A Bibliometric Mapping of Investment Portfolio Management Research

Darius Sabaliauskas  
Department of Information Systems  
Faculty of Fundamental Sciences  
Vilnius Gediminas Technical UniversitySauletekio al. 11, Vilnius, LT-10223, Lithuania  
e-mail: [dariussabaliauskas@stud.vilniustech.lt](mailto:dariussabaliauskas@stud.vilniustech.lt),

Jolanta Miliauskaitė  
Department of Information Systems  
Faculty of Fundamental Sciences  
Vilnius Gediminas Technical UniversitySauletekio al. 11, Vilnius, LT-10223, Lithuania  
e-mail: [jolanta.miliauskaite@vilniustech.lt](mailto:jolanta.miliauskaite@vilniustech.lt),

ORCID: 0000-0003-1237-3499

***Abstract* — *Investment portfolio management in the context of volatile financial markets demands innovative approaches that can effectively handle uncertainties, patterns, and trends.***

***This paper presents a comprehensive bibliometric analysis of research in the field of Investment Portfolio Management (IPM) to uncover its intellectual landscape, emerging trends, and key contributors. Utilizing advanced bibliometric techniques, we systematically review a vast collection of scholarly articles and related publications from Web of Science (WoS) databases. By employing co-citation analysis, keyword clustering, and network visualization, we map the intellectual structure of IPM research, identifying influential works, seminal authors, and prominent research themes.***

***Our analysis reveals the evolution of IPM research over time, highlighting pivotal publications that have shaped the discipline. We uncover clusters of research themes, ranging from traditional portfolio optimization and asset allocation to contemporary topics such as behavioral finance's impact on investment decisions and the integration of machine learning techniques in portfolio management. Furthermore, we identify the collaborative networks that have formed among researchers, institutions, and countries, illustrating the global reach and interconnectivity of IPM studies.***

***Through this bibliometric mapping, we provide a holistic view of the IPM research landscape, shedding light on its intellectual foundations and current directions. This study serves as a valuable resource for scholars, practitioners, and policymakers seeking to understand the trajectory of IPM research, pinpoint areas of innovation, and facilitate interdisciplinary collaborations. As the field continues to evolve, our bibliometric analysis offers insights into the dynamic nature of investment portfolio management research, enabling stakeholders to make informed decisions in a rapidly changing financial landscape***

***In conclusion, By providing a comprehensive overview of existing research, the survey contributes to understanding the current state of the field, fostering further research innovation, and inspiring new strategies to tackle the complexities of modern financial markets.***

***Keywords—Artificial intelligence, neural network, machine learning, portfolio optimization, financial predictions, investment, portfolio allocation, algorithms, mapping***

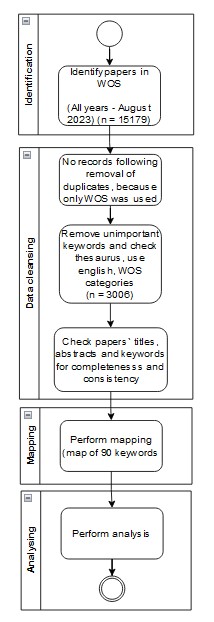
# I. INTRODUCTION

The field of investment portfolio management stands at the crossroads of finance, economics, and technology, navigating the intricate landscapes of risk and return, asset allocation, and market dynamics. Nowadays, there are a lot of attempts to find solutions for how to manage portfolios [1]–[3]. The profound impact of investment strategies [4]–[6] on wealth generation and financial stability has fueled an ever-expanding body of research aimed at deciphering the complexities inherent in this domain [7]. In the era of information explosion and data-driven decision-making, understanding the trajectory of investment portfolio management research [8], [9] has gained paramount importance [10]. This paper embarks on a bibliometric journey to unravel the underlying trends, contributors, and emerging themes within investment portfolio management research, drawing insights from the rich repository of the Web of Science (WoS) database. The structure of the paper is as follows. Section 2 describes the methodology of a systematic mapping of the Investment portfolio allocation and portfolio optimization approaches for Investment Portfolio Management. Section 3 presents the results of the systematic mapping. Section 4 discusses the results and Section 5 concludes the paper.

# II. SURVEY METHODOLOGY

A systematic mapping (SM) on methodologies and models for the *Investment Portfolio Management* (IPM) topic is conducted as proposed in [11], [12]. Its flow diagram is presented in Fig. 1.

It consists of four main steps as follows: identification, data cleaning, mapping, and analysis. The main results of the performed steps are presented in brackets. A detailed description of the steps is forwarding.



1. The flow diagram for our SM.

## Identification

In this step, the following activities are performed.

#### Research Questions

The main research question (MRQ) follows: (MRQ1) *What are the predominant methodologies and models employed in investment portfolio management research?* According to MRQ1, the following research questions (RQ) are defined:

* RQ-1: How has **topics** of IPM and its related models, changed over different time periods?
* RQ-2: Are certain methodologies and models more prevalent in specific **geographic regions** within the field of IPM research?
* RQ-3: What are the most frequently used **methodologies and models** of IPM research?

#### The List of Search Sources

The WoS databases were used based on the experience reported in[13]. Gusenbauer and Haddaway [14] have found that 14 of 28 academic search systems (ASS) are well-suited to SLR. Emphasizing the Computer Science research area, the main ASS are the following: ACM Digital Library, Bielefeld Academic Search Engine (BASE), ScienceDirect, Scopus, WoS, and Wiley Online Library. For this SM, we need ASS, which has a full download function of search results for bibliometric analysis in a particular reference management software. Moreover, another important attribute is the quality of the presented research. Consequently, we have found that WoS meets all those requirements.

#### Searching Query

We have identified three main groups of terms for the search as follows:

* Concepts related to investment portfolio;
* Concepts related to methods and models that are relevant in the context of investment portfolio management.

The final search query is developed based on WoS search requirements as follows:

("investment portfolio\*" OR "asset allocation\*" OR "diversification\*" OR "portfolio optimization\*" OR "portfolio allocation\*" OR " financial instrument\*" OR "portfolio management\*" OR "portfolio rebalancing\*" OR "securities\*" OR "financial prediction\*") AND ( "artifici\* intelligence\*" OR "machi\* learn\*" OR "quantitative method\*" OR "statistical\*" OR "algorithms\*" OR "mathematical\*" OR "heuristic\*" OR "simulation\*")

#### d) Searching by Query

In this step, we have performed a search using the defined search query in WoS. Moreover, we have limited the search by language and document type and by categories as presented in Table I. We have not limited research areas, since it is related to RQ-2.

1. results of identification and data cleansing steps

| Database | Document Type | Language | Categories | Search Results |
| --- | --- | --- | --- | --- |
| WoS | article OR proceedings papers OR review | English | Computer sciences Artificial Intelligence OR Computer science Information Systems OR Computer science Interdisciplinary Applications | 3009 |

## Data Cleansing

In this step, we have removed plural and unimportant keywords using the thesaurus function in the VOSviewer (<https://www.vosviewer.com/>) software tool for conducting and visualizing bibliometric networks [14]

Also, titles, abstracts, and keywords of all identified papers were checked for completeness and consistency.

## Mapping

A set of relevant papers obtained were transferred to VOSviewer.

To develop a keywords map on IPM, we have used keywords from the relevant papers. VOSviewer uses an automatic keyword identification technique [14]. It creates the keywords map according to the closeness and strength of existing links between found keywords calculating the number of papers, in which both keywords have occurred together. The size of the frames presents the occurrence of keywords.

Moreover, in this SM we concentrate on the keywords map coloring according to Average Publication Year (APY) which indicates the average publication year of the papers in which a keyword occurs [15]. Note that APY is a positive rational number.

## Analyzing

In this SM, the content was analyzed as follows:

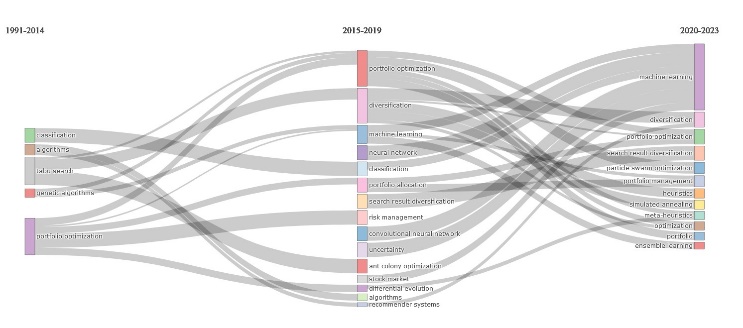
* Chronological visualization and analysis (RQ-1);
* Countries visualization and analysis (RQ-2);
* Keyword occurrence visualization based on APY and analysis (RQ-3).

The found keywords were analyzed quantitatively according to APY (RQ-3) as follows:

* newest keywords– those keywords’ occurrence is between 9 and 118, but the high is just two keywords deep learning(78) and machine learning(118) others are in intervals (9-38) (i.e., in the context of this SM ), but their APY is new [2019-2021] (see Fig. 4);
* older keywords – those keywords’ occurrence can be similar to the newest between 9 and 129 (genetic programming) and their APY is old [2011-2015] (see Fig. 4).

# III. MAIN RESULTS OF SM

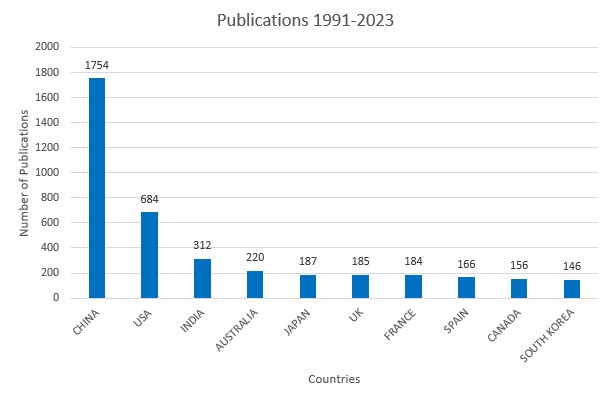
The changed topics and main keywords of the found publications on IPM (RQ-1) are presented in Fig. 2. The found publications were published in the period of 1991 – 2023 (August). The graphs show an increase of new keywords related to new models and methods for portfolio optimization in the publications on the analyzed topic, from 2015.



1. The changed topics over time of the found publications on IPM (RQ-1)

The distribution of the analyzed topic according to **countries** (RQ-2) is presented in Fig. 3. As can be seen, the ten most active countries are as follows: China, the United States, India, Australia, Japan, the United Kingdom France, Spain, Canada, and South Korea. Other countries are those which have less than 140 publications in WoS.

The developed keyword map, which presents the main **methodologies and models** being investigated in IPM (RQ-3) according to (APY), is presented in Fig. 4. Keywords colored in yellow are found in articles published in recent years (APY ~ 2020); green – APY ~ 2018; blue – APY ~ 2014 (see Fig. 4).

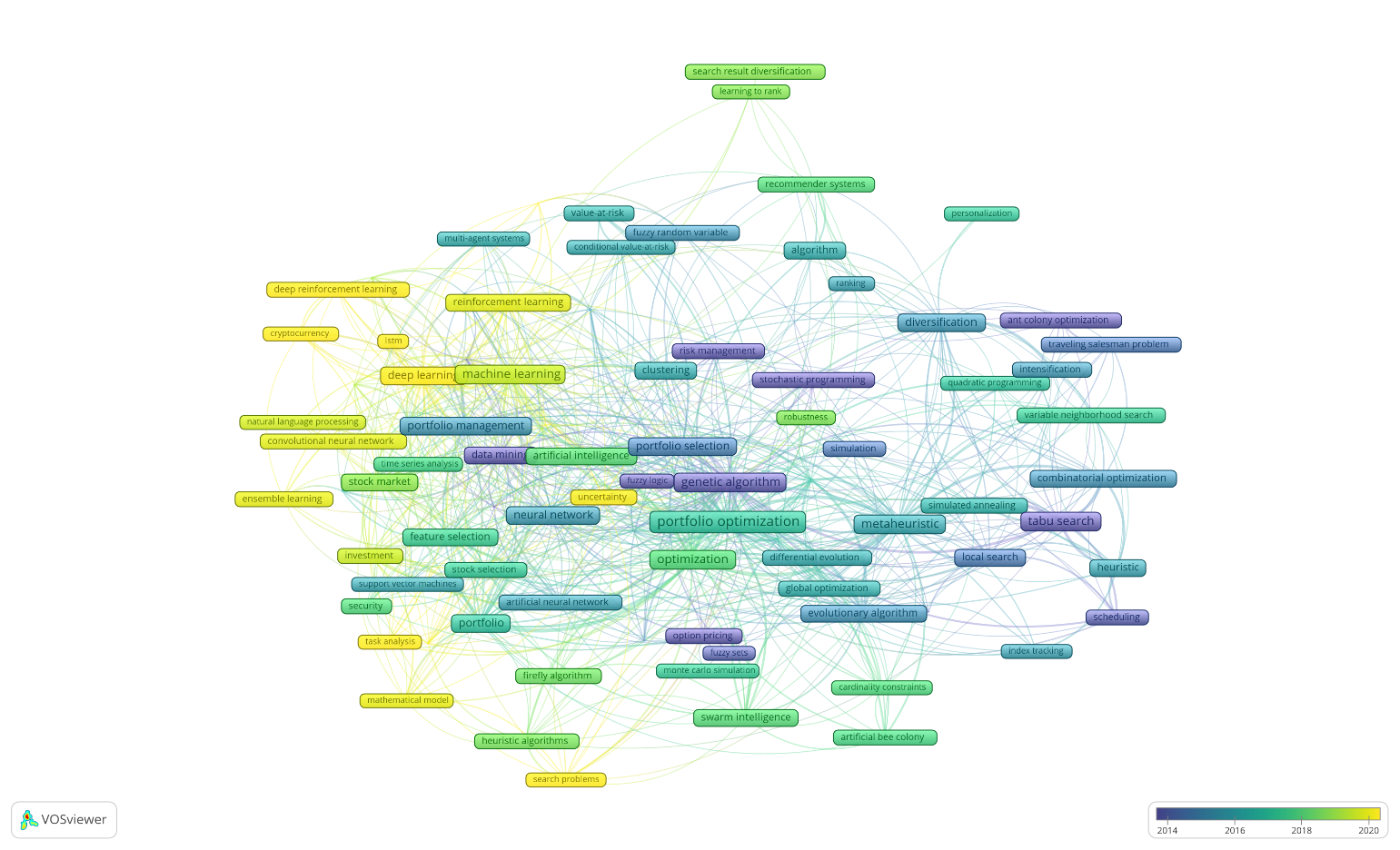


1. Countries analyzing IPM (RQ-2)

Moreover, the bigger frames present the more occurring keywords, and the smaller frames – less occurring keywords, which can be trends or unexpanded and forgotten topics depending on their color.

Yellow frames represent new keywords, i.e., their APY equals current years, which means novelty and new trends. Green and large frames present keywords that occur continuously in all the analyzed periods. Green and small frames present keywords that occur in the period of [2018; 2020]. Blue frames present older keywords.

So, we can see from Fig. 4, the newest keywords (APY (2019; 2021]; color – yellow), like deep learning, deep reinforcement learning, portfolio allocation, machine learning, reinforcement learning, search problems, mathematical model, uncertainty, lstm, convolutional neural network, sentiment analysis, etc.; continuous occurrence of keywords (APY (2017; 2018] color – green), like artificial intelligence, optimization, feature selection, firefly algorithm, swarm intelligence, cardinality constraints, artificial bee colony, swarm intelligence, forecasting, etc.; and older keywords (APY (2014; 2016] color – blue), like, fuzzy logic, data mining, portfolio management, metaheuristic, tabu search, combinatorial optimization, clustering, diversification, ant colony optimization, risk management, support vector machine, financial time series, multi-agent systems etc.



1. The keyword map.
2. results of classification of concepts to INVESTMENT PORTFOLIO AND MODELS, METHODS FOR PORTFOLIO MANAGEMENT

| APY | Investment portfolio | Methods and Models for Portfolio Management |
| --- | --- | --- |
| 2021 | --  task analysis |  |
| 2020 | cryptocurrency | deep learning, deep reinforcement learning, search problems, LSTM |
| 2019 | investment, natural language processing, extraction, portfolio allocation, uncertainty | sentiment analysis, machine learning, convolutional neural network, reinforcement learning, ensemble learning, mathematical model, unsupervised learning, feature |
| 2018 | diversification, stock market, algorithmic trading | firefly algorithm, heuristic algorithms, robustness, learning to rank, search result |
| 2017 | stock selection, forecasting, security, recommender systems, optimization, | quadratic programming, personalization, Monte Carlo simulation, variable neighborhood search, feature selection, time series analysis, artificial bee colony, cardinality constraints, swarm intelligence, artificial intelligence |
| 2016 | classification, multiobjective optimization, portfolio, portfolio optimization | support vector machines, metaheuristic, financial time series, differential evolution, conditional value-at-risk, clustering, iterated local search, multi-agent systems, algorithm, global optimization, value-at-risk, simulated annealing, particle swarm optimization |
| 2015 | Sharpe ratio, finance, diversification, intensification, asset allocation, artificial neural network, index tracking, ranking | evolutionary algorithm, combinatorial optimization, fuzzy random variable, portfolio management, neural network, evolutionary computation, heuristic |
| 2014 | Scheduling, portfolio selection | fuzzy sets, memetic algorithms, simulation, traveling salesman problem, local search, collaborative filtering |
| 2013 | data mining | genetic programming, ant colony optimization, genetic algorithm, stochastic programming |
| 2012 | Risk management |  |
| 2011 | financial prediction | tabu search |
| 2010 |  | Fuzzy logic |

According to the classification of concepts to Investment portfolio (Column 2), Methods and Models for Portfolio Management (Column 3), the newest found terms are presented in Table II. Note that years are APY (Column 1), but not the real year of the paper publication.

Based on Table II, we can see that the analyzed scope that was found in the Investment portfolio scope is as follows:

task analysis, investment, natural language processing, extraction, portfolio allocation, uncertainty diversification, stock market, algorithmic trading stock selection, security, recommender systems, classification, multiobjective optimization, portfolio, Sharpe ratio, finance, asset allocation, artificial neural network, index tracking.

The Investment portfolio *processes* are as follows: forecasting; portfolio selection, simulation; multi-criteria decision making; data mining; identification; diversification, intensification, risk management, optimization, portfolio optimization, ranking, scheduling, and financial prediction.

The found Methods and Models for Portfolio Management terms are as follows:

deep learning, deep reinforcement learning, search problems, LSTM, sentiment analysis, machine learning, convolutional neural network, reinforcement learning, ensemble learning, mathematical model, unsupervised learning, feature, firefly algorithm, heuristic algorithms, robustness, learning to rank, search result, quadratic programming, personalization, Monte Carlo simulation, variable neighborhood search, feature selection, time series analysis, artificial bee colony, cardinality constraints, swarm intelligence, artificial intelligence, support vector machines, metaheuristic, financial time series, differential evolution, conditional value-at-risk, clustering, iterated local search, multi-agent systems, algorithm, global optimization, value-at-risk, simulated annealing, particle swarm optimization, evolutionary algorithm, combinatorial optimization, fuzzy random variable, portfolio management, neural network, evolutionary computation, heuristic, fuzzy sets, memetic algorithms, simulation, traveling salesman problem, local search, collaborative filtering, genetic programming, ant colony optimization, genetic algorithm, stochastic programming, tabu search, Fuzzy logic.

# IV. DISCUSSION

In this research, we have applied SM to answer the main research question: (MRQ1) *What are the predominant methodologies and models employed in investment portfolio management research?* It was decomposed into the following sub-questions: (RQ-1) How has **topics** of IPM and its related models, changed over different time periods? (RQ-2) Are certain methodologies and models more prevalent in specific **geographic regions** within the field of IPM research? (RQ-3) What are the most frequently used **methodologies and models** of IPM research? The conducted SM on the IPM topic shows a significant increase in papers since 2015 in the period of 1991-2023 (till June) (RQ-1). It can be explained by the development of AI techniques and the increase in their application in the analyzed topic.

The analysis of countries investigating IPM (RQ-2) shows that the most active and influencing ten countries are as follows: China, the United States, India, Australia, Japan, the United Kingdom France, Spain, Canada, and South Korea.

In order to answer RQ-3, the mapping with VOSviewer has been performed and the keywords map based on keyword occurrence and APY have been developed. It shows that the newly occurring keywords in the scope of IPM are as follows: cryptocurrency, deep learning, LSTM, mathematical model, convolutional neural network, and ensemble learning.

The most visible and applicable ten methods in IPM are heuristic algorithms, artificial neural networks, collaborative filtering, multiobjective optimization, diversification, tabu search, firefly algorithm, evolutionary algorithm, personalization, and classification.

Summing up, there are a lot of Methods and Models for Portfolio Management Therefore, there is a need for more investigation, into which one or several are the best for Portfolio management.

# V. CONCLUSIONS

The presented SM briefly reviews the current state of the methodologies and models for investment portfolio management. The Methods and Models for Portfolio Management are found and summarized.

From the current SM, it is obvious that effective Methods and Models are an important topic, which should be analyzed further in more detail. Though there are approaches to IPM, there is a need for deeper analysis in future research and better usage of Methods and Models for the management of Investment portfolios.

# VI. BIBLIOGRAPHY

[1] N. Nazareth and Y. V. R. Reddy, “Financial applications of machine learning: A literature review,” *Expert Syst Appl*, vol. 219, Jun. 2023, doi: 10.1016/j.eswa.2023.119640.

[2] Y. Ma, R. Han, and W. Wang, “Portfolio optimization with return prediction using deep learning and machine learning,” *Expert Syst Appl*, vol. 165, Mar. 2021, doi: 10.1016/j.eswa.2020.113973.

[3] W. Chen, H. Zhang, M. K. Mehlawat, and L. Jia, “Mean-variance portfolio optimization using machine learning-based stock price prediction,” *Appl Soft Comput*, vol. 100, Mar. 2021, doi: 10.1016/j.asoc.2020.106943.

[4] Z. Tao and G. Gupta, “Stock Investment Strategies and Portfolio Analysis,” in *PROCEEDINGS OF ACADEMIA-INDUSTRY CONSORTIUM FOR DATA SCIENCE (AICDS 2020)*, G. Gupta, L. Wang, A. Yadav, P. Rana, and Z. Wang, Eds., in Advances in Intelligent Systems and Computing, vol. 1411. 2022, pp. 397–406. doi: 10.1007/978-981-16-6887-6\_32.

[5] N. Maknickienė and D. Sabaliauskas, “Investment portfolio analysis by using neural networks,” Vilnius Gediminas Technical University, May 2019. doi: 10.3846/cibmee.2019.028.

[6] J.-Y. Shyng, H.-M. Shieh, and G.-H. Tzeng, “An integration method combining Rough Set Theory with formal concept analysis for personal investment portfolios,” *Knowl Based Syst*, vol. 23, no. 6, pp. 586–597, Aug. 2010, doi: 10.1016/j.knosys.2010.04.003.

[7] T. K. Lee, J. H. Cho, D. S. Kwon, and S. Y. Sohn, “Global stock market investment strategies based on financial network indicators using machine learning techniques,” *Expert Syst Appl*, vol. 117, pp. 228–242, Mar. 2019, doi: 10.1016/j.eswa.2018.09.005.

[8] J. Choungsirakulwit and D. Sutivong, “Portfolio management of option-based investment in technology research and development,” in *6TH IEEE/ACIS INTERNATIONAL CONFERENCE ON COMPUTER AND INFORMATION SCIENCE, PROCEEDINGS*, R. Lee, M. U. Chowdhury, S. Ray, and T. Lee, Eds., 2007, p. 732+. doi: 10.1109/ICIS.2007.150.

[9] V. V Dombrovsky and E. A. Lashenko, “Robust control of linear systems with random parameters and multiplicative disturbances with application to the investment portfolio management,” in *SICE 2003 ANNUAL CONFERENCE, VOLS 1-3*, 2003, pp. 1116–1121.

[10] A. Gunjan and S. Bhattacharyya, “A brief review of portfolio optimization techniques,” *Artif Intell Rev*, vol. 56, no. 5, pp. 3847–3886, May 2023, doi: 10.1007/s10462-022-10273-7.

[11] K. Petersen, S. Vakkalanka, and L. Kuzniarz, “Guidelines for conducting systematic mapping studies in software engineering: An update,” *Inf Softw Technol*, vol. 64, pp. 1–18, Aug. 2015, doi: 10.1016/j.infsof.2015.03.007.

[12] M. K. Linnenluecke, M. Marrone, and A. K. Singh, “Conducting systematic literature reviews and bibliometric analyses,” *AUSTRALIAN JOURNAL OF MANAGEMENT*, vol. 45, no. 2, pp. 175–194, May 2020, doi: 10.1177/0312896219877678.

[13] D. Kalibatiene and J. Miliauskaite, “A Hybrid Systematic Review Approach on Complexity Issues in Data-Driven Fuzzy Inference Systems Development,” *INFORMATICA*, vol. 32, no. 1, pp. 85–118, 2021, doi: 10.15388/21-INFOR444.

[14] N. J. van Eck, L. Waltman, E. C. M. Noyons, and R. K. Buter, “Automatic term identification for bibliometric mapping,” *Scientometrics*, vol. 82, no. 3, pp. 581–596, Mar. 2010, doi: 10.1007/s11192-010-0173-0.

[15] N. Jan van Eck and L. Waltman, “VOSviewer Manual,” 2023.