

研究者同定とその応用に向けて

Towards a researcher identification and its application

高久雅生[†], 相澤彰子^{††}, 馬場康維^{†††}

Masao Takaku[†], Akiko Aizawa^{††}, Yasumasa Baba^{†††}

[†] 物質・材料研究機構 (National Institute for Materials Science)

^{††} 国立情報学研究所 (National Institute of Informatics)

^{†††} 統計数理研究所 (The Institute of Statistical Mathematics)

目次

- データベース統合による研究者同定
- 統計分野3学会の会員名簿の人手同定
 - 3学会・のべ3000人規模 vs KAKEN(15万人規模)
 - 同定結果
 - 統計研究者コミュニティの分析
- 機関リポジトリにおける著者IDサービスの展開
 - NIMS eSciDoc
 - ワンストップ型サービス
 - 多目的利用
 - 組織内外とのより良い連携強化を目指す

背景

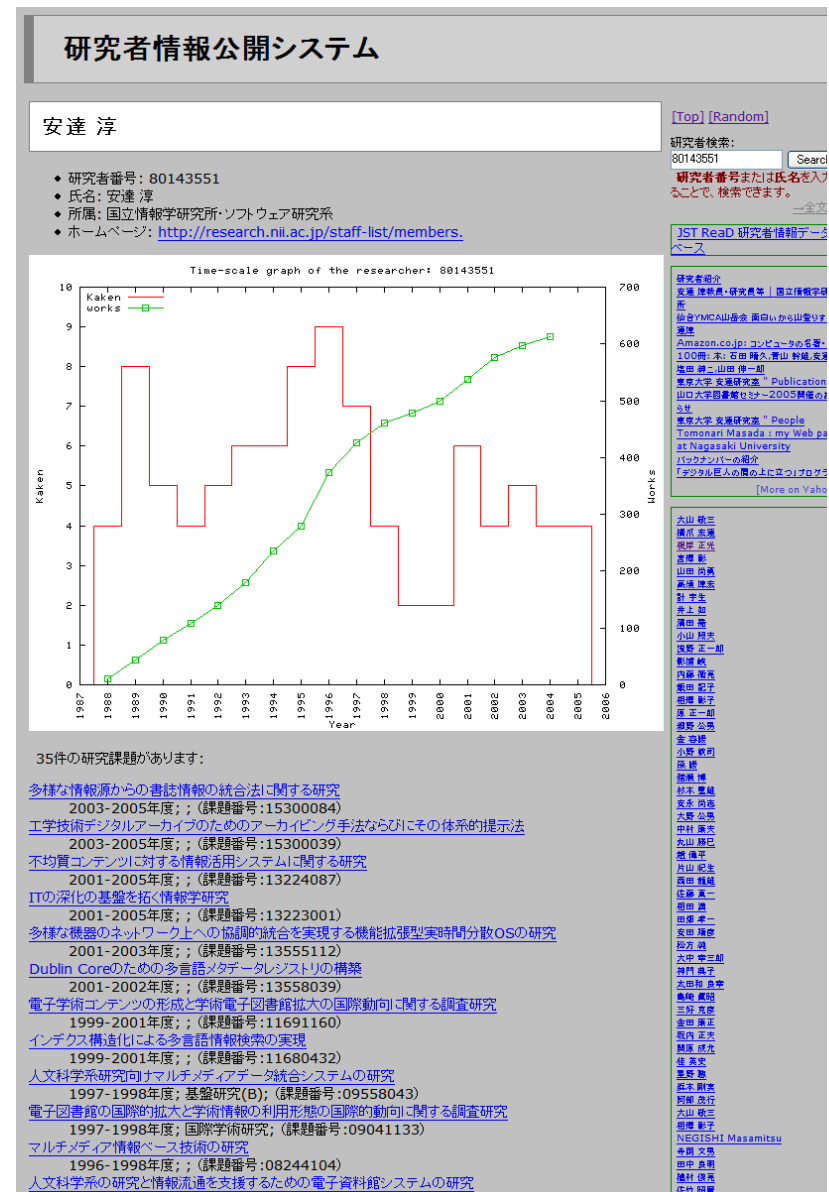
- 著者同定によるサービス展開
 - 古くから構想されてきた...
 - 書籍に関しては、著者名典拠 (Authority file) を活用したサービスを提供
 - 論文データベースに関しては規模などの点から困難...
- 研究者評価
 - 共著論文を (正しく?) 評価するには同定が必須
- 難しい点
 - 同姓同名
 - 研究者にひもづいた情報は頻繁に変わるため...
 - 組織改編、異動、研究分野、改姓・改名
 - 人手での同定を要する...

目的

- 氏名同定
 - 異なるデータベース間の統合(リンケージ)
 - 同姓同名問題の解決を目指す
- 研究者情報
 - 科研費研究者番号をもとにした基本情報セット
 - 研究者情報サーバ
- 統計関連3学会名簿との突合せ
 - 人手同定作業
 - 科研費基本情報との統合によるコミュニティ分析、共同研究分析

科研費データベースをもとにした 研究者情報サーバ

- 科研費データベースから再編成
 - 研究課題数: 247,745
 - 実績報告書+成果報告書: 1989-2004
 - 研究者数: 133,067
 - (代表者および分担者)
- 個人単位の基本情報
 - 氏名、所属組織名、職位
- 可視化
 - 時系列: 採択課題数と発表文献数
 - 共同研究ネットワーク
 - 過去の課題での共同研究関係
 - Google Mapsを用いたマッピング



統計分野研究者データセット

- 統計関連3学会：のべ3031件
 - 日本統計学会 (2003): 1,545名
 - 日本行動計量学会(2005): 1,070名
 - 日本計算機統計学会 (2002): 416名
- 氏名、所属組織の情報のみを同定に用いる

研究者同定(ステップ1)

氏名表記での一致候補を抽出

- 科研費データベース上の氏名表記と名簿の氏名表記をマッチさせ、一致したものを同定のための候補として抽出。
- 1,400 件の同定候補:
 - 日本統計学会: 697 (45%)
 - 日本行動計量学会: 519 (49%)
 - 日本計算機統計学会: 184 (44%)

研究者同定(ステップ2)

人手判定

- 統計名簿上の各会員に科研費研究者番号を付与

同定候補

統計3学会 名簿	科研費研 究者番号
BSJ0004	20024581
BSJ0005	50305313
JSS0012	70303047 09246528
BSJ0007	90184332
JSS0014	90184332
JSS0015	90132696

重複
(複数学会
所属者)

同定結果

統計3学会 名簿		科研費研 究者番号
BSJ0004	別人	20024581
BSJ0005	同一人物	50305313
JSS0012	別人	70303047 09246528
BSJ0007	同一人物	90184332
JSS0014	同一人物	90184332
JSS0015	同一人物	90132696

人手判定

研究者同定ツール

統計学分野研究者同定システム - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(T) ヘルプ(H)

戻る 進む 更新 中止 ホーム http://localhost/~masao/private/cvswork/jstat-dataset/app.cgi

統計学分野研究者同定システム

同定候補 (1-20):

- 再判定
- 野口 隆士 == 野口 隆士 再判定
- 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定
- 山田 隆士 == 山田 隆士 再判定

ページ: [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [33] [34] [35] [36] [37] [38] [39] [40] [41] [42] [43] [44] [45] [46] [47] [48] [49] [50] [65] [66] [67] [68] [69]

[候補統合へ](#)

統計学分野研究者同定システム - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(T) ヘルプ(H)

戻る 進む 更新 中止 ホーム http://localhost/~masao/private/cvswork/jstat-dataset/app.cgi?start=320

統計学分野研究者同定システム

同定候補 (321-340):

- 再判定
- 再判定
- 再判定
- 再判定
- 再判定
- 関西大学商学部 (JSS0098)
80202745 ☐ 同一人物 ☐ 別人
同定確定
- 青山学院大学経営学部 (JSS0393)
20303050 ☐ 同一人物 ☐ 別人
同定確定
- 京都学国大学人間文化学部人間関係学科 (BSJ0226)
60240628 ☐ 同一人物 ☐ 別人
同定確定
- 一橋大学経済学部 (JSS0400)
70154838 ☐ 同一人物 ☐ 別人
同定確定
- 豊田工業高等専門学校 (BSJ0229)
40112833 ☐ 同一人物 ☐ 別人
90216705 ☐ 同一人物 ☐ 別人
同定確定
- 東京工業大学大学院情報理工学系研究科数学・計算科学専攻 (JSS0402)
30154876 ☐ 同一人物 ☐ 別人
40312300 ☐ 同一人物 ☐ 別人
70016153 ☐ 同一人物 ☐ 別人
同定確定
- 中外製薬(株)臨床解析部 (BSJ0230)
60125391 ☐ 同一人物 ☐ 別人
同定確定
- 筑波大学社会医学系 (JSS0404)
80261808 ☐ 同一人物 ☐ 別人
同定確定

完了

- [redacted] 京都学園大学人間文化学部人間関係学科 (BSJ0226)

[redacted] 60240628 ☒ 同一人物 ☐ 別人

同定確定

- [redacted] 一橋大学経済学部 (JSS0400)

[redacted] 70154838 ☒ 同一人物 ☐ 別人

同定確定

- [redacted] 豊田工業高等専門学校 (BSJ0229)

[redacted] 40112833 ☐ 同一人物 ☒ 別人

[redacted] 90216705 ☒ 同一人物 ☐ 別人

同定確定

- [redacted] 東京工業大学大学院情報理工学研究科数学・計算科学専攻 (JSS0402)

[redacted] 30154876 ☐ 同一人物 ☒ 別人

[redacted] 40312300 ☐ 同一人物 ☒ 別人

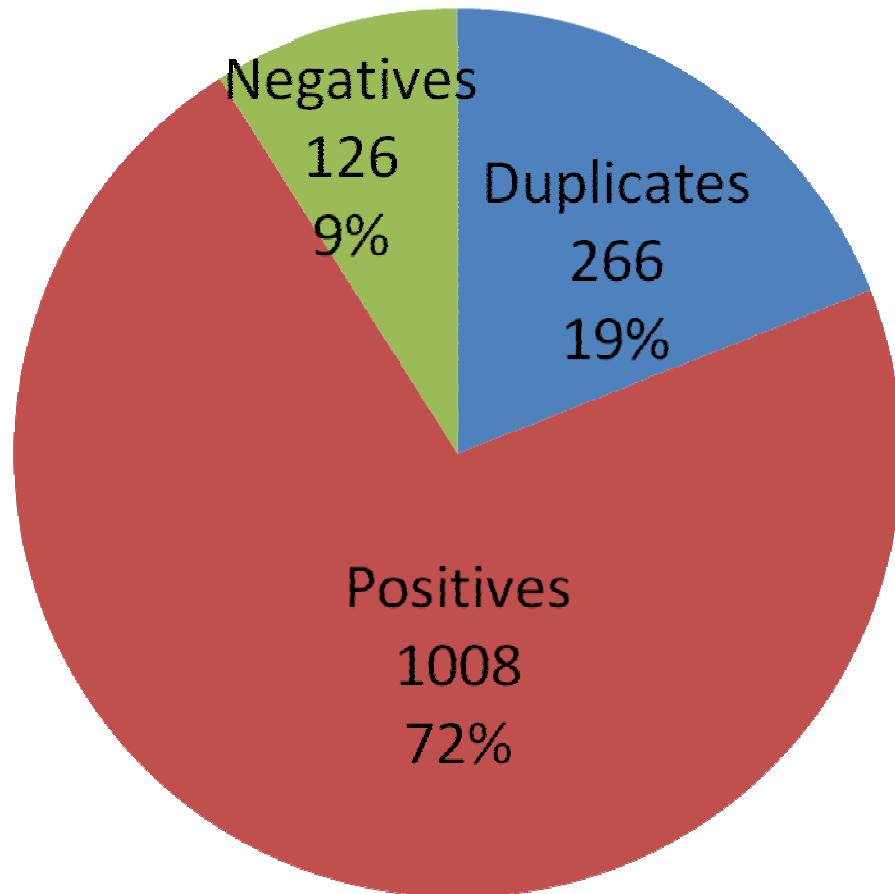
[redacted] 70016153 ☒ 同一人物 ☐ 別人

同定確定

研究者情報サーバ
へのリンク

同定結果

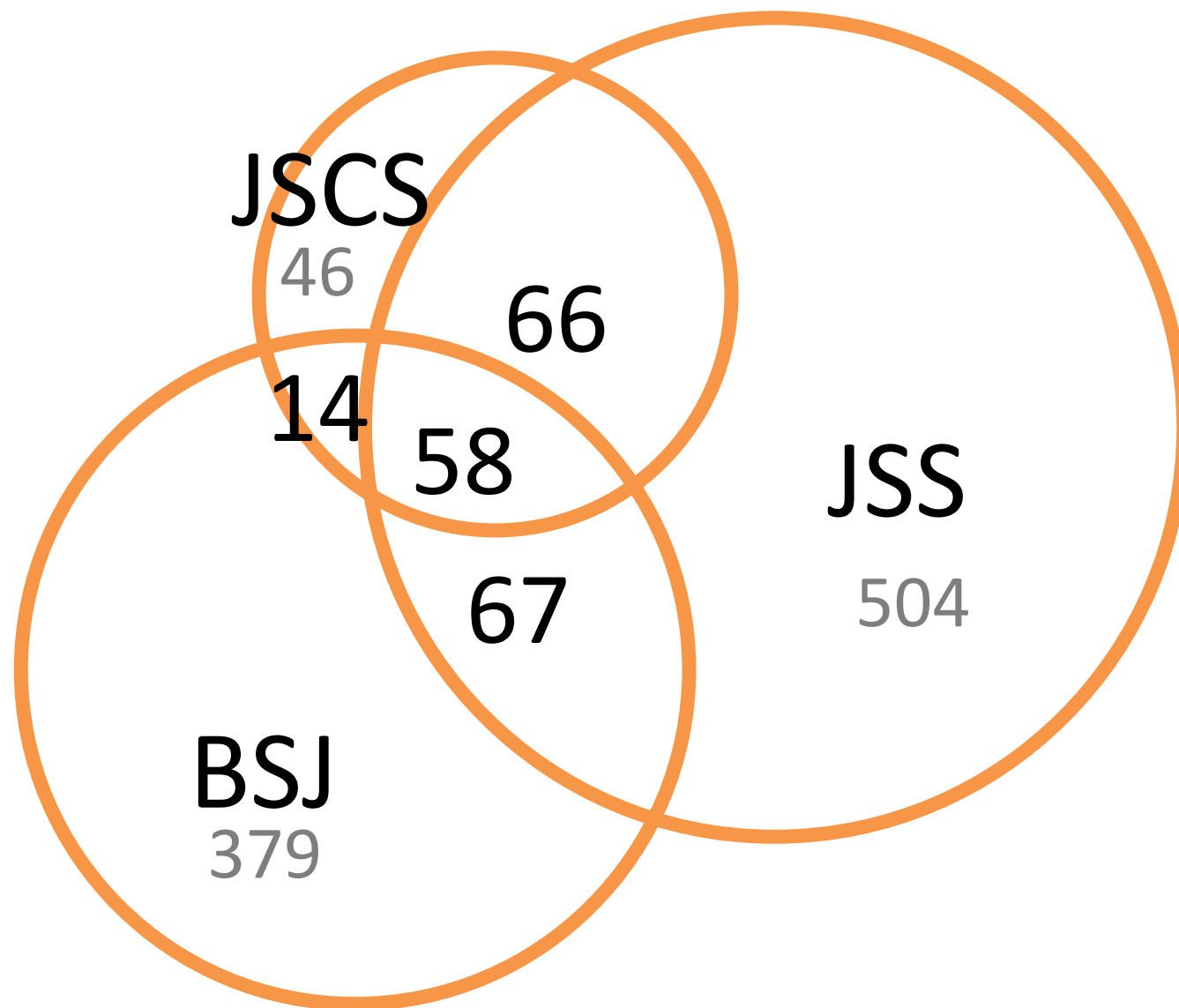
- 266件は重複(複数学会所属者)分
- 1008 名が科研費データベース中の少なくともレコードと同定



重複
(複数学会
所属者)

Members of societies		KAKEN IDs
BSJ0004	別人	20024581
BSJ0005	同一人物	50305313
JSS0012	別人	70303047 09246528
BSJ0007		
JSS0014	同一人物	90184332
JSS0015	同一人物	90132696

同定候補中の重複(複数学会所属者)



同姓同名研究者に関して

A) 科研費DB内での同姓同名

- 科研費データベースへの氏名表記マッチングにおいて複数IDとマッチし、かつ、別人であったもの。

→ 105 件 (同定調査対象中の10%)

Members of societies	KAKEN IDs	
JSS0035	Identical	308679
	Different	241595

B) 科研費DB外との同姓同名

- 科研費データベースに同姓同名の別人が含まれており、かつ、学会名簿上の会員は別人であったもの。

→ 126 件

Members of societies	KAKEN IDs	
BSJ0004	Different	024581
JSS0012	Different	79303047 55246528

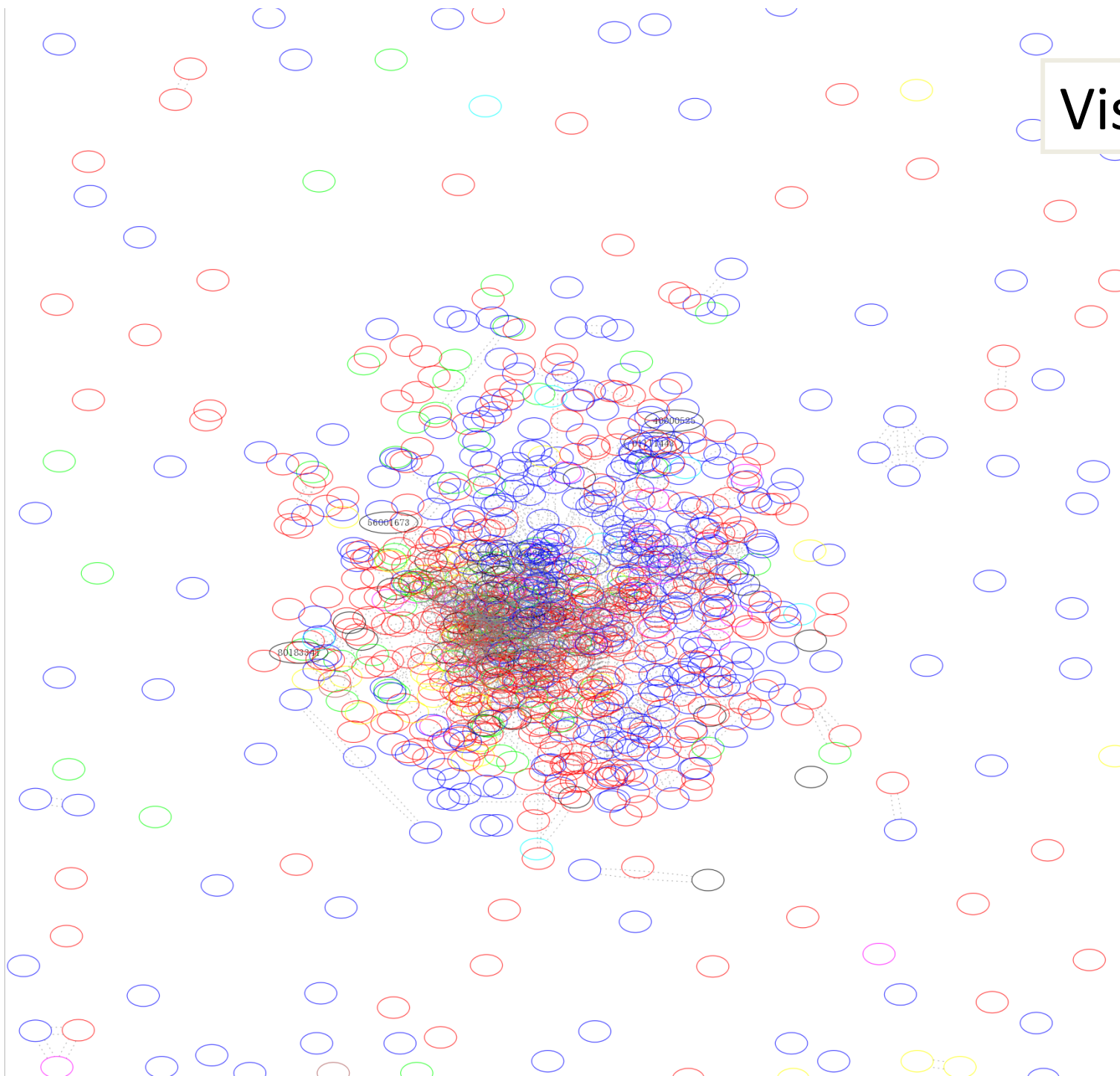
- (a) + (b) → 231件 (同定調査対象中の20%)

研究者ネットワーク

- 共同研究者ネットワークの可視化
 - 過去約15年間の共同研究関係を可視化
 - 統計分野コミュニティ
- 応用領域：
 - 同定のための可視化
 - 研究領域の把握のための可視化

Visualization

- JSS
- BSJ
- JSCS



目次

- データベース統合による研究者同定
- 統計分野3学会の会員名簿の人手同定
 - 3学会・のべ3000人規模 vs KAKEN(15万人規模)
 - 同定結果
 - 統計研究者コミュニティの分析
- 機関リポジトリにおける著者IDサービスの展開
 - NIMS eSciDoc
 - ワンストップ型サービス
 - 多目的利用
 - 組織内外とのより良い連携強化を目指す

機関リポジトリ

Institutional Repository (IR)

- オープンアクセス運動
 - 雑誌資料費の高騰に対するアンチテーゼ
- 電子図書館
 - 図書館によるデジタル資料の提供
- 大学等の高等教育機関による成果物の集積体
 - 大学広報および活動の可視性を高める運動
- 普及状況
 - 全世界: 668機関
 - [ROAR \(Registry of Open Access Repositories\)](#)
 - 国内の大学および研究機関: 87機関
 - <http://www.nii.ac.jp/irp/list/>

NIMS eSciDoc

- 物質・材料研究機構における機関リポジトリ
- 今年度から開発開始
- マックスプランク研究所との共同開発

The screenshot displays the NIMS eSciDoc web interface. The top navigation bar includes links for Logout, Masao, Takaku, About, FAQ, Help, and flags for Germany and the UK. A search bar with a 'Go' button and options for 'Include files', 'Advanced Search', and 'Organization Search' is present. The main content area shows the breadcrumb 'Actions: Home > Moderator Workspace > View Item'. On the left, a sidebar contains links for Home, Deposit Workspace, Full Submission, Easy Submission, and Quality Assurance Workspace. The main content area is titled 'Title: デジタルライブラリー、その先に見えるもの (Talk)' and includes sections for Item Details, Cite this Item as Alternative Title, Creators, Affiliations, Dates, free keywords, Language(s), and Rev. Method. The right sidebar contains links for Actions, Help, Modify, Create new revision, and Use as template.

Top Navigation: Logout Masao Takaku About FAQ Help

Search Bar: Go ☐ Include files [Advanced Search](#) [Organization Search](#)

Breadcrumb: Actions: Home > Moderator Workspace > View Item

Left Sidebar:

- [Home](#)
- [Deposit Workspace](#)
- [Full Submission](#)
- [Easy Submission](#)
- [Quality Assurance Workspace](#)

Main Content Area:

Title: デジタルライブラリー、その先に見えるもの (Talk)

Item Details: [View Release History](#) [View Revisions](#) [Files](#) [View Item Statistics](#) [Item Log](#)

Cite this Item as Alternative Title: no entry

Creators: Author: Tanifuji, Mikiko [1]
Author: Dreyer, Malte [2]
Author: Laurent, Romary [2]
Author: Takaku, Masao [1]
Author: Otsuka, Shingo [1]

Affiliations: 1: Scientific Information Office, Planning Division, National Institute For Materials Science
2: Max Planck Society Munich, DE

Dates: Published Online: 2008-11-27

free keywords: デジタルライブラリー, 機関リポジトリ, サブジェクトリポジトリ, Fedora, library system, deposit and archive scientific materials and research knowledge

Language(s): ja

Rev. Method: no entry

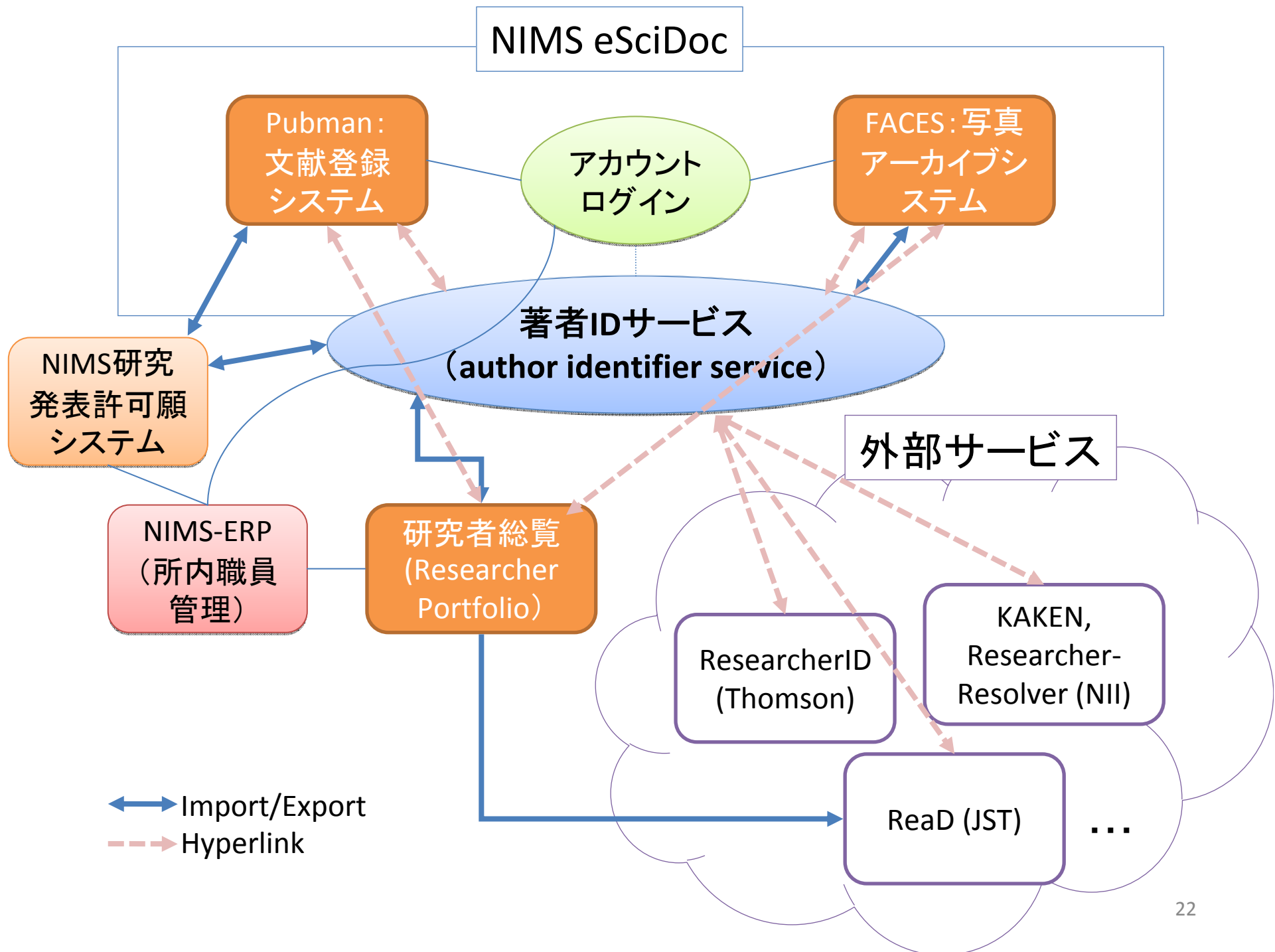
Right Sidebar:

- [Actions](#)
- [Help](#)
- [Modify](#)
- [Create new revision](#)
- [Use as template](#)

著者IDサービス

(Author identifier service)

- 現状：著者氏名は表記上のテキスト情報のみ
- 研究者総覧機能、著者単位のグルーピング・解析機能等で用いるための著者単位での基礎的なひも付けが必要
 - － 別サービスとして実装
 - 文献登録ツール、実験データ登録ツールなど複数のサービスと連携
 - 機関リポジトリだけでなく、所内研究者総覧・職員管理DB等との連携が重要に...
 - － 外部IDとの連携可能性を検討する
 - NII-KAKEN, JST-ReaD, ResearcherID.com, etc.



Spec. of author identifier service (draft)

- For each item (article level – local author ID)
 - Ex. “item/escidoc:1234/author:{1 .. n}”
- Global level author ID
 - Aggregating among overall systems (eSciDoc&other external systems incl. NIMS ID)
 - Ex. author1 = [
 - item/escidoc:1234/author:2,
item/escidoc:2345/author:1, ...,
 - :name:”高久雅生”, :name:”Masao Takaku”
 - ext:kaken:00000024, ext:nims:8717,
ext:read:1000302228, ext:researcherid:AA3957002654,
ext:xxx:zzz, ...,]
 - API: Create, Edit, Search/Retrieve, Export APIs
 - (Demo?)

まとめ

- 統計分野学会名簿と科研費DB研究者の同定
 - 3学会のべ3000人 vs 15万人
 - 1割ほどに同姓同名者
 - データベース統合の際のコスト見積もり
 - 共同研究ネットワーク
 - 研究コミュニティの可視化の可能性
- 機関リポジトリでの運用
 - 著者IDサービスへの構想
 - より良いデータベース統合の連携に向けて