

文献レビューにおける構造トピックモデルの活用 —心理学概念セルフ・コンパッションを題材として—

岡野裕仁 村上遥 有海春輝 野村理朗

京都大学

{okano.hirohito.38m, murakami.haruka.32m}@st.kyoto-u.ac.jp, aharuki0408@gmail.com,
nomura.michio.8u@kyoto-u.ac.jp

概要

数千件を超えるような、大規模な先行文献すべてを対象にレビューすることは困難を伴う。このよう
なとき、タイトルや抄録に対するトピックモデルの
適用が有効である。今日では、従来の古典的なトピ
ックモデルである潜在ディリクレ配分法 (LDA) 等
ではなく、論文情報に紐づいたメタ情報をモデリン
グ過程で有効活用できる構造トピックモデル (STM)
の有用性が示唆されつつあるが、心理学を含む人文
学における文献レビューでは STM はほとんど用いら
れていない。こうした研究動向を踏まえ、本研究で
は、「自分への思いやり」をさすセルフ・コンパッ
ションという心理学的概念を題材とし、STM を用い
たレビューを行った。結果、解釈可能な 38 のトピ
ックが、トピック割合の時間的推移と共に抽出され、
先行研究における主要なテーマを得ることができた。

1 はじめに

「困難な状況で自分に向けられる思いやり」をさ
すセルフ・コンパッション (self-compassion) は、
2003 年に初めて定義されて以来、心理学をはじめと
した様々な分野で研究が進められ、特に近年では研
究の増加が顕著である [1], [2], [3]。これまで、特性
的なセルフ・コンパッションの高さや、セルフ・コ
ンパッションを向上させる介入は、精神的健康 [4]
や 身体的健康 [5]、人生満足度 [6] など、様々なポ
ジティブな結果をもたらすことが示されてきた。今
日ではセルフ・コンパッションは人々の well-being
を高める重要な変数や介入ターゲットとしてみなさ
れ、多様なサンプル、分野、手法で研究されており
[2], [7]、現在では論文数は 4,000 件を超える。

研究の劇的な増加に伴い、特定のテーマや問題に
焦点を当てたセルフ・コンパッションに関する系統
的レビュー [8], [9], [10] やメタ分析 [4], [6], [7] も
徐々に増えてきている。しかし、これらのレビュー

は特定のトピックや問いに焦点を絞っており、先行
研究を全体的に俯瞰するものはほとんどない。先行
研究における主要な研究テーマやトピックを把握す
るためには、より大規模な文献サンプルに基づくス
コーピングレビュー (ある研究領域の基盤となっ
ている主要な概念、利用可能な文献やエビデンスを素
早くまとめるレビュー [11]) が必要である [12], [13]。

しかし、従来の方法でセルフ・コンパッション研
究のスクーピングレビューを行うためには、現在ま
での 4,000 件を超えるすべての文献を読んだう
えでレビューする必要がある、大きな困難を伴う。本
研究は、大規模な文献レビューにおけるこうした課題
点を解決する手段として近年有用性が注目されてい
る構造トピックモデル (structural topic model ; STM
[14]) を用いたスクーピングレビューを試みた。

2 関連研究

セルフ・コンパッション研究の文脈で自然言語処
理を活用したレビューは我々の知る限り Swami ら
[13] によるものだけである。彼らは、2020 年までに
出版されたセルフ・コンパッションに関する論文の
タイトル・抄録・キーワード中に出現する単語に対
して共起ネットワーク分析を適用し、文献における
主要なキーワードと関連性を可視化した。また、そ
れらのキーワードが 4 つのクラスターに分類される
ことを示した。Swami らによる研究はセルフ・コ
ンパッションの文献レビューにおける自然言語処理活
用の有望性を示した一方で、2 つの限界点が存在す
る。第一に、特定されたクラスター数が少なく、こ
れまでの研究文献にいかなる具体的トピックが存在
しているのか、詳細を把握することが困難である。
第二に、各単語の共起やトピックは時間経過と共に
変遷していくにもかかわらず、共起ネットワークの
導出過程で時間情報 (すなわち、論文の出版年) を
考慮していない。このような問題点を解決するた
めに、STM [14] の適用が有効であると考えられる。

トピックモデルとは、文書中に潜在するトピックを抽出する手法の総称であり、(a) 多数の文書の背景にあるトピックの抽出、(b) 各トピックで出現する確率の高い単語の特定、(c) 各文書において優勢なトピックの特定、(d) トピック同士の相関を可視化が可能となる。トピックモデルの強みは、単語どうしの共起の可視化に留まらず、人間が解釈可能な個々の具体的トピックを発見・記述できる点にある [15]。トピックモデリングによるスコーピングレビューは、心理学内外の学術分野で行われてきた (例: [12], [15], [16], [17], [18], [19])。例として、Kee ら [12] は、心理学的構成概念の 1 つマインドフルネスに関連する文献を対象に、トピックモデルを活用したスコーピングレビューを行っている。

従来最も広く使われてきたトピックモデルは、古典的な潜在ディリクレ配分法 (latent Dirichlet allocation; LDA [20]) であるが、LDA ではトピック間の相関を考慮できず、トピックの推定プロセスにおいて各文書に紐づいたメタ情報 (文献レビューでは、論文の出版年や掲載誌の学術情報等が考えられる) を有効に活用することもできない。一方、LDA の拡張である STM は、トピック間の相関を表現できる点、また文書のメタ情報を使用してトピック-メタ情報間の潜在的な関連を直接モデル化しつつ、より質の高いトピックを抽出できる点 LDA より優れている [21]。それゆえ近年では、トピックモデルによる文献レビュー手法は、自然科学を中心とした様々な分野において、LDA から STM に移行しつつある (例: [16], [21], [22], [23], [24], [25], [26]) が、人文学分野における活用は非常に限られている¹。

以上の研究動向を背景として、本研究は、心理学的構成概念の 1 つセルフ・コンパッションに関して、これまでに出版されてきた文献のタイトルと抄録に STM を適用することにより、文献に存在するトピックを抽出し、現存する文献における主要なトピックとその間の関連を明らかにすることを目的とし、同時に心理学・人文学における文献レビューにおける STM の有用性を検討した。また、モデリングの際に、2 つの変数 (論文の出版年と論文の掲載誌の学術分野) をメタデータとして投入することで、トピックとこれら 2 つの変数の関連を可視化した。

¹ LDA や STM 以外に利用可能な最新のトピックモデル手法として、深層学習に基づく BERTopic[35]や大規模言語モデルがある。しかし、前者は STM のようにメタ情報をトピック推定過程で適切に活用するための確立された方法論が存在せず、後者はハルシネーションが生じるリスクのコントロールが困難である。

3 データ収集とモデル構築

文献情報の収集 先行研究 [12]の手続きに従って行われた。2024 年 7 月 1 日に、Web of Science Core Collection を用いて、検索文字列 (TS=(self-compassion)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article)) で検索を行った。すなわち、英語の原著論文であり、文献のタイトル・抄録・キーワードのいずれかに self-compassion の語を含む論文が検索された。その後、抄録が空である論文を削除し、最終的な論文数は 4,395 件となった。

テキストの前処理 各論文について、先行研究 [12]と同様、論文のタイトルと抄録テキストを結合した。このとき、トピックモデリングにおいてノイズとなりうる語はストップワードとして削除した。

トピックモデリング R のパッケージ stm [14]を使用した。トピック普及率共変量として、Web of Science 上で入手可能なメタ情報である論文出版年 (4 次スプライン) と論文の掲載誌の学術分野を投入した。掲載誌の学術分野として、研究が特に多く発表されている分野を選び (*psychology, education, nursing, psychiatry, public health, health care, social work, family studies*)、各論文が各学術分野に分類されているかどうかを示すダミー変数として投入した。トピック総数 K は、ガイドライン [27] に従い、5 以上 150 以下のすべての整数を検討し、トピックの質を評価するための 4 つの指標 (意味的一貫性、排他性、変分下限、残差) をプロットした。それらの指標に基づいて最適な K は 41 であると結論づけた。

トピックの命名 トピックの命名手続きも、ガイドライン [27] に従った。本研究では、各トピックについて出現確率が高い 20 語を出力し、トピック割合が高い論文のタイトルと抄録を 50 件出力した。これらの単語と論文を確認し、トピックを命名した。明確なテーマを見出すことが難しく命名が困難であるトピックは分析対象から除外した [27], [28]。

可視化 トピック間の関連を可視化するために、各トピックの相関をネットワークグラフとして表示した。グラフを見やすくするために、ネットワークグラフにおけるコミュニティ検出手法 infomap [29] を用いてトピックをクラスタリングした。

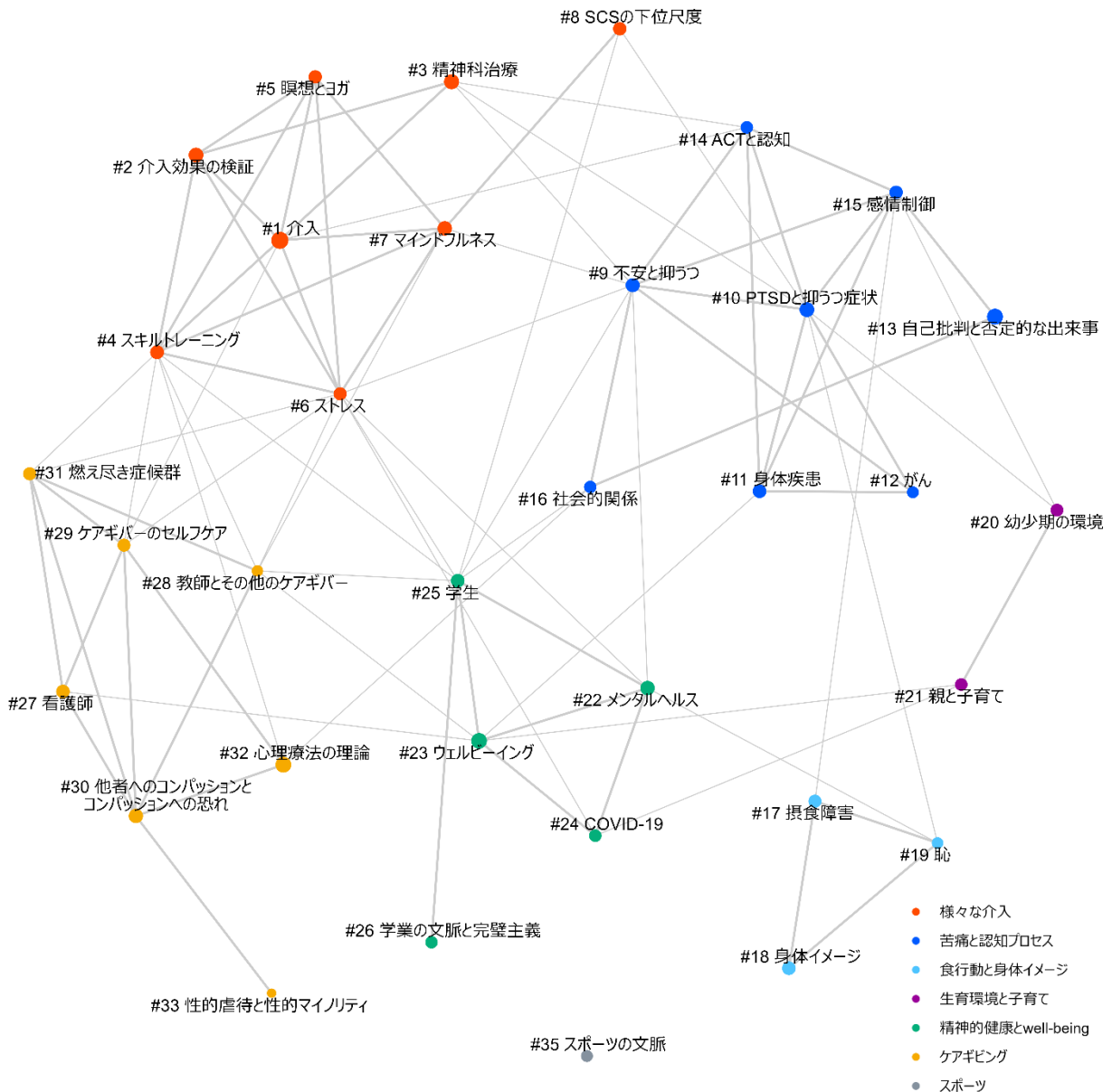


図 1 トピック間の相関を表すネットワークグラフ (ノードの色はクラスタに対応)

5 結果

抽出された 41 のトピックのうち、38 個を命名することができた。命名されたトピックのうち、研究手法と統計に関するトピックを除いたネットワークグラフを図 1 に示す。このグラフから 7 つのクラスターが検出 (図 1 の凡例を参照)、命名された。

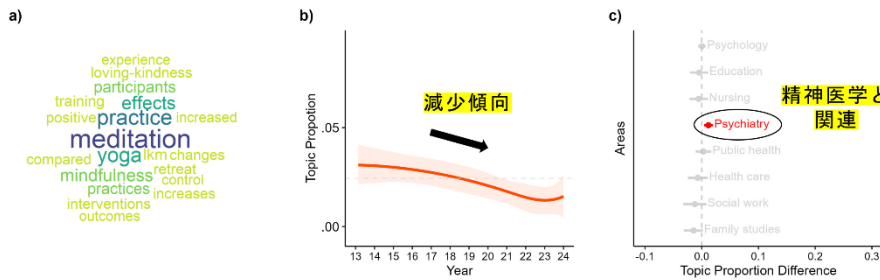
トピック割合と論文メタ情報 (出版年、掲載誌の学術分野) の関連に関しては、結果をすべて示すことは紙面の都合上できないため、トピックの推移が減少、増加している代表的なトピックとしてトピック #5 (瞑想とヨガ)、#24 (COVID-19) を図 2 に示す。

6 考察

本研究は、セルフ・コンパッションに関する 4,000 超の先行研究のスコopingレビューを目的とし、全論文のタイトルと抄録に対して STM を適用した。38 の命名可能なトピックが抽出され、それらは 7 つのクラスターに分類された。

各クラスターは、互いに正に相関するトピックから構成されている。クラスター 1 (様々な介入) は、多様なサンプル (トピック #1)、精神疾患を持つ人々 (#3)、医療関連職 (#4) 等のセルフ・コンパッションを高めることをめざす様々な介入に関するものであった。興味深いのは、東洋における伝統的

トピック#5 瞑想とヨガ



トピック#24 COVID-19

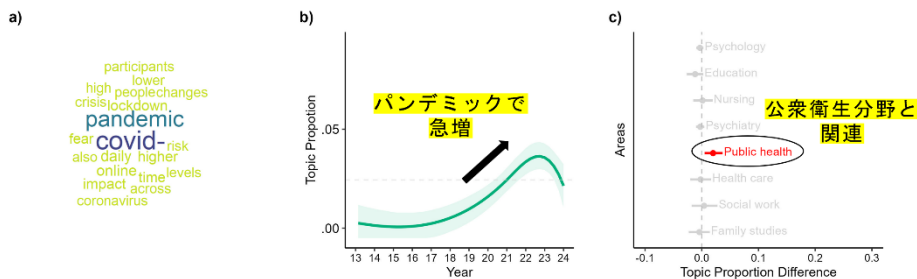


図 2 トピック#5 と #24 のワードクラウド (a)、トピック推移 (b)、掲載誌の学術分野 (c) との関連

実践である瞑想・ヨガに関連するトピック#5 が減少傾向にある点で (図 2)、セルフ・コンパッションを向上させる実践についての研究関心が瞑想やヨガ等の伝統的なものから、Mindful Self-Compassion Program[30]に代表されるようなパッケージ化された介入へと移行しつつあることを示唆している。

クラスタ 2 (苦痛と認知プロセス)、クラスタ 3 (食行動と身体イメージ)、クラスタ 5 (精神的健康と well-being) は、介入によるセルフ・コンパッションの涵養を通じて向上が目指される精神的/身体的健康、well-being と関連している。クラスタ 1 と同様、研究の対象は、一般サンプルのみならず、PTSD 患者 (#10)、身体疾患患者 (#11-12)、マイノリティ学生 (#25) など多岐にわたっていた。また、トピック#24 (COVID-19) は、2020 年に急激にトピック割合が上昇し、2023 年ごろ低下を始めている (図 2)。これらの時期は、それぞれ WHO による国際的パンデミックの発生宣言 (2020 年 1 月) と国際的に懸念される公衆衛生上の緊急事態の収束宣言 (2023 年 5 月) の時期と符合し、人々が社会的に孤立し不安が高まったパンデミック期において、心理学的な保護因子としてセルフ・コンパッションへの研究関心が向けられたことを示唆する。

クラスタ 4 (生育環境と子育て) は、ヒトの発達過程におけるセルフ・コンパッションの役割に焦点を当てている。幼少期に虐待等の不適切な養育を受

けたり愛着が不安定だった個人ほどその後のセルフ・コンパッションが低く、様々な精神病理のリスクが高まることが示されている[31], [32]。

クラスタ 6 (ケアギビング) は、看護師 (#27) や教師、その他の医療関連職 (#28) などの他者へのケアを行う職業者に関するトピックを中心に構成されている。これらの職業者は、しばしば共感疲労や燃え尽き症候群 (#31) といった問題を抱えやすい。先行研究は、セルフ・コンパッションがこのような問題に対する保護因子となることを示している[33], [34]。最後に、クラスター 7 (スポーツ) は、単一のトピック (#35) で構成されているクラスターであり、スポーツに関する研究は他のテーマとは比較的独立に研究されていることを示唆する。

おわりに

セルフ・コンパッションに関する大規模な先行研究文献を対象に STM を適用することで、これまでのセルフ・コンパッション研究に潜在している具体的なトピックを抽出することができた。STM による文献レビューは、自然科学分野では比較的多いものの、人文学分野においては非常に限られている。今後の研究では、心理学を含めた人文学の文献レビューにおける STM の使用が期待される。

参考文献

- [1] K. Neff, 'Self-Compassion: An Alternative Conceptualization of a Healthy Attitude Toward Oneself', *Self and Identity*, vol. 2, no. 2, 2003, doi: 10.1080/15298860309032.
- [2] K. Neff, 'Self-Compassion: Theory, Method, Research, and Intervention', *Annu Rev Psychol*, vol. 74, no. 1, pp. 193–218, Jan. 2023, doi: 10.1146/annurev-psych-032420-031047.
- [3] K. Bluth and K. D. Neff, 'New frontiers in understanding the benefits of self-compassion', *Self and Identity*, vol. 17, no. 6, pp. 605–608, Nov. 2018, doi: 10.1080/15298868.2018.1508494.
- [4] A. MacBeth and A. Gumley, 'Exploring compassion: A meta-analysis of the association between self-compassion and psychopathology', *Clin Psychol Rev*, vol. 32, no. 6, pp. 545–552, Aug. 2012, doi: 10.1016/j.cpr.2012.06.003.
- [5] W. J. Phillips and D. W. Hine, 'Self-compassion, physical health, and health behaviour: a meta-analysis', *Health Psychol Rev*, vol. 15, no. 1, 2021, doi: 10.1080/17437199.2019.1705872.
- [6] U. Zessin, O. Dickhäuser, and S. Garbade, 'The Relationship Between Self-Compassion and Well-Being: A Meta-Analysis', *Appl Psychol Health Well Being*, vol. 7, no. 3, 2015, doi: 10.1111/aphw.12051.
- [7] M. Ferrari, C. Hunt, A. Harrysunker, M. J. Abbott, A. P. Beath, and D. A. Einstein, 'Self-Compassion Interventions and Psychosocial Outcomes: a Meta-Analysis of RCTs', *Mindfulness (N Y)*, vol. 10, no. 8, pp. 1455–1473, 2019, doi: 10.1007/s12671-019-01134-6.
- [8] L. R. Tavares, P. Vagos, and A. Xavier, 'The role of self-compassion in the psychological (mal)adjustment of older adults: a scoping review', *Int Psychogeriatr*, vol. 35, no. 4, pp. 179–192, Apr. 2023, doi: 10.1017/S1041610220001222.
- [9] D. L. Cormier, K. C. Kowalski, L. J. Ferguson, A. D. Mosewich, T.-L. F. McHugh, and P. Röthlin, 'Self-compassion in sport: a scoping review', *Int Rev Sport Exerc Psychol*, pp. 1–40, Jan. 2023, doi: 10.1080/1750984X.2022.2161064.
- [10] C. R. Lathren, S. S. Rao, J. Park, and K. Bluth, 'Self-Compassion and Current Close Interpersonal Relationships: a Scoping Literature Review', *Mindfulness (N Y)*, vol. 12, no. 5, pp. 1078–1093, May 2021, doi: 10.1007/s12671-020-01566-5.
- [11] H. Arksey and L. O'Malley, 'Scoping studies: towards a methodological framework', *Int J Soc Res Methodol*, vol. 8, no. 1, pp. 19–32, Feb. 2005, doi: 10.1080/1364557032000119616.
- [12] Y. H. Kee, C. Li, L. C. Kong, C. J. Tang, and K.-L. Chuang, 'Scoping Review of Mindfulness Research: a Topic Modelling Approach', *Mindfulness (N Y)*, vol. 10, no. 8, pp. 1474–1488, Aug. 2019, doi: 10.1007/s12671-019-01136-4.
- [13] V. Swami, N. Andersen, and A. Furnham, 'A Bibliometric Review of Self-Compassion Research: Science Mapping the Literature, 1999 to 2020', *Mindfulness (N Y)*, vol. 12, no. 9, pp. 2117–2131, Sep. 2021, doi: 10.1007/s12671-021-01662-0.
- [14] M. E. Roberts, B. M. Stewart, and D. Tingley, 'stm: An R Package for Structural Topic Models', *J Stat Softw*, vol. 91, no. 2, 2019, doi: 10.18637/jss.v091.i02.
- [15] H. Suh, S. Kim, and D. Lee, 'Review of Perfectionism Research From 1990 to 2019 Utilizing a Text-Mining Approach', *Review of General Psychology*, vol. 25, no. 3, pp. 283–303, Sep. 2021.
- [16] S. Khanal and H. Zhang, 'Ten Years of China's Belt and Road Initiative: A Bibliometric Review', *J Chin Polit Sci*, vol. 29, no. 2, pp. 361–395, Jun. 2024, doi: 10.1007/s11366-023-09873-z.
- [17] X. Bai, X. Zhang, K. X. Li, Y. Zhou, and K. F. Yuen, 'Research topics and trends in the maritime transport: A structural topic model', *Transp Policy (Oxf)*, vol. 102, pp. 11–24, Mar. 2021, doi: 10.1016/j.tranpol.2020.12.013.
- [18] S. Liu, R.-Y. Zhang, and T. Kishimoto, 'Analysis and prospect of clinical psychology based on topic models: hot research topics and scientific trends in the latest decades', *Psychol Health Med*, vol. 26, no. 4, pp. 395–407, Apr. 2021, doi: 10.1080/13548506.2020.1738019.
- [19] A. Pisarevskaya, N. Levy, P. Scholten, and J. Jansen, 'Mapping migration studies: An empirical analysis of the coming of age of a research field', *Migr Stud*, vol. 8, no. 3, pp. 455–481, Sep. 2020, doi: 10.1093/migration/mnz031.
- [20] D. M. Blei, A. Y. Ng, and Jordan Michael I, 'Latent Dirichlet Allocation', *Journal of Machine Learning Research*, vol. 3, no. 4/5, pp. 993–1022, 2003.
- [21] R. Tamakloe and D. Park, 'Discovering latent topics and trends in autonomous vehicle-related research: A structural topic modelling approach', *Transp Policy (Oxf)*, vol. 139, pp. 1–20, Aug. 2023, doi: 10.1016/j.tranpol.2023.06.001.
- [22] D. C. Broadstock and X. Wang, 'District cooling services: A bibliometric review and topic classification of existing research', *Renewable and Sustainable Energy Reviews*, vol. 190, p. 113893, Feb. 2024, doi: 10.1016/j.rser.2023.113893.
- [23] B. Sheng, Z. Wang, Y. Qiao, S. Q. Xie, J. Tao, and C. Duan, 'Detecting latent topics and trends of digital twins in healthcare: A structural topic model-based systematic review', *Digit Health*, vol. 9, Jan. 2023, doi: 10.1177/20552076231203672.
- [24] A. A. Guenduez and N. Walker, 'A Literature Review of Covid-19 Research: Taking Stock and Moving Forward', *International Journal of Public Administration*, pp. 1–17, Apr. 2024, doi: 10.1080/01900692.2024.2332618.
- [25] B. Sheng, X. Chen, Y. Zhang, J. Tao, and Y. Sun, 'Structural topic model-based comparative review of human pose estimation research in the United States and China', *Multimed Tools Appl*, vol. 83, no. 20, pp. 59379–59412, Dec. 2023, doi: 10.1007/s11042-023-17923-0.
- [26] S. Kraus, S. Kumar, W. M. Lim, J. Kaur, A. Sharma, and F. Schiavone, 'From moon landing to metaverse: Tracing the evolution of Technological Forecasting and Social Change', *Technol Forecast Soc Change*, vol. 189, p. 122381, Apr. 2023, doi: 10.1016/j.techfore.2023.122381.
- [27] S. J. Weston, I. Shryock, R. Light, and P. A. Fisher, 'Selecting the Number and Labels of Topics in Topic Modeling: A Tutorial', *Adv Methods Pract Psychol Sci*, vol. 6, no. 2, 2023, doi: 10.1177/25152459231160105.
- [28] S. Ceh, F. Putze, and M. Benedek, 'D-I-WHAT? Identifying creative domains in do-it-yourself videos on YouTube.', *Psychol Aesthet Creat Arts*, Jul. 2023, doi: 10.1037/aca0000606.
- [29] M. Rosvall and C. T. Bergstrom, 'Maps of random walks on complex networks reveal community structure', *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, Jan. 2008, doi: 10.1073/pnas.0706851105.
- [30] C. Germer and K. Neff, *Teaching the Mindful Self-Compassion Program: A Guide for Professionals*. Guilford Press, 2019.
- [31] K. Garbutt, M. Rennoldson, and M. Gregson, 'Sexual Offending: Adverse Childhood Experiences, Shame, and Self-Compassion Explain the Variance in Self-Harm and Harm Towards Others?', *Sexual Abuse*, vol. 36, no. 6, pp. 662–691, Sep. 2024.
- [32] S. L. Bistricky, M. W. Gallagher, C. M. Roberts, L. Ferris, A. J. Gonzalez, and C. T. Wetterneck, 'Frequency of Interpersonal Trauma Types, Avoidant Attachment, Self-Compassion, and Interpersonal Competence: A Model of Persisting Posttraumatic Symptoms', *J Aggress Maltreat Trauma*, vol. 26, no. 6, pp. 608–625, Jul. 2017, doi: 10.1080/10926771.2017.1322657.
- [33] L.-C. Chu, 'Effect of compassion fatigue on emotional labor in female nurses: Moderating effect of self-compassion', *PLoS One*, vol. 19, no. 3, p. e0301101, Mar. 2024, doi: 10.1371/journal.pone.0301101.
- [34] K. Kabakçı, 'Examining the Relationship Between Nurses' Compassion and Professional Burnout', *J Psychiatr Nurs*, 2022, doi: 10.14744/phd.2022.29291.
- [35] M. Grootendorst, 'BERTopic: Neural topic modeling with a class-based TF-IDF procedure', Mar. 2022.