SCIENTIFIC Publishing Limited

Digital Technologies Research and Applications
http://ojs.ukscip.com/index.php/dtra

Article

Thematic Evolution of Artificial Intelligence Research in Chinese-Speaking Academia (2021–2025): A Bibliometric and Text-Mining Analysis Using VOSviewer and KH Coder

Sheng-Ming Wang and **Chuding Chen** *

Department of Design, National Taipei University of Technology, Taipei 106, Taiwan
* Correspondence: ccd123501501039@gmail.com or 501501039@qq.com

Received: 29 November 2025; **Revised:** 29 December 2025; **Accepted:** 7 January 2026; **Published:** 20 January 2026

Abstract: This study examines the visible thematic patterns of artificial intelligence (AI)–related research in Mainland China, Taiwan, Hong Kong, and Macao from 2021 to 2025. The analysis is based on Scopus-indexed journal articles and uses keyword frequency counts, VOSviewer density visualizations, and KH Coder co-occurrence networks. These methods are applied to describe how AI-related keywords appear across the four regions and how their distributions change during the five-year period. Across all datasets, terms such as machine learning, deep learning, and neural network appear frequently and occupy central positions in the visual outputs. The VOSviewer heatmaps show that regions with larger publication volumes display wider areas of keyword density, while regions with smaller datasets present more compact clusters. Beginning in 2024 and 2025, generative AI–related terms, including large language model and ChatGPT, become visible across all regions. The KH Coder networks illustrate that the four regions contain multiple clusters of co-occurring keywords, with differences in cluster size and distribution reflecting the underlying dataset scale and the topics present in each regional corpus. Overall, the results offer a descriptive account of how AI-related terms appear in the collected datasets and how their visible distributions vary among the four regions during the study period. The findings are intended to summarize observable patterns without inferring causal explanations or evaluating the significance of regional differences.

Keywords: Artificial Intelligence; Bibliometric Analysis; VOSviewer; KH Coder

1. Introduction

Artificial intelligence (AI) has become a central driver of technological and scientific advancement worldwide. With rapid developments in machine learning (ML) and deep learning (DL), AI technologies have transformed applications across healthcare, engineering, natural language processing and data-intensive sciences. Foundational studies such as LeCun, Bengio and Hinton have highlighted the transformative capabilities of deep learning and its role in reshaping global research directions [1]. More recently, the emergence of large-scale generative models has introduced new methodological paradigms and accelerated the expansion of AI-related research outputs, resulting in increasingly complex publication landscapes [2]. Despite this global growth, existing bibliometric studies indicate that AI research is unevenly distributed across regions, and comparative thematic mapping between culturally related but institutionally distinct regions remains underexplored [3]. This gap is particularly evident in Mainland China, Taiwan, Hong Kong and Macao, where AI research is highly active yet lacks systematic cross-regional comparison of thematic structures.

Governments, industries and academic institutions have responded to the rapid expansion of AI by introducing strategic initiatives and research programs aimed at strengthening innovation capacity. In parallel, academic research has increasingly adopted bibliometric methods to understand AI-related publication trends, identify thematic hotspots and analyze research structures. Bibliometric tools such as VOSviewer and text-mining methods have been widely employed to map scientific landscapes across domains including computer science, engineering and information technology [4]. Despite this growing methodological interest, existing studies often focus on single regions or narrow disciplinary domains. Only a small number of studies investigate AI research within individual parts of Mainland China, Taiwan, Hong Kong and Macao, and even fewer compare thematic structures across the four regions using standardized analytical tools. As a result, current literature still lacks a coherent, data-driven description of how AI-related research themes appear and differ among Mainland China, Taiwan, Hong Kong and Macao.

To address this gap, the present study applies a multi-method analytical framework that integrates keyword frequency analysis, VOSviewer heatmap visualization and KH Coder co-occurrence network analysis. These methods enable the extraction of visible thematic patterns, visualization of keyword density and identification of co-occurrence relationships across large publication datasets. By applying this analytical framework consistently across the four regions between 2021 and 2025, the study provides a structured and descriptive overview of AI-related thematic distributions. This approach offers a transparent and replicable way to observe similarities and differences across Mainland China, Taiwan, Hong Kong, and Macao, and contributes to establishing a clearer empirical understanding of the thematic landscape of AI research across these research contexts.

Based on the gaps identified above, the study addresses the following research questions:

RQ1. What prominent AI-related themes appear in the publications from Mainland China, Taiwan, Hong Kong and Macao between 2021 and 2025?

RQ2. How do the visible distributions of AI-related keywords differ across the four regions based on visualization outputs?

RQ3. How do AI-related themes observed in the four regions change during the five-year period?

2. Literature Review

2.1. Definition and Conceptual Scope of Artificial Intelligence (AI)

Artificial Intelligence (AI) refers to computational systems capable of performing tasks that traditionally require human intelligence, such as perception, reasoning, learning, problem-solving, and autonomous decision-making. Russell and Norvig conceptualize AI as the study of rational agents that perceive their environment and act to maximize the likelihood of achieving their goals, establishing a foundational framework for AI as both a theoretical and engineering discipline [5].

In contemporary research, AI is predominantly operationalized through machine learning (ML) and deep learning (DL). ML represents a data-driven paradigm in which algorithms improve performance through experience, forming the methodological backbone of modern AI [6]. Deep learning, a subfield of ML employing multi-layer neural networks, enables hierarchical feature extraction and has produced transformative advances in fields such as computer vision, speech recognition, and natural language processing [6]. The resurgence of neural network–based methods has redefined AI research trajectories and shaped global technological development.

Since the late 2010s, AI has experienced another paradigm shift through the emergence of large-scale "foundation models." These models—trained on massive, heterogeneous datasets—exhibit generalization capabilities across a broad range of downstream tasks. Bommasani et al. argue that foundation models constitute a new class of general-purpose AI systems with profound implications for science, industry, and society [7]. Their capabilities are exemplified by systems such as GPT-4, whose empirical evaluations demonstrate advanced reasoning and multimodal problem-solving abilities beyond prior AI benchmarks [8]. These developments underscore the rapid evolution of AI from task-specific systems to versatile, generalizable architectures, thereby influencing global research landscapes.

2.2. Definition and Foundations of Bibliometrics

Bibliometrics is the quantitative study of scholarly communication, emphasizing the statistical analysis of publications, citations, metadata, and textual features to map the intellectual structure and development of scientific fields. Donthu et al. define bibliometrics as a rigorous methodological framework that enables systematic examination of research output, influential scholarship, emerging themes, and disciplinary evolution [4]. Its increasing use reflects the rapid growth of scientific publications and the necessity for analytical tools that can synthesize complex patterns at scale.

Classic bibliometric foundations were laid by Price, who introduced the concept of exponential scientific growth and the cumulative nature of scientific knowledge, setting the conceptual basis for modern research evaluation [9]. Small later advanced the field through the development of co-citation analysis, demonstrating that citation patterns could reveal the intellectual organization of scientific specialties [10]. These foundational contributions shaped the trajectory of contemporary bibliometrics and established enduring analytical conventions.

Modern bibliometric techniques include citation analysis, co-authorship networks, keyword co-occurrence mapping, burst detection, clustering algorithms, and text mining approaches. Dedicated toolkits have also been developed to support comprehensive science mapping workflows, such as bibliometrix [11]. Co-word analysis is a foundational approach for mapping conceptual structures in scientific fields through patterns of term co-occurrence [12]. Zupic and Čater argue that keyword co-occurrence analysis is particularly valuable for identifying research fronts and thematic clusters in fast-evolving scientific fields [13]. In artificial intelligence research, bibliometric mapping is widely used to trace technological diffusion, comparative regional performance, and shifts in methodological dominance. For example, Agbo et al. conducted the first comprehensive bibliometric analysis of the smart learning environments field, mapping its intellectual structure, identifying established and emerging research hotspots, and pinpointing underdeveloped yet promising themes that require further integration with the core concept of smart learning [14]. De la Vega Hernández et al.'s global bibliometric mapping of artificial intelligence research shows how co-occurrence and network-based analyses can uncover conceptual clusters, evolutionary trajectories, and cross-national differences in research strengths [15].

In sum, bibliometrics provides a robust analytical foundation for examining temporal and regional variations in AI research—a core objective of this study.

2.3. KH Coder: Definition, Analytical Capabilities, and Prior Applications

KH Coder is an open-source software platform designed for quantitative content analysis, text mining, and linguistic pattern extraction. Developed by Higuchi, the software integrates natural language processing with statistical analysis to support tasks such as tokenization, morphological analysis, frequency computation, co-occurrence analysis, hierarchical clustering, and multidimensional scaling [16]. Its graphical co-occurrence network is one of its most widely utilized features, enabling researchers to visualize semantic structures and thematic clusters within large textual datasets. These capabilities make KH Coder particularly suitable for analyzing author-supplied keywords in bibliometric research, where identifying conceptual linkages and thematic convergence is essential.

KH Coder has been applied across a wide range of research domains. Kakouris and Georgiadis employed KH Coder to analyze entrepreneurship education literature, using co-occurrence networks to identify conceptual clusters and thematic evolution within the field [17]. Oe and Weeks used KH Coder to examine the global diffusion of Japanese "kawaii" cultural values, illustrating how text mining combined with bibliometrics uncovers deeper socio-cultural narratives embedded in large corpora [18].

More recently, KH Coder has been integrated into sophisticated domain-specific investigations. Sakai et al. applied text mining and bibliometric analysis using KH Coder to identify research topics in occupational medicine from 1990 to 2022, revealing long-term trends and thematic transitions within the discipline [19]. Nagasue et al. used KH Coder to systematically review passive cooling technologies in hot and humid climates, highlighting the tool's strengths in synthesizing dispersed technical literature into coherent thematic structures [20].

Collectively, these studies demonstrate KH Coder's effectiveness for reconstructing semantic networks, mapping conceptual domains, and identifying thematic evolution across diverse research areas. Its demonstrated reliability and adaptability justify its adoption in the present study for analyzing AI-related keyword corpora across distinct regional research ecosystems.

2.4. Positioning of the Present Study

Recent studies have increasingly employed bibliometric and visualization-based approaches to examine the structure and evolution of artificial intelligence research. Such work has mapped thematic clusters, identified emerging topics, and traced longitudinal trends across large publication corpora, often at a global scale or within specific application domains. For example, Deleca [21] conducted a bibliometric analysis of the Web of Science database, revealing sustained annual growth in AI-related publications and identifying major research trends and thematic clusters in local and regional studies between 2002 and 2023. Ekundayo and Chaudhry [22] analyzed a large corpus of AI-related publications, highlighting key thematic areas such as applications in education and emerging technologies, as well as patterns of international collaboration.

While these studies provide valuable overviews of AI research landscapes, they typically emphasize aggregate patterns or methodological innovations rather than region-specific comparisons. In contrast, the present study adopts a deliberately descriptive perspective, concentrating on the comparative visibility of AI-related themes across Mainland China, Taiwan, Hong Kong, and Macao. By focusing on author-supplied keywords and consistent journal-based data, this work complements existing surveys by offering a regionally differentiated mapping of AI research themes without introducing evaluative or causal claims.

3. Method

3.1. Data Source and Sample Construction

The literature search in this study followed a predefined and transparent set of criteria to ensure conceptual consistency and cross-regional comparability. Bibliographic records were retrieved from the Scopus database for the years 2021–2025 (2021, 2022, 2023, 2024, and 2025). Scopus is widely used in bibliometric research and has been compared with other major bibliographic sources in terms of coverage and citation indexing [23]. Prior studies have shown that database coverage and citation counts can differ across Scopus, Web of Science, and Google Scholar, which may affect comparative bibliometric results [24]. Publications were included only when the exact term "artificial intelligence" appeared in the Author Keywords field, thereby retaining articles that explicitly self-identified AI as a central thematic focus. To ensure consistency across records, the dataset was restricted to documents categorized as Article under Document Type and published in Journal sources under Source Type. Regional assignment was determined based on the affiliation country/territory reported in each Scopus record, limited to Mainland China, Taiwan, Hong Kong, and Macao.

Under these parameters, the search retrieved a total of 25,973 records for Mainland China, 2664 for Taiwan, 1753 for Hong Kong, and 407 for Macao. These totals represent the complete set of journal articles meeting the inclusion criteria (**Figure 1**). To facilitate the interpretation of **Figure 1**, the annual publication trends for each region are summarized below.

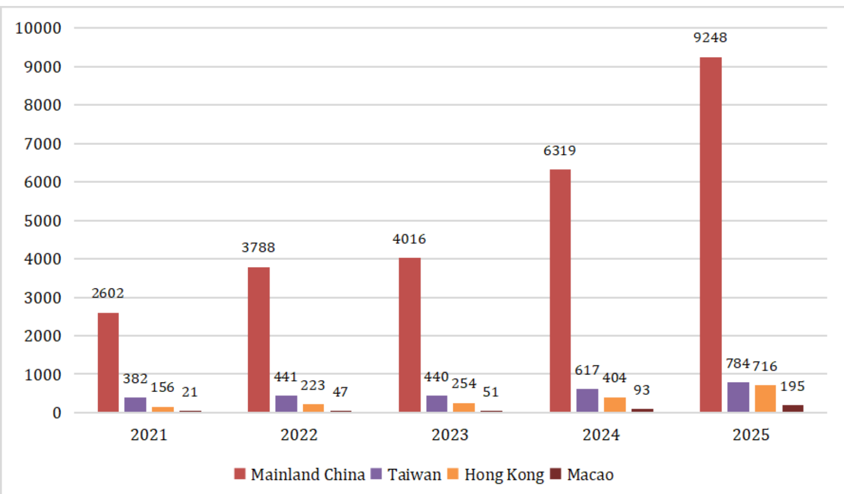

Figure 1. Annual AI-related Journal Articles by Region (2021–2025).

- Mainland China exhibits a sustained and accelerating increase in publication volume, rising from 2602 articles in 2021 to 3788 in 2022, 4016 in 2023, 6319 in 2024, and reaching 9248 in 2025.
- Taiwan shows a more moderate but steady growth trajectory, with annual counts increasing from 382 in 2021 to 441 in 2022, remaining relatively stable, 440 in 2023, and subsequently rising to 617 in 2024 and 784 in 2025.
- Hong Kong where publication numbers grow from 156 in 2021 to 223 in 2022, 254 in 2023, 404 in 2024, and 716 in 2025.
- Although Macao's output remains modest in volume, it shows a steady upward trend, rising from 21 publications in 2021 to 47 in 2022, 51 in 2023, 93 in 2024, and 195 in 2025.

The annual distribution of publication counts across regions documents the year-by-year composition of the dataset and provides the empirical basis for subsequent keyword frequency calculations, density visualizations, and co-occurrence analyses. The corpus therefore consists of four region-specific subsets of differing sizes, a factor that is relevant for interpreting the comparative visibility and density of thematic patterns in the later analytical sections.

3.2. Keyword Corpus Development and Descriptive Profiling

Author-provided keywords were extracted for every article in the final Scopus sample, forming the core corpus for hotspot identification and longitudinal comparison. In bibliometric research, author keywords are widely regarded as high-signal indicators of explicit topical intent, and their aggregation allows systematic mapping of dominant themes and emerging agendas. To reduce measurement noise and ensure cross-year and cross-regional comparability, the keyword corpus underwent a rule-based normalization process. Orthographic variants were unified by standardizing letter case; morphological variants such as singular/plural forms and near-duplicate expressions were aligned under a consistent counting convention; domain-specific synonym sets were reviewed and harmonized where conceptual equivalence was evident; and non-thematic generic items lacking discriminative value were removed. Following normalization, keywords were aggregated within each region–year subset and occurrence frequencies were computed. The ten most frequent keywords per subset were retained as indicators of explicit research hotspots and are reported in **Tables 1–4**. These Top-10 lists constitute the direct empirical basis for the integrated analysis in Section 3.3.

Table 1. Top-10 Author Keywords in Mainland China (2021–2025).

Rank	2021 (Freq.)	2022 (Freq.)	2023 (Freq.)	2024 (Freq.)	2025 (Freq.)
1	deep learning (176)	deep learning (332)	deep learning (373)	deep learning (482)	deep learning (701)
2	machine learning (142)	machine learning (270)	machine learning (305)	machine learning (428)	machine learning (646)
3	internet of things (67)	internet of things (86)	internet of things (75)	internet of things (108)	large language model (418)
4	COVID-19 (53)	convolutional neural network (72)	neural network (64)	explainable artificial intelligence (99)	generative artificial intelligence (405)
5	convolutional neural network (44)	neural network (67)	convolutional neural network (51)	ChatGPT (99)	explainable artificial intelligence (155)
6	neural network (36)	COVID-19 (45)	COVID-19 (39)	neural network (97)	ChatGPT (137)
7	convolutional neural networks (22)	big data (44)	attention mechanism (36)	generative artificial intelligence (87)	neural network (125)
8	reinforcement learning (20)	ant colony optimization (37)	computer vision (36)	convolutional neural network (81)	internet of things (109)
9	big data (18)	computed tomography (31)	magnetic resonance imaging (35)	large language model (73)	federated learning (91)
10	computed tomography (17)	transfer learning (29)	explainable artificial intelligence (32)	computer vision (55)	computer vision (82)

Table 2. Top-10 Author Keywords in Taiwan (2021–2025).

Rank	2021 (Freq.)	2022 (Freq.)	2023 (Freq.)	2024 (Freq.)	2025 (Freq.)
1	machine learning (49)	machine learning (69)	machine learning (59)	deep learning (67)	deep learning (80)
2	deep learning (43)	deep learning (67)	deep learning (51)	machine learning (53)	machine learning (78)
3	internet of things (27)	internet of things (22)	internet of things (20)	ChatGPT (25)	generative artificial intelligence (57)
4	convolutional neural network (16)	convolutional neural network (16)	convolutional neural network (15)	explainable artificial intelligence (23)	large language model (39)
5	Artificial Intelligence of Things (9)	Artificial Intelligence of Things (10)	explainable artificial intelligence (13)	internet of things (23)	ChatGPT (31)
6	COVID-19 (8)	explainable artificial intelligence (9)	neural network (10)	generative artificial intelligence (19)	explainable artificial intelligence (24)
7	convolutional neural networks (8)	transfer learning (7)	Artificial Intelligence of Things (10)	Artificial Intelligence of Things (17)	internet of things (23)
8	neural network (7)	artificial neural network (7)	edge computing (9)	convolutional neural network (13)	neural network (18)
9	edge computing (5)	COVID-19 (7)	cloud computing (8)	large language model (12)	convolutional neural network (16)
10	transfer learning (5)	big data (7)	random forest (7)	natural language processing (10)	natural language processing (11)

Table 3. Top-10 Author Keywords in Macao (2021–2025).

Rank	2021 (Freq.)	2022 (Freq.)	2023 (Freq.)	2024 (Freq.)	2025 (Freq.)
1	feature extraction (2)	COVID-19 (3)	machine learning (7)	machine learning (10)	generative artificial intelligence (19)
2	neural network (2)	machine learning (3)	deep learning (5)	ChatGPT (5)	large language model (11)

Table 4. Top-10 Author Keywords in Macao (2021–2025).

Rank	2021 (Freq.)	2022 (Freq.)	2023 (Freq.)	2024 (Freq.)	2025 (Freq.)
3	others ≤ 1	ant colony optimization (2)	natural disaster (3)	generative artificial intelligence (5)	machine learning (7)
4	others ≤ 1	internet of things (2)	natural hazard (3)	deep learning (4)	deep learning (6)
5	others ≤ 1	unsupervised learning (2)	storm surge (3)	Artificial Intelligence of Things (4)	neural network (5)
6	others ≤ 1	deep learning (2)	tropical cyclone (3)	explainable artificial intelligence (3)	ChatGPT (5)
7	others ≤ 1	supervised learning (2)	breast cancer (3)	neural network (2)	explainable artificial intelligence (5)
8	others ≤ 1	others ≤ 1	Artificial Intelligence of Things (3)	explainable AI (2)	internet of things (3)
9	others ≤ 1	others ≤ 1	climate change (2)	compact data (2)	AI literacy (3)
10	others ≤ 1	others ≤ 1	early warning (2)	data complexity (2)	Artificial Intelligence of Things (2)

Table 5. Top-10 Author Keywords in Hong Kong (2021–2025).

Rank	2021 (Freq.)	2022 (Freq.)	2023 (Freq.)	2024 (Freq.)	2025 (Freq.)
1	machine learning (21)	deep learning (24)	machine learning (28)	generative artificial intelligence (35)	generative artificial intelligence (104)
2	COVID-19 (11)	machine learning (23)	deep learning (16)	ChatGPT (28)	large language model (56)
3	deep learning (10)	internet of things (10)	neural network (7)	deep learning (27)	machine learning (31)
4	internet of things (6)	digital health (5)	internet of things (7)	deep learning (25)	deep learning (27)
5	gene expression programming (4)	federated learning (3)	ChatGPT (5)	large language model (17)	ChatGPT (26)
6	neural network (3)	COVID-19 (3)	generative artificial intelligence (5)	explainable artificial intelligence (13)	explainable artificial intelligence (12)
7	genetic algorithm (3)	convolutional neural network (3)	federated learning (5)	internet of things (6)	Artificial Intelligence Generated Content (11)
8	digital twin (2)	transfer learning (3)	computed tomography (4)	computer vision (6)	higher education (10)
9	smart city (2)	explainable artificial intelligence (3)	COVID-19 (4)	instructional design (4)	computer vision (8)
10	natural language processing (2)	industrial internet of things (2)	bibliometric analysis (4)	AI literacy (4)	artificial intelligence literacy (7)

Using the publication volumes in **Table 1** and the explicit hotspot profiles in **Tables 2–4**, this section delineates the observable thematic evolution of AI research across Mainland China, Taiwan, Hong Kong, and Macao from 2021 to 2025. The analysis addresses within-region longitudinal trajectories and cross-regional convergence and divergence under a shared global paradigm transition.

Mainland China exhibits a pronounced dual structure. The persistent dominance of deep learning and machine learning in **Table 2** across all five years, together with their steep frequency growth, indicates long-run methodological consolidation around core learning paradigms that scale in parallel with the rapid expansion of AI publication output. A structural breakpoint becomes visible after 2024, when generative-AI terms enter the Top-10 list and rise sharply by 2025, revealing accelerated thematic realignment toward LLM-centered research. At the same time, the continued appearance of internet of things suggests stability in AI-industry integration, whereas the disappearance of medical-imaging terms after 2023 implies a temporary reprioritization from sectoral applications toward paradigm-level innovation within the keyword hierarchy.

Taiwan displays a more gradual trajectory. **Table 3** shows a stable ML/DL core throughout 2021–2025 with moderate growth, aligning with Taiwan's steady increase in publication volume. The recurring presence of internet of things and Artificial Intelligence of Things indicates a durable specialization in AIoT-linked industrial and edge contexts. Generative-AI terms emerge from 2024 and expand further in 2025, yet they coexist with, rather than displace, the established ML/DL-AIoT backbone, suggesting responsive adoption under structural continuity.

Hong Kong and Macao, although smaller in absolute output, reveal distinctive late-period convergence. In Macao, **Table 4** shows that the 2021–2022 subsets are not robust for hotspot inference, 2023 features dispersed application-oriented keywords, and 2024–2025 exhibit rapid synchronization with global GenAI/LLM trajectories, accompanied by the emergence of AI literacy in 2025. In Hong Kong, **Table 4** indicates an early ML/DL-centered profile with short-lived COVID-19 salience in 2021, followed by a rapid shift to generative-AI dominance after 2024; the entry of higher education and AI literacy into the 2025 Top-10 list suggests localized coupling between GenAI diffusion and higher-education or pedagogical agendas.

3.3. VOSviewer-Based Visualization of AI-Related Keywords in Mainland China, Taiwan, Hong Kong, and Macao

VOSviewer is a software tool for constructing and visualizing bibliometric networks based on relationships such as citation, co-citation, bibliographic coupling, and co-authorship. It also provides text-mining functions for building and visualizing term co-occurrence networks [3,25].

3.3.1. Overview of Heatmap Visualization

This subsection presents the VOSviewer density visualizations generated from AI-related publications in Mainland China, Taiwan, Macao, and Hong Kong. The heatmaps display keywords according to their frequency and co-occurrence density within each regional dataset. Warmer colors indicate higher local density of keyword occurrences, while cooler colors represent lower density. The descriptions below report only the terms and density patterns that appear explicitly in the visualizations.

3.3.2. Mainland China Heatmap

In the heatmap for Mainland China (**Figure 2**), several high-density areas are visible.

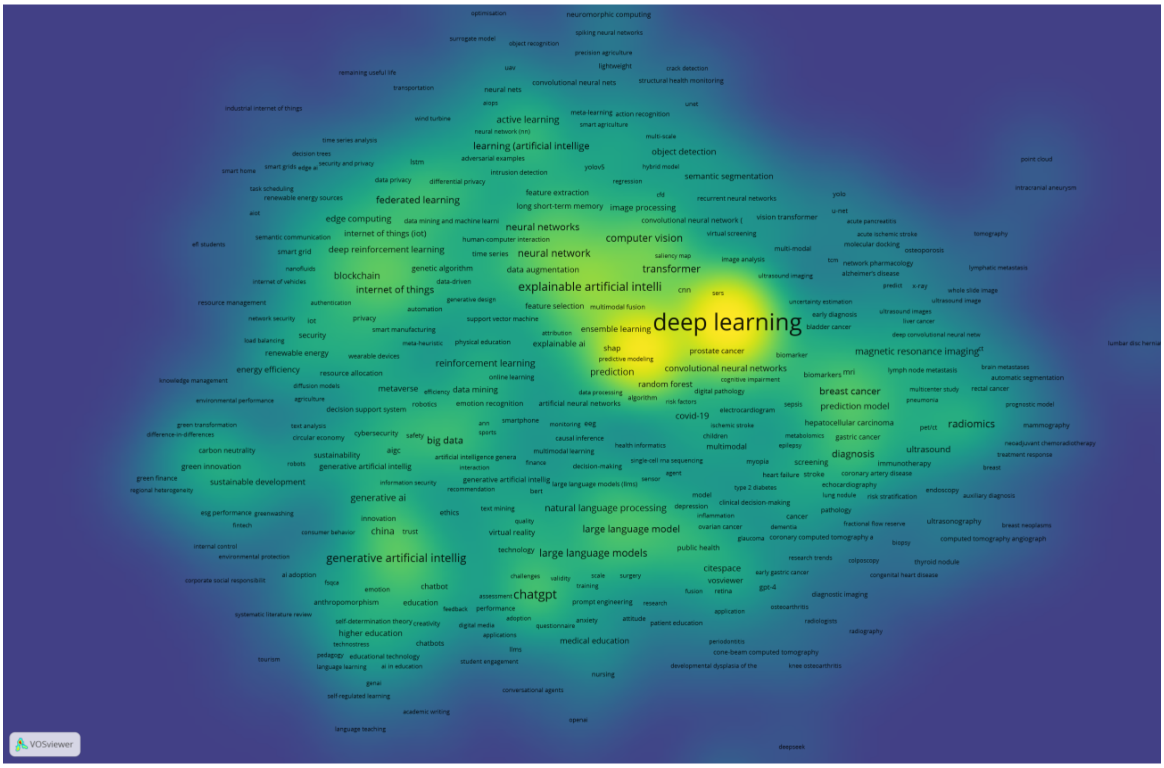

Figure 2. Mainland China Heatmap.

The term "deep learning" appears at the center of the main high-density region. Surrounding this term, other frequently displayed keywords include: convolutional neural network, neural network, computer vision, transformer, prediction, image processing.

Another set of visible keywords includes: internet of things, federated learning, reinforcement learning, edge computing.

Medical-related terms also appear in the visualization, including: magnetic resonance imaging, breast cancer, radiomics, ultrasound, COVID-19.

Terms associated with recent developments in AI can also be seen, such as: generative artificial intelligence, ChatGPT, large language model. These terms form multiple areas of density across the heatmap.

3.3.3. Taiwan Heatmap

In the Taiwan heatmap (**Figure 3**), "deep learning" and "machine learning" appear in high-density regions. Other visible terms near these areas include: convolutional neural network, neural networks, explainable artificial intelligence, radio mics, image recognition.

Additional keywords that appear in the heatmap include: natural language processing, internet of things, edge computing, blockchain. Several terms related to generative AI also appear: generative AI, ChatGPT, large language model. Educational and learning-related terms such as self-regulated learning and computational thinking are also visible in the visualization.

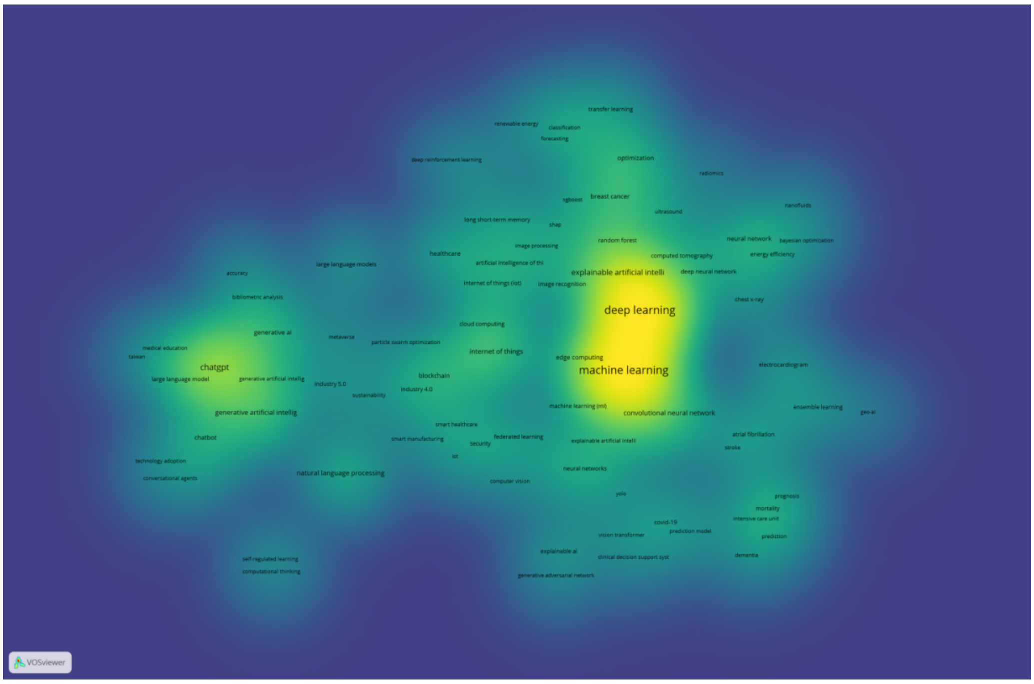

Figure 3. Taiwan Heatmap.

3.3.4. Macao Heatmap

The Macao heatmap presents several localized density areas (**Figure 4**). Three clusters of visible terms include:

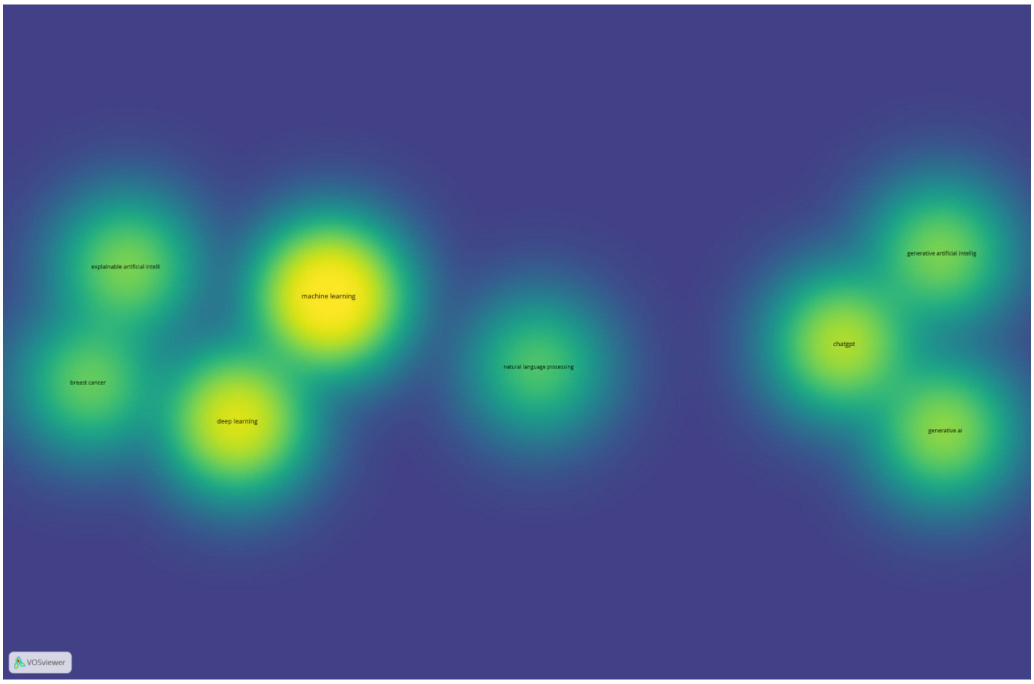

Figure 4. Macao Heatmap.

A cluster containing "machine learning", accompanied by "deep learning" and "explainable artificial intelligence." A cluster where natural language processing appears as a stand-alone visible term. A cluster containing generative-AI-related terms: ChatGPT, generative AI, generative artificial intelligence. A medical-related term, breast cancer, also appears in one of the density areas.

3.3.5. Hong Kong Heatmap

In the Hong Kong heatmap (**Figure 5**), "deep learning" and "machine learning" are displayed in high-density regions. Other terms visible near these areas include: neural network, explainable artificial intelligence, computed tomography, ultrasound.

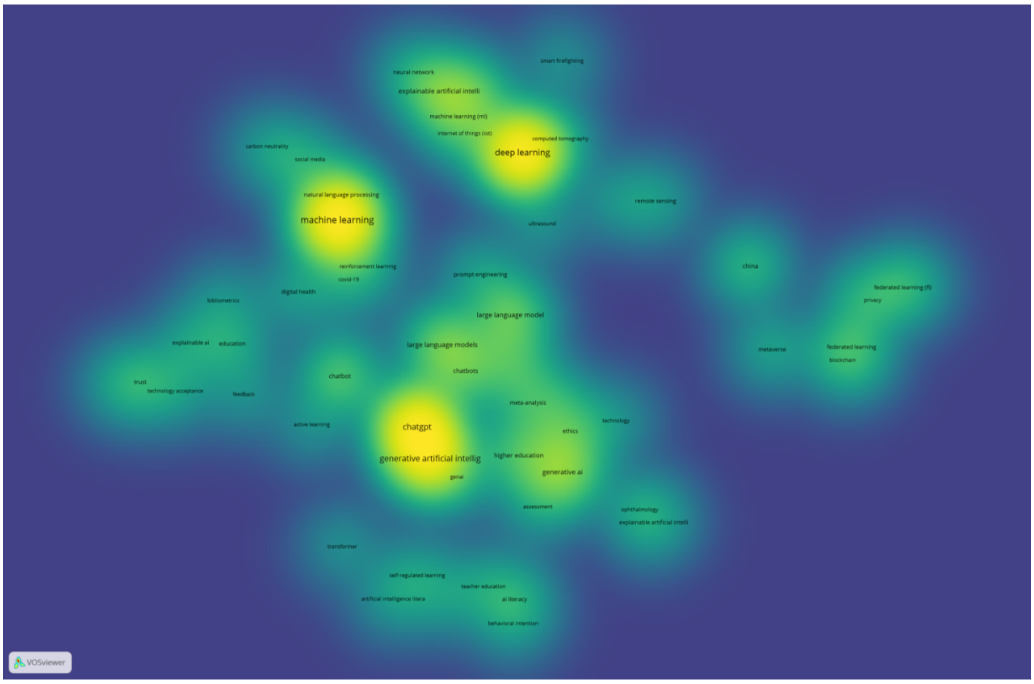

Figure 5. Hong Kong Heatmap.

Keywords appearing in other parts of the heatmap include: natural language processing, remote sensing, federated learning, blockchain.

Terms associated with generative AI form another area of density, including: ChatGPT, generative AI, large language models.

Additional visible terms include education, technology acceptance, trust, and AI literacy.

3.4. KH Coder Analysis

3.4.1. Mainland China

The co-occurrence network for Mainland China reveals several coherent clusters of frequently co-appearing keywords (**Figure 6**). These subgraphs reflect major thematic orientations in China's AI research ecosystem during 2021–2025. The following sections summarize five subgraphs that appear to carry the strongest thematic weight based on node size, density, and centrality.

- Subgraph 01—Core Machine Learning and Deep Learning Methods

This subgraph is centered on foundational AI methodologies, including deep learning, machine learning, neural network, image recognition, classification, and segmentation. The close distances among these nodes indicate consistent methodological co-occurrence, reflecting the continued prominence of deep learning paradigms in Mainland China's AI publications. Terms such as convolutional neural network and processing further reinforce the presence of vision-related tasks, suggesting that computer vision remains a stable and widely adopted application domain.

Overall, Subgraph 01 represents the methodological backbone of AI research in Mainland China. Its prominence is consistent with the country's large publication volume and its longstanding emphasis on algorithmic performance, optimization, and scalable model architectures. Although newer paradigms emerge in later years, the persistence of these keywords across 2021–2025 suggests continuity rather than displacement.

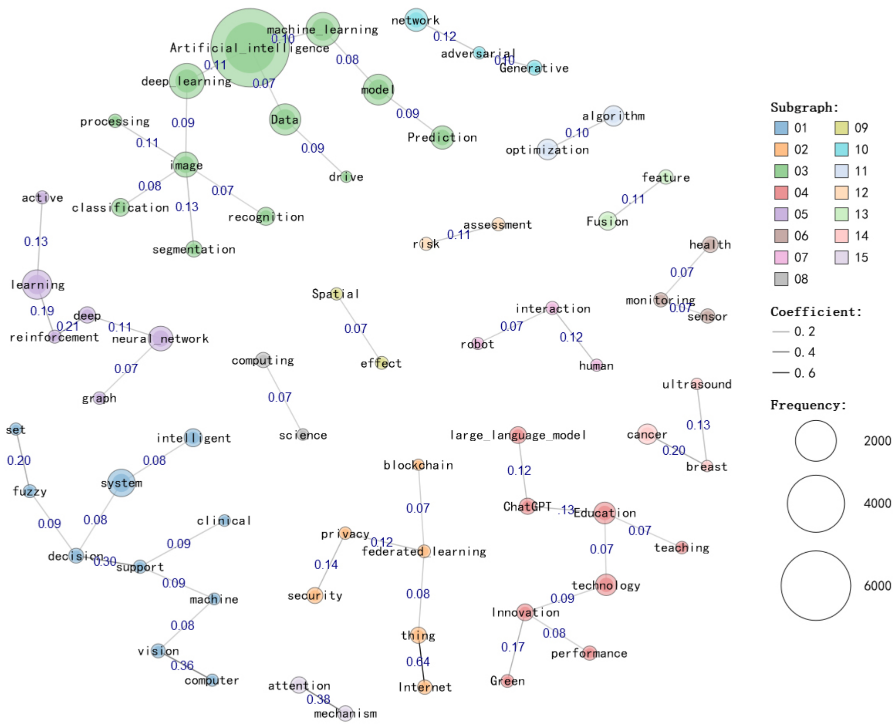

Figure 6. Co-occurrence network for Mainland China.

- Subgraph 02—Generative AI and Large Language Models

This subgraph contains keywords such as ChatGPT, generative, large language model, foundation model, and language. These terms form a concentrated cluster, indicating that generative AI has become a distinct and increasingly interconnected research theme. The presence of Education and teaching near this cluster suggests that one of the early application domains involves pedagogical or instructional scenarios, although such associations should be interpreted cautiously given the broad applicability of generative models.

The emergence of this subgraph aligns with the global diffusion of transformer-based architectures beginning in 2023–2024. In the context of Mainland China, this cluster's density suggests an expanding research interest in LLM-related topics, but the surrounding nodes also imply an exploratory stage, with attention distributed across multiple early application areas.

- Subgraph 03—Healthcare, Medical Imaging, and Diagnosis

Subgraph 03 highlights medically oriented keywords, including cancer, breast, ultrasound, tomography, and diagnosis. These nodes form a coherent group associated with AI-enabled medical imaging and clinical decision support. The spatial connection to safety and detection suggests continuing interest in accuracy, robustness, and risk assessment within clinical settings.

Although medical-imaging terms appeared more frequently in earlier years (2021–2023) within the hotspot tables, this subgraph indicates that medical applications remain present within the broader knowledge structure. The cluster's moderate density suggests that while medical AI is not the dominant thematic area, it continues to serve as a meaningful application domain that supports cross-disciplinary collaboration.

- Subgraph 04—Internet of Things, Edge Computing, and Industrial Applications

This subgraph includes Internet of Things, edge, inference, communication, industry, privacy, and security. The grouping reflects longstanding interest in the integration of AI with IoT systems. The presence of federated learning and privacy indicates attention to data governance issues associated with distributed architectures.

Compared with generative AI subgraphs, this cluster appears more established and methodologically mature. Its stability across multiple years suggests that Mainland China continues to prioritize AI deployment in industrial, manufacturing, and infrastructure settings—areas aligned with broader national strategies involving smart industry and digital transformation.

- Subgraph 05—Human–Computer Interaction, Social Acceptance, and Governance

This subgraph includes terms such as evaluation, interaction, social, governance, English, teacher, and higher. These keywords collectively point to topics related to human–AI interaction, user acceptance, education, and governance considerations. While the nodes are smaller compared to core methodology clusters, their presence reflects emerging interdisciplinary engagement around AI use in social and educational contexts.

Given the complexity of social-impact discussions, the connections within this subgraph likely represent early-stage engagement with issues of acceptance, evaluation, and governance rather than a fully established research field. Nevertheless, the appearance of these terms suggests expanding attention to human-centric and ethical dimensions.

Taken together, the Mainland China network suggests a research landscape where core machine-learning methods continue to dominate, while generative AI and its emerging applications begin to take shape as a notable new cluster. Medical applications, industrial IoT deployment, and elements of human-AI interaction appear as stable secondary themes. These observations provide a basis for comparison with other regions, where distinct research emphases may be visible despite shared global trends.

Across the five selected subgraphs, the co-occurrence network for Mainland China suggests a research structure characterized by continuity in core methodologies and gradual diversification into emerging thematic areas. Subgraph 01 indicates that deep learning and machine learning continue to function as the central methodological anchors, supporting a wide range of applications. Subgraph 02 shows that generative AI–related topics have begun to form a distinct thematic cluster, although the surrounding nodes suggest that such research may still be in an exploratory stage. Subgraph 03 demonstrates that medically oriented AI applications remain present but appear to be positioned as a secondary rather than dominant theme within the broader network. Subgraph 04 reflects the stable relevance of IoT and edge-computing applications, which align with long-term industrial and infrastructural development directions. Finally, Subgraph 05 illustrates early signs of engagement with human–AI interaction, governance, and educational aspects, though the density of the cluster indicates that this area is still emerging.

Taken together, the network structure suggests that Mainland China's AI research during 2021–2025 may be characterized by the coexistence of mature methodological foundations and emerging thematic experimentation. While core algorithms remain central, the increasing presence of generative AI, human-centric topics, and applied industrial or medical themes points to a gradual broadening of research interests. These patterns provide useful contextual grounding for cross-regional comparison in subsequent sections.

3.4.2. Taiwan

The co-occurrence network for Taiwan reveals several coherent clusters of frequently co-appearing keywords (**Figure 7**). These subgraphs reflect major thematic orientations in Taiwan's AI research ecosystem during 2021–2025. The following sections summarize five subgraphs that appear to carry the strongest thematic weight based on node size, density, and centrality.

- Subgraph 01—Core Machine Learning and Deep Learning Techniques

This subgraph contains machine learning, deep learning, neural network, regression, and classification, forming a dense methodological cluster. Their close proximity suggests that foundational algorithmic development continues to be a central theme in Taiwan's AI research. The cluster structure indicates strong internal coherence and sustained emphasis on computational model performance, training procedures, and predictive tasks.

Consistent with Taiwan's long-standing strengths in engineering and computer science, this subgraph reflects a stable methodological foundation rather than rapid thematic shifts. Its dominance indicates that ML/DL methods remain integral across multiple application areas.

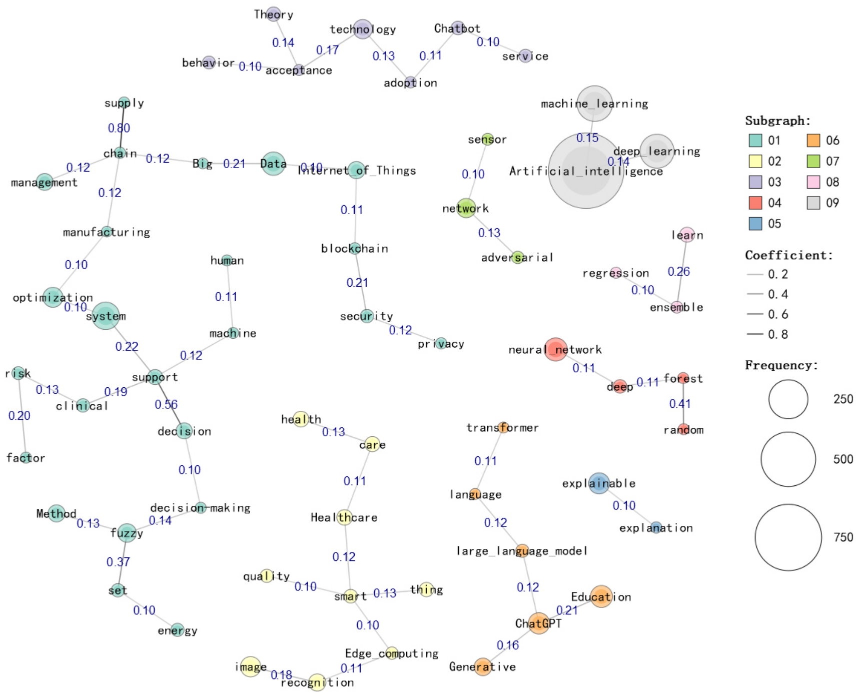

Figure 7. Co-occurrence network for Taiwan.

- Subgraph 02—Generative AI and Large Language Models

The second major cluster includes generative artificial intelligence, large language model, ChatGPT, language, and text. Their co-occurrence indicates that generative AI topics have begun to form a distinct research theme. Compared with Mainland China, the density of this cluster is moderate, suggesting emerging but still developing engagement with generative models.

The presence of language-related terms highlights the growing interest in natural language processing applications. However, the overall distribution suggests these topics may still be in the early stages of exploration within Taiwan's research environment.

- Subgraph 03—Healthcare and Medical Image Analysis

Subgraph 03 includes healthcare, disease, recognition, image, and support. These keywords cluster around AI-enabled medical imaging and diagnostic support tasks. Similar to Mainland China, medical applications appear as a recurring but not dominant theme in Taiwan's network.

The cluster's moderate density suggests sustained interest, potentially driven by Taiwan's strong biomedical research base, but the structure implies that such research likely coexists with—and depends heavily on—the methodological core represented in Subgraph 01.

- Subgraph 04—IoT, Edge Computing, and Smart Systems

This subgraph includes Internet of Things, edge computing, environment, system, and engineering. These nodes reflect Taiwan's sustained interest in integrating AI with IoT infrastructures, consistent with its industrial specialization in electronics, smart devices, and semiconductor-related technologies.

The presence of edge and system suggests that computational efficiency and deployment-oriented AI development—particularly for embedded or low-latency environments—remain important research goals.

- Subgraph 05—Explainable AI and Decision Support

This subgraph includes explainable artificial intelligence, support, evaluation, and decision. The grouping indicates that Taiwan's AI research places noticeable attention on interpretability, model transparency, and human-centered decision processes.

The cluster density is moderate, suggesting growing interest in the responsible deployment of AI, but its smaller size compared with core methodological clusters implies that this area is still an emerging field rather than a dominant research direction.

Together, the five subgraphs reveal a structure in which foundational ML/DL methods remain central, while generative AI, medical applications, IoT-related technologies, and interpretability issues appear as secondary but notable thematic areas. This section synthesizes these patterns to examine the broader trajectory of Taiwan's AI research.

Across the selected subgraphs, Taiwan's AI co-occurrence network suggests a research landscape marked by methodological stability and gradual thematic expansion. Subgraph 01 reflects continued reliance on core machine learning and deep learning techniques, forming the structural backbone of the network. Subgraph 02 indicates emerging engagement with generative AI and language-related applications, although the spread of keywords suggests that this area is still developing rather than consolidated. Subgraph 03 shows that healthcare and medical imaging applications maintain a consistent presence but tend to remain secondary in prominence. Subgraph 04 aligns with Taiwan's industrial strengths, highlighting ongoing interest in IoT and edge-computing systems. Subgraph 05 demonstrates attention to explainability and decision support, suggesting gradual expansion toward human-centered and responsible AI topics.

Overall, Taiwan's research structure appears to balance methodological continuity with incremental diversification. While foundational ML/DL techniques remain dominant, emerging themes such as generative AI and explainability suggest that Taiwan is cautiously integrating new global developments within its established research framework. These patterns provide a useful point of comparison with other regions evaluated in this study.

3.4.3. Macao

The co-occurrence network for Macao reveals several coherent clusters of frequently co-appearing keywords (**Figure 8**). These subgraphs reflect major thematic orientations in Macao's AI research ecosystem during 2021–2025. The following sections summarize five subgraphs that appear to carry the strongest thematic weight based on node size, density, and centrality.

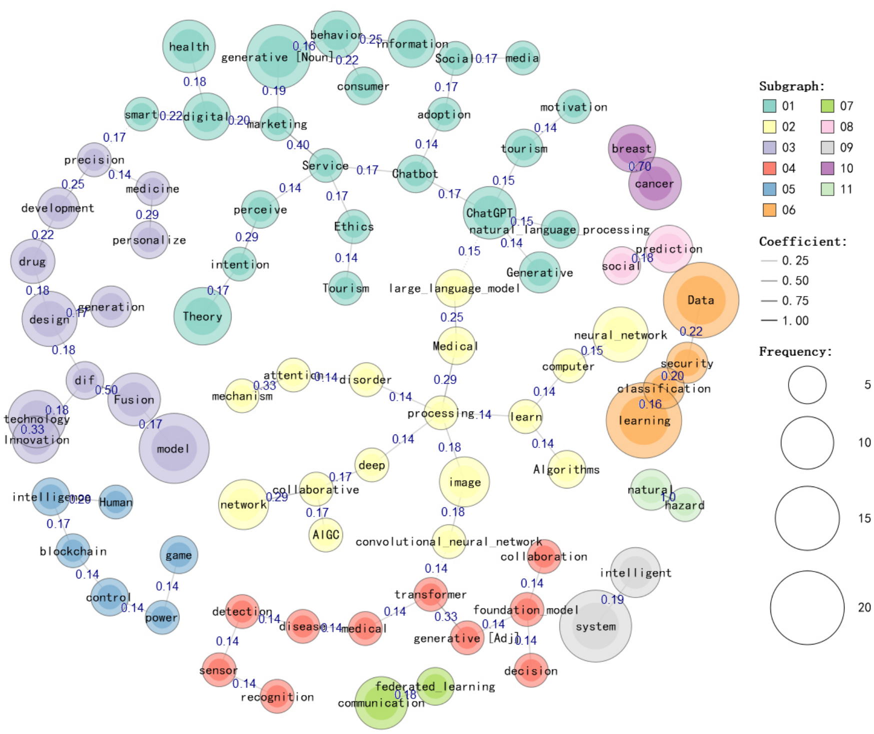

Figure 8. Co-occurrence network for Macao.

- Subgraph 01—Core Machine Learning and Deep Learning Methods

This subgraph contains machine learning, deep learning, training, model, and data, forming the central methodological cluster in Macao's network. Despite the smaller scale of research output, the prominence of these keywords suggests that Macao's AI research relies heavily on mainstream ML/DL frameworks.

The subgraph also shows connections to classification and accuracy, indicating an emphasis on performance optimization and predictive modeling. This aligns with broader global AI research practices, even within a smaller research ecosystem.

- Subgraph 02—Generative AI, LLMs, and Language-related Applications

Subgraph 02 includes generative, large language model, ChatGPT, text, and language. Similar to other regions, generative AI has emerged as an identifiable thematic cluster. However, the density of the cluster is moderate, suggesting that interest in LLM-based applications is present but still developing.

The presence of education as a neighboring node indicates early exploration of LLMs for instructional or pedagogical tasks, though such associations should be interpreted cautiously due to the broad applicability of generative models.

- Subgraph 03—Healthcare, Medical Diagnosis, and Biological Applications

This subgraph includes breast cancer, diagnosis, cell, biological, and detection. These keywords suggest that AI-assisted medical diagnosis constitutes a recognizable research direction within Macao's network, despite the limited scale of output.

The cluster is relatively compact, indicating that medical AI research may be concentrated within a small number of studies or collaborations, but nonetheless represents a recurring thematic orientation.

- Subgraph 04—Environmental and Disaster-related Applications

This subgraph contains climate, hazard, tropical cyclone, storm surge, and environment. The concentration of these terms suggests that environmental monitoring and disaster prediction represent a meaningful application area for AI research in Macao.

Given Macao's geographical location and susceptibility to typhoons, such thematic focus aligns logically with local needs. The density of this cluster suggests a coherent, application-driven research direction.

- Subgraph 05—AIoT and Smart Systems

This subgraph includes Artificial Intelligence of Things, IoT, sensor, system, and design. The cluster reflects interest in AI-enabled smart systems, particularly in the context of IoT integration. While smaller than the core methodological cluster, it appears as a stable and recurrent topic.

The presence of data complexity and compact data suggests that resource-efficient system design may be a relevant concern within these studies.

Together, these subgraphs indicate that although Macao's AI research community is relatively small, its thematic distribution spans methodological, generative, medical, environmental, and IoT-oriented domains. The next section synthesizes these patterns into a unified interpretation.

Across the five selected subgraphs, Macao's AI research network appears characterized by a compact but diversified thematic structure. Subgraph 01 reflects continued reliance on core ML/DL methodologies, forming the central foundation of the network. Subgraph 02 shows emerging engagement with generative AI topics, although the moderate density of this cluster suggests that such research is still at an early developmental stage. Subgraph 03 demonstrates that AI-assisted medical diagnosis, despite being produced by a limited number of publications, represents a consistent application area. Subgraph 04 highlights a distinctive focus on environmental and hazard-related applications, which may reflect local geographical conditions and practical needs. Subgraph 05 indicates ongoing interest in AIoT and smart-system design, albeit with smaller cluster size.

Overall, Macao's network structure suggests a research landscape that, while modest in scale, incorporates a range of themes consistent with broader global AI trends. The persistence of methodological and applied clusters indicates that research activities are distributed across both foundational and domain-specific topics. These characteristics provide a basis for comparison with the networks of other regions.

3.4.4. Hong Kong

The co-occurrence network for Hong Kong reveals several coherent clusters of frequently co-appearing keywords (**Figure 9**). These subgraphs reflect major thematic orientations in Hong Kong's AI research ecosystem during 2021–2025. The following sections summarize five subgraphs that appear to carry the strongest thematic weight based on node size, density, and centrality.

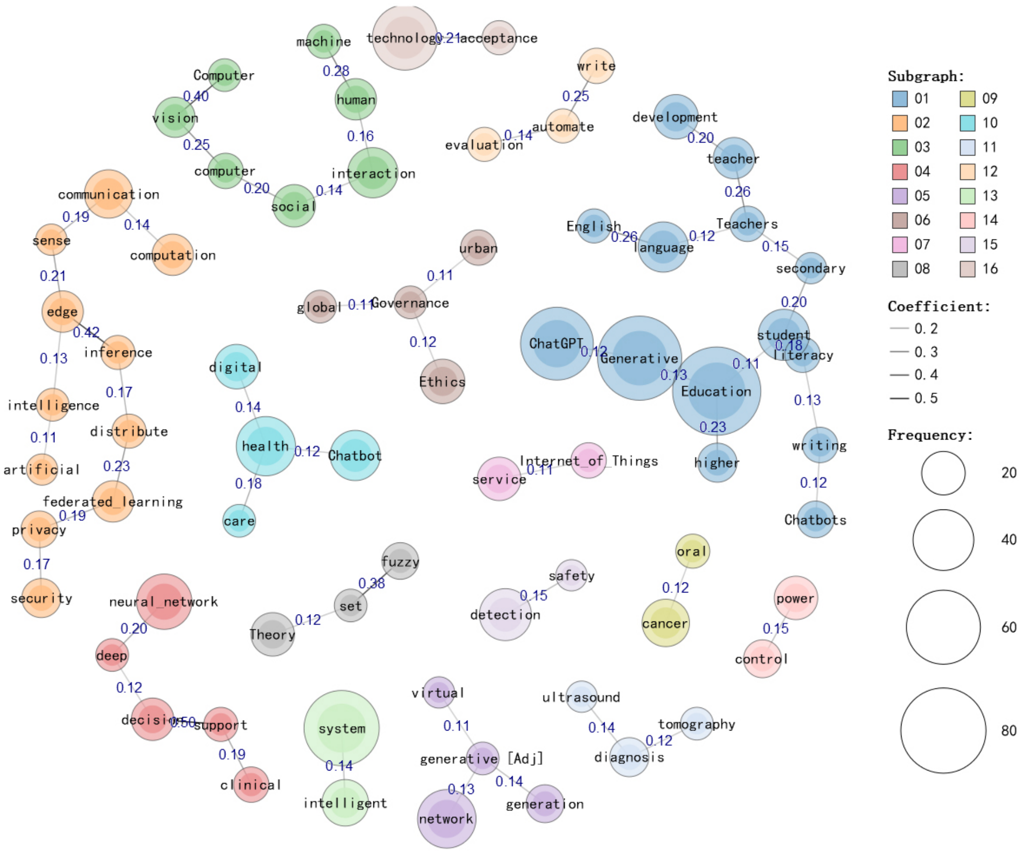

Figure 9. Co-occurrence network for Hong Kong.

- Subgraph 01—Machine Learning, Deep Learning, and Core Algorithms

This subgraph includes machine learning, deep learning, neural network, classification, and prediction. These nodes form a dense cluster that represents Hong Kong's methodological foundation in AI research.

The presence of model and training indicates a continued emphasis on algorithm performance and optimization. As in Mainland China and Taiwan, these methods appear to serve as the core computational tools supporting a wide range of downstream applications.

- Subgraph 02—Generative Artificial Intelligence and LLM-related Topics

This subgraph includes generative artificial intelligence, large language model, ChatGPT, prompt, and content. The density of this cluster is higher than that of Macao and Taiwan, suggesting that Hong Kong researchers show relatively concentrated interest in generative models and their associated applications.

The presence of education, higher, and learning near this cluster may indicate exploration of generative AI within instructional contexts. However, given the broad applicability of LLMs, such associations should be interpreted cautiously and not assumed to imply a dominant application field.

- Subgraph 03—Explainable AI and AI Ethics

Subgraph 03 contains explainable artificial intelligence, fairness, transparency, ethics, and trust. This constellation of terms indicates that Hong Kong may have a comparatively stronger emphasis on responsible AI, governance, and interpretability compared with other regions.

The density of this cluster suggests that these discussions may be more developed in Hong Kong's research environment, potentially linked to local academic priorities or interdisciplinary collaborations involving social sciences and public policy.

- Subgraph 04—Medical and Clinical Applications

This subgraph includes health, disease, medical, diagnosis, risk, and care. These keywords form a coherent cluster representing AI use in medical diagnosis, risk assessment, and clinical support.

While the cluster does not demonstrate the same density as similar clusters in Mainland China, it nonetheless indicates that medical AI continues to be a stable application area in Hong Kong's research output.

- Subgraph 05—Smart City, IoT, and Urban Technology

The final cluster contains Internet of Things, smart city, urban, sensor, environment, and mobility. These terms signal Hong Kong's interest in AI deployment within urban infrastructure and smart-city initiatives.

Compared with Taiwan and Mainland China, the cluster appears smaller but relatively well-defined, suggesting that urban technology is a consistent—though not dominant—research theme.

These five subgraphs collectively show that Hong Kong's AI research landscape includes methodological foundations, emerging generative AI topics, responsible AI considerations, medical applications, and urban technology. The following section synthesizes these observations into a broader interpretive analysis.

Across the selected subgraphs, Hong Kong's AI research network appears to balance established methodological themes with significant engagement in emerging areas. Subgraph 01 reveals that ML/DL techniques remain foundational, consistent with trends observed in other regions. Subgraph 02 suggests relatively strong interest in generative AI and LLMs, with cluster cohesion indicating that these topics may be more actively explored in Hong Kong than in Taiwan or Macao. Subgraph 03 highlights notable attention to explainability, fairness, and AI ethics, representing one of Hong Kong's distinctive thematic features. Subgraph 04 demonstrates the continued presence of medical AI applications, though at a moderate scale. Subgraph 05 reflects sustained interest in smart-city and IoT-related technologies, which aligns with Hong Kong's urban context and technological initiatives.

Overall, the network presents a diversified research structure, with emerging themes—such as generative AI and responsible AI—playing a more visible role. These characteristics distinguish Hong Kong's research orientation within the cross-regional landscape.

Across Mainland China, Taiwan, Hong Kong, and Macao, the co-occurrence networks collectively suggest both shared methodological foundations and region-specific thematic orientations. All four regions maintain ML/DL techniques as the methodological core, reflecting global AI research trends. Generative AI clusters appear across all regions beginning in 2024–2025, although their density varies: Mainland China and Hong Kong show more cohesive and concentrated structures, while Taiwan and Macao display more exploratory patterns. Medical AI appears as a recurring but secondary theme in all regions except Macao, where environmental and hazard-related applications form a distinctive cluster instead. IoT-related themes are present in each region but manifest with different emphases depending on local industrial and infrastructural priorities. Considerations related to explainability and responsible AI appear most prominently in Hong Kong.

Taken together, these patterns indicate that while the four regions share methodological continuity and respond similarly to global AI developments, their thematic emphases reflect local research ecosystems, infrastructural needs, and academic traditions. This provides important context for interpreting regional differences in AI research trajectories.

4. Discussion

The purpose of this study was to describe the visible thematic structure of AI-related research in Mainland China, Taiwan, Hong Kong and Macao from 2021 to 2025. The discussion presented in this chapter focuses exclusively on the patterns that emerged from the keyword frequencies, VOSviewer visualizations and KH Coder co-occurrence networks. The interpretations that follow remain closely aligned with the empirical evidence revealed by these results.

Across all four regions, the collected datasets showed a clear concentration of high-frequency keywords related to core AI methodologies. Terms such as machine learning, deep learning and neural network appeared prominently in each region and consistently occupied central positions in the visualization results. This consistency suggests that these three terms serve as foundational descriptors of AI research across Mainland China, Taiwan, Hong Kong, and Macao during the examined period. While the intensities and densities of these keywords varied among regions, the presence of these terms across all datasets demonstrates a shared reliance on common methodological vocabulary.

In addition to the core methodological terms, each region's heatmap and co-occurrence outputs contained additional keywords that differed in visibility and distribution. The VOSviewer visualizations illustrated how certain terms appeared in concentrated areas of higher density within each region. For example, the Mainland China heatmap displayed a larger and more continuous area of dense coloration surrounding high-frequency keywords, whereas the maps for Hong Kong and Macao presented more compact clusters and smaller regions of high intensity. This difference in visual scale does not imply theoretical or practical disparities. Instead, it reflects the variations in dataset size and publication volume that naturally produce distinct visualization patterns.

When observing the patterns revealed across the four regions, it becomes evident that all regions introduced generative AI related terms during the final years of the examined period. Words such as large language model, generative artificial intelligence and ChatGPT appeared in each dataset, although the density and concentration of these terms differed. Their consistent appearance indicates that generative AI became a visible part of the AI research vocabulary in all four regions. The VOSviewer heatmaps further confirmed this by depicting small but clear regions of color intensity around these terms. This provides a descriptive indication that generative AI had entered the thematic landscape by 2024 and 2025.

The KH Coder co-occurrence networks complemented the heatmap findings by illustrating which keywords tended to appear in proximity to one another. In Mainland China and Taiwan, core methodological terms were situated at central positions within the network, whereas newer generative AI terms occupied positions slightly further outward. In Hong Kong and Macao, the networks displayed smaller clusters with shorter distances among terms. These differences reflect the structural outcomes generated by datasets of varying size. The networks also visually confirmed that no region contained entirely isolated terms, since all major keywords appeared connected to at least one other term. This provides assurance that the datasets contained coherent and interpretable structures.

Together, the VOSviewer and KH Coder outputs provide a consistent and visually coherent representation of AI-related research themes across Mainland China, Taiwan, Hong Kong, and Macao. The descriptive patterns emerging from these outputs respond directly to the research questions posed in this study. The results provide a clear overview of the themes that appeared most prominently, the ways these themes differed in visual presentation across regions, and the emergence of generative AI terminology during the final two years of the study period. The interpretations offered in this chapter remain closely aligned with the visual evidence and avoid causal or explanatory statements beyond what the data directly presents.

5. Limitations and Future Research

Despite the consistency of the observed patterns across regions, several limitations of this study should be acknowledged. First, the analysis is based exclusively on Scopus-indexed journal articles. While this data source ensures peer-review stability, standardized metadata, and cross-regional comparability, it may not fully capture the breadth of AI research activities, particularly conference-oriented publications that are common in fast-evolving AI subfields. The findings should therefore be interpreted as reflecting visible thematic patterns within the journal literature rather than the entirety of AI-related research outputs.

Second, the study relies on author-provided keywords as the primary analytical unit. Although author keywords are widely used in bibliometric research to represent explicit thematic intent, they do not encompass the full semantic content of articles. As a result, some nuanced research topics or methodological details embedded in full texts may not be fully reflected in the keyword-based visualizations and co-occurrence networks.

Future research may extend the present work in several directions. One possible extension is the integration of multiple bibliographic databases, such as Web of Science or Dimensions, to examine whether the observed thematic patterns remain consistent across different indexing systems. In addition, incorporating conference proceedings may provide a more comprehensive representation of AI research dynamics, particularly in rapidly developing subfields. Methodologically, future studies could complement keyword-based approaches with alternative text-mining or citation-based techniques, such as topic modeling or co-citation analysis, to explore deeper semantic structures and intellectual linkages within AI research. Such approaches may offer additional perspectives on thematic evolution while building upon the descriptive foundation established in this study.

6. Conclusions

This study examined AI-related publications from Mainland China, Taiwan, Hong Kong and Macao between 2021 and 2025 using keyword frequency analysis, VOSviewer heatmaps and KH Coder co-occurrence networks. The aim of the study was to describe the visible thematic patterns within each region and to compare the similarities and differences in the way AI-related terms appeared across the collected datasets. The analysis addressed three research questions that guided the scope of the study.

The findings show that machine learning, deep learning and neural network appeared as the most visible and frequently occurring terms across all four regions. These terms occupied central positions in both the heatmaps and co-occurrence networks. This indicates that they functioned as the fundamental descriptors of AI research during the examined period. While the specific distribution patterns differed among regions, the presence of these terms remained consistent.

The second major finding concerns the visual structure of keyword distribution. The VOSviewer heatmaps revealed that larger datasets, such as those from Mainland China, produced wider regions of high density, while smaller datasets, such as those from Hong Kong and Macao, generated more compact clusters. These differences reflect the underlying data volume and do not imply substantive variation. Instead, they provide a clear depiction of how dataset size influences visualization outcomes.

A third finding relates to the emergence of generative AI terminology during the final years of the period. Terms including large language model and generative artificial intelligence appeared in all four datasets by 2024 and 2025. Their appearance in both the heatmaps and co-occurrence networks shows that generative AI had entered the research vocabulary across the four regions. Although the density and distribution of these terms differed, they remained visible and thematically connected within each region's visualization.

Overall, this study provides a descriptive overview of AI-related themes across the four examined research contexts. The combined use of keyword frequency analysis, VOSviewer and KH Coder produced a coherent and visually interpretable set of findings.

These findings illustrate how these themes appeared across the four datasets and describing how the thematic landscape evolved from 2021 to 2025. Future studies may build upon this descriptive foundation by incorporating additional datasets, expanding the time frame or applying explanatory methods to investigate underlying causes of the observed patterns.

Author Contributions

Conceptualization, C.C. and S.-M.W.; methodology, C.C. and S.-M.W.; software, C.C.; validation, C.C. and S.-M.W.; formal analysis, C.C.; investigation, C.C.; resources, C.C.; data curation, C.C.; writing—original draft preparation, C.C.; writing—review and editing, C.C. and S.-M.W.; visualization, C.C.; supervision, S.-M.W.; project administration, C.C. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable, as this study analyzed publicly available bibliographic data and did not involve humans or animals.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data used in this study were retrieved from the Scopus database under the authors' institutional access. Due to licensing restrictions, the raw exported Scopus records cannot be publicly shared. The search strategy and the processed data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
2. Xie, W.; Hu, A.; Xie, Q.; et al. Bibliometric analysis and review of AI-based video generation: research dynamics and application trends (2020–2025). *Inf. Retr. J.* **2025**, *28*, 130. [CrossRef]
3. Van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef]
4. Donthu, N.; Kumar, S.; Mukherjee, D.; et al. How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* **2021**, *133*, 285–296. [CrossRef]
5. Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 4th ed.; Pearson: Hoboken, NJ, USA, 2021; pp. 1–30.
6. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [CrossRef]
7. Bommasani, R.; Hudson, D.A.; Adeli, E.; et al. On the opportunities and risks of foundation models. *arXiv preprint* **2021**, *arXiv:2108.07258*. [CrossRef]
8. Bubeck, S.; Chandrasekaran, V.; Eldan, R.; et al. Sparks of Artificial General Intelligence: Early Experiments with GPT-4. *arXiv preprint* **2023**, *arXiv:2303.12712*. [CrossRef]
9. Price, D.J.D. *Little Science, Big Science*; Columbia University Press: New York, NY, USA, 1963; pp. 1–119.
10. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [CrossRef]
11. Aria, M.; Cuccurullo, C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975. [CrossRef]
12. Callon, M.; Courtial, J.-P.; Turner, W.A.; et al. From translations to problematic networks: An introduction to co-word analysis. *Soc. Sci. Inf.* **1983**, *22*, 191–235. [CrossRef]
13. Zupic, I.; Čater, T. Bibliometric methods in management and organization. *Organ. Res. Methods* **2015**, *18*, 429–472. [CrossRef]
14. Agbo, F.J.; Oyelere, S.S.; Suhonen, J.; et al. Scientific production and thematic breakthroughs in smart learning environments: A bibliometric analysis. *Smart Learn. Environ.* **2021**, *8*, 1. [CrossRef]
15. De la Vega Hernández, I.M.; Serrano Urdaneta, A.; Carayannis, E. Global bibliometric mapping of the frontier of knowledge in the field of artificial intelligence for the period 1990–2019. *Artif. Intell. Rev.* **2023**, *56*, 1699–1729. [CrossRef]
16. KH Coder. Available online: https://khcoder.net/en/ (accessed on 26 November 2025).
17. Kakouris, A.; Georgiadis, P. Analysing entrepreneurship education: A bibliometric survey pattern. *J. Glob. Entrep. Res.* **2016**, *6*, 6.
18. Oe, H.; Weeks, M. Kawaii culture in Japan: A bibliometric analysis and text mining approach based on pop-cultural diplomacy and transmission into global values. *BIRCI J. Humanit.* **2020**, *3*, 3990–4004. [CrossRef]
19. Sakai, K.; Nagata, T.; Mori, T.; et al. Research topics in occupational medicine, 1990–2022: A text-mining-applied bibliometric study. *Scand. J. Work Environ. Health* **2024**, *50*, 1–12. [CrossRef]
20. Nagasue, M.; Kitagawa, H.; Asawa, T.; et al. A systematic review of passive cooling methods in hot and humid climates using a text mining-based bibliometric approach. *Sustainability* **2024**, *16*, 1420. [CrossRef]
21. Deleca, C.; Nica, J.; Ionescu, S.; et al. Mapping the frontier: A bibliometric analysis of artificial intelligence applications in local and regional studies. *Algorithms* **2024**, *17*, 418. [CrossRef]
22. Ekundayo, T.; Chaudhry, S.A. Mapping the future: A bibliometric analysis of engagement trends in artificial intelligence within higher education. *Digit. Technol. Res. Appl.* **2025**, *4*, 1285. [CrossRef]

23. Visser, M.; van Eck, N.J.; Waltman, L. Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quant. Sci. Stud.* **2021**, *2*, 20–41. [CrossRef]
24. Martín-Martín, A.; Orduna-Malea, E.; Thelwall, M.; et al. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *J. Informetr.* **2018**, *12*, 1160–1177. [CrossRef]
25. VOSviewer. VOSviewer—Visualizing Scientific Landscapes. Available online: https://www.vosviewer.com/ (accessed on 9 November 2025).