applied sciences

Article

Criteria Analysis for the Selection of a Generative Artificial Intelligence Tool for Academic Research Based on an Improved Group DEMATEL Method

Constanta Zoie Radulescu [1,*] and Marius Radulescu [2]

[1] National Institute for Research and Development in Informatics, 011455 Bucharest, Romania
[2] "Gheorghe Mihoc-Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy, 050711 Bucharest, Romania; mradulescu.csmro@yahoo.com
* Correspondence: radulescucz@yahoo.com

Abstract: Generative Artificial Intelligence (GenAI) tools are transforming academic research by significantly enhancing efficiency, accuracy, and productivity. However, selecting the most appropriate GenAI tool requires careful evaluation of multiple, interdependent criteria. This paper makes two main contributions. First, it introduces IDEMATEL, an improved decision-making method that advances beyond the traditional DEMATEL approach. Unlike DEMATEL, which can encounter technical limitations when analyzing complex relationships, IDEMATEL ensures robust and reliable results by guaranteeing the necessary mathematical conditions for analysis in all cases. This enhancement makes IDEMATEL more broadly applicable and dependable for evaluating interrelated criteria. Second, the paper demonstrates the practical value of IDEMATEL by applying it to the selection of GenAI tools for academic research. Using this method, a comprehensive set of criteria—including functionality, ease of use, cost, data security, and community support—is systematically analyzed. The results provide researchers and decision-makers with clearer insights into how these factors interact and influence the selection process. By leveraging IDEMATEL, stakeholders can make more informed and confident choices, ensuring that the selected GenAI tools best meet the diverse needs of academic research.

Keywords: generative AI tools; academic research; multi-criteria evaluation; DEMATEL method; convergence to zero; weights

Academic Editors: António Correia, Rui Araújo and Vincent A. Cicirello

Received: 9 April 2025
Revised: 7 May 2025
Accepted: 9 May 2025
Published: 12 May 2025

Citation: Radulescu, C.Z.; Radulescu, M. Criteria Analysis for the Selection of a Generative Artificial Intelligence Tool for Academic Research based on an Improved Group DEMATEL Method. *Appl. Sci.* **2025**, *15*, 5416. https://doi.org/10.3390/app15105416

1. Introduction

While both general AI tools and GenAI tools support academic research, they differ in functionality. AI tools typically assist in organizing, retrieving, and analyzing information—such as finding related studies or visualizing citation networks—whereas GenAI tools go further by generating new content such as text, code, or images [1-4]. For instance, Semantic Scholar (Allen Institute for AI, Seattle, WA, USA) and Research Rabbit (ResearchRabbit Inc., Austin, TX, USA) help researchers discover and connect ideas, while (OpenAI, San Francisco, CA, USA), Jenni AI (Jenni AI, San Francisco, CA, USA) and Paperpal (CACTUS Communications, Mumbai, India) assist in creating academic writing. Tools like Perplexity AI (Perplexity AI, San Francisco, CA, USA) and Consensus (Consensus, Boston, MA, USA) synthesize content based on user queries. Although AI tools and GenAI tools serve different purposes in academic research, the line between them is often blurred, as they can complement each other.

To maximize the benefits of these AI tools, it is essential to establish clear selection criteria (factors). Researchers should consider criteria such as functionality, ease of use, cost, security, community support, and reputation that can have different levels of importance and can influence each other. These criteria can be used in a selection of GenAI tools based on a multi-criteria method.

While various multi-criteria and weighting methods can be applied to evaluate, selection and adoption of GenAI tools, there is a lack of studies focused on GenAI tools for academic research that consider causal relationships between evaluation criteria.

The Decision-Making Trial and Evaluation Laboratory (DEMATEL) method is a method capable of modeling causal relationships among criteria for analysis of criteria used in selection of a GenAI tool for academic research [5].

It identifies cause-and-effect relationships between criteria, facilitating better prioritization and decision-making. By visualizing these relationships, DEMATEL helps decision-makers understand how criteria affect the decision process. Using a structured framework based on DEMATEL ensures that the selection process is systematic and accurate.

The gap in analysis of causal relationships between evaluation criteria emphasizes the need for more sophisticated decision-support methods that model the influence and interaction between criteria. This paper aims to fill this gap by proposing an Improved Group DEMATEL method called IDEMATEL, that models the causal relationships between the criteria involved in the selection of GenAI tools for academic research. IDEMATEL offers a systematic approach to identify cause-and-effect relationships, with the advantage of ensuring that the successive powers of the initial direct relation matrix converge to zero, allowing for the computation of the total influence matrix. The paper provides theoretical foundations for evaluating whether the powers of a square matrix converge to zero using spectral radius and some matrix norms. The method constructs a causality diagram and digraph to visually represent the relationships among criteria. An application of the IDEMATEL method that involves the evaluation the criteria used to select a suitable GenAI tool for academic research: functionality and objectives, ease of use and accessibility, scalability and integration, data quality and security, cost and value, support and community, and reputation and reviews are developed and analyzed.

In summary, this paper provides

(1) A review of evaluation criteria relevant to the selection of GenAI tools for academic research.
(2) A IDEMATEL method that enhances the classical DEMATEL approach.
(3) An application of the IDEMATEL method to analyze interdependencies among criteria in selecting a GenAI tool for academic research.

By providing both theoretical contributions and practical insights, this study aims to support researchers and institutions in making informed, structured, and justifiable decisions when selecting GenAI tools for academic use.

The remainder of the paper is structured as follows. Section 2 presents recent research in the selection, adoption, and implementation of GenAI tools based on multi-criteria methods. A review of the criteria that should be considered when selecting GenAI tools for academic research is realized. Section 3 argues the reasons for selecting the DEMATEL method for evaluation and analysis of the criteria. The IDEMATEL method and the algorithm are proposed in Sections 4 and 5. An application of the proposed method for evaluation and analysis of criteria used in the selection of GenAI tools for academic research is presented in Section 6. Conclusions are given in Section 7.

2. Research on Evaluation and Analysis of Criteria for GenAI Tools Based on Multi-Criteria Methods

With the proliferation of GenAI tools, selecting the right one suitable for academic research can be a challenging task. It is important to evaluate these tools based on specific criteria to ensure they meet the needs of researchers and contribute effectively to the research process. In the specialized literature, this problem has been approached from several points of view and has been solved by various methods.

The papers [6-9] approach the topic from a conceptual or descriptive perspective, outlining general criteria for selecting or adopting the AI tools for academic research without employing a multi-criteria decision-making framework. These works do not propose weighting methods for establishing relationships between criteria.

The paper [8] outlines the general criteria for selecting artificial intelligence tools, with a particular focus on those used in publishing. They emphasize that researchers should consider ethical standards, regulatory compliance, and societal impact when choosing AI tools. The paper [6] provides an overview of AI tools applicable to academic research, offering guidance to educators on effectively integrating these technologies.

The paper [7] examines the role of AI in research, identifying two major limitations: the exclusion of commercial AI tools and the narrow focus on those designed for direct research queries. The study highlights the need to explore how paid AI tools and multi-system integrations can enhance research outcomes. Future research should assess the effectiveness of commercial AI, the impact of combining multiple AI tools, and the variability of AI-generated responses. These findings underscore the importance of critically selecting and evaluating AI tools for academic research.

The paper [9] is about the adoption of ChatGPT tool for education. The factors for adoption of ChatGPT include hedonic motivation, usability, perceived benefits, system responsiveness and relative advantage, social influence, facilitating conditions, privacy, and security. This paper does not use a weighting method.

2.1. Comparative Analysis of the Use of Multi-Criteria Decision-Making (MCDM) Methods in the Evaluation of GenAI Tools

A significant number of articles utilize the Analytic Hierarchy Process (AHP) to assign weights to various criteria. For example, studies such as [10–15] apply AHP (sometimes combined with Delphi) to assess tools based on qualitative, ethical, or user-experience criteria. However, these studies do not explore causal or dependency relationships between the criteria. Although the BWM or SWARA methods, which are more recent methods, have a much smaller number of pairwise comparisons, the AHP method is preferred.

The chapter [10] evaluates AI tools in universities by first reviewing relevant literature and then applying Fuzzy Multi-Criteria Decision-Making to prioritize them. Five key criteria: perceived ease of use, perceived usefulness, personalization, interaction, and trust, were weighted based on expert opinions and literature analysis. This study applies an intuitionistic fuzzy approach to enhance evaluation accuracy and prioritize AI tools effectively. This paper uses an AHP prioritization method but does not establish causal relationships between criteria.

The paper [11] examines whether restrictions or legislation should control GenAI, using ChatGPT as a case study. A systematic literature review and Analytic Hierarchy Process (AHP) were used to analyze 10 key ethical concerns, with expert panel evaluations. The criteria considered are copyright, legal, and compliance issues; privacy and confidentiality; academic integrity; incorrect citation practices; and safety/security concerns. The analysis showed that restriction was favored slightly over legislation. This paper only considers criteria related to restrictions or legislation that should control GenAI tools. No relationships are established between criteria. In the case study, only ChatGPT is analyzed.

The paper [12] develops a quality assessment framework by integrating the Delphi method and Analytic Hierarchy Process (AHP) for AI-generated digital educational resources. A systematic literature review identifies initial quality indicators across four key dimensions: content, expression, user aspects, and technical aspects. Expert feedback, gathered through two rounds of Delphi surveys, refines these indicators, while AHP assigns weight coefficients to determine their significance. The criteria considered are as follows: Content characteristics (Authenticity, accuracy, specifications, relevance, novelty, diversity, timeliness). Expression characteristics (Legitimacy, knowledgeability, logicality, comprehensible). User characteristics (Achievement, acquisition, compatibility, friendliness). Technical characteristics (Conciseness, stability, human analogy, security, big data). Findings highlight content quality as the most critical factor, followed by expression, with user and technical aspects also playing a role. Among second-level indicators, authenticity, accuracy, legitimacy, and relevance are prioritized. This paper only analyzes qualitative criteria. It does not analyze the relationships between the criteria.

The research [13] identifies and categorizes ethical concerns associated with ChatGPT. The criteria used in AHP method are as follows: Risk (infodemics and misinformation, bias response, plagiarism, privacy and confidentiality, academic integrity concern, risk hallucination-manipulation and mislead, safety and security concern). Reward (question answering, dissemination and diffusion of new information, streamlining the workflow, personalized learning, decrease teaching workflow, idea and text generation and summarization, increase productivity and efficiency). Resilience (appropriate testing framework, acceptable usage in science, co-creation between humans and ai, academic integrity policies, solidity ethical values, transform educative systems, higher-level reasoning skills). In this paper, only ethical criteria associated with ChatGPT are analyzed. The relationships between the criteria are not analyzed.

The study [14] identifies critical decision points for AI integration through a systematic literature review and prioritizes them using the Analytic Hierarchy Process (AHP). The results highlight adaptive organization structure, specialized AI teams, ethics oversight, governance, and innovation infrastructure as key factors. A sensitivity analysis confirms the robustness of the rankings, showing minimal impact from criteria weight changes. The study contributes to the discussion on workforce transformation in the AI era, providing insights for organizations navigating AI adoption. The paper addresses the issue of AI integration not GenAI selection. It does not analyze the relationships between the criteria. It is not oriented towards academic research.

The purpose of the study [15] is to comprehend and address the barriers which are impeding the implementation of GenAI Technologies, such as ChatGPT in the educational landscape. The criteria considered are Ethical concerns (risk of academic integrity, risk of biased outcomes, accessibility and inclusivity). Technological concerns (digital literacy disparity, technological infrastructure, limited up-to-date knowledge, upfront cost, integration with existing systems). Regulatory concerns (data privacy, copyright issues (ownership and licensing), cybersecurity issues). Societal concerns (widening digital divide, economic disparities, job displacement and skills gap, cultural sensitivity, social injustice and inequality, cognitive and emotional development, commercialization, use for malicious purpose). Trust issues (lack of transparency, explainability, accountability). Human values concerns (erosion of critical thinking, lack of focus on holistic development, lack of personal and emotional interaction). Psychological concerns (anxiety about technological competence, cognitive overload, weak self-efficacy, social isolation, resistance to change). The paper addresses the problem of GenAI implementation, not GenAI selection. AHP is used. The relationships between criteria are not analyzed.

Papers like [16,17] propose hybrid decision-making frameworks such as fuzzy TOPSIS or AHP + CoCoSo in commercial or industrial contexts (chatbot selection or telecom), but again, they do not target the academic research domain nor analyze interdependencies among evaluation criteria.

In the paper [16], a fuzzy multi-attribute decision-making framework to assist in selecting GenAI chatbots based on key features and performance is presented. It introduces a modified TOPSIS method adapted for an interval-valued hesitant Fermatean fuzzy (IVHFF) environment, improving the handling of uncertainty and hesitation in decision-making. The framework evaluates GenAI chatbots both statically and dynamically using four criteria: conversational ability, user experience, integration capability, and price. A case study comparing ChatGPT, Copilot (Microsoft Copilot, Redmond, WA, USA), Gemini (Google DeepMind, London, UK), Claude (Anthropic, San Francisco, CA, USA), and Perplexity demonstrates its effectiveness. Recommendations are provided for selecting and implementing chatbots in various applications. This paper does not analyze the relationships between criteria. It is not oriented towards academic research. GenAI chatbots are selected.

The paper [17] introduces a decision-making framework for selecting chatbots in the telecommunication industry, leveraging AI advancements to enhance customer support. The proposed method combines the Analytic Hierarchy Process (AHP) for attribute weighting and the Combined Compromise Solution (CoCoSo) for ranking chatbots within a single-valued neutrosophic set (SVNS) environment, effectively handling uncertainty. A case study validates the model's efficiency and applicability, with sensitivity analysis confirming its robustness. The findings suggest that the AHP-CoCoSo approach yields more realistic chatbot selection outcomes under uncertain conditions, offering a practical tool for telecom decision-makers. In this paper, the problem is about selecting chatbots in the telecommunication industry. The relationships between criteria are not analyzed. It is not oriented towards academic research.

Only a few studies [18–20] employ advanced methods such as DEMATEL, F-DEMATEL, or grey-DEMATEL to analyze causal relationships between criteria. Nevertheless, these are not focused on the selection of GenAI tools and not on academic research. Instead, they address issues related to the integration of AI in government services, energy management, and general barriers to GenAI adoption.

The study [18] explores the integration of ChatGPT-type models into government services, highlighting both their potential to improve efficiency and the risks they pose. To understand public acceptance, the researchers used Latent Dirichlet Allocation (LDA) to identify 15 key influencing factors and applied grey-DEMATEL and TAISM to analyze their causal relationships. The criteria considered are as follows: Data layer (Data capacity, Data quality). Technology layer (own technology, technology maturity, technology ethics). User layer (information literacy, perceived risk, trust). Service layer (meeting demands, perceived ease of use, relative advantage),. environment layer (old and new parallel, oversight and accountability, security guarantee, public officials' literacy). This paper addresses the problem of integration of ChatGPT-type models, not of the selection of a GenAI. It is not oriented towards academic research.

The paper [19] explores the role of GenAI in enhancing energy efficiency amid growing demand for sustainable energy solutions. Addressing concerns about uncontrolled use and ethical implications, it introduces an innovative Energy GenAI Technology Framework (EnGenAI) that integrates GenAI, energy efficiency, Business Ethics (BE), and Corporate Social Responsibility (CSR) principles. The methodology includes a scoping review, multi-criteria analysis, and the development of the framework. The paper uses the F-DEMATEL method to analyze their causal relationships. It is not oriented towards academic research.

The study [20] investigates barriers to GenAI adoption through two studies using a mixed-method approach. The first analyzes YouTube datasets using text mining, revealing key challenges like trust, anticipation, and surprise, along with barriers such as ethical, technological, and cost issues. The second study, based on an extensive literature review, identifies barriers like privacy, return on investment, and lack of infrastructure. The findings from both studies are compared, confirming the value of the mixed-method approach. The paper uses the F-DEMATEL and FAHP method to analyze their causal relationships. It is not oriented towards academic research.

In summary, these studies are presented in Table 1.

Table 1. Research about criteria and methods in evaluating the AI tools.

References	Problem	Multi-Criteria Method	Criteria	Domain
[10]	Evaluate AI tools that can be used in universities	Intuitionistic Fuzzy Multi-Criteria Decision-Making	Perceived ease of use, perceived usefulness, personalization, interaction, trust	Education
[11]	Whether to impose restrictions or legislate in the usage of GenAI	Analytic Hierarchy Process (AHP)	Copyright, legal, and compliance issues; privacy and confidentiality; academic integrity; incorrect reference and citation practices; safety and security concerns	Education
[12]	An evaluation index system for AI-generated digital educational resources	A combination of the Delphi method and AHP	Content characteristics, expression characteristics, user characteristics, technical characteristics	Education
[13]	Identifies and categorizes ethical concerns associated with ChatGPT	AHP	Risk, reward, resilience	Education
[14]	Selection of generative artificial intelligence GenAI chatbots	Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)	Conversational ability, user experience, integration capability, price	Various applications
[15]	A decision-making framework for chatbot selection in the telecommunication industry	Combined Compromise Solution (CoCoSo) and AHP single-valued neutrosophic sets	Security, speed, responsiveness, satisfaction, reliability, assurance, tangibility, engagement, and empathy	Telecommu nication industry
[16]	Identifying critical decision points and alternatives in integrating generative AI through a systematic literature review	AHP	Innovativeness, productivity enhancement from investment made, change in customer experience, data safety and ethics, organizational adaptability	Organizatio ns
[17]	The barriers which are impeding the implementation of Generative AI Technologies, such as ChatGPT in the educational landscape	Fuzzy AHP	Ethical concerns, technological concerns, regulatory concerns, societal concerns, trust issues, human values concerns, psychological concerns	Education
[18]	Integrating the Chat Generative Pre-Trained	LDA Grey-DEMATEL method and TAISM	Data layer, technology layer, user layer, service layer, environment layer	Government services

	Transformer-type (ChatGPT-type) model with government services			
[19]	GenAI a critical solution for improving energy efficiency amid increasing demand for sustainable energy at the operational and supply chain management levels	Fuzzy DEMATEL	Integrity, transparency and accountability, fairness and bias mitigation, privacy and data protection, compliance with laws and regulations, sustainability and environmental responsibility, employment and workforce	Energy and business
[20]	Adoption of GenAI tools	F-DEMATEL and FAHP	Ethical, technological, regulations and policies, cost, and human resources	Business, service organizations

The reviewed literature highlights a growing interest in the evaluation, selection, adoption and implementation of GenAI tools based on multi-criteria methods. However, the adoption of structured decision-making methodologies, especially those capable of managing complex criteria relationships, remains limited.

2.2. Criteria for Selecting GenAI Tools

Some criteria for selecting GenAI tools are:

Methodological and Ethical Considerations: The integration of AI tools in academic research requires careful consideration of methodological and ethical implications. This includes addressing issues related to authorship, accountability, and the role of human researchers [11,13,17,19,21,22]. The European Union, for example, has established frameworks to guide the responsible development and deployment of AI systems [8].

Functionality: The AI tools should be relevant to the research objectives and domain. AI tools should enhance research efficiency, accuracy, and quality across various stages of the research process, such as literature review, data analysis, and writing [22,23]. They should also support diverse research methods, including quantitative and qualitative data analysis [23]. The selection of AI tools should be based on their performance and suitability for specific research tasks [6,23-25]. AI tools designed for literature review must efficiently identify relevant sources [7].

Transparency and Accountability: It is important to ensure transparency in the use of AI tools, emphasizing the importance of human intelligence and critical thinking in the research process. This helps maintain the authenticity and credibility of academic work [17–19,21,26].

Ease of Use: The tools should have an intuitive interface, should be user-friendly, allowing researchers without technical skills to utilize them effectively [10,23,27]. This enhances accessibility, especially for those without advanced technical skills [3] tools with quick responses and user-friendly interfaces are preferred [27].

Adaptability and Innovation: AI tools should be adaptable to various stages of the research process, from literature review to data analysis and writing. They should also foster innovation while balancing ethical considerations [16,22,28].

Scalability: AI tools must be able to handle increasing amounts of data and complexity as research projects grow. This ensures that the tools remain useful and relevant as research demands evolve [14,16,22].

Data Quality and Security: High-quality data input and output are critical for reliable research outcomes. AI tools should be capable of processing complex data accurately, which is particularly important in fields like healthcare and pharmaceutical research [29].

Consider the privacy and security features of the AI tool, especially when handling sensitive data [11,12,17,18,30]. Verify that the AI tool provides accurate information based on credible sources. Ensure that the tool adheres to data security standards and ethical research practices. Data quality is important for systematic reviews [22].

Cost and Value: The cost of AI tools should be justified by the value they provide in terms of improved research outcomes and efficiency. This includes considering both initial investment and ongoing operational costs [14,20,22]. It is important to assess whether the AI tools are financially accessible to researchers. Availability of free versions, trials, or institutional licenses can be a deciding factor.

Ability to Respond to Complex Queries: The tool should efficiently respond to complex questions and offer advanced conversational support [14]. Consider both free and commercial options; the tool's value should justify the investment [7].

Support and Community: A strong support network and active user community can enhance the usability and troubleshooting of AI tools. This includes access to documentation, forums, and customer support [15,23]. Support and training are vital for responsible AI use, enabling researchers to maximize benefits and minimize risks [30].

Integration with Other Tools: Integration with other tools helps streamline the research workflow by automating data transfer and reducing manual effort [14,17]. Evaluate the ability of AI tools to integrate with other AI academic tools. Evaluate the AI tool's ability to integrate with other software programs and databases used in research.

Reputation and Reviews: The reputation of AI tools, as reflected in user reviews and academic literature, can provide insights into their reliability and effectiveness. Tools with positive feedback and proven track records are preferable [22,23]. Positive reviews and a strong reputation can indicate that the tool is well regarded and likely to deliver on its promises.

Selecting a GenAI tool for academic research involves a thoughtful evaluation process that considers several criteria to ensure the chosen GenAI tool effectively supports research activities such as data analysis, idea generation, and information synthesis. The GenAI tool should align with research objectives and be financially accessible with free versions or trials available. It must provide accurate information based on credible sources and efficiently respond to complex questions while adhering to data security standards and ethical research practices. Transparency in how it generates information and makes decisions is important, as well as the ability to avoid biases and integrate with other academic tools. Regular updates and feedback from the academic community are also essential in evaluating the tool's effectiveness and utility in practice.

Selecting GenAI tools for academic research requires careful consideration of criteria such as accuracy, transparency, ethical compliance, adaptability to specific research needs, user-friendliness, cost-effectiveness, and access to training resources. When selecting GenAI tools for academic research, certain criteria are more important due to their direct impact on research quality, efficiency, and ethical compliance. Accuracy and reliability are essential to ensure that outputs align with credible sources, preventing misinformation and maintaining the integrity of scholarly work. Transparency in data sources and decision-making processes is important for replicability and trustworthiness in research outcomes. Ethical compliance must be prioritized to address biases, privacy concerns, and authorship accountability, ensuring adherence to academic standards. User-friendliness enhances accessibility for researchers with varying technical expertise, while cost-effectiveness ensures equitable access to both free and paid solutions. Support and training resources are vital for responsible AI use, enabling researchers to maximize tool benefits while minimizing risks. When choosing a GenAI tool for academic research, it is important to consider criteria such as functionality and objectives, ease of use and accessibility, scalability and integration, data quality and security, support and community, and reputation and reviews. These criteria are interconnected and collectively influence the effectiveness of AI tools in research.

Some criteria hold more weight depending on the context. For instance, data quality is paramount for systematic reviews, while ease of use may be more critical for undergraduate researchers. The links between these criteria—such as how functionality influences data quality or how ease of use affects accessibility—highlight the need for a holistic evaluation when selecting GenAI tools.

An analysis of the links between these criteria reveals their interdependence.

3. The Selection of DEMATEL Method for GenAI Tools Criteria Analysis

The aim of this section is to show the reasons for selecting the DEMATEL method for criteria analysis of GenAI tools. Subjective weighting methods in multi-criteria decision-making (MCDM) are essential for determining the relative importance of criteria based on decision-makers' preferences. Popular methods are the Analytic Hierarchy Process (AHP) [31,32] Analytic Network Process (ANP) [33], Decision-Making Trial and Evaluation Laboratory (DEMATEL) [5], and the Best Worst Method (BWM) [34]. A summary of the criteria weighting methods is presented in [35].

DEMATEL considers interdependence and causal relationships, offering both short-term and long-term improvement perspectives when it is integrated with AHP [36].

The paper [37] compares and proposes hybrid models combining AHP, BWM, MACBETH, and DEA for multi-criteria decision-making. The paper [38] compares the results of AHP, limited AHP, and BWM, finding that BWM's results are comparable to standard AHP while limited AHP is generally inferior. The choice of method and number of comparisons significantly influences the resulting priority vector [38]. A comparison of AHP and BWM is presented in the paper [39]. The strategy of the paper [40] is to compare ANP with DEMATEL-ANP for the selection of a location.

A comparison of subjective weighting methods—AHP, ANP, DEMATEL and BWM—for core mechanism, features, and strengths and weaknesses is presented in Table 2.

Table 2. A comparison of a set of subjective weighting multi-criteria methods from core mechanism, features, and strengths and weaknesses.

Method	Core Mechanism	Features	Strengths	Weaknesses
AHP	Hierarchical pairwise comparisons using Saaty's 1–9 scale to derive weights.	- Handles qualitative/quantitative criteria - Widely validated - Robust consistency checks	- Relatively intuitive and widely used	- Exponential growth of comparisons - Rank reversal issues - Can be subjective
ANP	Extends AHP to account for interdependencies among criteria using feedback loops.	- Models complex interrelations - Flexible for networked systems	- Models complex dependencies	- Computationally intensive - Requires expertise in network modeling - Complex to implement and interpret
DEMATEL	Builds causal relationships between criteria using influence matrices.	- Identifies cause–effect relationships - Visualizes system dynamics	- Visualizes complex relationships - Can identify key drivers	- Subjective causality assessments - Limited to medium-sized systems - Can be difficult to validate

BWM	Compares all criteria against predefined "best" and "worst" anchors.	- Requires fewer comparisons - Ensures high consistency	- More consistent than AHP - Requires fewer comparisons than AHP	- Sensitive to outlier judgments - Limited flexibility for adding new criteria - Can be more abstract for decision-makers

These methods differ in scale used, the number of comparations required and the dependency consideration of criteria. The comparison is presented in Table 3.

Table 3. Comparison of a set of subjective weighting multi-criteria methods for number of comparations, criteria dependency, and scale used.

Method	Number of Comparisons Required	Criteria Dependency Consideration	Scale Used
AHP	$n(n-1)/2$	None	Saaty's 1–9 Scale
ANP	$n(n-1)/2$	Interdependencies	Saaty's 1–9 Scale (or extensions)
DEMATEL	$n(n-1)$	Causal Relationships	Integer Scale (e.g., 0–4 or 0–5)
BWM	$2n-3$	None	1–9 Scale

The selection of DEMATEL (Decision-Making Trial and Evaluation Laboratory) method over other multi-criteria weighting methods like AHP, ANP and BWM is determined in principle by the causal relationship analysis, visualization of complex systems, and flexibility of application. DEMATEL excels in analyzing causal relationships between criteria, helping to identify which factors have the most influence on others. This method generates visual representations (cause–effect diagrams) that clearly illustrate the relationships between criteria. This helps decision-makers understand the dynamics of the system and prioritize critical factors more intuitively. While AHP is good for hierarchical structures and consistency checks, it does not handle interdependencies well. ANP accounts for interdependencies but is more complex and computationally intensive compared to DEMATEL. BWM reduces the number of comparisons but does not explore causal relationships between criteria. DEMATEL stands out because it focuses on analyzing causal relationships and interdependencies among criteria, offering a holistic approach to decision-making.

The DEMATEL method is a valuable tool for analyzing complex systems and decision-making processes. It has been applied to prioritize factors affecting AI adoption in banking [41] and identify challenges in implementing AI in public manufacturing sectors [42]. The DEMATEL method has been successfully used in various fields, including waste recycling analysis, project selection, and e-learning program evaluation [43].

In the case of the addressed problem of prioritizing and analyzing the importance of the evaluation criteria used in choosing a GenAI tool for academic research, the choice of the DEMATEL method is justified by various criteria such as performance, cost, user interface, and integration capabilities that are often interconnected.

4. The IDEMATEL Method

The DEMATEL (Decision-Making Trial and Evaluation Laboratory) method, was created at the Battelle Memorial Institute in Geneva by Gabus and Fontela. This systematic approach was designed to analyze complex decision-making issues. It involves creating a directed graph (digraph) to illustrate the causal relationships among the criteria involved in a decision-making problem. In this graph, the criteria are represented as nodes, while the directed edges depict the causal connections between them.

DEMATEL combines expert knowledge with mathematical techniques to develop a cause-and-effect matrix, allowing for the classification of criteria into cause-and-effect groups. This method identifies the most influential criteria within the decision-making framework, which are subsequently prioritized for further analysis.

The traditional DEMATEL method involves calculating the total influence matrix by summing the powers of an initial direct relation matrix, assuming that these powers converge to zero. However, research made in ref. [44] revealed that this convergence does not always occur, which can invalidate the calculation of the total influence matrix. See also [45]. To address this issue, in ref. [44], a Revised DEMATEL method is proposed to ensure convergence and enhance result reliability. However, the normalized direct-relation matrix proposed in ref. [44] is not equal to the corresponding matrix from the traditional DEMATEL method.

The IDEMATEL method, proposed in this paper, provides very general conditions on the direct-relation matrix such that the sequence of powers of the normalized direct-relation matrix from the traditional DEMATEL method is convergent to zero. Note that the above-mentioned convergence to zero allows for the computation of the total influence matrix.

In Appendix A, we shall prove a theorem that will help us to give conditions for the convergence to zero of powers of a square matrix. This theorem will be used for the proposal of a new method, called IDEMATEL, that yields result identical to those of the original DEMATEL method for a large class of direct-relation matrices.

5. The IDEMATEL Algorithm

The IDEMATEL method will be used in the analysis of criteria used in the selection of a GenAI tool for academic research. The algorithm of this method is presented in the following. The algorithm aims to analyze and visualize the causal relationships between a set of criteria. It combines expert opinions with mathematical calculations to create a causal diagram.

Step 1. Form a group of experts $E = \{E_1, E_2, \dots, E_p\}$ who possess knowledge and experience relevant to the criteria being evaluated.

Step 2. Define a comprehensive set of criteria (factors) $C = \{C_1, C_2, \dots, C_n\}$ that will be analyzed for their causal relationships.

Step 3. Each expert E_k evaluates the influence of each criterion C_i on every other criterion C_j and generates the initial direct-relation matrix $\mathbf{D}^{(k)} = \left(d_{ij}^{(k)}\right); i, j = 1,2, \dots, n; k = 1,2, \dots, p$. The entry $d_{ij}^{(k)}$ represents expert k's assessment of the influence of criterion C_i on criterion C_j. The principal diagonal entries of matrix $\mathbf{D}^{(k)}$ are equal to zero: $d_{ii}^{(k)} = 0$. The assessments are made using a predefined DEMATEL scale: 0: No influence, 1: Very low influence, 2: Low influence, 3: High influence and 4: Very high influence.

Step 4. Combine the individual direct-relation matrices from all experts into an aggregate direct-relation matrix $\mathbf{D} = \left(d_{ij}\right)$. This is achieved by averaging the corresponding elements $\mathbf{D}^{(k)}$ across all expert matrices:

$$d_{ij} = \frac{1}{p} \sum_{k=1}^{p} d_{ij}^{(k)}. \tag{1}$$

Step 5. Normalize the aggregate direct-relation matrix $\mathbf{D}$ to create a normalized direct-relation matrix $\bar{\mathbf{D}} = \left(\bar{d}_{ij}\right)$. Denote $Q = \{z \in \mathbb{C}: |z| = 1\}$. Let p and s be defined as follows:

$$p = \max\left(\max_{1 \le i \le n} \left(\sum_{j=1}^{n} d_{ij}\right); \max_{1 \le j \le n} \left(\sum_{i=1}^{n} d_{ij}\right) \right), \tag{2}$$

$$s = \begin{cases} p & \text{if } \det\left(z\mathbf{I} - \frac{1}{p}\mathbf{D}\right) \neq 0 \text{ for every } z \in Q \\ p + \varepsilon & \text{if } \det\left(z\mathbf{I} - \frac{1}{p}\mathbf{D}\right) = 0 \text{ for some } z \in Q \end{cases} \tag{3}$$

where ε is a small positive number and $\mathbf{I}$ is the identity matrix.

Build the $n \times n$ normalized initial direct-relation matrix $\bar{\mathbf{D}} = \left(\bar{d}_{ij}\right)$ as

$$\bar{d}_{ij} = \frac{d_{ij}}{s} \text{ for all } i, j \tag{4}$$

In matrix $\bar{\mathbf{D}}$, all principal diagonal elements are equal to zero.

Some details for computing parameter s will be given in the following.

Consider the characteristic polynomial of the matrix $\frac{1}{p}\mathbf{D}$:

$$f(z) = \det\left(z\mathbf{I} - \tfrac{1}{p}\mathbf{D}\right), z \in \mathbb{C}. \tag{5}$$

Note that f is a polynomial function with real coefficients of degree n, that is

$$f(z) = \sum_{q=0}^{n} a_q z^q, \tag{6}$$

with $a_q \in \mathbb{R}$. The largest eigenvalue of the matrix $\frac{1}{p}\mathbf{D}$ is the spectral radius of this matrix. It is equal to the greatest absolute value of the roots of polynomial f. One can easily see that $\rho\left(\frac{1}{p}\mathbf{D}\right) < 1$ if and only if polynomial f has no roots on Q.

In order to see that f has a root in Q we shall check if the absolute value of f vanishes on Q. This is equivalent to the fact that the function

$$g(t) = \left(\sum_{q=0}^{n} a_q \cos(qt)\right)^2 + \left(\sum_{q=0}^{n} a_q \sin(qt)\right)^2, t \in [0,2\pi), \tag{7}$$

has a root in the interval $[0,2\pi)$. We shall define s as follows:

$$s = \begin{cases} p & \text{if } g(t) \neq 0 \text{ for every } t \in [0,2\pi) \\ p + \varepsilon & \text{if } g(t) = 0 \text{ for some } t \in [0,2\pi) \end{cases}. \tag{8}$$

Step 6. From Theorem 1 (Appendix A), it follows that the sequence of powers of the initial direct-relation matrix $\bar{\mathbf{D}}$ converges to zero. This implies that the series that defines the total influence matrix: $\boldsymbol{T} = \bar{\boldsymbol{D}} + \bar{\boldsymbol{D}}^2 + \ldots = \sum_{m=1}^{\infty} \bar{\boldsymbol{D}}^m$ converges. Hence, the total influence matrix $\mathbf{T} = \left(t_{ij}\right)$ is calculated as:

$$\mathbf{T} = \bar{\mathbf{D}}(\mathbf{I} - \bar{\mathbf{D}})^{-1}. \tag{9}$$

The total influence matrix $\mathbf{T}$ captures the complete picture of how criteria influence each other, considering both direct and indirect effects. It serves to produce the causal diagram map.

Step 7. The row sums and column sums of the total influence matrix $\mathbf{T}$ are calculated in the vectors $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$. The entries of $\mathbf{a}$ and $\mathbf{b}$ are calculated as follows:

$$a_i = \sum_{j=1}^{n} t_{ij}; i = 1,2,\ldots,n;\ b_j = \sum_{i=1}^{n} t_{ij}; j = 1,2,\ldots,n, \tag{10}$$

The entries of the importance vector denoted by $\mathbf{r}^+ = (r_1^+, r_2^+, \ldots, r_n^+)$ and the entries of the relation vector denoted by $\mathbf{r}^- = (r_1^-, r_2^-, \ldots, r_n^-)$ are calculated as follows:

$$r_i^+ = a_i + b_i \text{ and } r_i^- = a_i - b_i,. \tag{11}$$

The element r_i^+ represents the overall influence of criterion Ci, encompassing both the influence it exerts and the influence it receives. A higher r_i^+ value indicates greater importance within the system.

Meanwhile, the element r_i^- determines the causal nature of the criterion. If r_i^- is positive, C_i has a stronger influence on other criteria than it receives, placing it in the cause group. Conversely, if r_i^- is negative, C_i is more influenced by other criteria, categorizing it within the effect group.

By plotting the vectors $\mathbf{r}^+$ and $\mathbf{r}^-$, the data can be visualized in a causality diagram, which illustrates the relationships between criteria. In this diagram:

The horizontal axis (Importance) represents $\mathbf{r}^+$, indicating the overall significance of each criterion.

The vertical axis (Cause/Effect Relation) corresponds to $\mathbf{r}^-$, distinguishing between cause-and-effect groups.

Criteria with positive $\mathbf{r}^-$ values belong to the cause group, meaning they exert more influence than they receive. Conversely, criteria with negative $\mathbf{r}^-$ values fall into the effect group, indicating they are more influenced by other factors.

Step 8. The criteria weights are calculated in the vector $\mathbf{w} = (w_1)$. The entries of w are calculated as

$$w_i = {r_i^+}\Big/{\sum_{k=1}^{n} r_k^+}. \tag{12}$$

Step 9. Determine a threshold value to generate the Impact Relations Map (IRM), a directed graph representing the influence between criteria. In some cases, incorporating all values from matrix **T** into the IRM may result in an overly complex diagram, making decision-making difficult. To simplify the representation, a threshold is applied to exclude negligible influences. Only criteria with influence values in matrix **T** exceeding the threshold are included in the IRM, ensuring a clearer and more meaningful visualization of key relationships.

Two methods are considered for computing the threshold: the arithmetic mean method and the arithmetic mean—standard deviation (MSD) method.

Arithmetic Mean Method

In this approach, the threshold is set as the arithmetic mean μ of matrix **T**. The entries of the matrix $\mathbf{G} = (g_{ij}); i,\ j = 1,2,\ldots,n$ are computed as

$$\mu = \left(\sum_{j=1}^{n}\sum_{i=1}^{n} t_{ij}\right)/n^2;\ g_{ij} = \begin{cases} 1 & \text{if } t_{ij} \geq \mu \\ 0 & \text{if } t_{ij} < \mu \end{cases}. \tag{13}$$

The resulting binary matrix **G** is then used for constructing the IRM and digraph.

Arithmetic Mean—Standard Deviation (MSD) Method

This method refines the threshold by incorporating both the arithmetic mean μ and the standard deviation σ of matrix **T**. The threshold is calculated as $\theta = \mu + \sigma$.

The resulting matrix is then used for IRM construction and digraph design.

The algorithm is also presented in Figure 1. Figure 1 illustrates the step-by-step process of the IDEMATEL method, including the construction of the initial direct-relation matrix, normalization, calculation of the total influence matrix, and derivation of the cause–effect relationships among criteria. The algorithm ensures convergence and provides a comprehensive framework for analyzing interdependencies in multi-criteria decision-making problems.

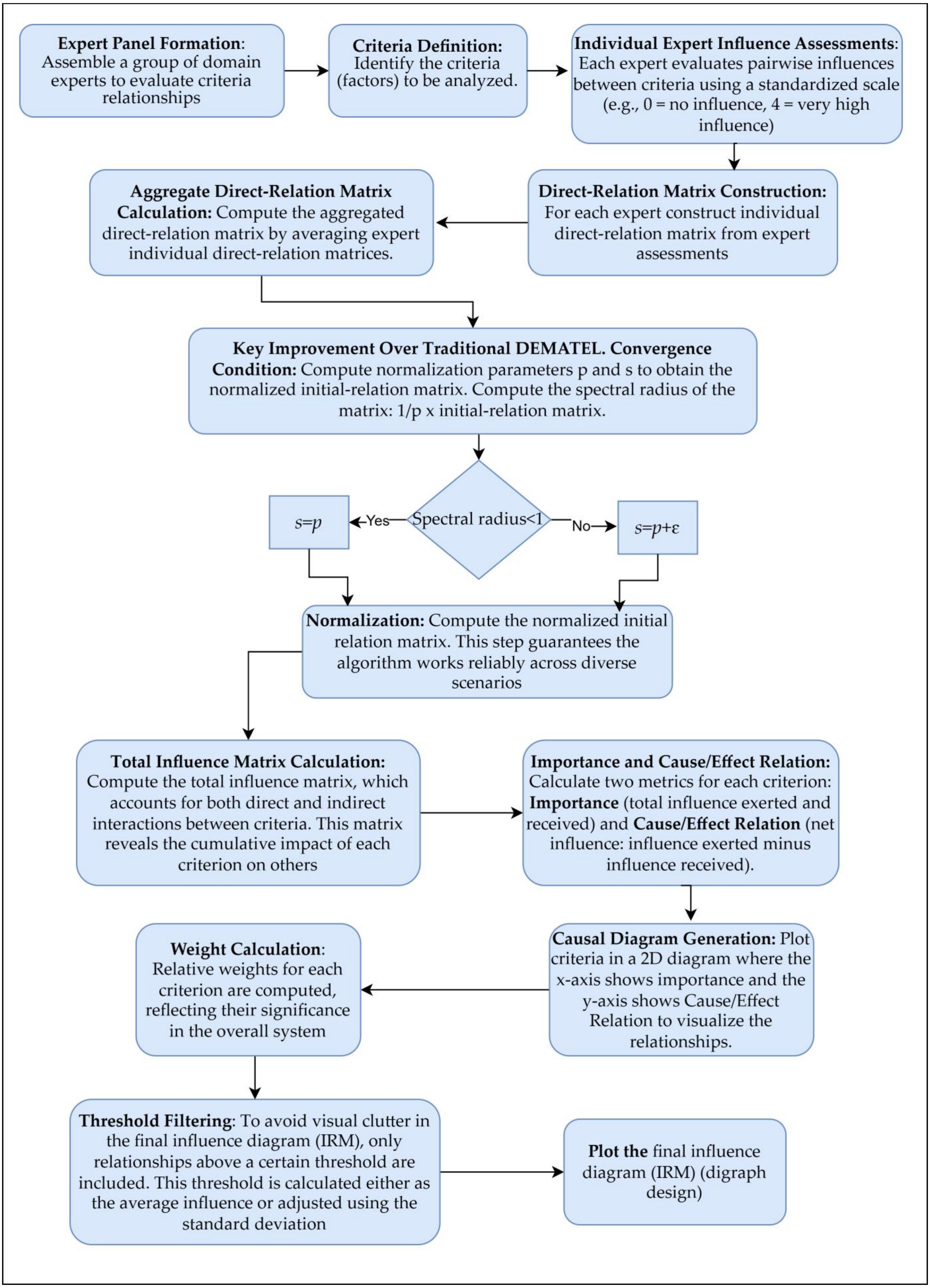

Figure 1. Flowchart of the IDEMATEL algorithm.

6. An Application of the IDEMATEL Method for Evaluation and Analysis of Criteria Used in Selection of GenAI Tools for Academic Research

The framework based on the IDEMATEL method is applied to analyze cause-and-effect relationships among a set of criteria for the selection of a suitable GenAI tool for academic research.

A group of three experts, with solid knowledge and experience in the field was chosen. These experts select a set of important criteria for GenAI tools, based on their experience and a review of the literature. The criteria selected are Functionality and Objectives, Ease of Use and Accessibility, Scalability and Integration, Data Quality and Security, Cost and Value, Support and Community and Reputation and Reviews (Table 4).

Table 4. Key criteria for selecting GenAI tools in academic research, including their symbols and detailed descriptions.

Criterion	Symbol	Description
Functionality and Objectives	FO	Alignment of the tool's features with specific research tasks (literature review, data analysis, writing, etc.).
Ease of Use and Accessibility	UA	User-friendliness for researchers with varying technical expertise; intuitive interfaces; compatibility across platforms.
Scalability and Integration	SI	The tool's ability to handle increasing data/demands and integrate with existing research workflows and systems.
Data Quality and Security	QS	Ensuring data accuracy, reliability, and protection against breaches; maintaining privacy.
Cost and Value	CV	Balancing the cost (subscriptions, usage charges) with the value provided (efficiency, accuracy, improved outcomes).
Support and Community	SC	Availability of technical support, training resources, and a community of users for assistance and best practices.
Reputation and Reviews	RR	Overall reputation of the AI tool and vendor; feedback from other users.

The criteria are interconnected, with some having a direct influence on others (e.g., functionality depending on data quality). The importance of each criterion can vary depending on the specific research context (e.g., data quality is paramount for systematic reviews).

Using the IDEMATEL scale, each expert performed evaluations of each criterion in relation to all other criteria. The $\mathbf{D}^{(k)}$ (initial direct-relation matrix for each expert k) is built. Then, based on Equation (1), the matrix $\mathbf{D}$ is obtained (Table 5).

Table 5. Initial direct-relation matrix $\mathbf{D}$ representing the influence scores among the criteria for selecting GenAI tools.

Criteria	FO	UA	SI	QS	CV	SC	RR
FO	0.000	3.000	3.000	4.000	4.000	2.000	3.000
UA	2.667	0.000	2.000	3.000	2.000	2.000	2.333
SI	2.667	2.000	0.000	2.000	2.000	2.000	4.000
QS	3.333	3.000	3.000	0.000	3.000	2.333	2.667
CV	3.667	2.000	2.000	3.000	0.000	4.000	3.333
SC	1.000	2.000	2.000	2.000	2.000	0.000	3.667
RR	2.000	2.000	2.000	2.000	2.000	2.000	0.000

Each cell (*i,j*) shows the direct influence of criterion *i* on criterion *j*, scored by experts using the IDEMATEL scale. The diagonal values are zero since a criterion does not influence itself. This matrix forms the basis for further normalization and analysis in the IDEMATEL method.

Note that in matrix $\mathbf{D}$, the maximum sums of row elements (19) are equal to the maximum sums of column elements (19). In this case, we have $p = 19$.

The characteristic polynomial of matrix $\frac{1}{p}\mathbf{D}$ is

$f(z) = \det\left(z\mathbf{I} - \frac{1}{p}\mathbf{D}\right) = z^7 + (7.078 \times 10^{-16})z^6 - 0.3896z^5 - 0.1674z^4 - 0.0313z^3 - 0.003\,z^2 - 0.00015z - 2.988 \times 10^{-6}$. Since the coefficient of z^6 is extremely small (approximately zero), the polynomial can be simplified to:

$f(z) = z^7 - 0.3896z^5 - 0.1674z^4 - 0.0313z^3 - 0.003z^2 - 0.00015z - 2.988 \times 10^{-6}$

The maximum value of the roots of the characteristic polynomial is equal to 0.8070 (approximately 0.80699641). This represents the largest eigenvalue of the matrix or the spectral radius of matrix $\frac{1}{p}\mathbf{D}$. Since $\rho\left(\frac{1}{p}\mathbf{D}\right) \cong 0.80699641 < 1$, it follows that the powers of $\frac{1}{p}\mathbf{D}$ tend to zero. In this case, we have $s = p = 19$ and $\bar{\mathbf{D}} = \frac{1}{p}\mathbf{D}$. The normalized initial direct-relation matrix $\bar{\mathbf{D}}$ is displayed in Table 6. Each value of Table 6 represents the normalized direct influence of one criterion on another, scaled between 0 and 1 according to the IDEMATEL method. The normalization ensures comparability across criteria and is a key step in deriving the total influence matrix for further analysis.

Table 6. The normalized initial direct-relation matrix $\bar{\mathbf{D}}$.

Criteria	FO	UA	SI	QS	CV	SC	RR
FO	0.000000	0.157895	0.157895	0.210526	0.210526	0.105263	0.157895
UA	0.140351	0.000000	0.105263	0.157895	0.105263	0.105263	0.122807
SI	0.140351	0.105263	0.000000	0.105263	0.105263	0.105263	0.210526
QS	0.175439	0.157895	0.157895	0.000000	0.157895	0.122807	0.140351
CV	0.192982	0.105263	0.105263	0.157895	0.000000	0.210526	0.175439
SC	0.052632	0.105263	0.105263	0.105263	0.105263	0.000000	0.192982
RR	0.105263	0.105263	0.105263	0.105263	0.105263	0.105263	0.000000

Since the powers of $\bar{\mathbf{D}}$ tend to zero, it follows that the matrix $\mathbf{T}$ can be computed. The total influence matrix $\mathbf{T}$ is displayed in Table 7.

Table 7. Total influence matrix $\mathbf{T}$ showing the cumulative effects among criteria for GenAI tool selection.

Criteria	FO	UA	SI	QS	CV	SC	RR
FO	0.615483	0.700611	0.700611	0.809385	0.777539	0.672040	0.868314
UA	0.586781	0.429056	0.524294	0.617193	0.552236	0.529353	0.664188
SI	0.592153	0.529982	0.434743	0.582269	0.558179	0.536459	0.743810
QS	0.710256	0.653214	0.653214	0.580363	0.687035	0.634654	0.794699
CV	0.730663	0.624210	0.624210	0.728502	0.563766	0.715220	0.838054
SC	0.456101	0.467019	0.467019	0.509966	0.488738	0.380419	0.651403
RR	0.493836	0.463589	0.463589	0.508177	0.487105	0.470331	0.480049

This matrix reveals the overall impact pathways within the system and is essential for identifying the most influential and dependent criteria in the decision-making process.

The non-zero diagonal entries in Table 7 indicate the presence of indirect self-influence resulting from the interconnected structure of the criteria, while the zero diagonal entries in Tables 5 and 6 reflect the absence of direct self-influence by definition.

The sums of elements from the rows (vector **a**) and from columns (vector **b**) of matrix **T** are calculated. Then, the $\mathbf{r}^{+}$ (importance vector), $\mathbf{r}^{-}$ (relation vector), and **w** (criteria weights) are determined (Table 8).

Table 8. The importance vector, relation vector, and criteria weights IDEMATEL.

Criteria Symbols	Criteria	a	b	$\mathbf{r}^{+}$	$\mathbf{r}^{-}$	$\mathbf{r}^{+}$ Rank	w
FO	Functionality and Objectives	5.14398	4.18527	9.32926	0.95871	1	0.15893
UA	Ease of Use and Accessibility	3.90310	3.86768	7.77078	0.03542	6	0.13238
SI	Scalability and Integration	3.97760	3.86768	7.84528	0.10991	5	0.13365
QS	Data Quality and Security	4.71343	4.33586	9.04929	0.37758	2	0.15416
CV	Cost and Value	4.82463	4.11460	8.93922	0.71003	3	0.15229
SC	Support and Community	3.42067	3.93847	7.35914	−0.51781	7	0.12537
RR	Reputation and Reviews	3.36667	5.04052	8.40719	−1.67384	4	0.14322

This table summarizes the influence and dependency of each criterion in the selection of GenAI tools, highlighting their relative importance and causal relationships. The weights reflect the criteria's overall impact on decision-making, guiding prioritization in academic research tool evaluation.

A comparison of the IDEMATEL results with DEMATEL method and the Revised DEMATEL results obtained by Lee is presented in Table 9. Because the maximum sum of the rows in matrix **D** (19) equals the maximum sum of the columns (19), the Revised DEMATEL method proposed by Lee uses ε = 0.001 to calculate the normalized initial direct-relation matrix $\bar{\mathbf{D}}$. In this case, the matrix $\bar{\mathbf{D}}$ from Lee's method is not equal to the matrix $\bar{\mathbf{D}}$ from the original DEMATEL method. The results obtained using the IDEMATEL method are identical to those from the DEMATEL method, with the important distinction that IDEMATEL guarantees the convergence condition, allowing the total influence matrix **T** to be computed for any normalized initial direct-relation matrix $\bar{\mathbf{D}}$.

Table 9. A comparison of IDEMATEL results with DEMATEL method and Revised DEMATEL results obtained by Lee.

Criteria Symbols	Criteria	IDEMATEL		DEMATEL		Revised DEMATEL (LEE)	
		$\mathbf{r}^{+}$	$\mathbf{r}^{-}$	$\mathbf{r}^{+}$	$\mathbf{r}^{-}$	$\mathbf{r}^{+}$	$\mathbf{r}^{-}$
FO	Functionality and Objectives	9.32926	0.95871	9.32926	0.95871	9.32671	0.95845
UA	Ease of Use and Accessibility	7.77078	0.03542	7.77078	0.03542	7.76866	0.03541
SI	Scalability and Integration	7.84528	0.10991	7.84528	0.10991	7.84313	0.10989
QS	Data Quality and Security	9.04929	0.37758	9.04929	0.37758	9.04683	0.37748
CV	Cost and Value	8.93922	0.71003	8.93922	0.71003	8.93679	0.70984
SC	Support and Community	7.35914	−0.51781	7.35914	−0.51781	7.35713	−0.51767
RR	Reputation and Reviews	8.40719	−1.67384	8.40719	−1.67384	8.40490	−1.67340

By mapping the vectors $\mathbf{r}^{+}$ and $\mathbf{r}^{-}$ the data are displayed in a causality diagram (Figure 2). The causality diagram visualizes the causal relationships among the criteria.

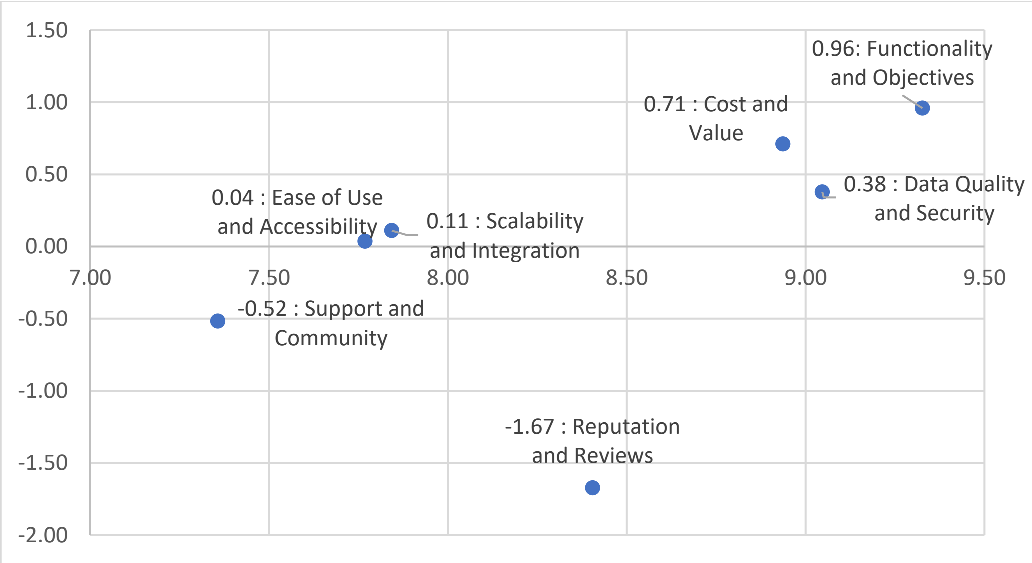

Figure 2. The causality diagram.

A causality diagram (Figure 2) visually represents the causal relationships among criteria. This type of diagram helps to illustrate which criteria act as causes (influencers) and which as effects (influenced), making complex interdependencies easier to interpret and analyze.

In the context of selecting a GenAI tool for academic research, the results obtained from the IDEMATEL method reveal how different criteria influence each other. Driving (cause) criteria are "Functionality and Objectives" ($\mathbf{r}^- = 0.958$), "Cost and Value" ($\mathbf{r}^- = 0.710$) and "Data Quality and Security" ($\mathbf{r}^- = 0.377$). "Functionality and Objectives" represents the core capabilities of a GenAI tool, such as generating text, performing data analysis, or assisting with literature reviews. This criterion is important because it defines the primary purpose and use cases the GenAI can support in academic research. GenAI tools like ChatGPT 4, Perplexity, Gemini, and Claude are the most influential as they shape how researchers interact with information, generate content, and explore academic ideas. Their ability to understand queries, synthesize responses, and provide structured outputs makes them important in academic research.

The cost-effectiveness and overall value of the GenAI tool are significant. Researchers typically have tight budgets, so understanding the value proposition relative to the price is essential. The pricing and accessibility of GenAI tools impact adoption. While ChatGPT 4 and Claude have free and premium versions, tools like Paperpal, Grammarly (Grammarly Inc., San Francisco, CA, USA), and Trinka (Trinka AI, Mumbai, India) rely on subscription-based models. The value proposition of GenAI tools is determined by their accuracy, integration, and return on investment for researchers.

As academic research often involves sensitive data or unpublished findings, the security and reliability of the GenAI tool's data processing capabilities are important. Additionally, the accuracy of the output produced by the GenAI (data quality) is central to ensuring valid research outcomes. Tools like Consensus, Semantic Scholar, and Perplexity ensure access to high-quality, peer-reviewed research. However, privacy concerns arise, especially with tools like ChatGPT 4 and Gemini, which handle sensitive research data.

The balanced criteria (neutral Influence) are "Ease of Use and Accessibility" ($\mathbf{r}^- = 0.035$) and "Scalability and Integration" ($\mathbf{r}^- = 0.110$). Ease of use plays a more supportive role compared to the driving factors. However, GenAI tools should still be user-friendly to accommodate researchers with varying technical expertise. GenAI tools like Jenni.ai, Research Rabbit, and Yewno Discover (Yewno, Inc., Redwood City, CA, USA) are designed for accessibility, but their effectiveness depends on user experience and integration with academic workflows.

The ability to scale the GenAI tool for larger datasets or complex tasks (such as data mining or machine learning model creation) and integrate with existing research workflows (e.g., citation management software) is important but secondary to core functionalities. Many tools rely on integration with Zotero (Center for History and New Media, George Mason University, Fairfax, VA, USA), Mendeley (Mendeley Ltd., London, UK), Google Scholar (Google LLC, Mountain View, CA, USA), and other platforms. Semantic Scholar and Perplexity depend on external databases, making them reliant on scalable infrastructure.

The dependent (effect) criteria are "Reputation and Reviews" ($\mathbf{r}^- = -1.673$) and "Support and Community" ($\mathbf{r}^- = -0.518$). The reputation of the GenAI tool and user reviews tend to be influenced by the tool's performance (functionality, value, data security, etc.). Positive experiences with GenAI tools lead to better reviews and word-of-mouth recommendations in academic communities. The credibility of GenAI tools is still evolving, with many researchers remaining skeptical about AI-generated content. Tools like Grammarly, Trinka, and Paperpal have strong reputations, whereas newer tools like Jenni.ai and Research Rabbit are still gaining trust.

The level of support and community around the tool is important, especially for troubleshooting, learning how to use advanced features, or collaborating on research. However, it is largely dependent on the performance and user base of the tool. GenAI adoption in research is influenced by user feedback, academic acceptance, and continuous improvements. Tools like Consensus and Research Rabbit evolve based on community engagement, making them more dependent than influential.

The most important criteria (with higher importance scores) are "Functionality and Objectives" ($\mathbf{r}^+ = 9.33$), "Data Quality and Security" ($\mathbf{r}^+ = 9.05$), "Cost and Value" ($\mathbf{r}^+ = 8.94$), and "Reputation and Reviews" ($\mathbf{r}^+ = 8.40$).

The least important criteria (with lower importance scores) are "Scalability and Integration" ($\mathbf{r}^+ = 7.84$), "Ease of Use and Accessibility" ($\mathbf{r}^+ = 7.77$), and "Support and Community" ($\mathbf{r}^+ = 7.36$).

While community and reviews are important, they depend on other factors and should improve naturally if the tool performs well in other key areas. Academic researchers require tools that can generate accurate and contextually relevant content for papers, presentations, and reports; assist with data analysis (e.g., conducting statistical analysis, creating models); and aid in the literature review by summarizing articles, extracting key information, and identifying research gaps. For researchers, particularly those with limited funding (e.g., graduate students or researchers at small institutions), cost is a significant consideration. GenAI tools in academic research must comply with data privacy regulations and produce high-quality outputs. Tools that store or process research data should ensure that sensitive information is protected. While reputation and support are dependent factors, they are still essential for long-term success.

The IDEMATEL analysis reveals not only which criteria are most important for GenAI tool selection in academic research, but also how these criteria influence one another. For instance, "Functionality and Objectives" emerges as both a highly influential and highly important criterion, underscoring that a GenAI tool must align closely with the specific research needs it is intended to address. This finding suggests that, regardless of other features, a tool lacking core research functionalities is unlikely to be adopted widely in academic settings.

Similarly, the criterion "Data Quality and Security" is identified as a key driver. In academic research, where data integrity and privacy are paramount, tools that do not ensure robust data protection or output reliability may be unsuitable, even if they excel in other areas. The high importance of the criterion “Cost and Value” reflects the budget constraints typical in academia, highlighting the need for solutions that offer a strong balance between price and performance.

Interestingly, “Reputation and Reviews” and “Support and Community” are found to be more dependent (effect) criteria. This indicates that user satisfaction and community engagement are largely shaped by the tool’s core features, security, and value, rather than being independent decision factors themselves. Therefore, efforts to improve a tool’s reputation or community support should focus first on enhancing its fundamental capabilities.

Based on this analysis, the recommendations are to choose a tool that excels in core features for academic research, ensures high data quality and security standards, and offers cost-effective solutions with good value.

These results have practical implications for institutions and researchers selecting GenAI tools. Decision-makers should prioritize tools that demonstrate strong performance in functionality, data quality, and cost-effectiveness. While ease of use and integration are important for adoption, they are secondary to the tool’s ability to deliver reliable, secure, and relevant research outputs.

For example, an institution with strict data privacy requirements should weigh “Data Quality and Security” more heavily, possibly even above cost considerations. Conversely, for smaller research groups with limited funding, “Cost and Value” may take precedence, provided minimum standards for functionality and security are met.

Ultimately, the IDEMATEL framework provides a structured approach to balancing these priorities and understanding the trade-offs involved in tool selection.

Compared to traditional decision-making methods such as AHP or classic DEMATEL, IDEMATEL offers enhanced insights by explicitly capturing both direct and indirect influences among criteria. The close alignment of IDEMATEL results with those of the Revised DEMATEL (Lee) confirms the robustness of the method, while the causal analysis adds interpretive depth. Unlike simple ranking methods, IDEMATEL’s causality diagram helps decision-makers visualize not just which criteria matter most, but how improving one area (such as functionality) can have cascading positive effects on others (such as reputation and user community).

Table 10 presents a quantitative comparison of five prominent GenAI tools used in academic research, evaluated according to the IDEMATEL criteria. For the “Cost and Value” criterion, actual pricing information is provided. For all other criteria, the tools are rated on a 1–5 scale, where higher values indicate better performance or suitability. Please note that the scale allows for decimal values (e.g., 4.3 or 4.7) to reflect nuanced differences in performance among the tools.

Table 10. Quantitative comparison of GenAI tools.

Criterion	ChatGPT (OpenAI)	Perplexity	Scite.ai (Scite (New York, NY, USA)	Gemini (Google DeepMind)	Claude (Anthropic)
Functionality and Objectives	4.5 (Conversational AI for text/code generation, no citations)	4.8 (Real-time citations + file analysis)	5 (Smart citation context analysis)	4.7 (Multimodal AI for reasoning across text, images, audio)	4.6 (Data analysis, and ethical automation)
Cost and Value	$20/month for Plus (GPT-4/4.5), free basic tier	$20/month (Pro plan), free limited use	$10–$50/month (Institutional rates, academic/enterprise access needed)	Free tier + enterprise pricing	Custom pricing

Data Quality and Security	4 (Pre-2023 data, no source verification, security respected)	4.9 (Links to real sources, source citations)	5 (1.2B+ verified citations, very strong on safety)	4.5 (Multimodal accuracy)	4.7 (Content moderation)
Ease of Use and Accessibility	4.5 (Web, app, and API interfaces, user-friendly design)	4.8 (Browser extensions, simple interface)	4.6 (intuitive dashboard; browser extensions)	4.7 (Integrated with Google ecosystem)	4.5 (Web, API, and enterprise solutions, Multilingual support)
Scalability and Integration	4 (API for developers, plug-ins for various apps, API rate limits)	4.3 (Collaboration features, Sonar API access)	4.8 (Integrates with Zotero, EndNote, browser plug-ins)	5 (Native integration with Google services)	4.4 (API and enterprise workflow integration)
Reputation and Reviews	4.6 (Widely adopted, 100M+ users)	4.8 (Speed, accuracy, and citation transparency)	4.9 (Highly regarded among researchers for citation analysis, context, and integration)	4.7 (Recognized for multimodal prowess, trusted infrastructure, Google credibility)	4.5 (High ethical standards, top-rated for safety)
Support and Community	4 (OpenAI forums, support channels, Community forums)	4.5 (Priority support for Pro, active community)	4.7 (Institutional webinars, documentation, responsive support; growing reputation in academia)	4.8 (Google Cloud support, large user base)	4.6 (Dedicated enterprise support)

In order to reduce the complexity of the criteria relations, a threshold value is set to filter out negligible effects. The threshold calculated for **T** matrix with the arithmetic mean methods is μ = 0.599. Only criteria whose influence value in matrix **T** is higher than the threshold value can be chosen and converted into the IRM. The digraph for this threshold is illustrated in Figure 3.

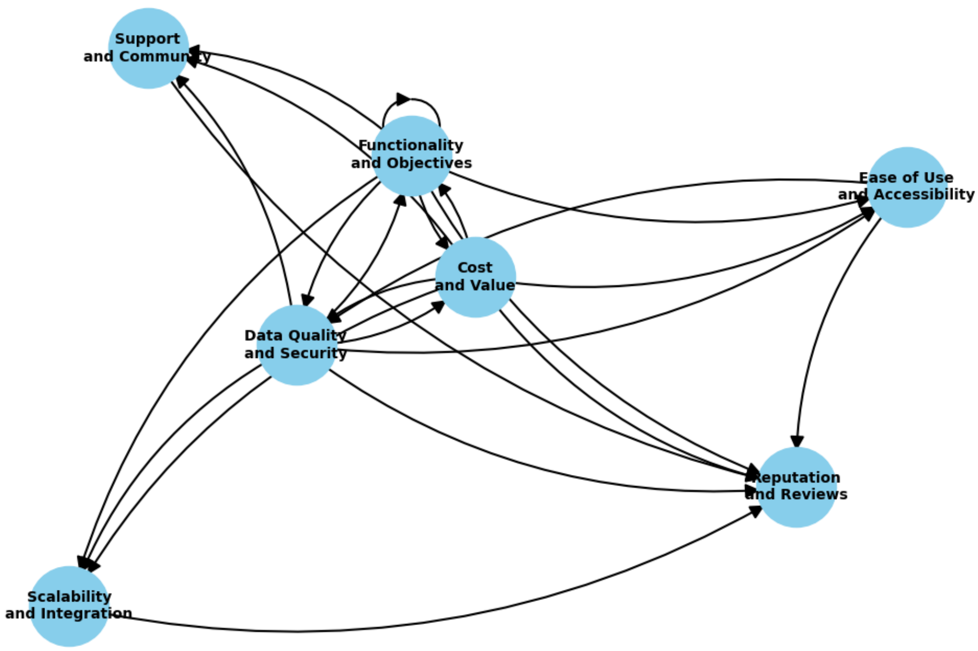

Figure 3. The digraph for μ = 0.599.

The digraph visualizes the significant causal relationships among the criteria for selecting GenAI tools in academic research. Each node represents a specific criterion, and each directed arrow (edge) indicates that one criterion has a direct influence on another, as determined by the threshold value ($\mu = 0.599$). The arrows point from the influencing criterion (cause) to the influenced criterion (effect). Criteria with many outgoing arrows, such as “Functionality and Objectives” and “Data Quality and Security”, act as key drivers in the system. They influence multiple other criteria, suggesting that improvements or changes in these areas are likely to have a broad impact. Criteria with many incoming arrows, such as “Reputation and Reviews”, are more dependent on other factors. Their status is shaped by changes in other criteria. This digraph helps to identify which criteria should be prioritized in decision-making, as influencing them can lead to widespread improvements across the system.

The threshold value of 0.599 was selected based on the distribution of influence values in the total influence matrix. This value effectively filters out weaker, less significant relationships, allowing the digraph to highlight only the most meaningful and impactful connections between criteria. By using this threshold, the resulting digraph avoids excessive complexity and focuses the analysis on the dominant causal pathways, making the visualization more interpretable and actionable. We also tested other threshold values and found that 0.599 provides a balanced representation, neither omitting important relationships nor overcrowding the diagram with minor influences.

The IDEMATEL method offers several key advantages over traditional approaches for analyzing criteria in the selection of GenAI tools. First and foremost, it enables a comprehensive mapping of both direct and indirect relationships among criteria, providing a holistic view of how improvements in one area can cascade through the system and affect others. This causality analysis allows decision-makers to identify not only the most influential criteria but also the pathways through which these criteria impact overall outcomes, supporting more strategic prioritization and resource allocation.

A significant technical advantage of IDEMATEL is its robust approach to calculating the total influence matrix. Unlike classical DEMATEL methods, which may fail or produce unreliable results if certain mathematical conditions are not met, IDEMATEL introduces a generalized convergence condition. This ensures that, regardless of the initial structure of the direct relation matrix, the algorithm can always compute the total influence matrix reliably. As a result, IDEMATEL guarantees the stability and applicability of the analysis across a wide variety of real-world scenarios, making it a more versatile and dependable tool for complex decision-making environments such as GenAI tool selection in academic research.

By combining these methodological strengths, IDEMATEL not only deepens the understanding of criteria interdependence but also ensures that the analytical process is mathematically sound and broadly applicable. This added value is particularly important in academic settings, where decision quality and methodological rigor are paramount.

While this study provides valuable insights into the selection of GenAI tools for academic research, certain limitations should be acknowledged. The evaluation involved a small group of experts, which may influence the generalizability of the findings. Additionally, the criteria and their assessments were based on expert judgment combined with literature review, introducing an element of subjectivity that is common in decision-making research. The specific context and tools considered may not capture the full diversity of academic environments or GenAI applications.

These considerations open opportunities for future research to expand the expert panel, incorporate empirical validation, and explore a broader range of criteria and tool types. Nevertheless, the current analysis offers a robust foundation for understanding key factors and their interrelationships in GenAI tool selection.

7. Conclusions

The integration of Artificial Intelligence (AI) tools into academic research has revolutionized the way scientific researchers approach research subjects. These tools enhance productivity, streamline data analysis, and facilitate the discovery of new insights. Recent research highlights the transformative potential of GenAI tools in academic research, offering benefits such as increased efficiency in literature reviews, writing, data analysis, and peer review processes. One major drawback of using generative AI tools in academic research is the risk of inaccuracies, biases, and lack of transparency in source selection, which can compromise the reliability and credibility of researchers' work.

The paper presents an analysis of the criteria used in the selection and adoption of GenAI tools for academic research. The choice of the DEMATEL method to analyze the influences between criteria and the calculation of the criteria weights is justified in comparison with the commonly used weighting methods AHP, ANP, and BWM.

An improved group DEMATEL method, called IDEMATEL, is then proposed to ensure that the powers of the normalized direct relation matrix converge, which guarantees the computation of the total influence matrix under any circumstances. The IDEMATEL method is applied to analyze a set of criteria involved in the selection of GenAI tools for academic research, based on evaluations made by a group of experts.

The results highlight that Functionality and Objectives should be prioritized, as they are the main drivers of GenAI tool selection. Additionally, Cost and Value, as well as Data Security, are considerations that can directly impact the decision-making process. Reputation and Community Support play secondary roles, primarily depending on the tool's performance and user experience. These insights can help researchers and institutions focus their evaluation efforts and resources on the most influential factors, ultimately enhancing research productivity and quality.

DEMATEL helps to understand which criteria are the key drivers (causes) and which are more dependent (effects). By identifying the "cause" criteria (ex: functionality, cost, etc.), DEMATEL enables decision-makers to focus resources on the most influential factors. While relying on expert judgements, the process of pair-wise comparisons and the mathematical formulation of DEMATEL helps reduce the subjectivity and allows for a more objective analysis than simply brainstorming or intuition.

The results of this study have important implications for both individual researchers and institutional decision-makers. By clarifying the causal relationships among selection criteria, institutions can develop more effective procurement policies and support structures, ensuring that investments in GenAI tools yield the greatest impact. The causal analysis also suggests that improvements in core functionalities and data security can have positive ripple effects on reputation and community support, contributing to a more sustainable and trusted research ecosystem.

While this study provides a robust framework, several limitations should be acknowledged. The analysis was based on the judgments of a small group of experts, which may affect the generalizability of the results. The selected criteria and their assessments, though grounded in the literature and expert opinions, may not capture all relevant factors or reflect the diversity of academic disciplines and institutional contexts. Additionally, the evaluation focused on a specific set of GenAI tools and may not fully represent the rapidly evolving AI landscape.

To build on these findings, future research should explore the application of IDEMATEL in diverse academic domains and organizational settings, as the relative importance of criteria may vary across fields such as the social sciences, natural sciences, or humanities. Integrating IDEMATEL with fuzzy logic or other uncertainty-handling techniques could further enhance its ability to manage ambiguous or imprecise expert judgments. Comparative studies involving larger and more varied expert panels would strengthen the evidence base and refine the evaluation framework.

Author Contributions: Conceptualization, C.Z.R. and M.R.; methodology, M.R. and C.Z.R.; software, C.Z.R.; validation, C.Z.R.; resources, M.R.; writing—original draft preparation, C.Z.R. and M.R.; writing—review and editing, C.Z.R. and M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, and further inquiries can be directed to the corresponding author.

Acknowledgments: The first author would like to acknowledge technical support from the project PN 23 38 01 01 "Contributions to the consolidation of emerging technologies specific to the Internet of Things and complex systems".

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Some Conditions That Imply the Powers of a Square Matrix Tend to Zero

This subsection discusses conditions under which the powers of a square matrix converge to zero. The spectral radius of a matrix, which is determined by its eigenvalues, plays a critical role in understanding whether the powers of the matrix will converge to zero. If the spectral radius is strictly less than one, the powers of the matrix will tend to zero. However, calculating the spectral radius can be computationally expensive, so alternative sufficient conditions are often used. A sufficient condition for ensuring that the powers of a matrix converge to zero is the existence of a matrix norm such that the norm of the matrix is less than one. Two specific norms are discussed: the 1-norm and the ∞-norm.

If $\mathbf{A}$ is a square matrix denote by $\sigma(\mathbf{A})$ the set of eigenvalues of $\mathbf{A}$. The number $\rho(\mathbf{A}) = \max_{\lambda \in \sigma(\mathbf{A})} |\lambda|$ is called the spectral radius of the matrix $\mathbf{A}$. A necessary and sufficient condition that the sequence of powers of $\mathbf{A}$ be convergent to zero is that $\rho(\mathbf{A}) < 1$. Unfortunately, the computationally effort for finding the spectral radius is great. That is why one can use instead condition $\rho(\mathbf{A}) < 1$ only sufficient conditions that imply the convergence of the sequence of powers of matrix $\mathbf{A}$. A sufficient condition that implies that the sequence of powers of a $n \times n$ matrix $\mathbf{A}$ is convergent to zero is the existence of a norm on $\mathbb{R}^n$ such that $\|\mathbf{A}\| < 1$. On $\mathbb{R}^n$ consider the following norms:

- the l^1 norm that is $\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$ and
- the l^∞ norm that is $\|\mathbf{x}\|_\infty = \max_{1 \le i \le n} |x_i|$.

If $\mathbf{A} = (a_{ij})$ is a $n \times n$ matrix and on $\mathbb{R}^n$ consider the l^1 norm then the norm of the linear application generated by the matrix A is:

$$\|\mathbf{A}\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|.$$

In the case on $\mathbb{R}^n$ we consider the l^∞ norm then the norm of $\mathbf{A}$ is:

$$\|\mathbf{A}\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|.$$

Let $s = \max(\|\mathbf{A}\|_1, \|\mathbf{A}\|_\infty)$.

Theorem A1. *Let* $\boldsymbol{A} = (a_{ij})$ *be a* $n \times n$ *matrix,* $\varepsilon > 0, s = max(\|\boldsymbol{A}\|_1, \|\boldsymbol{A}\|_\infty)$, $\boldsymbol{B}_1 = \frac{1}{s}\boldsymbol{A}$, $\boldsymbol{B}_2 = \frac{1}{s+\varepsilon}\boldsymbol{A}$. *Let* $Q = \{z \in \mathbb{C}\colon |z| = 1\}$. *Then, the following assertions hold:*

(i) *If* $det(z\boldsymbol{I} - \boldsymbol{B}_1) \neq 0$ *for every* $z \in Q$ *then the powers of matrix* $\boldsymbol{B}_1$ *tend to zero.*

(ii) *If* $det(z\boldsymbol{I} - \boldsymbol{B}_1) = 0$ *for some* $z \in Q$ *then the powers of matrix* $\boldsymbol{B}_2$ *tend to zero.*

Here, we denoted by **I** the unit matrix.

Proof of Theorem A1. In order to prove assertion (i) suppose that $\det(z\mathbf{I} - \mathbf{B}_1) \neq 0$ for every $z \in Q$. This implies that matrix $\mathbf{B}_1$ has no eigenvalues on the boundary of the unit disk, hence $\rho(\mathbf{B}_1) < 1$. Thus, the powers of the matrix $\mathbf{B}_1$ tend to zero.

In order to prove assertion (ii) suppose that $\det(z\mathbf{I} - \mathbf{B}_1) = 0$ for some $z \in Q$. Then, $\rho(\mathbf{B}_1) = 1$. Since $\mathbf{B}_2 = \frac{s}{s+\varepsilon}\mathbf{B}_1$ it follows that $\rho(\mathbf{B}_2) = \frac{s}{s+\varepsilon} < 1$, hence the powers of the matrix $\mathbf{B}_2$ tend to zero. □

References

1. Panda, S.; Kaur, D.N. Exploring the role of generative AI in academia: Opportunities and challenges. *IP Indian J. Libr. Sci. Inf. Technol.* **2024**, *9*, 10–18231. https://doi.org/10.18231/j.ijlsit.2024.003.
2. Kalota, F. A primer on generative artificial intelligence. *Educ. Sci.* **2024**, *14*, 172. https://doi.org/10.3390/educsci14020172.
3. Akpan, I.J.; Kobara, Y.M.; Owolabi, J.; Akpan, A.A.; Offodile, O.F. Conversational and generative artificial intelligence and human–chatbot interaction in education and research. *Int. Trans. Oper. Res.* **2025**, *32*, 1251–1281. https://doi.org/10.1111/itor.13522.
4. Andersen, J.P.; Degn, L.; Fishberg, R.; Graversen, E.K.; Horbach, S.P.; Schmidt, E.K.; Sørensen, M.P. Generative artificial intelligence (GenAI) in the research process: A survey of researchers' practices and perceptions. *Technol. Soc.* **2025**, *81*, 102813. https://doi.org/10.1016/j.techsoc.2025.102813.
5. Gabus, A.; Fontela, E. *World Problems: An Invitation to Further Thought Within the Framework of DEMATEL*; Battelle Geneva Research Center: Geneva, Switzerland, 1972.
6. Pinzolits, R. AI in academia: An overview of selected tools and their areas of application. *MAP Educ. Humanit.* **2024**, *4*, 37–50. https://doi.org/10.53880/2744-2373.2023.4.37.
7. Danler, M.; Hackl, W.O.; Neururer, S.B.; Pfeifer, B. Quality and effectiveness of AI tools for students and researchers for scientific literature review and analysis. In Proceedings of the dHealth 2024 Conference, Vienna, Austria, 7–8 May 2024; pp. 203–208. https://doi.org/10.3233/SHTI240038.
8. Šarlauskienė, L.; Dagytė, S. Criteria for selecting artificial intelligence tools. *Innov. Publ. Print. Multimed. Technol.* **2024**, 100–106. https://doi.org/10.59476/ilpmt2024.100-106.
9. Al-kfairy, M. Factors impacting the adoption and acceptance of ChatGPT in educational settings: A narrative review of empirical studies. *Appl. Syst. Innov.* **2024**, *7*, 110. https://doi.org/10.3390/asi7060110.
10. Aydogmus, H.Y.; Aydogmus, U. Evaluation of artificial intelligence tools for universities with fuzzy multi-criteria decision-making methods. In *AI Adoption and Diffusion in Education*; Sart, G., Sezgin, F.H., Eds.; IGI Global Scientific Publishing: Hershey, PA, USA, 2025; pp. 153–178. https://doi.org/10.4018/979-8-3693-7949-3.ch006.
11. Bukar, U.A.; Sayeed, M.S.; Razak, S.F.A.; Yogarayan, S.; Sneesl, R. Decision-making framework for the utilization of generative artificial intelligence in education: A case study of ChatGPT. *IEEE Access* **2024**, *12*, 95368–95384. https://doi.org/10.1109/ACCESS.2024.3425172.
12. Huang, Q.; Lv, C.; Lu, L.; Tu, S. Evaluating the quality of AI-generated digital educational resources for university teaching and learning. *Systems* **2025**, *13*, 174. https://doi.org/10.3390/systems13030174.
13. Bukar, U.A.; Sayeed, M.S.; Razak, S.F.A.; Yogarayan, S.; Sneesl, R. Prioritizing ethical conundrums in the utilization of ChatGPT in education through an analytical hierarchical approach. *Educ. Sci.* **2024**, *14*, 959. https://doi.org/10.3390/educsci14090959.
14. Agarwal, A. Optimizing employee roles in the era of generative AI: A multi-criteria decision-making analysis of co-creation dynamics. *Cogent Soc. Sci.* **2025**, *11*, 2476737. https://doi.org/10.1080/23311886.2025.2476737.

15. Gupta, S.; Kaur, S.; Gupta, M.; Singh, T. AI empowered academia: A fuzzy prioritization framework for academic challenges. *J. Int. Educ. Bus.* 2024, *ahead-of-print*. https://doi.org/10.1108/JIEB-06-2024-0071.
16. Ilieva, G. Extension of interval-valued hesitant Fermatean fuzzy TOPSIS for evaluating and benchmarking of generative AI chatbots. *Electronics* **2025**, *14*, 555. https://doi.org/10.3390/electronics14030555.
17. Chakrabortty, R.K.; Abdel-Basset, M.; Ali, A.M. A multi-criteria decision analysis model for selecting an optimum customer service chatbot under uncertainty. *Decis. Anal. J.* **2023**, *6*, 100168. https://doi.org/10.1016/j.dajour.2023.100168.
18. Yang, L.; Wang, J. Factors influencing initial public acceptance of integrating the ChatGPT-type model with government services. *Kybernetes* **2024**, *53*, 4948–4975. https://doi.org/10.1108/K-06-2023-1011.
19. Fontoura, L.; de Mattos Nascimento, D.L.; Neto, J.V.; Caiado, R.G.G. Energy Gen-AI technology framework: A perspective of energy efficiency and business ethics in operation management. *Technol. Soc.* **2025**, *81*, 102847. https://doi.org/10.1016/j.techsoc.2025.102847.
20. Gupta, R.; Rathore, B. Exploring the generative AI adoption in service industry: A mixed-method analysis. *J. Retail. Consum. Serv.* **2024**, *81*, 103997. https://doi.org/10.1016/j.jretconser.2024.103997.
21. Dergaa, I.; Chamari, K.; Żmijewski, P.; Saad, B. From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biol. Sport* **2023**, *40*, 615–622. https://doi.org/10.5114/biolsport.2023.125623.
22. Oyelude, A. Artificial intelligence (AI) tools for academic research. *Libr. Hi Tech News* **2024**, *41*, 18–20. https://doi.org/10.1108/lhtn-08-2024-0131.
23. Burger, B.; Kanbach, D.K.; Kraus, S.; Breier, M.; Corvello, V. On the use of AI-based tools like ChatGPT to support management research. *Eur. J. Innov. Manag.* **2023**, *26*, 233–241. https://doi.org/10.1108/EJIM-02-2023-0156.
24. Olu-Ajayi, R.; Alaka, H.; Sunmola, F.; Ajayi, S.; Mporas, I. Statistical and artificial intelligence-based tools for building energy prediction: A systematic literature review. *IEEE Trans. Eng. Manag.* **2024**, *71*, 14733–14753. https://doi.org/10.1109/TEM.2024.3422821.
25. Ekundayo, T.; Khan, Z.; Nuzhat, S. Evaluating the influence of artificial intelligence on scholarly research: A study focused on academics. *Hum. Behav. Emerg. Technol.* **2024**, *2024*, 8713718. https://doi.org/10.1155/2024/8713718.
26. Polonsky, M.; Rotman, J. Should artificial intelligent agents be your co-author? Arguments in favour, informed by ChatGPT. *Australas. Mark. J.* **2023**, *31*, 91–96. https://doi.org/10.1177/14413582231167882.
27. Iorliam, A.; Ingio, J.A. A comparative analysis of generative artificial intelligence tools for natural language processing. *J. Comput. Theor. Appl.* **2024**, *1*, 311–325. https://doi.org/10.62411/jcta.9447.
28. Zawacki-Richter, O.; Marín, V.; Bond, M.; Gouverneur, F. Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *Int. J. Educ. Technol. High. Educ.* **2019**, *16*, 39. https://doi.org/10.1186/s41239-019-0171-0.
29. Bhattamisra, S.; Banerjee, P.; Gupta, P.; Mayuren, J.; Patra, S.; Candasamy, M. Artificial intelligence in pharmaceutical and healthcare research. *Big Data Cogn. Comput.* **2023**, *7*, 10. https://doi.org/10.3390/bdcc7010010.
30. Michalak, R. Fostering undergraduate academic research: Rolling out a tech stack with AI-powered tools in a library. *J. Libr. Admin.* **2024**, *64*, 335–346. https://doi.org/10.1080/01930826.2024.2316523.
31. Saaty, T.L. A scaling method for priorities in hierarchical structures. *J. Math. Psychol.* **1977**, *15*, 234–281. https://doi.org/10.1016/0022-2496(77)90033-5.
32. Saaty, T.L. *The Analytic Hierarchy Process*; McGraw-Hill Press: New York, NY, USA, 1980.
33. Saaty, T.L. Decision Making with Dependence and Feedback: The Analytic Network Process; RWS Publications: Pittsburgh, PA, USA, 2001.
34. Rezaei, J. Best-worst multi-criteria decision-making method. *Omega* **2015**, *53*, 49–57. https://doi.org/10.1016/j.omega.2014.11.009.
35. Rădulescu, C.Z.; Rădulescu, M. Group decision support approach for cloud quality of service criteria weighting. *Stud. Inform. Control* **2018**, *27*, 275–284. https://doi.org/10.24846/v27i3y201803.
36. Wu, H.H.; Tsai, Y.N. An integrated approach of AHP and DEMATEL methods in evaluating the criteria of auto spare parts industry. *Int. J. Syst. Sci.* **2012**, *43*, 2114–2124. https://doi.org/10.1080/00207721.2011.564674.
37. Tavana, M.; Soltanifar, M.; Santos-Arteaga, F.J.; Sharafi, H. Analytic hierarchy process and data envelopment analysis: A match made in heaven. *Expert Syst. Appl.* **2023**, *223*, 119902. https://doi.org/10.1016/j.eswa.2023.119902.
38. Srđević, Z.; Srđević, B.; Ždero, S.; Ilić, M. How MCDM method and the number of comparisons influence the priority vector. *Comput. Sci. Inf. Syst.* **2022**, *19*, 251–275. https://doi.org/10.2298/CSIS210410051S.

39. Tan, Y.; Wang, X.; Liu, X.; Zhang, S.; Li, N.; Liang, J.; Hu, D.; Yang, Q. Comparison of AHP and BWM methods based on ArcGIS for ecological suitability assessment of Panax notoginseng in Yunnan Province, China. *Ind. Crops Prod.* **2023**, *199*, 116737. https://doi.org/10.1016/j.indcrop.2023.116737.
40. Muerza, V.; Milenkovic, M.; Larrodé, E.; Bojovic, N. Selection of an international distribution center location: A comparison between stand-alone ANP and DEMATEL-ANP applications. *Res. Transp. Bus. Manag.* **2024**, *56*, 101135. https://doi.org/10.1016/j.rtbm.2024.101135.
41. Bharti, S.S.; Prasad, K.; Sudha, S.; Kumari, V. Prioritisation of factors for artificial intelligence-based technology adoption by banking customers in India: Evidence using the DEMATEL approach. *Appl. Finance Lett.* **2023**, *12*, 2–22. https://doi.org/10.24135/afl.v12i2.623.
42. Sharma, M.; Luthra, S.; Joshi, S.; Kumar, A. Implementing challenges of artificial intelligence: Evidence from public manufacturing sector of an emerging economy. *Gov. Inf. Q.* **2022**, *39*, 101624. https://doi.org/10.1016/j.giq.2021.101624.
43. Alinezhad, A.; Khalili, J. DEMATEL method. In *New Methods and Applications in Multiple Attribute Decision Making (MADM)*; International Series in Operations Research & Management Science; Springer: Cham, Switzerland, 2019; Volume 277, pp. 215–221. https://doi.org/10.1007/978-3-030-15009-9_15.
44. Lee, H.-S.; Tzeng, G.-H.; Yeih, W.; Wang, Y.-J.; Yang, S.-C. Revised DEMATEL: Resolving the infeasibility of DEMATEL. *Appl. Math. Model.* **2013**, *37*, 6746–6757. https://doi.org/10.1016/j.apm.2013.01.016.
45. Chen, C.Y.; Huang, J.J. A novel DEMATEL approach by considering normalization and invertibility. *Symmetry* **2022**, *14*, 1109. https://doi.org/10.3390/sym14061109.