

ASIA DATA WEEK 2017 参加報告

参加者：大谷周平（JPCOAR 広報普及作業部会員・琉球大学附属図書館）

日程：2017年11月20日（月）～24日（金）

場所：LOTTE Hotel Jeju

35 Jungmungwangwang-ro72beon-gil, Saekdal-dong,

Seogwipo, Jeju-do, 大韓民国

イベントページ：<http://www.asiadataweek.org>



写真 1 イベント会場

1. 概要

2017年11月20日（月）～24日（金）にかけて韓国済州にて ASIA DATA WEEK 2017 (AWD) が開催された。AWD は KISTI (Korea Institute of Science and Technology Information) が主催するイベントで、ビッグデータや IoT (Internet of Things) などのテーマとあわせて研究データ管理に関連したセッションも複数設けられていた。筆者は研究データのメタデータをテーマとしたセッションで JPCOAR の活動及び JPCOAR スキーマに関するプレゼンテーションを行った。各セッションの概要や KISTI 関係者との情報交換について報告する。

2. プレゼンテーション概要

「Overview of JPCOAR and JPCOAR Schema : A brand-new metadata schema for Open Science in Japan」と題して、日本のリポジトリの現状、JPCOAR の概要や研究データに関する取組を中心としてタスクフォースや作業部会の活動紹介、JPCOAR スキーマについて策定の経緯・コンセプト・具体的な記述例を報告した。

プレゼンテーション資料

<http://id.nii.ac.jp/1458/00000045/>

質疑では以下が寄せられた。

Q. 日本で 500 を超える機関リポジトリが構築された要因は何か。

A. CSI 事業、学位規則改正による博士論文のインターネット公開、IRDB や JAIRO Cloud 等のインフラが整備されていることを挙げた。

Q. junii2 や JPCOAR スキーマは各大学の機関リポジトリに対して強制力があるのか。

A. IRDB を中心としたメタデータ流通の状況を説明しデファクトスタンダードになっていること、IRDB が JPCOAR スキーマによるハーベストを開始しても当面は junii2 によるハーベストも平行して行うことを説明した。

Q. JPCOAR スキーマと DataCite Schema の相互運用性について

A. 例えば contributorType は JPCOAR スキーマ独自要素だが、語彙については DataCite Schema に準じており、相互運用を念頭に置いている。

後述するように韓国でも研究データ管理のためのメタデータを検討しているので、メタデータの普及や国際的な相互運用性の部分に関心を持っていただいたようである。



写真 2 メタデータセッション会場

3. 各セッション概要

Metadata Forum for Research Data

筆者の報告に先立ち、韓国における研究データのメタデータ検討ワーキンググループの活動が報告された。このワーキンググループは2016年に活動を開始し、リポジトリ管理者、研究者および大学図書館員計13名で構成されている。国際的な研究データのためのメタデータについて調査を行うとともに、韓国における研究データ管理のメタデータ標準化の活動を行っている。2017年3月にはワーキンググループで作成したメタデータ標準が韓国情報通信技術協会ⁱの規格 TTA.KO-10.0976 として採用されている。TTA.KO-10.0976 はコレクション・データセット・ファイル・リポジトリの4つについてメタデータを定義しており、DCMI・韓国ダブリンコア・re3data.org・DataCite が関連情報源として挙げられている。データセットにおけるメタデータ要素例を表2に示す。その後、国際的な動向として DataCite Schema の4.0、4.1における変更点も紹介された。

ID	D10				
요소 이름	Contributor	표시 상수	Contributor	요소 유형	Property
요소 설명	▫ 데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함				
활용 예시	▫ EPSRC - Engineering and Physical Sciences Research Council				
관련 출처	▫ DataCite-Property (ID 7, Contributor) 참조. https://doi.org/10.5438/0010				

ID	D10				
要素名	Contributor	表示	Contributor	要素タイプ	Property
要素説明	データセットの収集、生産、予算支援などに関係する人物や機関等を記述する。				
活用例	EPSRC - Engineering and Physical Sciences Research Council				
関連情報源	DataCite-Property (ID 7, Contributor) を参照. https://doi.org/10.5438/0010				

表1 TTA.KO-10.0976 の要素例と日本語訳。日本語訳は筆者によるもの。

Seminar on Research Data Management and Data Management Plan

研究データの定義や特徴から始まり、Data Management Plan (DMP) の概要や国際的な動向紹介と医学・海洋学・生態学・鉄道分野における韓国での実践例についての報告が行われた。

その中で韓国におけるオープンアクセスの状況について医学系を中心とした報告も行われた。韓国国内では44個の機関リポジトリが設置（うち8個が医学系の機関リポジトリである）され、その多くは Dspace で構築されている。また、リポジトリとは別に KoreaMed Synapseⁱⁱ という電子ジャーナルプラットフォームで医学ジャーナルの公開・アーカイブ、メタデータ流通促進がはかられており、DOI 付与も行っていることが紹介された。

Forum on Data Leadership and Stewardship

研究データ管理において、国内で先行している分野（極地学・気象学）の事例報告、政策面などにおける国際的な動向やオーストラリアの事例が紹介された。国外におけるオープンサイエンスのコミュニティ活動のベストプラクティスとして JPCOAR の活動も紹介された。これから韓国で研究データ管理の体制を整えていくなかで、必要とされるデータセンターなどのインフラ、オープンサイエンス促進のコミュニティ活動、政策レベルの取組の必要が議論されていた。

4. KISTI 関係者との情報交換

フォーラム前後に KISTI で研究データ管理を推進する関係者と情報交換を行った。インフラ面について 2 点印象に残ったので紹介したい。

1 つはデータリポジトリと研究データ公開システムである。2012 年に P-Cube というデータリポジトリシステム (Fedora ベース) を開発 (2014 年に DataNestⁱⁱⁱ という名称に変更) している。現在は各機関に設置された DataNest からメタデータを集約し研究データ公開プラットフォームとなる National Open Data Service (NODS) を開発しているとのことであった。各機関に設置された DataNest では柔軟にメタデータスキーマを設定でき、ハーベストと検索機能のためにメタデータ間のマッピングをそれぞれ行う運用を予定しているとのことであった。NODS は正確な時期は未定だが 2018 年には公開されると思われる。

The screenshot shows the DATANEST website interface. At the top, there is a search bar with the text '통합검색' and 'GO', and a '상세검색' link. The main header includes the DATANEST logo and the KISTI logo (한국과학기술정보연구원). Below the header, there is a 'LOGIN' section with fields for '이메일 주소' and '패스워드', and a 'Login' button. The main content area features a large banner with the text 'WELCOME TO DATANEST RESEARCH DATA REPOSITORY'. Below the banner, there are three navigation links: 'WHAT IS DATANEST?', 'HOW TO USE', and 'HOW TO SUBMIT'. The main content area is divided into several sections: 'ORGANIZATIONS' with a list of items, 'Browse for Collections' with a pie chart showing 93.7%, 'NEWS', and 'RECENTLY ADDITION'. The left sidebar contains 'HOME', 'BROWSE', 'MY DATANEST', and 'D-DRIVE' sections.

図 1 Datanest 画面

もう1つは研究プロジェクト管理システムである。KISTI による National Science & Technology Information(NSTI)^{iv}というシステムが構築されており、研究者情報や研究設備などの情報が登録されているだけでなく、各研究助成金の情報も集約されているとのことであった。また、研究成果についても論文だけでなく特許や開発されたソフトウェアなども登録されている。NSTI は 2006 年に開始され長期間運用されており、研究者にも定着したシステムということである。研究データ管理を推進していく上ですでにこのような情報が集約されている意義は大きいと考える。

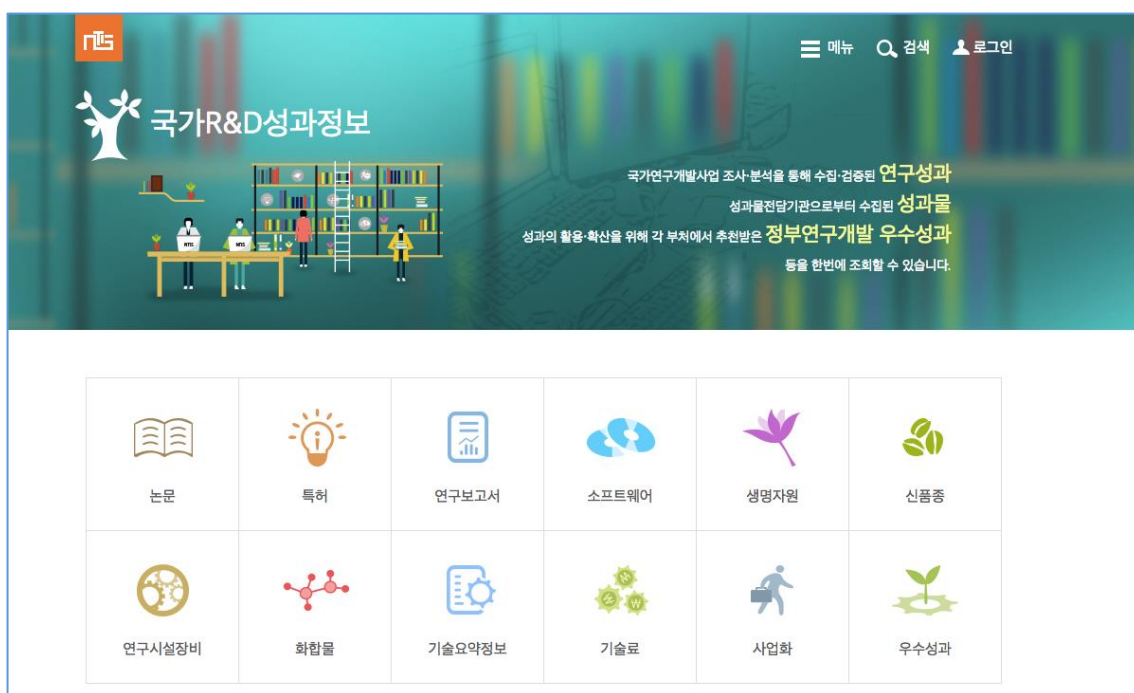


図 2 NTIS の研究成果検索画面

ADW は 2017 年が初めての試みとのことであり、研究データ管理に関するセッションは国内向けセッションであった。参加者はほぼ韓国の研究者でしめられ報告・質疑は韓国語を中心として行われた。しかし、2018 年以降の開催も検討しており、その場合は研究データ管理のセッションを JPCOAR も含めたアジアのオープンサイエンス・オープンアクセスコミュニティとも連携した国際的な企画として展開していきたいということであった。

5. 所感

筆者が現地で情報交換をした範囲ではあるが、参加者はすでに研究データ管理を実践している主題研究所に所属する研究者や研究支援業務に携わる人が中心で、大学からの参加者はごくわずかに留まっていた。一方でインフラ面の整備や助成機関への働きかけなどは進んでおり、今後 KISTI を中心として研究データ管理・オープンサイエンスへの取組が進展していく機運は感じられた。前述した NODS の公開なども含めて今後の動向に注目するとともに情報交換を行っていきたい。

6. 謝辞

プレゼンテーション作成にあたって、メタデータ普及タスクフォースメンバー各位および研究データタスクフォースの南山泰之氏より情報提供およびアドバイスをいただいた。この場を借りて感謝申し上げます。また、出張事務を担当いただいた JPCOAR 事務局および国立情報学研究所学術コンテンツ課の皆さまにも感謝申し上げます。

以上

i http://committee.tta.or.kr/data/standard_view.jsp

ii <https://synapse.koreamed.org>

iii <http://kdr.kisti.re.kr/cube/index.do>

iv <https://www.ntis.go.kr/en/GpIndex.do>