

13th International Digital Curation Conference 参加報告

2018年4月5日
鹿児島大学 西菌由依
筑波大学 松野渉

1. 会議概要

会議名：13th International Digital Curation Conference ¹

日程：2018年2月19日(月)～22日(木)

会場：NH Collection Barcelona Tower, Barcelona

2. 参加目的

International Digital Curation Conference (IDCC)は、研究データの共有や活用をテーマに、英国 Digital Curation Centre (DCC)が毎年開催している国際会議である。研究データの共有・利活用はわが国の学術情報流通においても喫緊の課題であり、国際的な取り組みとの協調的な進展が求められている。今後のわが国における幅広い議論の為、本件に関して諸外国における最新の動向を把握することは非常に重要であると考えられる。本年のIDCCへは上記のような目的を持ち、日本から筆者ら二名が参加した。

3. 概要

会議は本会議とワークショップから構成され、本会議には欧米やアジアから約180名の参加があった。今回の会議テーマは“Beyond FAIR - from principles to practice to global join up”（試訳「FAIRを超えて：原則から実践に向けてグローバルに連携を」²）とされ、研究データ管理に関する幅広い実践的な議論が行われた。以下、今後の日本における取り組みに参考となると思われるものを中心に報告する。



(会場の様子)

¹ <http://www.dcc.ac.uk/events/idcc18>

² 「FAIR データ原則」については以下を参照:

<https://www.force11.org/group/fairgroup/fairprinciples>

3-1. 本会議 (2月20~21日)

2日間にわたって開催され、基調講演3本、研究発表3本、発表27本(パラレルセッション)、デモンストレーション4種、ポスターセッションが行われた。セッションのテーマは以下のとおり(末尾の*は筆者参加を示す)：

- ・ Enabling and Measuring FAIR Data
- ・ Cross-Institutional and National Data Services *
- ・ More than Data *
- ・ Repository Services
- ・ Ethics and Legal
- ・ Data Management Planning *
- ・ Data Creation and Re-Use Studies
- ・ Data Policy and Guidance *

3-1-1. Complexities of Digital Preservation in a Virtual Reality Environment, the Case of Virtual Bethel (研究発表)

本発表では、近年普及が進む仮想現実空間(Virtual Reality(VR))を構成するデータの保存に関する取り組みについて紹介があった。題材として取り上げられたのは米国インディアナ州の”Bethel Church”をVR上で”保存”するプロジェクト”Virtual Bethel”である。同教会はアメリカにおけるアフリカ系アメリカ人コミュニティにとって非常に重要な建造物であったが、近年売却され、別の施設になる予定であった。そこでインディアナ大学の図書館情報学系、メディアアート系の学生・教員らによりこのプロジェクトが立ち上がったとのことである。このプロジェクトはインディアナ歴史協会とIUPUI(インディアナ大学・パデュー大学インディアナポリス校)からの資金提供を受けている。

実装されたVRデータの保存について既存のワークフローやポリシーは存在しないことから、このプロジェクトでは様々なアプローチが検討されているようだが、特定の技術やファイル形式に依存する事なく、将来に渡ってユーザーのVRへのアクセスを保証する為、計画のさらなる具体化が必要とされているようだ。当該データは現在クラウドストレージに保存されており、近い将来のうちにインディアナ大学のDSpaceに取り込まれるとの事だが、その後の長期保存については他のオプションも検討されている。

VRに限らず、全く新しい技術やそれに伴うデータが生じた際、その保存や利活用をどのようなプロセスで検討するかという点において、この報告は非常に興味深いものであった。

Virtual Bethel <https://comet.soic.iupui.edu/bethel/>

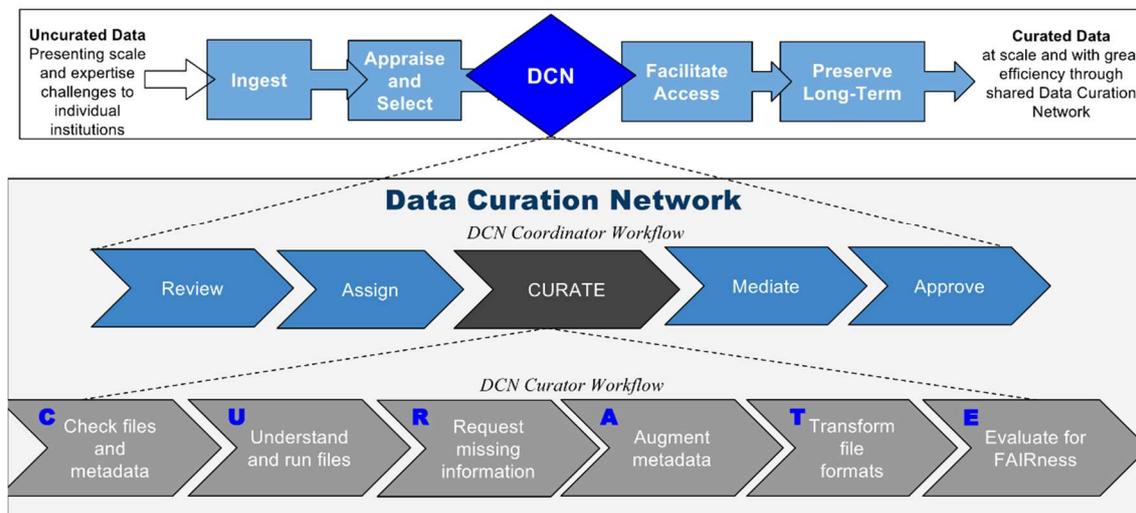
3-1-2. Data Curation Network: A Cross-Institutional Staffing Model for Curating Research Data (“Cross-Institutional and National Data Services”セッション)

機関にとって研究データ管理に必要なスキルや専門知識を有した人材の確保が課題となっているが、本発表では、複数機関が連携して研究データ管理に関する人的資源を共有する取り組み"Data Curation Network (DCN)"について紹介があった。ミネソタ大学やミシガン大学等の米国東部 6 大学³による計画段階（2016 年～2017 年）を経て、参加機関を拡大し（8 大学⁴およびデータリポジトリ Dryad）、実装段階（2018 年 5 月～3 年間）に入るとのことである。実装計画は、人的資源の共有、特定の技術に依存しない投稿ワークフロー、標準化されたキュレーション手続き、これらが一体となっており、3 年間の助成期間終了後も、DCN が、必要なキュレーションサービスを利用できるモデルとして普及していけるように、持続可能性を考慮したアプローチを取っている。

各参加機関は、自機関のデータ管理サービスに研究者から投稿されたデータセットを、必要に応じて各機関のリポジトリシステムと統合されている DCN へ送り、専門的なキュレーションを得ることができる。DCN 全体で様々な主題知識に精通した 19 名のデータキュレーターを擁しており、CURATE の略称で呼ばれる標準化されたキュレーション手続きが行われる。

日本でも人材の確保は重要な課題であり、本取り組みの動向については継続的に注視していく必要があると思われる。

DCN Workflow



図：DCN ワークフロー（スライド資料より）

Data Curation Network <https://sites.google.com/site/datacurationnetwork/>

³ The University of Minnesota, Cornell University, Penn State University, the University of Illinois, the University of Michigan, Washington University in St. Louis.

⁴ 前述の 6 大学に Duke University, Johns Hopkins University の 2 大学が加わった。

3-1-3. DMPRoadmap: new features in DMPonline and DMPTool (デモンストレーション)

米国発のデータ管理計画作成ツール"DMPTool"と英国発の同ツール"DMPOnline"との間で進められていた、両ツールのコード統合作業 (DMPRoadmap) が完了し、新機能の紹介が行われた。新機能の主な例は以下のとおり：(i)各機関等が提供しているガイダンスの情報が表示されるようになった、(ii)利用者が、機関内の他者に代行権限を付与できるようになった、(iii)機関管理者が、機関ロゴや機関の各種関連情報 (例えば RDM ポリシー) へのリンクを設定できるようになった、(iv)機関管理者が、構成員に対してデータ管理計画へのフィードバックサービスを提供できるようになった (提供するかどうかは選択可能)。

- ・ DMPRoadmap <https://github.com/DMPRoadmap/roadmap>
 - ・ DMPTool <https://dmp.cdlib.org/> ※2018/02/27 付で新コード反映 (ver.3 リリース)
 - ・ DMPOnline <https://dmponline.dcc.ac.uk/> ※2018/03/08 付で新コード反映
- ※両ツールは統一コードの下、今後も各々サービスを継続する。

3-1-4. The Data Deposit Recommendation Service (DDRS) (デモンストレーション)

オランダの Data Archiving and Network Services (DANS)より、特定分野 (人文科学分野) の研究者に向けたデータリポジトリ推薦サービス"The Data Deposit Recommendation Service (DDRS)"について紹介があった。選択肢から国や研究分野を選択すると、登録に適したデータリポジトリを提示し、さらにこのサービスを介して、選択したデータリポジトリへのデータ登録手続きを進めることもできる。既存のデータリポジトリディレクトリ re3data.org の情報をベースにして開発されており、情報量が膨大な re3data.org を、特定分野の研究者にとって使いやすい形に調整して提供している事例と言える。

DDRS: <https://ddrs-dev.dariah.eu/ddrs/>

3-1-5. A Landscape Survey of #ActiveDMPs ("Data Management Planning" セッション)

データ管理計画 (DMP) を、現在の静的な文書から動的で機械処理可能なものへ転換することで、研究者にとって負担となるものではなく、研究者やその他のステークホルダーにとってより有益なものにしようとする、"ActiveDMP"への関心が高まっている。発表では、"ActiveDMP"の要件・優先事項とその解決方法について情報収集活動を行っているグループや、各種技術基盤、機械処理可能な DMP のための 10 のルール (提言) についての紹介があった。現在、"ActiveDMP"に関する Web サイトも立ち上がっている。

- ・ RDA - Active Data Management Plans Interest Group
<https://www.rd-alliance.org/groups/active-data-management-plans.html>
- ・ RDA - DMP Common Standards Working Group

- <https://www.rd-alliance.org/groups/dmp-common-standards-wg>
- ・ RDA - Exposing Data Management Plans Working Group
<https://www.rd-alliance.org/groups/exposing-data-management-plans-wg>
- ・ FORCE11 - FAIR DMP Working Group <https://www.force11.org/group/fairdmp>
- ・ ANDS - Data Management Plans (DMPs) Interest Group
<https://www.ands.org.au/partners-and-communities/ands-communities/dmps-interest-group>
- ・ Miksa, Tomasz, Simms, Stephanie, Mietchen, Daniel, & Jones, Sarah. (2018). Ten simple rules for machine-actionable data management plans (preprint) (Version preprint). <http://doi.org/10.5281/zenodo.1172673>
- ・ ActiveDMP <http://activedmps.org/>

3-1-6. Advancing Policy and Changes for Graduate Data Management (“Data Policy and Guidance” セッション)

研究者に研究データ管理を根付かせるためには、リバタリアン・パートナーリズムの概念を取り入れた戦略を立てるべきだとの提案がなされた。機関・教員・学生それぞれの優先事項や懸念等を理解した上で、研究者にとってより望ましい選択肢をデフォルトとするような制度設計を行い、研究者をサポートする手順を設計していくことが重要であると述べられた。

3-1-7. The impact on authors and editors of introducing Data Availability Statements at Nature journals (“Data Policy and Guidance” セッション)

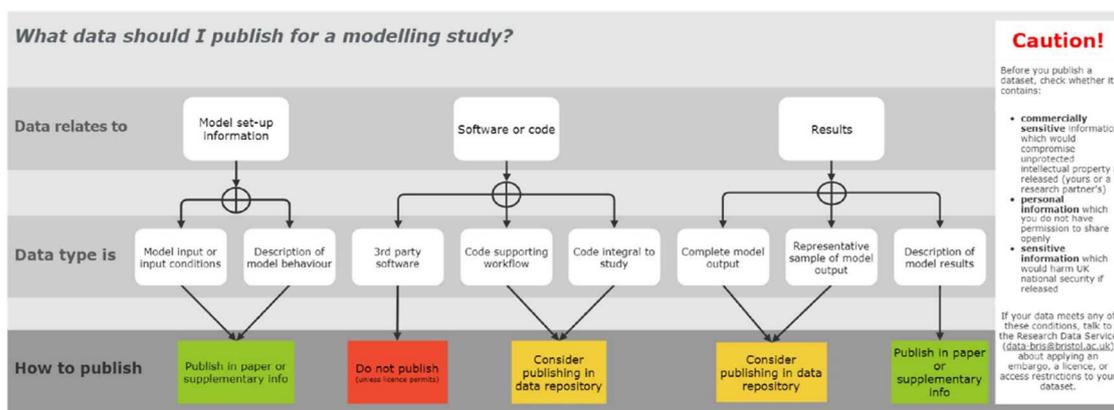
Springer Nature 社より、同社における研究データポリシー標準化の取り組みについての紹介と、論文への「データ利用可能性ステートメント (data availability statements : DAS)」記載義務化が出版社に与える影響についての調査結果報告があった。近年、多くの出版社が、論文の根拠データがどのように利用可能であるかを示す DAS の記載を要求するようになっているが、同社では研究データポリシーと DAS の類型化を行いながら DAS の記載要求を進めているとのことであった。同社がジャーナル編集スタッフに対し調査を行ったところ、DAS 記載義務化にともない、著者が記載した内容の確認、類型ごとに定める書式への整形等の編集作業が必要となり、これらの作業によって編集コスト（所要時間）が増加することが判明した。しかし、DAS 記載の重要さとそれがもたらす便益を考慮すれば妥当な範囲であるとの見解が示された。

3-1-8. Disciplinary data publication guides (“Data Policy and Guidance” セッション)

ブリストル大学より、特定分野（工学分野）の研究者に向けた簡潔なデータ出版図解が

イドを作成した事例について紹介があった。データ出版の基準や各種データポリシーは分野により異なるが、国内外のガイドラインでは分野による違いが十分にカバーされていないこと等を背景に、研究者からの要望を受けて作成したとのことである。データ出版を4つの要素（論文の種類、データの分類、データの種類の、出版の仕組み）に分け、要素間の関係性について概念モデルを作成した上で、決定木を作成した。研究者からは概ね肯定的反応を得ているとのことであった。

課題の一つとして、作成にあたっては当該分野に関する多くの予備知識を要することを挙げていたが、今後も、他分野向けガイドの作成や内容充実といった取り組みを行っていききたいとのことであった。



図：決定木最終版（スライド資料より）

3-2. ワークショップ（2月19, 22日）

本会議の前後2日間に、以下に挙げた6件のワークショップが開催された（末尾の*は筆者参加を示す）：

- Curating for Reproducibility: Producing High Quality Data and Code for Transparent and Reproducible Research
- EOSC as a 'skills commons' for developing research data stewardship skills at scale *
- How Research Institutions and Libraries can help deliver the European Open Science Cloud *
- Train the Trainer Workshop: How do I create a course in research data management? *
- Supporting FAIR Data Principles with Fedora
- A Research Data Shared Service in a localised, international context—what could/should it mean for you? *

3-2-1. EOSC as a ‘skills commons’ for developing research data stewardship skills at scale

欧州委員会（EC）が、既存の基盤をベースにした統合的な研究データ基盤「欧州オープンサイエンスクラウド（European Open Science Cloud : EOSC）」の構築を進めているが、本ワークショップでは、データ管理に関わるスキルの獲得における EOSC の役割について議論が行われた。EOSC のサービスには教育の提供も含まれるが、参加者からは、研究者への教育を行う研究支援者への教育や、研究データ管理におけるスキルギャップを解決するための教育等に優先的に取り組んでほしいとの要望が寄せられた。関連して、EOSC が提供する教材への FAIR 原則適用が可能かどうかという議論も行われた。

3-2-2. Train the Trainer Workshop: How do I create a course in research data management?

本ワークショップでは、研究データ管理に関する教育を行う研究支援者に向けた、研修設計の講義とグループワークが行われ、計 32 名の参加があった。

はじめに、研修実施に際しての方法論や教授法についての説明が行われ、以下のような実践的な助言があった：(i)20 分を目途に講義と実習を切り替えたり講義テーマを変化させること、(ii)学習スタイルの 4 類型（聴覚型、視覚型、コミュニケーション型、運動感覚型）を意識するとともに、それらを組み合わせること、(iii)研究データ管理に関する教育は成人教育であり、成人教育の特徴と留意点について認識しておくこと。

続いて、研究データリテラシーのコア・コンピタンス（Provide, Identify, Scope, Plan, Store, Protect, Evaluate, Manage）⁵と、データ管理に必要な諸スキル⁶、ならびに両者の関係について説明があり、これらを踏まえた上で、研修の所要時間やターゲットに応じた内容構成を検討するよう助言があった。なお、グループワークでは、グループごとに所要時間やターゲットを設定して研修プログラム案を作成したが、議論の過程において、研究データ管理に関わる様々な概念の理解が参加者間で異なることも度々生じた。研究データ管理に関する教育に携わる者同士での共通認識を深めていくことの重要性を実感することとなった。

3-2-3. A Research Data Shared Service in a localised, international context-what could/should it mean for you?

研究データ共有サービス(RDSS)はオープンデータ・オープンサイエンスにおいて最も必要とされるサービスのひとつと考えられる。このワークショップでは RDSS の構築に際し

⁵ 以下に挙げる情報リテラシーの各種プロセスモデル等を統合したもの：Big6（<http://big6.com/>）、SCONUL による 7 つの柱（<https://www.sconul.ac.uk/sites/default/files/documents/coremodel.pdf>）、米国議会図書館による DPOE カリキュラム（<http://www.digit.alpreservation.gov/education/curriculum.html>）

⁶ <http://data-forum.blogspot.jp/2008/12/rdmf2-core-skills-diagram.html>

て、国際的にどのようなニーズが存在するのか、どのような利用シーンが想定されるのかを多様なバックグラウンドを持つ参加者同士のディスカッションによって検討した。

ワークショップを主催したのは JISC が現在推進している RDSS の開発事業に関する国際的なニーズを探ることが本ワークショップの主目的だったと考えられる。

ワークショップを通して明らかになったのは、環境によってオープンデータ・オープンサイエンスに関するワークフローにおける RDSS の位置付けは非常に多様だということである。一例を挙げれば、RDSS において最も重要な要素としてデータ保存の為にストレージとそのセキュリティを挙げる参加者と、Archivematica や Figshare などのコンポーネントの統合レイヤーとして RDSS を位置づける参加者の間で議論が交わされた。

筆者は一連のワークショップを通じ、国際的なデファクト・スタンダードとしての RDSS の構築には非常に多くの要素を（少なくともプラグインレベルで）含む必要があり、その開発はかなりの困難を伴うのではないかという感想を持った。

4. 所感

本会議ではテーマに「原則から実践へ」と掲げられていた通り、諸外国における研究データの共有・利活用についての実践事例が数多く発表されていた。こういった取り組みは決して大規模大学や多くの研究者を擁する研究機関に限られた事ではなく、比較的小さなコミュニティを形成している大学や研究機関からも多くの報告・発表がなされていた事が印象的であった。無論、会議の性質によるバイアスを無視するわけにはいかないが、諸外国、特に欧米圏では今日的な研究データ管理と研究データの共有・利活用が研究サイクルの中で普遍的な一要素となっている事が感じられた。その上で、共通の研究データ基盤をどのように活用していけるかといった議論や、フィードバック機能実装をはじめとした DMP 作成ツールの改修、分野別データリポジトリ推薦サービス、複数機関による人的資源共有の取り組みといった、現在提供中の研究データ管理サービスを実運用面から充実させるための幅広い取り組みが行われている。

翻ってわが国の状況を見ると、オープンサイエンス・研究公正推進の観点から、研究データ管理を重要課題として捉える意識は高まっており、助成団体によるデータ管理計画提出の義務化や、研究データ基盤の構築、教育活動、人材育成の取り組み等が徐々に行われつつある。研究データ管理が研究者の研究サイクルの中に根付いていく過程で、今回の会議で見られたような実践的な取り組みは必須となり、今後、国内でも日本の研究環境や個別の課題に根差した実践事例が出てくると予想される。しかし、これらの取り組みはまだ緒に就いたばかりである。研究データ管理やその利活用推進には、ステークホルダー間でそれを行う意義についての共通認識が深まる事が不可欠であり、今後、大学図書館業界だけでなく、幅広く大学や研究機関を巻き込んだ議論がより一層必要だと考えられる。

最後に、筆者らが今回の会議で見聞した先行諸事例は今後国内で取り組むべき課題を検討する上で示唆に富んでおり、わが国にとって多くの有益な情報を含むものであったと言

える。今後、国内での研究データ管理・利活用が進んだ段階で、国内で得られた経験や知見を本会議のような場で共有し国際的な貢献へ繋げて行くことが望まれる。

以上