픈데이터 활용사례와 비즈니스 네트워크 분석 : 데이터와 기술 분야를 중심으로”, 「Journal of the Korean Data Analysis Society」, 17(5): 2525-2536.

이재원 (2016), “공공데이터 포털 국내외 동향 분석”, 「ICT Veiwer」, 한국 정보화 진흥원.

정유철 (2016), “오픈 데이터 플랫폼 동향”, 「정보처리학회지」, 23(5): 53-63.

한국데이터진흥원 (2017), “민간데이터 연계와 CKAN의 활용”, 「Data Issue Report」, 2017-02(106).

한국정보화진흥원 (2017), “오픈데이터 해외 동향 : Digital Economy Act 2017”, 「Global Open Data, Now 오픈데이터 해외주요동향」, 2017(2):5-6, 대구: 한국정보화진흥원.

한은영 (2014), “영국 오픈데이터 정책의 특징 및 시사점”, 「정보통신방송정책」, 26(23).

함유근 (2015), 「이것이 빅데이터 기업이다」, 삼성경제연구소, 193.

행정자치부 (2016), 「공공데이터 개방 표준」(행정자치부고시 제 2016-46호) 2014년도 10월 제정, 2016년도 12월 5차 개정, 행정자치부.

행정자치부 보도자료 (2017.05.29.), “대한민국, 공공데이터 개방 세계 5대 강국 우뚝”.

공공데이터 포털 세종 데이터 실록, https://data.sejong.go.kr/ 참조 (접속일: 2017.8.31.).

디지털데일리 (2016.04.22.), “강릉서 가장 인기 업종은? 네이버, 지역정보 빅데이터 공개”, http://www.ddaily.co.kr/news/article.html?no=142615 참조 (접속일: 2017.08.17.).

디지털데일리 (2016.07.20.), “네이버 데이터 랩, 쇼핑 빅데이터 공개”, http://www.ddaily.co.kr/cloud/news/article.html?no=145503 참조 (접속일: 2017.08.17.).

서울시 전자정부 (2015.1.23.), https://seoulsolution.kr/ko/content/시민의-알권리-행정정보-공개-자동시스템화 참조.

수원 공공데이터 포털, http://data.suwon.go.kr/pubr/data/viewDataSetSelfList.do?gubun=&mId=0101000000 참조 (접속일: 2017.8.31.).

신한카드블로그 (2016.2.23.), “신한카드・제주도・한국은행, 공동 빅데이터 협업 추진 “, http://www.shinhancardblog.com/202 (접속일: 2017.2.21.).

아이슬란드 농업 자문 센터(Icelandic Agricultural Advisory Center), http://www.porthonnun.is/en/portfolio/icelandic-agricultural-advisory-centre 참조 (접속일: 2017.08.17.).

IT DAILY (2017.04.26.), “솔트룩스, AI 기반 데이터 클라우드 공개…데이터 100억건 무료 개방”(접속일: 2017.05.24.)

조선일보 (2015.11.09.), “오바마의 '열린 정부'… 캐나다의 '서비스 캐나다'“, http://biz.chosun.com/site/data/html_dir/2015/11/08/2015110802578.html 참조 (접속일: 2017.08.21.).

한겨레 21 (2016.03.22.), “와글이 만난 몽상가들- 캐나다 ‘오픈노스’ 설립자 제임스 매키니”, http://h21.hani.co.kr/arti/society/society_general/41380.html 참조 (접속일: 2017.09.13.).

Brett Goldstein, Lauren Dyson (2013), 「Beyond Transparency: Open Data and the Future of Civic Innovation Open Data」, Code for America Press.

Naghman Waheed, Martin Mendez-Costabel (2017.03.15.), “Enterprise Geospartial Platform – A perfect fusion of cloud and open source technologies”, MONSANTO, Strata + Hadoop World.

Renata Máchová (2017), “Evaluating the Quality of Open Data Portals on the National Level”, *Journal of Theoretical and Applied Electric Commerce Research* [online], 12(1).

Winn, Joss (2013), “Open Data and the Academy: An Evaluation of CKAN for Research Data Management“, IASSIST 2013, 5-7, 14-17.

AWS Public Datasets, https://aws.amazon.com/ko/public-datasets/ 참조 (접속일: 2017.8.31.)

CC BY 4.0 license, https://creativecommons.org/licenses/by/4.0/ 참조

CKAN, “CKAN Portfolio – data.gov.au”, https://CKAN.org/portfolio/data-gov-au/ 참조 (접속일: 2017.09.09.)

CKAN을 사용하는 데이터 포털 목록, https://CKAN.org/about/instances/ 참조.

Cyrille CHAUSSON (2014.10.31.), “Data.gouv.fr: the French Open Platform for Open Data”, https://joinup.ec.europa.eu/community/epractice/case/datagouvfr-french-open-platform-open-data 참조 (접속일: 2017.09.29.).

DATA SHIFT, http://civicus.org/thedatashift/ 참조 (접속일: 2017.08.17.).

DATA.GOV (2013.01.28.), “CKAN: the Horizon for Data.gov 2.0”, https://www.data.gov/blog/CKAN-horizon-datagov-20 참조 (접속일: 2017.05.12.).

Data.gov.sg (2017), Singapore Open Data Licence, https://data.gov.sg/open-data-licence 참조 (접속일: 2017.05.17.).

Europeandataportal e-learning module, https://www.europeandataportal.eu/elearning/en/module2/#/id/co-01 참조 (접속일: 2017.08.17.).

Free our data, http://www.freeourdata.org.uk 참조 (접속일: 2017.08.17.).

Google Trends Datastore, http://googletrends.github.io/data/ 참조 (접속일: 2017.8.31.)

iCON 브리프 (2016. 03.09), “캐나다 오픈데이터 현황”, http://icon.ndsl.kr/i_trend/icon_briefDetail.jsp?record_no=5261&trendType=O 참조 (접속일: 2017.08.21.).

InfoWorld (2016.10.17.), “4 Google data sets to kickstart machine learning”,

https://www.infoworld.com/article/3131515/artificial-intelligence/4-google-data-sets-to-kickstart-machine-learning.html 참조 (접속일: 2017.08.24.)

Krasin I. et al (2017), “Open Images: A public dataset for large-scale multi-label and multi-class image classification,”, https://github.com/openimages 참조.

Link Digital, “Federal Government Open Data Portal”, https://linkdigital.com.au/content/datagovau-0 참조 (접속일: 2017.05.25.)

Linking Paths (2014.03.06.), “Icelandic Open Data, lambs and maps”, http://weblog.linkingpaths.com/post/78767700473/icelandic-open-data-lambs-and-maps 참조 (접속일: 2017.08.17.).

Mark Wainwright (2013), “US government’s data.gov to use CKAN”, https://CKAN.org/2013/02/04/us-data-gov-to-use-CKAN/ 참조 (접속일: 2017.5.19.).

Microsoft DataMarket https://datamarket.azure.com 참고

Monsanto Geospatial Platform Engineer Job offer, https://jobs.monsanto.com/job/st-louis/geospatial-platform-engineer/769/5100033 참조 (접속일: 2017.09.25.).

Nikhil Pahwa (2012.03.30.), “US &India Governments Launch Open Government Platform”, MEDIANAMA, https://www.medianama.com/2012/03/223-open-government-platform/ 참조 (접속일: 2017.08.21.).

Nikhil Pahwa (2011.06.06.), “Data.Gov.In To Launch; Use Cases: Powercuts.in; Challenges”, MEDIANAMA, https://www.medianama.com/2011/06/223-data-gov-in-to-launch-use-cases-powercuts-in-challenges/ 참조 (접속일: 2017.08.21.).

NuCivic (2017), DKAN을 사용하는 데이터 포털 목록, https://github.com/NuCivic/DKAN-sites/ 참조 (접속일: 2017.5.12.).

Open Data Charter (2015), “Towards an International Open Data Charter”, http://opendatacharter.net/history/ 참조 (접속일: 2017.08.02.).

Open Data Commons, “Making Your Data Open: A Guide”, https://opendatacommons.org/guide/ 참조 (접속일: 2017.09.12.).

Open Data Commons, “Open Database License”, https://opendatacommons.org/licenses/odbl/ 참조 (접속일: 2017.09.12.).

Open Data Commons, “Public Domain Dedication and License(PDDL)”, https://opendatacommons.org/licenses/pddl/ 참조 (접속일: 2017.09.12.).

Open Government Platform(OGPL), https://github.com/opengovplatform/opengovplatform-beta 참조 (접속일: 2017.05.12.).

Open Images Dataset (2017.07.20.), Open images dataset, https://github.com/openimages/dataset 참조

Open Knowledge International (2015), 전세계 오픈 데이터 포털 리스트(Data Portals), http://dataportals.org 참조 (접속일: 2017.5.14.).

Open Knowledge International, “What is Open Data ?”, 「Open Data Handbook」, http://opendatahandbook.org/guide/ko/what-is-open-data/ 참조 (접속일: 2017.09.12.).

Open Knowledge International, https://okfn.org/about/ 참조 (접속일: 2017.08.17.).

OpenDataSoft (2015), “A COMPREHENSIVE LIST OF 2600+ OPEN DATA PORTALS AROUND THE WORLD”, https://www.OpenDataSoft.com/a-comprehensive-list-of-all-open-data-portals-around-the-world/ 참조 (접속일: 2017.05.12.).

Press Information Bureau, Government of India (2012.03.30.), “Shri Kapil Sibal Launches ‘Open Government Platform’ Developed by India and US Initiative will Spur Citizen Engagement and Interface with Government”, http://pib.nic.in/newsite/erelease.aspx?relid=82025 참조 (접속일: 2017.08.21.).

Sudheendra Vijayanarasimhan and Paul Natsev (2016.09.28.), “Announcing YouTube-8M: A Large and Diverse Labeled Video Dataset for Video Understanding Research “, Google Research Blog (접속일: 2017.08.30.)

The Economic Times (2011.06.04.), “New Govt portal Data.Gov.In launch next month to ease information search”, https://en.wikipedia.org/wiki/Data.gov.in#cite_note-2 참조 (접속일: 2017.08.21.).

Tracxn (2016.8), “Tracxn BlueBox : Data as a Service “, 「Data as a Service Report」, California: Tracxn.

USAspending.gov, http://www.usaspending.gov 참조 (접속일: 2017.08.17.).

W3C (2009), “W3C guidelines for Open Data”, https://www.w3.org/TR/gov-data/ 참조 (접속일: 2017.08.03.).

Where Does My Money Go?, http://www.wheredoesmymoneygo.org/ 참조 (접속일: 2017.08.17.).

YouTube-8M Dataset https://research.google.com/youtube8m/ 참조

1. 본 보고서는 과학기술정보통신부의 출연금으로 수행한 DB산업 육성 사업의 결과입니다.
2. 본 보고서의 내용을 발표할 때에는 반드시 과학기술정보통신부 DB산업 육성 사업의 결과임을 밝혀야 합니다.
3. 본 보고서 내용과 관련한 문의는 한국데이터진흥원 정책기획실 (02-3708-5364)로 연락해 주시기 바랍니다.