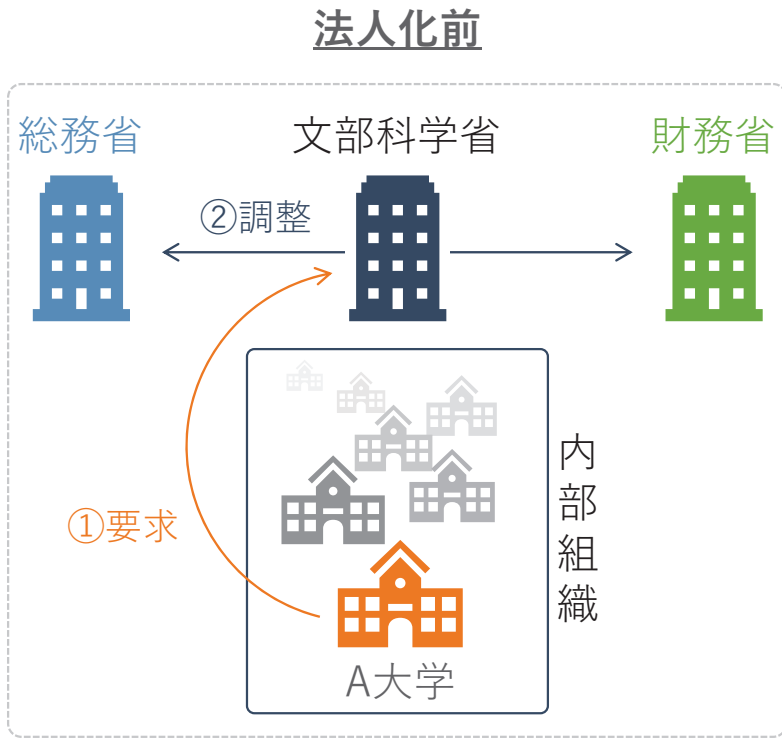


九州・沖縄オープンユニバーシティ 研究支援人材WG合同研修 2025/7/8

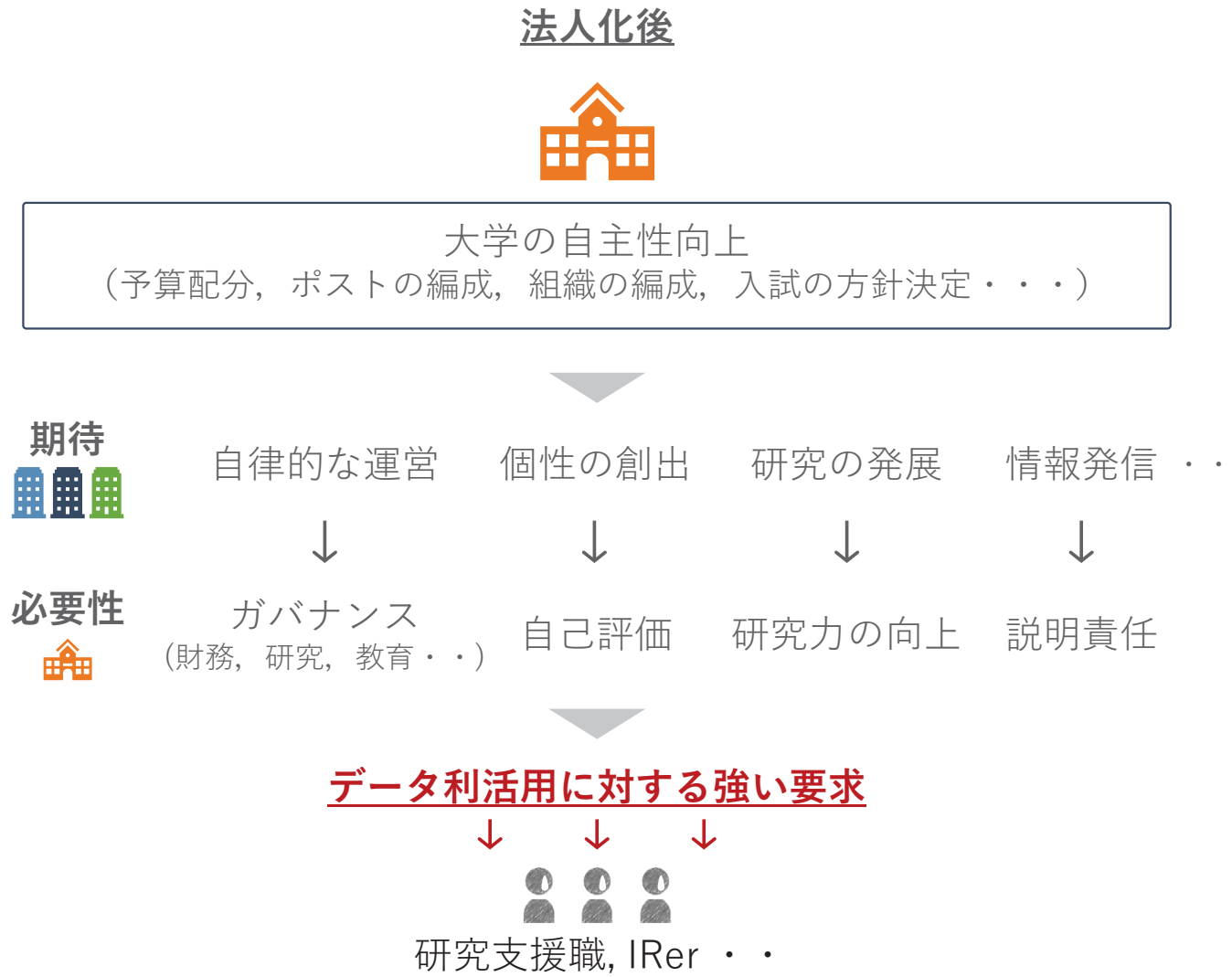
## 書誌情報が持つポテンシャルとその限界

九州工業大学 山本 鉦

# 国立大学を取り巻く環境の変化とデータ利活用の意義



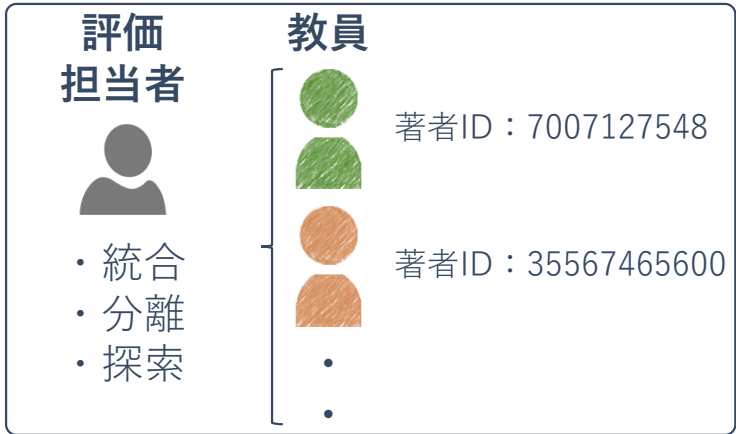
大学の望みが直ぐに叶えられない



# Contents

- 01 書誌情報に基づく分析の位置づけ
- 02 学外への研究力の発信
- 03 研究分野を問わない評価指標の開発
- 04 組織の研究力分析手法の確立
- 05 シーズニーズマッチングシステムの開発
- 06 まとめ

学外への研究力の発信（仕組み）

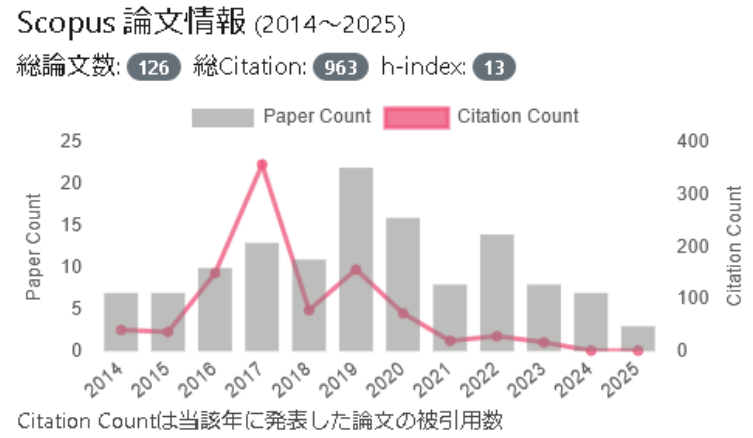
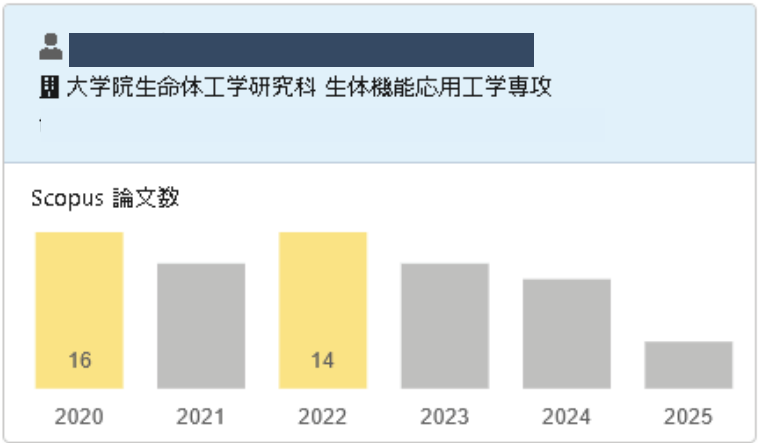


Scopus著者IDの整理（年1回）

56796331100  
35567465600  
.  
著者IDリスト



集計結果を個人単位で表示



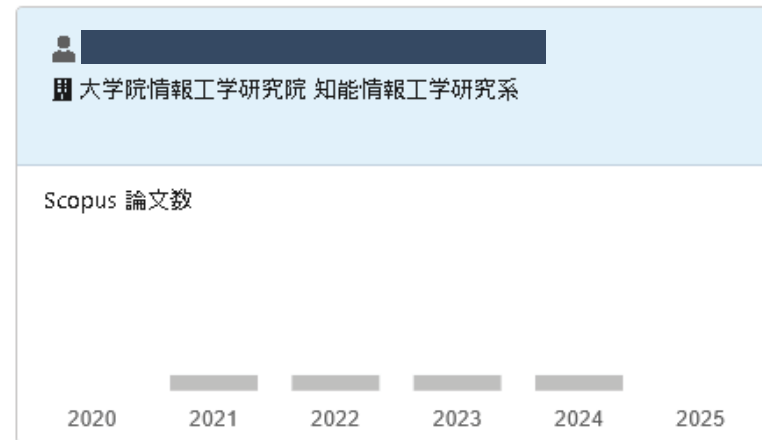
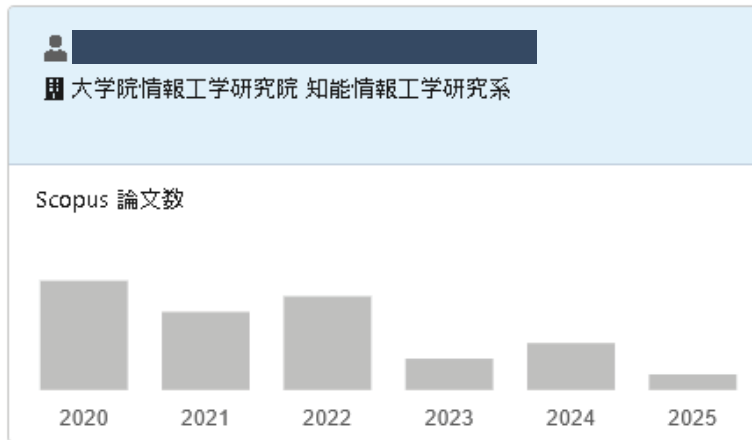
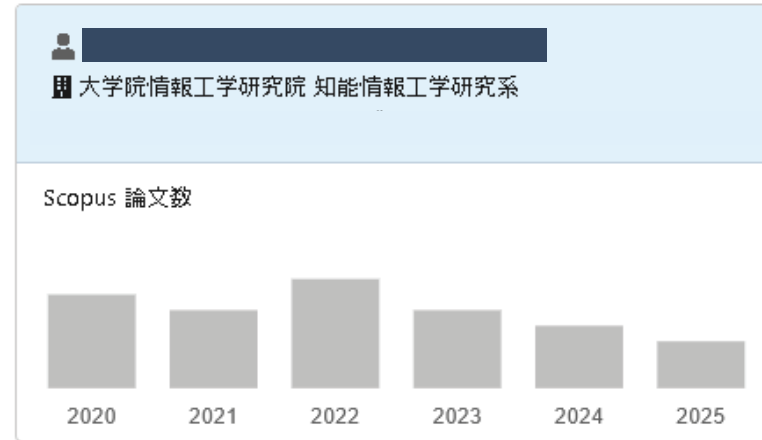
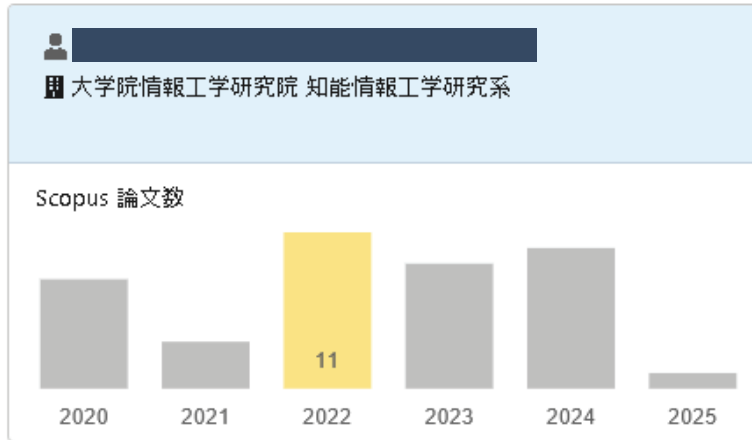
各教員の入力に頼らず最新情報を表示可能 → 他大学でも利用

## 研究者情報



検索結果

市一社



同一所属組織内での比較

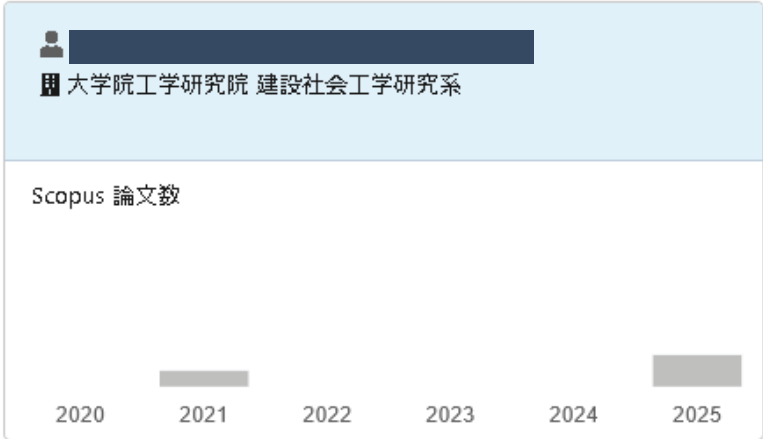
# 学外への研究力の発信（注意点）

🔍 検索条件 【キーワード検索】

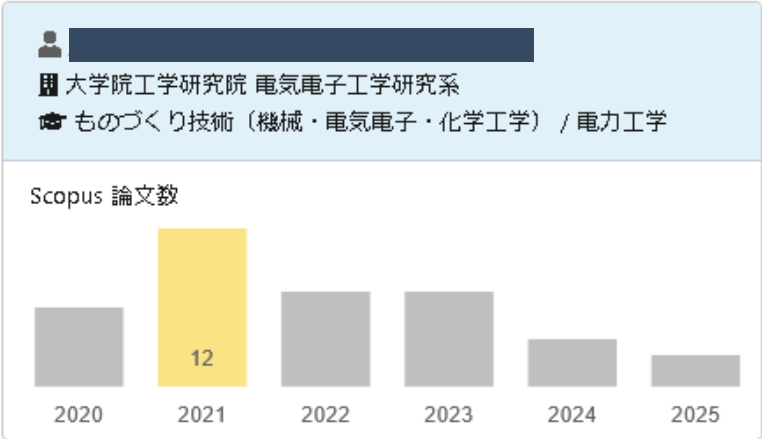
ヒートポンプ 🔍 検索する

AND検索 OR検索 📌 使い方

各年のScopus論文数のグラフは上限を10報としています。  
 Scopus掲載論文以外も含めた業績は、各研究者の詳細をご確認ください。  
 Scopusにおける名寄せの状況などにより、本来の論文数と差異が生じる場合があります。



建築学系



電気電子工学系

研究分野毎の特性を考慮した指標が必要



# データの取得から分析まで

学術文献DB : Scopus  
 大学数 : 37 (国内)  
 著者ID数 : 25,679  
 対象期間 : 10年間  
 論文区分 : Article  
                   Review  
                   Conference Paper  
 取得年月 : 2024/12  
 更新頻度 : 年1回

論文数 : 471,534

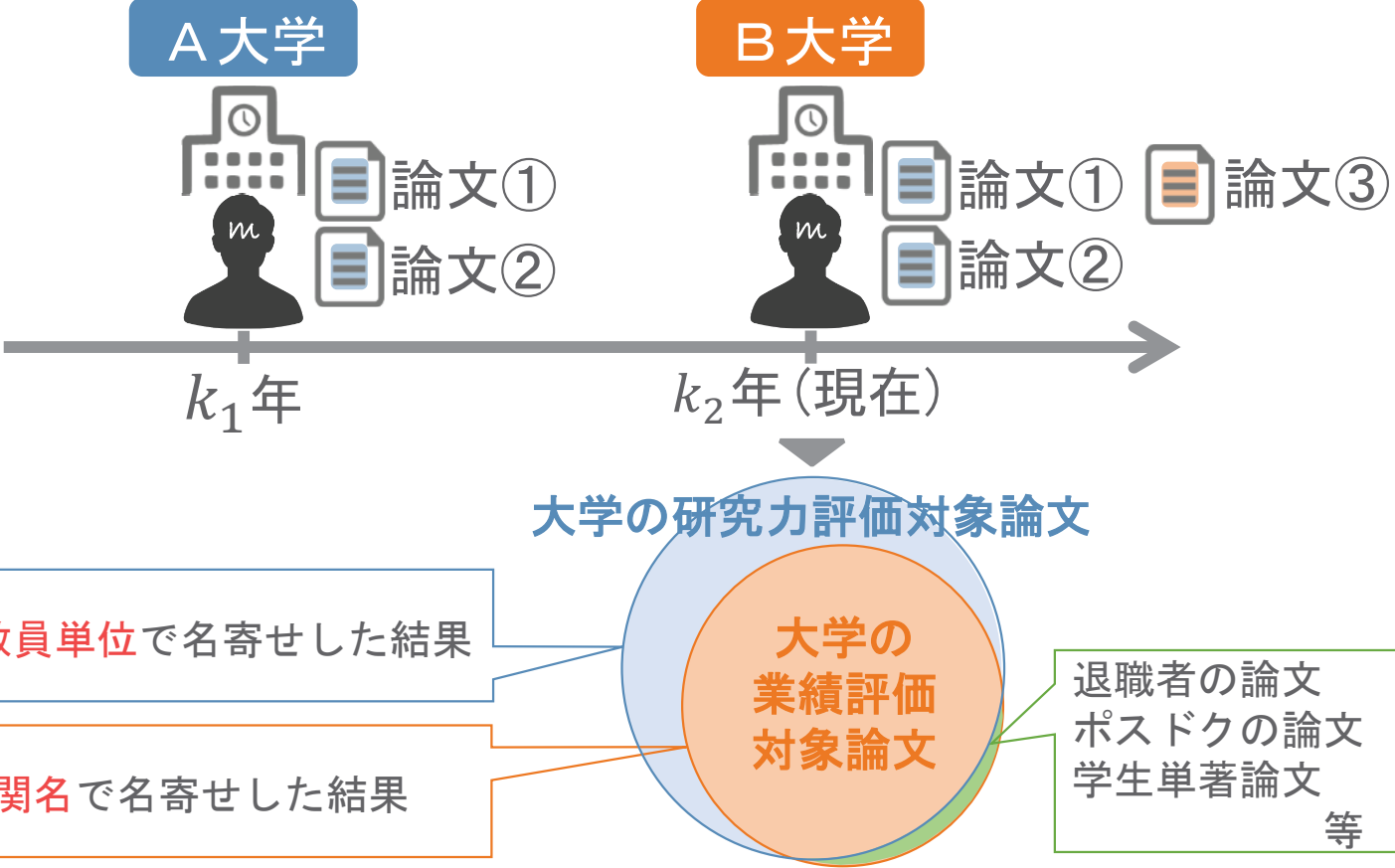
- 出版年
- 総被引用数
- 単年毎の被引用数 (max : 出版年から10年)
- FWCI
- FWCIに基づくパーセンタイルランク
- **ASJC27**, 334分野のコード
- DOI



データセットの違い

$$\sum_{k=k_1}^{k_2} \sum_{i=1}^N (\alpha_{i.})_k = \text{大学の現在の研究力} \neq \text{大学の研究業績}$$

$N$  教員数  
 $(\alpha_{i.})_k$  ある教員の論文数



研究力分析のためのデータセット作成は難しいが、改善のための施策立案には不可欠

## 正規化論文数の計算方法

- 前提条件 1: Scopusへの収録論文数が0の教員は研究分野が未知のため除外  
 2: 論文数は共著者数で割る(共著者数2なら1/2本: 分数カウント法)  
 3: **1人の教員が対象期間の論文執筆に要した労力を1マンパワーと定義**

仮定 : 教員  $i$  の分野  $j$  への投入マンパワーは論文数に比例する

- ① 分野  $j$  の平均的な論文の出易さ

$$P_j = \frac{\text{分野 } j \text{ の全論文数}}{\text{分野 } j \text{ への全投入マンパワー}} \quad \leftarrow \bar{X}$$

- ② 教員  $i$  の分野  $j$  における正規化論文数

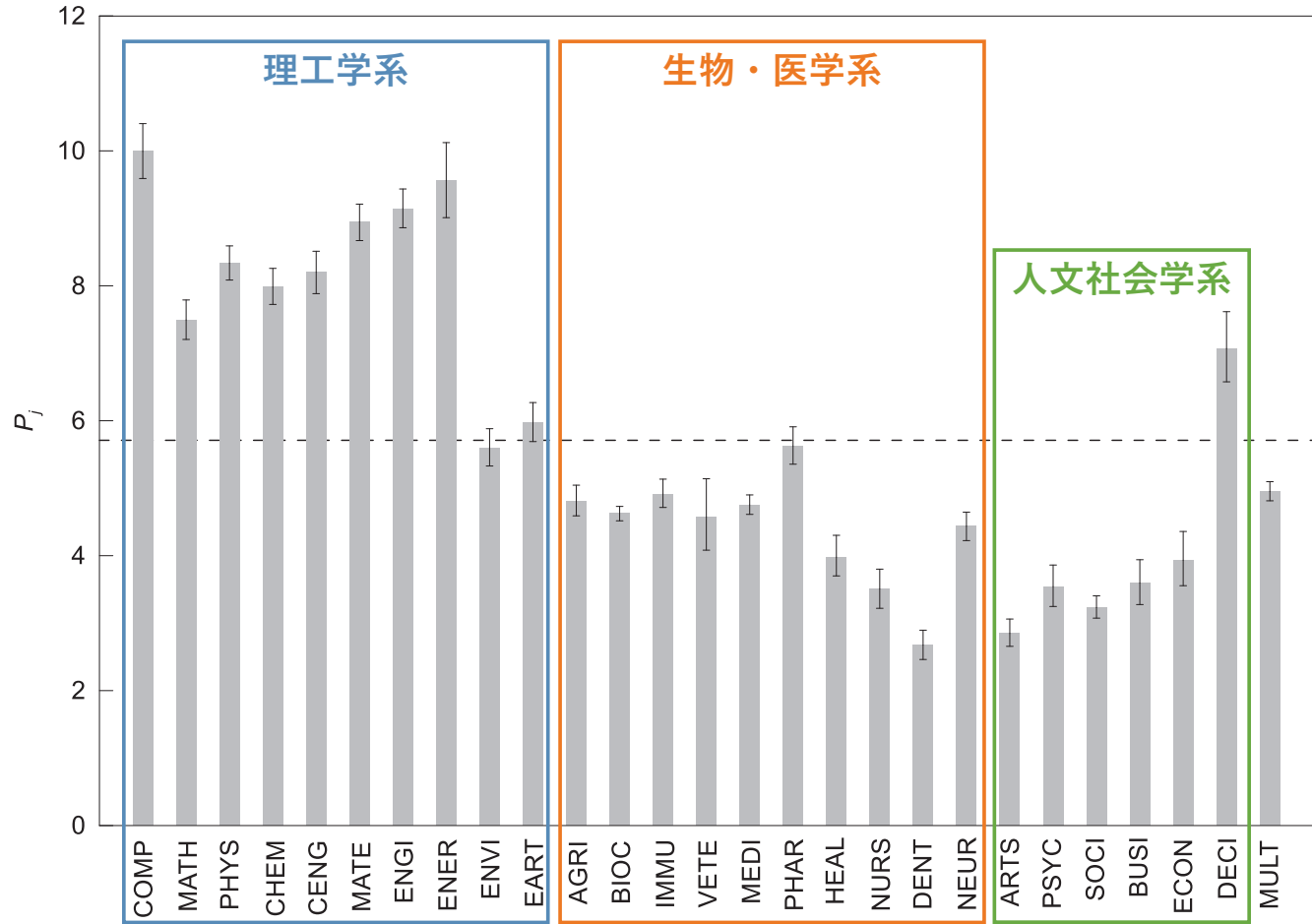
$$Y_{ij} = \frac{\text{教員 } i \text{ の分野 } j \text{ における論文数}}{P_j} \quad \leftarrow \frac{X_i}{\bar{X}}$$

- ③ 教員  $i$  の正規化論文数

$$Y_{i\cdot} = Y_{ij} \text{ の全分野合計}$$

**全教員の正規化論文数の平均は1 (1以上: 平均よりも論文を多く発表と解釈)**

平均論文生産性 (10年間で分数カウントで1人平均何本の論文を發表するか)



COMP : Computer Science  
MATH : Mathematics  
PHYS : Physics and Astronomy  
CHEM : Chemistry  
CENG : Chemical Engineering  
MATE : Materials Science  
ENGI : Engineering

ENER : Energy  
ENVI : Environmental Science  
EART : Earth and Planetary Sciences  
AGRI : Agricultural and Biological Sciences  
BIOC : Biochemistry, Genetics and Molecular Biology  
IMMU : Immunology and Microbiology  
VETE : Veterinary

MEDI : Medicine  
PHAR : Pharmacology, Toxicology and Pharmaceutics  
HEAL : Health Professions  
NURS : Nursing  
DENT : Dentistry  
NEUR : Neuroscience  
ARTS : Arts and Humanities

PSYC : Psychology  
SOCI : Social Sciences  
BUSI : Business, Management and Accounting  
ECON : Economics, Econometrics and Finance  
DECI : Decision Sciences  
MULT : Multidisciplinary

分野単位で見ると計量書誌学で周知の傾向と類似

正規化前後の学内順位差

ID	論文数 (正規化前)	論文数 (正規化後)	順位 (前)	順位 (後)	
2***0553200	219	11.3	3	1	
6***6550800	158	7.9	6	2	
3***4317200	185	7.5	4	3	
3***9704885	122	7.0	15	4	
<b>5***5354100</b>	<b>27</b>	<b>6.9</b>	<b>296</b>	<b>5</b>	<b>社会学系</b>
6***3334900	132	6.3	12	6	
3***5742002	137	6.2	11	7	
<b>8***1575300</b>	<b>38</b>	<b>5.9</b>	<b>212</b>	<b>8</b>	<b>心理学系</b>
9***9783498	155	5.8	7	9	
・					
・					
<b>1***3706700</b>	<b>262</b>	<b>3.6</b>	<b>1</b>	<b>24</b>	<b>宇宙物理学系</b>
・					
・					

順位(前) : 正規化前の論文数を基にした学内順位  
 順位(後) : 正規化論文数を基にした学内順位

より公平な評価に資する結果

計量書誌学での一般的な正規化手法

~~$$X' = \frac{X - \text{Min}}{\text{Max} - \text{Min}}$$~~  
 範囲(0~1)を固定

$$\frac{X_i}{\bar{X}}$$

- 論文  $i$  の被引用数 : FWCI, CNCI
- 研究者  $i$  の論文数 : 2014~2022
- 研究者  $i$  の被引用数 : 2023~  
など

$X_i$ が属するグループの平均値 (基準値)

↓  
 基準値の精度の確保  
 ↓  
 基準値の信頼性の確認

Moed, HF. et al. in Handbook of quantitative science and technology research (2004) 19-50  
 Waltman, L. et al. J. Informetrics. 5, 37-47 (2011)  
 Abramo, G., & D'Angelo, C. A., J. Scientometrics. 101, 1129-1144 (2014)

$P_j$ の国内平均（推定値）との比較：2023年データ（Article, Reviewのみ）

グループ	人数シェア	大学数	研究者数	層化抽出人数
1	11%	2	3,708 (15%)	941 (11%)
2	19%	9	11,385 (45%)	1,591 (19%)
3	20%	11	5,900 (23%)	1,753 (20%)
4	50%	14	4,268 (17%)	4,268 (50%)
総計	100%	36	25,261 (100%)	8,553 (100%)

高  
生産性  
低

層化抽出法に基づく国内の平均論文生産性（推定）

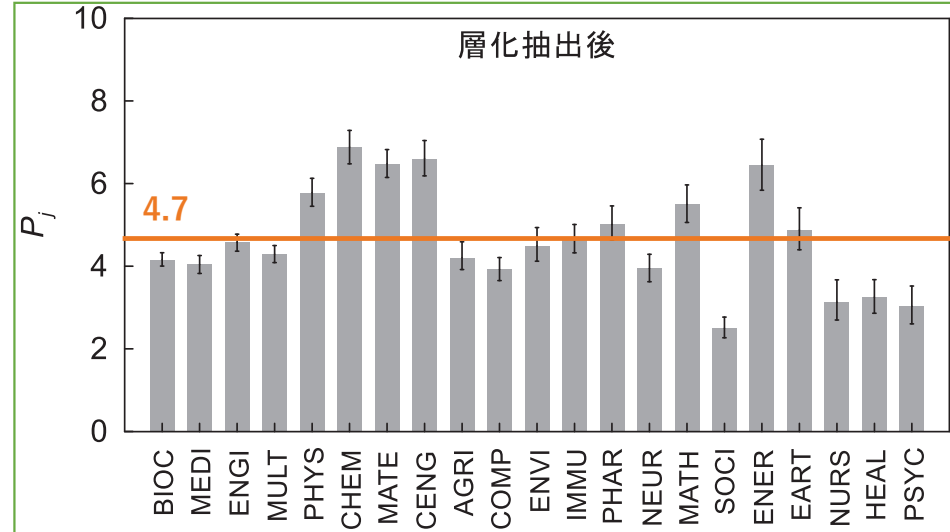
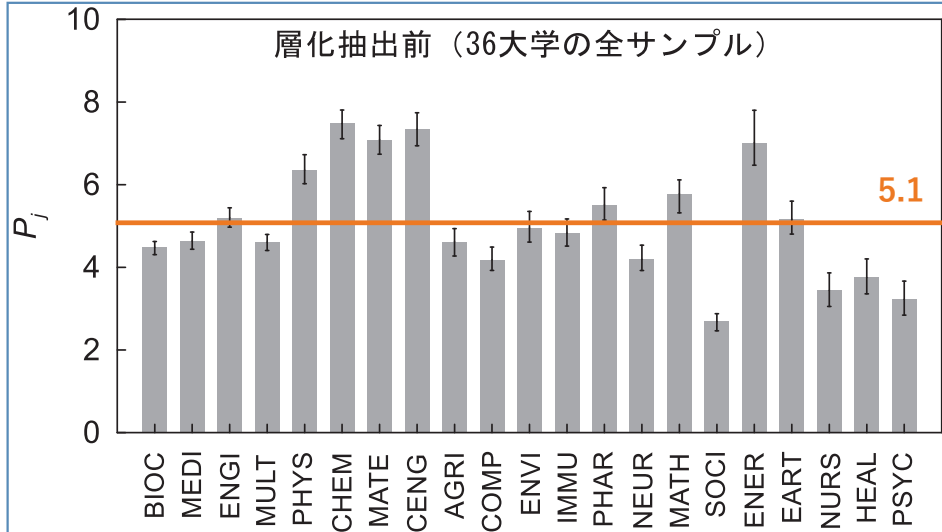
↓

平均は減少，分野間の相対的な大小関係は変わらず

論文シェアに基づく大学のグループ分けと，各グループの人数シェア

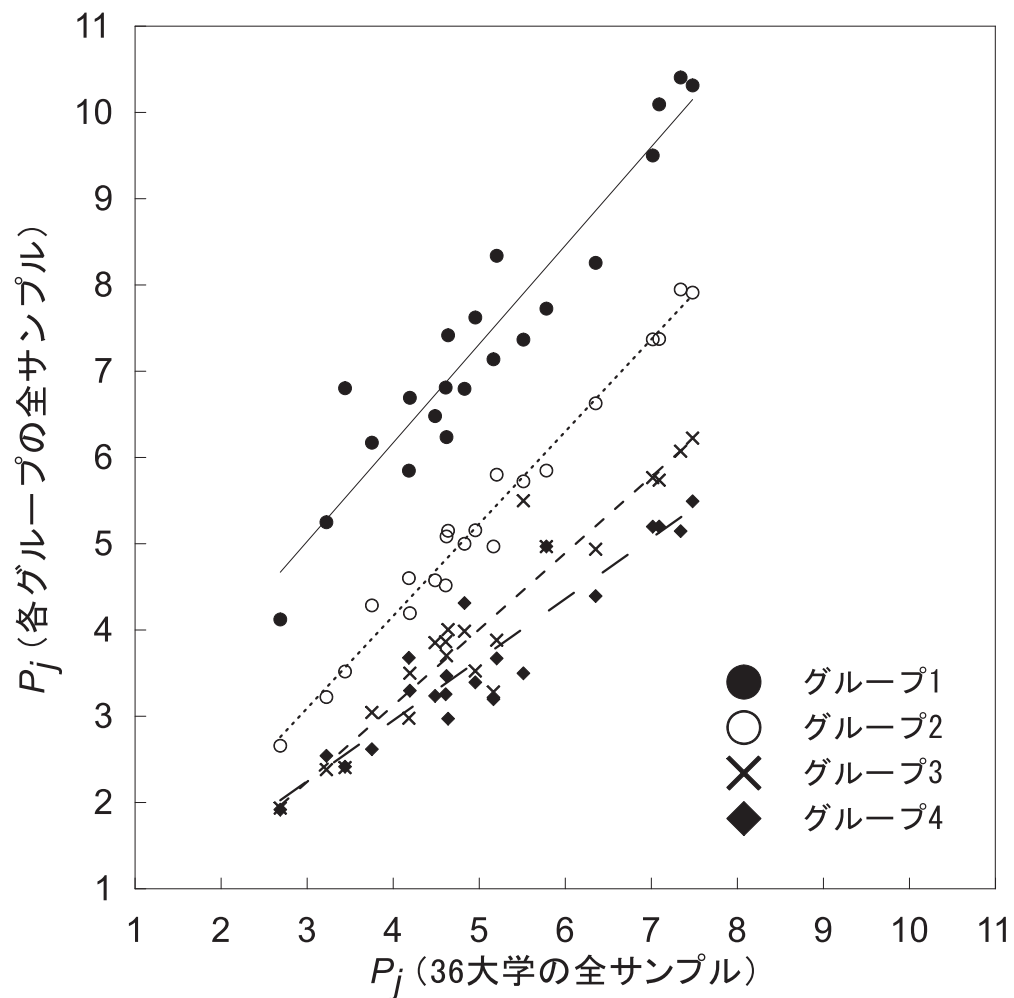
約34%

村上昭義, 伊神正貫, 科学技術・学術政策研究所, Tech. Rep. (2018)



複数大学の $P_j$ は高い方に偏向するが，分野間の相対的な大小関係は類似

## 相対的な評価を行ううえでのデータ収集



$P_j$ の分野間の大小関係はグループ間で類似



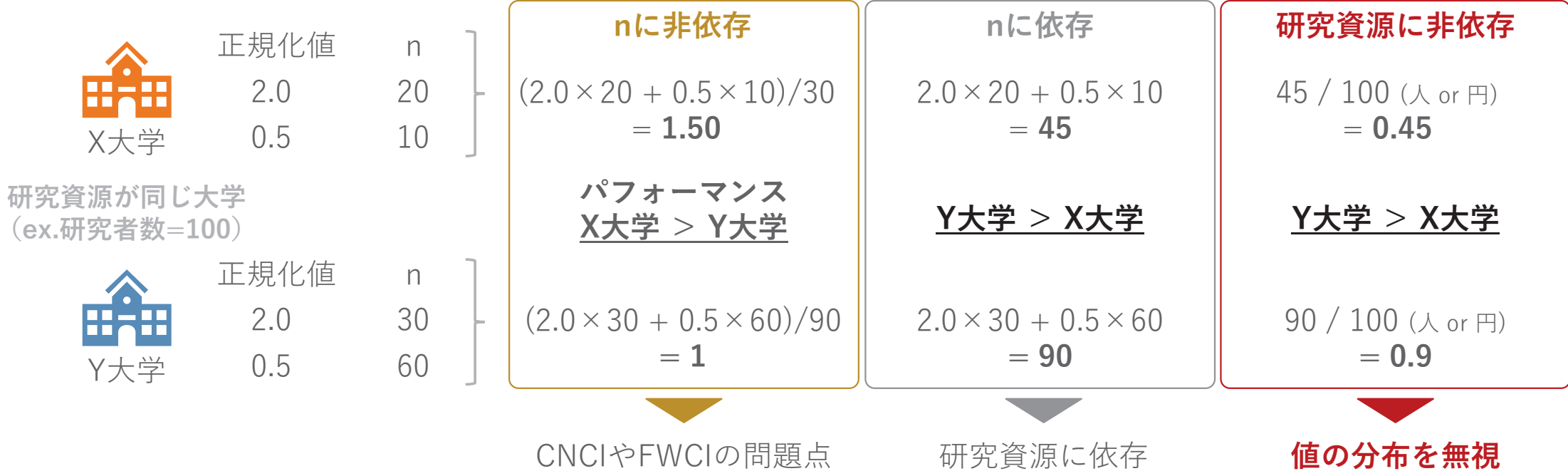
サンプルの偏りが全体の $P_j$ の  
大小関係に与える影響は小さい



正規化論文数による相対的な評価には  
大きな影響を及ぼさない

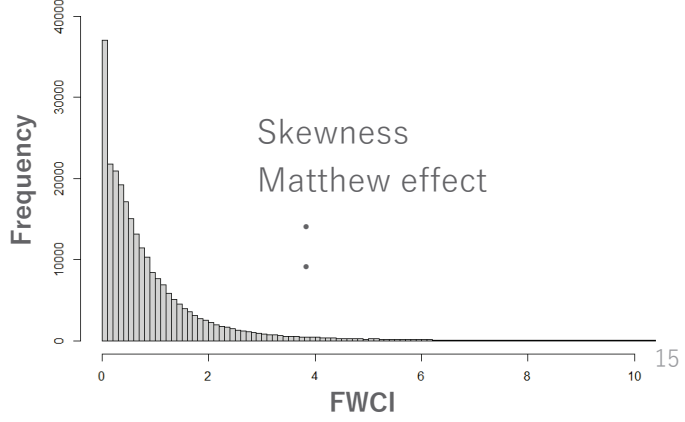
信頼性を確保するために国内全大学のデータを収集する必要は無い

個(人・論文)の評価から組織の評価へ



分析結果の意味を正確に把握すること (リテラシー) が重要

Seglen, P. O., J. Am. Soci. Inf. Sci. 43, 628-638 (1992)  
 Waltman, L. et al. J. Informetrics. 5, 37-47 (2011)  
 Perc, M., J. Royal Soc. Interface. 11, 20140378 (2014)  
 Abramo, G., & D'Angelo, C. A., J. Scientometrics. 101, 1129-1144 (2014)  
 Abramo, G., & D'Angelo, C. A., J. Informetrics. 10, 646-651 (2016)  
 Waltman, L., et al. J. Informetrics. 10, 671-674 (2016)



## パーセンタイルランクを活用した指標（先行研究）

$$R(6) = \sum_{k=1}^l k \cdot p(k)$$

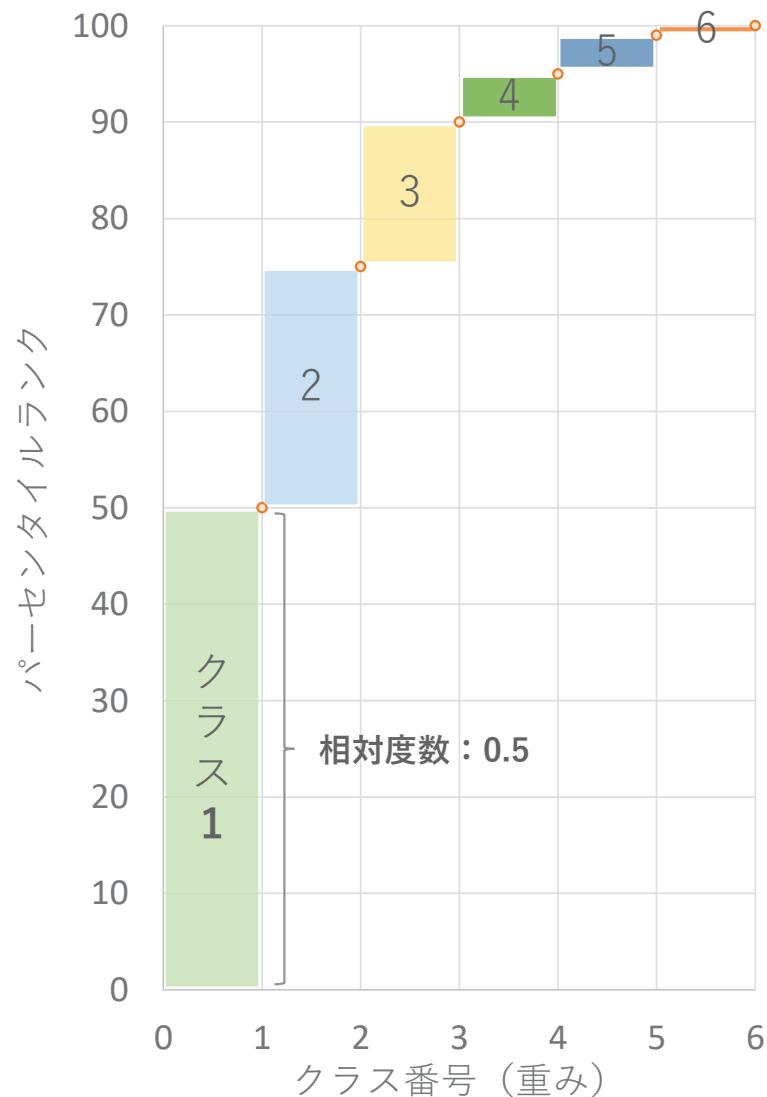
クラス番号（重み）

クラスkの相対度数

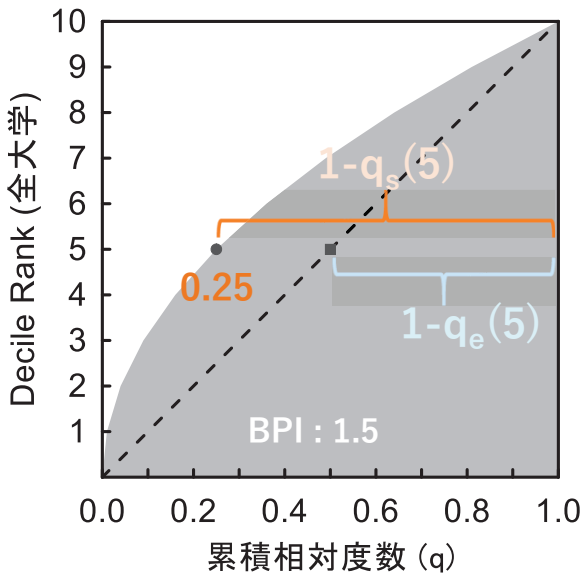
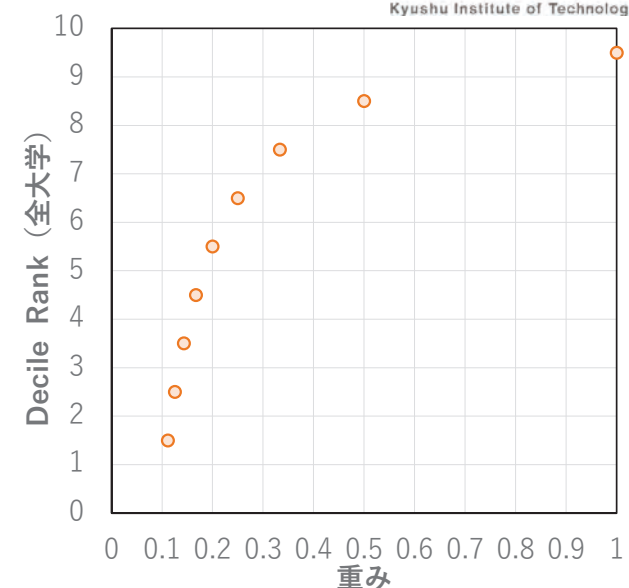
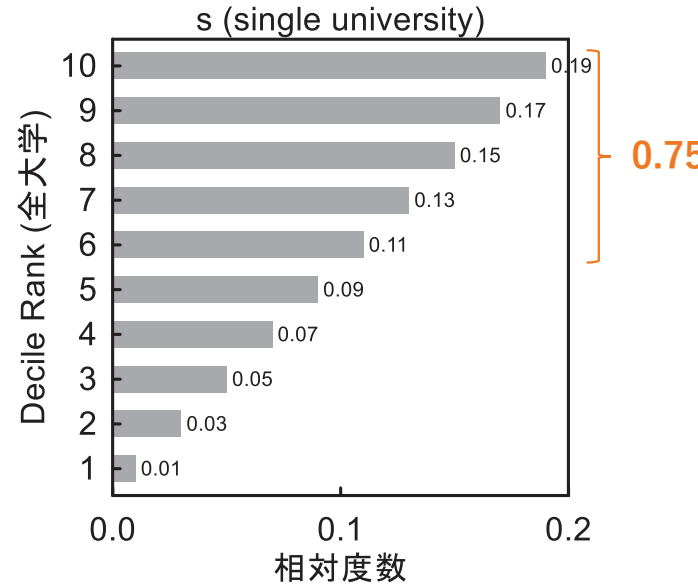
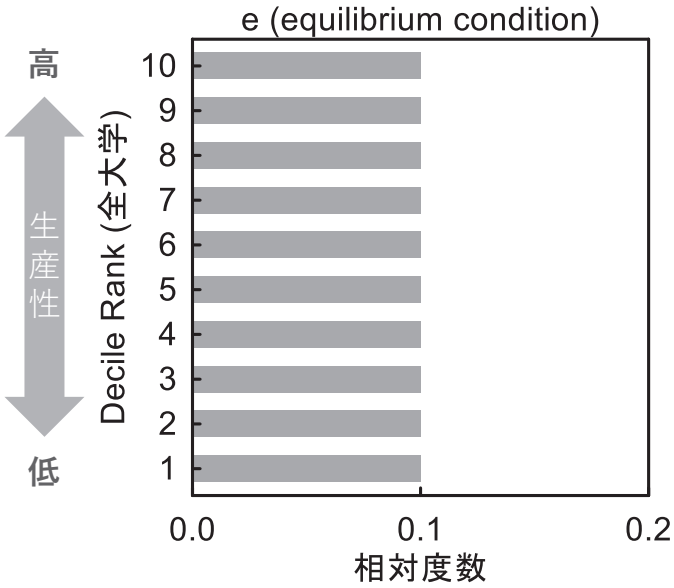
- ・ 被引用数を活用した研究者評価のための指標
- ・ 被引用数に基づくパーセンタイル（WoS）を活用
- ・ 6つに分割（50, 75, 90, 95, 99パーセンタイルで区切る）  
→ 重みは非線形に増加 → 卓越する論文に大きな重み

- ▼
- ・ 分割方法が恣意的
  - ・ 分割数は変更不可

期待値： $0.5 \times 1 + 0.25 \times 2 + 0.15 \times 3 + 0.05 \times 4 + 0.04 \times 5 + 0.01 \times 6 = 1.91$



パーセンタイルランクを活用した指標 (BPI: Better Performance Index)



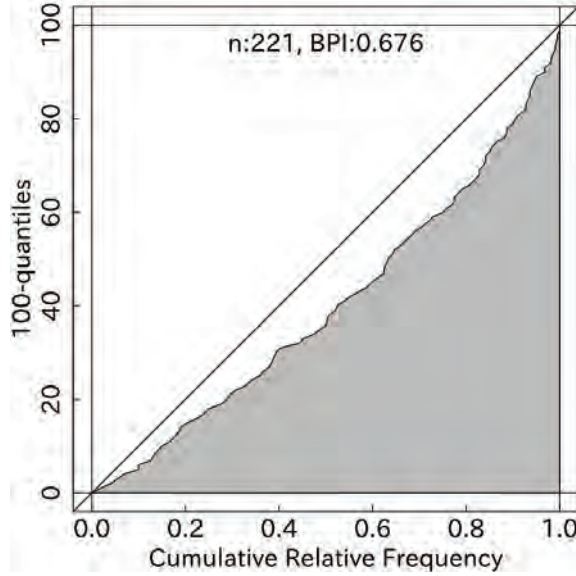
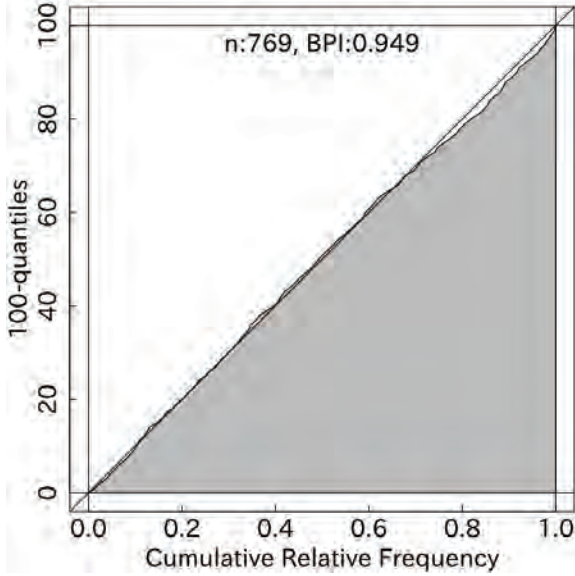
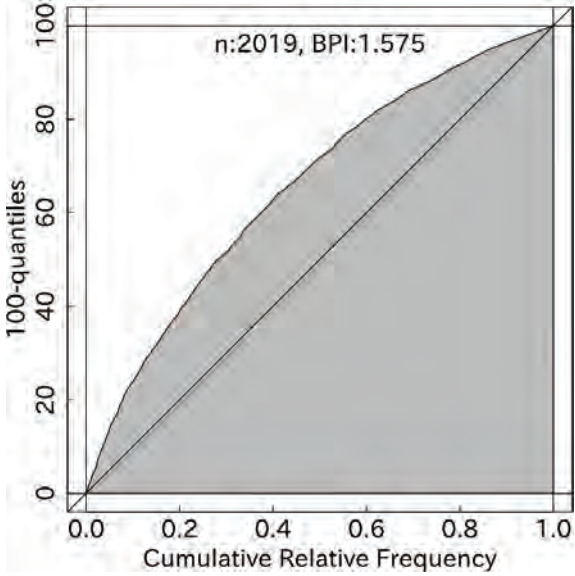
$1-q_s$  : より上位のランクに位置する教員比率 (実際)  
 $1-q_e$  : より上位のランクに位置する教員比率 (期待)

$$BPI(l) = \frac{1}{l-1} \sum_{k=1}^{l-1} \frac{1-q_s(k)}{1-q_e(k)} \quad \leftarrow \text{期待値: 1}$$

$1-q_s(k)$  に対する重み  $1/[1-q_e(k)]$  は,  $k$ が増えるとともに非線形に増加する.

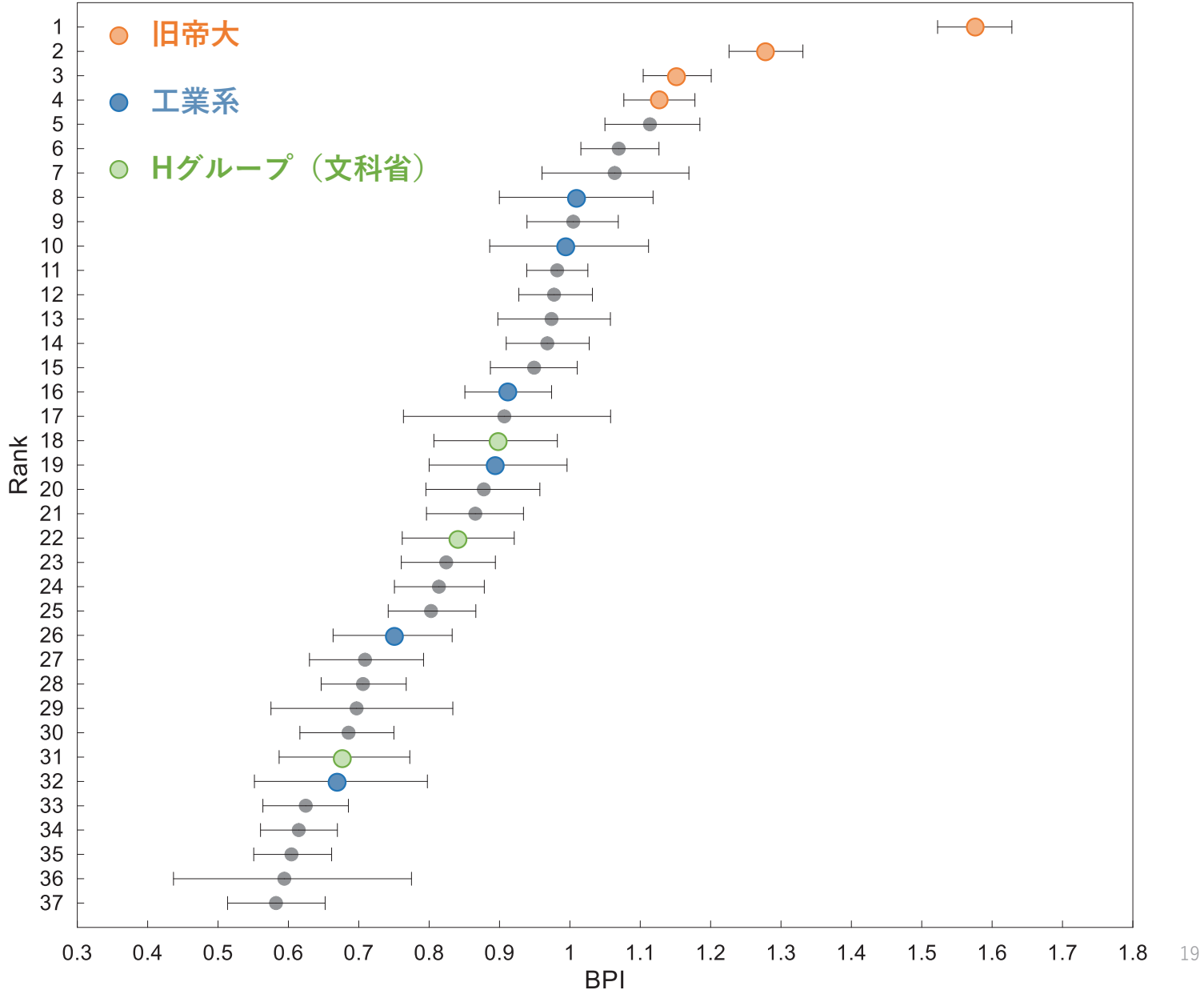
特徴: 上位のランクに位置する教員比率の大きさを重視

# 大学間の比較（全体）



正規化指標 × BPI → 研究分野分布や組織の規模の違いを超えた研究力の比較

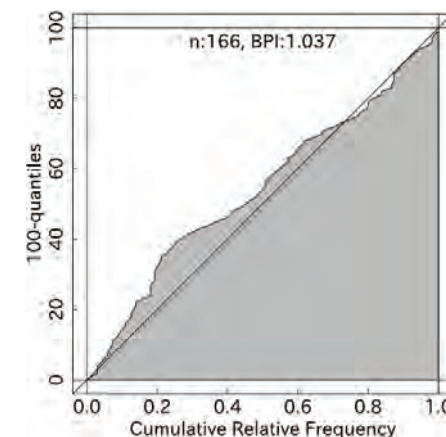
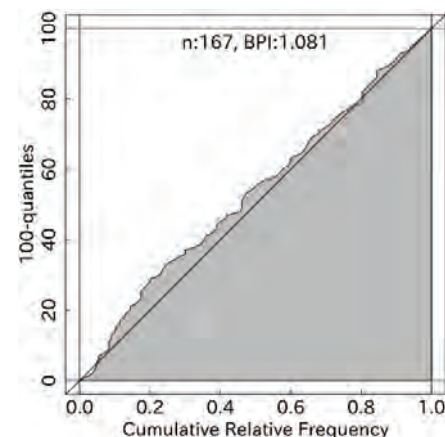
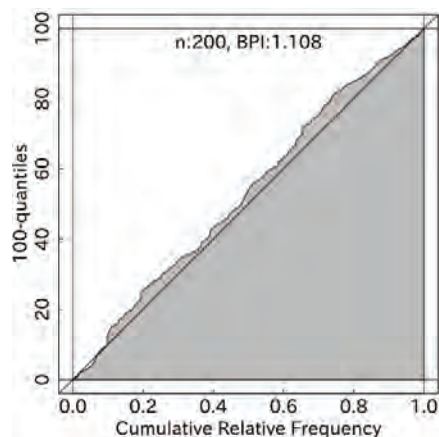
# 大学間の比較（全体）



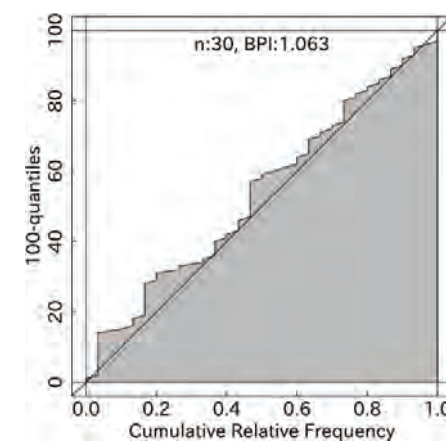
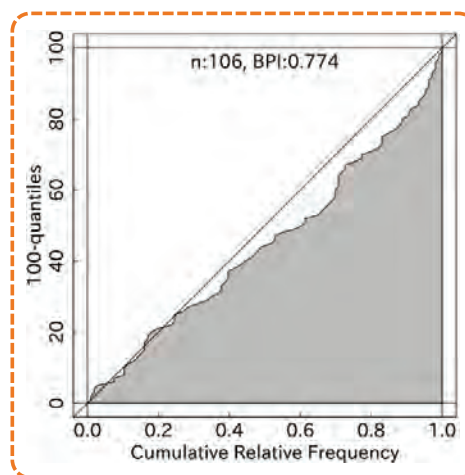
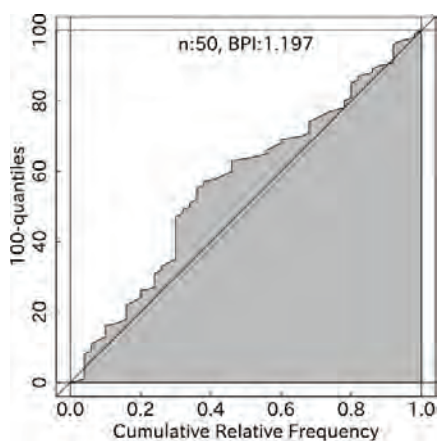
研究分野分布が近い大学と比較  
 (ASJC27は分類の粒度が粗いため)

# 大学間の比較 (学系)

工学系



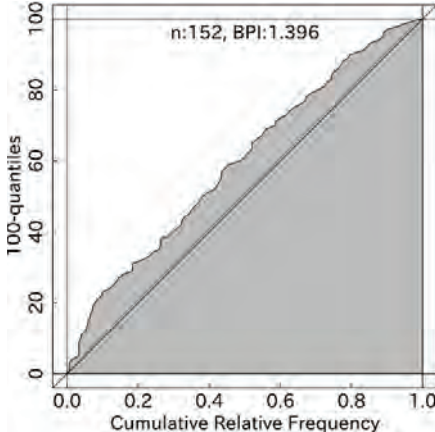
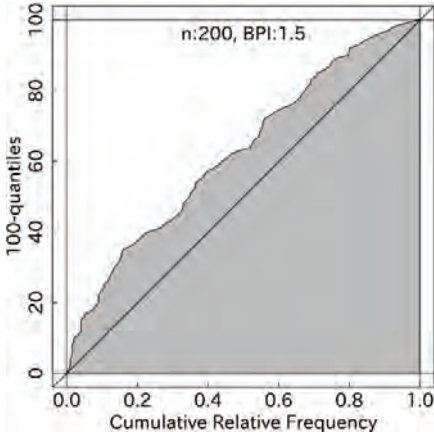
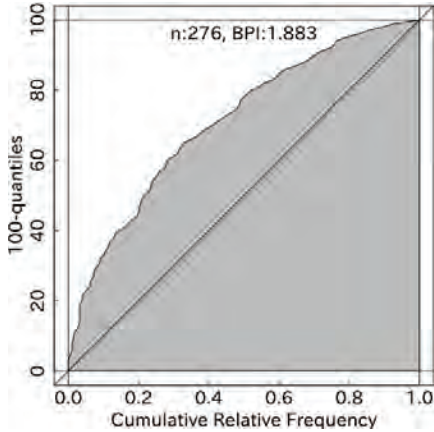
情報学系



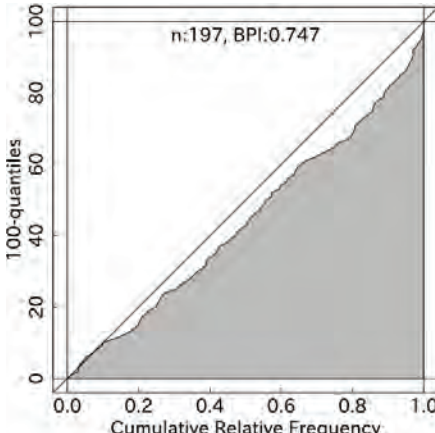
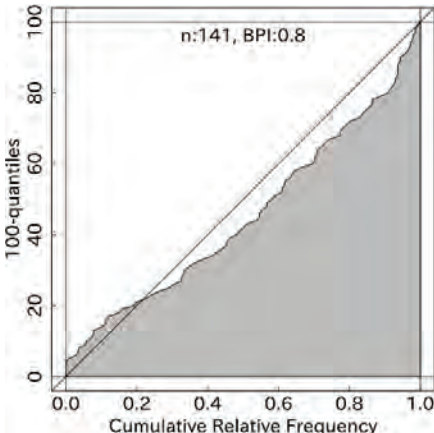
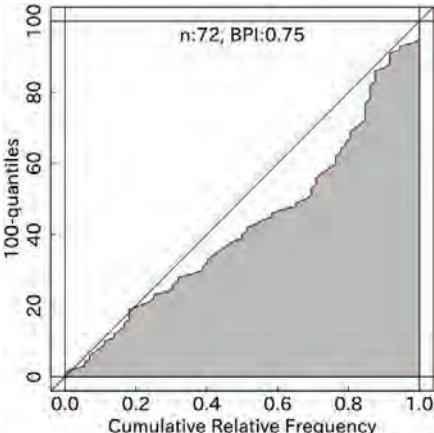
改善の余地あり？

# 大学間の比較 (学系)

基礎医学系

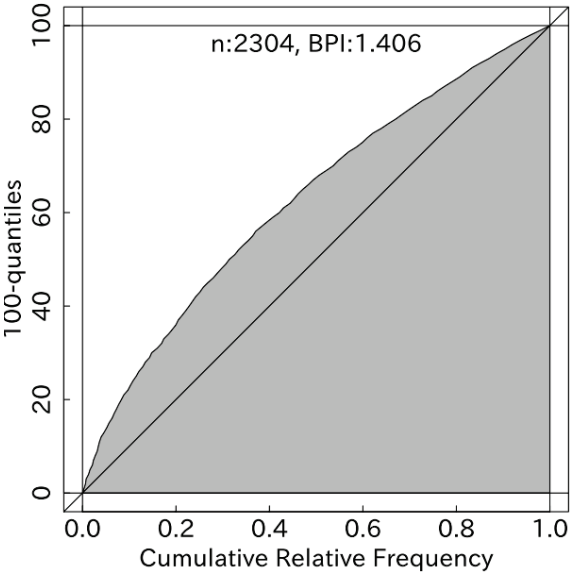
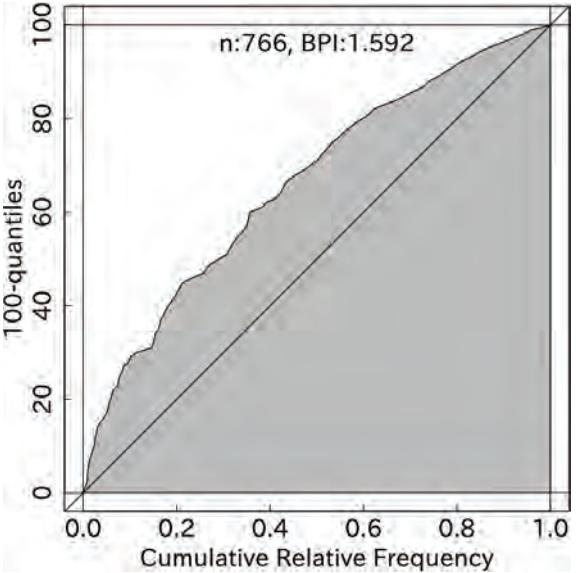
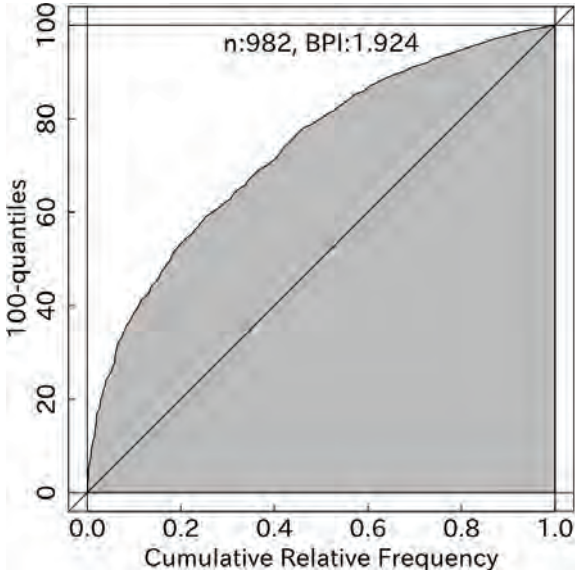


臨床医学系



臨床医学系はどの大学もScopus収録論文数が少ない傾向

# 大学間の比較（全体：国際）

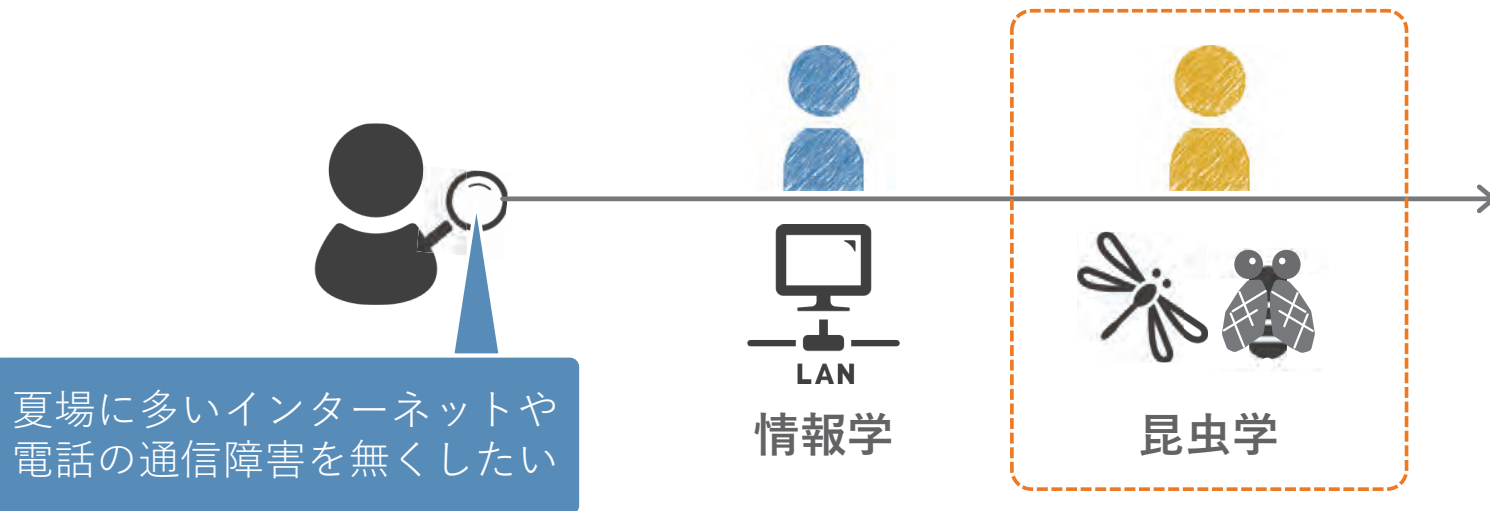


※ 分析データ内での相対評価のため、対象年や対象大学の違いによりBPIが異なる

継続的な改善：適切なベンチマーク対象の選定が必要

## コンセプト

- ・ キーワードに文字揺れがあっても探索できること
- ・ 在職中の教員を対象に探索すること
- ・ 学習データの更新が容易であること
- ・ 思いもよらない研究者を探索できること



## 概要

- ・このAIは公開データのみ使用して構築
- ・クラウドは活用せず，AIやデータは九州工業大学のサーバ内に存在
- ・利用履歴は非保持



  
大学の担当者が  
年に一回更新



## 特徴①

複数大学の複数分野の研究者を探索することが可能



### 2025年7月：24大学+2機構

- ・ 帯広畜産大
- ・ 室蘭工業大
- ・ 東北大
- ・ 新潟大
- ・ 静岡大
- ・ 富山大
- ・ 名古屋大
- ・ 名古屋工業大
- ・ 奈良先端大
- ・ 岡山大
- ・ 九州大
- ・ 九州工業大
- ・ 佐賀大
- ・ 長崎大
- ・ 熊本大
- ・ 大分大
- ・ 宮崎大
- ・ 琉球大
- ・ 名古屋市立大
- ・ 北九州市立大
- ・ 産業医科大
- ・ 豊田工業大
- ・ 立命館大
- ・ 日本文理大

## 特徴②

日本語でも英語でも研究者を探索することが可能

## 機械学習

人工知能の手法の一つ、データの中の**規則性やルール**を見つけ出し**予測や判断**を行う手法

### 教師あり学習

学習データに**正解ラベル**を付けて学習する手法

#### 回帰

数値を予測するアルゴリズム

#### 分類

正解ラベル通りにグループ化を行うアルゴリズム

### 教師なし学習

学習データに**正解ラベル**を付けずに学習する手法

#### クラスタリング

似た特徴を持つデータをグループ化するアルゴリズム

#### 次元圧縮

データのサイズを小さくするアルゴリズム

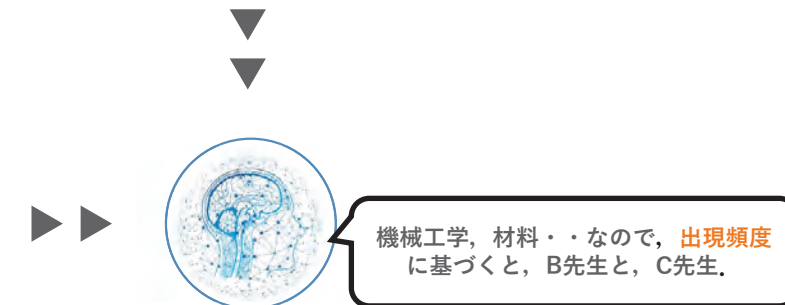
### 強化学習

教師あり学習のように、与えられた正解ラベルをそのまま学習するのではなく、将来的に問題を解決した際に得られる**利益が最大になるように学習**する手法。

## 学習

文章 (学習データ)		研究者
論文1	Progress of Photothermal and Photoacoustic Science and Technology - Semiconduc . . .	A教授
論文2	Photoacoustic and Photothermal Phenomena in Semiconductors North Holland T. I . . .	
論文3	薄膜シリコン系太陽電池の最新技術 シーエムシー出版、2009年7月、pp.107-117 碓 . . .	
論文4	シンポジウムシリーズ2「分離技術のシーズとライセンス技術の実用化」 分離技術会 . . .	
論文5	粉体・微粒子分析 テクニック事例集、第10章第15節 液中乾燥法を用いて調製した農 . . .	
論文6	多孔質マイクロカプセルに担持したLIX抽出剤の銅の抽出過程のFI-IRによる . . .	
論文7	マイクロ/ナノカプセルの新規調製と次世代製品開発技術 (株)技術情報協会 塩盛弘 . . .	
論文8	Dynamics and Control of Advanced Structures and Machines -Contributions from . . .	B教授
論文9	Mechanicsmechanical engineering and Model-Based Control of Smart Materials and . . .	
論文10	Encyclopedia of Thermal Stresses Springer Y. Tanigawa, R. Kawamura Analytical . . .	
論文11	Encyclopedia of Thermal Stresses Springer R. Kawamura Nonhomogeneous plate . . .	
論文12	設計する際には、機械や製品が壊れないように高すぎる応力が発生している箇所は . . .	C助教
論文13	JSMEテキストシリーズ 演習 材料力学 -Problems in Mechanics of Materials- 丸善 . . .	
論文14	材料力学は強度設計には欠かせない基礎科目です。しかし、大学・高専でこの . . .	
:	:	

## 推定



任意の文章をコピー

システムに入力

20~30秒程度で  
探索結果を表示

関係しそうな研究者を  
大学・分野横断的に  
探索することが可能



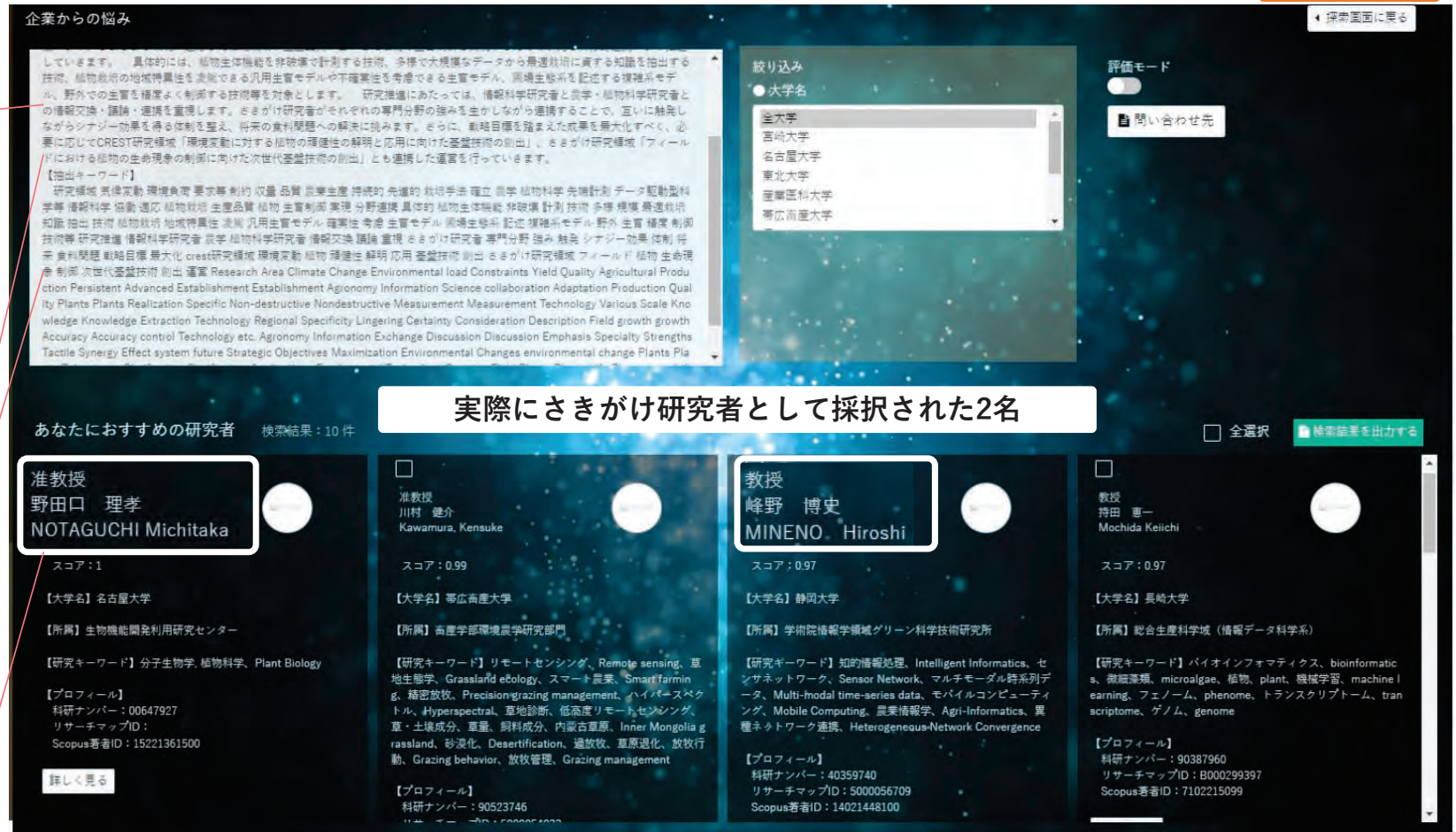
実際に探索する様子をご覧ください



入力した文章

自動で抽出された  
キーワード(AI目線)

最適な人(スコア=1)から順に  
10人を表示



実際にさきがけ研究者として採択された2名

准教授  
野田口 理孝  
NOTAGUCHI Michitaka

スコア: 1  
【大学名】名古屋大学  
【所属】生物機能開発利用研究センター  
【研究キーワード】分子生物学、植物科学、Plant Biology  
【プロフィール】  
科研ナンバー: 00647927  
リサーチマップID:  
Scopus著者ID: 15221361500

准教授  
川村 健介  
Kawamura, Kensuke

スコア: 0.99  
【大学名】帯広畜産大学  
【所属】畜産学部環境農学研部門  
【研究キーワード】リモートセンシング、Remote sensing、草地生態学、Grassland ecology、スマート農業、Smart farming、精密放牧、Precision grazing management、ハイパースペクトル、Hyperspectral、草地診断、低高度リモートセンシング、草・土壌成分、重量、飼料成分、内蒙古草原、Inner Mongolia grassland、砂漠化、Desertification、過放牧、草原退化、放牧行動、Grazing behavior、放牧管理、Grazing management

教授  
峰野 博史  
MINENO Hiroshi

スコア: 0.97  
【大学名】静岡大学  
【所属】学術院情報学領域グリーン科学技術研究所  
【研究キーワード】知的情報処理、Intelligent Informatics、センサネットワーク、Sensor Network、マルチモーダル時系列データ、Multi-modal time-series data、モバイルコンピューティング、Mobile Computing、農業情報学、Agri-Informatics、異種ネットワーク連携、Heterogeneous-Network Convergence  
【プロフィール】  
科研ナンバー: 40359740  
リサーチマップID: 5000056709  
Scopus著者ID: 14021448100

教授  
持田 恵一  
Mochida Keichi

スコア: 0.97  
【大学名】長崎大学  
【所属】総合生産科学域 (情報データ科学系)  
【研究キーワード】バイオインフォマティクス、bioinformatics、微細藻類、microalgae、植物、plant、機械学習、machine learning、フェノーム、phenome、トランスクリプトーム、transcriptome、ゲノム、genome  
【プロフィール】  
科研ナンバー: 90387960  
リサーチマップID: 8000299397  
Scopus著者ID: 7102215099

## このシステムの限界

- ・ 詳しく調べると探索された研究者が最適ではない
- ・ 探索された研究者が、類似の共同研究を他社と実施している
- ・ 探索された研究者の現時点の忙しさが分からない



**最終的にはコーディネータやURAのみなさんの協力が不可欠**



本システムはマッチング**支援**システムであることにご注意ください

### 相談したい研究者を探したい時に（企業・技術者）



#### 事例1

研究開発の打合せの際、困ったことを解決できそうな研究者がいないか、瞬時に当たりをつけ、その場で次の計画を立てやすいので重宝している。

#### 事例2

将来の研究計画を作成する際に、具体的に研究者を探索できるので便利。

### 支援企業からの相談で困った時に（財団・CD）



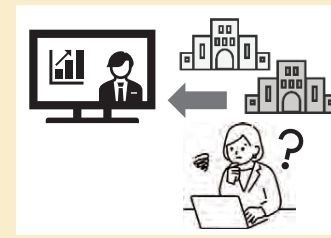
#### 事例3

解決が困難な相談がきたときにも、このシステムで解決の糸口を支援企業へ示すことができるので、ありがたい。

#### 事例4

地場の中小企業は近くの大学の研究者を探したいとのニーズが高く、IQUESTECHがあれば、相談したい大学も絞り込めるので重宝している。

### 専門家の意見を取り入れたい時に（新聞・テレビ局）



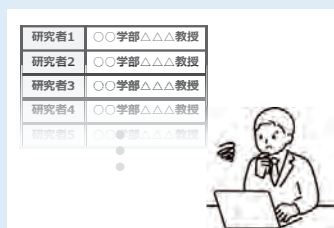
#### 事例5

番組制作する際専門家の意見を取り入れたいが、どんな専門家がいるのか、瞬時にあたりをつけ大学へアプローチすることができる。地元の大学で絞込みができるのも有難い。

#### 事例6

番組制作する上で一番手間な作業が、番組内容に精通した専門家を探すこと。IQUESTECHを活用することで、作業時間が大幅に削減できるので、大変助かっている。

### 難しい相談や曖昧な相談で困った時に（大学・CD）



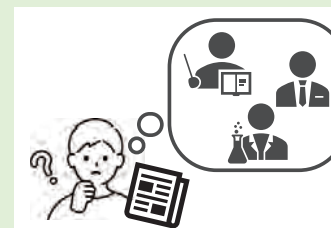
#### 事例7

大学へ相談に来られた企業様と話している途中で研究者を探索し、実際にどのような研究者がいるか、具体的に示すことで、相談者の表情が変わった。

#### 事例8

民間財団などの研究助成事業を研究者へ紹介する際、IQUESTECHで対象者を選別できるので助かる。

### 研究者を支援し、外部資金を獲得したい時に（大学・URA）



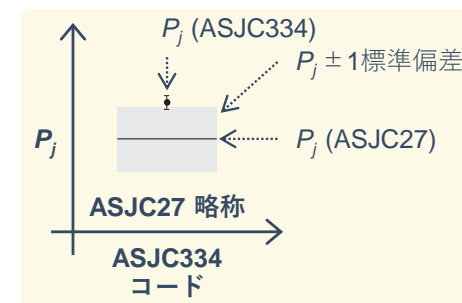
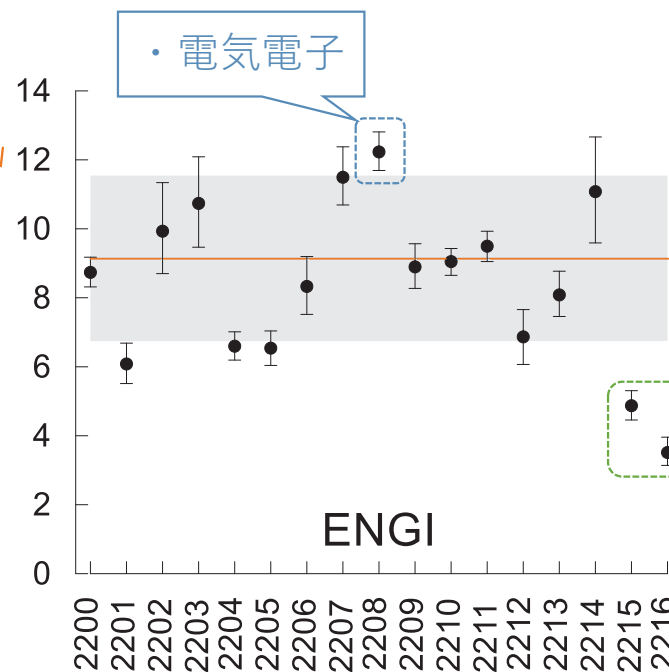
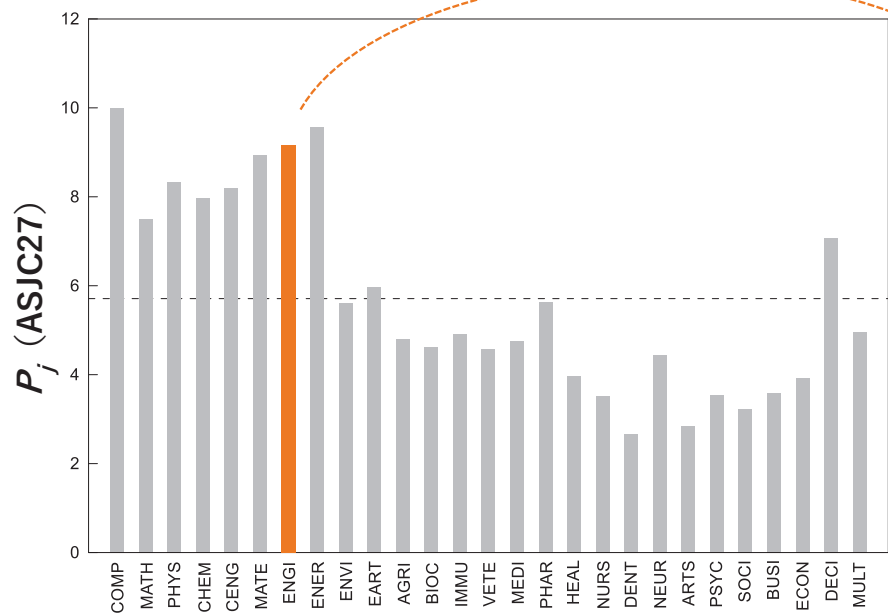
#### 事例9

公募型研究事業の評価委員として、研究者を推薦して欲しいと依頼された際、想定していた研究者も含め、瞬時に複数名を選出することができ大変助かった。

#### 事例10

これまで把握できていなかった研究者が出てくるので、支援の幅が広がる。

# 今後の取り組み：正規化に活用する研究分野分類の再考



大学	役職	論文数 (Scopus)	ASJC27		ASJC334	
			正規化論文数	学内のパーセンタイルランク	正規化論文数	学内のパーセンタイルランク
A大学	教授	21	0.90	59.3	1.49	77.4
B大学	助教	16	0.69	49.2	1.02	63.8
C大学	准教授	25	0.88	58.3	1.28	72.0
D大学	教授	23	1.04	64.7	1.60	79.6

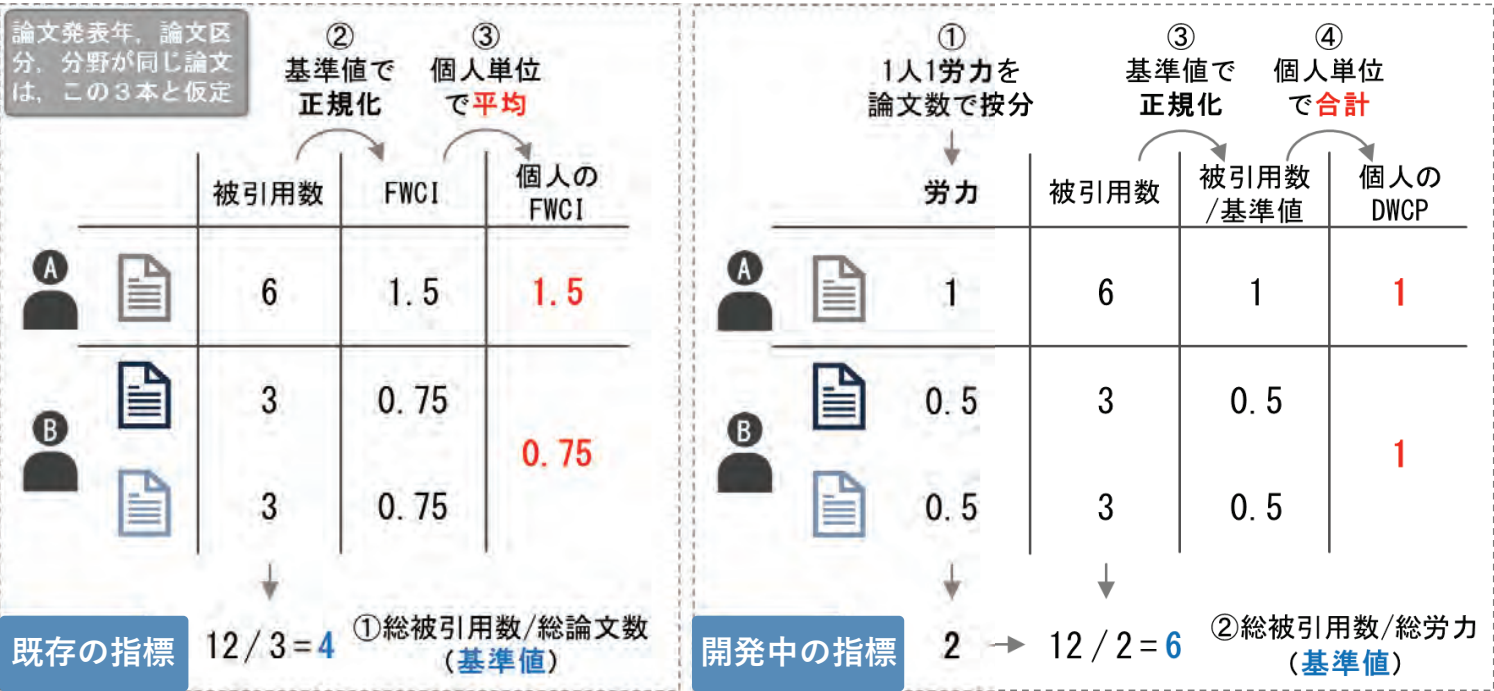
より細かな研究分野分類を選択することを検討

# 今後の取り組み：インプットを考慮した被引用数の正規化

## 正規化被引用数（FWCIやCNCI）の問題点

「論文」そのものの「注目度」を比較することを念頭にした指標。

「人」の「注目度」として活用するうえでは・・・ → **効率(inputとoutputのバランス)の考慮**



効率を考慮した被引用数の正規化指標の開発が必要

# 書誌情報に基づく分析の位置づけ

## 教育研究活動を評価できるデータ

	自機関	他機関	
書誌情報	◎	○	→ 自機関を超えた評価が可能
研究費	◎	△	→ 自機関内での評価が中心
指導学生数	△	×	→ 自機関内での評価も厳しい可能性
・			
・			

← 相対的な評価

論文数 ≈ 研究の活動度  
被引用数 ≈ 研究の注目度

【課題】異なる分野や年，組織規模での比較

↓  
正規化指標，パーセンタイルランクの活用

### 【限界】

- 単一指標では評価できない → 複数の指標の組み合わせ方
- ScopusやWoSでの未収録論文は対象外 → 国内雑誌に収録される論文の含め方
- 論文を発表しない分野は評価できない → 人文社会学系分野の成果物の扱い方

↓  
ピアレビューが重要

# まとめ



- ・ 定量評価だけに基づく判断は危険
- ・ ピアレビューが重要
- ・ 可能な限り正しく分析することが必要
  - 的確なニーズの把握
  - 正確なデータセットの作成
  - 再現可能な分析作業 (検証可能性: verifiability)
  - データリテラシーの涵養