

Open Repositories 2017 参加報告

九州大学附属図書館・林豊
ORCID:0000-0001-7761-3444

2017年6月、オープンアクセスリポジトリ推進協会（JPCOAR）メタデータ普及タスクフォースの用務で、国際会議「Open Repositories 2017」（OR2017）に参加した。以下に概要を報告する。

1. 会議概要

概要	オープンアクセスリポジトリ等に関する国際会議（毎年開催、今回で12回目）
日程	2017年6月27日（火）～30日（金） ※6月26日にはプレイベントとして、クイーンズランド大学及びグリフィス大学のツアーが開催された（不参加）
会場	クイーンズランド工科大学ガーデンズポイントキャンパス（ワークショップ及びポスターセッション）、ヒルトンブリズベン（本会議）（オーストラリア・クイーンズランド州ブリズベン）
会議テーマ	Open : Innovation Knowledge Repositories
ハッシュタグ	#OR2017
参加人数	335人（27か国） ※日本からは筆者の他に、九州大学の池田大輔准教授、国立情報学研究所（NII）の林正治特任助教及び林洋平氏が参加
参加目的	日本の機関リポジトリのための新しいメタデータスキーマ「JPCOARスキーマ」に関するポスター発表、及びオープンサイエンスに関する最新動向に関する情報収集（メタデータスキーマ普及タスクフォースの用務）
ウェブサイト	http://or2017.net
公開資料	プログラム： https://or2017.net/program/ ポスター： https://or2017.net/electronic-poster-presentations/ プレゼン： https://or2017.net/presentation-template/ 動画： https://or2017.net/filmed-conference-sessions/

2. ワークショップ（6月27日）

本会議前日には、クイーンズランド工科大学ガーデンズポイントキャンパスの“Cube”と呼ばれる建物で、計11件のワークショップ（有料；昨年のOR2016では無料だった）が開催された。筆者は次の2つのワークショップに参加した。

2.1. ReDBox and Mint Workshop: Research Data Management for your Organisation

https://www.conftool.net/or2017/index.php?page=browseSessions&form_session=280

2000年に設立された非営利企業Queensland Cyber Infrastructure Foundation (QCIF)が開発している、研究データ管理のためのオープンソースソフトウェアReDBox、及び典拠管理ツールMintについてのワークショップ。ソフトウェアのインストールを含めた操作実習が行われたが、会場のネットワークの問題でダウンロードに非常に時間がかかりインストールまでたどり着けない参加者が大半だったため、Web版インタフェースで操作感を体験することになった。

<両ソフトウェアの概要>

ReDBox	<ul style="list-style-type: none">● Java ベースのオープンソースソフトウェア。2010年に開発が開始され、最新バージョンはv1.9（v2.0が開発中）。 http://www.redboxresearchdata.com.au/● QCIFが有償サポートを提供している。● 導入機関<ul style="list-style-type: none">➢ http://www.redboxresearchdata.com.au/users➢ http://libguides.newcastle.edu.au/c.php?g=265586&p=1775740● 主な機能は、データマネジメントプランの作成、研究データのメタデータの登録（研究者が登録し、管理者がレビューを行う）、Research Data Australia等へのパブリッシュ。● 研究データ自体は外部ストレージに保管する。● 様々なメタデータフォーマットに対応している。その一つとして RIF-CS が紹介された。 http://www.ands.org.au/guides/rif-cs-explained● メタデータ入力画面では地図上の範囲選択によって地理情報（geolocation）を入力する機能があった。
Mint	<ul style="list-style-type: none">● 著者名典拠や統制語彙を管理するツール。● ReDBoxのメタデータ入力画面で、“Name lookup”機能を使ってMintのデータを参照することができる。

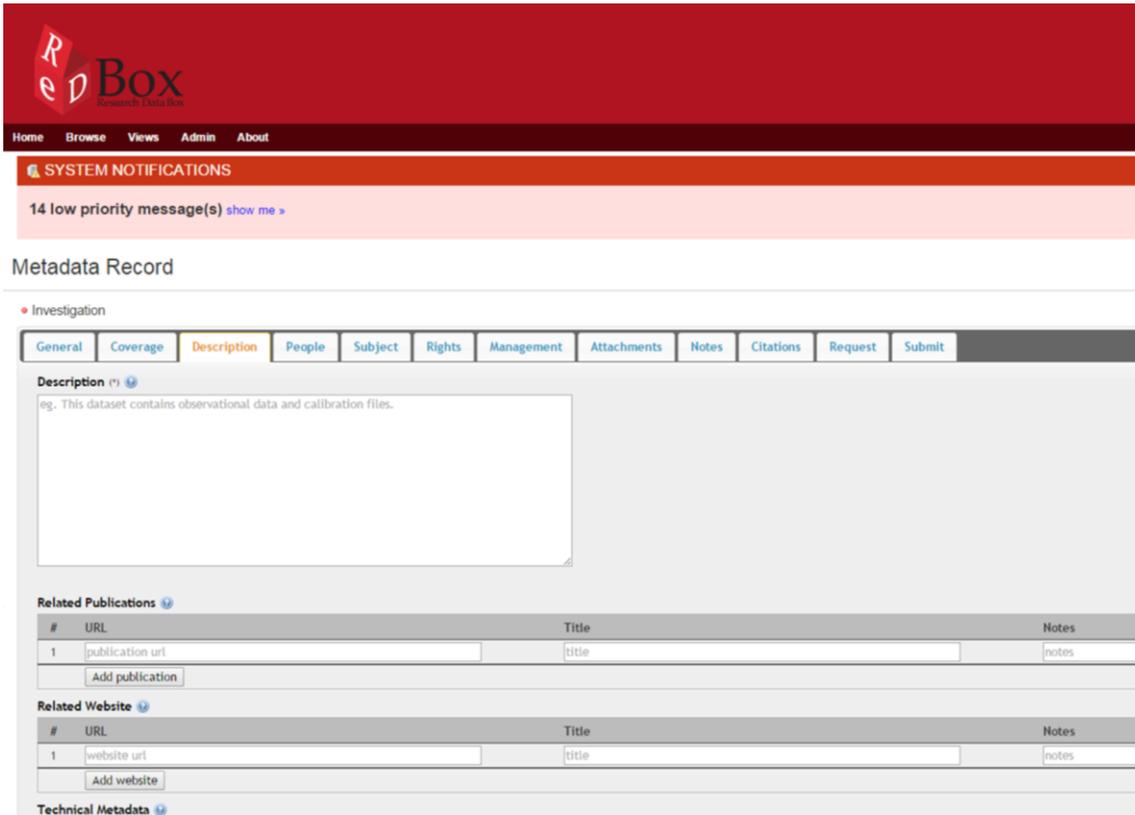


図 1 : ReDBox のメタデータ入力画面



写真 1 : ReDBox ワークショップ風景



写真 2 : クイーンズランド工科大学ガーデズポイントキャンパス (右手奥が “Cube”)

2.2. Introduction to the International Image Interoperability Framework (IIIF)

https://www.conftool.net/or2017/index.php?page=browseSessions&path=adminSessions&form_session=286

スタンフォード大学の Tom Cramer 氏ら IIIF コミュニティの中心人物による、IIIF の入門的なワークショップ。講義を聞きながら実際に手を動かす (URL のパラメタを修正してアクセスしてみる等) ことで API の意味を体感できるようになっていた。ワークショップで使用されたスライドは IIIF の基本がよくまとまっているため、ぜひご覧いただきたい。

<http://bit.ly/or17-iiif-1>

<IIIFの意義>

図書館等ではデジタル化資料等の「画像」 (image-based content) をウェブで公開しているが、その多くはサイロに閉じ込められてしまっており、必ずしも利用者が満足するかたちで利用できるようにはなっていない。画像をIIIFという国際規格に対応した方法(*)で公開することによって、例えば、利用者は自身の好みのビューワーを選択できるようになり、異なるサイトにある画像を同一画面内で比較したり、画像を重ね合わせて表示したり (バラバラになってしまった元は同一の資料や、同

じ地域を描いた現代の地図と古地図、等）、画像の一部分にアノテーション（注釈）をつけることが容易になる等、より自由に画像を利用しやすくなる。

(*)「IIIF対応」の明確な定義があるわけではないようだが、IIIF APIのうちImage API及びPresentation APIを実装することが最低限の条件と言えるようだ。

<IIIFとは>

ワークショップでは、IIIFはフレームワークであり、その意味は“A Community that develops Shared APIs, implements them in Software, and expose interoperable Content”であると繰り返し説明された。もちろんIIIFの中核にあるのはAPI群（後述）であるが、APIを実装したソフトウェア（Image API対応のサーバ、Universal ViewerやMirador等のビューワー、DSpace-GLAMのようなりポジトリでの実装等）、APIを通じて公開された相互運用性の高い画像（現在約3.35億件以上）、そしてこれらを実現する世界中のコミュニティをも含めた包括的な概念だということである。

<IIIFのAPI>

名称	概要	実装方法・実装例
Image API	画像を取得するためのパラメタ付きURLの仕様	ミドルウェア（IIP Image Server、Loris server等）の導入
Presentation API	画像のマニフェストファイル（画像のURL等のメタデータが記されている）のための仕様	マニフェストファイル（JSON-LDフォーマット）の作成・公開
Search API	画像のアノテーションの検索のための仕様（ビューワー内で「Ctrl+F」を実現するイメージ）	BL、Wellcome Trust、NCSU等の34以上の機関で導入 https://wellcomelibrary.org/item/b1802922x
Authentication API	画像のアクセス制限のための仕様	BL、Wellcome Library、Gettyで導入
A/V API（策定中）	音声・映像のための仕様	BLの“Save Our Sounds”プロジェクト（予定）

<IIIF 対応ビューワーの例>

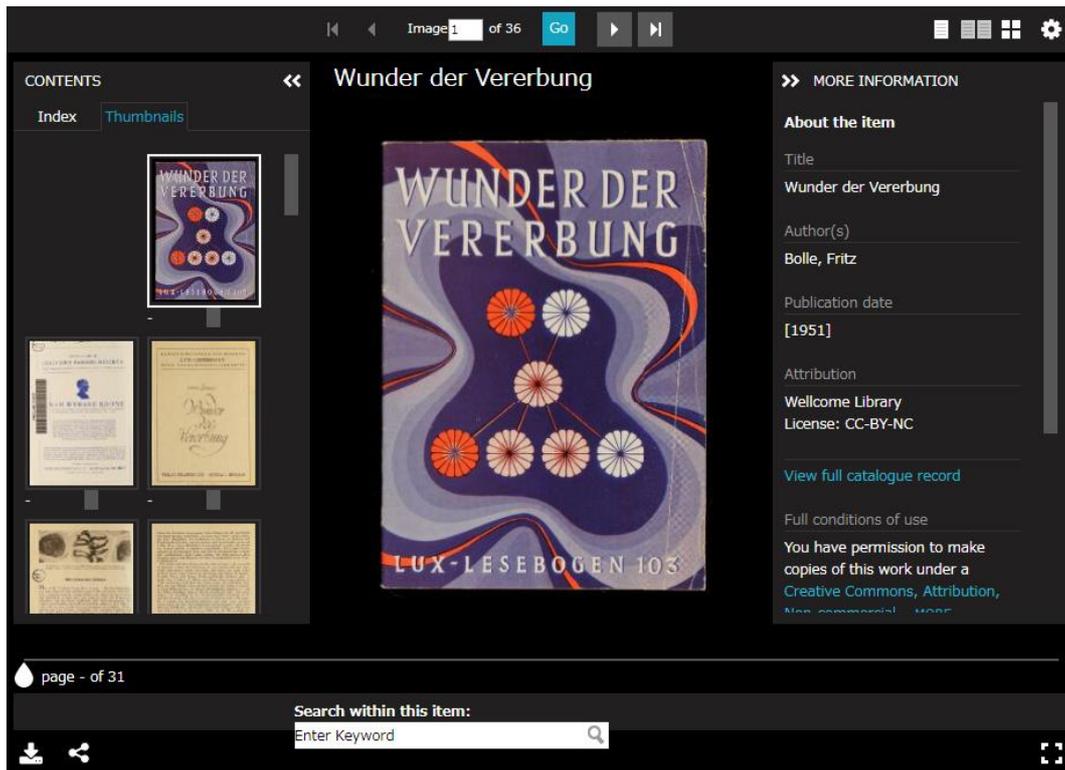


図 2 : Universal Viewer (Search API に対応した検索窓も備えている)

<http://universalviewer.io/examples/?manifest=http://wellcomelibrary.org/iiif/b17502792/manifest>

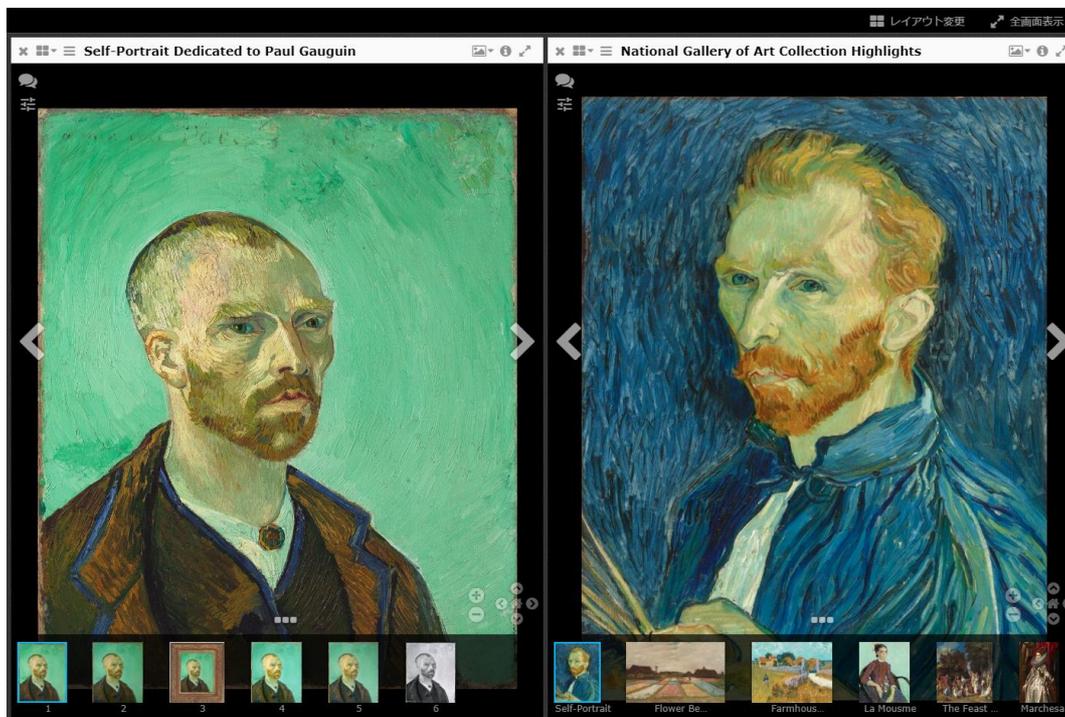


図 3 : Mirador (異なるサイトから複数の画像を並べて表示している)

<http://projectmirador.org/demo/>

3. ポスター発表（6月27日夜）

ワークショップ終了後、引き続き同じ建物内でポスターセッションが行われた（“1 Minute Madness”のようなプレゼンタイムはなし）。

筆者は、JPCOAR メタデータ検討タスクフォース（2016年度）及びメタデータ普及タスクフォース（2017年度）で開発中の、日本の機関リポジトリのための新しいメタデータスキーマ「JPCOAR スキーマ」に関するポスターを発表した。

<https://or2017.net/wp-content/uploads/2017/06/131.pdf>

今回は、会場壁面に広がるディスプレイにポスターの画像が表示される「デジタルポスター」形式が採用され、日本から紙のポスターを持参する必要がなかったのが大変ありがたかった。しかし、ポスターとポスターの間に全くスペースがないため、通常のようにポスターの前に立って説明することが難しかった。また、ポスターセッションの時間帯はウェルカムレセプション（ドリンクと軽食が振る舞われる）を兼ねており、会場はなかなかの喧騒であった。ポスターを見てくれている方がいたら、後方から見守りつつ、頃合いを見て“What do you think about my poster?”等と話しかけるといった感じになった。

合計で7名程度の方とお話しただろうか。ポスターで著者要素（jpcoar:creator）の記述例を紹介していたこともあり、会話の取っ掛かりとしてその説明を求められることが多かったが、複数の文字種で記述が必要になる点を興味深く感じていたようであった。残念ながらスキーマに対する具体的なフィードバックは得られなかったが、オーストラリアでも同じようにメタデータスキーマが課題になっているという話を聞いた。また、やはり日本の機関リポジトリの数の多さについては大変驚かれた（ときには「重要なのはリポジトリの数じゃない！」と張り合われたりもした）。

お話しした方のひとりに、筆者が2016年2月に機関リポジトリ推進委員会の用務として訪問したビーレフェルト大学のFriedrich Summann氏がいた。同氏は世界中のオープンアクセス文献のアグリゲーターであるBASEの開発者であり、訪問時には日本の機関リポジトリのメタデータの品質について要望を受けた。今回、「JPCOAR スキーマが普及すればあの時の要望に応えることができるようになる」と報告できたのは、とても嬉しく、誇らしいことだった。

日本からは他に、九州大学の池田大輔准教授が「Automatically Extracting Keywords from Documents for Rich Indexes of Searchable Data Repositories」というポスターを発表されていた。

<https://or2017.net/wp-content/uploads/2017/06/311.pdf>

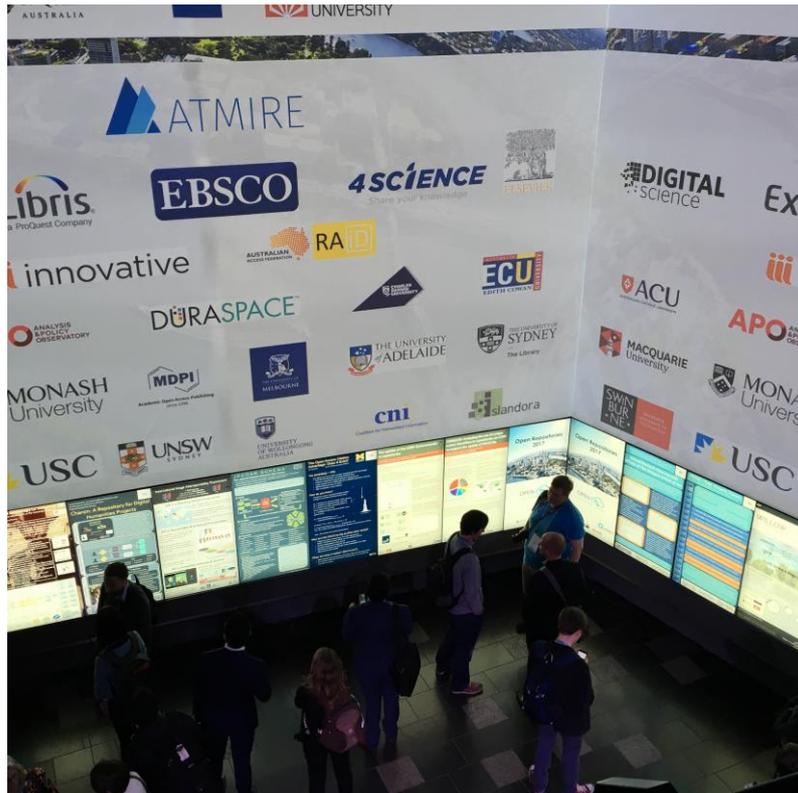


写真 3 : ポスターセッション会場 (左から 4 番目に筆者らのポスターが掲示されている)

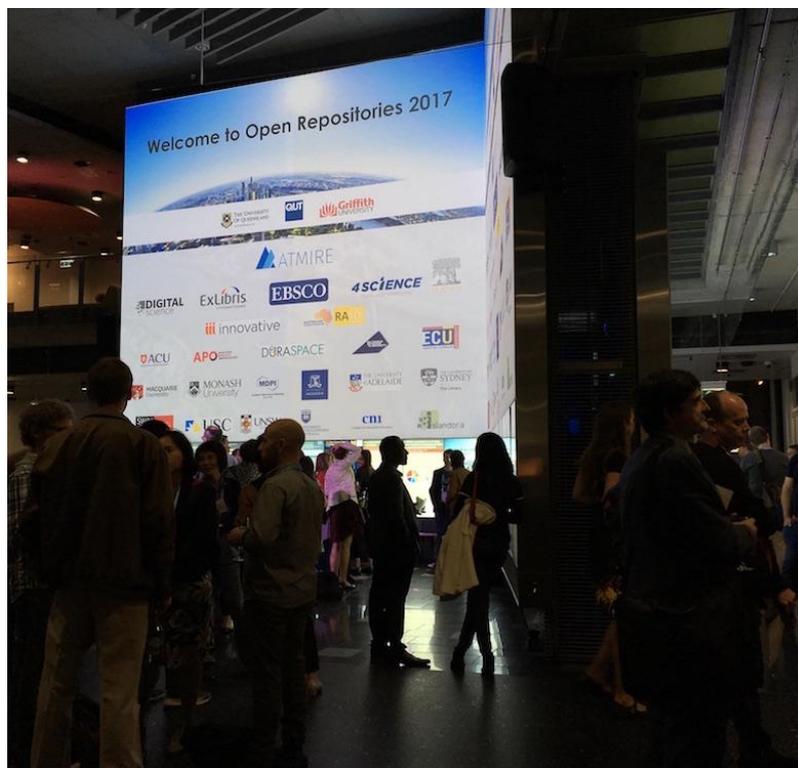


写真 4 : 3 つの面で計 49 件のポスターが展開された

4. 本会議 (6月28日~30日)

3日間の本会議は、フィールズ賞数学者として名高いケンブリッジ大学のウィリアム・ティモシー・ガワーズ教授の基調講演で幕開けとなった。彼は、Elsevierへのボイコット運動やarXivを利用したオーバーレイジャーナル「Discrete Analysis」の刊行等、オープンアクセス関連の活動でも有名である。

その後は、最大5会場のパラレルセッションで、計124本の発表が行われた。拝聴した全ての発表について紹介することはできないが、特に印象強く残ったものについて概要をまとめたい。詳しくは、公開されている各発表のスライドをご覧ください (動画があるものはURLを付す)。

4.1. Open Science as a Service: for repositories and research communities (Pedro Principe, 6/28, 24x7 Presentations 1)

欧州のOpenAIREによる、30か月間(2017年1月~2019年6月)の新しいプロジェクト「OpenAIRE Connect」の紹介。11のパートナー機関(CNR、ARC、ミーニョ大学、CERN、JISC等)を迎え、Research Community Dashboard(データプロバイダ向けのダッシュボード機能)及びCatch-All-Notification Broker(いわゆるブローカーサービス)を開発する。

<https://www.openaire.eu/connect>

4.2. DSpace-CRIS: new features and contribution to the DSpace mainstream (Andrea Bollini, 6/28, DSpace IG 1)

イタリアの4Scienceが開発したオープンソースソフトウェアDSpace-CRISの紹介。DSpaceのデータモデル(Researchers, Projects)を拡張してCRIS機能を追加したものである。最近の取り組みとして、Signposting対応(HTTPヘッダーのTyped Linksで関連する情報源を記述する)、ORCID連携(データの相互取込)、CKANアドオンによる研究データ対応、IIIF対応等が紹介された。DSpace-CRISの導入機関としては香港大学が有名だが、現在では世界中で80以上の導入事例があると聞き、その増加ぶりに驚いた。

<https://dspace-cris.4science.it/> (デモサイト)

4.3. Next Generation Repositories Session (6/28, General Track 7)

COARのNext Generation Repositories WG(日本からはNIIオープンサイエンス基盤研究センター・山地一禎センター長が参加)の紹介と現状報告。WGでは、主要なユースケースを洗い出した上で、必要機能、概念モデル、技術・アーキテクチャを検討した後、最終的には勧告(recommendation)をリリースし、その実装をサポートしていくという。

ここでも、リソースレベルでの相互運用性の好例としてIIIFに言及されていたのが非常に印象的であった。また、Signpostingについても紹介され、DOIリゾルブの際に論文著者のORCID iDを

「Link: <http://orcid.org/****>; rel="author"」と HTTP ヘッダーで発信する例が紹介された。

<https://www.coar-repositories.org/activities/advocacy-leadership/working-group-next-generation-repositories/>

【動画】 <https://www.youtube.com/watch?v=SsWdSxxEWS0>

4.4. Maximize the Visibility and Impact of Open Access and Other Articles through integration of publisher APIs (6/28, General Track 7)

次世代リポジトリのセッションの後半では、フロリダ大学と Elsevier のパイロットプロジェクトの紹介がなされた（会場は緊張感が漂っていた）。大学側は、ScienceDirect API を利用して（Atmire 社が DSpace 用のプラグインを開発）、学内研究者の論文のメタデータを機関リポジトリに取り込み、そこから ScienceDirect へのアクセスを提供する。他方、ScienceDirect 側では過去 3 年間の論文については著者最終稿を閲覧可能にするというものである（Embedded accepted manuscript という、ScienceDirect のサイトには遷移せずにポップアップウィンドウで著者最終稿を表示する機能もある）。著者最終稿のファイルが大学側に提供されるわけではなさそうだが、リポジトリでフルテキスト検索ができるようである。フロリダ大学は Elsevier にアクセス統計を提供するらしく、それがこのパイロットプロジェクトにおける Elsevier 側のうまみだろうか。

同様のパイロットをカタール大学が行なっている。フロリダ大学では CHORUS とも同種のパイロットを行なっている（千葉大学も）。

<http://ufdc.ufl.edu/ufirg>

<https://www.elsevier.com/solutions/sciencedirect/support/institutional-repository>

4.5. Open Access policy 3 years in, has researchers behavior changed? (6/29, General Track 13)

南オーストラリア大学の事例報告。2014 年に OA 方針を策定（アクセプトから 1 か月以内に著者最終稿をリポジトリに登録。出版から 12 か月以内のエンバーゴは許容）。2015 年から Ex Libris の Alma でリポジトリを運用しているという点が非常に興味深い。Scopus や WoS からメタデータを収集・登録することで研究者のメタデータ入力負担を軽減しているようだ。Alma の API を利用して学内開発したというメール送信機能を、最も重要なワークフロー改善だと評価していた。報告の後半では OA 率等の数字が披露され、例えば 2016 年出版の論文では 34%程度が OA、8%程度がエンバーゴという結果になっている。

<http://search.ror.unisa.edu.au/>

【動画】 <https://www.youtube.com/watch?v=060aplD1w-M>

4.6. The LA Referencia Software and the Brazilian Portal of Scientific Open Access Publications (oasisbr) (6/29, General 14)

IBICT（Brazilian Institute of Information in Science and Technology）の運営するブラジルの

リポジトリポータル oasisbr の報告。oasisbr は、ラテンアメリカのリポジトリネットワークである LA Referencia (OpenAIRE と連携している) のブラジルノードでもある。約 180 万件のコンテンツのうち、46%が Article で、続く 28%が修士論文というのが興味深い (博士論文は 9%、卒業論文は 5%)。LA Referencia にきれいなメタデータを送るためにバリデーションを行っているということで、oasisbr の管理画面が多数紹介された。

<http://oasisbr.ibict.br>

4.7. The repository as an interactive tool (6/29, General 15)

オランダのライデン大学の事例。DigiTool、Memorix、Gridminer、E-depot、DSpace 等のデジタルアーカイブシステムを Islandora に統合した。統合後のシステムは、OAIS、Data Seal of Approval、HORA (Higher Education Reference Architecture)、DERA (Digital Ecosystem Reference Architecture) に準拠しているという。Passive repository から interactive research tool への進化ということで、同システムを活用した、Sino-Malay literature のテキストデータマイニング、北朝鮮ポスターコレクションのアノテーション (IIIF) という 2 つのプロジェクトが紹介された。

<https://digitalcollections.universiteitleiden.nl/>

4.8. DSpace for Cultural Heritage: adding support for images visualization, audio/video streaming and enhancing the data model (Andrea Bollini et al., 6/29, DSpace IG 3)

イタリアの 4Science が開発しているオープンソースソフトウェア DSpace-GLAM の紹介。DSpace/DSpace-CRIS のアドオンであり、DSpace のデータモデルを拡張 (Persons, Families, Fonds, Events, Places, Concepts,) している。IIIF 対応したデジタル化画像の表示 (Universal Viewer ベース。Presentation API 及び Search API (hOCR) に対応し、Authentication API もまもなく対応) の他、アーカイブ資料 (ISAD(G)、EAD 対応) の取扱や、CKAN 対応によるデータ分析等の機能が紹介された。

<https://www.4science.it/en/dspace-glam/>

<https://dspace-glam.4science.it/>

5. Gala Dinner（6月29日夜）

本会議 2 日目夜には、ブリズベン川のほとりにあるクイーンズランド博物館を会場に、Gala Dinner（無料；昨年の OR2016 では有料だった）が開催された。閉館後のナイトミュージアムという格好で、ミュージアムショップで買い物をしたり、開催中の展示（Gladiators: Heroes of the Clossium 等）を無料で見ることもできた。アボリジニによる演奏やウォンバットに触れるようなエンターテイメントで歓迎していただいた。

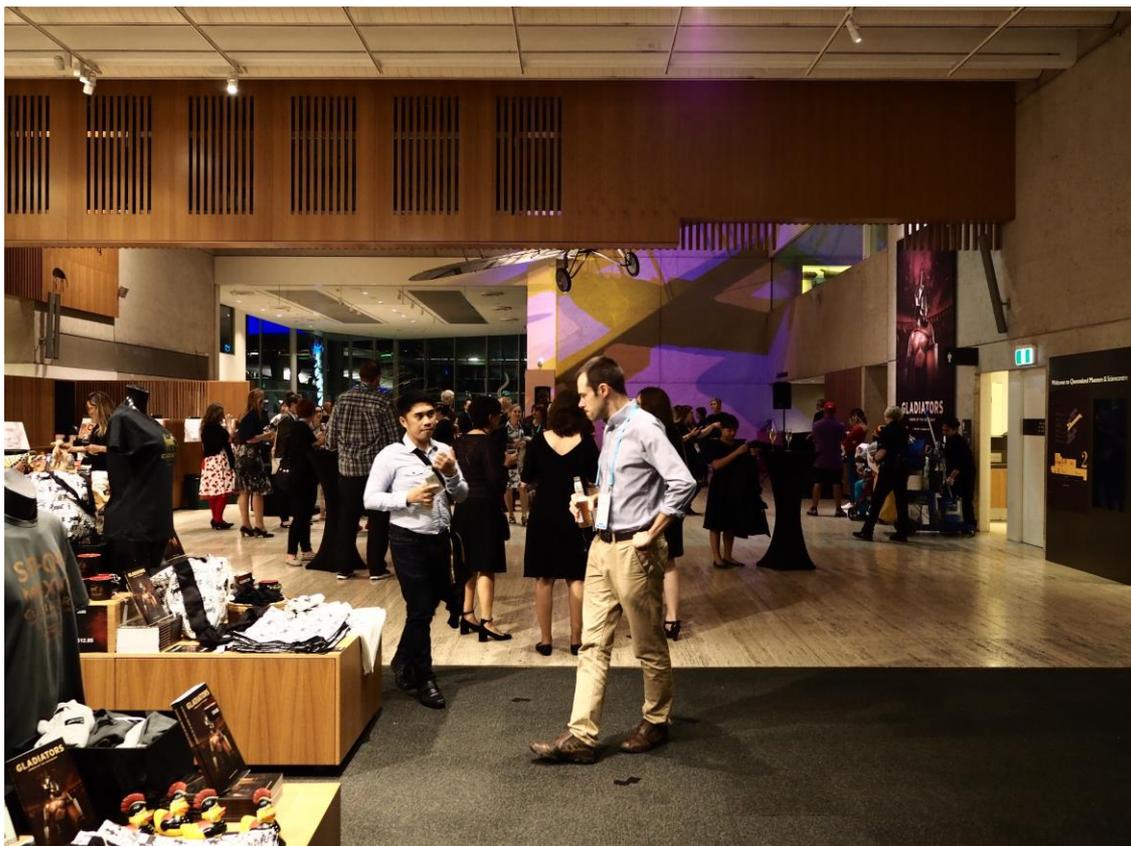


写真 5 : Gala Dinner@クイーンズランド博物館の風景（ミュージアムショップ付近）

6. 所感

当初は筆者が参加する予定ではなく、急遽代打としての参加ではあったが、同じ国際会議に2年連続で参加できたのは非常に貴重な経験になった。昨年お会いした方との再会も多く、あれっきりだったのにも関わらず顔を覚えていてくださったことに感激した。このような海外出張についてはなるべく多くの方にチャンスをとという考え方もある一方で、同じ人間が継続的に海外の場に出ていき交流の機会を重ねていくということの大切さが少し実感できた。

以下、印象に残っていることを3点述べたい。

(1) 一年前の OR2016 と比較して、文献、研究データ、そしてデジタル化画像（特に IIF）の3つのテーマが拮抗しているという印象を受けた。昨年は記憶に残らなかったデジタル化画像がリポジトリコミュニティで存在感を放っているのは衝撃であった。それを代表するのが 4Science 社の DSpace-GLAM だろう。そもそもリポジトリシステムとデジタルアーカイブシステムはデジタルアセットマネジメントシステムとしてくくることもでき、不思議なことは何もないのかもしれない。

(2) COAR や OpenAIRE の関係者の話を聞いていると、学術コミュニケーションが一極集中（寡占化）していくことに対する強い疑義と、リポジトリによって作られる世界に対する信念が堂々と語られ、はっとさせられる。リポジトリによる学術コミュニケーションは一極集中による問題点を回避できるが、一方で、分散したリポジトリは各々がサイロに陥りがちでもある。世界中に分散するリポジトリをどう連携させ、うまく組み合わせでひとつのグローバルなインフラを構築していくのか。Next Generation Repositories の議論の本質はそこにあると感じた。そのこと自体に新しさはないのかもしれないが、ここで議論される相互運用性は、従来のようにメタデータだけに留まらず、コンテンツ自身や、その上に広がるアノテーション等の付加価値をも射程に入れている。

(3) 日本では NetCommons をベースにした JAIRO Cloud の普及により、一般のリポジトリソフトウェアへの関心が薄れている。しかし、海外では DSpace、Fedora、Hydra 等のソフトウェアは継続的に進化を続けているし、DSpace-CRIS や DSpace-GLAM のように従来のリポジトリの役割を様々な方向へ拡張するような動きもある。こうした最新動向もウォッチしていくことが今後の JAIRO Cloud の成長にとっても必要だと感じる。日本ではナマの情報を得る機会が少ないため、OR のような場は貴重である。

今回の会議では、スケジュール管理用の iOS/Android アプリが提供されていた (CrowdCompass を利用)。アプリにはあらかじめ会議のプログラムが入力されており、数回タップするだけで自分のスケジュールに発表単位で登録できる。同じ時間帯の複数の発表を迷っているときはとりあえず両方登録しておける点など、全体的に使い勝手が良かった。

最後に、JPCOAR メタデータ検討タスクフォース（2016 年度）及びメタデータ普及タスクフォース（2017 年度）メンバー各位、とりわけポスター制作作業を中心的に担っていただいた岡山大学附属図書館の大園隼彦氏に感謝申し上げたい。また、出張事務をご担当いただいた JPCOAR 事務局及び国立情報学研究所（NII）学術コンテンツ課の皆さまにも感謝申し上げます。

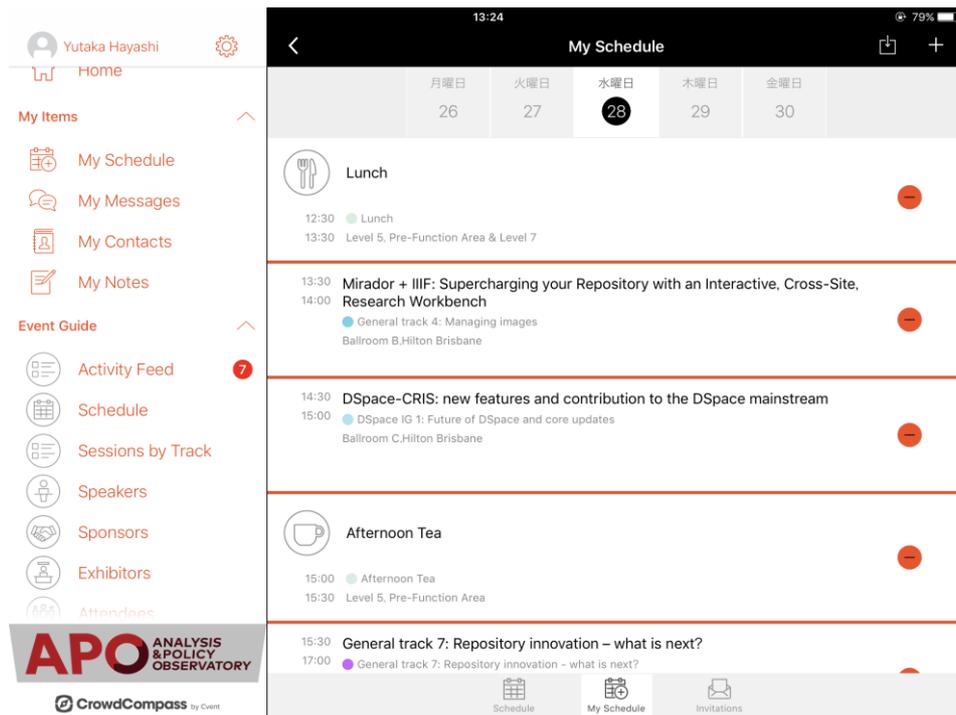


図 4 : スケジュール管理アプリ (CrowdCompass)

次回の Open Repositories 2018 は、2018 年 6 月 4 日～7 日に、米国モンタナ州ボーズマン (モンタナ州立大学) で開催される予定である。来年もまた JPCOAR としての成果を世界に披露できることを期待しつつ、筆者も貢献を重ねていきたい。■



写真 6 : 次回 OR2018 (というよりモンタナ州ボーズマン) の宣伝ブース

<http://www.or2018.net/>