

NISTEP 定点調査における自由記述の感情分析に関する基礎的考察

01900990	東京海洋大学	*渡部 大輔	WATANABE Daisuke
	大阪大学	標葉 隆馬	SHINEHA Ryoma
	防災科学技術研究所	松本 拓郎	MATSUMOTO Takuro
	デジタル庁	中川 尚志	NAKAGAWA Takashi
01002750	政策研究大学院大学	大山 達雄	OYAMA Tatsuo

1. はじめに

研究業績の評価に基づく資源配分を行う上で、研究者の意識を把握した上で、制度設計を行うことが重要である。これまでに筆者らは、文部科学省科学技術・学術政策研究所 (NISTEP) による「科学技術の状況に係る総合的意識調査」(NISTEP 定点調査) の自由記述を対象として、共起ネットワークや対応分析などテキストマイニングを通じて、研究者の資源配分に関する意見の特徴を計量的に分析した[1]。その結果を踏まえて本研究では、同調査を用いて研究者の意識について感情分析を用いて分析することで、賛否両論の意見に対する特徴を職位別の違いを中心として考察する。

2. 分析の概要

対象データとして、NISTEP 定点調査の第3期(2016-20年)におけるアンケート調査の結果並びに自由記述のデータベース(以下、DB2020)を用いる[2]。質問事項に関する概要と自由記述の質問番号は、表1の通りである。職位としては、「社長・役員、学長等クラス」(以下、学長等)、「部課長、教授クラス」(教授等)、「主任研究員、准教授クラス」(准教授等)、「研究員、助教クラス」(助教等)に分類されており、「その他」はサンプルが少ないことから対象外とする。

感情分析のツールとしては、「エモレポ」(第1.3.3版, 2021年12月23日発行)を用いる[3]。同ソフトは、自然言語理解技術に基づく約1,000億文の大規模学習データを駆使し、書き手の感情(81種類)やテキストの話題・分野(980種類)に分類・抽出が可能であり、商品やサービスの改善に向けた「苦情情報」や強みとして打ち出すための「好評情報」などを直感的に把握することが可能である。

3. 指数を用いた傾向分析

アンケート調査では、各質問において「不十分」から「十分」の6段階尺度の選択形式で回答されており、10点満点の指数に変換された数値を用いる。つまり指数が高いほど、満足度が高いこととなる。2020年における各パートの指数の平均値は図1の通り、

人材と改革・機能において学長等とそれ以外で傾向が異なる。調査時点間の比較として、2016年と2020年の指数の平均値の差は図2の通り、いずれのパートで全ての職位で悪化しており、助教等とそれ以外で傾向が異なる。助教等は指数が高い一方、時点間での減少が大きいことが特徴である。

表1 質問事項の概要

パート	項目	略称	質問数	自由記述
1	大学・公的研究機関における研究人材	人材	15	Q115
2	研究環境及び研究資金	環境・資金	11	Q211
3	学術研究・基礎研究と研究費マネジメント	研究・マネ	9	Q309
4	産学官連携とイノベーション政策	連携・政策	19	Q419
5	大学改革と機能強化	改革・機能	6	Q506
6	社会との関係深化と推進機能の強化	社会関係	9	Q609

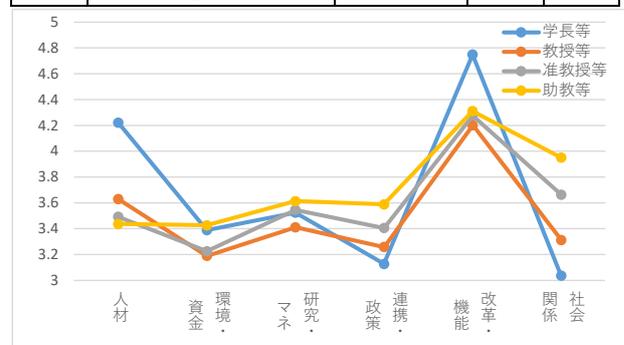


図1 各パートの指数の平均値 (2020年)

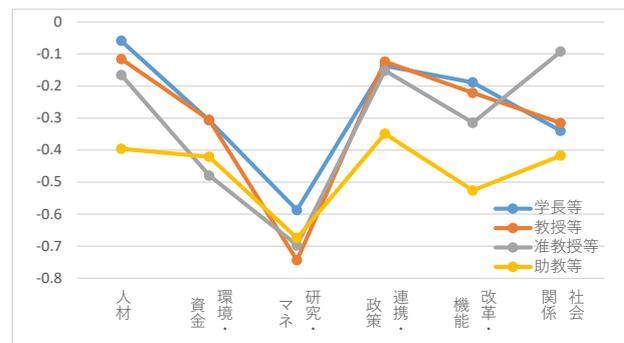


図2 各パートの指数の平均値の差 (2016→2020年)

4. 自由記述を用いた感情分析

4.1 文字数に関する基礎統計

各パートにおける自由記述の文字数は、図3の通りであり、全体の文字数は約162万語となる。人材と環境・資金における記述が多いことが分かる。職位別の割合若干変化はあるものの、学長等20:教授等40:准教授等25, 助教等15である。

4.2 感性に関する分析

感性情報出力の細かさについて、文章全体の主旨となりえる感性をひとつだけ出力する「結論」を用いる。感性区分としては、中分類の「苦情」「要望」「好評」「質問」「その他」を用いる。「苦情」に分類された意見の割合は図4の通りであり、人材が最も多く、いずれの職位も70%以上となった。それ以外のパートは職位による違いが見られるものの、学長等・教授等と准教授等・助教等はそれぞれの傾向が似ている。

4.3 指数と自由記述の比較分析

前節で得られた満足度(図1・2)と不満足度(図3)の結果を比較する。改革・機能に着目すると、学長等は指数が最も高く(図1)、減少幅が最も少ない(図2)ことから、肯定的であることが分かる。更に、図4では苦情の割合が最も低いことから、指数と自由記述で同様に傾向が見られる。

一方、人材に着目すると、図1と図2ともに他のパートと比べても悪い値ではないものの、図4では全ての職位において苦情の割合が最も高いことから、両者に違いが見られる。このことから、自由記述による感情分析は、指数だけでは捉えられない結果が得られることが分かる。

4.4 資源配分に関する分析

研究費に関する分析として、「資金」「研究資金」「研究費」「予算」の4語を抽出し、感性毎に出現率を図示すると図5のようになる。助教等ではいずれも15%程度となり、その他の職位と比べて出現率が高い。感性間の比較として、准教授等は「好評」が最も高い一方、他の職位は「要望」が最も高い。このことから、研究費に関する用語は苦情と要望とともに好評の意見を述べるのにも用いられており、その出現率に大きな違いが見られないことが分かる。

5. おわりに

本研究では、DB2020に対する感情分析を通じて、調査項目や職位に着目した研究者の意識の違いを考察し、賛否両論の意見について自由記述における用語の出現率から考察した。今後、所属機関等の属性や質問項目との対比を含めた分析とともに、各設問の意見変更の理由を用いた分析を行う必要がある。

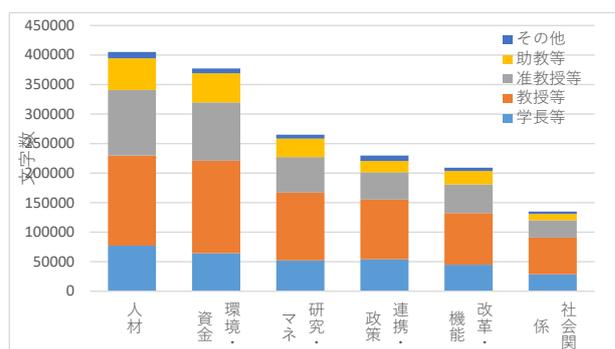


図3 自由記述の文字数

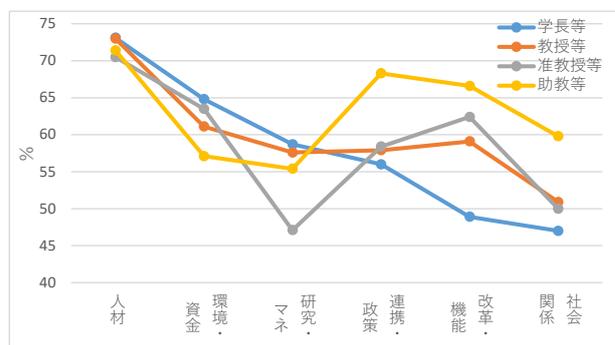


図4 自由記述における「苦情」の割合

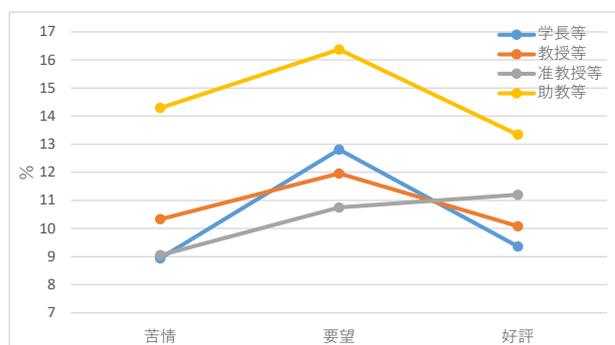


図5 自由記述における研究費に関する用語の出現率

謝辞

本研究はSciREX 共進化実現プログラム(第IIフェーズ)「研究業績の評価に基づく資源配分効果の動的な特性分析のための基礎研究」の一環として実施された。

参考文献

- [1]渡部大輔, 楠木亮介, 松本拓郎, 大山達雄: NISTEP 定点調査における自由記述のテキストマイニング手法を用いた基礎的分析, 日本オペレーションズ・リサーチ学会 2021 年秋季研究発表会, 2-E-5, 2021.
- [2]文部科学省科学技術・学術政策研究所「科学技術の状況に係る総合的意識調査(NISTEP 定点調査 2020) データ集」, NISTEP REPORT, No.190, 2021.
- [3]株式会社言語理解研究所: 感情分析レポート「エモレポ」, <https://www2.abx2.net/>, 2021/12/29 アクセス