

第3回 SPARC Japan セミナー2019

「実践 研究データ管理」

JPCOAR による研究データマネジメント人材の育成と 研究データに関する取組について

結城 憲司

(北海道大学附属図書館/JPCOAR 研究データ作業部会)

講演要旨



本報告では、オープンアクセスリポジトリ推進協会（JPCOAR）が取り組んできた研究データマネジメント人材の育成（RDM 教材、研修）の他、研究データに関する取組（データベースレスキュープロジェクト、RDM 事例形成プロジェクト、JPCOAR スキーマ等）について紹介する。その他、大学 ICT 推進協議会研究データマネジメント部会（AXIES-RDM 部会）や国立大学図書館協会オープンアクセス委員会における研究データに関する諸活動についても触れる。



結城 憲司

北海道大学附属図書館研究支援課長。2018年4月から現職。2018年4月からJPCOAR運営委員会委員、2019年4月からJPCOAR研究データ作業部会主査。その他、大学ICT推進協議会研究データマネジメント部会（AXIES-RDM部会）に参加。国立大学図書館協会（JANUL）オープンアクセス委員会には事務局として関与。

今回はオープンアクセスリポジトリ推進協会（JPCOAR）研究データ作業部会主査という立場で、JPCOAR が行ってきた人材育成と研究データに関する幾つかの取り組みについてご説明します。また、大学 ICT 推進協議会（AXIES）の RDM 部会の活動にも関与しており、国立大学図書館協会の OA 委員会でも研究データの取り組みを行っていますので、それについてもお話しします。

JPCOAR の取り組み

JPCOAR は、国公私立の図書館、主に大学図書館あるいはそれ以外の機関の図書館等が加盟している団体です。設立は2016年7月で、2020年1月14日時点で

は629機関が参加しています。主に大学等が使っている JAIRO Cloud という機関リポジトリのシステムを国立情報学研究所（NII）と共同運営しています。

2019年から2021年にかけて、「オープンアクセスリポジトリ戦略」として五つの戦略を立てています。その1番目の「オープンサイエンスの推進に寄与するため、研究データの公開、流通に関する先導的な取組を行う」が今回に関わるものです。こちらは私が担当している研究データ作業部会が行っています。

内閣府から2018年6月に出された「統合イノベーション戦略」は、オープン・アンド・クローズ戦略に基づき幾つかの提案がされており、それについて JPCOAR は図書館としてどのようなことができるかと

いうことで、三つの見解を述べています。まずは、JAIRO Cloud により研究データの公開に取り組めるのではないかと、また、これまで整備してきた JPCOAR スキーマ (次世代のメタデータのスキーマ) により研究データ等に対応していくことができること、それから、RDM トレーニングツールの開発・普及・利用促進についてもやっていけるのではないかとということです。

人材育成の取り組みとしては、教材開発を行い、三つの教材を作ってきました。最初に作った教材が「RDM トレーニングツール」です。こちらは研究データ管理に関する基礎的な知識を習得するためのツールとなっています。図 1 がその内容ですが、結構詰め込んだ形です。それまで日本語の教材がなかったので、初めて作ったツールということで、いろいろ弱点はあったのではないかと思います。

それらのツールをもう少し理解しやすくしようと作ったのが、「オープンサイエンス時代の研究データ管理」です。MOOC 用に公開された動画付きのトレーニングツールで、4 週にわたって受講する形で内容の再編を行っています (図 2)。

そして、「研究データ管理サービスの設計と実践」という研究支援者をターゲットとした教材も公開しました。それまでは、誰が対象者なのかよく分からない教材だというコメントを頂いていたので、そのあたりに力点を置いて作っています。特徴は、研究データのライフサイクルに基づいて教材を構成していることです。例えば図書館の職員なのか、URA の方なのか、研究推進の方なのか、情報系の方なのか、それぞれに

立場によってどこで研究者に関与できるのかという視点で教材が理解できるようになったのではないかと思います。図 3 がその中身です。研究前の支援、研究中の支援、研究後の支援、日常的な支援ということで整理しました。

これまでに作った RDM 教材は、土木学会の eラーニングに収録されたり、「統合イノベーション戦略」や「研究データリポジトリ整備・運用ガイドライン」の中で取り上げられたりしています。

ここからは、今年度行ってきたこととなります。まずは学認 LMS という NII が開発しているラーニング・マネジメント・システムで、二つの RDM 教材を提供し、昨年 9 月から 12 月にかけて 22 機関が参加して試験運用を行いました。その報告会を 12 月に開催し、いろいろなご意見を頂いたので、今後反映させていきたいと思っています。ただ、この学認 LMS のシステムは開発に若干時間がかかりそうということで、実際の運用はもう少し先になりそうです。また、参加機関による研究データ管理に関する取り組みの情報共有を行い、他館の状況を把握できたという別のメリッ

章	内容
第1章 導入	研究データ管理の重要性が増している背景や、研究データ、研究データ管理の定義について学ぶ。
第2章 データ管理計画 (DMP)	効果的なデータ管理に欠かせないデータ管理計画に関し、作成義務化の動向や、その構成要素について学ぶ。
第3章 保存と共有	研究データの研究期間中の保管や長期保存に関する留意点について学ぶ。また、研究データの共有に関して、その意義や検討すべき点、共有方法について学ぶ。
第4章 組織化、文書化、メタデータ作成	研究データを長期的に管理・活用するために欠かせない、一定のルールに則ったデータの組織化や、データについて説明する文章やメタデータの作成について学ぶ。
第5章 法・倫理的問題	研究データをめぐる著作権や、再利用を促進するためのライセンスの仕組みについて学ぶ。あわせて、センシティブデータを取り扱う上での留意点や、研究倫理についても学ぶ。
第6章 ポリシー	国や産学連携、機関、学会等が、研究データの保存や共有を求めるポリシーを策定する例が増えており、これらの動向及びポリシー策定の要件について学ぶ。
第7章 研究データ管理サービスの検討	学習者が自機関での研究データ管理サービスを構築していくためのステップを学ぶ。

(図 1)

週	内容
第1週: 研究データ管理とは	<ul style="list-style-type: none"> 本講義の背景とねらい 研究データ管理の重要性が増している背景 研究データ、研究データ管理の定義 効果的なデータ管理に欠かせないデータ管理計画 データ管理計画をめぐって国内外の動向 データ管理計画の実践
第2週: 保存・共有・文書化	<ul style="list-style-type: none"> 研究データの保管や長期保存・共有・再利用 研究データの組織的な管理 データを文書にまとめる
第3週: メタデータ・法倫理的問題	<ul style="list-style-type: none"> データにメタデータを付与する データの引用 研究データをめぐる著作権再利用促進のためのライセンス 研究不正・研究倫理 センシティブデータ
第4週: 運用に向けて	<ul style="list-style-type: none"> ポリシーとは 研究データ管理サービスとは? サービスの設計～学習者が自機関での研究データ管理サービスを構築するためのステップ

(図 2)

章	内容
第1章: 序論	研究データ管理に関する基礎的な知識を学ぶ ・ 研究データ管理とは何か、研究データ管理の重要性が増している背景 ・ 研究データ管理サービスとは何か
第2章: サービス設計	自機関の実情に合ったサービスを設計するために必要なことを学ぶ ・ 組織づくりや機関としての組織やポリシーの立案 ・ Data Asset Framework (データ資産フレームワーク) の活用
第3章: 研究前の支援	研究者が研究を開始する前段階で、どのような支援が求められているかについて学ぶ ・ 研究データ管理サービスとは何か、研究データ管理サービスが求められるDMP (データ管理計画) の作成支援
第4章: 研究中の支援	研究者が研究を実施している最中に求められる支援について学ぶ ・ 研究データの組織化やメタデータの作成、データの分析や可視化の支援 ・ 研究データの保存や共有の支援、研究倫理について学ぶ
第5章: 研究後の支援	研究終了後に、得られたデータを公開することについて学ぶ ・ データの公開に関する留意点や、データの公開と研究データの可視化の支援 ・ 論文発表の際に必要とされるデータの取扱い、データ管理計画の可視化
第6章: 日常的な支援	日常的な研究データ管理サービスについて学ぶ ・ 研修の実施方法、ポータルサイトの構築と更新すべき情報、日々の業務の実施体制 ・ 信頼性の高い利用状況やアクセスログの活用

(図 3)

トもあったと思っています。

もう一つ、RDM 教材の取組みとして、若手研究者向けの教材作成を行っています(図4)。受講した方は分かると思いますが、既存の教材は受講にかなり長い時間がかかる教材で、あれを研究者の皆さんに受けてくれというのは少し無理があると思います。また、研究者の皆さんであれば、ピンポイントで知りたいことが学べればいいわけです。例えば研究データの引用をどうすればいいのか、データのマネジメント計画をどのように作ればいいのかということをピンポイントで教えてくれる教材がないと先生方に受けてくれとは言えないので、もう少し分割して若手研究者向けに作ってみようという取組んでいます。千葉大学の方で先行して開発中の教材があるということで、そちらを参考に作っています。今年3月には試行版を作成し、来年度中にブラッシュアップして公開できるようにしていきたいと思っています。

また、JPCOAR では研修事業を行っています(図5)。こちらは JPCOAR の参加機関の方が対象です。来年度の計画としては、基礎研修としてオープンアクセス新任担当者研修(仮称)を行う予定です。これまでも JPCOAR では機関リポジトリ新任担当者研修を行ってきましたが、もう少し幅広に内容を盛り込み、年に2回程度開催したいと思っています。それから専門研修としては、今年10月から新 JAIRO Cloud が稼動することになっているので、操作説明会を秋以降、全国5会場程度で開催する予定です。JPCOAR セミナーは、来年開催できるかどうか不確かですので、参考程度にご覧ください。

若手研究者向けの教材作成

- 今回の教材のターゲットは若手研究者
- 既存の教材は受講に時間がかかりすぎ、研究者は利用しにくい
- 既存のRDM教材を分割して(マイクロコンテンツ化)、若手研究者向けに再構成
- 先行している千葉大学で開発中の教材等も参考に
- 2019年度の目標は
 - 現在の支援者向け教材をベースに、研究者が行うタスクと支援者が行うタスクを仕分ける
 - 他の「研究者向け教材」から、現在の支援者向け教材に不足しているコンテンツを洗い出す
 - 2020年3月末までに試行版を作成予定

(図4)

JPCOAR の研究データに関する取組みを幾つか紹介します。図6の「失われていく過去の研究データ」は、研究データ作業部会のメンバーである京都大学の天野さんが調査したものです。この調査によると、科研費等の助成を受けて作成されるデータベースのうち、半分ぐらいが適切に継続管理されていないということです。既に失われてしまっているようなデータもあり、そういったデータを救済する必要があるのではないかとこのコンセプトの下に考えているのが、データベースレスキュープロジェクトです。管理していた研究者がいなくなると、どうしても継続が困難になるということがあります。大学によってはデジタルアーカイブが受け皿になる可能性もありますが、機関リポジトリがその受け皿になる可能性を模索するプロジェクトです。2019年度の活動として、作業部会員の所属する機関においてこのプロジェクトを進めました。ただ、実際には公開まで持っていくには少々時間がかかるというのが正直なところです。こちらからアプローチしても、そのデータがどこかのハードディスクにはあるのだけれどもどこかよく分からず、なかなかデータの

2020年度のJPCOARの研修(計画中)

- JPCOAR参加機関対象の研修
- 基礎研修
 - オープンアクセス新任担当者研修(仮称)
 - これまで開催してきた「機関リポジトリ新任担当者研修」をリニューアル
 - 学術情報流通の概論、OA出版、RDM、オープンサイエンスの動向等の内容を盛り込む
 - 年2回開催予定
- 専門研修(他にも検討中)
 - JPCOARセミナー
 - JPCOARスキーマ、OAI-PMH、IIF、DOI、ORCID等の学術コミュニケーション要素技術の解説、夏頃開催予定
 - 新JAIRO Cloud操作説明会
 - 秋以降、全国5会場程度開催予定

(図5)

失われゆく過去の研究データ

URLが生きている/移動している: 47-62件 (9-7件 行方不明)

メンテナンス状態: 24-34件 (19-21件 不安定, 1-2件 不可用, 3-5件 非オープン(OAのみ、有料、要登録))

「KAKEN: 科学研究費助成事業データベース」で平成21~23年度採択分(文系)を調査した結果

2016年及び2019年に調査 (京都大学・天野氏による調査)

科研費の助成を受けて作成されたデータベースの半分が適切に継続管理されていない

(図6)

救済には至らないケースもあります。

図 7 は過去に取り組んだ事例です。「矢内原忠雄文庫植民地関係資料データベース」は琉球大学の事例です。これは 2017 年度に JPCOAR の研究データタスクフォース（当時）として取り組んで公開されたものです。「北の息吹：日本の魅力的なワイルドフラワー500種」は本学の事例です。部局のホームページで公開していた研究データが、その先生がいらっしゃらなくなる予定のため公開の維持が難しいということで図書館に相談に來られ、機関リポジトリが受け皿として公開した事例です。「ピラネージ画像データベース」は機関リポジトリというよりはデジタルアーカイブでレスキューした、東京大学の事例です。

次に、RDM 事例形成プロジェクトの取り組みをご紹介します。こちらは JPCOAR と、後で述べる AXIES-RDM 部会との連携プロジェクトになります。目的は RDM に関する事例を実際に積み重ねていくことで、2019 年度に参加したい機関を募集し、2020 年、2021 年と 2 年間かけて蓄積していくプロジェクトです。前に述べた教材でも、日本国内の事例が少なく海外の事例ばかりではないかということをおっしゃる方がいらっしゃいました。まだ大学の事例が多くないので、これを積み上げていき、次に続く大学が参考にできるようにものを公開していきたいと考えています。JPCOAR としては、研究データの公開、発信等の部分を中心に担当する予定です。

もう一つの JPCOAR の取り組みとして、「JPCOAR スキーマ」があります（図 8）。現在、日本の機関リポジトリでは「junii2」という、どういった基準でデー

タを記述するのかということを決めているメタデータスキーマがあるのですが、その次世代のものを開発して、2017 年 10 月に公表しています。目的はオープンサイエンスおよびオープンアクセス方針への対応で、研究データに対応したメタデータスキーマとなっています。これまでの「junii2」では、論文に必要なメタデータはあったのですが、研究データとなるともう少し幅広い項目が必要になってきますので、それらを追加したり、国際的な相互運用の向上やメタデータ構造の修正を行うなどしています。JPCOAR スキーマへの対応を始めており、ジャパンリンクセンター、CiNii Articles、医中誌、新 IRDB、次に今年 10 月に公開する新 JAIRO Cloud という形で公開が順次進んでいます。今後は JPCOAR スキーマが主流が変わっていくのではないかと考えています。

その構造を表したのが図 9 の図です。各国内の機関リポジトリから IRDB の方に JPCOAR スキーマでデータが送られ、各種サービスに提供されて、各利用者が利用できるようになります。

(図 7)

(図 8)

(図 9)

AXIES-RDM 部会の取り組み

次に、AXIES の RDM 部会の取り組みについてお話しします。この部会のメンバーである NII の船守先生から、こちらの団体は情報系のセンターの教員が中心の組織なので図書館の人も入ってもらいたいという話があり、私が入らせていただいて取り組みに協力しています。

AXIES は情報系のセンターの教員を中心に運営されている組織で、その中に研究データマネジメント (RDM) 部会があります。主な取り組みとしては、2019 年 5 月に「学術機関における研究データ管理に関する提言」を公開しており、2020 年 1 月に「大学における研究データ管理に関するアンケート (雛形)」の公開をしています。

「学術機関における研究データ管理に関する提言」については、図 10 にあるような資料が公開されています。研究データの管理・公開について非常に具体的に書かれているので、いろいろ疑問に思っていた部分に分かりやすくなっていると思います。例えば、研究データを誰が管理するのかという話については、それは研究者だということが明確に書かれており、それを管理するための環境等を機関としてきちんと提供しないとけないということが書かれています。管理という言葉にすると、教員は上から全部管理されるのかと拒否反応を示す場合が多いので、その辺は丁寧に説明する必要があるのではないかと思います。また、公開に関してもいろいろ誤解がありますので、その辺でも参考になる資料ではないかと思います。

もう一つの「大学における研究データ管理に関する

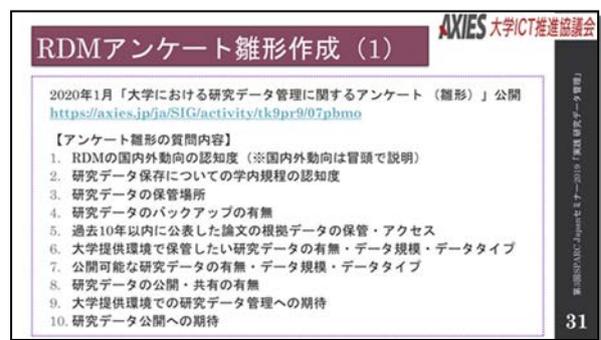


(図 10)

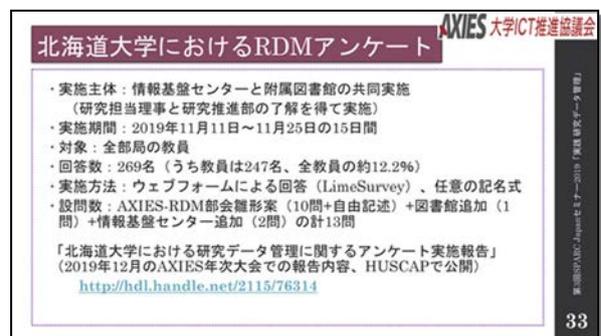
アンケート (雛形)」の作成が、私がメインで協力している取り組みです (図 11)。つい先日、1 月下旬に雛形が公開されました。質問は 10 問程度と、割とシンプルにまとめています。これ以上質問数が多くなると回答率が下がるという意見があり、必要最小限のものとして 10 問に絞っています。こちらは元々、名古屋大学で実施されたものを多少修正して、幾つかの大学でアンケートを実施させていただきました。

2019 年 8 月に雛形案を作成、2019 年 10~12 月に京都大学、広島大学、大阪府立大学、本学でアンケートを実施、2019 年 12 月に福岡で開催された AXIES 年次大会でアンケートの実施結果を報告、2020 年 1 月に雛形公開、2020 年 3 月に全体的な結果を情報処理学会 IOT 研究会で報告というスケジュールで取り組んでいます。

北海道大学で実施した事例が図 12 です。情報基盤センターと附属図書館が共同で、研究担当理事や研究推進部の方にも話を通した上で実施しています。実施期間は昨年 11 月 11 日~11 月 25 日の 15 日間です。全部局の教員を対象としており、回答数は 269 名 (うち



(図 11)



(図 12)

教員は247名、回収率は全教員の12.2%)でした。ウェブフォームで実施したのですが、AXIES-RDM部会の雛形の10問の他に、図書館で1問、情報基盤センターで2問追加しました。

特徴的なのは、任意の記名式にしたことです。なぜかという、これを入り口としたいと思っていたからです。名前を挙げてくれる方には個別にお問い合わせすることもありますということを明記した上で、来年度インタビュー調査を行いたいと思っています。その中で研究データに関心をお持ちの先生方に話をし、実際に研究データの公開等につなげていければと思っています。アンケートの集計結果については本学の機関リポジトリ HUSCAP で公開していますので、そちらを見ていただければと思います。

簡単に説明しますと、研究データの10年保存ルールについては、65%の方が知っていると回答しており、割と認知度が高かったです。科研費の申請をするときに研究倫理教材を受けさせられるので、その中で認知度が上がっているのではないかと思います。

研究データの保管場所については複数回答を認めているのですが、一番多いのはやはり個人のパソコンで88%、研究室のサーバーやストレージが49%でした。

機関によるRDM環境が必要だと回答した方は62%でした。それに対して、研究データの公開ニーズについては32%の方しか必要だと回答しておらず、やはり公開に関しては少し抵抗感が強いということが分かりました。これに基づいて、個別の先生にいろいろ聞いていきたいと思っています。

国大図協 OA 委員会の取組み

続いて、国立大学図書館協会(国大図協)のオープンアクセス委員会の取組みを説明します。国大図協は、全国86の国立大学、放送大学、大学共同利用機関等の図書館を会員とする団体です。ここでもオープンサイエンスや研究データ管理に関する取組みを行っており、2019年3月に取組みの文書を公表していますし、6月の岡山の総会研究集会において、「大

学図書館のオープンサイエンスに向けた取組み」というテーマで報告および議論を行っています。そして2020年1月に「オープンサイエンスの推進に向けた協会の行動計画」を公表しました。この中で、まず取り組むこととして、アドボカシー活動の推進、人材の育成、先導的事業の推進について、短期と中期に分けて計画しています。

国大図協の中にオープンアクセス委員会というものがありません。こちらはオープンアクセスだけではなく、オープンサイエンスの研究データのことを扱っているので、活動を少し紹介します。

昨年度、機関リポジトリの再定義を行いました。「統合イノベーション戦略」の中で機関リポジトリという言葉が急に出てきたわけですが、当初、大学図書館の皆さんにとっては、機関リポジトリはそれぞれの大学の研究者の皆さんの論文、あるいはそれぞれの大学で発行している紀要、あるいは博士論文を公開する場所だという認識でした。しかし、「統合イノベーション戦略」で求められているのはもっと幅広いもので、通常、商業流通されている学術文献、あるいは大学の活動で生み出される学術文献以外に、研究データやデジタル化された研究資料があります。これまでデジタルアーカイブとして図書館がいろいろ公開してきた画像や資料も機関リポジトリの中に含まれてもいいのではないかと考えられます。その辺の線引きはいろいろシステム的な問題があるのですぐには統一できないと思いますが、将来的には分け隔てなく収録していく必要があるのではないかと考えた下、機関リポジトリの再定義を行いました。

もう一つ、昨年取り組んだ事項として、「オープンサイエンスに向けての国立大学図書館が担う具体的な役割」というものがあります。内容としては、①各会員館の役割、②会員館共通の取組み、③協会としての取組みという三つのレベルを設けて、それぞれどんなことに取り組んでいくのかということを一覧記しました。特に①会員館の役割については、各大学の事情や規模が違うので、こういうことは取り組めるのでは

ないだろうかということを一覧した上で、それを取捨選択し、それぞれの大学で取り組んでいただく形にしております。

その背景を説明したのが図 13 の図です。大学の図書館の皆さんに、もっとオープンサイエンスに向けて取り組んでほしい、こう変わっていくのだということを図示しています。旧来は大学図書館というと資料の収集や保存、提供など紙ベースのものを中心に行っていて、その後、電子ジャーナルが出てきたら電子ジャーナルに対応していたと思います。ただ、2005 年ごろから、それだけでは駄目で、オープンアクセスの推進や機関リポジトリの取り組みを行う必要があるだろうということで、それぞれの大学の研究成果を公開する取り組みを行ってきました。そして今度は、研究データのことも段階を踏んで行っていく流れとなっています。

今年度の取り組みについて簡単に説明します。一つは、「研究者に対する研究データ調査の項目リスト（案）」の作成です。大学図書館がこのオープンサイエンスや研究データ管理に取り組む上で何ができるかということは、このオープンアクセス委員会の中で議論になりました。研究者の皆さんで、オープンサイエンスにそこまで取り組まれている方はまだ多くはないと思いますが、まずはその実態を知ろうということです。方法としては、①アンケート調査、②個別インタビュー調査という 2 段階を想定しています。以前、オープンアクセスをやっている時代にもこのような取り組みをやってきたので、それを参考に組み立てています。



(図 13)

①アンケート調査については、AXIES の作ったアンケートの雛形を利用推奨しています。こちらを使って次の個別インタビュー調査につなげるような形で名前をお聞きした上で、研究データに関心を持っているような研究者を捕捉します。②個別インタビュー調査については、AXIES の雛形のようなものがないので、オープンアクセス委員会の方で雛形を用意して公開する予定です。

もう一つ必要ではないかと議論になっているのが、教員向けに研究データのオープン化のメリットを示すような資料です。インタビュー調査に行ったときに必要ではないかということです。そこで、「研究データのオープン化の具体例とそのメリット（仮称）（案）」を作っています。ただ、論文を公開することのメリットは割とスムーズに説明できるのですが、研究データを公開することのメリットとしては、引用されることや、科学のさらなる発展につながることも言えると思いますけれども、実感としてメリットがあるということをどれだけ研究者の皆さんに伝えられるかというのは難しいところです。今たたき台を作っているところですが、難航しています。また、もう一つこれで行いたいと思っているのは、誤解を解くことです。先ほど言いましたように、研究データ管理というと、管理されてしまうのかとか、オープンというと、全部オープンにしなければいけないのかなどと思っている先生もいらっしゃるの、そのあたりの誤解も解きたいと思っています。こちらについては 3 月末までにまとめる予定です。

今年度のこの二つの取り組みについては、公開までにはもう少し時間がかかるかと思えます。公開したら国大図協でお知らせしたいと思っています。

●林 JPCOAR の取り組みをいろいろご紹介いただきました。私は国の研究所にいるのでこれに入っているのですが、JPCOAR の会員機関として国の研究機関は

どのくらい入っているのでしょうか。

●**結城** 機関数は把握していませんが、作業部会にも研究所の方に何人か参加していただいています。どうしても大学を主体として話をしてしまうので、大学とは違った視点でいろいろお話を頂けるのは非常にありがたいです。特に研究と近いところにいらっしゃる人が多いので、大変参考になる意見を頂いています。