

テキストマイニングによるジオパーク研究動向の分析

－『日本のジオパークに関連する文献』2005～2014年を中心に－

Analysis on the trends of geo-park studies using text-mining :
Focused on literature list of geo-park Japan 2005-2014

肖 鋁¹

¹北海道大学大学院文学研究科

Kun Xiao¹

¹Graduate School of Letters, Hokkaido University

Kita 10, Nishi 7, Kita-ku, Sapporo, Hokkaido, Japan 060-0810

キーワード：テキストマイニング，日本，ジオパーク，研究分野，視点

Keywords : Text-mining, Japan, Geo-park, Research field, Perspective

抄録

本稿の目的は、今日に至るまでの日本におけるジオパーク研究の歩みを概観するとともに、その研究上の特徴を明らかにすることにある。具体的には、日本地理学会ジオパーク対応委員会のホームページに掲載された『日本のジオパークに関連する文献』を研究の対象として取り上げ、2005年から2014年に至るまでの期間にそこに掲載されたジオパークに関する研究論文（合計103編）の標題を分析し、まずは研究分野の分類から検討した。またテキストマイニングの手法を用いながら、日本におけるジオパーク研究に見出される動向の特徴を検討することにある。

1. はじめに

本稿の目的は、日本地理学会ジオパーク対応委員会のホームページに掲載された『日本のジオパークに関連する文献』¹（以下、『文献』と略す）の研究論文等について分析し、日本におけるこの10年間のジオパーク研究動向について考察することである。

分析対象としたのは、2005年から2014年に『文献』に掲載されたジオパークに関わる研究論文等の103編である。

最初に、研究論文等を研究分野別の分類基準に基づき分類し、この10年間の研究分野を検討し、次に、2005年から2014年の10年間を5年ごとに区切った上で、それぞれを同様に研究分野別の分類基準により分類し、この10年間における研究分野の変化を検討する。さらに、今回は、研究論文等の標題をテキストマイニングで分析し、この10年間の研究動向を分析することにも挑戦する。具体的な目標としては、テキストマイニングを活用することで、研究分野別の分類基準に基づいた分類結果から検討される傾向以外の傾向を見出すこ

とができるのか否かについて検討する。

構成は次のとおりである。Ⅱの「研究分野の検討」において、研究論文等を読み、その内容に従って、研究分野の分類基準で分類した結果と検討した傾向を示す。その上で、Ⅲの「テキストマイニング分析」において、テキストマイニングを活用して分析した結果と、Ⅱの研究分野の検討と比較するとともに、新たな動向を見出すことができるのか否かについて議論する。

2. 研究分野の検討

2.1. 分類基準

最初に、研究論文、掲載雑誌等を読み、論文内容に従って、研究分野から分類した。研究分野別分類基準には次のような項目が含まれる。

- | | |
|---------|-----------|
| I 地質学 | VI 地域開発学 |
| II 地理学 | VII 博物館学 |
| III 教育学 | VIII 地球科学 |
| IV 観光学 | IX ほか |
| V 経営学 | |

2.2. 10年間の研究分野（2005～2014年）

研究分野別分類基準に基づいた分類結果は、上位5位に注目してみると、1位が地理学（32本）、2位が地域開発学（22本）、3位が教育学（16本）、4位が地質学（11本）、5位が観光学（10本）であった。なお、（ ）の中の本数は掲載論文数を示す。

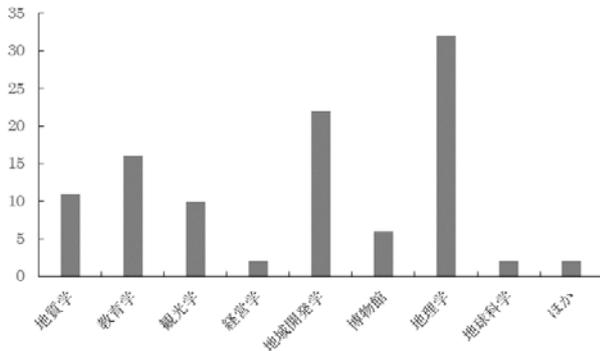


図1 研究分野のヒストグラム（2005～2014年）

このように、研究分野別に基づいた分類結果からは、この10年間、研究分野では、地理学、地域開発学、観光学、教育学、地質学がジオパークに注目されてきたことを確認できた。

2.3. 研究分野の変化（2005～2009年、2010～2014年）

次に、2005年から2014年の10年間を、2005～2009年、2010～2014年というように5年ごとに区切った上で、研究分野別で分類し、期間ごとの結果を比較することで、前述した10年間の傾向にはどのような変化があったのかについて検討する。

研究分野について、地質学は15%→7%と減少している。教育学は23%→13%と減少している。観光学は27%→5%と減少している。地域開発学は4%→29%と増加している。地理学は15%→30%と増加している。なお、それぞれの割合は、2005～2009年、2010～2014年のそれぞれの期間におけるそれぞれの研究分野から扱う掲載論文数の割合を示す。観光学の割合が大幅減少であるのに対し、逆に地域開発学の割合は増加になっている。

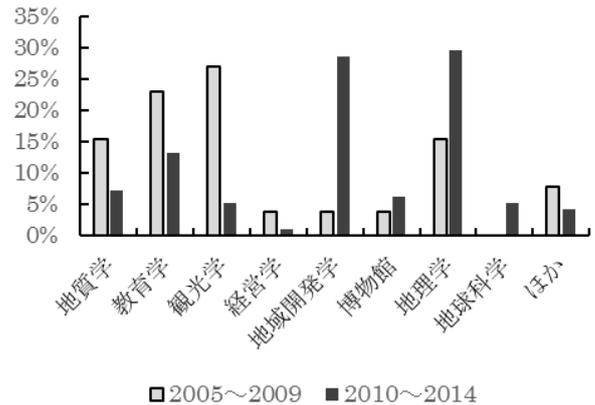


図2 研究分野の変化（2005～2014年）

3. テキストマイニング分析

3.1. テキストマイニングについて

テキストマイニングは、コンピュータを活用して、テキストの中から有用な情報を探し出す技術である。これは、文書が大量であればあるほど、有効な手段となる。文書の解析にコンピュータを活用し、一定のルールに従って重要だと判断される情報を探し出せるならば、大量の文書であっても、客観的に重要な情報、つまり有用な情報を得られる可能性がある。

3.2. テキストマイニングの手続き

今回は、研究の最初の段階として、分析の方向性を定めることを目的として、研究論文等の標題に注目し、テキストマイニングで分析した。手続きは次のとおりである。最初に、2005～2014年の研究論文等の標題のテキストファイルと、2005～2009年、2010～2014年、という5年ごとの研究論文等の標題のテキストファイルを作成した。次に、それぞれのファイルを、形態素解析辞書「MeCab」を使用して解析した。具体的には、標題を単語に分解した後、ファイルごとに、出現した単語の頻度のヒストグラムを作成した。

3.3. 分析結果

以下、本項では、まず検討の対象とした期間全体（2005～2014年）を通じての日本のジオパーク研究に見られる特徴を明らかにしたい。そして、その後、検討対象とする期間を5年という年数単位で機械的に2つの時代（2005～2009年、2010～2014年）に区分し、より細分化されたそれぞれ

の期間にみだされる特徴を概観することにした。

(1) 総合 (2005～2014年)

「図3」からは、「活用」というキーワードが、検討対象として期間全体を通じて日本のジオパーク研究における最も重要なキーワードであったことが確認できる。

また、「ジオパーク活動」や「事例」が、「地域振興」や「大地」よりも上位に来ることも、日本のジオパーク研究は現場の実態を重視される特徴を示している。「糸魚川ジオパーク」や「山陰ジオパーク」も上位に位置したが、それはジオパークの名声と歴史に関わるものからである。「糸魚川ジオパーク」と「山陰ジオパーク」は日本において最も歴史が古い世界ジオパーク、研究対象としては当然なことである。

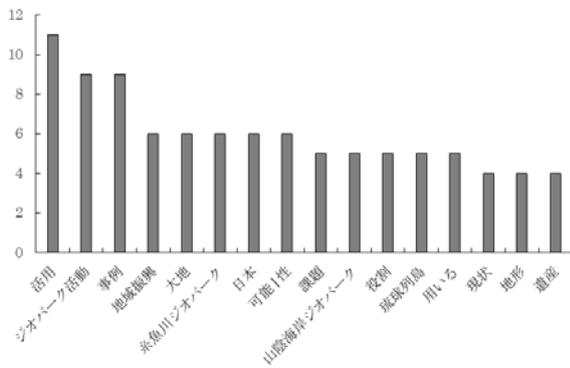


図3 キーワードのヒストグラム (2005～2014年)

(2) 2005～2009年

「2005～2009年」については、「図4」から、「可能性」や「地質遺産」、あるいは「糸魚川イオパーク」や活用、考察などといったキーワードが上位にあり、このことから、当該期間におけるジオパークに関する研究の特徴は地質学的研究に見出されるものと考えられる。

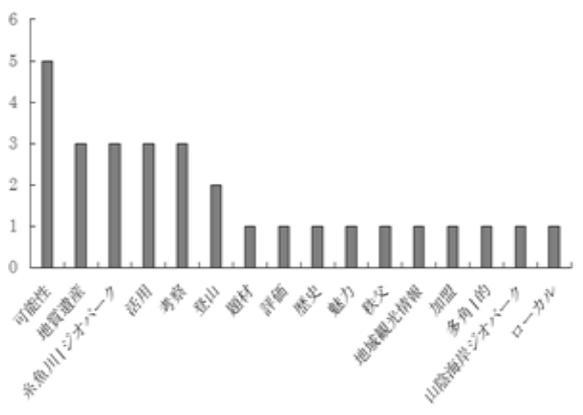


図4 キーワードのヒストグラム (2005～2009年)

(3) 2010～2014年

次の「2010～2014年」の特徴として、「図5」からは、ジオパークの研究の視点は地質学分野よりも地域開発学分野へと移行したことがうかがえる。キーワードにおいて、「事例」や「活用」、「ジオパーク活動」や「地域振興」といったキーワードが上位になった。

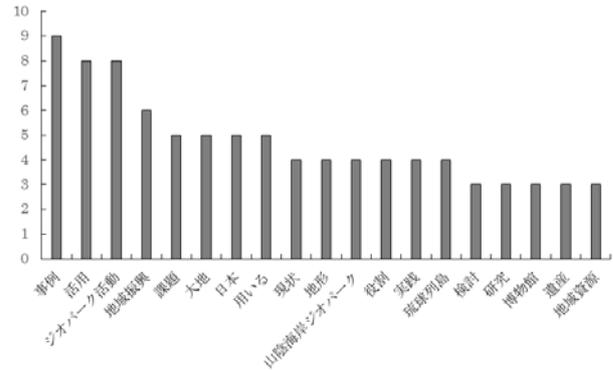


図5 キーワードのヒストグラム (2010～2014年)

4. おわりに

本稿では、『文献』の研究論文等を分析し、2005年から2014年までの日本のジオパーク研究の特徴を検討してきた。研究論文等を読んで、研究分野に基づいて分類した結果から特徴を検討するとともに、標題をテキストマイニングで分析した結果からも動向を検討してきた。その結果からは、次の二点が明らかになった。

1. 日本のジオパーク研究は、この10年間、研究分野として、地理学、地域開発学、観光学、教育学、地質学がジオパークに注目されてきたことを確認できた。2005年から2009年までは、教育学や地質学分野の研究は多かった。2010年から2014までは多くの地理学や地域開発学、観光学分野の学者がジオパークを注目のした。

2. ジオパーク研究に関するキーワードについて、「活用」、「ジオパーク活動」、「事例」、「地域振興」、「大地」がよく出てきた。日本ではジオパーク研究は最初の地質学と教育学の視点から地理学、地域開発学の視点へ移行したと考えられた。

本稿での検討の対象になったデータは非常に限られたものであり、したがって、そこから導き出された結論も試論ないしサンプル的小括に過ぎないが、今後の研究に対する一つの方向性と課題を示すものといえよう。

注

- 1 日本地理学会ジオパーク対応委員会ウェブサイト
<https://sites.google.com/site/ajggeopark/home/references>
2016年6月12日閲覧.

謝辞

本研究では,平成27年度苗場山麓ジオパーク学術研究助成金の一部を利用しました.

参考文献

- [1]石田基広 [2008] 『Rによるテキストマイニング入門』森北出版株式会社.
[2]金 明哲 [2009] 『テキストデータの統計科学入門』岩波書店.
[3]松村真宏・三浦麻子 [2012] 『人文・社会科学のためのテキストマイニング』誠信書房株式会社.

Abstract

The purpose of this paper is to investigate the trends in the studies of geo-park in Japan through analyzing research papers related to geo-park in literature list of geo-park Japan, 2005~2014(103papers). This study first identified the overall trends among the research papers from the results obtained through summarizing and analyzing them. The first, From research field to analyze the papers. Then analysis on the trends of geo-park studies using text-mining.

(受付日 : 2016年8月13日, 受理日 : 2016年8月26日)



肖 鋨 (しょう こん)

現職 : 中国・襄陽職業技術学院大学教員

北海道大学大学院文学研究科博士後期課程在学.
専門は人文地理学, ジオパーク関連の研究を行っている.