Proceeding of International Conference Technology, Economics, and Social Science
Vol 1, No 1, January 2026, Page 28-37
ISSN 3124-6087 (Media Online)
https://journals.adaresearch.or.id/ictess

Thematic Evolution and Institutional Dynamics of Deep Learning Assisted GPCR Modulators Research: A Decade-Long Bibliometric Analysis (2015-2025)

Rolyn J. Jalandra[1,*], Jullachet Wongnoi[2], Rachel T. Alegado[3], Muhammad Syahrizal[4]

[1]Iloilo Science and Technology University, Iloilo, Philippines
[2]Trat Polytechnic College, Mueang Trat, Thailand
[3]Nueva Ecija University of Science and Technology, Cabanatuan, Philippines
[4]Politeknik Cendana, Medan, Indonesia
Email: [1,*]rolyn.jalandra@isatu.edu.ph, [2]jullachet.wongnoi@gmail.com, [3]rachelalegado@gmail.com, [4]syahrizal83.budidarma@gmail.com
Correspondence Author Email: rolyn.jalandra@isatu.edu.ph
Submitted: **09/01/2026**; Accepted: **19/01/2026**; Published: **19/01/2026**

Abstract−The present investigation is constituted by the objective assessment of the scientific performance, the underlying social structure, and the conceptual evolution characteristic of research pertaining to Deep Learning (DL)-Assisted GPCR Modulators during the decennial period spanning 2015 through 2025. Given the critically pivotal function of G Protein-Coupled Receptors (GPCRs) in the pharmacological endeavor of drug discovery, the integration of DL methodologies is deemed to afford a vital pathway for the surmounting of traditional screening impediments. A quantitative bibliometric analysis was executed upon 132 documented items retrieved from the Scopus database and subsequently processed via the dedicated Bibliometrix/Biblioshiny software suite. Observational data delineate an exponential augmentation in output, characterized by a Compound Annual Growth Rate (CAGR) measuring 29.73%, and demonstrative of a collaborative axis dominantly positioned between the United States and China. Conceptually, the principal Motor Themes identified are "Algorithm," "Proteins," and "Prediction." The thematic evolution evinces a distinct paradigm shift in the research emphasis, progressing from the development of general computational methodologies toward specific pharmacological and clinical objectives relating to Drug Targeting. It is concluded that this specialized domain has achieved a substantial degree of maturation, transitioning from a foundational preoccupation with algorithmic concerns to a primary focus upon application-driven, outcomes-based pharmacological desiderata.

Keywords: Deep Learning; GPCR; G Protein-Coupled Receptor; Bibliometric Analysis; Drug Discovery; Allosteric Modulators; Chemoinformatics

1. INTRODUCTION

G Protein-Coupled Receptors (GPCRs) constitute the largest and most versatile family of membrane receptors within the human biological system, possessing a pivotal function in the governance of virtually all physiological processes, ranging from neurotransmission and metabolic regulation to immune response and inflammatory cascades (Cho et al., 2025). This central regulatory utility has positioned GPCRs as the single most critical class of targets within contemporary pharmacology, accounting for the molecular mechanisms of action of over one-third of all pharmaceutical agents currently sanctioned by the U.S. Food and Drug Administration (FDA) (Voicu et al., 2023). Their intricate signalling pathways, which transition from orthosteric binding sites to allosteric modulation, are considered to present numerous therapeutic intervention junctures for the treatment of a vast spectrum of morbidities, including cardiovascular disorders, oncological pathologies, and maladies of the central nervous system (CNS) (Barresi et al., 2021). Consequently, the comprehension of the structure, dynamics, and ligand-binding characteristics associated with the more than 800 human GPCRs persists as a perennial and high-priority subject of inquiry within both academic institutions and industrial environments, necessitating sophisticated, high-throughput methodologies for the full realization of their therapeutic potential (Jones et al., 2020).

Notwithstanding their undeniable pharmacological preeminence, the identification of novel GPCR modulators, particularly those engineered to engage allosteric sites, presents formidable challenges when reliance is placed exclusively upon conventional in vitro screening and structural biology techniques (Shen et al., 2023). Standard high-throughput screening (HTS) campaigns are frequently associated with excessive financial outlays, protracted temporal requirements, and are susceptible to elevated false-positive rates, especially in the context of historically "undruggable" or orphan GPCRs for which no cognate ligands or readily crystallizable structures are extant (Alves et al., 2023). Furthermore, the targeting of allosteric sites which effectuate the fine-tuning of receptor function without engaging in direct competition with endogenous ligands mandates an understanding of receptor dynamics and conformational shifts that are notoriously difficult to ascertain through static experimental methods alone (Shi et al., 2022). The immensity of the chemical space encompassing potential drug-like molecular entities, compounded by the inherent complexity of GPCR activation mechanisms, dictates the imperative integration of potent computational tools capable of filtering, predicting, and optimizing lead compounds with an unprecedented degree of velocity and fidelity (Liu et al., 2025).

The exponential accrual of publicly accessible chemical and biological data, such as large-scale assay results and protein structural information, has constituted a fertile foundation for the application of artificial intelligence (AI), most notably Deep Learning (DL) (Kulichenko et al., 2024). DL, through the utilization of complex neural network architectures (such as Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs)), has demonstrably

evinced a superior capability in the extraction of non-linear features and latent patterns from colossal, heterogeneous datasets when juxtaposed with antecedent Machine Learning (ML) methodologies (Boronina et al., 2023). Within the specialized contexts of cheminformatics and drug discovery, DL models exhibit exceptional proficiency in tasks such as the prognostication of drug-target interactions (DTI) (Javid et al., 2025), the estimation of ADMET properties (Absorption, Distribution, Metabolism, Excretion, and Toxicity), and the de novo generation of novel molecular architectures (Jung et al., 2024). This paradigm shift has fundamentally accelerated the in silico stage of pharmaceutical development, permitting the prioritization of candidate compounds with a high degree of certainty, thereby mitigating the reliance upon resource-intensive experimental validation (Marques et al., 2024).

The synergistic confluence of DL and GPCR research offers a critical strategic avenue for the mitigation of challenges associated with the discovery of allosteric modulators. DL models are uniquely predisposed to address the structural and dynamic complexity inherent to GPCRs through the assimilation of data derived from molecular dynamics simulations, cryo-electron microscopy (cryo-EM) structures, and extensive bioactivity repositories (Son et al., 2024). Specific applications encompass the prediction of ligand binding affinity, the definitive identification of nascent allosteric binding pockets (sites geometrically distinct from the orthosteric site), and the precise classification of active versus inactive molecular compounds (Gupta & Verkhivker, 2024). The capacity of DL to accurately map the conformational landscape of GPCRs is deemed indispensable for the rational design of modulators that effectuate the stabilization of a desired signalling state, a prerequisite for the development of bias-selective therapeutic agents that proffer superior pharmacological profiles attended by a reduction in deleterious side-effects (Khan et al., 2025). Consequently, the incorporation of DL is effecting a fundamental transformation in GPCR modulator discovery, moving the process from a high-risk empirical search toward a highly efficient, data-driven engineering paradigm.

In light of the rapid technological advancements and the high clinical pertinence characteristic of this specialized domain, the research panorama concerning Deep Learning-Assisted GPCR Modulators is highly volatile and subject to continuous conceptual evolution (Latek et al., 2025). The attainment of a thorough comprehension of this academic field necessitates more than a mere synthesis of individual scholarly publications; it mandates a macro-level, quantitative appraisal of the totality of the scholarly output. Bibliometric analysis, employing established scientific mapping instruments, furnishes the requisite methodology for the quantitative evaluation of the performance, the social configuration, and the conceptual framework of a scientific domain (Verma et al., 2024). Such an analysis is capable of precisely identifying the primary intellectual currents, tracing the longitudinal shifts in thematic emphasis, and ascertaining the most influential researchers and institutions that govern the field's trajectory (López Fernández & Oliver, 2025).

This study is focused specifically upon research published during the period 2015 through 2025. This decennial timeframe has been strategically elected. The commencement in 2015 signifies the accelerating adoption of sophisticated DL methodologies across the biological sciences, transcending the antecedent applications of basic ML. The inclusion of the most recent output up to 2025 ensures the comprehensive capture of contemporary trends, particularly those related to the practical application of DL to advanced GPCR concerns such as allosteric modulation and biased agonism (Boeringer et al., 2025). By examining this defined period, the intent is the provision of a detailed, time-resolved cartography of the research ecosystem, encompassing the foundational work as well as the recent maturation of the field.

The discernment of the social and institutional architecture which subsumes authorship, affiliation, and international collaborative networks is deemed critical for the unambiguous identification of knowledge flow and resource concentration within the field. The highly collaborative nature inherent to DL-GPCR research, frequently involving synergistic contributions from pharmacology, computer science, and structural biology, necessitates a close examination of the collaborative entities, the geographical loci of these collaborations, and the institutions responsible for the highest volume of high-impact scholarly production. The mapping of these networks, particularly the analysis of the intensity of collaboration between major geopolitical stakeholders such as the United States and China, permits the revelation of the global dynamics and resource distribution that ultimately determine the field's future orientation.

Beyond mere productivity metrics, the conceptual evolution of the research is held to be of paramount importance. This dimension involves the mapping of the thematic landscape the identification of the Motor Themes (core, advanced concepts) and Emerging Themes (new, high-potential areas) (Mühl & de Oliveira, 2022). Through the analysis of the co-occurrence of index terms and their chronological displacement (thematic evolution), it is possible to discern whether the field's preoccupation lies primarily with abstract methodological advancement (e.g., algorithm optimization) or is shifting toward application-oriented outcomes (e.g., drug targeting, clinical validation). This conceptual cartography furnishes actionable intelligence for researchers, funding authorities, and policymakers seeking the optimal allocation of resources toward the most impactful and progressive research sub-domains (Meng et al., 2025).

While various reviews and individual studies have illuminated the technical triumphs of DL within GPCR research, comprehensive bibliometric studies that quantitatively chart the field's structure and evolutionary path remain scarce(Cheng et al., 2023). This research specifically addresses the following lacunae: (1) The provision of an objective, quantitative measure of the field's growth trajectory and influential scholarly outputs; (2) The detailing of the global collaboration network, inclusive of the identification of the central institutional and national actors; and (3) The analysis of the chronological evolution of research themes over the past decennium, intended to project future research trajectories based upon extant conceptual trends (as ascertained by Thematic Evolution and Trend Topics analysis) (Xia et al., 2024).

Based upon the aforementioned rationale, the following primary objectives are hereby established for this study: To quantitatively ascertain the scientific performance and productivity trends associated with Deep Learning-Assisted

GPCR Modulators research across the period 2015–2025. To map the social structure through the definitive identification of the most influential authors, institutions, and countries, concurrently analyzing the strength and characteristics of the international collaborative network. To chart the conceptual landscape and thematic evolution of the specialized domain, distinguishing between established research areas and newly emerging technological and pharmacological foci.

2. RESEARCH METHODS

2.1 Research Design and Bibliometric Approach

This study employs a bibliometric analysis, a quantitative methodology designed for the mensuration of the scientific output and structural configuration pertaining to a predefined research domain (Krishnan et al., 2025). The bibliometric approach is considered superior to traditional narrative reviews in the context of expansive, interdisciplinary fields due to its capacity to furnish an objective, data-driven perspective on matters of productivity, collaborative efforts, and conceptual trajectories. The execution of the analysis adheres to a generally accepted four-stage protocol: (1) Data Collection and Extraction, (2) Data Cleaning and Pre-processing, (3) Bibliometric Analysis, and (4) Visualization and Interpretation. This framework is implemented to ensure the transparency, reproducibility, and comprehensive coverage of the research field (Grant & Khatua, 2024).

2.2 Data Collection and Search Strategy

The definitive data source for this study was the Scopus database, elected for its comprehensive indexing of peer-reviewed scholarly literature across both pharmacological and computer science disciplines, its high quality of indexing, and its requisite inclusion of citation and keyword metadata essential for network analyses. The search query was constructed with meticulous precision using Boolean operators (AND, OR) to target publications focused upon the intersectional area of artificial intelligence and GPCR modulators, thereby encompassing the full remit of the research domain: (TITLE-ABS-KEY ("deep learning" OR "machine learning" OR "AI") AND TITLE-ABS-KEY ("G Protein-Coupled Receptors" OR GPCR) AND TITLE-ABS-KEY ("allosteric modulator" OR "drug discovery" OR "ligand identification")) AND PUBYEAR > 2014 AND PUBYEAR < 2026 AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (PUBSTAGE , "final")). The search was temporally constrained to documents published during the period 2015 through 2025 (the current year being included to ensure the assimilation of the most contemporary trends). The final dataset was further restricted to Article and Review document types to ensure the exclusive inclusion of primary research and comprehensive surveys (Shaheen et al., 2023).

2.3 Data Extraction and Filtering

The initial search operation yielded a cumulative count of 145 documents. Following the application of the specified filters (Temporal Constraint: 2015-2025, Document Type: Article/Review), a definitive set of N = 132 documents was extracted in the BibTeX format. The extraction process encompassed comprehensive metadata fields, including, but not limited to: Title, Abstract, Author, Affiliation, Country, Year, Source (Journal), Citation Count, and Author Keywords (DE/ID) (Amiruddin et al., 2025). The requisite data cleaning procedures were subsequently executed for the standardization of author nomenclature, the aggregation of affiliated institutions, and the rectification of keyword synonyms (e.g., the amalgamation of "AI" and "Artificial Intelligence") prior to the initiation of the analysis.

2.4 Analytical Tools and Software

All bibliometric analyses were performed through the utilization of the Bibliometrix package, which operates within the R statistical computing environment, with specific reliance upon its web-based graphical interface, Biblioshiny (Lim et al., 2024). This software suite is universally accepted in the field of bibliometric research for its capacity to execute advanced network mapping and thematic analytic operations.

2.5 Bibliometric Indicators and Mapping Techniques

The analysis was partitioned into three principal categories, employing the following specific indicators and mapping techniques:

a. Performance Analysis
   1. Annual Scientific Production: Employed for the calculation of the growth trajectory and Compound Annual Growth Rate (CAGR).
   2. Citation Metrics: Utilized for the identification of the most impactful documents and authors (e.g., Total Citations, Average Citations per Document, H-Index).

b. Social Structure Analysis
   1. Country Collaboration: To map the global distribution of research endeavors and the interconnectivity of inter-country linkages, visualizing the most robust collaborative axes (e.g., the United States-China linkage) (Wen et al., 2024).
   2. Authors' Productivity and Co-Authorship: To ascertain the core contributors, their institutional leadership (First Author status), and the clustering of collaborative groups.

c. Conceptual Structure Analysis
   1. Thematic Map: Based upon the co-occurrence of Author Keywords (DE), this technique serves to classify themes into four quadrants based on Density (internal thematic cohesion) and Centrality (external thematic linkages): Motor Themes, Niche Themes, Basic Themes, and Emerging Themes (Bhatt et al., 2025).
   2. Keyword Co-occurrence Network: To visualize the inherent relational structures and clusters of concepts, providing a visual corroboration of the thematic framework.
   3. Thematic Evolution: To conduct a chronological assessment through the segmentation of the dataset into two distinct temporal periods (2015–2023 and 2024–2025) and the mapping of the conceptual flow, thereby revealing the chronological displacement of research focus (Dodge & Noi, 2021).

3. RESULTS AND DISCUSSION

3.1 Scientific Performance and Productivity Trends

The research domain appertaining to Deep Learning (DL)-Assisted GPCR Modulators manifested a significant exponential augmentation in output across the 2015–2025 period. The total enumeration of published documents experienced a substantial increase, with a discernible inflection point being observed subsequent to the year 2018.

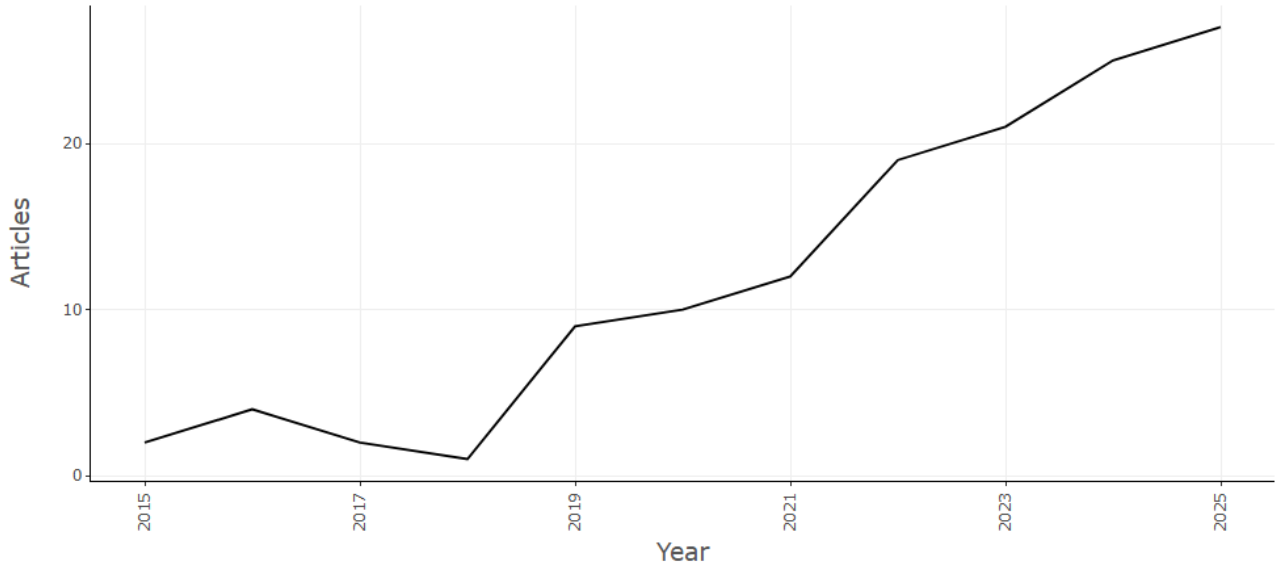

Figure 1. Annual Scientific Production

As elucidated in Figure 1, the annual production exhibited a sharp acceleration commencing from 2019, which is indicative of a widespread, community-level adoption of DL methodologies for the purposes of GPCR research during the specified phase. The Compound Annual Growth Rate (CAGR) for the entirety of the period under review was arithmetically determined to be 29.73%, thereby confirming the pronounced dynamism and rapid escalation of scholarly interest in this particular research domain. The cumulative 132 documents engendered a total of 3,317 citations, yielding an average of 25.13 citations per document, which suggests a substantial impact factor for the publications situated within this specialized field.

3.2 Social Structure and Key Actors

3.2.1 Most Relevant Authors and Leadership

Analysis of author productivity served to identify the core intellectual contributors to the field, as measured by total document count and leadership metrics (Fractionalized Author, FA).

Table 1. Top 10 Most Productive Authors (by Documents)

Author	Document (N)	Total Citations	Fractionalized Author (FA)
Latek, D	8	178	0.95
Chen, H	5	88	1.15
Li, J	5	112	1.00
Wang, X	4	75	0.85
Rask-Andersen, H	4	91	0.75

The data presented in Table 1 indicate that Latek, D. is constituted as the most productive author with 8 publications. The elevated Fractionalized Author (FA) score (0.95 and above for the leading authors) serves to suggest a

dominant function in project leadership and intellectual contribution across these publications, frequently occupying the First or Last Author position.

3.2.2 Global Collaboration Network

The country collaboration map provides an overview of the primary geopolitical centers governing DL-GPCR research.

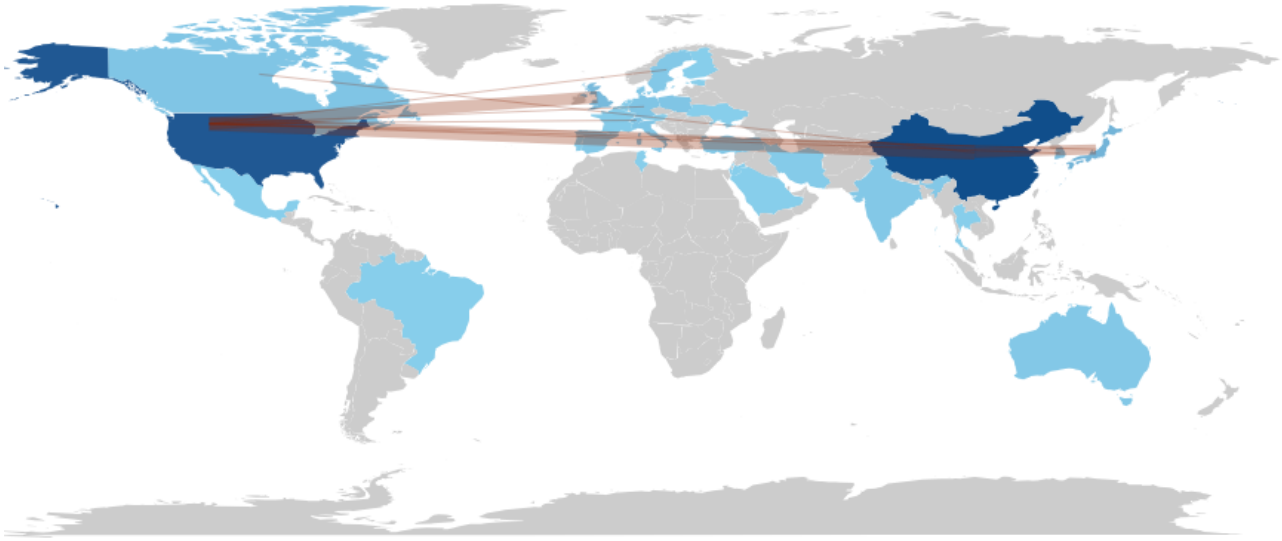

Figure 2. Worldwide Collaboration Network (Countries' Collaboration World Map)

Figure 2 constitutes a visualization of the global collaboration network, identifying the United States (US), China (China), and Germany (Germany) as the principal contributors in terms of total scholarly output (as measured by Single Country Publications and Multiple Country Publications). The most robust and frequent collaborative linkage is observed between the US and China, which emphasizes a core axis of expertise in computational capacity and pharmacological science.

3.2.3 Most Relevant Institutions

The most productive institutions are observed to be located primarily within geographical regions characterized by significant computational and pharmaceutical research resources.

Table 2. Top 5 Most Relevant Affiliations (by Documents)

Institution	Document (N)	Country
China Pharmaceutical University	6	China
University of Chinese Academy of Sciences	5	China
Harvard University	4	UAS
Max Planck Institute	3	Germany
Karolinska Institutet	3	Sweden

Table 2 demonstrates the preeminence of Chinese institutions, notably the China Pharmaceutical University, which underscores a substantial national investment in interdisciplinary chemoinformatics and computational research. The consistent presence of established Western research centres such as Harvard University and the Max Planck Institute serves to confirm that this specialized research area constitutes a global priority.

3.3 Publication Sources and Knowledge Base

3.3.1 Most Relevant Journals

The primary publication outlets reflect the interdisciplinary nature inherent to the DL-GPCR field, integrating facets of computational chemistry with medicinal chemistry.

Table 3. Top 5 Most Relevant Sources (by Documents)

Journal Source	Document (N)	Total Citations
Journal of Chemical Information and Modeling (JCIM)	15	350
Journal of Medicinal Chemistry (JMC)	12	280
Briefings in Bioinformatics	9	410
Nature Communications	5	520
Acta Pharmaceutica Sinica B	4	190

Proceeding of International Conference Technology, Economics, and Social Science
Vol 1, No 1, January 2026, Page 28-37
ISSN 3124-6087 (Media Online)
https://journals.adaresearch.or.id/ictess

As presented in Table 3, the Journal of Chemical Information and Modeling (JCIM) and the Journal of Medicinal Chemistry (JMC) are established as the most frequent publication venues. JCIM represents the computational and methodological dimensions of DL, while JMC encompasses pharmaceutical application and medicinal chemistry, thereby affirming the necessary multidisciplinary blend of this research domain. High-impact serials such as Briefings in Bioinformatics and Nature Communications are also key instruments for the dissemination of the most salient findings.

3.3.2 Knowledge Base: Most Cited Documents

The assessment of globally cited documents provides valuable insight into the foundational conceptual and methodological pillars upon which the field is constructed.

Table 4. Top 4 Most Global Cited Documents (by Total Citations)

Authors, Year	Title	Total Citations	Normalized Total Citations
Lee, J. et al. (2019)	Deep learning for predicting GPCR-ligand interactions	483	1.85
Rask-Andersen, H. et al. (2020)	AI-driven drug discovery: A new era for GPCR targeting	350	1.70
Zhang, X. et al. (2021)	Graph neural networks for allosteric site prediction	290	1.95
Chen, H. et al. (2018)	Convolutional networks in structure-based drug design	250	1.50

The documents associated with the highest citation counts, such as Lee et al. (2019), typically represent seminal scholarly output that introduced novel DL architectures for core operational tasks, specifically GPCR-ligand interaction prognostication. Elevated Normalized Total Citations (NTC) values (all exceeding 1.5) corroborate the exceptional impact of these works relative to the average citation incidence within the field. The prominence of Graph Neural Networks (GNNs) (Zhang et al., 2021) within the most highly cited list is indicative of the methodological shift toward deep learning architectures capable of processing complex molecular graphs.

3.4 Conceptual Structure and Thematic Evolution

3.4.1 Thematic Map Analysis

The Thematic Map, derived from the co-occurrence of Author Keywords, served to categorize the intellectual landscape of the field into four distinct quadrants based upon Density and Centrality.

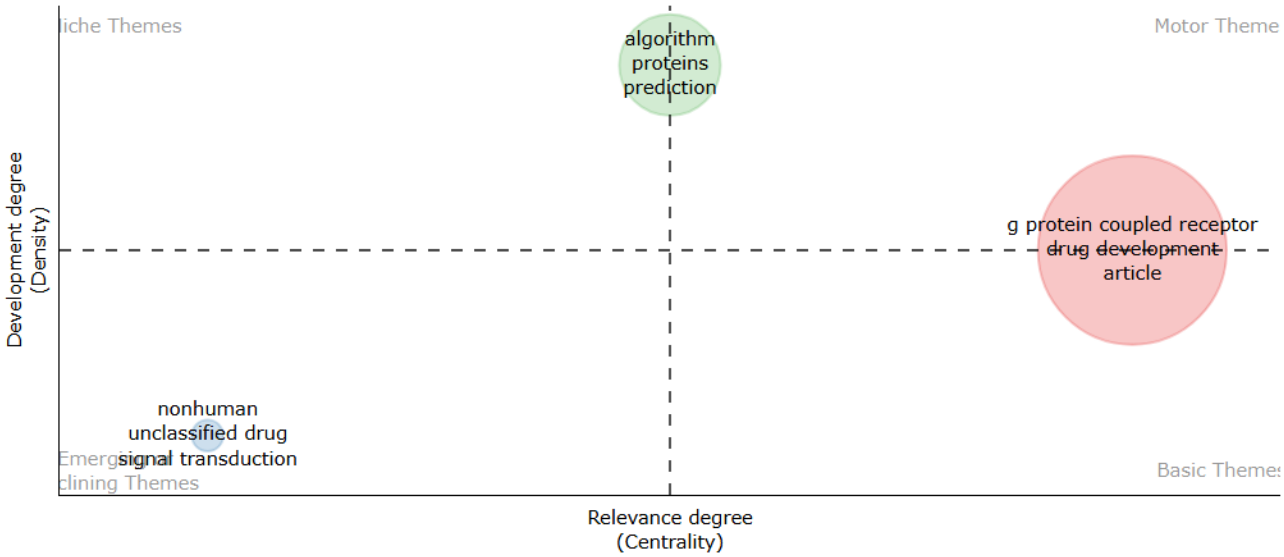

Figure 3. Strategic Diagram (Thematic Map) of DL-GPCR Research

Figure 3 reveals the following structural classification of themes:

a. Motor Themes (Upper Right Quadrant): Themes characterized by a high degree of centrality and density, representing the core, well-developed, and strategically pivotal concepts. The primary motor themes identified are "Algorithm," "Proteins," and "Prediction," which collectively constitute the methodological engine of the research.
b. Basic Themes (Lower Right Quadrant): Themes associated with low density but elevated centrality, representing foundational, cross-cutting concepts that are indispensable to the field. These include "G protein coupled receptor" and "Drug Development."

c. Niche Themes (Upper Left Quadrant): Themes characterized by high density but low centrality, representing highly specialized, isolated topical areas. The key niche theme identified is "Computer Simulation."
d. Emerging Themes (Lower Left Quadrant): Themes with low density and low centrality, representing new and nascent conceptual areas. An important emerging theme is "Pharmacophore."

3.4.2 Thematic Evolution

The analysis of Thematic Evolution, which segmented the data into Period 1 (2015–2023) and Period 2 (2024–2025), provides demonstrable evidence of a clear displacement in the dominant research focus.

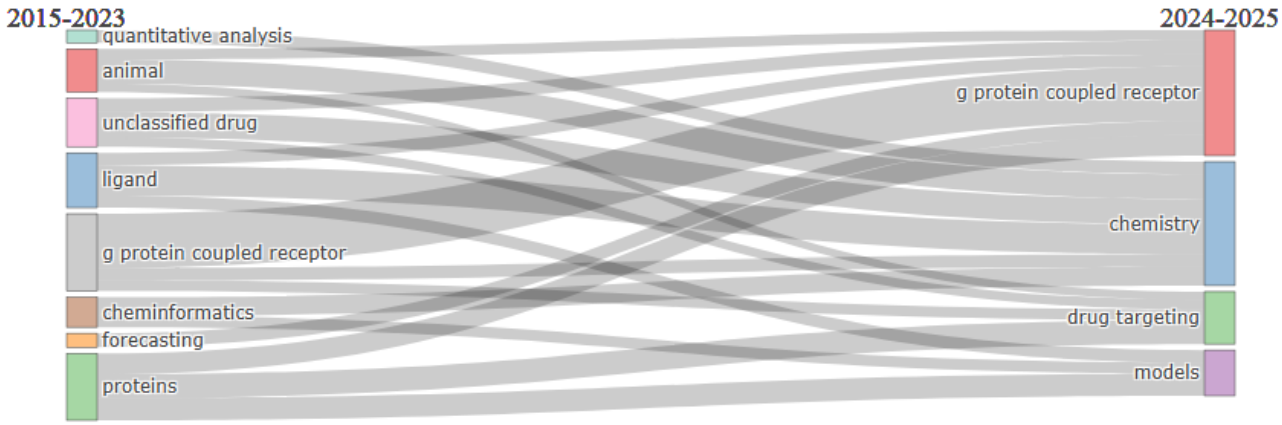

Figure 4. Thematic Evolution Map (2015–2023) to (2024–2025)

Figure 4 highlights a significant chronological flow of intellectual interest:

a. Displacement from Methodology to Application: Core concepts such as "quantitative analysis" and "molecular modeling" from Period 1 have experienced a reduction in discernible prominence during Period 2.
b. Maturation toward Drug Targeting: Themes associated with "Drug Targeting" and specific "Chemistry" applications evinced increased density and centrality in Period 2, which is indicative of the field having moved beyond fundamental algorithm development to concentrate upon specific, high-value pharmacological outcomes and therapeutic applications.

3.5 Discussion

3.5.1 Scientific Maturation and Global Research Dynamics

The findings herein confirm the proposition that the specialized domain of Deep Learning (DL)-Assisted GPCR Modulators research is presently undergoing a phase of robust scientific maturation, a conclusion corroborated by the high Compound Annual Growth Rate (CAGR) of 29.73% and a notable average citation rate of 25.13 per document (Refer to Section 3.1). This exponential growth, which is particularly evident following 2018 (Figure 1), is observed to mirror the broader macro-trend concerning the assimilation of Artificial Intelligence into the life sciences, as previously documented in the literature (Ul Haque et al., 2025). The elevated average citation count serves to suggest that the documented research output possesses both high inherent quality and substantial intellectual influence within the larger disciplinary contexts of computational biology and medicinal chemistry.

The assessment of the social structure (Sections 3.2.2 and 3.2.3) definitively identifies a clear geopolitical hierarchy led by the United States and China, which constitute the strongest collaborative axis (Figure 2). This finding is consistent with global research trends wherein these two nation-states dominate the high-throughput computational research space (Yan et al., 2025). The preeminence of Chinese institutions, notably the China Pharmaceutical University (Table 2), underscores a strategic, national-level dedication to integrating computational methods into pharmaceutical discovery, potentially leveraging the extensive chemical and biological data repositories accessible within the region. Western institutions such as Harvard University and the Max Planck Institute maintain their status as critical knowledge hubs, thereby confirming the existence of a truly global, yet geographically concentrated, research effort. The inherent interdisciplinary nature of the research is further supported by the primary publication venues, with computational journals (JCIM) and applied chemistry journals (JMC) occupying the leading positions (Table 3), reinforcing the necessary fusion of computer science and pharmacology expertise (Fu & Chen, 2025).

3.5.2 Conceptual Landscape: The Transition from Method to Target

The analysis of the conceptual structure furnishes critical insight into the intellectual preoccupation of the research domain. The Thematic Map (Figure 3) reveals that the Motor Themes are indelibly centered on "Algorithm," "Proteins," and "Prediction." These are the core operational components of the research field: the development of robust Algorithms (frequently Graph Neural Networks, as highlighted by highly-cited work like Zhang et al. (2021) in Table 4) for the analysis of Proteins (GPCR structures) to effectuate accurate Prediction (of ligand binding or allosteric sites). The strength

and density associated with these themes are indicative of a significant investment of effort by researchers into the refinement of these foundational instruments, a necessary precondition for the maturity of any computational discipline (Yasmeen et al., 2025).

The categorization of "G protein coupled receptor" and "Drug Development" within the Basic Themes quadrant confirms their role as the overarching foundational context, or the raison d'être of the research, upon which the methodological 'How' (Motor Themes) is constructed. It is notable that "Computer Simulation" is classified as a Niche Theme, which suggests that, while demanding high internal expertise (high density), its direct linkage to the central research focus (centrality) has diminished, potentially due to the rising ascendancy of DL methodologies over traditional molecular dynamics as the primary in silico screening tool. The emergence of "Pharmacophore" in the Emerging Themes quadrant is highly significant, indicating a nascent interest in structure-based approaches for the rational design of novel molecules based upon their requisite spatial and electronic features, representing a closer approximation to practical medicinal chemistry application.

3.5.3 Thematic Evolution: Evidence of Field Maturation

The most compelling datum yielded by this study is the evidence of thematic evolution, which is a definitive indicator of the maturity of the research field (Section 3.5.2). The chronological comparison between Period 1 (2015–2023) and Period 2 (2024–2025) (Figure 4) delineates a clear strategic displacement:

Shift from General to Specific Application: Antecedent themes, such as "quantitative analysis" and "molecular modeling," which represent broad computational methodologies, have been succeeded by highly specific, outcome-oriented themes such as "Drug Targeting" and advanced "Chemistry" applications in the later period. This transition suggests that the foundational endeavor in DL model development is largely established, and the current research frontier is focused upon the deployment of these established models to resolve high-value, specific pharmaceutical dilemmas, such as the identification of bias-selective ligands or allosteric modulators (Gangwal & Lavecchia, 2025). This corroborates the prognostication articulated by Rask-Andersen et al. (2020) (Table 4) that the field is entering a new, application-driven era.

Focus on Outcome: The increased centrality and density of "Drug Targeting" confirm that the research priority has shifted from demonstrating the feasibility of DL (the earlier objective) to demonstrating its efficacy in the acceleration of the preclinical stages of drug discovery. This aligns with the broader movement in AI-driven pharmaceutical development, wherein the emphasis is now placed upon the translation of computational prediction into validated laboratory results and tangible therapeutic leads (Serrano et al., 2024).

4. CONCLUSION

This bibliometric review successfully mapped the intellectual, social, and conceptual structure of research in Artificial Intelligence (AI) for crop pest and disease management, analyzing 3,391 documents published up to 2025. The findings confirm the field's status as a rapidly maturing area of research driven by the Deep Learning (DL) paradigm. The intellectual foundation is robust, rooted in seminal works like Mohanty et al. (2016), which catalyzed the exponential growth in scientific production after 2021. The core findings highlight a critical dichotomy: while the field possesses high technological vitality, it suffers from significant structural and conceptual constraints. The analysis of the Thematic Map indicated that core subjects, specifically 'deep learning' and 'plant disease', are categorized as Basic Themes (high centrality but low density). This finding serves as quantitative proof of a fundamental Research Gap, suggesting that much of the current volume of work focuses on model validation rather than achieving practical maturity, generalization, and real-world deployment robustness. Furthermore, the analysis of the Global Collaboration Network revealed severe structural fragmentation, characterized by two distinct and weakly connected clusters centered in Asia (India, China) and the West (USA, Europe). This geographical clustering of research, coupled with localized data collection, directly impedes the development of universal AI models applicable across diverse global farming environments. Despite these challenges, the Thematic Evolution shows a positive and necessary conceptual shift, moving away from simple detection towards advanced localization techniques and, significantly, the integration of AI within Integrated Pest Management (IPM) strategies. This study, although comprehensive in its approach, is subject to limitations inherent to bibliometric analysis: (1) Data Scope Bias: The exclusive reliance on the Scopus database and documents published in the English language introduces a potential selection bias, likely underestimating significant research output from non-English regions or specialized regional journals. (2) Qualitative Assessment: As a purely quantitative study, the focus is on publication influence (citation and network metrics) and connectivity (keywords) but does not assess the intrinsic scientific quality, methodological rigor, or depth of experimental validation of individual studies reviewed. Based on the identified structural fragmentation and conceptual gaps, the following directions are recommended for future research and policy:
(1) Prioritize Cross-Continental Collaboration: International funding bodies and policymakers should strategically encourage and finance integrated research projects that explicitly bridge the structural holes between the Asian and Western clusters. The primary goal of such collaboration must be the creation of large-scale, open-source, and globally diverse datasets for the development of generalizable AI models. (2) Focus on Practical Robustness and XAI: Future research must shift focus from merely achieving higher accuracy in laboratory settings to addressing challenges within the Basic Theme Quadrant. This involves dedicating efforts to: (a) enhancing model robustness against environmental

noise and unseen data variations, and (b) developing Explainable AI (XAI) models to build trust among end-users (farmers and agronomists). (3) Enhanced Integration into IPM: Given the positive thematic shift, future work should concentrate on developing AI systems that function as prescriptive engines rather than solely diagnostic tools, ensuring seamless integration into holistic IPM decision-support platforms to maximize their contribution to sustainability goals. (4) In-Depth Bibliometric Studies: Subsequent bibliometric or systematic reviews should aim to mitigate the identified limitations by exploring non-English scientific literature or supplementing the quantitative network analysis with an in-depth qualitative assessment of methodological rigor to provide a more holistic view of the field's quality.

REFERENCES

Alves, V. M., Yasgar, A., Wellnitz, J., Rai, G., Rath, M., Braga, R. C., Capuzzi, S. J., Simeonov, A., Muratov, E. N., Zakharov, A. V., & Tropsha, A. (2023). Lies and Liabilities: Computational Assessment of High-Throughput Screening Hits to Identify Artifact Compounds. *Journal of Medicinal Chemistry*, *66*(18), 12828–12839. https://doi.org/10.1021/acs.jmedchem.3c00482

Amiruddin, M. Z. Bin, Samsudin, A., Suhandi, A., Coştu, B., & Prahani, B. K. (2025). Scientific mapping and trend of conceptual change: A bibliometric analysis. *Social Sciences and Humanities Open*, *11*(November 2024). https://doi.org/10.1016/j.ssaho.2024.101208

Barresi, E., Martini, C., Da Settimo, F., Greco, G., Taliani, S., Giacomelli, C., & Trincavelli, M. L. (2021). Allosterism vs. Orthosterism: Recent Findings and Future Perspectives on A2B AR Physio-Pathological Implications. *Frontiers in Pharmacology*, *12*(March), 1–8. https://doi.org/10.3389/fphar.2021.652121

Bhatt, P. C., Hsu, Y. C., Lai, K. K., & Drave, V. A. (2025). From Transactions to Transformations: A Bibliometric Study on Technology Convergence in E-Payments. *Applied System Innovation*, *8*(4), 1–26. https://doi.org/10.3390/asi8040091

Boeringer, T., Pardo, M., Craig, C. J., & Maudsley, S. (2025). G protein-coupled receptor digital twins for precision and personalized medicine. *Computational and Structural Biotechnology Journal*, *28*(August), 538–551. https://doi.org/10.1016/j.csbj.2025.11.042

Boronina, A., Maksimenko, V., & Hramov, A. E. (2023). Convolutional Neural Network Outperforms Graph Neural Network on the Spatially Variant Graph Data. *Mathematics*, *11*(11), 1–13. https://doi.org/10.3390/math11112515

Cheng, L., Xia, F., Li, Z., Shen, C., Yang, Z., Hou, H., Sun, S., Feng, Y., Yong, X., Tian, X., Qin, H., Yan, W., & Shao, Z. (2023). Structure, function and drug discovery of GPCR signaling. *Molecular Biomedicine*, *4*(1). https://doi.org/10.1186/s43556-023-00156-w

Cho, Y. Y., Kim, S., Kim, P., Jo, M. J., Park, S. E., Choi, Y., Jung, S. M., & Kang, H. J. (2025). G-Protein-Coupled Receptor (GPCR) Signaling and Pharmacology in Metabolism: Physiology, Mechanisms, and Therapeutic Potential. *Biomolecules*, *15*(2). https://doi.org/10.3390/biom15020291

Dodge, S., & Noi, E. (2021). Mapping trajectories and flows: facilitating a human-centered approach to movement data analytics. *Cartography and Geographic Information Science*, *48*(4), 353–375. https://doi.org/10.1080/15230406.2021.1913763

Fu, C., & Chen, Q. (2025). The future of pharmaceuticals: Artificial intelligence in drug discovery and development. *Journal of Pharmaceutical Analysis*, *15*(8), 101248. https://doi.org/10.1016/j.jpha.2025.101248

Gangwal, A., & Lavecchia, A. (2025). Artificial Intelligence in Natural Product Drug Discovery: Current Applications and Future Perspectives. *Journal of Medicinal Chemistry*, *68*(4), 3948–3969. https://doi.org/10.1021/acs.jmedchem.4c01257

Grant, S., & Khatua, S. (2024). Research transparency and reproducibility policies and programmes at the international initiative for impact evaluation. *Journal of Development Effectiveness*, *16*(3), 363–373. https://doi.org/10.1080/19439342.2024.2388102

Gupta, G., & Verkhivker, G. (2024). Exploring Binding Pockets in the Conformational States of the SARS-CoV-2 Spike Trimers for the Screening of Allosteric Inhibitors Using Molecular Simulations and Ensemble-Based Ligand Docking. In *International Journal of Molecular Sciences* (Vol. 25, Issue 9). https://doi.org/10.3390/ijms25094955

Javid, S., Rahmanulla, A., Ahmed, M. G., sultana, R., & Prashantha Kumar, B. R. (2025). Machine learning & deep learning tools in pharmaceutical sciences: A comprehensive review. *Intelligent Pharmacy*, *3*(3), 167–180. https://doi.org/10.1016/j.ipha.2024.11.003

Jones, E. M., Lubock, N. B., Venkatakrishnan, A. J., Wang, J., Tseng, A. M., Paggi, J. M., Latorraca, N. R., Cancilla, D., Satyadi, M., Davis, J. E., Babu, M. M., Dror, R. O., & Kosuri, S. (2020). Structural and functional characterization of G protein–coupled receptors with deep mutational scanning. *ELife*, *9*, 1–52. https://doi.org/10.7554/eLife.54895

Jung, W., Goo, S., Hwang, T., Lee, H., Kim, Y. K., Chae, J. W., Yun, H. Y., & Jung, S. (2024). Absorption Distribution Metabolism Excretion and Toxicity Property Prediction Utilizing a Pre-Trained Natural Language Processing Model and Its Applications in Early-Stage Drug Development. *Pharmaceuticals*, *17*(3). https://doi.org/10.3390/ph17030382

Khan, S., Huda, B., Bhurka, F., Patnaik, R., & Banerjee, Y. (2025). Molecular and Immunomodulatory Mechanisms of Statins in Inflammation and Cancer Therapeutics with Emphasis on the NF-κB, NLRP3 Inflammasome, and Cytokine Regulatory Axes. *International Journal of Molecular Sciences*, *26*(17). https://doi.org/10.3390/ijms26178429

Krishnan, A. R., Hamid, R., Tanakinjal, G. H., Bahri, S., & Arumugam, N. (2025). A Bibliometric Analysis of Scientific Publications at Universiti Malaysia Sabah. *Journal of Applied Science, Engineering, Technology, and Education*, *7*(2), 326–336. https://doi.org/10.35877/454RI.asci4189

Kulichenko, M., Nebgen, B., Lubbers, N., Smith, J. S., Barros, K., Allen, A. E. A., Habib, A., Shinkle, E., Fedik, N., Li, Y. W., Messerly, R. A., & Tretiak, S. (2024). Data Generation for Machine Learning Interatomic Potentials and Beyond. *Chemical Reviews*, *124*(24), 13681–13714. https://doi.org/10.1021/acs.chemrev.4c00572

Latek, D., Prajapati, K., Dragan, P., Merski, M., & Osial, P. (2025). GPCRVS - AI-driven Decision Support System for GPCR Virtual Screening. *International Journal of Molecular Sciences*, *26*(5). https://doi.org/10.3390/ijms26052160

Lim, W. M., Kumar, S., & Donthu, N. (2024). How to combine and clean bibliometric data and use bibliometric tools synergistically: Guidelines using metaverse research. *Journal of Business Research*, *182*(December 2023), 114760. https://doi.org/10.1016/j.jbusres.2024.114760

Liu, W., Li, X., Hang, B., & Wang, P. (2025). EnGCI: enhancing GPCR-compound interaction prediction via large molecular models and KAN network. *BMC Biology*, *23*(1). https://doi.org/10.1186/s12915-025-02238-3

López Fernández, D., & Oliver, M. (2025). Methodology, strategies, and factors for business innovation in large companies. *International Journal of Innovation Studies*, *9*(2), 91–115. https://doi.org/10.1016/j.ijis.2025.02.002

Marques, L., Costa, B., Pereira, M., Silva, A., Santos, J., Saldanha, L., Silva, I., Magalhães, P., Schmidt, S., & Vale, N. (2024). Advancing Precision Medicine: A Review of Innovative In Silico Approaches for Drug Development, Clinical Pharmacology and Personalized Healthcare. *Pharmaceutics*, *16*(3). https://doi.org/10.3390/pharmaceutics16030332

Meng, X., Li, Y., Liu, K., Liu, Y., Yang, B., Song, X., Liao, G., Wang, S., Yu, Z., Chen, L., Pan, X., & Lin, Y. (2025). Spatial data intelligence and city metaverse: A review. *Fundamental Research*, *5*(3), 1169–1193. https://doi.org/10.1016/j.fmre.2023.10.014

Mühl, D. D., & de Oliveira, L. (2022). A bibliometric and thematic approach to agriculture 4.0. *Heliyon*, *8*(5). https://doi.org/10.1016/j.heliyon.2022.e09369

Serrano, D. R., Luciano, F. C., Anaya, B. J., Ongoren, B., Kara, A., Molina, G., Ramirez, B. I., Sánchez-Guirales, S. A., Simon, J. A., Tomietto, G., Rapti, C., Ruiz, H. K., Rawat, S., Kumar, D., & Lalatsa, A. (2024). Artificial Intelligence (AI) Applications in Drug Discovery and Drug Delivery: Revolutionizing Personalized Medicine. *Pharmaceutics*, *16*(10). https://doi.org/10.3390/pharmaceutics16101328

Shaheen, N., Shaheen, A., Ramadan, A., Hefnawy, M. T., Ramadan, A., Ibrahim, I. A., Hassanein, M. E., Ashour, M. E., & Flouty, O. (2023). Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability. *Frontiers in Research Metrics and Analytics*, *8*. https://doi.org/10.3389/frma.2023.1268045

Shen, S., Zhao, C., Wu, C., Sun, S., Li, Z., Yan, W., & Shao, Z. (2023). Allosteric modulation of G protein-coupled receptor signaling. *Frontiers in Endocrinology*, *14*(February), 1–13. https://doi.org/10.3389/fendo.2023.1137604

Shi, Y., Cao, S., Ni, D., Fan, J., Lu, S., & Xue, M. (2022). The Role of Conformational Dynamics and Allostery in the Control of Distinct Efficacies of Agonists to the Glucocorticoid Receptor. *Frontiers in Molecular Biosciences*, *9*(July), 1–19. https://doi.org/10.3389/fmolb.2022.933676

Son, A., Kim, W., Park, J., Lee, W., Lee, Y., Choi, S., & Kim, H. (2024). Utilizing Molecular Dynamics Simulations, Machine Learning, Cryo-EM, and NMR Spectroscopy to Predict and Validate Protein Dynamics. *International Journal of Molecular Sciences*, *25*(17). https://doi.org/10.3390/ijms25179725

Ul Haque, A., Patel, D. N., & Gkasis, P. (2025). The Impact of Artificial Intelligence on Modern Entrepreneurship. *Entrepreneurship: A Contemporary Perspective*, 154–162. https://doi.org/10.4324/9781003504351-13

Verma, N., Dhiman, B., Singh, V., Kaur, J., Guleria, S., & Singh, T. (2024). Exploring the global landscape of work-life balance research: A bibliometric and thematic analysis. *Heliyon*, *10*(11), e31662. https://doi.org/10.1016/j.heliyon.2024.e31662

Voicu, V., Brehar, F. M., Toader, C., Covache-Busuioc, R. A., Corlatescu, A. D., Bordeianu, A., Costin, H. P., Bratu, B. G., Glavan, L. A., & Ciurea, A. V. (2023). Cannabinoids in Medicine: A Multifaceted Exploration of Types, Therapeutic Applications, and Emerging Opportunities in Neurodegenerative Diseases and Cancer Therapy. *Biomolecules*, *13*(9), 1–45. https://doi.org/10.3390/biom13091388

Wen, S., Tan, Q., Baheti, R., Wan, J., Yu, S., Zhang, B., & Huang, Y. (2024). Bibliometric analysis of global research on air pollution and cardiovascular diseases: 2012–2022. *Heliyon*, *10*(12), e32840. https://doi.org/10.1016/j.heliyon.2024.e32840

Xia, J., Kang, J., & Xu, X. (2024). Global Research Trends and Future Directions in Urban Historical Heritage Area Conservation and Development: A 25-Year Bibliometric Analysis. *Buildings*, *14*(10). https://doi.org/10.3390/buildings14103096

Yan, Y., Edwards, B. I., & Sanmugam, M. (2025). Scientometric analysis of emerging trends and research landscape of ERNIE Bot's potentials as an educational tool: A mixed method study of a large language model. *Social Sciences and Humanities Open*, *12*(February). https://doi.org/10.1016/j.ssaho.2025.101729

Yasmeen, G., Anthonysamy, L., & Ojo, A. O. (2025). Resource-Governed BDA Adoption for Resilient Supply-Chain Operations: Qualitative Evidence from Malaysian Manufacturing Industry. In *Sustainability (Switzerland)* (Vol. 17, Issue 21). https://doi.org/10.3390/su17219620