IASSIST 컨퍼런스 참가
결과보고서

2014. 7.

정보서비스팀

차 례

Ⅰ. 출장개요 ····· 1

Ⅱ. 컨퍼런스 참석결과 ····· 3

Ⅲ. Arofan gregory 면담결과 ····· 23

Ⅳ. 시사점 및 제언 ····· 27

<참고1> 컨퍼런스 세부일정 ····· 30

<참고2> DDI 연합(DDI Alliance Group) ····· 31

<참고3> MTNA & OPENDATAFORGE ····· 35

<참고4> 컨퍼런스 참석 사진자료 ····· 39

Ⅰ. 출장개요

1. 출장 목적

○ 자료정보처리 기술 및 자료제공의 국제적인 동향 파악
- 40th IASSIST* Annual Conference(2014) 참석
. 주제 :「Aligning Data And Research Infrastructure」

○ DDI 전문가와의 개별면담을 통한 세부기술 습득

IASSIST

◈ **IASSIST**
(International Association for Social Science Information Service & Technology)

■ 주로 **학교, 정부, 비영리단체에 소속된 정보기술 및 자료 전문가들이 모인 국제조직**으로 사회과학 분야의 조사, 교육을 지원하기 위한 데이터처리 기술 및 자료를 공유하기 위해 설립.

■ 매년 컨퍼런스를 개최하며 이번 2014년 컨퍼런스는 40번째임.
- 40th Annual Conference(2014) : Aligning Data And Research Infrastructure
39th Annual Conference(2013) : Data Innovation_ Increasing Accessibility, Visibility And Sustainability

2. 출장지, 기간 및 출장자

□ 출장지 : Ryerson University, 캐나다 토론토

□ 출장기간 : 2014. 6. 2. ~ 6. 8. (5박7일)
* IASSIST 회의기간 : 2014. 6. 3. ~ 6. 7. (4일)

□ 출장자 : 통계정보국 정보서비스팀 6급 최유성

3. 출장일정 및 업무 수행내용

□ 출장일정

월.일 (요일)	출발지	도착지	업무수행내용
6.2(월)	인천	토론토	출국/도착
6.3(화)	토론토		IASSIST Workshops 자료 수집 및 Conference Opening .Data Management & Curation: Lessons from Government, Academia, and Research .Teaching an introductory workshop in digital preservation
6.4(수) ~ 6.6(금)			IASSIST Conference 참석(6.4~6.6) .Integrated Data Discovery and Access .Teaching Data Management and Statistical Literacy .Data Without Boundaries .Next-generation DDI .Energizing the DDI Standard through Tools .Integrated Data Discovery and Access Aforan Gregory 면담(6.5, 6.6)
6.7(토)~ 6.8(일)	토론토	인천	입국

□ 업무 수행내용

○ 2014 IASSIST Annual Conference 참가

- Research Data Management, Professional Development, Data Developers and Tools, Research Infrastructure관련 세부주제 논의

○ DDI 전문가'Arofan Gregory' * 개별 면담

* Metadata technology North America(MTNA)는 DDI와 SDMX등의 메타데이터 표준을 도입한 응용프로그램 개발 및 데이터 관리 솔루션을 제공하는 IT업체이며 Arofan Gregory는 MTNA의 Manager임

- 한국 통계청의『국가통계 마이크로데이터 통합시스템 구축』사업을 설명하고 한국에 맞는 DDI버전 및 관련 도구, 시스템 구축 시 주의사항 등에 대해 세부적인 기술자문 요청

Ⅱ. 컨퍼런스 참석결과

1. 주요내용

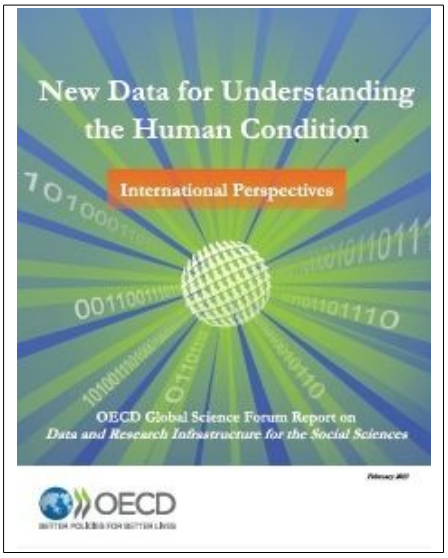
□ 주제 : 데이터의 정렬과 연구인프라

- “Aligning Data And Research Infrastructure”

○ 2014 IASSIST conference 주제는 “데이터의 정렬과 연구인프라”로 국제 및 학계 요구 사항에 부합하는 데이터의 정렬과 데이터 공유 인프라에 대해 논의

- ‘13년 OECD 세계 과학 포럼 보고서 “인간의 조건을 이해하기위한 새로운 데이터”에는 최근에 나타난 새로운 형태의 사회과학 자료*에 대해 국제적인 데이터 공유와 협업을 강조

* 전자상거래, 인터넷 검색, 소셜 네트워킹 등을 통해 수집 된 새로운 형태의 데이터



- 이러한 사회적 변화에 대응하기 위해 **오픈 데이터(Open data), 아카이빙(Archiving), 복제나 재사용이 허용되는 장기적 데이터 큐레이션(Curation)및 기반확충**을 각 기관의 **사례중심**으로 논의

- 4가지 부문의 ‘Stream’ 으로 구성
  ① **연구자료의 통합 및 관리**(RDM : Research Data Management)
  ② **연구기반환경**(RI : Research Infrastructure),
  ③ **자료전문가**(DP : Data Profession),
  ④ **자료의 발견과 툴**(DDT : Data Discovery and Tools)

2. 컨퍼런스 세부내용

< Stream별 세부주제 >

Stream		주제
RDM	1	Building Data Collections
	2	Next-generation DDI
	3	RDM across boundaries and disciplines
	4	Enhancing Restricted Access
	5	Encouraging Access to Confidential Data
	6	Innovative Approaches to Promoting Transparency in Research
	7	Research Environments
	8	RDM Activites within CESSDA
RI	1	Data Without Boundaries
	2	Data Service Infrastructure for the Social Sciences and Humanities
	3	Collaboration and Networked Infrastructure
	4	Metadata Portal for the Social Sciences
	5	National Data Managment Policies
	6	Data Curation and Workflow
DP	1	Developing Meaningful Data Support Roles and Services
	2	Teaching Data Management and Statistical Literacy
	3	Integrated Data Discovery and Access
	4	Trust and Data Sharing
	5	Planning and Assessing Research Data Services
	6	Supporting Data Use
DDT	1	Big Picture Metadata
	2	Community Source meets Open Source
	3	RDM Tools and Services
	4	Harmonization, Thesauri and Indexing
	5	Integrated Data Discovery and Access
	6	Energizing the DDI Standard through Tools

- RDM : Research Data Management
- RI : Research Infrastructure
- DP : Data Profession
- DDT : Data Discovery and Tools

※ 날짜별 컨퍼런스 상세일정 <참고 1>

RDM 1 Building Data Collections

□ 역사 데이터 복원 및 보존(영국의 사례)

• 발표자 : Arne Wolters(영국 데이터 아카이브) 외 3인

- 역사인구조사 마이크로데이터(EEHCM) 프로젝트는 영국의 데이터 아카이브, 영국통계청, 스코틀랜드 국립기록(NRS)과 북 아일랜드 통계연구기관(NISRA)의 공동프로젝트

- 두 단계로 구성되었으며 첫 번째 단계는 영국의 1961년, 1966년, 1971년과 1981년 인구조사의 복원 및 보존의 추진, 두 번째 단계는 각 인구조사 데이터의 공동사용을 위한 데이터 샘플을 작성

- 의의는 천 년 전 국립기록 백업 테이프를 복원, 역사적 데이터를 성공적으로 DB화하여 역사연구의 복원에 새로운 사례를 남김

□ 국제사면위원회 자료의 디지털화 프로젝트 : "Coding + Metadata = Datasets"구성의 예

• 발표자 : Amy Barton(Purdue University) 외 3인

- (미)국제사면위원회(AI-USA), 퍼듀(Purdue) 대학의 연구자의 공동 프로젝트로 AI-USA의 초기 기록을 보존을 위해 국제사면위원회 연구 데이터의 디지털화를 추진

- AI-USA의 1974년 이후의 '긴급행동공지' 관련 데이터 소스를 병합, 추가 메타 데이터를 개발, Nvivo 코딩을 통해 데이터를 DB화 시킨 사례

RDM 2 Next-Generation DDI

□ 모델기반의 차세대 DDI에 대한 논의 : Building Model-Based DDI 4

• 발표자 : Steve McEachern(호주 데이터 아카이브)외 4인*

* Arofan Gregory, Open Data Foundation 포함

- DDI는 조직을 위한 구조화된 메타 데이터의 설명서임 지금까지 XML 스키마(Schemas)기준의 두 가지 사양 DDI Codebook, DDI Life-cycle이 있으며 추가적으로 통제어휘(Controlled vocabularies)과 RDF 어휘(RDF vocabularies)를 개발 이들은 모두 DDI 연합에서 관리되고 있음

- 데이터의 DDI적 접근으로의 유용성은 ①풍부한 콘텐츠 ②메타데이터를 전 생성주기(Life-cycle)에 걸쳐 재사용 ③기계로 실행가능(Machine-actionability) ④자료의 정리 및 관리, 보관의 유용성 ⑤종적 데이터의 지원과 비교를 들 수 있음.

- **(DDI4란)** 안정적이면서도 유연한 사양을 필요로 하는 새로운 환경적 요구에 부합하고 생성주기의 범위를 확장, 다른 표준과의 통합을 향상시키기 위해 개발되고 있음

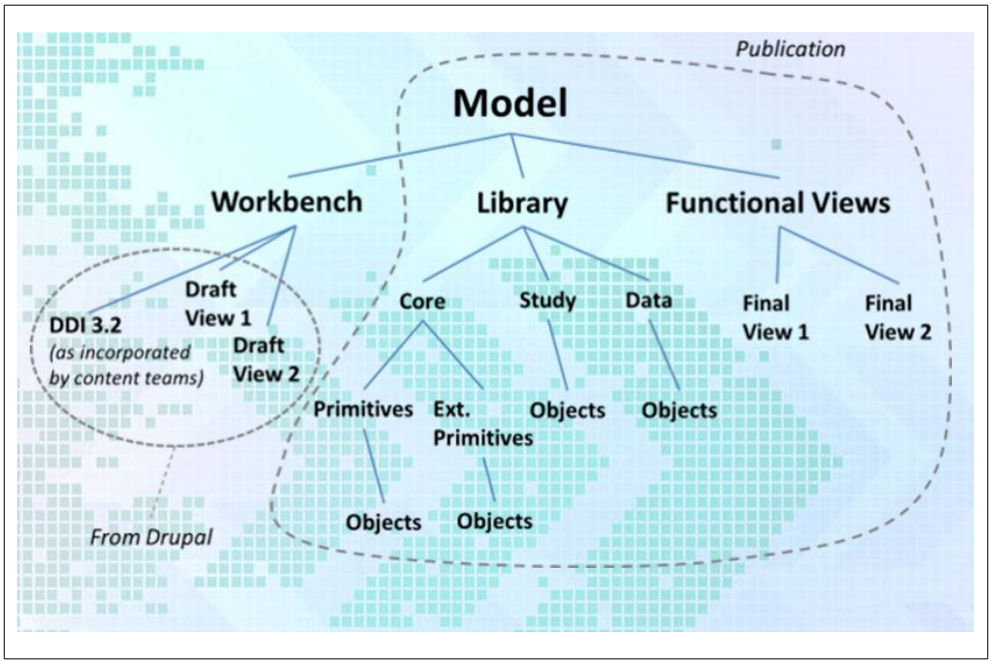
  1) **구조적 목표, 설계목표(Architecture Goals)** : 다른 분야 표준과의 의사소통을 향상, 모델의 기술적 표현과 관련한 유연성, 효율적인 개발과 유지관리

  2) **내용적 목표(Content Goals)** : ①서로 다른 유형의 데이터를 처리 할 수 있도록 데이터 캡처, 수집, 소스의 추출 ②샘플링, 조사의 실시, 가중치, Paradata에 대한 새로운 내용 추가 ③GSBPM, GSIM, SDMX, CDISC, Triple-S등 다른 메타데이터 기준과 통합하는 관리 계획

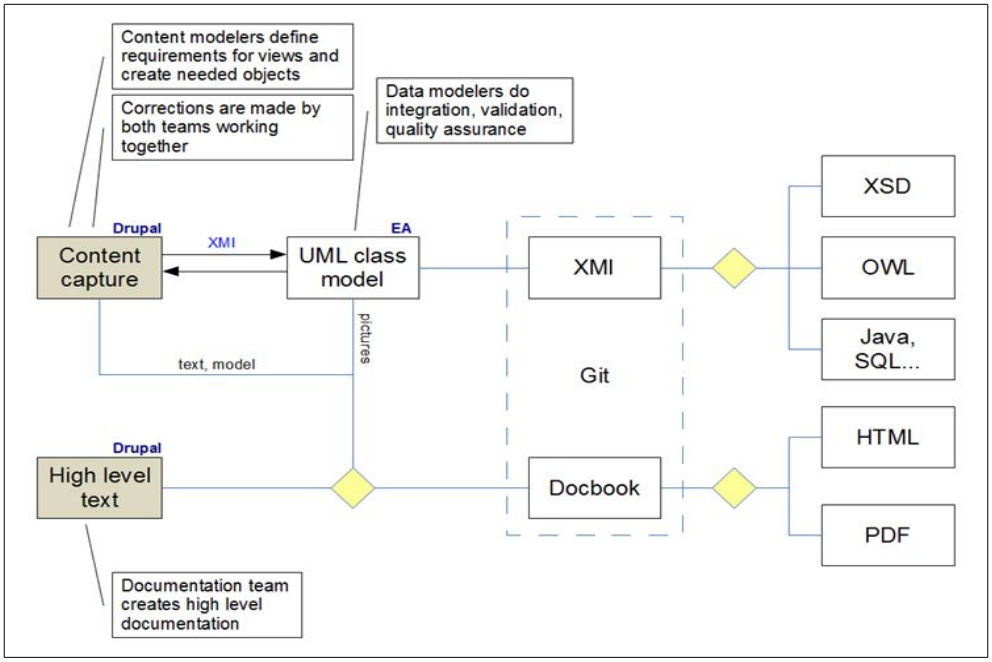
- **(진행상황)** 2012년 Dagstuhl 트레이닝 센터(Computer science research center in Germany)에서 DDI4 요구사항과 설계원칙을 최초로 정하고 '13년 10월 Wadern(독일), '13년 12월 EDDI(파리), '14년 4월 NADDI(밴쿠버), '14년 6월 IASSIST(토론토)에서 토론을 진행

DDI 4 : 모델 기반의 기술적 프레임워크

(DDI4는 주제별로 구성된 개체(Objects)의 라이브러리로 조직화)



XSD, PDF형식을 DDI products(XML형식)로 생성하기 위한 변환모듈



□ 연구 재단에서 기업 데이터 서비스를 구축(FRB 사례)

• 발표자 : San Cannon(미연방준비제도이사회,FRB)

- 데이터는 금융 정책, 금융 안정, 감독, 소비자 보호 및 경제 연구 등의 기능, 광범위한 FRB의 임무를 완수하는데 중요한 역할을 하지만 분산된 데이터 관리, 분산된 의사결정 및 조달, 극히 제한적인 데이터 공유로 금융위기를 겪음

- OCDO(The Office of the Chief Data Officer)는 데이터 거버넌스, 데이터 관리 및 데이터 통합을 강화하기 위해 기업의 초점과 의무를 명확하게 설정, 포스트 금융위기 회의의 데이터 요구에 대응하기 위해 2013년 5월 구축

- 연구기능 지원, 사용자의 폭을 넓혀 데이터 공유를 강화하고 다양한 분석 작업이 가능한 서비스를 제공

OCDO by FRB service portpolio

RDM 8 RDM Activites within CESSDA

□ CESSDA 보관소와 연구 데이터 관리 활동 소개

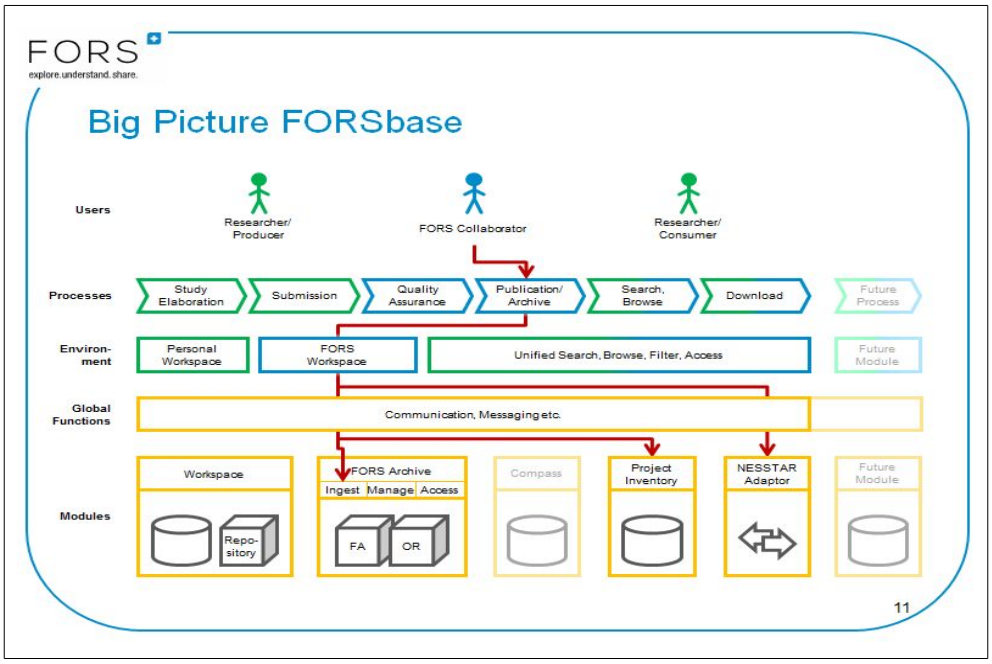
• 발표자 : Laurence Horton(GESIS)외 6인

- 유럽 사회 과학 데이터 보관소위원회(CESSDA)는 유럽의 다양한 데이터 인프라를 개발하기 위해 공동으로 작업하는 유럽국가 기록 보관소 협회임 회원국 데이터 아카이브와 함께 연구 데이터 관리 교육 및 지원을 추진

- 유럽 각국(독일, 핀란드, 영국, 덴마크, 노르웨이, 스위스 등)은 CESSDA 멤버로서 사회 과학 데이터 공유 요구 사항을 설명, 데이터의 재사용을 위해 데이터를 보관하고 연구 데이터 관리 지원하는 각국의 서비스를 소개

FORs* :

Developing data management support services in Switzerland



* FORs : 스위스의 사회과학연구재단 데이터 관리 Tool(DDI Codebook 기반)

RI 1 Data Without Boundaries

□ 분산된 데이터 소스에 대한 Microdata Computation Centre(MiCoCe)

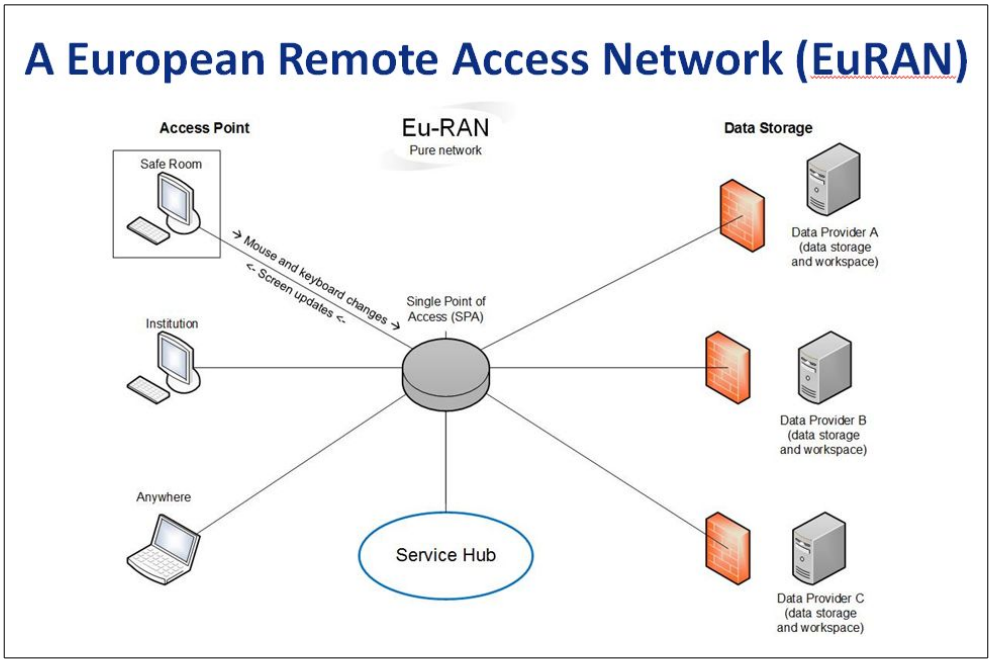
• 발표자 : David Schiller

* Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB)

- 경계없는 데이터 (DwB) 프로젝트는 비밀 마이크로 데이터에 액세스 할 수 있는 유럽의 원격 액세스 네트워크(EuRAN)를 제안 유럽 RAN 내에서 중앙 서비스 허브가 연구자 및 연구 프로젝트를 지원하는 다른 서비스를 호스팅 할 수 있도록 구축

- 그 중 대표적인 사례는 Microdata Computation Centre(MiCoCe) 이며 이것은 분산 된 데이터 소스의 분석을 가능하게 함 EuRAN과 MiCoCe는 공공 및 민간 부문의 데이터에 접근할 수 있는 인프라로 작동

EuRAN 구조



□ 유럽 원격접근 네트워크(European Remote Access Network)의 시작

• 발표자 : Anja Burghardt, FDZ/BA/IAB

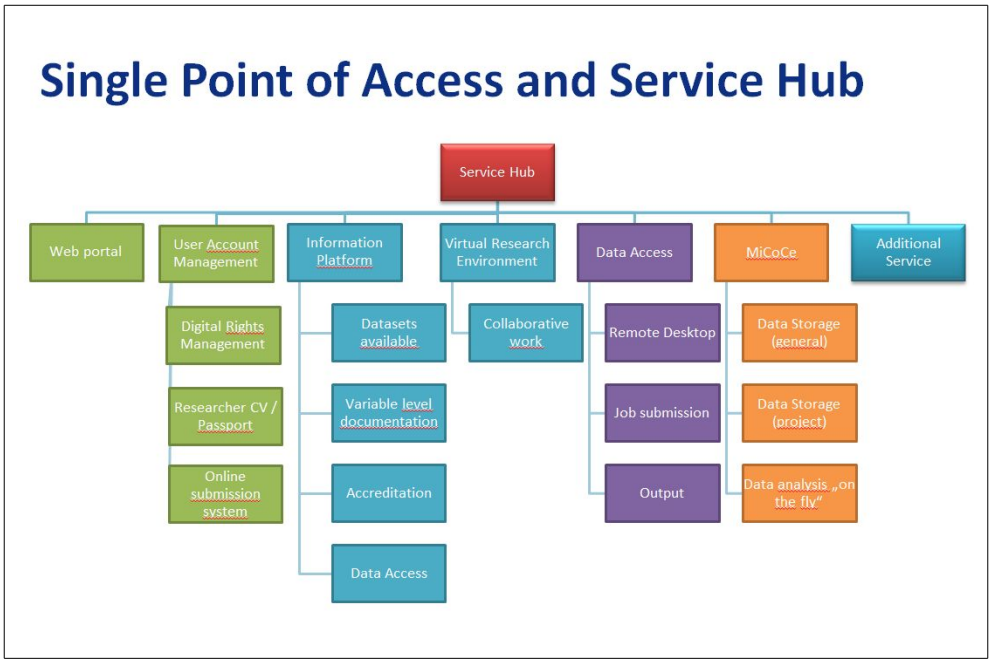
- 영국의 데이터 아카이브(UKDA) 와 독일 연방 고용청 (BA)의 연구데이터센터(FDZ)의 기밀데이터 액서스 사례는 EuRAN와 SPA의 첫 번째 모범사례

- SPA와 EuRAN는 분산 데이터 소스를 분석 할 수 있는 도구에 액세스 할 수 있는 보안 인프라임

□ Single Point of Access (SPA): 원격 액세스 네트워크(EuRAN)를 위한 서비스 허브

• 발표자 : Anja Burghardt(FDZ/BA/IAB) 외 1인

SPA



- 액세스 (SPA)의 중앙 집중화 된 단일 지점은 쉽게 여러 분산 된 조직의 네트워크 지점에 도달하고 이 네트워크 내에서 다른 데이터 소유자의 데이터로 프로젝트를 실행 연구자의 작업을 용이하게 함

RI 2 Data Service Infrastructure for the Social Sciences and Humanities

□ 인문사회과학을 위한 데이터 서비스 인프라(DASISH) 소개

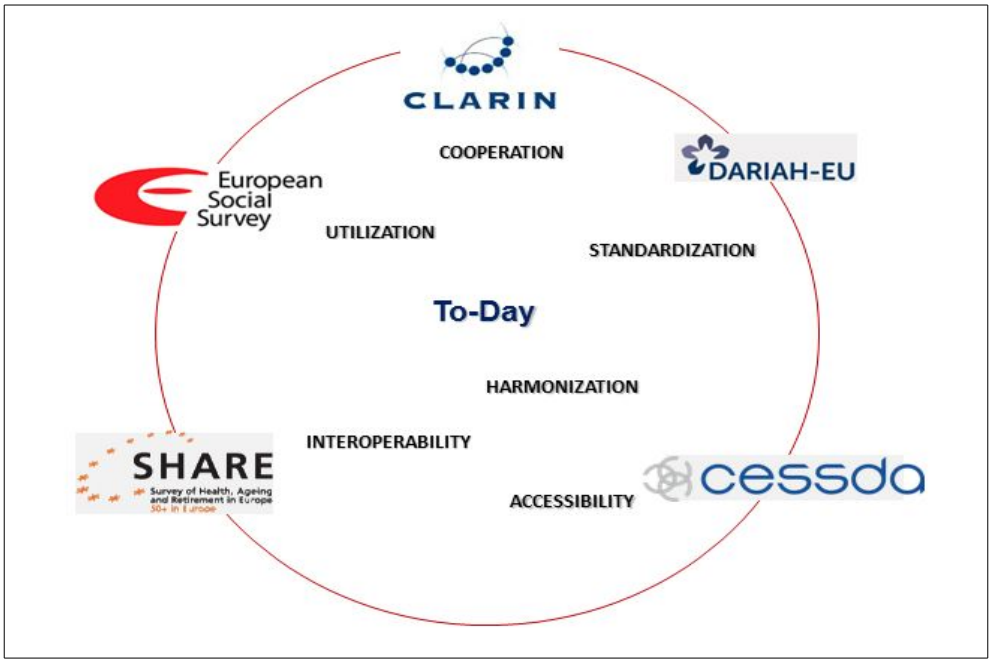
• 발표자 : Michelle Coldrey(스웨덴 국립데이터 서비스, SND) 외 5인

- DASISH는 인문 사회과학 분야에 있어 SSH[1]내의 5ESFRI[2] 인프라에 대한 클러스터 프로젝트임

1) Social Sciences and Humanities (SSH) projects : 2010년 유럽 2020전략 문제를 해결하기 위한 핵심요소로 EU가 추진하는 사회과학 및 인문학 연구 프로젝트

2) CESSDA : Council of European Social Science Data Archives
CLARIN : Common Language Resources and Technology Infrastructure
DARIAH : Digital Research Infrastructure for the Arts and Humanities
ESS : European Social Survey
SHARE : Survey of Health, Ageing and Retirement in Europe

Cluster DASISH



- 국제협력 강화를 위한 솔루션 제공, 13개 유럽국가 18개 기관과 파트너를 채결하고 디지털 자원, 도구, 서비스에 대한 가시성(visibility)과 재사용성(re-usability)을 강화함을 목표

RI 5 National Data Managment Policies

□ 국가의 데이터 관리 정책 : 데이터 아카이브 및 기관 저장소에 미치는 영향의 국가 간 비교

• 발표자 : Steven McEachern(호주 데이터 아카이브) 외 2인

- 최근 각 국가정부의 “Open Data, Open Access” 정책의 발전은 데이터 관리 정책 및 절차에 대한 강조와 함께 증가하고 있으며 이것은 국제적인 추세임

- 연구 자금제공자(funder)은 연구기관이 효율적으로 데이터를 이해, 관리, 지원하도록 요구하며 때론 연구 보조금의 지급을 조건으로 연구의 시작에서 결과까지 데이터의 관리계획을 강조

(사례1) 국가의 데이터 관리 정책 : 독일의 사례

• 발표자 : Alexia Katsanidou(GESIS)

- 독일연구협회(DFG)의 1998년 게시에 의하면 출판에 사용된 공적인 투자를 받은 데이터는 보존되어야 하고 적어도 10년 동안은 볼 수 있도록 해야 한다고 규정
- DFG는 데이터 관리를 충분히 고려하는 연구자를 장려, 데이터 공유를 촉진하는 가이드라인 발표
- EU는 데이터 작성에 더 높은 투명성을 강조하고, 많은 국가들이 데이터 공유를 위한 응용프로그램에 대한 투자를 진행
- 독일 고등교육기관은 연구데이터 관리문제를 해결하는데 적극적인 역할을 해야 함

(사례2) IFDO*의 연구 자금제공자(Research Funder)의 데이터 정책 조사

• 발표자 : Vigdis Namtvedt Kvalheim

* IFDO : International Federation of Data Organizations for social science

- 데이터는 연구 인프라가 작동하는 핵심적 요소이며 기존액서스 메커니즘은 더 이상 사회환경 변화에 적용될 수 없음을 인식

- 데이터 공유에 대한 강력한 정치적 의지를 법률로 규정하고 데이터 액세스 및 데이터 공유인프라 구축을 지원, 강화하기 위한 새로운 자금 조달 계획을 수립함

- 연구 자금제공자의 데이터 정책에 대한 IFDO조사는 현재 기관 연구데이터 정책에 대한 국가별 정보를 수집 공식데이터 정책에 초점, 데이터 공유의 질이나 내용에 대한, 자금 결합의 유형 등을 조사 2012년부터 2013년 1월까지 90개국을 대상으로 조사

- 연구 자금제공자 정책은 ①데이터 공유의 정책이 명시적으로 규정되어 있고 저장의 위치가 분명한 경우②데이터 공유의 정책은 명시적이나 실행이 불분명한 경우③정책이 규정되어 있지 않은 경우로 나뉘어짐

- 데이터의 공유가 이루어지지 않는 이유로 언급된 것은 ①공유 문화의 부족, ②확립된 정책이 없음, ③ 인프라 및 자금 부족 등

IFDO의 DPDS(Data Policy Description Sheet)서식

Topic	Nr.	Topic Item
Background information	1	Name of funder
Background information	2	Homepage
General policy	3	General conditions
General policy	4	Data Management Plan (DMP) for Proposal
General policy	5	Data Timeframe
General policy	6	Guidance
General policy	7	Compliance/Monitoring
General policy	8	Funding / Costs
General policy	9	Scope of policy
Standards/Documentation	10	Documentation Requirements
Standards/Documentation	11	Data Standards
Standards/Documentation	12	Metadata Standards
Access and preservation	13	Data Preservation
Access and preservation	14	Scope of preservation provisons
Access and preservation	15	Data Access / Sharing
Access and preservation	16	Data Access / Sharing incentives
Access and preservation	17	Data Sharing Rights (IPR)
Access and preservation	18	Data Embargo / Data Retention
Access and preservation	19	Data Sharing requirements / timeframe
Access and preservation	20	Designated Data Repository
Access and preservation	21	Data Repository Supported
Access and preservation	22	Institutional (data repository) Requirements
Publications	23	Open Access to Publications
Publications	24	Publication Repository Specified
Publications	25	Publication Repository Supported
Resources/References	26	Date of policy
Resources/References	27	Policy link
Resources/References	28	Policy link

DDT 1 Big Picture Metadata

□ 식별자 시스템을 이용한 연구자와 연구자료의 연결

• 발표자 : Elizabeth Newbold(영국도서관)

- ORCiD[1] 및 DataCite Interoperability Network(ODIN) 프로젝트는 자신의 연구 자료와 연구자, 저자와 기여자를 연결하는 사람들과 데이터 세트에 대한 지속적인 개방과 상호 운용이 식별자를 사용하는 것을 목적으로 만들어진 영국 국립 도서관, CERN, ORCiD, DataCite[2], Dryad[3], arXiv 및 호주 국립 데이터 서비스 간의 공동 프로젝트

1) ORCiD : Open Researcher and Contributor ID, 과학 및 기타 학술저자를 식별하는 숫자코드
2) DataCite : 데이터 인용(Cite)를 위한 비영리단체
3) Dryad : 기본 과학 출판물 범용 저장소

- 인문 사회 과학 (HSS) 및 에너지 물리학 (HEP)등 여러 분야에 걸쳐 식별자 체계를 구현하기 위해 공통적 요소를 밝히고, 식별자 시스템에서 제공하는 오픈 API를 기반으로 한 부가 서비스를 소개

□ DDI handbook : DDI 개요와 적용의 추천사례 모음 프로젝트

• 발표자 : Joachim Wackerow, GESIS

- DDI 핸드북 프로젝트의 소개 DDI를 사용한 프로젝트의 모범사례들을 정리하고 웹 싸이트에 게시(https://ddi-handbook.org)

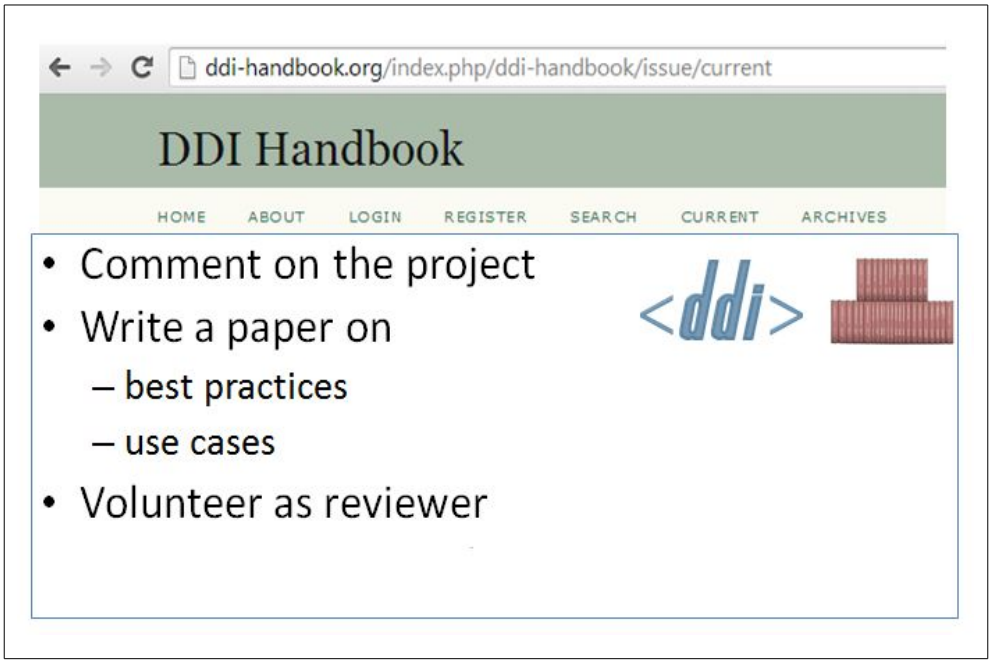
- DDI는 데이터 사용자와 생산자가 증가할수록 사용량이 증가하며 그 사용자, 생산자의 분야가 다양해서 실제 DDI 적용의 모습도 매우 다양 이러한 현상은 GESIS뿐 아니라 모든 영역에서 나타나는 현상임

- 처음 DDI 적용에 어려움이 따르는 것은 DDI를 원하는 영역에서 사용하기 위한 최선의 방법이 제시되지 않았기 때문이며 균일한 구조(homogeneous structure)에서 유용하고 접근하기 쉬운 정보를 제공하기 위한 프로젝트를 시도

- 기존의 DDI Working Paper들은 적용범위가 규정되지 않고, 균일구조로 이루어지지 않은 한계가 존재 이러한 문제를 해결하기 위한 초기회의 2013년 11월 GESIS 에서 갖고 전체적인 전략과 개념토론을 거쳐 독립적 온라인 저널을 오픈

- 컬렉션의 기본 구조는 DDI 라이프 사이클과 정렬로 구성됨 목표는 DDI 규격을 사용하는 모든 조직과 개인을 위한 자원의 공유된 내용을 생산함으로써 DDI 커뮤니티에 참여하도록 지원함 좋은 방법은 편집자 리뷰들로 구성된 팀에 의해 검토하고 전용 웹 사이트에 게시되어지는 것임

DDI Handbook (https://ddi-handbook.org)



DDT 2 Community Source meets Open Source

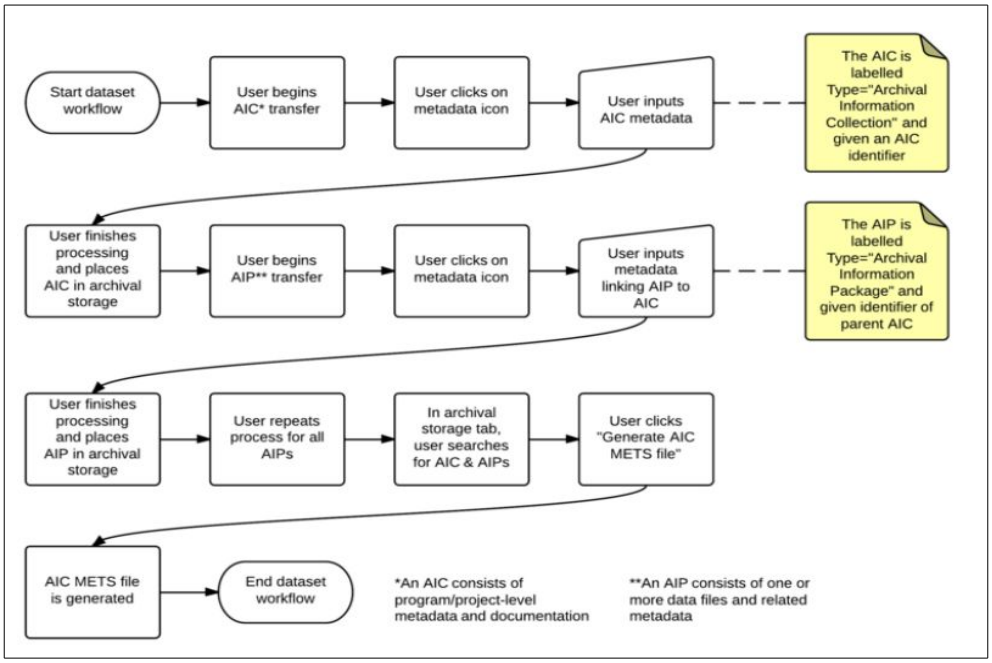
□ 커뮤니티 소스(Community Source)와 오픈 소스(Open Source)의 결합 = 협력적 오픈소스(Collaborative Open Source)의 사례

• 발표자 : Carla Graebner, Simon Fraser University 외 3인

- 브리티시 컬럼비아 도서관(University of British Columbia Library), 앨버타 대학 도서관(The University of Alberta Library), 사이먼 프레이저 대학 도서관(Simon Fraser University Library)은 각 대학 사이트들은 새로운 디지털 저장소 사업에 대한 요구 사항을 충족하기 위해 디지털 컬렉션 및 아카이브 소프트웨어 플랫폼 로컬 오픈 소스 개발자들과 협력

- 사이먼 프레이저 대학은 Research Data Repository (RaDar)를 개발, 브리티시 컬럼비아의 대학은 캐나다 북극 데이터 네트워크를 개발하고 있는 cIRcle과 앨버타 대학과의 통합을 포함하는 디지털 보존 프로그램을 만드는 것을 추진

AIP Workflow



DDT 3 RDM Tools and Services

□ 연구자료 관리를 위한 도구(Tool)및 서비스 소개

• 발표자 : Carly Strasser(California Digital Library)외 3인

- 캘리포니아 디지털 도서관의 캘리포니아 대학의 큐레이션 센터(UC3) 그룹과 그들의 파트너는 연구자들이 처리, 공유 및 데이터 세트의 보관을 개선하는 데 도움이 되는 도구를 개발 그 도구들은 데이터 생명주기(Life-cycle)*에 따라 서로 다른 중점역할을 수행

* 데이터 생명주기 : Plan → Collect → Describe → Analyze → Preserve → Share

캘리포니아 대학의 큐레이션 센터 (UC3) 그룹 개발 도구(Tool)

1) (Plan) DMPTool 2.0 : 데이터 생성, 편집, 재사용 등 데이터 관리 계획 도구
2) (Describe, Share) EZID : 식별자 및 관련 메타데이터 관리, 식별자 확인
3) (Collect, Describe, Analyze, Preserve, Share) DataUp 2.0 : 관리 및 데이터 공유
4) (Preserve, Share) Merritt, dash : 저장소, 디지털 자산 및 액세스
5) (Preserve) Web Archiving Service : 데이터의 유지 및 저장(Web)

DDT 5 Integrated Data Discovery and Access

□ 사회과학의 통합데이터 포털 da|raSearchNet 소개

• 발표자 : Tanya Friedrich, 독일 사회과학연구소(GESIS*) 외 2인

* GESIS : Leibniz Institute for the Social Sciences in Germany

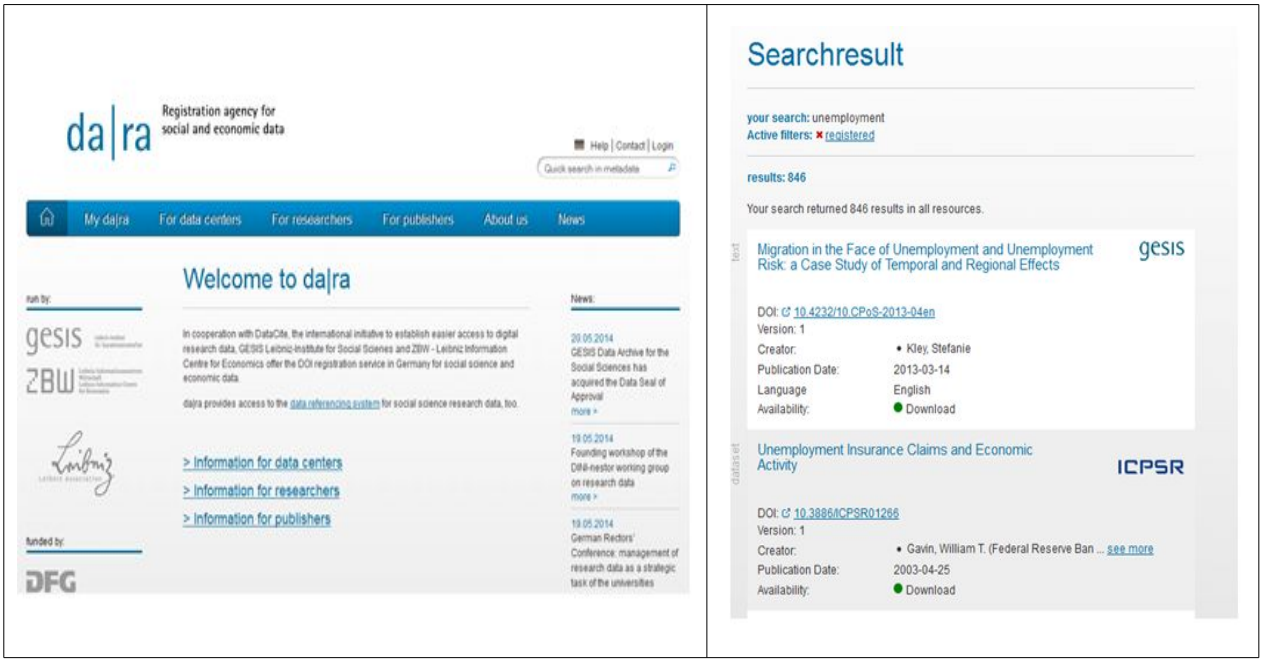
데이터 공유 인프라(Data sharing infrastructure)

- 연구데이터의 편리한 검색은 데이터 공유 인프라에 크게 의존함 하지만 데이터 저장소 및 메타데이터 서비스의 현재상황은 같은 분야에서 조차 일관성이 없고 따로 흩어져 있음

- 프로젝트 da|raSearchNet에서 사회과학 분야 내에서 데이터 공유 육성을 목표로 통합 검색 인프라를 설계하고 구현하는 프로젝트를 수행

- 주요특징으로는 ①데이터 세트의 인용에 대한 포괄적 인 메타데이터 (영역별) ②독일과 미국 데이터에 초점 (GESIS 및 ICPSR) 및 기본검색기능이 강조

da|ra 메인화면과 자료검색화면



- da|ra- 사회과학 데이터 시스템의 완성으로 30개 이상의 데이터 공급자와 20000 데이터 파일 및 기타 리소스가 하나의 데이터 베이스에 저장 이용자가 메타 데이터를 통합하고 응용 프로그램을 검색 할 수 있는 기능 구축

□ 사회과학데이터와 건강데이터의 연결(덴마크 사례)

• 발표자 : Amalie Nielsen(덴마크 데이터 아카이브)

- 보관, 보급 및 연구 데이터의 재사용은 새로운 주제에 부차적인 분석을 할 수 있는 기회를 향상 시킴 DDI-Lifecycle 표준을 사용하여 메타 분석을 할 수 있고 서로 다른 연구 분야의 데이터 자료를 비교 분석 할 수 있음 덴마크의 데이터 아카이브(Danish data Archive)에 사회과학 데이터, 건강 데이터가 있으며 이러한 종류의 데이터를 서로 연결함으로써 공중 보건의 사회적 영향에 대한 지식의 새로운 형태를 얻을 수 있음

- 덴마크 데이터 아카이브는 건강 설문 조사와 문화 습관 설문 조사와 같은 고전 사회 조사에 링크 할 수 있는 있으며 이와 함께 보건 의료 시스템 및 레지스트리, 데이터베이스의 사용이 용이 데이터 소재가 다른 경우(다른 목적을 위해 다른 연구자나 기관에 의해 수집)는 복잡 비교방식 사용

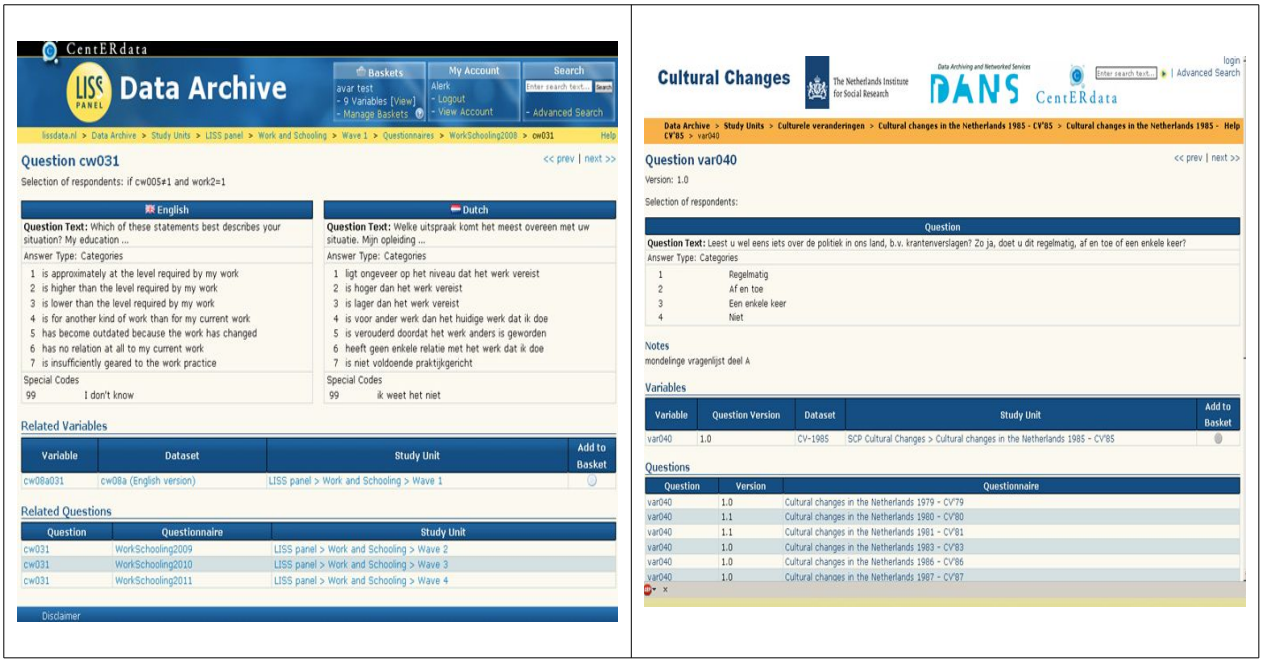
□ Questasy : DDI3기반 종단면 데이터의 문서화와 도구(네덜란드사례)

• 발표자 : Edwin De Vet, CentERdata*

* CentERdata는 네덜란드 Tilburg대학의 연구 기관으로 패널조사 및 자료관리, 양적자료분석(Quantitative Analysis), 소프트웨어 솔루션을 제공하고 있음

- Questasy는 CentERdata*에서 개발 DDI3 기반으로 데이터 보급 툴로 CakePHP의 작성 및 MySQL 데이터베이스를 사용 종단면 분석연구뿐 아니라, 사용자 정의 데이터 세트의 생성을 지원함

Questasy를 이용한 LISS 와 SCP 데이터



□ 복합 연구자료에 대한 DDI 적용 메타데이터 포털의 구축(독일사례)

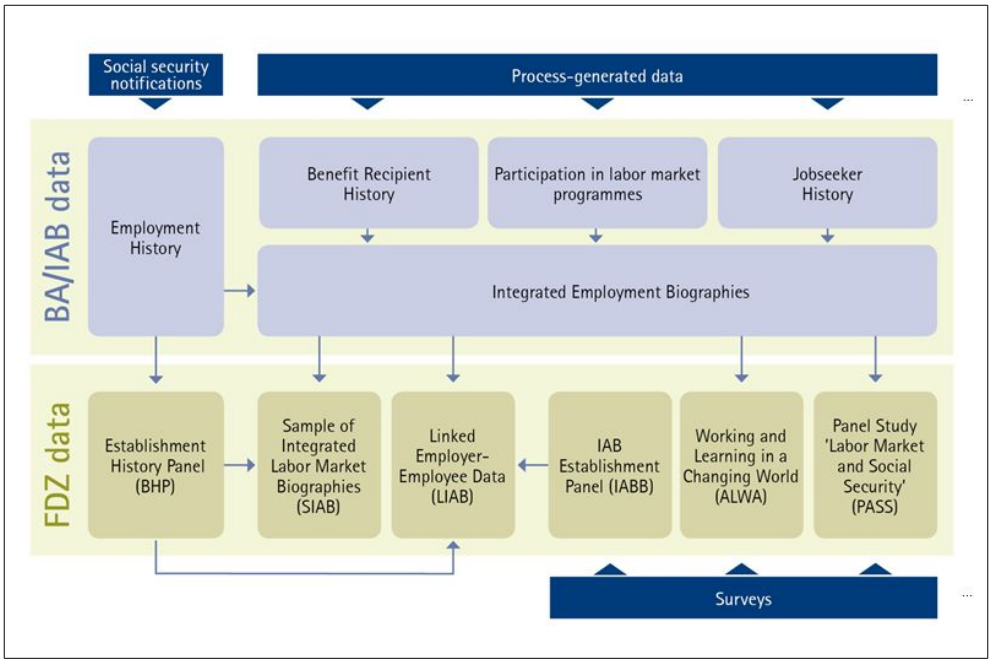
• 발표자 : David Schiller, 고용연구소(IAB*) 외 2인

* IAB : Institute for Employment Research

- 독일 고용 연구소(IAB)에 있는 연방 고용 기관의 연구 데이터 센터(FDZ)는 과학 커뮤니티에 대한 서로 다른 유형의 데이터를 제공함 예를 들어 설문 조사 및 등록 데이터 사이의 레지스터 데이터, 조사 데이터와 연결된 데이터가 있음 하지만 사용 가능한 메타데이터 도구는 연구 데이터 센터의 연구 데이터의 다양성과 부서 내에서 데이터 문서 및 부서 간의 협력의 흐름을 고려하지 않아 이용에 제한

- 이러한 문제를 해결하기 위하여 TBA 21 Assessment SYSTEME GmbH(독일), OPIT Consulting Kft(헝가리)와 2012년에 Colectica (미국) 기술 응용 프로그램을 사용하여 모든 종류의 데이터에 대한 메타 데이터를 표준화 처리, 공동으로 국제 개발 프로젝트를 시작 Colectica를 이용하여 DDI 3.2 적용한 사례

IAB 데이터 리소스의 복합구조



※ 보고서에 실리지 않은 Topic의 원문 발표 자료는 별도첨부

Ⅲ. Arofan Gregory 면담결과

Arofan gragory

◈ (소속) 북미 데이터 관리 및 DDI, SDMX 메타데이터 표준 기반의 소프트웨어 개발, 컨설팅 전문 회사 Metadata technology North America(MTNA)소속으로 과학기술 출판산업, 전자 상거래 및 국제 무역 분야 경력이 있으며, IT 업계에서는 20년 이상 종사하였음

◈ (활동) SDMX 프레임 워크와 DDI, GSIM 등의 메타데이터 표준을 적용하는 응용프로그램을 홍보, 교육, 지원하는 컨설턴트 활동이 주된 업무이며 특히 소속회사인 MTNA가 DDI 연합의 준회원기관으로 Arofan은 DDI 연합의 Working Group(「Qualitative Data Model Working Group」, 「Technical Committee」)에서 활동 중

A. DDI 연합(DDI alliance group) 회원으로의 활동

○ DDI 연합의 회원들은 자발적인 회비(Membership Fee)로 조직을 운영하고 있는 것으로 알고 있다. 공식적인 영리활동이 있다면 어떤 것들이 있나?

➡ DDI 표준을 적용한 **응용프로그램을 개발**하거나 **DDI 관련 워크샵, 컨퍼런스 등을 개최**하고 교육자료 등을 출판한다. INSEE, ICPSR, GESIS, ABS 등 약 40개의 회원기관과 함께 **공동 프로젝트를 진행하면서 얻는 수입(연구 보조금)**등 도 있다.

○ DDI 연합의 회원으로서 국제적으로 주로 하는 역할은?

➡ DDI 연합은 DDI 표준을 지속적으로 개발, 홍보하여 전 세계 연구자들이 사회과학 자료의 생산, 접근, 이용이 용이하게 하는 것을 목표로 한다. 나는 DDI 연합의 워킹그룹에 속해 **DDI 표준을 정하는 회의에 참석하거나 DDI를 교육, 컨설팅**하는 역할을 한다.

○ DDI 연합에서 정회원과 준회원의 차이점이 있다면 무엇이며, 향

후 한국 통계청이 DDI 적용하여 통합시스템을 구축할 경우 관리와 운용에 있어 DDI 연합의 정회원이 될 필요가 있다고 생각하나? 있다면 어떤 점이 가장 큰 도움이 될 수 있나?

➡ ① DDI 연합의 정회원과 준회원의 차이점은 **의사결정의 투표권이 있는지 여부**이다. 정회원은 **연간 $2,500를 회비**를 내고 연합의 운영에 관여(의장을 선출, 예산의 지출의결 등)하거나 DDI 표준을 정하는 기술적인 결정에 투표권을 갖는다. 준회원은 회비를 내지 않고 의결권은 없지만 연합내의 **워킹그룹(Working group)에 소속**되어 DDI 표준을 정하는 기술적인 의견을 제시 할 수 있다.

② 한국 통계청이 DDI 표준을 적용한다면 DDI 연합에 가입하는 것은 필요하다고 생각한다. **호주통계청(ABS)은 국가적인 차원에서 정회원으로 가입**한 경우인데 정회원이 아니더라도 준회원으로 가입한다면 여러 **연구기관, 학계와 정기적인 회의를 통해 네트워크를 형성** 할 수 있고 최신의 기술과 정보를 공유할 수 있으며 다양한 워킹그룹에서 한국 통계청의 의견을 제시할 수도 있다.

※ DDI 연합관련 추가적인 설명은<참고2>

B. 한국 통계청에서 DDI 표준 적용방안

○ 한국처럼 분산형 통계제도인 나라에서 새롭게 DDI가 적용 될 때 개별 통계작성기관에서 반대할 경우 그 갈등을 어떤 방식으로 해결해 나갔는지 모범사례가 있다면?

➡ 아직까지 국가적인 차원에서 특정분야와 상관없이 모든 데이터를 통합한 프로젝트는 경험해 보지 못했다. 한국의 사업에 대해 자세히 알게 되면 사례를 알려주고 컨설팅 할 수 있다. 추후 한국의 사업계획서를 영문으로 작성해서 보내주길 바란다.

○ 현재 한국의 조사자료는 각 조사시스템에서 조사가 완료 된 후 최종적으로 Metadata 및 Microdata가 저장되는 방식으로 관리되고 있다. 이러한 환경에서 한국 통계청 내부에서는 DDI Codebook이 DDI lifecycle 보다 현실적으로 DDI를 적용하기 쉬울 것 같다는 의견이 다수인데 당신의 의견은 어떠한가?

➡ 각 버전별로 장단점은 있으나 현재 추세는 **DDI 2.0 대를 적용한 시스템을 3.0대로 변환하는 사업**을 추진 중에 있다. 예를 들어 도쿄대학의 경우 DDI Codebook 버전을 NESSTAR을 이용하여 구축하였지만 현재 ODF를 이용하여 Life-cycle 버전으로 전환하는 프로젝트를 진행하고 있으며 **DDI 연합에서는 현재 DDI4에 대해 논의가 진행** 중이다.

○ DDI를 적용하여 운용 중 한국통계청의 현실과 필요성에 맞게 각각의 element를 확장하거나 개별 element의 의미를 변형시켜 사용하는 것이 가능한가? 예를 들어 교육의 경우 국제표준분류가 있지만 나라마다 교육시스템(제도)가 다른 경우 같은 element일 경우라도 의미가 다른데 이런 경우 해결방법은?

➡ Element의 의미가 다르다 할지라도 Metadata에 라벨을 붙여서 자료이용자가 알 수 있게 남겨두면 된다.

○ 현재 DDI 연합에서 파악하고 있는 DDI버전 별 적용 비율은? DDI 3.0 기반의 대표적인 예로 미국 NORC data enclave와 캐나다 RDC Data Libration Initiative 로 파악하고 있음. 3.0 버전의 또 다른 사례와 2.0 버전의 참고할 만한 모범사례는?

➡ 3.0대 버전을 적용한 경우가 더 많다. 2.0대 버전을 적용한 경우는 스위스의 사회과학연구재단의 경우를 들 수 있다. 현재는 2.0적용 시스템을 3.0대로 변환하기도 한다.

C. OpenDataForge

○ OpenData Foundation INT.에 대하여 소개를 요청하면? 또한 OpenDataForge는 DDI Lifecycle 버전의 Tool인데 ODaF에서는 DDI Codebook버전의 Tool이 개발되어 있나? 없다면 개발계획은?

➡ ① OpenData Foundation[1]은 MTNA[2](Metadata technology North America)의 자회사이다. MTNA는 DDI, SDMX등의 메타데이터 표준을 적용한 응용프로그램을 개발하거나 데이터 관리 솔루션을 제공하는 IT컨설팅 회사이며 OpenData Foundation는 특히 DDI 메타표준을 기반으로 한 응용프로그램 개발부문을 담당하고 있다. Metadata Technology[3] 도 MTNA의 자회사이며 SMDX 표준을 적용한 소프트웨어 개발 및 컨설팅 부문을 담당

1) OpenData Foundation : http://www.openmetadata.org/
2) MTNA : http://www.mtna.us/
3) Metadata Technology : http://www.metadatatechnology.com/

② ODF 는 SledgeHammer, Caelum, Asmurex 로 구성된 툴 모음을 의미하며 DDI 버전을 가리지 않는다. **Codebook과 Life-cycle 버전 모두 ODF를 활용**할 수 있다. 만약 한국에서 ODF를 사용하여 시스템을 구축한다면 **ODF를 한국어 버전으로 제공 할 수도** 있다.

○ 국제적으로 OpenDataForge 이용하여 개발된 시스템은 얼마나 되나?

➡ 최근 NORC, 캐나다 통계청, 뉴질랜드 통계청과의 프로젝트에서 사용하였고 일본 도쿄대학교의 DDI Codebook에서 Life-Cycle 버전으로의 변환 프로젝트를 진행하고 있다.

○ OpenDataForge는 무료로 사용이 가능한가. 비용을 지불한다면 어떤 기준으로 요금이 책정되며 그 수준은?

➡ Edition 마다 다르다. 무료로 사용 할 수 있는 Freeware와 유료인 Pro/Enterprise 가 있으며 유료 edition인 경우 **1명의 유저가 평생 사용할 수 있는 비용은 $280 정도** 이다. Freeware는 테스트 버전 정도로 생각 할 수 있다.

※ OpenDataForge 관련 추가적인 설명은<참고3>

D. 추후 기술자문 절차

○ 향후 한국통계청에서 동 사업과 관련하여 지속적으로 기술자문을 받고 싶다. 공식적인 자문절차 및 방법은?

➡ 기본적인 컨설팅은 e-mail을 통해 가능하지만 **영문으로 작성된 사업계획서를 검토**한다면 한국통계청에 실질적으로 도움이 될 만한 컨설팅을 할 수 있을 것 같다. **DDI 관련 교육을 위하여 한국을 방문할 의사**도 있다. 최근에 독일에서 DDI 관련 교육을 속성으로 진행하였는데 일주일 정도 걸렸다.

IV 시사점 및 제언

□ 컨퍼런스 참석

○ “Data acccess, Data sharing, Data archiving” 이번 컨퍼런스에서 재차 강조된 목표. 구체적으로는 기관 간 협력(Cooperation) , 데이터의 표준화(Standardization), 조화&일치(Harmonization), 접근가능성(Accessibility), 상호운용성(Interoperability), 활용성(Utilization)의 증대가 언급됨

- 질 높은 메타데이터와의 결합을 통한 데이터 재생산의 다양한 사례들을 접할 수 있었고 그 결과 데이터의 생성과정(Lifecycle) 중 데이터의 생성과 분석 중심의 사고가 데이터의 관리와 공유로

전환되는 계기가 될 수 있었음

- 그런 측면에서 접근가능성, 표준화, 활용성이 부족한 데이터를 국제메타 표준을 적용해 시스템으로 구축하는 「마이크로데이터 통합시스템 구축」사업은 데이터의 공유와 접근성을 강조하는 세계적인 추세에 발맞춘 좋은 사례가 될 수 있을 것임

○ IASSIST 컨퍼런스에서 대부분의 참가자들은 서로 네트워크가 형성되어 있었으며 아시아권에서는 한국을 제외하고 일본인이 참석하였는데 많은 참가자들과 이미 알고 있었음 한국 통계청은 40회를 맞는 데이터 처리 및 관리, 제공의 국제적인 컨퍼런스에 이번에 처음으로 참석하게 된 것은 뒤늦은 감이 있음

- '14년 4월에 개최한 북미DDI 컨퍼런스(NADDI), '14년 12월 영국에서 개최하는 유럽DDI 컨퍼런스(EDDI)*는 IASSIST와 함께 DDI를 표준으로 하는 기관들이 대부분 참가하는 회의임

* EDDI(Annual European DDI User Conference) : 유럽 국가들의 DDI 표준 확산과 정보공유를 위한 컨퍼런스로 '14년은 6회째 임

- 향후 마이크로데이터 통합시스템을 DDI 표준을 적용하여 구축하면 IASSIST나 EDDI, NADDI 등에 주기적으로 참가하여 국제적인 기술과 사례를 접하고 인적 네트워크를 형성할 필요

□ Arofan Gregory 면담

○ DDI 관련 최신정보의 획득과 세계적인 흐름의 이해의 폭을 넓히고 국제적 네트워크 형성을 위해 DDI 연합에 예산을 수반하지 않는(회비없음) 준회원으로 한국 통계청이 참여하는 것을 적극적으로 검토할 필요 있음

○ DDI Codebook(2.X)을 적용할 것인가 Life-cycle(3.X)을 적용할 것인가 여부는 이미 「마이크로데이터 통합시스템 구축 1단계」사

업이 Codebook(2.X)을 기반으로 착수되어 논의의 필요가 없을지 모르나, 세계적인 추세(Codebook(2.X)에서 Life-cycle(3.X)로의 변환, DDI4에 대한 개발의 진행)를 감안했을 때 Life-cycle(3.X)로의 변환 혹은 새로 개발될 DDI4로의 변환을 염두 해 두어야 할 필요가 있으며 DDI4는 다른 메타데이터의 표준도 함께 고려되고 Life-cycle(3.X)보다 이용을 더욱 쉽게 한다는 측면에서 DDI4의 개발완료시 마이크로데이터 통합시스템에 적용을 고려해 볼 만함

○ 응용프로그램(Tool) 사용여부는 자체개발한 Tool을 사용하되 'Nesstar', 'DSD', 'Colectica', 'OpendataForge'의 무료버전을 이용한 테스트 등을 통해 Tool 사용 여부에 대한 검토를 지속할 필요

참고 1 컨퍼런스 상세 일정

일	시간	주제
6월2일(화) Workshop	9:00-12:00	Data Management &Curation: Lessons from Government, Academia, and Research Teaching an introductory workshop in digital preservation Data Visualization and R openICPSR: Why, When, What and How? Using open source tools for creation and sharing of DDI-Lifecycle metadata
	13:30-16:30	Advanced SDA Usage for Data Librarians Reshaping Data with R Introduction to Terra Populus: Integrated Data on Population and Environment Introduction to QGIS
6월3일(수) conference □	09:20-10:40	Plenary 1 : Barbara Entwisle
	11:00-12:40	A. Building Data Collections RDM B. Data Service Infrastructure for the Social Sciences and Humanities RI C. Integrated Data Discovery and Access DDT D. Teaching Data Management and Statistical Literacy DP
	13:40-15:20	E. Data Without Boundaries RI F. Harmonization, Thesauri and Indexing DDT G. Next-generation DDI RDM H. RDM across boundaries and disciplines RDM
	15:40-17:20	I. Collaboration and Networked Infrastructure RI J. Developing Meaningful Data Support Roles and Services DP K. Enhancing Restricted Access RDM
6월4일(목) conference □	09:20-10:40	Plenary 2 : Seamus Ross
	11:00-12:40	L. Encouraging Access to Confidential Data RDM M. Energizing the DDI Standard through Tools DDT N. Integrated Data Discovery and Access DP O. Trust and Data Sharing DP
	14:00-15:00	P. Big Picture Metadata DDT Q. Community Source meets Open Source DDT R. Innovative Approaches to Promoting Transparency in Research RDM S. RDM Tools and Services DDT
	15:20-16:40	Pecha Kucha
	15:20-17:40	Posters
	19:00-22:00	IASSIST Banquet at the Steam Whistle Brewery
6월5일(금) conference □	09:00-10:20	T. Metadata Portal for the Social Sciences RI U. National Data Managment Policies RI V. Planning and Assessing Research Data Services n DP W. Research Environments RDM
	10:40-12:00	Plenary 3 : Myron Gutman
	13:00-14:40	X. Data Curation and Workflow RI Y. RDM Activites within CESSDA RDM Z. Supporting Data Use DP
	15:00-16:40	Conference wrap-up!

DDT = Data Discovery and Tools
DP = Data Profession
RDM = Research Data Management
RI = Research Infrastructure

참고 2	DDI 연합(DDI Alliance Group)

□ 설립목적

○ DDI(Data Documentation Initiative)는 사회과학 데이터를 설명하는 Metadata에 대한 XML형식의 표준이며, DDI Alliance는 DDI 사양을 개발 및 배포, 국제표준화 하기위한 비영리단체로 궁극적으로 데이터의 생산과 접근, 이용을 용이하게 하여 DwB*를 실현하는데 기여함

* 국경없는 데이터 프로젝트(DwB) : Data without Boundaries, 유럽지역 내의 공식 마이크로데이터에 대한 국가간 접근성 강화를 위한 상호협력 이니셔티브

□ 조직

○ 집행이사회(Executive Board) : Alliance의 정책결정과 감독의 중심

① 의장(Chair) : 투표권이 있는 기관의 대표에 의해 선출된 집행이사회의 의장으로 연례회의를 주도하거나 집행이사회의 의장으로서 역할을 함

② 부의장(Vice Chair) : 투표권이 있는 기관의 대표에 의해 선출된 집행이사회 부의장

③ 집행이사(Executive Director): Alliance의 행정과 사무관리 책임자이며 사무국(secretariat)을 운영

○ 과학이사회(Scientific Board) : Alliance의 과학 기술적 의결의 중심으로 상설위원회인 과학위원회(Technical Committee)가 있음 과학위원회는 DDI 설명서(Specification)의 내용을 업데이트, 유지, 수정하는 역할을 함

○ 주최기관(Host Institution) : 조직운영의 기반을 제공하고 연합에 대한 직접적인 재정 및 법적 책임이 있는 기관 연합을 대신하여 각종 지적 재산권(라이센스, 등록 상표권 등)을 보유 현재

호스트 기관은 미시간 대학(University of Michigan)임

○ 워킹그룹(Working Groups) : 과학이사회(Scientific Boards)나 집행이사회(Executive Board)의 의결에 의해 설립되며 지정된 목표나 목적을 일정기간동안 수행하며 현재 10개의 워킹그룹이 운영되고 있으며 구성은 회원기관의 대표자(Member Representatives)가 되나, 회원기관에 속하지 않은 개인이 될 수도 있음

< Working Groups >

(2014년 4월 기준)

Group	연구
▪Administrative Data Group	조사자료의 DDI 적용에 대한 연구
▪Controlled Vocabularies[1] Working Group	통제어휘에 대한 연구, 확인된 DDI Element에 대한 통제어휘의 개발
▪Experimental Data Working Group	실험자료의 DDI 적용에 대한 연구
▪Qualitative Data Model Working Group	질적자료 데이터 교환을 위한 XML형식의 체계(scheme) 연구
▪RDF Vocabularies[2] Working Group	RDF 어휘의 연구
▪Survey Design and Implementation Working Group	조사설계 및 구현작업 그룹
▪Technical Committee	DDI의 실질적인 내용에 대한 요구에 따라 DDI specification을 강화하거나 유지, 창조 다른 메타데이터의 표준이나 커뮤니티 기반의 요구사항도 고려
▪Tools Catalog Group	DDI 표준의 사용을 촉진하기 위하여 DDI 작업을 도와주는 소프트웨어 및 관련 도구를 설명
▪Web Site Development Group	DDI 웹 싸이트(Site) 개발 그룹
▪DDI Developers Community	DDI 표준을 기반으로 한 소프트웨어 개발 그룹

1) **통제어휘(Controlled Vocabularies)** : 통제어휘는 웹상의 콘텐츠를 조직하고, 콘텐츠 객체와 연결된 메타데이터를 나타내기 위해 용어를 배정 통제어휘집의 모든 어휘는 분명하고 중복되지 않아야 하며 같은 용어가 다른 문맥에서 사용될 때에는 그 명칭을 명백하게 제한함

2) **RDF어휘(RDF Vocabularies)** : 웹상에 존재하는 기계 해독형(machine-understandable) 정보를 교환하기 위한 XML형식의 어휘로 상이한 메타데이터 간의 어의, 구문 및 구조에 대한 공통적인 규칙을 지원하며 메타데이터간의 효율적인 교환 및 상호호환을 목적으로 W3C(World Wide Web Consortium)에서 제안

※ Arofan Gregory는 「Qualitative Data Model Working Group」과 「Technical Committee」에 속하여 활동하고 있음

□ 운영

○ **(회원가입)** 개인이 아닌 단체 혹은 기관으로 가입하며 기관은 영리단체 비영리단체를 구분하지 않음

- 회원기관(Member Institutions) : 회비[1](Membership fee)를 납부하고 각 의사결정에 투표권이 있는 기관으로 회원대표(Member Representatives)는 회원기관의 대표자로 투표권[2]을 행사하며 하나 이상의 워킹그룹(Working Groups)에 속하여 다양한 의견을 제시

1) 회비(Membership fee)는 연간 $2,500임
2) 의장, 부의장 선출권과 과학이사회, 집행이사회의 의결권

- 준회원기관(Associate Memberships) : 회비를 납부하지 않는 기관으로 투표권은 없으나 각종회의나 워킹그룹(Working Groups)에 참석하여 의견을 제시할 수 있음

○ **(재정)** 회원기관의 회비와 주최기관 및 회원기관과의 계약에 의한 프로젝트 연구 보조금, 집행이사의 기금마련 활동으로 재정적 자립을 유지하며 추가적으로 프로그램 라이센스, 워크샵, 심포지엄, 출판 수입 등이 있음

○ **(연구활동)** 실질적인 DDI 사양(Specification)의 개발은 워킹그룹 중 하나인 과학위원회(TC)의 연구, 토론을 통해 이루어지며 구체적인 연구개발의 절차*가 연합의 정관으로 규정되어 있어 DDI 사양(Specification)의 변경은 과학이사회(Scientific Board)의 투표에 의해 확정됨

* 연구개발의 절차 규정 : "Standards Development and Review Process and Procedure"

□ 소속회원

○ 36개의 회원기관과 4개의 준회원기관이 참여 중(2014년 4월 기준)

< Member Institutions & Associate Members >

회원 (36개 기관)	대학 (11)	• University of Alberta, Canada • University of California, Berkeley • Cornell University (CISER) • University of Toronto Libraries • University of Toronto Scholars Portal • University of Washington, Center for Studies in Demography & Ecology (CSDE) • University of Guelph • Stanford University • University of Minnesota, Minnesota Population Center • Massachusetts Institute of Technology (MIT) • Survey Research Operations, University of Michigan
	비영리 단체 연구기관 국제기구 (14)	• World Bank, Development Data Group (DECDG) • Institute for Quantitative Social Science (IQSS) at Harvard University • Institute for Social and Economic Research (ISER) • Institute for the Study of Labor (IZA) • Inter-university Consortium for Political and Social Research (ICPSR) • Food and Agriculture Organization (FAO) • GESIS - Leibniz Institute for the Social Sciences • Research Data Centre of the German Federal Employment Agency, Institute for Employment Research (IAB) • Roper Center • German Institute for International Educational Research (DIPF) • German Socio-Economic Panel Study (SOEP) • Swiss Foundation for Research in Social Sciences (FORS) • Epidemiology France, Aviesan - ITMO Santé Publique • Centro De Investigaciones Sociologicas (CIS), Spain
	정부기관 (11)	• Australian Bureau of Statistics (ABS) • Statistics New Zealand • Swedish National Data Service (SND) • Eurostat • National Institute of Statistics and Economic Studies (INSEE) • Norwegian Social Science Data Service (NSD) • United Kingdom Data Archive • Australian Data Archive (ADA) • Danish Data Archive • Data Archiving and Networked Services (DANS), The Netherlands • Finnish Social Science Data Archive
준회원 (4개 기관)		• Centre for Longitudinal Studies, Institute of Education, University of London • Colectica • Open Data Foundation • U.S. Bureau of Labor Statistics

참고 3	MTNA & OPENDATAFORGE

□ Metadata technology North America(MTNA)

○ SDMX 및 DDI 메타데이터의 표준을 활용하도록 돕는 응용프로그램 개발 및 컨설팅 전문업체로 각국의 통계청, 연구기관, 중앙은행, 데이터 아카이브, 국제기구 등과 메타데이터 표준을 기반으로 한 오픈데이터 네트워크 구축 사업을 진행하고 있으며 DDI 연합의 준회원으로 등록되어 있음

○ (개발프로그램) OpenDataForge와 OM Survey Manager[1], Data Management Plan (DMP) Editor[2]등을 개발

1) OM Survey Manager : DDI Life-cycle기반 조사 및 변수 메타 데이터를 관리 할 수 있는 오픈 소스 데스크톱 응용 프로그램
2) Data Management Plan (DMP) Editor : 데이터 관리 계획을 준비하는 연구자를 지원하는 오픈 소스 데스크톱 응용 프로그램

○ (서비스 및 컨설팅)

①전문가 자문 및 기술 지원 : 기업 데이터 관리 전략 수립, 메타 데이터 표준과 관련된 모범 사례 채택, 데이터 관리 인프라에 대한 기술 사양 개발, 통계 및 과학 데이터 관리 기술

②응용프로그램 개발 :통계와 과학적인 데이터 관리 솔루션, 솔루션 생산, 보관, 보급, 분석을 포함, 웹 기반 데이터 전송 솔루션 (웹 서비스, 포털, 보안 환경), 표준 기반의 데이터

③교육 및 지원 : 데이터 문서 이니셔티브 (DDI), 통계 데이터 및 메타 데이터 교환 (SDMX), 일반 통계 비즈니스 프로세스 모델 (GSBPM), 일반 통계 정보 모델 (GSIM), 일반 데이터 및 메타 데이터 관리 지원

- 그 외 클라우드 기반 인프라의 구축, 호스팅 및 관리 등이 있음

○ (주요수행사업)

< MTNA의 데이터 시스템 구축관련 프로젝트 >

	프로젝트	내용
1	META-SSS ICPSR NORC	- NORC, ICPSR[1]와 협력으로 미국 일반사회조사[2](GSS)와 미국 선거관리연구[3](ANES)에 대한 웹 포털의 새로운 메타데이터의 개발을 지원하며 연구의 전 진행과정에 걸쳐 메타데이터의 재사용과 새로운 메타데이터 기반의 워크플로우 및 모범사례를 공동개발 - 2014년 중반 완료 예정
2	SNZ CMS Statistics New Zealand TATAURANGA AOTEAROA	- 뉴질랜드 통계청과 함께 새로운 분류관리시스템(CMS)을 구축 기존 데이터베이스를 교체하여 새로운 데이터 관리체계로 통합 하는 프로젝트이며 2014년 2분기에 완료될 예정
3	NORC Data Enclave NORC at the UNIVERSITY of CHICAGO	- NORC의 데이터 Enclave(민감한 데이터 세트에 대한 보안 원격 접근을 가능하게 하는 가상환경)의 설립, 유지, 보수 및 개발을 지원
4	BRECKENHILL Canada THE•AIM•GROUP	- 캐나다 통계청의 CRDCN 데이터 저장소 관리도구 개발의 기술지원 - CRDCN는 캐나다 연구 데이터 센터 네트워크로서 RDC 중앙 데이터 저장소에서 설문조사 데이터 셋 및 문서파일을 관리하고 CRDCN을 통해 배포
5	NSF SED Editor NORC NSF	- NORC 및 NSF NCSES[4]와 공동으로 SED(Survey of Earned Doctorate) 조사에 대한 데이터 및 문서의 자동화 및 DDI 요소 메타데이터 의 수집을 용이하게 하기위한 응용프로그램을 구축 그 결과는 NSF 데이터 저장소에 저장 - NSF-SED Editor는 RCP(원격 복사 프로토콜) 응용프로그램이며 OpenMetadata 프레임워크를 활용 2013.6월 완료

1) ICPSR(The Inter-University Consortium for Political and Social Research) : 1962년에 설립된 미국 최대 사회과학데이터보관소로써 1400년대 부터 세계 각 국에서 수집된 5,800여개의 조사자료(survey data)와 100,000여개의 수치자료(numeric data)를 소장, 현재 전 세계 500여 개의 대학이 컨소시엄으로 구성되어 있음

2) GSS(General Social Survey) : 미국의 대규모 사회조사로 NORC에 의해 진행되며 미국의 인구통계학적 특성 및 주민의 태도에 대해 조사함 면접조사 18세 이상을 대상으로 1972년이후 매년 실시

3) ANES(American National Election Studies) : 미국선거관리연구기관으로 미국 대통령선거 전후에 미국 유권자의 태도를 조사

4) NCSES (National Center for Science and Engineering Statistics)

○ **(협업기관 및 고객)** NORC[1] , Colectica기관 등과 전략적 제휴를 맺고 있으며 주요 프로젝트 고객으로는 ABS(호주통계청), CRDCN[2], DDI Alliance, NSF[3], 뉴질랜드 통계청 등이 있음

1) NORC(National Opinion Research Center) : 1941년에 설립된 미국 시카고 대학의 사회연구기관 미국의 주요한 사회조사 및 국제적인 프로젝트를 다수 진행

2) CRDCN(Canadian Research Data Centre Network) 2000년 이후 구성된 캐나다 연구 데이터 센터 네트워크로 캐나다 통계청과 각 연구기관이 협력하여 캐나다의 사회, 경제, 건강에 대한 연구를 진행

3) NSF(National Science Foundation) : 미국국립과학재단

□ OPEN DATA FORGE

○ ODF는 DDI 메타데이터 표준을 촉진하기 위해 MTNA에 의해 개발된 데스크톱 및 클라우드 기반 솔루션의 모음

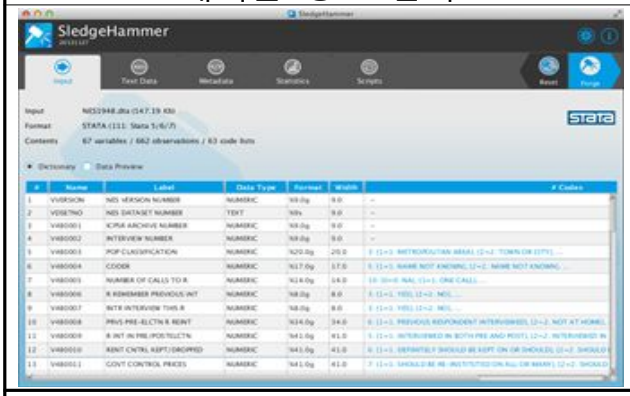
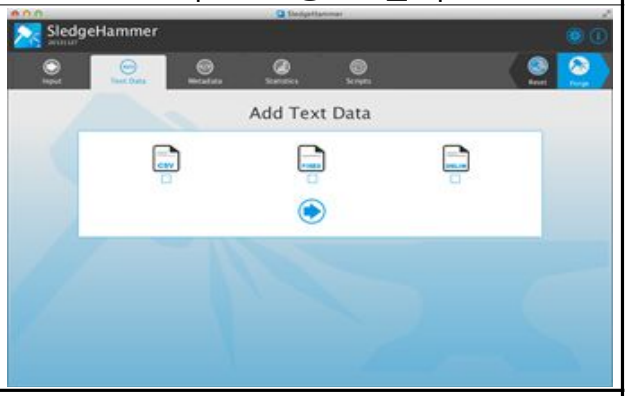
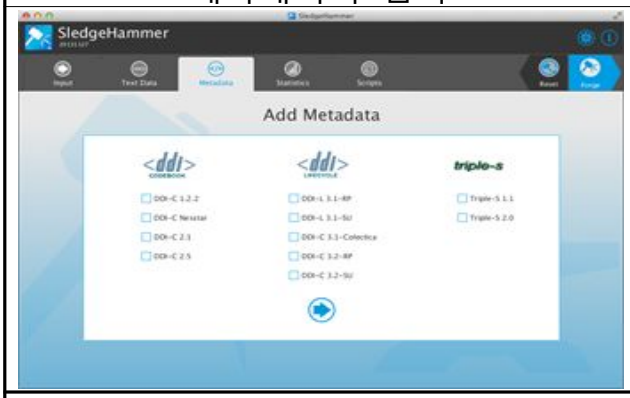
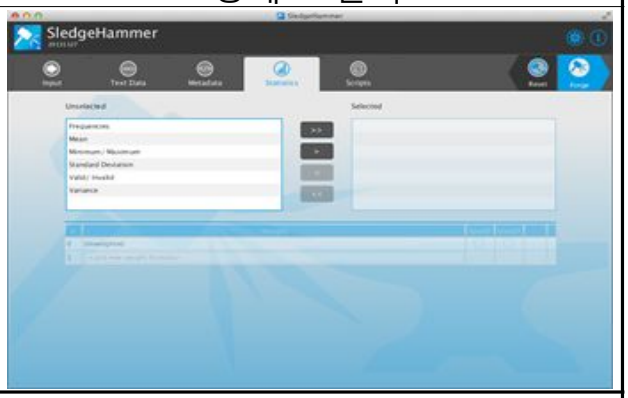
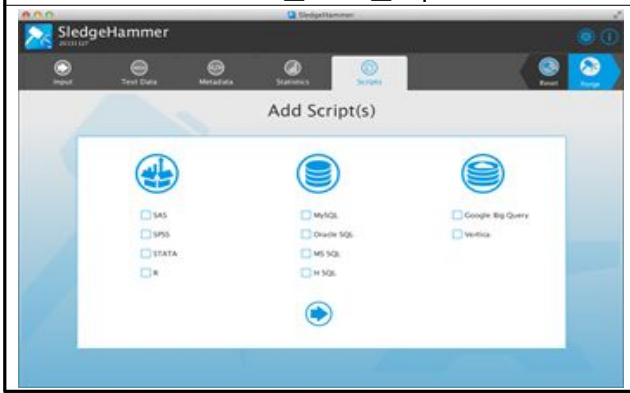
- 독점적 형태의 데이터 및 메타데이터를 잠금 해제하여, 통계 및 과학적 데이터의 변형 및 가공을 용이하게 만드는 장점이 있고 ①SledgeHammer(데이터/메타데이터의 변환) ②Caelum(보고/ 출판) 및 ③Asmurex(ASCII 다중 레코드 Exctraction 도구)로 구성

○ (비용)

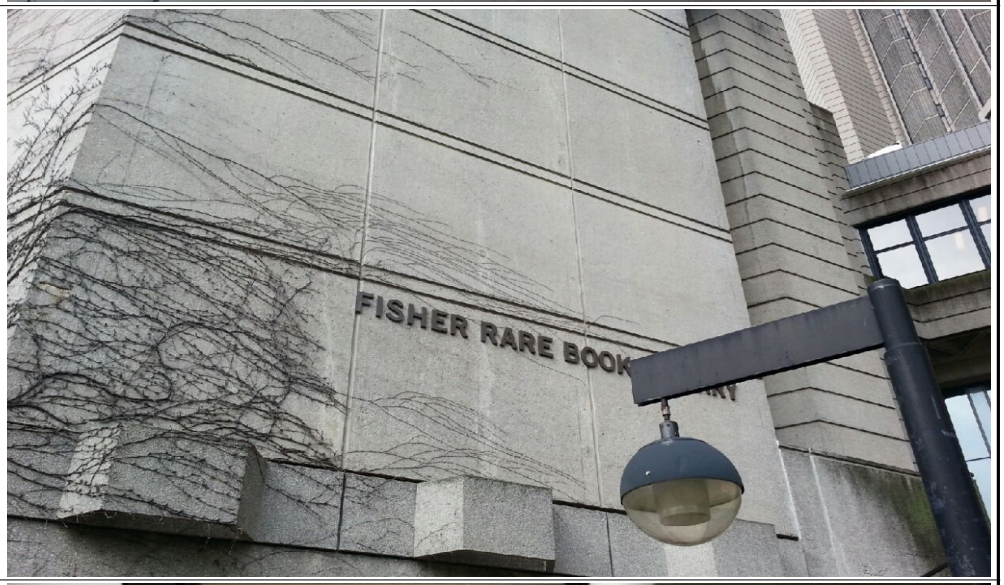
Edition	Freeware	Pro/Enterprise
관측범위	100 변수 / 1,000 관측치 500 변수 / 5,000 관측치	데이터 사이즈에 제한이 없음
특징	-	가중 통계, 데이터베이스 형식 및 스키마의 추가 분류 자동 조정 등
비용	무료	$29/3개월 or $99/연간 1명의 유저(User) 영구사용 : $279

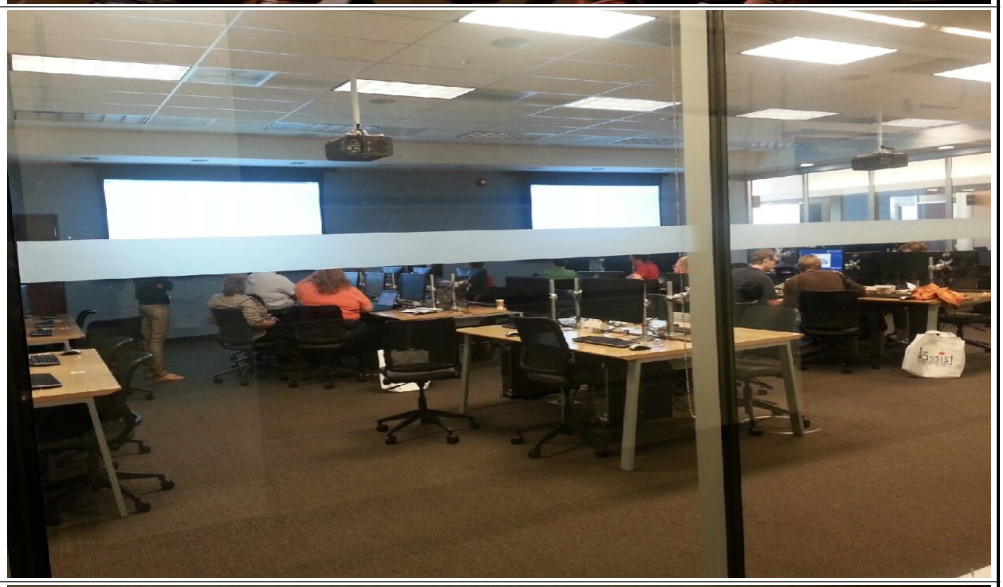
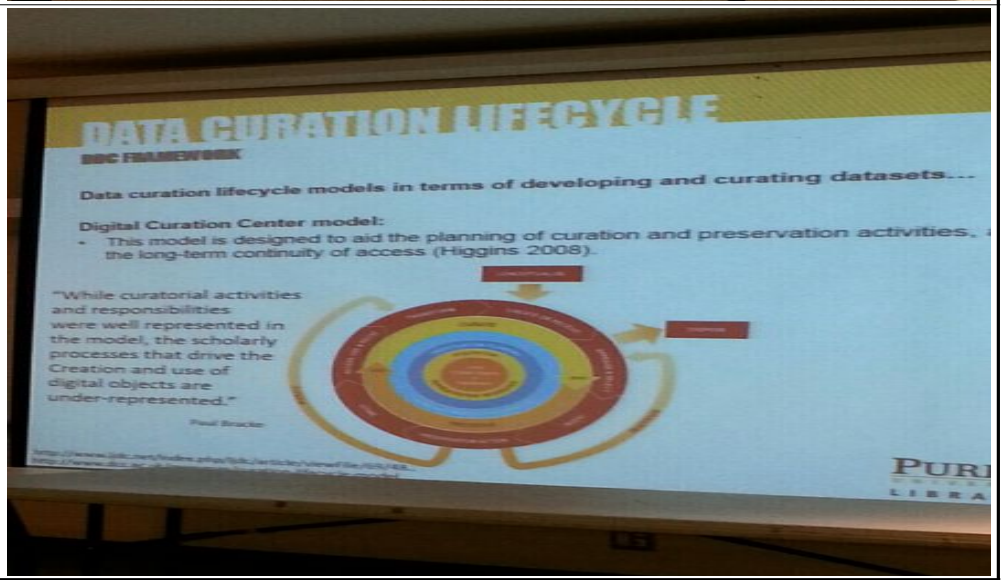
□ 화면구성

DataForge_SledgeHammer 화면구성

초기화면	자료 입력
테이블 정보 입력	**텍스트 정보 입력**
	
메타데이터 입력	**통계표 출력**
	
스크립트 입력	
	

참고 4	컨퍼런스 참석 사진자료

토론토 공항의 온타리오 자연사 박물관 홍보물	
워 크 샵 장 소 인 토론토 대학 도서관	
IASSIST 컨퍼런스 개시 현수막	

컨퍼런스 기조 연설 시작 전	
워크샵 진행중	
컨퍼런스 세션 진행중