

研究データへの DOI 登録実験プロジェクト 報告書

2015年10月8日

ジャパンリンクセンター
研究データへの DOI 登録実験プロジェクト

目次

1. 概要.....	1
2. 公募による参加機関募集.....	2
2.1 公募予告.....	2
2.2 公募説明会.....	2
2.3 公募.....	2
2.4 選考.....	3
3. プロジェクト実施方式の検討.....	4
3.1 プロジェクトの体制.....	4
3.2 プロジェクト計画のレビュー.....	5
3.3 プロジェクト実施スケジュール.....	5
3.4 各参加機関の概要、実施体制、準備段階で認識している課題の共有.....	7
3.4.1 独立行政法人 科学技術振興機構.....	7
3.4.2 大学共同利用機関法人 情報・システム研究機構 国立極地研究所.....	10
3.4.3 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所（地球環境情報統融合プログラム（DIAS-P））.....	14
3.4.4 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 学術基盤推進部.....	21
3.4.5 産業技術総合研究所.....	23
3.4.6 情報通信研究機構.....	25
3.4.7 千葉大学附属図書館.....	29
3.4.8 独立行政法人 物質・材料研究機構.....	33
3.4.9 理化学研究所脳科学総合研究センター神経情報基盤センター.....	36
4. DOI テスト登録.....	39
4.1 DOI テスト登録まとめ.....	39
4.2 DOI テスト登録要領.....	39
4.3 DOI テスト登録計画.....	45
4.4 DOI テスト登録計画の検討・承認.....	48
4.5 DOI テスト登録実施報告.....	49
4.6 DOI テスト登録を通じて抽出された課題のガイドラインへの落とし込み.....	130
4.7 DOI テスト登録を通じて抽出された課題のシステム面での対応.....	130
5. DOI 本登録.....	135
5.1 DOI 本登録要領.....	135
5.2 DOI 本登録計画.....	136
5.3 DOI 本登録計画の検討・承認.....	138

5.4 DOI 本登録実施結果.....	138
6. 総括.....	139
6.1 国立研究開発法人 科学技術振興機構.....	139
6.2 大学共同利用機関法人 情報・システム研究機構 国立極地研究所.....	142
6.3 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所.....	147
6.4 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所.....	150
6.5 国立研究開発法人 産業技術総合研究所.....	152
6.6 国立研究開発法人 情報通信研究機構.....	155
6.7 千葉大学附属図書館.....	159
6.8 国立研究開発法人 物質・材料研究機構.....	162
6.9 国立研究開発法人 理化学研究所 脳科学総合研究センター神経情報基盤センター	166
7 会議体開催実績.....	169
7.1 ミーティング等開催状況.....	169
7.2 対話・共創の場.....	171
7.3 報告会.....	172
8 プロジェクト参加者名簿.....	173
9 附録資料.....	174
9.1. 附録資料 1 研究データへの DOI 登録実験プロジェクトについて【予告】.....	174
9.2. 附録資料 2 公募要領.....	174
9.3. 附録資料 3 「研究データへの DOI 登録実験プロジェクト」参加機関募集 公募説明 会 質疑応答内容.....	174
9.4. 附録資料 4 各参加機関から提出された「DOI テスト登録計画」.....	174
9.5. 附録資料 5 各参加機関から提出された「DOI テスト登録計画」.....	174
9.6. 附録資料 6 JaLC における研究データの登録手順書.....	174
9.7. 附録資料 7 各参加機関から提出された「DOI 本登録計画」.....	174

1. 概要

ジャパンリンクセンター（以下「JaLC」という）は、国内の各機関が保有する電子的学術コンテンツに DOI を登録するサービスである。2012 年 3 月に DOI 登録機関（Registration Agency ; RA）に認定されて以来、主にジャーナル論文に対して DOI の登録を行ってきた。2014 年 12 月に JaLC 新システムをリリースし、ジャーナル論文に加えて、ジャーナル、予稿集、研究報告書、書籍、研究データ、大学コース(e-learning) へと DOI 登録コンテンツを拡大した。

このうち、研究データへの DOI 登録に関しては、世界的にもいまだ様々な課題の検討が進められている段階であり、十分なノウハウが蓄積されているとは言いがたい。そのため、JaLC において研究データへの DOI 登録を開始するのにあたり、研究データへの DOI 登録に特有の課題抽出とその解決、運用方法の確立、DOI の活用方法などの検討を行い、我が国における研究データへの DOI 登録の仕組みを参加機関とともに新たに構築することを目的として実験プロジェクトを実施した。

本報告書は、2014 年 10 月から 2015 年 9 月にかけて 1 年間実施してきた取り組みなどを漏れなくとりまとめたものである。

2. 公募による参加機関募集

2.1 公募予告

公募による参加機関の募集に先立ち、2014年8月20日に公募の予告を行った。ジャパンリンクセンターwebサイトに、「研究データへのDOI登録実験プロジェクトについて【予告】」（附録資料1）を掲載した。主要な内容は、(i) 公募によりプロジェクト参加機関を募ること、(ii) プロジェクト参加機関の主体的要件、(iii) プロジェクト実施期間、(iv) 費用についてである。

2.2 公募説明会

次のとおり、公募説明会を開催した。

「研究データへのDOI登録実験プロジェクト」参加機関募集 公募説明会	
日時：	2014年9月17日（水）15時00分～16時00分
会場：	独立行政法人 科学技術振興機構 東京本部 別館 2階セミナー室
内容：	・公募要領（附録資料2） ・プロジェクト計画書 ・「参加申し込み書」記入要領
参加機関数：	3機関

当日の質疑応答内容を取りまとめ、ジャパンリンクセンターwebサイトに公開した（附録資料3）。

2.3 公募

2014年9月10日（水）9:00～2014年9月30日（火）16:00の公募期間を経て、表2-1.の応募を得た。

表 2-1. 応募機関一覧

項番	機関名	プロジェクト責任者		
		所属	役職	氏名
1	独立行政法人 科学技術振興機構	バイオサイエンスデータベースセンター	研究員	榎田 達矢
2	大学共同利用機関法人 情報・システム研究機構 国立極地研究所	北極観測センター	特任准教授	矢吹 裕伯
3	大学共同利用機関法人 情報・システム研究機構 国立情報学研究所	学術基盤推進部	准教授	山地 一禎
4	大学共同利用機関法人 情報・シス	コンテンツ科	准教授	北本 朝展

	テム研究機構 国立情報学研究所	学研究系		
5	独立行政法人 産業技術総合研究所	情報技術研究 部門	総括研究主幹	小島 功
6	独立行政法人 情報通信研究機構	統合データシ ステム研究開 発室	室長	村山 泰啓
7	千葉大学附属図書館	利用支援企画 課	課長	杉田 茂樹
8	独立行政法人 物質・材料研究機構	企画部門科学 情報室	室長	谷藤 幹子
9	独立行政法人 理化学研究所 脳科学総合研究センター神経情報 基盤センター	神経情報基盤 センター	センター長	山口 陽子

2.4 選考

平成26年度ジャパンリンクセンター運営委員会(第3回)(2014年10月2日(金)開催)において、プロジェクト参加機関の選考が行われた。選考の結果、応募した9機関の採択が決定された。

3. プロジェクト実施方式の検討

キックオフミーティング（2014年10月27日（月）開催）にて、プロジェクト参加機関が集まり、(i) プロジェクト体制の検討、(ii) プロジェクト計画のレビュー、(iii) プロジェクト実施スケジュールの確認、(iv) 各参加機関の概要、実施体制、準備段階で認識している課題の共有を行った。

3.1 プロジェクトの体制

本プロジェクトの体制は、次のとおり決定された。

表 3-1. プロジェクトの体制

メンバー	担当
プロジェクトリーダー (武田英明 国立情報学研究所教授)	進捗管理、報告
サブリーダー (村山泰啓 独立行政法人 情報通信研究 機構 室長)	プロジェクトリーダーの進捗管理、報告を支援 する
プロジェクト参加機関	<ul style="list-style-type: none">・研究データへの DOI 登録について、運用ポリシーおよび登録フローを検討する。・DOI 登録のテストを実施する。・本プロジェクト実施報告書を作成する。
JaLC 運営委員会	<ul style="list-style-type: none">・本プロジェクト参加申請の審査を行う。・本プロジェクトの報告を受け、JaLC 運営方針を検討し決定する。
対象コンテンツの拡大検討分科会	<ul style="list-style-type: none">・本プロジェクトに対して助言を行う。
JaLC 事務局 (JST)	<ul style="list-style-type: none">・JaLC 運営委員会、対象コンテンツの拡大検討分科会、本プロジェクトの運営サポートを行う。・JaLC システムの運用を行う。

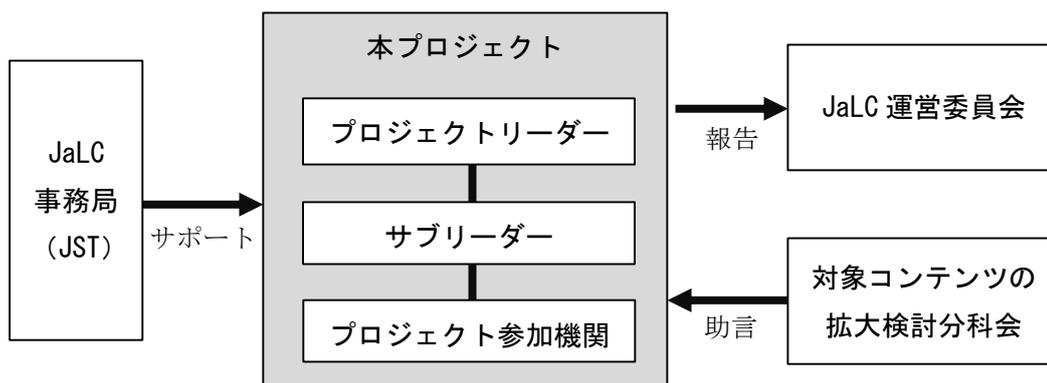


図 3-1. プロジェクト全体組織図

3.2 プロジェクト計画のレビュー

本プロジェクトの計画（事務局案）のレビューを行い、公募時のもので承認された。

3.3 プロジェクト実施スケジュール

本プロジェクト実施のスケジュール（事務局案）も承認された。

参加者が 40 名程度にもおよぶ大きな会議であることから、キックオフミーティングの段階で、1 年間分のプロジェクトミーティングや、チェックポイントとなる中間報告会、懇親会などのスケジュールをあらかじめ決めたことが成功要因のひとつであった。

表 3-2. プロジェクト実施スケジュール

日時	形態	項目	担当	内容	備考
2014 10/27	月 会合	第 1 回ミーティング	全 G	<ul style="list-style-type: none"> 委員長挨拶、計画書説明 各グループ計画説明 スケジュール検討 勉強会講師割り当て 第 2 回ミーティングまでの予定（テスト計画、ポリシー作成） 	
12/10	水 会合	DataCite 等勉強会	テスト実施 G 希望 G	<ul style="list-style-type: none"> DataCite のメタデータスキーム 先行事例の紹介（各分野で海外先行事例を紹介、たとえば WDS や海洋データなど） 	JaLC2 リリースが近いので日程きびしいが、作業に取りかかるためには早い開催が望ましい。事務局だけが説明する場ではなく、ノウハウを持ち寄る場として開催したい。
10/27 -12/26	作業	登録テスト計画作成	テスト実施 G	<ul style="list-style-type: none"> 登録テスト計画作成 	
12/27 -1/23	作業	登録テスト計画検討	全 G	<ul style="list-style-type: none"> メール等による情報共有、議論 	

10/27 -1/23		作業	ポリシー検討	各G	<ul style="list-style-type: none"> ・課題抽出 ・ポリシー検討 ・メール等による情報共有、議論 	
2015 1/28	水	会合	第2回ミーティング	全G	<ul style="list-style-type: none"> ・登録テスト計画決定 ・ポリシー作成計画検討 ・「対話・共創の場」担当決め 	
2/27	金	会合	JaLC 対話・共創の場（第2回「研究データに対するDOI登録」）	全G	<ul style="list-style-type: none"> ・外部（学会、図書館等）からの意見を取り入れる場。いくつかのグループから議論の導入として活動内容を発表。 ・懇親会 	
1/29 ～		作業	登録テスト	テスト実施G	<ul style="list-style-type: none"> ・登録テスト ・結果レポート作成、共有 	
		作業	ポリシー検討	全G	<ul style="list-style-type: none"> ・課題抽出 ・ポリシー検討 ・メール等による情報共有、議論 	
		会合	ポリシー検討	全G	<ul style="list-style-type: none"> ・必要に応じて他グループと議論 	
		作業	登録計画作成	登録実施G	<ul style="list-style-type: none"> ・登録計画作成 ・メール等による情報共有、議論 	
4/20	月	会合	第3回ミーティング	全G	<ul style="list-style-type: none"> ・登録テスト総括 ・ポリシー検討 ・実登録計画決定（・中間報告書目次確認、役割分担）？ 	
		作業	実登録	登録実施G	<ul style="list-style-type: none"> ・登録テスト ・結果レポート作成、共有 	
		会合	ポリシー検討	全G	<ul style="list-style-type: none"> ・必要に応じて他グループと議論 	
		作業	中間報告書作成	担当G	<ul style="list-style-type: none"> ・中間報告書作成 	
6/		会合	DataCite と情報共有	運営委員会事務局	IDF の定例会に合わせて DataCite、CrossRef と情報共有、意見交換。	アイデア段階。DataCite のミーティング（8月ごろ？）では遅すぎると思料
7/3	金	会合	中間報告会	全G	外部関係者（学会、図書館等）とともに、これまでの検討およびテストの結果共有、今後の方向性の議論、確認。	「第3回対話共創の場」とするか？
8/5	水	会合	第4回ミーティング	全G	<ul style="list-style-type: none"> ・実登録総括 ・中間報告書確認 ・最終報告書目次確認、役割分担 	
		作業	最終報告書作成	担当G	<ul style="list-style-type: none"> ・最終報告書作成 ・メール等による事前共有、確認 	
9/25	金	会合	第5回ミーティング	全G	<ul style="list-style-type: none"> ・最終報告書確認 ・懇親会 	
12/3		会合	最終報告会	全G	IDF の定例会・アウトリーチミーティングを東京に誘致し、半日程度を本プロジェクト報告会にあてる。	アイデア段階

3.4 各参加機関の概要、実施体制、準備段階で認識している課題の共有

3.4.1 独立行政法人 科学技術振興機構

表 3-4-1-1. 参加機関概要

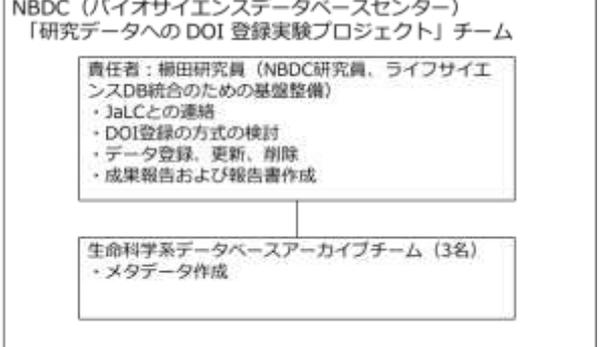
プロジェクト責任者	バイオサイエンスデータベースセンター 榎田達矢
実施内容	<p>NBDC がサービスを提供している「生命科学系データベースアーカイブ」に収録されているデータセットに対して DOI を付与することを検討しています。</p> <p>「生命科学系データベースアーカイブ」は、国内のライフサイエンス研究者が生み出したデータセットをわが国の公共財としてまとめて長期間安定に維持保管し、データ説明（メタデータ）を統一して検索を容易にすると共に、利用許諾条件などの明示を行うことで、多くの人々が容易にデータへアクセスしダウンロードを行えるようにするサービスです。</p> <p>(http://dbarchive.biosciencedbc.jp/)</p> <p>① DOI 登録するデータの分野：生命科学 データ種類：表形式データ他 登録件数：90 データベース</p> <p>② 登録した DOI の活用方法：データセットの二次利用の促進</p>
実施体制	

表 3-4-1-2. 想定される主な検討課題

アクセスの持続性の保証	<p>② DOI 登録後のデータ取り下げの取り扱いについて</p> <p>「生命科学系データベースアーカイブ」は、各研究機関から寄託していただいたデータベースのデータについて、長期間安定に維持保管し、データ説明（メタデータ）を統一して検索を容易にすると共に、利用許諾条件などの明示を行うことで、多くの人々が容易にデータへアクセスしダウンロードを行えるようにする NBDC のサービスです。</p> <p>NBDC データベース寄託規約のなかで、データベースを公開した後、「寄</p>
-------------	---

の種別ごとの取り扱いの検討について

合のバージョン管理は行っていません（修正、追加前後のデータの URL は同一のものを使用）。DOI 登録後にデータ追加や修正等が行われた場合の対応は本プロジェクト内で検討する予定です。

また、いわゆるビックデータと呼ばれるようなデータ量の多いデータについては、現在、生命科学系データベースアーカイブの対象になっていないので、必要に応じて生命科学系データベースアーカイブの運用ルールと連動してその取扱いを検討したいと考えます。

表 3-4-1-3. プロジェクト計画に対する意見等

参考：データベース（Open TG-GATEs）およびデータ（病理情報）のメタデータの例

Database Description		Original website information	
General information of database		Database maintenance site	
Database name	Open TG-GATEs	URL of the original website	http://www.ninbio.go.jp
Alternative name		Creation start date	2011/02/01
Creator	Creating Name: Toxicogenomics Project / Toxicogenomics Informatics Project Creating Affiliation: National Institute of Biomedical Innovation, National Institute of Health Sciences, and 15 pharmaceutical companies	Last update date	2017/07/18
Contact address	Toxicogenomics Informatics Project National Institute of Biomedical Innovation 7-8-8, Tsukuba-ku, Utsunomiya-city, Chiba 305-8565, Japan TEL: 81-72-441-9920 Email: open@nimbio.go.jp	URL of the portal site	http://www.ninbio.go.jp
Database classification	Toxicogenomics Database	Whole data download	-
Organism	Taxonomy Name: Homo sapiens Taxonomy ID: 9606	Reference database	-
Database description	Toxicogenomics Project (TOP) is a government/private companies collaborative project started by the National Institute of Biomedical Innovation, the National Institute of Health Sciences, and 15 pharmaceutical companies in 2002. After 5 years of the project, 153 chemicals were administered to rats or exposed to rat and human primary cultured hepatocytes, and the gene expression profiles in the liver and kidney of the animal or in the cultured cells were comprehensively analyzed by microarray. As a result, a high-quality large-scale toxicogenomics database with the systems to analyze the gene expression data and predict the safety of candidate chemicals has been developed (TG-GATEs, Toxicogenomics Project Genomics Analysis Toxicity Evaluation system). Toxicogenomics Informatics Project (TOPI) is TOPI's subproject started by the National Institute of Biomedical Innovation, the National Institute of Health Sciences, and 13 pharmaceutical companies in 2007. After 5 years of the project, more than 70 safety chemicals were developed by using TG-GATEs. In addition, data acquired in experiments and analysis that relationships are included in TG-GATEs. Open TG-GATEs is a toxicogenomics database based on the public information to utilize related results of TOPI and TOPI2, and releases the data of 170 compounds stored in TG-GATEs in Open TG-GATEs. It is possible to search toxicogenomics data by compound name or pathological finding. It is also possible to download gene expression data associated with chemical IDs such as pathological findings as a GSE file. GSE is one of the file formats that expresses gene expression data from data generated from microarray experiments.	Query search	Not available
Features and content of utilization of database	Searching toxicogenomics data by compound name or pathological finding Downloading gene expression data associated with phenotype data such as pathological findings as a GSE file	Web service	Not available
Background and funding	<ul style="list-style-type: none"> Health and Labour Sciences Research Grant (H14-Tsukuba-01 and H19-Tsukuba-01) Collaborative research funds from participating companies 	URL of data archive	-
References	Article title: Author name(s): Journal: Published ID:	Need for user registration	-
		Pathological items	
		Data detail	
		Data name	Pathological items
		Description of data contents	The list of pathological findings obtained from a liver test.
		Data file	File name: 1001_020001_pathology.zip File URL: http://www.ninbio.go.jp/ftp/1001_020001_pathology.zip File size: 28 KB
		Simple search URL	http://www.ninbio.go.jp/pathologysearch/1001_020001_pathology
		Data acquisition method	All generated pathology data were obtained by optical microscopy and then performed microarray analysis of microarrayed by optical microscopy.
		Data analysis method	-
		Number of data entries	12,981 entries

日本語版、英語版のどちらをDOI登録するか(JaLC or DataCite)？

3.4.2 大学共同利用機関法人 情報・システム研究機構 国立極地研究所

表 3-4-2-1. 参加機関概要

プロジェクト責任者	北極観測センター 矢吹裕伯
実施内容	<p>①DOI 登録するデータの分野：地球科学・環境分野 データ種類： 北極域における実験及び観測データ（気象、水文、大気、海洋、雪氷、生物、地震、宙空） 南極域における実験及び観測データ（気象、水文、大気、海洋、雪氷、アイスコア、生物、固体地球物理、宙空） 登録件数：北極域:約 50 件、南極域：約 100 件</p> <p>②登録した DOI の活用方法 登録されたデータの DOI は、科学論文を作成する場合において引用を行う。DOI を付与する事はデータの恒久的なデータのリンクをデジタル的に確保することであり、それにより科学論文の透明性と説明責任を明確にすることができる。また科学論文の公開とデータの公開は第三者による研究結果の検証を可能にし、研究方法の向上と妥当性の確認を推進することができる。</p> <p>データへの DOI 付与は、データの作成者、および管理するデータセンターの寄与が明確になり、評価基準として用いることができる。将来的には、引用 DOI 数を統計処理することによりデータ作成者やデータセンターの寄与を明確にすることで、研究者およびデータセンターの評価の仕組みを構築する。</p>
実施体制	<p>研究統括（進捗を管理・報告） 国立極地研究所・副所長 教授 榎本 浩之 実施責任者：(DOI 付与に関して取りまとめを行う) 国立極地研究所・北極観測センター 特任准教授 矢吹裕伯</p> <ul style="list-style-type: none"> ・研究データへの DOI 付与について、運用ポリシーおよび登録フローを検討する。 ・DOI 付与する研究データの登録基準を検討する。 ・DOI 付与に関して取りまとめを行う。 <p>実施担当者（DOI 付与に関して検討を行う） 国立極地研究所極域データセンター・副センター長 准教授 岡田雅樹</p>

	<ul style="list-style-type: none"> ・ DOI 登録データを持続的に維持管理する体制について検討する ・ 関連する研究機関および研究プロジェクトとの連携について検討する。 <p>国立極地研究所・北極観測センター 特任研究員 杉村 剛 国立極地研究所・北極観測センター 特任研究員 照井 健志</p> <ul style="list-style-type: none"> ・ 研究データへの DOI 付与について、運用ポリシーおよび登録フローを検討する。 ・ DOI 付与する研究データの登録基準を検討する。 ・ DOI 付与に関してシステムを構築する。 ・ DOI 付与に関してデータ提供者へのチュートリアルを実施する。
--	--

表 3-4-2-2. 想定される主な検討課題

<p>アクセスの持続性の保証</p>	<p>① 有期のプロジェクトで保有するデータの扱いについて</p> <p>有期プロジェクトで保有するデータについて、そのプロジェクトを管轄する機関が持続的なアクセスを保障すべきであり、そのような手段を準備しなくてはならない。</p> <p>もしそのようなプロジェクトもしくはその上位の機関が持続的なアクセスを保証できない場合は、持続的なアクセスを保障する機関もしくはデータセンター等へデータを送り、持続的なアクセスを確保する必要がある。ここで有期プロジェクトの代表機関でなくても、有期プロジェクトが保有する研究データが関連するデータセンターもしくは機関がデータの持続的なアクセスを保障すればよいということになる。</p> <p>② DOI 登録後のデータ取り下げの取り扱いについて</p> <p>データの取り下げが行われた場合でも、そのデータは利用されていると考えるべきである。取り下げを行った場合でも一度 DOI を付与してしまったデータに関しては、持続的なアクセスを保障すべきであり、ランディングページにて、取り下げが行われたことを記載し、DOI はそのまま残すべきである。</p>
<p>DOI 登録する対象データの粒度について</p>	<p>国立極地研で取り扱うデータセットの中での粒度について</p> <p>基本的に、自動観測で長期的に取得するものと、マニュアル観測によるスナップショットに観測データがある。自動観測で長期的に取得する観測データは、自動的に転送されてくるもの、現地での年 1 回ないし 2 回の割合でデータ回収を行うものなどがある。通常自動的に転送されてくるデータに関しては、年単位で 1 つのデータセットとすることが多い。また現地で回収してくるようなデータの場合は、回収毎に 1 つのデータセットとする</p>

	<p>場合や、年単位で 1 つのデータセットとすることが多い。これらのデータセットはシリーズものとして扱い、年単位もしくは、データ取得ごとのデータセットがそのシリーズの一連のデータセットとする。</p> <p>【データの条件】</p> <p>我々が扱うデータは観測データが主であるが、DOI の登録対象とすべきデータの条件は特になく、すべての観測データが対象となる。</p> <p>ただし、それらのデータは、品質管理されたものであり、基本的なメタ情報、取得方法、観測エラー等を記載して、第 3 者による査読が行われたデータとする。</p>
DOI のランディングページの要件の検討について	<p>データの公開、非公開のケースを認めるのでしょうか？</p> <p>この考えが理解できません。DOI 付与を行うデータは公開というのが原則ではないのでしょうか？非公開にする場合があるとすると、それはどのような想定があるのか、教えてほしい。</p> <p>極地研では DOI 付与すべきデータは公開するという方針なので、この件に関して現在意見はない。</p>
研究データの種別ごとの取り扱いの検討について	<p>DOI 登録後にデータ追加が行われる場合について：</p> <p>【その 1】</p> <p>どのような場合が想定されるが不明であるが、基本的にデータ追加された場合は、そのデータセットは、異なるバージョンのデータセットとして扱う必要がある。ただし、この場合 DOI は単一のものであり、そのリンク先では、異なるバージョンのデータセットを参照できるようにする必要がある。</p> <p>【その 2】</p> <p>データが追加された場合は、データセットとしてはあくまでも別のバージョンのデータセットとなるので異なる DOI を付与する必要がある。</p> <p>【その 3】</p> <p>データの粒度と非常に関連するが、長期的に取得している観測データセットにデータが追加された場合と考えると。データセットはシリーズとして別名をつけ年々データ増えることになり、年単位ごとに DOI を付与するという考えも必要かも知れない。</p>

表 3-4-2-3. プロジェクト計画に対する意見等

DOI を付与する観測データセットは持続的なアクセスを保障することになる。DOI 付与を行うデータセットに求められる重要な条件は品質である。DOI 付与はあくまでも仕組みであり、その品質に関しては関与しないかもしれないが、DOI を付与されたデータセットはやはり品質管理されていることが必要である。DOI を付与されたデータの信頼性

を高め、今後一層の DOI 利用推進を進めるためには、DOI 付与するデータの、基本的な条件（ガイドライン）を決める必要があると考えられる。
この DOI 付与するデータに求められる、基本的なガイドラインは必要である。

3.4.3 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所（地球環境情報統融合プログラム (DIAS-P)）

表 3-4-3-1. 参加機関概要

プロジェクト責任者	国立情報学研究所 北本朝展
実施内容	<p>「地球環境情報統融合プログラム (DIAS-P)」は、地球規模課題、特に気候、水循環、生物多様性・農業・水産業に関する科学知の深化と公共的利益創出のための情報基盤のパイロットシステムの構築を目指すものである。様々な分野の利用者（ステークホルダー）が、超大容量で多様なデータ・情報を協働して統融合し、新たな価値を創出できる情報基盤（ワークベンチ）のプロトタイプを構築するとともに、地球規模課題解決に向けて、科学的先端性を持続的に発揮し、実利用によって公共的利益を実現できる運用体制を設計・提案する。この研究活動の一環として、データの公開および共有は重要なミッションの一つであるため、JaLC 実験プロジェクトに参加することにより、データに関する研究基盤を整備することを目指す。</p> <p>①データの分野： 地球科学データ：地上観測データ(75)、衛星観測データ(92)、気象予測モデル(34)、気候変動予測モデル(6)、その他社会データ(9)</p> <p>DIAS ではデータ提供者が指定する単位で識別子を付与する仕組みを採用しているが、異なる識別子を与える根拠となるのは、主にデータの種別とデータポリシーの違いである。DIAS-P においては 4000 万ファイルにも達するデータの規模を考えれば、そのすべてに識別子を与えることは非現実的であるため、このように粗い粒度で識別子を与える方式を採用した。とはいえ、現在の方式ではデータセットの識別子が 200 個強にとどまっており、相互運用性およびデータサイテーションの観点から最適な方式になっているかどうかは評価しなければならない。</p> <p>また、DIAS では国際標準のメタデータ、および国際的に利用されているオントロジーなどを取り込みながら、独自のメタデータ形式を定義している。この形式は主に地球環境データへの適用を意図した形式であるが、分野内の相互運用性あるいは分野間の相互運用性を考えた場合、さらに改善する余地が生まれる可能性もある。</p> <p>②登録した DOI の活用方法</p>

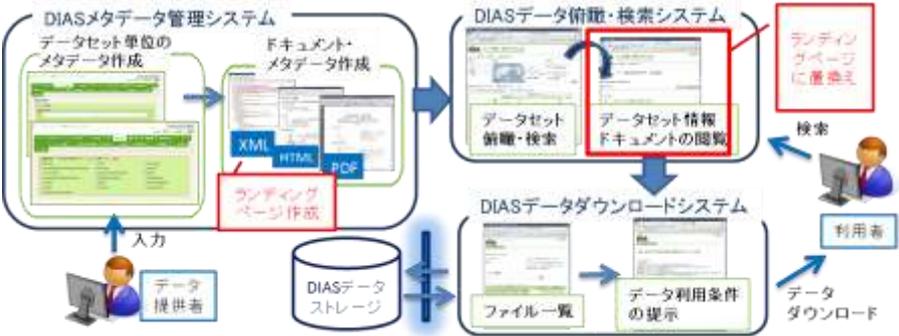
	<p>DIAS-P では既にメタデータを活用した検索サービスを提供しているが、他のシステムで生み出されたメタデータも含めたサービスを実現することを考えれば、先進的なサービスの可能性は大きく広がっていく。この点に関しては、日本での DOI 付与状況なども踏まえつつ、メタデータサービスの構築に向けた新しいアイデアを検討する計画である。</p>
実施体制	<p>DIAS-P では、メタデータに関する検討グループが4つ存在し、個々のグループでは以下の課題に関する研究を進めている。</p> <ol style="list-style-type: none"> 1.メタデータ形式の研究グループは、標準的なメタデータの調査や研究分野のニーズに合わせた改良などの問題に取り組んでいる。 2.メタデータに利用する専門用語・オントロジーの研究グループは、世界で広く利用されているオントロジーの収集や統合、活用などの問題に取り組んでいる。 3.メタデータ登録システムの研究グループは、定義したメタデータ形式を地球環境の研究者が簡便に入力するための登録システムを構築し運用している。 4.メタデータ検索システムの研究グループは、登録したメタデータを使って必要なデータにアクセスするための検索システムを構築し運用している。 <p>これらの研究は、主に東京大学地球観測データ統融合連携研究機構（EDITORIA）および京都大学情報学研究科が中心となって進めており、さらに DIAS 長期運用体制については海洋開発研究機構（JAMSTEC）でも検討を進めている。そこで、これらのグループにオブザーバとしての参加を要請して協力を仰ぎつつ、国立情報学研究所が JaLC プロジェクトを取りまとめるという体制で、実験プロジェクトを実施する。</p> <div style="display: flex; justify-content: space-around; align-items: flex-start; margin-top: 20px;"> <div style="text-align: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 5px;">国立情報学研究所</div> <div style="border: 1px solid gray; padding: 5px; background-color: #f0f0f0;">JaLCプロジェクトとりまとめ</div> </div> <div style="text-align: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 5px;">東京大学地球観測データ 統融合連携研究機構</div> <div style="border: 1px solid gray; padding: 5px; background-color: #f0f0f0;">データアーカイブ・オントロジー</div> </div> <div style="text-align: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 5px;">海洋開発研究機構</div> <div style="border: 1px solid gray; padding: 5px; background-color: #f0f0f0;">長期運用体制の検討</div> </div> <div style="text-align: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 5px;">京都大学情報学研究科</div> <div style="border: 1px solid gray; padding: 5px; background-color: #f0f0f0;">メタデータ形式・登録システム</div> </div> </div>

表 3-4-3-2. 想定される主な検討課題

運用フロー 取りまとめ 機関について	<p>【アーカイブ対象データ】</p> <ul style="list-style-type: none"> ・DIAS は、地球環境に関わる課題解決のため、その課題にかかる政策決定者も参加した協働プロジェクトを支援する情報基盤を提供し、場合によっては協働プロジェクトの成果の社会実装先となる場合もある。従って、
--------------------------	--

	<p>DIAS はこれら協働プロジェクトが必要とするデータ、及び、同プロジェクトが収集・作成したデータを優先的にアーカイブする。すなわち、DIAS がプロジェクト支援のために自ら作成したデータのほか、他のデータセンターが提供するデータやプロジェクト参加機関が収集・作成したデータがアーカイブの中心となる。</p> <ul style="list-style-type: none"> ・それに加えて、総合科学技術・イノベーション会議 CSTI の方針として、「我が国の地球観測計画」に掲げられた全てのデータセットのメタデータを DIAS に登録し、分野横断的な検索を可能とする。また、GEOSS ポータル及び DIAS/CEOS 水ポータルにおいて、メタデータを介した国際的なデータの相互流通にも寄与する。 <p>【DOI 登録】</p> <ul style="list-style-type: none"> ・以上、DIAS にアーカイブされたデータのうち、DIAS-P が DOI を付与するのは、第一義的には DIAS-P が自ら作成したデータと、協働プロジェクトのもとで収集又は作成された他機関のデータである。他機関が提供したデータの著作権はデータ作成者に帰属するので、DOI 付与の可否並びに DOI 付与の粒度は、データ作成者との合意のもとに行われる。 ・超大容量データの収集・解析、リアルタイム処理等の環境を研究者が個々に用意するのは大変である。このため、DIAS は他機関が提供する超大容量の再解析データや長期予測データのアーカイブ、並びに、観測データや短期予測データのリアルタイム・アーカイブと解析環境を提供している。現時点ではこのような目的によりアーカイブされている他機関の提供データへの DOI 付与については白紙である。 ・そのほか DIAS にメタデータのみ登録されている他機関保有データへの長期運用 DIAS による DOI の付与については、メタデータ提供機関の中には希望するところも少なくないと思われるが、現時点では白紙である。
<p>アクセスの持続性の保証</p>	<p>① 有期のプロジェクトで保有するデータの扱いについて</p> <p>【DIAS の継続性】</p> <ul style="list-style-type: none"> ・DIAS は、地球環境データの統融合システムおよびその運用体制を指すものである。 ・DIAS は、現在は第2期プロジェクトの4年目にあたり、これを DIAS-P と呼んでいる。 ・DIAS-P は、2015 年度末までの有期プロジェクトであり、現在の東京大学 EDITORIA を代表研究機関とする研究開発プロジェクトは終了する。 ・2016 年度からは長期運用体制を整えた長期運用 DIAS が発足する予定であり、その運用主体は 2014 年 12 月末に文科省が決定する見込みである。長期運用 DIAS は有期プロジェクトではなく、継続性のある業務となる予定

	<p>である。</p> <p>【実データ】</p> <ul style="list-style-type: none"> ・ DIAS-P は、RECCA、GRENE-ei、環境省 S-8、創生プロ分野 D など気候変動適応に関わる有期プロジェクトの成果を受入れることを文部科学省の方針としている。これら受入れデータのうち、超大容量の高解像度なシミュレーション結果など再計算可能なデータについては、一定保存期間後の削除もありうるが、シミュレーションの入力データとなる観測データは、永続的な保存を想定しており、データ DOI の付与を想定している。その他共用性の高い再解析データ、予測データについて保存／廃棄ポリシーを定める予定。 ・ 不幸にも長期運用 DIAS が終了する場合には、再計算不可能な観測データをパッケージ化し、国立国会図書館ほか主要なデータセンターに寄託するなどの方策を取ることが考えられる。 <p>【メタデータ】</p> <ul style="list-style-type: none"> ・ DIAS がメタデータのみを登録している他機関のデータセットについて、その機関での保有が困難になった場合の DIAS への受入れポリシーを定める予定。また、これにより受入れたデータのうち、そのデータ作成者がデータの品質管理等の責務を果たせなくなった場合の取扱いについてポリシーを定める予定。 <p>② DOI 登録後のデータ取り下げの取り扱いについて</p> <p>【データの取り下げ】</p> <ul style="list-style-type: none"> ・ DIAS ストレージにファイルが格納されているデータセットを DIAS では DOI 登録対象と考えている。 ・ DOI 登録後にデータ提供者が研究データの公開を取り下げた場合でも DIAS はメタデータの公開は継続することを想定している。 ・ 運用上はメタデータにデータ提供者、データ公開を取り下げた理由を記述して DOI のランディングページは維持されるものと想定している。
DOI 登録する対象データの粒度について	<p>【データセット】</p> <p>DIAS が公開しているデータは、原則としてデータ提供者の判断でデータセットとしての粒度を決めてデータセット ID を付与することになっている。データセットの粒度は常に検討課題となっている。</p> <p>データセットによりデータ量には差がある。1 データセット内のファイル数が 10 未満のものから数千万のものまであり、論理的なデータの識別と物理的なデータの識別には大きなギャップがある。</p> <p>DIAS からは同一データセット内のデータの利用規約は同一であることを</p>

	<p>求めている。</p> <p>データセットごとにメタデータが作成される。データセット単位のメタデータを DIAS ではドキュメントメタデータと呼んでいる。</p> <p>【DOI 登録の単位】</p> <p>DOI を付与するデータセットの粒度も現実的な対応としては、現在 DIAS が公開しているデータセットと 1 対 1 に対応する方向だと考えている。</p> <p>DOI に対応したデータセットのランディングページは DIAS がデータセットごとに作成しているドキュメントメタデータから作成されるものを想定している。</p>
<p>DOI のランディングページの要件の検討について</p>	<p>【DIAS におけるメタデータ管理】</p> <ul style="list-style-type: none"> • DIAS ではデータセットごとにデータセット ID を付与し、DIAS メタデータ管理システムがメタデータを管理している。各データセットの担当者(データ提供者など)は担当メタデータの更新ができる。 • メタデータは ISO19139 に基づいた XML 形式で管理され、その内容は HTML および PDF でも出力される。 <p>【DIAS 俯瞰・検索システム】</p> <ul style="list-style-type: none"> • メタデータ管理システムにより作成されたメタデータは DIAS 俯瞰・検索システムで検索可能である。 <p>【ランディングページ】</p> <ul style="list-style-type: none"> • メタデータ管理システムが出力するドキュメントメタデータを基にしたウェブページが担うことが想定できる。例えば、DIAS 俯瞰・検索システムにおける各データセットのメタデータ閲覧ページ (例: http://dias-dss.tkl.iis.u-tokyo.ac.jp/ddc/viewer?ds=MIRAI_CTD) が考えられる。 • データの公開・非公開はメタデータ中のオンライン情報にファイルダウンロードへのリンクとして記述している。従って DOI のランディングページを DIAS メタデータ管理システムに対しての変更を反映する形式にしておく限り、維持管理は問題なく実施できると考えられる。 
<p>研究データ</p>	<p>【DIAS データダウンロードシステム】</p>

<p>の種別ごとの取り扱いの検討について</p>	<ul style="list-style-type: none"> ・DIAS で公開されているデータはDIAS データダウンロードシステム(以下、ダウンロードシステム)におけるウェブインタフェースを通してダウンロード可能である。 ・DIAS からデータを公開するに当たり、データファイルへの利用者のアクセスに許可申請が必要か否かをデータ提供者が指示する。 <p>ダウンロードシステムではデータセットごとにファイル一覧を管理しており、利用者は一覧の中からファイルを選択してダウンロードすることができる。(要ユーザ登録、ポリシーへの同意)</p> <ul style="list-style-type: none"> ・ダウンロードシステムではデータセットごとに対応するダウンロード URL を割り当てる。そのアドレスを DIAS ドキュメントメタデータのオンライン情報に記述することで現状のデータセット ID とデータセット内のファイルリストとの関連付けを行っている。 <p>【ランディングページにおけるデータ】</p> <ul style="list-style-type: none"> ・ランディングページとしてはデータそのものではなく、当該データセットのメタデータ閲覧ページを想定している。データファイルダウンロードへのリンクはメタデータ中のオンライン情報の項目に記載されている。 <p>【データファイルの管理】</p> <ul style="list-style-type: none"> ・データ追加、変更、削除はオンライン情報のリンク先として記述されているダウンロードシステムで行う。 ・データファイルの検索機能はダウンロードシステムが担う。 <p>【データの取り扱いに関する議論】</p> <ul style="list-style-type: none"> ・データの追加、変更等が行われた場合でもデータセットとしての同一性が認識できれば同一の DOI で参照するのが現実的だと考えられる。その場合、ランディングページには、データの追加変更記録を記述（ドキュメントメタデータの改訂）することが望ましい。 ・データセット単位でのメタデータ閲覧ページをランディングページとする場合、データ量の多寡は問題とはならないと考えられる。
--------------------------	---

表 3-4-3-3. プロジェクト計画に対する意見等

<ul style="list-style-type: none"> ・DOI の粒度に関しては、何を基準とするかが重要な点となる。短期的にはデータ提供者の都合が優先されるようなインセンティブ設計となってしまう、データサイテーションによる利用状況把握などの長期的な観点は疎かになりがちである。データ提供者に長期的な観点からの DOI 付与を検討してもらうよう、働きかけるロジックが必要である。 ・サイテーションのための DOI とアクセシビリティのための DOI という二つの異なる考え方のギャップは残ると考えられ、そのギャップを埋めるためのデータベースやサービスも検討課題になるかもしれない。

- ・ジャーナルがデータ DOI の表示を論文採択の条件とするよう働きかけることも重要な課題である。これは個々のプロジェクトの課題というよりは、JaLC 全体の課題と言える。
- ・ジャーナルがデータ DOI を採用すると並行して、データ DOI を用いてデータの利用状況をどのように把握するか、といったサービスも検討しておく必要がある。これも JaLC 共通の課題である。

3.4.4 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 学術基盤推進部

表 3-4-4-1. 参加機関概要

プロジェクト責任者	学術基盤推進部 山地一禎
実施内容	<p>① DOI 登録するデータの分野、データ種類、登録件数</p> <p>国立情報学研究所が運用する IRDB システムを介して、データに DOI を付与する。データの種別としては、IRDB と連携する機関リポジトリ、または、サブジェクトリポジトリが保有する、研究データ等が該当する。具体的な種別と件数は以下のとおりである。</p> <p>テキストデータ：約 100 件 画像データ：約件 1000 件 音声データ：約 200 件</p> <p>② 登録した DOI の活用方法</p> <p>DOI を付与したデータの引用に活用されることを期待する。</p>
実施体制	<p>現在登録を予定している下位団体</p> <ul style="list-style-type: none"> ・ 情報学研究データリポジトリ (NII) ・ NTCIR 言葉コーパスリポジトリ (NII) ・ 学術資源リポジトリ (学術資源リポジトリ協議会) <p>その他、大学・研究機関からの要請により、参加階機関が参加する可能性あり</p>

表 3-4-4-2. 想定される主な検討課題

運用フロー 取りまとめ 機関について	<p>国立情報学研究所の学術機関リポジトリデータベース (IRDB) がハーベストした研究データについて、DOI を登録する方式を検討する。</p> <ul style="list-style-type: none"> ・ IRDB のハーベスト対象の再定義 ・ 登録申請フローの確立 ・ IRDB と JaLC のメタデータマッピングの検証
アクセスの	② DOI 登録後のデータ取り下げの取り扱いについて

<p>持続性の保証</p>	<p>「DOI 登録後に研究データの取り下げが行われた場合、持続的なアクセス保証を維持できる」というのが、どのようなユーザエクスペリエンスを想定しているのかがわからない。</p> <p>データ自体がインターネット空間から削除される場合には、データ提供者の要望を尊重する運用方針をとることになると考えられる。</p> <p>その場合でも DOI へのアクセス後に、利用者には何らかのランディングページを提供するという観点からは、JaLC のリゾルバがそのような機能を提供する方法もある。ポリシーだけではなく、利用者に何をどう担保すべきか、という観点から機能面とポリシー面の双方から議論がされるべきである。</p>
<p>DOI 登録する対象データの粒度について</p>	<ul style="list-style-type: none"> ・基本的にはデータ登録者が引用を期待する粒度に DOI を付与する。 ・下位団体の持つ研究データのメタデータと、IRDB と JaLC のメタデータマッピングの検証を行う。
<p>DOI のランディングページの要件の検討について</p>	<p>非公開のデータに DOI を付与することを検討するという意味が分からない。</p>
<p>研究データの種別ごとの取り扱いの検討について</p>	<p>データリポジトリ内におけるバージョン管理機能（バージョンごとのランディングページ付与機能）などについて検討する。</p>

3.4.5 産業技術総合研究所

表 3-4-5-1. 参加機関概要

<p>プロジェクト責任者</p>	<p>情報技術研究部門 小島功</p>
<p>実施内容</p>	<p>産総研の保有する研究成果のデータベース「産総研データバンク」のうち、地質および材料系を中心としていくつかについて、j aLC や関連委員会等の協力を得ながら DOI を付与する。この時、DOI を与えたデータベースについてその引用の仕方等を検討しつつ DOI を付与していく。</p> <div style="display: flex; justify-content: space-around;"> <div data-bbox="411 757 783 1104"> </div> <div data-bbox="807 757 1257 1070"> <p>地質・地質系データバンク</p> <ul style="list-style-type: none"> ● 地質図、衛星データ（画像）、岩石や鉱物などの情報、地質調査結果や関連文献など ● 複数のデータベースに対し、地図上で統合的に検索・表示することが可能 </div> <div data-bbox="807 1093 1257 1406"> <p>物質・材料系データバンク</p> <ul style="list-style-type: none"> ● 有機化合物のスペクトルデータベースや熱物性データベースなど ● 複数のデータベースに対し、化学式、周期律表、物質名等を使った統合的検索が可能 </div> </div>
<p>実施体制</p>	<p>情報技術研究部門を技術中心体制として、産総研全体でデータバンクを所掌する事務局の国際標準推進部、その元でデータバンク各事業を遂行する各部門、特に地質調査情報センターおよび計量標準総合センターなどのデータ作成部門と連携して実施する。</p> <div style="text-align: center;"> <p>担当</p> <pre> graph TD A[産総研データバンク] --- B[情報通信・エレクトロニクス分野(担当)] A --- C[国際標準推進部] B --- D[情報技術研究部門] A --- E[コンテンツ提供研究ユニット] A --- F[地質調査情報センター] A --- G[計量標準] C --- H[事務局] </pre> </div>

表 3-4-5-2. 想定される主な検討課題

<p>アクセスの持続性の保証</p>	<p>③ DOI 登録後のデータ取り下げの取り扱いについて</p> <ul style="list-style-type: none"> ・独立行政法人等の研究所は中期計画に縛られるので、持続性の保証について厳密には担保はできない。 ・データの取り下げの意図はないものの、維持コストが出なくなる可能性はあるので、データの引き取りやアーカイブ等の制度が併せて提供されるとありがたい。(国会図書館とか?)
<p>DOI 登録する対象データの粒度について</p>	<ul style="list-style-type: none"> ・研究所内部でも検討中であり、現場の意見や技術的な難度、メタデータ設計の手間等の勘案してプロジェクトの中で議論できれば良い。 ・また、同じような事情の研究機関等が多いと思われるので、事情を共有・交換して互換性の高い解法を定めたい。 ・データベース全体に doi を与え、その中でさらに名前解決をして個々のデータに対する識別が与えられるような標準的な仕組み (RDA?) があると良い。
<p>DOI のランディングページの要件の検討について</p>	<ul style="list-style-type: none"> ・検討中で本プロジェクト内で議論を行いたい。 ・産総研としては独自の landing page を保有・提供する必要を感じているので、そういう事情を共有できる組織があればありがたい。 また、組織の都合を維持しつつ landing page をホスティングしてくれるような環境があるとありがたい。 ・一方、産総研では、CKANを運用しつつあるが、このページを landing page にして、個々のメタデータ表示項目をDOI の表示項目にするとということを検討している。このような手法の是非についても議論・検討したい。 
<p>研究データの種別ごとの取り扱いの検討について</p>	<ul style="list-style-type: none"> ・データの単位によって扱いが大きく異なるので未議論 ・Doi を付与することによって生じるコンテンツ作成側の負担について (内部でも doi をきちんと使わないといけない) ・メタデータ項目について 必須が多すぎないか? 等の議論 (まだ始まったばかり)

3.4.6 情報通信研究機構

表 3-4-6-1. 参加機関概要

プロジェクト責任者	統合データシステム研究開発室 村山泰啓
実施内容	<p>①DOI 登録するデータの分野、データ種類、登録件数</p> <ul style="list-style-type: none"> ・超高層物理学分野 ・地磁気データ、電離層データ、オーロラ画像データ、衛星データ ・約 10 件(実験プロジェクト中) <p>②登録した DOI の活用方法</p> <p>データ引用の実現を目指し、学术论文にデータの DOI を記載することの紹介・普及活動を行っていく。将来的には、データ提供者・データセンターが、その活動に見合った還元・対価が得られるような仕組みの構築を目標とする。</p>
実施体制	<p>【取りまとめ機関（主なメンバー）】</p> <p>◎ 情報通信研究機構・統合データシステム研究開発室 村山泰啓、木下武也、渡邊堯</p> <p>【下位機関（主なメンバー）】</p> <p>(a) 京都大学大学院理学研究科・地磁気世界資料センター 能勢正仁、小山幸伸、家森俊彦</p> <p>(b) 情報通信研究機構・電離圏宇宙天気世界資料センター 石井守</p> <p>(c) 国立極地研究所・オーロラ世界資料センター 門倉昭</p> <p>(d) 宇宙航空研究開発機構・宇宙科学衛星資料センター 篠原育</p> <p>【責任者の役割】</p> <ul style="list-style-type: none"> ・取りまとめ機関の責任者である村山は、JaLC との連絡を行い、DOI のデータ登録・更新・削除のテストを実施する。 ・下位機関の責任者は、4 機関を代表して能勢がその任に着き、村山と密に連絡をとりながら DOI 登録の実務を担当する。 ・各下位機関は、取りまとめ機関の統合データシステム研究開発室が用意する内部データ処理システムを通して、各機関が保持するデータの DOI 登録を JaLC に対して行う。

表 3-4-6-2. 想定される主な検討課題

<p>運用フロー 取りまとめ 機関につい て</p>	
<p>アクセスの 持続性の保 証</p>	<p>③ DOI 登録後のデータ取り下げの取り扱いについて 本プロジェクトでは、DOI に紐づけられる URL にはランディングページを配置するため、取り下げられたデータは、取り下げた事実、理由などをを明記する HTML ページを配置して DOI 自体は破棄しない運用が考えられる。</p>
<p>DOI 登録す る対象デー タの粒度に ついて</p>	<p>データ種別 1 つあたり 1 つの DOI を付与、とする案を現在、検討中。 (例：「地磁気 DST 指数」に 1 つの DOI、など)</p>
<p>DOI のラン ディングペ ージの要件 の検討につ いて</p>	<ul style="list-style-type: none"> ・ランディングページ例(下図) ・バージョンing：なお、データのバージョンが更新された際にはバージョンingの経歴をランディングページに明記する案を考えている。(バージョンingも DB で管理する)

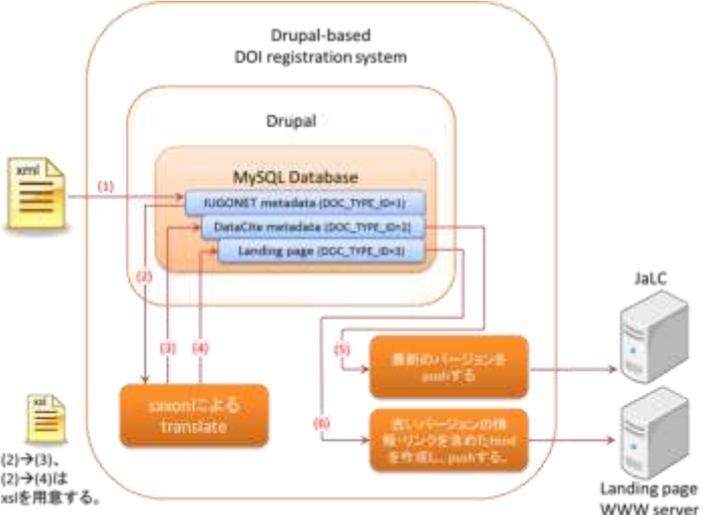
	
<p>研究データの種別ごとの取り扱いの検討について</p>	<ul style="list-style-type: none"> ・ 時間的に変化する自然現象を、時間（準）連続に測定するデータが対象（Dynamic Data）。 ・ 時間経過に即したデータ追加は原則、データの同一性に影響しないものとして管理。（バージョンニングしない）。 ・ データ取得・処理や補正、誤差の修正などはバージョン更新と考えバージョン管理する（下図）。  <p>(2)→(3)、(2)→(4)は xslを用意する。</p>

表 3-4-6-3. プロジェクト計画に対する意見等

前述のように、JaLC へのデータ登録のためのプリプロセス用システムを開発している。このシステムの JaLC 用メタデータ生成部分の詳細仕様を確定するため、JaLC で受け付けるメタデータスキーマ、メタデータファイルの受け付け方法（インターフェイス）の詳

細をなるべく早期に確定・公開してほしい。

3.4.7 千葉大学附属図書館

表 3-4-7-1. 参加機関概要

プロジェクト責任者	利用支援企画課 三角太郎
実施内容	<p>1.一般的な研究データ登録フロー：図書館と関連部署との連携した運営フローを想定。</p> <p>2.外部資金の助成による研究成果の登録フロー：学内の研究推進部門と連携した運営フローを想定。助成機関の OA ポリシーや研究不正対応のガイドライン（注2）を踏まえる必要あり。</p> <p>3. 観測データ等のアーカイブ：関連部署と連携した運営フローを想定。</p> <p>いずれもメタデータは、機関リポジトリへ登録の上、DOI を付与することを想定している。しかし、現状の機関リポジトリの規格では、データフォーマットに対応できない。</p> <p>本プロジェクトで得た知見については、本学ばかりでなく全国の大学図書館と共有し、学術情報基盤のさらなる充実につなげたいと考えている。そのため以下を実施する。</p> <p>1) システムの検討</p> <p>国内外の先進事例等の情報を探りながらメタデータのスキーム、システム仕様、運用フローについて検討をすすめたい。DOI の登録システムについては、どのように進めるか未定であるが、本プロジェクトを通して最適な方式を探りたい。</p> <p>2) 学内の研究データの実態調査</p> <p>本プロジェクトの一環として、学内において研究データがどのように生産され、どのように蓄積、保管、共有されているのか実態を調査し、研究データへの DOI 付与のデマンドについて検討する。</p> <p>3) 成果の共有</p> <p>国公立大学図書館協力委員会と国立情報学研究所による機関リポジトリ推進委員会では、研究データへの対応にむけて、本格的な検討を開始している。研究データ検討の WG に本学からもメンバーを派遣しており、そちらとも密に連携しながら成果の共有をはかりたい。</p>

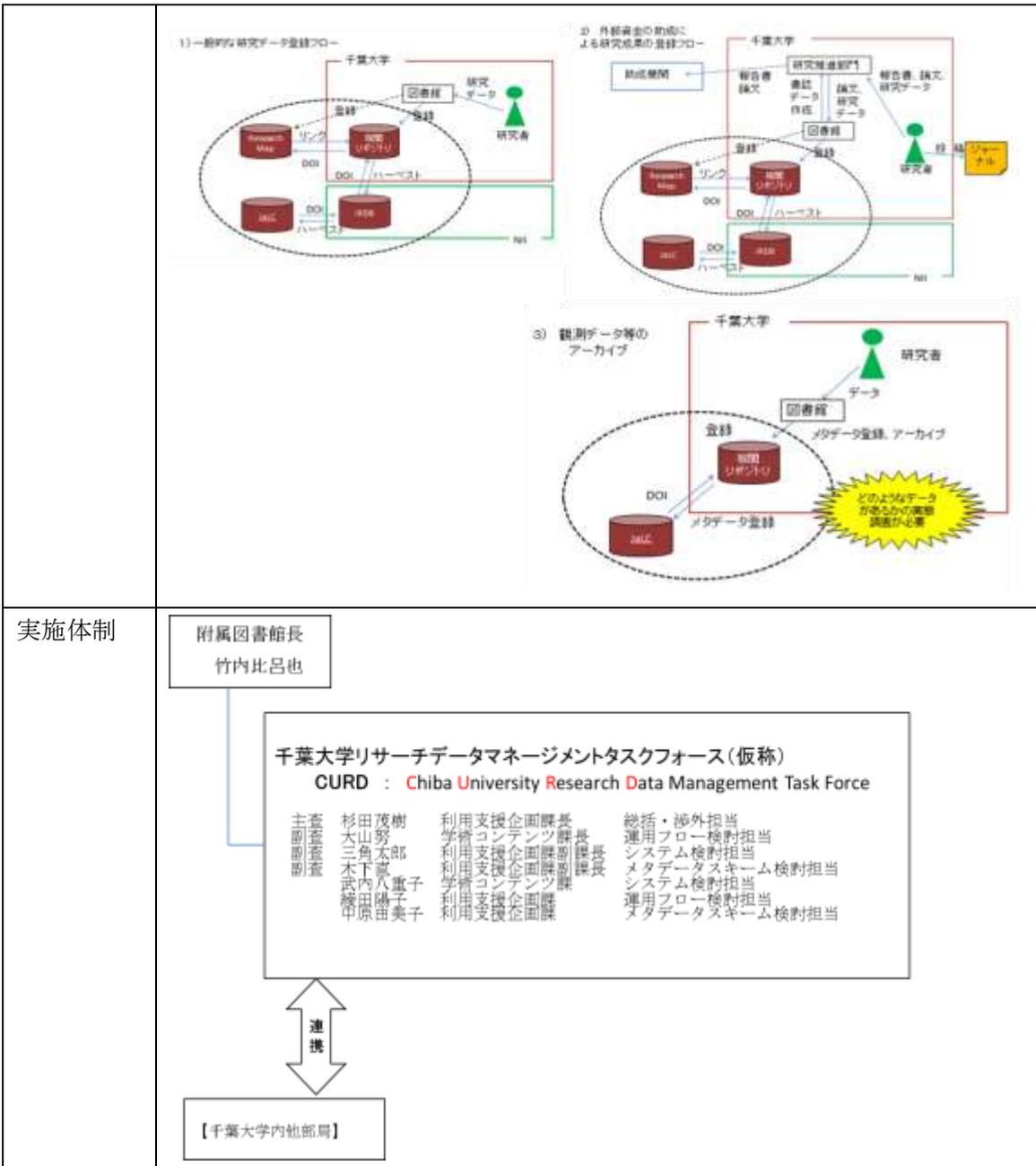


表 3-4-7-2. 想定される主な検討課題

<p>アクセスの持続性の保証</p>	<p>② DOI 登録後のデータ取り下げの取り扱いについて</p> <p>ポリシーをどのように定めるか、そしてそのポリシーをどのように運用に落とし込み、研究者に周知するかが課題である。データ登録後のデータ取り下げの問題は、機関リポジトリでも、既に起きている問題である。機関リポジトリで登録したデータは、通常、自機関内のサーバーだけでなく、OAI-PMHによるハーベスティングで、メタデータが IRDB 等の別サーバーにも流れているため、取り下げデータを単純に削除するのではなく、削除履歴を記した</p>
--------------------	---

	<p>データと差し替える（所謂「墓標」）。</p> <p>研究データの場合も同様の作業が必要になると考えているが</p> <p>1) 運用面については</p> <p>① データ登録ポリシーを作成し、そのなかに削除の条件も明記する。ポリシーは、「研究活動における不正行為への対応等に関するガイドライン」（注）とも整合性をもたせる。</p> <p>② ポリシーとは別にポイントをまとめた、リーフレットも作成し、研究者へ周知する。</p> <p>2) 作業面については</p> <p>① 取り下げデータは、非公開スペースを用意し、所定の期間、保管する。非公開スペースへのアクセスについても、データ登録 ポリシーに定める。</p> <p>② 削除履歴を記したデータを作成し、取り下げデータと差し替える。</p> <p>なお、登録および管理についても、研究者本人が登録するのか、機関内の担当部署が登録管理するのか、フローの整理が必要である。</p> <p>注）「研究活動における不正行為への対応等に関するガイドライン」（平成 26 年 8 月 26 日文部科学大臣決定）（PDF:1238KB）</p> <p>http://www.mext.go.jp/b_menu/houdou/26/08/_icsFiles/afieldfile/2014/08/26/1351568_02_1.pdf</p>
<p>DOI 登録する対象データの粒度について</p>	<p>データ粒度は、研究分野、データの種類により変える必要がある。例えば、生物種の記載データは粒度も明快であるが、観測データの場合、データは連続的であり、粒度をどのように設定するか判断は容易ではない。またデータ粒度ばかりでなく、データサイズも大きな問題であり、特にリモートセンシング等の観測データでは、極めて巨大なものになる可能性があり、その保存をどのように考えるかも検討しなければならない。</p> <p>本学は総合大学であり、所属研究者の研究分野も多岐にわたり、理工系・医学系から社会科学系・人文系にいたるまで、多くの研究者が所属している。その特色を生かし、学内の研究者、特に大型の外部資金獲得者を中心にヒアリングを行い、それぞれの研究分野におけるデータの扱い方や必要な粒度について検討する。</p>
<p>DOI のランディングページの要件の検討について</p>	<p>ランディングページに記載する内容を定めるためには、それぞれのデータに対するメタデータの記述のスキームを定めなければならないし、分野により、データの種類により必要なエレメントは異なってくるであろう。</p> <p>大学図書館は、機関リポジトリ等でメタデータ管理のノウハウは蓄積してきている。問題はメタデータのフォーマットであると考えている。フォーマットが、どの程度の種類必要になるかを</p> <p>① 海外の先行事例調査</p>

	<p>② 所属教員へのヒアリング等を通して検討する。</p>
<p>研究データの種別ごとの取り扱いの検討について</p>	<p>データ登録ポリシー、メタデータ・フォーマットにも深く関わる問題であるが、</p> <ul style="list-style-type: none"> ・追加時にメタデータに更新履歴を記せば良いのか？ ・更新前のデータはどのように扱うのか？ ・どのようなデータの場合はデータ追加を認めるのか？ ・追加作業は誰が行うのか？（研究者本人？担当部署？） <p>など、検討すべき課題は多く、分野によって異なることが予想される。研究分野の粒度も今回検討しなければならない課題であり、メタデータの記述ルール、データの保存ルールをどこまで細分化するか、海外の先進事例なども参照しながら検討したい。実際のメタデータの記述については、研究者以外への入力も想定し、マニュアルを整備したい。</p>

表 3-4-7-3. プロジェクト計画に対する意見等

研究データについては、1)オープンアクセス、2)研究倫理の二つの方向から、組織としての研究データ管理が求められている。前者はオープン化を指向するものであるから本質的にオープンである。一方で後者は成果発表時点のデータの保管（凍結）が求められ、むしろクローズな運用が求められる。両者はアプローチは大きく異なるが、いずれにおいても、メタデータを記述し識別子として DOI を付与することが、安定的なデータ管理のためには重要である。

本プロジェクトには、大学図書館として、また総合大学として参画したいと考えている。特にメタデータ入力や本文データのコンテンツ管理などの業務において、大学図書館として何ができて何ができないのかを明確にし、大学における運用のモデルを構築したい。そのために、今回のプロジェクトの成果をもとに、DCC 等の海外関連機関や、海外の研究図書館への実地調査も行いたい。

3.4.8 独立行政法人 物質・材料研究機構

表 3-4-8-1. 参加機関概要

プロジェクト責任者	企画部門科学情報室 谷藤幹子
実施内容	<p>「福島除染デジタルライブラリー」</p> <ul style="list-style-type: none"> ・物質・材料研究機構(NIMS)と日本原子力研究開発機構(JAEA)の協同研究プロジェクト「Cs 脱離機構解明と脱理法の開発」で収集したデータ（およそ 1000 件）を DOI の付与対象とする。 ・本プロジェクトの目的は以下の通りである。(1)セシウムの吸着のメカニズムの研究。(2)セシウムの脱離機構の研究。(3)上記の研究成果を生かして、Cs の吸着した汚染土壌の減容化手法の開発。特に除染作業で発生する 3×107m3 の土砂の減容化に向けて、実用可能な手法の開発を目指している。本プロジェクトは、復興予算で運営されており、「プロジェクトの成果の社会への還元」が強く求められている。 ・本プロジェクトの成果を実際の除染作業に生かすには、研究成果を学術ジャーナルで英語の論文として発表するだけでなく、日本語での情報を広く流通する必要がある。このため、研究プロジェクトでの成果物を、NIMS で運用するセルフアーカイブシステム"imeji"および"PubMan"を利用して、「福島除染デジタルライブラリー」として公開することを予定している。また、それらのデータのうち、恒久的に流通するのがふさわしいものに DOI を付与する予定である。今回の研究データ登録実験プロジェクトで実施するのは、この DOI 付与の部分となる。 ・本研究プロジェクトでは、福島県相馬郡飯舘村において、ため池の底を含む様々な場所から土壌を採取し、これらの土壌を深さ方向と粒径に分類して放射能を測定している。このような深さをふくむ位置情報、粒径と関連づけた放射能濃度、エネルギースペクトル等を数値(csv)およびグラフ(png)、地図や航空写真と放射能濃度や線量情報を合成した図(png)を公開しリンクを張る。また様々な減容化技術の特徴をまとめた日本語のマニュアル類を公開する。特に DOI を付与することにより、再利用しやすい形で研究成果の公開を行う。登録件数は、約 1000 件程度を予定している。
実施体制	<p>本プロジェクトは、以下の体制で行う。</p> <ul style="list-style-type: none"> ・アプリケーション担当： 田邊浩介 <p>収集した研究データに対して DOI を付与するためのアプリケーションの開発と運用を行う。</p>

<ul style="list-style-type: none"> ・研究プロジェクト連絡担当：河野健一郎 <p>JAEA をはじめとする研究プロジェクト参加機関との連絡・交渉を行う。</p> <ul style="list-style-type: none"> ・プロジェクト進行総括担当：谷藤幹子、サンヨト フェニー <p>本プロジェクトの全体の進行管理を行う。</p> <p>また、研究データ登録に用いるアプリケーションであるセルフアーカイブシステム"imeji""PubMan"については、開発元の Max Planck Digital Library と協力し、追加機能の開発や運用を行う。</p>

表 3-4-8-2. 想定される主な検討課題

運用フロー 取りまとめ 機関について	<p>福島を除染に関するデータは社会的な影響が大きいので、取り下げる可能性がないものに限って DOI 登録を予定している。様々なデータが存在する中で、社会的にも広く共有される価値があり、取り下げる可能性がないようなデータのみに対して、DOI 付与を予定している。</p> <table border="1"> <thead> <tr> <th>データの種類</th> <th>アクセス範囲</th> <th>変更の可能性</th> <th>DOI付与</th> </tr> </thead> <tbody> <tr> <td>プロジェクト内部用データ</td> <td>プロジェクト構成員のみ</td> <td>取り下げを含めて自由</td> <td>No</td> </tr> <tr> <td>一般公開用データ</td> <td>制限無し</td> <td>最新のデータに更新の可能性あり</td> <td>No</td> </tr> <tr> <td>恒久保存用データ</td> <td>制限無し</td> <td>社会的な影響も考慮して取り下げや差し替えの可能性がきわめて低いデータ。</td> <td>Yes</td> </tr> </tbody> </table>	データの種類	アクセス範囲	変更の可能性	DOI付与	プロジェクト内部用データ	プロジェクト構成員のみ	取り下げを含めて自由	No	一般公開用データ	制限無し	最新のデータに更新の可能性あり	No	恒久保存用データ	制限無し	社会的な影響も考慮して取り下げや差し替えの可能性がきわめて低いデータ。	Yes
データの種類	アクセス範囲	変更の可能性	DOI付与														
プロジェクト内部用データ	プロジェクト構成員のみ	取り下げを含めて自由	No														
一般公開用データ	制限無し	最新のデータに更新の可能性あり	No														
恒久保存用データ	制限無し	社会的な影響も考慮して取り下げや差し替えの可能性がきわめて低いデータ。	Yes														
アクセスの 持続性の保証	<p>① 有期のプロジェクトで保有するデータの扱いについて</p> <p>② DOI 登録後のデータ取り下げの取り扱いについて</p>																
DOI 登録する対象データの粒度について	<p>以下の種類のデータのエントリーを予定している。</p> <p>(1)福島で撮影した写真。imeji で管理予定(http://imeji.org)。写真をまとめたアルバム単位で管理する。</p> <p>(2)線量率を地図上に表示したデータ。imeji で管理予定。ため池ごとなどのアルバムで管理予定。</p> <p>(3)日本語で書かれた、除染で発生した土壌の減容化の手法のマニュアル。pubman で管理予定(http://pubman.mpdl.mpg.de/)。</p>																
DOI のランディングページの要件の検討について	<p>保存・配信するデータフォーマットが多様になると思われるため(JSON, XML, テキスト、…)、HTTP での Content-Negotiation の対応は必須になると考える。</p> <p>また、フォーマットが明示的に指定されなかった場合、どのようなファイルフォーマットでの取得を希望するかをたずねるインターフェースが必要</p>																

	になる。
研究データの種別ごとの取り扱いの検討について	データの量については、社会的な影響等を考慮して厳選したデータのみを登録予定なので、問題となるほど多くない予定。またデータの追加についても発生しないように、注意をする予定。

表 3-4-8-3. プロジェクト計画に対する意見等

事務連絡をもっとしっかりしてもらいたい。

- ・キーパーソンには確実に連絡すること。
 - ・日時の変更がある場合は、どのように変更されたか、確実にわかるようにすること。
- 黙って変更された時間のみを連絡するのはだめです。

3.4.9 理化学研究所脳科学総合研究センター神経情報基盤センター

表 3-4-9-1. 参加機関概要

プロジェクト責任者	理化学研究所脳科学総合研究センター神経情報基盤センター 山口陽子
実施内容	<p>当センターでは、脳・神経科学に関する様々な情報ならびに研究データを体系的に捉え、整理・発信する為の仕組み作りに取り組んでいる。これらの活動の中で運用している 12 分野のプラットフォーム（データベース）ならびポータルサイト計 15 サイトでは、サイトごとに WS 予稿集、研究成果報告書、学会発表原稿（ポスター、プレゼンテーション）等の各種ドキュメントや学术论文の付随データ（実験入力データ、実験結果データ、数理モデル、オンラインデモ）、解析ツール、シミュレータ等の様々な研究データについて、メタデータと共に収集・管理・公開している[1]。また、当センターではこれらのデータベースの基盤システムとして XooNIps を使用しており、技術開発としてこのソフトウェアの保守開発ならびにフリーソフトウェアとしての公開も実施している[2]。</p> <p>本課題では、これらのデータベースにて公開している 18,000 個以上の研究データに対して各データベースのデータ運用ポリシーに沿った形で DOI を登録すると共に、これらの基盤システムである XooNIps に対しても DOI の円滑な登録・管理・相互運用に向けた機能の検討ならびに必要な応じた機能改修を行う。</p> <p>これらの成果として、各種ドキュメントや研究データに対して DOI を付与し公開することで、持続可能性を向上させることができ、文献引用時の URI として利用しやすくなることが期待できる。また、基盤システム XooNIps に対する機能検討ならびに改修の恩恵は、当センターの運用するデータベース群だけでなく、学術情報リポジトリとして慶應義塾大学、近畿大学をはじめ 10 以上の全国の図書館でも採用されている[3]ことから、ニューロインフォマティクス研究分野に留まらず幅広い場面においても今後の DOI の活用・用途が広まることが期待できる。</p> <p>[1] INCF Japan Node ポータルサイト http://www.neuroinf.jp/ [2] XooNIps 公式サイト http://xoonips.sourceforge.jp/ [3] XooNIps 研究会 https://niji.brain.riken.jp/xoonips/</p>

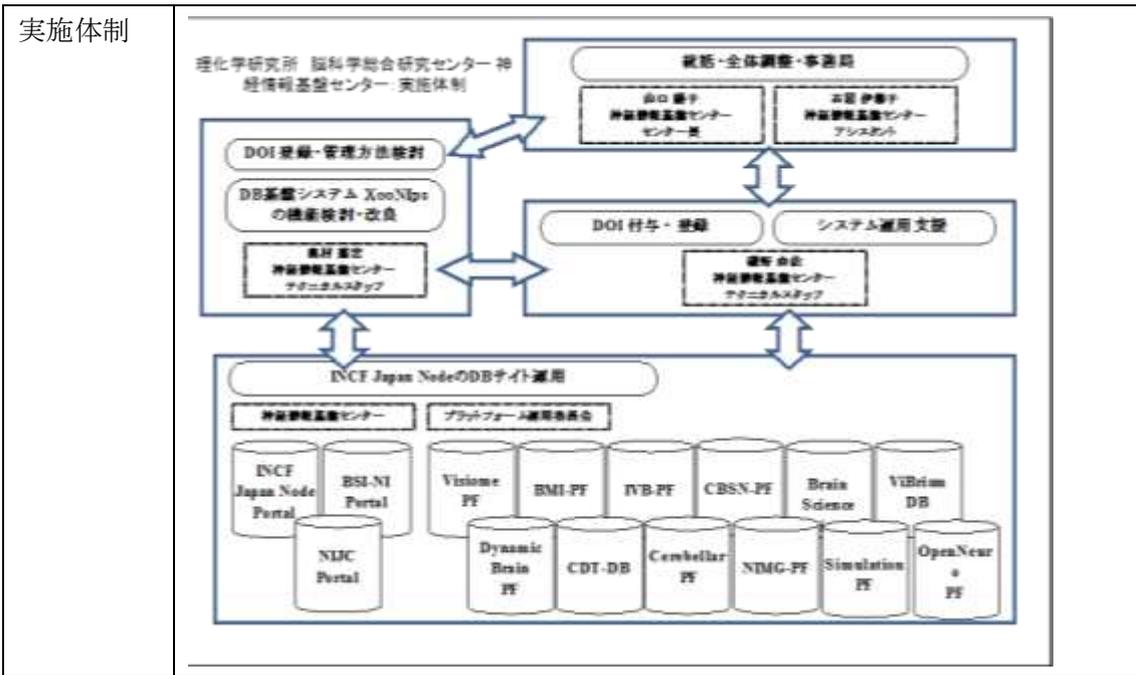


表 3-4-9-2. 想定される主な検討課題

<p>アクセスの持続性の保証</p>	<p>② DOI 登録後のデータ取り下げの取り扱いについて 取り下げが行われたことを明示した上でアクセスが保証されるようにしたい。登録したメタデータが参照できる必要はある。</p>
<p>DOI 登録する対象データの粒度について</p>	<ul style="list-style-type: none"> ・単独で意味をなすものについて登録したい。 データセットとしてのみ意味のあるものについてはそのセットについて登録したい。 ・XooNips に関してはデータベースに登録する時点で、アイテムごとに ID を発行しているのでそれに対応させる形で DOI を付与させたい。 <p>XooNips 以外のデータベースに関しては別途個別に検討する必要がある。</p>
<p>DOI のランディングページの要件の検討について</p>	<p>当組織においては基本的に非公開なものはないが、閲覧ユーザを制限する場合はありうる。</p> <p>XooNips では図のように登録時にメタデータを同時登録している。</p> <p>データの登録は承認制であり、承認後は基本的に登録者であっても内容の変更はできない。</p>

	
<p>研究データの種別ごとの取り扱いの検討について</p>	<ul style="list-style-type: none"> ・ 同じデータと見なされる場合は改版が判るようにしたい。 データ量が多い場合は基本的にはデータセットとしてまとめて一つとして扱いたいが、データの種類によっては枝版がつけられるとよい。 ・ 脳のデータであれば、ひとつの個体の脳全体のデータとその断面一枚ごとのデータ、脳波データであれば、一連の測定で記録された全体データと個別のデータ、などを関連付けられるとよい。

4. DOI テスト登録

4.1 DOI テスト登録まとめ

すべてのプロジェクト参加機関が DOI テスト登録を実施した。参加機関別の DOI 登録件数を表 4-1-1 に示す。

表 4-1-1. DOI テスト登録の結果（参加機関別登録件数）

	プロジェクト参加機関	JaLC DOI	DataCite DOI の場合
01	国立研究開発法人 科学技術振興機構	72	0
02	国立極地研究所	0	20
03	国立情報学研究所 地球環境情報統融合プログラム (DIAS-P)	23	5
04	国立情報学研究所 学術基盤	0	0
05	産業技術総合研究所	8	3
06	情報通信研究機構 ほか国内 WDC4 機関	4	0
07	千葉大学附属図書館	7	0
08	物質・材料研究機構	14	0
09	理化学研究所 脳科学総合研究センター神経情報基盤センター	12	0
—	参加機関の特定ができない DOI	1,019	36

4.2 DOI テスト登録要領

DOI テスト登録は、次に示す「DOI テスト登録要領」に従って行われた。

1. DOI テスト登録概要

(1) DOI テスト登録の趣旨

DOI 登録テストは、来年度初めに予定している研究データに対する DOI 本番登録に備え、必要な作業の確認、課題抽出、練習等を行うための作業です。そのため、プロジェクトにおいて DOI の実登録を行う機関は、原則としてテスト登録にご参加ください。なお、本登録に間に合うよう実施するためにスケジュール上で期間を限っていますが、テスト環境は継続して保持するため、その後も必要があればテストを行っていただいてもかまいません。

(2) DOI テスト登録の実施概要と注意事項

① DOI テスト登録の前提条件

ジャパンリンクセンターのテスト環境) を利用します。通常、外部ネットワークからテスト環境へのアクセスは許容していませんが、テスト期間に限って、事前に申請のあった IP アドレスからのアクセスのみを許容しています。DOI を実登録する環境とは異なる環境です。そのためログイン ID/PWD をプロジェクト参加機関毎に 1 つ払い出します。

② テスト登録インターフェース

テスト登録参加機関は、JaLC の登録インターフェース (web フォーム/web xml/API xml) を用いて、DOI のテスト登録ができます。

③ 登録した DOI は有効

DOI テスト登録は、本番の場合と同様に、LHS への登録まで行うため、DOI は有効となり、DOI からのリダイレクトができるようになります)。

④ テスト登録した DOI は全削除

テスト期間終了後は、テスト登録した DOI を全て削除します。

⑤ テスト登録した DOI 名の露出は禁止

テスト登録した DOI 名を印刷物に印刷したり、インターネットに掲載したりして、世の中に露出させないでください。

⑥ 使用する prefix (全機関共通)

a) JaLC DOI の場合

テスト登録用に、JaLC DOI の 5 つの prefix を払い出します。

[test prefix for JaLC DOI]

- 10.14977
- 10.14978
- 10.14979
- 10.14980 (※)
- 10.14981 (※)

JaLC DOI のテスト登録の際には、この中のいずれかをご利用いただきます。各機関が同じ prefix を用いますが、10.14977~10.14979 は下記「⑧ suffix について」のとおり suffix の先頭 3 文字を機関毎に使い分けることにより、DOI の重複を避けることができます。

使用する prefix は共通でも、機関毎にログイン ID が異なるため他機関により登録が上書されることはありません。ただし、他機関からの閲覧は可能です。

(※) suffix の先頭 3 文字を機関毎の指定文字列にすることが難しい場合には 10.14980、10.14981 をお使いください。この 2 つの prefix は「⑧ suffix について」の suffix ルールを適用しません。この prefix を使う時には事前に事務局に連絡をお願いします。

3/3(火)からテスト登録した研究データの JaLC-DOI のリゾルブが有効となります。その後、有効化は毎週火・木の週 2 回の運用となります。

b) DataCite DOI の場合

DataCite では、テスト登録用の prefix として「10.5072」を用いることになっております。

[test prefix for DataCite DOI]

- 10.5072

DataCite DOI のテスト登録の際には、10.5072 をご利用ください。

⑦ DataCite へのデポジット可能時期

平成 27 年 4 月上旬（予定）から、JaLC から DataCite へのデポジット（*）が可能となります。それまでの期間は、DataCite へのデポジットのフロー（【参考 1】DOI テスト登録のデータフロー参照）のうち、プロジェクト参加機関から JaLC までのデポジット（*）を確認していただくことになります。

⑧ suffix について

suffix の最初の 3 文字は、下表に示すようにプロジェクト参加機関の番号を付してください。

	プロジェクト参加機関	(例) JaLC DOI の場合	(例) DataCite DOI の場合
01	NBDC/JST	10.14977/01.～	10.5072/01.～
02	国立極地研究所	10.14977/02.～	10.5072/02.～
03	国立情報学研究所 地球環境情報統融合プログラム (DIAS-P)	10.14977/03.～ 10.14978/03.～ 10.14979/03.～	10.5072/03.～
04	国立情報学研究所 学術基盤	10.14977/04.～	10.5072/04.～
05	産業技術総合研究所	10.14977/05.～	10.5072/05.～

06	情報通信研究機構 ほか国内 WDC4 機関	10.14977/06.~ 10.14978/06.~	10.5072/06.~
07	千葉大学附属図書館	10.14977/07.~	10.5072/07.~
08	物質・材料研究機構	10.14977/08.~	10.5072/08.~
09	理化学研究所 脳科学総合研究センター 神経情報基盤センター	10.14977/09.~	10.5072/09.~

⑨テスト実施期間

平成 27 年 1 月 29 日～平成 27 年 5 月 26 日

研究データ メタデータ項目拡張 (【参考 4】参照) に対応した新しいメタデータの XML スキーマ、サンプル XML ファイルを平成 27 年 4 月 9 日に提示しました。

また、新しいメタデータでの登録は平成 27 年 4 月 20 日に開始します。下位互換があるので新しいメタデータの開始以降も従来のメタデータでの登録ができます。

⑩本番登録

本番登録は平成 27 年 6 月より開始します。

【参考 1】 DOI テスト登録のデータフロー

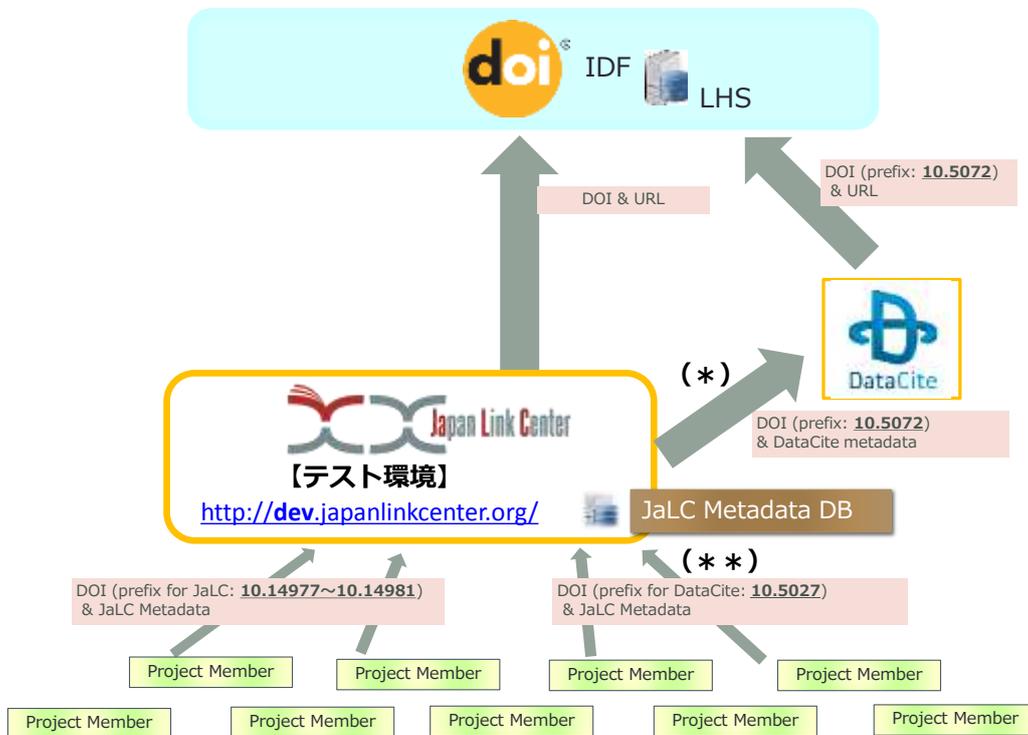
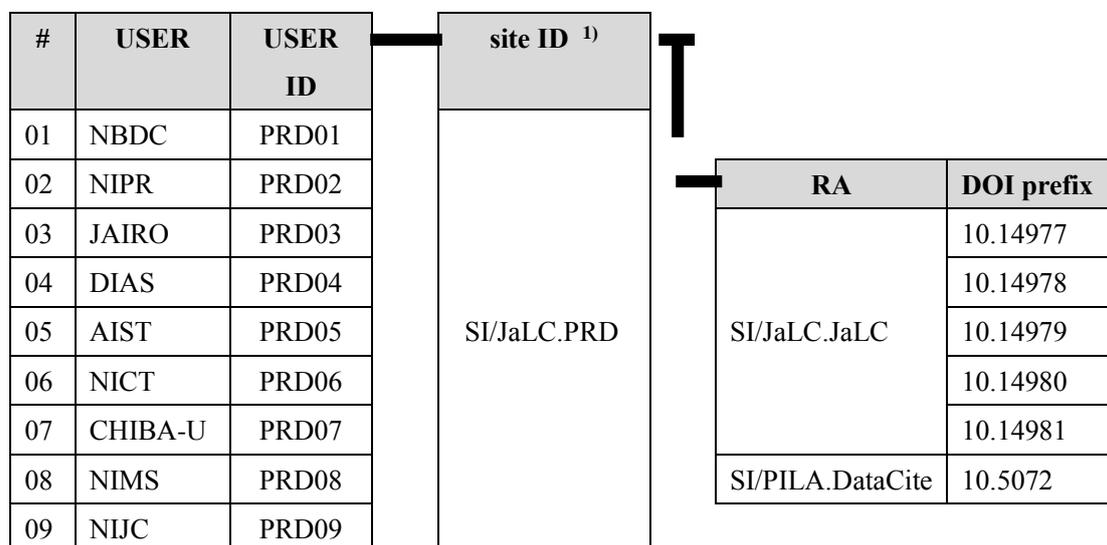


図. DOI テスト登録のデータフロー

【参考 2】 テスト登録における USR ID、site ID、DOI prefix の関係



【参考 3】 テスト登録と本番登録の対比

項目	テスト登録	本番登録
使用システム	テスト環境	本番環境
アカウント	テスト用	本番用
JaLC DOI 有効化	可能	可能
DataCite DOI 有効化	平成 27 年 4 月以降（予定）可能	可能
DOI prefix	テスト用	本番用

【参考 4】 研究データ メタデータ項目拡張

現在の JaLC システムの研究データ登録用のメタデータ項目は、DataCite Metadata Schema V3.0 に準拠して設計されている。しかし、その後、最新の DataCite Metadata Schema V3.1 がリリースされたため、JaLC システムでもこれにあわせ、メタデータ項目の拡張を行う。

拡張される項目を「参考資料」に示す。

¹⁾ サイト ID は、データの出典元・収集元を示す情報で JaLC へのデータの登録機関の名称・システム名を示します。JaLC 正会員毎に固有の文字列を JaLC 事務局が設定します。命名規則として、"SI"+"登録機関名" +"."+"サービス名"を基本法則としています。JaLC にデータ登録時には、サイト ID を必ず指定します。Web 画面では選択必須、XML ファイルでは入力必須項目です。

2. DOI テスト登録計画書の検討

各参加機関からご提出いただいた DOI テスト登録計画書のとりまとめを「別紙」に示す。テスト計画の内容等を検討のうえ、プロジェクトリーダーが、DOI テスト登録実施の可否を判断する。

3. DOI 登録テスト実施後の報告の観点

DOI テスト登録後の報告は、主として次の観点を取り入れていただきたい。

- ・ ワークフロー的ではうまく行くように見えたが、システム接続の際に明らかになった問題点
- ・ データの粒度の異なるケースの DOI 登録
- ・ Prefix をプロジェクト構成組織毎に分けた場合の長所/短所等

4.3 DOI テスト登録計画

各参加機関から提出された「DOI テスト登録計画」を以下にとりまとめる。なお、各機関から提出された原本は「附録資料 4」として添付する。

	1. テスト方針	2. テスト登録対象データ	3. 方法	4. 体制	5. 環境	6. スケジュール	7. 開始基準および終了基準	8. 仮定と除外、制限	9. テストの成果物	10. テスト報告書	11. その他															
1	NBDC/JST NBDC がサービスを提供している「生命科学系データベースアーカイブ」に収録されているデータセットに対して DOI を付与するテストを行う。 JaLC の登録インターフェースを使用し、すでに与えられている生命科学系データベースアーカイブのメタデータが DOI 登録にもそのまま活用できるか、不足の項目があるかなどを確認し、本格運用開始に備えることを目的とする。	◆データの種別：ライフサイエンス系データ ◆件数：92 データベース (428 データセット) ◆データ管理責任者：NBDC ◆テスト対象とする理由：DOI を付与することにより、データセットの二次利用が効率的に促進されるかどうかに関心がある。	◆RA：DataCite、JaLC の両方を希望 ◆登録手順：生命科学系データベースアーカイブに収録されている 100 件のデータセットを順次登録予定。 ◆JaLC との連絡、DOI 登録の方式の検討、データ登録、更新、削除、成果報告および報告書作成 ◆JaLC の登録インターフェース種類 (XML/Web)：未定	責任者：榎田研究員 (NBDC 研究員、ライフサイエンス DB 統合のための基盤整備) JaLC との連絡、DOI 登録の方式の検討、データ登録、更新、削除、成果報告および報告書作成 生命科学系データベースアーカイブチーム (3 名) メタデータ作成		準備 (登録システムの確認他)：2/1 ~ 2/6 データ登録・修正等 のテスト：2/7 ~ 3/24 報告書作成：3/25 ~ 3/31	開始基準： Suffix のデザイン決定 終了基準： 100 件のデータセットに対する DOI 付与。DOI が付与できないデータセットに対するその理由や状況の整理。	特になし	テスト報告書を作成予定。	実施状況 (方法など)、実施結果 (登録件数など)、評価、課題等	特になし															
2	国立極地研究所 データアーカイブに登録されたメタデータが DOI を付与するために必要なメタデータに対応するかの検討を行う。 また長期的な運用を図る上で各データに付与する suffix をつけるためのガイドラインを策定する。これらの作業を行うと共に、JaLC への登録システムの構築を行う。 また今回 DataCite メタデータが扱える時間及び空間的な粒度を検討する。	データの種別： ◆北極地域の定点観測で取得されている自動気象観測データを対象とする。 これらのデータは長期取得されているデータであるので、時間的粒度をどのよう区分するかの検討するのには有益なデータセットである。 ◆北極地域の定点観測地点群 (107 地点) で取得された気象観測データセットを対象とする。 このデータは、107 地点で取得された約 50 年間の複数要素をもつ気象観測データセットであり、空間的な粒度を検討するのには有益なデータセットである。	RA (DataCite) ◆登録手順：具体的な登録手順は、下記のとおりである。 データアーカイブシステムに登録されたメタデータより、DataCite のメタデータへ XMLformat 変換を行い登録準備用の XML メタデータの作成を行う、その後データに対して Suffix を付与し、実登録用のメタデータを作成する。 登録は、当初においては手動による XML 登録を進めるが、DOI 登録テスト期間中に Suffix を付与できた段階で、自動登録を行えるシステムの構築を行う。 ◆使用する JaLC の登録インターフェース種類 (XML/Web) の指定 ◆DOI 付与に関して取りまとめを行う予定。	研究統括 (進捗管理・報告)： 国立極地研究所・副所長 教授 榎本 浩之 実施責任者：(DOI 付与に関して取りまとめを行う) 国立極地研究所・北極観測センター 特任准教授 矢吹裕伯 ◆suffix をつけるためのガイドラインを策定 ◆研究データへの DOI 付与について、運用ポリシーおよび登録フローを検討する。 ◆DOI 付与する研究データの登録基準を検討する。 ◆DOI 付与に関して取りまとめを行う。 実施担当者 (DOI 付与に関して検討を行う) 国立極地研究所・極域データセンター 准教授 岡田 雅樹 国立極地研究所・北極観測センター 特任研究員 杉村 剛 国立極地研究所・北極観測センター 特任研究員 照井 健志 ◆suffix をつけるためのガイドラインを策定 ◆研究データへの DOI 付与について、運用ポリシーおよび登録フローを検討する。 ◆DOI 付与する研究データの登録基準を検討する。 ◆DOI 付与に関してシステムを構築する。	極地研では、今回テストに参加する、ADS システムの改良を行い、JaLC 2 への登録を行う。 基本的に、システムは、LINUX 上で開発された独自システムを用いる。登録環境およびハードウェア構成は下記の通りである。仮想サーバーを利用することで、単一の物理サーバー上に複数の仮想環境を用意し、様々なケースを実装することが可能である。	2 月下旬：DOI 用メタデータ作成システムの検証 3 月上旬：DOI 登録テスト開始 3 月中旬：DOI 登録テスト収量	開始基準： ◆研究データへの DOI 付与における運用ポリシーおよび登録フロー確立 ◆suffix をつけるためのガイドラインの策定が完了 ◆DOI 作成システムの構築 メタデータ作成システムの構築およびそのメタデータ検証 終了基準： ・JaLC2 への登録完了	特になし	◆研究データへの DOI 付与における運用ポリシーおよび登録フロー ◆suffix をつけるためのガイドライン テスト登録に利用したデータに関して下記の件に関して報告予定 ◆時間的粒度の問題点 ◆空間的粒度の問題点	◆研究データへの DOI 付与における運用ポリシーおよび登録フロー ◆suffix をつけるためのガイドライン テスト登録に利用したデータに関して下記の件に関して報告予定 ◆時間的粒度の問題点 ◆空間的粒度の問題点	特になし															
<div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px;"> <p>XML conversion</p> <p>ADS metadata → pre DataCite Metadata → DataCite Metadata → JaLC</p> <p>Suffix を付与</p> <p>当初は手動登録→最終的には自動登録</p> </div> <div style="border: 1px solid black; padding: 5px;"> <p>登録環境</p> </div> <div style="border: 1px solid black; padding: 5px;"> <p>ハードウェア構成</p> </div> </div>																										
3	"国立情報学研究所 地球環境情報統融合プログラム (DIAS-P)" DIAS-P において公開するデータセット約 200 個の中から典型的なデータセット数個を選び、識別子を付与する粒度・スライス・コレクション・多言語等にかかわる問題を検討するとともに、既存のメタデータからの変換やプロジェクト内での運用方式などを検証する。	データセットとしては、単独のデータセットであること、データの階層構造のテストを行うこと、日本語・英語の 2 言語でメタデータを持つことなどの点から、以下のデータセットを対象とする。	◆ RA (DataCite または JaLC) : JaLC に登録する。ただし日本語と英語のメタデータの共存に関する検討が必須。 ◆登録手順：メタデータを XML ファイルとして用意し、アップロードする。 ◆テストするパターン (テストケース)：新規登録、更新 ◆使用する JaLC の登録インターフェース種類 (XML/Web) の指定：Web ◆データタイプ、機能、組織などにより、テスト実施を分割する：該当なし"	"国立情報学研究所：取りまとめ 東京大学地球観測データ統融合連携研究機構：データセット選定、識別子 京都大学情報学研究所：メタデータ 海洋開発研究機構：運用体制検討"	"既存のメタデータ管理システムとの連携システムを独自開発する。既存の XML ファイルを XSLT で変換する予定。 テスト規模が小さいため、外部業者利用の予定はなし。"	"2 月：識別子の付与方法の仮決定。 3 月：メタデータの変換および JaLC ウェブインターフェース登録テスト。"	"開始基準：識別子付与規則の決定、メタデータフォーマット変換完了。 終了基準：全メタデータがエラーなく登録完了。"	"多言語メタデータの登録において、本書とは異なる形式を用いる可能性がある。 識別子は長期運用を指して決定するが、現段階ではあくまで仮決定にとどまり、今後の方針変更もありうる。"	特になし	"テスト報告書では、識別子の付与に関する方針、特に粒度・スライス・コレクション・多言語等に関する問題を述べ、その過程で得られた知見と課題を中心にまとめる。 ツールについては現段階では本書システムは構築しないため、成果としては報告しない予定。 また、プロジェクト内のワークフローとして構築する必要上、プロジェクト内部のシステム連携にも検討課題が多く、この部分についても成果がまとまれば報告する。"	特になし															
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 20%;">AGURAM 全球都市域マッピング地上情報データベース</td> <td style="width: 10%;">空間データ</td> <td style="width: 10%;">1</td> <td style="width: 10%;">東京大学 地球観測データ統融合連携機構</td> <td style="width: 50%;">単独のデータセットとしてのテストケース</td> </tr> <tr> <td>DIAS 衛星データセット</td> <td>衛星画像データ</td> <td>6</td> <td>宇宙航空研究開発機構</td> <td>種類としての階層構造を持つデータセットのテストケース</td> </tr> <tr> <td>AWCI 流域データセット (バングラデシュ・メグナラハ)</td> <td>時系列データ</td> <td>7</td> <td>DIAS Office</td> <td>地域的な階層構造を持つデータセットのテストケース</td> </tr> </table>												AGURAM 全球都市域マッピング地上情報データベース	空間データ	1	東京大学 地球観測データ統融合連携機構	単独のデータセットとしてのテストケース	DIAS 衛星データセット	衛星画像データ	6	宇宙航空研究開発機構	種類としての階層構造を持つデータセットのテストケース	AWCI 流域データセット (バングラデシュ・メグナラハ)	時系列データ	7	DIAS Office	地域的な階層構造を持つデータセットのテストケース
AGURAM 全球都市域マッピング地上情報データベース	空間データ	1	東京大学 地球観測データ統融合連携機構	単独のデータセットとしてのテストケース																						
DIAS 衛星データセット	衛星画像データ	6	宇宙航空研究開発機構	種類としての階層構造を持つデータセットのテストケース																						
AWCI 流域データセット (バングラデシュ・メグナラハ)	時系列データ	7	DIAS Office	地域的な階層構造を持つデータセットのテストケース																						

4	"国立情報学研究所 学術基盤"	データに関して、国立情報学研究所が運用している IRDB を介した DOI の付与機能ならびにプロセスを検証する。また、国立情報学研究所が OSS として提供しているリポジトリアプリケーション WEKO における DOI の付与機能ならびに利用機能におけるワークフローについて検証する。	<p>◆歴史的資料写真データ 使用範囲：公開データ 件数：1000 件程度 データ管理責任者：学術資源リポジトリ協議会 テスト対象とする理由：データリポジトリのみが機関内に併設される場合の機能的要件等に関する検証</p> <p>◆科学衛星観測データ 使用範囲：公開データ 件数：1000 件程度 データ管理責任者：金沢大学 テスト対象とする理由：文献リポジトリとデータリポジトリが機関内に併設される場合の機能的要件等に関する検証</p>	<p>◆ RA JaLC での機能検証を先行させ、実験中にシステム開発が準備できれば DataCite についても検証を進める。</p> <p>◆登録手順 国立情報学研究所の IRDB を介した DOI の付与</p> <p>◆テストするパターン リポジトリシステム内における Prefix 登録 (リポジトリ運用機関) リポジトリシステム内における DOI の付与 (リポジトリ運用機関) リポジトリから IRDB へのハーベスティング (リポジトリ運用機関&NII) IRDB におけるエラーチェック (NII) IRDB から JaLC へのハーベスティング (NII, JST) DOI からリポジトリへのリダイレクト (リポジトリ運用機関&NII)</p> <p>◆使用する JaLC の登録インターフェース種類 (XML/Web) の指定 IRDB から JaLC への OAI-PMH によるハーベスティング</p> <p>◆ データタイプ、機能、組織などにより、テスト実施を分割する場合は明記 テストするパターンの項を参照</p>	"NII：IRDB、WEKO における必要な機能の開発 リポジトリ運用機関：機関内調整、リポジトリシステムのアップグレード"	◆NII：IRDB ◆リポジトリ運用機関：WEKO"	"2月：システム準備、機関内調整 3月：IRDB を介した DOI の付与実験"	◆WEKO→IRDB： メタデータスキーマ junii2 によりメタデータの提供が可能になった時点で実験を開始。DOI を付与したサンプルデータで適切なリダイレクトを確認して実験終了。 ◆ IRDB→JaLC： OAI-PMH によるリポジトリからのメタデータ取得が可能になった時点で実験を開始。JaLC からのハーベス後、データ提供元で適切なリダイレクトを確認して実験終了。"	特になし	テスト報告書	実施内容、実施結果、成果、課題、等	特になし
5	産業技術総合研究所	"産業技術総合研究所の公開データベース (産総研データバンク) のうち、1) 有機化合物のスペクトルデータベース (SDBS) と 2) 熱物性データベースについて、doi を付与するテストを行い、コンテンツ提供者の意図がシステムも含め実現可能かどうか検証してフィードバックを行う。 また並行して、他の研究機関と情報や方針を共有して整合性のあるポリシー等の策定に有効な知見を得る。"	"テスト方針記載の2つのデータベース (進捗によってはテスト期間中に追加変更等の可能性あり) "	<p>◆ RA (DataCite) (メタデータは日本語も含む可能性があり記述可能性を要確認)</p> <p>◆登録手順 Prefix 以降の識別子の決定→テスト登録</p> <p>◆テストするパターン (テストケース)： 上記2つのデータベース。ただし登録単位は異なる。</p> <p>◆使用する JaLC の登録インターフェース種類 (XML/Web) の指定 XML/Web の両方の可能性がある</p> <p>◆データタイプ、機能、組織などにより、テスト実施を分割する場合は明記等 現在のところ想定なし"</p>	情報技術研究部門 小島が準備・調整およびとりまとめ、各データベースの管理者が参画する。	特段のものは用意していない。	"2月末 Prefix 以下の識別子の決定方法の検討と問題点の抽出と解決 ~3月中旬 上記方針に沿って登録するメタデータ情報の整備 ~3月末 テスト登録 一部は前倒して実施"	データベース全体でなく、その中の個別の情報ごとに doi を与えることを検討しているため、その方針の確定が作業開始に必要	"コンテンツ提供者の意図の反映度合い。例えば SDBS では各スペクトルデータ毎に doi を付与する一方で現状で各スペクトルには間接的にしかアクセスできないのでこの実現の可能性によってはテスト内容等に変更の可能性がある。"	報告書	内容未定	"要望： ◆JaLC の新登録システムのテストではなく、doi を付与するテストなので、メタデータの整備やその正当性の確認だけでなく、doi の登録単位の決定、Prefix 以下の識別子の策定、Landing Page の整備あるいは既存のものの変更の必要性など、doi を付与するにあたって必要となる組織の業務ワークフロー全体に対する情報共有・支援等の場であってほしい。 ◆同じく、Prefix 以下の識別子の与え方について組織内での管理のコストやシステムの支援の必要性についての懸念がある。他の組織で簡便な手法あるいはシステムを保有しているか、あるいは JaLC で保有・提供されているのなら利用・共有を検討したい。 検討事項： ◆データベース全体とその中の個別データの扱い。 ◆ (今回のテストで) サブ組織が別箇に Prefix を取得できるかどうか。またその時 RA が変わっていきいか。 ◆ Landing Page の形式やホスティング等について組織間で情報を共有したい。"
6	"情報通信研究機構 ほか国内 WDC4 機関"	<p>◆参加グループで保有・作成するメタデータ (IUGONET メタデータスキーマで作成したもの) からリポジトリ XML ファイルが問題なく作成できるかどうかをテストする。</p> <p>◆作成されたリポジトリ XML ファイルを JaLC2 システムに登録する。登録画面からの XML ファイルの手動アップロードおよび HTTP 共通インターフェースの両方をテストする。</p> <p>◆参加グループが作成した Landing Page URL と doi が結び付けられているか確認する。"</p>	<p>◆データの種類：地磁気変動データおよび電離圏観測データ</p> <p>◆件数：数件</p> <p>◆データ管理責任者：能勢正仁(京都大学理学研究科)、石井守・山本和憲(情報通信研究機構)</p> <p>◆テスト対象とする理由：本運用では、地磁気データおよび電離圏観測データを始めとする超高層物理学分野で用いられるデータを登録する予定である。そのため、上記のデータでテストを開始することは本運用へのスムーズな移行のために重要である。"</p>	<p>◆ RA: DataCite/JaLC (両者の prefix 希望)</p> <p>◆登録手順：上記1. のテスト方針に従って登録を行う。</p> <p>◆テストするパターン(テストケース)： 峰山・信楽地磁気観測所における地磁気変動データのメタデータ XML ファイル 稚内、国分寺、山川、大宜味観測所における電離圏観測データのメタデータ XML ファイル</p> <p>◆使用する JaLC の登録インターフェース種類 (XML/Web) の指定： 登録画面からの XML ファイルの手動アップロードおよび HTTP 共通インターフェースの両方</p> <p>◆データタイプ、機能、組織などにより、テスト実施を分割する場合は明記： 地磁気データについては京都大学理学研究科から、電離圏観測データについては情報通信研究機構からテストの実施を行う。ただし、テストの進行具合により、他機関(国立極地研究所、宇宙航空研究開発機構)から別の超高層物理学データのテスト実施を行うこともあり得る。"</p>	<p>"取りまとめ機関[主なメンバー] 情報通信研究機構・統合データシステム研究開発室 [村山泰啓、木下武也、是津耕司、渡邊聡]</p> <p>データ登録機関[主なメンバー] (a) 京都大学大学院理学研究科・地磁気世界資料センター [能勢正仁、小山幸伸、家森俊彦] (b) 情報通信研究機構・電離圏宇宙天気世界資料センター [石井守、山本和憲、加藤久雄] (c) 国立極地研究所・オーロラ世界資料センター [門倉昭] (d) 宇宙航空研究開発機構・宇宙科学衛星資料センター [篠原育]</p> <p>外部業者 株式会社情報試作室 [相良毅]</p> <p>役割分担 ◆取りまとめ機関の責任者である村山は、JaLC との連絡や報告を行う ◆データ登録機関の責任者は、4 機関を代表して能勢がその任に就き、村山および各機関の担当者と共に連絡をとりながら DOI 登録・更新・削除のテスト実務を担当する。</p>	"情報通信研究機構・統合データシステム研究開発室に内部データ処理システムを設置する。 京都大学をはじめとする各データ登録機関は、このシステムにアクセスして DOI 登録テストを実施する。"	"2月上旬～下旬：地磁気データの登録テストおよび内部データ処理システムの改善 2月下旬：参加グループ内でのミーティング 3月上旬～下旬：地磁気データ・電離圏観測データの登録テストおよび内部データ処理システムの改善、可能であれば他の超高層物理学データの登録テスト"	"開始基準：基本的に、本運用でも引き続き doi 付与を行いたい種類のデータについてテストを行う。 終了基準：上記1. のテスト方針の基準を満たした場合。"	特になし	テストが無事に成功した場合には、テストで用いた doi-URL の対応付けをそのまま本運用に引き継ぐ。	"テスト方針に沿って実施した内容やその結果を報告書に記述する。課題などは適宜メッセージングや JaLC のミーティングなどで報告・議論を行い、情報共有や積極的な解決を図る。課題の解決も報告書に記述する。"	<p>"必要な情報 ◆コンテンツ情報削除した場合、その後の対処方法(JaLC 側で「入力された doi に対応する URL はない」旨の表示が出るようになる?)。 ⇒物理削除ではなくダミー化され、削除ページにリダイレクトされることとなります。 http://japanlinkcenter.org/html/DOI_Not_Found.html (1/27 JST 回答)</p>



					◆各データ登録機関は、取りまとめ機関の統合データシステム研究開発室が用意する内部データ処理システム(メタデータの変換・Landing Pageの生成システム)を通して、各機関が保持するデータのDOI登録・更新・削除のテストを行う。内部データ処理システムの開発は、テスト実務担当者(主に村山、木下、小山、能勢)と株式会社情報試作室と協同で行う。																																																				
7	千葉大学附属図書館	"研究データ管理のプラットフォームは、基本的には附属図書館で運用している機関リポジトリシステムを想定している。本テストでは、機関リポジトリシステムを研究データマネジメントのプラットフォームとして使用する際の課題点を抽出、解決策を検討し、最終的には実運用の業務フローの確立を目指す。本学では既存の研究データのデータベースをもつわけではなく、研究成果のオープンアクセス化の対象の一部として、研究データをとらえている。そのためDOIの登録テストについても、研究データに限定せず、論文等のDOI付与のフローもあわせて検討する。特に1)登録ポリシー 2)運用フロー 3)メタデータスキーマについての検討を行う。"	"本テストで計画しているテスト登録は以下の二つのタイプである。 1) 既存の機関リポジトリシステム搭載のデータセット 現在、運用中の機関リポジトリシステム搭載のデータのうち(表1)、資料種類がデータセットの約5万2千レコードのうちの代表的なもの100件程度へのDOI付与を行い、データ登録へ要する業務量の見積りも行う。 2) 研究プロジェクトの成果物の登録シミュレーション 現在、本学のアカデミック・リンク・センターで実施されている研究プロジェクトの成果である研究データについて、機関リポジトリ上でのデータ保存・公開・DOI付与のシミュレーションを行う。本テスト内で、実際にDOIを付与するかどうかは、シミュレーションの検討結果によって決定する。アカデミック・リンク・センターのプロジェクトをとりあげるのは附属図書館サイドでハンドリングが容易なためである。アカデミック・リンク・センターは、アクティブ・ラーニングについての実践的な研究をミッションとしているが、本テストの実施母体である附属図書館と連携して研究を行っているため、プロジェクト研究活動の全体像を附属図書館サイドでも把握しているためである。 実際の登録時には、大学内で進められている研究プロジェクトについて、学内の関連部署とも連携しながら、成果の登録・保存・公開の業務を進めることになるが、本テストではプロジェクトのスタートから終了までの研究サイクルの各フェーズについて、実際にどのような業務を行うべきかを検討し、課題の抽出をはかる。 研究データ以外の生産物については、本テストの対象外であり、本テスト内ではDOIは付与しないが、業務フローとしては不可分のものであると予想されるため、検討の対象として含める。 具体的に想定している研究データとしては、画像、動画、数値データを想定し、その他に公表論文、講演資料、各種報告書を想定している。"	"◆RAはJaLCをもちいる ◆登録手順 <事前準備> ①登録ポリシーの策定(登録対象、登録情報等) ②運用フローの策定(プロトタイプ) ③テスト計画の策定(登録レコード選択、作業者・確認者の決定) ④メタデータスキーマの調整(機関リポジトリ側) ⑤登録環境の準備 <登録作業> ①JaLC登録テスト ②登録データの検証 ③作業に要した業務量(作業時間等)の検証 ④運用フローの見直し ◆使用するJaLCの登録インターフェース種類(XML/Web)の指定 XML、Webの双方のインターフェースをテストしたいと考えている。 ◆テストするパターン(テストケース)登録対象データにもとづきテストを行う。"	"主査 杉田茂樹(利用支援企画課長) 総括 渉外担当 副査 大山努(学術コンテンツ課長) 運用フロー検討担当 副査 三角太郎(利用支援企画課副課長) システム検討担当 副査 木下直(利用支援企画課副課長) メタデータスキーマ検討担当 武内八重子(学術コンテンツ課) システム検討担当 綾田陽子(利用支援企画課) 運用フロー検討担当 中原由美子(利用支援企画課) メタデータスキーマ検討担当"	登録作業は千葉大学附属図書館内の特定の業務用のPCからの作業を予定(SINET接続)している。 2/21-2/22 事前準備(手続、環境設定等) 2/23-3/9 登録テスト 3/10-22 検証 3/23-30 報告書作成"	"開始基準 登録開始日 終了基準 登録データの検証終了日"	テスト登録時には、同一データへの付与テストを複数回繰り返すことが予想されるが、実運用時には、テスト登録時のデータをクリアしてから、実登録に移行することは可能か(DOIも付与しなおす)。	"成果物として計画しているのは下記の三点である。 ①登録ポリシーサンプル ②運用フローサンプル ③メタデータスキーマサンプル 以上は、他大学でもサンプルとして参考にすることが可能な標準的なものの作成をめざす。"	"報告書の構成予定は以下のとおり 1)概要 2)背景 3)目的 4)環境 5)計画 6)結果 7)実験結果の検証・考察 8)まとめ 9)今後の課題 10)参考情報"																																															
					<table border="1"> <thead> <tr> <th>データタイプ</th> <th>レコード数</th> <th>パーセンテージ</th> </tr> </thead> <tbody> <tr> <td>Journal Article(学術雑誌論文)</td> <td>25,175</td> <td>28.10%</td> </tr> <tr> <td>Thesis or Dissertation(学位論文)</td> <td>894</td> <td>1.00%</td> </tr> <tr> <td>Departmental Bulletin Paper(紀要論文)</td> <td>7,497</td> <td>8.40%</td> </tr> <tr> <td>Conference Paper(会議発表論文)</td> <td>30</td> <td>0.00%</td> </tr> <tr> <td>Presentation(会議発表用資料)</td> <td>120</td> <td>0.10%</td> </tr> <tr> <td>Book(図書)</td> <td>35</td> <td>0.00%</td> </tr> <tr> <td>Technical Report(テクニカルレポート)</td> <td>35</td> <td>0.00%</td> </tr> <tr> <td>Research Paper(研究報告書)</td> <td>803</td> <td>0.90%</td> </tr> <tr> <td>Article(一般雑誌記事)</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Preprint(プレプリント)</td> <td>12</td> <td>0.00%</td> </tr> <tr> <td>Learning Material(教材)</td> <td>210</td> <td>0.20%</td> </tr> <tr> <td>Data or Dataset(データ・データベース)</td> <td>52,223</td> <td>58.40%</td> </tr> <tr> <td>Software(ソフトウェア)</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Others(その他)</td> <td>2,438</td> <td>2.70%</td> </tr> <tr> <td>合計</td> <td>89,472</td> <td></td> </tr> </tbody> </table> <p>表1 千葉大学機関リポジトリ登録コンテンツ</p>	データタイプ	レコード数	パーセンテージ	Journal Article(学術雑誌論文)	25,175	28.10%	Thesis or Dissertation(学位論文)	894	1.00%	Departmental Bulletin Paper(紀要論文)	7,497	8.40%	Conference Paper(会議発表論文)	30	0.00%	Presentation(会議発表用資料)	120	0.10%	Book(図書)	35	0.00%	Technical Report(テクニカルレポート)	35	0.00%	Research Paper(研究報告書)	803	0.90%	Article(一般雑誌記事)	0	0.00%	Preprint(プレプリント)	12	0.00%	Learning Material(教材)	210	0.20%	Data or Dataset(データ・データベース)	52,223	58.40%	Software(ソフトウェア)	0	0.00%	Others(その他)	2,438	2.70%	合計	89,472					
データタイプ	レコード数	パーセンテージ																																																							
Journal Article(学術雑誌論文)	25,175	28.10%																																																							
Thesis or Dissertation(学位論文)	894	1.00%																																																							
Departmental Bulletin Paper(紀要論文)	7,497	8.40%																																																							
Conference Paper(会議発表論文)	30	0.00%																																																							
Presentation(会議発表用資料)	120	0.10%																																																							
Book(図書)	35	0.00%																																																							
Technical Report(テクニカルレポート)	35	0.00%																																																							
Research Paper(研究報告書)	803	0.90%																																																							
Article(一般雑誌記事)	0	0.00%																																																							
Preprint(プレプリント)	12	0.00%																																																							
Learning Material(教材)	210	0.20%																																																							
Data or Dataset(データ・データベース)	52,223	58.40%																																																							
Software(ソフトウェア)	0	0.00%																																																							
Others(その他)	2,438	2.70%																																																							
合計	89,472																																																								
8	物質・材料研究機構	計画変更のため検討中																																																							
9	"理化学研究所 脳科学総合研究センター 神経情報基盤センター"	JaLC Web 登録画面よりデータ登録のテストを行い、問題点、課題等を確認し本番につなげていくものとする。	"神経情報基盤センターが管理する Web サイト、および各プラットフォーム(PF) Web サイトに、登録されているデータの中から対象データを選択する。 具体的には以下のコンテンツを予定している。括弧内のカテゴリ表記については現時点での見込みである。 ◆PF 紹介ポスター(汎用データ) ◆ニューズレター(汎用データ) ◆場所細胞とグリッド細胞記事(ジャーナルアーティクル) ◆Simulation PF オンラインシミュレーション(e-learning) ◆動向調査報告書(ジャーナルアーティクル) ◆Visiome PF Model(研究データ) 本番に向けて、登録データに対するカテゴリ設定、必要なメタデータの抽出、登録の各手法に役立てていく。"	"◆RAとしてはJaLCを使用する。 ◆登録手順、テストパターン等はデータ選定とともに検討する。 ◆JaLC Webを用いた登録を行う。"	"基本的に神経情報基盤センターで行う。各PFにある研究データについてはそれぞれの担当者、あるいはその権利者に対してそれらの登録について確認したうえで登録を行う。"	"◆公開データベース: XooNips、WordPress ◆データの抽出: ブラウザによるリストアップ、Excelによる一覧作成 ◆JaLC Web フォームへのブラウザによる手作業入力"	"1月末までの登録前 期間: 具体的なテスト実施要領の確立 2月中: 全対象データの登録実施 3月: 登録作業についての問題点の洗い出し、実施結果の取りまとめ"	"開始基準 ◆対象データの具体的な選定 ◆対象データがもつメタデータとJaLC登録カテゴリ選定および項目マッピング ◆DOI命名規則決定 終了基準 ◆全対象データの登録完了、もしくは登録ができなかった際の問題点の特定"	本テストは公開中データに対するものであるため、特に存在しない。	"◆登録データ一覧 ◆テスト手順書 ◆テスト結果一覧 ◆今後の課題・問題点等の一覧"	テスト成果物をまとめたものを実施結果報告書とする。	本計画は、実施状況等により適宜実施計画の変更等ともなるものである。																																													

4.4 DOI テスト登録計画の検討・承認

ミーティング（第2回）（2015年1月28日（水）開催）において、DOI テスト登録計画の検討を行った。

その結果、すべてのテスト計画が承認され、DOI テスト登録を開始することとなった。

4.5 DOI テスト登録実施報告

各参加機関から提出された「DOI テスト登録報告書」を以下にとりまとめる。なお、各機関から提出された原本は「附録資料 5」として添付する。

		1. テスト実施方針	2. テスト登録対象データと登録対象件数	3. テスト登録方法	4. 結果	5. 結果の検証と考察	6. 今後の課題	7. 本番登録の意向等	8. その他、参考情報等
1	NBDC/JST	NBDC がサービスを提供している「生命科学系データベースアーカイブ」(https://dbarchive.biosciencedbc.jp/) に収録されているデータセットに対して DOI を付与するテストを行う。JaLC の登録インターフェースを使用し、すでに与えられている生命科学系データベースアーカイブのメタデータが DOI 登録にもそのまま活用できるか、不足の項目があるかなどを確認し、本格運用開始に備えることを目的とする。	テスト登録対象データ： 生命科学系データベースアーカイブに収録されているデータセット全件（約 500 データセット（101 データベース）） 登録対象件数（登録実験済の数）：74 データセット（17 データベース）	登録のインターフェースは、Web 入力、XML 一括入力の両方を使用した。 RA は、JaLC を選択。	生命科学系データベースアーカイブですでに整備されているメタデータが、新規作成、修正することなく、ほぼそのままの形で JaLC の研究データ登録に活用できることを確認した（表 1）。 一方で、JaLC の必須以外のメタデータ項目には、生命科学系データベースアーカイブのメタデータ項目にないものもある（例、寄与者）。 ランディングページは、生命科学系データベースアーカイブで用意する既存のページを活用した。例、 doi.org/10.14977/01.rgp_gmap98_detail 生命科学系データベースアーカイブを対象とした場合、登録のインターフェースとして、Web 入力よりも XML 一括入力での効率的に作業が行えることを確認した。	生命科学系データベースアーカイブの既存のメタデータセットが DOI 登録の必須メタデータに活用が可能であり、また既存の各データセットの説明ページがランディングページとして利用できることを確認した。生命科学系データベースアーカイブのデータセットに対する JaLC の DOI 登録の本番作業は、6. 今後の課題に記述した各課題を解決後、可能になると考える。 RA は、当面、原則として JaLC を利用する見込みである。日本語メタデータの取扱の利便性や信頼性および今後の日本語を対象としたサービスの拡張性を考慮した結果、生命科学系データベースアーカイブのデータセットに対する DOI 登録では、JaLC が適当であると考えた。また、DataCite で DOI 登録を行う際に必要なメタデータ項目の確認、および DataCite が提供するメタデータを対象とする検索サービスである Metadata Search (http://search.datacite.org/) などをテストし、JaLC と比較して、DataCite で DOI を取得することでライフサイエンス系データの二次利用が促進される可能性があるかどうかを検討したが、DataCite を利用する明らかなメリットは見出せなかった。	Suffix のデザインをどうするか。人間が可読なもの（データセットの名称を含む Suffix）にするか否か（シリアル番号）について検討中。そのメリット、デメリットを再評価し、デザインの方針を確定したい。 JaLC のメタデータ項目の作成者および寄与者のそれぞれにどの情報を入力すべきか結論が出ていない。JaLC のメタデータ項目の作成者には、オリジナルデータの作成者の情報を入力すればよいのか、また、（生命科学系データベースアーカイブ固有の問題として）オリジナルデータからアーカイブ用にデータを作成した人物（主に NBDC スタッフ）や部署名の情報をどのように取り扱うか結論が出ていない。他機関の実例も参考にしたいと考えている。JaLC のガイドラインでも方針を示してほしい。 データ更新に伴い、新 DOI の付与や新旧データの保管が必要になった場合への対応が決まっていな。現在、生命科学系データベースアーカイブではデータの更新があった場合、更新後のデータのみアクセス可能になっている。データの更新等により新 DOI の付与や新旧データ両方の保管が必要になった場合、生命科学系データベースアーカイブの運用の基本方針の変更、ランディングページの作成など新たな作業の発生による作業工程の見直し、システムの改良が必要になる可能性がある。	希望する。ただし、上述の課題の検討が終了するまで、しばらくの間、本番登録は実施できない見通しである。	
2	国立極地研究所	データアーカイブに登録されたメタデータが DOI を付与するために必要なメタデータに対応するかの検討を行う。また長期的な運用を図る上での各データに付与する suffix をつけるためのガイドラインを策定する。これらの作業を行うと共に、JaLC への登録システムの構築を行う。また今回 DataCite メタデータが扱える時間及び空間的な粒度を検討した。DOI 付与検討時に、国立極地研究所では、suffix の作成指針の検討した。 suffix の作成指針 (AAA) + (.) + (YYYY) + (XXXXXX) として、 全ての文字は数字とする。 AAA: 数字で極地研内の各部署ごとに数字を割り振る。 YYYY : 登録年 XXXXXX : 自由とする。	下記 8 件 同じタイトルで異なる期間のデータ (1) Meteorological Data at Japanese Ny-Alesund Observatory, Svalbard in 1992 and 1993. (2) Meteorological Data at Japanese Ny-Alesund Observatory, Svalbard in 1994 and 1995. (3) Meteorological Data at Japanese Ny-Alesund Observatory, Svalbard in 1996 to 1998. (4) Meteorological Data at Japanese Ny-Alesund Observatory, Svalbard in 1999 and 2000. (5) Meteorological Data at Japanese Ny-Alesund Observatory, Svalbard in 2001 and 2002. (6) Meteorological Data at Japanese Ny-Alesund Observatory, Svalbard in 2003 to 2006. 複数の geolocation を持つデータ (7) Analyses of Ice Core Data from Various Sites in Svalbard Glaciers from 1987 to 1999 (8) Baseline Meteorological Data in Siberia (BMDS) Version 5.0	国立極地研究所では、今回 DOI の登録を行うシステムは、北極域データアーカイブシステム (ADS) を用いた。ADS では、メタデータ管理用システム ADS メタデータマネージメントシステム (ADS-AMS) に DOI の登録機能を実装した。テスト登録では、ADS に登録された、国内の北極域研究者が作成を行ったメタデータを用いてテスト登録を行った。DOI 登録用の JaLC2 メタデータの作成のために、AMS で利用しているメタデータスキーマを JaLC2 のメタデータスキーマを用いてマッピングを行った。AMS では、マッピング結果による、XML コンバージョンフィルターを作成した。実際に DOI 付与を行うデータについてのみ登録者（管理者）が XML ファイルを生成できるシステムの構築を行った。 <システム概要> ADS-AMS のシステム概要は、図 1 のとおりである。 国立極地研究所の北極域データアーカイブ (ADS) は、一元的なメタデータ管理とデータアーカイブ機能を持ったシステムである。全てのメタデータは、ADS メタデータマネージメントシステム (ADS-AMS) により管理されおり、ADS-AMS では登録されるメタデータおよび実データは全て研究者の手によって作成される。ADS-AMS では、メタデータ登録時に、ADS メタデータスキーマに従って、メタデータのエラーチェックを行い、エラー発生時にはユーザーにフィードバックを行う。ユーザーはフィードバック結果に従い、メタデータの該当箇所を修正し、登録を行わなくてはならない。 また国立極地研究所・北極域データアーカイブ（データセンター）におけるデータライフサイクルの担当者は表 1 のとおりである。 ADS-AMS では、国立極地研究所 DOI 付与ポリシーに従い、管理者のみが DOI を付与・登録を行う機能の構築を行った。今回 ADS では DOI のテスト登録のために AMS へ DOI 付与・登録機能を実装した（図 2、図 3）。これらの画面は管理者のみがアクセスが可能で、登録データを選択し、DOI 付与を行うことができる。 ADS における全ての登録データには、メタデータに記載された情報をもとに図 4 のランディングページが自動的に作成される。このページを通じて詳細なメタデータや実データ、実データのクイックルック、付属のドキュメント等にアクセスできる。 DOI 付与がなされたデータの引用方法は、ランディングページでは、図 5 のように表示される。 一方 DOI 付与されていないデータの場合は図 6 のように表示される。	全てのメタデータは、ADS への登録時に、ADS メタデータスキーマに従ってエラー処理されメタデータが登録されるので、DOI 登録のためのメタデータの不具合は生じないシステムである。よって特に問題は発生しなかった。ただし、JaLC2 のシステムの不備により、ADS の持つメタデータを修正して登録しなくてはならない状況が生じた。（subject が複数登録できない、description の文字数が 1000 文字限定） 今回のメタデータの登録により下記のテスト DOI を付与した。 (1) doi:10.5072/02.001.2013101009 (2) doi:10.5072/02.001.2013101008 (3) doi:10.5072/02.001.2013101006 (4) doi:10.5072/02.001.2013101005 (5) doi:10.5072/02.001.2013101003 (6) doi:10.5072/02.001.2013101002 (7) doi:10.5072/02.001.2013101311 (8) doi:10.5072/02.001.2013110702	今回 DOI の XML スキーマに従って、システムの構築を行ってきたが、登録時のエラーに対して発行されるエラーメッセージが分かりづらい。メタデータのエラー項目が XML のタグに対して日本語表記であり、エラーがどこで生じているのか分かりづらい。JaLC2 の DOI 登録システムが、XML スキーマに対してまだエラーを持っており、登録システムのバグ fix の必要性があると考えられる。	今回利用した JaLC2 研究メタデータスキーマで必須項目でない部分で利用していない項目があり、これらを今後利用することを考慮し ADS のメタデータのスキーマを見直す必要がある。（例：研究者 ID, FUND 等） ADS で利用するメタデータは、科学データのメタデータの流通促進のため、利用する vocabulary を基本的に国際的に利用されている vocabulary (NASA/GCMD/Keyword, ISO 等) に限定している。よって FUND 等にも国際的に流通可能な vocabulary を用いて利用を検討する。	国立極地研では、今後本番登録を行うために、所内での DOI 付与ポリシー等について検討を始めている。 【要望、報告事項】 1. 一部のメタデータ項目について、日英両方で記入できるようにしてほしい。例、publisher, rights. 2. DataCite で DOI 登録しようとしたら、Web 入力、XML 一括入力ともに「DOI(PREFIX)が登録されていません」と表示されて登録できない例があった（2015 年 4 月 16 日）。 3. XML 一括入力で、挙動がおかしな例があった。 例 1. XML ファイルによる一括登録（バッチなし）すると、「JaLC エラー データ不整合エラーです。 ログイン」と表示される。ただし登録はできている。続けて作業を使用したら、再ログインする必要があった。 例 2. XML ファイルによる一括登録（バッチなし）すると、「JaLC エラー データ不整合エラーです。 ログイン」と表示される。例 1 と違い、登録もできなかった。ただし、Web 入力でも登録ができた。	

3	国立情報学研究所地球環境情報統合プログラム (DIAS-P)	DIAS-P において公開するデータセット約 200 個の中から典型的なデータセット数個を選び、識別子を付与する粒度・スライス・コレクション・多言語等にかかわる問題を検討するとともに、既存のメタデータからの変換やプロジェクト内での運用方式などを検証する。	データセットとしては、単独のデータセットであること、データの階層構造のテストを行うこと、日本語・英語の2言語でメタデータを持つことなどの点から、表1のデータセットを対象とする。	既存の XML 形式メタデータを JaLC 形式メタデータに変換し、これをウェブおよび API 経由で登録する方法を用いる。メタデータの変換は現在のところ手動で行っているが、近い将来には XSLT による変換機能を開発する予定である。サフィックスについては、DIAS-P のデータセット番号を流用し、DIAS.<データセット番号>という形式に定めた。ただし階層構造を持つデータセットについては、DIAS.<データセット番号>.<枝番>のように、独自の枝番を付与することも可能とした。ランディングページについては、既存の DIAS 俯瞰検索システムの対応データページを用いた。	必須項目を中心に一部の任意項目を追加したメタデータを作成し、ウェブおよび API 経由で行ったテスト登録は成功した。また doi.org によるレゾルブ機能も、JaLC のプレフィックスについては動作を確認したが、DataCite のプレフィックスについてはまだ動作を確認できていない。	現在は DIAS のメタデータ項目の中から JaLC の必須項目を中心に登録しており、今後このメタデータをさらに詳細化するについては、メタデータ登録の具体的な利点との兼ね合いに基づき決める必要がある。	データセットの変更あるいは取り下げに関するワークフローはまだ検討が進んでいない。現状の DIAS メタデータ登録システムと、システム内のワークフローを考慮しながら統合を進めていく必要がある。また JaLC 側の多言語対応機能にはまだ不十分な部分も残っているが、多言語への対応は DataCite にはない JaLC 独自のニーズとして、引き続き改良に取り組む必要があると考える。ランディングページについては既存のウェブページを用いたが、今後はここを入口とした関連データ発見機能などの充実に課題である。	本番登録は行わず、テスト用プレフィックスでの登録を引き続き行う。	
4	国立情報学研究所 学術基盤	データに関して、国立情報学研究所が運用している IRDB を介した DOI の付与機能ならびにプロセスを検証する。また、国立情報学研究所が OSS として提供しているリポジトリアプリケーション WEKO における DOI の付与機能ならびに利用機関におけるワークフローについて検証する。	歴史的資料物写真データ 使用範囲：公開データ 件数：1000 件程度 データ管理責任者：学術資源リポジトリ協議会 テスト対象とする理由：データリポジトリのみが機関内に併設される場合の機能的要件等に関する検証 科学衛星観測データ 使用範囲：公開データ 件数：1000 件程度 データ管理責任者：金沢大学 テスト対象とする理由：文献リポジトリとデータリポジトリが機関内に併設される場合の機能的要件等に関する検証	RA : JaLC での機能検証を先行させ、その後 DataCite にも登録。□登録手順 : 国立情報学研究所の IRDB を介した DOI の付与□テストするパターン リポジトリシステム内における Prefix 登録 (リポジトリ運用機関) リポジトリシステム内における DOI の付与 (リポジトリ運用機関) リポジトリから IRDB へのハーベスティング (リポジトリ運用機関&NII) IRDB におけるエラーチェック (NII) ・ IR に登録されていないプレフィックス ・ selfDOI 要素の値が変更されている ・ JaLC 資料種別が変更されてしまう Niltype 変更 ・ selfDOI 要素の ra 属性種別(JaLC、CrossRef、DataCite)に存在しない ・ selfDOI 要素に ra 属性が存在しない ・ selfDOI 要素の値にマルチバイト文字が含まれている ・ selfDOI の形式ではない (サフィックスの使用できない文字を持つ場合も含む) ・ selfDOI 要素の値の長さが 300 文字を超えている ・ 既に削除されている selfDOI (削除 selfDOI が存在するのに selfDOI を持つ有効なアイテムが存在しない) ・ selfDOI 要素で必須とされる fullTextURL 要素が存在しない ・ 設定されている ra 属性では登録できない Niltype ・ 設定されている ra 属性における必須要素 xxx が存在しない ・ 設定されている ra 属性で当該要素に許される数を超えている ・ 設定されている ra 属性において当該要素の lang 属性は eng しか許されない IRDB から JaLC へのハーベスティング (NII、JST) DOI からリポジトリへのリダイレクト (リポジトリ運用機関&NII) データ登録までのフロー 1. 各リポジトリから OAI-PMH により IRDB にハーベスト 2. JST より IRDB をハーベストしデータ登録 テスト登録のための開発システム ・ 機関リポジトリソフトウェア WEKO の改良 ・ IRDB における JaLC 登録機能の改良*	必要とするシステム開発も完了し、リソルブに至るまでの一連の動作を検証できた。	データとしてのアイテムタイプの検証 実験中、データに対してコンテンツ型を Book としてメタデータを生成し、JaLC システムに送信したことがあった。今回はデータ DOI の実験であったために、JaLC 側で問題を検出できた。IRDB では、データ以外のコンテンツのメタデータも集約し、OAI-PMH により JaLC システムがハーベストする。したがって、実運用では今回のようなメタデータ付与のミスは発見することができない。リポジトリ運用側に慎重な操作が必要とされる。 JaLC 研究データの Creator の扱い 歴史的資料物写真データの DOI 登録において、JaLC 研究データのメタデータの必須項目である Creator が不足しているものがあった。写真で電子化した資料の Creator は、電子化史料と同様に資料現物の作成者とするのが妥当と考えられるが、作成者が不明な場合は空欄になり DOI を登録できなくなることが考えられる。このケースについて、どのように対応すべきかのガイドラインがあると望ましい。	1 機関複数 Prefix とする場合の準会員登録時の確認内容・申請書 従来は 1 機関 1 Prefix を想定した運用をしてきたが、例えば図書館が運用する機関リポジトリとは別に、学部や研究室、研究プロジェクトでデータリポジトリが運用されることが容易に想定される。これらのデータリポジトリは、将来的に他機関や他組織で運用が引き継がれる可能性がある。リポジトリの引越しを円滑に行い DOI の恒久性を維持するためには、同一機関内においても、予めリポジトリ毎に異なる Prefix をしておく方がよい。こうした同一機関において準会員機関が複数の Prefix を発行するためのワークフローを用意する必要がある。リポジトリが他機関に引越しする場合は異なり、同一機関内の複数のリポジトリのコンテンツが統合される場合には、1 つのリポジトリソフトが複数の Prefix を管理できる機能が必要である。今回実験した WEKO においても、現時点ではそうした機能は備えていないため、必要に応じて実装する必要がある。	実施予定	
5	産業技術総合研究所	材料データや地理空間データなど複数のデータセットを対象として、データの粒度の違いなど幾つかの場合を投入することで、doi の登録に必要な技術課題や決めるべきポリシーの課題等、コンテンツ作成者にとっての課題を明らかにする。	産総研の既公開データベースの一部 (SDBS および熱物性データベース)	一つの doi の空間を複数の担当者が別個に使うために、suffix の先頭にデータベース名を入れるなど(例 10.14977/05.tdbs-23732)して、名前の衝突を避けるようにした。Web インターフェイスによるものと XML 投入による方法を試みた。 SDBS については粒度の異なるデータを投入し、熱物性データベースについてはランディングページを試験的に構築できる等、産総研側のシステムを改造した。	既に報告済みの内容も含め概要を示す。 ・ 粒度の異なるパターンについて登録を行いアクセスを確認した。 ・ (同じデータに対し) JaLC と Datacite の両方で登録アクセスを確認した。 ・ 物質 1 つの場合に対するランディングページを試作してアクセスを確認した。	基本的には要求するパターンにおいて登録が可能であることが検証できた。 メタデータ項目が個人向けになっている。組織で登録する場合の具体例があるとありがたい。 項目が不明なものが散見された。ガイドラインや説明が充実するとありがたい。 ・ Related_Content 等 その他細かいフィードバックや機能要求 ・ JaLC の Web 入力と XSD エレメント関連表との対応関係の分かりにくさ ・ 登録メタデータのダウンロード機能、適切なエラーメッセージの表示等 ・ Creator タグなどでスキーマがあいまい。 ・ XSD に対する改善要求など(Date 文字列→Int 等) 産総研側のシステムに改造の必要性がある ・ ランディングページの生成 ・ 検索結果をポップアップで表示するスタイルの改修 ・ スペクトルに到達するための CGI 改修 ・ doi をシステム内で表示させるなど doi をシステム内部で利用する方法 ・ doi から得られる URL にパラメータが入るために不測のリンク切れ等が考えられるため、アクセスの保証を行うための方法など	doi の登録に対応した Landingpage の作成や、suffix の登録管理 (重複のチェック) 等を支援できるソフトウェアあるいは外部サービスがあるとありがたい。 地理空間データについては残念ながら取り組むことができなかった。また、産総研のオープンデータカタログと doi ランディングページとの連携については引き続き検討していきたい。	あり (ただし時期等不確定要素大)	

6	<p>“情報通信研究機構ほか国内WDC4機関”</p>	<p>参加グループで保有・作成するメタデータ(IUGONETメタデータスキーマで作成したもの)からリポジトリリクエストXMLファイルが問題なく作成できるかどうかをテストする。 作成されたリポジトリリクエストXMLファイルをJaLC2システムに登録する。登録画面からのXMLファイルの手動アップロードおよびHTTP共通インターフェースの両方をテストする。 参加グループが作成したLanding Page URLとdoiが結び付けられているか確認する。</p>	<p>データの種類：地磁気変動データ 件数：数件データ 管理責任者：能勢正仁(京都大学理学研究科) テスト対象とする理由：本運用では、地磁気データおよび電離圏観測データ、オーロラ観測を始めとする超高層物理学分野で用いられるデータを登録する予定である。そのため、上記のデータでテストを開始することは本運用へのスムーズな移行のために重要である。</p>	<p>作成したDOI登録システム、およびシステム内のデータの流れを図1に示した。システムの中核サーバーは、情報通信研究機構が運営するWDS/IPO(World Data System/ International Program Office)に設置されており、図1では中央上の黄色で示している。また、Landing Pageをホストするためのウェブサーバーも情報通信研究機構・WDS/IPOに設置されており、中央下の黄色で示している。時間は上から下に進行する。 1. JaLCからDOI prefix(10.14977/06、10.14978/06、10.5072/06)が割り当てられる。 2. 各データセンターは、DOI suffixの作成をDOI登録システムサーバーに対してリクエストする。 3. DOI登録システムサーバーは、リクエストを受けた時刻をシードとして13文字からなるランダムな文字列を生成し、これを「Unique & Random DOI suffix」としてデータセンターに通知する。 4. データセンターは受け取ったDOI suffixを、登録しようとしているデータのメタデータファイルに埋め込む。このメタデータは、IUGONETメタデータスキーマで記述されている。 5. DOI suffixを埋め込んだメタデータを、DOI登録システムサーバーに対して送信する。 6. DOI登録システムサーバーは、受け取ったメタデータの情報を元にして、(1)JaLCメタデータスキーマで記述したメタデータ、(2)Landing Pageを作成する。 7. DOI登録システムサーバーは、6(1)で作成したメタデータをJaLCに送信する。送信は、Web APIを利用している。 8. DOI登録システムサーバーは、6(2)で作成したLanding Pageをウェブサーバーへ転送する。4で作成するメタデータファイルの中には、Landing Pageに掲載するリンク情報を埋め込むことになっており、各データセンターが管理するデータやドキュメントのウェブページへの到達が可能となる。 9. DOI登録システムサーバーでは、各データセンターから受け取ったメタデータファイル中のDOI情報を元に新旧バージョンの管理を行っている。更新された日付情報や旧バージョンのLanding Pageへのリンクも最新バージョンのLanding Pageに表示されるようになっている。 図2の左は、DOI登録システムサーバーのトップ画面である。各データセンターの実務者は、DOI登録の全ての作業(DOI suffixの申請、IUGONETメタデータスキーマによるメタデータのアップロードなど)をここから行うことになる。図2の右は、今回のテスト登録で作成されたLanding Pageの一例を示している。データの簡単な説明、データ引用の際の表示方法、観測点の情報、地図上の位置が表示されている。また、その下にはデータセンターのデータやドキュメントへのリンクが掲載されており、ユーザーはこのリンクから実データへのアクセスが可能となる。最下部には、更新された日付情報や旧バージョンのLanding Pageへのリンクが表示されている。</p>	<p>テスト登録においては、上記のシステムは正常に動作した。今回は、下記の3つのDOIの登録を行い、Landing Pageの生成、そこへのDOIのアドレス解決が正常に行われていることを確認した。 Mineyama (sample 1) doi:10.14977/06.54c861b824d66 → http://www.info-proto.com/wds/10.14977_06.54c861b824d66.html Mineyama (sample 2) doi:10.14977/06.54e444fd7ca3 → http://www.info-proto.com/wds/10.14977_06.54e444fd7ca3.html Dst index doi:10.14977/06.550ea56ddb03b → http://www.info-proto.com/wds/10.14977_06.550ea56ddb03b.html</p>	<p>システムの構築やJaLC Web APIとのエラーコード対応などに労力がかかったものの、全般的なDOIのテスト登録を実施し、満足する結果が得られた。 今後より多くの機関が登録システムを利用することになると思われるため、登録者や登録システム開発者が各自トラブルシューティングを行うためのツールの整備に期待したい。たとえば、登録しようとするメタデータの整合性検証を行い、問題があればその箇所を特定できる対話的なツールが提供されれば、利用者の利便性が大幅に高まると思われる。”</p>	<p>【データセンター・DOI登録者の課題】 ・DOI suffixには意味を持たせずにrandomな文字列を採用したが、その妥当性・利便性・文字列の長さなどの検討。 ・データのアップデートが行われた際の過去のデータ保存。およびその過去のデータとDOIやLanding Pageとの対応付け。 ・DOI登録システムサーバーのメンテナンス。特にメタデータスキーマがアップデートされた際に、新しいバージョンへの対応に労力がかかる。 【JaLCの課題】 ・DOIリゾーブが有効になるまでの時間をもう少し(せめて半日から1日程度まで)短縮できないでしょうか。 【研究データDOI全体の課題】 ・DOI引用文化の・習慣の浸透。研究データのDOIも付与しただけでは意味がなく、論文に引用されて初めて価値が出てくる。</p>	<p>引き続き本番登録へ参加したい。</p>	<p>(1) DOI suffix について以下の論文誌の12ページ目の第一パラグラフにDOI suffixの方針についていろいろな例を紹介している。 http://eosps0.gsfc.nasa.gov/sites/default/files/ea_pdf/Sept_Oct_2012_color_508.pdf (2) Citationのフォーマットについて以下の記事では、データ引用の書き方をいろいろなフォーマットで表示している。(Figures 1 and 2)。データへのアクセス日を表記するように求めているフォーマットもある。この記事自身のreference (further information)では、""Retrieved""として、文献にアクセスした日付を表記している。 http://www.dcc.ac.uk/resources/how-guides/cite-data-sets</p>
7	<p>千葉大学附属図書館</p>	<p>研究データ管理のプラットフォームは、基本的には附属図書館で運用している機関リポジトリシステムを想定している。本テストでは、機関リポジトリシステムを研究データマネジメントのプラットフォームとして使用する際の課題点を抽出、解決策を検討し、最終的には実運用の業務フローの確立を目指す。本学では既存の研究データのデータベースをもつわけではなく、研究成果のオープンアクセス化の対象の一部として、研究データをとらえている。そのためDOIの登録テストについても、研究データに限定せず、論文等のDOI付与のフローもあわせて検討する。特に1)登録ポリシー 2)運用フロー 3)メタデータスキーマについての検討を行う。</p>	<p>本テストで計画しているテスト登録は以下の二つのタイプである。 1) 既存の機関リポジトリシステム搭載のデータセット 現在、運用中の機関リポジトリシステム搭載のデータのうち(表1)、資料種類がデータセットの約5万2千レコードのうちの代表的なもの100件程度へのDOI付与を行い、データ登録へ要する業務量の見積りも行う。 2) 研究プロジェクトの成果物の登録シミュレーション 現在、本学のアカデミック・リンク・センターで実施されている研究プロジェクトの成果である研究データについて、機関リポジトリ上のデータ保存・公開・DOI付与のシミュレーションを行う。本テスト内で、実際にDOIを付与するかどうかは、シミュレーションの検討結果によって決定する。アカデミック・リンク・センターのプロジェクトをとりあげるのは附属図書館サイドでハンドリングが容易なためである。アカデミック・リンク・センターは、アクティブ・ラーニングについての実践的な研究をミッションとしているが、本テストの実施母体である附属図書館と連携して研究を行っているため、プロジェクト研究活動の全体像を附属図書館サイドでも把握しているためである。実際の登録時には、大学内で進められている研究プロジェクトについて、学内の関連部署とも連携しながら、成果の登録・保存・公開の業務を進めることになるが、本テストではプロジェクトのスタートから終了までの研究サイクルの各フェーズについて、実際にどのような業務を行うべきかを検討し、課題の抽出をはかる。研究データ以外の生産物については、本テストの対象外であり、本テスト内ではDOIは付与しないが、業務フローとしては不可分のものであると予想されるため、検討の対象として含める。具体的に想定している研究データとしては、画像、動画、数値データを想定し、その他に公表論文、講演資料、各種報告書を想定している。</p>	<p>1) RAはJaLCをもちいる 2) 登録手順 ＜事前準備＞ ① 登録ポリシーの策定(登録対象、登録情報等) ② 運用フローの策定(プロトタイプ) ③ テスト計画の策定(登録レコード選択、作業者・確認者の決定) ④ メタデータスキーマの調整(機関リポジトリ側) ⑤ 登録環境の準備 ＜登録作業＞ ① JaLC登録テスト ② 登録データの検証 ③ 作業に要した業務量(作業時間等)の検証 ④ 運用フローの見直し 3) 使用するJaLCの登録インターフェース種類(XML/Web)の指定 XML、Webの双方のインターフェースをテストしたいと考えている。 4) テストするパターン(テストケース) 登録対象データにもとづきテストを行う。</p>	<p>(i)登録インターフェース 当初は、WebとXMLの双方をテストする計画であったがWebのみの個別登録にとどまっている。本登録時には、数千件のオーダーのデータアップロードが必要であるが、今回登録時にEXCELにて作成したメタデータについて、インポートに失敗した。データを視認した限りでは不備は見えないため、改行コード、区切記号等が原因ではないかと考えている。その原因について解決できていない。しかし実験では、メタデータ仕様等に検討を特化することとしたため、XML登録については、本登録時までテストを先送りすることとした。 (ii)テスト登録 コンテンツについて、今回は本学の機関リポジトリシステムのデータセット中の最大のコレクションである、萩庭コレクション(植物標本)にフォーカスを絞ることとし、複数件について登録テストを実施した。他の学術雑誌掲載論文の附属データについて、昨年度より論文本部の登録要請時に、附属データについても登録可否をあわせて問い合わせをはじめたが、現時点ではほとんど登録が進んでいない。</p>	<p>今回登録テストをおこなった、植物標本データについて、採集者/標本作成者/デジタルファイル作成者等の複数の人間が作成にかかわっているが、それをメタデータへどのように記述するかで苦慮した。実験データごとに、メタデータの記述方法に大きなバリエーションはあることが予想されるが、最低限の入力ガイドラインは必要であろう。</p>	<p>千葉大学内の機関リポジトリのメタデータスキーマの拡張(現時点ではJunii2準拠で未拡張)、および、NII IRDBおよびJaLCへのマッピングテーブルの作成とハーベスティングの実験</p>	<p>萩庭コレクションについては約5万件の大規模の植物標本のコレクションであり、貴重なサンプルも多く、識別子の付与をめざしたい。カレントの研究データの登録については、さらなる運用体制の検討が必要であり、データ登録にとどまらず、大学としてオープンサイエンスの推進をめざした、ポリシーメーカーからの検討が必要と考えている。</p>	<p>JaLCの事業からは若干外れるが、欧米ではFOSTER等、研究データマネジメントのためのトレーニングキットが整備されつつある。本実験では、比較的大規模で、研究データのメタデータの取り扱いのノウハウをもった機関/実験グループが参加しているが、次のステップとして、登録者の拡大を目指すのであれば、運用組織の整備の他、人材育成も検討する必要があるのではないかと。 (参考) Foster: https://www.fosteropen science.eu/</p>

8	物質・材料 研究機構	研究データの登録は、研究成果の公表を補完し、機構として長期的に保存、適切な利用条件のもとに公開することを目的に、NIMS デジタルライブラリー（セルフアーカイブシステム）を利用するものとし、 ・登録者は公開する成果物（論文、データ、補足資料）のメタ情報について、登録時に、可能な範囲で責任を持ってメタ情報を整理し、登録する ・公開される登録情報の有効期間は、予測や設定不可能な場合も含み、原則として利用者の判断、責任で利用する ・登録される成果物は、登録者および機構の研究成果として図書館機能を提供する科学情報室にて管理するものとするという基本案の実現性、有効性を検証する。なお研究データの登録とは、機構でなされた（あるいは機構と共同研究協定の元に行われた）研究成果を対象として、 ①科学情報室が管理する機構のサーバで運用する。 ②利用者すなわち機構で研究に携わる職員が、データを登録する。必要に応じて科学情報室が代理登録をするなどの支援を行う。 ③利用者は公開・非公開のアクセス条件の設定を行う。 ④セルフアーカイブシステムに登録された論文・データ・関連資料等は、NIMS デジタルライブラリー内で相互リンクを張り、情報の関係性を示すことを基本的な運用原則とし、その実効性を評価することを目標とする。 なお本テストの成果は、当事者へのヒアリングによって評価し、機能性やユーザーインターフェース改良への意見などは次期システム更新時に反映することを旨とする。	画像データ等、約 150 点（電子顕微鏡で撮影した画像など）	<p>(1) 実験対象とするデータおよびシステム NIMS では研究データのセルフアーカイブシステムとして、「NIMS eSciDoc」（以下「eSciDoc」）を運用している。eSciDoc は論文等の文書系アーカイブサービス PubMan と、画像データ系アーカイブサービス Imeji からなり、本実証実験は Imeji を登録実験の対象とするが、2つのサービス間の連携（相互リンク等）の検証のため、DOI 付与の検討は両サービスでの運用を想定し、登録実験は Imeji で行う。</p> <p>(2) アプローチ □ JaLC に DOI の取得リクエストを送信し、結果を Imeji に返す機能を持つ中継用アプリケーションを用意する。 □ セルフアーカイブでの登録者（研究データの作成者）を同定する手法として、同アプリケーションには、職員認証（職員コードとの対応付け）に加え、ORCID との対応付けを可能とする機能を用意する。</p> <p>(3) ワークフロー（表1） 登録者は機構内の LDAP アカウントで中継アプリケーションにログインし、アップロードした研究データに対して DOI 付与の対象とするデータを選択し、付与する。付与の結果は Imeji で確認することができ、DOI または URL として引用することを利用方法として想定する。</p> <p>(4) データフロー（表1） 中継アプリケーションにログインし、研究データを登録、うち選択するデータに DOI を振ると同時に、登録者名の同定、ならびに登録者と研究データの紐付けを行う。ログイン中の登録者の ORCID 番号を取得し、DOI 付与時に研究データのメタデータとともに、JaLC に送信する。</p> <p>(5) 中継用アプリケーションの開発と機能、および eSciDoc との連携本実証実験では、下記の e 以外の実装と検証を行った。 a. 登録者自身の機構内 LDAP アカウントを用いたログイン b. 登録者の LDAP アカウントと ORCID アカウントの紐付け（ORCID の WebAPI を使用） c. 研究データのアップロードとメタデータの追加・編集 d. アップロードした研究データに対する JaLC DOI の登録者自身による選択的付与（JaLC の WebAPI を使用） e. アップロードした研究データとその DOI を eSciDoc 上で公開する機能（eSciDoc の WebAPI を使用）</p>	中継用アプリケーションと JaLC、ならびに ORCID の WebAPI を用いて、中継用アプリケーションにアップロードした研究データに対する ORCID の紐付けと、DOI の取得、ならびに中継用アプリケーションへの保存を行うことができた。これは、管理者がメタデータの付与と公開作業を行う運用（大学での例など）と異なるワークフローであり、DOI の指し示すデータの同一性を損ねる可能性（DOI が指すデータは常に同じとは限らない）という課題を残している。このため、研究データへの DOI の本番登録においては、本課題であると認識している。	「テスト実施方針」で述べたとおり、機構のセルフアーカイブシステムへの登録は、登録するデータの選定、データの質、データの公開範囲も含め、登録者が各自の責任で行うことを前提としており、セルフアーカイブ公開後のメタデータの編集やデータの取り下げは、登録者が自ら行えるようになっていく。これは、管理者がメタデータの付与と公開作業を行う運用（大学での例など）と異なるワークフローであり、DOI の指し示すデータの同一性を損ねる可能性（DOI が指すデータは常に同じとは限らない）という課題を残している。このため、研究データへの DOI の本番登録においては、本課題であると認識している。	「結果の検証と考察」で述べたセルフアーカイブシステムの運用方針、ワークフローおよびデータフローについて、次の課題を残した。 (1) セルフアーカイブのルール ① セルフアーカイブする登録者は、ORCID 番号を取得していることを登録時の必須条件とするか。（研究者総覧 SAMURAI の次バージョンでは、対象職員が ORCID 番号を取得していることを前提とする予定で、ここでの業績リストとセルフアーカイブされたアイテムを同期する仕組みを想定） ② セルフアーカイブしたアイテムを取り下げる必要がある時、DOI 付与済みアイテムである場合のルール（ランディングページを表示するか） (2) DOI 付与のルール ① DOI が指すアイテムは常に同じものであるとするか否か ② DOI は公開設定アイテムだけに付与するか (3) デジタルライブラリーシステム eSciDoc の仕様変更 ① アイテムの登録、公開・非公開設定、CC ライセンス設定周りの機能改良 ② 登録済みアイテムの DOI 表示 ③ 更新用 WebAPI の実装	「今後の課題」で挙げた課題が解決した後に、研究データの本番登録を実施する予定である。
9	"理化学研究所 脳科学総合研究センター 神経情報基盤センター"	JaLC Web 登録画面よりデータ登録のテストを行い、問題点、課題等を確認。 当センターでは国内の大学研究所と共同研究として web データベースを開発しており、それをプラットフォーム(省略して PF)と呼ぶ。現在 13 の PF があり、それぞれがコンテンツを登録して公開している。 http://www.neuroinf.jp/modules/program/index.php?content_id=2 今回は、これら PF の主たるコンテンツ管理システムである XooNlps を用いて運営する PF に限定して、サンプルを選んで登録テストを実施する。この検証結果は当センターが管理する大半の研究データへの登録体制構築に利用が可能となると期待される。	DOI テスト登録計画とは異なり、本登録テストのスコープである研究データについてのみ登録を行った。XooNlps が定義するコンテンツカテゴリーの中で、研究データに利用されているカテゴリーである data、model、tool、stimulus、simulator について網羅するようにデータを選定することにした。Visiome PF から 8 件、ダイナミックブレイン PF から 1 件、合計 9 件の研究データを選択した。	<ul style="list-style-type: none"> RA としては JaLC を使用 ランディングページは XooNlps がアイテム登録ページを設定した。同ページには、カテゴリー毎に、一定のメタデータの記載が用意されており、DOI の記入項目と一定の対応をつけることが可能である。 JaLC Web を用いた手入力により登録 	登録したデータについて表 1 に示す。登録した DOI は表 2 の登録確認 URL によって、元の登録 URL に遷移することが確認できた。	メタデータの入力には基本的にあらかじめ XooNlps に登録されているメタデータをそのまま転記するポリシーだったが、たとえば追加情報の項目が 1000 文字を超える場合の登録方法も問題であった。現状のシステムでは 1000 文字を超える文章をペーストした際に 1000 文字に丸められるが、登録後にメタデータを確認しようとしても登録したメタデータをシステム上で確認、ダウンロードすることができなかった。追加情報の文字数制限はシステムの改修により 5000 字に変更されるということだが、そもそも登録すべき文字数が、その入力項目の制限文字数を超過している場合にどうすべきかを決めなければならない。 文字数制限の問題は他の項目にもいえることであるので JaLC としての指針があればそれをもとに考えたい。 権利の記述は自由記述と URI がどちらも必須になったことも問題であった。登録するコンテンツによっては、自由記述による権利記述のみがあり、権利 URI の記載がないものもあり、そのようなデータの場合に必須である権利 URI の項目についてどうすべきかという問題があったので、どちらかのみの記述でも許されることが望ましいといえる。 Web への手入力による登録の場合、サイト ID、出版者のように固定的な内容である場合や、言語設定のようにデフォルトで設定しておきたい項目がある場合などにも逐一再入力、再設定しなければならなかったり、連番に対する採番を入力にたよることが手間であった。 本登録をするにあたっての最大の問題点はメタデータ登録にあたって、現状では JaLC サイト登録後に査読フェーズがなく、そのまま DOI 登録されてしまうということであった。品質保証を考えると、登録以前に査読を行う必要がある。そのためコンテンツ毎のメタデータ作成者、そのメタデータ公開を管理する PF の運営責任者、さらに DOI 登録者である当センターの間で査読のためのやりとりが発生することとなる。将来大量コンテンツの登録に向けた障害となることが予想され、解決が必要である。	確実に効率のよい登録のためには前項の査読に関わる問題をクリアしたうえで、自動化、半自動化についても検討していく必要がある。	当方としては本登録もぜひ参加したいと思っている。

各参加機関の「DOIテスト登録報告書」において言及されている表および図を以下にとりまとめる。

NBDC/JST

メタデータ項目(主なもの)	JALIC	JALICメタデータの値
生命科学系データベースアーカイブ		
データベース名称		
データ名	title	図2 map02 - 異温帯連続観測情報
データURL	url	http://dbarchive.biosciencedbc.jp/jp/rgp-gmap06/data-2.html
データ内容の説明		
データファイル		
作成者氏名	creator (ja)	香島高章, 倉田ひかり, 長村吉典, 矢野高樹, 佐々木卓治*
作成者氏名	creator (en)	Yoshiaki Harushima, Nori Kurata, Yoshiaki Nagamura, Masahiro Yano, Takuji Sasaki
作成者所属	affiliation	社団法人農林水産先端技術研究所, 農業生物資源研究所, 社団法人農林水産先端技術研究所, 農業生物資源研究所, 農業生物資源研究所
連絡先		
データベースURL		
データベース分類		
生物種		
データベースの説明		
データベースの特長・有用性・活用方法		
予算的資金・プロジェクト	fund	
種別等・アクセス		
署名名		
雑誌名/掲載年月/号		
Pubmed ID		
利用許諾	rights	クリエイティブ・コモンズ 表示 1 日本
利用許諾URL	rights_url	http://dbarchive.biosciencedbc.jp/jp/rgp-gmap06/sc.html
データベース運用場所		
オリジナルサイト		
運用開始年月日	publication_date	1998/1/17
最終更新年月日		
DOI		10.14977/05.rgp_gmap06_detail
publisher		国立研究開発法人 科学技術振興機構 Japan Science and Technology Agency (JST)

表 1. メタデータ項目の比較と値の一例

国立極地研究所

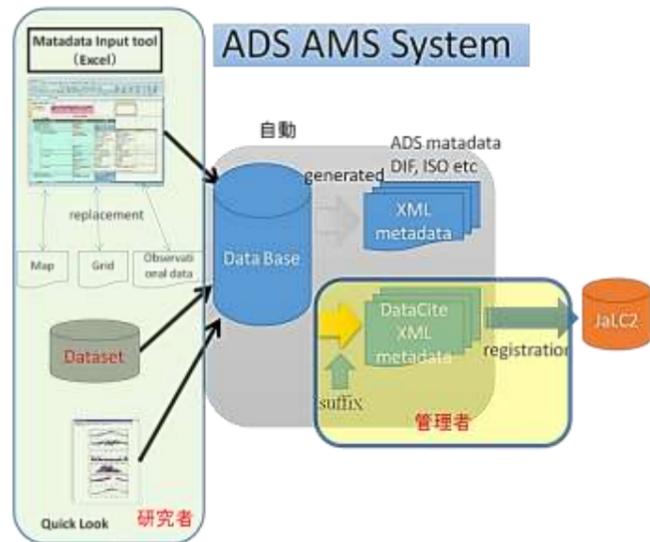


図 1 : ADS-AMS のデータ管理システムの概要



図 2 : ADS-AMS の DOI 登録表示画面。現在の登録データに対して、DOI の付与状況が示される。管理者は DOI 付与メタデータに対してこの画面より操作を行う。



図 3 : ADS-AMS の DOI 付与画面。管理者はこの画面より DOI の付与を行う。データ登録者への e-mail の連絡や、twitter への配信機能も有する。



図 4 : 登録データのランディングページ。

▶ Dataset citation
Hironori Yabuki, Hotaek Park, Rikie Suzuki, Tetuo Ohata, Haruko Kawamoto, 2011, Baseline Meteorological Data in Siberia (BMDS) Version 5.0, 5.00, Arctic Data archive System (ADS), Japan, doi: 10.5072/02.001.2013110702

図 5 : DOI 付与がなされたデータの引用方法についてのランディングページでの表示

▶ Dataset citation
Nozomu Takeuchi, Sota Tanaka, Masaya Miyairi, 2013, o of Suntar Khayata Glacier No.31, Eastern Siberia, July 2012, GRENE project, 1.00, Arctic Data archive System (ADS), Japan, https://ads.nipr.ac.jp/dataset/A20131206-023

図 6 : DOI 付与されていないデータの引用方法についてのランディングページでの表示

データライフサイクル	実務者/責任者
データ作成	研究者
データ保存	データセンター
メタデータ作成	研究者(必要があればデータセンター)
メタデータ修正	研究者(必要があればデータセンター)
データ修正	研究者(必要があればデータセンター)
データ公開化	データセンター管理者
データ廃棄(非公開化)	データセンター管理者(研究者が申請あり)
DOI登録	データセンター管理者

表 1 : 国立極地研究所 北極域データアーカイブ (データセンター) におけるデータライフサイクルの担当者

国立情報学研究所
地球環境情報統融合プログラム (DIAS-P)

AGURAM 全球都市域マップ 地上情報データベース	空間データ	1	東京大学 地球観測データ統融合連携機構	単独のデータセットとしての テストケース
AWCI 流域データセット (パン グラデシュ・メグナ流域)	時系列データ	7	DIAS Office	地域的な階層構造を持つデー タセットのテストケース

表 1 : テスト登録対象データセット

情報通信研究機構
他国内 WDC4 機関

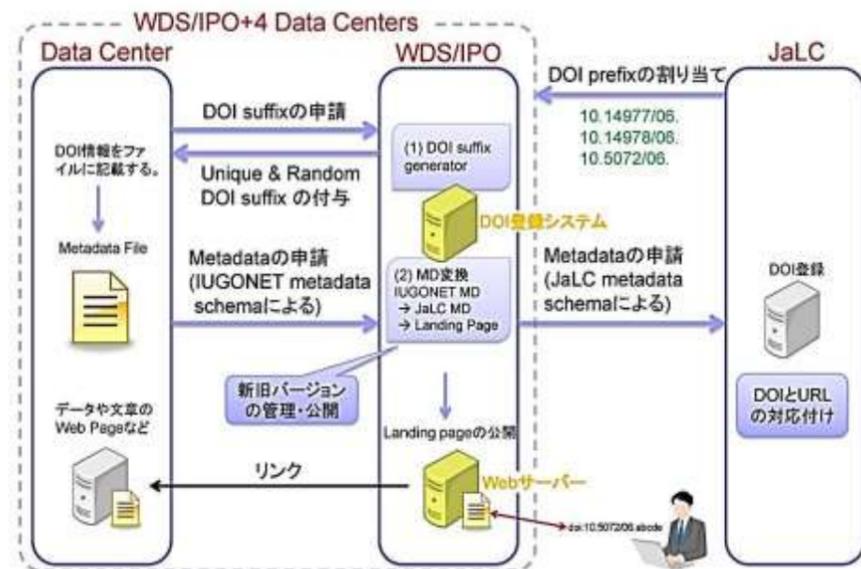


図 1: 作成した DOI 登録システム、およびシステム内のデータの流れ。
情報通信研究機構が中央に示した WDS/IPO を運営しており、DOI 登録システムサーバー
(中央上)および Landing Page のウェブサーバー(中央下)を管理している。

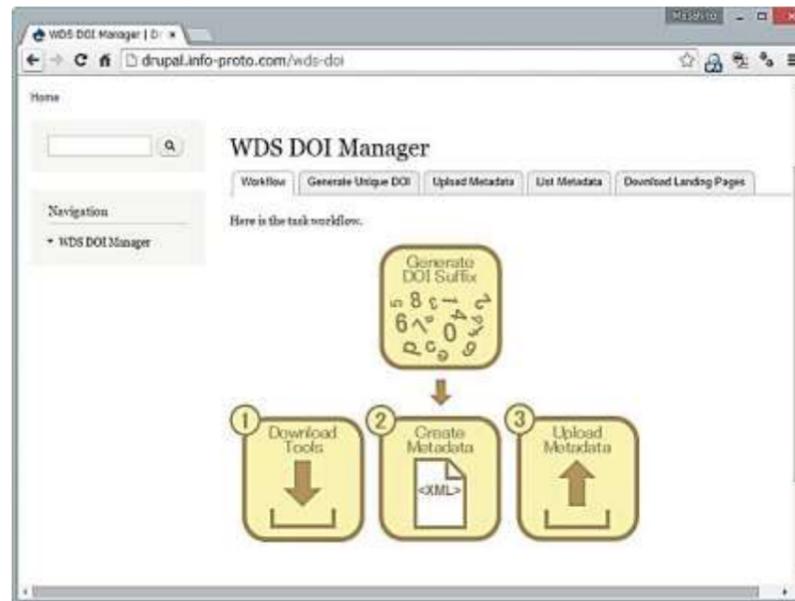
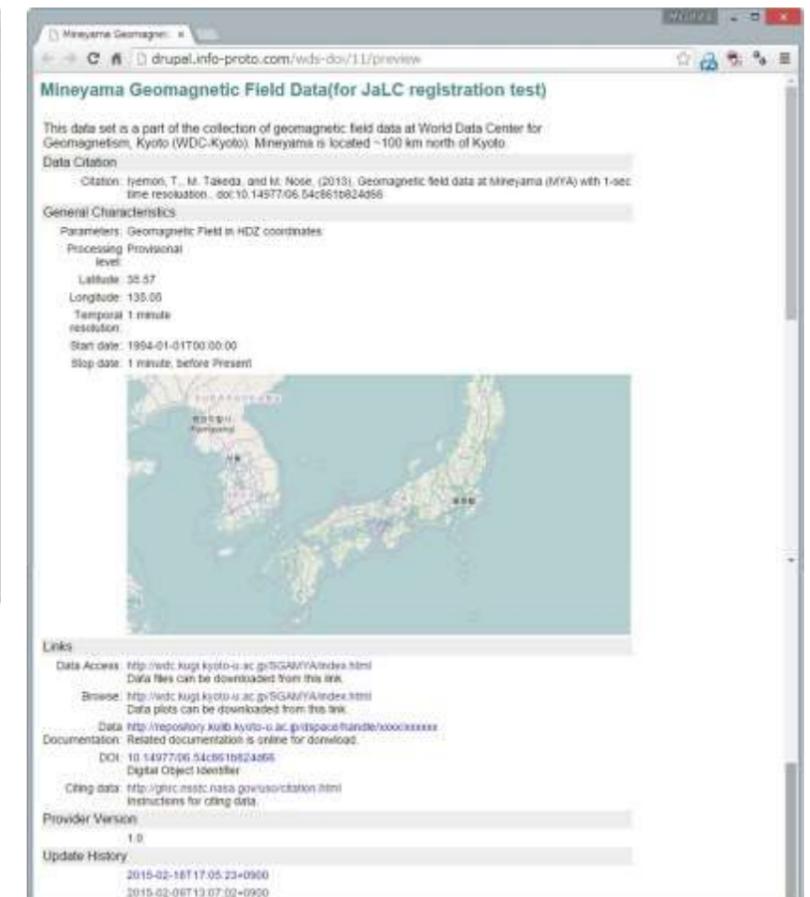


図 2: (左) DOI 登録システムサーバーのトップ画面。各データセンター
はこのインターフェースからすべての DOI 登録作業を行う。(右)テスト登
録で作成された Landing Page の例。データ引用の際の表示方法や観測
点の情報などの他に、下の方にデータセンターのデータやドキュメント
へのリンクが掲載されている



千葉大学附属図書館

データタイプ	レコード数	パーセンテージ
Journal Article (学術雑誌論文)	25,175	28.10%
Thesis or Dissertation (学位論文)	894	1.00%
Departmental Bulletin Paper (紀要論文)	7,497	8.40%
Conference Paper (会議発表論文)	30	0.00%
Presentation (会議発表用資料)	120	0.10%
Book (図書)	35	0.00%
Technical Report (テクニカルレポート)	35	0.00%
Research Paper (研究報告書)	803	0.90%
Article (一般雑誌記事)	0	0.00%
Preprint (プレプリント)	12	0.00%
Learning Material (教材)	210	0.20%
Data or Dataset (データ・データベース)	52,223	58.40%
Software (ソフトウェア)	0	0.00%
Others (その他)	2,438	2.70%
合計	89,472	

表 1 : 千葉大学機関リポジトリ登録コンテンツ

物質・材料研究機構

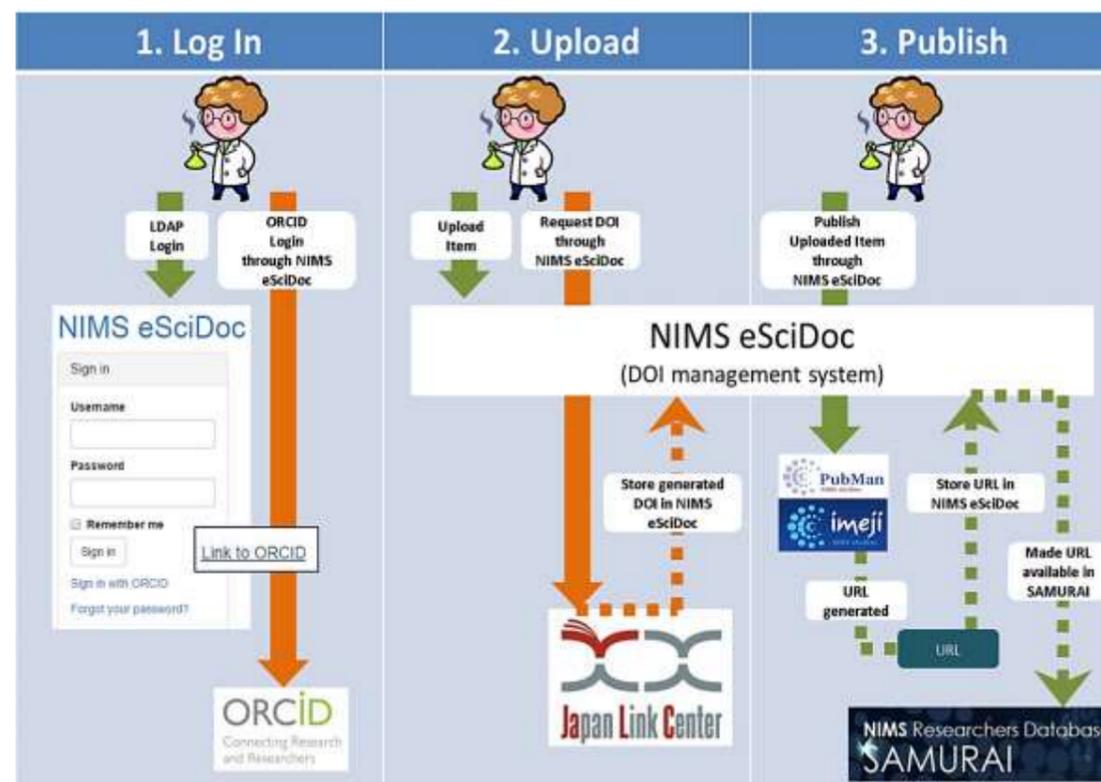


表 1 NIMS eSciDoc における DOI 付与のワークフローとデータフロー

理化学研究所脳科学総合研究センター神経情報基盤センター

URL	DOI
1 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=7111	10.14977/09.visiome7111
2 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=6894	10.14977/09.visiome6894
3 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=6448	10.14977/09.visiome6448
4 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=6411	10.14977/09.visiome6411
5 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=4619	10.14977/09.visiome4619
6 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=6125	10.14977/09.visiome6125
7 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=1647	10.14977/09.visiome1647
8 https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=3467	10.14977/09.visiome3467
9 https://dynamicbrain.neuroinf.jp/modules/xoonips/detail.php?item_id=10143	10.14977/09.db10143

表 1. テスト登録したデータ

1 http://doi.org/10.14977/09.visiome7111
2 http://doi.org/10.14977/09.visiome6894
3 http://doi.org/10.14977/09.visiome6448
4 http://doi.org/10.14977/09.visiome6411
5 http://doi.org/10.14977/09.visiome4619
6 http://doi.org/10.14977/09.visiome6125
7 http://doi.org/10.14977/09.visiome1647
8 http://doi.org/10.14977/09.visiome3467
9 http://doi.org/10.14977/09.db10143

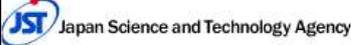
表 2. テスト登録したデータの登録確認 URL

各参加機関から中間報告会にて提出された「DOIテスト登録中間報告書」は以下のとおりである。



科学技術振興機構
バイオサイエンスデータベースセンター
の取り組み

2015.07.03
JaLC研究データDOI登録実験 中間報告会
櫛田達矢 (NBDC/JST)



1

目次

1. テスト実施方針
2. 生命科学系データベースアーカイブとは
3. クリエイティブ・コモンズ (CC)
4. テストの方法
 - 登録対象データと登録対象件数
 - テスト登録方法
5. 結果
6. 課題と検討事項
7. 本番登録の予定

2



© 2015 NBDC/JST Licensed Under CC 表示 2.1 日本

1. テスト実施方針

- NBDCがサービスを提供している「生命科学系データベースアーカイブ」
 (<https://dbarchive.biosciencedbc.jp/>)
 に収録されているデータセットに対してDOIを付与するテストを行う。JaLCのDOI登録インターフェースを使用し、生命科学系データベースアーカイブのメタデータがDOI登録にもそのまま活用できるか、不足の項目があるかなどを確認し、本格運用開始に備えることを目的とする。

2. 生命科学系データベースアーカイブ

データセットを**長期安定に維持保管**し、**データ説明（メタデータ）を統一**して検索を容易にすると共に、**利用許諾条件などの明示**を行うことで、**多くの人が容易にデータへアクセスしダウンロード**を行えるようにするサービス



アーカイブデータベース一覧 (ヘルプ)

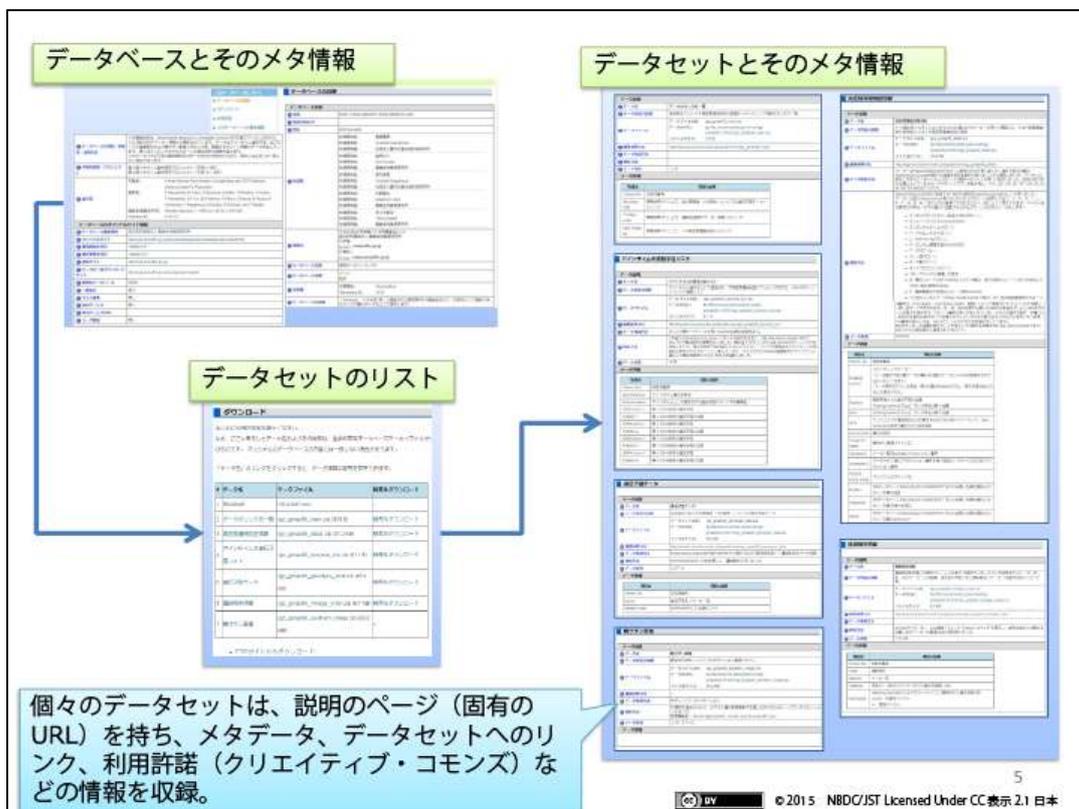
全 102 件 (1 から 10 件)

データベース	データベース提供場所	代表種	データベースID	生物種	要約 (キーワードも太字表記)	利用許諾
NBDK NihkgjRDF	ダウンロード オンラインサイト	科学技術振興機構	-	在来種	日本化学物産協会 (JCA) のデータベース	CC 表示 継承
RGP caps	ダウンロード (検索結果 オンラインサイト)	農業生物資源研究所	在来種	RNA-マーカー	イネの多量発現遺伝子に特異的なDNA-マーカー (EST-マーカー) よびNCAPs-マーカーのデータベース	CC 表示 継承 改訂
RGP estmap2001	-	-	-	-	-	-

102件のデータデータベース、約500件のデータセットを収録 (2015年6月現在)。

© 2015 NBDC/JST Licensed Under CC 表示 2.1 日本

57



3. クリエイティブ・コモンズ (CC)

再利用と共有を促進する仕組み

CCライセンスの種類
 作品の利用（再配布やリミックス作品の公開、実演等）のための条件は4種類あります。

 表示 作品のクレジットを表示すること	 非営利 営利目的での利用をしないこと
 改変禁止 元の作品を改変しないこと	 継承 元の作品と同じ組み合わせのCCライセンスで公開すること

↓

4種類のマークを組み合わせることで6種類の利用条件を選択することができる



CC-BY ジャパンクリエイティブ・コモンズ・ジャパン (<http://creativecommons.jp/licenses/>)

生命科学系データベースアーカイブで用いられる標準的なライセンス

6
© 2015 統合データベース講習会 Licensed Under CC表示 2.1 日本

4. テストの方法

- 登録対象データと登録対象件数
 - 生命科学系データベースアーカイブに収録されているデータセット全件（約500データセット（102データベース））
 - 登録対象件数（登録実験済の数）：74データセット（17データベース）
- テスト登録方法
 - 登録のインターフェースは、Web入力、XML一括入力の両方を使用した。
 - RAは、JaLCを選択。

実施体制

NBDC（バイオサイエンスデータベースセンター）
「研究データへの DOI 登録実験プロジェクト」チーム

責任者：櫛田研究員（NBDC研究員、ライフサイエンスDB統合のための基盤整備）

- JaLCとの連絡
- DOI登録の方式の検討
- データ登録、更新、削除
- 成果報告および報告書作成

生命科学系データベースアーカイブチーム（3名）

- メタデータ作成

5. 結果 (1/2)

- 生命科学系データベースアーカイブのメタデータが、新規作成、修正することなく、ほぼそのままの形でJaLCの研究データ登録に流用できることを確認した(表)。

9

 © 2015 NBDG/JST Licensed Under CC 表示 2.1 日本

表. メタデータ項目の比較と値の一例

メタデータ項目 (主なもの)		JaLCメタデータの値
生命科学系データベースアーカイブ	JaLCのメタデータ	
データ名	title	RGP gmap98 - 遺伝子型データ
データURL	url	http://dbarchive.biosciencedbc.jp/jp/rgp-gmap98/data-4.html
作成者氏名	creator (ja)	春島嘉章、倉田のり、長村吉晃、矢野昌裕、佐々木卓治*
作成者英名	creator (en)	Yoshiaki Harushima, Nori Kurata, Yoshiaki Nagamura, Masahiro Yano, Takuji Sasaki
作成者所属	affiliation	社団法人農林水産先端技術研究所、農業生物資源研究所、 社団法人農林水産先端技術研究所、農業生物資源研究所、 農業生物資源研究所
利用許諾	rights_url	http://dbarchive.biosciencedbc.jp/jp/rgp-gmap98/lic.html
運用開始年月日	publication_date	1998/1/17
	DOI	10.14977/01.rgp_gmap98_genotype_data
	publisher	国立研究開発法人 科学技術振興機構 Japan Science and Technology Agency (JST)

10

 © 2015 NBDG/JST Licensed Under CC 表示 2.1 日本

5. 結果 (2/2)

- ランディングページは、生命科学系データベースアーカイブのデータセットの「説明ページ」が適当であることを確認した。

11

© 2015 NBDG/JST Licensed Under CC表示 2.1 日本

ランディングページの例 (ダウンロード)

RGP gmap98

データベース名

遺伝子型データ

データセット名

このデータベースについて

- データベースの説明
- ダウンロード
- 利用詳細
- このデータベースの更新履歴

データ説明	
データ名	遺伝子型データ
データ内容の説明	各染色体の全ての分離体 (186個体) についての遺伝子型データ
データファイル	データファイル名: rgp_gmap98_genotype_data.zip データのURL: ftp://ftp.biosciencedbc.jp/archive/rgp-gmap98/LATEST/rgp_gmap98_genotype_data.zip ファイルサイズ: 48.4 KB
閲覧検索URL	http://togodb.biosciencedbc.jp/togodb/view/rgp_gmap98_genotype_data
データ取得方法	NipponbareとKasalathの掛け合わせから得た186のF2雑種体を用いて連鎖解析を行った結果。
解析方法	MAPMAKER/EXP 3.0を使用して、連鎖解析を行いました。
データ件数	2,277 件
データ詳細	
項目名	項目の説明
Chrom. No.	染色体番号
Locus	遺伝子座名 (マーカー名)
Genetic code	MAPMAKERによる遺伝コード

DOI

[10.14977/01.rgp_gmap98_genotype_data](https://doi.org/10.14977/01.rgp_gmap98_genotype_data)

ランディングページのURL
(データURL)

<http://dbarchive.biosciencedbc.jp/jp/rgp-gmap98/data-4.html>

12

© 2015 NBDG/JST Licensed Under CC表示 2.1 日本

ランディングページの例 (簡易検索)

簡易検索URL	http://togodb.biosciencedbc.jp/togodb/view/rgp_gmap98_genotype_data
データ取得方法	NipponbareとKasalathの掛け合わせから得た189のF2植物体を用いて連鎖解析を行った結果。
解析方法	MAPMAKER/EXP 3.0を使用して、連鎖解析を行いました。

English Japanese

RGP gmap98 - 遺伝子型データ NBDC - アーカイブトップ - ヘルプ

データベースの説明 | データ項目の説明 | 利用許諾 | ヘルプ

データのリンク先一覧 | 高密度連鎖地図情報 | アイソゲムの遺伝子座リスト | 遺伝子型データ | 連鎖順序情報

< 利用者の方へ >

- ダウンロード (Download:の右) する前に利用許諾を注意深くお読み下さい。ダウンロードを実行することによって、本利用許諾を承諾したものと見なします。

Find All Search Clear Advanced search

Download: Selected as CSV All

Entry	Chrom. No.	Locus	Genetic code
Show	1	R637	HHBHAAHAAADHAAHHHHHHBHBAHHHDBAHHHHHAHDBH
Show	1	C181	HHBHAAHAAABHAAHHHHHHBHBAHHHDBAHHHHHAHDBH
Show	1	C632	H.BHAH...ABHAAHH.HHSHHBAH. HHBAHHHHHHHAHSHHBAHAAHHDBHAAHHDBHAAHHHHH
Show	1	S1442	H.B.AH...ABHAAHH.HHSHHBAH.HHBAHH. HABSHHBAHAAHHHHHHBAHAAHHDBHAAHHHHH
Show	1	C970	HHBHAAHAAABHAAHHHHHHBHBAHHHDBAHHHHHAHDBH

13

©2015 NBDC/JST Licensed Under CC表示 2.1 日本

ランディングページの例 (利用許諾)

RGP gmap98

このデータベースについて

- データベースの説明
- ダウンロード
- 利用許諾**
- このデータベースの更新履歴

遺伝子型データ

利用許諾

データ説明 本データベースの利用許諾

データ名 本データベースは、以下で定める利用許諾に基づき提供されています。本利用許諾は、本データベースの利用における前提条件、及び利用者が遵守する条件を定めています。

データID 本データベースの利用許諾は、クリエイティブ・コモンズ 表示-継承3.0 日本での改訂版に基づきます。本データベースのクレジットは、「RGP: A HIGH-DENSITY RICE GENETIC MAP 8 農畜生物資源研究機関 (Revised) under CC表示継承3.0 日本版」です。利用にあたり必ず表示してください。

簡易検索 クリエイティブ・コモンズ 表示-継承3.0 日本版に基づきます。具体的な検索手順はこちらをご覧ください。

データ市 本データベースにおいて、以下の条件が適用される場合があります。

解析方法

データ作

データ詳

本利用許諾に基づいて利用する際に従うべき条件

- 本データベースの全部または一部、あるいは二次的変換物の提供を目的として、本データベースの作成者のクレジットを侵害してはなりません。
- 本データベースの全部または一部を複製して複製された二次的変換物は、この利用許諾の下で提供されなければなりません。
- 本利用許諾で許諾されていない事項については、以下のデータベースの作成者に連絡をとり、利用許諾を定める必要があります。

データベースの作成者連絡先

〒300-8602 茨城県つくば市緑園4-1-0
農畜生物資源研究機関
センター
E-mail: rmap@nrc.affrc.go.jp
〒565-0871 大阪府吹田市
E-mail: kst@nrc.affrc.go.jp

14

©2015 NBDC/JST Licensed Under CC表示 2.1 日本

6. 課題と検討事項

- Suffixのデザインをどうするか。人間が可読なもの（データセットの名称を含むSuffix）にするか否か（シリアル番号）について検討中。

[10.14977/01.rgp_gmap98_genotype_data](https://doi.org/10.14977/01.rgp_gmap98_genotype_data)

Prefix

Suffix

- JaLCのメタデータ項目の作成者および寄与者のそれぞれにどの情報を入力すべきか結論が出ていない。
- データ更新時の新DOIの付与や新旧データの保管が必要になった場合へ対応が決まっていない。

15

 © 2015 NBDC/JST Licensed Under CC表示 2.1 日本

7. 本番登録の予定

- 前述の課題を解決した時点で、登録を開始したい。

16

 © 2015 NBDC/JST Licensed Under CC表示 2.1 日本

国立極地研究所 中間報告

矢吹裕伯^{1,2}

- 1:国立極地研究所 北極観測センター
- 2:海洋研究開発機構

2013年7月3日
研究データへのDOI登録実験プロジェクト中間報告書

近年の動向

- オープンサイエンスの概念
 - 2013年6月:英国におけるG8科学大臣会合
 - 研究論文だけでなく研究データのオープン化
- その分野の知見を集約した国際的なデータベースの構築
 - 知見の集約
 - 情報ソースのハブ化
 - 教育・検索コストの低価格化
- 研究機関ごとに、独自のデータベースシステムの構築
 - 低価格で高性能なIT技術によって実現
 - 研究機関で独自に集約したデータの集約化
 - 資金を投入した研究成果の公共性の向上



- 属人性の高さ
 - データの管理、収集について、個々の裁量に依存
- 再利用性の低下
 - フォーマットの変換、所有権の確認に、時間が掛かる
- 意識の低さ
 - データ出版・公開に対する評価基準がない

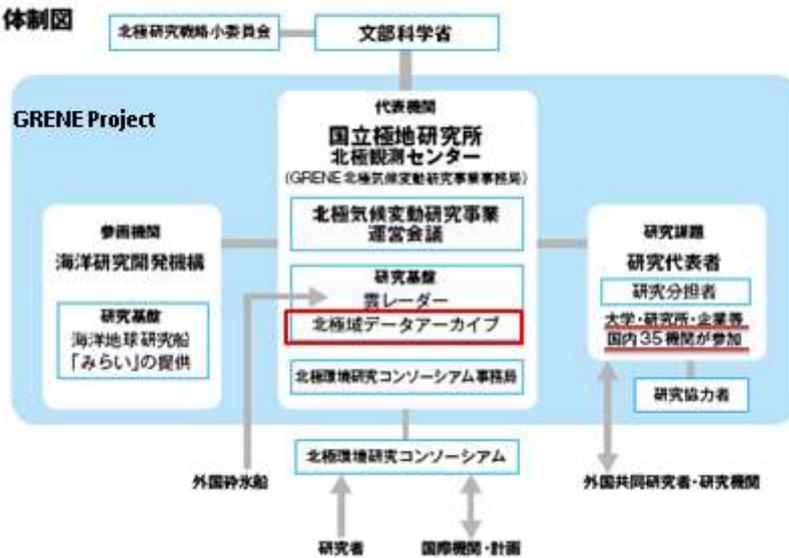


データの死蔵化

- 日の目を見ない観測データ・解析データ
- 他研究による被引用数の低下

データ出版・公開のインセンティブの確保が必要

体制図



ABS Arctic Data archive System GRENE北極事業(北極気候変動研究プロジェクト)

北極域における環境変動、気候システムの理解

データ共有 相互利用

データ基盤による データ相互流通

各分野間でこれらの様々データを共有することで効率よく解析を行う必要がある。
モデル研究者及び**観測研究者**によるデータの相互利用

5 国立極地研究所

ABS Arctic Data archive System GRENE北極事業(北極気候変動研究プロジェクト)

北極域における環境変動、気候システムの理解

データ公開基盤としてだけでなく
 北極研究の共同研究解析プラットフォーム
 としての位置づけ

データ基盤による データ相互流通

各分野間でこれらの様々データを共有することで効率よく解析を行う必要がある。
モデル研究者及び**観測研究者**によるデータの相互利用

6 国立極地研究所

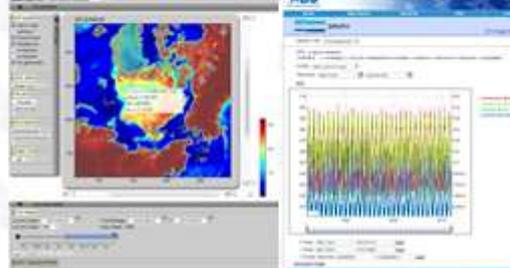
ADS Webアプリケーション群

極KIWA 研究データ検索・管理システム



<https://ads.nipr.ac.jp/>

VISION オンライン可視化アプリケーション



極域環境監視モニター
・可視化された衛星データを準リアルタイムでWebで提供する



ブラウザ上でグリッドデータや時系列データを可視化し、解析を可能にするオンラインで可視化するWebアプリケーション

7  国立極地研究所

ADSの利用方法



8  国立極地研究所

ADSメタデータ

メタデータの記載は全てエクセルシートを用いて行う。エクセルシートではマクロが組んであり、入力補助を行っている。



メタ情報規格

- 国際規格: **ISO19139**
- 関連プロジェクトの規格: DIAS, GCMD DIF, DataCite
=各々の必須項目は必ず書き、マッピングに対応
- ADS独自規格: ADS

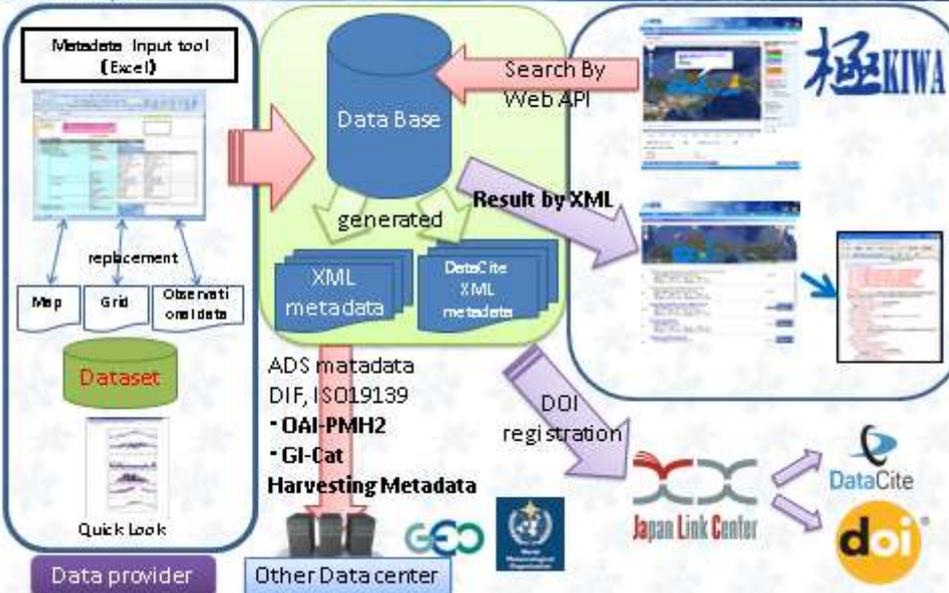
関連する規格	カテゴリ	集数
DIAS	地球科学全般	85
DIF (Global Knowledge Portal)	GCMDの定義	55
DataCite	科学データ間の関連	18

シート	シート名	項目	集数(必須項目)
A	Title & Personnel	データセットの基本情報	32 (22)
B	Management	サイトにおけるデータ表示、関連するプロジェクトなど	22 (5)
C	Constraint	権利に関する項目	18 (5)
D	Keywords	キーワード	14 (5)
E	Period	時間に関する項目	11 (2)
F	Location	空間に関する項目	17 (4)
F-1	Point	地点情報 ポイントデータ	20
F-2	Line	地点情報 ラインデータ	8
F-3	Grid	モデルデータ	17

この部分: データの種類により入れ替え(粒度の細かいデータセットを記述できる)

9 国立極地研究所

ADS Arctic Data archive System



The diagram illustrates the ADS workflow:

- Metadata Input tool (Excel)**: Used for entering metadata, with options for **Map**, **Grid**, and **Observation data**.
- replacment**: A process step connecting the input tool to the database.
- Data Base**: The central storage for metadata.
- generated**: The process of creating metadata files, including **XML metadata** and **DataCite XML metadata**.
- Search By Web API**: A method to query the database.
- Result by XML**: The output of a search query.
- Dataset**: The actual data being archived, with a **Quick Look** option.
- Data provider**: The source of the data.
- Other Data center**: External data sources.
- ADS metadata**: Includes **DIF, ISO19139**, **OAI-PMH2**, and **GI-Cat**.
- Harvesting Metadata**: The process of pulling metadata from external sources.
- DOI registration**: The process of assigning a Digital Object Identifier to the data.
- Japan Link Center** and **DataCite** are the organizations involved in the DOI registration process.

10 国立極地研究所

ADS Arctic Data archive System

Registration of Digital Object Identifier

ADSメタデータコントロール画面

ADSのデータランディングページ

11 国立極地研究所

研究データへのDOI付与のメリット

データ出版 (Data Publication) データ引用 (Data Citation)

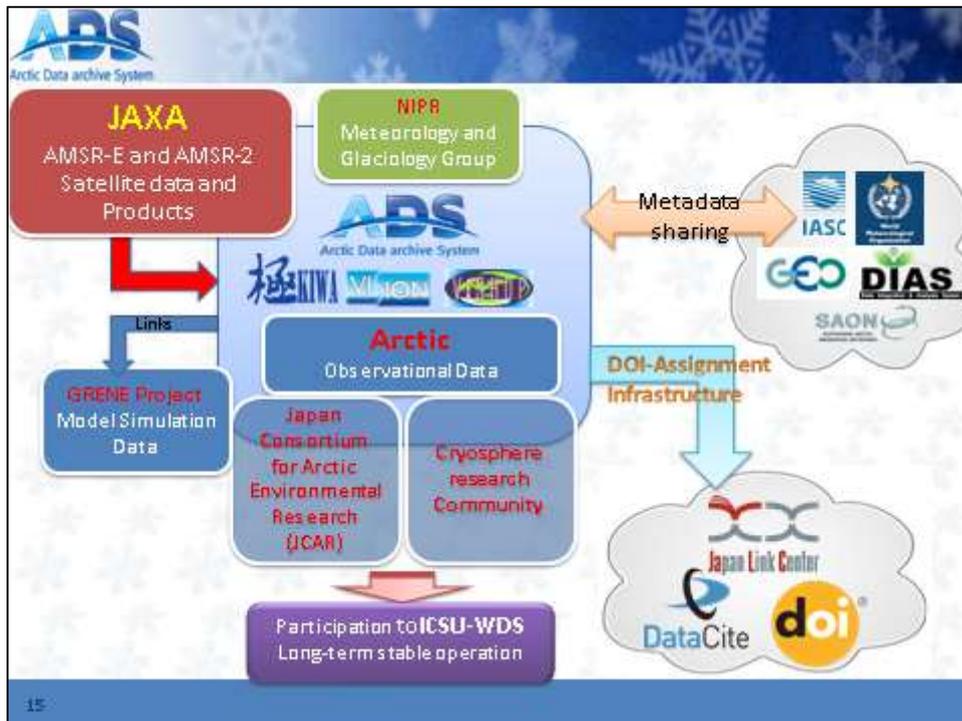
- データ利用者のメリット
 - 論文に引用されるデータにアクセスができる。
 - 論文の引用から新たなデータへのアクセスが可能
- データ提供者のメリット =>
 - データ公開が業績・評価の対象
 - データの利用数、被引用数を調べられる=>被引用論文数と同じ
- データセンターのメリット
 - データの整形等を通して2次利用が可能になり、データベースへの利用が可能になる。(統合的データベースの構築)
 - データ提供者へのインセンティブの確保

12 国立極地研究所

- DOIは 単なる恒久的なパーマントリンクを確保する“しくみ”。
- データセンター
 - **研究・観測データの品質が確保**されることが重要
 - データセンター自体の品質にかかわる
- DOI発行者の責務
 - 公開・出版されるデータの **品質の確保**
 - DOIを発行するということは研究・観測データの品質の確保が **最重要課題** => データセンターにとって重要なアクション

- 現在:
 - 観測データ報告(編集・査読)
 - JARE-data reports
 - NIPR arctic data reports
- 検討中(案)
 - Data Reportsを発展させ査読制度を設ける
 - 品質を確保する
 - これらの査読を通過したデータに関してDOIの付与を行う。
 - **極地研独自のデータ品質管理**
 - 極地研のデータ出版として data reportsの活用(CrossRef?)
 - データリポジトリとして **ADS**の活用(DataCite)
 - 学術論文に利用されたデータの登録(DOI付与)
 - 他データジャーナルに投稿されたデータ(DOI付与)





15

ADS
Arctic Data archive System

ありがとうございました。

URL: <https://ads.nipr.ac.jp>
 Contact : ads-info@nipr.ac.jp

16

データDOI実験中間報告

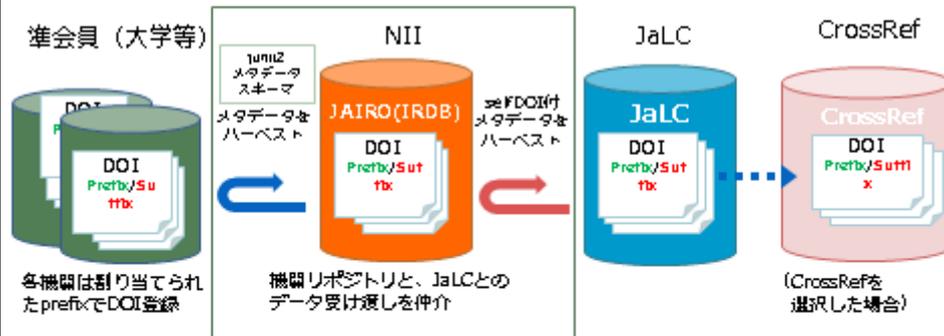
2015年7月3日 国立情報学研究所 山地一禎

テスト方針

- ▶ NIIが運用しているIRDBを介したDOIの付与機能ならびにプロセスを検証
- ▶ NIIがOSSとして提供しているリポジトリアプリケーションWEKOにおけるDOIの付与機能の検証

IRDBを介したDOIの付与機能

- ▶ 全てのIRのメタデータがNIIのIRDBに集約
- ▶ IRDBがJaLCと通信してDOIを登録



3

国立情報学研究所では、ソフトウェアWEKOをベースに機関リポジトリのシステム環境を構築・提供しています。

テスト登録対象データ

- ▶ 歴史的資料物写真データ
 - ▶ 使用範囲：公開データ
 - ▶ 件数：1000件程度
 - ▶ データ管理責任者：学術資源リポジトリ協議会
- ▶ 科学衛星観測データ
 - ▶ 使用範囲：公開データ
 - ▶ 件数：1000件程度
 - ▶ データ管理責任者：金沢大学

4

国立情報学研究所では、ソフトウェアWEKOをベースに機関リポジトリのシステム環境を構築・提供しています。

テストパターン

- ▶ リポジトリシステム内におけるPrefix登録
 - ▶ リポジトリ運用機関
- ▶ リポジトリシステム内におけるDOIの付与
 - ▶ リポジトリ運用機関
- ▶ リポジトリからIRDBへのハーベスティング
 - ▶ リポジトリ運用機関&NII
- ▶ IRDBにおけるエラーチェック
 - ▶ NII
- ▶ IRDBからJaLCへのハーベスティング
 - ▶ NII、JST
- ▶ DOIからリポジトリへのリダイレクト
 - ▶ リポジトリ運用機関&NII

5

国立情報学研究所では、ソフトウェアWEKOをベースに
機関リポジトリのシステム構築を構築・提供しています。

結果と課題

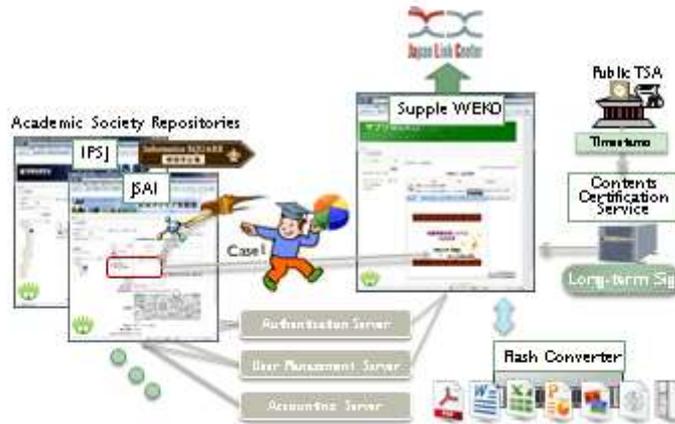
- ▶ 特に問題なく実験を完了
- ▶ 課題
 - ▶ Suffixの機械的な付与と人的な付与
 - ▶ メタデータ付与のミス防止
 - ▶ 1機関複数Prefixとする場合の準会員登録基準

6

国立情報学研究所では、ソフトウェアWEKOをベースに
機関リポジトリのシステム構築を構築・提供しています。

今後1

- ▶ 情報学広場サービスにおけるサプリメントデータに対するDOIの付与



7

国立情報学研究所では、ソフトウェアWEKOをベースに
機関リポジトリのシステム環境を構築・運用しています。

今後2

- ▶ 機関リポジトリに保存されるデータに対するDOIの付与
 - ▶ 千葉大三角さんから報告

8

国立情報学研究所では、ソフトウェアWEKOをベースに
機関リポジトリのシステム環境を構築・運用しています。

DIASにおける取組み

絹谷弘子

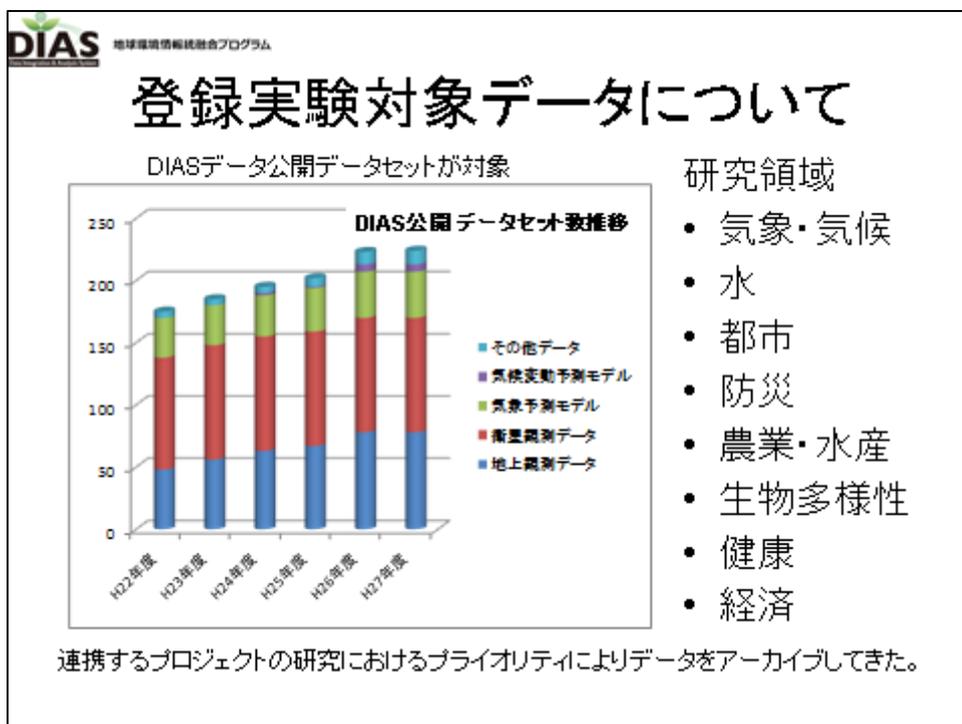
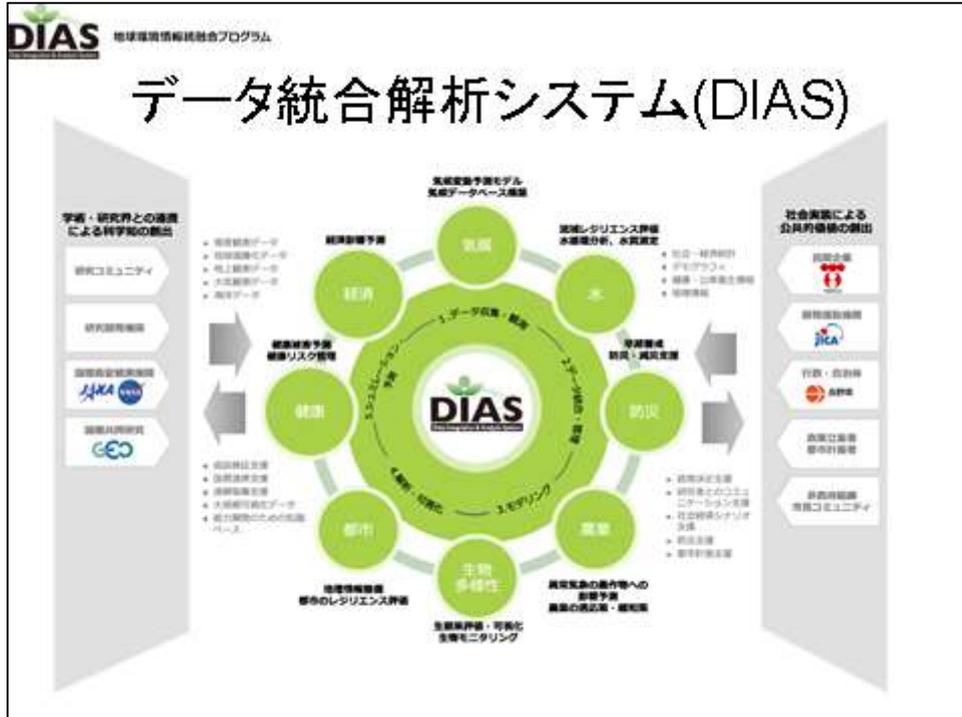
東京大学 地球観測データ統合連携研究機構

2015/07/03

ジャパンリンクセンター 研究データへのDOI登録
実験プロジェクト 中間報告会

概要

- ❖ DIASとJaLC 研究データDOI登録実験
- ❖ DIASにおけるDOI付与
- ❖ JaLCプロジェクトへの期待



登録実験

DIASは永続的な組織体制が未確定なためこのプロジェクト期間中はDOI実登録は実施しない

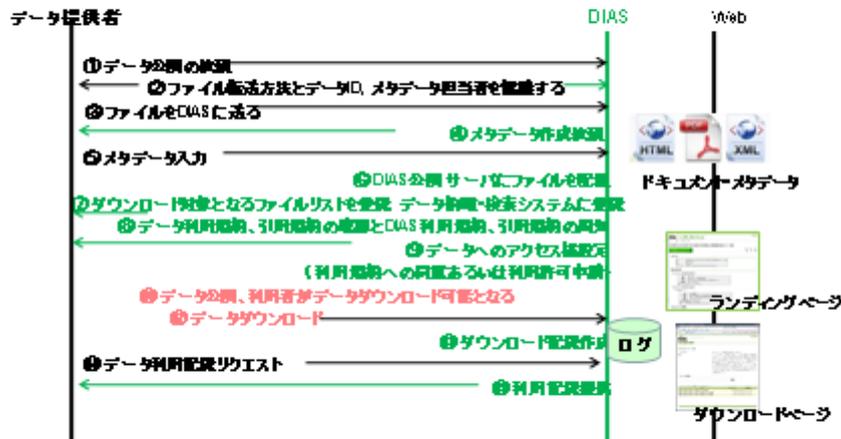
- テスト方針
 - 識別子を付与する粒度・スライス・コレクション・多言語等にかかわる問題を検討
 - 既存のメタデータからの変換やプロジェクト内での運用方式などを検証
- テスト方法
 - 識別子付与規則決定
 - 対象データセット選定
 - DIASメタデータとの対応項目確認
 - Webインタフェース、XMLファイルアップロード、API利用の3つの方法を実験し、いずれも問題なく登録できた
- 実験結果
 - JaLC, DataCiteへの登録方法は理解した
 - DIASメタデータとJaLCメタデータの項目対応も可能(必須項目について確認した)
- 今後の課題
 - 必要なランディングページ的设计と現状の改訂方針作成
 - DIASが公開するデータのDOI付与の手続きとワークフロー作成
 - DIASにメタデータがあり、DIAS以外が公開するデータへのDOI付与についての調査

DIASデータ公開プロセス(現状)

- ❖ 公開データは平均10件/年
- ❖ メタデータ作成重点項目
 - ❖ 問合せ先(個人、組織やプロジェクトの継続を重要視)
 - ❖ 収録期間と地理空間範囲、キーワード(シソーラスの利用推進)
 - ❖ データに関するオンライン情報のリンク
 - ❖ データ利用規約とデータ引用規約
 - ❖ これらは、データ提供者の編集に加えDIASメタデータ管理者が確認して編集する
- ❖ メタデータの品質チェック
 - ❖ 問合せ先やオンライン情報を可能な限りリンク切れしないように保守している
 - ❖ 表記ゆれや緯度経度、グリッドの表記等の統一化に努める
- ❖ データ訂正時等にデータ利用者情報をデータ提供者に通知できるように利用記録を管理している

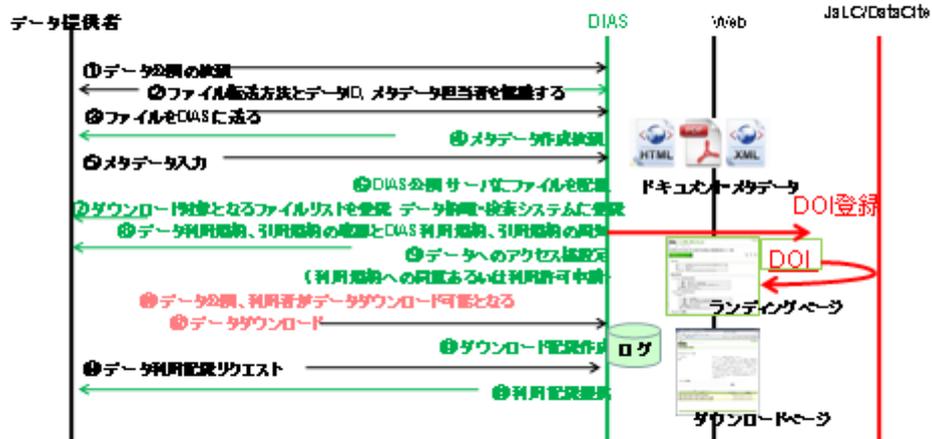
DIASデータ公開プロセス

現状



DIASデータ公開プロセス

DOI導入



DIASオープンサイエンス勉強会の活動

- ❖ JaLCへの参加を含め、オープンサイエンスのトレンドにDIASがどう応えるべきかを検討する会
- ❖ メタデータ、運用体制設計チームを中心としたDIASメンバーに国立環境研メンバーが加わり、すでに8回開催した
 - ❖ 対話と共創の場をきっかけとして国立環境研が参加
- ❖ 勉強会での議論をベースに、JaLCに対しても不具合の報告やガイドラインの方向性に対して、積極的に提言を行ってきた

ガイドラインへの要望(1)

- ❖ ガイドラインを継続的にアップデートする仕組みを構築して欲しい
 - ❖ データDOIと論文DOIの違いについては未だ明確ではない部分がある
 - ❖ 世界的な状況の変化や研究環境の変化に合わせてガイドラインの見直しが必要となる
 - ❖ 研究データの現場からの意見を集約する場も引き続き必要
- ❖ 多様な分野のデータ提供者が見てわかる「用語集」を追加してほしい
 - ❖ 例:ドメインメタデータ、JaLCメタデータ、データリポジトリ、ランディングページ、データのライフサイクル、ワークフロー

ガイドラインへの要望(2)

- ❖ JaLC会員が主語となる行為に義務・許可・推奨の違いがわかるようにしてほしい
 - ❖ 先行事例に基づいて推奨している部分が大部分ではあるが、shall(強い強制)、will(弱い強制)、may(許可)などの区別が文章でわかるとよい
 - ❖ ランディングページに記載する内容に義務・許可・推奨項目を明確に示してほしい
- ❖ JaLCがDOI登録機関に対して提供するサービスの範囲と責任を明らかにしてほしい

国内の研究データを国内でアーカイブし、世界をリードするためにJaLCのDOI活動に期待したい

産業技術総合研究所

渡邊 宏 (計量標準総合センター 物質計測標準部門)

山下雄一郎 (計量標準総合センター 物質計測標準部門)

小島 功 (情報・人間工学領域 研究戦略部)

概要

- 産総研の研究データベースの現状
- doiテストプロジェクトへ参加
- 実験内容・参加DB紹介
- まとめ

産業技術総合研究所の研究データベース

研究情報公開データベース

- http://www.aist.go.jp/aist_j/aist_repository/riodb/index.html
 - Web上で公開している研究成果のデータベース
 - 上記リストで約50+
 - 他にも研究ユニットが研究プロジェクト内で独自に作成・提供している
 - 多くはWebサービスとして検索や可視化などを提供



本プロジェクトへの参加

産総研におけるオープンデータ推進の中の1活動

- 産総研データバンク構想：データベースの高度化による利用促進
 - 研究DBのLOD(Linked Open Data)化、外部DBとの連携
 - ツールの高度化(分散SPARQL検索など)
 - CKANカタログの構築・インターフェイスの高度化
 - データベースの統合・最適化
 - doiの付与の検討
- doiテスト
 - 内部で呼びかけ：材料系の2つのDBからご協力いただく
 - 有機化合物のスペクトルデータベース (SDBS)担当：渡邊宏
 - <http://sdfs.db.aist.go.jp/>
 - 熱物性データベース：担当：山下雄一郎
 - <http://tpds.db.aist.go.jp/>
 - 全産総研的な動きにすべく、引き続き内部アピールや協力拡大を図っていく必要がある

試験検討項目

1. Doiのsuffix(10. 0.14977/05.XXXXXX) の決め方
2. Doiを付与する単位
3. Landing Page の自動作成など、対応するシステムの改修の必要性等

1. doiの決め方

- 産総研内で複数の組織が別々に登録すると予想
 - 組織で統一してdoiを与えるような体制には現在ない。
 - プロジェクトなりDB担当者が行う可能性が高い。
 - 名前の衝突等を避ける必要がある
 - 名前決めの都合はデータの保有者が一番分かっている？
- 今回は仮に以下のようにポリシーで決めてみた。
 - (prefix)/(もしあれば) サブ組織名) -(DB名)- (識別子)
 - 例 : <http://doi.org/10.14977/05.tdbs-23732>
 - 例2 : <http://doi.org/10.14977/05.gsj-aster-xxxx>
 - DB名でsuffix空間を分け、以降はDBの担当者が決める。
 - 自主管理ルールなので他の担当者が間違える可能性は残る
 - 将来サブ組織が必要ならその前に入れる。
 - 別ドメインのある組織があるから(gsj/nmij)

2.doiを付与する単位

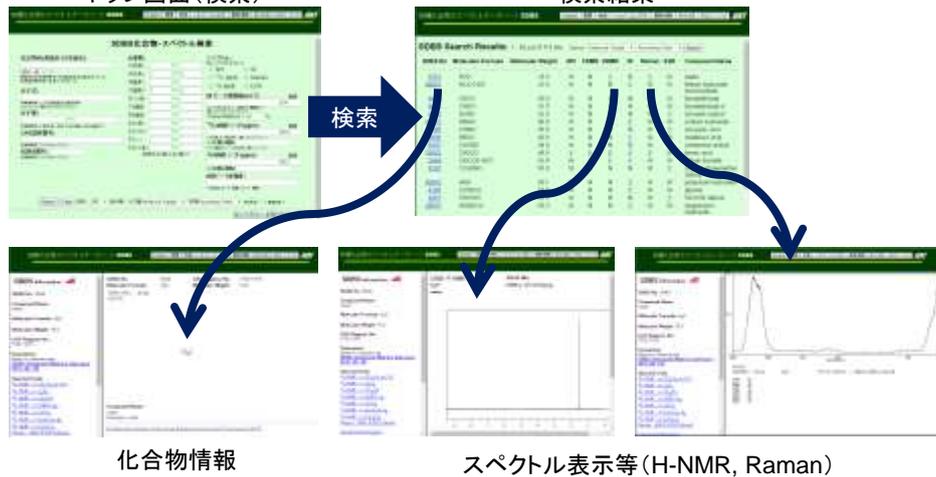
- できるだけ細かい方が引用には便利
 - 検索の結果でもあるのでシステムの挙動と対応が取れないといけない。
 - Landing Pageやメタデータの準備も増える。
- 同時に、DB全体でも参照できると良い

→いくつかの単位で試験を行って挙動を確認する

SDBS(有機化合物のスペクトルデータベース)

<http://sdfs.db.aist.go.jp/>

- 産総研の研究情報DBで最も大規模かつ著名なDB (1997~Web公開)



試験内容と結果

次の三種類のものを登録する。

1. SDBSのWebサイト： ○
2. sdb sno=4544 化合物「水」の化合物情報： ○
3. 水の1H NMRスペクトル HR2014-02904NS:

doiとそれによって解決されるURLの組み合わせテスト

1. SDBS番号（SDBSの中で使われる・使ってほしい番号）をdoi suffixに入れる・入れない、など
 - 05.sdb s-4544-HR20xxxxxx
 - 05.sdb s-HR20xxxx
2. URLがcgiになるのでいくつかのパターンをテスト
 - http://sdb s.db.aist.go.jp/sdb s/cgi-bin/direct_frame_disp.cgi?sdb sno=4544
 - http://sdb s.db.aist.go.jp/sdb s/cgi-bin/img_disp.cgi?disptype=disp3&imgdir=hpm&fname=HR201402904NS&sdb sno=4544

結果：

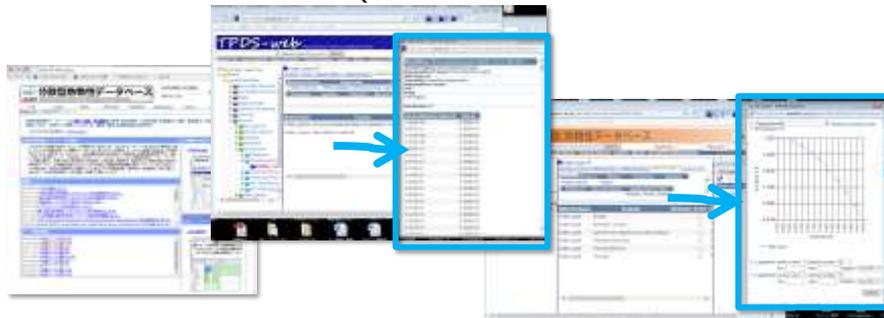
- doi:ネーミング等での問題はない。
- システム側：スペクトルデータに直接到達できるcgiを設ける必要がある。

3.Landing Pageの生成などシステムの対応 分散型熱物性データベース

<http://tpds.db.aist.go.jp/>

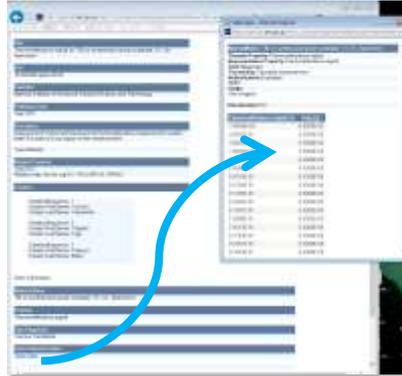
物質の熱物性値（熱伝導率、比熱容量、密度、表面張力、蒸気圧など）のデータベース

- 組織に跨った分散データベース
- ポップアップでデータ表示(検索結果&ナビゲーション結果)



試験内容と結果

- 検索結果の物質にLanding Pageをプログラムから生成する
 - JaLCスキーマに相当するデータ+αを表示
 - Landing Pageに埋め込まれたリンクから実データをアクセス & ポップアップ
 - 1物質・URLで試験→アクセス可
 - ポップアップ表示によるDB
 - doiからアクセスした場合の情報の断片化が懸念
 - 表示システムの再検討必要？



CKAN? as a Landing page provider

- CKANのページをLanding pageに使う
 - CKAN : オープンデータでのカタログサイトのデファクト
 - メタデータ項目=Dublin core
 - -> これをJaLC/DataCiteのメタデータ項目に拡張(可能)
 - まだ作業中(未テスト)
 - CKANそのものが内部にしか公開されていない。
- Dspace?
 - こちらの方が拡張性および機能が高い
 - doiがなくても、handleでそもそも十分？



まとめ

基本的には意図した通りのdoiの付与が可能と思われる。

doiを付与すること(登録など対JaLCの作業)に加え、
所内のさまざまな業務やシステム改修が想定され、その手間の軽減が重要。

- Landing page の作成と公開
- 対応システムの改修・システム内でのdoiの利用
- suffixの登録管理などの内部業務フロー
- 導入できるソフトウェアが提供されるとありがたい。
 - たぶんどの組織でもニーズはおおむね共通と思われる。
 - 独自の作り物は長期的な維持管理が問題なので、できれば既存のソフトのアドオンなりが望ましい。
- アウトソースでも良い
 - クラウド等のホスティングサービスで提供されてもよいが、landing pageがホスティング先に作られるので、ドメインだけでもaistである必要がある。
 - 上記の業務も含めいっそ全部特定のデータジャーナルにアウトソース？

引き続き検討をしつつ正式な登録へ進んでいきたい。

- 所内体制の整備

参加機関の取組み： WDS国際プログラムオフィス+国内4WDC

能勢正仁¹、小山幸伸²、家森俊彦¹、村山泰啓³、木下武也³、
渡辺 堯³、石井守⁴、山本和憲⁴、加藤久雄⁴、
門倉昭⁵、田中良昌⁵、篠原育⁶

¹ 京大大学院理学研究科地磁気世界資料センター

² 新領域融合研究センター

³ 情報通信研究機構・世界資料システム国際プログラムオフィス

⁴ 情報通信研究機構

⁵ 国立極地研究所

⁶ 宇宙航空研究開発機構

2015/07/03

DOI登録実験プロジェクト 中間報告会

1

WDC (World Data Center)

- 1957年7月-1958年12月：国際地球観測年(IGY)実施、66か国、約4000観測点
 - 気象・地磁気・極光・電離層・大気光・太陽活動・宇宙線・ロケット・人工衛星・緯度・経度・氷河・海洋・地震・重力・大気放射能
- 1957年-1958年：World Data Centerの創設
 - IGY期間中に汎地球的に取得された地球物理・太陽データを保存するため。
- 2007年：創設50周年、12ヶ国、51ヶ所のデータセンター。
 - 分野も土壌、気候、生物多様性、環境などを含むようになった。



日本国内は7ヶ所

World Data Centers

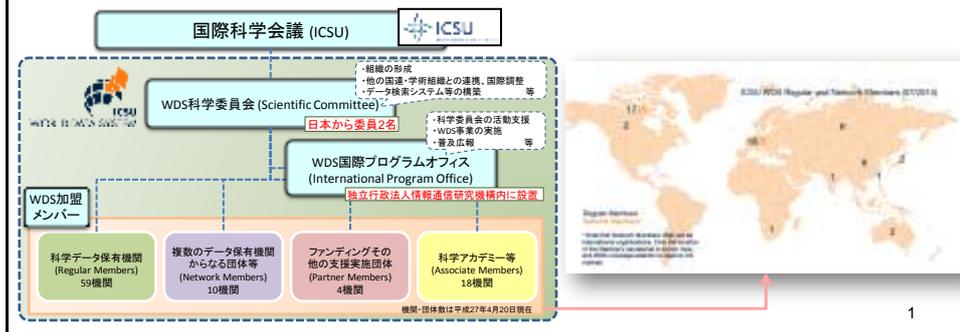
April, 2007

1

WDCからWDS (World Data System)へ

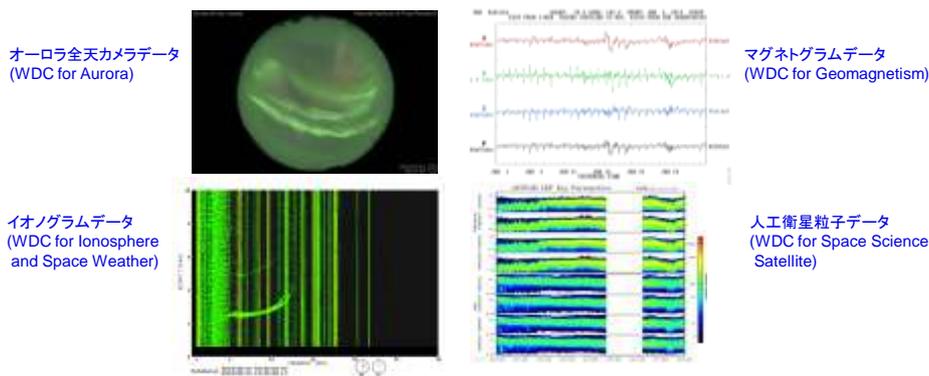


- 50年が経過した後、International Council for Science (ICSU、国際科学会議)傘下のデータ関連組織の見直しが行われた。
- 2009年より、World Data Systemとして新たに発足。(WDC+FAGS=WDS)
 - 自然科学分野だけでなく、今後は人文・社会科学系のデータセンターも加える。
 - 分野横断型データ利用に対処するため、メタデータやデータ利用システムの整備。
 - データ出版・データ引用への対応。最新のIT技術への対応。
- WDS/International Program Office (WDS/IPO)
 - 情報通信研究機構に設置が決定。2011年度より業務を開始。

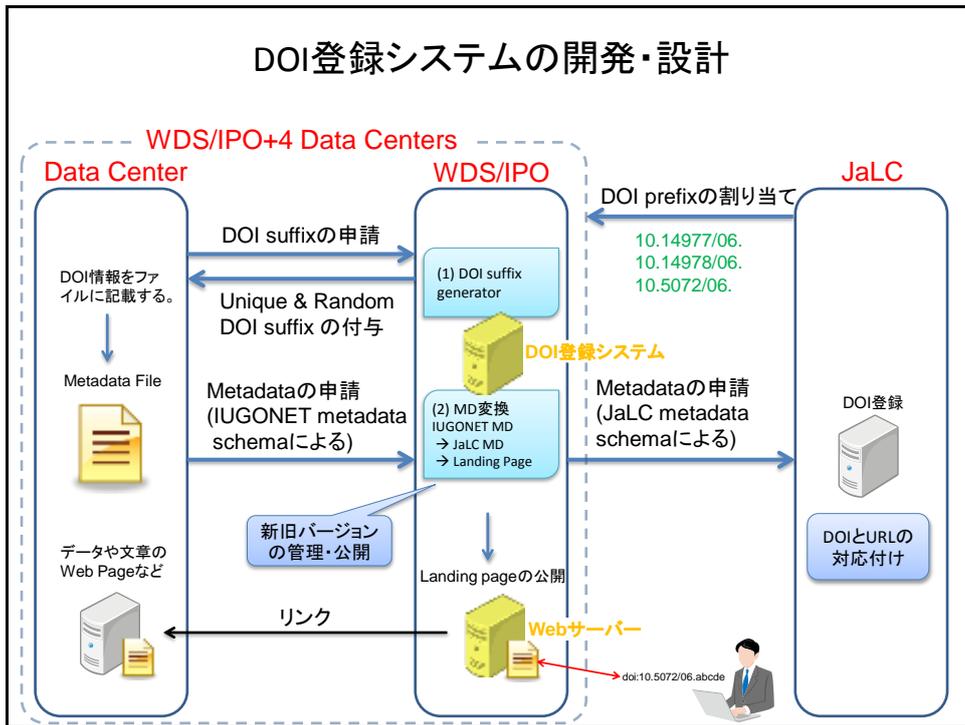


国内World Data CenterにおけるデータベースへのDOI付与

- WDS国際プログラムオフィスと国内の4ヶ所のWDCで、科学データへのDOI付与について2013年8月より話し合いを行ってきた。
 - WDC for Aurora (国立極地研究所)
 - WDC for Geomagnetism (京都大学)
 - WDC for Ionosphere and Space Weather (情報通信研究機構)
 - WDC for Space Science Satellites (航空宇宙開発研究機構)



DOI登録システムの開発・設計



稼働中のDOI登録システム

doi:10.14977/06.54c861b824d66

← DOI登録(メタデータ)ページ

↓ 自動作成されるLanding Pageの例

デモンストレーション (1)

- DOI登録システム
(1) Unique DOI generator → (2) Upload Metadata → (3) メタデータの確認
→ (4) Landing Page (Preview, Publish) → (5) To JaLC (Validate, Publish)



デモンストレーション (2)

- テスト登録したDOIのリゾルブ
(1) dx.doi.org/10.14977/06.550ea56ddb03b → (2) ランディングページ
→ (3) データページ



まとめ

- WDS国際プログラムオフィスと超高層科学データを扱う国内の4ヶ所のWDCで、JaLCの研究データへのDOI登録実験プロジェクトに参加してきた。
 - WDC for Aurora (国立極地研究所)
 - WDC for Geomagnetism (京都大学)
 - WDC for Ionosphere and Space Weather (情報通信研究機構)
 - WDC for Space Science Satellites (航空宇宙開発研究機構)
- drupalを土台とした「DOI登録システム」を開発した。
 - 各データセンターは、(1)メタデータ(IUGONET MDスキーマ)を作成してアップロードする、(2)ランディングページからのリンク先ページを用意する、ことで登録ができる。
 - データの更新は、新たなメタデータを作成し、再アップロードする。変更履歴はランディングページに掲載されていく。
 - DOI postfixはランダム文字列を採用している。(桁数は今後議論する)
- DOIテスト登録は地磁気の時系列データ(以下の3種類)について行った。概ね期待通りの結果が得られ、本登録への移行はスムーズに行えると思込んでいる。
 - dx.doi.org/10.14977/06.550ea56ddb03b → Dst index
 - dx.doi.org/10.14977/06.54c861b824d66 → Mineyama (sample 1)
 - dx.doi.org/10.14977/06.54e444fdc7ca3 → Mineyama (sample 2)

千葉大学

千葉大学附属図書館
三角太郎

1

1. テスト実施方針

研究データ管理のプラットフォームは、基本的には附属図書館で運用している**機関リポジトリシステム**を想定している。本テストでは、機関リポジトリシステムを研究データマネジメントのプラットフォームとして使用する際の課題点を抽出、解決策を検討し、最終的には実運用の業務フローの確立を目指す。本学では既存の研究データのデータベースをもつわけではなく、**研究成果のオープンアクセス化**の対象の一部として、研究データをとらえている。そのためDOIの登録テストについても、研究データに限定せず、論文等のDOI付与のフローもあわせて検討する。特に1)登録ポリシー 2)運用フロー 3)メタデータスキーマについての検討を行う。

2

実験には図書館として参加

大学図書館の使命は、**学術情報の流通**です。論文だけでなく、研究データも含めた多様なコンテンツに対応していきたいと考えています。

図書館として何ができるのか？それを探るために、実験に参加させていただきました。

1

機関リポジトリの運用指針

【運用指針の例】

千葉大学学術成果リポジトリ運用指針

平成 17 年 2 月 1 日
附属図書館運営委員会制定
平成 22 年 3 月 30 日
情報化推進企画室図書館専門部会改正

(千葉大学学術成果リポジトリ)

1. 千葉大学附属図書館は、千葉大学(以下「本学」という。)において作成された電子的な学術研究成果を収集し、千葉大学学術成果リポジトリ(以下「リポジトリ」という。)に**恒久的に蓄積・保存**し、学内外に無償で発信・提供することにより、**本学の学術研究の発展に資するとともに、社会に対する貢献**を果たすものとする。

(登録)

2. (登録対象となる学術研究成果は以下の要件を満たすものとする。

- (1) **学術的な研究の成果**であること。
- (2) 本学においてその主要な部分が作成されたものであること
- (3) 電子的フォーマットで作成されていること
- (4) ネットワークを通じて配信できること

1

2. テスト登録対象データと登録対象件数

本テストで計画したテスト登録は以下の二つのタイプ。

- 1) 既存の機関リポジトリシステム搭載のデータセット
- 2) 研究プロジェクトの成果物の登録シミュレーション

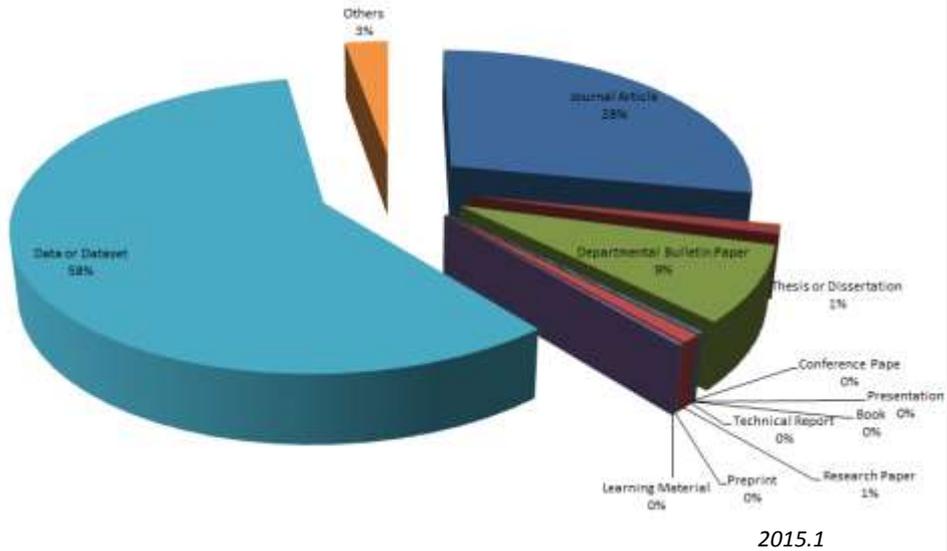
1

1) 既存の機関リポジトリシステム搭載のデータセット

現在、運用中の機関リポジトリシステム搭載のデータのうち(表1)、資料種類がデータセットの約5万2千レコードのうちの代表的なもの100件程度へのDOI付与を行い、データ登録へ要する業務量の見積りも行う。

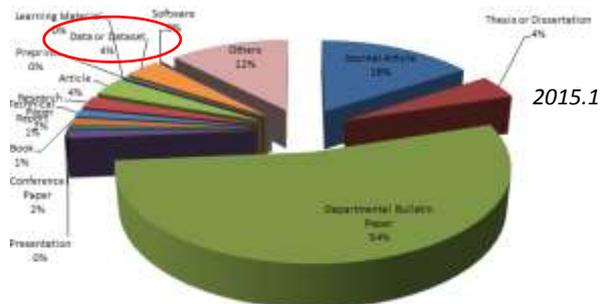
1

千葉大学機関リポジリのコンテンツ構成



1

あらためて機関リポジリの現状はどうなっているのか



日本の機関リポジリ中のデータ 52,421件
 そのうち千葉大学のコンテンツ 52,223件

99.6%

1

e-Science Project

CSI事業 2008～2009年度

実証実験として、コンテンツの作成者側と利用者側のそれぞれの立場から、e-Scienceデータ（主として千葉大学リモートセンシング研究センターの衛星画像）へ試行的にメタデータを付与できる環境を用意し、そのメタデータの評価をとおして、異種データの格納が可能なユーザコミュニティの需要に応えるメタデータ形式の検討を行い、人類の知的生産物の管理ツールとして機関リポジトリが位置づけられることをめざす。

この画像、あなたなら、どう検索しますか？



"Data-driven"

データの使い方を利用者に委ねることで研究の可能性を無限に広げてゆくのが、e-Scienceの魅力。でも、アクセスできなければ、その可能性は眠ったまま。

「最高の検索キーは、利用者が知っている」これが私たちの出発点です。

Folksonomy が、e-Science を変える。
Social Tag が、データに命を奪える。
「みんな」で動く e-Science、リポジトリから始まります。

教員と図書館のコラボレーション

2008年、IRコンテンツにソーシャルタグを付与するシステムを開発。

「メタデータから漏れていたキーワードが、タグによって補われていた」
「他のタグからの連想で、新たなタグが付与された」
実証実験によって確認された効果の一例です。

2009年は「教員と図書館の連携」を固りながら、実証実験を拡大します。

e-Science推進構築のためのデータ・コレーション 総務課元の大塚京樹 代表取締役、千葉大学、津田塾大、北海道大、金沢大、大阪大、九州大



2) 研究プロジェクトの成果物の登録シミュレーション

現在、本学のアカデミック・リンク・センターで実施されている研究プロジェクトの成果である研究データについて、機関リポジトリ上でのデータ保存・公開・DOI付与のシミュレーションを行う。本テスト内で、実際にDOIを付与するかどうかは、シミュレーションの検討結果によって決定する。アカデミック・リンク・センターのプロジェクトをとりあげるのは附属図書館サイドでハンドリングが容易なためである。アカデミック・リンク・センターは、アクティブ・ラーニングについての実践的な研究をミッションとしているが、本テストの実施母体である附属図書館と連携して研究を行っているため、プロジェクト研究活動の全体像を附属図書館サイドでも把握しているためである。

1

2) 研究プロジェクトの成果物の登録シミュレーション

実際の登録時には、大学内で進められている研究プロジェクトについて、学内の関連部署とも連携しながら、成果の登録・保存・公開の業務を進めることになるが、本テストではプロジェクトのスタートから終了までの研究サイクルの各フェーズについて、実際にどのような業務を行うべきかを検討し、課題の抽出をはかる。研究データ以外の生産物については、本テストの対象外であり、本テスト内ではDOIは付与しないが、業務フローとしては不可分のものであると予想されるため、検討の対象として含める。具体的に想定している研究データとしては、画像、動画、数値データを想定し、その他に公表論文、講演資料、各種報告書を想定している。

1

実験結果

(i)登録インターフェース

当初は、WebとXMLの双方をテストする計画であったがWebのみの個別登録にとどまっている。本登録時には、数千件オーダーのデータアップロードが必要であるが、今回登録時にEXCELにて作成したメタデータについて、インポートに失敗。データを視認した限りでは不備は見えないため、改行コード、区切記号等が原因ではないかと考えている。その原因について解決できていない。しかし実験では、メタデータ仕様等に検討を特化することとしたため、XML登録については、本登録時までテストを先送りすることとした。

1

実験結果

(ii)テスト登録

コンテンツについて、今回は本学の機関リポジトリシステムのデータセット中の最大のコレクションである、萩庭コレクション(植物標本)にフォーカスを絞ることとし、複数件について登録テストを実施した。他の学術雑誌掲載論文の附属データについて、昨年度より論文本部の登録要請時に、附属データについても登録可否をあわせて問い合わせをはじめたが、現時点ではほとんど登録が進んでいない。

1

実験結果

植物標本データについて、採集者/標本作成者/デジタルファイル作成者等の複数の人間が作成にかかわっているが、それをメタデータへどのように記述するかで苦慮した。実験データごとに、メタデータの記述方法に大きなバリエーションはあることが予想されるが、**最低限の入力ガイドライン**は必要。

1

実験結果

研究プロジェクトの成果物の登録シミュレーションについては、検証が十分にできなかった。本センターで扱うデータは、基本的に利用者の個人情報絡む。論文、報告書等にもちいる場合には抽象化・統計化し、データの解析・保存はスタンドアローンのPCで行うなどの、十分な注意をはらっている。しかし**研究データとして保存する場合に、どの段階のデータを残すべきか**は、まだまだ議論が必要。おそらく社会科学系のデータについては同様の問題が生じるが、調査結果を統計的に分析して公表する場合に、分析した元データが個人情報を十分に抽象化してから統計分析を行っているのか、統計分析を行った結果として個人情報が抽象化されたのであって、元データは個人情報が特定できる状態なのか？

1

あらためて図書館の特徴を考えて見る

- 学術情報の**アーカイブ**をミッション
 - 大学内でアーカイブの文化をもつ組織は図書館・博物館・記念室ぐらい。
- 書誌情報(メタデータ)作成のノウハウ
 - まともな図書館員ならある程度の書誌作成のスキルはもっている
- 中身を**理解できなくても扱うことができる**
 - まったく**理解できない**専門的な数学書の書誌を作ったり、**読めない**外国語の書籍の書誌を作ったりを**業務としてこなす**ようトレーニングされている

1

あらためて、図書館から どうアプローチするかを考えてみる

- 1) 最初に研究ありき
 - 最終的な研究成果物である論文があつて、その添付物としての研究データがある
- 2) 最初にデータありき
 - 博物館標本の画像・計測データとか観測データとか、まずはデータがあつて、それがみな論文になっているとは限らない

1

アプローチ1

1) 最初に研究ありき

- 最終的な研究成果物である論文があつて、その添付物としての研究データがある

図書館員にはわかりやすいアプローチだが、しかし、いつまで論文が最終的な研究成果物であるかもわからない。研究データ流通の構造そのものが大きく変わろうとしているのでは？

論文は研究データのメタデータ？

1

アプローチ2

2) 最初にデータありき

- 博物館標本の画像・計測データとか観測データとか、まずはデータがあつて、それがみな論文になっているとは限らない

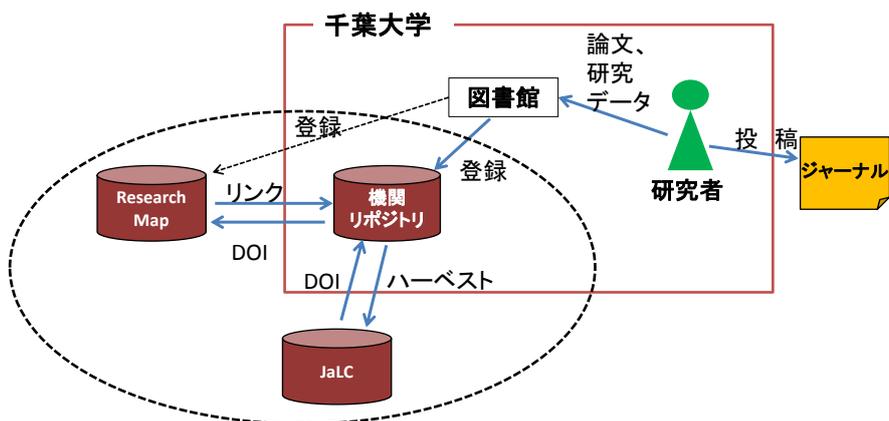
テキスト情報が付与されていないもののメタデータは困難。

ノウハウがあれば、ある程度は・・・。

1

課題：業務フロー検討案

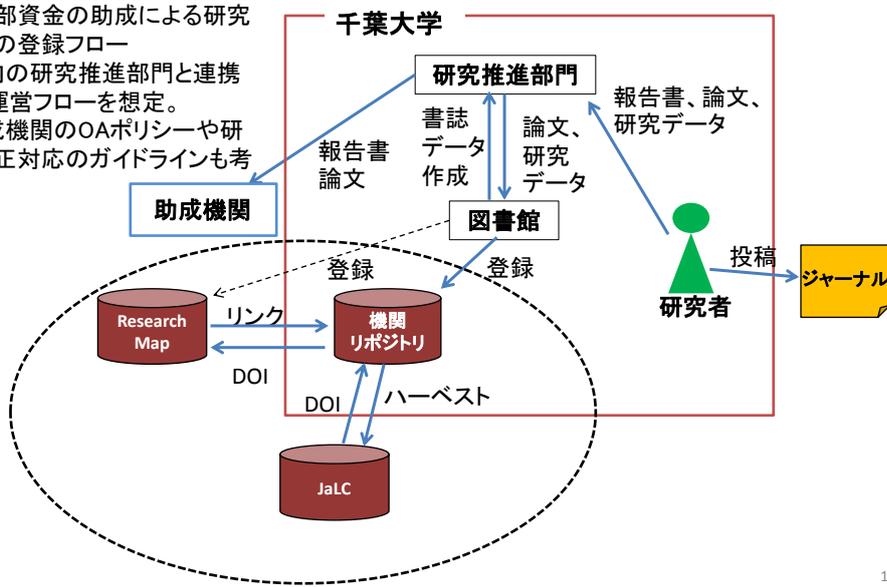
- 1) 一般的な研究データ登録フロー
 図書館と研究者と連携した運営フローを想定。



1

課題：業務フロー検討案

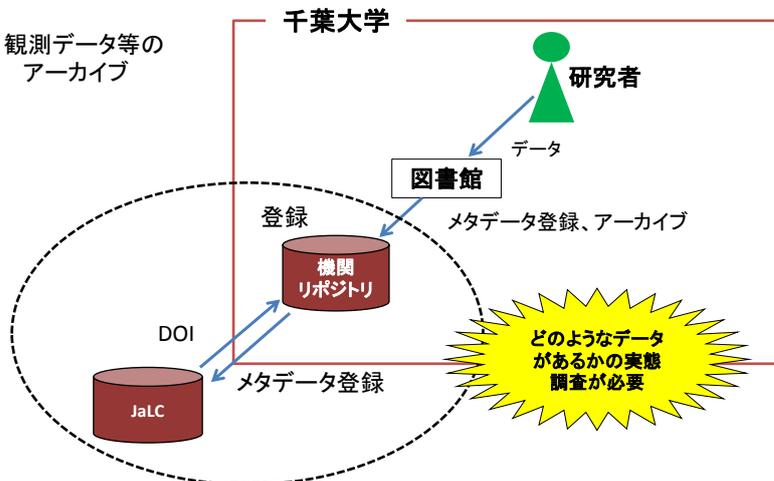
- 2) 外部資金の助成による研究成果の登録フロー
 ・学内の研究推進部門と連携した運営フローを想定。
 ・助成機関のOAポリシーや研究不正対応のガイドラインも考慮



1

課題：業務フロー検討案

3) 観測データ等の
アーカイブ

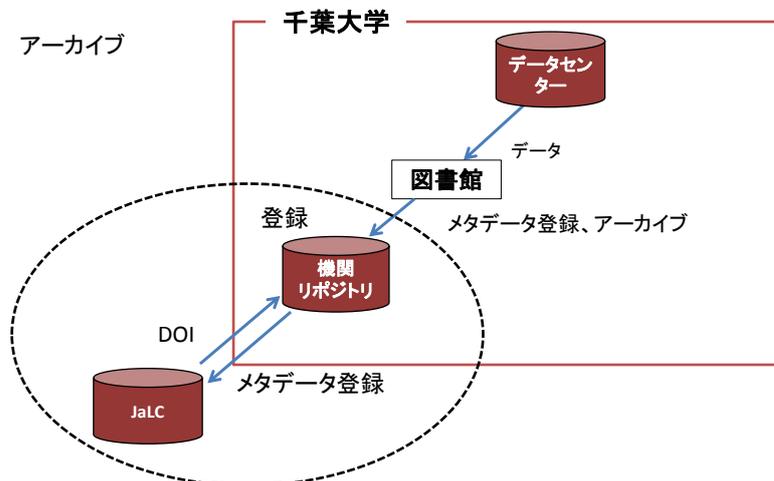


研究データそのものを機関リポジトリで
アーカイブすることは現実的ではない
かもしれないが、機関として

1

課題：業務フロー検討案

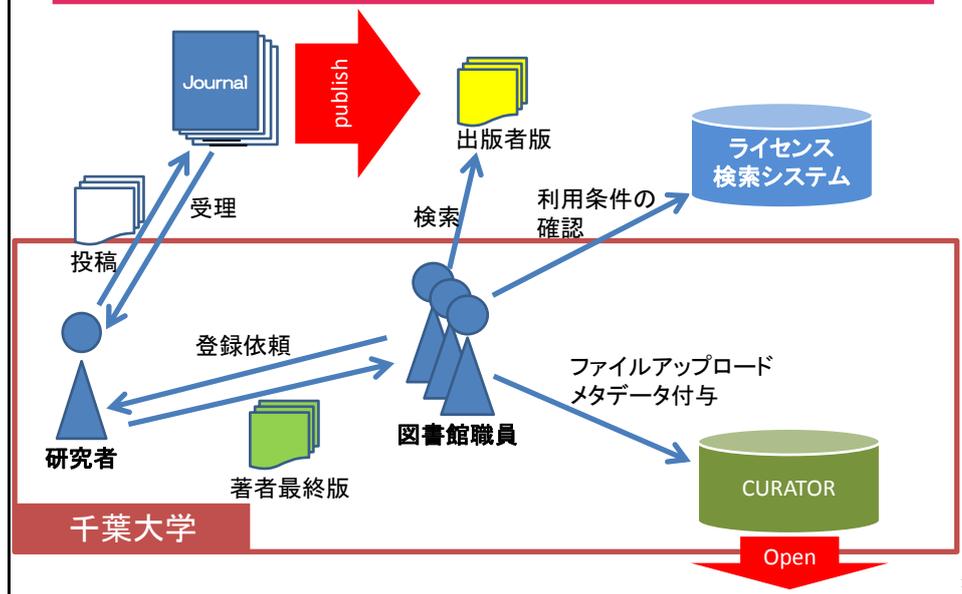
3) アーカイブ



研究データそのものを機関リポジトリでアーカイブすることは現実的ではない
かもしれないが、機関としてリソースのカタログを作りDOIを付与することは重
要ではないか？

1

機関リポジトリ登録フロー(図書館依頼)



作業負荷

- 実はメタデータ付与、ファイルアップロードよりも、**ライセンス処理**や**研究者との連絡調整の前処理**の負荷のほうがはるかに重い(研究データの著作権は?)。
- メタデータ付与やファイルアップロードは、手順が整理できれば、それほど大変ではない。
- 本来は、研究者が自ら発信するものだが、待っていたら集まらないので、現実には図書館員が研究者に個別にコンタクトして、**登録発信を促している**例が多い。
- **研究業績データベース**との連携は一部の大学ではリンク機能を実装済み。ただし**標準ができていない**(機関リポジトリ側には著者IDが入っていない、業績データベース側はそもそもメタデータ標準がない)ので、**システム継承は力技**(ResearchMapの利用は拡大するか?)。

課題：データマネジメントのポリシー

日本の大学図書館で作成しているところは、おそらくまだないが、何を集めて何を集めないのか？データ移行をどう考えるか？データ粒度をどう考えるか？などなど検討しなければならないことだらけ。

論文より明らかに複雑。そもそも**データの定義**からはじめなければならない。

個々の分野では参考例があるが、総合的に参考にする例がない。

1

課題：メタデータスキーマ

○機関リポジトリの国内のスキーマ: junii2

<http://www.nii.ac.jp/irp/archive/system/junii2.html>

○ JaLCのスキーマ

○ DataCite

<https://www.datacite.org/>

いずれを用いるにせよ、まともにやろうと思ったらシステムの改修が必要

1

スキーマだけでは品質は保持できない

入力(記述)マニュアルは必要！

コンテンツのどこをどのように見て、どのように転記して、どのように転記するかは、図書館員では判断できない。

しかし、**マニュアルがあれば**、それを守るのが図書館員。

1

課題：システム

- Dspace、CKAN等、何を選ぶのか？
- ディスク容量はどれぐらい必要か？
- メタデータスキーマは？
- DOI付与システムは？
- 既存の機関リポジトリシステムでどこまで対応可能か？
- ワークスペースとオープンスペースとアーカイブスペースは一緒に良いのか？(机と書架と書庫)

1

大学図書館全体のとりくみ

機関リポジトリ推進委員会

<https://ir-suishin.repo.nii.ac.jp/>

「大学共同利用機関法人情報・システム研究機構国立情報学研究所と国公私立大学図書館協力委員会により設立、「機関リポジトリを通じた大学の知の発信システムの構築」に関する事項を企画・立案し、学術情報の円滑な流通及び発信力の強化にかかる活動を推進することを目的とする。

1

オープンサイエンス班の平成28年度活動計画

- 【1】OAポリシーの策定支援ツールの開発
- 【2】RDMトレーニングツールの開発
- 【3】研究データ対応メタデータスキーマの検討
- 【4】ケーススタディによる研究データ管理
ノウハウの蓄積

1

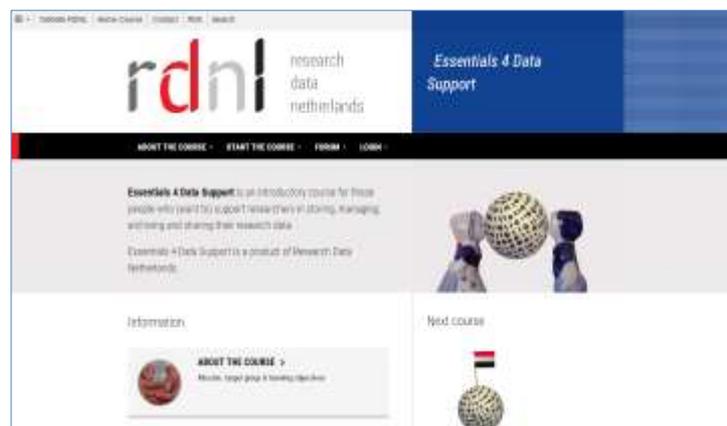
RDMトレーニングツールの例



<http://rdmrose.group.shef.ac.uk/>

1

RDMトレーニングツールの例



<http://datasupport.researchdata.nl/en/>

1

RDMトレーニングツールの例



<http://datalib.edina.ac.uk/mantra/>

1

Thank You

1

機関リポジトリおよび セルフアーカイブシステム「NIMS eSciDoc」を 介したDOI取得・登録の実証実験報告

2015年7月3日
国立研究開発法人物質・材料研究機構
企画部門科学情報室
田辺 浩介



1

NIMS eSciDocとは

- 物質・材料研究機構(NIMS)で運用する
セルフアーカイブ
– 文書系アーカイブシステム「PubMan」と
画像系アーカイブシステム「imeji」で構成



PubMan <http://pubman.nims.go.jp>



imeji <http://imeji.nims.go.jp>



2

NIMS eSciDocの利用

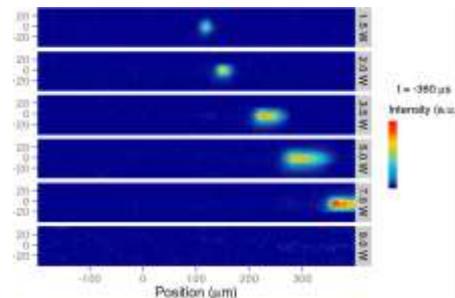
- 機構の全職員が利用可能
 - 研究職、エンジニア職、事務職など
- 登録と公開は原則として、利用者自身の責任
において行う
- システムの運用と管理は科学情報室が行う



2

研究データ登録のテスト実施方針(1)

- 画像系アーカイブシステム「imeji」に画像データを登録し、JaLC WebAPIを用いてDOIを付与
 - 電子顕微鏡で撮影した画像など約150点
 - ファイバーフェーズ
(右側画像参照)
 - フラレンナノ
ウィスカー



2

研究データ登録のテスト実施方針(2)

- 研究データと他のサービス間の相互リンクを付与し、情報の関係性を示す
 - PubMan・imeji間の相互リンク
 - ORCID、研究者総覧「SAMURAI」など



SAMURAIでのPubManとのリンクの例

- 本実験プロジェクトでは、画像アップロードとメタデータ登録は研究者が、DOIの取得は科学情報室が行う



5

テスト登録方法

- 「DOI管理システム」を利用してDOIを登録
 - JaLCからDOIを取得し、結果をセルフアーカイブシステムに送信する機能を持つ
 - セルフアーカイブでの登録者(研究データの作成者)を同定する手法として、ORCIDとの対応付けを可能とする機能も備える



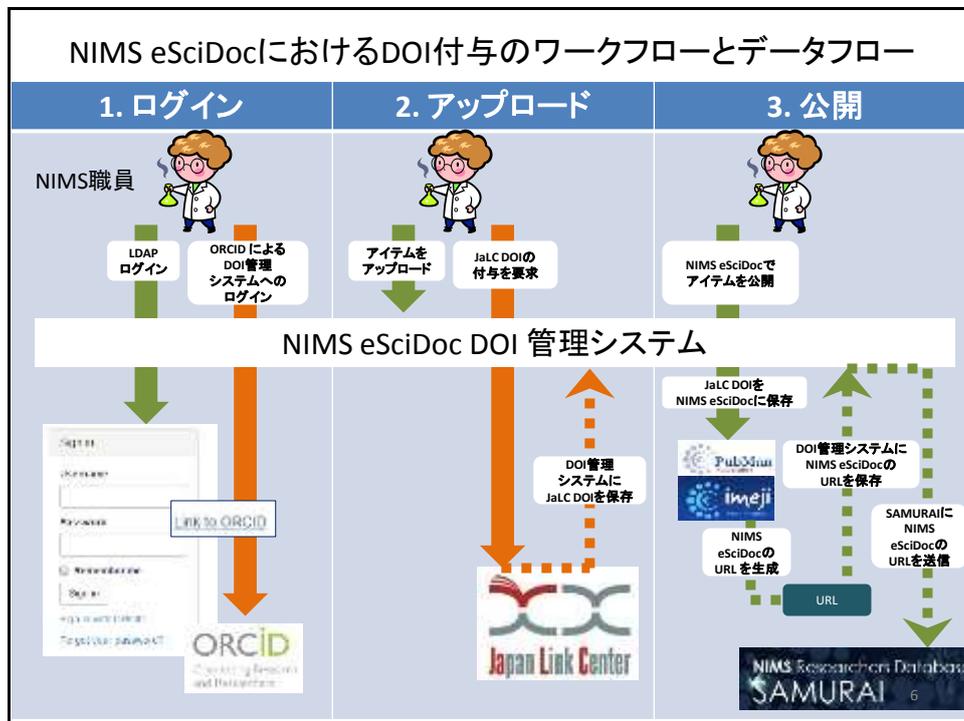
3

DOIのフォーマット

- 機関プレフィックス/nims.通し番号
– 例: 10.11503/nims.1001
- DOIの文字列に組織名やORCID番号を含める
という意見もあった
 - DOIのみで登録者の情報がわかるようにしたい、
という要望
 - 登録者の情報はJaLCやCrossRefのWebAPIで取得
できるので見送り



6



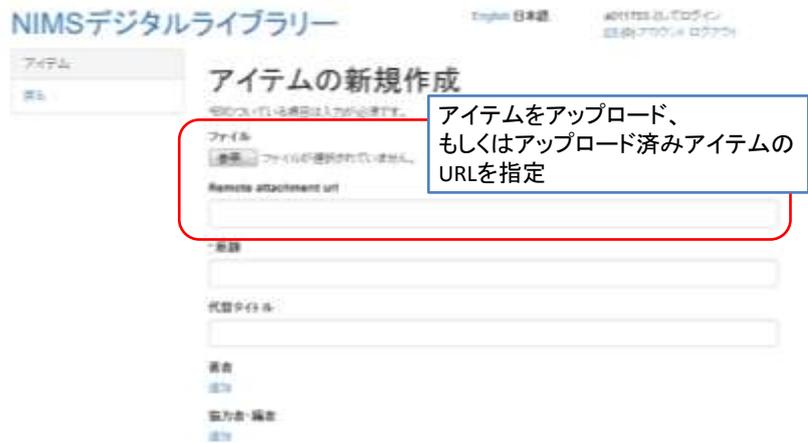
DOI管理システム



- 図書館管理システムNext-L EnjuをベースにDOI付与機能を追加
- PubMan・imejiを補完するサブシステムとして動作
- セルフアーカイブ利用者(研究者・図書館員)がログイン可能



アイテムの登録画面



アイテム詳細画面とDOIの取得リンク

English 日本語

アイテム

アイテムの表示

アイテムは正常に更新されました。

次へ 前へ 戻る 検索語: 検索 条件を詳しく

[M] テストB

著者: 協力者・編者: 出版者:

形態: 電子オンラインリリース / 画像画像

言語: 日本語

出版社:

DOIを取得する

NIMS 6

JaLC DOIの取得画面

English 日本語 ログイン

NIMSデジタルライブラリー

JaLC DOI

検索

JaLC DOIの取得

User: 8011755

Manifestation: マテリアル

DOI: 10.14877/00.0000.3447

URL: http://www.nims.go.jp/manifestations/3447/

登録する

© 2015 National Institute for Materials Science. Powered by NIMS eDoc and HOD L. Eiga.

NIMS 6

DOI取得成功

ライブラリー

English 日本語

2011753 としてログイン
詳細 (0) アカウント ログアウト

User:2011753

Manifestation:テストB

Status:success

DOI:10.14977/08.nims.3447 (JaLCテスト環境)

```
XML response: <?xml version="1.0" encoding="UTF-8" standalone="yes" ?><root><head>
<totalSize>1</totalSize><okcount>0</okcount><ngcount>0</ngcount></head><body><result><step>1</step>
<resultstatus>0</resultstatus><doi>10.14977/08.nims.3447</doi></result></body></root>
```



6

テスト結果

- 達成できた点
 - JaLCのWebAPIを用いたDOIの付与
 - ORCIDのWebAPIを用いた、研究データのアップロード者に対するORCIDの紐付け
- 達成できなかった点
 - セルフアーカイブシステム上で研究データとDOIを体系的に表示する機能
 - imejiのアイテム更新用WebAPIの不具合等、複合的な理由によりテスト実施に至っていない



6

テストの考察

- 登録データの選定・品質保証・メタデータ付与・公開範囲の設定は、利用者各自の責任で行うことになっている
- DOIの管理も同じルールでよいか検討が必要
 - 以下の場合、DOIはセルフアーカイブ上でどう扱うべきか
 - 利用者が新しい版としてデータをアップロードした
 - 利用者がメタデータを編集した
 - 利用者がデータを取り下げた



6

今後の課題(1)

- DOI付与のルールの検討
 - セルフアーカイブにおいて、どのアイテムをDOIの付与対象とするか
 - 付与にあたってスクリーニングを必要とするのか



6

今後の課題(2)

- ワークフローの検討
 - ORCIDを登録に必須とするかどうか
 - セルフアーカイブしたアイテムを取り下げた場合のランディングページ的设计
- セルフアーカイブシステム自体の改良
 - アイテム更新用WebAPIの修正
 - DOIの表示、ライセンスの表示



理化学研究所
脳科学総合研究センター
神経情報基盤センター (NIJC)
における、研究データへのDOI登録実験プロジェクトへの取り組み

磯野由依、奥村嘉宏、山口陽子
理化学研究所脳科学総合研究センター
神経情報基盤センター

1.
神経情報基盤センター(NIJC)紹介

• 神経情報基盤センター

(Neuroinformatics Japan Center: NIJC) について

- 理化学研究所、脳科学総合研究センター内の1部署
- 国内のニューロインフォマティクス研究に関する活動拠点
- INCf:ニューロインフォマティクス国際統合機構(カロリンスカ研究所、スウェーデン)の日本ノードの事務局としての役割をもつ

• Neuroinformatics とは

ニューロインフォマティクスは脳神経科学と情報科学・技術の接点であり、脳(神経)科学を発展させるために、多様で膨大なデータを総合的に整理、分析する新しい研究分野。ニューロインフォマティクスは次の3つの技術要素を開発し、脳科学を総合的に整理して行く。



1. データベース&データ共有

一次データ(正常/疾病状態にあるあらゆる生物種や組織標本からの、ゲノム、分子、構造、細胞、ネットワーク、システム、及び行動のレベルでの実験及び実験条件)、オントロジー、メタデータの共有する標準と仕組みの開発。

2. 解析のためのツール開発

データを操作し、管理するツールの開発。神経科学分野では必要に応じて目的毎に特別な分析ツールやアルゴリズムを共同で設計、開発。

3. 計算論的モデル

データを使って検証可能な脳の構造や機能の計算論モデルの作成。神経科学で得られた実験のデータセットとニューロインフォマティクスによって開発された数理モデルの結果を比較する。

プラットフォーム(PF)について(承前)

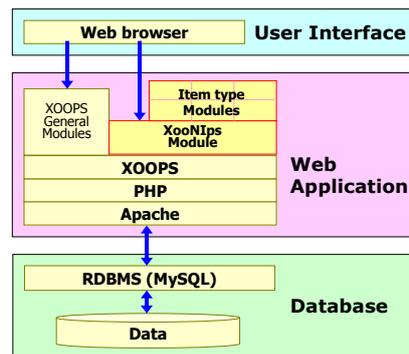
- ニューロイメージングPF <https://nimg.neuroinf.jp>
 - » 脳活動イメージング
- ダイナミックブレインPF <https://dynamicbrain.neuroinf.jp>
 - » 脳の行動原理、細胞、局所回路、脳全域、行動、実験的、計算論
- 包括脳PF <https://cbsn.neuroinf.jp>
 - » 脳科学、神経科学、行動学、遺伝学
- マウス表現型データベース <http://www.mouse-phenotype.org>
 - » 脳科学、神経科学、行動学、遺伝学
- シミュレーションPF <http://sim.neuroinf.jp>
 - » モデルシミュレーション、仮想マシン
- 脳科学辞典 <http://bsd.neuroinf.jp>
 - » 脳科学一般
- ViBrism データベース <http://vibrism.neuroinf.jp>
 - » 3D-ISMデジタル脳
- BSIニューロインフォマティクス <http://bsi-ni.brain.riken.jp>
 - » 理化学研究所脳科学総合研究センター(BSI)の研究成果公開

データベースについて

- いくつかのPFはDB基盤システムXooNIps上にデータベースを構築している

XOOPS(eXtensible Object Oriented Portal System)
PHPで記述されたオープンソースのCMS

XooNIps
XOOPSにニューロインフォマティクス(NI)のための拡張を加えたデータベース基盤システム

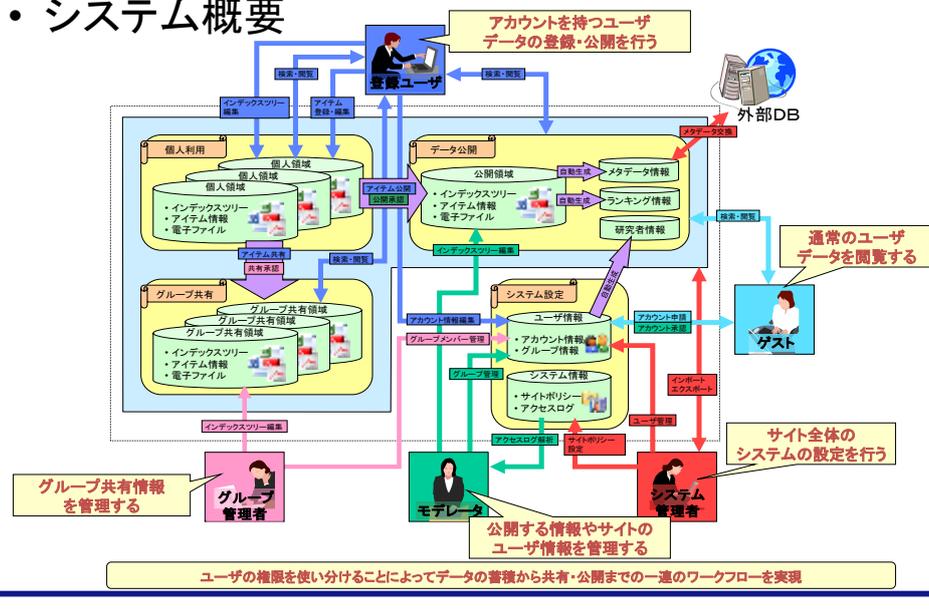


• XooNIPSとは

- インターネット上で多様な電子データ(ファイル, URL, 書籍・論文等)をメタデータと共に共有を図るためのWebデータベース基盤システム



• システム概要



• XooNlpsの特徴

1. データベース構築の柔軟さ

CMSを元として開発されているため、モジュールによる拡張が容易で、だれでもシステムを立ち上げることができる

2. 扱うデータの多様性、拡張性

多様な種類のデータを扱うためのアイテムタイプ(メタデータ形式)が標準で備えられている。さらに、新たなデータ形式を扱うための拡張方法も準備されている

3. メタデータの流通機能

標準プロトコルのひとつであるOAI-PMH(Open Archive Initiative-Protocol for Metadata Harvesting)を実装しているため、OAI-PMHを備えたデータベースとメタデータを介して協調させることができる

• XooNlpsの導入例

- ニューロインフォマティクス分野にとどまらず、様々な大学・学術研究機関の機関リポジトリとしても採用されている

• 学術情報リポジトリ

- 青森県立保健大学リポジトリ A-plus(アプラス)
- 旭川医科大学 学術成果リポジトリ AMCoR (Asahikawa Medical College Repository)
- 近畿大学 学術情報リポジトリ (Kinki University Academic Resource Repository)
- 慶應義塾大学 学術情報アーカイブ KOARA (KeiO Academic Resource Archive)
- 埼玉県地域共同リポジトリ SUCRA (Saitama United Cyber Repository of Academic Resources) - 埼玉大学, 文教大学, 城西大学, 埼玉女子短期大学, 跡見学園女子大学, 国立女性教育会館, 駿河台大学, 淑徳大学国際コミュニケーション学会, 埼玉純真短期大学, 埼玉県立大学, 共栄大学
- 聖学院学術情報発信システム SERVE (Seigakuin Repository for academic archive)
- 多摩大学 学術情報リポジトリ Tama蔵(たまくら)
- 同志社女子大学 学術リポジトリ
- 東邦大学 学術リポジトリ
- 奈良県立図書館 奈良県地域史料目録
- 奈良大学リポジトリ
- 別府大学 機関リポジトリ BUILD (Beppu University Information Library for Documentation)
- 別府大学 地域連携プログラム BUNGO (Beppu UNiversity and Oita reGional society co-Operation)
- 桃山学院大学リポジトリ STARS (St.Andrew's University Repository System)

2.

研究データへのDOI登録実験プロジェクトへの取り組み

- 参加目的
- 公開運用中のDBに登録されている研究データを世に広めるための手段としてDOIを付与
 - 研究成果データ公開に向けたモチベーションの向上
 - 研究成果へのアクセシビリティの向上
 - 運用者に求められる要件の見極め
 - 品質・データ持続性の担保等
- 実運用におけるDOI登録方法の確立
 - DOIに正しく登録できることの確認
 - 各DBの研究データのメタデータ互換性、マッピング
 - DBにおけるデータ公開からDOI登録までのスキームの確立
 - そのためのシステムの検討

DOI登録支援システムの検討

各PFの研究データにDOIを付与するまでの流れを支援するシステム

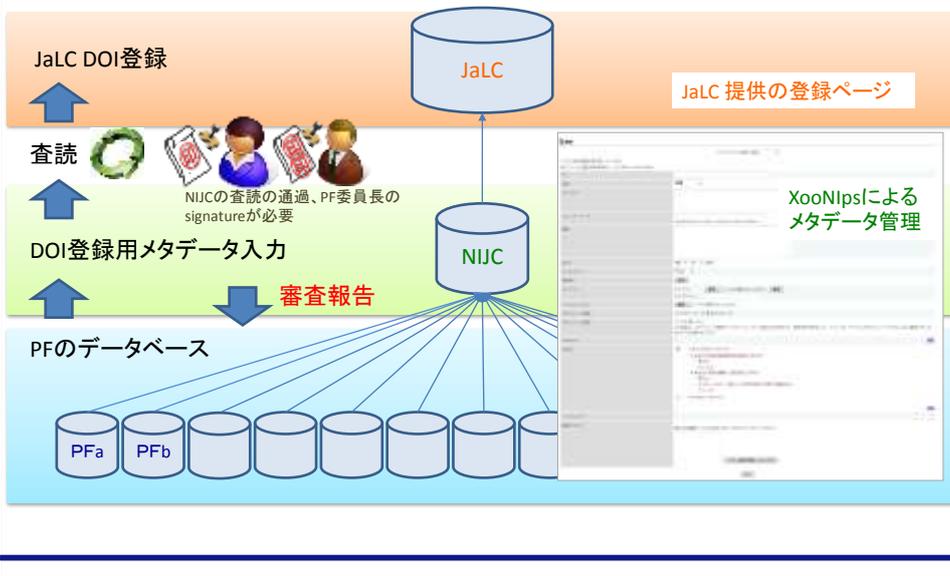
機能要件:

- 各PFの運用者からメタデータ提供を受ける
- DOI登録ポリシーに沿うかどうかを確認する(査読)
 - データの種類・粒度
 - メタデータ
 - 永続性の有無
- 一括登録
- 付与したDOI一覧の管理



- XooNips のもつデータ登録・査読・管理機能を活用
 - DOIメタデータ収集用のアイテムタイプの定義

DOI登録の流れ



- **現状**

- 各PFにアンケートを実施し、DOIに対する意識調査、DOI登録候補コンテンツを幾つかあげるよう依頼
 - 本登録への希望がPFより集まっている
- 登録支援システムの試用開始
 - DOI本登録までは至っていない

- **今後の展開**

- DOI登録支援システムの運用を通じて
 - データ登録担当者、承認担当者などのシステム利用者からのフィードバックを得て、問題点や改善点を洗い出し、より洗練されたものへとシステムを改良することにより、DOI登録までの円滑なスキームを確立していく
- 本プロジェクト中に何件か本登録を実施する
- これにより得られた知見は、今後も継続して各方面へフィードバックしていきたい

ありがとうございました

4.6 DOI テスト登録を通じて抽出された課題のガイドラインへの落とし込み

参加機関から事前に提示された「想定される主な検討課題」や、DOI テスト登録を経て抽出された課題について、ミーティング等で検討し、別冊の『JaLC における研究データへの DOI 登録ガイドライン』としてとりまとめた。

4.7 DOI テスト登録を通じて抽出された課題のシステム面での対応

また、DOI テスト登録を経て抽出されたシステム面での課題は、次表のとおり対応した。

表 4-7-1. DOI テスト登録を通じて抽出された課題のシステム面での対応

区分	タイトル	内容	発生	発生日	対応状況	対応日	回答
1		(別表 A 参照)					
2	不具合	ファイルアップロードでログアウト	研究データ実験 PJ ミーティング時	2015/1/28	検討課題	2015/1/29	"ファイルフォーマットの確認の仕組み導入には工数がかかるのですぐの対応は困難です。
3	要望	英語画面	研究データ実験 PJ ミーティング時	2015/1/28	対応済	2015/2/27	対応表を作成して配布しております。(2/27JST)
4	質問	ブラウザのバージョン	研究データ実験 PJ ミーティング時	2015/1/28	対応済	2015/1/28	"IE11, Safari 44.0 は対応しています。Firefox は崩れる場合があります。"
5	要望	フィールド名とエレメントの対応	研究データ実験 PJ ミーティング時	2015/1/28	対応済	2015/2/27	対応表を作成して配布しております。(2/27JST)
6	要望	エラーメッセージ	研究データ実験 PJ ミーティング時	2015/1/28	検討課題	2015/1/28	システムのつくりのため、すぐに対応は困難です。今後の検討課題とさせていただきます。(1/28JST)
7	要望	Web アップロード時のエラーメッセージについて	研究データ実験 PJ ミーティング時	2015/4/20	検討課題	2015/7/30	今後の検討課題とさせていただきます。
8	要望	XML エラーメッセージについて	研究データ実験 PJ ミーティング時: DIAS 真板様	2015/6/3	検討課題	2015/7/30	今後の検討課題とさせていただきます。
9	要望	データ登録・登録データ確認に関する要望一式	理化学研究所 磯野様	2015/4/14	検討課題	2015/4/16	Web 画面につきましては、論文等の登録画面と共通の作りとしており研究データの入力に必ずしも最適でない部分があるかと思えます。頂きましたご要望につきましては、システムの改修が必要となりますので、今後のエンハンスの際の参考とさせていただきます。
10	要望	データ登録に関する要望	理化学研究所 磯野様	2015/4/14	対応せず	2015/4/16	画面からの入力時は、1 言語での入力か、複数言語での入力かの判別が難しいため、記述言語をプルダウンで入れて頂く形としております。
11	要望	Web での入力について	研究データ実験 PJ ミーティング時: 千葉大学付属図書館様	2015/4/20	検討課題	2015/7/30	今後の検討課題とさせていただきます。
12	要望	データ登録時のタイムアウトについて	理化学研究所 磯野様	2015/4/14	検討課題	2015/4/16	現行のタイムアウト時間の設定は、JaLC 正会員が論文等を画面から登録する際に支障のない程度として 30 分と定めた経緯がございます。タイムアウトを避ける方法として、例えば、全ての入力項目を一気に入力するのではなく、必須項目のみ最初に入力して一旦登録を行い、その後(あまり間をおかず)入力内容を呼び出して入力し更新(変更)するような入力方法も可能です。
13	要望	メタデータ拡張仕様の Excel ファイルが欲しい	極地研 矢吹様	2015/4/17	対応済	2015/4/17	研究データのメタデータ設計書のエクセルファイルをメールで指定の場所にアップロードしました。

14	質問	JaLC2Web 画面の遷移	XML の登録のときにシステムエラーとなったが、その後確認したらデータは正常に登録できていた	研究データ実験 PJ ミーティング 時: NBDC 様	2015 /4/20	対応 せず	2015 /4/20	Web 画面の遷移の際に、ブラウザの戻るボタンを押すとシステムエラーという表示がでることがあります。ブラウザの戻るボタンをサポートしていないので、お手数ですが、画面内にある「前画面に戻る」などのボタンを押して操作してください。
15	要望	DOI のリゾルブについて	データを登録したあと、すぐにリゾルブが働かを確認できない。DOI 有効化処理が現在週 2 回のことだが、もっと回数を増やして欲しい。	研究データ実験 PJ ミーティング 時: 産総研様	2015 /4/20	対応 済	2015 /7/30	実験環境では、週 2 回の運用としております。本番環境では日次運用を行っております。
16	要望	"	デポジット頻度の増加。	研究データ実験 PJ ミーティング 時: DIAS 真板様		検討 課題	2015 /7/30	本番環境では、毎日 DOI 登録を行います。DOI 登録のタイムラグ短縮につきましては今後の検討課題とさせていただきます。
17	要望	XSD (スキーマ) について	今配布されているスキーマだと要素が必須のものをチェックしていない。要素のチェックを厳密にやるようなスキーマを用意して欲しい。	研究データ実験 PJ ミーティング 時: NICT 様	2015 /4/20	対応 中	2015 /7/30	厳密にチェックを行うスキーマを検討中です。(7/30JST)
18	不具合	研究データの旧メタデータスキーマで登録したデータにおいて代替識別子が画面で確認できない。	10.14977/09.db10143<http://dev.japanlinkcenter.org/jalc/linkSakiAnnaiOto/bookMetaDoiSearch/index?key=NlJCO12215&doi=10.14977/09.db10143>は代替識別子が繰り返し項目となつてからの登録ですが 10.14977/09.vision7111 はそれらが繰り返し項目になる以前の登録でこちらは登録したはずの代替識別子やサイズなどの情報を web でも api でも確認できないのです。	理化学研究所: 磯野様	2015 /5/22	対応 済	2015 /6/5	旧メタデータスキーマで登録されたレコードを新メタデータスキーマで登録されたテーブルにコピーしてなかったため。5/29 に実験環境のレコードをコピーして解消しました。
19	不具合	JaLC2Web 画面で研究データの権利 URI 入力が必要になっている	権利情報を入力する場合、テキストと URI の両方が必須となっておりますが、もともとのメタデータがテキストのみの場合にどうするべきなのかが判りませんでした。	理化学研究所: 磯野様	2015 /5/22	対応 済	2015 /6/30	設計書では画面の方が必須、XML メタデータ項目表では任意となっていた。⇒ 画面を任意に変更する。6/12 実験環境を修正しました。6/30 本番環境を修正しました。
20	要望	JaLC2Web 画面でのデータ登録時に余計なスペースを入ると、検索時にそのスペースを入れないとヒットしない	登録時に "Neuroinformatics Japan Center" などと前後に余計なスペースを入れてしまうとトリミングされないので検索時それをキーとした場合に見つからないということがありました。できれば、DOI 登録時のバリデーションで登録項目の前後のスペースが、デフォルトでトリミングされると嬉しいのですが。(前後のスペースが重要な意味をもつことがあるかもしれないので追加のパラメータが必要かもしれませんが)	理化学研究所: 磯野様	2015 /5/22	検討 課題	2015 /7/30	今後の検討課題とさせていただきます。
21	要望	XML ファイルによる一括登録画面	リクエスト種別のラジオボタンを「登録」「更新」を区別した方がいいのではないかと。ヒューマンエラーの回避の意味で。	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /4/20	対応 せず	2015 /4/20	CrossRef や DataCite など DOI 登録の新規登録と更新の区別を行っていません。「更新」を区別しないことでデポジット側も修正が容易です。
22	要望	登録処理	登録の処理時間が長い。Dry run または XML ファイルのバリデーションだけを行うモードが欲しい。	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /4/20	対応 済	2015 /7/30	処理時間については、オンラインからバッチ起動をしているため、時間がかかっています。XML ファイルのバリデーションについては、XSD によるチェックの厳密化を検討中です。
23	要望	登録処理	リザルト画面がシンプル過ぎる。もっと情報を出して欲しい。	研究データ実験 PJ ミーティング 時: DIAS 真板様		検討 課題	2015 /7/30	大量に DOI 登録を行うケースがあることを踏まえ、結果画面はシンプルにしています。
24	要望	不正データの振る舞い	エラーの原因が画面に表示されない。XML ファイルを開かないとエラーの原因が解らない。	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /4/20	対応 中	2015 /7/30	大量に DOI 登録を行いエラーが大量に発生する場合にも対応できるようにするため、現在はファイルにて提供しています。
25	質問	ウェブ IF と XML	ウェブインターフェイスから XML ファイルをアップロードすると <request_kind> は無視されるが、API 経由の時は <request_kind> が参照される	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /5/22	対応 済	2015 /6/5	ご指摘のとおり、Web 画面から XML ファイルをアップロードすると、画面での指定の方が優先され、XML ファイル側の指定は無視されます。現状では、API 経由での XML ファイルアップロード時と同じ XML ファイルを Web 画面からのアップロード時にも流用可能な形としております。
26	要望	IF	<head> は、サーバの動作を指定するもので必要ではないと考えられる	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /5/22	対応 済	2015 /6/30	今後の検討課題とさせていただきます。
27	不具合	エラーメッセージ	「筆頭著者」等という、仕様書に定義されていない用語が使用されている。	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /6/3	対応 せず	2015 /7/30	今後のドキュメントの見直しの中で修正等の対応を行います。
28	不具合	IF	仕様が直感的でない。同名の属性なのに仕様が異なり、混乱する。	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /6/3	対応 予定	2015 /7/30	今後の検討課題とさせていただきます。
29	不具合	設計書	PDF 版仕様書と XML スキーマ版仕様が一致していない。	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /6/3	検討 課題	2015 /7/30	PDF 版と XML スキーマ版の仕様の版の違いかと思えます。今後のドキュメントの見直しの中で正しくなるよう対応いたします。
30	要望	IF	サーバのテスト、デバッグの実施。	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /6/3	検討 課題	2015 /7/30	再テストを実施中です。もしお気づきの不具合がございましたらお手数ですがご連絡頂けると助かります。
31	要望	IF	登録 -> 削除 -> 同じ DOI で登録、が可能。共時的にはユニークだが、通時的には二重。サーバ上に削除記録を残しておくべきでは？エンドユーザの記憶力を信用してはいけない。意図せずに「使い直し」してしまう可能性。過去に使用歴のある DOI で「新規登録」の場合は、警告やエラーが出るようなシステムが望ましい。そもそも、	研究データ実験 PJ ミーティング 時: DIAS 真板様	2015 /6/3	対応 せず	2015 /7/30	DOI の削除時は、DB 上では削除記録を付与してレコードを残しています。また、間違っただけに使っている DOI を更新してしまった場合は、元に戻す更新デポジットを行うことでリカバリーできます。

		簡単に削除できてしまっているのか？				
--	--	-------------------	--	--	--	--

【別表 A】 DIAS 様から頂いた質問・要望への回答

項番	要素・属性	内容 (質問・要望など)	備考 (問題点・定義の整理)	回答
1	sequence 属性 (content/creator/contributor/affiliation_name 要素)	【入力値について】 xs:string 型で定義されており、備考欄に「同一 XML ファイル内でユニークにする」とあるが、ユニークであれば問題ないか？数値として、順列関係は評価されないか？あるいは、もし順序の評価だけに使うのであれば、sequence 属性を使わずとも、タグの順番で代用できないか？独自拡張の必須指定とするならば、定義と用途を明確にしていきたい。	【JaLC 独自属性 (必須指定)】 ・ JaLC (research_data_v1.21.xsd)の定義 <xs:attribute name=""sequence"" type=""xs:string"" use=""required""/>	content に関してはユニークな半角数字であれば問題ありません。creator/contributor/affiliation_name に関しては、筆頭にしたいものに""1""を、それ以降順番に""2""、""3""、…としてください。 XML 中の記述方法として""2""のあとに""1""が来ても構いません (""1""を筆頭として登録します)。 現行の仕様は、ジャーナルアフィリエイトなどのコンテンツと同様としたためこのようになっております。 Sequence 属性使わずタブの順番で代用する方式については、今後の検討課題とさせていただきます。
2	doi 要素	【表記ルールについて】 DataCite では 10 (ディレクトリサーバー番号) で開始される xs:token 型のパターンとして細かく定義されているが、JaLC では xs:string 型の自由記述となっている。両者のマッピングに問題はないか？	【JaLC と DataCite で仕様が異なる】 ・ DataCite 3.1 の定義 <xs:simpleType name=""doiType""> <xs:restriction base=""xs:token""> <xs:pattern value=""10Y.#/+""/> </xs:restriction> </xs:simpleType>	マッピングは問題ありません。 JaLC では、DOI プレフィックスが JaLC に登録されているものと一致しているかの確認を行った上で、DataCite ヘドボジットしています。DOI のサフィックスは自由記述ですが、DataCite で扱える文字は下記となっております。 0-9 a-z A-Z - (dash) . (dot) _ (underscore) + (plus) : (colon) / (slash)
3	titles 要素 subtitle 要素	【DataCite の titleType 属性への対応要望】 JaLC では DataCite の titleType のうち subTitle のみ採用されているが、TranslatedTitle、AlternativeTitle についても対応して欲しい。右の欄に記載した通り、AlternativeTitle を用いて、略称/正式名称の使い分けが可能。また、TranslatedTitle を用いて、日本語と英語のタイトルのうち、正式なほうを title に記述し、翻訳版のほうを TranslatedTitle に記述するという活用ができる。	【JaLC と DataCite で仕様が異なる】 ・ DataCite 仕様であれば、DIAS メタデータの内容がカバーされる ・ DataCite 3.1 では、titleType として、AlternativeTitle、Subtitle、TranslatedTitle の 3 種類が定義。DIAS のメタデータ内でも AlternativeTitle という項目が定義されており、タイトルに略称が使われていた場合には正式名称、タイトルに正式名称が使われている場合はその略称、といった用途で使われている。	JaLC のジャーナルアフィリエイト等のコンテンツとの兼合いで現状の仕様となっております。 titleType 属性の追加は、現在進めているメタデータ項目の拡張では、改修範囲が大きくなるため困難です。 今後の検討課題とさせていただきます。
4	subject 要素	【subject の活用について】 DataCite の注釈にない、DIAS のメタデータ内に記述された、分類や検索用のキーワード等を、この subject 要素に当てはめることを検討しているが、問題はないか？ また、JaLC では、DOI の検索の際、この subject 要素の利用を想定しているか？	JaLC と DataCite で機能的には等価。しかし、JaLC 側には subject 要素に対する記述がなく、用途が不明。 ・ DataCite 3.1 の subject 要素に関する注記 "Subject, keywords, classification codes, or key phrases describing the resource."	subject については、DataCite の注釈にない使い方では構いません。 また、現状の JaLC では、subject 要素の検索はサポートしておりません。今後の検討課題とさせていただきます。
5	"	サブジェクト<subject>が 1 言語 1 レコードしか入らない	研究データのサブジェクトに<lang=en>で複数データを登録しようとする「サブジェクト記述言語が重複しています。」というエラーが表示される。	現状のエラーチェックが 1 言語 1 レコードとなっております。6/30 に本番環境・実験環境とも複数データに対応できるよう対応しました。
6	creator/ contributor 要素 の type 属性	【質問・提案・指摘】 この属性の値の person と institute がどの要素と対応づけるか不明。必須要素の names 内の値か？それとも任意要素の affiliation の有無か？また、type 属性の値が、person か institute の選択になるのであれば、自由記述型ではなく、スキーマ内で xs:restriction を使って制限したほうが良いのでは？また、DataCite カラムに記載されている 2.2: nameIdentifier が対応するのは、正しくは、この type 属性ではなく id_code 要素のほうでは？	【JaLC 独自属性 (任意指定)】 資料中、type の横に person, institute との記載があるが、research_data_v1.21.xsd では、自由記述で定義されている。	person と institute は、type 属性の属性値です。 (例) <creator sequence=""1"" type=""person""> <creator sequence=""2"" type=""institute""> スキーマ内での制限については、今後のスキーマのバージョンアップの際に考慮したいと思います。また 2.2: nameIdentifier の対応するエレメントは、ご指摘の通り id_code が正しいです。修正します。
7	"creator/ contributor の first_name 要素"	【質問】 DIAS メタデータでは、名称を苗字と名前に分けておらず、last_name, first_name, prefix, suffix の分類が正確にはできない。従って、first_name にフルネームを入れるような運用になるが、問題ないか？ また、データ作成者の情報が、役職名/機関名しかない場合、任意要素の affiliation ではなく必須要素の names>first_name 要素に入力せざるをえないが、これも問題ないか？また、こうした運用をした場合、DataCite とのマッピングにも影響はないか？	【JaLC 独自拡張要素 (必須指定)】 DIAS メタデータでは、データ作成者の情報として、個人名/役職名/機関名などがあり、役職名や機関名が一つだけしか入力されていない例もある。	first_name にフルネームを入れて頂く形で構いません。 また、データ作成者の情報の入力については、役職名/機関名を first_name に入れて頂くのは構いませんが、DataCite へは creatorName としてマッピングされます。
8	"publisher 要素 (publisher_name, location)"	【publisher 要素に記載すべき内容について質問・要望】 publisher には JaLC 正会員としての情報 (我々の場合、DIAS) を記述すれば良いか？それとも、publisher の記載内容は、我々の運用に一任されているのか？ また、publisher_name は lang 属性をもつが、単	【location は JaLC の独自要素 (任意指定)】 ・ publisher_name に lang 属性は定義されているが publisher, publisher_name 共に単一要素のため、日本語・英語の併記はできない。 ・ DataCite でも単一必須要素として定義さ	publisher の記述・運用については、現状では利用機関にお任せしている状況です。ただ、記述内容の明確化に向け、研究データ実験 PJ の中でポリシー検討を提案したいと思います。 また publisher_name に記述する言語ですが、DataCite の DOI を取得する場合には、本要素が DataCite にデボジットされますので、英語表記での記述をお奨めしま

		一要素のため、日本語・英語の併記ができない。 publisher_name を複数記述可の要素に拡張することは可能か？もし現在の仕様のままだと、DataCite でも publisher 要素は単一の必須要素となっているため、日本語の publisher_name を DataCite にそのままマッピングするような運用になると思われる。 なお、JaLC 独自要素として location(値として国名コードをとる)が追加されている。大抵の場合、JPN になると思われるが、どのようなケースを想定して独自拡張したのか、説明を追加して欲しい。	れている (nonemptycontentType 型で定義されているので、空白も不可)。すると、マッピングルールはどうなる？ 資料の備考欄には、「データ所有者 (アーカイブを含む)、研究機関」との記載がある。	す。
9	"	<publisher>が一つしか入れられない		研究データの現在の仕様では 1 つとなっています。(DataCite も 1 つ)
10	"	項目名の表記として、出版者というのは研究データに関しては違和感がある		Publisher (出版者) については、研究データの所有者・所有機関やアーカイブ機関を想定しております。
11	"	Publisher と Creator についてどちらに何を入れたらよいか迷った。例を示して欲しい。		今後のドキュメント整備の中で考えます。
12	contributor_type 属性	【要望】 DataCite の仕様を理解している機関は contributorType の値を適切に入力するかもしれないが、そうでない機関は、JaLC で自由記述となっているため、思い思いの値を入力するケースも考えられる。入力値にバラつきが発生するのを避けるため、JaLC のスキーマにも、DataCite の xs:enumeration を反映して欲しい。	【DataCite の contributorType との関係が不明瞭】 ・ JaLC(research_data_v1.21.xsd) xs:string 型 (自由記述) で定義 ・ DataCite (datacite-contributorType-v3_1.xsd) xs:enumeration の 22 の値が定義。	contributorType の値については、DataCite の DOI を取得する場合は、DataCite で定める 22 の値が必要となります。一方、JaLC の DOI を取得する場合は、特に制限は設けておりません。 以上のことから、DataCite-DOI 取得用スキーマと JaLC-DOI 取得用スキーマとを分けてご提供した方が良く考えております。
13	variation 要素	【variation の活用について】 これは異版コンテンツを記述するための JaLC 独自要素だと思われるが、どういう使い方をする要素なのか (どういう値の入力を想定しているのか) が不明。せつかくの独自要素なので、想定している用途や具体的な事例など、説明を追加したほうが良いのでは？	【JaLC 独自要素 (任意指定)】	ご指摘ありがとうございます。 Variation 要素につきましては、論文等における出版版 / 著者版のような情報を想定したものにします。マニフェスト更新時での説明追加を検討します。
14	related_content の type 属性	備考欄に "DataCite: relationType P27 から DOI、URL のみ適用" とあるが、relationIdentifierType の間違いでは？ また、値が DOI、URL に決まっているのであれば、xs:string 型ではなく、スキーマで制限したほうが良いと思われる。	JaLC では DataCite の relationIdentifierType の一部のみ対応	ご指摘ありがとうございます。relatedIdentifierType が正しいです。修正いたします。 type 属性の属性値を DOI または URL で制限する件については、今後のスキーマバージョンアップの際に検討したいと思います。
15	related_content の relation 属性	DataCite の relationType には非対応ということか？また、自由記述で条件付き必須属性になっているが、この仕様だと、ただ related_content 要素を使って DOI や URL のリストを追加したい場合であっても、relation 属性にひとつひとつの関係性に関するテキストを入力しなければならない。こうした事情は想定しているか？ なお DIAS では、dataCite の relationType のうち、"References" などを使って、URL のリストを追加する、といった運用を考えている。	JaLC では、DataCite の relationType に非対応で、relation 属性は自由記述をとる。しかし、必須指定となっており、実運用を考えると、扱いづらい。	DataCite の relationType は JaLC では relation_list の related_content の relation 属性が該当します。 <relation_list> <related_content type=""URL""relation=""References""> (参照先 URL1<related_content> <related_content type=""URL"" relation=""References""> (参照先 URL2<related_content> </relation_list> のような記載になるかと考えます。"
16	date 要素の type 属性	【要望】 DataCite の仕様を理解している機関は dateType の値を適切に入力するかもしれないが、そうでない機関は JaLC で自由記述となっているため、思い思いの値を入力するケースも考えられる。入力値にバラつきが発生するのを避けるため、JaLC のスキーマにも、DataCite の xs:enumeration を反映して欲しい。	【DataCite の dateType と関係が不明瞭。】 ・ JaLC(research_data_v1.21.xsd) xs:string 型 (自由記述) で定義 ・ DataCite (datacite-dateType-v3.xsd) xs:enumeration の 9 の値が定義。	コメントありがとうございます 他のところでも書きましたが、DataCite-DOI 取得用スキーマと JaLC-DOI 取得用スキーマとを分けてご提供した方が良く考えております。
17	resource_type 要素の type 属性	【要望】 DataCite の仕様を理解している機関は resourceTypeGeneral の値を適切に入力するかもしれないが、そうでない機関は JaLC で自由記述となっているため、思い思いの値を入力するケースも考えられる。入力値にバラつきが発生するのを避けるため、JaLC のスキーマにも、DataCite の xs:enumeration を反映して欲しい。	【DataCite の resourceTypeGeneral と関係が不明瞭。】 ・ JaLC (research_data_v1.21.xsd) xs:string 型 (自由記述) で定義 ・ DataCite (datacite-resourceType-v3.xsd) xs:enumeration の 14 の値が定義。	コメントありがとうございます。 他のところでも書きましたが、DataCite-DOI 取得用スキーマと JaLC-DOI 取得用スキーマとを分けてご提供した方が良く考えております。
18	rights 属性	【質問・相談】 DIAS には、日本語と英語、プロジェクトの利用規約 (複数のデータで共通のルール)、データ提供者による利用規約 (そのデータのみルール) など、データに関する複数の利用規約が存在する。その場合どのように記述したら良いか？日本語・英語の併記について、lang 属性の拡張は一つの案になると考えられる。また、JaLC 的な考え方にすれば、multiple_resolution_priority 要素や、sequence 属性などを追加して、複数の利用規約に優先順位をつけるアプローチなども考えられるが、そうした検討の予定はあるか？	DataCite、JaLC ともに lang 属性が定義されていないため、日本語・英語を区別して併記できない。	rights 要素は、その研究データを参照・利用する人向けにライセンスポリシーを記述する部分と考えておりますが、rights に記載すべき内容に関しては、研究データ実験 PJ 中でのポリシー検討を提案したいと思います。 なお rights 要素 に関しては、メタデータ拡張で複数記述できるようにする予定です。また言語につきましては、rights は複数の要素を自由に DataCite にデフォットできますので、日本語は日本語のまま、英語は英語のまま (UTF-8 で) 記述してあれば、そのまま DataCite ヘドボットされます。
19	<description>の文字数制限について	項番 95 の "description" の文字数ですが、現状で 1000 だが、4000~5000 文字に拡張して欲しい。		研究データのメタデータ拡張対応にて文字数制限を 5000 文字に拡張しました。(2/20 メタデータ拡張決定、6/30 本番環境適用済 JST)
20	description 要素の type 属性	【質問・相談】 DIAS のメタデータ内にある、データの概要・データの処理情報・品質情報の 3 種類のテキスト情報を description 要素に入力することを考えて	【DataCite の descriptionType と関係が不明瞭】 ・ JaLC (research_data_v1.21.xsd) xs:string 型 (自由記述) で定義	DataCite の仕様書によると、該当するものが無い場合は "Other" とするようです。 Other description information that does not fit into an existing category. / Use for any other description type.

		いる。しかし、データ処理情報・品質情報については descriptionType(DataCite) の 5 種類中にも直接対応がとれるものがない。ゆえに、JaLC の descriptionType に品質情報というタイプを追加することは可能か？ また、JaLC からの提案として、DataCite 側の仕様にも、品質情報というタイプを追加するよう提案することは可能か？	・ DataCite (datacite-descriptionType-v3.xsd) xs:enumeration の 5 の値が定義。	
21	signature 要素	【要望】 JaLC 独自要素であるが、資料中に説明がなく、どのような内容を記述すべきか、用途が不明。説明を追加して欲しい。”	【JaLC 独自要素 (任意)】	コメントありがとうございます。 signature 要素はコンテンツの品質保証の署名等に使うことを想定しております。今後の改定の中で用途等の説明を追記していきます。
22	geolocation_point 要素	【要望】 右の備考欄に示す通り同じ値を使っているの で、JaLC の xml サンプルは、DataCite の xml サンプルを引用したものと思われるが、DataCite では値をスペースで区切っている一方、JaLC ではスペースで区切っていない。おそらく、DataCite の仕様から” 31.233” は緯度、” -67.302” は経度 (マイナスなので西経) を表しているかと推測するが、それを理解した上で JaLC のサンプルを作成しているか？位置情報の記述方法については、JaLC の資料内にも、十分な説明が無い。xs:string 定義になっているので、緯度経度の記述順序や、北緯/南緯・東経/西経を+で表現するのかなど、この要素について、あまり検討されていないのでは？もう一度、きちんと検討したほうが良いと思われる。	【JaLC と DataCite で型が違う】 【JaLC 側の位置情報の表記ルールが不明瞭】 ・ JaLC の定義(research_data_v1.2.xsd) xs:string 型 ・ JaLC のサンプル(research_data_v1.2.xml) <geolocation_point>31.233-67.302</geolocation_point> ・ DataCite の定義 (スペース区切りのダブル型) <xs:simpleType name="point"> <xs:restriction base="listOfDoubles"> <xs:minLength value="2"/> <xs:maxLength value="2"/> </xs:restriction> </xs:simpleType> ・ DataCite のサンプル <geolocationPoint>31.233 -67.302</geolocationPoint>	ご指摘ありがとうございます。 ご指摘のとおりと考えますので、サンプルの方は修正します。また仕様書やスキーマの方も修正・説明の追記を行います。
23	geolocation_box 要素	上記の geolocation_point 要素と同様の問題。box では、二つの point を記述するので、その順序なども含めて、もう一度、きちんと検討したほうが良いと思われる。	【JaLC と DataCite で型が違う】 【JaLC 側の位置情報の表記ルールが不明瞭】 ・ JaLC の定義(research_data_v1.2.xsd) xs:string 型 ・ JaLC のサンプル(research_data_v1.2.xml) <geolocation_box>41.090-71.032 42.893-68.211</geolocation_box> ・ DataCite の定義 (スペース区切りのダブル型) <xs:simpleType name="box"> <xs:restriction base="listOfDoubles"> <xs:minLength value="4"/> <xs:maxLength value="4"/> </xs:restriction> </xs:simpleType> ・ DataCite のサンプル <geolocationBox>41.090 -71.032 42.893 -68.211</geolocationBox>	項番 17 と同様に、修正・追記を行います。
24	geolocation_place 要素	lang 属性がなく、繰返しもできないので、日本語・英語を区別して併記できない。従って現在の仕様だと、日本語の地名がそのまま DataCite にマッピングされる運用になってしまう。lang 属性の追加、繰返し可能への拡張を検討してほしい。	・ JaLC、DataCite 共に lang 属性がなく、繰返しが許されない ・ JaLC の日本語の地名はそのまま DataCite にマッピングされる？	DataCite-DOI を取得される場合は、英語表記をお奨めします。 lang 属性の追加は、現在進めているメタデータ項目の拡張では、改修範囲が大きくなるため困難です。今後の検討課題とさせていただきます。
25	multiple_resolution_priority 要素	JaLC 独自要素だが、説明が少ないので、よくわからない。どのような用途を想定しているのか、説明を追加して欲しい。複数の content 要素の優先順位をつけるための要素か？もしそうであれば、sequence 属性との関係は？ また、値が 1~999 の数値をとるのであれば、xs:string ではなく、xs:integer 型などで定義したほうが良いのでは？	・ JaLC 独自要素 (任意)	マルチプルレゾリューションは、全く同一のコンテンツが異なる場所にある場合に使う用途のものになります。 <例> 「機関リポジトリにある文献を NDL 国立国会図書館でも収集・公開しているときに、同じ DOI に対して機関リポジトリの URL と NDL の URL の 2 箇所をリゾルブ先とすることができる。」 研究データでこのようなニーズがどの程度あるか不明ですが、仕組みとして研究データでも使えるようにしています。ただし、マルチプルレゾリューションにできるのは、JaLC-DOI のときだけで、DataCite-DOI はサポートしていません。
26	sequence のケタ数について	8 桁で入力したときと、7 桁で入力したときとでエラー内容が異なる。		本事象は不具合でした。 6/30 に不具合を修正いたしました。
27	著者について	結果から推定すると、筆頭著者の有無、重複確認は str 型で比較データ格納時は int 型に変換して実行される アプリケーション側で意図的に変換しているのではなく、str 型のまま SQL 文を組み立てた後で、DB サーバが SQL 文を解釈する時に int として扱われるのではないか。		ご指摘の通り、重複確認は String 型で比較しているため、「01」と「1」を区別する実装となっております。そのため、それぞれの名が登録されます。

5. DOI 本登録

5.1 DOI 本登録要領

DOI 本登録は、『JaLC における研究データの登録手順書』(附録資料 6) に従って行った。

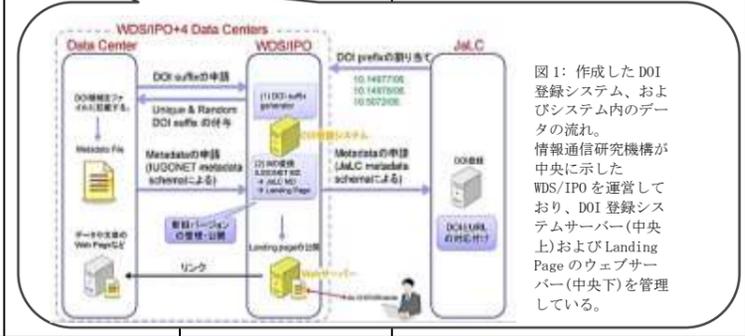
DOI 本登録は、ジャパンリンクセンターシステムの本番環境にログインして、各機関固有の本番用の DOI prefix を使用して行った。ここで使用したサイト ID や、DOI prefix は、本プロジェクト終了後には、ジャパンリンクセンターの正式な会員として使用を継続するものである。

5.2 DOI 本登録計画

各参加機関から提出された「DOI 本登録計画」を以下にとりまとめる。なお、各機関から提出された原本は附録資料 7 として添付する。

		1. 本登録実施方針	2. 本登録対象データ	3. 方法	4. 体制	5. 環境	6. スケジュール	7. 中長期的な DOI 登録方針	8. 本登録した DOI の活用方針
1	NBDC/JST	研究データへの DOI 登録実験プロジェクトに引き続き、NBDC が管理運営している「生命科学系データベースアーカイブ」のデータセットを本登録の対象と考えている。ただし、データ更新等に伴う新 DOI の付与および新旧両方のデータの保管などへの対応が決まっていないため、直ちに本登録に移行することは困難と考える。	生命科学系データベースアーカイブ (https://dbarchive.biosciencedbc.jp/) に収録されているデータセット全件 (約 500 データセット (101 データベース))。	XML 一括入力を利用。RA は JaLC。NBDC 単独で使用できる prefix 1 件の払い出しを希望。	【責任者】 生命科学系データベースアーカイブチーム DOI 担当者: JaLC との連絡、DOI 登録の方式の検討、データ登録、更新、削除、各種報告 【生命科学系データベースアーカイブチーム (3 名)】 メタデータ作成			生命科学系データベースアーカイブ (https://dbarchive.biosciencedbc.jp/) に収録されているデータセット全件 (約 500 データセット (101 データベース)) を対象として DOI 登録を行う。	DOI 登録したデータの二次利用の促進
2	国立極地研究所	極域研究では、環境変化の実態把握それらもたらす様々な変動メカニズム及び将来予測結果等の科学的知見のオープンアクセスが求められている。極域の環境変動システムを明らかにするためには、学際的な研究を通して、多種多様なデータや研究成果を集積し、かつこれらを用いて得られた科学的な知見を社会へ適切に公開していくことが必要である。そのためには品質が管理されたデータセットを公開すると共に、適切に引用されなくてはならない。国立極地研究所では、適切に品質が管理されたデータに対して DOI 付与を行い、データ登録者へのメリットを確保すると共に、データアーカイブ (データセンター) の長期的運用を確保する。	2015 年 9 月末までは ①DOI 登録するデータの分野: 地球科学・環境分野 データ種類: 北極域における実験及び観測データ (気象、水文、大気、雪氷) DOI 登録件数: 10 件程度 2015 年 10 月以降下記の部分の登録を順次行っていく。 北極域における実験及び観測データ (気象、水文、大気、海洋、雪氷、生物、地震、宙空) 南極域における実験及び観測データ (気象、水文、大気、海洋、雪氷、アйсコア、生物、固体地球物理、宙空)	【本登録のために使用するシステムの概要】 国立極地研究所で DOI テスト登録を行った北極域データアーカイブシステム (ADS) を用いる。 <システム概要> ADS-AMS のシステム概要は図 1 のとおりである。 国立極地研究所の北極域データアーカイブ (ADS) は、一元的なメタデータ管理とデータアーカイブ機能を持ったシステムである。全てのメタデータは、ADS メタデータマネージメントシステム (ADS-AMS) により管理されており、ADS-AMS に登録されるメタデータおよび実データは全て研究者の手によって作成される。ADS-AMS ではメタデータ登録時に ADS メタデータスキーマに従って、メタデータのエラーチェックを行い、エラー発生時にはユーザーにフィードバックを行う。ユーザーはフィードバック結果に従い、メタデータの該当箇所を修正し、登録を行わなくてはならない。 ADS-AMS では、国立極地研究所 DOI 付与ポリシーに従い、管理者のみが DOI を付与・登録を行う機能を持つ。ADS における全ての登録データには、メタデータに記載された情報をもとにランディングページが自動的に作成され、このページを通じて詳細なメタデータや実データ、実データのクイックルック、付属ドキュメント等にアクセスできる。 (RA(JaLC, DataCite)の別も含む) この部分をどのようにするのか? ADS に登録されたデータは、JaLC2 システムを経由して DataCite に登録する。なお、本プロジェクトには直接影響しないものの、JaLC 加入後は研究データ以外への DOI 付与も予定しており、JaLC2 及び CrossRef へ登録するデータがある点、念のため記載する。	【研究統括 (進捗管理・報告)】 国立極地研究所・副所長 教授 榎本 浩之 【実施責任者: (DOI 付与に関して取りまとめを行う)】 国立極地研究所・北極観測センター 特任准教授 矢吹裕伯 ・ suffix をつけるためのガイドラインを策定 ・ 研究データへの DOI 付与について、運用ポリシーおよび登録フローを検討する。 ・ DOI 付与する研究データの登録基準を検討する。 ・ DOI 付与に関して取りまとめを行う。 【実施担当者 (DOI 付与に関して検討を行う)】 国立極地研究所・極域データセンター 准教授 岡田 雅樹 国立極地研究所・北極観測センター 特任研究員 杉村 剛 国立極地研究所・北極観測センター 特任研究員 照井 健志 国立極地研究所・宙空圏研究グループ 特任教員 田中 良昌 国立極地研究所・情報図書室 事務職員 南山 泰之 (国立極地研究所では単一のサイト ID を希望する。)	極地研では、今回テストに参加する、ADS-AMS を行い、JaLC 2 への登録を行う。	【本登録実施期間 (～2015 年 9 月末) における実施スケジュール】 6 月下旬: DOI 付与の基準の確認 7 月中: システム改良 (本登録への対応修正) 9 月: DOI 登録完了"	国立極地研究所では、既存のデータ報告書 (Data reports) を発展させ、査読制度を設け、適切な品質管理を行い、品質管理後のデータに関して DOI 付与を行う方針である。当初の計画では北極域の気象水文に限定されるが、中長期的には、北極、南極域における実験及び観測データに対して DOI を付与する方針である。また、学術論文に利用されたデータや他データジャーナルに掲載されたデータについても DOI を付与することも検討する。	登録されたデータの DOI は、科学論文を作成する場合において引用を行う。DOI を付与する事はデータの恒久的なデータのリンクをデジタル的に確保することであり、それにより科学論文の透明性と説明責任を明確にすることができる。また科学論文の公開とデータの公開は第 3 者による研究結果の検証を可能にし、研究方法の向上と妥当性の確認を推進することができる。DOI を付与したデータの引用の積極的な取り組みや Data Citation Index 等の評価基準等のサービスが確立することが必要では有るが、データへの DOI 付与は、データの作成者、および管理するデータセンターの寄与が明確になり、評価基準として用いることができる。将来的には、引用 DOI 数を統計処理することによりデータ作成者やデータセンターの寄与を明確にすることで、研究者およびデータセンターの評価の仕組みを構築する。
3	国立情報学研究所 地球環境情報統合 プログラム (DIAS-P)	本登録は実施しない (永続的な識別子の付与体制が整っていないため)。							
4	国立情報学研究所 学術基盤	国立情報学研究所が運用している IRDB を介した DOI の付与を行う。	実験時のリポジトリ (歴史的資料物写真データ、科学衛星観測データ) から登録を開始する予定。登録件数についても、実験時と同様で各 1000 件オーダーのコンテンツが対象となる。その他、準会員の要望に応じて参加機関数は変化する可能性がある。	実験同様、NII の IRDB を介して登録する。現在想定している上記の 2 つのリポジトリは日本語でメタデータを管理しているため、当面は RA : JaLC の登録のみとなる予定である。	・ NII : IRDB の運用、WEKO における必要な機能の開発 ・ リポジトリ運用機関 (複数): コンテンツ登録"	・ NII : IRDB ・ リポジトリ運用機関: WEKO、DSpace など	・ IRDB の改修は完了 ・ 歴史的資料物写真データ及び科学衛星観測データの DOI 登録開始時期については調整中	・ 準会員向けガイドラインの改訂 (研究データの記載を追加) を年内に予定 ・ 国立情報学研究所がとりまとめる JaLC 準会員 (大学等) の判断による"	特になし
5	産業技術総合研究所	テスト登録の経験と結果を踏まえてぜひ本登録に進みたいが、以下の課題面ある。 ①技術的・ポリシー的な未解決事項等が残り、コンテンツホルダが直接作業するにはまだ障壁が高いように思える。 ②研究所内での長期的な方針 (あるかどうかも含め) 十分議論できていない。 従って、技術的解決や他機関での事例を見て課題を解消しつつ、概ねの方向性を研究所内で合意しながら進めて行きたい。	検討中	未定	現体制を維持しつつ、他機関の進行状況や決定状況などを随時教えてもらいながら、可能な部分を進めていく、という体制が取れるとありがたい。なお、産総研内でのデータベースの運用管理体制が検討中なので、その結果も踏まえた形になる予定	未定	未定	産総研内で適切な登録ポリシー等を定めて統一的に取り組むようにしたい。	

6	<p>情報通信研究機構 他国内 WDC4 機関</p>	<p>・参加機関で保有する観測データに対応して、IUGONET メタデータスキーマでメタデータを作成する。 ・テスト登録の際に開発した DOI 登録システムを利用し、JaLC へ Web API を経由して DOI の登録、および Landing Page の生成・公開を行う。 ・参加機関は、観測データ・IUGONET メタデータなどを管理し、DOI の継続性を保証する努力を行う。</p>	<p>【地磁気変動データ】 登録データの内容： 地上で観測された地磁気変動の時系列データ、およびそのデータを用いて作成した地磁気指数の時系列データ。 現在稼働している地磁気観測所は約 150 ヶ所、観測時間分解能の種類は主に 1 時間、1 分、1 秒の 3 種類。 DOI 登録予定件数： 数件～10 件 (2015 年 9 月末までの実施期間中) 【電離圏観測データ】 登録データの内容： 国内 4 拠点で観測された電離層観測データ。1 拠点につき、イオノグラム画像、手動解析値、自動解析値のデータがある。現在も継続して観測中で、観測時間分解能は基本的に 15 分。 データ管理責任者： 石井守、加藤久雄、山本和憲(情報通信研究機構) 粒度： 観測所ごと・データフォーマットごとに 1 つの DOI を割り当てる。 DOI 登録予定件数： 12 件 (2015 年 9 月末までの実施期間中)</p>	<p>テスト登録期間中に開発した DOI 登録システムをアップデートしながら、登録を行う。DOI 登録システム、およびシステム内のデータの流れを図 1 に示した。詳細は、「DOI テスト登録報告書」に記載している。 DataCite 経由の登録(DataCite が管理する DOI の付与)を予定している。 現在のところ、複数の DOI prefix の払い出しを希望している。WDC の各機関(4 機関)に DOI prefix を一つずつという使い分けを予定している。</p>	<p>【体制】 取りまとめ機関[主なメンバー] 情報通信研究機構・統合データシステム研究開発室 [村山泰啓、木下武也、是津耕司、渡邊堯] データ登録機関および関係機関[主なメンバー] (a) 京都大学大学院理学研究科・地磁気世界資料センター[能勢正仁、家森俊彦] (b) 情報通信研究機構・電離圏宇宙天気世界資料センター[石井守、山本和憲、加藤久雄] (c) 国立極地研究所・オーロラ世界資料センター[門倉昭、田中良昌] (d) 宇宙航空研究開発機構・宇宙科学衛星資料センター[篠原育] (e) 情報・システム研究機構・新領域融合研究センター[小山幸伸] (f) 株式会社情報誌作室[相良毅] 【役割分担】 取りまとめ機関の責任者である村山は、JaLC との連絡や報告を行う。データ登録機関の責任者は、(a)-(d) の 4 機関を代表して能勢がその任に就き、村山および各機関の担当者と密に連絡をとりながら DOI 登録・更新・削除のテスト実務を担当する。 各データ登録機関は、取りまとめ機関である情報通信研究機構が運営する DOI 登録システム(メタデータの変換・Landing Page の生成システム)を通して、各機関が保持するデータの DOI 登録・更新を行う。DOI 登録システムサーバーおよび Landing Page ウェブサーバーのメンテナンスは、主に村山、木下、能勢、小山、相良が協同で行う。 【サイト ID】 サイト ID は情報通信研究機構へ発行済みの 1 つで差支えない。現在のところ、新たに複数のサイト ID を申請する予定はない。</p>	<p>図 1 の中央に示したように、情報通信研究機構が運営する WDS/IPO(World Data System/International Program Office)に DOI 登録システムサーバーと Landing Page ウェブサーバーを設置する。京都大学をはじめとする各データ登録機関は、DOI 登録システムサーバーにアクセスして DOI の登録を行う。</p>	<p>6 月 10 日： 参加機関の実務担当者間でグループミーティング 6 月下旬～8 月上旬： 各機関で DOI 登録の実施 8 月下旬： 参加機関の実務担当者間でグループミーティング、DOI 登録方法やシステムの改善について話し合い 9 月上旬～9 月下旬： 各機関で DOI 登録の実施 9 月末： 最終報告書の作成</p>	<p>プロジェクト期間が終了後も、引き続き DOI 登録を行う。データのアップデートに応じて、既に登録した DOI や Landing Page のアップデートも行い、メンテナンスを継続する。 DOI を登録するデータは、World Data Center として各機関がこれまでに長期にわたって保管・利用してきたものであり、今後もデータベースとして一般に公開していくものであるため、DOI 継続性は必要最小限の努力で保証できると考えている。</p> <p>研究データのトレーサビリティのため、論文引用と同様にデータ引用の研究文化・習慣を啓蒙していこうと考えている。 DOI を用いたデータ引用により、各機関がこれまでに多大な労力を費やしてきたデータベース整備・管理に対して正当な評価が行われ、それに見合った対価が支払われることを期待している。</p>
7	<p>千葉大学附属図書館</p>	<p>本学機関リポジトリに登録済のデータ・データセットへの DOI 付与を行う。</p>	<p>本学機関リポジトリに登録済のデータ・データセットのうち、主に蕨庭標本を登録対象(約 5 万件)を対象と考えているが、作業量的に膨大なため、本実験期間中で全登録は困難であると考えている。</p>	<p>システムとしては本学機関リポジトリを使用予定。 登録は XML により一括。RA は JaLC の予定。"</p>	<p>本学附属図書館員が作業予定。</p>	<p>7 月：作業計画策定 8 月～：作業開始"</p>	<p>今後、研究データの継承は、科学研究の成果の継承と同義となることが予想される。本実験には図書館として参加した。最先端の研究成果に対する DOI 付与はそれぞれの研究グループで行うとしても、長期のデータ保存は組織的な対応が必要となってくるであろうし、研究機関内におけるデータ保存のセーフティネットに図書館はなりうるのではないかと、本実験を通してそのように思察しはじめているところである。その上でどのようなデータ受け入れ体制を敷くことを検討したい。</p>	<p>機関リポジトリを通して DOI を公開する予定。</p>
8	<p>物質・材料研究機構</p>	<p>研究データの登録は、テスト登録の結果明らかとなった課題の解決を前提とした上で行う。また、研究データの登録は、研究成果の公表を補完し、機構として長期的に保存、適切な利用条件のもとに公開することを目的に、NIMS デジタルライブラリー(セルフアーカイブシステム)を利用するものとし、 ・登録者は公開する成果物(論文、データ、補足資料)のメタ情報について、登録時に、可能な範囲で責任を持ってメタ情報を整理し、登録する ・公開される登録情報の有効期間は、予測や設定不可能な場合も含み、原則として利用者の判断、責任で利用する ・登録される成果物は、登録者および機構の研究成果として図書館機能を提供する科学情報室にて管理するものとする という原則のもとに行う。 なお研究データの登録とは、機構でなされた(あるいは機構と共同研究協定の元に行われた)研究成果を対象とし ① 科学情報室が管理する機構のサーバで運用する、 ② 利用者すなわち機構で研究に携わる職員が、データを登録する。必要に応じて科学情報室が代理登録をするなどの支援を行う、 ③ 利用者は公開・非公開のアクセス条件の設定を行う、 ④ セルフアーカイブシステムに登録された論文・データ・関連資料等は、NIMS デジタルライブラリー内で相互リンクを張り、情報の関係性を示すことを基本的な運用原則とする。</p>	<p>未定(機構内で募集予定)</p>	<p>登録先の RA は JaLC のみとする。複数の DOI prefix は使用しない予定である。</p>	<p>・図書館機能を持つ科学情報室にて運用を行う ・将来 DOI の付与・管理を組織で分散して行う可能性も残し、複数のサイト ID を希望する</p>	<p>未定(テスト登録報告書で挙げた課題の解決を優先) 未定(テスト登録報告書で挙げた課題の解決を優先)</p>	<p>未定</p>	<p>公的研究機関として、政府のオープンサイエンス基本方針等を念頭に、研究成果の保存・管理・公開に関し、研究の現場で実践可能なスキームを構築し、運用することを目指す。この過程で DOI の有効性を研究環境に還元する。</p>
9	<p>理化学研究所 脳科学総合研究センター 神経情報基盤センター</p>	<p>テスト登録の経験をもとに JaLC Web 登録画面よりデータ登録を行う。 テスト登録時には検証できなかった、共同研究契約を結んでいる研究機関(プラットフォーム(以下 PF)) http://www.neuroinf.jp/modules/program/index.php?content_id=2 における DOI 登録ポリシーに基づいた対象データの選定、それに基づく当センターと PF 間における登録すべきデータ、メタデータを吟味し最終的に当センターの査読のうえでデータ登録を行う。</p>	<p>PF それぞれのデータベース上に登録されている研究データを対象とする。粒度は原則的にそれぞれのデータベースに登録されているデータの粒度と一致させるものとする。 対象件数は各 PF のデータベース内の研究データのうち 3 件程度ずつを対象とする。13 の全 PF および当センター内部のデータを合わせて 45 件程度を予定している。これによりテスト環境では実施しなかった XoonNips 以外のプラットフォームも含まれることになる。</p>	<p>・ RA としては JaLC を使用 ・ JaLC Web を用いた手入力により登録"</p>	<p>当センターが研究データへの DOI 登録全般について、9 月までの本登録の方法を企画し、多機関の協力作業となる登録について、連絡方法、作業方法を設定する。最終的な登録内容についても責任をもつ。 登録データ選定：各 PF に依頼する。選定コンテンツが登録に適するものかどうかは当センターの設定する基準にて判定する。 登録：当センターがメタデータ記入ファイルの本番登録システムとは別に作成し、PF に提供する。まず PF にて記入をし、それを当センターに送付してもらって不適切な記載について修正を求める。それらを繰り返し、問題が解決した後、signer として選定した人の承認を受ける。登録システムへの入力作業は当センターが実施し、間違いがないかどうかのダブルチェックを行う。</p>	<p>・公開データベース： XoonNips、WordPress、FileMaker、MySQL など ・データの抽出： ブラウザによるリストアップ、Excel による一覧作成 ・ PF、当センター間での Excel 受け渡し環境 ・ JaLC Web フォームへのブラウザによる手作業入力"</p>	<p>準備ができれば以下の手順で作業を行う ・登録ポリシーに基づいた研究データをもつ PF から本登録参加を募集 ・登録データとそのメタデータ情報を Excel によりやりとりし査読を行う ・本登録開始 ・本登録まとめ"</p>	<p>今回 9 月までに実施する本登録の結果をふまえて、大量登録に適した方法を再検討したい。現状では DOI 登録を Web からの手入力によって行っているが、本格的な運用のためには査読の問題をクリアしたうえで、半自動化、自動化のためのツール開発が必要と考えている。 DOI の付与により、コンテンツ利用の場合の引用方法を明確にする。このことによりコンテンツの知財としての尊重が改善すると期待する。またこのことによりコンテンツ提供の促進効果も期待する。</p>



5.3 DOI 本登録計画の検討・承認

ミーティング（第5回）（2015年6月3日（水）開催）において、DOI 本登録計画の検討を行った。

その結果、すべての本登録計画が承認され、DOI 本登録を開始することとなった。

5.4 DOI 本登録実施結果

プロジェクト実施期間内では、次のとおり2機関がDOIの本登録を行った。

機関	DOI	RA	タイトル
国立研究開発法人 情報通信研究機構	10.17591/55838dbd6c0ad	DataCite	Mesospheric wind velocity data (30min. mean) observed with MF radar at Poker Flat, Alaska
国立極地研究所	10.17592/001.2015093001	DataCite	The GRENE-TEA Model Intercomparison Project (GTMIP) stage 1 forcing dataset

その他の参加機関については、プロジェクト実施機関内に実際にDOIの本登録をするまでには至らなかった。DOI 本登録の際には、組織的なオーソライズの必要性があることや、大量データを取り扱うための独自システムの整備をする必要がある。そのために時間を要していることが主要な理由であった。

6. 総括

各参加機関からの「研究データへの DOI 登録実験プロジェクト最終報告書」を以下に示す。

6.1 国立研究開発法人 科学技術振興機構

「研究データへの DOI 登録実験プロジェクト」報告書
科学技術振興機構 バイオサイエンスデータベースセンターの取り組み

2015 年 9 月 14 日

NBDC/JST

1. テスト実施方針

NBDC がサービスを提供している「生命科学系データベースアーカイブ」(<https://dbarchive.biosciencedbc.jp/>) に収録されているデータセットに対して DOI を付与するテストを行う。JaLC の DOI 登録インターフェースを使用し、生命科学系データベースアーカイブのメタデータが DOI 登録にもそのまま活用できるか、不足の項目があるかなどを確認し、本格運用開始に備えることを目的とする。

2. 生命科学系データベースアーカイブとは

データセットを長期間安定に維持保管し、データ説明（メタデータ）を統一して検索を容易にすると共に、利用許諾条件などの明示を行うことで、多くの人が容易にデータへアクセスしダウンロードを行えるようにするサービス

3. テストの方法

3. 1. 登録対象データと登録対象件数

生命科学系データベースアーカイブに収録されているデータセット全件（約 500 データセット（102 データベース））

登録対象件数（登録実験済の数）：74 データセット（17 データベース）

3. 2 テスト登録方法

登録のインターフェースは、Web 入力、XML 一括入力の両方を使用した。

RA は、JaLC を選択

4. 結果

生命科学系データベースアーカイブのメタデータが、新規作成、修正することなく、ほ

ぼそのままの形で JaLC の研究データ登録に流用できることを確認した (表)。

表. メタデータ項目の比較と値の一例

メタデータ項目 (主なもの)		JaLCメタデータの値
生命科学系データベースアーカイブ	JaLCのメタデータ	
データ名	title	RGP gmap98 - 遺伝子型データ
データURL	uri	http://dbarchive.biosciencedbc.jp/jp/rgp-gmap98/data-4.html
作成者氏名	creator (ja)	春島嘉草、倉田のり、長村古晃、矢野昌裕、佐々木卓治*
作成者英名	creator (en)	Yoshiaki Harushima, Nori Kurata, Yoshiaki Nagamura, Masahiro Yano, Takuji Sasaki
作成者所属	affiliation	社団法人農林水産先端技術研究所、農業生物資源研究所、社団法人農林水産先端技術研究所、農業生物資源研究所、農業生物資源研究所
利用許諾	rights_url	http://dbarchive.biosciencedbc.jp/jp/rgp-gmap98/lic.html
運用開始年月日	publication_date	1998/1/17
	DOI	10.14977/01.rgp_gmap98_genotype_data
	publisher	国立研究開発法人 科学技術振興機構 Japan Science and Technology Agency (JST)

ランディングページは、生命科学系データベースアーカイブのデータセットの「説明ページ」が適当であることを確認した (図)。

RGP gmap98 データベース名

遺伝子型データ データセット名

追記予定

このデータベースについて

- データベースの説明
- ダウンロード
- 利用許諾
- このデータベースの更新履歴

データ説明

データ名	遺伝子型データ DOI: 10.14977/01.rgp_gmap98_genotype_data
データ内容の説明	各染色体の全ての分離遺伝子 (186個体) についての遺伝子型データ
データファイル	データファイル名: rgp_gmap98_genotype_data.zip データのURL: http://dbarchive.biosciencedbc.jp/archive/rgp-gmap98/LATEST/rgp_gmap98_genotype_data.zip ファイルサイズ: 48.4 KB
簡易検索URL	http://logodb.biosciencedbc.jp/logodb/view/rgp_gmap98_genotype_data
データ取得方法	NipponbareとKasalathの掛け合わせから得た186のF2個体を用いて連鎖解析を行った結果。
解析方法	MAPMAKER/EXP 3.0を使用して、連鎖解析を行いました。
データ件数	2,277 件

データ詳細

項目名	項目の説明
Chrom. No.	染色体番号
Locus	遺伝子座名 (マーカー名)
Genetic code	MAPMAKERによる遺伝コード

DOI: [10.14977/01.rgp_gmap98_genotype_data](https://doi.org/10.14977/01.rgp_gmap98_genotype_data)

ランディングページのURL (データURL): <http://dbarchive.biosciencedbc.jp/jp/rgp-gmap98/data-4.html>

図. ランディングページの例

5. 課題と検討事項

- Suffix のデザインをどうするか。人間が可読なもの (データセットの名称を含む Suffix) にするか否か (シリアル番号) について検討中。そのメリット、デメリットを再評価し、デザインの方針を確定したい。
- JaLC のメタデータ項目の作成者および寄与者のそれぞれにどの情報を入力すべきか結論が出ていない。他機関の実施例も参考にしたいと考えている。JaLC のガイドラインでも方針、事例を示してほしい。
- データ更新時の新 DOI の付与や新旧データの保管が必要になった場合へ対応が決まっていない。

6.2 大学共同利用機関法人 情報・システム研究機構 国立極地研究所

研究データへの DOI 登録実験プロジェクト最終報告書

大学共同利用機関法人 情報・システム研究機構 国立極地研究所 矢吹裕伯

1. 参加の背景と目的

国立極地研究所北極域データアーカイブ(ADS)では、現在様々な北極域に関連する観測データを収集・公開している。我々が求めるデータアーカイブは、データ取得者が公開を目指して各自でデータを登録して、それを検索しやすく利用しやすいものである。しかし、データ取得者による積極的なデータ登録が増加しないことが現状である。それを打破するためには、データ登録者へのインセンティブの提供が必要である。よって、国立極地研究所北極域データアーカイブ(ADS)では DOI を付与することで、データ提供者へのインセンティブ提供することを目的として本プロジェクトへの参加を行った。

本プロジェクトでは、現在公開している、データアーカイブに登録されたメタデータが DOI を付与するために必要な情報に対応するかの検討を行う。また長期的な運用を図る上での各データに付与する suffix をつけるためのガイドラインを策定する。これらの作業を行うと共に、JaLC への登録システムの構築を行う。また今回 DataCite メタデータが扱える時間及び空間的な粒度を検討する。

さらに、国立極地研究所のデータマネジメント委員会、図書委員会との連携を通して、DOI 付与のための体制の検討を行う。

2. 実施方針

- ・ 北極地域の定点観測で取得されている観測・モデルデータを対象とする。
- ・ ADS で公開する観測データに対して、ADS メタデータスキーマを用いて DOI 登録用のメタデータを作成する。
- ・ ADS では JaLC への登録のために WebAPI を用いての DOI 登録システムを構築すると共に、DOI に対応したランディングページの作成・公開を行う。

3. 結果

3. 1 DOI 付与システムの構築

構築した DOI 登録システム、及び ADS システム内でのデータの流れを図 1 に示す。

構築したシステムの登録までの流れ

1. データマネージャーが、DOI を付与する ADS メタデータを指定する
2. システムは、ADS メタデータより、DataCite のメタデータへ XMLformat 変換を行い登録用の XML メタデータの自動作成を行う
3. ADS のメタデータの登録 ID をもとに自動的に Suffix を付与し、実登録用のメ

タデータを作成する。

4. データマネージャーは確認後、WebAPI を等して JaLC への登録を行う。
5. DOIが付与されたデータはADSのシステム上でランディングページが表示される。(図2)

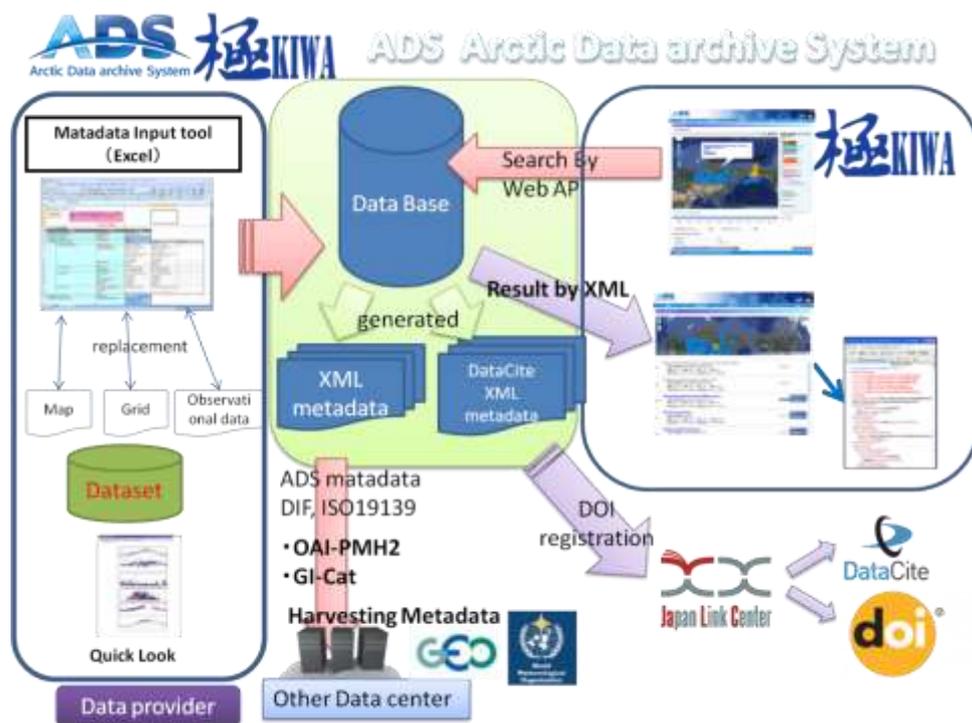


図1 : ADS システム内でのデータフロー。ADS では全てデータはメタデータベースによって管理されており、DOI を付与するメタデータを選択するときのみ、データマネージャーが関与する。



図 2 ; ADS で公開されるランディングページ

3. 2 DOI 付与方針

メタデータ： ADS の XML 形式メタデータを用いて JaLC 形式メタデータに変換を行った。ADS メタデータスキーマおよび JaLC メタデータスキーマのマッピングを行い、自動変換機能を実装した。

DOI サフィックス： 国立極地研究所では今後の DOI 付与を考慮して”XXX.YYYYYYYYYY” とすることとした。”.” 以前の”XXX” は 3 ケタ極地研の部局で使用するものとして YYYYYYYYYY の 10 ケタを任意の文字列で作成することとした。

ADS ではシステム上で用いられている、メタデータ ID を流用することとした。

データポリシー： 国立極地研究所では DOI 付与ポリシーを検討中であるが、DOI 付与は研究所の責務によるところが大きいので、品質管理済みのデータに対して DOI 付与を行うということで、共通理解を得ている。

4. 今後の方針

国立極地研では、データに対する DOI 付与が可能になることを前提として、データジャーナルの発刊を検討している。国立極地研究所の DOI 付与ポリシーにより品質管理を行ったデータについてのみ DOI 付与を行う予定である。この品質管理プロセスをデータセンターで行うには、多大な負担を伴う。その負担を軽くするために、国立極地研究所で発行している”Data Reports”を発展させて、査読プロセスを入れたデータジャーナル

の発刊を予定している。当初は北極域を限定としたデータジャーナルとする。
一方、様々なデータジャーナルが存在するが、現在のデータジャーナルでは投稿するときに DOI 付与が条件となっているデータジャーナルも多く、そのようなデータに対しても DOI を付与する予定である。

5. 問題点及び今後への期待

現在学術論文ではサイテーションインデックスという、論文を評価する仕組みがある。一方研究データに対しては、このような評価の仕組みがない。長期的に、研究データの DOI 付与を推進するためには、データサイテーションインデックス等のサービスの拡充は必要であり、今後の JaLC のサービス展開が必要である。また我々 DOI 付与側としても、研究データへの DOI 付与を推進し、データ DOI という文化をサイエンスコミュニティーに根付かせる必要があり、今後とも JaLC への積極的協働が必要である。

6. 本登録

本登録を 2015 年 9 月 30 日に行い、日本で 2 番目のデータ DOI を付与した。

登録 DOI : doi:10.17592/001.2015093001

本データは、**Earth System Science Data** に投稿された、下記のデータ論文に付随するデータに対して付与されたものである。

”Sueyoshi, T., Saito, K., Miyazaki, S., Mori, J., Ise, T., Arakida, H., Suzuki, R., Sato, A., Iijima, Y., Yabuki, H., Ikawa, H., Ohta, T., Kotani, A., Hajima, T., Sato, H., Yamazaki, T., and Sugimoto, A.: The GRENE-TEA Model Intercomparison Project (GTMIP) stage 1 forcing dataset, *Earth Syst. Sci. Data Discuss.*, 8, 703-736, doi:10.5194/essdd-8-703-2015, 2015.”

JaLC への本申請と同時並行的に投稿を行っていたデータ論文に対しての DOI 付与になる。現在本論文は、**Discussion Paper** の段階であり最終稿の段階で、上記付与したデータ DOI が引用される予定である。

[Home](#)
[Applications](#)
[About Us](#)
[FAQ](#)
[Links](#)

[Login](#)

Data Summary

The GRENE-TEA Model Intercomparison Project (GTMP) stage 1 forcing dataset

← Back
View Meta Data Record
Download
Document
ADS Home

Select Version: 1.00

Summary

Here, the authors describe the construction of a forcing dataset for Land Surface Models (including both physical and biogeochemical models, LSMs) with eight meteorological variables for the 33 year period from 1978 to 2013. The dataset is intended for use in a model intercomparison (MIP) study, called GTMP, which is a part of the Japanese-funded Arctic Climate Change research project. In order to prepare a set of site-fitted forcing data for LSMs with realistic yet continuous entries (i.e. without missing data), four observational sites across the pan-Arctic region (Fairbanks, Tiksi, Yaktutsk, and Kevic) were selected to construct a blended dataset using both global reanalysis and observational data. Marked improvements were found in the diurnal cycles of surface air temperature and humidity, wind speed, and precipitation. The datasets and participation in GTMP are open to the scientific community (<https://ads.nipr.ac.jp/gtmp/gtmp.html>). This data includes four data sets: Salin.LD data: 34-year, reanalysis-based (ERA-Interim): 95-min data; L1 data: L1 data merged with the GRENE-TEA site observations; Extended data: 20-year detrended driving data for spin up based on L1 (1990-1999)/PAR/LAI data: MODIS based monthly data at GRENE-TEA site (2007-2011, monthly)

Dataset citation

Tetsuo Sueyoshi, Kazuyuki Saito, Shin Miyazaki, Junko Mori, Takeshi Ise, Hazuki Arakida, Rikie Suzuki, Atsushi Sato, Yoshihiro Iijima, Hiromori Yabuki, Hiroki Iwawa, Takeshi Ohta, Ayumi Kotani, Tomochiro Hajima, Hisashi Sato, Takeshi Yamazaki, Atsuko Sugimoto. 2014. The GRENE-TEA Model Intercomparison Project (GTMP) stage 1 forcing dataset. 1.00. Arctic Data archive System (ADS), Japan. [doi: 10.17592/001.2015093001](https://doi.org/10.17592/001.2015093001)

Spatial coverage

N: 71.59° S: 62.25° E: 129.62° W: -147.49°

Temporal coverage

1978-08-31 - 2014-01-02

License

CC BY 2.1 JP

This data is licensed under a [Creative Commons Attribution 2.1 JP License](https://creativecommons.org/licenses/by/2.1/jp/)

DATA CONTRIBUTOR & GROUP

- Tetsuo Sueyoshi (NIPR)
- Kazuyuki Saito (JAMSTEC)
- Shin Miyazaki (NIPR)
- Junko Mori (NIPR)
- Takeshi Ise (I Kevic)
- Hazuki Arakida (RIKEN)
- Rikie Suzuki (JAMSTEC)
- Atsushi Sato (NIED)
- Yoshihiro Iijima (JAMSTEC)
- Hiromori Yabuki (JAMSTEC)
- Hiroki Iwawa (NAE0)
- Takeshi Ohta (Nagoya Univ.)
- Ayumi Kotani (Nagoya Univ.)
- Tomochiro Hajima (JAMSTEC)
- Hisashi Sato (JAMSTEC)
- Takeshi Yamazaki (I Takhu)
- Atsuko Sugimoto (I Hokkaido)

PARAMETERS

- Surface Winds
- Air Temperature
- Humidity
- Shortwave Radiation
- Longwave Radiation
- Precipitation Amount
- Photosynthetically Active Radiation
- Vegetation Index

← Back

Copyright 2012, National Institute of Polar Research. All right reserved. [Data Policy Privacy Policy](#)

図 3 : ADS で付与した doi:10.17592/001.2015093001 のランディングページ。

146

6.3 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所

研究データへの DOI 登録実験プロジェクト最終報告書

国立情報学研究所

目的

国立情報学研究所では、学術機関において運用されている機関リポジトリやサブジェクトリポジトリに対する支援事業を行っている。各リポジトリに登録されたコンテンツのメタデータは、OAI-PMH と呼ばれるプロトコルを活用して、NII が運用する IRDB と呼ばれるデータベースに集約されると共に、CiNii などの論文検索エンジンに提供される。

現在のところ、IRDB のコンテンツ種別としては、学術雑誌論文や紀要論文などの文献が主である。データ・データセットに分類されるものは全体の 5% に満たないが、オープンサイエンスに向けての動きが加速化する中で、研究データを公開するプラットフォームとしての機関リポジトリの重要性は、より増してくるものと考えられる。各機関で分散型に運用されているリポジトリに登録された研究データの再利用を促進する上では、効率的に DOI を付与するフレームワークの構築が不可欠となる。

一般的な JaLC の利用では、研究データを管理するデータベースなどから個別に JaLC のインターフェースにアクセスし、DOI の登録を実施する。これに対してリポジトリでは、メタデータが集約される IRDB を活用することで効率的な DOI の登録が実現できる。そこで本実験では、研究データに対して IRDB を介した DOI の付与機能ならびにプロセスを実現することを目的とした。

方法

IRDB を介した DOI 付与のためのシステム構成を図 1 に示す。実験では、リポジトリシステムとして、国立情報学研究所が開発する WEKO を対象とし、DOI の自動付与機能を実装した。

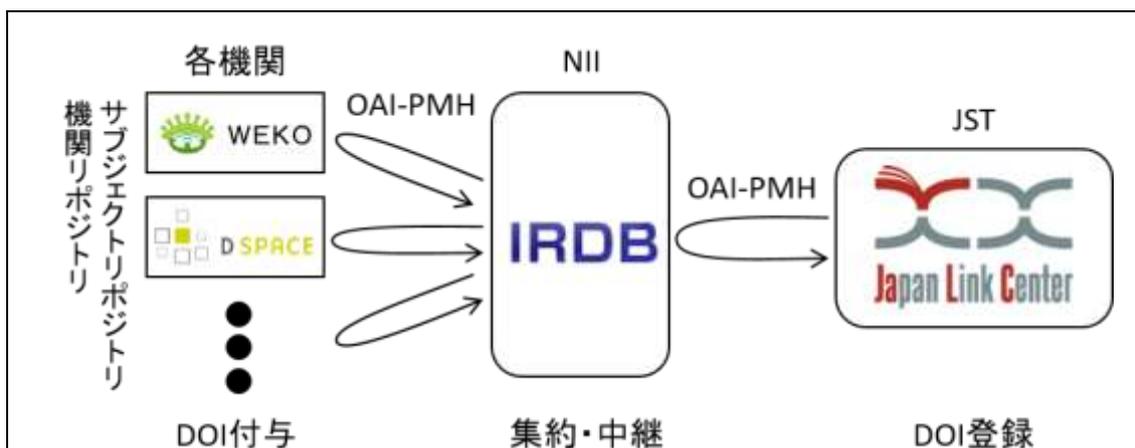


図 1 IRDB を介した DOI 付与の全体構成

各機関のリポジトリ内で付与された DOI は、国立情報学研究所が規定する junii2 メタデータフォーマットの形式で、OAI-PMH プロトコルにより IRDB に集約される。IRDB においてメタデータ形式のエラーチェックをクリアしたコンテンツに付与された DOI のみが、JaLC に提供される。実験で対象としたリポジトリとデータの内容は、表 1 の通りである。

表 1 IRDB を介した DOI 付与実験対象データ

データ管理機関	学術資源リポジトリ協議会	金沢大学 通信情報工学研究室
リポジトリソフト	WEKO	WEKO
件数	1000 件程度	1000 件程度
データ種別	歴史的資料物写真データ	科学衛星観測データ

結果および考察

JaLC における研究データメタデータ仕様と junii2 のマッピング

IRDB には junii2 として規定したスキーマに従ってメタデータが集約される。DOI の登録を完了するためには、junii2 から JaLC のメタデータ仕様に情報をマッピングする必要がある。メタデータの意味の対応だけでなく、言語属性や繰り返し属性の有無、文字数の制約条件などについても矛盾が生じないように精査し、必要に応じて JaLC における研究データメタデータ仕様の変更を行った。

IRDB ならびに WEKO の機能開発

IRDB では、ジャーナルアーティクル型での DOI 登録機能をすでに実装していたために、今回の拡張では、前項で決定したマッピング規則に従ってエラーチェック機能を実装することが主とした開発内容であった。WEKO についてもジャーナルアーティクル型

における仕様を踏襲し、junii2の「Data or Dataset」として登録されたコンテンツに対して、JaLCかDataCiteのRAを選択してDOIを付与する機能を実装した。

連結実験

表1にあげた2つのリポジトリの実コンテンツを用いた実験サイトを構築し、図1の一連のフローが実現できることを確認した。不備のあるデータを投入することで、WEKOやIRDBにおけるエラー処理も適切に実行できていることも確認した。

準会員申請時の申請書の記載法の見直し

本実験プロジェクトの過程で、研究データの管理体制が変更された場合にも柔軟に対応する方策として、1つの機関に複数のPrefixを与える運用について議論が進められた。この運用に対応するために、国立情報学研究所が準会員に対して複数のPrefixを発行する条件について精査を行い、申請書の記載法の見直しを行った。

研究データへの DOI 登録実験プロジェクト最終報告書

2015年9月14日

国立情報学研究所・地球環境情報統融合プログラム (DIAS-P)

1. プロジェクトの目的

DIAS-P において公開するデータセット約 200 個の中から典型的なデータセット数個を選び、識別子を付与する規則や粒度などについて方針を決定するとともに、ポリシー策定やワークフローの整備などの観点から、既存の DIAS システムも DOI に対応できるよう改善する。

2. JaLC へのテスト登録

データセットとしては、単独のデータセットであること、データの階層構造のテストを行うこと、日本語・英語の 2 言語でメタデータを持つことなどの点から、以下のデータセットを対象とした。そして、JaLC のテスト用プレフィックスを用いて、ウェブおよび API 経由で JaLC に対するテスト登録を行った。ただし、DIAS-P はプロジェクトとしては今年度末で終了することから、データの永続性が保証できないという理由で本登録を見送った。

AGURAM 全球都市域マップ地上情報データベース	空間データ	1	東京大学 地球観測データ統融合連携機構	単独のデータセットとしてのテストケース
AWCI 流域データセット (バングラデシュ・メグナ流域)	時系列データ	7	DIAS Office	地域的な階層構造を持つデータセットのテストケース

項目ごとの結果は、以下の通りである。

- メタデータ：既存の XML 形式メタデータを JaLC 形式メタデータに変換した。現在は人手により作業を行っているが、将来的には XSLT を用いた変換機能を実装する予定である。
- DOI サフィックス：DIAS-P で既に定めたデータセット番号を流用し、サフィックスは DIAS.<データセット番号>という形式とした。ただし階層構造を持つデータセットでは、DIAS.<データセット番号>.<枝番>のように独自の枝番も付与できることとした。
- ランディングページ：既存の DIAS 俯瞰検索システムのドキュメントページを用いた実験を行った。ただし DOI の正式導入前のため、ランディングページ上に DOI を表示する等の機能開発は行わなかった。

- データポリシー：既存の DIAS データポリシーを再検討し、DOI 時代のデータ公開・共有の実態や要望を反映したものになるよう、精査と改善を進めているところである。

3. コミュニティの形成

JaLC 実験プロジェクトに対する DIAS の意見を集約するとともに、意見をまとめるための基礎知識を共有するために、「DIAS オープンサイエンス勉強会」を結成し、これまでに 10 回の勉強会を開催してきた。この勉強会では、JaLC 実験プロジェクトの進捗報告にとどまらず、世界のデータリポジトリやデータサイテーションに関する動向調査や、DOI の付与規則とメタデータ形式、バージョン管理などのテーマに関して、発表と議論を行ってきた。こうした活動は、勉強会メンバーの間でオープンサイエンスに関する知識と意識を高めるために有効であったと考えている。

こうした活動を継続するためにも、研究データを取り巻く諸問題とその解決策、あるいはオープンサイエンスを対象とした議論を深めることのできるコミュニティは、JaLC 実験プロジェクトの終了後も何らかの形で維持していくべきと考える。研究データへの DOI 付与は世界中でまさに展開が加速しているところであるため、今回のガイドラインは最終版ではなく、今後も世界の動向に合わせてシステムや仕組みを継続的に更新していく必要がある。こうした世界的な動向に、日本は残念ながら遅れてしまっているが、このまま遅れていていいわけではない。研究データを尊重する文化の担い手となるコミュニティを広げていく努力が必要ではないかと考える。

4. 今後の課題

データセットの変更あるいは取り下げに関するワークフローはまだ検討が進んでいない。現状の DIAS メタデータ登録システムと、システム内のワークフローを考慮しながら統合を進めていく必要がある。また JaLC 側の多言語対応機能にはまだ不十分な部分も残っているが、多言語への対応は DataCite にはない JaLC 独自のニーズとして、引き続き改良に取り組む必要があると考える。ランディングページについては既存のウェブページを用いたが、今後はここをサイトへの入口とした付加価値サービス、例えば関連データの発見機能や、データの切り出し機能などを展開していくことが重要な課題である。

テスト最終報告書

2015年9月16日

国立研究開発法人 産業技術総合研究所

小島 功/渡邊 宏/山下雄一郎/的野 晃整/馬場 哲也

概要：産業技術総合研究所（産総研）が整備・公開している研究情報データベースについて、doiを与える実験を行うことで、登録上のノウハウなどの知見を取得すると同時に、システム上の課題や制度上の課題など本格運用に対して解決が必要な課題を明らかにする。

内容：以下の2つの公開データベースについて doi を与え、いくつかの検討点について試験・検証をおこなった。

- 有機化合物のスペクトルデータベース <http://sdfs.db.aist.go.jp/>
- 分散型熱物性データベース <http://tpds.db.aist.go.jp/>

まず、doiの suffix の与え方について試験的にポリシーを決めて登録し問題がないか検証した。次に、DOI を与える単位を変えた時にデータに到達できるかどうかと、Landing Page の関連するシステムの改修の必要点について確認した。

さらに、産総研内で試験整備中のオープンデータカタログ CKAN についてそのメタデータ表示ページを Landing Page として利用可能かどうか検討した。

結果：

1. **Suffix の与え方：**産総研のデータベースの分散管理の現状から、ばらばらに登録されることが想定されるので、名前の衝突が起こらないようなポリシーを検討した。

- Prefix/(サブ組織名)-(DB名)-(DB担当者が自由に決める識別子)

DB名で空間を分ける。

例：<http://doi.org/10.14977/05.tdfs-23732>

識別子以降は担当者が分野やDBの特性に応じて決める。

例：05.sdfs-4544-hr20xxx (4544はSDFS内での固有の識別子として著名)

産総研の中に別ドメインのサブ組織がある(gsj/nmij)ので、その記載を担保する

例：<http://doi.org/10.14977/05.gsj-aster-xxxx>

結果：重複登録などの問題は（数が少なかったこともあり）起きなかった。

2. **doiを付与する単位：**細かい方が引用には便利だが、Landing Pageの準備の手間や検索の結果としてデータが表示される場合など、システムの挙動との対応が必要なため、SDFSにおいて1) サイト全体でdoiを一個与える場合、2) 化合物単位、3) スペクトル単位などの単位で試験を行った。

結果：データに直接到達できる cgi がない場合は、URL の記載が難しく困難な点が判明

特に、複数のフレームを使って1データを表示する状況を一つの URL で表現できない。

3. **Landing Page の自動生成**：現行のシステム（熱物性データベース）を改修して、データにアクセスすると Landing Page を自動生成するプログラムを作り、これからデータ本体にアクセスできるようにした。

結果：1 物質 1 URL でアクセス可能であったが、元のシステムがポップアップによるデータ表示のため、複数表示すると Landing Page との対応が不明確になる点が判明

4. **CKAN のページを Landing Page に使う**：オープンデータのデファクトカタログなために適用性が高い。

結果：CKAN 自体が未公開・テスト中のために残念ながら十分な試験は未実施。

まとめ：

1. JaLC のシステムの利用など、doi の命名・登録側については概ね問題がないと考えられる。
 - 登録システム(Web/XML)
 - Suffix の決め方などのポリシー
2. 既存のデータベースシステムの改修が必要な場合が多い。
 - 特に、フレームやポップアップなどを利用した既存の表示機能
3. 業務フローと対応付ける必要がある
 - 登録管理・確認などの公開と関連した業務。

議論：

1. doi を付与する周辺での様々なシステム改修や業務システムの構築が必要：これらを実現するソフトウェアを一時的に構築・運用することはできるが、長期の維持管理・発展の点で問題なので、クラウド上での一括したサービスとして提供されるか、利用者コミュニティが確実に存在するオープンソースなどで提供されるとありがたい。
2. データジャーナルとして上記の業務も含めたサービスが提供されれば、これら的一切がアウトソースでき、利点が大きいものとして望まれる。
3. JaLC のサービスとしての魅力が求められる。JaLC に登録することで研究者が享受できるメリットを訴求する必要がある。特に、高度な検索や便利なデータ管理と言った、doi に基づくデータの流通と整備を促進するような機能を持つサービスの展開が期待される。

正式登録：

基本的に、システム環境や運用体制など、所内の環境・体制づくりが重要であるため、上記課題を検討しつつ所内の環境・体制づくりを進め、正式登録を検討していきたい。一方で、これらをすべてアウトソースするサービスの登場も期待される。

以上

最終報告書

2015年9月14日

(情報通信研究機構ほか国内WDC4機関)

1. 実施方針

- 参加機関で保有する観測データに対応して、IUGONETメタデータスキーマでメタデータを作成する。
- 開発したDOI登録システムを利用し、JaLCへWeb APIを経由してDOIの登録、およびLanding Pageの生成・公開を行う。
- 参加機関は、観測データ・IUGONETメタデータなどを管理し、DOIの継続性を保証する努力を行う。

2. 得られた結果

2.1. DOI登録システム

作成したDOI登録システム、およびシステム内のデータの流れを図1に示した。システムの中核サーバーは、情報通信研究機構に設置されており、図1では中央上の黄色で示している。また、Landing Pageをホストするためのウェブサーバーも情報通信研究機構に設置されており、中央下の黄色で示している。時間は上から下に進行する。

1. JaLCからDOI prefixが割り当てられる。
2. 各データセンターは、DOI suffixの作成をDOI登録システムサーバーに対してリクエストする。
3. DOI登録システムサーバーは、ランダムな文字列を生成し、これを「Unique & Random DOI suffix」としてデータセンターに通知する。
4. データセンターは受け取ったDOI suffixを、登録しようとしているデータのメタデータファイルに埋め込む。このメタデータは、IUGONETメタデータスキーマで記述されている。
5. DOI suffixを埋め込んだメタデータを、DOI登録システムサーバーに対して送信する。
6. DOI登録システムサーバーは、受け取ったメタデータの情報を元にして、(1) JaLCメタデータスキーマで記述したメタデータ、(2) Landing Pageを作成する。
7. DOI登録システムサーバーは、6. (1)で作成したメタデータをJaLCに送信する。送信は、Web API(JaLC2システム外部提供インタフェース)を利用している。
8. DOI登録システムサーバーは、6. (2)で作成したLanding Pageをウェブサーバーへ転送する。4.で作成するメタデータファイルの中には、Landing Pageに掲載する

リンク情報を埋め込むことになっており、各データセンターが管理するデータやドキュメントのウェブページへの到達が可能となる。

- DOI登録システムサーバーでは、各データセンターから受け取ったメタデータファイル中のDOI情報を元に新旧バージョンの管理を行っている。更新された日付情報や旧バージョンのLanding Pageへのリンクも最新バージョンのLanding Pageに表示されるようになっている。

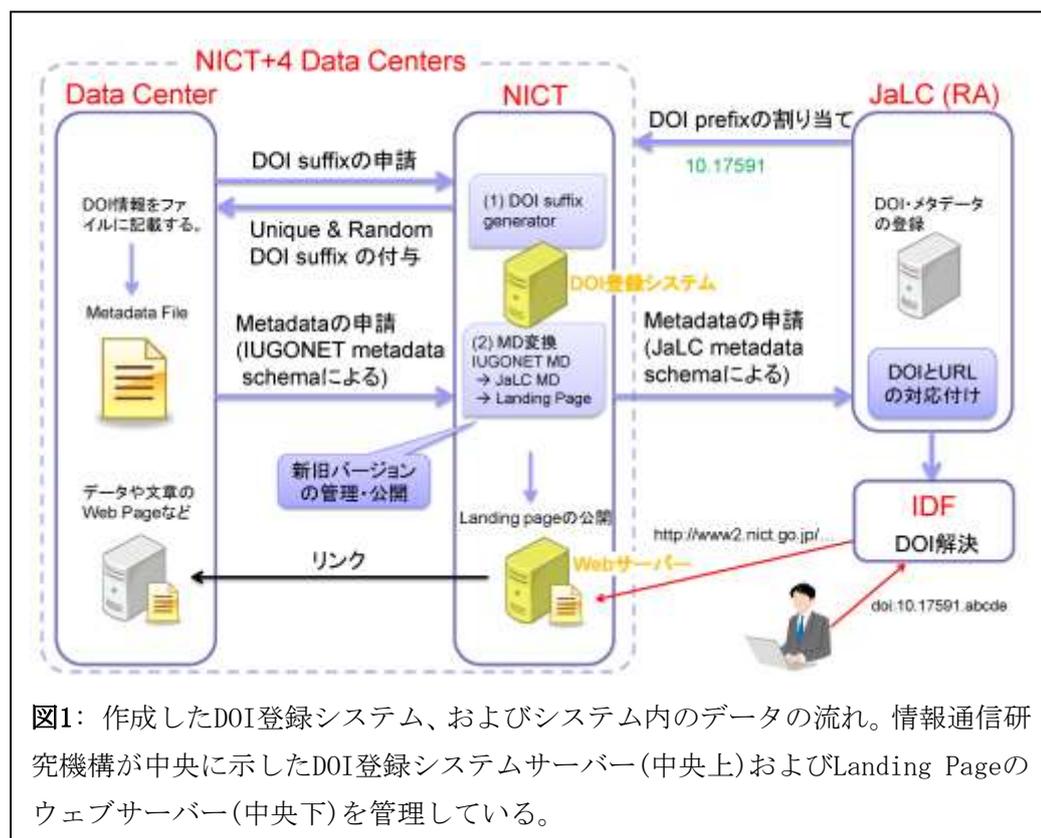


図1: 作成したDOI登録システム、およびシステム内のデータの流れ。情報通信研究機構が中央に示したDOI登録システムサーバー(中央上)およびLanding Pageのウェブサーバー(中央下)を管理している。

2.2. 日本初のデータDOIとデータ引用

本登録を2015年7月6日に行い、日本初のデータDOIを生成した。このDOIは、

doi:10.17591/55838dbd6c0ad

であり、その中身は、Mesospheric wind velocity data (30min. mean) observed with MF radar at Poker Flat, Alaskaとなっている。(図2左)

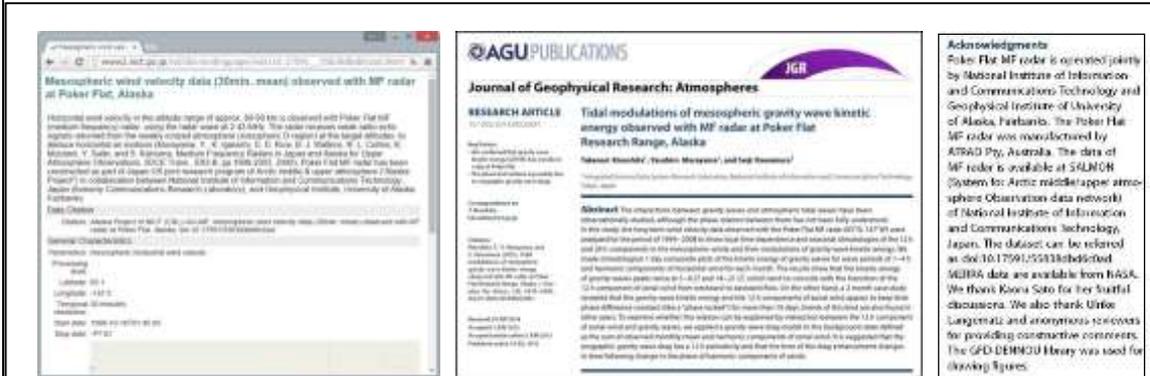


図2: (左) 日本初のDOIであるdoi:10.17591/55838dbd6c0adのランディングページ。(中) 日本初のデータ引用の論文。(右) その論文中のAcknowledgmentsで、用いたデータのDOIを示している。

また、このDOIは、Acknowledgmentsにデータを利用したことを明記する形で学术论文に引用されており、日本初のデータ引用論文の例にも貢献している。このデータ引用を行っている論文は以下のとおりである。(図2中、右)

Kinoshita, T., Y. Murayama, and S. Kawamura (2015), Tidal modulations of mesospheric gravity wave kinetic energy observed with MF radar at Poker Flat Research Range, Alaska, *J. Geophys. Res.*, 120, doi:10.1002/2014JD022647.

3. 中長期的なDOI登録方針

プロジェクト期間が終了後も、引き続き地磁気変動データや電離層観測データのDOI登録を行う。データのアップデートに応じて、既に登録したDOIやLanding Pageのアップデートも行い、メンテナンスを継続する。

DOIを登録するデータは、World Data Centerとして各機関がこれまでに長期にわたって保管・利用してきたものであり、今後もデータベースとして一般に公開していくものであるため、DOI継続性は必要最小限の努力で保証できると考えている。

4. 本登録したDOIの活用方針

研究データのトレーサビリティのため、論文引用と同様にデータ引用の研究文化・習慣を啓蒙していこうと考えている。

DOIを用いたデータ引用により、各機関がこれまでに多大な労力を費やしてきたデータベース整備・管理に対して正当な評価が行われ、それに見合った対価が支払われることを期待している。

— 以 上 —

6.7 千葉大学附属図書館

最終報告書

千葉大学附属図書館

1. 主旨

今回の実験には大学図書館の立場で参加した。主目的は、本学機関リポジトリ保有のデータセットへの DOI 付与の可能性の検討であったが、もう一つの目的は、研究データの流通に大学図書館がどのような貢献ができるかを探ることであった。

オープンアクセスのムーブメントは、論文を中心とする学術成果へのアクセスのオープン化からはじまった。雑誌論文等の学術資料の電子化オープンアクセス化は、伝統的な図書館業務における学術資料の収集、公開の流れの延長上に位置付けることが可能であり、図書館員にとっても、図書館にとっても、業務の流れのなかに位置付けやすいものであった。しかし現在急速に進みつつある「オープンサイエンス」は、根本的なスキームの見直しが必要であり、かつ先行する欧米では図書館が中心的な役割を担っている。国内の大学図書館では具体的な動きはまだ出てきていないが、今後、積極的に参画していくべきであると考えている。本実験は、研究データの流通促進の重要なキーである DOI を通して、大学図書館としてどのようなスタンスをとるべきかを検討する良い機会であると考え、本実験に参加した。

2. 実験方針

研究データ管理のプラットフォームは、基本的には附属図書館で運用している機関リポジトリシステムを想定した。本実験では、機関リポジトリシステムを研究データマネジメントのプラットフォームとして使用する際の課題点を抽出、解決策を検討し、最終的には実運用の業務フローの確立を目指した。本テストで計画したテスト登録は以下の二つのタイプである。

- 1) 既存の機関リポジトリシステム搭載のデータセット
- 2) 研究プロジェクトの成果物の登録シミュレーション

研究プロジェクトは、附属図書館と連携して研究を行っているカデミック・リンク・センターのものを想定した。

3. 結果および考察

1) データセットへのテスト登録について

コンテンツについて、今回は本学の機関リポジトリシステムのデータセット中の最大のコレクションである、萩庭コレクション（植物標本）にフォーカスを絞ることとし、複数件について登録テストを実施した。他の学術雑誌掲載論文の附属データについて、昨年度より論文本部の登録要請時に、附属データについての登録可否をあわせて問い合わせることとしたが、現時点ではほとんど登録が進んでいない。

2) 一括登録用 XML ファイルの作成

当初は Web インターフェースからの個別登録と XML による一括登録の双方をテストする計画であったが、結果として個別登録にとどまった。本登録時には、数千件オーダーのデータアップロードが必要であるが、本学の機関リポジトリのシステムでは、メタデータを所定の XML 形式に変換・出力する機能はない。手作業での XML 形式の一括登録用のデータ作成を試みたが、数千件オーダーのデータ作成は困難であり、何らかのデータコンバートのシステム開発が必要との結論にいった。本登録には機関リポジトリからダイレクトではなく、NII 経由での登録が現実的であろう。

3) メタデータスキーマ

植物標本データについて、採集者/標本作成者/デジタルファイル作成者等の複数の人間が作成にかかわっているが、それをメタデータへどのように記述するかで苦慮した。実験データごとに、メタデータの記述方法に大きなバリエーションはあることが予想され、効率的に入力し、かつ品質を保つためには、あらかじめ入力ガイドラインを準備することが必要であろう。

4) 研究プロジェクトの成果物の登録シミュレーション

研究プロジェクトの成果物の登録シミュレーションについては、検証が十分にできなかった。本センターで扱うデータは、基本的に利用者の個人情報絡む。論文、報告書等にもちいる場合には抽象化・統計化し、データの解析・保存はスタンドアローンの PC で行うなどの十分な注意を払っている。しかし研究データとして保存する場合に、どの段階のデータをどのように残すべきか議論が必要である。

4. 今後の課題

1) データマネジメントのポリシー

国内の大学図書館で具体的な形で定めているところは、おそらくまだない。何を集めて何を集めないのか？データ移行をどう考えるか？データ粒度をどう考えるか？など、論文の管理よりも明らかに複雑である。そもそもデータの定義からはじめる必要がある。個々の分野では参考とすべき例はあるが、総合的に参考にする例はない。将来的に、対象とする研究データの範囲を大学全体、研究機関全体に拡大していくためには、それぞれの大学、研究機関でデータマネジメントのポリシーを定める必要がある。

2) 研究データ対応メタデータスキーマの検討

研究成果の流通のためには、メタデータおよび識別子は極めて重要である。しかし特に研究データのメタデータは、海外事例などをみても多様かつ複雑である。まずは国内外の研究データ用のメタデータスキーマを我が国の事情にあわせて整理し、国内の機関リポジトリの標準である junii2 のスキームでどのように扱うかを検討、最終的には図書館員向けの運用マニュアルの整備および junii2 の拡張案提案を行うことが必要になる。

3) 図書館員のトレーニング体制の構築

先行する欧米では、データマネジメントプランの作成支援を図書館が行うケースが多く、対応のための図書館員向けのトレーニングツールも、多数公開されている。国内の大学においても、コンテンツを組織的かつ長期的に管理している部署は、図書館以外はほとんどない。今後の持続可能なデータ管理体制を大学内で担う部署としては、大学図書館が最適であり、そのために図書館員の組織的なトレーニング体制を構築すべきである。

4) 機関リポジトリ推進委員会

機関リポジトリを通じた学術情報の流通及び発信力の強化をめざして、国立情報学研究所と国公私立大学図書館協力委員会により、機関リポジトリ推進委員会が設置されている。活動の主体は図書館となっているが、研究データの流通・促進を組織的継続的に展開するためには機関リポジトリ推進委員会がキープレーヤーとなるであろうし、密に連携していくべきである。

「研究データへの DOI 登録実験プロジェクト」最終報告書

平成 27 年 9 月 18 日

物質・材料研究機構 報告者：田辺浩介、谷藤幹子

1. プロジェクトの実施方針

本プロジェクトでは、機構での研究成果（論文・データ・補足資料）の公表を補完し、機構として長期的に保存、適切な利用条件のもとに公開することを目的に、機構のセルフアーカイブシステム（NIMS デジタルライブラリー）を用いた職員による登録～公開までのワークフローや実現性・有効性についての検証を行った。

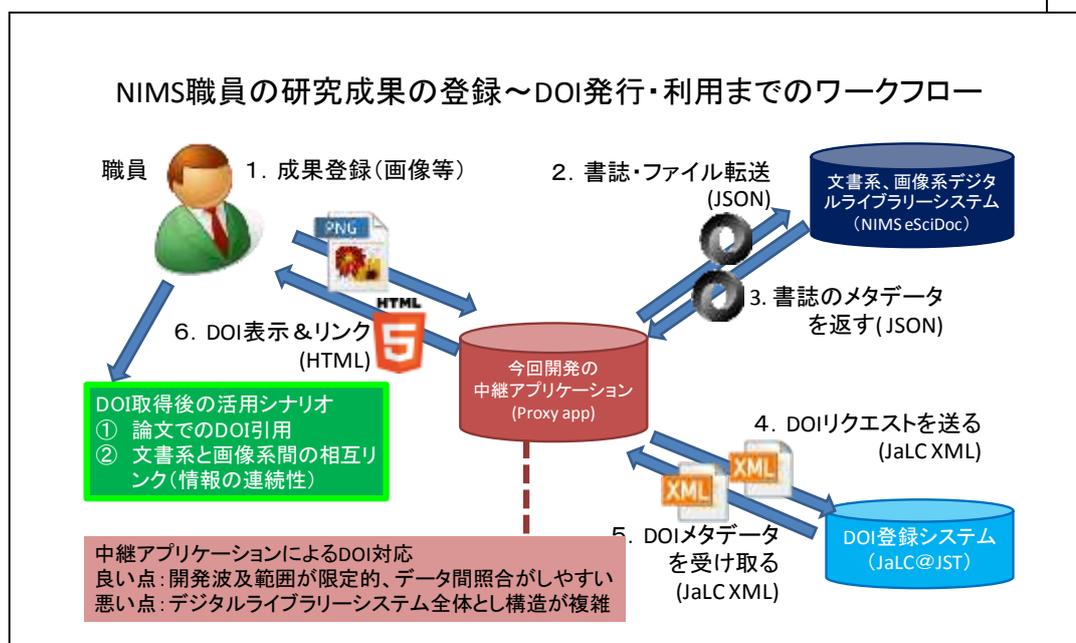
2. 登録対象データ

- ・ 画像データ等、約 150 点（電子顕微鏡で撮影した画像など）

3. プロジェクトでの成果

(1) ワークフローにみる実現性

セルフアーカイブシステムの中で DOI 発行・登録・リンク機能を付加することはせず、図書管理システムの外装として中継アプリケーションを開発した。将来のカスタマイズ等への柔軟性を考慮しオブジェクト指向で進める予定である。



(2) アプリケーションによる処理系の実現性

本アプリケーションは、職員による①セルフアーカイブ（ファイルのアップロード）、②当該ファイルの URL、③職員の ORCID 番号を JaLC に送信し、JaLC WebAPI を用いて DOI を取得する機能を持つ。研究データの場合、論文と異なりファイルそのものに著作者の記述がない場合があるため、著作者の同定の目的で ORCID のような著者識別が重要であると考えている。セルフアーカイブされた研究成果について、ORCID(人)と DOI (著作ファイル) の紐付けを行えることを確認した。また、機構で別に開発を進めている次世代研究者総覧システム (Ninja) においても、論文情報と研究者情報の名寄せの手段として本アプリケーションの利用を予定している。セルフアーカイブシステムとしてデータと DOI を体系的に表示する（取り込み等）機能においては継続課題としている。

(3) 実環境での実効性検証で検討した DOI 付与ルール

DOI の文字列生成と認識の容易さ、JaLC 側で著者名やジャーナル名、ORCID などのメタデータを DOI に関連づけることができることを前提として、次のルールを用意した。

- ・ 機構名を入れる
- ・ 連番を使う
- ・ バージョン情報は含めない
- ・ データや物質名などの名称は含めない
- ・ 区切り文字はピリオドのみを使う

例) 10.11503/nims.99999

(4) ほか、DOI 付与の場面

本プロジェクト期間中、当初予定していた DOI 付与とは別に、論文に付随する研究データ (Supplementary materials) をジャーナルが示す選択肢 (ジャーナルサイトか、機関リポジトリに登録するか) で後者の事例があり、論文出版時に記述されている研究データの URL は、DOI ではなく、セルフアーカイブの URL として対応した。後日、アーカイブ済み本データに DOI を付与する予定である。なおセルフアーカイブで公開されている研究データは複数のファイルで構成されているが、今回は個々のファイルではなくデータセットとして扱い、このセットに対して DOI を付与する予定である。

3. 今後の課題

運用化にあたっては、図書館員を介さない「研究者自らのセルフアーカイブ」に伴う検討課題が多く抽出され、本機構所轄委員会やヒアリングを通して解決、または次の課題を残している。

- ・ ファイル登録者は、ORCID 番号を取得している職員であることを前提条件とするか
- ・ 登録済みファイルのメタデータを変更した際に、新たに DOI を付与するか、DOI

付与後のメタデータは編集不可とするべきか。

- ランディングページの掲載情報（更新履歴も含み）

4. 参考

- (1) imeji <http://imeji.nims.go.jp>
- (2) 次世代研究者プロフィールサービス SAMURAI から Ninja へ.
<http://dx.doi.org/10.1241/johokanri.58.107>
- (3) equilibrium Magnetic Response of Anisotropic Superparamagnetic Nanoparticles and Possible Artifacts in Magnetic Particle Imaging.
<http://dx.doi.org/10.1371/journal.pone.0118156>
- (4) Data set used for the article Nonequilibrium magnetic response of anisotropic superparamagnetic nanoparticles and possible artifacts in magnetic particle imaging on PLOS ONE. <http://pubman.nims.go.jp/pubman/item/escidoc:1728217:1>

4. 謝辞

本実証実験は研究員の協力を得て実現に至った。（敬称略）この場を借りて御礼申る。

宮澤 薫一、先端材料プロセスユニット、フラーレン工学グループ（データ提供）
轟 眞市、先端フォトニクス材料ユニット、応用フォトニック材料グループ（データ提供）

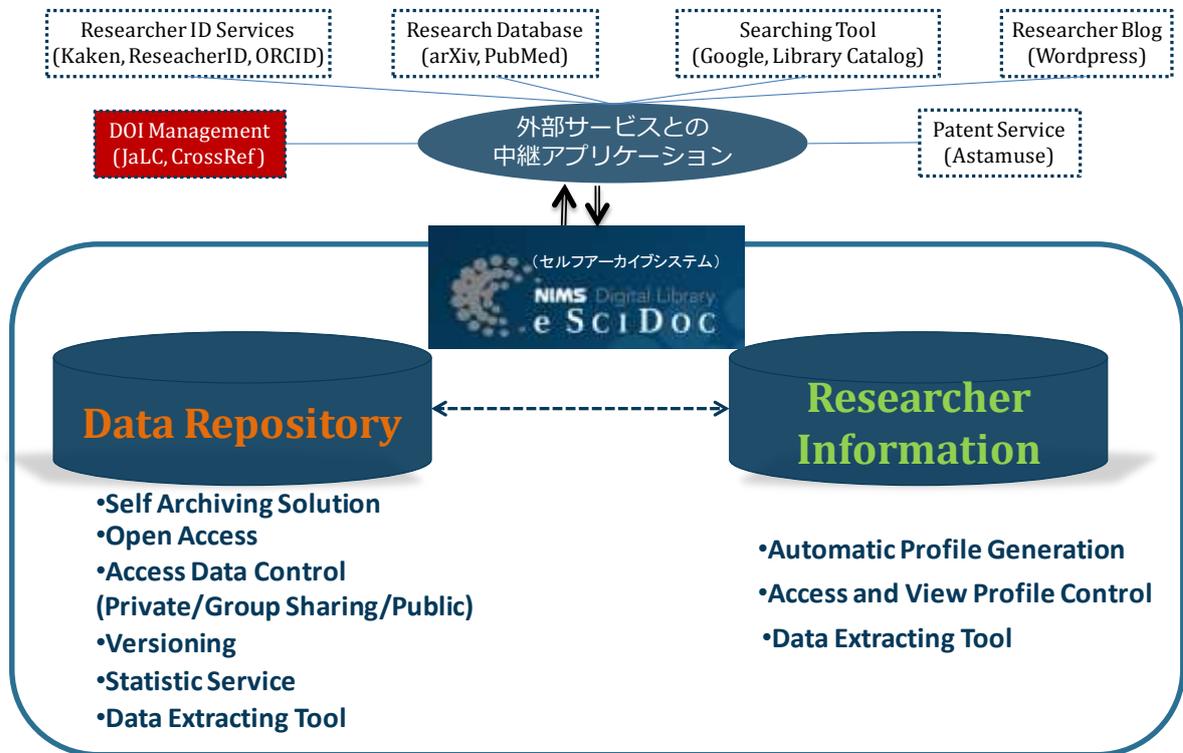
小澤 忠司、超伝導物性ユニット、材料開発グループ

末原 茂、理論計算科学ユニット、材料特性理論グループ

中西 尚志、MANA 独立研究者

以上

セルフアーカイブシステム(NIMS eSciDoc)全体構想におけるDOI管理システムの位置づけ



研究データへの DOI 登録実験プロジェクト参加最終報告書

国立研究開発法人理化学研究所

脳科学総合研究センター

神経情報基盤センター

1. 参加の背景と目的

神経情報基盤センター(以下 NIJC)では、ニューロインフォマティクス国際統合機構(以下 INCF)の日本ノードの活動の一環として、脳科学に関する web データベースの構築運営を支援し全国の大学研究所と協力して研究データを含む種々のデータを web 上で公開している。DOI を導入することで、研究成果データ公開に向けたモチベーションの向上、研究成果へのアクセシビリティの向上が期待でき、NIJC としては書誌情報の DOI 登録を開始している。

研究データ登録開始にあたっては、登録データの備えるべき属性、既存の各データベースに存在するメタデータと JaLC の DOI メタデータとの関連性、データの公開から DOI 登録までに必要な作業スキームの検討を第一の目的とした。さらに、NIJC 管理の各種データベース運営委員会(プラットフォーム委員会)との連携を通して DOI 付与のための協力体制の促進についても検討する。

2. NIJC における研究データへの DOI 登録実験プロジェクト

2.1. 方針と実施の概要

NIJC で管理する web データベースは分野毎に構成されており、各々プラットフォーム委員会がコンテンツ管理を行っている。このため、登録コンテンツの候補の選定とそのメタデータの準備をプラットフォーム委員会が行うこと、NIJC はメタデータの編集環境を提供し、プラットフォーム側と NIJC で段階的に査読を行うこととした。プラットフォーム委員会が用意するメタデータを一旦 NIJC のメタデータ管理システムに収集し、査読が済んだ段階で JaLC の DOI 登録システムに登録する。DOI のメタデータ内の signature はプラットフォーム委員長が書き込み、JaLC の登録システムへの登録実施の承認を NIJC センター長が行う。

テストプロジェクトとしてのおもな作業としては以下の通り。

- 1) 各プラットフォーム委員会に登録希望の研究データに関する情報の提出を依頼する。
- 2) 提出された情報について NIJC がデータの種類・粒度、メタデータ、永続性、アクセシビリティ、ランディングページについて確認し、登録の可否の判断と不可の場合は改善の必要な点を伝える。
- 3) 登録可となったものについて、各データの作業者が NIJC の DOI メタデータ管理システムへの登録を行い、各データベースの責任者が signature をメタデータ項目のひ

とつとして登録する。

4) 3の項目が完了した段階でNIJCが内容を再度査読し、NIJCセンター長が承認するものについてNIJCがJaLCのDOI登録システムに登録する。

ここで、NIJCのメタデータ登録システムは、今回のテストプロジェクトの実施にあたって新たに開発したもので、NIJCが管理するINCF Japan Node portalのサービスの1つとして追加された。XOOPS上のデータベースモジュールの一つであるXooNIpsのver.4.0.4用いている。新たなアイテムタイプを簡易に追加できるようになっており、これを利用し、今回JaLCのDOI登録システムに登録する際に必要なメタデータと、NIJCで必要と判断したメタデータを登録し管理するシステムを構築した。さらに、XooNIpsのアイテム登録に備わるワークフロー管理機能を利用して、DOI登録の査読のためのワークフローの導入も行った。

2.2 進捗状況

2.2.1. 進捗状況

登録対象のデータベースは13種類あるがその半数から登録希望の問い合わせがあった。各データベースにおいて永続性、アクセシビリティ、ランディングページなどの問題を検討したところ、いくつかの検討すべき点が見つかった。

登録要件を満たしたものとして、現在「無脊椎動物脳プラットフォーム(IVB-PF)」と「脳トランスクリプトームデータベース(BrainTx)」において、4件のコンテンツを対象に登録作業が進行している。いずれも作業ステップの3までがほぼ終了し、メタデータの最終的な確認を行っている。これらが終了すれば、問題なく本登録に進むと想定される。

2.2.2. 登録作業進行中の登録対象データ

● 無脊椎動物脳プラットフォーム (IVB-PF)

<https://invbrain.neuroinf.jp/modules/newdb5/detail.php?id=549>

● 脳トランスクリプトームデータベース (BrainTx)

<http://www.cdtdb.neuroinf.jp/CDT/ReferCDInformation.do?cdid=CD00010>

<http://www.cdtdb.neuroinf.jp/CDT/ReferTemporal.do?br=cerebellum&cdIdCh=CD00010.1>

<http://www.cdtdb.neuroinf.jp/CDT/BrowseImage.do?cdid=CD00010&stage=07>

2.3. まとめ

プラットフォーム委員会からは期待以上にDOI登録への関心が高く、テスト登録への参加希望も多かった。ただし、登録を推進する上でいくつか解決すべき課題も見つかった。今回は、制限された時間の中での登録のためのワークフローを検討することに目的

を絞って、4件のコンテンツ登録の進行となった。現状のワークフローにより、各データベース上の研究データを DOI 登録することは可能であるが、より円滑に登録作業を進めるためには以下の点について改善していきたい。

- NIJC としてのガイドライン、ポリシーを明示し、各データベースの管理者が DOI 登録すべきデータとその要件についての理解を深めていく。
- DOI 登録までのワークフローとその進捗状況を各データベース管理者が把握しやすくすることで、登録に対するモチベーションを上げることを検討する。
- 現在手入力であるデータベースから NIJC の DOI メタデータ管理システムへのメタデータ登録を自動・半自動化することで、省力化と転記ミスをなくすための方法を検討し実装していく。
- 同様に NIJC の DOI メタデータ管理システムから JaLC の DOI 登録システムへの登録も JaLC の DOI 登録システムが求める形式の XML データに変換する方法を検討し実装していくことで、省力化と転記ミスを図る。

以上

7 会議体開催実績

7.1 ミーティング等開催状況

<第1回>

日時：平成26年10月27日(月) 9:30～12:00

場所：国立研究開発法人科学技術振興機構 東京本部別館2階セミナー室

- 次第：(1)プロジェクトリーダー挨拶
(2)参加者紹介
(3)計画書説明
(4)活動スケジュール検討
(5)各グループ活動計画（各グループ代表）
(6)DataCite等勉強会
(7)第2回ミーティングまでの進め方検討

参加者数：37名

<勉強会>

日時：平成26年12月10日(水) 15:00～18:00

場所：国立研究開発法人科学技術振興機構 東京本部 住宅棟1F会議室

- 次第：(1)挨拶
(2)JaLCの概要
(3)JaLC2のメタデータ（DataCiteとの対比）、データデポジット方法
(4)API説明
(5)DataCite、CrossRef、IDF会議報告
(6)ライフサイエンス系データにおけるDOI関連事例調査
(7)休憩
(8)今後の進め方について、ディスカッション
(9)事務連絡

参加者数：41名

<第2回>

日時：平成27年1月28日(水) 13:00～15:00

場所：国立研究開発法人科学技術振興機構 東京本部別館2階セミナー室

- 次第：(1)プロジェクト進捗確認
(2)DOIテスト登録計画書の検討・承認
(3)ポリシー検討の進め方
(4)「ジャパンリンクセンター活用のための対話・共創の場」（第2回）の開催(案)

(5)プロジェクト実施状況の公表について

(6)事務連絡

参加者数：36名

<第3回>

日時：平成27年3月13日(金) 10:00～12:00

場所：国立情報学研究所（東京都千代田区一ツ橋2-1-2）12階 1208,1210 会議室

次第：(1)「ジャパンリンクセンター活用の為の対話・共創の場（第2回）～研究データへのDOI登録～」開催報告

(2)ポリシー検討に関する議論

(3) DOI テスト登録の際に気が付いた課題の共有

(4)その他

参加者数：34名

<第4回>

日時：平成27年4月20日(月) 15:00～17:30

場所：国立研究開発法人科学技術振興機構 東京本部別館2階会議室A-2

次第：(1)DOI テスト登録実施状況の確認と課題等の共有

(2)今後のスケジュール

(3)ポリシー検討に関する議論

(4)その他

参加者数：37名

<第5回>

日時：平成27年6月3日(水) 15:00～17:30

場所：国立研究開発法人 科学技術振興機構 東京本部別館2階会議室A-2

次第：(1)研究データに対するDOI登録ガイドライン検討

(2)DOI テスト登録報告と本登録計画

(3)DOI 本登録開始要領

(4)今後のスケジュール

(5)その他

参加者数：29名

<第6回>

日時：平成27年8月5日(水) 15:00～17:30

場所：国立研究開発法人 科学技術振興機構 東京本部別館2階セミナー室

- 次第：(1)本プロジェクトにおける最終成果物とりまとめの方向性
(2)研究データに対する DOI 登録ガイドライン検討
(3)活動状況
(4)その他

参加者数：35名

<第7回>

日時：平成27年10月8日(木) 15:00～17:30

場所：国立情報学研究所 12階 1208, 1210 会議室

- 次第：(1)各参加機関からの最終報告
(2)本プロジェクトにおける最終成果物
(3)本プロジェクトの最終報告会
(4)その他

参加者数：38名

7.2 対話・共創の場

<対話・共創の場>

日時：平成27年2月27日(金) 15:00～18:00

場所：国立研究開発法人科学技術振興機構東京本部別館 2階会議室 A-1,A-2

- 次第：(1)開会のご挨拶
(2) ジャパンリンクセンターの概要
(3) プロジェクトの概要
(4) プロジェクトの進捗状況
(5) 参加機関の取組み (1) DIAS プロジェクトの取組み
(6) 参加機関の取組み (2) 国立極地研究所・北極域データアーカイブの紹介と今後の取組み
(7) 休憩
(8) Research Data Alliance の活動状況
(9) オープンサイエンスの推進と Data Citation Principles
(10) 対話・共創
(11) 閉会のご挨拶

参加者数：69名

集会報告：福山 樹里 「ジャパンリンクセンター活用の為の対話・共創の場(第2回)～研究データへの DOI 登録～」 情報管理(2015)58 巻2号 145 頁, doi:
10.1241/johokanri.58.145

7.3 報告会

<中間報告会>

日時：平成 27 年 7 月 3 日(金) 14:00～18:00

場所：国立研究開発法人科学技術振興機構東京本部別館 2 階会議室 A-1,A-2

次第：(1) 開会のご挨拶

(2) プロジェクトの概要

(3) 参加機関の取組み (9 機関)

(4) 休憩

(5) 「研究データへの DOI 登録ガイドライン(案)」報告

(6) 対話・共創

(7) 閉会のご挨拶

参加者数：70 名

集会報告：小林 賢 「ジャパンリンクセンター「研究データへの DOI 登録実験プロジェクト」中間報告会」情報管理(2015)58 巻 6 号 485 頁, doi: 10.1241/johokanri.58.485

<最終報告会>

日時：平成 27 年 12 月 3 日(金) 9:00～12:00

場所：国立研究開発法人科学技術振興機構東京本部地下 1 階大会議室

(IDF DOI Outreach meeting の中で開催)

参加者数：130 名

8 プロジェクト参加者名簿

No.	機関名	所属	氏名	
1	科学技術振興機構	バイオサイエンスデータベースセンター	櫛田 達矢	
2	国立極地研究所	北極観測センター	矢吹 裕伯	
		国際北極環境研究センター	水澤 美帆	
		宙空圏研究グループ	岡田 雅樹	
		極域データセンター	門倉 昭	
		北極域データアーカイブ	杉村 剛 照井 健志	
		GRENE 北極事務局	熊谷 宏靖 上曾 由紀江	
		情報図書室	南山 泰之	
3	国立情報学研究所	学術ネットワーク研究開発センター(コンテンツ科学研究系)	山地 一禎	
		学術基盤推進部学術コンテンツ課	高橋 菜奈子 加藤 寛士 前田 朗 丸山 寛 高木 正弘	
4	国立情報学研究所	コンテンツ科学研究系	北本 朝展	
	海洋研究開発機構	地球情報基盤センター 統合地球情報研究開発部	角田 晋也 西村 一 柏野 祐二	
	東京大学	生産技術研究所	絹谷 弘子	
		空間情報科学研究センター	小野 雅史	
	京都大学		中原 陽子 李 吉屹 清水 敏之 吉川 正俊	
		国立環境研究所	地球環境研究センター 地球環境データベース推進室	白井 知子 曾 継業 佐藤 智弘 塚田 康弘
			地球環境研究センター 陸域モニタリング推進室	真板 英一 山尾 幸夫
	5	産業技術総合研究所	情報・人間工学領域 研究戦略部	小島 功
			計測標準研究部門	山下 雄一郎 渡邊 宏
	6	情報通信研究機構	統合データシステム研究開発室	村山 泰啓 木下 武也 渡邊 堯 國武 学
電磁波計測研究所 宇宙環境インフォマティクス研究室			山本 和憲 加藤 久雄 西岡 未知	
電離圏宇宙天気世界資料センター			石井 守	
情報利活用基盤研究室			是津 耕司	
		ユニバーサルコミュニケーション研究所 情報利活用基盤研究室	赤星 祐平	
京都大学大学院理学研究科		地磁気世界資料センター	能勢 正仁 家森 俊彦	
国立情報学研究所		新領域融合研究センター	小山 幸伸	

	(株)情報試作室		相良 毅
	宇宙航空研究開発機構	宇宙科学衛星資料センター	篠原 育
	国立極地研究所	宙空圏研究グループ	田中 良昌
7	千葉大学附属図書館	利用支援企画課	杉田 茂樹 三角 太郎 綾田 陽子 中原 由美子 木下 直
8	物質・材料研究機構	企画部門科学情報室	谷藤 幹子 田辺 浩介 河野 健一郎
9	理化学研究所脳科学総合研究センター神経情報基盤センター	神経情報基盤センター	山口 陽子 奥村 嘉宏 磯野 由依
	JaLC 運営委員会	国立情報学研究所	武田 英明
		科学技術振興機構	水野 充 小賀坂 康志
		国立国会図書館	原田 久義 伊東 敦子
	JaLC 分科会	筑波大学附属図書館	真中 孝行
		国立国会図書館	徳原 直子 福山 樹里 小澤 弘太
		国立情報学研究所	大向 一輝
	JaLC 事務局	科学技術振興機構	中島 律子 加藤 斉史 余頃 祐介 小林 賢 岩崎 泰久 針谷 和夫 高築 浩一 宮村 あゆみ

9 附録資料

<web 版にて、全文を掲載>

- 9.1. 附録資料 1 研究データへの DOI 登録実験プロジェクトについて【予告】
- 9.2. 附録資料 2 公募要領
- 9.3. 附録資料 3 「研究データへの DOI 登録実験プロジェクト」参加機関募集 公募説明会
質疑応答内容
- 9.4. 附録資料 4 各参加機関から提出された「DOI テスト登録計画」
- 9.5. 附録資料 5 各参加機関から提出された「DOI テスト登録計画」
- 9.6. 附録資料 6 JaLC における研究データの登録手順書
- 9.7. 附録資料 7 各参加機関から提出された「DOI 本登録計画」