

平成 27～28 年度 国立情報学研究所実務研修報告書
欧州原子核研究機構 (CERN; スイス) におけるオープンアクセスに関する長期研修

提出日:平成 29 年 3 月 21 日

所属:附属図書館利用支援課資料サービス掛専門職員

氏名:大村 明美

1. 期間 平成 28 年 3 月 11 日～平成 29 年 3 月 2 日

2. 場所 欧州原子核研究機構 (CERN)

3. 概要

CERN が運営する高エネルギー物理分野の文献情報データベース Inspire について、日本人研究者の執筆論文を中心に著者情報の品質管理業務に従事することにより、オープンアクセスに関する先導的な知識と技術を習得する。具体的には、研究者同定のための国際的な識別子 ORCID (Open Researcher and Contributor ID) 付与に関連する調整、高エネルギー加速器研究機構 (KEK) 等との協力、他のデータベース (arXiv 等) との連携を行う。

また、研修先の CERN の部署が担当している SCOAP3 (Sponsoring Consortium for Open Access Publishing in Particle Physics; 高エネルギー物理学分野の査読付きジャーナル論文のオープンアクセス化を目的とした国際連携プロジェクト) をはじめとするオープンアクセス事業に関する情報を収集する。

さらに、研修期間中に欧州の他の教育研究機関を訪問して、欧州におけるオープンアクセスに関する情報を収集する。

4. 具体的な内容

(Inspire 研修)

- Inspire 全般に関する知識の習得
 - データ構造
 - ✓ 論文レコード HEP、著者レコード HepName、機関レコード Institution、雑誌レコード Journal、実験レコード Experiment 等からなる。
 - ✓ これらのレコードが関連付けられて著者ごとの執筆論文プロフィール Author Profile が生成される。
 - レコード/資料種別ごとのメタデータ記述方法、処理ツールの使用方法

(Inspire 日常業務)

論文レコード HEP の著者、著者所属機関、共著者、分野、関係する実験等のメタデータを自動分析することにより、著者の表記の揺れを吸収して、著者ごとの執筆論文プロフィール Author Profile が生成される。これが正確に機能するためには、HEP が正確にキュレートされていなければならない。

また、Author Profile は最終的に著者レコード HepName にリンクしなければならないが、そのためには HepName を正確かつ可能な限り最新状態にしておく必要がある。

さらに、論文レコード HEP の引用文献情報を充実させることによって、関連する論文間でリンクが形成される効果がある。

- 論文レコード HEP の処理業務 (3 種類のツールによる更新・新規登録)

- ツール「BibEdit」による個別更新・新規作成
- ツール「Multiedit」による複数レコードの一括更新
- スクリプトによる PDF からのメタデータ自動作成
- 著者情報 (HepName、Author Profile 等) の充実
 - 著者レコード HepName の更新・新規作成および Author Profile とのリンク形成
 - 日本関係機関レコード Institution の更新・新規作成
 - 日本関係実験レコード Experiment の更新・新規作成

(SCOAP3 対応)

SCOAP3 論文のデータを正確にキュレートすることにより、正確な機関別統計出力を可能にした。

- 日本人著者による SCOAP3 論文の Inspire データ更新 (特に著者所属機関の正確化)
- 日本関係の Inspire 機関データ Institution の更新・新規作成、サンプル調査 (本文 PDF データとメタデータの不整合チェック)

(Inspire 日本人著者名寄せ) *別添詳細報告書 1 を参照

Inspire 日常業務著者情報の充実について、科学研究費助成事業データベース (KAKEN) との照合 (著者レベルおよび論文レベル) によって作業をさらに進めた。

- KAKEN データに基づく KAKEN ID、著者漢字表記の追加
- CiNii Dissertations 等の調査に基づく学位情報の追加
- Researchmap や所属機関が提供する研究者情報データベース、研究者個人のウェブサイト等の調査に基づく所属機関履歴情報の追加

(会議等)

- CERN/NII/KEK 遠隔会議 (4/15, 5/12, 6/2, 6/27, 11/24, 1/13, 2/16)
- 安達副所長との打ち合わせ (10/14; CERN)
 - Inspire 日本人著者データ名寄せ作業の現状について報告した。
- KEK 菊谷教授との打合せ (2/6-8)
 - Inspire の質向上のために日本側で可能な協力について検討した。

(その他)

- Library Desk 業務
- CERN Library 見学 (4/1) (高エネルギー加速器研究機構 (KEK) 研究協力部国際企画課長・原文絵氏来訪)
- Inspire システムプレゼンテーション聴講 (4/11, 18, 25)
- Inspire プレゼンテーション (OA 担当者) 聴講 (5/2)
- Book Presentation (著者による書籍紹介) 聴講 (5/3)
- CERN Exhibition 見学 (5/9)
- Book Presentation (Proceedings of ThEC13: Thorium Energy for the World) 聴講 (6/13)
- Inspire システムプレゼンテーション聴講 (6/20)
- Library Science Talk (Alexander Grossmann "New perspectives in scientific publishing") 参加 (6/20 WTO)
- International and Interdisciplinary workshop on Active Data Management Plans 参加 (6/28-30 CERN)
- Inspire システムプレゼンテーション聴講 (7/11)
- CERN 実験施設見学 (CMS) (7/28)

- Library Science Talk 参加(9/12, ITU)
Courtney Mumma, Internet Archive (San Francisco, United States) "Capturing moments in time - Collecting our world's memory together"
- Presentation 聴講(9/28, CERN)
Johnny Mariethoz (RERO) "From MARC-XML to JSON-LD (=JavaScript Object Notation for Linked Data): A New Invenio Data Model for Bibliographic Objects and Beyond"

5. 出張 *別添詳細報告書 2 を参照

- 1) iPRES (International Conference on Digital Preservation) 2016 参加(10/3-6, Bern)
ベルンで開催された iPRES 2016 に参加し、デジタル情報の長期保存に関する情報収集を行った。
- 2) ETH Zürich 訪問調査およびワークショップ参加(10/19, Zürich)
ETH Zürich を訪問し、研究データ管理に関する調査を行った。また、同大学の大学院生・若手研究者を対象として開催された研究データ管理に関するワークショップ("Data Management in Research – Why and How?")に参加した。
- 3) 英国大学図書館等の訪問調査(10/26-11/3)
University of Edinburgh, National Library of Scotland, University of St. Andrews, University of Sheffield, University of Leeds, University of Oxford, University of Cambridge を訪問し、デジタルコレクションの管理(IIIF への対応を中心に)、研究データ管理に関する調査を行った。
- 4) Bibliothèque nationale de France 訪問調査(11/14-16, Paris)
BnF を訪問し、IIIF の取り組みについて、すでに導入している海外機関(Bibliothèque nationale de France、Europeana)の事例を調査した。
- 5) ドイツ研究機関等の訪問調査(2017/1/24-28, Regensburg, Munich)
University Library Regensburg, Bavarian State Library, Max Planck Digital Library を訪問し、ドイツにおけるオープンアクセスの現況(ゴールド OA の状況、グリーン OA の状況、データリポジトリの構築等)を調査した。
- 6) IDCC17 参加(2017/2/19-24, Edinburgh)
デジタルキュレーションに関する国際会議である IDCC17 (12th International Digital Curation Conference) に参加して情報収集を行った。

6. 本研修成果の本学への活用方法・フィードバック

欧州原子核研究機構(CERN)が運営する高エネルギー物理分野の文献情報データベースとして定評があり、同分野の研究者が世界的に多く利用している Inspire の質向上プロジェクトに参加した。我が国の国立情報学研究所(NII)と高エネルギー加速器研究機構(KEK)が CERN と国際連携協定を結び、共に進める事業であり、NII を通じて日本の大学図書館職員として、CERN の国際的な環境の中に実際に身を置き、国際プロジェクトに関わることができたことは大変貴重な経験となった。また、他に欧州の高等教育・学術研究機関を多数訪問して、オープンアクセスの現況に関する情報収集・意見交換を行うことができた。

担当者と意見を交わす中で共通していた認識は、オープンアクセス、オープンデータの意義が国際的に認められつつあること。しかし一方で、様々な課題を一度に解決できる決定的な方法はなく、各機関と

も互いに情報を密に共有して学び合い、試行錯誤しつつも現実的な取り組みによってサービス向上に努めていることがわかった。

研究データ管理や研究データリポジトリの運用に関して本学でも多くの課題があるが、本学の現況に適した対策を着実に進める上で、今回の研修で得た知見を活用して貢献したい。

以上

平成 28 年度 国立情報学研究所実務研修 報告書
 高エネルギー物理分野文献情報データベース Inspire
 日本人著者情報充実プロジェクト報告

提出日:平成 29 年 3 月 15 日

所属:京都大学附属図書館

氏名:大村 明美

1. 欧州原子核研究機構(CERN)¹

- 英語名 European Organization for Nuclear Research、フランス語名 Organisation Européenne pour la Recherche Nucléaire
- 世界最大規模の素粒子物理学の研究所。CERN Convention により 1954 年に発足。
(参加国)
- メンバー国(member states): 設立時の参加国はベルギー、デンマーク、フランス、ドイツ、ギリシャ、イタリア、オランダ、ノルウェー、スウェーデン、スイス、英国、ユーゴスラビアの 12 カ国。その後オーストリア、スペイン、ポルトガル、フィンランド、ポーランド、チェコスロヴァキア、ハンガリー、ブルガリア、イスラエル、ルーマニアが参加し、現在 22 カ国。²
- 協力国(associate members): セルビア、キプロス、トルコ、パキスタン、ウクライナ、インド
- オブザーバ国(observer states): European Commission、日本、ロシア、UNESCO、アメリカ
(研究施設)
- スイスのジュネーブ郊外、フランスとの国境地帯に所在。主なキャンパスは Meyrin(スイス領とフランス領にまたがるメインキャンパス)と Prévessin(フランス領)。
- 全周 27km の円形加速器、大型ハドロン衝突型加速器(Large Hadron Collider; LHC)が国境を横断して地下に設置されている。さらに CMS、ATLAS、ALICE 等の検出器が設置されている。

2. 文献情報データベース Inspire <http://inspirehep.net/>

2.1. 概要

- 高エネルギー物理分野の文献情報データベース(論文件数 120 万件以上)。オープンアクセス論文は本文データも提供している。
- 高エネルギー物理分野の研究機関である CERN、DESY(Deutsches Elektronen-Synchrotron; ドイツ電子シンクロトロン)、Fermilab(フェルミ国立加速器研究所; 米国)、IHEP(High Energy Physics of the Chinese Academy of Sciences; 中国科学院高能物理研究所)、SLAC(SLAC 国立加速器研究所; 米国)の協力により国際共同運営されている。
- 担当者向けヘルプ³がインターネット上に整備されている。今後 1 年以内にシステム移行を予定している。

¹ 本実務研修は、CERN、国立情報学研究所、高エネルギー加速器研究機構の間で締結された国際連携協定(Collaboration Agreement for the development of advanced information services, for the benefit of the Japanese and global High-Energy Physics community; 2015 年 10 月締結)に基づき、大学図書館職員を派遣するものである。

² ユーゴスラビアは 1961 年脱退。チェコスロヴァキアは 1993 年チェコ、スロヴァキアとして再参加。

³ Knowledge Base <https://labs.inspirehep.net/internal-help/knowledge-base/>

2.2. データベース構造とレコード管理ツール

- Inspire を構成するレコードコレクションとして、HEP(論文)、HepNames(著者)、Institutions(機関)、Conferences(会議)、Experiments(実験)、Journals(雑誌)等があり、レコード間にリンクが形成されている。すべてのレコードはMARC形式で記述されている。HEPが自動分析され、著者ごとの執筆論文プロフィール Author Profile が生成されている。
- HEP へのレコード追加は、arXiv や出版社等のサーバからの自動ハーベスト(プレプリント版のレコードと出版社版のレコードは自動的に統合)、利用者からのデータ提出、職員による手入力により行われる。
- レコード管理ツールとして、Bibedit(レコードの新規登録、更新、削除等)、Bibmerge(重複レコードを並列表示して比較しながら統合)、Multiedit(複数レコードの同一項目を一括して同一値に更新)、Batch uploader(レコードの一括更新・追加)等がある。
- この他、様々な検索オプションやレコード抽出オプションがある。

3. Inspire による著者情報の管理

3.1. HEP

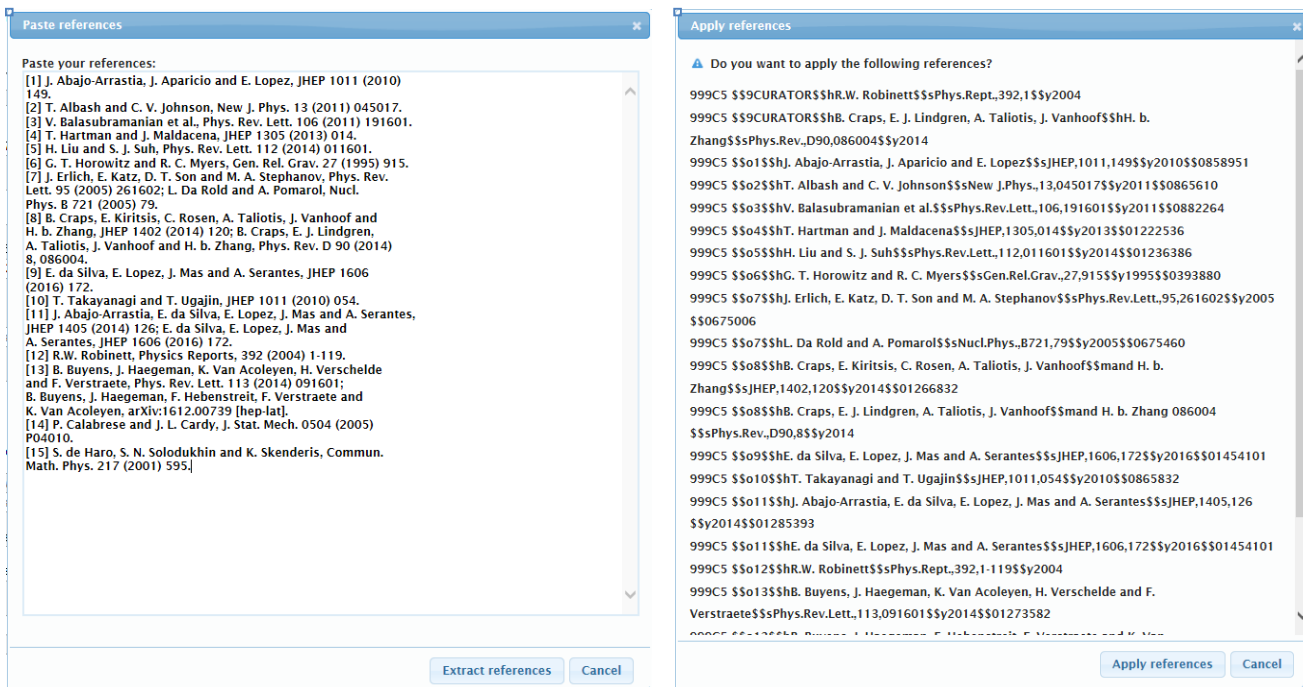
- 著者名は、サブフィールド 100__a(責任著者)、700__a(共著者)に記述されている。所属機関は同様に 100__u、700__u に記述されている。
- Refextract 機能等により、本文 PDF ファイル等から引用文献情報を抽出し、MARC 形式のメタデータへ変換することができる。

(例) HEP (Bibedit による編集画面)

record revision: 2016.06.16 17:52:37 florian

001		378081
005		20160616175237.0
0247	2	DOI
	a	10.1143/PTP.93.67
035	9	arXiv
	a	oai:arXiv.org:gr-qc/9410007
	z	oai:arXiv.org:gr-qc/9410007
035	9	DESY
	z	D94-19367
035	9	DESY
	z	D95-17072
035	9	SPIRETeX
	a	Tanaka:1994yh
035	9	KEKSCAN
	a	94-10-246
035	9	ADS
	a	1995PThPh..93..67T
037	9	arXiv
	a	gr-qc/9410007
	c	gr-qc
037	a	KUNS-1295
041	a	language (if not English)
084	2	PACS
	a	pacs
100	a	Tanaka, Takahiro
	j	ORCID:
	m	email
	u	Kyoto U.
210	a	title_variation - extra words/phrasings for title
242	a	Translated title (English)
245	a	Decoherence on quantum tunneling in the multidimensional wave function approach

(例) 本文 PDF ファイル等から得た引用文献情報(テキスト)を MARC 形式のメタデータへ変換



3.2. HepNames

- 氏名、氏名の別標記、所属機関(履歴)、学位取得に関する情報、研究分野、各種研究者識別 ID (arXiv、ORCID 等) が記述されている。

(例) HepName(ウェブサイト表示)

Takahiro Tanaka (田中, 貴浩) (Kyoto U., Yukawa Inst., Kyoto)
[\[Author Profile\]](#) [\[Google\]](#) [\[Students\]](#) [\[arXiv\]](#) [\[ADS\]](#)

PhD Institution: [Kyoto U.](#)
Undergrad: [Kyoto U.](#)
Email: tanaka@yukawa.kyoto-u.ac.jp
URL: <https://kyouindb.iimc.kyoto-u.ac.jp/e/pE7dU>
Field: HEP-TH, GR-QC
Author Profile: [Takahiro.Tanaka.1](#)
Inspire ID: INSPIRE-00162441

Institutional History:

Institution	Rank	Start Date	End Date
Kyoto U., Yukawa Inst., Kyoto	SENIOR	2008	
Kyoto U.	JUNIOR	2003	2008
Kyoto U., Yukawa Inst., Kyoto	JUNIOR	2000	2003
Osaka U.	PD	1995	2000
JSPS, Tokyo	PD	1993	1995
Kyoto U.	PHD		
Kyoto U.	UG		

(例) HepName (MARC 形式)

```

035_ $$0$hepire$$hepire-00162441
035_ $$0BAI$$aTakahiro.Tanaka.1
100_ $$aTanaka, Takahiro$$bVOLATILE:exnumeration (e.g. III or Jr.)$$cVOLATILE:title (e.g. Sir)$$dVOLATILE:dates$$gVOLATILE:active, retired. $$qTakahiro
Tanaka
371_ $$aKyoto U., Yukawa Inst., Kyoto$$mtanaka@yukawa.kyoto-u.ac.jp$$rSENIOR$$s2008$$tVOLATILE:end year$$zCurrent
371_ $$aKyoto U.$mVOLATILE:public email$$rJUNIOR$$s2003$$t2008$$zVOLATILE:current
371_ $$aKyoto U., Yukawa Inst., Kyoto$$mVOLATILE:public email$$rJUNIOR$$s2006$$t2005$$zVOLATILE:current
371_ $$aOsaka U.$mVOLATILE:public email$$rPD$$s1995$$t2000$$zVOLATILE:current
371_ $$aJSPS, Tokyo$$mVOLATILE:public email$$rPD$$s1993$$t1993$$zVOLATILE:current
371_ $$aKyoto U.$mVOLATILE:public email$$rPHD$$sVOLATILE:start year$$tVOLATILE:end year$$zVOLATILE:current
371_ $$aKyoto U.$mVOLATILE:public email$$rUG$$sVOLATILE:start year$$tVOLATILE:end year$$zVOLATILE:current
400_ $$aVOLATILE:variant name
595_ $$aVOLATILE:hidden note$$mVOLATILE:contact email$$oVOLATILE:old email
65017 $$2hepire$$aHEP-TH
65017 $$2VOLATILE:hepire$$aGR-QC
667_ $$aVOLATILE:nonpublic note
670_ $$aVOLATILE:source$$2013-12-19
678_ $$aVOLATILE:keyword
680_ $$iVOLATILE:public note
693_ $$dVOLATILE:end date$$sVOLATILE:experiment$$sVOLATILE:start date$$zVOLATILE:current
701_ $$aVOLATILE:Name of advisor$$gVOLATILE:degree (PhD, master, diploma...)$$iVOLATILE:hepire id of advisor
711_ $$aVOLATILE:conf attended (cnum)
8564_ $$uhttps://kyouin-db.imac.kyoto-u.ac.jp/wp/E7dU
880_ $$a 田中, 孝浩
909CD $$oai:hepirehep.net:986620$$p$hepire:HepNames
961_ $$x2000-11-15
961_ $$x2009-10-23
970_ $$aHEPNAMES-5244
    
```

(例) HepName (Bibedit 編集画面)

record revision: 2016.04.13 18:17:54 aholikamp

001 986629

005 20160413181754.0

035_ 9 INSPIRE
a INSPIRE-00162441

035_ 9 BAI
a Takahiro.Tanaka.1

100_ a Tanaka, Takahiro
b numeration (e.g. III or Jr.)
c title (e.g. Sir)
d dates
g active, retired...
q Takahiro Tanaka

371_ a Kyoto U., Yukawa Inst., Kyoto
m tanaka@yukawa.kyoto-u.ac.jp
r SENIOR
s 2008
t end year
z Current

371_ a Kyoto U.
m public email
r JUNIOR
s 2003
t 2008
z current

371_ a Kyoto U., Yukawa Inst., Kyoto
m public email
r JUNIOR
s 2000
t 2003
z current

371_ a Osaka U.
m public email
r PD
s 1995
t 2000
z current

371_ a JSPS, Tokyo
m public email
r PD
s 1993

Display options: References, Authors, Others, Curator view

3.3. Author Profiles

- 論文レコード HEP の著者、著者所属機関、共著者、分野、関係する実験等のメタデータを自動分析することにより、著者の表記の揺れを吸収して、著者ごとの執筆論文プロファイル **Author Profile** が生成される。
- **Author Profile** には、HepName の他、論文から分析した著者所属機関、共著者、分野、コラボレーション（関係する実験グループ）、キーワード、出版物の種別、分野、引用文献情報等が表示されている。
- **Author Profile** は HepName とリンクされていなければならない。

(例) Author Profile (ウェブサイト表示) <https://inspirehep.net/author/profile/Takahiro.Tanaka.1>

(例) Author Profile の識別子 BAI が HepName (035 サブフィールド) に格納され、両者がリンクされている。

4. 日本人著者情報の充実

4.1. 実施前の問題点

- 著者ごとの執筆論文プロフィール Author Profile と著者レコード HepName との間のリンク関係が形成されていない。
- HepName が未作成である、または情報が不十分である。
- Author Profile に別人の論文が混入したり、含まれるべき論文が漏れたりしている。

4.2. 作業手順

- 1) 著者レコード HepName を正確かつ可能な限り最新の情報に更新する(必要に応じて新規レコードの登録または既存レコードの更新)。
 - 所属機関履歴、学位取得情報等の更新
 - アルファベット表記が同姓同名の著者の識別を容易にするために、漢字表記を追加(880_a)
- 2) HepName と著者ごとの執筆論文プロフィール Author Profile とのリンク形成
- 3) Author Profile のクリーニング
 - 別人の論文を切り離し、当該著者の論文を追加
 - 同一著者に対して複数の Author Profile が存在している場合は統合
 - 当該著者の Author Profile が生成されていない場合は新規作成

(例) 当該著者の論文を追加、別人の論文を切り離し

Tanaka, Takahiro

Profile Name Search

View Profile Manage Profile Manage Publications Help Open Tickets

Papers (278) Papers removed from this profile (88) Tickets (0) Data

Select All | Select None | Invert Selection | Hide successful claims

Yes, those papers are by this person. No, those papers are not by this person Assign to another person

Paper Short Info	Author Name	Affiliation	Date	Experiment	Actions
<input type="checkbox"/> 1. Fermi large area telescope observations of Markarian 421: The missing piece of its spectral energy distribution LAT and MAGIC Collaborations (A.A. Abdo et al.).	Tanaka, T.	Not assigned	2011-08-02	MAGIC	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 2. Chronology protection and quantized field in generalized Misner space T. Tanaka, W.A. Hiscock (Montana State U.).	Tanaka, T.	Montana State U.	1994-07-01	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 3. Chronology protection and quantized fields: Nonconformal and massive scalar fields in Misner space T. Tanaka, W.A. Hiscock (Montana State U.).	Tanaka, T.	Montana State U.	1994-12-16	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 4. Variational perturbation approach to resonance state wave functions T. Tanaka (Tsukuba U.).	Tanaka, T.	Tsukuba U.	1997-06-25	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 5. Progress in 1-KeV - 2-KeV spectroscopy using a YB(66) monochromator J. Wong (LLNL, Livermore), M. Froba (LLNL, Livermore & Hamburg U.), E. Tamura (LLNL, Livermore), M. Rowen, Z. Rek (SLAC, SSRL), T. Tanaka (Tsukuba U.).	Tanaka, T.	Tsukuba U.	1997-01-01	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 6. Optimized perturbation theory for wave functions of quantum systems T. Hatsuda (Tsukuba U.), T. Kunihiro (Ryukoku U.), T. Tanaka (Tsukuba U.).	Tanaka, T.	Tsukuba U.	1996-12-11	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 7. Guiding optics system for LEBRA FEL user facility T. Tanaka, K. Hayakawa, Y. Hayakawa, I. Sato (Nihon U., Narashino).	Tanaka, T.	Nihon U., Narashino	2004-08-01	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 8. Performance of the FEL Linac at Nihon University T. Tanaka, K. Hayakawa, K. Yokoyama, Y. Matsubara, K. Sato, I. Kawakami (Nihon U., Tokyo), S. Fukuda, S. Ohsawa, S. Anami (KEK, Tsukuba).	Tanaka, T.	Nihon U., Tokyo	1998-03-01	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 9. Geometrical effect of target crystal on PXR generation as a coherent X-ray source Y. Hayakawa, K. Hayakawa, M. Inagaki, T. Kuwada, K. Nakao, K. Nogami, T. Sakai, I. Sato, Y. Takahashi, T. Tanaka. 2010	Tanaka, T.	Not assigned	2010-11-12	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 10. A Novel Approach in Constraining Electron Spectra in Blazar Jets: The Case of Markarian 421 M. Ushio, L. Stawarz, T. Takahashi, D. Paneque, G. Madejski, M. Hayashida, J. Kataoka, Y.T. Tanaka, T. Tanaka, M. Ostrowski.	Tanaka, T.	Not assigned	2010-09-14	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 11. Study of the Spectral and Temporal Characteristics of X-Ray Emission of the Gamma-Ray Binary LS 5039 with Suzaku T. Takahashi, T. Kishishita, Y. Uchiyama, T. Tanaka, K. Yamaoka, D. Khangulyan, F.A. Aharonian, V. Bosch-Ramon, J.A. Hinton.	Tanaka, T.	Not assigned	2008-12-17	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 12. Suzaku Observations of Tycho's Supernova Remnant T. Tamagawa, A. Hayato, S. Nakamura, Y. Terada, A. Bamba, J.S. Hiraga, J.P. Hughes, U. Hwang, J. Kataoka, K. Kinugasa et al.	Tanaka, T.	Not assigned	2008-05-22	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person
<input type="checkbox"/> 13. Development of double-sided silicon strip detectors (DSSD) for a Compton telescope S. Takeda, S. Watanabe, T. Tanaka, K. Nakazawa, T. Takahashi, Y. Fukazawa, H. Yasuda, H. Tajima, Y. Kuroda, M. Onishi et al.	Tanaka, T.	Not assigned	2007-09-01	N.A.	<input checked="" type="checkbox"/> Yes, this paper is by this person. <input checked="" type="checkbox"/> No, this paper is not by this person <input type="checkbox"/> Assign to another person

(例) 同一著者に対して複数の Author Profile が存在している場合は統合

➤ 情報源として利用したデータベース等

- KAKEN 科学研究費助成事業データベース <https://kaken.nii.ac.jp/index/>
- 研究者リゾルバー <http://rns.nii.ac.jp/>
- Researchmap <http://researchmap.jp/search/>
- 学位論文データベース Cinii Dissertations <http://ci.nii.ac.jp/d/> 他
- JAIRO <http://ju.nii.ac.jp/> 他、各大学の機関リポジトリ
- 各大学の研究者業績データベース等
- 研究者の個人ウェブサイト等

4.3. Inspire-KAKEN データベース間の照合

- Inspire と科学研究費助成事業データベース(KAKEN)との照合(著者レベルおよび論文レベル)を行い、KAKEN ID を HepName、HEP レコードに追加した。
- KAKEN ID を利用することにより、著者情報充実のためのレコード更新作業が迅速化した。

1) NII より KAKEN データを CERN へ提供

1-1) KAKEN データベースから抽出した著者クラスター(サイテーション付)

(KAKEN レコードの構造)

- Project の researchers に氏名と ID を含む。
- Project の reports に publications を含む。
- Project は research fields を持つ。
- サイテーションを含む著者情報を抽出可能
 - 著者 216,533 件(全研究分野)
 - サイテーション 延べ 3,330,472 件(異なり数 2,538,913 件)
 - ほとんどのレコードに DOI を含まない。

1-2) 高エネルギー物理分野のプロジェクト

高エネルギー物理分野のプロジェクトは、以下の科研費研究分野に分類されている。

- "素核宇宙融合による計算科学に基づいた重層的物質構造の解明",
 - "原子が切り拓く極限量子の世界—素粒子的宇宙像の確立を目指して—",
 - "多彩なフレーバーで探る新しいハドロン存在形態の包括的研究",
 - "先端加速器 LHC が切り拓くテラスケールの素粒子物理学~真空と時空への新たな挑戦",
 - "先端加速器 LHC が切り拓くテラスケールの素粒子物理学~真空と時空への新たな挑戦",
 - "物理",
 - "物理学",
 - "素粒子・核・宇宙線",
 - "素粒子・原子核・宇宙線・宇宙物理",
 - "核・宇宙線・素粒子",
 - "素粒子・原子核・宇宙線・宇宙物理(理論)",
 - "素粒子・原子核・宇宙線・宇宙物理(理論)",
 - "素粒子・原子核・宇宙線・宇宙物理(実験)",
 - "素粒子・原子核・宇宙線・宇宙物理(実験)".
- 全研究分野 1,003 件のうち、関連するのは 14 研究分野
 - Particle/Nuclear/Cosmic ray/Astro physics

1-3) 氏名のマッチング

サイテーションに研究者の (1) 姓 (フルスペル) および (2) 名のフルスペルまたはイニシャル+ “.” が含まれている必要がある。

1-4) 高エネルギー物理分野の著者クラスターの状況

- 著者 2,012 件、共著者 3,624 件 (異なり数)
- サイテーション 89,245 件 (延べ数。異なり数は 74,552 件)
 - うち DOI を含むもの 16,017 件 (延べ数。異なり数は 7,186 件)

2) CERN 側でアルファベット姓名によるデータ一括照合 (7 月実施)

(結果および手作業の内容)

・照合不完全 (7-8 月完了)

- フルスペルで複数件一致 27 件 `more_than_one.txt`
 - 同一人物を確定のうえ (必要に応じて `HepName` を新規作成)、情報追加
- イニシャルで 1 件一致 30 件 `initials.txt`
 - 同一人物を確定のうえ (必要に応じて `HepName` を新規作成)、情報追加
- 一致しないが `Author Profile` (論文レコード) が存在 396 件 `but_has.txt`
 - `HepName` を新規作成

・照合成功

- フルスペルで 1 件一致 1,306 件 (KEKAN ID を一括投入済)
- (内訳)

- Author Profile と未リンク 358 件(8 月完了)
- Author Profile とリンク済 948 件(12 月完了)
- ※HepName の確認・情報追加

一括照合では同姓同名別人の可能性を排除できないため、既存情報を元に同一人物であることを確認のうえ、漢字姓名、経歴、情報源等の追記および必要に応じて Author Profile とのリンク作成を手作業で行う。

3) CERN 側で、KAKEN データと Inspire データの論文 DOI を照合(9 月実施)

(結果および手作業の内容)

KAKEN データの論文	74,577 件
うち、DOI あり	7,186 件
うち、DOI 一致せず	2,574 件 ⁴
DOI 一致	4,612 件
うち、- 少なくとも 1 件以上の著者名が Hep レコードの著者名(100/700 サブフィールド)	
1 件に一致する論文	4,106 件
	→Inspire 論文情報に KAKEN ID を一括投入済
- 少なくとも 1 件以上の著者名が Hep レコードの著者名(100/700 サブフィールド)	
複数に一致する論文 ⁵	119 件(more.log)
	→手作業による確認、情報追加(1 月完了)

※ 今後、論文レベルで著者名のみでのマッチングをする際に、KAKEN ID を持たない著者について特に注意を要する(名寄せできていない可能性がある)。

⁴ DOI が一致しなかった 2,574 件のうち、Inspire 収録対象分野の論文は 122 件。Inspire に既存レコードが無いが、手作業で確認。

(結果)

KAKEN データに存在するが Inspire に存在しない DOI を持つ論文	49 件
DOI が無効かつタイトルが一般的過ぎてタイトルによる同定ができなかったもの	1 件
Inspire に既存レコードが存在しているが、DOI が入力されていなかった	
(または DOI に変更があり、変更前の DOI が入力されていた)もの(DOI 追加済)	72 件

⁵ Collaboration による論文等の場合、ファーストネームがイニシャルのみの著者名表記でその数も膨大であるため。

(例) KAKEN ID が投入された HepName (ウェブサイト表示および Bibedit 編集画面)

INSPIRE HEP

Welcome to **INSPIRE**, the High Energy Physics information system. Please direct questions, comments or concerns to feedback@inspirehep.net.

HEP :: **HEPNAMES** :: INSTITUTIONS :: CONFERENCES

Hikonojo Orihara (織原 彦之丞) (Tohoku Inst. Tech.)
[\[Author Profile\]](#) [\[Google\]](#) [\[Students\]](#) [\[arXiv\]](#) [\[ADS\]](#)

PhD Institution: [Tohoku U.](#)
 Undergrad: [Tohoku U.](#)
 Email: orihara@cyric.tohoku.ac.jp
 URL: <http://researchmap.jp/read0168566/?lang=english>
 Author Profile: [H.Orihara.1](#)

Institutional History:

Institution	Rank	Start Date	End Date
Tohoku Inst. Tech.	SENIOR	2003	
CYRIC, Tohoku U.	SENIOR	1978	2003
Tohoku U.	JUNIOR	1969	1978
Tohoku U.	PHD		1969
Tohoku U.	UG		1964

001		994792
005		20161013162447.0
035	<input type="checkbox"/> 2	script
	<input type="checkbox"/> 9	BAI
	<input type="checkbox"/> a	H.Orihara.1
035	<input type="checkbox"/> 2	kaken_import.py
	<input type="checkbox"/> 9	KAKEN
	<input type="checkbox"/> a	00004432
100	<input type="checkbox"/> a	Orihara, Hikonojo
	<input type="checkbox"/> b	numeration (e.g. III or Jr.)
	<input type="checkbox"/> c	title (e.g. Sir)
	<input type="checkbox"/> d	dates
	<input type="checkbox"/> g	active, retired...
	<input type="checkbox"/> q	Hikonojo Orihara
371	<input type="checkbox"/> a	Tohoku Inst. Tech.
	<input type="checkbox"/> m	public email
	<input type="checkbox"/> r	SENIOR
	<input type="checkbox"/> s	2003
	<input type="checkbox"/> t	end year
	<input type="checkbox"/> z	current
371	<input type="checkbox"/> a	CYRIC, Tohoku U.
	<input type="checkbox"/> m	public email
	<input type="checkbox"/> r	SENIOR
	<input type="checkbox"/> s	1978
	<input type="checkbox"/> t	2003
	<input type="checkbox"/> z	current
371	<input type="checkbox"/> a	Tohoku U.
	<input type="checkbox"/> m	public email

(例) KAKEN ID が投入された HepName が表示されている Author Profile

INSPIRE HEP

Welcome to **INSPIRE**, the High Energy Physics information system. Please direct questions, comments or concerns to feedback@inspirehep.net.

HEP :: **HEPNAMES** :: INSTITUTIONS :: CONFERENCES :: JOBS :: EXPERIMENTS :: JOURNALS :: HELP

Orihara, H.

Profile Name Search

View Profile Manage Profile Manage Publications Help Open Tickets

PERSONAL INFORMATION PUBLICATIONS AND OUTPUT STATS

Personal Details (HepNames)

Name Hikonojo Orihara (織原 彦之丞)
 Current Institution Tohoku Inst. Tech.
 E-mail orihara@cyric.tohoku.ac.jp
 Links <http://researchmap.jp/read0168...>
 Identifiers BAI: H.Orihara.1
 KAKEN: 00004432

Period Rank Institution
 2003 SENIOR Tohoku Inst. Tech.
 1978 – 2003 SENIOR CYRIC, Tohoku U.
 1969 – 1978 JUNIOR Tohoku U.
 1969 PHD Tohoku U.
 1964 UG Tohoku U.

HepNames Record Update Details

Name Variants
 Orihara, H. (64)
 Orihara, H (1)

Publications Datasets External

- Structure of the unbound N-11 nucleus by the (He-3, He-6) reaction
- Proton single particle strength in F-19 measured via the O-18(d, n) reaction
- Gamow-Teller matrix elements in light nuclei and (p, n) cross sections at E p = 35 MeV
- Isovector part of optical potentials studied through analog transitions in the (p, n) reaction at 35-MeV
- 0+ to 0- transition in the O-16 (p, n) F-16 reaction at 79 MeV
- Analog transitions in sd- and f-shell nuclei and the isovector part of optical potentials studied by the (p, n) reaction at 35 MeV
- High energy gamma-ray production from Be, C, and Al targets with 65 MeV He-3 bombardment
- Study of the early stage of the rapid proton process
- The 17, 18 O(p, n) 17, 18 F reaction at E p = 35 MeV
- Structure of light proton-rich nuclei on the drip-line

Click here to see all

Co-Authors

Hajime OHNUMA.1 (39)
 Takao Nakagawa.1 (28)
 K.Miura.2 (25)
 T.Nitzeki.1 (22)
 S.Kubono.1 (21)
 H.Kawashima.1 (16)
 K.I.Maeda.1 (16)
 K.Ishii.1 (16)
 Y.Fuchi.1 (16)
 T.Nomura.2 (14)
 more

Papers

	All papers	Single authored
All papers	65	0
Book	0	0
ConferencePaper	11	0
Introductory	0	0
Lectures	0	0
Published	49	0
Review	0	0
Thesis	0	0
Proceedings	0	0

Citations Summary

65 papers found, 52 of them citeable (published or arXiv)

	Citeable papers	Published only
Number of papers analyzed:	52	49
Number of citations:	426	421
Citations per paper (average):	8.2	8.6
h _{np} index [?]	12	12

Breakdown of papers by citations:

	Citeable papers	Published only
Renowned papers (500+)	0	0
Famous papers (250-499)	0	0
Very well-known papers (100-249)	0	0
Well-known papers (50-99)	0	0
Known papers (10-49)	16	16
Less known papers (1-9)	34	31
Unknown papers (0)	2	2

Click here to view statistics without self-citations or RPP

Warning: The citations count should be interpreted with great

(例) KAKEN ID が投入された HEP (ウェブサイト表示および Bibedit 編集画面)



Welcome to **INSPIRE**, the High Energy Physics information system. Please direct questions, comments or concerns to feedback@inspirehep.net.

HEP :: HEPNAMES :: INSTITUTIONS :: CONFERENCES :: JOBS :: EXPERIMENTS :: JOURNALS :: HELP

Information | References (24) | Citations (1) | Files | Plots

Gamow–Teller matrix elements in light nuclei and (p, n) cross sections at E p = 35 MeV

H Orihara, A Terakawa, K Itoh, H Suzuki, K Kumagai, Y Kikuchi (CYRIC, Tohoku U.), G.C Jon (Taiwan, Inst. Phys.), K Ishii (Tohoku U. (main)), T Niizeki (Tokyo Kasei U.), H Sagawa (Aizu U.), H Ohnuma (Chiba Inst. Tech.) [Hide](#)

2002 - 6 pages


Phys.Lett. B539 (2002) 40-45
(2002)

DOI: [10.1016/S0370-2693\(02\)02036-1](https://doi.org/10.1016/S0370-2693(02)02036-1)

Abstract (Elsevier)
The (p, n) cross sections for the 0 + → 1 + excitation are studied experimentally at E p = 35 MeV in eight nuclei with T z = (N - Z)/2 = 1, 14 C, 18 O, 22 Ne, 26 Mg, 30 Si, 34 S, 38 Ar and 42 Ca. It is found that the peak (p, n) cross sections around θ ≈ 30° are proportional to the corresponding β + -decay rates of the T z = -1 partners of the T = 1 multiplets. We propose a prescription to extract GT matrix elements in light nuclei from (p, n) cross sections using such proportionality. The method is applied to determine an unknown B (GT) matrix element in 26 Al, which is energetically inaccessible by β + -decay experiments.

Keyword(s):
Author supplied: [Gamow–Teller matrix elements](#) | [\(p, n\) reaction](#) | [Spin-isospin excitation](#) | [TOF measurement](#) | [\(p, n\) reaction](#)

Record added 2016-05-21, last modified 2016-10-13



Welcome to **INSPIRE**, the High Energy Physics information system. Please direct questions, comments or concerns to feedback@inspirehep.net.

HEP :: HEPNAMES :: INSTITUTIONS :: CONFERENCES :: JOBS :: EXPERIMENTS :: JOURNALS :: HELP

record revision: 2016.10.13 16:25:57 omura

Record

1463160 x Rec

Search

Submit Cancel

Switch to: Read-only

Tickets

[New ticket!](#)

Fields

Add

Delete selected

View

Holding Pen

Undo/Redo

History

Ready

Special symbols Help

Display

References

Authors

Others

Curator view

001	1463160	
005	20161013162557.0	
<input type="checkbox"/> 0247	<input type="checkbox"/> 2	DOI
	<input type="checkbox"/> 9	Elsevier
	<input type="checkbox"/> a	10.1016/S0370-2693(02)02036-1
<input type="checkbox"/> 035	<input type="checkbox"/> 9	INSPIRE TeX
	<input type="checkbox"/> a	Orihara:2002dry
<input type="checkbox"/> 041	<input type="checkbox"/> a	language (if not English)
<input type="checkbox"/> 084	<input type="checkbox"/> 2	PACS
	<input type="checkbox"/> a	pacs
<input type="checkbox"/> 100	<input type="checkbox"/> a	Orihara, H
	<input type="checkbox"/> i	KAKEN-00004432
	<input type="checkbox"/> m	orihara@cync.tohoku.ac.jp
	<input type="checkbox"/> u	CYRIC, Tohoku U.
	<input type="checkbox"/> v	Cyclotron and Radioisotope Center, Tohoku University, Sendai 980-8578, Japan
<input type="checkbox"/> 210	<input type="checkbox"/> a	title_variation - extra words/phrasings for title
<input type="checkbox"/> 242	<input type="checkbox"/> a	Translated title (English)
<input type="checkbox"/> 245	<input type="checkbox"/> a	Gamow–Teller matrix elements in light nuclei and (p, n) cross sections at E p = 35 MeV
<input type="checkbox"/> 260	<input type="checkbox"/> c	2002
	<input type="checkbox"/> t	published
<input type="checkbox"/> 300	<input type="checkbox"/> a	6
<input type="checkbox"/> 520	<input type="checkbox"/> 9	Elsevier
	<input type="checkbox"/> a	The (p, n) cross sections for the 0 + → 1 + excitation are studied experimentally at E p = 35 MeV in eight nuclei with T z = (N - Z)/2 = 1, 14 C, 18 O, 22 Ne, 26 Mg, 30 Si, 34 S, 38 Ar and 42 Ca. It is found that the peak (p, n) cross sections around θ ≈ 30° are proportional to the corresponding β + -decay rates of the T z = -1 partners of the T = 1 multiplets. We propose a prescription to extract GT matrix elements in light nuclei from (p, n) cross sections using such proportionality. The method is applied to determine an unknown B (GT) matrix element in 26 Al, which is energetically inaccessible by β + -decay experiments.
<input type="checkbox"/> 540	<input type="checkbox"/> a	terms of use, e.g CC License
	<input type="checkbox"/> u	url of file specifying access rights
<input type="checkbox"/> 541	<input type="checkbox"/> a	Elsevier
	<input type="checkbox"/> c	batchupload
<input type="checkbox"/> 542	<input type="checkbox"/> f	Elsevier Science B.V.
	<input type="checkbox"/> u	url of copyright statement
<input type="checkbox"/> 595	<input type="checkbox"/> a	hidden note
<input type="checkbox"/> 65017	<input type="checkbox"/> 2	INSPIRE

5. 今後の課題

5.1. 機関情報の整合性維持

- 現時点では、日本に所在する機関の **Institutions** レコードは 1,000 件余り存在し、他国と比較しても件数が多い。これは、レコード作成レベル(機関全体か部局単位か、あるいはキャンパス単位か)や改組・名称変遷等の要因によるものと考えられる。しかし、レコード作成・更新時期により判断に揺れがあったり、レコードの変遷・親子関係が複雑になったりしているため、整合性維持が困難になっており、これに対応することが課題となっている。

5.2. 研究者とのコミュニケーション

- 著者自身あるいは関係する研究者との協力が今後重要になると考えられる。現在、CERN と高エネルギー加速器研究機構や日本高エネルギー物理学研究者会議との協力推進の動きが加速しており、今後研究者から直接情報を得ることにより、さらなるデータベースの質向上が期待される。
- Inspire は、現行システムでも利用者とのコミュニケーションツールを持つが、新システムではさらに充実する予定である。

(例) Inspire が持つ利用者とのコミュニケーションツール

The screenshot displays the Inspire user profile for Yoshinobu Takaiwa. The page includes navigation tabs for 'View Profile', 'Manage Profile', 'Manage Publications', and 'Help'. Two red boxes highlight the 'arXiv' and 'ORCID' sections. The 'arXiv' section prompts the user to log in to manage their profile. The 'ORCID' section indicates that the profile is not connected to an ORCID account and offers a suggestion feature. The 'HepNames' section provides a detailed overview of the user's career, including their name (Yoshinobu Takaiwa), current institution (Tsukuba Univ. Tech.), email, and a table of past positions.

Period	Rank	Institution
2006	STAFF	Tsukuba Univ. Tech.
1986 – 2006	STAFF	KEK, Tsukuba
1981 – 1986	PD	Fermilab
1977 – 1981	PD	Tsukuba U.
1972 – 1977	PHD	Tokyo U. of Education
1968 – 1972	UG	Tokyo U. of Education

(例)利用者(論文著者)が Inspire の Author Profile と arXiv のアカウントをリンクすることができる

arXiv.org

Log in to **INSPIRE** via arXiv.org

Use your arXiv account to log in to INSPIRE. A successful login will take you directly to INSPIRE and will allow you to explore and use all the services on INSPIRE.

arXiv and INSPIRE have collaborated to implement this login mechanism in order to provide better integrated services. Information about your user account and papers will be shared between arXiv and INSPIRE. In particular, information about which papers you have authored will be exchanged so that your publication record will be complete and accurate on both systems.

If you've never logged onto arXiv.org:

Registration is required to submit or update papers.

If you're already registered:

Enter your **username or e-mail address**,

and **password**.
 (Forgot your password? [Try this.](#))

Safeguarding your arXiv account: This login page is hosted by arXiv.org. You can confirm this by looking at the web location shown by your browser. You should never enter your arXiv username and password on any other site. We provide this login service for our trusted partner INSPIRE in order to verify your identity. INSPIRE will not know your arXiv password.

[Contact](#)

(例)利用者(論文著者)が Inspire の Author Profile と ORCID のアカウントをリンクすることができる

ORCID

INSPIRE-HEP

has asked for the following access to your ORCID Record



Read your ORCID record

- Allow this permission until I revoke it.
*You may revoke permissions on your account settings page.
Unchecking this box will grant permission this time only.*

This application will not be able to see your ORCID password, or other private info in your ORCID Record. [Privacy Policy](#).

Already have an ORCID id? [Sign In](#)

As per ORCID's [terms and conditions](#), you may only register for an ORCID id for yourself.

First name *

Last name

以上

平成 28 年度 国立情報学研究所実務研修 報告書
iPRES 2016 参加報告

提出日：平成 28 年 10 月 16 日

所属：京都大学附属図書館

氏名：大村 明美

1. 開催概要

iPRES 2016: how to preserve digital information over the long term

13th International Conference on Digital Preservation

開催日・開催地: 2016 年 10 月 3 日-6 日; ベルン (スイス)

参加者: 339 名 (30 カ国; 参加人数上位からスイス、米国、英国、ドイツ、オランダ、スウェーデン、デンマーク、カナダ、日本[以下省略])

ウェブサイト: <http://www.ipres2016.ch>

プログラムおよびプロシーディング:

http://www.ipres2016.ch/frontend/index.php?folder_id=353

写真: iPRES 2016 Photo Gallery

<https://www.flickr.com/photos/ipres2016/albums/with/72157672176057472>

2. 開催趣旨

iPRES は、第 1 回北京 (2004 年) 以来毎年開催されている、デジタル情報の長期保存に関する国際会議である。ベルンで開催された第 13 回 iPRES 2016 では、30 カ国から 300 人以上が参加し、国 (国立図書館) レベル、機関 (大学図書館、コンソーシアム等) レベル、研究者レベル等さまざまな立場から、活発な議論が交わされた。

3. 基調講演概要

3.1. Challenges and Opportunities for Digital Preservation

Dr. Robert E. Kahn (CEO, Corporation for National Research Initiatives, Reston (USA))

デジタルデータの長期保存のためには、技術変化に対応可能な、特定の技術に依存しない設計思想が不可欠である。その要素として、デジタル情報の構造とその識別手段、構造要素の解釈の一貫性確保、セキュリティが確保された相互接続性、指定したレベルのメタデータ抽出、過去の蓄積データへの簡素なアクセス、実装システムのパフォーマンスの評価基準の決定、実装システムのパフォーマンスの評価、さらに設計思想の管理・向上に必要な社会的な仕組みの構築について述べられた。

また、こうしたデータ構造、識別子の管理運営が、従来の米国・政府関係機関主体から、国際的なコミュニティへ移行していく方向性についても言及された。

Ref.) 16 November 2005, the United Nations-sponsored World Summit on the Information Society (WSIS), held in Tunis <http://www.itu.int/net/wsis/basic/about.html>

DONA Foundation (2014-) <https://dona.net> (Handle システム、また将来は関連するデジタルアーキテクチャコンポーネントを管理する国際組織。2014年1月20日にスイス、ジュネーブの非営利組織として設立)

3.2. Net-based and Networked – Challenges for the Conservation of Digital Art

Sabine Himmelsbach (Director, Haus der elektronischen Künste Basel, Basel (CH))

静止状態で保存可能な従来の芸術作品と異なり、リアルタイムに動作する情報源がインターネット上で複雑にからみあうインターネットアートは、技術の変化により再生が不可能になる危険性があり、新しい保存手法の開発が喫緊の課題となっている。本講演では、House of Electronic Arts Basel における実例が紹介された。

3.3. Digital memories, digital futures – the interplay between humans and technology

Dr. David Bosshart (CEO, Gottlieb Duttweiler Institut für Wirtschaft und Gesellschaft, Rüschlikon/Zürich (CH))

現代社会においては、リアルタイムの情報検索やソーシャルメディアによる情報拡散にみられるように、公と私の境界が失われつつあり、新しいメディアによって新しいコミュニケーションスタイルが生まれ、情報の形自体も変化している。こうした様々な事象と、それから生じる未来に対する疑問が指摘された。

基調講演のほか、論文発表 35 件、ポスター発表 18 件、パネルディスカッション 3 件、ワークショップおよびチュートリアル 18 件が各日並行して行われた。

以下、特に研究データ管理、長期保存と識別子の観点から注目した論文発表、ポスター発表、ワークショップ、チュートリアルの概要を報告する。

4. 論文発表

4.1. 責任あるデータ管理と持続可能性

Will Today's Data Be Here Tomorrow? Measuring The Stewardship Gap

研究助成を受けたプロジェクトで生成されているデータ量と、実際に再利用可能な状態で適切に管理されているデータ量の乖離は認識されているが、あるべき管理状態の評価基準と、現状の把握を目的とした調査の結果報告がなされた。

What makes a Digital Steward: A Competency Profile Based On The National Digital Stewardship Residencies

データ管理者に必要な技能・責任・知識、その訓練プログラムについて論じられた。

Ref.) What makes a digital steward: A competency profile based on the National Digital Stewardship Residencies <https://osf.io/zndwq/>

Digital Preservation through Digital Sustainability

持続可能なデータ保存に不可欠な条件が提示された。すなわち、データの綿密性が確保されていること、構造が明確であること、十分なメタデータが付与されていること、

保存場所が明示されていること、オープンライセンス表示の仕組みがあること、暗黙知が共有されていること、関係者の参加が十分に果たされていること、適切な管理体制があること、これらを支える複数の資金があることである。

Building a Future for our Digital Memory: A National Approach to Digital Preservation in The Netherlands

デジタルデータの可視性向上、利用性向上、持続的長期保存実現のための全国的なインフラ整備を目的として、オランダで行われた全国的な共同プロジェクトの事例が紹介された。既存の仕組みを活かす手法をとった結果、既存のすべての部署が何らかの形でデジタル保存に関わることとなった。

4.2 品質とリスク評価

Identifying Barriers To File Rendering In Bit-level Preservation Repositories: A Preliminary Approach

保存時のファイル形式を決定する上で考慮すべき問題として、ビットレベルの保存について論じられた。

Practical Analysis of TIFF File Size Reductions Achievable Through Compression

TIFF データの最も効果的な圧縮手法を考察したプロジェクトが紹介された。

4.3. 研究データの保存

Precise Data Identification Services for Long Tail Research Data

データの再利用を可能とするためには、研究データからの分析結果の再現性や比較可能性を確保しなくてはならない。しかし、膨大な研究データは通常一部サブセットのみをダウンロード利用する一方で、絶え間ないデータ更新も行われる。これに対処する手段として、データのバージョン管理・タイムスタンプの付与、タイムスタンプを付与した抽出クエリの保存による、動的なデータ引用手法が提示された。

Ref.) Recommendations of the Working Group on Data Citation - Revision Oct. 20 2015

<https://www.rd-alliance.org/recommendations-working-group-data-citation-revision-oct-20-2015.html>

CERN Services for Long Term Data Preservation

高エネルギー物理学研究データの長期保存サービスについて、データそのもの、ソフトウェア、メタデータ、データの処理・再分析・再利用に必要な環境の保存の取り組みが紹介された。

Ellipse – Long-term and Permanent Protection and Accessibility of Geodata

ジオデータの長期保存事例が報告された。

4.4. 拡張性と相互接続性

Navigating through 200 Years of Historical Newspapers

200年に及ぶスイスの新聞アーカイブデータ閲覧用インターフェースの開発事例と、機能拡張を目指す今後の展望について紹介された。

4.5. フレームワークと管理の実例

Conceptualising Optimal Digital Preservation and Effort

ニュージーランド国立図書館によるデジタル保存のあり方の概念化およびその実現に向けて必要な取り組みについて紹介された。

Copyright in Cultural Heritage Institutions: Snapshot of the Situation in Switzerland from a librarian perspective

スイスにおける著作権をめぐる諸問題の現状が報告され、孤児作品や大量デジタル化に関する新しい法整備の動きについても言及された。

4.6. ウェブアーカイブ

Preserving Websites of Research & Development Projects

研究プロジェクトのウェブサイトには、実験に用いられたソフトウェアやテストデータセット、灰色文献、ニュース、配布資料等様々な貴重な情報が含まれているが、プロジェクト終了後適切に管理保存されている状況とは言い難い。EUの研究助成FP4からFP7までの助成を受けたプロジェクトについて行われた、ウェブサイトの所在を自動で探し出し保存する取り組みについて報告された。

5. ポスター発表

5.1. Assigning Rights Statements to Legacy Digital Collections

マイアミ大学図書館において、これまでに電子化された120コレクション5万2千点のデータについて、個々（コレクションレベルではなく、アイテムレベル）の権利状況を評価して権利記述を付したプロジェクトの事例が報告された。

権利状況判断の対象となる情報はメタデータとし（データ自体を確認することはしない）、判断基準は RightsStatements.org のパラメータを使用した。判断に際しては、コレクション入手時の状況（購入・寄贈）を調査することも役立った。判断基準の明確化・一貫性維持のため、マトリックスを作成した（プロジェクトの進行に従い、更新も行った）。権利記述には、メタデータ2項目を当てた（1つは自館の必要に応じて自由に記述できる項目、もう1つは RightsStatements.org に準拠した、標準化された機械可読表記）。

Ref.) RightsStatements.org. Europeana and Digital Public Library of America

<http://rightsstatements.org/page/1.0/?language=en>

情報があまりに少ないために権利状況を判断することができない場合（いわゆる孤児作品）については、“copyright undetermined”とした。将来的に追加情報が得られた場合は、その内容に応じて権利記述も更新する必要がある。

5.2. To Act or Not to Act – Handling File format Identification Issues in Practice

チューリッヒ工科大学では、データファイルのフォーマットを認識するツールをデータリポジトリへ組み込んでおり、そのワークフロー等が紹介された。

6. ワークショップ

6.1. Innovative Practices in Digital Preservation, Data Life Cycle Management (DLCM): A Swiss Initiative

スイスで行なわれている DLCM project の活動内容が紹介され、その成果に対するワークショップ参加者の議論・意見交換が行われた。

Ref.) DLCM <http://www.dlcm.ch>

6.1.1. プロジェクト概要

スイス国内の大学・研究機関が参加する、研究データのライフサイクル管理に関する3年間（2015年9月-）のプロジェクトで、5つのサブグループから構成される。

Track 1 DMP, Policies & Guidelines ポリシーとガイドライン

Track 2 Active Research Data Management アクティブな研究データ管理

Track 3 Publication and Long-term Preservation 出版と長期保存

Track 4 Training & Expertise 訓練と専門知識

Track 5 Awareness & Communication 意識向上とコミュニケーション

6.1.2. 成果の紹介と意見交換

（ポリシー）

ポリシーのテンプレートが紹介された（参考_DLCM_Policy_Outline.pdf）。

実際の機関内での承認手順（ジュネーブ大学の例）としては、今後1-2年程度以内の承認を目指し、まず関係部署間（図書館や研究支援関連部署等）で草案に対する議論を行い、その結果を学長へし、委員会等の審議にかけるとのことである。

（研究データの全ライフサイクルにおける管理）

研究者本人が研究データについて様々な観点から理解・判断・評価するための詳細なチェックリスト、その結果をもとに、研究者が研究データについて記述する様式のテンプレートが紹介された（参考_DLCM_Checklist.pdf, 参考_DLCM_DMP_Template.pdf）。チェックリストは計画書等を作成するための本人用控えであり、特に提出を求めている。研究の進行により変更も生じるため、定期的に更新する必要がある。

（サービス）

ビジネスモデル分析手法が紹介された。既存部署と連携して必要とされているサービスの分析・提供を行う。

（研究データの評価）

短期保存の場合は柔軟な運用が可能な識別子が、長期保存の場合は恒久的識別子が必要であることが指摘された。

7. チュートリアル

7.1. Persistent Identifiers for Digital Cultural Heritage

データ識別子の基本的要件について論じられ、現在運用されている識別子 (DOI, Handle, URN, URN:NBN, PURL, URI, ARK 等) の概要が説明された。

識別子は単純に技術の問題というよりも、その実施を支えるポリシーと体制の問題であり、社会的な基盤の重要性が指摘された。また、データ識別子の信頼性の条件が提示され、相互接続性の重要性についても言及された。

8. 所感

京都大学では、研究不正対策の一環として研究データ保存規則を定めたが、これは研究データのライフサイクル全体を通じた管理のうち、一部のみ（論文の根拠データの10年間保存）を規定したものである。今後これを研究データのライフサイクル全体の管理に対応できるものへ拡張していく必要がある。今回のカンファレンスでは研究データをめぐる諸問題全体を概観するとともに、大学における研究データ管理の実例に関する情報を収集することができ、有意義であった。

研究者だけでなく図書館員の声も多く聞くことができたが、ある図書館員の“**We guide; researchers decide**”という言葉が印象に残った。研究データ管理をめぐって図書館に期待・要求する声が高いことはプレッシャーになるかもしれないが、従来の図書館のイメージや役割を変えるチャンスでもあり、様々な事例を参考にしつつ、それぞれの文化に適した、最も肯定的に受け入れられるアプローチを探ることが重要だと感じた。

以上

平成 28 年度 国立情報学研究所実務研修 報告書
ETH Zürich における研究データ管理に関する調査

提出日：平成 29 年 2 月 10 日

所属：京都大学附属図書館

氏名：大村 明美

1. ETH Zürich (Swiss Federal Institute of Technology Zurich) 概要

- 1855 年設立、16 部局からなるスイス連邦工科大学
- 学生数約 26,500 名、教員数約 500 名、職員数約 8,000 名 (130 カ国)
- 図書館 ETH-Bibliothek (www.library.ethz.ch/en)
 - 1855 年設立、中央図書館と 5 つの関連図書館で構成 (他に独立した学部図書館が存在する)
 - アナログ資料 780 万点、デジタル資料 45 万 7000 点
 - 利用者数 30,490 名、貸出件数 301,000 件 (2013 年統計)
 - 資料購入予算 1,610 万 CHF (うち電子資料購入予算 640 万 CHF)
 - 職員数 275 名 (うちフルタイム 183 名)

2. ETH Zürich 研究データ管理担当者との懇談

日時: 2016 年 10 月 19 日 9:30-12:00

対応者:

Dr. Matthias Töwe (Team leader, digital curation) 研究データ管理サービス
 Dr. Franziska Geisser (Data archivist) ETH Data Archive、Research Collection
 Mr. Mohamed El-Saad (Data archivist) docuteam packer
 Dr. Ana Sesartic (Data management planning and teaching) 研究データ管理計画策定支援

2.1. 図書館が提供する主なサービス

- Knowledge Portal ディスカバリーシステム
- ETH E-Citations 研究者の業績情報データベース
- ETH E-Pics 画像管理システム
- ETH E-Collection 機関リポジトリ (論文フルテキスト)
- **ETH Data Archive** 機関リポジトリ (研究データ)
- Open Access OA 出版相談
- **Digital Curation** 研究データ管理

2.2. 現在進行中の主なプロジェクト

- **E-Publications** 現行の ETH E-Citations、ETH E-Collection、ETH Data Archive を統合
- E-Lending 学外利用者への電子書籍の貸出

- ETHorama 所蔵資料の地図による所在表示

2.3. 研究データ管理サービス

2.3.1. 組織体制

- ETH Zürich における研究データ管理サービスは、図書館の Customer Services 部門の Digital Curation チーム (5 名) により提供されている (フルタイム換算で図書館職員 3.9 名)。さらに、図書館 IT チーム職員 1.8 名 (フルタイム換算) も加わる。
- 数年間の準備期間を経て、2010-2011 年にプロジェクト開始、2014 年以降実質的な活動を開始した。
- 当初、所蔵資料のデジタルデータ保存対策が課題だったが、現在は研究データに主眼を置き、より包括的なデジタルデータ管理を行うことを目的としている。

2.3.2. サービス提供の方針

- 大学全体の IT 部門とも協力し、ソフトウェア関連案件は図書館が、ハードウェア関連案件は IT 部門が担当する。また、研究データ管理のうち、進行中の研究プロジェクトのデータ (active data) は IT 部門が、研究終了後の成果データは図書館が担当する。
- 資料電子化 (通常内部でスキャン。適切な機材がない場合外注) 等の業務は図書館の Archive チームが担当し、デジタルデータのマイグレーション等は Data Curation チームが行う。
- 現在、研究者の業績情報データベースとして ETH E-Citations を、論文フルテキスト用の機関リポジトリとして E-Collection を、研究データ用の機関リポジトリとして ETH Data Archive を運用している。これらの検索インターフェースとして Knowledge Portal を提供している。なお、論文のサプPLEMENTデータ等、サイズが比較的小さなデータは、論文フルテキストリポジトリ E-Collection に格納している。
- 当該研究分野で広く利用されている分野リポジトリが存在する場合は、機関リポジトリよりも分野リポジトリへの登録が優先される。大学図書館が提供しているサービスのみで研究者のニーズがすべて満たせるわけではないことを認識した上で、学内外のソリューションを広く俯瞰して、最適な選択をするようにしている。
- 図書館職員がすべての研究分野に精通しているわけではないことを認識したうえで、図書館員がデータ管理の観点からアドバイスを提供し、研究者が主体となって決定するスタンスを取っている。研究分野の専門知識を必要とするメタデータ等は、研究者が提供する必要がある。 (“Librarians guide; researchers decide.”)

2.4. 研究データリポジトリ ETH Data Archive および新統合プラットフォーム Research Collection

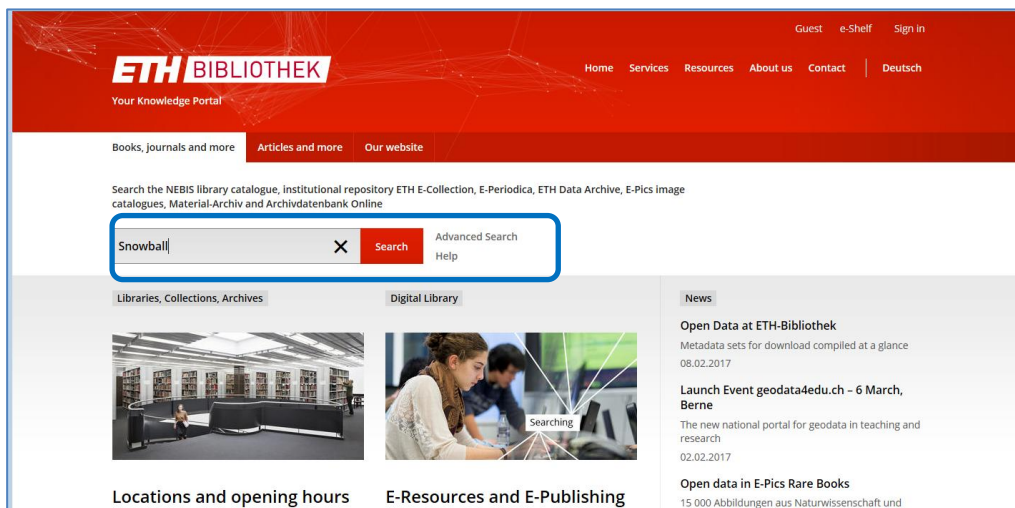
2.4.1. ETH Data Archive <http://www.library.ethz.ch/en>

- データアーカイブシステム Ex Libris Rosetta にデータを格納し、検索インターフェースとして Knowledge Portal (Primo) を提供している。
- 通常 Rosetta にアップロードするとエラーになるファイル種別のデータも、現時点の現実的な解決策として、当該ファイルを zip または tar に圧縮してパッケージとして保存

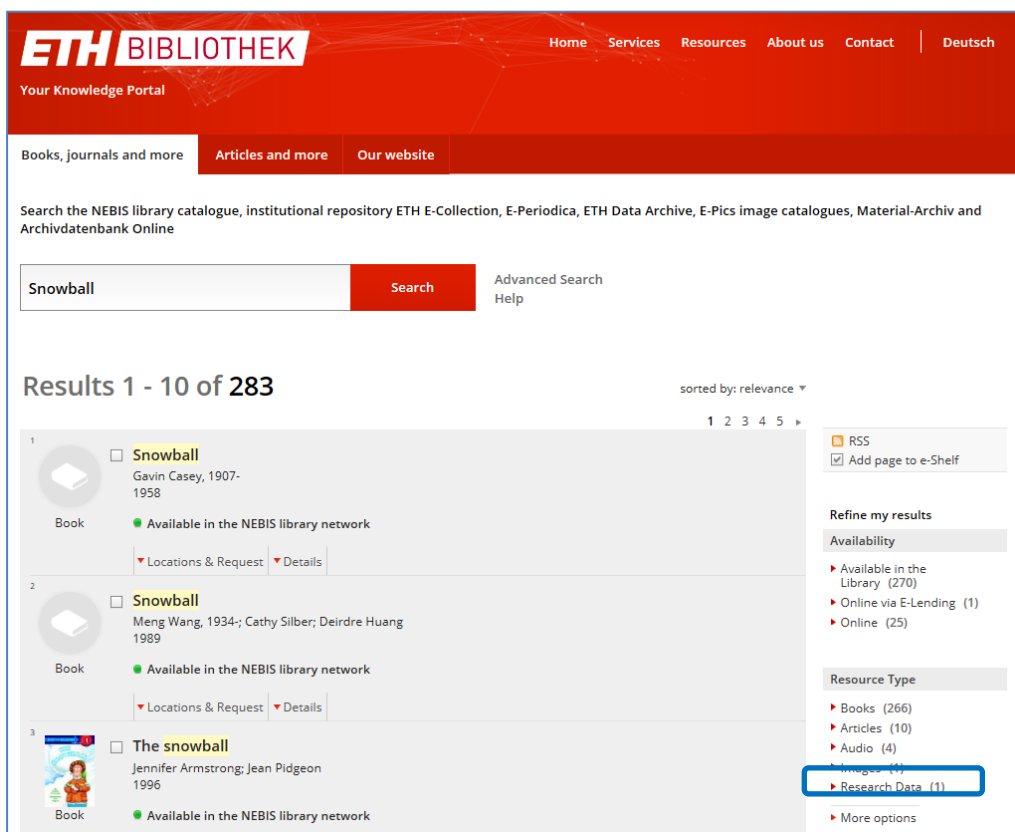
している。将来的に、デジタルデータの長期保存上の必要性に応じて、データ変換等の対応を考慮する。

(検索例)

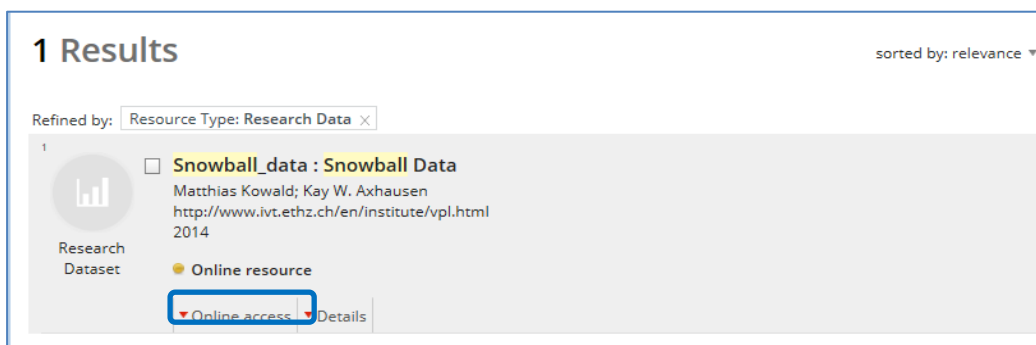
- 1) <http://www.library.ethz.ch/en> の検索窓から、「snowball」を入力。



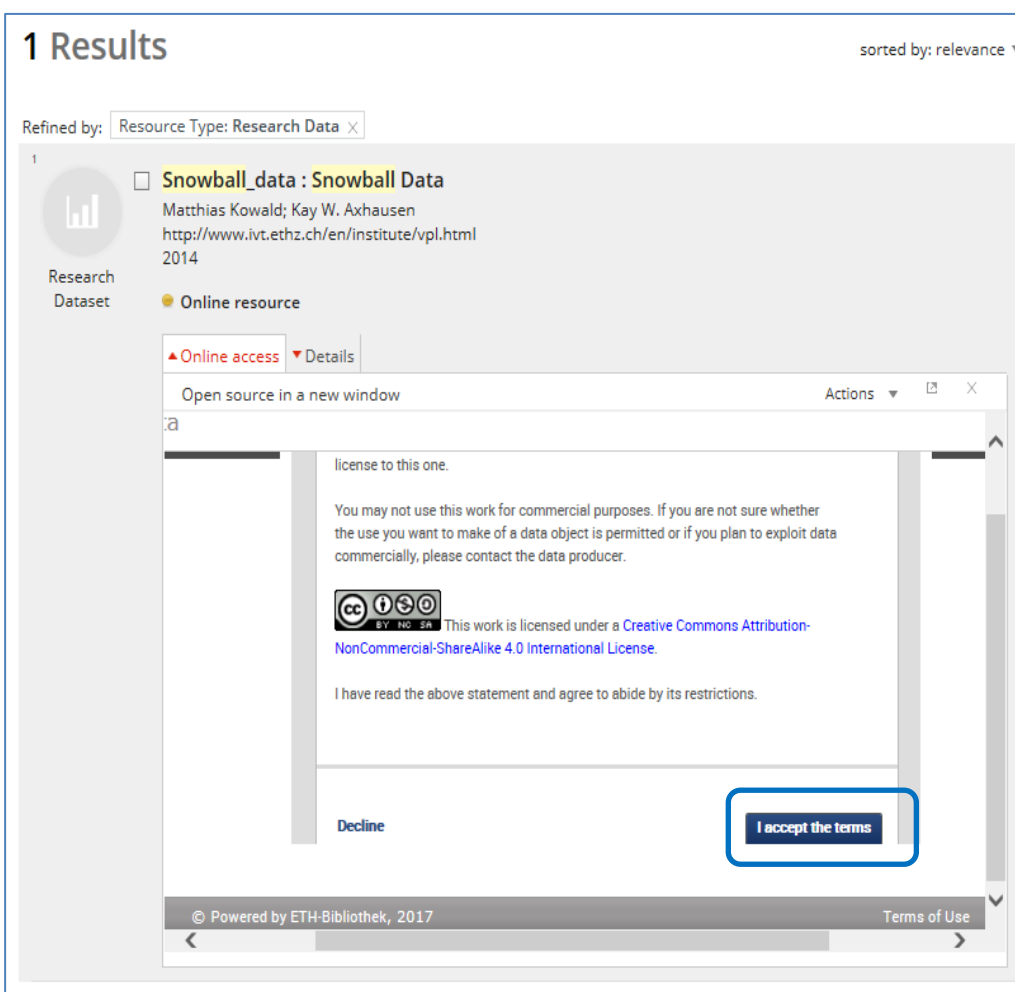
- 2) フィルターで研究データのものに絞り込む (Resource Type – Research Data を選択)。



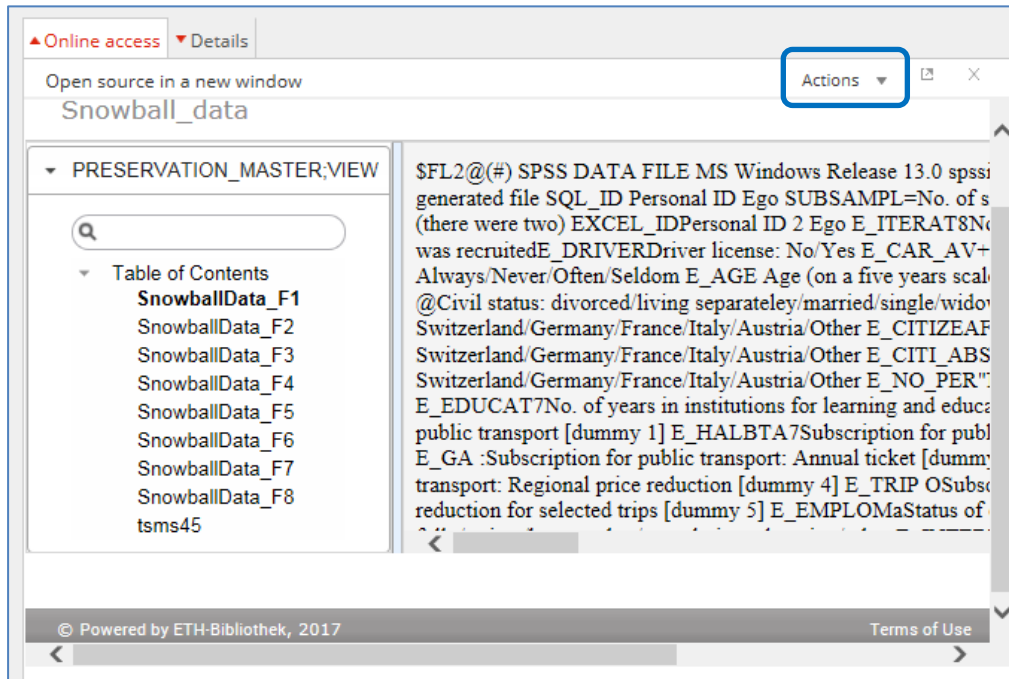
- 3) 研究データの詳細を表示する。



4) Online access をクリックし、データ利用条件の確認画面で同意すると、ファイルが開く。



5) Actions をクリックすると、データ取得方法が複数表示される。



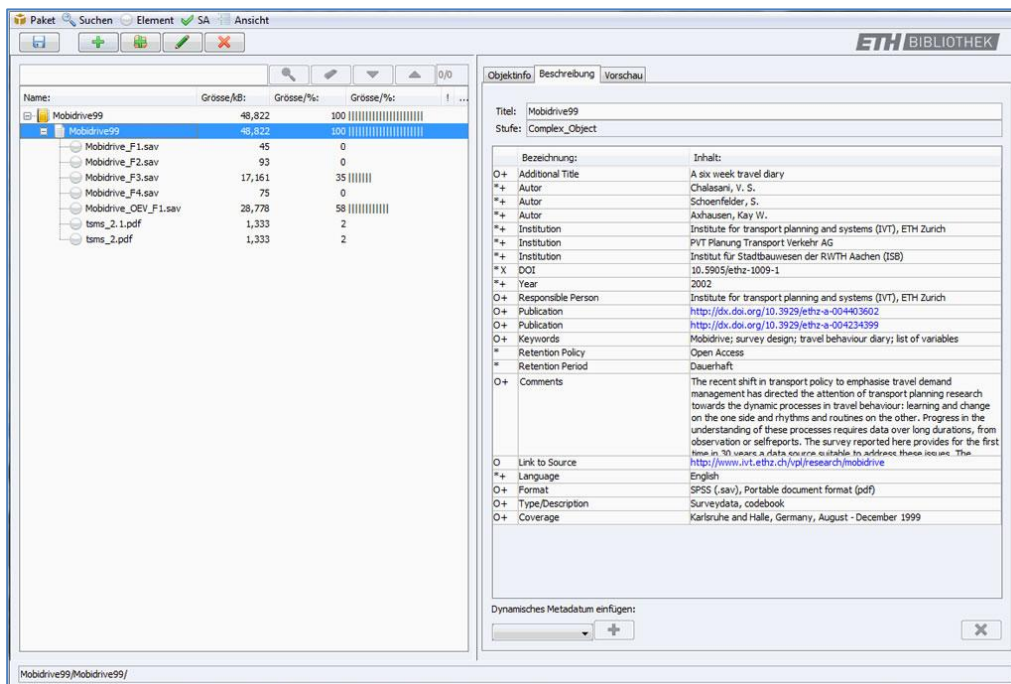
2.4.2. ETH Data Archive への研究データ登録手順

- 研究者がウェブサイトで自分のアカウントにログインしてデータを提出し、図書館職員が技術的な確認を行った後、確定して公開する。複数のデータファイルを頻繁にアップロードする研究者に対しては、より簡便なデータ処理が可能な支援ツールソフトウェア **docuteam packer** を提供している。
- リポジトリには **DROID**、**JHOVE** がプラグインとして組み込まれており、データファイル自体が持つファイル情報から技術的なメタデータを抽出している。

2.4.2.1. 研究データ登録支援ツール docuteam packer

<https://blogs.ethz.ch/innovethbib/en/2016/09/28/docuteam-packer-einfaches-tool-fuer-die-wissenschaft/>

- **docuteam packer** は、リポジトリへの研究データ登録を支援するオープンソースソフトウェア。
- 研究者は自分の PC にこのソフトウェアをインストールし、ソフトウェアの指示に従い必要項目を入力してメタデータを作成し、複数ファイルをリポジトリに登録することができる。
- メタデータおよび DOI の付与レベルを、フォルダ単位（複数のデータに対してまとめて 1 件）またはファイル単位から選択できる。
- 画面例



2.4.3. 新統合プラットフォーム Research Collection

現在、E-Publications プロジェクト (<http://dx.doi.org/10.3929/ethz-a-010668112>) を進めており、現行の ETH E-Citations、ETH E-Collection、ETH Data Archive を、2017 年に Research Collection として統合する（プラットフォーム名はプロジェクト名とは異なる Research Collection となる予定）。

- Research Collection は、DSpace で構築する。Research Collection の運用開始後も、ETH Data Archive は並行運用を予定している。

2.5. 研究データ管理計画策定支援

- サービスがまだ広く知られていないこと、研究データのオープンアクセスが現時点で義務化されていないことから、相談件数は月平均数件程度にとどまっている。
- Swiss National Fund が研究データ管理計画の提出を義務化する動きもあり、今後の需要の高まりも勘案しつつ、現在は少人数体制でベストプラクティスを模索している。
- サービス周知のために、アウトリーチ活動も行っている。
- 研究者のニーズは多岐にわたっているため、個別、直接に対応している。
 - 研究分野に特有の要件に対応する必要がある。
 - 研究分野で国際共同研究がよく行われるかどうかによっても、研究者の理解度に差異が見られる。
 - ETH Zürich の構成員は国際的であり、文化的背景も多様である。同僚研究者との peer consultation を好んだり、自ら主体的に取り組むことを好んだりするなど、研究者の気質に則したアプローチを取る必要がある。
- データについて最も熟知しているのは作成した研究者自身であり、研究者自身がメタデータ（のもとになる情報）を記述しなくてはならない。図書館職員ができること、でき

ないことを率直に明確にして、研究者とのコミュニケーションを図り、互いに学び合う姿勢が重要である。

3. ETH Zürich 研究データ管理支援ワークショップ

Workshop “Data Management in Research – Why and How?”

日時: 2016年10月19日 13:15-17:00

場所: ETH Zürich, main building, room F 33.1

対象者: 大学院生、若手研究者

言語: 英語

説明者: Matthias Töwe、Ana Sesartic、大学全体の IT 部門職員 1 名

出席者: 22 名 (学外図書館職員等も含む)

3.1. 概要

- レクチャー形式の説明の間に、3回のグループディスカッションが行われ、グループごとに出た意見をまとめ、全体へ報告する形式が取られた。
 - 研究データ管理とは何か、なぜそれが自分にかかわりがあるのか
 - 研究データ管理に関する規則
 - 研究データの長期保存
 - Active data の管理
 - 大学の IT 部門が提供するサービスの紹介
- 大学の IT 部門 ID-SIS では、進行中の研究プロジェクトのデータ (active data) の管理、ソフトウェアソリューション、データ管理プラットフォーム openBIS (<https://openbis-eln-lims.ethz.ch/>) 等のサービスを提供している。

3.2. グループディスカッションの内容

(Exercise 1)

- 自己紹介、自分にとってのデータ管理とは何か
- 研究中のデータ管理をどうするか、研究終了後データをどう共有・公開するか

(Exercise 2)

- 他人が作成したデータを理解するためには何が必要か (自分で作ったデータですら時間が経つとわからなくなる ; 異なるバージョンのファイルが多数存在し、違いを判別できない)
- 長期保存の「長期」とは具体的にどのくらいか
- 自分が持っているいちばん古いファイルはどれくらい以前のものか、今も使うのか (定期的に更新しているファイルは、「使えない古いファイル」になることはない)
- ❖ 実際に使う可能性と、再利用に備えるための手間との間のバランスを取る必要がある
- ❖ 研究者とサービス提供側はお互いに学ぶことができるし、そうしなければならない

(Exercise 3)

- どのようなデータ管理ツールを利用しているのか (Dropbox – sharing、Mendeley – 文献管理等)
- どのような管理ルール (ファイルの名付けルール等) を定めているのか (管理ルールが研究グループの中で完全に共有されているわけではない; 一貫性がない)

3.3. ワークショップのまとめ

- 自分がすべきことは何か? / 今すぐはじめよう / できるだけシンプルなやり方で、でもそのやり方を常に疑え (“Think about what you do. / Start early. / Keep it as simple as possible – but distrust it.”)

4. 所感

ETH Zürich では積極的な研究データ管理の取り組みが行われているが、現在進行中のものも多くある。そのような状況で、大学が提供できるサービスと既存のサービス (分野リポジトリ等) を広く俯瞰して、利用者にとって現実的な解決策を見つけていこうという姿勢が印象的である。

また、図書館職員がすべての研究分野に精通しているわけではないことも認識したうえで、図書館職員と研究者がお互いの知識を補完しあい、研究者が主体となって決定するスタンスを取り、それを研究者にも明確に伝えている点は注目に値する。

ワークショップでは、グループディスカッションに多くの時間を割いており、参加者が自分のこととして考えることにより、単なる座学よりも効果が高まり、今後の具体的な行動につながっていくのではないかと感じた。

5. 参考資料

- ETH Library <https://www.library.ethz.ch/en/>
トレーニングコース
<https://www.library.ethz.ch/en/Services/Training-courses-guided-tours/Training-courses>
ワークショップ
<https://www.library.ethz.ch/en/Services/Training-courses-guided-tours/Workshops>
オープンアクセス
<https://www.library.ethz.ch/en/ms/Open-Access-an-der-ETH-Zuerich>
- ETH Library Digital Curation Office
<http://www.library.ethz.ch/en/ms/Digital-Curation-at-ETH-Zurich>
- ETH Zürich IT Security Initiative <http://itsecurity.ethz.ch/en/>
- ETH Scientific IT Services <https://sis.id.ethz.ch/>
- E-Publication プロジェクト ポスター <http://dx.doi.org/10.3929/ethz-a-010668112>
- Blog post on bridging the curation gap between research and libraries:

<http://blogs.ethz.ch/innovethbib/en/2016/07/25/fallstudie-zu-curation-gap-zwischen-forschungs-und-bibliothekspraxis/>

(全文) Sesartić, A.; Fischlin, A.; Töwe, M., “Towards Narrowing the Curation Gap—Theoretical Considerations and Lessons Learned from Decades of Practice”, *ISPRS Int. J. Geo-Inf.* 2016, 5, 91. <http://doi.org/10.3390/ijgi5060091>

- docuteam packer
<https://blogs.ethz.ch/innovethbib/en/2016/09/28/docuteam-packer-einfaches-tool-fuer-die-wissenschaft/>
- Workshop 資料
各種スライド <http://www.slideshare.net/ETH-Bibliothek>
Data Management Checklist <http://bit.ly/rdmchecklist>
Recommended File Formats for Archiving
<http://www.library.ethz.ch/en/Media/Files/File-formats-for-archiving>

以上

平成 28 年度 国立情報学研究所実務研修 報告書
 英国におけるデジタルコレクション管理および研究データ管理の現況調査

提出日：平成 29 年 2 月 18 日

所属：京都大学附属図書館

氏名：大村 明美

(訪問日程・機関)

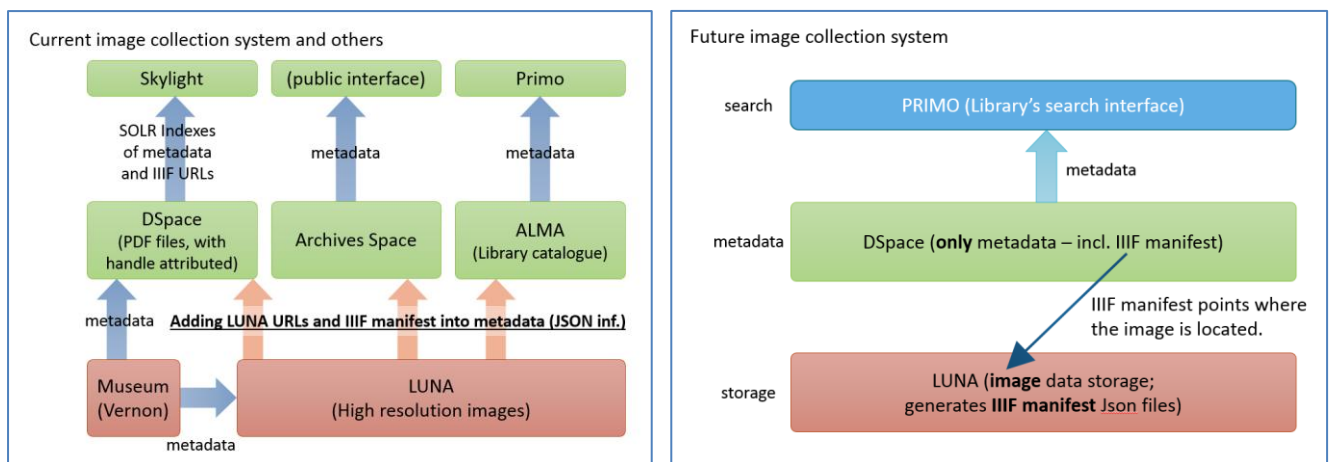
10 月 26 日 (水)	University of Edinburgh	Claire Knowles (Acting Head of Digital & Library Digital Development Manager) Scott Renton (System Developer) Cuna Ekmekcioglu (Senior Research Data Officer) Stuart Macdonald (Research Data Management Service Coordinator)
10 月 27 日 (木)	National Library of Scotland	Gill Hamilton (Digital Access Manager)
10 月 28 日 (金)	University of St Andrews	Alice Crawford (Digital Humanities Research Librarian) Patrick McCann (Applications Developer, Research Computing) Anna Clements (Assistant Director, Digital Research) Federica Fina (Data Scientist, Research Data Management)
10 月 31 日 (月)	University of Sheffield	Jez Cope (Research Data Manager)
11 月 1 日 (火)	University of Leeds	Rachel Proudfoot (Research Data Management (RDM) Advisor) Graham Blyth (Research Data Management Team Leader)
11 月 2 日 (水)	University of Oxford	Emma Stanford (Digital Support and Community Engagement Officer, Bodleian Digital Library Systems and Services) Mel Mason (Software Engineer, Bodleian Digital Library Systems and Services) John Southall (Bodleian Data Librarian and Subject Consultant for Economics, Sociology and Social Policy & Intervention)
11 月 3 日 (木)	University of Cambridge	Huw Jones (Digital Content Unit) Domniki Papadimitriou (Picture Library Co-ordinator Digital Content Unit) Dr Marta Teperek (Research Data Facility Manager) Rosie Higman (Research Data Adviser)

1. デジタルコレクション管理

1.1. University of Edinburgh

1.1.1. The University of Edinburgh Collections¹ 開発の経緯

- 2013年、前システムでは複数サーバで管理されていたデジタルコレクションを統合することを念頭に、関係する全学の図書館職員が集まりプロジェクトを開始した。
- 図書館資料（紙媒体、オープンデジタル資料）の他、大学史資料、教材、libraryが扱う楽器や絵画といった美術資料等、様々な種別の資料のデジタル画像を対象とする。
- 2014年5月DSpaceに統合、インターフェースとしてSkyLight²を使用し、各コレクション専用のウェブサイトを作成している。
 - <http://collections.ed.ac.uk/> Collections
 - <https://exhibitions.ed.ac.uk/> Exhibitions
 - <http://www.scottishgovernmentyearbooks.ed.ac.uk/> Scottish Government Yearbooks
 - <http://collections.ed.ac.uk/art> Art Collection
 - <http://collections.ed.ac.uk/mimed> Musical Instruments Museums
 - <http://collections.ed.ac.uk/calendars> Calendars
 - <http://collections.ed.ac.uk/iconics> Iconic Items
- 統合プロジェクト進行中にIIIF準拠の必要性が認識されるようになり、2015年12月からIIIFへの対応を開始。LUNA³を導入して高精細画像を格納している。画像は、既存のTIFFデータからJPEG2000へ変換した。
- 将来的には、図書館のディスカバリーシステムDiscoverEd⁴ (Primo)との統合を検討しているが、どの資料レベルのメタデータを流し込むか等の問題があり、未実現。



(図はインタビュー内容をもとに筆者が作成)

¹ The University of Edinburgh Collections <http://collections.ed.ac.uk>

Reusing repository technology for Cultural Heritage and Special Collections

https://figshare.com/articles/Reusing_repository_technology_for_Cultural_Heritage_and_Special_Collections/1089651

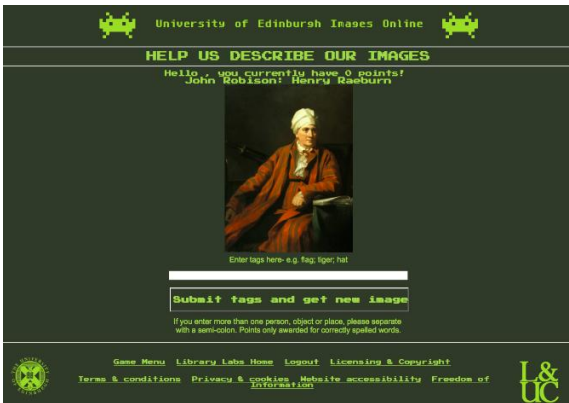
² SkyLight: Expose your collections <https://github.com/skylightui/skylight>

³ LUNA and IIIF <http://www.lunaimaging.com/iiif/>

⁴ DiscoverEd http://discovered.ed.ac.uk/primo_library/libweb/action/search.do?vid=44UOE_VU1

1.1.2. レガシーデータのメタデータ整備

- システム構築に当たっては、ひとまずデータを搭載して公開できるようにすることを優先し、その後メタデータの追加・向上を行うこととした。メタデータが不十分な、過去に電子化したレガシーデータについては、資料を再撮影して画像データを追加したり、従来のメタデータに、専門家以外の人にも理解しやすい一般向け記述を追加するなどの対応も行っている。
- 限られたリソースで事業を進めるために、利用者がメタデータ（タグ情報）追加に参加するプロジェクト Library Labs Games を企画した。6セッションで4652件のタグが付与された。⁵



Material	Oil on Canvas/Oil Paint
Dimensions	127 x 101.6 cm
Type	Painting
Subject	John Robison (b.1739, d.1805)
Collection	Art Collection
Classification	oil ; portrait ; paintings 1600-1800 ; painting ; painting
Accession Number	EU0010
Zoomable Image(s)	
Tags:	<input checked="" type="checkbox"/> Seated <input checked="" type="checkbox"/> Chair <input checked="" type="checkbox"/> Robes <input checked="" type="checkbox"/> Stripes <input checked="" type="checkbox"/> Hat <input checked="" type="checkbox"/> Collar <input checked="" type="checkbox"/> Ruffle <input checked="" type="checkbox"/> Book <input checked="" type="checkbox"/> Turban <input checked="" type="checkbox"/> Telescope
Add more tags at Library Labs Games (Create a login at Edinburgh Friend Account)	
BACK TO SEARCH RESULTS	

1.1.3. デジタルコレクションの再利用と共有

- The University of Edinburgh Collections では画像ダウンロード機能もあるが、利用者に積極的に周知していない。利用許諾条件に従った利用が確保されるように、利用者教育をする必要があると考えている。

(例) Art collection: Dr O Charnock Bradley (1936)

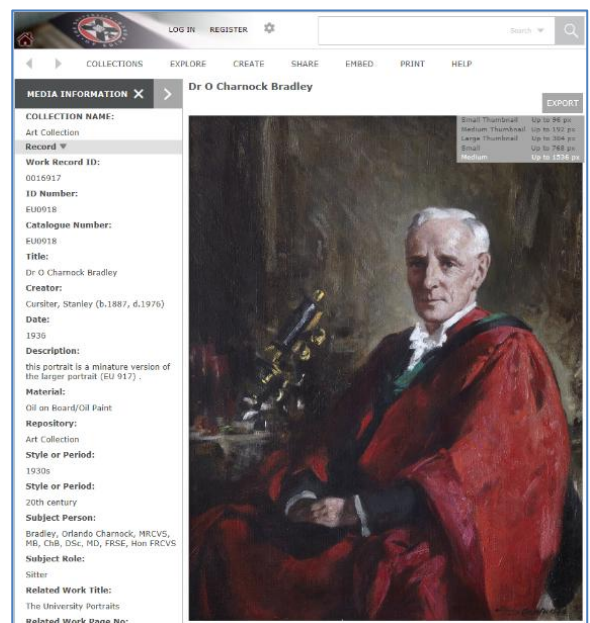
<http://collections.ed.ac.uk/art/record/20088?highlight=bradley> > zoomable image をクリック

<http://images.is.ed.ac.uk/luna/servlet/detail/uoart~1~1~4290~100397#>

画面右上 Export からサイズを選択してダウンロード

- IIIF の仕組みを利用して、同じテーマの資料を持つ図書館と共同で、サブジェクトポータルを構築する取り組みも行われている。

(例) ヨーロッパの楽器データのサブジェクトポータル MIMO: Musical Instrument Museums Online



⁵ Raising the Curtain or: getting unique and rare collections out to the public

<https://www.era.lib.ed.ac.uk/handle/1842/10415>

LibraryLabs Pop Up at Central <http://libraryblogs.is.ed.ac.uk/popuplibrary/>

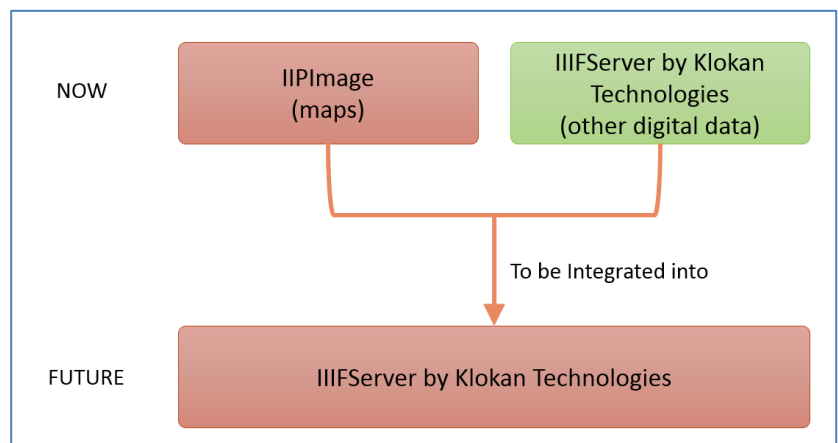
1.2. National Library of Scotland

1.2.1. National Library of Scotland 概要

- 前身は 1682 年に設立された法律図書館 Advocates Library で、1710 年には Copyright Act によりスコットランドの法定納本図書館となった。1925 年、National Library of Scotland Act が制定され、NLS となる。
- 2016 年 9 月、Glasgow に動画コレクション（フィルムおよびデジタル）に重点を置いた分館 Kelvin Hall を開館。
- 蔵書数約 2,400 万点（印刷資料 1500 万点、マニュスクリプト 700 万点、その他ポスター、絵画、動画、音声資料等）
- 2013 年 4 月以降、電子出版物も納本対象となった。⁶
- 地図資料は世界 7 指に入る 12 万点のコレクションを持つ。⁷
- 2016 年 11 月、研究や教育、リソースへのアクセス強化に関する連携協定を University of Edinburgh と締結。

1.2.1. デジタルコレクションと IIIF 対応

- National Library of Scotland Digital Gallery <http://digital.nls.uk/gallery.cfm>
 - project ごとに画像が並べられている（ブラウザ）
 - google のサイト内検索で検索
- 2025 年までに蔵書の 3 分の 1 をデジタル資料とすることを目指している（年々ボーンデジタルの資料が増加し、印刷出版物が減少傾向にあることも背景）。
- Hatti Trust に他館所蔵資料の電子化済データがあれば、当該データへリンクする。
- 2016 年 4 月 IIIF コンソーシアムに加盟。現在、IIIF サーバをテスト中。
- IIIF 対応の手順
 - 1) 資料を電子化し、マスターデータとして TIFF 画像を作成する
 - 2) TIFF 画像を JPEG2000 へ変換（kakadu⁸を使用—大量のファイルを高速に処理できる）
- 現在画像サーバとして、地図データには IIPImage⁹（IIIF 未準拠）、それ以外には IIIFServer¹⁰（IIIF 準拠）を使用。
- 今後、新サーバを導入して IIPImage を IIIFServer へ統合、JPEG2000 画像の作成、IIIF manifest の作成、インターフェースの開発等を予定。



（図はインタビュー内容をもとに筆者が作成）

⁶ National Library of Scotland Legal deposit <http://www.nls.uk/about-us/what-we-are/legal-deposit>

⁷ National Library of Scotland Map images <http://maps.nls.uk/>

⁸ kakadu software <http://kakadusoftware.com>

⁹ IIPImage <http://iipimage.sourceforge.net>

¹⁰ IIIFServer by Klokan Technologies <http://www.iiifserver.com>

- 資料が持つ文化特有の問題（例：ページ繰りの方向の違いー本を右から左へ読む日本語、ヘブライ語等）を共有して IIF の仕様に反映させることが重要。また、IIF に準拠しながら、かつ学内認証システム（LDAP 等）と連携させてアクセスを制御（大量ダウンロードを制限等）する等新しい機能も、現在 IIF メンバーで検討されている。欧米だけでなく、世界各国から多くの機関が参加して積極的に声を上げてほしい。IIF International Conference（2016 年はニューヨーク開催¹¹）等の機会もある。

1.3. University of St Andrews

1.3.1. 概要

- 1413 年設立（スコットランドで最初、英語圏で 3 番目）、学生数 8,400 名（120 カ国）
- St Andrews の人口約 17,000 人の半数が大学の学生・教職員で、大学の建物は明確にキャンパスとして区分されておらず、町の中に溶け込んでいる。
- 1976 年図書館設立

1.3.2. デジタルコレクション管理

- Digital Collections
<https://www.st-andrews.ac.uk/library/specialcollections/collections/digitalcollections/>
- 担当職員 2 名
- 2012-2013 年に構築、現在 312 のデジタルオブジェクトからなる（800GB）。
- フレームワークとして Islandora¹² を使用（ビューワーの性能を評価して選択）。
- Drupal（ユーザーインターフェース）、Fedora Commons（ストレージ）、Solr（検索システム）で構成される。
- 電子化の手順
 - 1) 資料保存を目的として、または利用者からの電子化依頼を受けて、資料を電子化（スキヤニング）（図書館内に reprographics technicians - 専任職員が 2 名（パートタイム）；Bookeye scanners 2 台設置）
 - 2) 電子化したデータを Digital Assets Register（スプレッドシート）に記入
 - 3) Digital Assets Register の中から、ウェブサイトで公開するものを決定
 - 4) 著作権の有無をチェックし、あればインターネット公開できるように許諾を得る
 - 5) フリーソフト GIMP¹³ を使用して画像を公開用に加工
 - 6) zip に圧縮して Islandora へアップロード
 - 7) メタデータを付与（目録レコードからの加工）
 - 8) デジタルコレクションを公開
 - 9) 広報（Library blog, library news letter, “Echoes from the Vault blog”, newsletter of our Digital Humanities Network 等）
- 現在、Europeana への参加を検討中（メタデータの変換など、必要な条件を満たせるか、確認中）

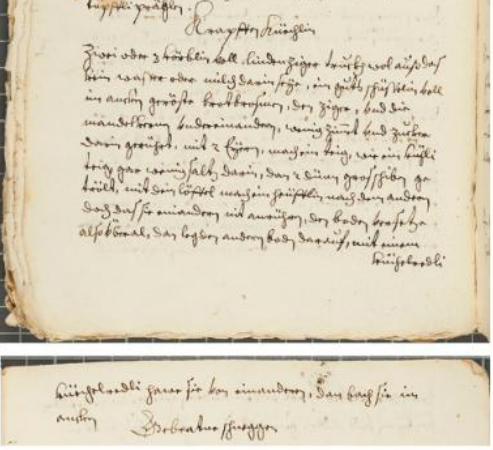
¹¹ Access to the World's Images - New York 2016 <http://iif.io/event/2016/newyork/>

¹² Islandora <http://islandora.ca/>

¹³ GIMP <https://www.gimp.org/>

1.3.3. 電子化プロジェクト例


- Swiss Renaissance recipe book. c.120 recipes in two distinct hands
 - スイスのドイツ語のレシピ本。実際にレシピどおりに料理し、ブログで公開(右図)¹⁴。
- Library Receipt Books (Borrowing Registers)¹⁵
 - 図書貸出簿の閲覧リクエストが多いことがきっかけとなり立ち上げたプロジェクト。
 - 18世紀の図書貸出簿(貴重資料、マニユスクリプト)を電子テキスト化して、検索可能にする(University of Oxford や University of Cambridge では図書館資料の館外持ち出しが禁止されていたが、St Andrews は貸出していたからこそできるプロジェクト)。
 - これを、1747-1898年に大学に所属していた学生、教職員の biographical record (電子テキスト化中、未公開)、および Library Catalogues¹⁶ とリンクすれば、いつどの研究者がどのような資料を閲覧して研究していたかがわかり、その研究者への理解を深めることができる。
 - 限られた予算・人手を行うために、crowdsourcing solution としてボランティアを利用(2016年6月パイロットプロジェクトを実施)。関心のある図書館職員(受付待機時等にバックグラウンド業務として作業)、Friends of the Library (多くは退職した卒業生で、図書館に関心を持つ人々。自宅で時間がある際に行う)が参加。大学の e-mail ニュースレター (in the Loop) で募集することも検討。




The recipe for Kropffen cakes

” Press out two or 3 baskets of [soft goat’s cheese], so no water or milk remains in them, a good bowl of breadcrumbs fried in butter, and mix together with the cheese and almonds, add a little cinnamon and sugar, make a dough with 2 eggs, like a cake dough, add a little salt, then take two large thin sheets, using a spoon make small heaps, and so that they are not touching spread the balls out, then put on the other balls, using a kitchen wheel separate them from each other, then bake them in butter.

There were two attempts at the second recipe. Kropffen cakes. This is one of the few recipes in the book to include cheese, in this case goat’s cheese. It sounded like it might be interesting, mixing the cheese with cinnamon and sugar, and crunchy breadcrumbs. The instructions were quite vague, particularly as to quantities and technique. For the first try the goat’s cheese mix was made up and then eggs were added to make this into dough. This produced hard gritty little rock cakes. That didn’t seem quite right.



Rock buns (first go)



¹⁴ HISTORICAL COOKING: WEEK 2 – MILK SOUP AND KROPFFEN CAKES

<https://standrewsrarebooks.wordpress.com/2016/10/20/historical-cooking-week-2-milk-soup-and-kropffen-cakes/>

¹⁵ Library Receipt Books (Borrowing Registers)

<https://arts.st-andrews.ac.uk/digitalhumanities/fedora/repository/digitalcollections%3A1862>

18世紀の18冊を対象とする。この時期にはスコットランドの歴史の中でも多くの書籍が出版され、研究者の需要が高いことがわかっていた。

¹⁶ Library Catalogues

<https://arts.st-andrews.ac.uk/digitalhumanities/fedora/repository/digitalcollections%3A4824>

図書貸出簿に記載されている資料名の省略形を調査可能。貴重資料、電子化済。

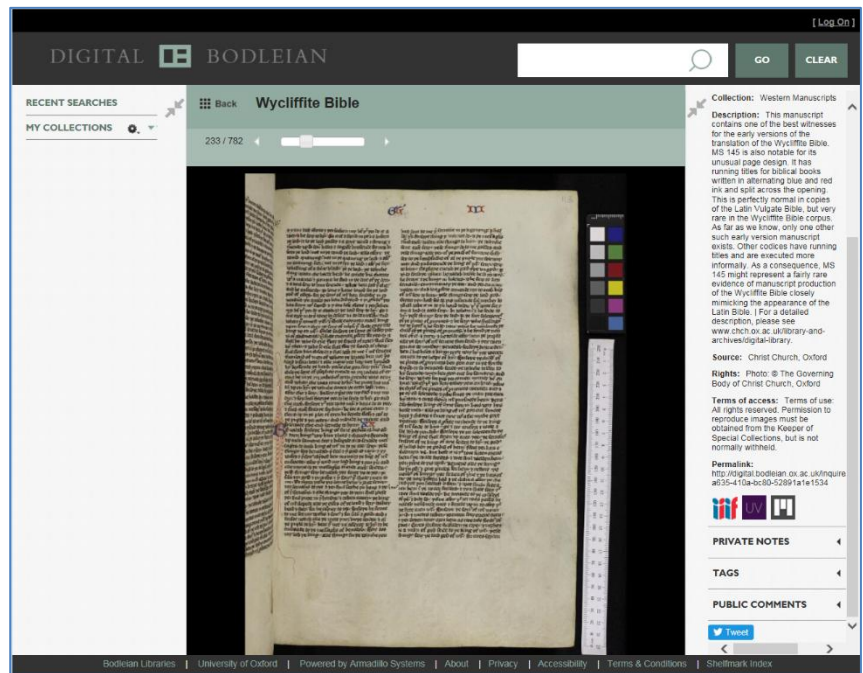
1.4. University of Oxford

1.4.1. 資料の電子化

- 通常、外部資金を獲得して行っている。
- 対象資料の決定に際しては、壊れている資料を優先し、教員の推薦、キュレーターや時には利用者からのリクエストを考慮するが、利用者のオンデマンドによる電子化サービスは行っていない。
- 電子化作業の管理に Goobi¹⁷ を使用。
- 過去に電子化したデータのメタデータの整備（メタデータが不十分）、インフラの維持管理（ソフトウェア更新等）が課題。

1.4.2. Digital Bodleian <http://digital.bodleian.ox.ac.uk/>

- 2015 年公開。過去に図書館が電子化した、諸々の事情で複数サーバに分散していたデータを統合した。今後も順次統合予定¹⁸。独自のウェブサイトを持っているカレッジの図書館もあるが、自館内で電子化しても、通常データ処理は Bodleian Library で行っている)。
- 原資料の目録データと画像データは一部リンク済み。
- フレームワークとして iNQUIRE¹⁹、イメージサーバとして IIPImage を使用。
- Private notes, tags, public comments を保存する機能は、iNQUIRE によるもの (IIIF のアノテーション共有機能ではない)。
- コレクショントップレベルおよび個々の画像データレベルに対して IIIF manifest を作成している (IIIF ロゴマークから IIIF manifest へリンク)。



1.5. University of Cambridge

1.5.1. 資料の電子化

- 図書館の Digital Content Unit では、個々の電子化プロジェクトの都度資料に適した高価な機材を導入(更新)して、貴重資料を慎重に丁寧に電子化している(大量・スピード電子化は対象としない)。

¹⁷ Modules for the digital library: Goobi <http://gdz.sub.uni-goettingen.de/en/entwicklung/workflow-goobi/>

¹⁸ <http://www.bodleian.ox.ac.uk/subjects-and-libraries/collections/digital> (Plus 以降のコレクション)

¹⁹ iNQUIRE <http://www.inquireresearch.co.uk/>

- 脆い資料に対応出来る機材（角度調整可能な台や気圧を利用したページ抑え）
- 立体資料を多方向からカメラ撮影して 3D データを作成することが出来る機材
- 数メートル四方の大きなサイズの資料を撮影出来る機材（資料台を動かしてグリッドごとに撮影。少しずつオーバーラップして撮影した画像をつなぎ合わせる。カメラの色彩調整等が狂う心配がない）
- 化学物質が含まれる 19-20 世紀の写真資料等の方がそれより古い資料よりも脆い。
- 資料の材質等に合わせて照明を調整する（例えば金を使用したものだと通常の照明では褪色して見えるので、調整が必要）。
- 資料の 1 ページ、一部のみを電子化した場合は、公開していない。資料全体がそろってこそ意味・価値を持つと考えるスタンスから、資料ごとの電子化・公開が原則。
- 既存の外部システムが利用可能な場合は、独自開発はせず、それを利用する。3D データは、[sketchfab](#)²⁰（大学のアカウントを作っている。フリー、データ制限なし）にアップロードして、それを Digital Library の画面に埋め込んで公開。動画についても同様に [YouTube](#)²¹ を使用。
- 研究プロジェクトとして研究者を巻き込んで電子化の助成金を獲得している。
 - 撮影したデータは研究者と一緒に確認し、研究者が専門的なメタデータを付与する。
 - 対象資料には、プロジェクトとして意味を持つ資料群を選定し、資料群としての新しい価値を付与するよう努めている。
 - 教員にとって、自らの研究成果になること、図書館が提供するインフラを利用できることがインセンティブとなっている。
 - 研究データの長期保存につながる（教員が独自で立ち上げたウェブサイト等は、短期間で消滅してしまうことが多い）。
 - 教員の協力を得たり、教員のプロジェクトの情報を得たりするためには、図書館から直接働きかけるよりも、図書館と協力して恩恵を感じた研究者が、ほかの研究者に図書館の活動を知らせてくれるほうが効果的である。

1.5.2. Cambridge Digital Library <https://cudl.lib.cam.ac.uk/>

- Digital Library のビューワーは、自館で [OpenSeadragon](#)²² を利用して開発。
- 画像は TIFF 形式。
- レガシーデータにも必要に応じて遡及的にメタデータを付与している（研究者やキュレーターの協力を得ている）。
- 3 種類のメタデータスキーマを採用。
 - TEI (Text Encoding Initiative)²³ マニュスクリプト
 - EAD (Encoded Archival Description)²⁴ アーカイブ資料

²⁰ Sketchfab Cambridge Digital Library <https://sketchfab.com/CamDigLib>

²¹ YouTube Cambridge University Library (the UL) <https://www.youtube.com/user/CamUniLib>

²² OpenSeadragon <https://openseadragon.github.io>

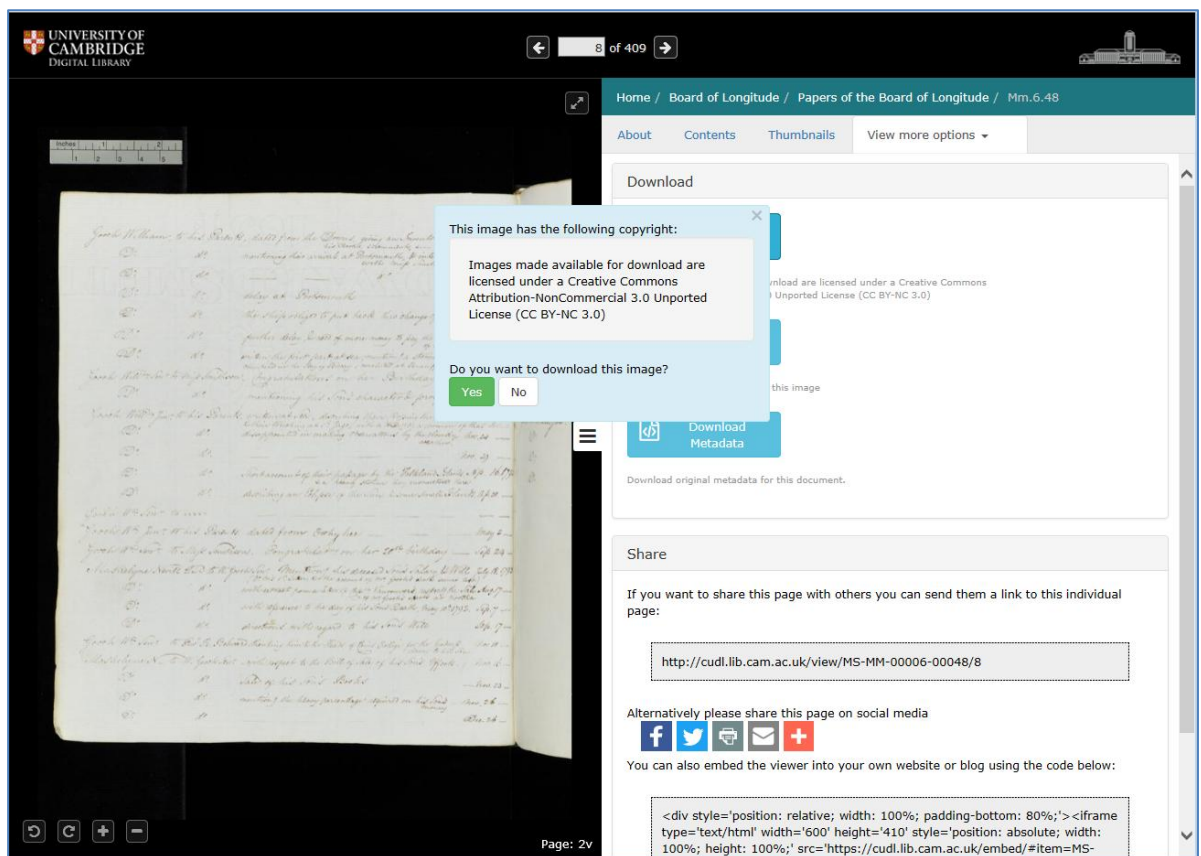
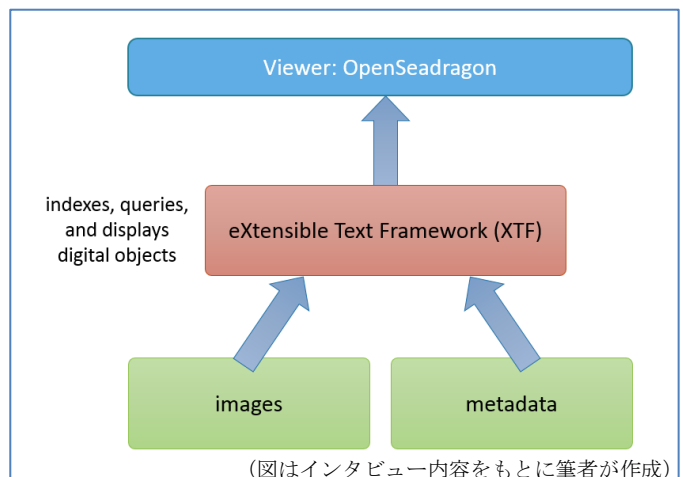
²³ TEI: Text Encoding Initiative <http://www.tei-c.org/index.xml>

²⁴ EAD: Encoded Archival Description <https://www.loc.gov/ead/>

(例) <https://cudl.lib.cam.ac.uk/view/MS-MM-00006-00048/1>

➤ METS (Metadata Encoding & Transmission Standard)²⁵ 刊本

- XTF²⁶ を使用して、画像とメタデータを処理。
- IIIF 対応を検討中 (TIFF を JPEG2000 に変換した場合にデータが劣化することを懸念しており、問題がないか技術担当が検討中)。
- 現在システム独自の永久識別子を付与しているが、研究データとして DOI を付与する案も出ており、今後検討する可能性がある (ただし優先順位は低い)。
- 画像の一括大量ダウンロードはできないようにしている。ただし、メタデータの一括ダウンロードは研究者からの要望が多いため、今後可能にすることも検討している。



²⁵ METS: Metadata Encoding & Transmission Standard <http://www.loc.gov/standards/mets/>

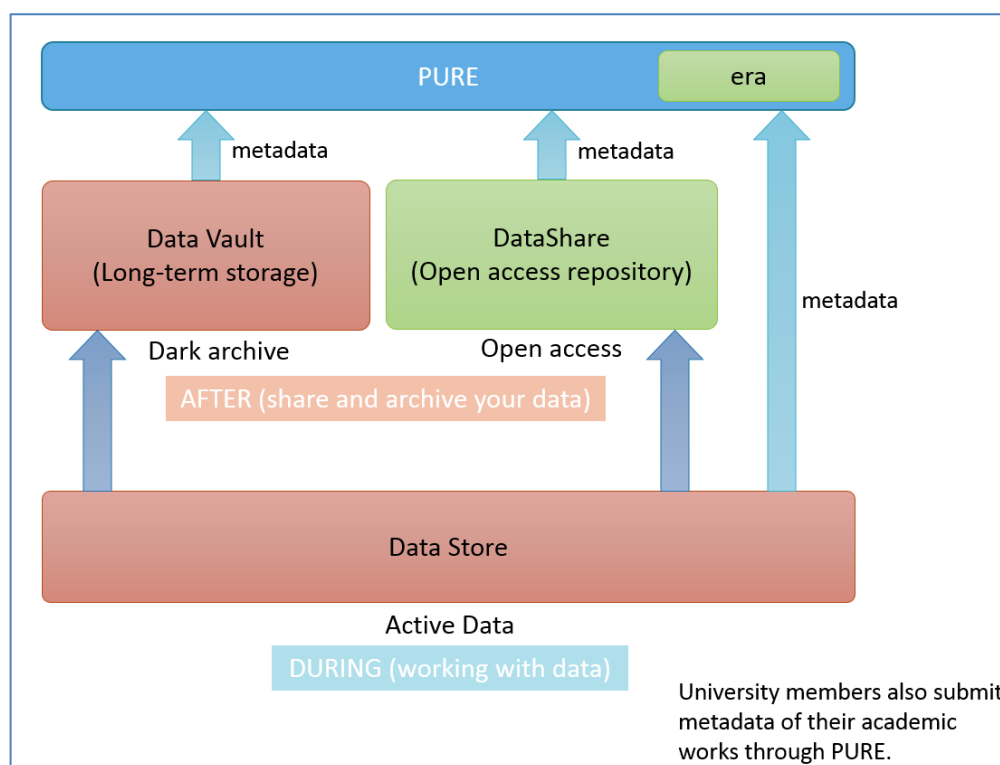
²⁶ eXtensible Text Framework (XTF) <http://xtf.cdlib.org/xtf/>

2. 研究データ管理

2.1. University of Edinburgh

2.1.1. システム構成

- 研究業績管理システムとして、Pure²⁷ を使用（非公開、学内のみ）。教員は研究成果のメタデータの登録を義務付けられている。
- Pureに登録された業績のうち公開可に指定されたデータが、機関リポジトリ era (Edinburgh Research Archive)²⁸ として公開されている。その他、現在進行中の研究活動情報が Edinburgh Research Explorer²⁹ でも公開されている。
- オープンアクセスの研究データリポジトリとして DataShare³⁰ が、ダークアーカイブ用の研究データ（研究終了後）リポジトリとして DataVault³¹ が提供されている。



(図はインタビュー内容をもとに筆者が作成)

- オープンアクセスの DataShare ではなく、アクセス制限の可能な DataVault にアーカイブする教員が多い。
- 通常、DataShare 等のリポジトリにデータが登録された際に DOI が付与され、Pure にもその DOI が格納される。DOI 未付与のメタデータが Pure に登録された場合、Pure の機能を使用して DOI を付与することもできる（ただし公開データに対してのみ）。

²⁷ <https://www.elsevier.com/solutions/pure>

²⁸ era: Edinburgh Research Archive <https://www.era.lib.ed.ac.uk/>

²⁹ Edinburgh Research Explorer <http://www.research.ed.ac.uk/portal/>

³⁰ DataShare <http://datashare.is.ed.ac.uk>

(メタデータ例) <http://datashare.is.ed.ac.uk/handle/10283/2132?show=full>

³¹ Jisc Data Vault <http://libraryblogs.is.ed.ac.uk/jiscdatavault/>

2.1.2. 研究データ管理サービス

(組織体制)

- 図書館だけでなく、学内関係部署と連携してサービスを提供している。
 - 全学の IT 担当部署がインフラ整備を、図書館の IT 担当チームがソフトウェア整備を担う。
 - リサーチ・アドミニストレーター、IT 担当職員、各部署のデータ・コーディネーターとも連携。
 - フォーラムや SharePoint 等を立ち上げ、研究スタッフや学生、研究支援・IT 担当職員等と情報共有。
 - 学内の記録管理担当、研究支援室等にも研究データ管理に関する情報を共有。
 - セキュリティや法律に関わる案件については、大学の研究支援・法務担当部署が対応。
- 図書館からは MANTRA によるトレーニングを修了した職員がサービス提供に当たっているが、実際に業務を行っているのは 2 名のみ。トレーニング修了後も自信を持てず躊躇する職員が多い（自信を持たせるには時間がかかる）。
- データライブラリアンは、データに関わる業務（データ分析、カタログニング、インデクシング等）に広く携わり、図書館業務や IT 知識が求められるが、現在データライブラリアンとして活躍している人に聞くと、「やるべきことをやってきたらデータライブラリアンといわれるようになった」“accidental data librarian”が多いという。

(対象)

- 3 つのレベルを意識して行っている。
 - 研究部門等の責任者: まず責任者に理解を深めてもらい、その上で責任者から管理下の研究者へ周知してもらう
 - 研究グループの代表者 Principal Investigator: 助成金獲得の条件として研究データ管理等が求められていることに注意を喚起
 - 研究スタッフ、大学院生等

(内容)

- 全般的な知識（MANTRA によるオンライン教育等）、分野独自の留意事項
- 「自分は研究データを持っていない」という研究者でも、研究データが何であるかの認識がないだけで、実際には持っていることがほとんど。まず何が研究データであるか知ってもらう必要がある。
- 研究データ管理計画策定のツールとして、DMPOnline³² を活用。DMPOnline は継続的に新機能追加等が行われており、最近では分野、機関用にカスタマイズされたテンプレートを選択したり、ガイダンスを参照したりすることができるようになっている。
- 利用者からの需要が多いのは、研究プロジェクト進行中のアクティブ・データの保存、研究終了後のデータ保管（特に倫理や個人情報に関わる管理に注意を要するデータの取り扱い）等。

(形態)

- レクチャー形式、事前予約（参加者のプロフィールを事前に確認して、より適した内容を準備）
- オンデマンドのワークショップ（多忙な教員に人気）
- 研究者を待たせないのが重要。サービス提供を予告しておきながら実施開始が遅れると、待ちくたびれて別のサービスへ流れていってしまう。

³² DMPOnline <https://dmponline.dcc.ac.uk/>

- 利用者からのフィードバックを得るためにウェブフォームも提供しているが、むしろインフォーマルな場で率直な意見を聞くことができている。

2.2. University of St Andrews

2.2.1. システム構成

- 研究業績管理システムとして Pure を使用。
- 機関として ORCID に参加し、研究者に ORCID の利用を推奨。Pure を通して ORCID と連携している。³³
- Pure には研究業績のメタデータだけでなく、研究データ本体も登録している。DOI も Pure で付与する。³⁴
- Pure に入力されたデータのうち、研究者本人が公開可に指定したデータが St Andrews Research Portal として公開される。³⁵
- DSpace による機関リポジトリも提供している。³⁶ スコットランドの機関間協力のひとつ、Scottish Digital Library Consortium (SDLC)が提供している、リポジトリのホスティングサービス (University of Edinburgh が開発) を利用している。
- 教員はウェブインターフェースからメタデータを入力する (図書館職員が確認してから公開する)。
- データをアップロードする際、ウェブインターフェースでは5分を超えるとタイムアウトになるため、巨大ファイルには対応できず、別途処理する必要がある (全体のデータ量としては、アップロード制限はない)。
- 現在、アクティブデータのためのアーカイブの設置を IT 担当と検討中 (現状では、ほとんどの研究者が個人のデバイスや Dropbox 等のクラウドサービスにデータを保存している)。

2.2.2. 研究データ管理サービス

(組織体制)

- 図書館が主導し、IT 担当、研究支援担当、倫理委員会、財務担当等の職員と連携している。
- 例えば倫理上注意を要する研究データの扱いについては、学内の倫理委員会が相談に応じる等、内容によって関連部署が連携する。その際も、サービス提供がシームレスに行われるよう留意している。

(内容)

- 研究データ管理計画策定に際しては、DMPOnline を使っている。
- 図書館職員はサービス提供に際して、どのような要件に注意すべきか等、経験を積んでいく必要がある。研究プロジェクトによって事情が異なるため、定型を準備するのは難しい。
- 特に需要の高いサービスは、ストレージ、研究データ管理計画策定、長期保存。

(研究者の意識向上のための工夫)

³³ ORCID Member Organizations <https://orcid.org/members>

³⁴ Pure – Research Information System <https://www.st-andrews.ac.uk/staff/research/pure/>

³⁵ St Andrews Research Portal <https://www.st-andrews.ac.uk/staff/research/pure/faq/portal/>

³⁶ St Andrews Research Repository <https://research-repository.st-andrews.ac.uk/>

- 教員の間での周知を徹底し理解を深めるために、さらなる取り組みの必要性を感じている（部局へ説明に赴く、イベントやミーティングを開催する、巡回相談デスクを設ける、方針について広報する）。
- 職員と大学院生と一緒に研究データ管理のコースを教える。
- 新入生に簡単なイントロダクションを提供する。新入生の意識が高まると、その指導教員も触発されて、積極的に Pure に入力するようになる。

（長期保存）

- 研究データ保存の際には、データから研究結果を再現できるようにすることが不可欠だが、オープンアクセスにするためにオリジナルデータを変換しなければならないこともあり、両者のバランスを取る必要がある。

2.3. University of Sheffield

2.3.1. システム環境

- 研究終了後の成果をオープンアクセスで公開する研究データリポジトリとして ORDA: Online Research Data³⁷ (figshare for institutions³⁸ で構築) を、アクティブデータやアクセス制限が必要なデータ用のファイルストレージサービスとして Google Apps for Education³⁹ を提供している。研究グループ等で独自のシステムを持っている場合もあるが、全学サービスへ統合する方向。
- ORDA では、まず研究者がデータ（メタデータを含む）をアップロードし、図書館員が確認後公開する。研究者が入力すべき情報は、極力項目を少なくしている（入力インセンティブを高めるため）。
今後、分野独自のメタデータ項目も追加する方向で検討中（XML ファイルとしてアップロードする等）。
- ORDA にアップロードする研究データのファイル種別については、すべてのファイル種別を把握できないことから、そのままの形式で保存している。ただし、確実な長期保存を担保するために、Digital Preservation Manager がデータ変換も含めたデータ保存プロジェクトを検討中。
- ORDA では、永久保存ではない、例えば 10 年保存のデータにも DOI を付与している。データを削除する場合は、ランディングページを残し、データの削除理由や状況について説明を追記する予定（実例はまだない）。DOI 付与の粒度は研究者が決める。
- 研究データ保存先の選択優先順位としては、まずは適した分野/全国的リポジトリ、なければ ORDA。分野リポジトリや全国的なリポジトリにデータを登録した場合も、メタデータを ORDA にも登録する。
- 現在進行中のプロジェクトとして、データ保存用に Rosetta を導入（ORDA のデータのバックアップ・ダークアーカイブ用）。

2.3.2. 研究データ管理サービス

³⁷ ORDA - The University of Sheffield Research Data Catalogue and Repository

<https://www.sheffield.ac.uk/library/rdm/orda>

³⁸ Figshare / University of Sheffield <https://sheffield.figshare.com/>

³⁹ Google Apps for Education <https://edu.google.com/products/productivity-tools/>

(機関としての方針)

- 研究データ管理に際しては、助成機関の要件、法的義務（データ保護、情報開示）、契約義務等を考慮しなくてはならない。
- 機関に求められているのは、ポリシーの制定、基盤整備、トレーニングやガイダンスの提供。

(組織体制)

- 図書館が主導的な役割を果たす（従来から研究者と密接な関係を持ち、利用者サービスを提供してきたことから、最も適切であると考えている）。図書館は情報管理、学術出版、メタデータの専門知識により貢献し、IT 担当の CICS は技術的な専門知識でアクティブデータのストレージ等インフラを整備し、研究支援担当の RIS は助成金や契約その他の法的専門知識をもって研究データ管理計画策定支援やコスト面の相談に応じ、Research Software Engineering はデータの長期保存や研究成果再現等、ソフトウェア面の支援を提供する。より詳細な情報を持つ部署から利用者へ直接回答してもらうか、図書館が継続して窓口となるかは、内容の複雑さにもより、ケースバイケースだが、シームレスなワン・ストップ・サービスを目指している。
- 2016年1月、所属が散在していた Research Support Unit を統合。ユニット責任者1名、Research Data Manager 1名、Open Access Manager 1名で運営。
- トレーニング提供時には、managers2名のほか、3名程度の図書館職員が臨時に加わる。
- 各部局には、Departmental Library Coordinator (DLC) が存在し、窓口となる。
- 他部署の職員と、戦略的な優先順位を共有して協力する必要がある。したがって、従来のサブジェクト・ライブラリアンのような研究分野の専門知識よりも、専門分野の情報源に関する知識、コミュニケーション能力、教える能力が求められる。
- トップダウン、ボトムアップ双方の、様々なコミュニケーション方法を利用している。

(対象)

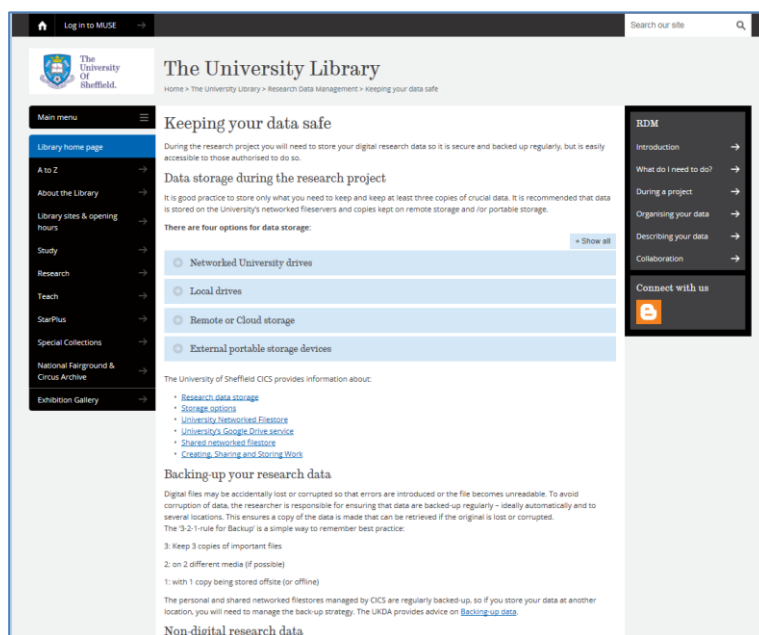
- 大学院生は新しい情報を学ぶ上で障壁が少ない傾向にあるが、すでに実績のある研究者は、新しい義務が生じることに抵抗を感じやすい。研究分野により傾向があるというよりも、個人のパーソナリティの問題。対応する際は、対面でのトレーニングが効果的。

(研究データ管理計画策定)

- 研究データ管理計画策定には DMPOnline を使用しているが、大学独自のテンプレートやガイドラインは未整備。今後追加予定。

(形態)

- 研究者の理解を深めるためにワークショップを開催したり、新大学院生向けのガイダンスや部局での説明会等を行っている。
- 電子メール、ニュースレター、twitter、対面のミーティング等も行う。



- 体系的な情報を提供するウェブサイトも必要。ウェブサイト⁴⁰では、研究段階ごとの実用的な説明を提供（例えば、研究プロジェクト進行中にやるべきこととして、データのバックアップやファイルの名付け規則等について具体的に説明）。

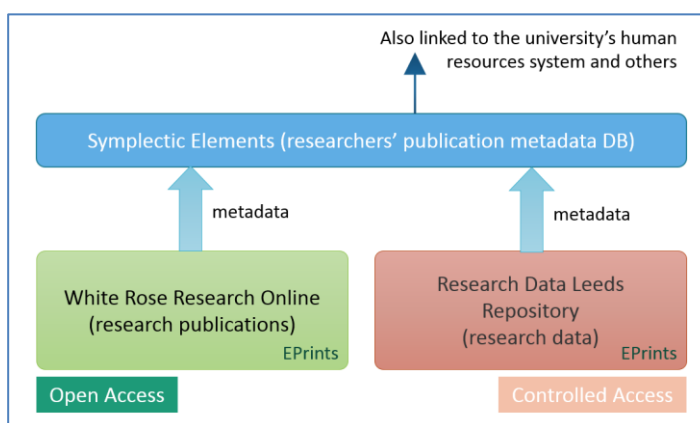
(研究者の意識向上のための工夫)

- 提供する情報は可能な限りシンプルにするようにしている。
- 研究データ管理を行うのが当然と考える環境、文化を作ることが重要。強制することはできないし、強制しても研究者は応じない。
- 図書館に研究データ計画作成を依頼してくる研究者もいるが、対応しない。研究データについて知っているのは、研究者自身のみであり、拙いものであってもまず研究者自身が計画案を作成し、図書館職員がそれをチェックして、研究者にフィードバックする必要がある。
- 図書館が相談窓口として認識され、常に連絡しやすい状況を提供することが重要。
- コミュニケーション方法の好みは研究者により異なるし、情報量によってもふさわしい方法が異なる。
- カフェなど、人目につくところで教員とミーティングするのも一つの方法。他の教員も目にする機会が増え、図書館員の活動を知ってもらえる機会になる。

2.4. University of Leeds

2.4.1. システム構成

- 教員の業績情報管理システムとして、Symplectic Elements⁴¹を使用。助成金管理システムや人事管理システムとリンクさせている。
- オープンアクセスの論文リポジトリとして、Universities of Leeds, Sheffield and York 3 大学によって開発した White Rose Research Online (WRRO)⁴²を使用。これとは別に社会学分野のプロジェクト用のリポジトリもある。いずれもソフトウェアは EPrints⁴³を使用。将来的に統合を考えている。
- 研究データリポジトリとしては、Research Data Leeds Repository⁴⁴を運用している。オープンアクセスだけでなく、アクセス制御も可能。
- DOI を付与している。データに別のバージョンが作成された場合は、別データと認識されるべきと判断されれば、別の DOI を付与する（場合によっては、正誤表を元のバージョンに付して対応することもある）。
- 大学全体の IT 担当部署から、アクティブデータ用に OneDrive Storage も提供されている（取り扱いに注意を要するデー



(図はインタビュー内容をもとに筆者が作成)

⁴⁰ The University Library Research Data Management <https://www.sheffield.ac.uk/library/rdm/index>

⁴¹ Symplectic Elements <http://symplectic.co.uk/products/elements/>

⁴² White Rose Research Online <http://eprints.whiterose.ac.uk>

⁴³ EPrints <http://www.eprints.org/uk/>

⁴⁴ Research Data Leeds Repository <https://library.leeds.ac.uk/research-data-deposit>

タの保存は不可。バックアップ用としての使用も不可)。

2.4.2. 研究データ管理サービス

(組織体制)

- The Engineering and Physical Sciences Research Council (EPSRC) が 2015 年から研究データのオープンアクセスを義務化した影響が大きく、準備が急速にすすめられた。⁴⁵
- Jisc の助成金を得て、プロジェクト RoaDMaP⁴⁶ を開始した。
- 2016 年 10 月からコアサービスとして位置づけられ、中央図書館に Research Data Management チームが統合された。Service Research Support Team が連携してサービスを提供している。
- 数年前から、従来のサブジェクト・ライブラリアン制に変えて、リエゾン・ライブラリアン制を導入している。
- 研究データ管理サービスを提供するために、図書館職員は研究プロセスについても知る必要がある。また、全学、全国、国際的な協力・連携も必要。

(研究者の意識向上のための工夫)

- 研究データ管理計画策定には DMPOnline を利用しているが、必ずしも DMPOnline の利用を積極的に勧めているわけではない（「また新しいシステムの使い方を覚えなくてはならない」などと否定的に受け取られないように配慮している）。簡単な計画は Word 等で作成し、複雑な計画は DMPOnline が便利。いずれの場合も、研究者自身が手段を選択する（助成機関の指定があればそれに従い、なければ自分で選ぶ）。
- 研究データ管理の理解を研究者に深めてもらうためには、ひとりひとり時間をかけて、対面で進める必要がある。会話をすることで理解が深まる。
- 研究者にとって、研究の継続性を確保するというメリットがあることを説明する。「研究データは持っていない」「自分のデータを見たがる人などいない」「私にしかわからないデータだ」という研究者こそ、研究データ管理の知識が必要。
- 学生と研究者を同時に対象としてワークショップやプレゼンテーションを行う（学生に向けて話しているように見せて、実はシニア研究者にも理解を深めてもらう）。
- 研究プロジェクト進行中から、メタデータ作成の情報源となる、研究者自身によるデータに対するドキュメンテーションが重要となる。研究者に対して「メタデータ」という言い方をすると誤解を招く傾向がある（「メタデータ」=図書館員がつくるものという思い込み）ので、ドキュメンテーションという言葉を使っている。

(他機関との連携)

- Universities of Leeds, Sheffield and York 3 大学で連携して行う研究データ管理トレーニングプログラム構築（コンテンツ作成）プロジェクト RDMRose⁴⁷ に取り組んでいる。
- Universities Leeds, Sheffield, York 3 大学で協力して、MANTRA よりも初級者向けのトレーニングツールを作成している。

⁴⁵ How can universities meet the expectations in the EPSRC research data policy?

<https://www.jisc.ac.uk/blog/how-can-universities-meet-the-expectations-in-the-epsrc-research-data-policy-27-mar-2015>

⁴⁶ RoaDMaP <https://library.leeds.ac.uk/roadmap-project>

⁴⁷ RDMRose <http://www.shef.ac.uk/is/research/projects/rdmrose>

2.5. University of Oxford

2.5.1. システム構成

- オープンアクセス論文リポジトリとして **ORA (Oxford Research Archive)**⁴⁸ を、オープンアクセス研究データリポジトリとして **ORA-Data**⁴⁹ を提供している。アクティブデータのストレージとしては、**Hierarchical File Server (HFS)**⁵⁰ を提供している。他にもカレッジレベルで複数のシステムが存在する。
- **ORA-Data** は施行期間として利用料を無料としているが、現在いくつかのビジネスモデルを検討しており、持続可能なコスト負担プランを導入する予定。

2.5.2. 研究データ管理サービス

(組織体制)

- この 10 年間で、英国では分野ごとの **National Data Centers** が充実し、各大学もインフラ整備の努力を行ってきた。最近の動向としては、全国的な **UK Data Archive** の構築を目指す動き。論文の裏付けとなる研究データの保存等が機関としての責務として求められるようになってきている。学術出版のあり方そのものも変わってきており、出版社も助成機関も、提供されるサービスもシステムも、常に過渡期の状態にある。
- 図書館が中心となり、研究支援担当、IT 担当、各カレッジの事務部、倫理・法務担当部署等と協力している。中でも図書館が研究者にとって最も来訪しやすいことから、出発点として適切と考えられている。複数の部署が連携しつつも、利用者の視点からはワン・ストップ・サービスとして認識されることが重要。
- 「研究データ保存に焦点を置いた従来のサブジェクト・ライブラリアンの業務」を拡大して、「研究データのライフサイクル全般を視野に入れたデータ・ライブラリアンの業務」として認識されている。
- 「データライブラリアン」にはさまざまなバックグラウンドの人がいる。キャリアとして最初から「データ・ライブラリアン」である人はまだ少ない（米国では、「データセットに関する専門的な知識を持ったライブラリアン」という狭義の意味で使われることが多いが、英国では異なる）。
- 研究データ管理サービスのコア・チームは 4-5 名（図書館職員+ IT 担当）。
- 研究データ管理サービスに関わる全学の図書館職員が **RDM Delivery Group**⁵¹ として毎月討議を行っている。
- 図書館職員と研究者に対して同時にトレーニングを行い、知識を得た図書館職員がさらに別の研究者に対しトレーニングを提供する。
- 経費は通常予算から支出しており、特別予算がついているわけではない。
- 研究データ管理経費を、大学側が研究者に求める動きも出てきている（データリポジトリの有料化）。研究者が助成金を申請する際に研究データ管理経費を含めることを認める助成機関もある⁵² が、現在は過渡期であり、大学と助成機関の間で互いに経費負担を求め合っている状態。また、研

⁴⁸ ORA: Oxford Research Archive <https://ora.ox.ac.uk/>

⁴⁹ ORA-Data <http://researchdata.ox.ac.uk/preserving-your-data/ora-data/>

⁵⁰ Hierarchical File Server (HFS) <http://help.it.ox.ac.uk/hfs/index>

⁵¹ RDM Delivery Group <http://researchdata.ox.ac.uk/rdm-delivery-group/>

⁵² Funder requirements <http://researchdata.ox.ac.uk/funder-requirements/>

究データ作成のための経費は認められても、長期保存にかかる経費に対してはまだ関心が向けられていない。

(形態)

- ウェブサイト⁵³ で包括的・体系的な情報提供を行い、電子メールや対面でのディスカッション、ワークショップなどで具体的な個人レベルの相談を行っている。
- 利用者の要望にすぐに応じられるようにすることが大切。期限直前に相談に訪れることが多い。

(研究者の意識向上のための工夫)

- 研究者に対し、研究データ管理を義務化して強要するのではなく、メリットを理解してもらい自発的に研究データ管理を行う環境を作っていく。例えば、データ共有の必要性（他者の利益）を最初から強調するよりも、データ保存によって研究者自身がメリットを得ることを強調する。

2.6. University of Cambridge

2.6.1. システム構成

- 研究者の業績成果管理システムとして Symplectic Elements⁵⁴ を使用し、ORCID と連携（研究者が選択）。Symplectic Elements のデータをもとに、公開用の研究者プロフィール VIVO⁵⁵ を構築中。
- アクティブデータのストレージとして、OneDrive（1TB）を提供。⁵⁶
- 論文・データリポジトリとして Apollo⁵⁷（DSpace で構築）を提供。DSpace ではデータは原則公開になる仕様だが、アクセス制御を細かく管理できる別システムへの移行も検討中。また、経費負担のあり方として、共通負担金の徴収、データ量に応じた課金等を検討中。
- 研究データの登録先としては、分野、全国リポジトリが優先され、無い場合に Apollo に登録することとしている。しかし、学内、国内に様々なシステムが乱立している現状が持続可能でないことは認識されている。
- 現在、Jisc により 2 つのパイロットプロジェクトが進められている。
 - Jisc R&D Project: Research data shared service⁵⁸ 英国の全国版共同リポジトリ構築
 - Jisc R&D Project: UK research data discovery⁵⁹ 英国のリポジトリポータルサイト構築

2.6.2. 研究データ管理サービス

(組織体制)

- 図書館が主導する。図書館が研究終了後の研究成果のオープンアクセスを、大学の IT 担当部署が研究プロジェクト進行中のアクティブデータの管理を担当。

⁵³ Bodleian Data Library <https://www.bodleian.ox.ac.uk/data>

⁵⁴ Elements <http://www.research-information.admin.cam.ac.uk/what-information-available/elements>

⁵⁵ Symplectic Open Profile Services VIVO <http://symplectic.co.uk/services/open-profile-services/>

⁵⁶ OneDrive – University of Cambridge <http://www.uis.cam.ac.uk/initiatives/ees/onedrive>

⁵⁷ Apollo – University of Cambridge Repository <https://www.repository.cam.ac.uk/>

⁵⁸ Jisc R&D Project: Research data shared service - Advancing research data management through collaboration <https://www.jisc.ac.uk/rd/projects/research-data-shared-service>

⁵⁹ Jisc R&D Project: UK research data discovery <https://www.jisc.ac.uk/rd/projects/uk-research-data-discovery>
UK Research Data Discovery Service – Alpha Site <http://ckan.data.alpha.jisc.ac.uk/dataset>

- 担当図書館職員のバックグラウンドも多岐に渡る（データライブラリアンとしての博士保持者、経理事務やデータ分析の経験者、ライフサイエンス分野の研究者等）。これまでの実務経験の結果、現在データライブラリアンとして勤務する。

(研究者の意識向上のための工夫)

- 部局ごとに“Data champion”（通常は研究者。大学院生や経験豊富な事務職員も含む）を任命する。
60
 - 部局の中で、研究データ管理の必要性、知識を提唱して広めてもらう。調査時点で 44 名。
 - Data champion を努める研究者にも、大学への貢献として経歴に記載できるメリットがある。
 - 研究分野独自のソフトウェア管理を行う場合もある。
- 研究者間で、研究データ管理サービスを受けた経験を宣伝してもらうと効果的。

(形態)

- 定期で行うワークショップ、ニュースレター、部局事務部への情報共有、関心を持つ研究者を対象を絞った情報提供、ウェブサイト上の FAQ の充実等。
- ウェブサイトの相談窓口情報のページに担当図書館職員の写真を掲載して、利用者が連絡しやすい雰囲気を作っている。

(その他)

- イベントごとに利用者からのフィードバックを Google form で集めて分析し、以降のイベントの質向上につなげている。

⁶⁰ Data Champions <http://www-library.ch.cam.ac.uk/data-champions>

3. 所感

デジタルコレクション管理について、特に印象的だったのは University of St Andrews の図書貸出簿電子化プロジェクトである。利用者の需要から資料価値を高めるプロジェクトを考案し、クラウドリソースの手法により着実に進行させている点に、学ぶべきことは多い。

IIIF 対応については、いずれの機関でも、まだ試験的に行っていたり対応を検討していたりする状況であることがわかった。複数の機関で言及されたように、こうした早い時期から日本の機関も参加して意見を表明していくことが重要であると考えられる。

研究データ管理サービスについては、いずれの機関でも、図書館が利用者としての研究者と近い関係を築いてきたことを活用して、サービスを提供している状況が興味深い。研究データ管理サービスがいかに必要とされているかは、従来の図書館サービス窓口と並んで、IT 相談窓口が設置されている状況に端的に現れている。

システム構築に関して気になったのは、研究データリポジトリ等の持続可能な管理のために、University of Oxford 等で経費負担の検討が行われていることである。

英国では研究データ管理に関する全国リソースが充実しているが、それに加えて、例えば University of Sheffield では自機関の環境により適合させたわかりやすいウェブサイトを利用者に提供するなど、きめ細かな取り組みが行われている。また、University of Leeds では研究者にとって研究データ管理への心理的なハードルを低くする具体的な工夫が見られ、参考にできる。

以上

平成 28 年度 国立情報学研究所実務研修 報告書
フランス国立図書館および Europeana における IIF 導入状況調査

提出日：平成 29 年 2 月 13 日

所属：京都大学附属図書館

氏名：大村 明美

日時： 2016 年 11 月 15 日

場所： フランス国立図書館

対応者：Emmanuelle Bermes

Deputy director for services and networks in charge of technical
and scientific matters, Bibliothèque nationale de France

Antoine Isaac

R&D Manager, Research & Development, Europeana

1. フランス国立図書館における IIF 導入状況

1.1. フランス国立図書館の画像データベース Gallica <http://gallica.bnf.fr/>

- 2004 年に Gallica Digital Library Charter を定め、フランス国立図書館だけでなく、300 館に及ぶ提携館との共同事業による所蔵資料の電子化を進めている。研究図書館として稀覯本・絶版本に重点を置いている。提携館とは電子化費用を折半し、技術的支援を提供している。¹
- さらに、OAI-PMH によりフランス国内の図書館のデジタルコレクションのメタデータをハーベストし、Gallica で検索可能にしている。
- 登録件数 3,654,758 件 (1,519TB) (うちメタデータハーベストのみ 217,293 件)、提携館約 300 館
- アクセス数 15,980,981 回 (2015 年)²
- 2015 年 10 月バージョンアップにより、閲覧性・検索性が向上した。

1.2. Galica の IIF 対応状況

- 現在、IIF 準拠のテストを行っている。すでに IIF manifest を作成・公開しているが、広く公表はしていない。ビューワーの IIF 準拠、アノテーション共有機能等は未対応。

¹ Gallica, shared digitization

http://www.bnf.fr/en/professionals/national_cooperation/a.gallica_shared_digitization.html

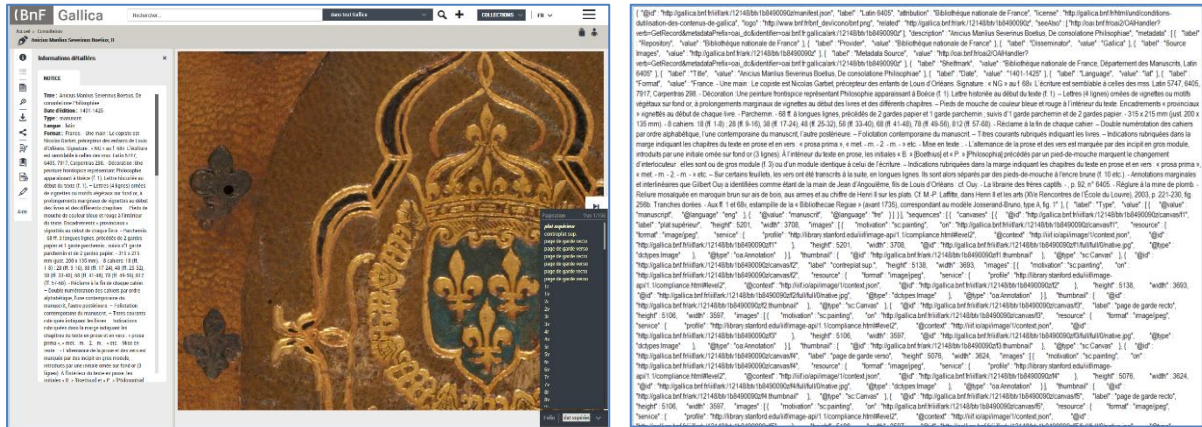
² Annual report 2015 – Key facts and figures

http://webapp.bnf.fr/rapport/pdf/rapport_2015_synthese_en.pdf

Gallica (1997–2016) De la bibliothèque de « l'honnête homme » à celle du Gallicanaute

<http://bbf.enssib.fr/consulter/bbf-2016-09-0048-005>

- 電子化の歴史が長いこともあり、ファイル形式が多岐にわたる。このため、あらかじめ画像を JPEG2000 に変換しておくのではなく、画像表示リクエストの都度、サーバ内で画像を JPEG2000 へ変換して表示している（一定期間キャッシュを保持）。
- IIIF が普及して画像の閲覧性が高まれば、利用者が画像データをサーバからダウンロードする動機を減らすことにつながるのではないかと期待している。
- フランス国内でも IIIF は（関連する研究者レベルでは知られていても、少なくとも機関レベルでは）まだよく知られておらず、様々な検討が始まったばかりの状況。
- Gallica の画面例と IIIF manifest の JSON ファイル例³



2. Europeana の IIIF 対応状況

2.1. Europeana

- ヨーロッパの図書館、博物館、文書館等が所蔵する資料の電子データへの統合的アクセスを提供（OAI-PMH によりデータ提供元からメタデータを収集し、公開）している。
- 事務局はデン・ハーグ所在（今回は IIIF 担当の Antoine Isaac 氏がフランス国立図書館内に最近設置された Europeana の拠点に滞在中、フランス国立図書館を訪問）
- 登録件数 54,253,500 件（再利用可 38.6% / 制限有 16.2% / 許諾要 45.2%）
- 44カ国の 50 アグリゲーター、3,521 機関がデータを提供
- アクセス数 4,486,261 回（2016 年）⁴
- メタデータスキーマは Europeana Data Model (EDM)、その他詳細仕様は Europeana Publishing Guide⁵（最新版は v1.4）にて公開。
- 例えば、フランス国立図書館はダブリンコア準拠のベーシックなメタデータを Europeana へ送付しているが、Europeana はそれを Europeana Data Model へ変換して使用している。

³ Gallica 画面例 <http://gallica.bnf.fr/ark:/12148/btv1b8490090z>

Gallica の IIIF manifest の JSON ファイル例

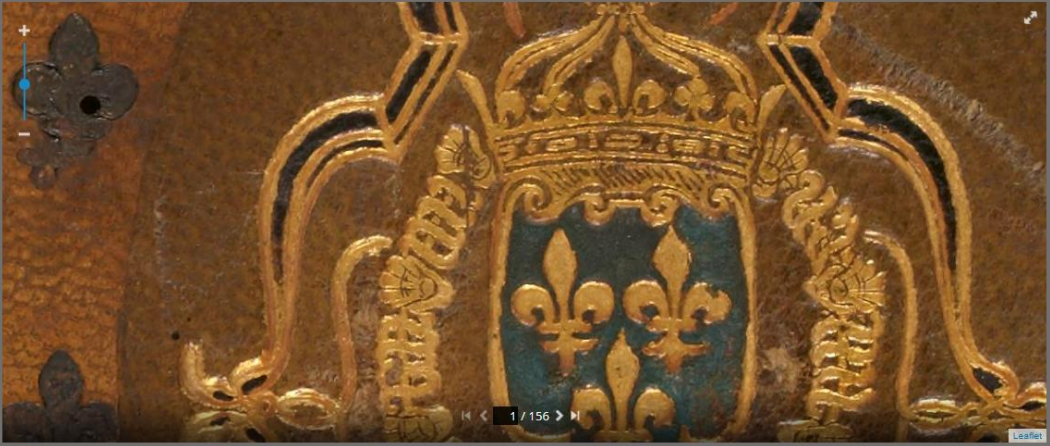
<http://gallica.bnf.fr/iiif/ark:/12148/btv1b8490090z/manifest.json>

⁴ Europeana Statistics Dashboard <http://statistics.europeana.eu/>

⁵ Europeana Publishing Guide <http://pro.europeana.eu/publication/publication-policy>

2.2. IIIF 対応状況

- Europeana は基本的にメタデータのみを持っており、データ本体を閲覧する際には参加機関のサイトへとぶ。したがって、参加機関各々で IIIF に対応する必要がある。現在のところ、約 10 機関が対応済（フランス国立図書館, Austrian National Library, University of Oxford 等）。
- Europeana 参加機関が IIIF に対応するための手順は以下のとおり。
 - 各機関が自身のサーバを IIIF に対応し、IIIF manifest を作成・設置する
 - メタデータの所定の項目に IIIF manifest の所在 URL を格納する
- Europeana 画面で見たフランス国立図書館所蔵資料詳細表示



The screenshot shows the Europeana Collections interface. At the top, there are navigation links for 'Collections', 'Explore', 'Exhibitions', and 'Blog'. A search bar is visible on the right. Below the navigation, there is a breadcrumb trail: 'Return to Home / Item'. The main content area displays a large image of a manuscript page with a golden initial 'D' and a blue background. Below the image, there is a 'Media Metadata' section with the following information:

Title Anicius Manlius Severinus Boetius , De consolatione Philosophiae. [Paris, BnF, MSS Latin 6405] | Boethius

Description Le volume, qui porte au f. 68v la signature de Nicolas Garbet, précepteur des trois fils de Louis d'Orléans , a peut-être servi de livre d'étude aux jeunes princes, à l'exemple du ms. Bnf., Latin 5747 (Salluste, Jugurtha). Les descriptions des différents exemplaires du De consolatione cités dans les inventaires de la librairie de Charles d'Orléans sont trop succinctes pour identifier, de façon définitive, notre manuscrit à l'un d'eux. Il est, en revanche, certain que le ms. Latin 6405 fit partie des collections de Jean d'Angoulême . Il est répertorié dans l'inventaire de la bibliothèque du duc fait en 1467 après son décès : « Le livre de Boece de Consolation en texte et parchemin, commandant ou second feuillet : ad superius elementum , et finissant ou derrain : Philosophie. Deo Gracias » (Arch. nat., P 1403, n° 38 et 39, d'après Ouy, p. 57, n° 29). Le second feuillet commence, en effet, par les mots : « ad superius elementum [esset...]. Le f. 68v

On the right side of the metadata section, there are links for 'FIND OUT MORE' (View at National Library Of France), 'CAN I USE IT?' (Free Re-use, Public Domain Marked), and 'SHARE' (Facebook, Twitter, Pinterest, YouTube, Tumblr). A 'FEEDBACK' button is located at the bottom right of the page.

5. 所感

フランス国内でもまだ IIIF そのものがよく知られておらず、議論が始まったばかりとのことであった。したがって、新しい機能の付加や実装等について、日本からも積極的に議論に加わって声を上げ、言語・文化特有の要件を明らかにし、IIIF の機能を深化させていくことが期待されている。

以上

平成 28 年度 国立情報学研究所実務研修 報告書
ドイツにおけるオープンアクセスの現況調査¹

提出日:平成 29 年 2 月 27 日

所属:京都大学附属図書館

氏名:大村 明美

1. University Library of Regensburg <http://www.uni-regensburg.de/library/>

1.1. 概要

- 大学 (University of Regensburg (UR) <https://www.uni-regensburg.de/index.html.en>) 設立 1962 年、学生数約 21,000 名。
- 部局図書館がなく、大学図書館がひとつの組織に集約されている。所蔵 (非電子) 資料約 400 万冊。

1.2. オープンアクセス推進の取り組み

- 国際的・全国的・地域的取り組みの主導・協力
 - OA2020 学術雑誌のオープンアクセスへの迅速・円滑な転換を目指す。
 - OA2020-DE ドイツ国内での OA の枠組み作り、移行の準備、交渉の整備等。
 - Alliance of German Research Organisations
 - Regensburger Verbundklassifikation <https://rvk.uni-regensburg.de/>
 - COAR² (オープンアクセスリポジトリ連合) に参加
- 電子リソース管理ツールの開発と他機関への提供
 - Elektronische Zeitschriftenbibliothek <http://rzblx1.uni-regensburg.de/ezeit/>
 - 契約情報管理ツール。ドイツ語圏で約 600 機関が利用中。
 - Datenbank-Infosystem (DBIS) <http://dbis.uni-regensburg.de/>
 - データベース検索ツール。約 300 機関が利用。
- オープンアクセス方針
 - 憲法に規定された学問の自由の観点から、大学当局が研究者に対して OA を義務化するポリシーは制定できないと認識。研究者本人の自主的な選択による。
- 機関リポジトリ University of Regensburg Publication Server <https://epub.uni-regensburg.de/>
 - CC-BY ライセンスの OA 論文は 10,000 件以上、相互運用性、利用のしやすさ、質などで 2015 年のドイツ語圏 OA リポジトリランキング 1 位を達成。³
 - 論文データだけでなく、研究データも同じリポジトリシステムに格納しているが、現在のところ巨大なデータセット等はまだ扱っておらず、論文のサプリメントデータ等に対応している。
 - ソフトウェアは EPrints で OpenAIRE (EU の助成研究の成果に関する OA 推進プロジェクト) の仕様に準拠。

¹ 本報告書は、JUSTICE より調査に参加された森嶋桃子氏 (慶應義塾大学メディアセンター本部) と砂押久雄氏 (東京工業大学研究推進部情報図書館課) と作成した JUSTICE への報告資料に基づく。

² COAR <https://www.coar-repositories.org/>

³ Open Access Repository Ranking <http://repositoryranking.org/>

Open Access Repository Ranking: Team <http://repositoryranking.org/team/>

- BASE (Bielefeld Academic Search Engine: 約 5,000 の OA 学術ウェブリソースを対象とした検索エンジン)、Google Scholar、RePEc (Research Papers in Economics: 経済学分野の論文アーカイブ) のインデックス対象となっている。
- 研究データのライフサイクルを考慮した管理計画の策定、データの頒布、プライバシーなど課題は多い。
- 部局や研究者個人のウェブサイト内に、API 経由で取得したリポジトリのデータを埋め込み表示することが可能 (研究成果可視化のツールとして、リポジトリが利用されている)。
- OA 出版のプラットフォームとしての役割も果たす。
 - 図書館の例 HEIMATFORSCHUNG (郷土史) <http://www.heimatforschung-regensburg.de/>
 - 雑誌の例 COPAS (Current Objectives of Postgraduate American Studies) <http://copas.uni-regensburg.de/index>
- ORCID との連携を進めている (UR が ORCID に機関として参加し、研究者に ORCID の利用を推奨する。研究者の承認が得られた場合、UR 側のシステムが ORCID の持つ研究者情報を取得できるようになる)。
- 研究データ管理支援
 - トレーニングに MANTRA を利用。
 - 研究データは分野によって研究者の反応に隔たりがある。
 - アウトリーチの取り組み (研究者を招いたランチ、トレーニングセッション、広報等)
- APC 管理
 - 研究者用ウェブツール (ログインには ID/PW 要) <http://oa.uni-regensburg.de/>
 - 研究者がメタデータを入力することにより、APC が還付される。
 - 還付の上限は 2,000 ユーロ (DFG⁴ 助成)、責任著者であること、質が保証されていること、フル OA 誌であることが条件。DFG の助成条件がフル OA であることもあり、ハイブリッドのデータを網羅的には収集していない。
 - INTACT⁵ による APC のビジュアル化 (OpenAPC) <https://epub.uni-regensburg.de/openapc.html>
- SCOAP3 への参加
- 雑誌のライセンス
 - Alliance Licences や各種コンソーシアムに参加。
 - DEAL
 - APC 込みの OA オフセットモデルについて 6 ヶ月間 Elsevier と協議したが頓挫し、60 大学 (UR を含む) が 2017 年よりキャンセル。複数年契約の機関 (MPDL 含む) はまだ契約期間が残っているため、購読可能。
 - 学術流通の問題 (図書館のニュース) が新聞の一面になるなんてめったにないことで、国民の注意を喚起した意義がある。現状では既存の出版社の権威やインフラに頼らなければいけない部分はあるが、国レベルでの取り組みを続けることが重要。
 - 現時点では研究者からネガティブな反応はない。しかし、今後購読停止期間が長くなり読めなくなった論文数が増えた際の反応が心配。この問題に対して研究者の理解を得ることが重要。

⁴ DFG: Deutsche Forschungsgemeinschaft (German Research Foundation; ドイツ研究振興協会) <http://www.dfg.de/en/>

⁵ INTACT <https://www.intact-project.org/>

2. Bayerische Staatsbibliothek (BSB) <https://www.bsb-muenchen.de/en/>

2.1. 概要

- 州立図書館、研究図書館、バイエルン地域の納本図書館としての役割を持つ。
- 蔵書数 1,000 万冊以上。約 20,000 のインキュナブラ等貴重書も多く所蔵し、電子化に積極的に取り組む。
- 大学、専門大学、研究機関、美術館といった枠組みを超えた立場で、研究者の需要を把握することに努めている。

2.2. オープンアクセス推進の取り組み

- ドイツにおける助成の特徴
 - DFG は連邦・州政府から予算を得て、全国の学術機関に対して、(州ごとの予算配分比率が決まっているわけではなく)プロジェクト評価に基づいて助成金を配分している。学術に関する現状に精通した専門家からの勧告等を受けており、戦略的な投資を行っている。
 - 長期プロジェクトに対しては初期立ち上げに投資するが、その後の継続的な経費については基本的に機関の自己負担が求められる。
- コンソーシアムによるライセンス
 - Bayern-Konsortium www.bayern-konsortium.de
 - バイエルン州のコンソーシアム。BSB が事務局を担当し、ミュンヘン工科大学やミュンヘン大学含む 11 の総合大学、19 の専門大学、10 の地域研究図書館 (BSB の分館) が加盟。
 - GASCO <https://www.hbz-nrw.de/produkte/digitale-inhalte/gasco>
 - 2000 年に始まった、ドイツ語圏の国を横断したワーキンググループ (ドイツ、オーストリア、スイス、ルクセンブルク)。
 - ナショナルライセンス <http://www.nationallizenzen.de/angebote>
 - 約 140 のバックファイルコンテンツを買切。維持費なし、恒久アクセス権、ホスティング権利あり。約 700 の大学で登録して利用中。
 - Priority Initiative "Digital Information" <http://www.allianzinitiative.de/en/>
 - カレントコンテンツへのオプトインモデル。一定のエンバーゴ期間の後で OA 化 ("Moving wall" 方式)。Springer、Wiley、Elsevier は対象外。
 - コミュニティベースのライセンス
 - Scientific Information Services (Fachinformationsdienste für die Wissenschaft :FID)
 - Special Subject Collections System (SSG)
 - DEAL
- OA のインフラ構築
 - DeepGreen <https://deepgreen.kobv.de/en/project-deepgreen/>
 - DFG の助成により、エンバーゴの後、機関／分野別リポジトリにおいて出版社稿の無料公開をめざす 2 年間のプロジェクト。システムティックに出版社からデータを集め、自動的にレポジトリに登録する試み。
 - FID に基づき、分野ごとのリポジトリを維持している。ソフトウェアは Dspace。2017 年中に研究データを格納するセクションを構築予定。同じく FID に基づき、ゴールド OA 誌をホスティングしている。ソフトウェアは OJS を利用。⁶

⁶ 例: reviews in history <http://www.recensio.net/front-page>

- 研究データに関する取り組み
 - BSB は所属研究者をもたないため、大学が主導。
 - 国レベルでのリサーチデータのインフラが必要という提言がでている (Rat für Informationsinfrastrukturen (Council for Information Infrastructures)⁷; 現在英語に翻訳中)。
 - Zentrum für Elektronisches Publizieren (Center for Electronic Publishing (ZEP)) <https://www.bsb-muenchen.de/en/zep/>
 - PKP LOCKSS <https://pkp.sfu.ca/pkp-lockss/>
 - 研究データ関連のネットワーク
 - CESSDA Consortium of European Social Science Data Archives <https://cessda.net/>
 - RDA Research Data Alliance <https://www.rd-alliance.org/groups>
 - DINI / NESTOR Working Group on Research Data http://www.forschungsdaten.org/index.php/AG_Forschungsdaten
 - ドイツ国内のコミュニティ <http://www.forschungsdaten.org/index.php/GelbeSeitenProtokolle>

3. Max Planck Digital Library (MPDL) <https://www.mpdl.mpg.de/en/>

3.1. 概要

- マックス・プランク協会 (Max-Planck-Gesellschaft zur Förderung der Wissenschaften; 英語表記は Max Planck Society for the Advancement of Science) は、ドイツにおける科学研究を推進・振興するための機関。自然諸科学から人文・社会諸科学に至る全学問分野についてドイツ各地に 80 以上の研究所を擁する。予算の大半は連邦政府および諸州政府によってまかなわれている。280 名の Director、13,000 名の研究者、10,000 名のスタッフが所属。管理本部、IT 本部、情報本部 (MPDL) の 3 つの中央組織を持つ。
- MPDL が情報と関連サービスを提供 (物理的な中央図書館は存在しない)。MPDL の中央予算で電子ジャーナル需要の 90% をカバーしており、残りは各研究所個別の予算で対応。
- 15,000 タイトルの電子ジャーナル、650,000 タイトルの電子ブック、150 のデータベースを契約。10 のレポジトリシステム (出版者、プロジェクト交流、研究データ等。MPG の機関レポジトリである MPG.PuRe を含む) を擁する。

3.2. オープンアクセス推進の取り組み

- OA2020⁸
 - 現在 OA は理念としては理解が広まっているが、実際的な活動としてはまだ十分に行われていない。しかし実際には、出版市場にはオープンアクセスモデルへ移行する資金が十分に存在する。
 - これまで「研究者が OA に積極的になるべき」という考え方で OA への転換が進められてきたが、結局研究者の判断に依存することになるため、なかなか進捗してこなかった。OA2020 では、研究者を変えようとするのではなく、出版システムのビジネスモデル自体を転換することにエネルギーを向けている。

⁷ Empfehlungen zum Management von Forschungsdaten

<https://www.bmbf.de/de/empfehlungen-zum-management-von-forschungsdaten-3036.html>

⁸ Open Access 2020 <https://oa2020.org/>

- MPDL としては全大学を説得することは考えていない。出版市場を健全化するのが目的であって、小規模な大学や(例えば出版物が国内を中心に流通している等の)人文系分野の研究者が現状で問題がないのであれば変える必要はない(必要性のある機関が加わればよい)。
- OA が進むことにより出版社の利益率が下がることは出版社も以前から分かっており、経営判断を行っている。例えば Elsevier はすでにそれに備えて、コンテンツ層以外に、Pure 等、ターゲットを絞ったミドルウェア製品等の開発に投資を始めている。Springer はもともと他社と比較して利益率が低いこともあって OA モデルへ移行しても大きく利益率を下げることにならないこと、また真っ先に移行すれば”研究コミュニティの味方”としての評判を確立できることもあり、OA モデルに積極的に対応している。
- 出版社が出す情報をそのまま受け取るだけでは対等な交渉は難しい。新しいモデルを提案するためには、自らデータを分析して根拠を提示していくことが必要。データ分析は MPDL の強みでもあるが、毎年データ分析の度に新しい驚きがある。
- APC 管理
 - MPDL 所属の著者については、約 2,200 タイトル以上を対象に、MPDL で一括して APC を管理・支払を行っている。
 - 一括処理により著者の負担が減り、結果的に OA が進展する。著者が論文提出時にフォームのドロップダウンメニューで MPS を選ぶ出版社(Wiley、IOP 等)が多い。オフセットの場合は IP レンジ内からのアクセスや MPS のメールアドレスがキーとなる。
 - MP として統一的な論文提出システムを構築する案もあるが、各研究所の自治性を考慮するとまだその時ではない。論文提出を把握することは、論文がリジェクトされた場合もその情報を把握することにつながり、研究者としてはセンシティブな問題でもある。
- MPDL の出版物レポジトリ MPG.PuRe <http://pubman.mpdl.mpg.de/pubman/faces/HomePage.jsp>
 - 登録件数 285,360 件のうち、30%がフルテキスト、そのうち 15%が OA (他機関と比較すると、登録データ数は多いが、OA の割合は低め。しかし絶対値としては多い)。登録は義務ではなく、推奨。
 - システム開発は PubMan/eSciDoc を使用して 2006 年に内部で行った。MPG 以外でも利用されている(例: 物質・材料研究機構(NIMS))。ハードウェアはゲッティンゲンのコンピュータセンターでホスティングしている。
 - Horizon2020 の助成を受けた研究プロジェクトの出版物は OA にし、かつ OpenAIRE⁹に登録する義務がある。OA に対応するよう研究者を説得する上で、こういった助成機関の姿勢等の外部からの圧力や明確なポリシーが効果的。
 - データの長期保存を担保するためには、保存する際のデータフォーマットを慎重に検討したり、データ閲覧のためのソフトウェアや環境も合わせて保存したりする等の対応が必要だが、そのためのポリシーやガイドラインは未策定。現時点では、データが失われる前にとにかく収集するところに重点を置いている。なお、Bielefeld University ではデータの長期保存担保の要件確認のために、Data Seal of Approval¹⁰ を参考にしている。

⁹ OpenAIRE <https://www.openaire.eu/>

¹⁰ Data Seal of Approval <http://www.datasealofapproval.org/en/>

4. 所感

今回の訪問調査では、大学図書館として研究者に貢献する Regensburg University Library、州立図書館、研究図書館、納本図書館として多層的なサービスを提供する BSB、ドイツ各州に研究所を持つ Max-Planck-Gesellschaft に統合的な情報サービスを行っている Max Planck Digital Library を訪問した。

それぞれ異なる特徴を持つこれらの機関で担当者から直接話を聞き、意見交換をする中で、ドイツでは日本と比べて学術システムが分散的な傾向にある一方で、各機関の積極的な努力により、地域や分野ごとのコンソーシアムや共同プロジェクト等が行われていることがわかった。また、APC 管理等、オープンアクセスをめぐる様々な課題に対して、早い時期から現実的な取り組みを行ってきたこと、ドイツ語圏独自のリポジトリ評価ランキングが作成されていることなどが、非常に興味深い。機関の姿勢として特に印象に残ったのは、「出版社と対等に交渉するために、自らデータを分析して裏付けをした提案をする必要がある」、という指摘である。こうした姿勢から学ぶところは大きい。

以上

平成 28 年度 国立情報学研究所実務研修 報告書
12th International Digital Curation Conference (IDCC17) 参加報告

提出日:平成 29 年 2 月 28 日

所属:京都大学附属図書館

氏名:大村 明美

1. 開催概要

12th International Digital Curation Conference (IDCC17)

開催日・開催地: 2017 年 2 月 20 日-23 日; エディンバラ(英国)

参加者: 約 260 名

ウェブサイト: <http://www.dcc.ac.uk/events/idcc17>

プログラム: <http://www.dcc.ac.uk/events/idcc17/programme>

2. 開催趣旨

IDCC は、デジタルキュレーションの専門家、教育関係者、データ作成者、利用者等の立場、多分野の観点から、デジタルキュレーションについて意見を交わす国際会議である。2017 年の会議は、デジタルキュレーションを機関におけるワークフローの中に組み込み、FAIR 原則に準拠したデータ管理を実現し、柔軟なインフラを構築することに焦点を置く。

3. 基調講演概要

3.1. "Rich Information Hides in Missing Data"

Dr Maria Wolters, Reader in Design Informatics, University of Edinburgh

健康管理分野の研究者としての立場から、保存対象から漏れがちなデータがいかに重要な情報を含んでおり、データセット本体の質の向上に有用であるかということが、EU の助成を受けた研究プロジェクト Help4Mood¹ や ForgetIT プロジェクト² 等を例に説明された。

3.2. "Data-Driven Museums" Alice Daish, Data Scientist, British Museum

大規模博物館の運営を例に、データ収集・分析を問題解決へ結び付けるために、別々の場所に隔離・分散されたデータをアクセス可能にし (de-siloed and accessible) 意思決定者の手に渡すための様々な取り組みが紹介された。

3.3. "Towards Automating the Data Analytics Process"

Professor Chris Williams, School of Informatics, University of Edinburgh

英国の Alan Turing Institute におけるデータ分析プロセス自動化プロジェクトの進捗状況紹介。データマイニングに必要な労力の 80%がデータの前処理 (data wrangling) に費やされているが、それを改善し省力化するための対話型ツールを開発する試み。

4. 発表

4.1. Curation Communities (Practice papers)

¹ Help4Mood <http://help4mood.info/site/default.aspx>

² ForgetIT <http://www.forgetit-project.eu/en/start/>

4.1.1. “Is Democracy the Right System? Collaborative Approach to Building an Engaged RDM Community”

Marta Teperek

University of Cambridge における研究データ管理は、当初少人数のチームで研究データ管理ポリシーについて説明を行うことで教員の意識向上を図っていたが、なかなか効果が見られなかった。そのため、より「民主的な」手法に転換し、利用者の需要を積極的に反映することを意図した Research Data Management Project や Data Champions (4.1.2. で詳述) の取り組みを開始した。

Research Data Management Project では、ボランティアのサービス提供メンバーを学内で広く募集し、図書館職員だけでなく研究に携わる様々な部署からの参加を得た。メンバーは 4 つのワーキンググループを結成し、2 ヶ月ごとに開催する全体ミーティングで活動状況を報告することとした。このプロジェクトの結果として、研究者のデータ共有行動に変化が見られた。以前は研究終了後にかけてこみでようやく共有することが常態であったが、現在では早い段階でデータ共有を行う(メタデータも研究者側から提供)ことが浸透してきた。「民主的な」手法により、効果的なサービス提供に不可欠な関係者コミュニティ、信頼関係が構築されたと言える。

4.1.2. “Creating a Community of Data Champions” Rosie Higman

University of Cambridge では、各部署の研究者、大学院生、職員らを Data Champion に任命し、研究データ管理サービスを提供する実践コミュニティを構築した。こうした取り組みには多くの時間を要するが、分野独自の研究データ管理知識を蓄積できる、全学で知識の共有ができる、研究データ管理の意識向上に取り組む場を増やすことができる、といった効果が得られた。当初予定していなかった活動に範囲が広がる等の事象もあるため、今後長期的な評価を行う必要がある。

4.1.3. “Data Curation for Community Science: CHIME Pilot Study” Andrea Copeland

Indiana University と Purdue University による、市民サイクリストが収集する動画データを使ったコミュニティ・データアーカイブの構築プロジェクト。データをキュレーションするには膨大な労力がかかるため、データ収集時から再利用やキュレーションに耐える質を確保するための基準を設定し、データ収集の担い手であるサイクリストにトレーニングを行った。

4.1.4. “When Scientists Become Social Scientists: How Citizen Science Projects Learn About Volunteers”

Peter Darch

研究にボランティアを巻き込んでデータキュレーションを進める試みとして、Galaxy Zoo³プロジェクトが紹介された。研究チームのボランティアに対する理解を深める重要性と、そのための様々な手法が示された。

4.2. Sensitive Data

4.2.1. “Introducing Safe Access to Sensitive Data at the University of Bristol” Debra Hiom

University of Bristol における、取り扱いに注意を要するデータの安全管理に関する紹介。2014 年 7 月から運用開始した研究データ用のリポジトリ data.bris⁴ は原則オープンアクセスとしてきたが、アクセス管理機能を求める声もあったため、2014 年 11 月から Data Access Task & Finish Group による検討が行われた。その結果、4 つのデータアクセスレベル (Open / Restricted / Controlled / Closed) を採用することとし、具体的な対応策を明確にし、トレーニングやガイダンスを行った。プロジェクトを進める中で、学内で合意を得るためには時間を要することを認識し、他部署との関係構築に努める必要があることがわかった。

4.2.2. “Orchestrating Cloud Infrastructure to Manage Sensitive Data” Sebastian Karcher

³ Galaxy Zoo <https://www.galaxyzoo.org/>

⁴ data.bris <https://data.bris.ac.uk/data/>

取り扱いに注意を要するデータの管理にクラウド資源の活用を試みる、QDR⁵ による取り組み。クラウドサービスのセキュリティ評価基準として CAIQ⁶ Self-Assessment が紹介された。

4.2.3. “Developing on Ethical Framework for Curating Social Media Data” Elizabeth Hull

ソーシャルメディアデータのキュレーションに際しての倫理的取り扱い基準を検討するプロジェクト。STEP (sensitive, transparent, expectation, platform) フレームワークとその原則(データ共有の価値の分析、責任の所在、継続的な検討の必要性)が提示された。

4.3. Workflows & Data Quality

4.3.1. “Encouraging and Facilitating Laboratory Scientists to Curate at Source” Cerys Willoughby

質の高いメタデータを研究者から得るために、データ入力インターフェースを変えることによって得られる効果について検討。具体的には、入力のハードルを下げるためにテンプレートを提供すること、適切な質問設定をすること、回答内容が抑制されないように初期値を入れないこと、研究データ管理計画策定により早い段階でのキュレーションを奨励すること、研究データ管理教育と支援の提供等の対策が挙げられた。

4.3.2. “Connecting data publication to the research workflow: a preliminary analysis” Angus Whyte

RDA-WDS Data Publishing Workflows group⁷ による、研究者のデータ公開プロセスの精査を通して実現する、オープンデータを推進するために効果的な支援サービスや基準の検討。

4.3.3. “Frictionless Data: Making Research Data Quality Visible” Jo Barratt

Open Knowledge International による研究データの質を見える化する試み、Frictionless Data⁸ の紹介。パイロットプログラムとして、DM4T (Data Management for TEDDINE)⁹ や PNNL (Active Data Biology at Pacific Northwest National Laboratory)¹⁰ が紹介された。

4.4. Curriculum & Training

4.4.1. “Choose your own Research Data Management Guidance” Alex Boll

GW4 Research Data Services Group¹¹ により開発された、研究データ管理トリアージツールの紹介。研究データ管理計画についてのガイダンスをゲームのスタイルで提供する。トリアージツールはシンプルな質問にシンプルな回答がすぐに得られるもので、既存のツールを補完するものとして有用と考えられる。

4.4.2. “Library Carpentry: building software skills for librarianship” Jeremy Cope

図書館職員は目録や電子リソース、統計といった多くのデータを扱い、研究者にサービスを提供する上で必要な様々なスキルを持っているが、プログラミングレベルの技術的知識は IT スタッフに頼る場面が多い。Library Carpentry プロジェクトでは、図書館職員向けのソフトウェア関連技術のトレーニングを構築することを目的とする。2015 年 11 月、City University of London においてパイロット版を初めて行い、これまで

⁵ Qualitative Data Repository <https://qdr.syr.edu/>

⁶ The Consensus Assessments Initiative Questionnaire (CAIQ)
<https://cloudsecurityalliance.org/star/self-assessment/>

⁷ RDA/WDS Publishing Data Workflows WG
<https://www.rd-alliance.org/groups/rdawds-publishing-data-workflows-wg.html>

⁸ Frictionless Data <http://frictionlessdata.io/>

⁹ DM4T <https://github.com/frictionlessdata/pilot-dm4t>

¹⁰ PNNL <https://github.com/frictionlessdata/pilot-pnnl>

¹¹ GW4 <http://gw4.ac.uk/>

に4大陸8カ国においてワークショップを19回行い、300人以上の図書館職員が参加した。開発した教材はウェブサイト¹²でも公開されている。スケジュールの調整、講師の手配等が課題となっている。

4.4.3. “Researcher training in spreadsheet curation” Gene Melzac

研究データとして使われることが多いスプレッドシートデータのキュレーションに関するトレーニングプログラム。既存のプログラムと提供形態を変え、ワークショップ形式等で十分な時間を割り振って実施している。

4.4.4. “Evaluating the Effectiveness of Data Management Training: DataONE’s Survey Instrument”

Chung-Yi Hou

DataONE¹³が開発した、研究データ管理を学ぶための教材の効果を評価して継続的に向上させるツール Education Evaluation instrument (EEVA) の紹介。

4.5. FAIR

4.5.1. “Are the FAIR Data Principles Fair?” Alastair Dunning & Jasmin Bohmer

- FAIR データ原則を様々な観点から評価し、現状では遵守されている割合がそれほど高くないこと、原則そのものの尺度に客観性がないものもあること、詳細度に差があること、分野によっては適合しにくいこと等が指摘された。

4.5.2. “Assessing the FAIRness of Datasets in Trustworthy Digital Repositories: A Proposal” Peter Doorn

オランダの Data Archiving and Networked Services (DANS)¹⁴ によるリポジトリのガイドライン策定の取り組み。2006年に5つの基本原則(インターネット上で発見可能であること、アクセス可能であること、利用できる状態であること、信頼性のあるデータであること、識別子が付与されていること)が策定され、ここから Data Seal of Approval¹⁵ の16ガイドラインが生まれた。2014年のFAIR データ原則との相互補完性についても言及された。

5. ポスター発表

5.1. “Integrating Research Data into Digital Preservation Workflows for Libraries & Archives: Mapping the Landscape” Amber Leahey and Grant Hurley, Scholars Portal, Ontario Council of University Libraries

カナダの Ontario Council of University Libraries (OCUL)による、研究データ管理サービスを既存のデータ保存サービスのワークフローに融合して、一貫したサービスを提供する取り組み。ワークフローとそれを支援するサービスがわかりやすく図示されている。

5.2. “Parallel Auditing of University of Oxford and Cambridge’s Institutional Repositories” Edith Halvarsson & Sarah Mason, Bodleian Libraries, Oxford – Lee Pretlove & Somaya Langley, Cambridge University Library

University of Oxford と University of Cambridge による、機関リポジトリ ORA/ORA-data、Apollo を同時に対象としたシステム適正化プロジェクト。リポジトリに関わる職員への聞き取り調査や処理手順の精査等を行うことにより、次世代の機関リポジトリやデータ長期保存のあり方を検討する。

6. ワークショップ

6.1. “A postcard from the future: tools and services from a perfect DMP world”

¹² Library Carpentry <https://librarycarpentry.github.io/>

¹³ DataONE www.dataone.org

¹⁴ DANS <https://dans.knaw.nl/nl>

¹⁵ Data Seal of Approval <http://www.datasealofapproval.org/en/>

研究データ管理計画を実際に役に立つ、FAIR データ原則に沿ったものにし、機械可読にするために、どのような研究データ管理計画策定支援ツールが理想的であるかが検討された。また、他のツールやサービスと統合することで既存の研究データ管理フローに組み込むことも検討された。

6.2. “OpenAIRE services and tools for Open Research Data in H2020”

H2020 Open Research Data initiative で求められている要件が示され、OpenAIRE が提供している支援ツールやサービスが紹介された。また、H2020 プロジェクトにおける FAIR データ原則の意義、データリポジトリ選択基準、DMPOnline で利用できる OpenAIRE のテンプレートとガイダンス、研究データと他のデータとのリンク等についても触れられた。

7. 所感

本会議に出席して、大学図書館における研究データ管理サービス、プライバシーや研究倫理が絡むデータの取り扱い、データの長期保存等、新しい需要に応じて生まれたサービスでありながら、既存のワークフローと融合して統合的なサービスとして提供していく取り組みを知ることができた。機関リポジトリと商業出版社がせめぎ合う論文のオープンアクセスとは異なり、研究データに関して機関リポジトリがオープンアクセス化に果たす役割が絶対的に大きいことを実感できた。

出席者の多くは図書館職員で、今回始めて出席した人が大半を占めたが、研究者も参加しており、休憩中に話をした参加者の中には、例えば、以前英国で活動していたが現在マレーシアの大学で大学院生の指導にあたる研究者がおり、研究データ共有の環境を整える必要性が大学当局から理解を得られない、と嘆く声も聞いた。こうした国際会議の場で各国の参加者から直接話を聞くことで、それぞれに異なる事情と解決しようとする取り組みがわかり、興味深かった。

以上