

# 1.研究者リゾルバーとJAIRO著者名検索 2.KAKENデータベースの機能拡張

蔵川 圭

国立情報学研究所  
学術コンテンツサービス研究開発センター

第13回図書館総合展 2011/11/9(水)～11/11(金) 会場 パシフィコ横浜

# 1.研究者リゾルバーとJAIRO著者名検索

# 研究者リゾルバーの目的

## 基本機能

- 日本の研究者を対象としたWeb上の識別子
  - あわせて、日本の研究組織を対象とした識別子
- Web上の研究者リソースのリンクングハブ



研究者リゾルバー

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

## 取り扱う情報

- 研究者の基本情報と研究業績リストをベースとしたリンクマネジメント

## 応用

- Linked Data Webのための、日本の研究者と組織のURI表現
- 名前解決するWebサービス
- NIIの学術関連サービスへの密接なデータ連携
- Web上の名前典拠としての利用

登録主体による位置づけ

| 業績<br>研究者 | 機関によるとりまとめ                                   | 研究者によるとりまとめ                 | 機械による自動処理                |
|-----------|----------------------------------------------|-----------------------------|--------------------------|
| 機関による登録   | 研究者ディレクトリ<br>ReaD<br>IR<br>KAKEN<br>研究者リゾルバー | 機関にあるホームページ                 |                          |
| 研究者による登録  |                                              | ResearcherID<br>Researchmap |                          |
| 機械による自動処理 | Cinii 著者検索                                   |                             | Google<br>Google scholar |

3

# 研究者識別子スキーム

- 13ケタの番号によって研究者を表す
- 研究者リゾルバーIDと呼ぶ

ID (13ケタの番号)      説明

10000xxxxxxxxx

科研費研究者番号登録者のID割り当て  
xxxxxxxx は、科研費研究者番号 (8 桁)

200xxxxxxxxxxx

ReaD研究者コード登録者のID割り当て  
xxxxxxxxxxx は、ReaD研究者コード (10桁)

3xxxxxxxxxxxxxx

それ以外のID割り当て

- 一人の研究者に対して、複数の研究者リゾルバーIDが登録された場合、いずれかのIDを代表させて、エイリアスを張る

## URIによる表現

- <http://rns.nii.ac.jp/nr/xxxxxxxxxxxxxx>
- xxxxxxxxxxxxxxは研究者リゾルバーID

4

# 初期登録

- 呼び水として、科研費データベースKAKENのデータを利用して研究者を登録
- データ項目
  - 姓名(漢字、ヨミ、ローマ字)
  - 所属機関(最新)、部局、職名
  - 科研費研究者番号
  - (科研費研究課題の研究分野)
  - (科研費研究課題の研究キーワード)
- 2011/10/15現在、192,716件登録

5

# 研究者リソース

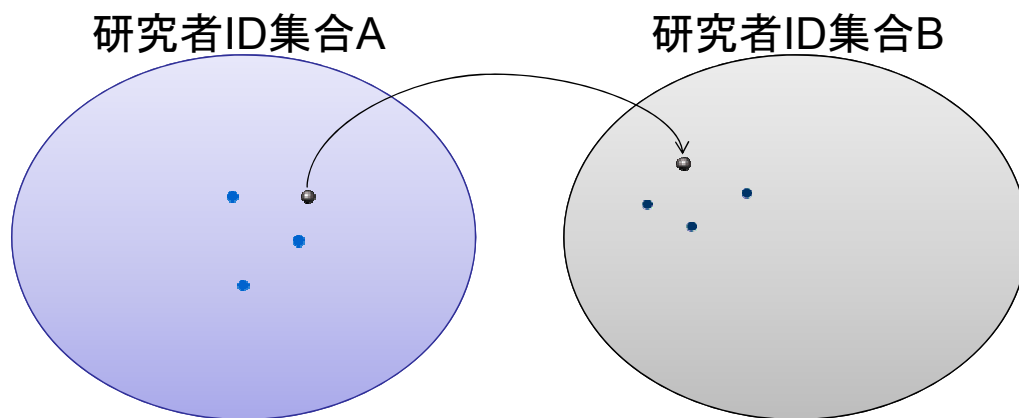
- Web上の研究者ID集合をもつデータベース  
たとえば、



6

# 異なる研究者リソース間のリンクング

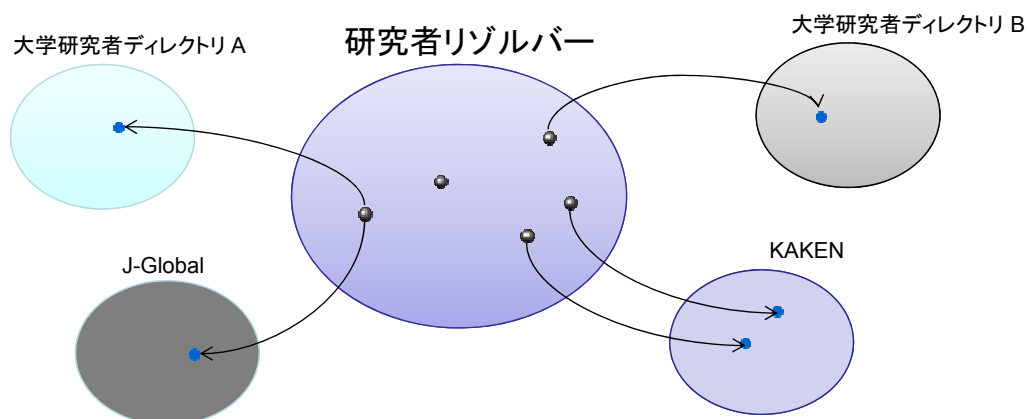
- 異なる研究者ID集合間のIDマッチング



7

# 研究者リソースリンクング

- 集合間のIDマッチング



8

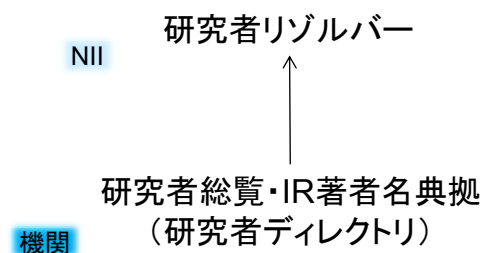
# クロールによるアプローチ

- 大学の研究者ディレクトリをクロール
  - 研究者ページのURLを特定
  - それぞれのURLに対して研究者名を特定
- ここでは簡単に、以下を同時に満たすとき、同一人物と判定
  - 漢字氏名の一致
  - 機関の一致
  - 機関内でユニーク名（同姓同名がない）
- クロールとリンクの実際
  - 226大学の研究者ディレクトリに対してクロール
  - 102,949件のURLを取得
  - 23,252件のリンク（リンク率23%、登録全体の12%）

9

# 機関のデータアップロードによるアプローチ

- データ
  - 研究者の基本情報と研究業績リストを収集
  - ResearcherID (Thomson Reuters)で利用されるデータアップロードXMLスキーマを日本語化して独自拡張
- 応用
  - 研究者リゾルバーIDと外部サービス識別子とのマッピングテーブルを構築
  - 外部の研究者リソースとの同定処理



10



# 検索リンク

The image shows a researcher's profile for 山中 伸弥 (Yamanaka, Shinya) with various IDs and affiliations. Below the profile, there is a list of search engines: Read, CINI, Google, CINI, Webcat Plus, Read, Google Scholar, and Google. Arrows point from the search engines to a list of search criteria: 検索(日本語姓名および所属による), 検索(日本語姓名による), and 検索(ローマ字姓名による). The list of search engines is also labeled with 検索(日本語姓名および所属による), 検索(日本語姓名による), and 検索(ローマ字姓名による). The page is from the NII National Institute of Informatics.

山中 伸弥  
Yamanaka, Shinya  
研究者ID: 1000010295694  
科研費研究番号: 10295694  
所属(2011年度のEAKENから): 京都大学/教授  
(注) この項目の値は科学研究費補助金データベースにおけるものらしい

検索(日本語姓名および所属による)  
検索(日本語姓名による)  
検索(ローマ字姓名による)

Read  
CINI  
Google  
CINI  
Webcat Plus  
Read  
Google Scholar  
Google

検索(日本語姓名および所属による)  
検索(日本語姓名による)  
検索(ローマ字姓名による)

科研費研究課題の研究分野  
内化学一般/発生生物学

NII National Institute of Informatics

13

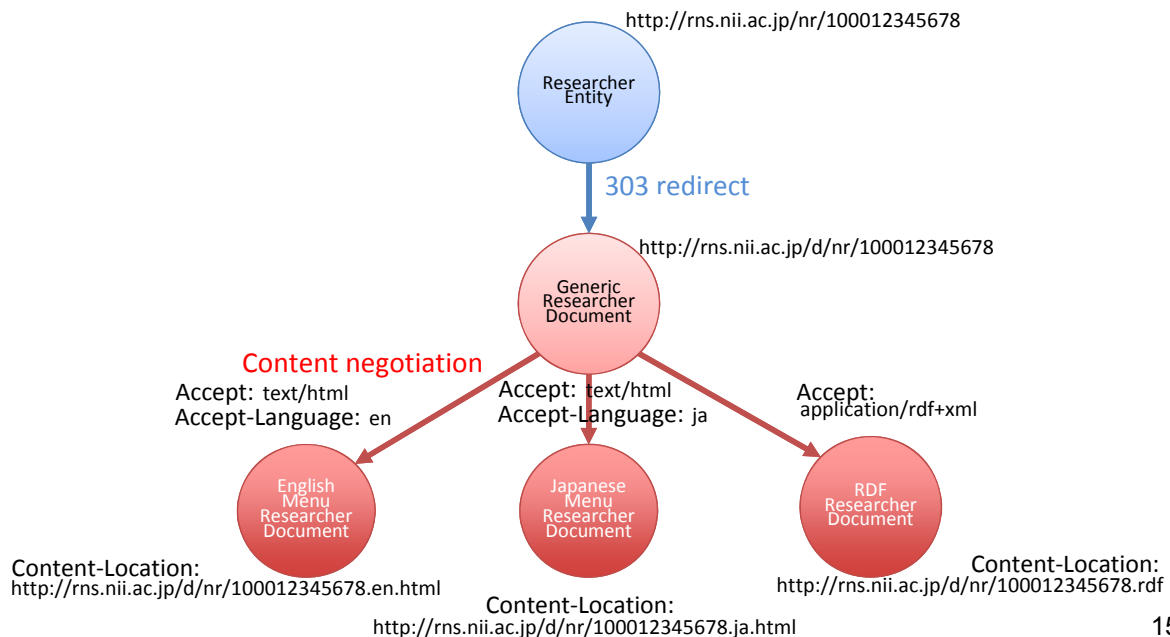
## 機械によるアクセス

- OpenSearchによる研究者検索
  - [http://rns.nii.ac.jp/html/opensearch\\_help.html](http://rns.nii.ac.jp/html/opensearch_help.html)
- Linked Data
  - 研究者URIと研究者ドキュメントURI
  - Dereference(転送)とコンテンツネゴシエーション
  - RDF/XMLによる情報公開
  - [http://rns.nii.ac.jp/html/linkeddata\\_help.html](http://rns.nii.ac.jp/html/linkeddata_help.html)

14

# 研究者URIとLinked Data

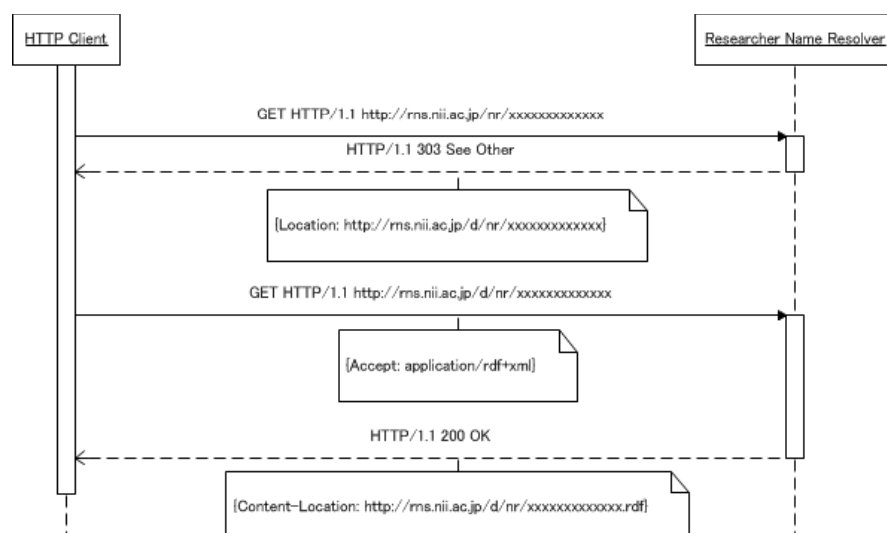
- 研究者URIと研究者ドキュメントURIを区別



15

## 研究者URIへのアクセス

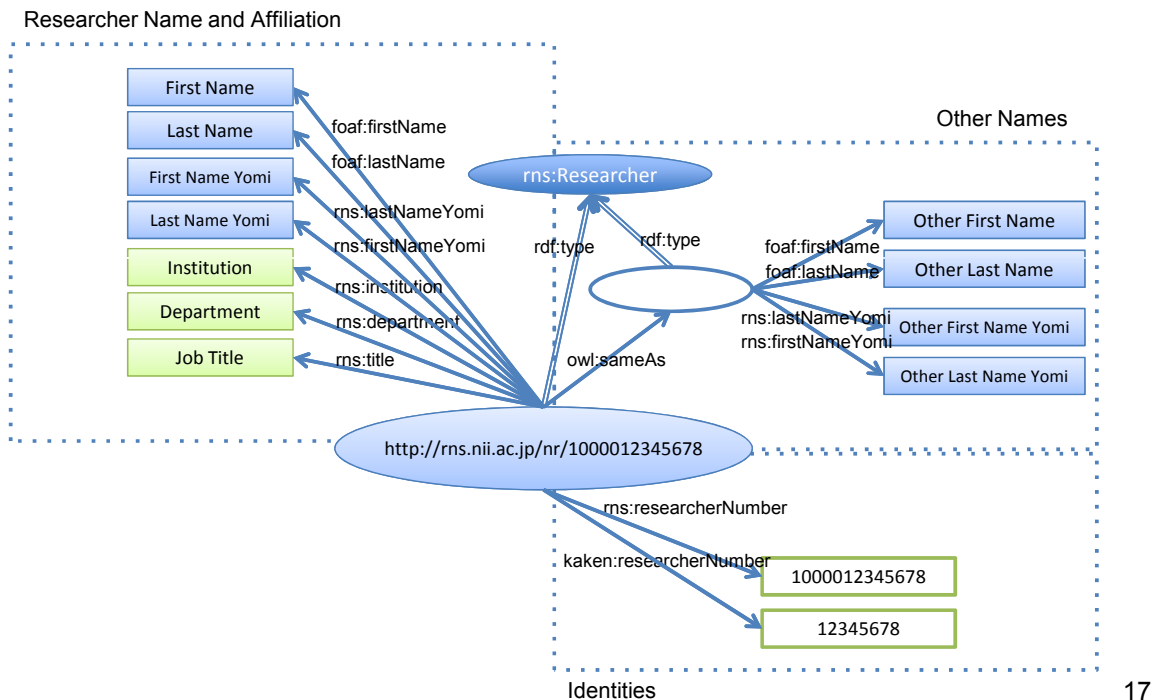
- 研究者URIから研究者ドキュメントへ転送 (dereference)



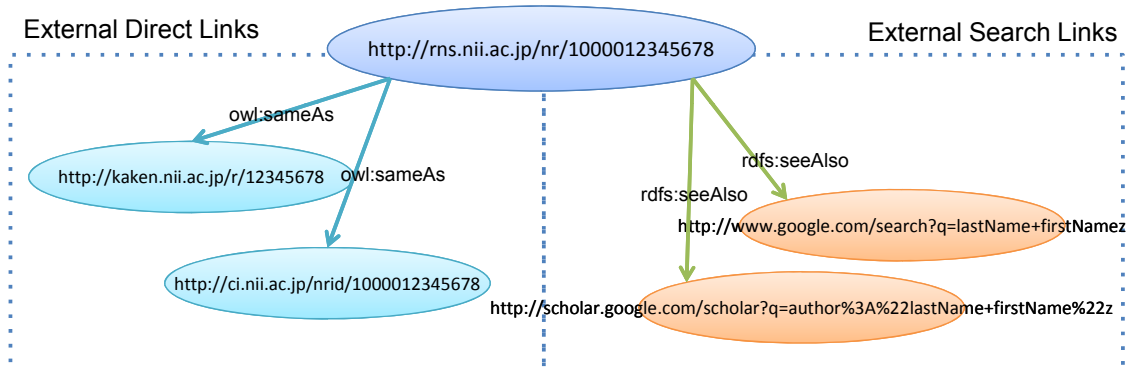
16



# RDFによる研究者情報(1/2)

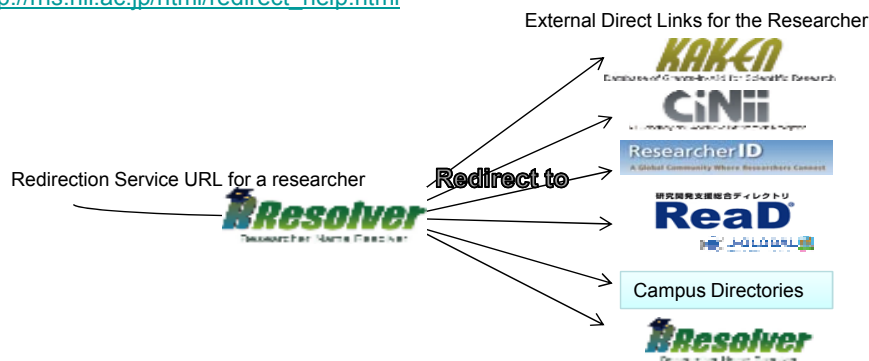


# RDFによる研究者情報(2/2)



# 研究者IDマッピングテーブル 関連ウェブサービス

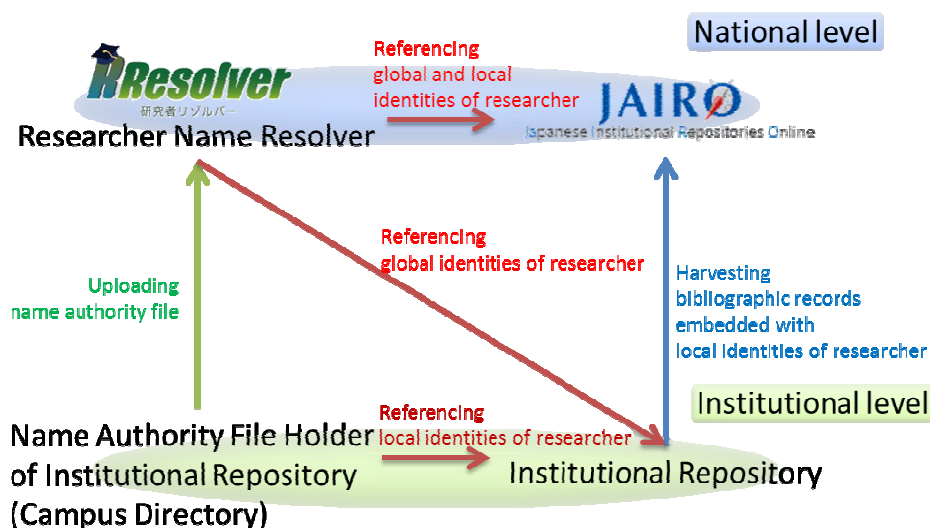
- 研究者IDマッピングテーブル検索
  - XML-RPCで既知のデータベースタイプと研究者ID、および知りたい研究者IDのデータベースタイプを指定すると、その研究者IDが返却される
  - [http://rns.nii.ac.jp/html/webservice\\_help.html](http://rns.nii.ac.jp/html/webservice_help.html)
- ダイレクトリンクのためのURLリダイレクトサービス
  - 既知のデータベースタイプと研究者ID、およびアクセス先のデータベースタイプを指定したURLでアクセスすると、転送される
  - [http://rns.nii.ac.jp/html/redirect\\_help.html](http://rns.nii.ac.jp/html/redirect_help.html)



19

# JAIRO著者名検索への応用

- 研究者リゾルバーはJAIROの著者名典拠
- 金沢大学、静岡大学、NIMSと実証実験中で、参加者募集中



募集中!

20

```
<?xml version="1.0" encoding="UTF-8" ?>
<OAI-PMH
  xmlns="http://www.openarchives.org/OAI/2.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2011-05-26T13:34:09Z</responseDate>
  <request metadataPrefix="junii12" verb="GetRecord">
    identifier="oai:ir.lib.shizuoka.ac.jp:10297/5644">http://ir.lib.shizuoka.ac.jp/dspace-oai/request</request>
  <GetRecord>
    <record>
      <header>
        <identifier>oai:ir.lib.shizuoka.ac.jp:10297/5644</identifier>
        <timestamp>2011-05-26T08:02:22Z</timestamp>
        <setSpec>hdi_10297_24</setSpec>
      </header>
      <metadata>
        <junii12
          xmlns="http://irdb.nii.ac.jp/oai"
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xsi:schemaLocation="http://irdb.nii.ac.jp/oai http://irdb.nii.ac.jp/oai/junii12.xsd">
          <title>Orientation-dependent epitaxial growth of GaAs by current-controlled liquid phase epitaxy</title>
          <creator>Mouleswaran, D.</creator>
          <creator id="http://rns.nii.ac.jp/nr/1000001133354">Koyama, T.</creator>
          <creator id="http://rns.nii.ac.jp/nr/1000048520242">Hayakawa, Yasuhiro</creator>
          <NDCD>459</NDCD>
          <description>
            The growth dependence of the selective epitaxial growth of Gallium Arsenide (GaAs) has been
            investigated to achieve a high-quality GaAs layer for application to X-ray detectors. Selective epitaxial growth was carried out
            on patterned GaAs substrates with [0 1 -2], [0 1 -1] and [1 1 -2] orientations, and their equivalent seed orientations by current-

```

```
orientations were suitable for application in a GAXS x-ray detector.</description>
<publisher>Elsevier</publisher>
<NIIttype>Journal Article</NIIttype>
<format>application/pdf</format>
<URI>http://hdl.handle.net/10297/5644</URI>
<fullTextURL>http://ir.lib.shizuoka.ac.jp/bitstream/10297/5644/1/110520001.pdf</fullTextURL>
<jtitle>Journal of Crystal Growth</jtitle>
<issn>0020248</issn>
<NCID>AA00696341</NCID>
<volume>321</volume>
<issue>1</issue>
<page>85</page>
<epage>90</epage>
<dateofissue>2011-04-15</dateofissue>
<language>eng</language>
<doi>info:doi/10.1016/j.jcrysgro.2011.02.026</doi>
<rights>Copyright © 2011 Elsevier B.V. All rights reserved.</rights>
<textversion>author</textversion>
</junit2>
</metadata>
</record>
</GetRecord>
</OAI-PMH>
```



- 機関を超えて正確にリストされる



## 2.KAKENデータベースの機能拡張

23

## 科研費とは

- 科学研究費補助金は、文部科学省と日本学術振興会による研究費助成制度
- 1939年の「科学奨学金」を起源とし、1965年に文部省（現文部科学省）が現行の科研費の制度を設立
- すべての分野と様々な研究活動の段階を網羅
- 我が国最大・唯一の国の研究費配分制度
- ほとんどの日本における研究者が申請登録  
約19.2万人がKAKENデータベースに登録されている
- 平成23年度は総計2600億円  
複数年度にわたる研究費使用を可能にするため、  
一部研究種目について「基金化」を実施
- 対応する諸外国の研究助成金制度運営組
  - NSF(National Science Foundation), USA
  - RCUK(Research Councils of UK), UK

科学研究費補助金  
Grant-in-Aid for Scientific Research



24

# 科学研究費補助金関連システム

- e-Rad(府省共通研究開発管理システム)
  - 文部科学省が運営
  - 研究開発管理関連ドキュメントの公開
  - 研究者番号の発行とユニークネス保証
- e-Rad連携・電子申請システム
  - 日本学術振興会(JSPS)が運営
  - 研究課題の申請受付・審査
- KAKEN
  - 国立情報学研究所が運営
  - 採択課題と研究成果の公開
  - <http://kaken.nii.ac.jp>



25

## KAKENにおける システムデザインの概念的枠組み

### 1. 検索

研究課題

研究者



### 2. 検索結果

研究成果



### 3. リンクをたどって閲覧

論文

Linked



26

# 検索結果一覧 (研究課題検索結果)

**検索条件**

**関連キーワード**

**検索結果**

研究課題名: 1013件 (1-20 表示)

1 2 3 4 5 6 ... 53

研究課題名: 1013件 (1-20 表示)

研究代表者: 1013件 (1-20 表示)

研究分担者・連携研究者・研究協力者: 1013件 (1-20 表示)

各年度報告書へのリンク

基本情報

研究概要

このページのURL: <http://kaken.nii.ac.jp/ja/p/17200007>

27

## 研究課題ページ

**研究課題名**

**研究代表者**

**研究分担者・連携研究者・研究協力者**

**各年度報告書へのリンク**

**基本情報**

**研究概要**

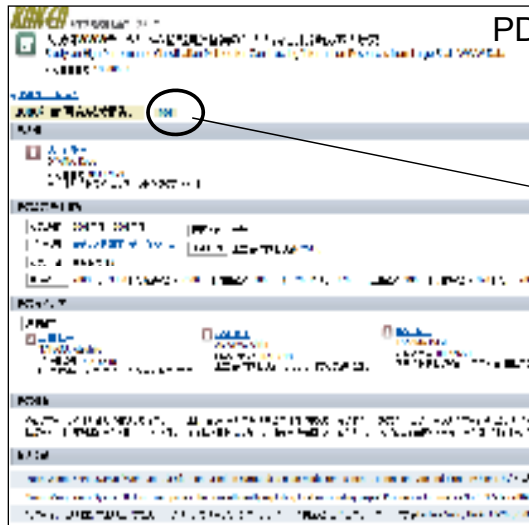
**キーワード**

このページのURL: <http://kaken.nii.ac.jp/ja/p/17200007>

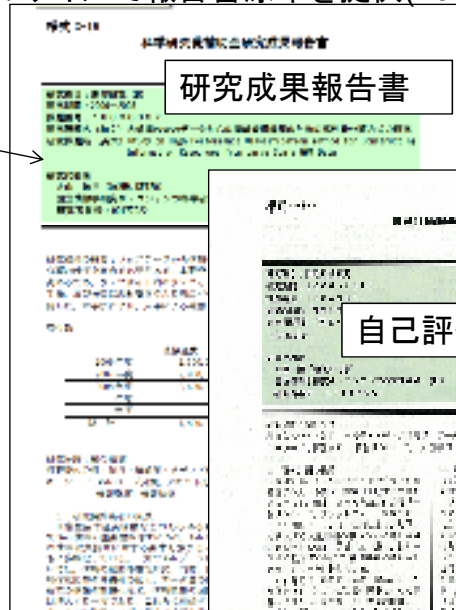
28



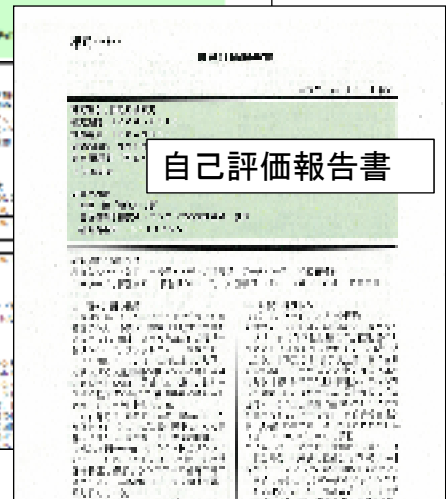
# 研究成果報告書,自己評価報告書



PDFファイルで報告書原本を提供(2008年度以降)



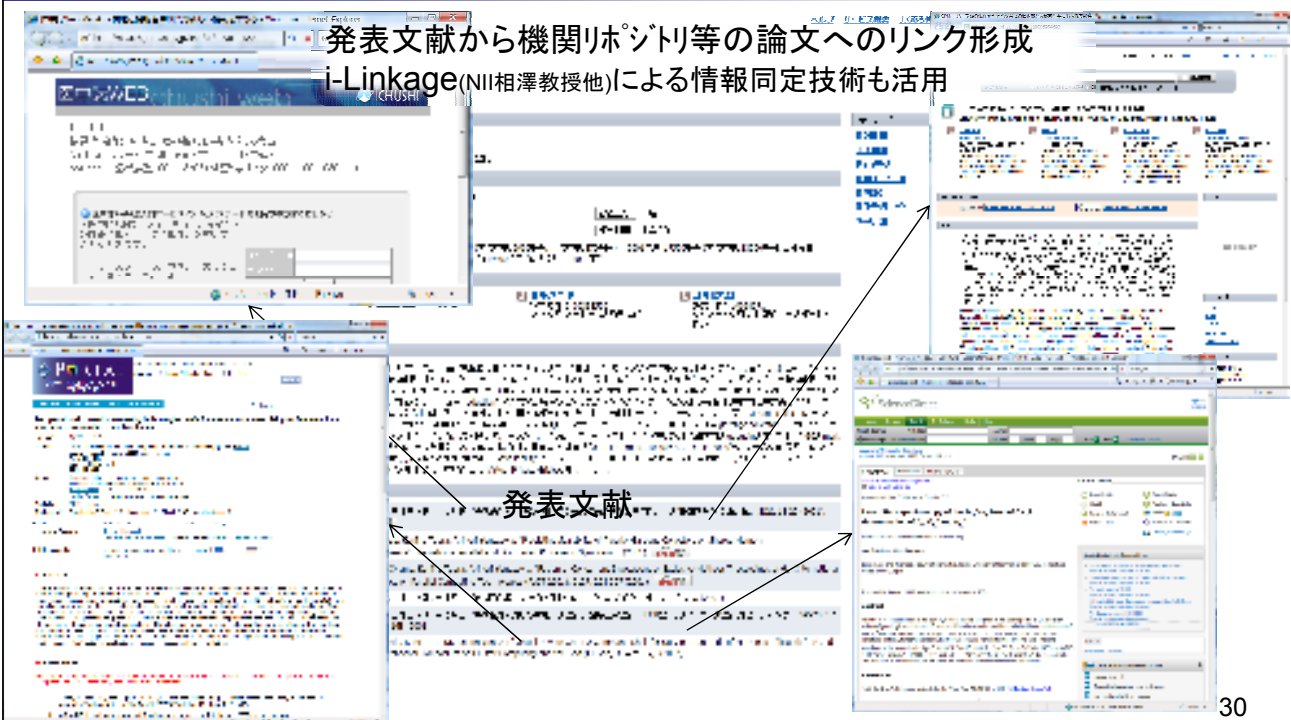
研究成果報告書



自己評価報告書

## 実績報告書ページ

発表文献から機関リポジトリ等の論文へのリンク形成  
i-Linkage(NII相澤教授他)による情報同定技術も活用



発表文献

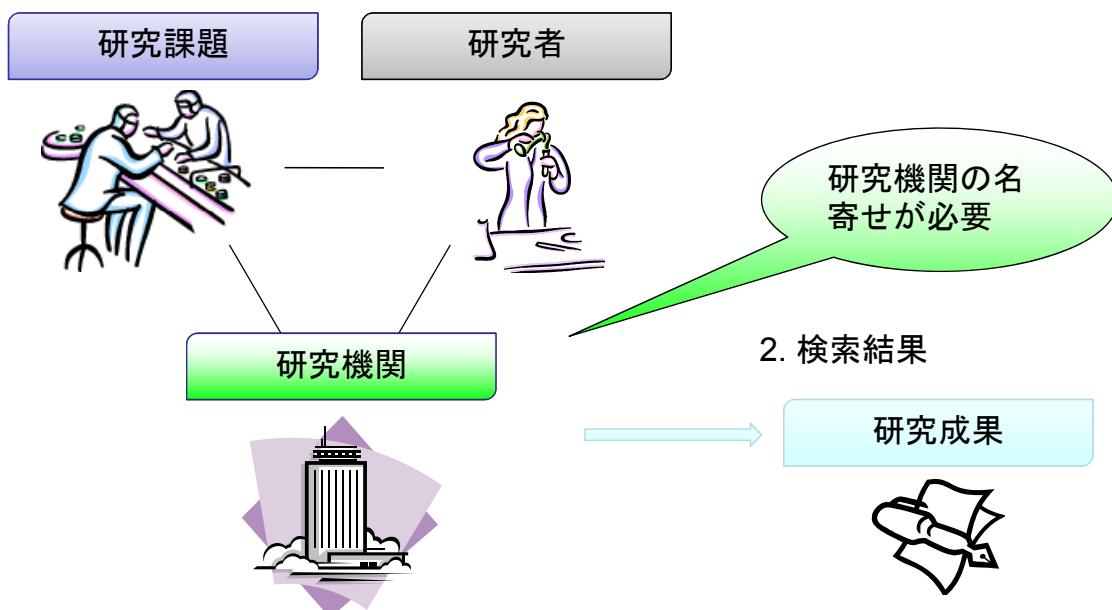
# 機械のためのAPI

- OpenSearchによる研究課題と研究者の検索
  - 検索結果一覧を返すフォーマット
    - Atom
    - RSS 1.0
- RDF/XMLによるデータ公開
  - 研究課題URI
    - <http://kaken.nii.ac.jp/p/xxxxxxx>  
(xxxxxxx is a project number)
  - 研究者URI
    - <http://kaken.nii.ac.jp/r/xxxxxxx>  
(xxxxxxx is a researcher number)
- SPARQLサーバ(予定)
- 詳しくは、CKANで
  - <http://thedatahub.org/dataset/kaken>

31

## 今度は、研究機関も対象に

### 1. 検索



32



# 実際のデータ(1/2)

- 研究課題実施機関

- 研究機関番号(文科省の定める5桁の数字)

同じ機関で、名称変更にも対応できる

【1973年】12102:東京教育大学(69)

【1974年】12102:東京教育大学(58)、12102:筑波大学(3)

【1979年】12102:筑波大学(204)

しかし、電話番号のように再利用されている大学がある・・・

【2002年】12103:図書館情報大学(12)

【2005年】12103:筑波技術大学(1)

当然のように、表記揺れもある・・・

【2008年】32612:慶應義塾大学、32612:慶応義塾大学

33

# 実際のデータ(1/2)

- 研究者の所属

- 機関番号なし

慶應義塾大学

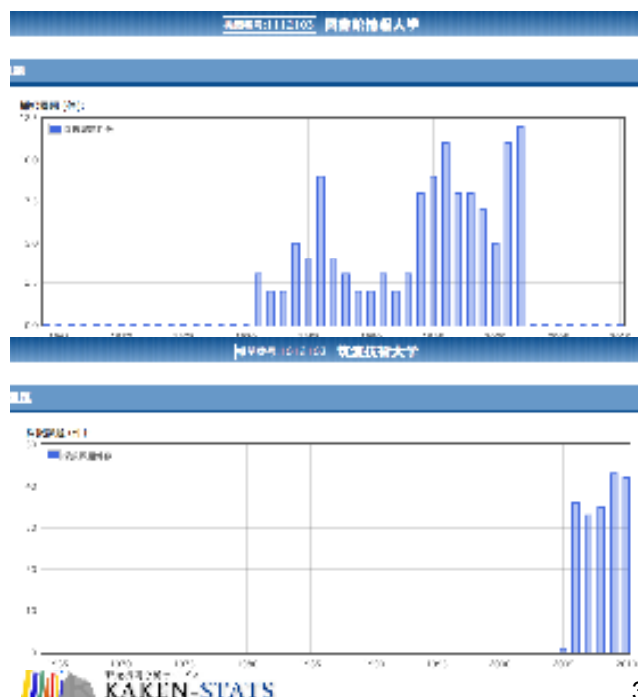
大量の表記揺れと誤表記

慶応大  
慶応大学  
慶応大学医学部  
慶應大学  
慶應義塾大学医学部  
慶應義塾大学医学部内科  
慶應義塾大学大学院  
慶應義塾大学言語文化研究所  
慶應義塾大学  
慶應義塾大学  
慶応義塾大  
慶応義塾大学  
慶應大  
慶應義塾大学

34

# 機関名名寄せができると(1/2)

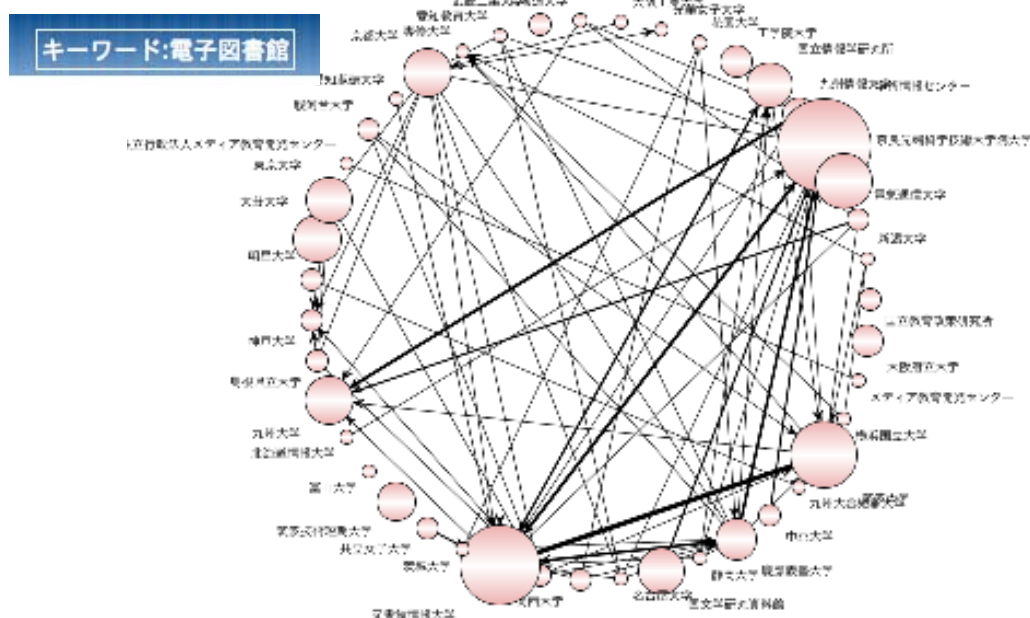
- 正確な検索結果
- 統計分析可能



35

# 機関名名寄せができると(2/2)

- 研究キーワードによる、研究機関ネットワーク分析



36

# まとめ

- 研究者リゾルバーとJAIRO著者名検索
  - 著者名名寄せ
- 科研費データベースKAKENの拡張
  - 機関名名寄せ
- 連絡先
  - kurakawa@nii.ac.jp
  - @keikurakawa
  - <http://facebook.com/kei.kurakawa>