平成 29 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）における研究データ整備に関する長期研修

提出日：平成 30 年 2 月 23 日
所属：東京大学附属図書館
氏名：横井慶子

1. 期間

平成 29 年 10 月 2 日～平成 29 年 12 月 26 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、日本人研究者の業績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、業績データ管理に関する先導的な知識と技術を習得する。さらに、研修期間中に欧州の他の教育研究機関、図書館の訪問や国際会議参加を通して、欧州におけるオープンアクセス、オープンデータへの取り組みに関する情報を収集する。

4. 具体的な内容

※別添報告書 1 を参照

(1) INSPIRE 日常業務

論文レコード HEP の収録情報（著者名、著者所属機関、共著者名、分野、実験等）のメタデータの自動分析および、著者の表記の揺れを吸収して、著者ごとの論文プロファイル Author Profile が生成される。このため HEP の正確さが、Author Profile が機能することにつながる。Author Profile は最終的に著者レコード HepNames にリンクせねばならず HepNames を正確かつ可能な限り最新状態にしておく必要がある。

- 著者情報（HepNames、Author Profile 等）の充実
- 論文データと HepNames の紐付け
- 著者レコード HepNames の更新・新規作成および Author Profile とのリンク形成

(2) ORCID 普及支援

- INSPIRE と ORCID の著者データの連携機能を説明するウェブページの翻訳
- INSPIRE と ORCID 間で論文データをインポート、エクスポートする手順書作成

(3) 会議等

- CERN/NII 遠隔会議(11/9,12/4)

(4) その他

- Topical group gathering 聴講(11/1)
- CERN 実験施設見学（ATLAS,LMF,CMS,CCC,CM14）(12/1)

5. 出張

※別添報告書 2-1,2,3 を参照

(1) ICOLC 参加（10/15-18）

プラハで開催された ICOLC2017 に参加し、オープンアクセス環境下における雑誌契約の国際的な現状に関して情報収集を行った。

(2) ドイツ国立科学技術図書館（TIB）訪問調査（12/19）

RADAR(Research Data Repository)および電子ジャーナル契約状況について情報収集を行った。

(3) Max Planck Digital Library (MPDL)訪問調査(12/21)

研究者の論文投稿状況の把握、管理手法およびオフセット契約締結における注意点などについて情報収集を行った。

6. 所感

INSPIRE の質向上プロジェクトでは、データ整備の効率化について考えさせられた。INSPIRE のデータを整備する上で必要となる根拠データを得るには、KAKEN や所属大学の公開する研究者情報、研究者個人のウェブサイトなど複数の情報源にあたる必要があり、1 件あたりの処理にかなりの時間を要した。CERN としては、将来的には ORCID と INSPIRE と連携させることで INSPIRE の質向上につなげたいとのことだった。この方策は今後も予想される論文数の増加や、共著者数の多い高エネルギー物理学分野において、極めて有効で効率的だと感じた。日本国内の機関リポジトリや研究者情報データベースでも、こういった方策をうまく取り入れ、研究者と図書館員や事務職員双方の負担を軽減しつつ、正確さと網羅性の向上につながればよいと感じた。

その他、上述の実作業以外の場面でも、貴重な経験や印象深く感じたことが主に 3 点あった。

1 つ目は、同僚とのディスカッションである。CERN スタッフが自発的にトピックを用意してプレゼンし、同僚とディスカッションする Topical group gathering に参加する機会を得た。自分の参加回のトピックは、SCOAP[3]始動前後における、対象誌の雑誌ウェブサイトおよび arXiv からのプレプリントのダウンロード数の変化だった。ディスカッションでは、目の前の現象に対して、参加者が自由に意見を出し合い、分析や解釈を試みた。「正解」のない状況で、参加者全員で思考する行為を通して、視野が広がって深い考察ができ、とても刺激的な経験となった。

2 つ目は、専門性と雇用関係である。MPDL 訪問時に対応された 3 名が、自己紹介で「MPDL 歴は〇年（いずれも 5 年以下）」と名乗られ、うち 2 名は有期雇用、無期雇用の方（契約担当）はもともとは航空関係の民間会社勤務とのことだった。また論文調査を行う Data Analysis team に求められるスキルは、PostgreSQL と PYTHON を用いたデータ分析能力のみで、図書館員ではなかった。「図書館員」という枠組みを超え、雇用する側と雇用される側の双方が、専門性の需給の一致するタイミングで雇用関係を結ぶというあり方がとても興味深かった。

3 つ目は、実験系物理学者の研究実態である。ATLAS の責任者である東京大学の教員と懇談する機会を数回得られた。実験に関わる全容をご説明いただき、国際的な研究グループの研究実態を垣間見ることができた。具体的には、各国の研究助成を受けているがゆえに、研究結果の公開に強

い義務感を感じていること、滞りなく実験を遂行することへの大きな責任感とプレッシャー、異なる考えをもつ外国人研究者との共同研究を進める上での配慮などである。間接的とはいえ、大学職員が支える教員の研究実態を知ることができ、とても参考になった。

いずれも日本国内の日常業務ではなかなか体験できないことで、このような貴重な経験ができたことは大変有意義であった。日本と欧州とでは大学・研究図書館の置かれている環境や下地が異なるため、今回学んだこと全てをそのまま日本での業務に反映させることは難しいと思う。しかし、そのまま反映できない部分も、大学図書館の課題を検討する上での参考となるものであり、日本の事情に合わせての改善、解決につなげられるよう活用したい。また、今回得られた人的つながりも大切にして、今後の大学図書館業務に貢献できるようにしたい。

以上

詳細報告 1

平成 29 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の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 30 年 2 月 23 日
所属：東京大学附属図書館
氏名：横井慶子

1. 欧州原子核研究機構（CERN）1*

下記数字は 2017/1/12 時点のもの

- 1954 年発足。世界最大規模の素粒子物理学の研究機関。英語名称は European Organization for Nuclear Research、フランス語名称は Conseil Européen pour la Recherche Nucléaire
- スイスのジュネーブ郊外とフランスの国境地帯に所在
- 加盟国：ヨーロッパ 22 カ国　※日本はオブザーバー国
- 年間予算：約 1,200 億円
- 職員：約 2,500 名
- ユーザー：12,816 名（世界 77 カ国）
- ユーザーの 60%は加盟国メンバー、日本は 294 名（2.3%）
- 全周 26.6km の大型加速器 LHC(Large Hadron Collider)ほか複数の加速器が地下 100m 付近に設置。ATLAS,CMS はじめ複数の検出器があり世界各国のユーザーが実験を行っている。

2. 高エネルギー物理分野の文献情報データベース

INSPIRE http://inspirehep.net

2.1. 概要

- 高エネルギー物理学分野の文献データベース
- 高エネルギー物理学分野の研究機関 CERN, DESY（Deutsches Elektronen-Synchrotron:ドイツ）, Fermilab（アメリカ）, IHEP（Institute of High Energy Physics Chinese Academy of Science：中国）, and SLAC（SLAC NATIONAL ACCELERATOR LABORATORY ：アメリカ）による国際共同運営
- 近々（1 年以内）にリプレイスの予定。リプレイス後はユーザーフレンドリーなインターフェースとなる。

[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 Japanese and global High-Energy Physics community; 2015 年 9 月締結にもとづき大学図書館職員を派遣するものである。

2.2 データベース構造

- HEP（論文）、HepNames（著者）、Institutions（機関）、Conferences（会議）、Experiments（実験）、Journals（雑誌）等のレコードがあり、それぞれの間でリンクが形成されている。HEP が自動分析され、著者ごとの業績プロファイル Author Profile が生成される。

3. INSPIRE による著者情報管理

3.1 Hep

論文レベルのメタデータ。管理ツール Bibedit にて、規定のフィールドを編集。

参考)

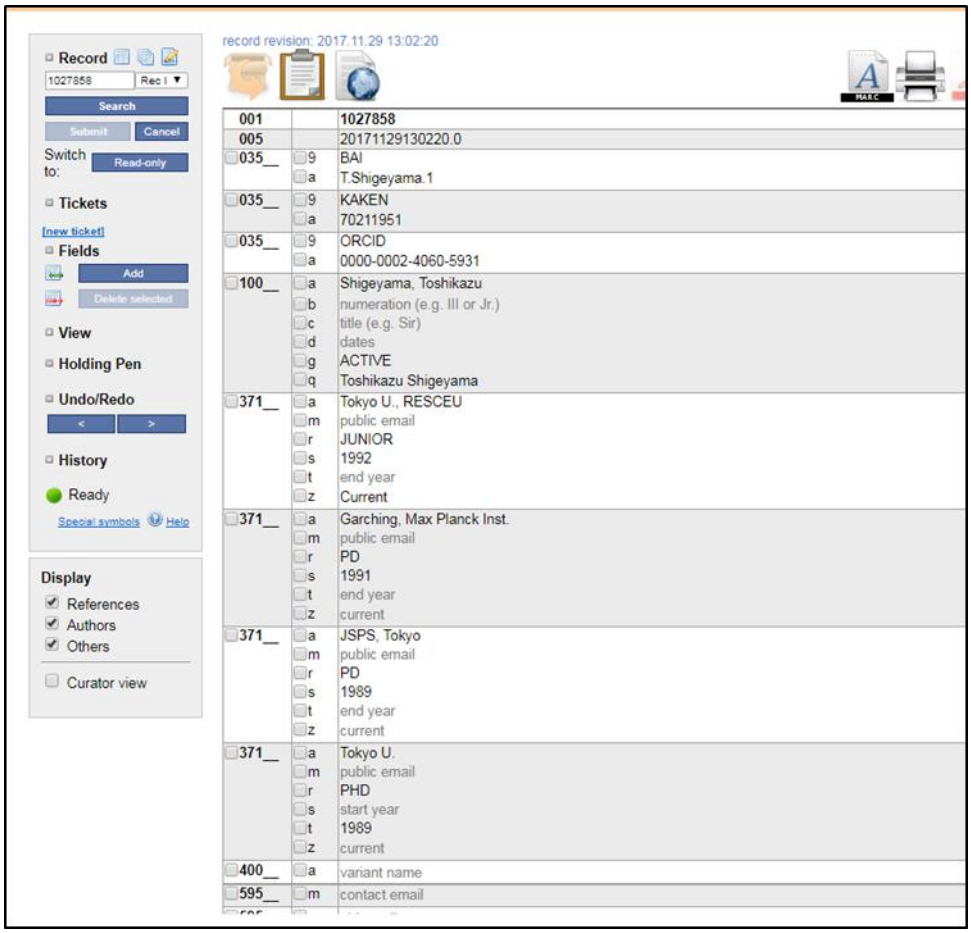
- HEP: Published paper/eprint curation
  https://labs.inspirehep.net/internal-help/knowledge-base/hep-publishedeprint-curation/
- Proceedings metadata table https://labs.inspirehep.net/internal-help/knowledge-base/hep-proceedings-metadata-table/

3.2 HepNames

著者情報のメタデータ。Bibedit にて、規定のフィールドを編集。

参考)

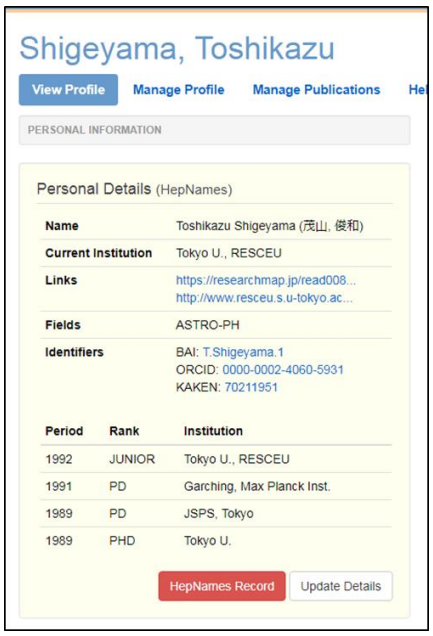
- HepNames: metadata table
  https://labs.inspirehep.net/internal-help/knowledge-base/hepnames-metadata-table/



HepNames の Bibedit 画面

3.3 Author Profile

- 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の自動分析および、著者の表記の揺れを吸収して、著者ごとの論文プロファイル Author Profile が生成される。
- Author Profile には、HepNames、論文から分析した著者所属機関、共著者、分野、関係する実験グループ、キーワード、出版物の種別、分野、引用文献情報当が表示されている。Author Profile と HepNames は、BAI をキーとしてリンクづけられていなければならない。ただしリンクづけられていない場合は、Author Profile に候補者の HepNames が表示される。



左：リンクづけられている例 http://inspirehep.net/author/profile/T.Shigeyama.1
右：リンクづけられていない例

4. 日本人著者情報の充実

4.1 作業実施前の問題点

- Author Profile と HepNames がリンク付けられていない
- HEPNames が未作成あるいは、情報に不足がある。
- Author Profile に別人の論文が混入、あるいは含まれるべき論文情報の欠如（別人の Author Profile に混入、あるいは Author Profile が重複して存在）

4.2. 作業手順(概要)

- 対象：「tokyo」を名称に含む機関（東大、東京理科大等）に在籍歴があり、Author Profile と HepNames がリンク付けられていない著者約 650 名

1）対象著者の特定（所属、移動履歴、漢字氏名）※詳細は「4.3. 作業手順(詳細)」
2) 1)で特定した情報にもとづき、Author Profile の論文情報の紐づけおよび、別人論文の切り離し、Author Profile と HepNames がリンク付ける。

3) 1)で特定した情報にもとづき、HepName の情報を充実させる。具体的には漢字氏名、移動履歴、各種 ID(ORCID, KAKEN)の追加

4.3. 作業手順(詳細)

対象著者の特定（所属、移動履歴、漢字氏名）する上での問題点

- HepNames に収録されている所属情報は、必ずしも現在の所属情報とは限らない。
- 2)の論文情報リンク付け時の根拠は、「論文掲載年」およびその時点の「所属機関」。

⇒ 事前に対象著者の移動履歴を把握しておく必要

⇒ 漢字フルネームを最初に特定しておくと、より正確な情報を効率的に取得できる

KAKEN ではアルファベットでヒットしなくても漢字氏名ではヒットする場合があったり、個人ウェブサイトの多くが日本語で作成されていたりするため。

<漢字氏名の特定>

下記ツールを用いて、アルファベット姓（＋アルファベット名/所属機関など）検索

1. CiNii　http://ci.nii.ac.jp/
   [長所] 検索対象が幅広く、院生でも特定できる場合あり
   [短所] ヒット数が多い（同一著者のものも重複あり）ので、絞込みが手間。日本語論文を書いていない場合はヒットしない。漢字氏名が登録されていない場合もある。
2. Researchmap　https://researchmap.jp/search/
   [長所] 収録対象の大部分が研究職に就いているので、対象候補を絞り込みやすい。個人差はあるが、在籍の変遷、主な業績も公開されているので、Author profile の情報と照合する際に便利。
   [短所] 院生や、既に現役を退かれてしばらく経つ研究者は見つかりにくい。
3. KAKEN　https://nrid.nii.ac.jp/en/index/
   [長所] 現役の研究者はおおむね見つかる。在籍の変遷歴や主たる研究業績、共同研究者などの情報も記載され、Author profile の情報と照合する際に便利。
   [短所] 科研費を取得しない人（院生、海外機関勤務の研究者）は収録されていない。アルファベットではヒットしない場合がある。
4. NDL-OPAC の博士論文検索　https://ndlopac.ndl.go.jp
   [長所] 古い博士論文も収録しており網羅性が高い。
   [短所] 最新 1-2 年の学位取得者情報は update されていない。検索時に大学名等を指定できず、検索結果一覧に表示されるのは論題と氏名のみなので、対象者の特定には検索結果一覧から 1 件ずつ詳細結果画面を開くことが必要。
5. CiNii Dissertation　http://ci.nii.ac.jp/d/
   [長所] 検索時に氏名のみならず、大学名や年も指定可能なので予め検索対象を絞り込める。検索結果一覧にも大学名が表示される。
   [短所] 最新 1-2 年の学位取得者情報は update されていない場合がある。古い学位論文情報を見つけられない。
6. JAIRO　http://jairo.nii.ac.jp/　あるいは、各大学の機関リポジトリ
   過去 1-2 年の学位取得者で NDL-OPAC や CiNii Dissertation が update されていない場合も、機関リポジトリでは既に公開されている場合もある。

7. Google

- Author profile 記載の論文に頻繁に登場する共著者は、指導教員である可能性あり。指導教員名を KAKEN や researchmap で特定し、その指導教員の研究室ウェブサイトを Google で見つけて、研究室構成員として著者の漢字氏名を特定できる場合がある。
- 「Author profile 記載の論題　AND 推測した漢字姓」で検索すると、業績を公開している著者個人のウェブサイトが見つかり、漢字姓名の特定につながることがある。
- 「漢字氏名 AND　所属機関名」で Google 検索すると個人ウェブサイト、SNS（facebook, Linkedin 等）、広報記事（講演プログラムに著者略歴記載あり）などの情報が見つかる場合がある。

5. 課題および改善案

- 高エネルギー物理学に関する知識

  自身が高エネルギー物理学に関する知識がないため、「著者」と「在籍履歴」のみを根拠にデータクリーニングを行った。同分野に関する知識が多少なりともあれば、より正確且つ効率的にクリーニング作業ができたと思われる。

- 情報源の不足、偏り

  在籍歴の変遷情報が複数のデータベースやウェブサイトに分散、あるいは存在しないことが多い。特に若い研究者や院生、研究活動をやめて別の業種に移った方は KAKEN や researchmap に登録されていないことが多く、情報源が乏しい。ただ一方で、個人ウェブサイトや Facebook, Researchgate などで情報発信している例が相当数あった。IT 業界等の別業種に移った方は Linkedin で情報発信している例が少なからずあった。最も特定が困難だったのは、現役を退かれ相当数の年数が経つ研究者（おそらく故人）で、情報源が皆無に等しかった。

  更改後の INSPIRE は ORCID との連携をより強化するとのことなので、現役研究者の論文書誌情報は自動的にクリーニングが進むと期待される。また現行 INSPIRE で ORCID との同期設定はできるので、研究者への周知と利用促進活動を行うことで、データクリーニングは進むことが期待される。

  著者の在籍歴の変遷については、著者が属する学会から一定年数分の名簿情報（漢字氏名、所属機関名のみ）の提供を得られれば、比較的古い時期の業績でも正確かつ効率的なクリーニングが可能になると思われる。

以上

詳細報告 2-1

平成 29 年度　国立情報学研究所実務研修　報告書

国際図書館コンソーシアム連合（ICOLC）2017 年秋季会合参加報告

提出日：平成 30 年 2 月 23 日
所属：東京大学附属図書館
氏名：横井慶子

1. 開催概要

日 程：2017 年 10 月 15 日（日） ～ 18 日（水）
場 所：プラハ（チェコ共和国）National Library of Technology（NTK）

2. 発表概要

10/15 （日）

場所：Clementinum Exhibition Hall
17:00-19:00
内容：レセプション
各コンソーシアム参加者と情報交換
Clementinum 内図書室の見学

10/16（月）

場所：National Technology Library Balling Hall
9:30-10:30 Open Access : EU/EUA Perspectives
- EU copyright reform and Open Access

Speaker:Julia Reda (EU Member of Parliament)

EU における著作権法改正の内容と、それが OA へ与える影響について説明があった。改正後は、著作権者の許諾がないリンクが規制対象となり得る。このため Researchgate のような研究者向け SNS では、サイト運営者がアップロードされたコンテンツの著作権の確認等が求められることになる。

- Germany and Open Access: The DEAL project

Speaker:Hildegard Schaeffler, Bayern‐Konsortium (Germany)

ドイツの Project DEAL の概要説明と、具体的な契約状況について報告があった。Springer Nature と Wiley とは合意に至ったが, Elsevie とはいまだ合意に至っていないとのことだった。

Project DEAL ウェブサイト：　https://www.projekt-deal.de/

10:50-12:00 Open Access ‐ Economic aspects
- OpenAPC project

Speaker:Dirk Pieper, Bielefeld University (Consortium)
Bielefeld 大学を中心に行われてきた Open APC プロジェクトについて

得られた情報からの分析結果が紹介された。2017 年 10 月時点では、117 機関 44,425 件の APC データが GitHub で公開されており、平均支払い APC 額は、Full OA 誌で 1,479 ユーロ、ハイブリッド誌で 2,493 ユーロとのことだった。

Open APC： https://www.intact-project.org/openapc/

- OA2020 and its feasibility in SELL countries

Speaker:Sandrine Malotaux, SELL (South European Library Link)

フランスなどの 6 参加国とオブザーバー2 か国とで構成される SELL（South European Library Link)の構成国にて、購読モデルからの flipping が実現可能かを検討した結果が報告された。算出根拠には 2015 年の購読費、および各国の著者が筆答著者になっている論文数が用いられた。その結果トルコ以外では、購読費よりも出版費が高く、APC が 1,200 ユーロであれば実現可能との結論が得られたとのことだった。

- Growing Urgency for a Global Transformation

Speaker:Ralf Schimmer, Max Planck Digital Library (Germany)

現在の学術雑誌購読モデルの限界を論じたうえで、Max Planck Institutes 生産の論文掲載状況を根拠に OA2020 実現の必要性について論じられた。

13:00-13:45 European OA regional principles and practices

- Nordic Collaboration on open access and licensing

Speaker:Britt‐Marie Wideberg, Annica Wentzel, BIBSAM (Sweden)

北欧のコンソーシアムを中心に開催されたワークショップの内容について報告された。Open APC Sweden、契約時に OA に関する部分について注意すべき事項のチェックリストなどを題材にディスカッションが行われたとのことだった。

- OA2020 Transformation Workshop follow‐up: Outcomes and results

Speaker:Colleen Campbell, Max Planck Digital Library (Germany)

OA2020 実現に向けてのロードマップ作成、必要なスキルの習得を目指して開催されたワークショップの報告がなされた。ワークショップの議題はデータ分析、戦略づくり、交渉、アグリーメントの管理、ワークフローの効率化などであった。検討する中で、個別には行っていても、それらがうまく組み合わせられず十分に機能していなかったことなどの気づきが報告された。

13:45-14:00 Conversion of academic journals to Open Access

- Converting Nordic Academic Journal to Open access

Speaker:Solveig Wikstrøm & Tore Nilsen, CERES (Norway)

北欧言語で書かれた人文社会科学系の学術雑誌をゴールド OA 化する取り組みについて事例報告が行われた。人文社会科学系の学術雑誌は、もともと政府の支援を受けていたが、2017 年以降は対象がゴールド OA 誌のみとなり、42 誌がゴールド OA 誌となった。ゴールド OA 化には政府の支援に加え、大規模大学からの拠出金も用いられた。

- Converting French Academic Journal to Open access

Speaker:Sandrine Malotaux, Couperin (France)

フランスの Ministry of Higher Education and Research が自国の出版社 EDP Science に対して行っている支援について報告された。この支援は、フランスのフラッグシップ誌の他社への移管防止と、ゴールド OA 誌転換を目的としている。Ministry of Higher Education and Research と参加機関がコストを負担することで、EDP Science とナショナルライセンス契約を結ばれている。参加機関は同社の全タイトルへアクセス可能で、筆答著者が参加機関に属している場合は論文を OA にできる。

14:50-16:00 Panel: Conversion of academic journals to Open Access

Susan Murray, Executive Director, African Journals Online (South Africa)

Journal Publishing Practices and Standards(JPPS)という、南アフリカでの取り組みが報告された。南アフリカの学術雑誌は、独自出版のものが多く、電子化も進んでおらず、ハゲタカ出版社の学術雑誌との誤解を受けることもある。大部分は Journal Impact Factor の収録対象にはなっていない。このため、JPPS では出版の質を基準に学術雑誌を 6 段階に評価している。具体的な評価基準は、ISSN などの基本情報の整備状況、著者への情報提供状況、刊行状況などである。

Leo Waaijers, Open Access Consultant (Netherlands)

出版の質を基準に図書館員等に各雑誌のスコアをつけてもらい学術雑誌の品質を測る、独自開発の仕組み Quality Open Access Market(QOAM)について紹介された。

Emilija Banionytė, Lithuanian Research Library Consortium

過去の事例にもとづき、ハゲタカ出版社の例が複数紹介された。たとえば、実在する学術雑誌の誌名や ISSN を騙り、偽物のウェブサイトを作成し、研究者からの投稿を誘導して APC をだまし取るといった手口などがある。

16:00-17:00

Vendor grill / discussion

Lucas Toutloff, Senior Project Manager, Cabell's

学術雑誌を独自の基準で評価して、ブラックリストとホワイトリストとして示す有料データベース Cabell's の概要が紹介された。会場からは信頼性や独自基準の妥当性などについて多く質問が寄せられた。

10/17（火）

場所：National Technology Library Balling Hall

8:30-10:00 Open source tools development for consortia

- CzechELib ERMS currently in development

Speaker:Jiri Jirat, CzeckElib (Czech Republic) Thomas Porquet, Couperin (France)

開発中の電子情報管理ツール CzechElib について、その機能概要が説明された。同システムでは既存の Verde や Coral、Alma などの電子情報管理ツールでは実現できていない、コンソーシアム向け機能などが盛り込まれるとのことであった。

- ezMeasure (Usage Stats tool – Couperin)

Speaker:Sandrine Malotaux, Couperin (France)

利用統計ツール ezPaarse と ezMesure について概要説明があった。前者はインストールして自機関の利用統計データを取得できる。後者は各機関の利用統計データをビジュアル化して表示できる。

- Electronic Resources Management System, ANKOS

Speaker:Sami Cukadar, ANKOS Board Member (Turkey)

トルコのコンソーシアム ANKOS の電子情報管理システムの概要説明と活用事例が報告された。活用事例としては契約交渉時に、過去の契約総額や契約数、一人当たりの論文利用数などが用いられているとのことだった。

- Strength in numbers: collaborative approaches to extending Jisc's library analytics services

Speaker:Jo Lambert, JISC (UK)

Journal Usage Statistics Portal (JUSP) の概要説明および利用事例報告があった。JUSP は雑誌、e-book、データベースの利用統計を一元化して示し、個々の図書館が行うデータ収集、分析の時間や労力を節約するものである、ということであった。

https://jusp.jisc.ac.uk/

10:00 -10:30

Battlefield Survey

Presenters:Dyveke Sijm, DEFF (Denmark) & Graham Stone, JISC (UK)

世界各国のコンソーシアムから提供された契約情報を分析した結果が示され、地域別にみると、欧米に比べてアジアは比較的契約額が抑えられていること等が指摘された。

11:00-12:00 Managing Consortial Collections

- South Africa experience

Speaker:Nora Buchanan

南アフリカの図書館のコレクション構築における課題について報告された。版元やベンダーから提供されるタイトルリストの情報が不正確なことが頻繁にあることが、特に問題視されていた。

- Joint Engagement in Managing Holdings

Speaker:Sami Hero, VP of Marketing and Customer Experience, Wolters Kluwer

Wolters Kluwer 社の概要と、同社の行っているタイトル管理の内容が紹介された。現状では Excel, HTML, KBART での提供を行っているが、将来的には図書館のニーズに応じた形での提供を実現にむけて取り組みたいとのことであった。

13:00-14:00 New consortial acquisition models

- Workflows in OA deals – the weakest link (Austria)

Speaker:Rita Pinhasi & Brigitte Kromp, KEMÖ (Austria)

論文を Open Access するにあたり発生する作業や金銭の処理について、ワークフローにもとづき紹介された。オフセットモデルの場合は、APC 支払額に応じて翌年の購読料が減額されるので、二重払いの防止につながるメリットがある一方、会計手続きの煩雑さが問題である。その点、Publish & Read モデルは購読料と APC を事前に一括払いするため会計手続きの煩雑さは軽減されるものの、支払額が増加するリスクがある。これらの点を考慮して効率的なワークフローのあり方を検討する必要性が説かれた。

- PALNI Shared Collections (Indiana, USA)

Speaker:Randy Dykhuis & Rick Burke for Kirsten Leonard, PALNI (USA)

中小規模私立大学で構成される、資料の共同利用のためのコンソーシアム PALNI について、概要と運用面での課題等について紹介された。予算やシステムの違いなどもあるが、それらは随時調整して今後の運用を検討したいとのことであった。

14:00-15:00 Consortial services – Return on Investment

- Case study: NEICON Russia

Speaker:Irina Razumova, NEICON (Russia)

1000 機関以上を抱えるロシアのコンソーシアム NEICON が直面している問題について予算的な事情を中心に説明された。以前は国のファンディングプロジェクトによる購読がおこなわれ、個別契約に比べて購読費を大幅に抑えることができた。だが一方でそのことにより各大学の権限が縮小され、直面するさまざまな窮状に個別での対応が難しくなっている現状が述べられた。

- Herding Cats & Getting to Yes

Speaker:Mihoko Hosoi, CDL (USA)

カリフォルニア大学は 10 のキャンパスで構成され、それぞれが強い権限をもつために、大学としての意見のとりまとめが難しいという背景がある。それを行うにあたり、ビジネス論理をベースにした考え方が紹介された。数値や正論だけを用いるのではなく、ストーリーテラー且つファシリテーターとなり、意見を平等に取り扱うことの大切さが述べられた。また、最終合意にいたるまでには、相反する考えや提案に対しても、No と突き返すのではなく、まずは「Yes, but…」といったん受け入れてから話を進めることの大切さも述べられた。

- From little to many: Stakeholders in the scientific information provision in the Netherlands

Speaker:Pim Slot, SURFmarket (Netherlands)

オランダにはコンソーシアム SURFmarket、大学図書館と国立図書館間のコンソーシアム UKB、オランダ大学協会 VNSU の三者が契約交渉に関与することになった。その中で SURFmarket は交渉材料となるデータ収集の役割を担っている、との報告がなされた。

15:15-16:00 Breakout sessions 1

Battlefield survey discussion Dyveke Sijm & Graham Stone, JISC (UK)

2017 年度契約の内容について、参加者間で意見交換を行い、2018 年度契約に向けての注意点を確認した。

16:15-17:00 Breakout sessions 2

Evidence based acquisition models for consortia Sharla Lair, LYRASIS (USA)

コンソーシアム契約を結ぶにあたっての注意すべき点を列挙し、それらをコンソーシアムのニーズにあわせてカスタマイズする必要性が論じられた。

10/18（水）

場所：National Technology Library Balling Hall

9:00-10:10 Central and Eastern Europe consortia: Issues

- First year of CzechELib Adam Der, EIS (Hungary)

電子資源の購入とアクセス確保のためのチェコのプロジェクト CzechELib は、EU operating program（2017 年-2022 年）の一環で 5000 万ユーロの予算がついている。これまでの活動としては、コンテンツの購読希望調査やシステム開発、標準的なアグリーメントの作成、契約交渉などを行ってきたとの報告がなされた。

また中東欧のコンソーシアムが抱える課題について、ハンガリー、リトアニア、クロアチアからそれぞれ報告があった。ハンガリーの場合は、コンテンツの増加と価格上昇により各参加館の負担額が増加傾向にあることが指摘された。リトアニアの場合は、財源がコンソーシアム参加館の会費と EU から補助されるプロジェクト資金にまかなわれているが、後者は 2021 年にはなくなるためその後の運営が課題として述べられた。クロアチアの場合は、ナショナルライセンスやディスカバリーサービスの導入など進めているが為替の動きによってコンテンツ整備に影響を及ぼすことがあるなどの問題点が示された。

3. 所感

初めての ICOLC 参加であったが、他の参加者は過去に複数回参加している方が多数いらっしゃり、互いによく知る関係のようであった。このように何度も直接顔を合わせて会話することで、連帯感が生まれ、率直で惜しみのない情報交換が可能になっているのかと感じ、ICOLC の重要性を強く感じた。

全体的にオープンアクセスに係るトピックが多く、各国それぞれの立場、制約下で、直面する状況に対して奮闘している様子が感じ取られた。

以上

詳細報告 2-2

平成 29 年度　国立情報学研究所事務研修　報告書

ドイツ国立科学技術図書館（TIB）における研究データ管理
および電子ジャーナル契約に関する取り組みの調査

提出日：平成 30 年 2 月 23 日
所属：東京大学附属図書館
氏名：横井慶子

1. ドイツ国立科学技術図書館（TIB: Technische Informations Bibliothek）概要（2016 年時点）

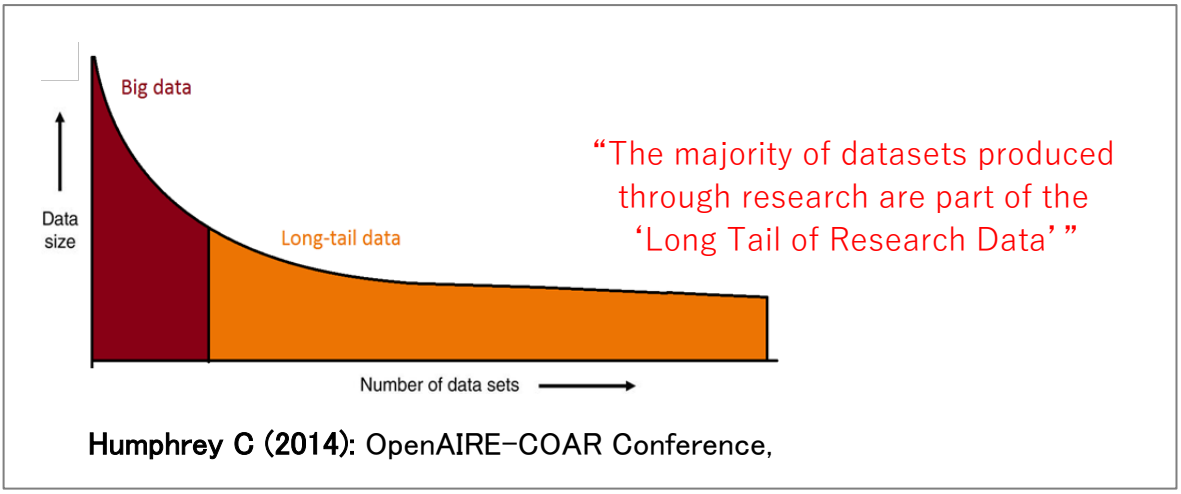
- 総予算 4,710 万ユーロ（資料費 1,750 万€、Engineering and sciences 1,400 万€）
- 第三者機関からの助成 150 万ユーロ、自己収入 150 万ユーロ
- 職員数（非常勤含む） 536 名、Apprentices（研修生のような位置付け）: 6 名
- 蔵書数（図書 580 万冊、マイクロ資料 340 万、電子資料（e-book 含）15 万、電子ジャーナル 43,000 誌、専門的な印刷版雑誌 15,000 誌、専門的なデータベース 5,600、専門的なビデオ資料 8、700(TIB AV-Portal: https://av.tib.eu/ で公開、特許および規格類 1,700 万）

2. 訪問調査

日　時：2017 年 12 月 19 日　9:00-14:00
対応者：(1) Irina Sens / Deputy Director
(2) Linna Lu / Subject specialist for Computer Science
(3) Angelina Kraft / Project manager of RADAR
(4) Dana Vosberg / License Manager

3. RADAR 概要

FAIR principles: Findable, Accessible, Interoperable, Reusable の方針のもと研究データの公開に力を入れている。研究データの中でも、一部のビッグデータよりもロングテールに位置するさまざまな分野のデータをアーカイブ、公開する RADAR: Research Data Repository プロジェクトを、2017 年 3 月から公開した。German Research Foundation (DFG)の助成を受け、2013-2016 年の 3 年間かけて準備された。



＜研究データの存在イメージ図＞

利用者としては研究者、図書館、研究機関、博物館、文書館などを想定している。基本的なサービスである Archival Service(5-15 年)と、拡大サービスである Data Publication(25 年)がある。使用料金は Archival: 年間使用料 500 € + 0.39 € GB data volume (正価)、Publication: 25 年間の利用保証で 6.37 € GB data volume (正価)と設定している。

データのステータスは Pending(6 ヶ月：編集可)、Review(6 ヶ月：編集不可、peer review 期間)、Archive(5-15 年：固定、RADA ID 付与)、Publish(25 年以下：固定、DOI 付与)と遷移する。データセンターはドイツ国内の Karlsruhe という都市にあり、10,000 スロット、2~8TB のテープがあり、8 つのロボットが動いている。

4. 電子ジャーナル契約

ドイツでは、地域別、国を超えてドイツ語圏の複数の国など、さまざまな形でコンソーシアムが組まれている。コンソーシアム参加機関、コンソーシアム契約の種類（出版社別）は年々増加している。ナショナルサイトライセンス契約にもとづき利用できていた電子ジャーナルコンテンツのダウンロード数は近年伸び悩んでいる。ナショナルサイトライセンスは 3 年単位で契約していたが、今後は延長せず、Open Access に力を入れる方針となった。

5. 所感

一連のプレゼンとディスカッションを通して、科学技術情報の収集と公開に対する TIB スタッフの使命感と、その実現のための行動力を強く感じた。たとえば、RADAR は、2010 年に企画、2012 年に DFG へ申請、2013 年 9 月に承認という流れで実現された、と伺った。企画段階において、研究データのおかれている状況をふまえ、ターゲットをロングテールに属する研究データと定めたこと、そしてプロジェクトの実現につなげた一連の流れが、とても戦略的であり、強い信念を感じた。

分野の違いに由来する研究データに関する話題も興味深かった。climate science, biology genetics などデータ共有によって研究が促進される分野では研究データ収集に対して協力的だが chemistry, architecture, computer science ではあまり協力的とはいえないというのが実情らしい。また、研究データの取り扱いには、その分野における専門知識が必要となる。その点については職掌にこだわらず、各図書館員が自身のもう一つの専門（スタッフは、基本的に図書館情報学以外の学位も有している）に応じて柔軟に協力し合って対応しているとのことだった。

詳細報告 2-3

平成 29 年度　国立情報学研究所実務研修　報告書

Max Planck Digital Library(MPDL)における
APC 管理および研究者投稿行動把握に関する調査

提出日：平成 30 年 2 月 23 日
所属：東京大学附属図書館
氏名：横井慶子

1. Max Planck Digital Library(MPDL)概要

ドイツ各地に 80 以上の研究所を擁するマックス・プランク協会（Max Planck Society for the Advancement of Science）に対して、情報と関連サービスを提供。OA2020 を主導し、最初にオフセットモデルを導入した。

2. 訪問調査

日　時：2017 年 12 月 21 日　10:00-14:00
対応者：

- Ralf Schimmer / Head of Information
- Kai Geschuhn / Innovative Services & Open Access
- Michael Schlachter / Open Access Projects
- Adriana Sikora / Knowledge Base Maintenance & Invoice Processing

3. 生産論文数の調査

3.1. 調査方法

Max Planck 所属研究者の論文掲載状況は、Scopus と Web of Science からデータを購入し、5 名から成る Data Analysis team が分析を行っている。PostgreSQL と PYTHON を用いてデータ分析ができるスキルが求められ、うち 2 名は博士号の学位を取得している。調査対象は Max Planck 所属研究者に限らず、スイスや韓国の研究機関、マサチューセッツ工科大学、カリフォルニア工科大学などの海外の機関も含まれるが世界中全ての機関というではない。論文生産の対象機関の判定は、Corresponding Author の所属機関を基準に行っている。

3.2. 調査から判明したこと

SCOAP[3] の影響で、2014 年以降 Max Planck 所属研究者の OA 論文数が増加している。2016 年には springer compact の契約が始まり、Springer での OA 論文数も増えており、それに対応し Springer 購読誌への掲載論文数は減少している。一方 Elsevier の場合は、論文数自体には大きな変化は見られないが、Max Planck 所属研究者生産論文全体に占める割合は減少傾向にあり、注目に値する。論文の 80%は 20 の出版社の雑誌に掲載されており、これは Max Planck 所属研究者に限らず、世界の他機関の研究者に関しても同様にいえることである。

<MPDL の目標>

Springer(2016-2018)、 RSC(2017-2018)、 Taylor & Francis(2017-2019)とは既にオフセット契約を結び、他にも 3 社と 2018 年以降にオフセット契約を結ぶ見込みである。この対象を、2020 年までに論文の 80%を掲載する上位 20 社全てに拡大し、OA を推進することが目標である。

4. オフセット契約を結ぶ上でのリスク、注意点

MPDL がオフセット契約を結ぶ際には、出版社と複数回の交渉を重ね、その都度、購読コスト、購読額の年間上昇率、APC ディスカウント率、APC 前払い対象論文数といった要素を何パターンも検討して、最終的な合意に達した。OA 論文の生産数が予定数を超えると予定以上の額を支払うリスクがある。このため 10 年単位で対象出版社の雑誌に掲載された自機関研究者の論文数の推移を見て平均値を割り出し、適切な予測を立てることが重要である。また出版社との交渉にあたっては、その出版社の雑誌掲載論文全体のうち、自機関研究者の論文数が占める割合を把握しておくことも交渉を進める上で重要である。

5. 研究者視点で見た論文 OA 化（投稿時の画面遷移）

なるべく研究者の手間を省略し、投稿時に考え込む時間や、選択のためのクリックが少なくスムーズに OA 論文を投稿できるようにしたいという考えである。出版社ごとに異なるが、1)投稿時のプルダウンメニューで機関名を選択、2)IP レンジで機関名を特定、3)メールアドレスから機関名を特定するといった方法で、画面上に MPDL と出版社間にもとづく OA 条件、APC 支払い条件が表示される。例えば PLoS の場合は、自動的に「APC は Max Plank が支払う」といった説明が表示されるだけで、研究者はそれ以上何も行う必要はない。ただし出版社によってプロセスは異なり、研究者自身に複数の項目の入力を求める場合もある。

copyright については、研究者には CC-BY を選択してほしいが、copyright に関する説明部分は読み飛ばされがちで、意識的に CC-BY を選択してくれる研究者は少ないという印象。Springer の場合は CC-BY のみなので、自動的に CC-BY が適用されて問題ないが、Taylor& Francis の場合は CC-BY-NC-ND、 CC-BY、RSC の場合は CC-BY-NC、CC-BY のいずれかを選択する形式なので、研究者が CC-BY を選択するように、理解を深めてもらうための努力が必要。

6. 所感

Data Analysis team をつくり新たに 5 名を雇用してまで、生産論文数調査を徹底する MPDL の姿勢、そしてその結果の活用状況を知り、同調査の必要性を強く感じた。それと同時に、同調査の精度を高めるには高い IT スキルが必要だが、現場スタッフに新たにそれを習得させるのではなく、既にそのスキルをもつ人を雇用するという対応も興味深かった。

Max Planck 生産論文数の掲載誌の出版社に関する説明の中では、Monthly Notices of the Royal Astronomical Society（Max Planck 論文掲載数が 2 番目に高い雑誌）が Wiley から OUP へ移管したために、OUP の比重が大きくなった、との話があった。またオフセット契約では、契約期間中に APC の値上げをしないといった取り交わしは行なっていないとの話もあった。出版社との契約のあり方を考える上で、さまざまな不確定要素があることがわかり、オフセット契約を結ぶ際の難しさを感じた。