



# junii2ガイドラインと SUREメタデータ

XooNlps研究会2011関西ワークショップ  
(2011.12.21・大阪市立大学)

杉山 智章



静大図書館

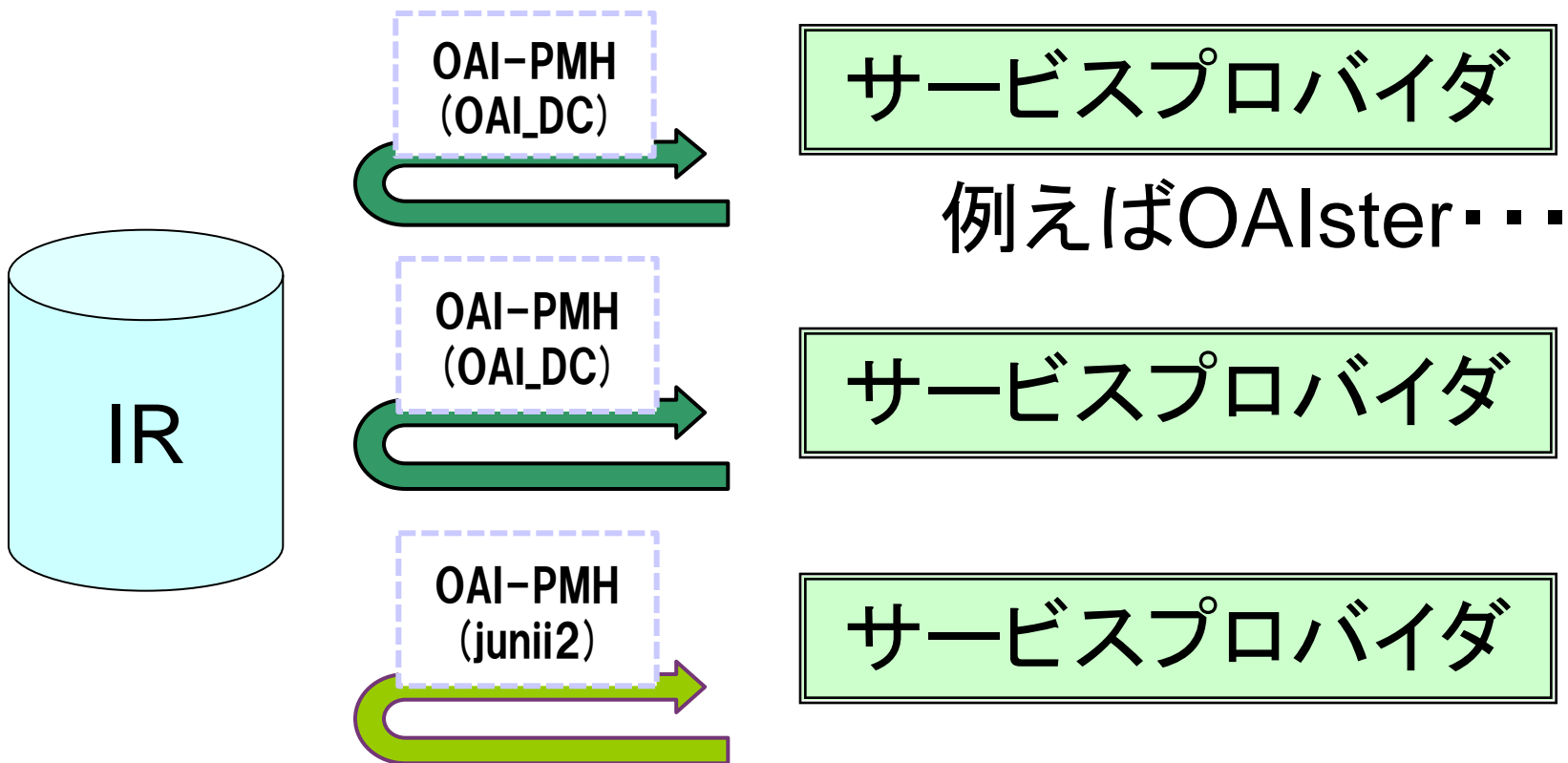
# 今日お話しすること

- 対象: IR担当者に
- 内容: 書誌情報のメタデータのお話し
  - junii2ガイドラインの目的と要点
  - 静岡大学学術リポジトリ(SURE)のメタデータ



# junii2ガイドラインの目的

# OAI-PMH

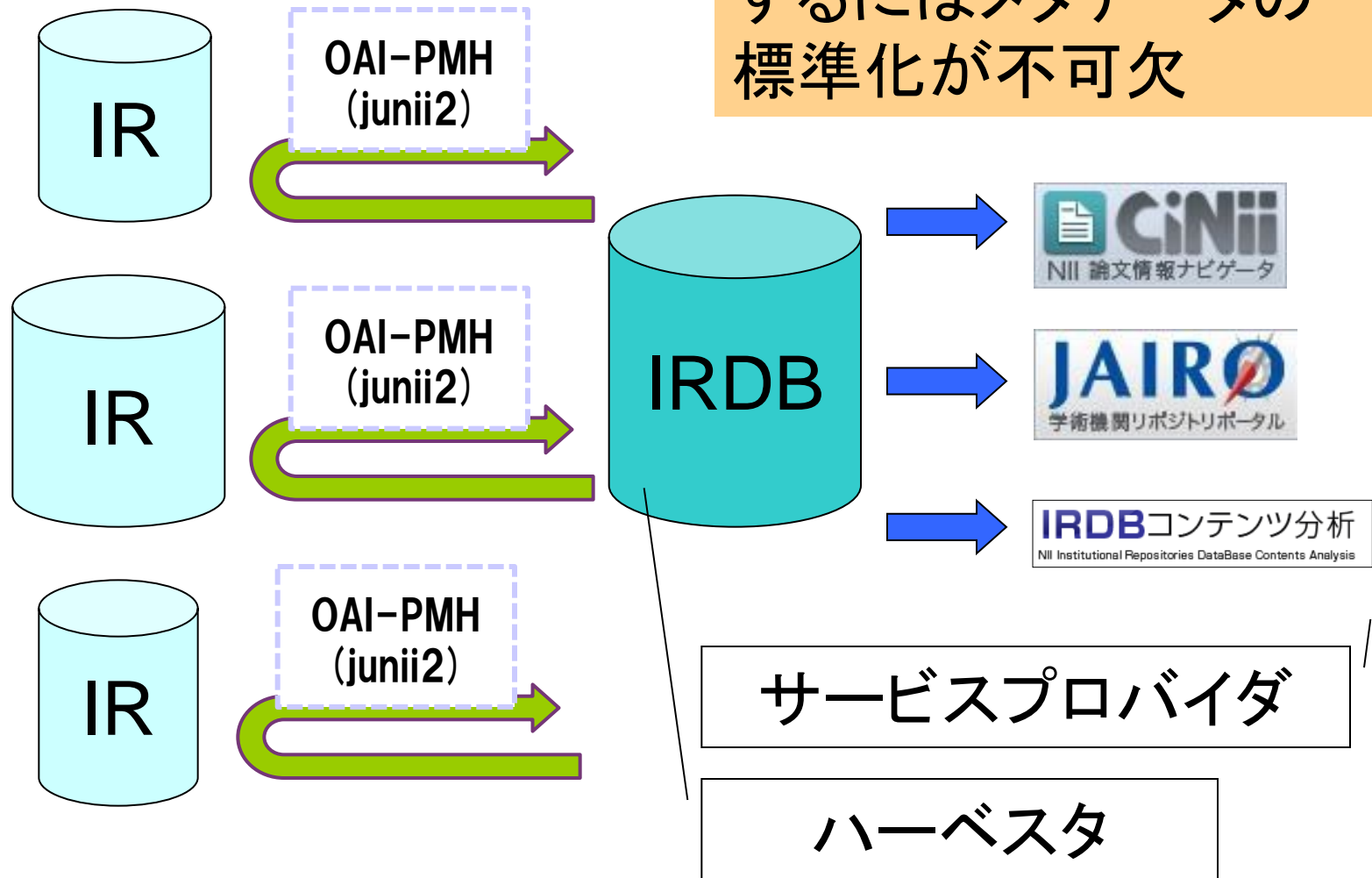


Harvest(ハーベスト)

`request?verb=ListRecords&metadataPrefix=oai_dc`

# IRとNIIサービス

高次なサービスを提供するにはメタデータの標準化が不可欠

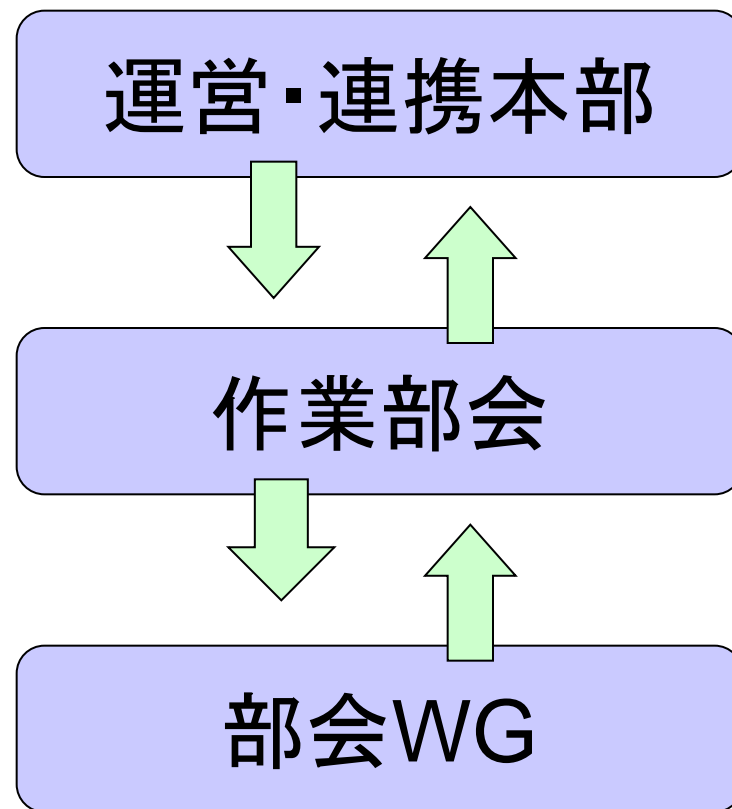


サービスプロバイダ

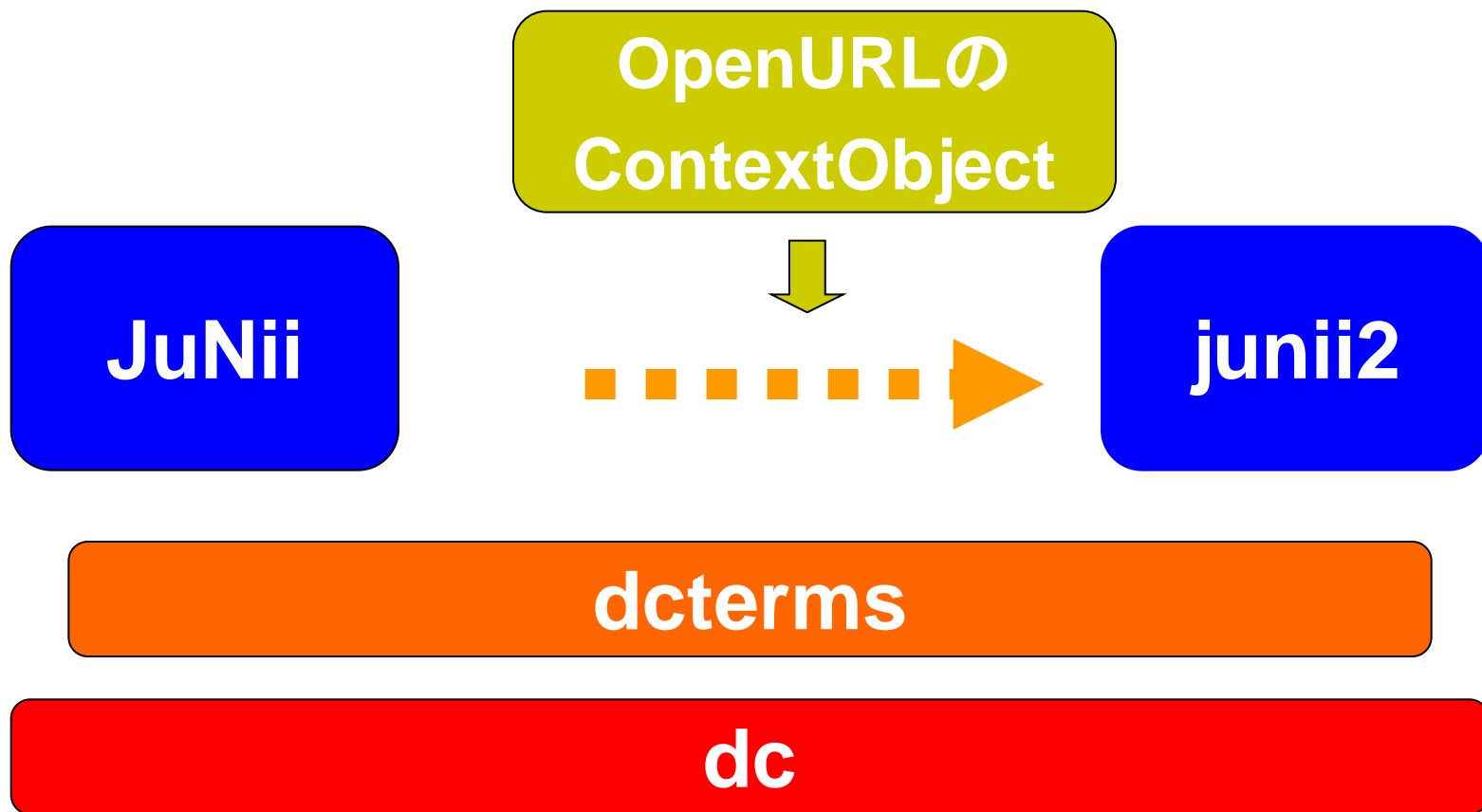
ハーベスタ

# 学術コンテンツ運営・連携本部図書館 連携作業部会WG

- 平成21年度6WG設置
  - WG2 IRシステム検討
    - タスク2 メタデータ指針
      - junii2ガイドライン作成
- WGについては22、23年度も改編されながら設置されている



# JuNiiからjunii2 (想像図)





# junii2

- Schema (スキーマ)

<http://irdb.nii.ac.jp/oai/junii2.xsd>

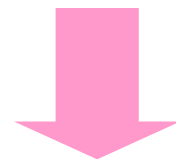
- ドキュメント

[http://www.nii.ac.jp/irp/archive/system/pdf/junii2\\_elements\\_guide\\_ver2.pdf](http://www.nii.ac.jp/irp/archive/system/pdf/junii2_elements_guide_ver2.pdf)

[http://www.nii.ac.jp/irp/archive/system/pdf/type\\_NII.pdf](http://www.nii.ac.jp/irp/archive/system/pdf/type_NII.pdf)

# 何のためのメタデータ？

大切なことは幅広く利活用されること



さまざまなweb資源とつながることができること

**OAI-PMH**はその方法のひとつ

**junii2**は(NIIとのための)書式

**junii2ガイドライン**は使用説明

# 作成上の留意点(結果として)

- NIIサービスにとって ○ (マル)
  - CiNii, JAIRO, IRDBコンテンツ分析 など
- oai\_dcにダムダウンが ○ (マル)
  - DRIVERガイドライン など
- OpenURLで解決が ○ (マル)
  - AIRway など
- 書誌情報として ○ (マル)
  - SIST02 など



各要素の説明 +  $\alpha$

# 共通して留意すること

- “junii2の使用方法=IRのメタデータ入力方法”  
ではない
- 正規化
  - 英数/記号については半角(1byte)文字とすべき。
  - スキームが定められているものについて、大文字・小文字の違いも注意しなければいけない。記述要素については大文字・小文字は原則として原表現のままよい。

# NIIttype

- ガイドラインのターゲットは主に論文系（CiNii連携のために）

Journal Article	15.3%
Departmental Bulletin Paper	52.2%
Article	4.3%
計	71.8%

IRDB 本文あり(2011.11)

# タイトル

- 学術コンテンツ登録システムのシリーズや特集の記述手引はガイドラインとしては採用せず。
  - 書誌情報の書式として標準化されていない
- タイトル関連情報がある場合は「△:△」で区切って使用する。並列タイトルはalternativeを使用。

# 作成者, 公開者, 寄与者(1)

- 特にcreatorについて、別名をjunii2に出力しない。(IRでのデータの持ち方は・・・)
  - 書誌情報として誤り、名寄せができない

(junii2ガイドライン作成の出発点)



# 作成者, 公開者, 寄与者(2)

## DRIVERガイドライン2.0

逆の語順の名前を使用し、構文は「姓,△頭文字(名)△接頭辞」とする。

たとえばJan Hubert de Smit は次のようになる:

```
<dc:creator>Smit, J.H.(John) de</dc:creator>
```

## Using Dublin Core - The Elements

Creators should be listed separately, preferably in the same order that they appear in the publication. Personal names should be listed surname or family name first, followed by forename or given name. When in doubt, give the name as it appears, and do not invert.

# 主題

- subjectは著者キーワード。(複数ある場合は繰り返す)
- NDC推奨としたのは、検索よりもむしろ主題統計などのため。

# 資源タイプ, フォーマット

- type・・・NIIttype以外の資源タイプをjunii2に出力した例は少数。
- format・・・MIMEメディアタイプ。ファイル拡張子から自動的に分かるので、機械的に解決することが望ましい。目録の資料形態(phys)ではない。

# 内容記述

- IR内では抄録と注記は別々に定義された項目で管理したほうがよい。
- IR内では学位論文情報の論文/課程の種類等や会議発表の会議名等は必要に応じて項目を定義したほうがよい。
  - ETD-MSなど、別フォーマットにも対応しやすい。  
メタデータは常に変換可能にしておく。

# 資源識別子

- URIは一度登録したら変えてはいけない。どうしても変えなければいけない場合は、URIはそのままリダイレクトされる仕組みを。
- fullTextURLはjunii2の特徴のひとつ。IRで管理されたコンテンツのURL。時間的地理的制約を加えずに公開されたもの。

# 著者版フラグ

- 本文(全文)コンテンツがある。
  - セルフアーカイブされた著者原稿 **author**
  - それ以外 **publisher**
- 本文がない。コンテンツが要旨などで本体でない。
  - **none**

# 掲載誌情報(1)

- 掲載誌情報が詳細に定義されているのも junii2の特徴。OpenURLでの解決、Google Scholarタグなどに対応できるなど、citation情報を正しく提供できるメリットがある。

# OpenURL(1)

## KEV Journal Formatの例(部分)

BaseURL?.....

&rft\_id=info:doi/10.1126/science.275.5304.1320

&rft\_id=info:pmid/9036860

&rft.issn=00368075

&rft.eissn=10959203

&rft.atitle=Isolation of a Common Receptor

&rft.jtitle=Science



# OpenURL(2)

&rft.aulast=Bergelson

&rft.aufirst=Jeffrey

&rft.date=1997

&rft.volume=245

&rft.issue=5304

&rft.spage=1320

&rft.epage=1323

rft・・・参照される対象

## 掲載誌情報(2)

- ISSNの-(ハイフン)はあってもなくてもよい。"X"は大文字が本来正しい。
- 合併号の区切り記号は "/" or "-"。OpenURLでの解決のため、CrossRefやPubMedのメタデータにあわせて"- "に。
- ボーンデジタルでページ数がない出版物のときは、記事番号(**artnum**)をspageを記入し、このときepageは使用しない。

# 日付

- 出版日付はdateofissuedを使用して、dateは主にIR登録(公開)日に使用。(DRIVERガイドライン2.0ではdateは出版日)
- 変更日、授与日、提出日、受理日・・・IR内部で管理される日付は、いろいろある。

# 言語

- ISO-639-2を推奨。

例:ドイツ語

ISO-639-1	de
ISO-639-2 (B)	ger
ISO-639-2 (T)	deu
ISO-639-3	deu (現代ドイツ語)

# 関係、ID、版に関する情報

- URI形式で使用する。
- DOI
  - URI・・・<info:doi/10.1000/7>
  - リンク先・・・<http://dx.doi.org/10.1000/7>
- PMID
  - URI・・・<info:pmid/19038271>
  - リンク先・・・  
<http://www.ncbi.nlm.nih.gov/pubmed/19038271>

# 情報源、範囲

- 論文がIRの主要なコンテンツの現状では、ほとんど使用されることはない。

# 権利

- 権利者からの指示がある場合、それに従うこと。  
→SHERPA/RoMEO, SCPJ参照

# 今さらながら・・・

- ガイドラインなどいちいち気にしなくても、正しいメタデータが出力できるIRシステムが理想。  
→ 著者によるセルフアーカイブのシステム
- 機械にできることは機械に。機械だけではできないことを人間がやればよい。  
→ IR担当者の役割



# junii2が直面する課題

- junii2策定当初から、“creatorについてNIIがハーベスト後に名寄せできるように”という意識はあった。
- CiNii著者名寄せにより問題が顕在化。ガイドラインでは1作成者1回のみ出力。
- 現在のjunii2では、別名や所属などを構造的に扱えない。

# NII実務研修(H22)での課題検討

- 検討1: junii2拡張
  - IRの著者に典拠機能
  - 機関内の研究者について典拠コントロール
- 検討2: junii3(仮)
  - RDFによるLinked Dataの実現
  - OAI-ORE(?)

# 研究者IDによる名寄せ

- H23年度WG2タスク2とCSI領域2プロジェクト「オープンアクセス環境下における同定機能導入のための恒久識別子実証実験」(代表:金沢大学)で進行中。
- junii2のバージョンアップ。

```
<creator id="{著者のURI}">姓, 名</creator>
```

研究者リゾルバーで解決



(おまけ) 静岡大学学術リポジトリのメタデータ

# SURE: 静岡大学学術リポジトリ

- メタデータを”いかに悩まないで済ませるか”にこだわっている。
- なるべくIDを用いてメタデータを管理
  - 業務を簡素化
  - データの正確性を向上
  - データ管理の効率化
  - webの世界へつながっていく

# IDをつかう(アイテム)

- DOI <http://dx.doi.org/{DOI}>
- PMID <http://www.ncbi.nlm.nih.gov/pubmed/{PMID}>
- NAID <http://ci.nii.ac.jp/naid/{論文ID}>
- NCID <http://ci.nii.ac.jp/ncid/{書誌ID}>
- KAKEN(P) <http://kaken.nii.ac.jp/p/{課題番号}>
- NDL-OPAC <http://opac.ndl.go.jp/recordid/{書誌ID}/>

# IDをつかう(個人)

- IRにとって重要なのは、その機関の研究者を網羅的に扱えること

- ORCID
- ResearcherID
- ISNI
- ReaD
- KAKEN(R)

機関内の研究者すべてを  
表現できるか？



UUIDで典拠テーブルを  
作成し、各IDをリンク

# IDをつかう(団体)

- 主に出版者に
  - SHERPA/RoMEO
  - SCPJ
  - NACSIS-CATのNAME



デモ

# おわりに

IR担当(あるいは大学図書館)にとって、メタデータを考えることは、これまでの図書館が積み重ねてきた資料組織法を、今後ますます発展していくweb上での学術情報に、どのように適用させサービスしていくのか？を考えることでもあると思います。