

生存圏ミッション研究のキーワード統計解析

上田 義勝^{1*}, 梅澤 俊明²

Statistical Keyword Analysis of Humanosphere Mission Research

Yoshikatsu Ueda^{1*} and Toshiaki Umezawa²

概要

生存圏ミッション研究は、生存圏研究所の中心となる重要な研究テーマである。研究所発足当初から継続されている重要な研究テーマから、新たに融合研究として生まれつつある研究テーマまで、幅広いミッションが存在している。本報告においては、このミッション研究の報告書の文章をキーワード解析する事で、研究の広がりについて解析を行った。この解析は生存圏研究所ミッション推進委員会として行ったものであるが、解析結果を資料として残す意味からも重要であると考え、まとめている。実際に解析を行ったのは平成 28 年度から令和 3 年度までの研究報告書を用いた。また、解析には R 言語を用いて行っている。

1. はじめに

生存圏研究所では、重要な研究ミッションとして 4 つのミッションを掲げている。これらは「環境診断・循環機能制御」、「太陽エネルギー変換・高度利用」、「宇宙生存環境」、「循環材料・環境共生システム」からなり、また 5 つ目の発展型ミッションとして「高品位生存圏」を定義し、研究成果の実相を含めた社会貢献を目指している。特にミッション 5 については、平成 28 年から新設されたミッションであり、社会とのつながりや国際化、物質・エネルギーの循環を、これまでよりも重視している。これらのミッション研究を平成 28 年から令和 3 年まで続けてきているが、その研究の広がりを確認する事で、これまでの研究所活動を見直し、今後研究所をより発展させるためのデータとして活用する事は非常に重要であると考ええる。

各ミッションの詳細な内容については生存圏研究所概要などに記載されているので割愛するが、それぞれのミッションでキーワードとなる文言が存在する。例えばミッション 1 では「環境」や「大気」、ミッション 2 では「バイオマス」や「マイクロ波」、ミッション 3 では「宇宙」や「プラズマ」、ミッション 4 では「木質材料」や「構造」など、それぞれのミッション研究を言い表す単語がある。我々はこのキーワードに着目し、年度ごとにその傾向がどう変化するか確認する事で、生存圏研究所の研究の広がりを表す事を試してみた。また、同様の研究の広がりとして、各ミッション研究の個別のテーマについてアンケートを取り、持続可能な開発目標（SDGs, Sustainable Development Goals）と、社会的共通資本（SCC, Social Common Capital）にどの程度当てはまるのかについても、調査を行った。この調査についてはミッション研究だけでなく、外部資金や国際共同研究/フラッグシップ研究についての調査し、その研究の広がりについて確認した。

2022 年 7 月 13 日受理.

¹〒611-0011 宇治市五ヶ庄 京都大学生存圏研究所 生存圏未来開拓研究センター
先端計測技術開発ユニット

²〒611-0011 宇治市五ヶ庄 京都大学生存圏研究所 森林代謝機能化学分野

* E-mail: ueda.yoshikatsu.4e@kyoto-u.ac.jp

2. 解析手法について

本章においては、解析に用いた R 言語¹⁾について紹介し、用いたプログラムについて解説する。R 言語はオープンソース・フリーソフトウェアの統計解析向けのプログラミング言語及びその開発実行環境であり、ベクトル処理による柔軟な処理を簡便に行なう事が可能である²⁾。

2.1 R 言語を用いたキーワード解析手法

我々はこの R 言語を用い、各ミッション研究の報告書内に含まれるキーワードを統計解析し、その解析結果を可視化する事で、年度ごとの傾向を確認する事を目的とした。R 言語のインストールは Windows 環境にて行い、機能提供ツールとして R tools³⁾、プログラムエディタとして R studio⁴⁾、また日本語分析のためのライブラリとして、MeCab⁵⁾もインストールした。また、解析終了後の可視化には、R 言語プログラムの wordcloud2⁶⁾パッケージを用いている。

プログラムの一例を図 1 に示す。日本語解析のための RMeCab ライブラリ、分析のためのパッケージとして tidyverse ライブラリを読み込む。また可視化のための wordcloud2 ライブラリと、可視化した後に画像として保存するための webshot、htmlwidgets ライブラリを読み込んだ後、テキスト化した文章をプログラムに沿って解析していく形となる。

```
library('RMeCab')
library('tidyverse')
library('wordcloud2')
library('webshot')
library('htmlwidgets')

sample <- RMeCabFreq("D:/2021/2021mission1.txt")

result <- sample %>%
  as_tibble %>%
  filter(!is.na(noun)) %>%
  filter(Freq>2) %>%
  mutate(noun=str_match(Term, '[A-Z0-9-9]+')) %>%
  filter(str_length(noun[,1])>2) %>%
  select(5, 4) %>%
  arrange(desc(Freq)) %>%
  filter(!noun[,1] %in% c("こと", "たち", "皆さん", "ため", "京都", "ミッション", "生存", "研究", "研究所", "京都大学", "大学", "京大", "of", "の", "top", "課題", "共同", "in", "and"))

my_graph <- result %>% wordcloud2(fontFamily='HG丸ゴシック-PRO', minRotation=0, maxRotation=0, size=0.8, shape='circle')
saveWidget(my_graph, "tmp.html", selfcontained = F)
webshot("tmp.html", "R1_mission1.png", delay = 5, vwidth = 1024, vheight = 1024)

write.csv(result, "R1_mission1.csv")
```

図 1：wordcloud2 による解析プログラム例

プログラム中で追加的な処理としては、例えば「京都大学」や「ミッション」など、解析では不要だが、カウントが多くなる一般的な単語を削除している。同様に個人名についても出来るだけ削除し、研究としてのキーワードの広がりやすくなる様に解析している。その他、可視化の際の可視化の形状 (circle) や解像度なども調整し、画像保存している。解析結果自体は可視化の他、csv ファイルとしてデータ保存も行った。

2.2 SDGs と SCC とのつながりについて

同じく R 言語を用いることで、各ミッションで設定されている個別の研究テーマが、社会的なテーマとどのようにつながるかについてテーマごとにアンケートを取り、そのつながりをミッション毎に確認している。また、大型外部資金や国際共同研究/フラッグシップ研究などについても同様にアンケートを取り、調査を行っている。こちらは統計解析ではなく、繋がりを可視化するために R 言語を用

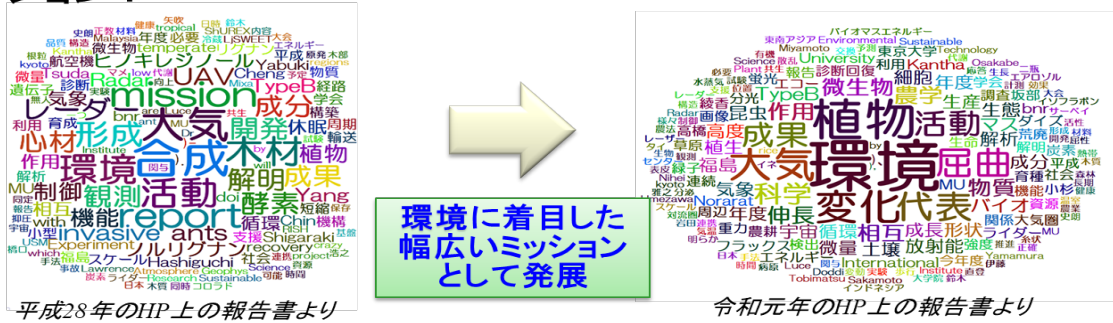
いた。具体的には stars 関数を用いることで、ミッション毎の広がりを半円のグラフとして表記した。

3. キーワード解析による生存圏ミッション研究の変遷

3.1 R 言語を用いたキーワード解析（平成 28 年度と令和元年度の比較）

図 2-5 に、各ミッションのキーワード解析結果（平成 28 年と令和元年）の変化について示す。文字の大きさは規格化していないため、各図同士の対応は取れていないが、文章中に含まれるキーワードの分布がそれぞれ変化している事がわかる。ミッション 1 においては、平成 28 年においては元々が大気やレーダー、木材といった本来の研究テーマの単語が多かったが、令和元年においては、環境という大きなキーワードを中心としたミッションに発展している事がわかる。ミッション 2 では、バイオマスに関連した研究が平成 28 年に多かったが、令和元年までにマイクロ波を利用した融合研究として発展をしてきたことが良く分かる結果となった。一方でミッション 3 においては、宇宙という大きなキーワードが中心ではあるが、その周辺となるキーワードが、プラズマ波動だけでなく、エネルギー利用などに向けた分野横断的な研究が進みつつある事が示唆された。ミッション 4 でも同様に木材という大きなキーワードを中心とするが、他のキーワードが数多くみられることで、同じく研究に広がりが見えている事が良く分かった。図 4 と 5 では、ミッション 5 の各サブミッション（ミッション 5-1 人の健康・環境調和（生理活性物質、電磁波、大気質）、ミッション 5-2 脱化石資源社会の構築（植物、バイオマス、エネルギー、材料）、ミッション 5-3 日常生活における宇宙・大気・地上間の連関性、ミッション 5-4 木づかいの科学による社会貢献（木造建築、木質住環境、木質資源・データベース、木づかいの変遷）におけるキーワードの変遷を示している。ミッション 5 は新設時の時点で幅広い研究テーマ設定が行われている事から、キーワードとして抽出される単語は多いが、それでも同様に研究の広がりが見える結果となった。

ミッション1



ミッション2

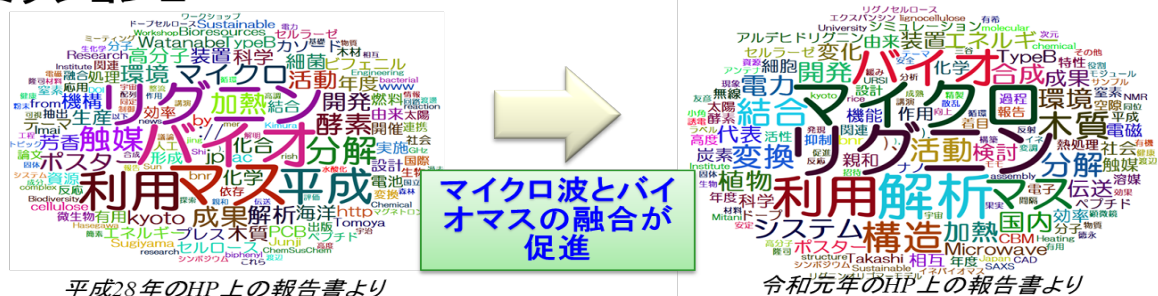
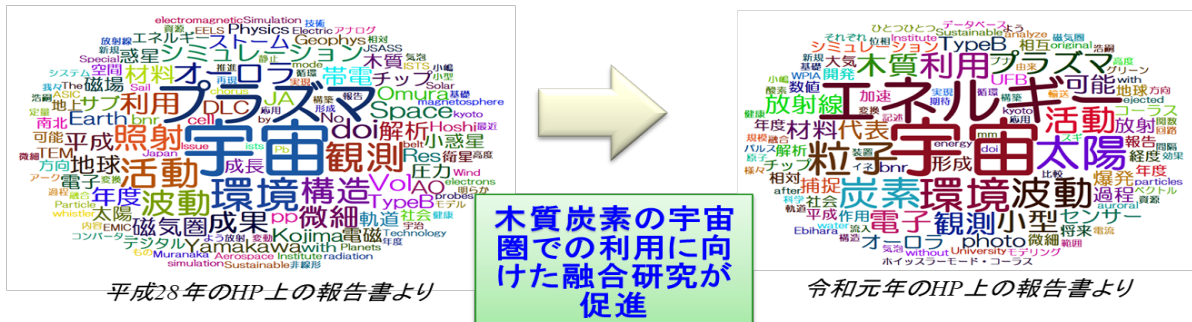


図2：平成28年と令和元年の比較（ミッション1と2）

ミッション3



ミッション4

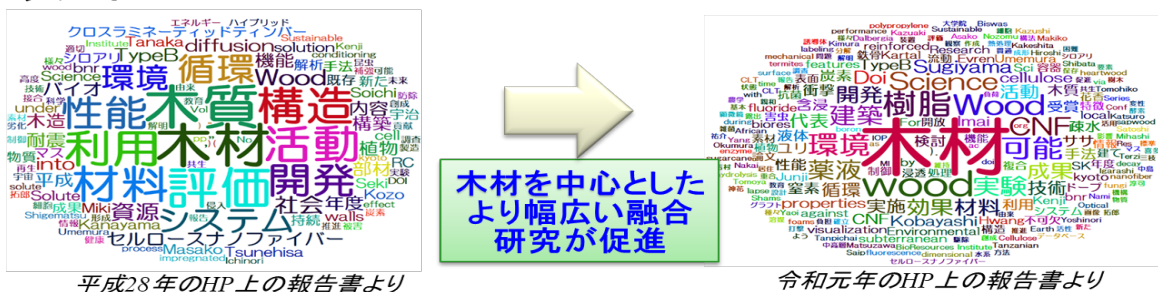
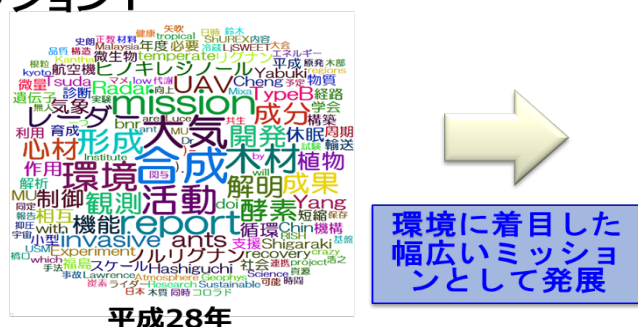


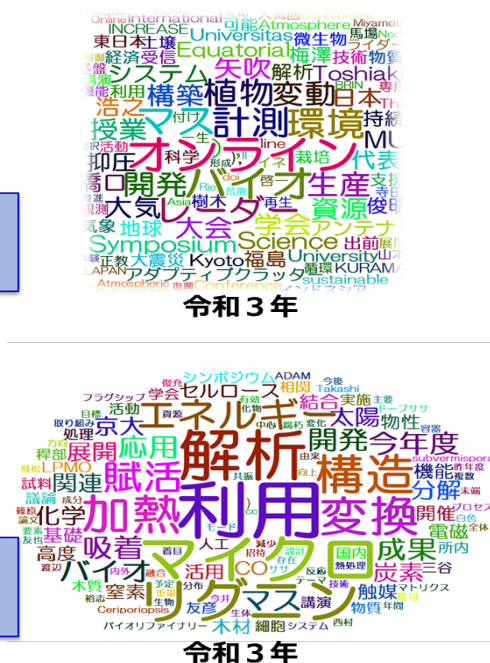
図3：平成28年と令和元年の比較（ミッション3と4）

図 6-9 では、さらに年度が経過した際の比較として、令和 3 年度と平成 28 年度の比較を行った。3.1 で述べた状況と同じものもある（ミッション 1 と 2）が、特にコロナ対策などでミッション 1 ではオンライン会議を積極的に利用した傾向も見えて取れる。ミッション 2 においては、令和元年と比較すると今度は利用と解析が中心となりつつあることから、成熟した研究テーマも増えているとも予想される結果となった。ミッション 3 においては図 3 と比較すると、キーワードがエネルギーから環境にシフトしつつあることから、宇宙圏の環境を見据えた研究テーマも増えてきている傾向がみられ。ミッション 4 では令和元年で非常に数多くの単語を使って解説されていた研究が、徐々に集約されて（「残留」や「圧延」、「応力」など）、テーマが選別されてきている様子も良く分かる結果となった。

ミッション1

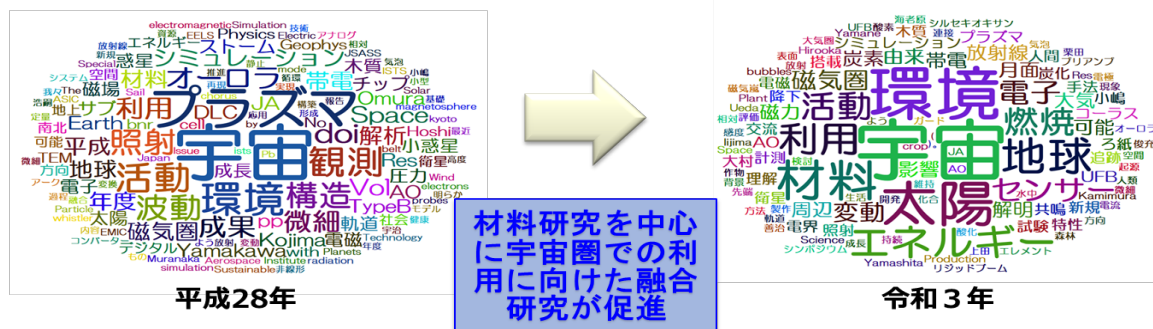


ミッション2



40

ミッション3



ミッション4

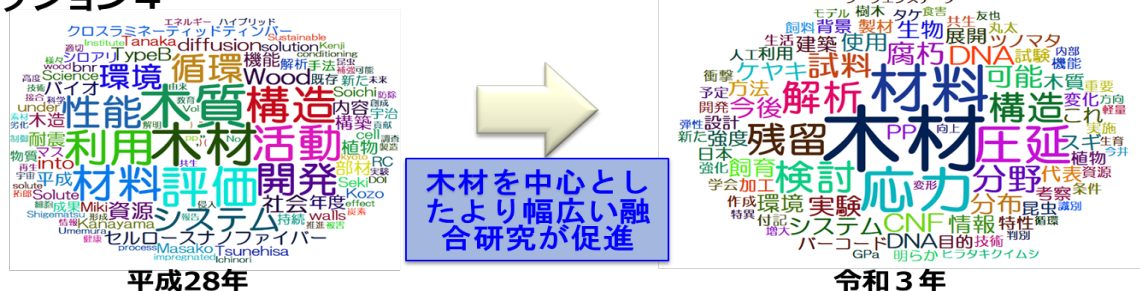


図7：平成28年と令和3年の比較（ミッション3と4）

ミッション5-1



ミッション5-2



図8：平成28年と令和3年の比較（ミッション5-1と5-2）

ミッション5-3



平成28年



宇宙圏から大気圏までの幅広い領域に展開している



令和3年

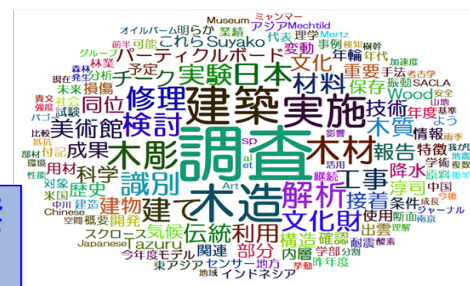
ミッション5-4



平成28年



文理融合や調査研究が発展している



令和3年

図9：平成28年と令和3年の比較（ミッション5-3と5-4）

3-4 平成28年から令和3年までの解析結果からみるキーワードの変遷について

3-1,3-2においては、視覚的にキーワードがどう変化したかを確認したが、文字の大きさは全体で規格化出来ているわけでは無いので、数値としても本節にて評価を行う。具体的な数値指標として、R言語の解析結果においてFrequency値（Freq）を用いた。各ミッションにおける代表的なキーワードを選び、令和3年（R3）、令和元年（R1）、平成28年（H28）における数値変化を表1に示す。

ミッション1に関しては、「大気」「MU」や「レーダー」のキーワードについては、一部減少はあるが、ほぼ毎年同程度の数値であった。一方、新しく抽出されてきているキーワードとして「福島」や「樹木」などがある。「福島」については、東日本大震災での支援研究が継続して研究が進んでいると考えられる。また、「樹木」については、「木材」がH28において頻度が高い一方で、R1以降数値が0になっていることから、研究分野が細分化しつつあることがわかる。ミッション2に関しては、「リグニン」や「マイクロ」「Microwave」などのこれまでの研究が継続して活発に行なわれている事がわかる。一方で「バイオマス」の数値が若干減りつつあるが、「構造」に関する研究や「セルロース」などのキーワードがR3にて再度増加しているため、ミッション1と同様、研究が広く細分化しつつあるのでは無いかと考えている。ミッション3では、これまでの継続研究として「宇宙」が継続して残っているが、特にR3にかけては、「太陽」「センサー」「大気」など、新しいキーワードが出てきている。その一方で「Plasma」というキーワードが若干減りつつあるが、これは分野の広がりというよりは、より大きな研究プロジェクトが進んでいるためでは無いかとも考えている。ミッション4においては、「木材」研究が中心である事が良く分かる傾向となっているが、新たなキーワードとして、「応力」や「材料」「構造」「CNF」など、具体的な研究課題が設定されつつあることが分かる。

ミッション5-1から5-4については、ミッション1から4をより社会的に広げ、国際的な研究として進めているため、より具体的なキーワードがいくつか出てきている。ミッション5-1においては、「洗浄」や「微細」など、環境工学できな研究発展がみられる。また、生存圏研究所の興味深い特色としては、「ウイルス」に関してはCOVID-19の影響もあり増加しているが、それと関係なく、過去か

ら着実に微生物研究をすすめてきている事も良くわかる。ミッション 5-2 については、ミッション 2 と同じ傾向だが、特に「社会」や「エネルギー」のキーワードが増加しており、ミッション 5 としての特徴を良く表している。ミッション 5-3 についてもミッション 3 と同傾向であるが、「観測」を中心とした研究が広く行われ、また「データ」からも大規模データを用いた研究が進んでいる事もわかる。ミッション 5-4 については、文理融合としての「木彫」「文化財」「伝統」などが上昇傾向にあり、生存圏研究所の幅広い研究の広がりを良く表している。

表 1：各ミッションの代表的なキーワードの Freq 値の変化

Mission 1				Mission 2				Mission 3				Mission 4			
	R3	R1	H28		R3	R1	H28		R3	R1	H28		R3	R1	H28
レーダー	10	2	14	リグニン	12	15	14	宇宙	15	12	11	木材 (wood)	29	37	21
大気	8	8	8	マイクロ	12	14	7	太陽	11	2	3	応力	21	0	0
MU	7	8	14	Microwave	12	17	8	センサー	6	0	0	材料	21	4	7
福島	6	6	3	構造	10	9	0	大気	4	0	0	構造	13	3	10
樹木	5	0	0	バイオマス	6	10	14	Plasma	3	7	10	CNF	9	10	0
微生物	5	5	3	セルロース	4	0	4	Space	2	14	12	Cellulose	0	10	4
木部	3	0	2					UFB	2	2	2				
木材	0	0	8												
Mission 5-1				Mission 5-2				Mission 5-3				Mission 5-4			
	R3	R1	H28		R3	R1	H28		R3	R1	H28		R3	R1	H28
物質	29	16	9	バイオマス	21	20	13	観測	25	19	16	建築	16	8	6
活性	27	18	12	植物	19	17	5	大気	13	16	8	木材 (wood)	12	18	17
遺伝子	18	12	9	リグニン	11	14	8	地磁気	9	2	0	木彫	12	0	0
洗浄	15	0	0	エネルギー	10	11	6	電離	9	4	4	文化財	9	4	0
微細	15	0	0	社会	9	0	4	データ	8	6	6	伝統	8	4	6
ウイルス	11	7	9	細胞	9	9	0	衛星	7	5	2	構造	6	5	7
				マイクロ	4	11	9								

3-4 SDGs と SCC とのつながりについて(令和 3 年度)

令和 3 年度の個別研究テーマについて、持続可能な開発目標 (SDGs) と、社会的共通資本 (SCC) の分類を Stars として図 10 と図 11 に示す。これらは SDGs と SCC の項目毎に、ミッション研究と関連したものかどうかを扇状に示したものである。ミッション 1 から 4 の傾向とミッション 5-1 から 5-3 の傾向については、おおよそ同じような傾向となった。水関連 (微細気泡) や文理融合など、一部発展して新しい項目が増えている部分もあるが、大きく目標や項目が増えた研究テーマはあまりはっきり見られなかった。その一方で、大型外部式や国際共同研究/フラッグシップ研究については、ミッション研究を補う部分も多いからか、項目の系統がミッション研究とは違っており、双方が補い合う事で、生存圏研究所の幅広さが見える結果となった。

SDGs目標とミッション系統に対する関係

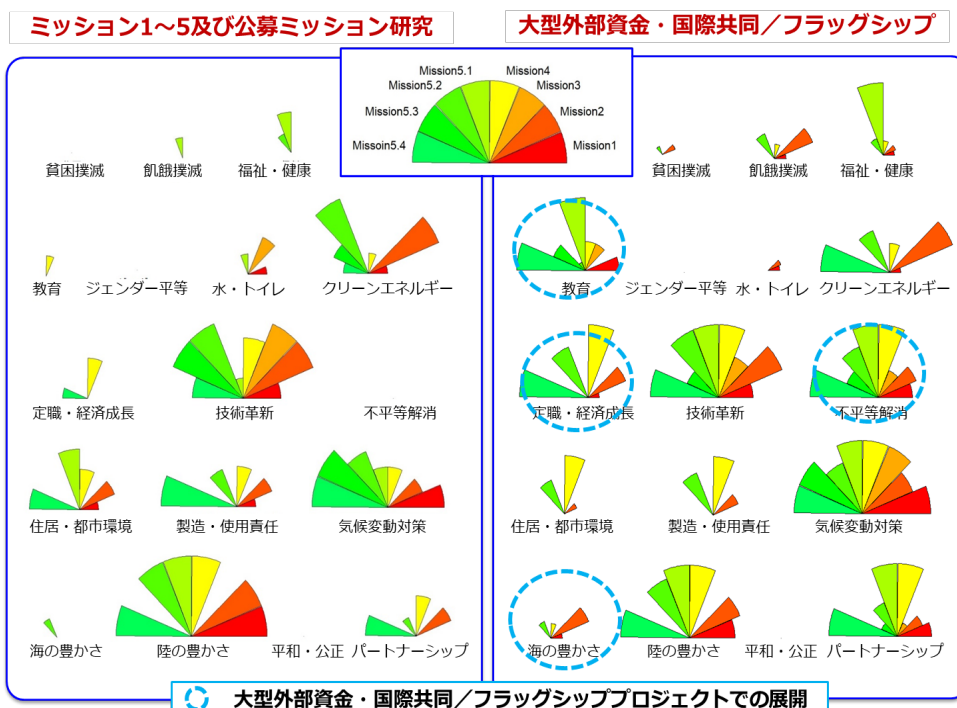


図 10：令和 3 年度におけるミッション研究の広がり（SDGs）（点線で囲んだ部分は新たに広がりを持った箇所）

SCC項目とミッション系統に対する関係

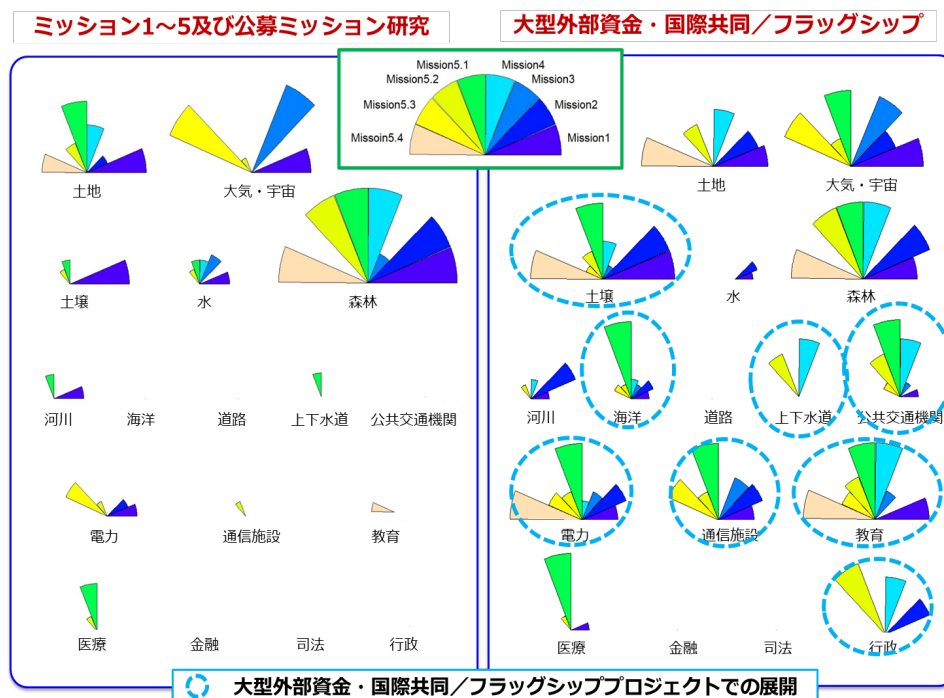


図 11：令和 3 年度におけるミッション研究の広がり（SCC）（点線で囲んだ部分は新たに広がりを持った箇所）

4. まとめ

生存圏研究所のミッション研究の発展の経緯を、キーワード解析を行う事で示し、継続されてきた研究の重要性と、その研究のひろがりを示すことが出来た。平成 28 年度から令和 3 年にかけて、各ミッションで中心となるキーワードを中心に、新しい単語として融合研究や新研究テーマが発生している事が良く分かった。また SDGs や SCC との関係性においては、ミッション研究単体での広がりには 1 年では示しきれなかったが、外部資金や国際共同研究などで補完的に広く研究を行っている事が分かり、生存圏研究所の深いポテンシャルを示すことが出来ている。尚、この解析は生存圏研究所ミッション推進委員会として行ってきたものであるが、今後の委員会への引継ぎも含め、手法として生存圏科学の発展に役立てば幸いである。また、プログラムについては研究所内の共有サーバに保存し、今後使用していただける様にしたい。

参考文献

- 1) R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- 2) Widemann, B.T.y., Bolz, C.F., and Grelck, C., The Functional Programming Language R and the Paradigm of Dynamic Scientific Programming. In: *Trends in Functional Programming* (Loidi, H.W. and Peña, R. eds.), vol. 7829, 2013, pp. 182-197, Springer Berlin, Heidelberg.
- 3) R Tools (ver4.2) <https://cran.r-project.org/bin/windows/Rtools/>
- 4) R Studio <https://www.rstudio.com/>
- 5) 形態素解析システム MeCab <https://www.mlab.im.dendai.ac.jp/~yamada/ir/MorphologicalAnalyzer/MeCab.html>
- 6) wordcloud2 package <https://cran.r-project.org/web/packages/wordcloud2/>

著者プロフィール



上田 義勝 (Yoshikatsu Ueda)

＜略歴＞ 2002 年京都大学情報学研究科単位取得退学／2004 年京都大学博士（情報学）取得／2002 年京都大学宙空電波科学研究センター 助手／2004 年京都大学生存圏研究所 助手／2007 年同助教より現在に至る

＜研究テーマと抱負＞水中の微細気泡の特性計測とその応用利用。また福島県における支援研究など。＜趣味など＞合気道（五段位）、犬の散歩



梅澤 俊明 (Toshiaki Umezawa)

＜略歴＞ 1980 年京都大学農学部林産工学科卒業／1982 年京都大学大学院農学研究科修士課程修了（農学博士）／同年京都大学木材研究所（現生存圏研究所）助手／1993 年同助教授／2005 年同教授現在に至る。＜研究テーマと抱負＞環境保全と調和した持続的社会の基盤となる植物バイオテクノロジー研究。＜趣味など＞散歩と乱読、IPA をたまに飲むことなど。