

日本の水資源行政の国際化の変遷(2006年-2017年) —単語ベクトルと単語共起の分析から—

Changes in Globalization on Japanese Water
Resources Administration (2006-2017)
- Based on Word Vectors and Word Co-occurrence Analysis -

杉 浦 政 裕^{*}
Masahiro SUGIURA

要 旨

本論文は、日本の水資源行政の国際化の変遷(2006年-2017年)の中に、国際協力から輸出ビジネスへの繋がりが見られることを、テキストマイニングにより確認したものである。テキストマイニングによる確認には、特定単語「協力」と共起関係にある語群の時系列変化、「ビジネス」と共起関係にある語群の時系列変化、および特定単語「協力」の類義語群の時系列分析の結果を組み合わせることを取り入れている。そして、単語のベクトル化により抽出した類義語群の時系列変化と単語の共起ネットワーク分析による共起関係の時系列分析の結果を組み合わせることで解釈することにより、気づきの確からしさを高められることも示唆している。

キーワード：水資源行政、テキストマイニング、Word2vec、共起分析

^{*} すぎうらまさひろ、九州国際大学現代ビジネス学部、sugiura@cb.kiu.ac.jp

1 序論

日本の水資源行政の国際化は、「海外社会資本事業への我が国事業者の参入の促進に関する法律（海外インフラ展開法）」（2018年）の施行に見られるように、相手国への輸出ビジネスへという流れが作られている。海外インフラ展開法の施行前に、従来の国際分野における国際協力という取り組みの流れの中で、どのように輸出インフラを扱ってきたのかを知ることが、とても興味深い。そこで、本論文では、日本の水資源行政を担う政府が発行する水資源白書や水循環白書の国際関連の章を分析対象テキストとし、テキストマイニングによる時系列分析により、日本の水資源行政における国際化の流れが、相手国への協力から輸出ビジネスへ繋がってきたのかを探っていくことにした。

これまでテキストマイニングによる時系列分析は、杉浦（2019）も言うように、例えば福原（2006）、杉浦（2010）により、Weblogや新聞記事を分析対象テキストとして、各分析対象テキストに現れる単語の関係を時系列分析する研究が行われてきた。また、対象分野を絞り込んだテキストからWord2vecにより単語のベクトル表現を行い、特定単語の類義語を複数抽出し、複数の類義語を時系列分析するという研究は、杉浦（2019）により行われている。ここで言うWord2vecとは、A. Mnih and K. Kavukcuoglu(2013) や T. Mikolov et al(2013) が提案しているとおり、もっとも簡単なニューラル言語モデルの1つである。

水資源白書の国際関連の章を分析対象テキストとした場合、特定単語「協力」の類義語群の時系列変化は、国際協力あるいは建設業界支援に関することを表していることは、すでに杉浦（2019）により確認されている。しかし、杉浦（2019）は、協力から輸出ビジネスへのつながりについて、分析していない。また、特定単語「協力」と共起関係にある単語の時系列変化の面からも分析をしていない。

本論文の目的は、杉浦（2019）の報告では、未着手である部分（日本の水資源行政の国際化は相手国への協力から輸出ビジネスへと繋がっている、という

ことへの気づきが得られること)を、特定単語「協力」と共起関係にある語群の時系列変化、特定単語「ビジネス」と共起関係にある語群の時系列変化、および特定単語「協力」の類義語群の時系列分析の結果を組み合わせることで、確認することである。

2 分析方法

特定単語「協力」及び特定単語「ビジネス」と共起関係にある単語の時系列変化、特定単語「協力」の類義語群の時系列変化を分析することから、日本の水資源行政の国際化は相手国への協力から輸出ビジネスへと繋がっている、ということへの気づきが得られることを、以下に示す分析方法で確認する。

分析の手順は以下のとおりである。

- 2-1 特定単語「協力」と共起関係にある単語の時系列変化を分析することで、「協力」と結びつきの近い単語の時系列変化を把握する。そして、「協力」を進めるためのコンテンツを読み取る。
- 2-2 特定単語「ビジネス」と共起関係にある単語の時系列変化を分析することで、「ビジネス」と結びつきの近い単語の時系列変化を把握する。そして、「ビジネス」を進めるためのコンテンツを読み取る。
- 2-3 特定単語「協力」の類義語群の時系列変化を分析することで、「協力」と意味の近い単語を把握する。そして、「協力」を進めるためのコンテンツの読み取り、2-1の分析結果を補完する。なお、2-3については、杉浦(2019)の分析結果と整合性を図るため、杉浦(2019)の分析方法と分析結果を採用する。

【2-1及び2-2の分析環境】

- (1) OS : Windows10 Pro

- (2) 形態素分割（分かち書き）：WordMiner1.1（日本電子計算機株式会社）
- (3) Jaccard係数を用いた共起ネットワーク分析：KH Coder 3
- (4) 分析対象テキスト：表1に示すとおり、杉浦（2019）と分析対象となるテキストと整合を図るため、『水資源白書』（2006年－2014年）及び『水循環白書』（2015年－2017年）の各全文のうち国際関連の記述がある章のみを抽出したテキストを分析対象テキストとして選定する。また、選定理由は、杉浦（2019）も言及するように、分析対象テキストを国際関連について記述されている章に絞り込むことにより、分析精度の向上を期待できるためである。

【2－1及び2－2の分析手順】

- (1) 分析対象テキスト（四期十二年間分）について形態素分割（分かち書き）をする。
- (2) 分析対象テキスト（四期十二年間分）について、期毎にJaccard係数を用いた共起ネットワーク分析を行う。
- (3) Jaccard係数を用いた共起ネットワーク分析を行う際に用いるパラメータは、単語の最小出現頻度＝10回、Jaccard係数＝0.16。なお、分析の対象となる品詞は名詞のみとする（ただし、地名、人名、組織名など固有名詞は分析対象の名詞に含まない）。

【2－3の分析環境】（杉浦（2019））

- (1) OS：Windows10 Pro
- (2) 使用言語：Python3.6.6
- (3) 使用ライブラリ：gensim-Word2vec
- (4) Word2vecのアルゴリズム：skip-gram
- (5) 形態素分割（分かち書き）：WordMiner1.1（日本電子計算機株式会社）
- (6) 度数分布図作成：Excel2019（マイクロソフト社）、Excel2007（マイク

ロソフト社) & エクセル統計2010 (株式会社社会情報サービス)

- (7) 分析対象テキスト：第1期として水資源白書(2006年－2008年)の国際関連の章、第2期として水資源白書(2009年－2011年)の国際関連の章、第3期として水資源白書(2012年－2014年)の国際関連の章、そして第4期として水循環白書(2015年－2017年)の国際関連の章(表1参照)

【2－3の分析手順】(杉浦(2019))

- (1) 分析対象テキスト(四期十二年間分)について形態素分割(分かち書き)をする。
- (2) Word2vecのパラメータであるsizeとwindowについては、size = 100、window = 16に設定する。
- (3) Word2vecを使って、類義語及び類義語の類似度を計算し、各期上位30語計120語を抽出する。
- (4) 抽出した120語から数字など特徴のない単語や重複する単語を除去し、45語を選定する。横欄に45語、縦欄に各期の計算された類義語の類似度を記入し表を作成する。そして、作成した表のデータをコレスポンデンス分析を行う。

表1 分析対象テキスト

期	年	分析対象テキスト	
第1期	2006	平成18年度日本の水資源(水資源白書)、国土交通省編	第11章 水資源に関する国際的な取組み
	2007	平成19年度日本の水資源(水資源白書)、国土交通省編	第12章 水資源に関する国際的な取組み
	2008	平成20年度日本の水資源(水資源白書)、国土交通省編	第12章 水資源に関する国際的な取組み
第2期	2009	平成21年度日本の水資源(水資源白書)、国土交通省編	第12章 水資源に関する国際的な取組み
	2010	平成22年度日本の水資源(水資源白書)、国土交通省編	第10章 水資源に関する国際的な取組み
	2011	平成23年度日本の水資源(水資源白書)、国土交通省編	第10章 水資源に関する国際的な取組み
第3期	2012	平成24年度日本の水資源(水資源白書)、国土交通省編	第8章 水資源に関する国際的な取組み
	2013	平成25年度日本の水資源(水資源白書)、国土交通省編	第8章 水資源に関する国際的な取組み
	2014	平成26年度日本の水資源(水資源白書)、国土交通省編	第8章 水資源に関する国際的な取組み
第4期	2015	平成27年度水資源施策(水循環白書)、内閣官房編	第8章 国際的な連携の確保及び国際協力の推進
	2016	平成28年度水資源施策(水循環白書)、内閣官房編	第8章 国際的な連携の確保及び国際協力の推進
	2017	平成29年度水資源施策(水循環白書)、内閣官房編	第8章 国際的な連携の確保及び国際協力の推進

(出所) 杉浦 (2019)

3 分析結果及び考察

特定単語「協力」および特定単語「ビジネス」と共起関係にある単語の時系列変化を分析した結果を、図1から図4に示す。図1は第1期(2006年－2008年)の共起ネットワーク分析結果、図2は第2期(2009年－2011年)の共起ネットワーク分析結果、図3は第3期(2012年－2014年)の共起ネットワーク分析結果、そして図4は第4期(2015年－2017年)の共起ネットワーク分析結果である。

特定単語「協力」の類義語群の時系列変化を分析については、杉浦(2019)の分析結果を一部表現修正して提示する。表2に、各期の特定単語「協力」の類義語上位30語及び類義語の類似度一覧を示す。表3に、特定単語「協力」の類義語45語の類似度の経年変化を示す。そして図5に、特定単語「協力」の類義語45語と各期の関係(コレスポンデンス分析結果)を示す。

図1から図4に示す特定単語「協力」および特定単語「ビジネス」と共起関係にある単語の時系列変化を分析した結果を順に見ていこう。

図1に示された第1期(2006年－2008年)の分析結果には、「協力」と共起関係が強い単語はとして、「開催」「資源」「パートナーシップ」が現れている。

図2に示された第2期(2009年－2011年)の分析結果には、「協力」と共起関係が強い単語は、「評価」「設立」「採択」が現れている。

図3に示された第3期「2012年－2014年」の分析結果には、「協力」と共起関係が強い単語として、「環境」「技術」が現れている。また、「ビジネス」と共起関係が強い単語として、「フォーラム」「企業」が現れ始めた。

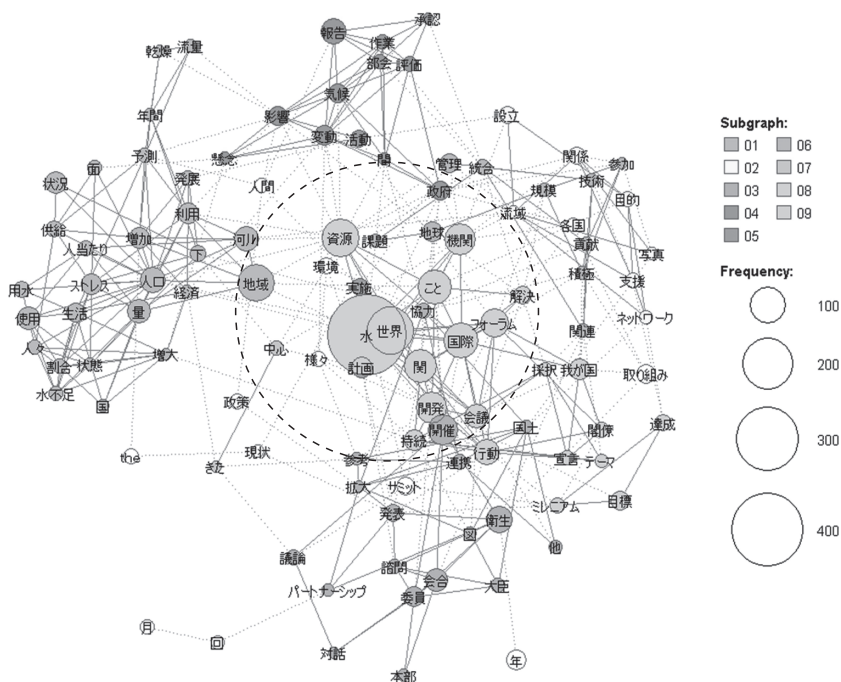
図4に示された第4期(2015年－2017年)の分析結果には、「協力」と共起関係が強い単語として、「災害」「活用」が現れている。そして「輸出ビジネス」と共起関係が強い単語は、「セミナー」「案件」「展開」と変化してきた。

次に、表2、表3及び図5を見ておこう。表2、表3及び図5は、杉浦(2019)が言うように、表2では、上位30語が、每期入れ替わっていることから、「協力」の使われる範囲が変化していることがわかる。

表3では、経年データから途中出現(以後消失なし)及び途中消失(以後出現なし)という特徴をもつ類疑義は、以下のとおりである。第2期から登場する類義語：維持、援助。第3期から登場する類義語：洪水、事業、発信。第4期から登場する類義語：参入、調査、理解。そして、第2期(2009年－2011年)及び第3期から消失する類義語：交流、省庁、様々、第4期から消失する類義語：プログラム、解明、懸念、貢献、成果、努力となっている。

全体を把握するために、表3の45語の類似度の経年変化をコレスポンデンス分析し、可視化した図5を見ておこう。第1期(2006年－2008年)周辺には「政策」「問題」「促進」といった語群、第2期「2009年－2011年」周辺には「声名」「呼びかけ」「交流」といった語群、第3期「2012年－2014年」周辺には「発信」「事業」「洪水」といった語群、そして第4期「2015年－2017年」の周辺には「参入」「質」「水災害」といった語群が出現している。

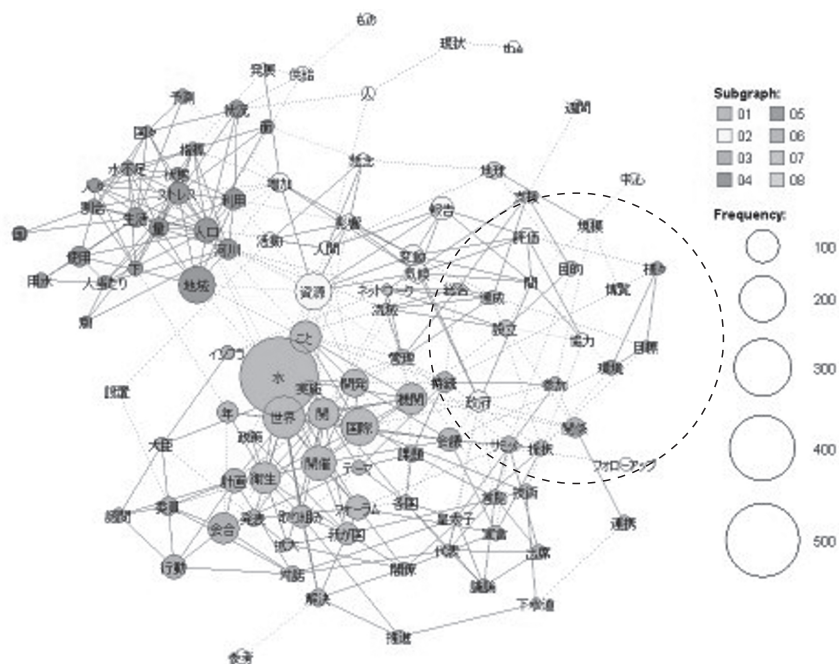
2006-2008



（出所）筆者作成

図1 第1期（2006年－2008年）特定単語「協力」および特定単語「ビジネス」と共起関係にある単語の時系列変化

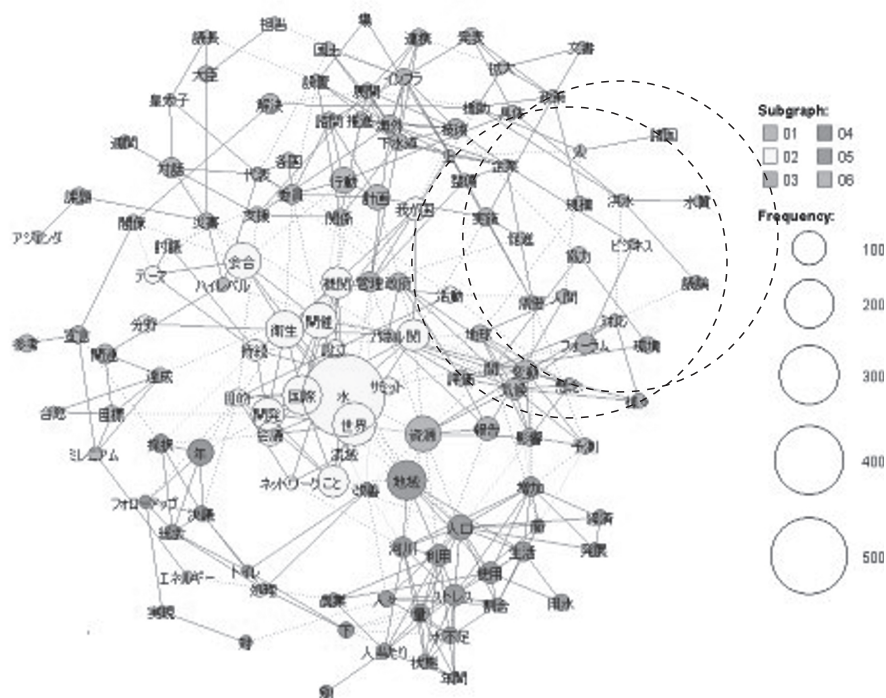
2009-2011



(出所) 筆者作成

図2 第2期(2009年—20011年)特定単語「協力」および特定単語「ビジネス」と共起関係にある単語の時系列変化

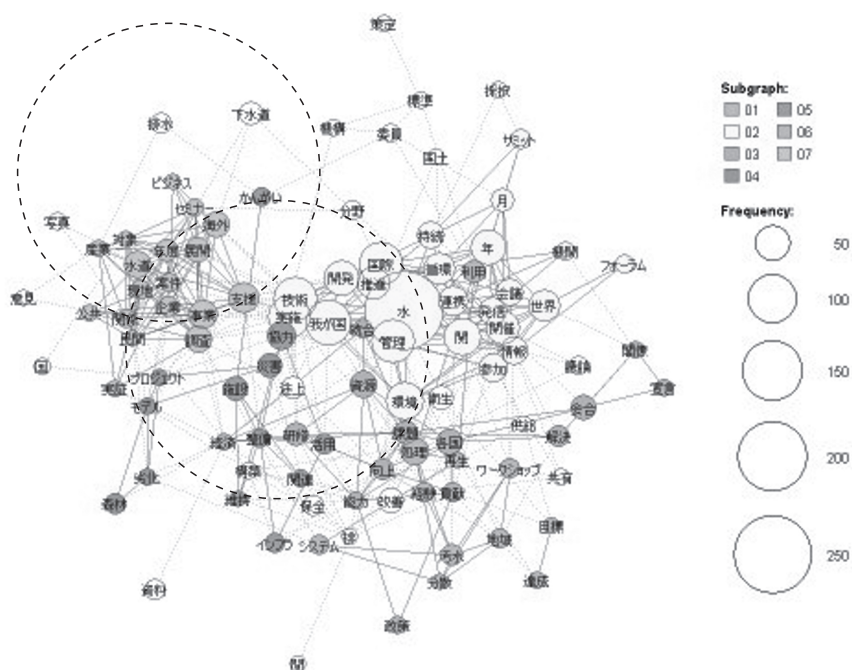
2012-2014



（出所）筆者作成

図3 第3期（2012年－2014年）特定単語「協力」および特定単語「ビジネス」と共起関係にある単語の時系列変化

2015-2017



(出所) 筆者作成

図4 第4期(2015年-2017年)特定単語「協力」および特定単語「ビジネス」と共起関係にある単語の時系列変化

表2 各期の特定単語「協力」の類義語上位30語及び類義語の類似度一覧（補正なし）

順	第1期 (2006年～2008年)	類似度	第2期 (2009年～2011年)	類似度	第3期 (2012年～2014年)	類似度	第4期 (2015年～2017年)	類似度
1	国際	0.9739	その他	0.9427	世界的	0.9696	図る	0.9630
2	環境	0.9688	促進	0.9421	高まる	0.9439	理解	0.9564
3	情報	0.9634	交流	0.9419	必要性	0.9418	適切	0.9563
4	貢献	0.9633	省庁	0.9353	よる	0.9411	維持	0.9548
5	経験	0.9628	経験	0.9304	貢献	0.9402	対して	0.9548
6	様々	0.9619	援助	0.9297	高まって	0.9371	資する	0.9541
7	積極的	0.9588	実施して	0.9210	積極的	0.9352	醸成	0.9528
8	他	0.9579	とも	0.9189	なか	0.9260	係る	0.9528
9	学術	0.9561	対して	0.9150	これまで	0.9189	下水	0.9511
10	つつ	0.9557	プログラム	0.9105	取組み	0.9130	通じた	0.9510
11	規模	0.9545	声明	0.9099	解明	0.9112	我が国	0.9505
12	管理	0.9535	行う	0.9050	きた	0.9058	○	0.9490
13	役割	0.9503	取組み	0.9040	様々	0.9004	ために	0.9489
14	果たす	0.9486	べき	0.9032	プログラム	0.9000	ソフト	0.9459
15	取組み	0.9468	関する	0.9007	①	0.8978	より	0.9458
16	フォローアップ	0.9465	きた	0.8982	経験	0.8916	水災害	0.9458
17	成果	0.9462	必要性	0.8979	現状	0.8863	国際	0.9454
18	及び	0.9461	政策	0.8915	事業	0.8845	活用	0.9454
19	ODA	0.9454	ため	0.8910	飲料水	0.8829	参入	0.9449
20	NGO	0.9443	環境	0.8868	水資源	0.8805	:	0.9447
21	技術	0.9434	国際的	0.8766	促進	0.8748	政府	0.9432
22	関連	0.9406	①	0.8750	される	0.8716	削減	0.9426
23	統合的	0.9401	実施した	0.8710	昭和	0.8662	洪水	0.9423
24	科学	0.9391	呼びかける	0.8697	援助	0.8645	対策	0.9422
25	議論	0.9348	執り行われた	0.8690	二国間	0.8625	質	0.9418
26	博覧会	0.9347	技術	0.8682	実施して	0.8615	高い	0.9416
27	ウ	0.9340	成果	0.8676	調査	0.8609	途上国	0.9407
28	政策	0.9339	ア	0.8676	懸念	0.8596	銀行	0.9403
29	国連	0.9339	推進	0.8665	発信	0.8575	経済	0.9398
30	きた	0.9327	努力	0.8652	問題	0.8549	諸国	0.9391

(出所) 杉浦（2019）を基に筆者修正

表3 特定単語「協力」の類義語 45 語の類似度の経年変化

番号	1	2	3	4	5
単語	プログラム	維持	飲料水	援助	下水
第1期(2006年～2008年)	0.7432				
第2期(2009年～2011年)	0.9105		0.8133	0.9297	
第3期(2012年～2014年)	0.9000		0.8829	0.8645	0.9511
第4期(2015年～2017年)		0.9548		0.9156	
番号	6	7	8	9	10
単語	我が国	説明	活用	環境	技術
第1期(2006年～2008年)	0.9230	0.8300		0.9688	0.9434
第2期(2009年～2011年)	0.7967	0.8648		0.8868	0.8682
第3期(2012年～2014年)	0.6679	0.9112		0.8517	0.7098
第4期(2015年～2017年)	0.9505		0.9454	0.8940	0.9000
番号	11	12	13	14	15
単語	銀行	経験	経済	懸念	呼びかける
第1期(2006年～2008年)	0.8656	0.9628	0.5725	0.7247	
第2期(2009年～2011年)	0.7404	0.9304	0.4559	0.7616	0.8697
第3期(2012年～2014年)	0.6212	0.8916	0.5902	0.8596	
第4期(2015年～2017年)	0.9403	0.8541	0.9398		
番号	16	17	18	19	20
単語	交流	洪水	貢献	国際	参入
第1期(2006年～2008年)	0.7564		0.9633	0.9739	
第2期(2009年～2011年)	0.9419		0.9402	0.6832	
第3期(2012年～2014年)		0.7982	0.8407	0.6964	
第4期(2015年～2017年)		0.9423		0.9454	0.9449
番号	21	22	23	24	25
単語	事業	質	取組み	省庁	推進
第1期(2006年～2008年)			0.9468	0.8696	0.9075
第2期(2009年～2011年)				0.9353	0.8665
第3期(2012年～2014年)	0.8845		0.9130		0.8120
第4期(2015年～2017年)	0.8500	0.9418			0.8523
番号	26	27	28	29	30
単語	水災害	水資源	世界的	成果	政策
第1期(2006年～2008年)		0.7673		0.9462	0.9339
第2期(2009年～2011年)		0.8060	0.8577	0.8676	0.8915
第3期(2012年～2014年)		0.8805	0.9696	0.6887	0.7507
第4期(2015年～2017年)	0.9458	0.7523			0.8388
番号	31	32	33	34	35
単語	政府	声明	積極的	促進	対策
第1期(2006年～2008年)	0.8388		0.9588	0.8544	
第2期(2009年～2011年)	0.6784	0.9099		0.9421	
第3期(2012年～2014年)	0.5995		0.9352	0.8748	
第4期(2015年～2017年)	0.9432		0.8851	0.9357	0.9422
番号	36	37	38	39	40
単語	調査	適切	途上国	努力	二国間
第1期(2006年～2008年)			0.8819	0.8642	
第2期(2009年～2011年)	0.8584		0.7139	0.8652	
第3期(2012年～2014年)	0.8609		0.6655	0.7463	0.8625
第4期(2015年～2017年)	0.8448	0.9563	0.9407		
番号	41	42	43	44	45
単語	発信	必要性	問題	様々	理解
第1期(2006年～2008年)			0.8702	0.9619	
第2期(2009年～2011年)		0.8979	0.8541	0.8645	
第3期(2012年～2014年)	0.8575	0.9418	0.8549	0.9004	
第4期(2015年～2017年)	0.8233		0.8264		0.9564

(出所) 杉浦 (2019) を基に筆者修正

対象テキスト=

第1期(2006年-2008年)水資源白書の国際関連の章

第2期(2009年-2011年)水資源白書の国際関連の章

第3期(2012年-2014年)水資源白書の国際関連の章

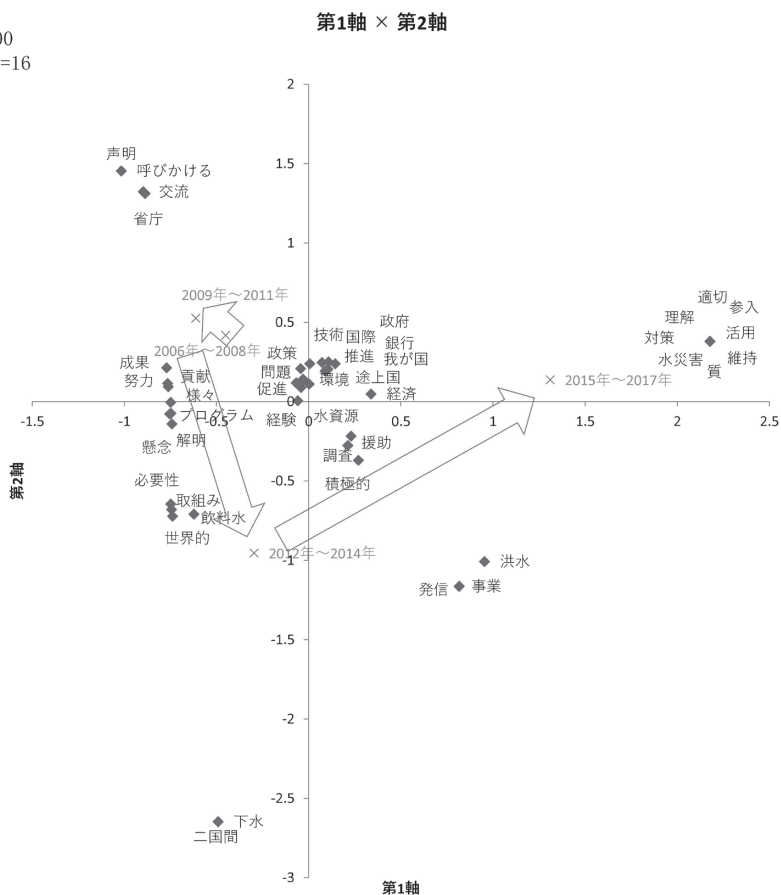
第4期(2015年-2017年)水循環白書の国際関連の章

特定単語=協力

sg=1

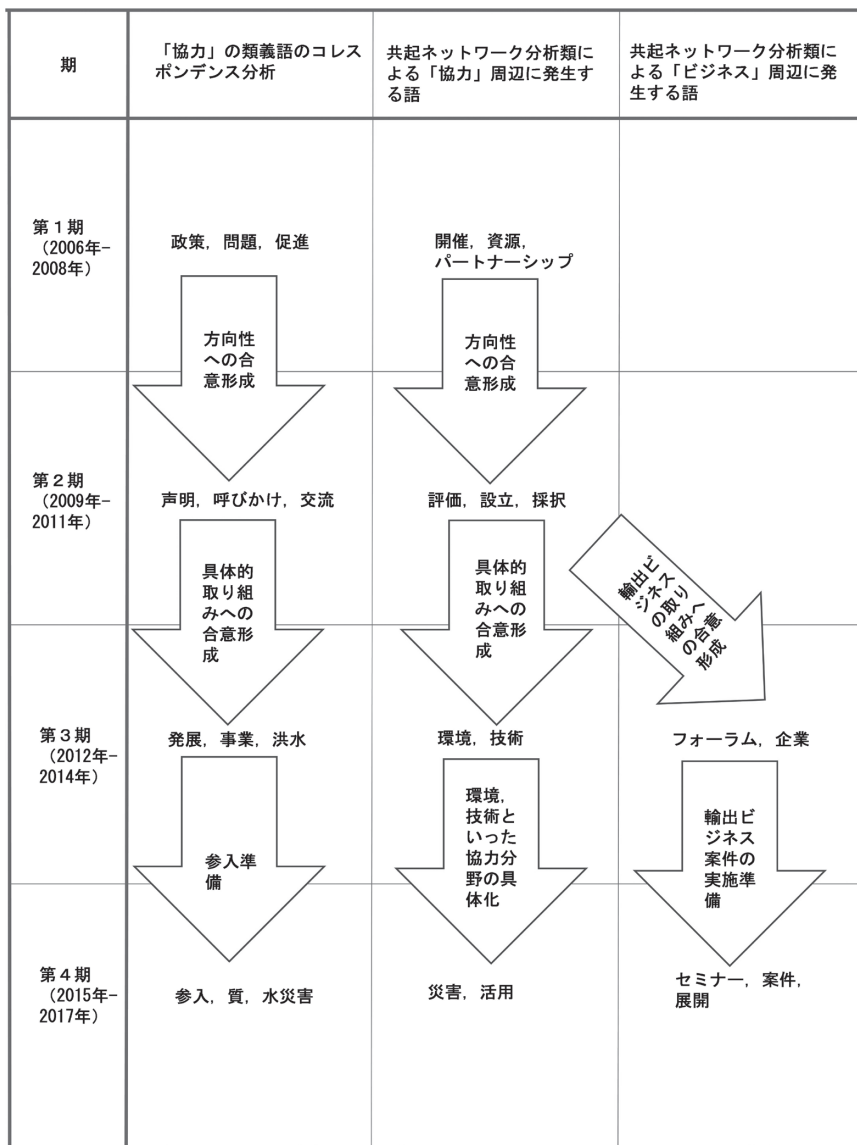
size=100

window=16



（出所）杉浦（2019）を基に筆者修正

図5 特定単語「協力」の類義語45語と各期の関係（コレスポンデンス分析結果）



(出所) 筆者作成

図6 日本の水資源行政の国際化の変遷 (国際協力と輸出ビジネスの関係)

4 結論

本研究では、日本の水資源行政の新たな国際化の流れを作った「海外社会資本事業への我が国事業者の参入の促進に関する法律（海外インフラ展開法）」（2018年）の施行前に、従来の国際分野においては国際協力という取り組みの流れの中で、どのように輸出インフラを扱い、どのように分派していったのかをテキストマイニングにより探ることを試みた。

そこで、日本の水資源行政を担う政府が発行する水資源白書や水循環白書の国際関連の章を分析対象として、特定単語「協力」及び特定単語「ビジネス」と共起関係にある単語の時系列変化、特定単語「協力」の類義語群の時系列変化を分析した。その結果、日本の水資源行政の国際化は、相手国への協力と並行して、輸出ビジネスへという流れができてきていることに気づくことができることを示唆することができた（図6）。また、図6に示すとおり、国際協力の方向性への合意形成の中で、輸出ビジネスの取り組むコンテンツに触れていることも示された。そして、協力として取り組む環境保全や災害対策といったコンテンツとは別に、輸出ビジネス案件の準備を進めていることに気づくことができることも示されていることがわかった。

以上のような成果を示した一方で、本研究による分析には、次のような課題も残っている。まず、共起関係を分析するために用いた特定単語の妥当性について、慎重に検討する必要がある、また、本研究の分析対象は、白書の国際関連の章のみであるため、他の政府発行文書も取り入れ分析することで、分析精度を高めることも検討しなければならないと考えられる。以上のような課題を残したものの、本研究の成果は、日本の水資源行政の国際化の流れの中で、輸出ビジネスの位置づけを客観的に示すことにより、過去の経緯を把握したうえで輸出ビジネスを考えることができ、輸出ビジネスの発展にも寄与するものと考えられる。

【参考文献】

- A. Mnih and K. Kavukcuoglu (2013) Learning word embeddings efficiently with noise-contrastive estimation. In Proceedings of NIPS, pp.2265-2273.T.
- Mikolov et al (2013) Efficient estimation of word representations in vector space. arXiv:1301.3781.
- 杉浦政裕 (2010) 「テキストマイニングによるインドネシアにおける水資源インフラストラクチャニーズの分析」『土木学会論文集F 4』, 66 (1), pp.27-35.
- 杉浦政裕 (2019) 「深層学習による日本の水資源行政の国際分野の動向分析」『経営戦略としてのテキストマイニング』, pp.91-115, 中央経済社.
- 福原知宏, 村山敏泰, 中川裕志, 西田豊明 (2006) 「Weblogから社会の関心を探る」『人工知能学会第20回全国大会』, 3D2-01.

