

# 機関リポジトリの メタデータ概論

平成22年度学術ポータル担当者研修  
NII会場 平成22年8月26日

静岡大学附属図書館 杉山智章

名古屋会場 北海道大学附属図書館 野中雄司 講師  
作成資料を再編したものです。

## この講義の目的

- 機関リポジトリの特徴のひとつである、**OAI-PMH**というプロトコル(通信手順)について、仕組みと役割を理解する。
- **内部メタデータ、外部提供メタデータ**の違いと、**oai\_dc**や**junii2**の**クロスウォーク**設計について、基本を理解する。

→リポジトリ担当者同士や、業者と基本的なコミュニケーションがとれるように。

# この講義の内容

- 1 基本理念
  - 1.1 何を考えてメタデータ設計すればよいか
  - 1.2 OAI-PMH
- 2 メタデータの設計方法
  - 2.1 機関リポジトリのメタデータ
  - 2.2 内部メタデータ設計
  - 2.3 外部へのメタデータ提供設計
- 補講 OAI-PMHを試してみる
- まとめ

1.1

何を考えてメタデータ設計  
すればよいか

## 機関リポジトリとは？（復習です）

クリフォード・リンチ(Clifford Lynch)：

大学がその構成員に提供する、大学とその構成員が創造したデジタル資料の管理や発信を行うために、**大学がそのコミュニティの構成員に提供する一連のサービス。**

## つまり機関リポジトリのサービスとは

- 機関内構成員に対するサービスである。

= (イコール)

- 機関内産出の研究成果を世界中からアクセスされやすくする。
  - これを可視性(visibility: ビジビリティ)の向上といいます。

# ビジビリティを向上させるには

- 自動的にページをインデクシングしてくれるGoogleなどの検索エンジンから見えるように（クローリングされるように）すること。
- 文献のメタデータを所定の様式で用意すると、そのデータを持って行って、自サービスにデータ登録し、サービス展開してくれるものがある。（NIIのJAIROなど）

→この講義はこちらの話

# 1.2

## OAI-PMH



# OAI-PMHとは？

- 複数のリポジトリのメタデータを収集し、それに基づいたサービスを提供するために開発されたデータ提供・収集用のプロトコル(通信手順)。
- HTTP通信を使用するため、アプリケーションに依存しない。
- Open Archives Initiative  
Protocol for Metadata Harvesting

- <http://www.openarchives.org/OAI/openarchivesprotocol.html>

- 日本語訳 <http://www.nii.jp/irp/archive/translation/oai-pmh2.0/>

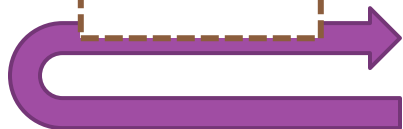
# OAI-PMHのイメージ

機関A  
リポジトリ

機関B  
リポジトリ

機関C  
リポジトリ

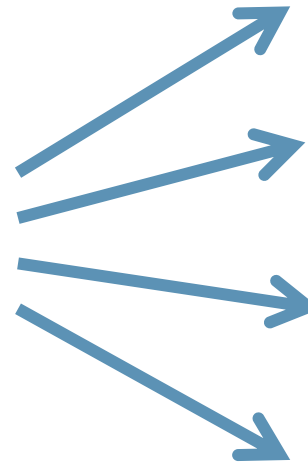
OAI-PMH



IRDBコンテンツ分析  
NII Institutional Repositories DataBase Contents Analysis



SCIRUS  
for scientific information only



SCOPUS

データプロバイダ

サービスプロバイダ

# リポジトリにとってのOAI-PMH

- 主にリポジトリ用の外部へのデータ提供を実現するためのもの。
- リポジトリを公開するということ  
= (イコール)

OAI-PMHで外部にデータ提供することです。

- あるサービスにデータ提供すると、その提供先がさらに別のサービスにデータ提供し、広がっていくことも。

## OAI-PMHで実現すること

- 機関リポジトリに収録されているメタデータを他のサービスが(勝手に)持って行って(刈り取って)くれて、サービスしてくれる。

ハーベスティング

例えば  
JAIRO, CiNii

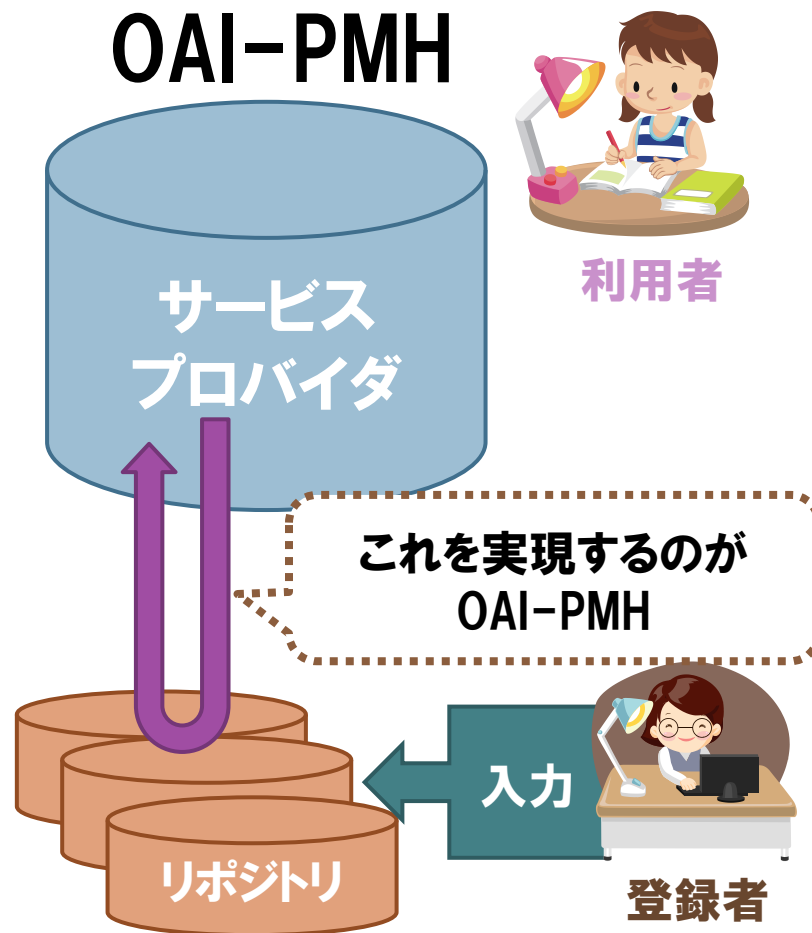
- データの更新や削除も考慮してくれる。

# NASCIS-CATとのモデル比較

## NACCSIS-CAT



## OAI-PMH



2.1

機関リポジトリのメタデータ

## 2つのメタデータ

- リポジトリシステム本体で使用するメタデータ  
(内部メタデータ)
  - 目録に相当する書誌情報
  - システムがコンテンツを管理するためのデータなど
- 外部への提供メタデータ
  - OAI-PMHなどで外部に提供するデータ

# 内部メタデータ

- コンテンツを登録する際に入力する書誌情報や、システムが自動的に付与する登録日などのデータ。 → 通常リポジトリ内のデータベースに登録される

The image shows two overlapping windows. On the left is the HUSCAP registration form, and on the right is a Microsoft Access database window.

**HUSCAP Registration Form (Left Window):**

- Header: 北海道大学学術成果コレクション HUSCAP Hokkaido University collection of Scholarly and Academic Papers
- Navigation: 記述 (selected), 記述, 記述, アップロード, 確認, ライセンス, 完了
- Section: 投稿: アイテムの記述
- Text: 投稿アイテムに関する情報を下のフィールドに入力してください。ほとんどのブラウザでは、タブキーによりフィールドを移動することができます。
- Fields: status\_origin, 担当者, status (none), Title, 出版年月日, (特論のみ)資料本体に収録されている出版年月日, (特論のみ)ページ数, 出版者, 引用情報(四書用), 雑誌名
- Buttons: Check Bot

**Microsoft Access Database (Right Window):**

- Table: 200812
- Fields: placement, qualifier, text\_value
- Records:

Field	Value
creator	紙谷, 五月
creator	野中, 雄司
creator	杉田, 茂樹
date	accessioned 2008-12-05T01:41:19Z
date	available 2008-12-05T01:41:19Z
date	issued 2008-12-01
identifier	url http://hdl.handle.net/2115/34975
description	ret the usage 特集 ファイタビリティ向上
description	abstract 機関リポジトリは独自の集客力を持たず、外部の情報サービスシステムに集客を
description	provence Submitted by 2 理字系 図書館 (huscap2@lib.hokudai.ac.jp) on 2008-12-05
description	provence Approved for entry into archive by 2 理字系 図書館(huscap2@lib.hokudai.ac.jp) on 2008-12-05
description	provence Made available in DSpace on 2008-12-05T01:41:19Z (GMT). No. of bitstreams: 1
language	iso jp
publisher	情報科学技術協会
relation	url http://www.infosta.or.jp/journal/200812.html#6
subject	機関リポジトリ
subject	リンク/リポジトリ
subject	OAI-PMH
subject	OAI-PMH
subject	Altway
subject	Institutional repository
subject	link-repo:her
title	機関リポジトリへのアクセス経路
title	alternative Access path to institutional repositories
type	alternative article (author version)
creator	alternative Koike, Satsuki
creator	alternative Nomaka, Yui
creator	alternative Sugita, Shigeji
subject	misc 007
citation	title 情報科学技術 三 The Journal of Information Science and Technology Assn
citation	issn 0913-3801
citation	volume 59
citation	issue 12
citation	page 610
citation	epage 614
citation	note AN10005857



# 外部への提供メタデータ

- 提供先に合わせた形式で、その都度、内部メタデータを自動変換して出力。

この XML ファイルにはスタイル情報が関連づけられていないようです。以下にドキュメントツリーを表示します。

```
<OAI-PMH xmlns:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2008-07-13T02:54:36Z</responseDate>
  <request identifier="oai:eprints.lib.hokudai.ac.jp:2115/34975" metadataPrefix="oai_dc" verb="GetRecord">
    http://eprints.lib.hokudai.ac.jp/dspace-oai/request
  </request>
  <GetRecord>
    <record>
      <header>
        <identifier>oai:eprints.lib.hokudai.ac.jp:2115/34975</identifier>
        <datestamp>2008-12-08T09:53:35Z</datestamp>
        <setSpec>hdl_2115_20083</setSpec>
        <setSpec>hdl_2115_1</setSpec>
      </header>
      <metadata>
        <oai_dc:dc xmlns:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
          <dc:creator>紙谷, 五月</dc:creator>
          <dc:creator>野中, 雄司</dc:creator>
          <dc:creator>杉田, 茂樹</dc:creator>
          <dc:date>2008-12-01</dc:date>
          <dc:identifier>http://hdl.handle.net/2115/34975</dc:identifier>
          <dc:description>特集 ファインズビリティ向上</dc:description>
        </oai_dc:dc>
        <dc:description>
          機関リポジトリは独自の集客力を持たず、外部の情報サービスシステムに集客を依存する。従って、より多く利用されるためには、主たる閲覧利用者層として想定される研究者の日常的な情報探索行動の中で、取載文献が自然に発見されアクセスされるような経路の形成と誘導が必要となる。これを目的として実装され、リンク・リゾルバのための追加情報源として機能するAIRwayシステムを概観し、現在の被利用状況について述べ、今後の展望についてまとめる。
        </dc:description>
        <dc:description>
          An institutional repository does not have its own power to attract large users, but commends user navigations to external service providers. Academics are assumed to be the primary users of institutional repository. Therefore, to gain their access, it is important to make the repository contents discoverable and accessible within their day-to-day research behavior. We have implemented "AIRway" system which is designed to be yet another knowledgebase for link-resolvers. This paper describes its functions and the usage.
        </dc:description>
        <dc:language>j</dc:language>
        <dc:publisher>情報科学技術協会</dc:publisher>
        <dc:relation>http://www.infosta.or.jp/journal/200812jhtm#6</dc:relation>
        <dc:subject>機関リポジトリ</dc:subject>
        <dc:subject>リンク・リゾルバ</dc:subject>
        <dc:subject>OAI-PMH</dc:subject>
        <dc:subject>OpenURL</dc:subject>
      </metadata>
    </record>
  </GetRecord>
</OAI-PMH>
```

OAI-PMH  
機械が読み取りやすい形式になっている

# 世界標準は顧慮せず,なおかつ顧慮する

- **内部**メタデータ設計は**自由**に！
  - リポジトリ上で見せたい情報や、内部でコンテンツを管理するためのデータを保持。
  - ハーベスティングしてほしいサービスに対応できるように情報を保持。
- **外部**への提供メタデータは**標準準拠**で！
  - OAI-PMHなどで外部へデータ提供するために標準準拠で！
  - サービスプロバイダによっていろいろな形式(**メタデータフォーマット**)の出力が求められる。
    - junii2, OAI\_DC, ETD-MS . . .

2.2

内部メタデータ設計

# 内部メタデータは自由

- 機関リポジトリはメタデータ項目、記述方法について、目録のような標準化された規則はない。
  - 扱うコンテンツの資料種別が多様だから・・・
- 設計は機関ごとにしなくてはならない。
  - 何を考慮して設計するか

# 設計で考慮すべきこと

- どのようなメタデータフォーマットでデータを外部に提供するのかをあらかじめ想定する。
- リポジトリ内部で保持する内部メタデータを資料種別ごとに明確に決定すること。

必要な情報を過不足なく定義すること

# 資料種別ごとに必要項目を考える

	雑誌論文	紀要論文	学位論文	会議発表	科研報告	授業教材
タイトル	●	●	●	●	●	●
著者	●	●	●	●	●	●
著者キーワード	●	●	●		●	
抄録・概要	●	●	●		●	
掲載誌情報	●	●				
学位授与機関・学位の種別 etc.			●			
会議名・開催地・日時				●		
課題番号・研究期間・種目 etc.					●	
授業年度・授業科目 etc.						●

# 学位論文の例

## HUSCAP

dc.title フラキストン「標本」史

dc.type theses (doctoral)

dc.subject.ndc 900

dc.description.degreeGrantor Hokkaido University (北海道大学)

dc.description.degreeLevel 博士

dc.description.degreeDiscipline 文学

## SURE

dc.title III/浴床によるGaPの液相成長

dc.contributor.author 杉浦, 敏文

dc.contributor.transcription スギウラ, トシフミ

dc.contributor.alternative Sugiura, Toshifumi

dc.subject.ndc 541

dc.description 静岡大学博士論文; 報告番号:甲第1号; 学位の種類:工学博士

dc.description.textversion publisher

dc.publisher 静岡大学

学位授与機関・学位の種類を内部で保持したいので追加

# 雑誌論文の例

## HUSCAP

**dc.type** article (author version)  
**dc.subject.ndc** 499  
**dc.citation.jtitle** International Journal of Pharmaceutics  
**dc.citation.issn** 0378-5173  
**dc.citation.volume** 376  
**dc.citation.issue** 1-2  
**dc.citation.spage** 99  
**dc.citation.epage** 103  
**dc.citation.doi** 10.1016/j.ijpharm.2009.04.032

## SURE

**dc.identifier.uri** <http://hdl.handle.net/10297/34>  
**dc.identifier.issn** 00225193  
**dc.identifier.eissn** 10958541  
**dc.identifier.jtitle** Journal of Theoretical Biology  
**dc.identifier.volume** 256  
**dc.identifier.issue** 4  
**dc.identifier.spage** 644  
**dc.identifier.epage** 654  
**dc.language.iso** eng  
**dc.relation.doi** 10.1016/j.jtbi.2008.10.024  
**dc.relation.pmid** 19038271  
**dc.relation.pmid** 19038271  
**dc.relation.pmid** 19038271

JAIROやCiNiiにデータ提供(junii2形式)するために、掲載誌情報を保持するよう項目を追加



# 内部メタデータの記述

- 書誌の記述方法について、NACISIS-CATコーディングマニュアルのような**統一規範はない**。
- NII「学術コンテンツ登録システム」の「II.データ記述マニュアル」を参考にしている機関もある。
  - <http://www.nii.ac.jp/nels/man/index.html>

「タイトルにルビが入っている場合はどうやって入力すれば？」  
などなど記述方法についての疑問

2.3

外部へのメタデータ提供設計

# 外部提供は標準準拠

- 外部提供メタデータには、標準化されたメタデータフォーマット仕様がある。
  - サービスプロバイダは、複数リポジトリのデータを収集してサービスをするため。

メタデータフォーマット	使用サービス等
OAI_DC	OAI_PMHで必須。OAIster等で使用。
junii2	JAIRO, CiNiiなどNIIで使用。
ETD-MS	学位論文用。NDLTDで使用。

# メタデータフォーマットの詳細情報

- OAI\_DC

- [http://www.openarchives.org/OAI/2.0/oai\\_dc.xsd](http://www.openarchives.org/OAI/2.0/oai_dc.xsd)
- <http://dublincore.org/documents/dces/>
- **DRIVERガイドライン(日本語訳あり)**  
<http://www.driver-support.eu/managers.html>

- junii2

- <http://irdb.nii.ac.jp/oai/junii2.xsd>
- <http://www.nii.ac.jp/irp/archive/system/junii2.html>
- **junii2ガイドライン**  
[http://www.nii.ac.jp/irp/archive/system/junii2\\_guide.html](http://www.nii.ac.jp/irp/archive/system/junii2_guide.html)

- ETD-MS

- <http://www.ndltd.org/standards/metadata/etdms/1.0/etdms.xsd/view>
- <http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html>

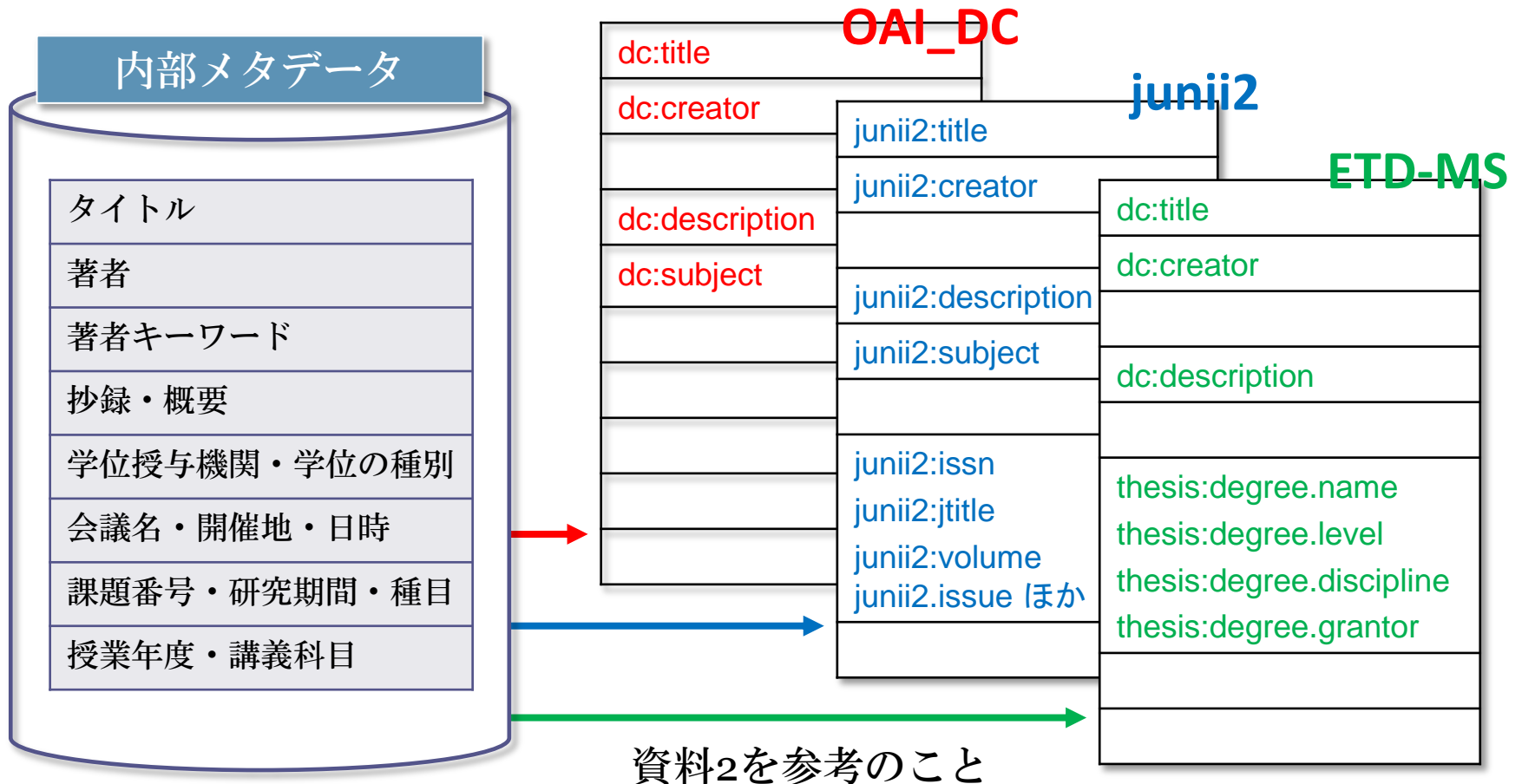
# 外部へメタデータを出力

- OAI\_DC, junii2, ETD-MSなどでメタデータを提供するために必要なこと。



- 内部メタデータから、外部提供メタデータへの変換ができるようにマッピングしておく必要がある。

# クロスウォーク (イメージ)



補講

OAI-PMHを試してみる

# OAI-PMH要求例

[Go](#)

静岡大学学術リポジトリ (SURE)のBaseURL

<http://ir.lib.shizuoka.ac.jp/dspace-oai/request>

**?verb=ListRecords**

データを全部ください

**&metadataPrefix=junii2**

junii2 形式で

**&from=2010-07-01**

2010年7月1日 から

**&until=2010-08-26**

2010年8月26日まで

登録・更新  
された

**&set=hdl\_10297\_20**

hdl\_10297\_20  
(創造院の雑誌論文)の集合



(応答例を確認しましょう)

# 要求と値

要求	説明	指定できる値
Identify	どんなりポジトリですか？	
ListMetadataFormats	どんな形式でデータを出力できますか？	i
ListSets	どんな集合がありますか？	r
GetRecord	データを1つください	m, i
ListIdentifiers	IDリストをください	m, s, f, u, r
ListRecords	データを全部ください	m, s, f, u, r

値	説明
<b>m</b> etadataPrefix	このデータ形式で
<b>s</b> et	この集合のデータを
<b>f</b> rom	この日付のデータから
<b>u</b> ntil	この日付のデータまで
<b>i</b> dentifier	このIDのデータを
<b>r</b> esumptionToken	次項のデータを

まとめ

# この講義の確認

外部データ提供用  
メタデータは  
標準仕準拠で！

クロスウォーク  
設定

機関  
リポジトリ

内部メタデータ  
は自由に！  
不足なく！

OAI-PMH  
(junii2)

OAI-PMH  
(OAI\_DC)

OAI-PMH  
(ETD-MS)

IRDBコンテンツ分析  
NII Institutional Repositories DataBase Contents Analysis



SCIRUS  
for scientific information only



# リポジトリを運用する上で

- 導入時には、内部メタデータのカスタマイズだけでなく、少なくともOAI-PMHに対応したOAI\_DC, junni2のクロスウォーク設定をする。
- 実際にメタデータ入力し、OAI-PMHの各フォーマットの応答画面を確認し、必要があれば調整する。
- 機関内だけで悩まないで、コミュニティに相談しましょう。みんな同じような経験を積んでいます。

# おわりに

サービスプロバイダにはどうやって  
ハーベスティングしてもらえばいいの？

次の「機関リポジトリの公開」講義で！！

機関内産出の研究成果を世界中から  
アクセスされやすくするために…