

Title	著者の属性情報と個人識別番号に基づく研究者の論文 生産履歴の分析
Author(s)	川島, 浩誉; 調麻, 佐志
Citation	年次学術大会講演要旨集, 32: 527-529
Issue Date	2017-10-28
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/14840
Rights	本著作物は研究・イノベーション学会の許可のもとに 掲載するものです。This material is posted here with permission of the Japan Society for Research Policy and Innovation Management.
Description	一般講演要旨

2 D 0 3

著者の属性情報と個人識別番号に基づく研究者の論文生産履歴の分析

○川島浩誉（NISTEP），調麻佐志（東京工業大学）

Introduction

Web of Science や Scopus に代表される論文データベースを集計し、書誌学的分析によって科学活動を量的に把握する調査は、日本においてもその他の国においても、数多くなされてきた。これらの計量書誌学において、著者の属性情報はこれまで変数としての活用が十分になされてこなかった。特に、この論文のこの著者は教員なのか大学院生なのか等の立場に関しては、論文の書誌情報として収録されないことから一定以上の規模の質問票調査を要し、先行例が必ずしも多くはない。

本研究では、Scopus において東京工業大学に所属する著者として出現する研究者を対象とし、同大学から提供を受けた名簿に依拠して職位を照合し、Scopus の著者 ID と繋げることによって、研究者＝著者に着目した論文生産構造を明らかにすることを目標とするものである。

本発表では、研究の途中段階として、まず名簿データと論文書誌データの結合によって作成したデータセットの統計的特徴について示し、その先にある論文生産構造の議論に関してどこまで踏み込みうるかについての検討を行う。

Materials and Methods

本発表で用いるデータセットは以下の手順で作成した。

1. Scopus に収録された論文書誌情報の集合より、著者の所属機関に東工大を含むものを抽出する。
2. 1 で抽出した論文書誌情報の部分集合から東工大に所属する著者の著者 ID を抽出する。
3. 2 で抽出した東工大に所属する著者の著者 ID を用いて、当該研究者が著者に名を連ねている論文の集合を Scopus から再抽出する。
4. 2 で抽出した著者に関して、東工大内の名簿と照合し、結合する。

この手続きにより、研究者としての履歴において、東工大に在籍して論文を公刊したことがある研究者に関して、当該研究者が著者として名を連ねている論文の集合とそれぞれの書誌情報（東工大の立場であるか否かは問わない）とそのうち東工大在籍時にはどのような立場であったかが情報として含まれるデータセットを作成した。作成工程を図 1 に示す。

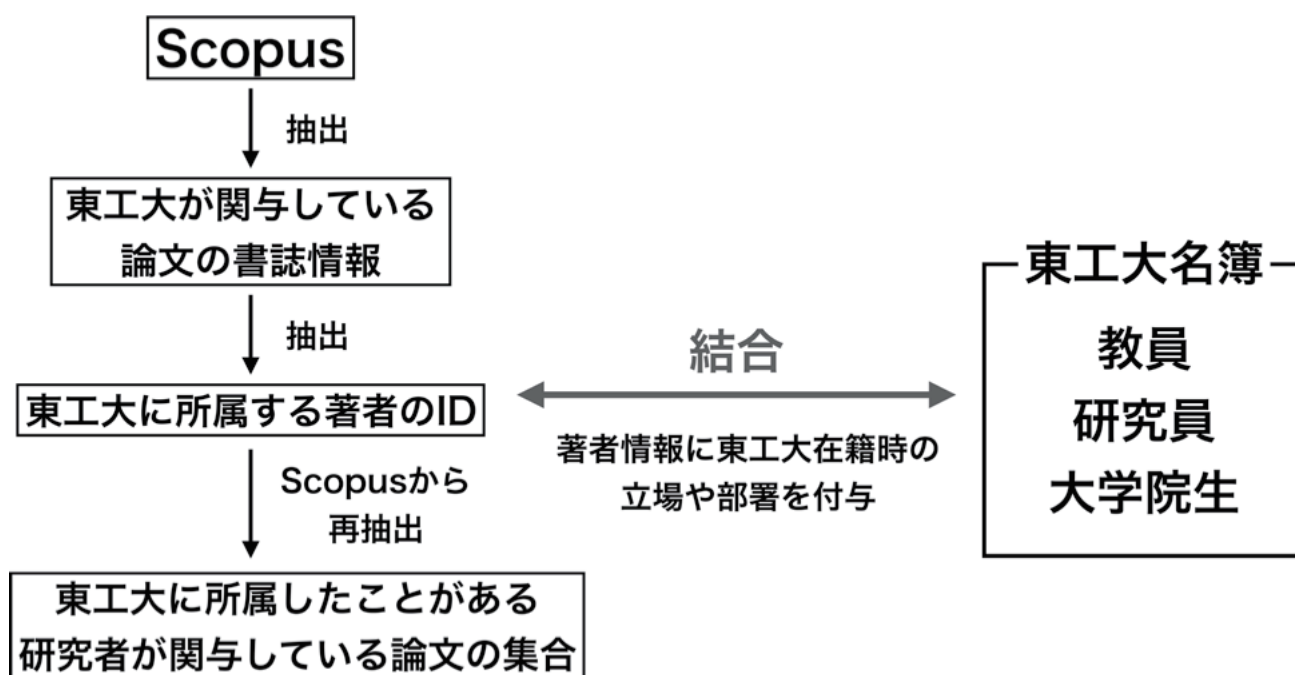


図1 使用するデータの作成

結合に用いたデータは、論文書誌データに関しては1996年から2014年の範囲でScopusのバルクデータから抽出を行い、名簿データに関しては、2002年から2014年の各種の在籍者名簿を用いた。相互の照会と結合の作業は人間が行い、目視およびwebや東工大が公開している研究者データベースを判断の材料とした。

Results

まず、結合作業の結果について述べる。Scopusから抽出された、東工大に所属する著者として論文を書いたことがある研究者は26276IDあり（これは真に正しい値ではなく、あくまでScopus内の著者IDによる数である）、そのうち、当該IDが著者に名を連ねている論文が多いものを優先した結果、3977IDに関して東工大在籍期間内の何らかの立場が判明した。表1は各研究者（著者ID）に関して何年分の立場が判明したかの内訳の分布である。ここで、判明しなかった年に関しては、当該IDがその年には、

- a. 東工大に在籍していない
- b. 在籍し名簿にも名前があるが結合ができていない
- c. 在籍しているが名簿に載っていない（学部生や名簿欠損に依る）

などの可能性がある。bに関しては結合作業上の問題である。

なお、データの結合に関しては、人間による作業であることから作業時間がボトルネックとなっており、結合が成立した人数を増やすためにはシステムティックな識別も補助的に用いる必要がある。本研究では現在、上記の人間による結合の結果に加えて、機械学習的な過程によっても結合作業を試みている。

また、表2に結合に成功した研究者（著者ID）の東工大における立場の内訳を示す。このとき、ひとりの研究者（著者ID）が複数年で同一の立場の場合は+1（ID）とカウントし、ひとりの研究者が准教授→教授と昇進していることが同定できている場合には、准教授に+1（ID）、教授に+1（ID）としている。

表 1 名簿と書誌情報の結合に成功した研究者の立場の判明期間の内訳

判明した期間の 年数（合計）	人数（ID 数）
13 (期間中全て)	666
12	61
11	68
10	76
9	75
8	113
7	146
6	194
5	680
4	320
3	559
2	782
1	237
合計	3977

表 2 名簿と書誌情報の結合に成功した研究者の立場の内訳

東工大における立場	人数 (ID 数)
教授	589
准教授（助教授）	522
専任講師	53
助教（助手）	688
特任教授	61
特任准教授（助教授）	54
特任講師	12
特任助教（助手）	114
研究員	316
技術専門職員（技官）	27
客員教員	124
特命・栄誉・名誉教授	183
非常勤講師	14
博士	1452
修士	1379
その他	112