

2024.3.15.

国際デジタルキュレーション会議（IDCC2024）等報告会

IDCC24 参加報告 及び OA等に関するインタビュー調査報告

有馬 良一（神戸大学附属図書館電子図書館担当 / JPCOAR研究データ作業部会）
佐桑 諒（神戸大学附属図書館社会科学系情報サービス係）

渡航の目的

- IDCC24参加による研究データ関連の情報収集
- スコットランドの大学へのインタビューによる即時OA対応先行事例の調査
 - 英国ではREF2021によりOAが実質義務化となっている
 - 研究助成機関（UKRI等）による即時OA義務化
 - 権利保持戦略の存在
 - 上記により英国のOA率は非常に高い

日程 (実際)

- 2/19 IDCC24 Workshop Day
(ワークショップへの参加 + Drinks receptionへの参加)
- 2/20 IDCC24 Conference Day
(発表の聴講 + Univ. Edinburghへのインタビュー調査)
- 2/21 IDCC24 Conference Day
(発表の聴講 ~~= Univ. Edinburghへのインタビュー調査~~)
- 2/22 University of Glasgow 訪問・インタビュー調査
- 2/23 ~~University of Strathclyde 訪問・インタビュー調査~~

- **IDCC : International Digital Curation Conference**

- DCCが主催するデジタルキュレーションの国際会議
- さまざまなステークホルダーが実践事例や研究成果を報告
- 今回（第18回）のテーマは「Trust through Transparency（透明性による信頼）」

- **DCC : Digital Curation Centre**

- Jisc等の補助を得て、2004年にデジタルキュレーションや長期保存の課題を解決するため設立
- デジタルキュレーションに関する研究、サービス開発、大学等への支援を行う

デジタル / データ キュレーションとは

- ・ 電子的な研究データの維持・保存・価値付加
- ・ デジタルキュレーションによる便益
 - 長期的な研究価値への脅威（改竄・不正利用, etc.）の軽減
 - 電子データの陳腐化のリスクの軽減（バージョン管理, etc.）
 - データの信頼性の担保
 - より広範な研究コミュニティでの共有・再利用の可能性
 - さらなる研究への発展性

会議プログラム

2/19	09:00 - 12:00	maRDMs	FAIR-IMPACT		PIDs
	13:00 - 16:00	maRDMs	UKRI	PIDs	FAIR・EOSC

2/20	10:45 - 12:15	Paper Session 1			
		DMPとデータスチュワード	キュレーションによる信頼と透明性の確立	キュレーションプロセス	
	14:15 - 15:45	Paper Session 2			
		キュレーションプロセス	共有と廃棄	信頼と透明性：基準と持続可能性	
	16:15 - 17:30	Lightning Talk Session 1			
		信頼のバランス、確立、取得	キュレーションプロセスの自動化への取組	透明性	

2/21	09:00 - 10:30	Paper Session 3			
		信頼と透明性：基準と持続可能性	プライバシーと透明性のバランス	キュレーションプロセス	
	11:00 - 12:30	Lightning Talk Session 3			
		研究とキュレーションによる信頼と透明性の確立	研究とキュレーションによる信頼と透明性の確立	普遍的な基準と特定の分野等に適応した基準への期待との間における信頼と透明性	

IDCC24 参加報告

【Workshop Day】

Elevating data management planning: Interoperability of RDM services through machine-actionability

- Tomasz Miksa, et al. Ten principles for machine-actionable data management plans.
<https://doi.org/10.1371/journal.pcbi.1006750>

午前：

- 上記 maDMPs 10原則から割り振られた 2 × 2 の原則について
「適当である」「現実的でない（削除）」「不足がある」等について議論

⇒ 基本的にすべての原則について、やや現実的ではない箇所があるものの、適当であるとの結論が各グループから出された

午後：

- DMPとRDMのプラットフォームをどのように結びつけるかについて
As a <stakeholder>, I want <something>, so that <something>
のかたちでユースケースを作成した

IDCC24 参加報告

【Paper Session】 Building trust and transparency into research and curation techniques, costs, and workflows
Cynthia H. Vitale. A model for calculating institutional costs for data management and sharing.

- 研究データ公開インフラのコストについての調査
- DMSについて費用を負担しているのは主に図書館とIT部局
- 図書館は人件費に最も費用をかけている
- IRをDMSプラットフォームとして使用することが研究者にとって最もコストが低い
- メソッドについて：<https://doi.org/10.29242/report.radsmethodology2023>
- DMSへの取組について：<https://doi.org/10.29242/radsdmsactivities2023>
- 支出についての調査：<https://doi.org/10.29242/report.radsexpense2024>

IDCC24 参加報告

【Lightning Talk】 Transparency

Lisa Olsso. Shining a light on data publication practices at Stockholm University.

- スウェーデンでは、2026年までにすべての研究データをFAIR原則に則った形式で（可能な限り）オープンにすることを目指している
- 2022年に査読付き論文の80%はOA化、オープンデータは39%程度
- なんらかのSuppl（補足資料）がついている論文は多い（46%）
- DOIが付与されたデータは10%程度

IDCC24 参加報告

【Lightning Talk】 Trust and transparency between universal standards and expectations and that of context and discipline specific adaptations
Federica Fina. The research data management journey of a postgraduate student at the University of St Andrews.

- St. Andでは2019年より博士課程院生にDMPを作成させるプロジェクトを開始した
- 2度のワークショップを行った後、D1の末でDMPを提出させる
- その後も研究データの公開や利用についてのワークショップを実施
- 学位論文と同時にその基礎データも提出を求める（提出率 82%）
- 理論的なワークショップはオンラインが、ライティングをともなうワークショップは対面が人気

IDCC24 参加報告

【その他の発表まとめ】

- 研究データの再利用者が、利用するデータについて追加テストを行うことは少ない
- データの質を上げるためにはキュレーター同士の連携とそのための可視化が必要
- 汎用リポジトリを比較してそれぞれの特徴を比較
- 信頼性の高いリポジトリは実際に信頼されているのか？どのようなときに不信などが生じるのか？
- 実験の再現ができない可能性が高い現状がある ⇒ データの透明性を促進する必要
- センシティブなデータを提供しているリポジトリの大半はリモート環境にデータを保存している
- 開発した質的データアーカイビングツールの目的と実際の利用について

IDCC24 参加報告

【Workshop Day】

Defining the criteria for assessing PID Policies and Services

- PID (Persistent Identifier ; 永続的識別子)
 - デジタルオブジェクトを恒久的かつ一意に識別するID
 - 様々な地域・分野・インフラ・サービス間における永続性・同一性を担保する必要性
- FAIR (Findable, Accessible, Interoperable, Reusable) 原則との関連性
 - EOSC (European Open Science Cloud) のPIDポリシー
 - EOSCによるCompliance Assessment Toolkit (CAT)
- PIDエコシステムにおける5つのアクターと役割
 - 枠組み・当局・サービス提供者・経営監督者・所有者
- 仮想のPIDサービスウェブサイトを用いてのグループディスカッション
 - 用意された良い例と悪い例を比較し、それぞれの問題点と改善点を議論
 - ポリシーやサービスに必要な要素や論点を具体的に確認

【Lightning Talk】 Balancing, building, acquiring trust

Monitoring for data access statements and open data at the University of St Andrews

- データアクセスステートメント
 - 〔データ利用可能な場合〕 DOI・データセットのURL
 - 〔データ利用不可の場合〕 その理由・アクセス要求方法
- データアクセスステートメントのモニタリング
 - Pure（機関データリポジトリ）のデータセットレコードで確認
 - 記述によってキーワードを付与して分類 → 様々な種類に対応可
 - 記述がない、不適切な場合はメールでリマインド → 非常に効果的
- データ公開率・データアクセスステートメント率ともに向上
 - データ公開率：66%（2021年） → 72%（2023年）
 - データアクセスステートメント率：34%（2017年） → 59%（2017年）

【Lightning Talk】 Balancing, building, acquiring trust

The data curation reading club as competence and trust builder

- 新入職員を含むRDMチーム全体の研修
- データキュレーションリーディングクラブを設立
 - 2週間毎に1時間半で開催
 - データキュレーションに関する記事について議論
 - 初めはコーディネーター、最近メンバーが記事を推薦
 - オンライン開催の他、ハイブリッドや対面も
 - 記事を読む時間がなくてもOK
- フィードバック（得られたこと）
 - データキュレーションの専門用語
 - 分野毎の重点の違いと専門家の見解
 - 同僚の専門性に対する理解と信頼性の高まり

【Lightning Talk】 Balancing, building, acquiring trust In sharing we trust. Taking advantage of a diverse consortium to build a transparent data service in Catalonia

- ・ カタルーニャの研究機関による大学サービスコンソーシアム
 - 22大学 + 22研究機関 [←13大学 + 40以上の研究機関 (1991年)]
- ・ DMPガイダンスの共同開発
 - eiNaDMP (無料のオンラインDMP支援ツール)
 - 1,300件以上のDMPsを作成、2,100名が利用登録
- ・ データリポジトリポリシーの共同作成
 - あらゆる分野に対応
 - 研究機関の関与により研究者の意見を反映
- ・ 協働の利点／課題
 - オーナーシップ・ガバナンス・専門性の共有／機関間のギャップ

IDCC24 参加報告

- 各プログラムのタイトルと抄録は以下のとおり
<https://www.dcc.ac.uk/idcc24/accepted-submissions>
- いくつかの発表資料は下記の Zenodo から参照可能
<https://zenodo.org/communities/idcc24/>

インタビュー調査

- [CWTS Leiden Ranking 2023](#)によると英国の大学はOA率上位をほぼ占めている（上位 18/20 が英国の大学。うち 6校がスコットランドの大学）

その理由として、

- REF2021がブロックグラントの算定の基準にOA率を入れている
- UKRIの助成条件として即時OAが課せられている
- 権利保持戦略の活用
- Jiscを中心とした出版社等との交渉

などが考えられる

インタビュー調査 (University of Glasgow)

- 創 立：1451
- 学部数：4 colleges (人文 / 生命 / 理工 / 社会), 23 schools
- 学生数：35,525
- 職員数：9,850 (うち教員：5,585)
- O A 率：91.4% (2018-2021) (うち GreenOA率：30.7%)

インタビュー調査 (University of Glasgow)

○ 即時OA義務化への対応

- 研究者は論文受理時に図書館にメールでそのことを知らせる必要がある。
Jisc Publication Routerを利用することもある
- 図書館ではリポジトリチームの6名を中心に通知された研究成果をIR上に登録している
- 通知のあった研究成果に対し、図書館では助成機関やREFの定めに準拠しているかを確認し、著者への助言を行ったり、定期的に報告書を作成したりしている
- OA義務化はUKRIのポリシー改正以前にBISからの要請によってはじまり、責任の所在やワークフローに大きな変化が起こった
- UKRIからの助成金は現状すべてOA化に必要な出版社への支払いに充てている
- ADCやAPCの値上がりは課題だが、できる限りのサポートは行いたい

インタビュー調査 (University of Glasgow)

○ DMP作成のサポート、研究データの管理など

- 基本的にはよく聞かれる質問などについてのガイドを作成することでサポートを行っている
- 特にDMPを作成時点でのサポート（研究データの保存先の決定など）に注力している
- 研究データ全体の95%程度はサードパーティのデータリポジトリに登録されており、大学のリポジトリに研究データを登録するのは大学院生がメイン
- 登録先の追跡やデータがどのように利用されているかの調査を今まきに行っている
- DMPマネジメントは大学院で必修 / RDMについてのEthical Reviewのトレーニングも重要
- データは助成機関の要請で10年ごとに見直しているが、ストレージの問題は大きい
⇒ データミニマイゼーションが喫緊の課題

インタビュー調査 (University of Edinburgh)

- 創 立：1583
- 学部数：3 colleges (人文社会 / 医学獣医学 / 理工), 21 schools
- 学生数：49,740
- 職員数：16,055 (うち教員：7,870)
- O A 率：90.8% (2018-2021) (うち GreenOA率：33.0%)

インタビュー調査 (University of Edinburgh)

○ 権利保持方針について

- [Research Publications & Copyright Policy \(2021\)](#)
- 方針策定のきっかけはUKRIによるOA義務化
- 学内の弁護士・学術コミュニケーションと協働し、弁護士の記載した文書で出版社へ通達
- 論文受理時にPure（研究者情報システム）にCC BYを付与した原稿を提出するよう依頼
- 出版社は現状本方針を受容しており、研究者へ当該方針に従う必要があることも連絡
- OA率は90%以上、4-5%程度は本方針の対象にも関わらずOA化できていないことが課題
- 権利保持方針によるOAの促進はうまくいっており、登録等の手間もあまりかかっていない

インタビュー調査 (University of Edinburgh)

○ その他の質疑

- ADCの新設などにはどう対応しているか？
 - ⇒ • 法外なAPCなどは助成金から支払えないと伝え、適切な出版社を選択するよう助言
 - 研究者が自身の研究費から支払った成果もPureに登録されるため把握はできている
- OA義務化にともなってストレージやIRの増強は行ったか？
 - ⇒ • OA義務化にともなったストレージの増設はあまりしていない
 - もともと研究データ用に15PB程度のストレージを持っていた
 - 公開用と機密情報用の2つのストレージがあり、極力前者の利用を推奨している
 - データ量の増大は今後の課題

インタビュー調査 (University of Strathclyde in Glasgow)

- 創 立 : 1796
- 学部数 : 4 faculties (工学 / 人文社会 / 理学 / 経営), 32 departments
- 学生数 : 24,213
- 職員数 : 4,535 (うち教員 : 2,010)
- O A 率 : 91.2% (2018-2021) (うち GreenOA率 : 49.8%)

○ 即時OA義務化への対応等

- OAの主担当は学術出版・研究データチームの7名
図書館の他チーム・研究室・研究科の代表者によるオープンリサーチチームも整備
- 研究者は研究情報システムPureを通じて、IRに研究成果等を登録
⇒PureとIRをつなぐコネクタを開発
- 研究者が登録したレコードは機関リポジトリ支援担当がチェック
- 研究成果の追跡にはPure、Scopus、WoS、Jisc Publication Routerを利用
- UKRIの助成金はAPCの支払いやRead & Publish契約の費用へ
- 権利保持戦略やR&P契約については、研究者からの反応が良好である一方、理解不足も
⇒研究者によるIFへのこだわりやワークフローの複雑化が原因という認識

- OA率は確かに高いが、ゴールドOAもかなりの部分を占めている → コストは？
 - 基本的には分野毎のデータリポジトリ、セーフティネットとしての機関リポジトリ
 - 権利保持戦略は組織的に
 - 研究者が助成機関と出版社の板挟みにならないように
- cf. 船守美穂 (2023) 「動向レビュー：即時オープンアクセスを巡る動向」
(<https://current.ndl.go.jp/ca2055>)
- 国際会議はコミュニケーションの場
 - 論文に加え、データが学術コミュニケーションのエコシステムを巡る争点に

所感

- 研究データ管理やDMPの作成は大学院で
- データキュレーションやデータリポジトリの知識習得が急務？
- 義務化すればOA率は上がる ⇒ 手間を増やさない工夫は必要
- 研究成果の収集はアクセプトのタイミングが一番いい？
- OA化にかかる資金の出所も検討する必要があるそう
- APCの値上がりやADC等への対応は……
- 学会や会議への参加は継続が大切（ぜひ応募を！）

