

平成21年度
学術ポータル担当者研修



機関リポジトリの公開

2009年8月5日-7日, 9月9日-11日

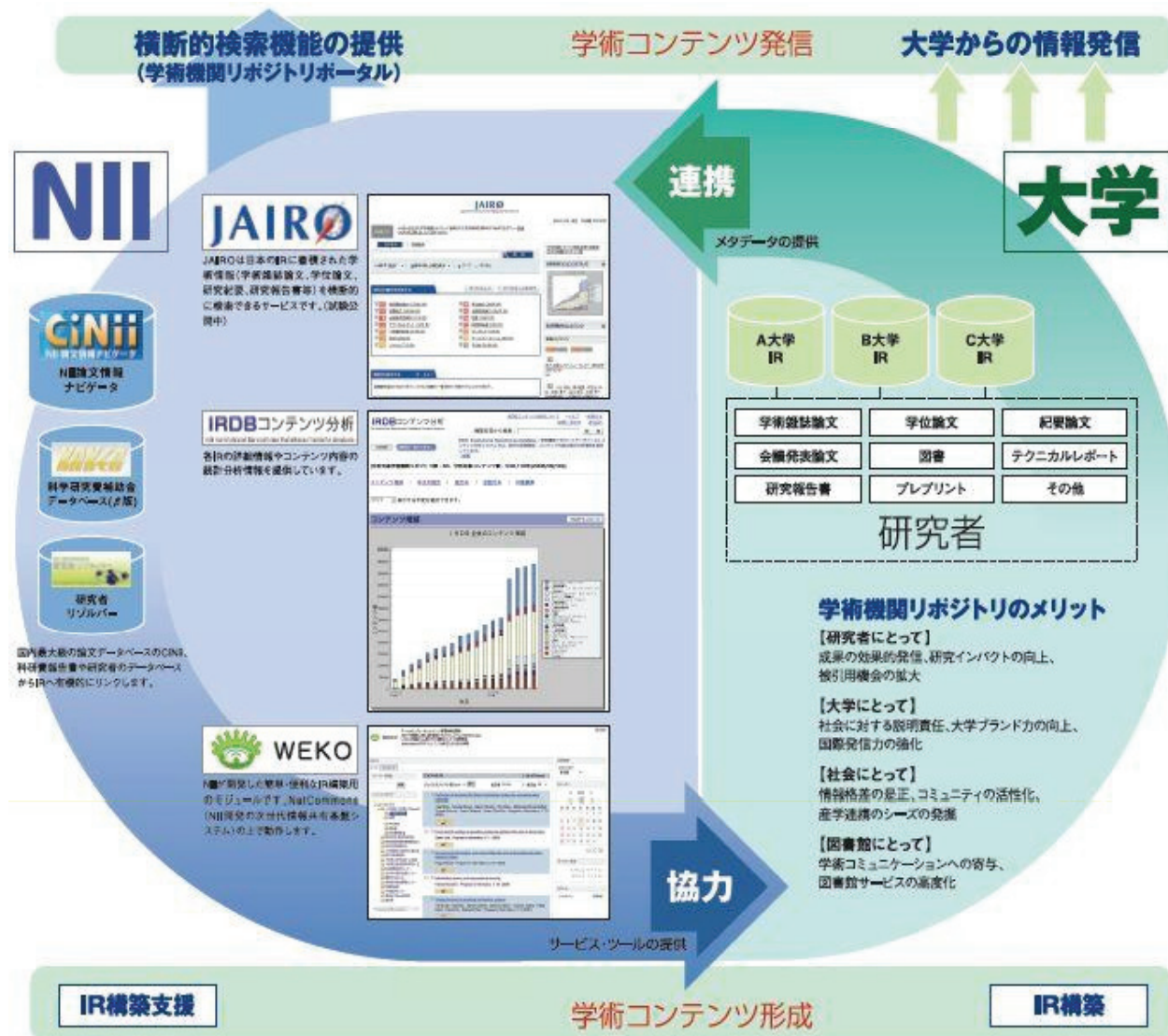


国立情報学研究所 学術基盤推進部 学術コンテンツ課
図書館連携チーム (IR/SPARC担当)

杉田いづみ, 内藤裕美子

<http://www.nii.ac.jp/irp/>
<http://www.nii.ac.jp/sparc/>
ir@nii.ac.jp

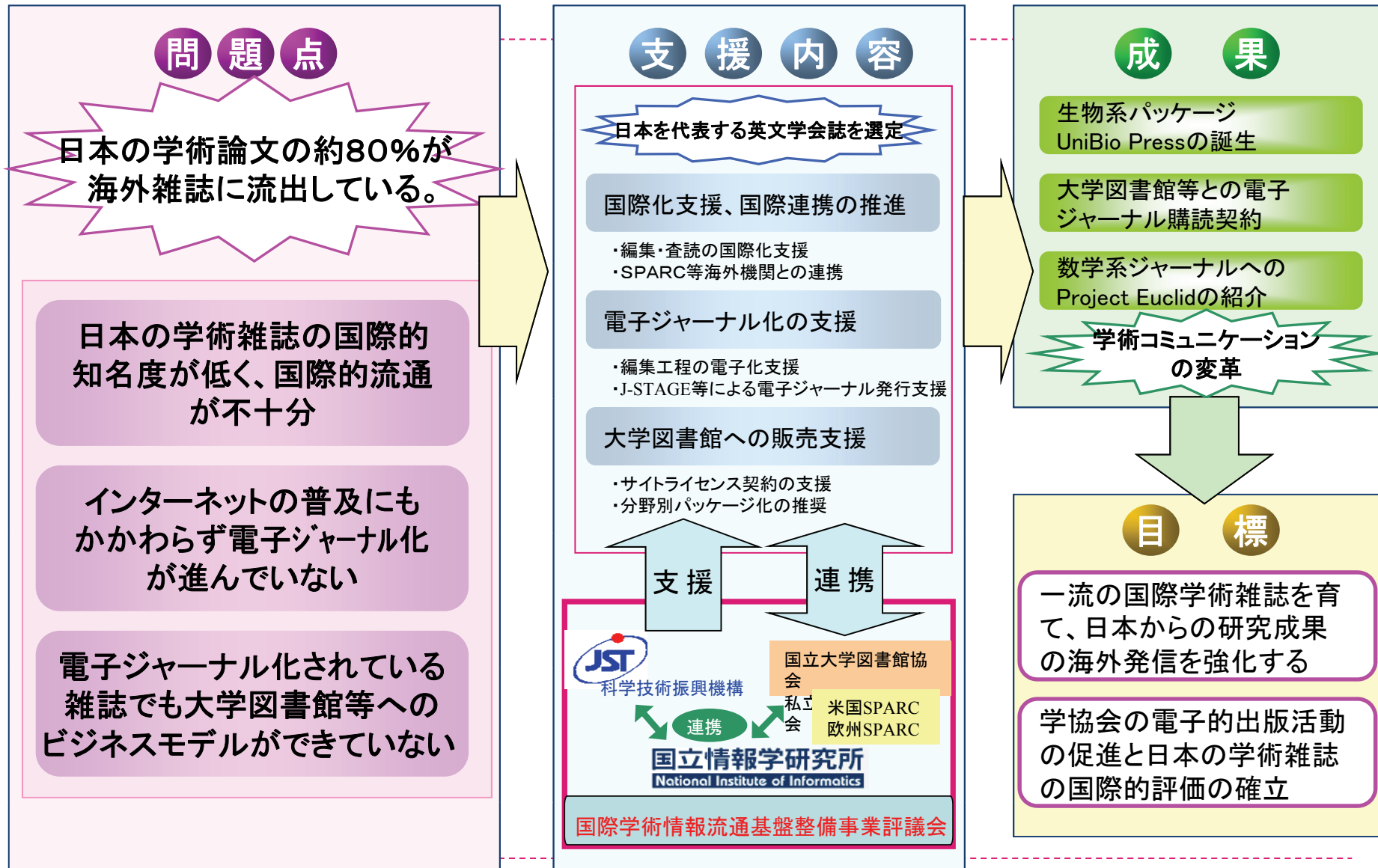
学術機関リポジトリ構築連携支援



NII学術機関リポジトリ構築連携の3つのサポート

- コンテンツ構築**
 - NII-ELSで電子化した研究紀要のコンテンツ提供（メタデータ+本文PDF）
実績：60機関，約600タイトル，約13万論文を提供
 - 国内学会誌本文コンテンツ（CiNii/ELS）の機関リポジトリに対する包括的利用許諾
 - SPARC JAPANパートナー誌へのリポジトリ・フレンドリ・ポリシーの推奨
- システム連携**
 - 国内学術機関リポジトリポータル”JAIRO”の提供
 - KAKEN-CiNii-IR連携関連文献のリンク形成
- コミュニティ形成**
 - 報告交流会、研修等の実施

国際学術情報流通基盤整備事業



このコマのねらい

▶ 研修概要によると

- ▶ 機関リポジトリ公開後，サービスプロバイダ（CiNii, JAIRO, AIRway, Google等）を使って集客する方法について講義を行う。
- ▶ また，サービスプロバイダとしてデータプロバイダに注意してもらいたい点（特にURIの永続性，削除履歴の保持）も述べる。

DP: Data Provider

2日目
「機関リポジトリ
のメタデータ概論」

メタデータ？

SP: Service Provider

3日目
「機関リポジトリ
の公開」

- ▶ 『メタデータ概論』のテキストを引用しつつ進めます

このコマの流れ

- ▶ 機関リポジトリ公開後の集客方法
 - ▶ クローラとハーベスタ
 - ▶ データプロバイダとサービスプロバイダ
 - ▶ サービスプロバイダに存在を知ってもらう
 - ▶ OAIデータプロバイダリストに登録しよう
 - ▶ 各種ディレクトリに登録しよう
 - ▶ NIIからハーベストされよう
- ▶ サービスプロバイダからデータプロバイダへのお願い
 - ▶ 失敗例に学ぶ重要ポイント
 - ▶ IRDBハーベストのシステム要件 まとめ
 - ▶ その他気をつけたいこと

クローラとハーベスタ

ビジビリティアップの基本

- 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。

– 構築後の設定でも十分対応可能

文献の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くれて、自サービスにデータ登録し、サービス展開してくれるサービスプロバイダー (JAIRO等) がある。

– 最初の設定が重要!

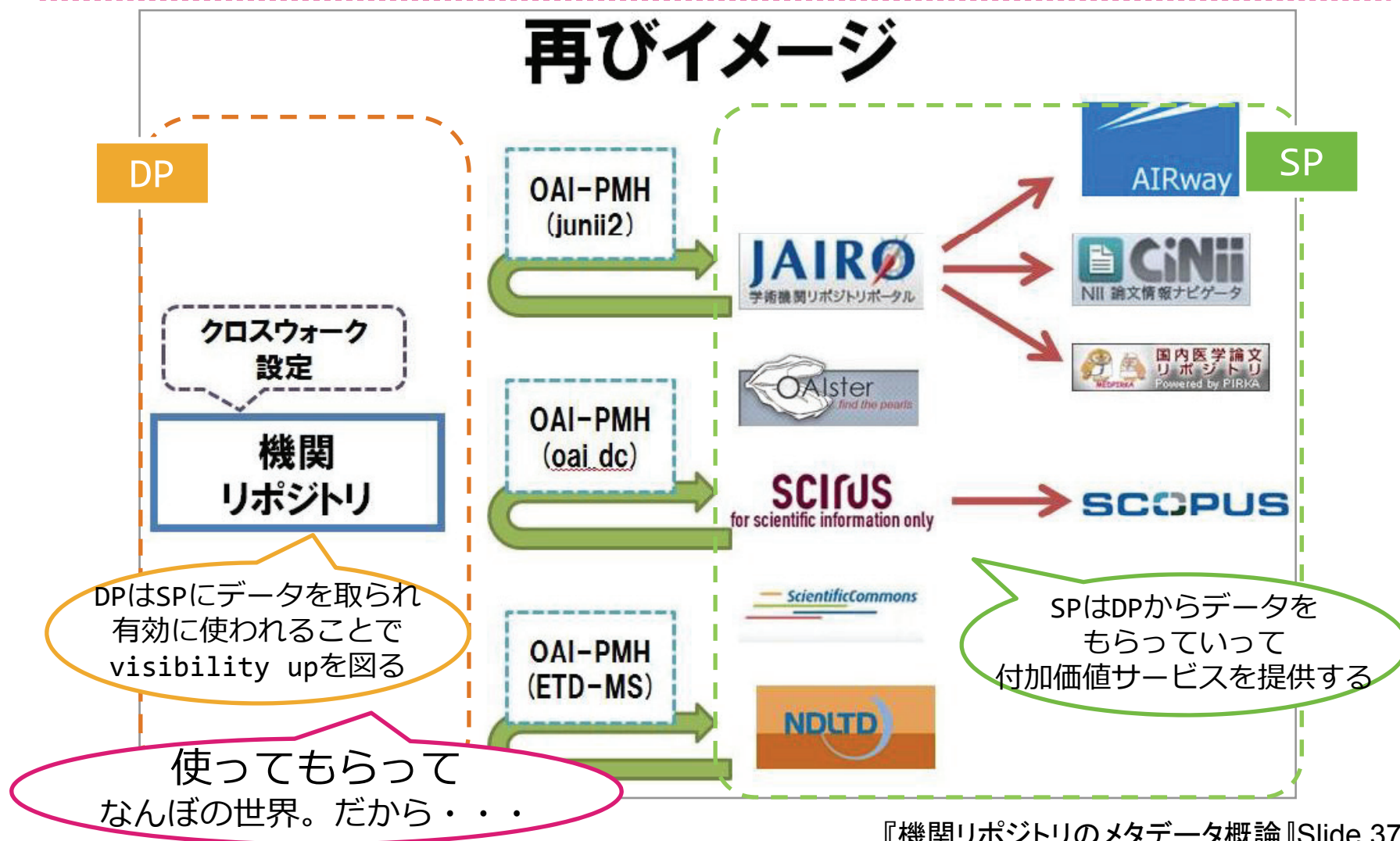
→ **今日はこちらの話**

クローラの話も少し

ハーベスタの話を中心に

『機関リポジトリのメタデータ概論』Slide.6

データプロバイダとサービスプロバイダ



『機関リポジトリのメタデータ概論』Slide.37

サービスプロバイダに存在を知ってもらう

- ▶ 公開の準備ができたなら
 - ▶ サービスプロバイダ等に登録依頼しましょう！

- ▶ まずは本家に
 - ▶ Open Archives Initiative

いちばん大事！
OAIデータプロバイダリストに登録する
→ちゃんとハーベストできるか
テストもできる登竜門的サイトです

- ▶ 業界への広報のために
 - ▶ OpenDOAR
 - ▶ ROAR

たしなみとして大事！
いろいろな所で使われています

IRを公開する
とは、このリストに
載せること

- ▶ 利用の拡大のために
 - ▶ JAIRO
 - ▶ OAIster

やっぱり大事！
国内ではJAIRO
国際的にはOAIster

具体的な登録方法は以下のサイトを参照

<http://drf.lib.hokudai.ac.jp/drf/index.php?startup>

OAIデータプロバイダリストに登録しよう-I

機関リポジトリを立ち上げたとき

どことどこに申し出るとよいか。

必須⁺

- OAIデータプロバイダリスト

IRを立ち上げて最初にすべきことは、OAIデータプロバイダリストへの登録です。これに載ると、世界中のサービスプロバイダがハーベストしにきてくれます。IRを公開するとはこのリストに載せることです。

OAIデータプロバイダリスト <http://www.openarchives.org/Register/BrowseSites>
方法は、

Registering as a Data Provider <http://www.openarchives.org/data/registerasprovider.html>

の一番下に、IRのOAI-PMHのベースURLを入れて送信するだけです。

(機械的なテストののちに登録されます。テストのみを受けることもできます。OAI-PMHとかベースURLとかわからなければ、導入業者にやってもらってもよいと思います)

具体的な登録方法は以下のサイト(DRF)を参照

<http://drf.lib.hokudai.ac.jp/drf/index.php?startup>

Startupのページから登録サイトにリンクあり

<http://www.openarchives.org/data/registerasprovider.html>

利用の拡大のために⁺

- Google

サイトの登録/削除 <http://www.google.com/addurl/?hl=ja&continue=/addurl>
からクロールを申し込みましょう。

- Google Scholar

「出版社向けサポート」 <http://scholar.google.com/intl/ja/scholar/publishers.html>
を参照のこと。「図書館向けサポート」ではありません。

- OAIster

「Steps to Becoming a Data Contributor」 <http://www.oaister.org/dataproviders.html>
を参照のこと。メールで申し込む方式に変わっています。

- JAIRO(NII)

方法は
Registering as
a Data Provider
の一番下にIRの
OAI-PMHのBaseURLを
入れて送信するだけ

OAIデータプロバイダリストに登録しよう-2

Registering as a Data Provider

<http://www.openarchives.org/data/registerasprovider.html>

試しにテストしてみよう！
確認するだけ(後で登録するかも)

Registering Your OAI-PMH 2.0 Repository

If you haven't read this page carefully, please go back and do so. If you are ready to register then enter your [baseURL](#) in the text box below and "submit". You'll receive [confirmation](#) shortly.

Validate only (may register later) Register this site

Please enter repository's base URL:

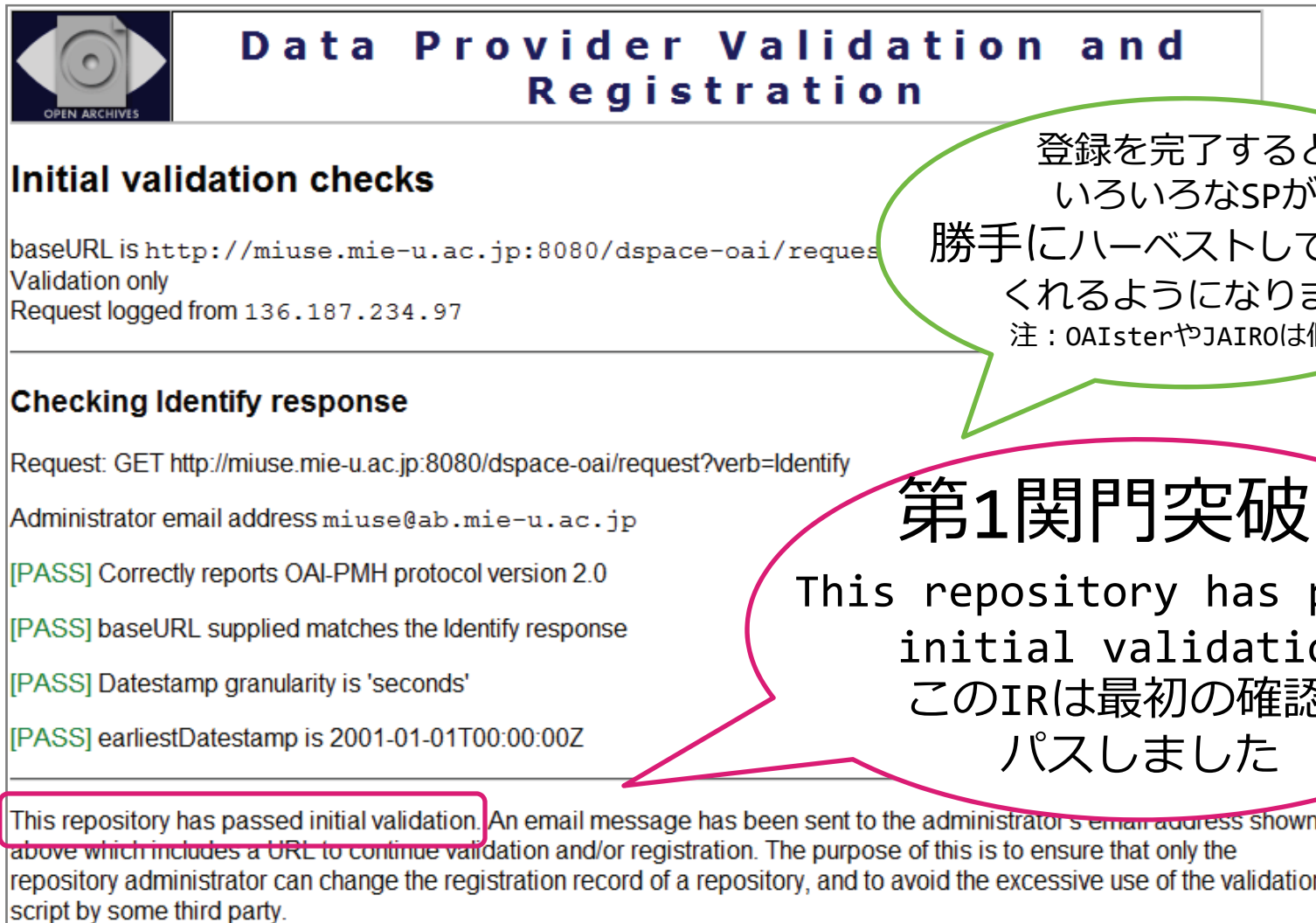
submit


Submitボタンをクリック！

BaseURL ?

SPからハーベストの要求を受付ける場所
設定によって異なる

OAIデータプロバイダリストに登録しよう-3



 **Data Provider Validation and Registration**

Initial validation checks

baseURL is `http://miuse.mie-u.ac.jp:8080/dspace-oai/request`
Validation only
Request logged from 136.187.234.97

Checking Identify response

Request: GET `http://miuse.mie-u.ac.jp:8080/dspace-oai/request?verb=Identify`
Administrator email address `miuse@ab.mie-u.ac.jp`

- [PASS] Correctly reports OAI-PMH protocol version 2.0
- [PASS] baseURL supplied matches the Identify response
- [PASS] Datestamp granularity is 'seconds'
- [PASS] earliestDatestamp is 2001-01-01T00:00:00Z

This repository has passed initial validation. An email message has been sent to the administrator's email address shown above which includes a URL to continue validation and/or registration. The purpose of this is to ensure that only the repository administrator can change the registration record of a repository, and to avoid the excessive use of the validation script by some third party.

OAIデータプロバイダリストに登録しよう-4

We received a request to validate (but not register) the following repository:

Base URL: <http://miuse.mie-u.ac.jp:8080/dspace-oai/request>
adminEmail: miuse@ab.mie-u.ac.jp

確認メールが
届きます

If you wish to continue the validation tests, please confirm by accessing the following URL:

<http://www.openarchives.org/Register/ValidateSite?c=WWSQ78QN>

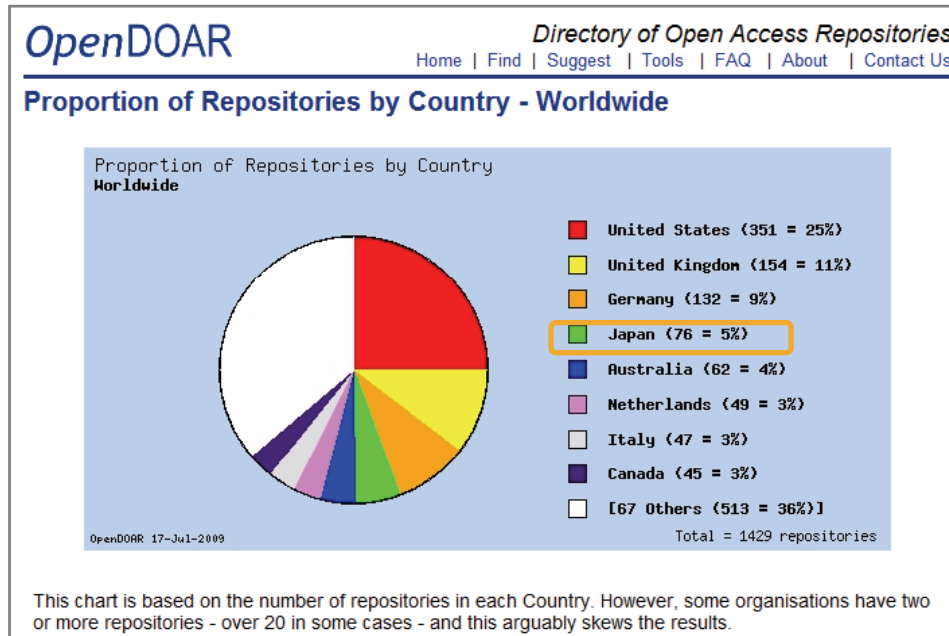
You need take no futher action if you do not wish to continue with the validation tests. The request was rec
event of difficulties please cont

続けるなら所定のURLにアクセスし
画面の指示に従って登録。
(続けなければ何もなくてOK)

The OAI validation service is sup,
0127308.

登録した結果 : <http://www.openarchives.org/Register/BrowseSites>

各種ディレクトリに登録しよう-I OpenDOAR



OpenDOAR Directory of Open Access Repositories
Home | Find | Suggest | Tools | FAQ | About | Contact Us

Search or Browse for Repositories

Recent Additions RSS1 Feed

Any Subject Area Any Content Type Any Repository Type
Any Country Any Language Any Software Search

Charts 20 per page Sort by: Repository Name New Query

To search the contents of the repositories listed in OpenDOAR, please see our [trial search page](#).

OpenDOAR Charts - Worldwide

Click on a thumbnail image to display the full-size chart and its key.

- Repositories by Continent
- Repository Organisations by Continent
- Repositories by Country
- Repository Organisations by Country
- Usage of Open Access Repository Software
- Open Access Repository Types
- Repository Operational Statuses
- Most Frequent Content Types
- Most Frequent Languages
- OpenDOAR Subjects
- Metadata Re-use Policy
- [Full Text] Data Re-use

Japan (76=5%)
世界第4位！でも
日本には100以上IRがあるはず

[University of Nottingham, UK](#)

Directory of Open Access Repositories (OpenDOAR)
申込画面 (Suggest a repository) <http://www.opendoar.org/suggest.php>

各種ディレクトリに登録しよう-2 ROAR

Browse Japan (ROAR) resulted in 70 repositories

View this result set as a chart: Export this result set as: [Comma-Separated Format](#) [Google Earth Overlay \(Thumbnails / Bars\)](#) ... other formats

Academic Research Repository at Institute of Developing Economies (685 records)

Running **DSpace**, based in **Japan** and is registered as **Research Institutional or Departmental**
 Registered on 2006-08-15
 Cumulative deposits: 685 total [\[table\]](#) [\[graph\]](#) [\[Preserv Profile\]](#)
 Daily deposits in last year: 18 days of 1-9, 0 days of 10-99, 1 days of 100+ [\[table\]](#) [\[graph\]](#) (PNG format) [\[interactive graph\]](#) (requires SVG format support)
[OAI Interface: Identify List Metadata Formats List Sets](#) [\[harvest status\]](#)
 100% freely accessible fulltext (* estimate)

The Institute of Developing Economies (IDE) aims at the expansion of trade relations and promotes economic cooperation with developing countries and regions, conducting high quality research work to meet the social needs.

AMCoR (Asahikawa Medical College Repository)

Not registered in Celestial:
 Either the OAI-PMH interface isn't working or it is awaiting action by a ROAR editor

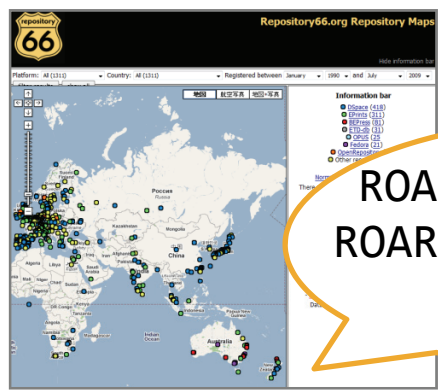
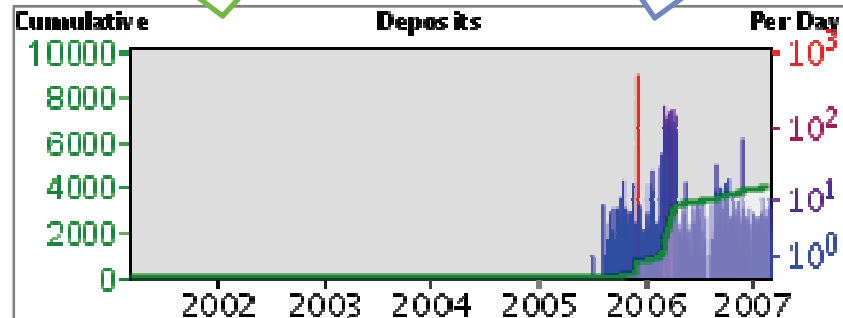
Running **Other softwares (various)**, based in **Japan** and is registered as **Research Institutional or Departmental**
 Registered on 2007-03-29
 No successful harvest yet.
[OAI Interface: Identify List Metadata Formats List Sets](#) [\[harvest status\]](#)
 100% freely accessible fulltext (* estimate)

Barrel (2527 records)

Running **DSpace**, based in **Japan** and is registered as **Research Institutional or Departmental**
 Registered on 2007-11-23
 Cumulative deposits: 2527 total [\[table\]](#) [\[graph\]](#) [\[Preserv Profile\]](#)
 Daily deposits in last year: 108 days of 1-

折れ線グラフ：
累積コンテンツ数

棒グラフ：
1日当たり登録数



ROAR登録国内IR = 70 →
ROARMAPやRepository66
などにも

Registry of Open Access Repositories (ROAR)

申込画面 (Register a Repository) <http://roar.eprints.org/index.php?action=add>

各種ディレクトリに登録しよう-3 NII-IRP機関リポジトリー一覧

国内の機関リポジトリー一覧

1. 北海道大学: [HUSCAP](#) / [EPrint Series of Department of Mathematics](#)
 2. 北海道教育大学: [北海道教育大学学術リポジトリ\(試験公開中\)](#)
 3. 室蘭工業大学: [室蘭工業大学学術資源アーカイブ](#)
 4. 小樽商科大学: [小樽商科大学学術成果コレクション Barrel](#)
 5. 帯広畜産大学: [帯広畜産大学学術情報リポジトリ](#)
 6. 北見工業大学: [北見工業大学学術機関リポジトリ KIT IR](#)
 7. 旭川医科大学: [旭川医科大学学術成果リポジトリ AMCoR](#)
 8. 弘前大学: [弘前大学学術情報リポジトリ](#)
 9. 岩手大学: [岩手大学リポジトリ](#)
 - (中略) -----
 107. 広島県大学共同リポジトリ: [HARP](#)
 108. 福井県地域共同リポジトリ: [ORFukui](#)
 109. 山口県大学共同リポジトリ: [維新\(試験公開中\)](#)
 110. 新潟県地域共同リポジトリ: [NIRR: Niigata Regional Repository\(試験公開中\)](#)
 111. 沖縄地域学リポジトリ: [ORION : Okinawa Repository Integrated Open-Access Network\(試験公開中\)](#)
- [日本における機関リポジトリ \(Open Access Japan\)](#)

2009/08/26時点で

111 IR!

IR名, URL等
(日英)を

ir@nii.ac.jp
にお知らせください

単独のIR以外に
共同リポジトリも
増えています

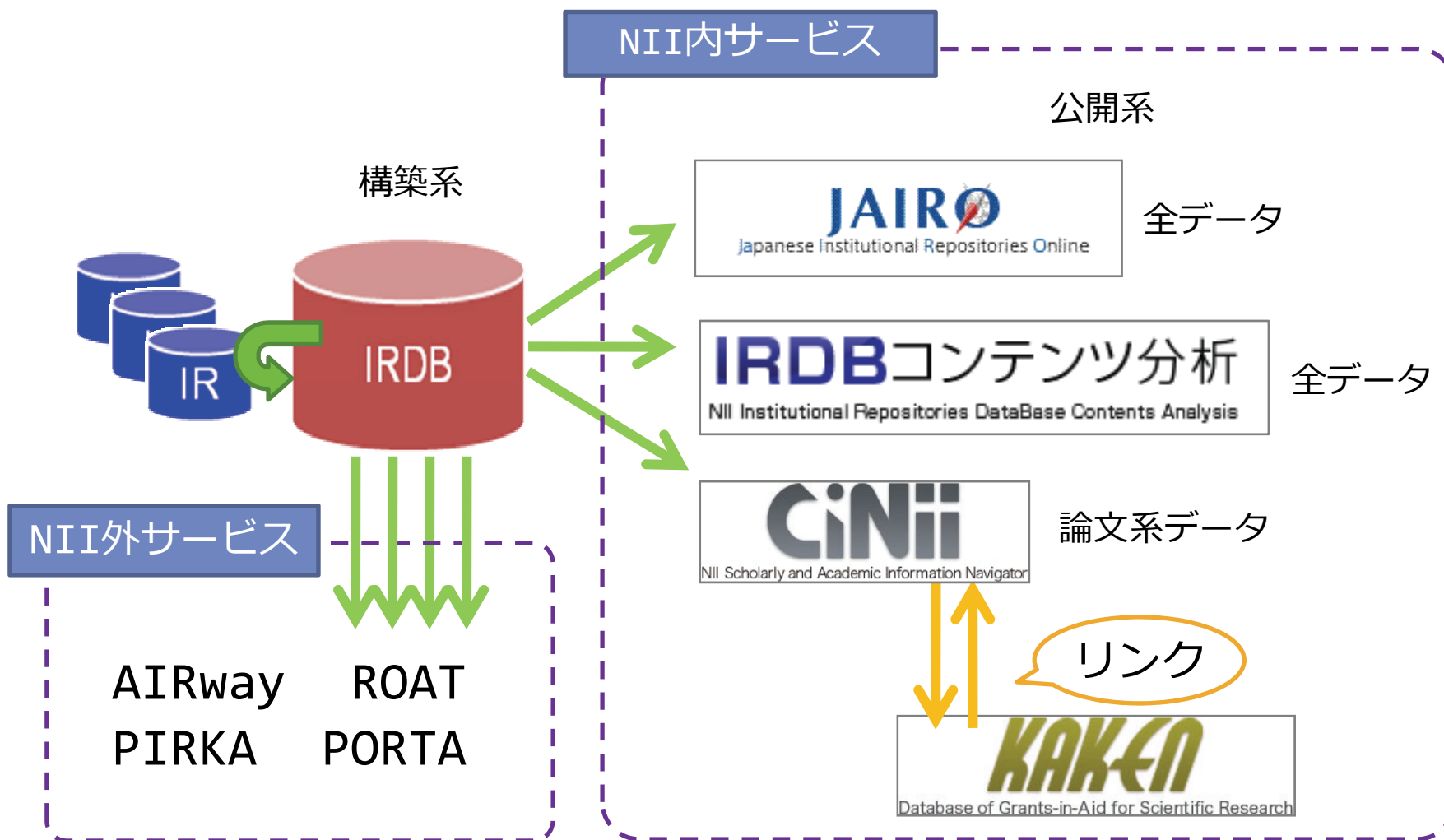
<http://www.nii.ac.jp/irp/list/>

NIIからハーベストされよう-I

▶ 何が嬉しいの？

- ▶ NIIからハーベストされると、国内のいろいろなサービスにメタデータが渡っていきます
 - ▶ NII内のサービス：
JAIRO, IRDBコンテンツ分析システム
CiNii, KAKEN...
 - ▶ NII外のサービス：
AIRway(リンクリゾルバのナレッジベース),
ROAT(機関リポジトリアウトプット評価システム),
PIRKA(双方向型医学系サブジェクトリポジトリ),
PORTA(国立国会図書館デジタルアーカイブポータル)

トピックス：NIIからハーベストされたデータの流れ



NIIからハーベストされよう-2

▶ どうすればいいの？

- ▶ 「国立情報学研究所 学術機関リポジトリデータベース 収集方針」に基づき、
- ▶ 「学術機関リポジトリデータベースへのデータ提供（申込書）」を担当までメールにてご提出ください。

▶ どんなふうにするの？

- ▶ Identifyコマンド(どんなIR?)では取得できない付加情報も収集し、付加価値サービスを提供します。
- ▶ NIIは特定のサイトからのみハーベストされています。
 - ▶ JAIROがDPとしてハーベストされると重複が生じるため、
「学術機関リポジトリデータベースからのデータ提供（申請書）」により申請があった場合のみ、データ提供します。

NIIからハーベストされよう-3

- ▶ 運用方式：単独／共同
- ▶ 基本情報：
 - ▶ IR名, URL, ソフトウェア
- ▶ 運用情報：
 - ▶ 機関名, BaseURL, ハーベスト頻度, 曜日, 設置種別, 公開日, 住所等
- ▶ 連絡担当者：
 - ▶ 所属, 氏名, 連絡先
 - ▶ 共同の場合
 - ▶ 個々の担当者/世話役担当者
- ▶ 通信欄：CiNii連携する？

| 学術機関リポジトリデータベースへのデータ提供（申込書） | |
|--|--|
| 「国立情報学研究所学術機関リポジトリデータベース収集方針」に基づき以下のとおり申し込みます。 | |
| 申込区分 | <input checked="" type="checkbox"/> 新規 <input type="checkbox"/> 変更 <input type="checkbox"/> 取消 |
| 運用方式 | <input type="checkbox"/> 単独 <input checked="" type="checkbox"/> 共同 |
| 【機関リポジトリ基本情報】 | |
| 機関リポジトリ名称（日）* | 国情リポジトリ |
| 機関リポジトリ名称（英）* | Kokujo Repository |
| 機関リポジトリURL（日）* | http://ir.kokujo.ac.jp/ |
| 機関リポジトリURL（英） | http://ir.kokujo.ac.jp/en/ |
| 使用ソフトウェア | DSpace |
| 【機関リポジトリ運用情報】 | |
| 機関コード | 9999 |
| 機関名称（日）* | 国情大学 |
| 機関名称（英）* | University of Kokujo |
| ハーベストBaseURL* | http://ir.kokujo.ac.jp/oi/request |
| SetSpec（共同の場合のみ） | hdl_123456789_13 |
| ハーベスト頻度* | <input checked="" type="checkbox"/> 週次：（火）曜日 ※水曜日以外 <input type="checkbox"/> 月次：（ ）日 ※1日～27日の間 <input type="checkbox"/> 年次：（ ）月（ ）日 |
| 機関規模 | D |
| 機関設置種別 | <input checked="" type="checkbox"/> 国立大学 <input type="checkbox"/> 公立大学 <input type="checkbox"/> 私立大学 <input type="checkbox"/> 短期大学 <input type="checkbox"/> 高等専門学校 <input type="checkbox"/> 大学共同利用機関 <input type="checkbox"/> その他（ ） |
| 郵便番号* | 101-8430 |
| 住所* | 東京都千代田区一ツ橋 2-1-2 学術総合センター21F |
| 試験公開日（YYYY/MM/DD） | 2007/2/1 |
| 本公開日（YYYY/MM/DD） | 2007/4/1 |
| 備考（日） | 機関リポジトリの紹介等 |
| 備考（英） | |
| 【機関リポジトリ連絡担当者情報】 | |
| 機関名・所属* | 国情大学附属図書館・機関リポジトリ担当 |
| 氏名* | 国情 太郎 |
| TEL*/FAX | 03-4212-2350 / 03-4212-2375 |
| E-mailアドレス* | kokujo-ir@kokujo-u.ac.jp |
| 【共同リポジトリ運営担当者情報】 | |
| 機関名・所属* | 〇〇大学附属図書館・機関リポジトリ担当 |
| 氏名* | 〇〇 太郎 |
| TEL*/FAX | 03-1234-5678 / 03-5678-1234 |

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

トピックス： NII内のサービス： JAIRO

▶ 日本のIRのコンテンツをまとめて探すなら

- ▶ ほぼ全てのメタ項目で検索可。
- ▶ 機関や資料種別で絞り込み可。
- ▶ 各IRのロゴを表示。
- ▶ よく利用されるコンテンツや新着コンテンツを表示。

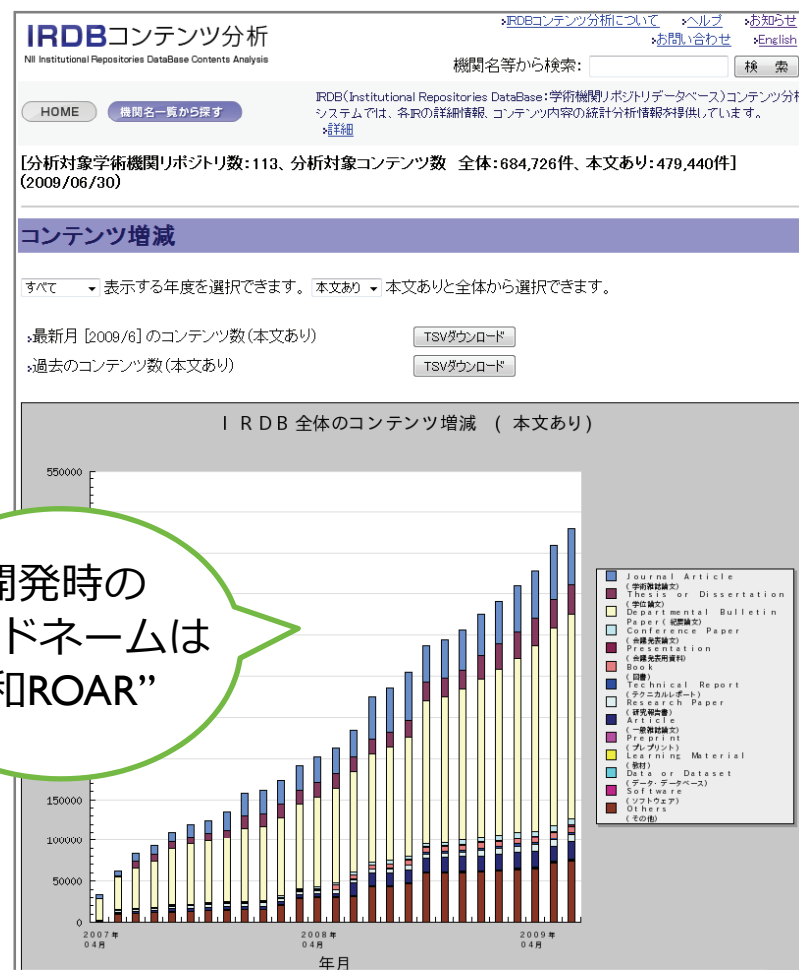
The screenshot shows the JAIRO website interface. At the top, it says 'JAIRO Japanese Institutional Repositories Online'. Below that, there's a search bar and a '検索' (Search) button. A sidebar on the right contains a bar chart and a section titled 'よく利用されるコンテンツ' (Popular Content). A green callout bubble points to this section, containing the text: 'Gyrocompass (羅針盤)のように 日本のIRの 現在と未来を 指し示したい'.

<http://jairo.nii.ac.jp/>

トピックス： NII内のサービス： IRDBコンテンツ分析システム

▶ 日本のIRの中身を知る

- ▶ 個々のIRの詳細情報がわかる。
IRのロゴや画面imageも登録できる。
- ▶ コンテンツ種別ごとの増減等が全体・機関別でみられる。
- ▶ IR担当者にバロメータとして活用してほしい！



<http://irdb.nii.ac.jp/analysis/index.php>

NIIからハーベストされよう-4

▶ 共同リポジトリで運用するなら

▶ 機関ごとにハーベストできるように設定したい

▶ SetSpecを分ける (DspaceやXoanIpsの機能)

□ 例-1 : 福井県地域共同リポジトリ (CRFukui)

- ・ハーベストBaseURL共通

`http://crf.flib.u-fukui.ac.jp/dspace-oai/request`

福井工高SetSpec=hd1_10461_2463, 福井工大SetSpec=hd1_10461_2458

▶ BaseURLを別にする (DSpaceのカスタマイズ)

□ 例-2 : 広島県大学共同リポジトリ (HARP)

- ・広島工業大学BaseURL

`http://harp.lib.hiroshima-u.ac.jp/oai-hit/request`

- ・広島市立大学BaseURL

`http://harp.lib.hiroshima-u.ac.jp/oai-hcu/request`

▶ BaseURLを別にする (仮想サーバ)

□ 例-3 : 岡山共同リポジトリ (O-AIR)

共同リポジトリ
については
領域2プロジェクト
ShaRe(シェア)のサイトも
ご参照ください

トピックス：共同リポジトリの事例 (H20報告交流会発表スライドから引用)

共同リポジトリ構築・公開まで

技術的なこと_2

共同リポジトリの参加機関をOAI-PMHのSetSpec毎に振り分ける選択ハーベスト(NIIの技術的サポート)

➡「JAIRO」において機関別にメタデータの登録が可能

1つのリポジトリシステムでありながら、あたかもそれぞれの機関がリポジトリシステムを立ち上げているかのように登録
(ある機関からの感想・・・面映ゆい感じもするが、誇らしくもある)

JAIRO機関一覧

| <input checked="" type="checkbox"/> は | |
|---|---|
| <input checked="" type="checkbox"/> 浜松医科大学 (742件) | <input checked="" type="checkbox"/> 一橋大学 (15,018件) |
| <input checked="" type="checkbox"/> 一橋大学(所蔵コレクション) (5,294件) | <input checked="" type="checkbox"/> 兵庫教育大学 (1,490件) |
| <input checked="" type="checkbox"/> 弘前大学 (1,491件) | <input checked="" type="checkbox"/> 広島経済大学 (955件) |
| <input checked="" type="checkbox"/> 広島工業大学 (187件) | <input checked="" type="checkbox"/> 広島国際学院大学 (79件) |
| <input checked="" type="checkbox"/> 広島国際大学 (29件) | <input checked="" type="checkbox"/> 広島市立大学 (247件) |
| <input checked="" type="checkbox"/> 広島女学院大学 (105件) | <input checked="" type="checkbox"/> 広島大学 (18,379件) |
| <input checked="" type="checkbox"/> 広島文化学園大学 (122件) | <input checked="" type="checkbox"/> 広島文教女子大学 (304件) |
| <input checked="" type="checkbox"/> 福井医療短期大学 (39件) | <input checked="" type="checkbox"/> 福井県文書館 (3件) |
| <input checked="" type="checkbox"/> 福井県立図書館 (32件) | <input checked="" type="checkbox"/> 福井工業高等専門学校 (230件) |
| <input checked="" type="checkbox"/> 福井工業大学 (1,458件) | <input checked="" type="checkbox"/> 福井大学 (1,200件) |
| <input checked="" type="checkbox"/> 福井大学 (1,047件) | <input checked="" type="checkbox"/> 文化女子大学 (96件) |

ハーベストの
申し込みも
機関ごと
に
お願いします

福井県
地域共同リポジトリ
(CRFukui) の
発表資料より



<http://www.nii.ac.jp/irp/event/2009/debrief/>

トピックス：NII内のサービス：CiNii-IR連携-I

▶ 論文という切り口で探す

▶ 目的

- ▶ CiNiiの本文到達率向上
- ▶ IRの利用機会向上

▶ 方法

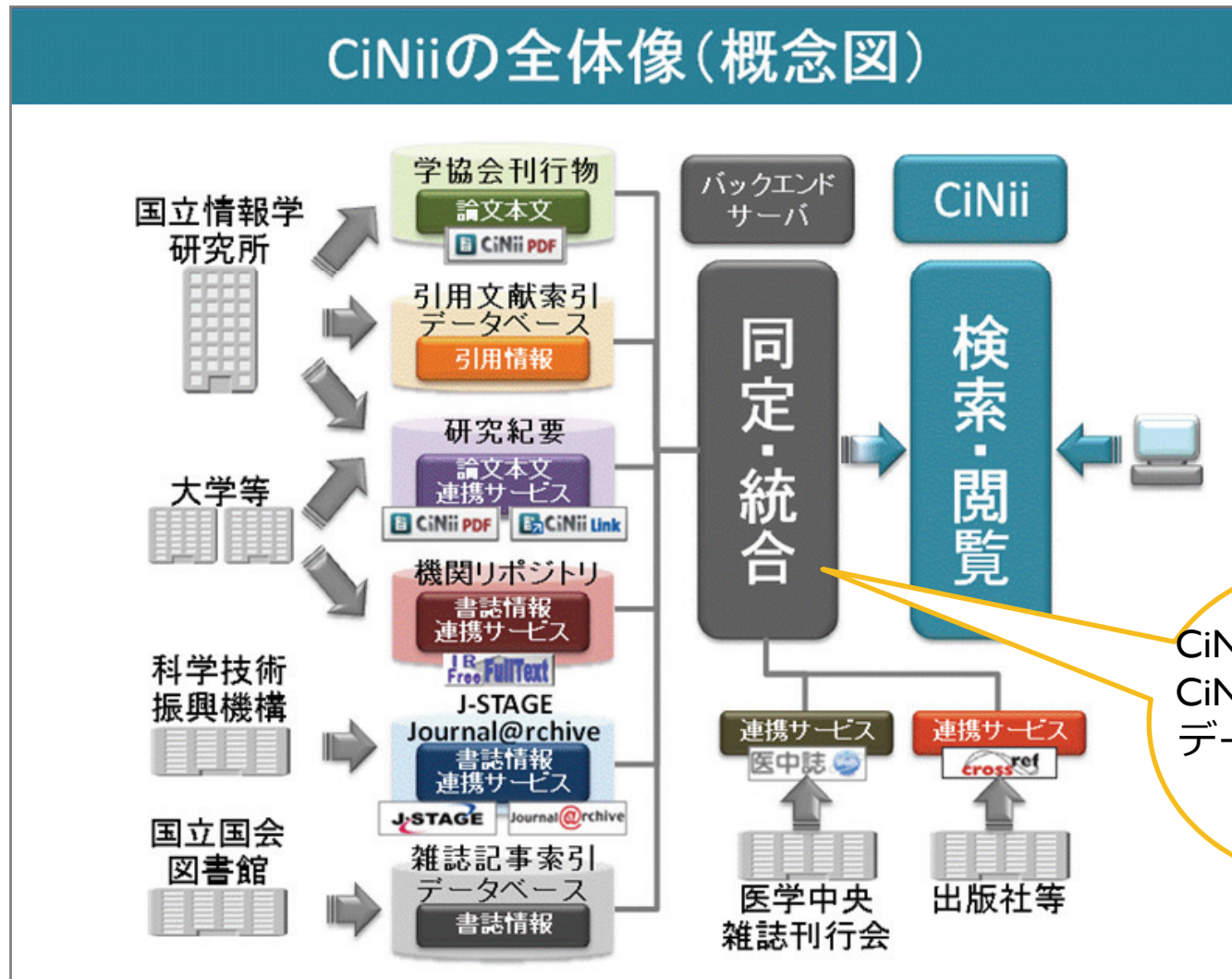
- ▶ 水曜日に以下の条件でメタデータを抽出同定

▶ 対象

- ▶ (1) NIItypeが論文系 (Journal Article, Departmental Bulletin Paper, Article) のもの
- ▶ (2) fullTextURLに値があるもの



トピックス： NII内のサービス： CiNii-IR連携-2



トピックス： NII内のサービス： CiNii-IR連携-3

□ データ数、増加数、更新頻度

| データベース名 | データ数 | 年間増加数 | 更新頻度 | 本文 | 料金 |
|--|-------------------------|-----------------------|-----------|----|----------|
| NII-ELS学協会刊 行物 | 約292万件 | 約18万件 | 日次 | ○ | 一部 有 |
| NII-ELS研究紀要 | 約86万件 | 約3.5万件 | 日次 | ○ | 無 |
| 引用文献索引 データベース | 書誌:約147万件 引用:約1593万件 | 書誌:約14万件 引用:約161万件 | 10回/年 | × | 引用 は有 |
| 雑誌記事索引 データベース | 約807万件 | 約40万件 | 週次 | × | 無 |
| 機関リポジトリ | 約23万件 | 不定 | 週次 | ○ | 無 |
| J-STAGE/ Journal@rchive | 約4万件 | 不定 | 数回/年 | ○ | 無 |
| CiNii合計* | 約1228万件 | 約70万件 | 日次 | | |

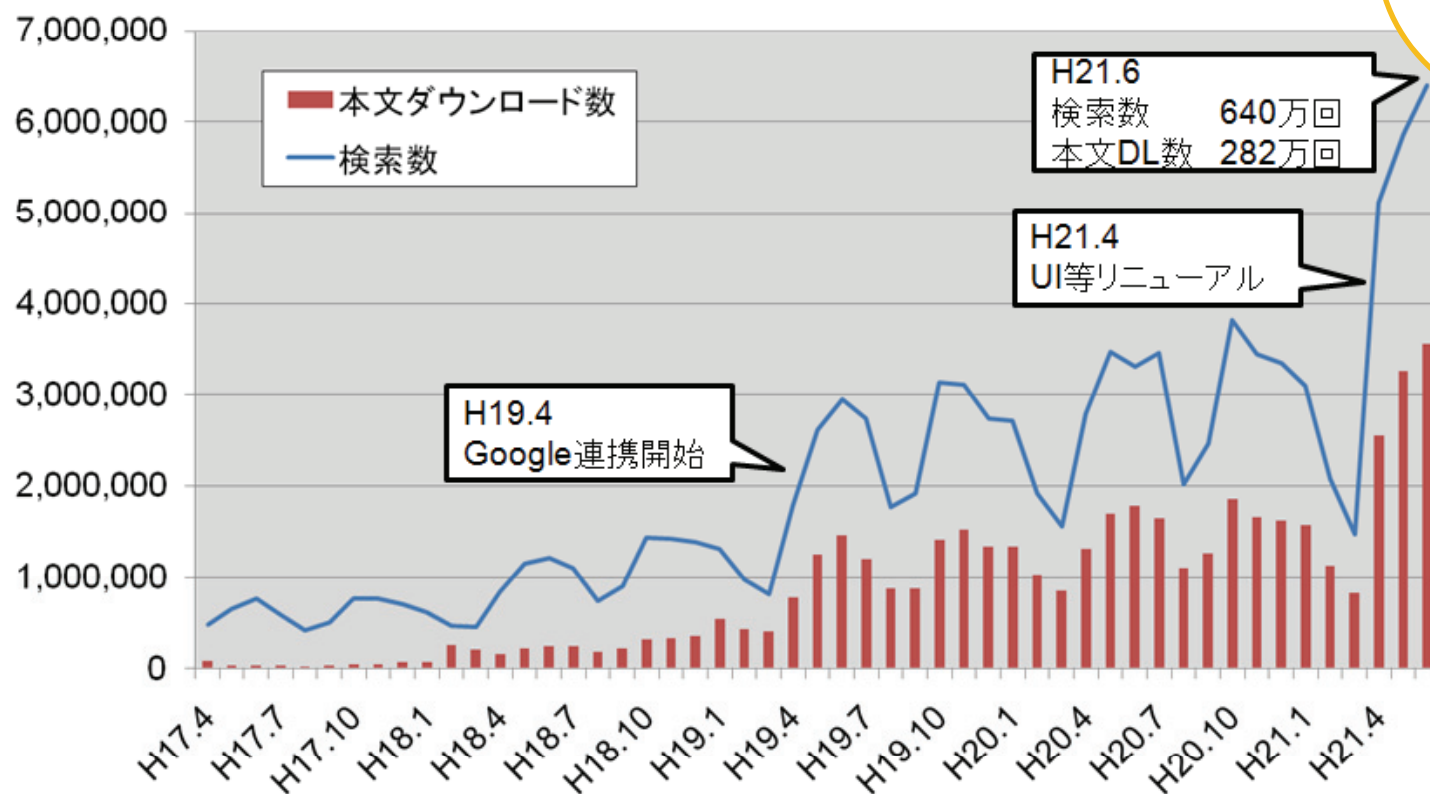
* 重複データが統合されるため、単純合計とは一致しない。

2009年6月末

トピックス： NII内のサービス： CiNii-IR連携-4

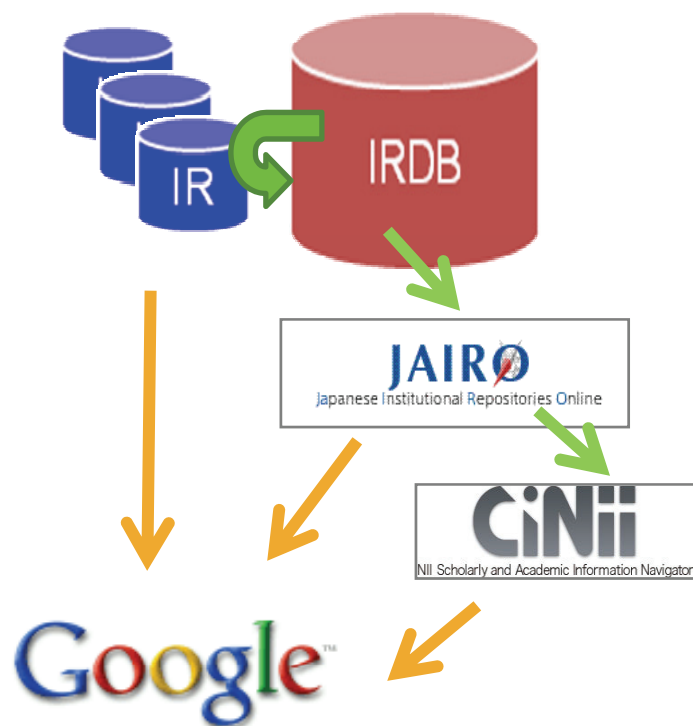
□ サービス開始(平成17年4月)からの利用回数の推移

- メインユーザ(大学生)の生活行動反映(夏季休暇中の低下など)
- 右肩上がりに増加



検索数は
JAIROの100倍以上
CiNii連携は
インパクト大

トピックス：Googleからクロールされるには-I



- ▶ Google
 - ▶ サイトの登録／削除
<http://www.google.com/addurl/?hl=ja&continue=/addurl>
からクローリングを申し込みましょう
- ▶ Google Scholar
 - ▶ 「出版社向けサポート」
<http://scholar.google.com/intl/ja/scholar/publishers.html>
を参照
 - ▶ 「図書館向けサポート」ではありません
- ▶ もっと有効にクロールされるために：
Googleサイトマップの登録
 - ▶ JAIROでサイトマップ登録済
 - ▶ CiNiiでも登録予定

トピックス：Googleからクローलされるには-2

- ▶ Googleサイトマップとは
 - ▶ Google サイトマップはウェブクロールシステム。
 - ▶ サイトマップを使用してGoogle のクローラにサイトの情報や場所を提供することで、クロールの精度と効率を高め、Google のインデックスにより多くのページを登録することができる。
 - ▶ サイトマップ形式のファイルをウェブサーバに置くことで、Google のクローラが、存在するページや最近更新されたページを認識し、その情報に従ってサイトをクロールすることができる。
 - ▶ Google サイトマップの基本的な使い方
 1. サポートされている形式（XML）でサイトマップを作成する。
 2. Google サイトマップ アカウントにサイトマップを追加する。
 3. サイトを変更した場合はサイトマップを更新する。

参考：Google Sitemaps 用XML自動生成ツール

<http://www.google-sitemaps.jp/>

トピックス：Googleからクローलされるには-3

▶ Googleサイトマップ用XMLのサンプル

```
<xml version="1.0" encoding="UTF-8">
<urlset xmlns="http://www.google.com/schemas/sitemap/0.84">
<url>
<loc>http://jairo.nii.ac.jp/0022/00009368</loc>
<lastmod>2009-08-02T06:34:07+09:00</lastmod>
</url>
<url>
<loc>http://jairo.nii.ac.jp/0022/00009368/en</loc>
<lastmod>2009-08-02T06:34:07+09:00</lastmod>
</url>
</urlset>
```

トピックス：NII外のサービス：AIRway (H20報告交流会ポスターから引用)

AIRwayとは

リンクリゾルバを使用し、文献データベースから電子ジャーナルだけではなく、機関リポジリに搭載された文献へ利用者を導く

ISI Web of Knowledge → SFX Services for this record → White Rose Research Online

AIRway平成20年度活動

(1)対応リンクリゾルバの拡大
ExLibris社SFXグローバルバージョンのソースターゲットに採用されました！

(2)対応機関リポジリの拡大

- ① IRDB(JAIRO対応全機関)をリンク解決対象に追加
- ② 国外15機関をリンク解決対象に追加

搭載レコード数約9万件 → 約27万件に大幅増加！

Repositories in AIRway

ただし以下の問題が浮き彫りに！

AIRwayの
リンク解決対象：
NIIからハーベスト
されているIRは
自動的に対象に

もう一つの
解決策
JEEPwayは
OAlsterが対象

トピックス：NII外のサービス：ROAT (H20報告交流会ポスターから引用)

ROAT[®]

Repository Output Assessment Tool
機関リポジリアウトプット評価システム

利用機関大募集!

お問合せ先：
千葉大学附属図書館
roat@ll.chiba-u.ac.jp

- 登録コンテンツ数が多くなくても、登録コンテンツの利用が多いリポジリはカッコいい!
※ ROATはリポジリに登録されたコンテンツの利用(アウトプット)回数による標準的な評価を目指します。
- となりのリポジリは何万アクセスと言っているけど、Google や Yahoo! などの巡回アクセスが多いらしい!
※ ROATは、そのようなアクセスを除去し、純粋な利用者によるアクセス数のカウントを目指します。
- リポジリが利用されているかは知りたいけど、システムをイジリたくない!
※ ROATは、プログラムをインストールする必要がありません。

次世代学術コンテンツ基盤共同構築事業 学術機関リポジリ構築連携支援事業
「機関リポジリ評価のための基盤構築」プロジェクト

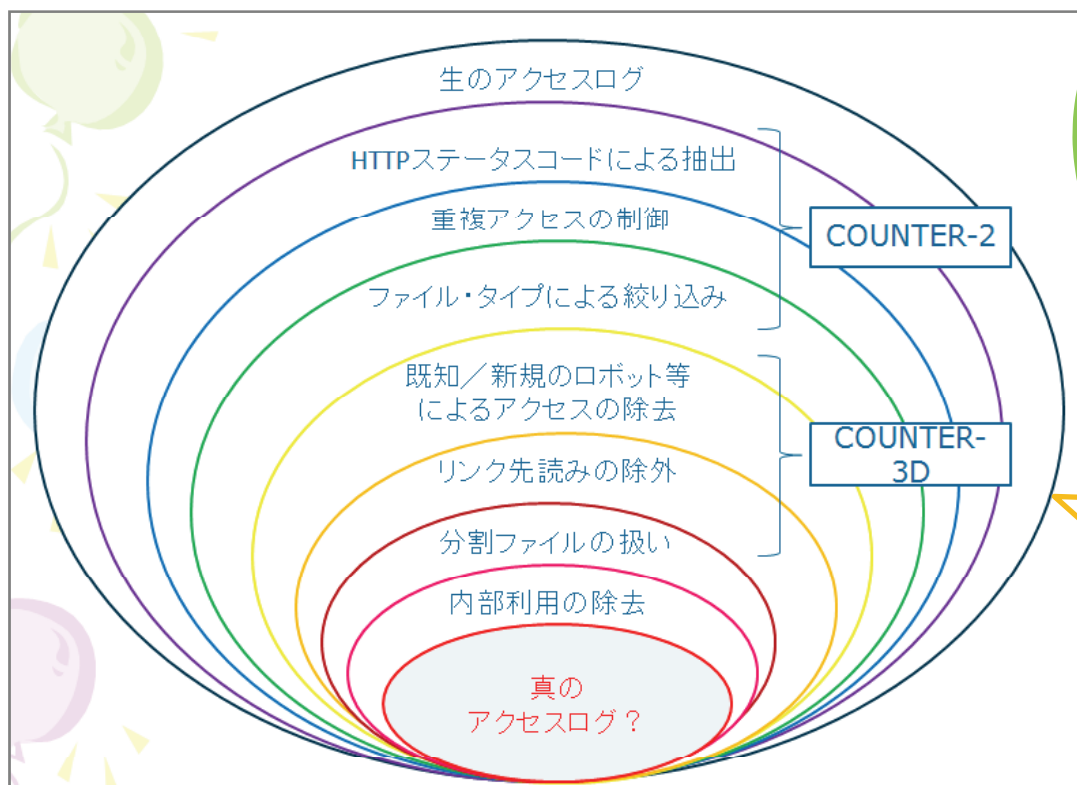
代表機関：千葉大学 / 連携機関：北海道大学, 東北大学, 金沢大学, 大阪大学

Webの
アクセスログと
JAIRO(IRDB)のメタ
データをマッチング

どのコンテンツが
どれだけアクセス
されたかを
統一的方法で
カウント

トピックス：NII外のサービス：ROAT (H19報告交流会佐藤義則先生発表資料から引用)

- ▶ 純粋な利用者によるアクセス
 - ▶ いわゆる精米(or玉ねぎの皮むき)



CSI委託事業を受託している機関にお願いしていること：IRで使用しているWebサーバ (apache, apache2, TomCat等) のアクセスログファイルを combined log format で蓄積／保存することを推奨

JAIROの統計もROATと同じくブラックリスト方式を取っています。

失敗例に学ぶ重要ポイント：アクセスの保証-I

▶ 機関リポジトリの信頼性 = アクセスの保証

▶ URIの永続性確保

▶ 失敗例-1：なんちゃってHandleの巻

- DSpaceをインストールしてコンテンツを登録してみた
- コンテンツに自動的に番号がついた

「このアイテムの引用には次の識別子を使用してください:

<http://hdl.handle.net/123456789/1>」

- クリックしたらエラーになる



CNRIが提供する
サービス。
年会費を払って
機関プリフィックス
を取得するか、
Handleを使わないよう
カスタマイズが必要。

失敗例に学ぶ重要ポイント：アクセスの保証-2

▶ 機関リポジトリの信頼性 = アクセスの保証

▶ URIの永続性確保

- ▶ 失敗例-2：データ一括置き換えでURIが変わっちゃったの巻
 - メタデータの構成を途中で変えることになった
 - 1件ずつ修正するのは非効率的なので、一括で置き換えることにした
 - 同一コンテンツを同一URIで登録できず全て変わってしまった

IRは
「Webサイトと違って
サーバが変わっても
URLが変わりません」
が売り文句の一つ！

JAIROなら
全件削除、
再ハーベストで
全件置き換えが可能
ですが・・・

CiNii連携している
場合、書誌の同定作業
の関係で、全件置き換えは
対応していません。

Googleに一旦クロール
されたデータは、キャッシュに
残ったりしてなかなか
削除されません。

失敗例に学ぶ重要ポイント：アクセスの保証-3

▶ 機関リポジトリの信頼性 = アクセスの保証

▶ 削除履歴の保持

- ▶ 失敗例-3：物理削除しちゃったの巻
 - 間違って登録したアイテムを削除して新しく登録したところJAIROで重複してしまった

OAI-PMHに対応すると...

- 機関リポジトリに搭載されている文献データを他のサービスプロバイダーが(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스팅」といいます

例えばJAIRO

- 自分たちからデータを送付する必要がない。
- リポジトリはデータを蓄積するのみ！
- **変更, 削除記録も考慮してくれる。**

SP側はメタデータに書かれてあることが全て。物理的にデータが削除されたらSPはデータを更新できない。結果、データが重複してしまう

データを削除する場合は、削除履歴を残す = 必ず「transient」か「persistent」にする

『機関リポジトリのメタデータ概論』Slide.9

IRDBハーベストのシステム要件 まとめ-I

▶ IRDBのハーベスタのリクエスト

▶ http :

```
//baseURL?verb=ListRecords&metadataPrefix=junii2&from=yyyy-mm-dd&until=yyyy-mm-dd
```

- ▶ junii2フォーマットで前回取得日付から今回取得日付までハーベスト
参照:コマンドとパラメタ 『機関リポジトリのメタデータ概論』Slide.14

▶ ハーベストの条件

- ▶ OAI-PMHの全ての仕様に準拠すること
- ▶ メタデータ・フォーマットとして junii2 をサポートすること
 - ▶ ハーベストされたメタデータは, URIにより同定され, baseURLもしくは SetSpec単位でデータベースに格納されます。
 - ▶ データベース上に同一のURIを持つメタデータが存在する場合は, 上書きします。

IRDBハーベストのシステム要件 まとめ-2

▶ URIを変更しない

- ▶ データの一括更新等でURIを変更してしまうケースが見られます。
- ▶ IRDBはデータを差分でハーベストしているため、URIを変更するとIRDB (JAIRO, CiNii) に同じメタデータが複数登録されてしまいます。
- ▶ 変更前のURIは各IRには存在しないので、JAIRO, CiNiiでリンク切れが発生します。

▶ メタデータを削除する場合は削除履歴を残す

- ▶ OAI-PMH2.0では、レコードの削除について3つのサポート基準（どれかを選択）を定義していますが、必ず「transient」もしくは「persistent」をサポートし、削除履歴を残すようにしてください。
 - ▶ no : 削除履歴を保持しない
 - ▶ persistent : 削除履歴を永続的に保持
 - ▶ transient : 削除履歴を保持するが永続性は保証しない

IRDBハーベストのシステム要件 まとめ-3

▶ CiNiiと連携する機関リポジトリ及びメタデータの条件

- ▶ (1) OAI-PMHの全ての仕様に準拠すること
 - ▶ (2) メタデータ・フォーマットとして junii2 をサポートすることの他に
 - ▶ (3) 差分更新に対応可能であること
 - ▶ (4) かつURIを変更しないことを確約いただく必要があります
- ▶ この条件はCiNiiからのリンク切れを防止するために設けさせていただいています。
- ▶ CiNii連携をしているIRのデータは、CiNiiで同定処理（機械処理・人手処理）を行っている関係で、全件置き換えはお受けしておりません。CiNii連携を希望する場合は特にご注意ください。

その他気をつけたいこと

- ▶ IRDBコンテンツ分析システムの各項目
 - ▶ 言語 "language"
 - ▶ NDC分布 "NDC"
 - ▶ 本文あり "fullTextURL"
- ▶ AIRwayで有効に活用されるために
 - ▶ 著者版フラグ "textversion" (特に論文系は要注意)
 - ▶ author:著者版
 - ▶ publisher:出版社版
 - ▶ none:上記以外
- ▶ CiNiiから正しくリンクするために
 - ▶ DOI "DOI"
 - ▶ 著者の別表記 "Creator"
 - ▶ . . .

「出力指針」検討中
ぜひご意見を
お寄せください!

さいごに

- ▶ 大切なのは
 - ▶ リポジトリのコンテンツのビジビリティを上げ、使われること。
- ▶ サービスプロバイダは
 - ▶ ビジビリティアップの手段として有効。
- ▶ 有効活用するには
 - ▶ 特性を知り、それに合わせたメタデータを出力すること。

こんなに**使われる**なら
リポジトリに登録しなきゃ損！

と言われるリポジトリを目指しましょう！そしてPRしましょう！

詳細情報は

学術機関リポジトリ構築連携支援事業
NII Institutional Repositories Program

HOME | ニュース | 事業について | 機関リポジトリ一覧 | 委託事業 | イベント情報 | ドキュメント | お問い合わせ

About Us
事業について

機関リポジトリとは、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です。
国立情報学研究所では、これまでのコンテンツ関連事業の成果を継承、拡充させ、次世代学術コンテンツ基盤の整備に資するために、各大学における機関リポジトリの構築とその連携を支援しています。

平成20年度CSI委託事業報告交流会(コンテンツ系)を平成21年7月9日(木)~7月10日(金)に開催しました。

News
ニュース

- 2009/07/16
平成20年度CSI委託事業報告交流会(コンテンツ系):開催報告を掲載
- 2009/06/29
平成20年度CSI委託事業報告交流会(コンテンツ系)発表者・発表タイトルの公開
- 2009/06/19
平成20年度CSI委託事業報告交流会(コンテンツ系)仮プログラムの公開
- 2009/06/19
平成20~21年度運用・実施に関するFAQ更新

Document
ドキュメント

- 基本文献
- コンテンツ提供
- 広報資料
- 委託事業各種様式
- 翻訳資料
- システム情報
- 調査・報告資料

- ▶ Webサイトをご覧ください
- ▶ お問い合わせ先
国立情報学研究所 学術基盤推進部
学術コンテンツ課 図書館連携チーム
(IR/SPARC担当)

ir@nii.ac.jp

システム情報は
ココ!

<http://www.nii.ac.jp/irp/>