

研究者同定とその応用 — 統計分野と材料科学分野を例として —

高久雅生^{*1}, 相澤彰子^{*2}, 馬場康維^{*3},
藏川圭^{*2}, 谷藤幹子^{*1}

- 1) 物質・材料研究機構
- 2) 国立情報学研究所
- 3) 情報・システム研究機構 新領域融合研究センター
/ 統計数理研究所

背景と目的

- 研究者同定による応用
 - 著者同定によるサービス展開
 - 書籍に関しては著者名典拠を活用したサービスが一般的
 - 論文データベースに関しては規模などの点から困難
 - 学術動向の把握:注目研究者の発見・抽出、分野の概覧
 - 研究者評価:共著論文における処理
- 困難な点
 - 同姓同名の別人、異名同人など
 - 研究者にひもづく情報の変更がたびたび発生
 - 改姓・改名、組織改編、異動、研究分野
 - 人手での同定を要する

研究者同定とその応用

- 複数データベース間での研究者同定
 - 研究者DB: KAKEN , ReaD, NIMS, 統計データセット, ResearcherID
 - 論文DB: CiNii, Web of Science, Scopus
- 研究者情報
 - 学術情報からの価値付加的な応用
 - 研究者情報の集約と交換

科研費IDを通じた研究者同定

- 科研費研究者番号を通じた同定
 - 原則として、一人一番号による背番号制
 - 異動や改姓による人物情報の変更に対応可能
 - 競争的資金取得用IDであるため、国内の大学・研究所系機関でのデファクトスタンダード
- KAKENデータベースにおいて集約
→ 研究者マスターデータとして採用

※科研費研究者データに関しては、時期による差異を吸収するため、2007年12月時点のデータを対象とした。

研究者同定・データセット

- 統計分野
 - 統計関連3学会・所属会員(延べ3,031名)
 - 日本統計学会(2003年) : 1,545名
 - 日本行動計量学会(2005年) : 1,070名
 - 日本計算機統計学会(2002年) : 416名
- NIMSデータセット
 - 物質・材料研究機構(NIMS)在職研究者(2009年)
 - 1,229名(定年制研究職、ポスドク等含む)

研究者同定の方法

- (氏名・所属情報のみを手がかりとする)
- 研究者名簿とKAKEN-DBに照合、氏名一致分を候補研究者として抽出
 - 統計分野: 3学会分の該当研究者:
 - 3,031 → 1,400 (46%)
 - 日本統計学会: 697件 (45%)
 - 日本行動計量学会: 519件 (49%)
 - 日本計算機統計学会: 184件 (44%)
 - NIMS: 1,229 → 348 (28%)
- 候補集合とKAKEN-DB等における登録内容を照合し、人手判定

研究者同定(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		同定結果	
統計3学会 名簿	科研費研 究者番号	統計3学会 名簿	科研費研 究者番号
BSJ0004	20024581	BSJ000	別人 20024581
BSJ0005	50305313	BSJ00	同一人物 305313
JSS0012	70303047 09246528	JSS0012	70303047 別人 09246528
BSJ0007	90184332	BSJ00	同一人物 184332
JSS0014	90184332	JSS00	同一人物 184332
JSS0015	90132696	JSS00	同一人物 132696

重複
(複数学会
所属者)

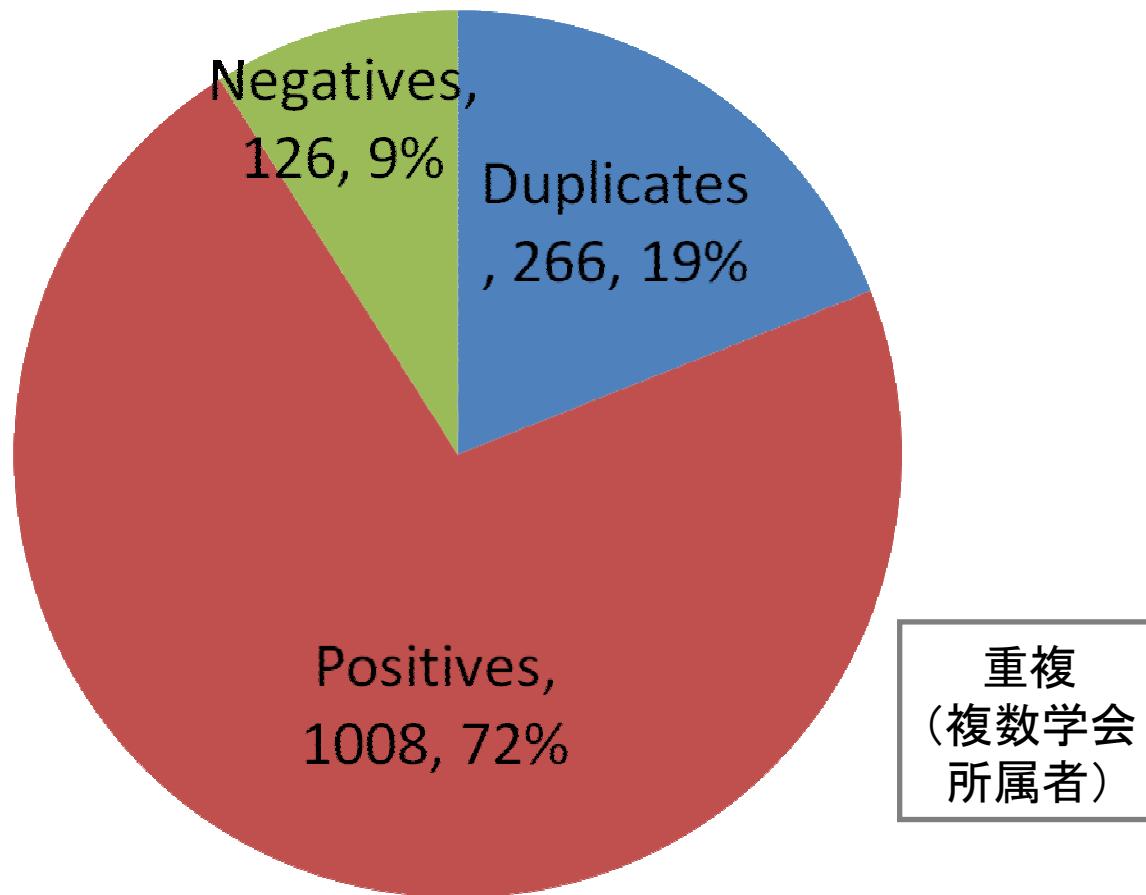
人手判定

- 京都学園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処理工学研究科数学・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](#)

「研究者情報サーバ」
へのリンク

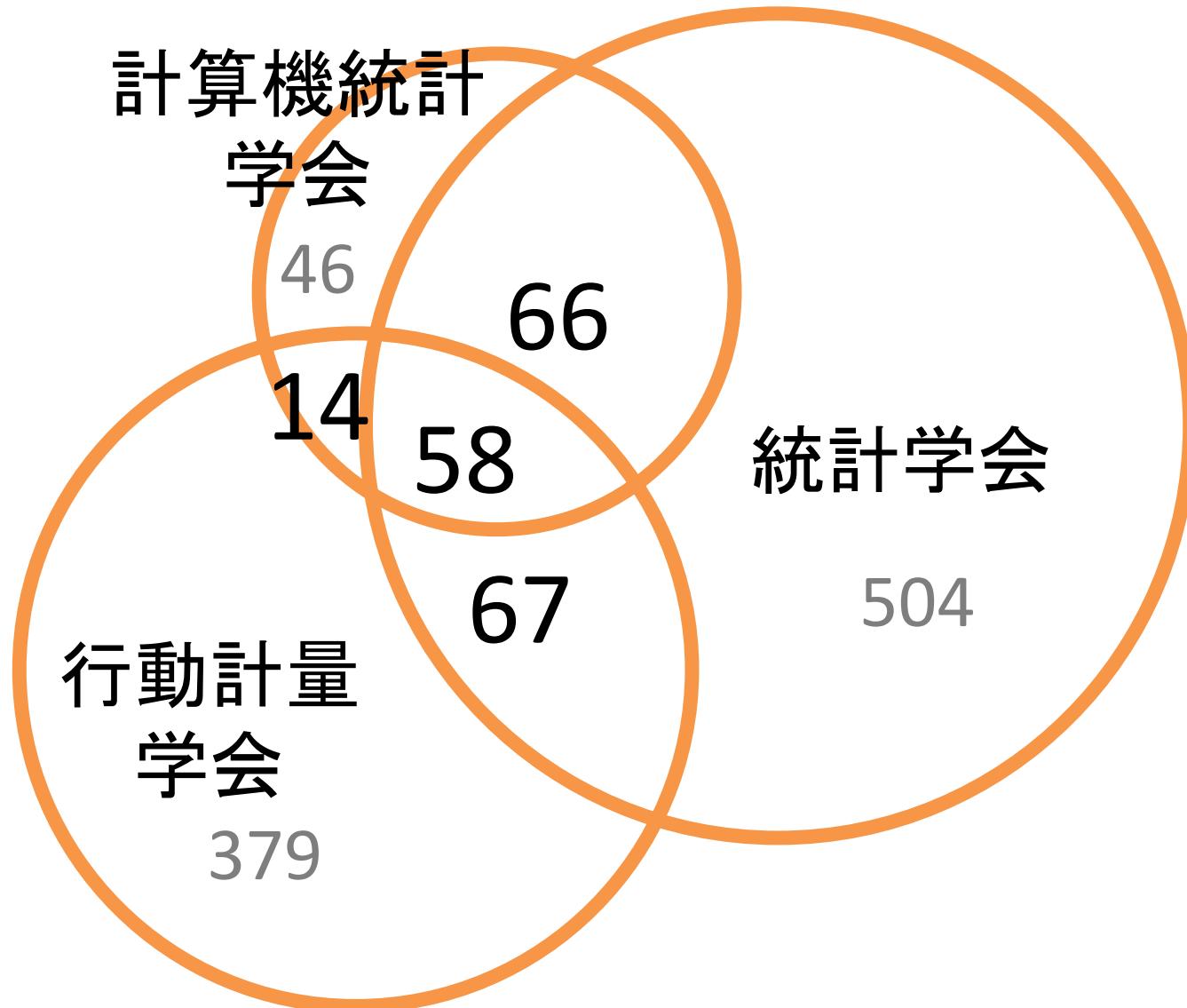
同定結果(統計分野)

- 266件は重複(複数学会所属者)分
- 1,008名が科研費データベース中の少なくとも1レコードと同定 (元データ集合の33.3%)



Members of societies	KAKEN IDs
BSJ0004	20024581
BSJ0005	同一人物 50305313
JSS0012	70303047 09246528
BSJ0007	同一人物 90184332
JSS0014	
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	同一人物 308679
JSS0035	同姓同名の別人 595

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
- (a) + (b) → 231件 (同定調査対象中の20%)

Members of societies	KAKEN IDs
BSJ	同姓同名の別人 581
JSS0012	70003047 同姓同名の別人 528

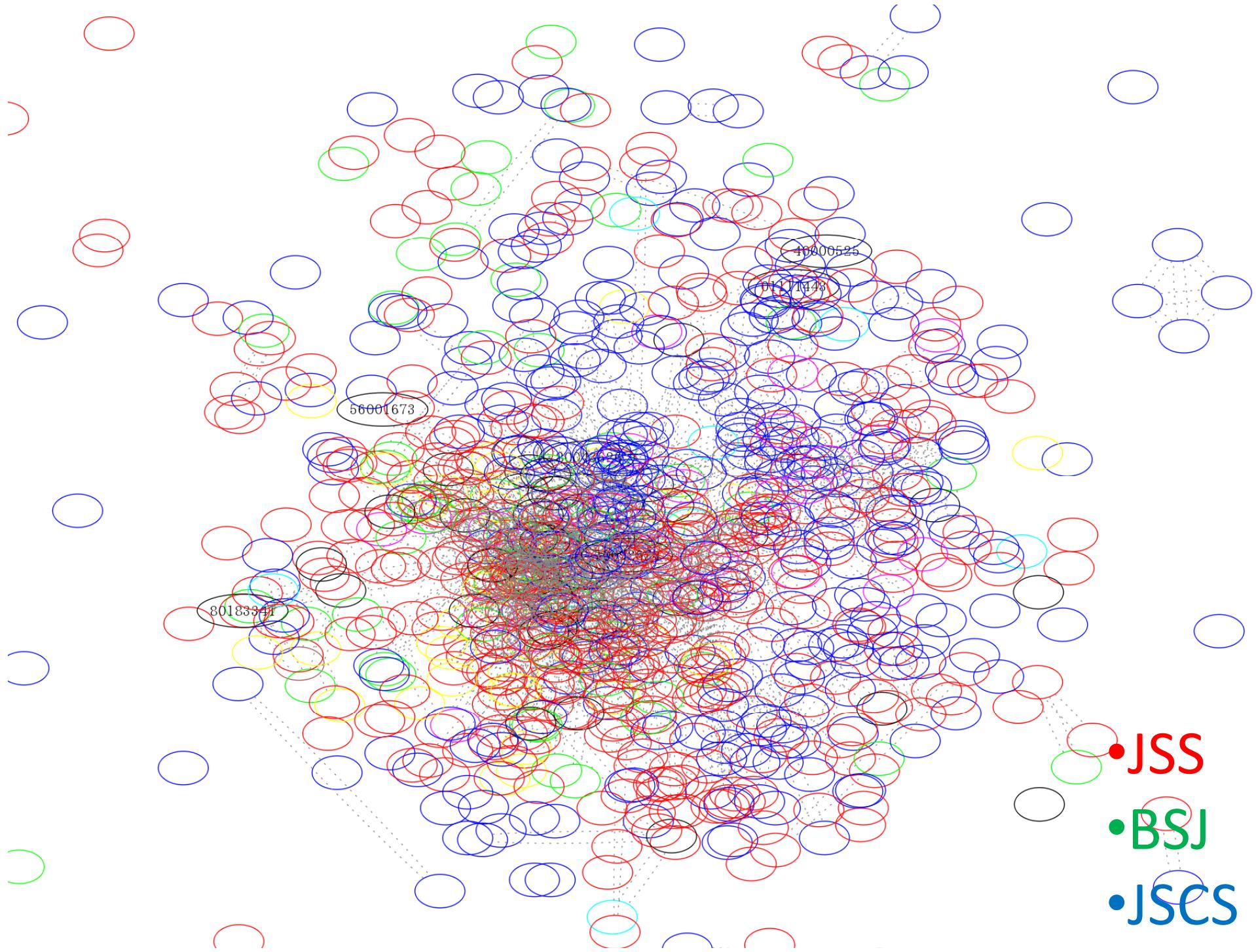
同定結果(NIMS)

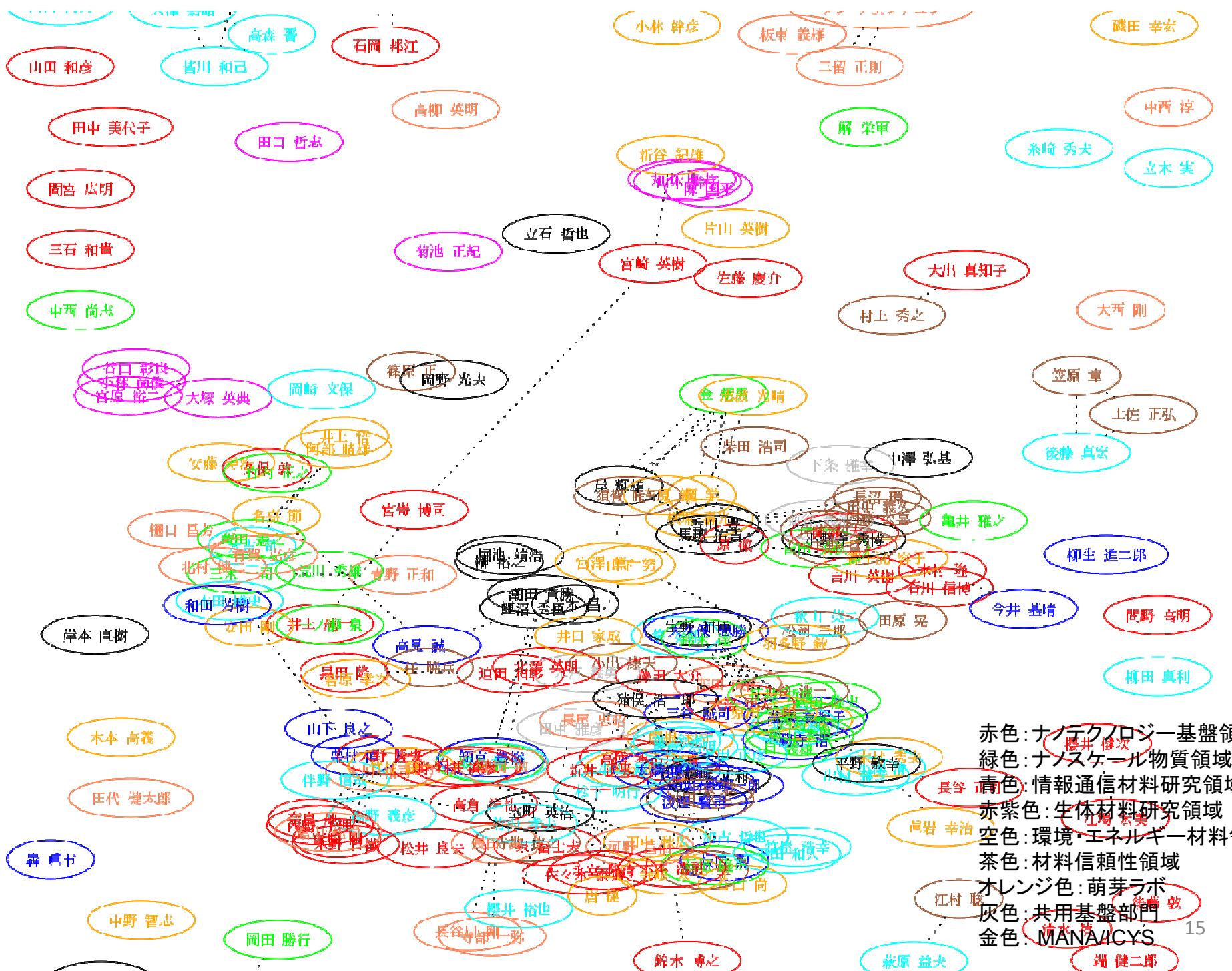
- 256名が科研費データベース中のレコードと
同定
 - 元データ集合の20.8%

※NIMSにおける科研費採択研究者は統計分野に比べると、やや少ない。
法人化以前の国研時代に科研費申請が行われていなかったため。

応用：研究分野の概観に向けて

- 研究者ネットワークの可視化
 - 科研費データベースに蓄積されている約15年分の研究者の共同研究関係
 - (研究代表者↔研究分担者)
 - 各研究者をノード、共同研究関係をエッジとするネットワークが可視化可能
- 統計分野3学会所属者の共同研究関係
 - 学会分野と共同研究関係の可視化
- NIMS所属研究者の共同研究関係
 - 所内部署単位の可視化
 - 分野・外部機関単位での共同研究可能性をさぐるマップ

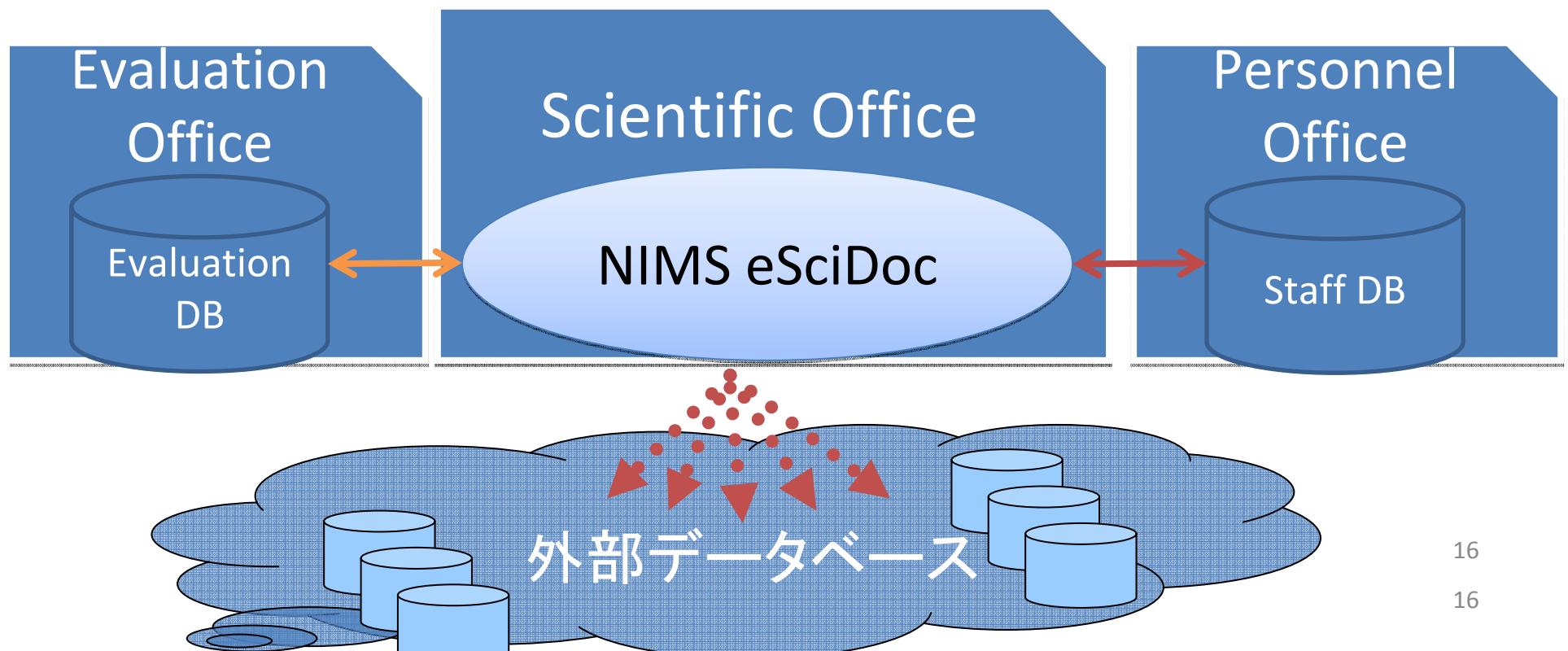




赤色: ナノテクノロジー基盤領域
緑色: ナノスケール物質領域
青色: 情報通信材料研究領域
赤紫色: 生体材料研究領域
空色: 環境エネルギー材料領域
茶色: 材料信頼性領域
オレンジ色: 萌芽ラボ
灰色: 共用基盤部門
金色: MANA/ICYS

応用：研究者情報の集約と交換に向けて

- NIMSデータセット
 - 機関リポジトリ ⇔ 外部データベース間での研究者情報の交換
 - 対象DB: ResearcherID, (研究者リゾルバ)



機関リポジトリにおける研究者総覧 (NIMS eSciDoc)

eSciDoc.CoNE English ▾

Researcher Portfolio

PhD Todoroki, Shin-ichi
National Institute for Materials Science, Optronic Materials Center



Researcher Profile	Current Position	Senior Researcher (National Institute for Materials Science)
	Current Position	Faculty Member (Optronic Materials Center)
	Researcher Homepage(s)	Researcher Blog
	Researcher ID	urn:cone:persons1008
	Research Fields	Computer Sciences: Free, Libre and Open Source Software; Information Science; Digital Libraries Engineering: Materials Engineering; Ceramic Engineering; Optical Engineering Informatics Library Science: Digital Libraries Physics: Materials Science
External references	WorldCat	Search for Todoroki, Shin-ichi
	Google Scholar	Search for Todoroki, Shin-ichi
Publications		

まとめ

- 研究者同定
 - 学術分野のヒト同定
 - 統計分野、材料研究所分野の数千人規模 vs 15万人規模
- 人手判定付加の軽減に向けて
 - 現在のデータセットで作成した判定結果を教師データとして学習判定できるようなツールの構築へ
- 研究者情報の一元的やり取りに向けて
 - NIMS ⇔ ResearcherID.com (Thomson Reuter), 研究者リゾルバ (NII), Read (JST), Scopus (Elsevier), ...
 - 1対1のサービス間やりとりだけでなく、研究者情報の相互互換的やりとりに向けた取り組みが必要となってくる