International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 10 | Oct - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Investment Analytics: A Bibliometric Review

Dr Richa Kalpesh Saxena[1], Aadi Jain[2], Aarit Shah[3], Aarya Chandak[4], Aaryan Lahoti[5], Abhinav Nair[6]

[1]Assistant Professor, Department of Marketing & Operations, Anil Surendra Modi School of Commerce, SVKM's Narsee Monjee Institute of Management Studies (NMIMS) Deemed-to-be-University, Mumbai, India,

[2, 3, 4, 5, 6] BBA, Anil Surendra Modi School of Commerce, SVKM's Narsee Monjee Institute of Management Studies (NMIMS) Deemed-to-be-University, Mumbai, India.

Abstract - This paper offers an expansive mapping of research on investment and analytics, examining how analytics impacts investment decisions and the value of investing in analytics. The review ranges through Scopus journal articles and revisits spanning 2017-2026 utilizing bibliometric techniques to categorize output, sources, countries, authors, and a keyword network. The reviewed corpus included 2,071 documents with an annual output ranging from 59 in 2017 to 483 in 2026 with 45.3% published in 2025 or 2026. The United States, China, and India lead mentions, while "investments" appears as the most common keyword. The literature covers analytics capability and firm performance, machine learning in financial decisions, supply chains and sustainability, and human-centered applications. The source section also indicates growth in artificial intelligence, ethics, and environmental, social, and governance domains. This paper reviews review methods and findings, then proposes research on realised value, adaptive appraisal under uncertainty, accountability, sustainability, and sectoral application. The findings are predominantly descriptive, the dataset is limited to English journal articles and reviews, and 2026 is incomplete.

Key Words: investment, analytics, bibliometric analysis, artificial intelligence, machine learning, sustainability.

1. INTRODUCTION

Investment decisions lie at the heart of organisational decision-making, as firms, financial institutions and governments must make choices about the allocation of limited resources. In the past, investment decisions were made based on financial information, managerial experience and professional judgement. However, the increasing availability of big, complex sets of data, combined with analytic processing, has changed the ways in which investment decisions can be supported and evaluated. This paper aims to establish links between analytics and investment in two ways: first, analytics is a tool which helps investors to assess the markets, estimate returns, evaluate risks and structure portfolios; second, analytics is something which has to be purchased every model which supports investment decisions has to be bought on the basis of a previous investment decision.

Sivarajah et al. (2017) reviewed the challenges of big data and the analytical methods used between 1996 and 2015. The study showed that organisations may invest heavily in big-data technologies without knowing their proper uses or value. Hence, analytics is not only a technical matter but also an investment choice. The paper has 1,736 citations, the most in the present data set. Similarly, Trieu (2017) reviewed empirical research on business intelligence and proposed a value chain linking investments to assets, impacts and organisational performance. However, the literature stayed scattered, as individual studies addressed different stages without an integrated view. Together, these studies suggest that realised value cannot be understood by the amount invested alone.

Later studies tested whether this value chain works. Müller et al. (2018) studied 814 companies between 2008 and 2014 and measured big data and analytics assets objectively. They found a 3–7% rise in firm productivity but it was only seen in IT-intensive or highly competitive industries. Hence, the payoff depends on industry and firm situation. The second role of analytics, supporting investment decisions, was shown by Ma et al. (2021). They used machine learning and deep learning models to predict stock returns and build portfolios using nine years of China Securities 100 Index data. Random forest portfolios performed best. However high trading turnover removed nearly half of the returns showing that good prediction does not always mean good investment and that every tool must be judged against its cost.

Spending on analytics and AI is very large now. The Stanford Institute for Human-Centred Artificial Intelligence (2026) reports that global corporate AI investment more than doubled in 2025 to US$581.69 billion which is made up of US$344.66 billion in private investment and US$214.44 billion in mergers and acquisitions. It also states that 88% of organisations used AI in 2025 compared with 55% in 2023. The International Monetary Fund (2026) estimates that AI-linked investment added about 0.5 percentage points to United States GDP growth in 2025. It also estimates that private AI investment could pass US$2 trillion worldwide in 2026. However, these figures show how much is spent and not how well it is spent.

The review was motivated by three developments. First, the field has grown rapidly. The authors saw 2,071 documents on investment and analytics appeared between 2017 and 2026 with 938 (45.3%) of them in 2025 and 2026. The leading author keyword shifted from big data to machine learning and data analytics, then to artificial intelligence (112 of 938 documents) hence so single-term searches can miss studies (Kamis, 2026). Second, expectations now extend beyond profit. Higher AI focus was linked with better environmental, social and financial performance but not clearly governance (Giordino et al., 2026). Weak governance and poor ESG data raised financing costs by 9-15% for emerging-market SMEs (Narayan et al., 2025). Ethics and ESG appear in 106 and 55 publications respectively mostly ranging from 2025-2026. AI also poses societal risks such as algorithmic bias (Galaz et al., 2021). Third, data-centre electricity use grew 17% in 2025 against 3% globally and is expected to rise from 485 TWh to about 950 TWh by 2030 (International Energy Agency [IEA], 2026), making investment a decision about energy, cost and sustainability.

The research on investment and analytics can be grouped into four main streams. The first stream looks at analytics as a costly choice that must be justified. (Mikalef et al., 2019,

2020) built the idea of big data analytics capability and linked it with firm performance. (Raguseo & Vitari, 2018) studied how investments in big data analytics affect the financial performance of firms, while (De Luca et al., 2021) found that many firms that invest in big data fail to reach the benefits that they expect. (Yalcin et al., 2022) added that business analytics systems are a major investment and showed how multi-criteria decision methods can help managers to choose between them. The second stream uses machine learning for financial decisions, such as stock prediction, sentiment analysis and portfolio building (Rouf et al., 2021; Sohangir et al., 2018; Vo et al., 2019) The third stream links analytics with supply chains, sustainability and ESG. (Kamble et al., 2020) studied agri-food supply chains, (Spieske & Birkel, 2021) looked at supply chain resilience and (Bindeeba et al., 2025) combined the results of 59 studies to examine how digital change affects ESG performance. The fourth stream covers human-centred uses for example in health care (Khalifa & Albadawy, 2024). Most of these streams agree that analytics creates value only under certain conditions, such as the type of industry, the skills of the firm and the quality of the data. However, they differ in the questions they ask, the theories they use and the words they choose, and the links between them are seldom stated.

Many reviews exist on parts of this field, but they are usually limited to one sector or one technology. (Kamble et al., 2020) looked at agri-food supply chains, (Nobanee et al., 2021) at big data in banking, (Seow, 2025) at ESG and machine learning, and (Khalifa & Albadawy, 2024) at health care. Hence, despite the large amount of research, few reviews, to the best of the authors' knowledge, look at the whole field of investment and analytics across sectors, countries and sources, or show how its language and themes have changed over time. Moreover, the field is still young, and its terms are not yet fixed. The present review aims to fill this gap. The primary goal is to describe the current status of academic research on investment and analytics from 2017 to 2026. The scope is limited to journal articles and reviews that are indexed in Scopus and written in English. The following questions guide the review. RQ1: What are the trends in publications and citations on investment and analytics? RQ2: What are the main themes, sources, authors and countries in this field? RQ3: How have the terms and themes changed over time, and what gaps remain for future research?

The review follows the PRISMA method (Moher et al., 2010). The authors searched the Scopus database on 11 September 2026 using “Investment AND Analytics”, which returned 5,252 records. After retaining only journal articles and reviews and removing records dated 2027 and two duplicate titles, 2,071 documents (1,874 articles and 197 reviews) published between 2017 and 2026 remained. The data were analysed using the R bibliometrix package (Aria & Cuccurullo, 2017) and VOSviewer for main information, annual scientific production, a three-field plot, country production and keyword co-occurrence. Bibliometric analysis applies statistical methods to publication data (Donthu et al., 2021) and suits large document sets. Annual output grew from 59 documents in 2017 to 483 in 2026, at an annual growth rate of 26.32%. The United States, China and India lead in country mentions, investments is the most common keyword (554 occurrences), and the leading author keyword has recently moved towards AI. The review benefits researchers, students, managers and policymakers by mapping a scattered field and pointing to future work. The paper proceeds with the literature review (Section 2), research methodology (Section 3), data analysis (Section 4) and a future research agenda (Section 5), and concludes with limitations in Section 6.

2. LITERATURE REVIEW

2.1 Overview of the research

A large and growing body of literature has examined how organisations use, and pay for, big data analytics and artificial intelligence (AI). A bibliometric analysis of the collected publications found 2,076 documents (1,878 articles and 198 reviews), all in English, published between 2017 and 2026 and cited 44,409 times. Output rose from 59 publications in 2017 to 455 in 2025, and 128 of the 198 reviews (65%) appeared in 2025 and 2026. Among the 433 review papers examined, 248 (57%) come from those two years, against 35 from 2017 to 2019.

Authors based in the United States appear on 456 publications, followed by China (270) and India (242). Sustainability (Switzerland) is the most common journal with 42 publications, and the review papers are spread over 347 journals. Topics have shifted: big data led the author keywords in 2017 to 2020, machine learning in 2021 to 2024, and artificial intelligence in 2025 and 2026 (112 of 938 publications). Generative AI terms appear in 54 publications, 47 of them from the last two years, and ethics rose from four publications in 2017 to 2020 to 83 in 2025 and 2026.

2.2 Themes shown by the network map

The network map helped identify five groups of connected keywords. The largest, in red, joins investments, big data, digital transformation, the internet of things, blockchain and sustainability. The blue group covers machine learning, deep learning, forecasting and financial markets. The green group covers human, humans, public health and covid-19. A smaller yellow group links artificial intelligence with systematic review, and a small purple group holds data privacy. Artificial intelligence sits between the red and green groups with links to both.

2.3 Big data analytics and business value

Much of the literature on big data analytics is concerned with how organisations gain value from it. Sivarajah et al. (2017), the most cited paper with 1,736 citations, reviewed the challenges and methods of big data analytics to support sound investment decisions. Trieu (2017) found that firms appear to gain value through a chain that runs from investments to assets, impacts and firm performance. Yalcin et al. (2022) treated business analytics as a significant investment and showed how multi-criteria methods help choose between alternatives. Decision making is a common thread, as 132 of the 433 reviews refer to it.

2.4 Machine learning, finance and investment

The blue group shows that machine learning is widely used in finance. Rouf et al. (2021) surveyed 145 works over the last 10 years on stock market prediction, finding that text analytics and ensemble had improved accuracy considerably. Nobanee et al. (2021) reported on 60 papers on big data in banking, with investment, credit risk and fintech among the themes. Garad et al. (2024) found that investments in information management promoted financial innovation where, for example, AI facilitates predictive analytics and the personalisation of customer-centric services. Stock market references appear in 118 publications; fintech 56. Taken together these papers show that machine learning has become ubiquitous for solving investment problems.

2.5 Sustainability, digital transformation and supply chains

The red group also links analytics to sustainability and operations. Sustainability, ESG or green topics appear in 142 review papers, and the internet of things in 75. Kamble et al. (2020) reviewed 84 papers on data-driven agri-food supply chains and proposed a framework in which visibility and resources drive analytics capability. Spieske and Birkel (2021) found that big data analytics is particularly suitable for improving supply chain resilience. Seow (2025) reviewed 57 articles on machine learning and ESG, and Bindeeba et al. (2025) found in a meta-analysis of 59 studies that digital transformation raises ESG scores, with the largest gains from process digitisation.

2.6 Health care and other human-centred areas

The green group comes mainly from health care, with 102 review papers mentioning health terms. Khalifa and Albadawy (2024) reviewed 30 studies and found that AI improves image analysis, speed and cost in diagnostic imaging. Mauro et al. (2024) found that the internet of things and AI have the greatest effect on administrative processes in health organisations, with staff skills as the main condition for success. Education (48 review papers), human resources (43), agriculture (46) and energy (43) also appear regularly.

2.7 Consumers and marketing

Consumer studies form a smaller part of the material. Retail, e-commerce, consumer and customer terms appear in 299 publications, and 59 review papers mention consumers or customers. Islam et al. (2024) screened 2,850 records, included 150 studies, and found that AI personalises the customer experience and helps anticipate consumer behaviour, while privacy remains a concern, with 84 publications on the topic. Mariani and Borghi (2021) analysed almost 50,000 online reviews of 19 hotels and found that reviews mentioning service robots were linked to higher ratings.

2.8 Methods used in the review literature

The bibliometric analysis found that at least 121 (28%) of the 433 review papers are systematic reviews, 41 of them using PRISMA, and at least 51 (12%) use bibliometric methods. Another 154 (36%) describe themselves as narrative, comprehensive or critical reviews, and some papers fit more than one type. Systematic and bibliometric work is rising, with 81 of the 121 systematic reviews and 37 of the 51 bibliometric studies from 2025 or 2026. Baker et al. (2020) show the bibliometric approach, mapping 25 years of the Review of Financial Economics into six clusters, and Bindeeba et al. (2025) show the meta-analysis approach.

All of these studies indicate that the field of analytics and AI is evolving rapidly and is very broad in its reach across disciplines. The biggest areas of interest include investments, forecasting, sustainability, and health care as a separate category. AI links these domains, and the development of generative AI, along with ethics and ESG, points to changing areas of interests.

3. RESEARCH METHODOLOGY

A systematic literature review (SLR) identifies, selects and synthesises research on a defined topic in a transparent and repeatable way (Paul & Criado, 2020). This study applies a systematic search and screening process to the Scopus database, reported according to the PRISMA framework (Moher et al., 2010), and analyses the resulting records using bibliometric methods (Donthu et al., 2021). The aim is to map the field of investment and analytics, covering its growth, sources, countries, authors and themes. It does not assess the quality of individual studies or synthesise their findings.

3.1 Article Selection, Inclusion, and Exclusion Criteria

A search of the database was conducted on 11th September 2026 using Scopus. The oldest paper related to our topic was published in 1974. We took a year range of 10 years, which was 2027-2017. This range was chosen to include ample data without the inclusion of outdated data.

The first stage of the PRISMA process involved the identification of relevant papers. A total of 5,252 records were identified from Scopus using the keywords:

"Investment AND Analytics"

The purpose of this stage was to establish an initial pool of records before applying our inclusion and exclusion criteria. This approach helps reduce the possibility of prematurely excluding potentially relevant literature.

After creating a pool of 5,252 records, the records go through the first inclusion and exclusion criteria which is based on document type. For the purpose of this study, the analysis was restricted to Articles and Review Papers. We excluded Books, Book Chapters and Conference Papers. This reduced the initial count from 5,252 records, down to 3,014 records. It excluded a total of 2,238 records that would have included books, Conference Papers etc. since they did not fit into the selected document type. This criterion was used to improve the consistency and academic relevance of the data set by targeting only articles and reviews. As a result, other publications that may not provide the same level of scholarly contributions or bibliographic information required by the intended analysis were excluded.

The next step saw the application of a journal filter to the remaining records. The rationale behind this criterion was that the final data set should contain journal-based scholarly publications. Journal literature was seen by the researchers as an appropriate publication type as it is a consistent publication format and contains the bibliographic information required to carry out the subsequent bibliometric analysis. It is therefore a way to ensure academic consistency within the field of study. Following on the document-type screening there were 3,014 records left. Of these 936 records were excluded on the basis of the journal criteria. This resulted in a reduction to 2,078 records. The journal filter helped further refine the dataset and ensured that the publications included in the analysis were consistent in terms of source type.

We also filtered the records on the basis of Language, only using Articles and Reviews written in English. The records we excluded included papers written in Chinese, Russian etc. We placed this criterion so that the analysis of the data would not be skewed or not be able to generate outputs because of Language restrictions.

During filtering out the data, we also found out which fields of work used similar keywords which we used while writing their research papers, articles and reviews. Some of the fields were "Business, Management and Accounting", "Computer Science", "Engineering", and "Social Sciences". This provided us with the information that our keywords are diversified between a lot of fields of work and is a well-researched area of interest.

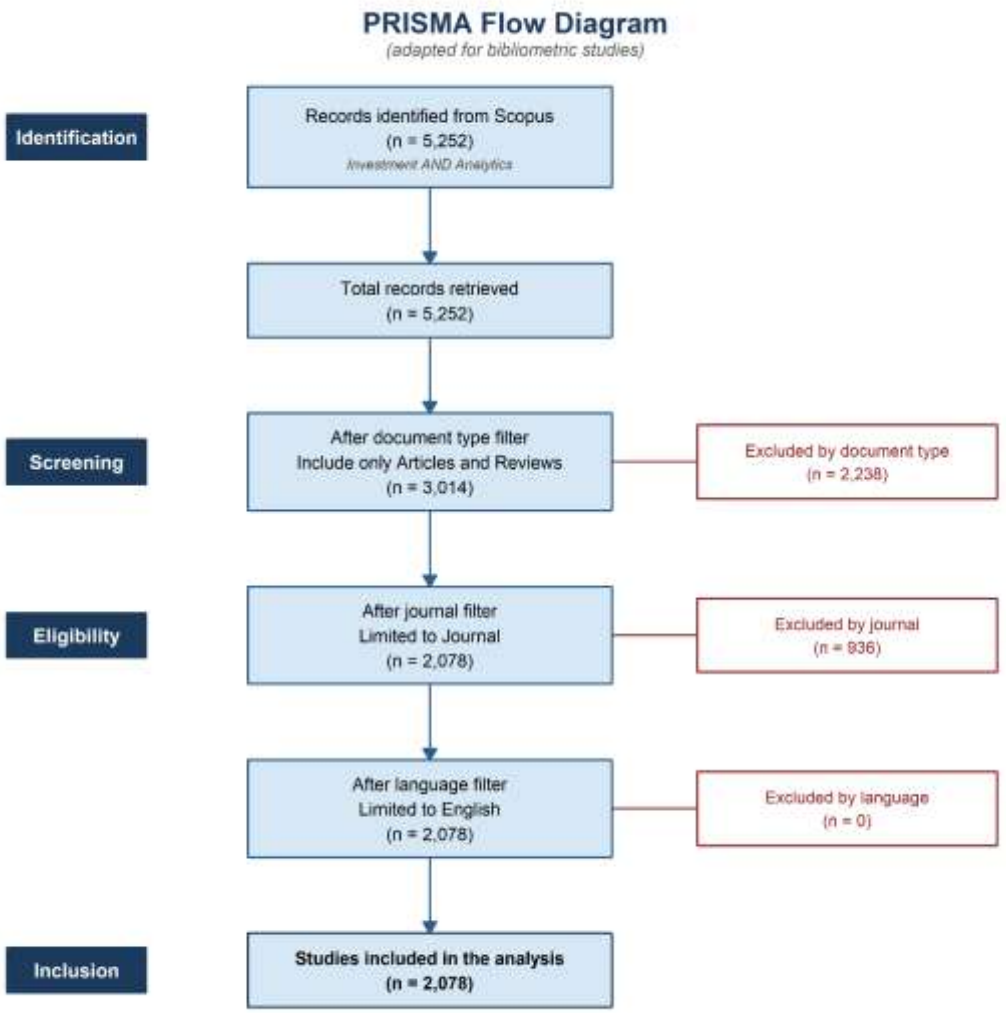


Fig -1: PRISMA flow diagram of the literature-selection process

The PRISMA flow diagram (Figure 1) was used to provide a clear and transparent representation of the literature-selection process. It documents both the number of publications retained and the number excluded at each stage, allowing readers to understand how the final dataset was constructed.

The overall process can be summarised as follows: 5,252 records were initially identified through Scopus. 2,238 records were excluded based on document type, leaving 3,014 records. A further 936 records were excluded through the journal filter, resulting in 2,078 records. Finally, the language criterion was applied, with zero records excluded, leaving all 2,078 records eligible for inclusion. These 2,078 records will be used to conduct our data analysis.

4. DATA ANALYSIS

4.1 Main Information about Data

The corpus was drawn from Scopus on 11 September 2026 and screened through the PRISMA framework (Moher et al., 2010). Identification returned 5,252 records; excluding books, chapters and conference papers removed 2,238 and a journal filter a further 936, leaving 2,078. Table 1 indicators are generated within the bibliometrix environment (Aria & Cuccurullo, 2017), following the staged science mapping workflow recommended for bibliometric studies (Zupic & Čater, 2015)The year 2027 is dropped, having returned only five papers and two duplicates.

Table -1: Main information for the corpus, from the Scopus export

Description	Results
Timespan	2017:2026
Sources (Journals, Books, etc.)	1190
Documents	2071
Annual Growth Rate %	26.32
Document Average Age	2.79
Average citations per doc	21.44
References	96997
Keywords Plus (ID)	8217
Author's Keywords (DE)	6922
Authors	7755
Authors of single-authored docs	215
Single-authored docs	223
Co-Authors per Doc	3.97
International co-authorships %	31.39
article	1874
review	197

At 1.74 documents per source the journal base is broad rather than concentrated, and at 3.97 co-authors per document the typical output is a small team.

Two features draw attention before the findings are read. The first is the gap between the PRISMA count and the analysed count, the difference of seven being the 2027 records and two duplicate titles. The second is citation concentration: an average of 21.44 per document sits alongside a heavily skewed distribution, with a single 2017 paper on big data challenges carrying 1,736 citations (Sivarajah et al., 2017). Section 4.5 should be read with that concentration in mind.

4.2 Annual Scientific Production

Table -2: Annual scientific production, 2017 to 2026

Year	Articles
2017	59
2018	69
2019	121
2020	138
2021	199
2022	156
2023	166
2024	225
2025	455
2026	483

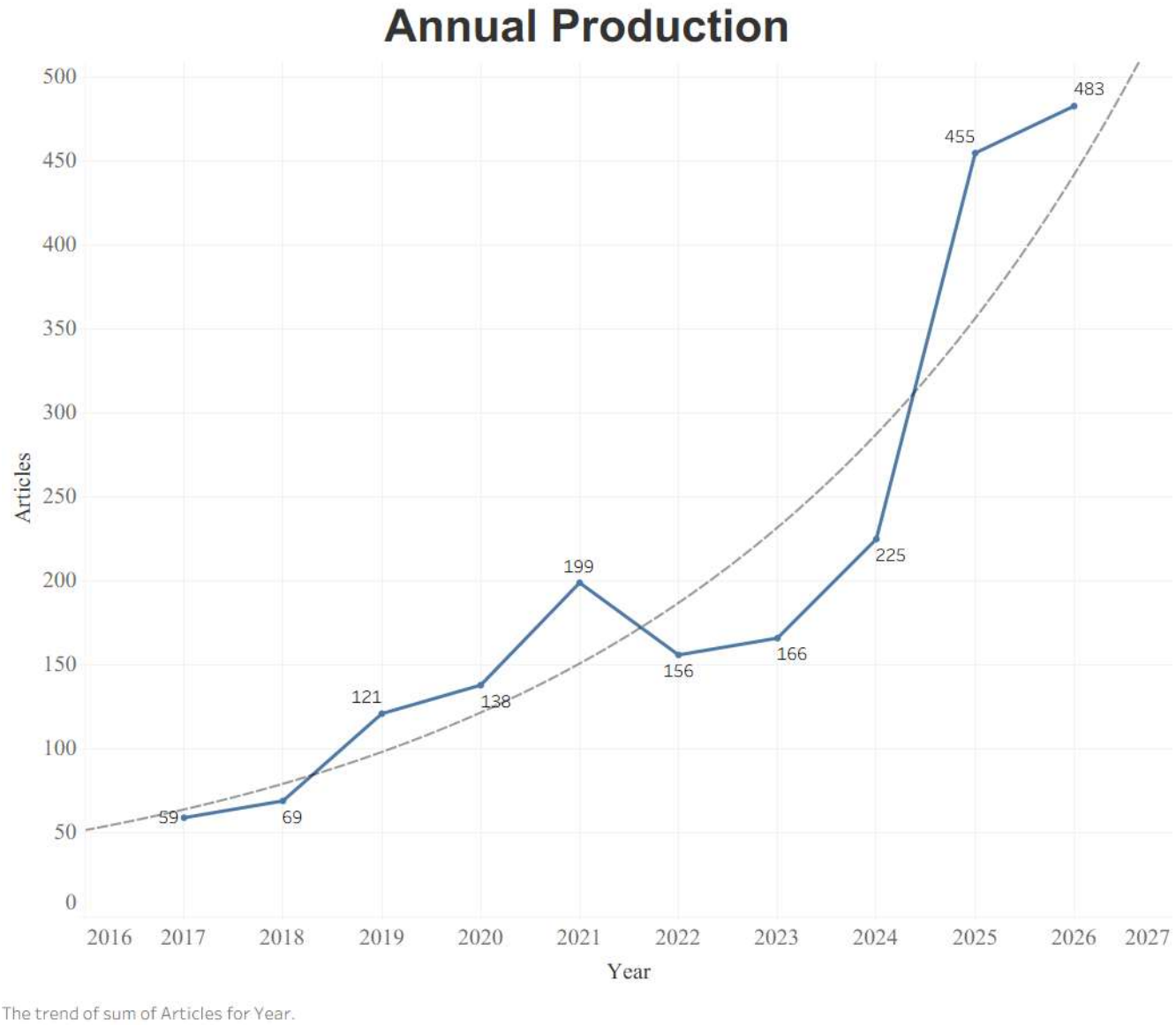


Fig -2: Annual scientific production with the fitted log-linear trend line

As Table 2 and Figure 2 show, the series runs from 59 documents in 2017 to 483 in 2026, a compound annual growth rate of 26.32%, and the ten values sum to 2,071, matching Table 1.

Growth is not uniform across the decade. The early phase adds 10 documents, a 16.9% increase, and is better described as a low base than a growth phase. The first substantial step comes in 2019, when output rises by 52 to 121, up 75.4%, followed by a second acceleration in 2021 at 199. The only contraction falls in 2022, where output declines by 43 to 156, down 21.6%, after which recovery is immediate, at 166 in 2023 and 225 in 2024.

The largest single-year gain occurs in 2025, where output rises by 230 documents to 455, up 102.2%, and 2026 adds a further 28 to 483, or 6.2%. The two most recent years carry 938 documents, 45.3% of the corpus.

An exponential trend fitted to the natural logarithm of article counts against year returns a Year coefficient of 0.215382 (SE 0.0256317, t = 8.40293, p < 0.0001) and an intercept of -430.273 (SE 51.8146, p < 0.0001), with an R-Squared of 0.898231. It is descriptive, not compared against alternative curve forms, indicating proportional rather than additive growth rather than testing any growth law.

Two documents are worth singling out, because they supply much of the vocabulary the keyword analysis returns. The 2017 paper on big data challenges is the earliest work in the corpus, supplying the challenge framing later analytics work builds on (Sivarajah et al., 2017), while the 2019 step is accompanied by the two studies that established the big data analytics capability construct, which recurs through the keyword structure in Section 4.5 (Mikalef et al., 2019, 2020).

4.3 Three-Field Plot

The three-field plot (Figure 3) maps the corpus through three linked fields, namely authors, author countries and merged keywords, generated through the bibliometrix environment (Aria & Cuccurullo, 2017). It follows the science mapping rationale of representing a field's structure as relationships between elements rather than a narrative classification (Chen, 2017). Each author connects to every country in which they publish, each pairing carrying that country's document keywords; circle size is frequency, line thickness the strength of association.

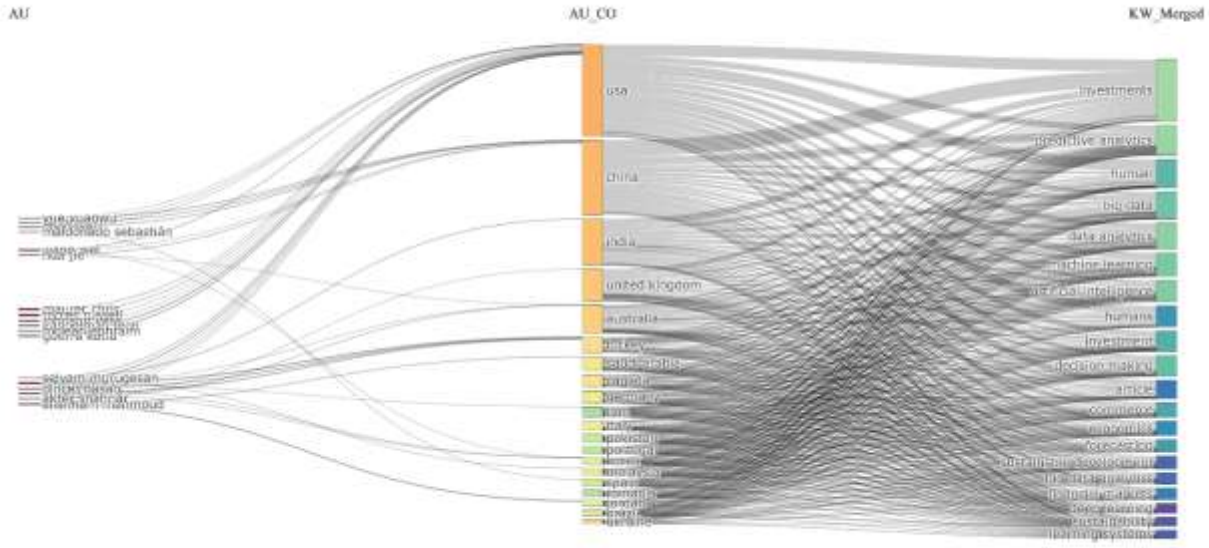


Fig -3: Three-field plot: authors (AU), author countries (AU_CO), merged keywords (KW_Merged)

Six of the twelve leading authors resolve to the United States (Table 3), five of them with no other affiliation.

Table -3: Author-country correspondences for the leading authors in Figure 3

Author as shown in plot	Documents in corpus	Countries of affiliation	Dominant country
Wang Wei	11	China (11), Taiwan (5), United States (4)	China
Johnson Vess	10	United States (10)	United States
Wang Hongwei	8	China (7), United States (4), Taiwan (2)	China
Yue Xuanwu	9	China (8), Hong Kong (6), United States (5)	China
Maurer Chris	9	United States (9)	United States
Torres Russell	9	United States (9)	United States
Chen Wei	5	China (5), Hong Kong (3), Singapore (2)	China
Kappelman Leon	6	United States (6)	United States
McLean Ephraim	6	United States (6)	United States
Guerra Katia	6	United States (6)	United States
Hua Jie	4	China (4), Australia (1)	China
Maldonado Sebastián	4	Chile (4), Belgium (2), Spain (1)	Chile

Maurer, Torres, Johnson, Kappelman and McLean are co-authors on the same pair of annual information-systems studies (Kappelman et al., 2018, 2021). The plot therefore renders one recurring team as five separate circles, their sizes counting teams as much as individuals. Maldonado's four papers sit in Chile and Belgium, and Yue Xuanwu's nine span China, Hong Kong and the United States.

Connection load is heaviest at the United States, then China and India. Chinese-affiliated work contributes the portfolio and visual analytics material on the right (Yue et al., 2020), Indian-affiliated work the data-driven supply chain and analytics capability studies (Fosso Wamba & Akter, 2019; Kamble et al., 2020). Multi-country papers sit at the junctions, as in the six-country study of systemic risk (Galaz et al., 2021) and the Norway and United States capability study (Mikalef et al.,

2020). Section 4.4 therefore examines geography before Section 4.5 turns to subject.

4.4 Country Scientific Production

Table -4: Country scientific production, with frequency counts and rank

Rank	Country	Frequency
1	USA	1024
2	CHINA	591
3	INDIA	542
4	UK	368
5	AUSTRALIA	332
6	UKRAINE	166
7	TURKEY	137
8	CANADA	133
9	MALAYSIA	132
10	SAUDI ARABIA	125
11	GERMANY	106
12	ITALY	106
13	BRAZIL	86
14	JORDAN	82
15	SOUTH KOREA	74

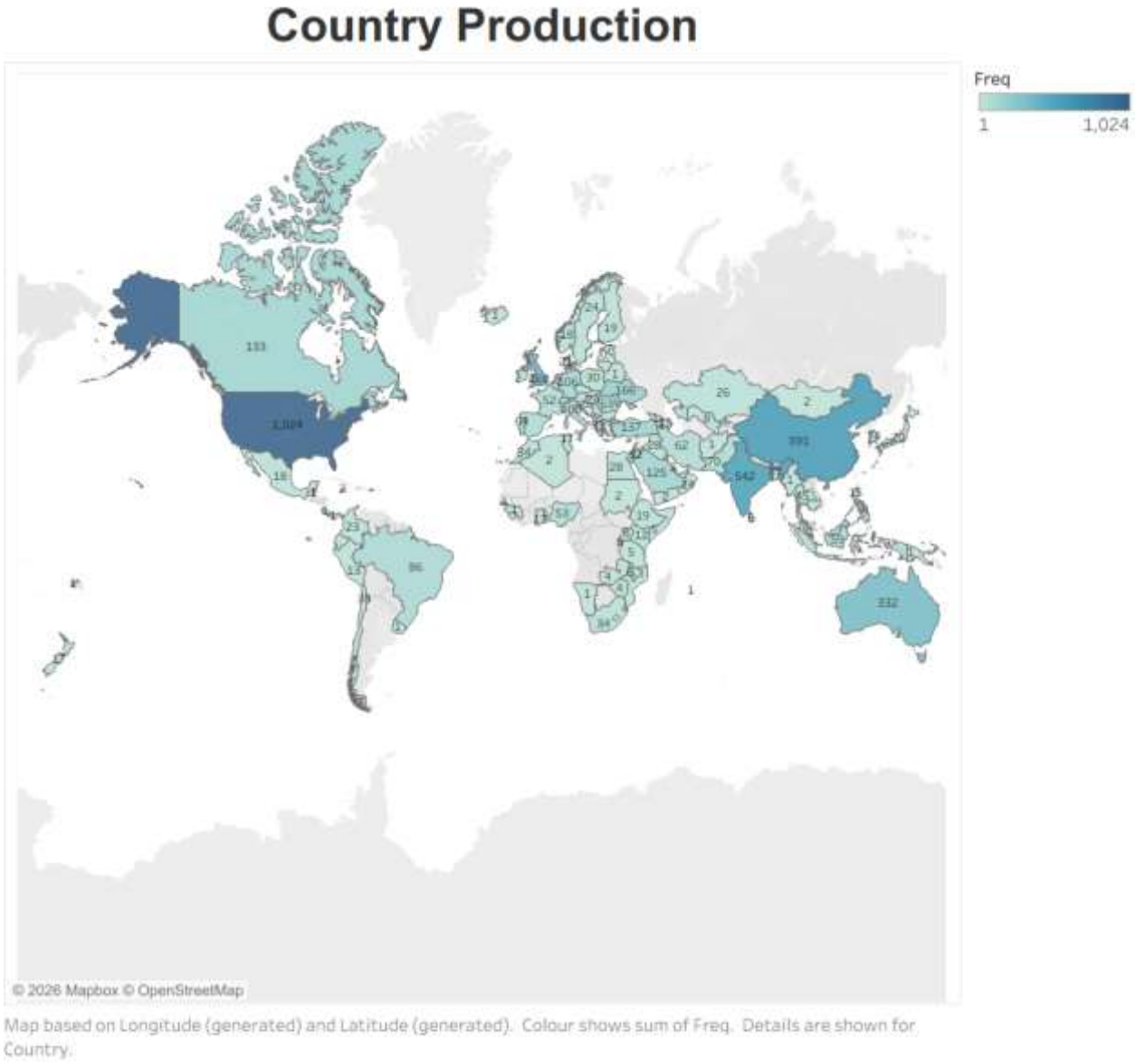


Fig -4: Country scientific production by number of country mentions

Table 4 covers 116 countries, and the frequencies sum to 5,602, exceeding the document count because a paper with co-authors in several countries counts once for each. The shares below are therefore shares of country mentions rather than of documents.

As Figure 4 shows, the United States contributes 1,024 mentions, 18.3% of the total and 49.4% of documents, followed by China at 591, India at 542, the United Kingdom at 368 and Australia at 332. Below them the pattern is less orderly: the top ten hold 63.4% of mentions, and sixty-three countries clear ten documents while twenty-five have exactly one. Ukraine's sixth place at 166 mentions, ahead of Turkey and Canada, departs from the pattern, and nothing in the supplied data accounts for it.

International co-authorship at 31.39% is consistent with 2.71 countries per document.

Output and citation weight point in different directions: the most-cited record is UK-affiliated (Sivarajah et al., 2017), although the United Kingdom ranks fourth by volume, and the two records behind it are India and United States collaborations (Kamble et al., 2020; Mikalef et al., 2020), with

a six-country study of systemic risk also in the top ten (Galaz et al., 2021). Cross-border work therefore carries the citation core while the largest national contributors carry the volume.

4.5 VOSviewer Comparative Analysis

Figures 5 to 8 give four co-occurrence maps of the same corpus, at four thresholds. Each circle is a keyword, sized by the number of documents carrying it; line thickness and distance give co-occurrence strength, colour gives cluster membership, and the shading overlay dates it. Only the surviving keywords change between maps.

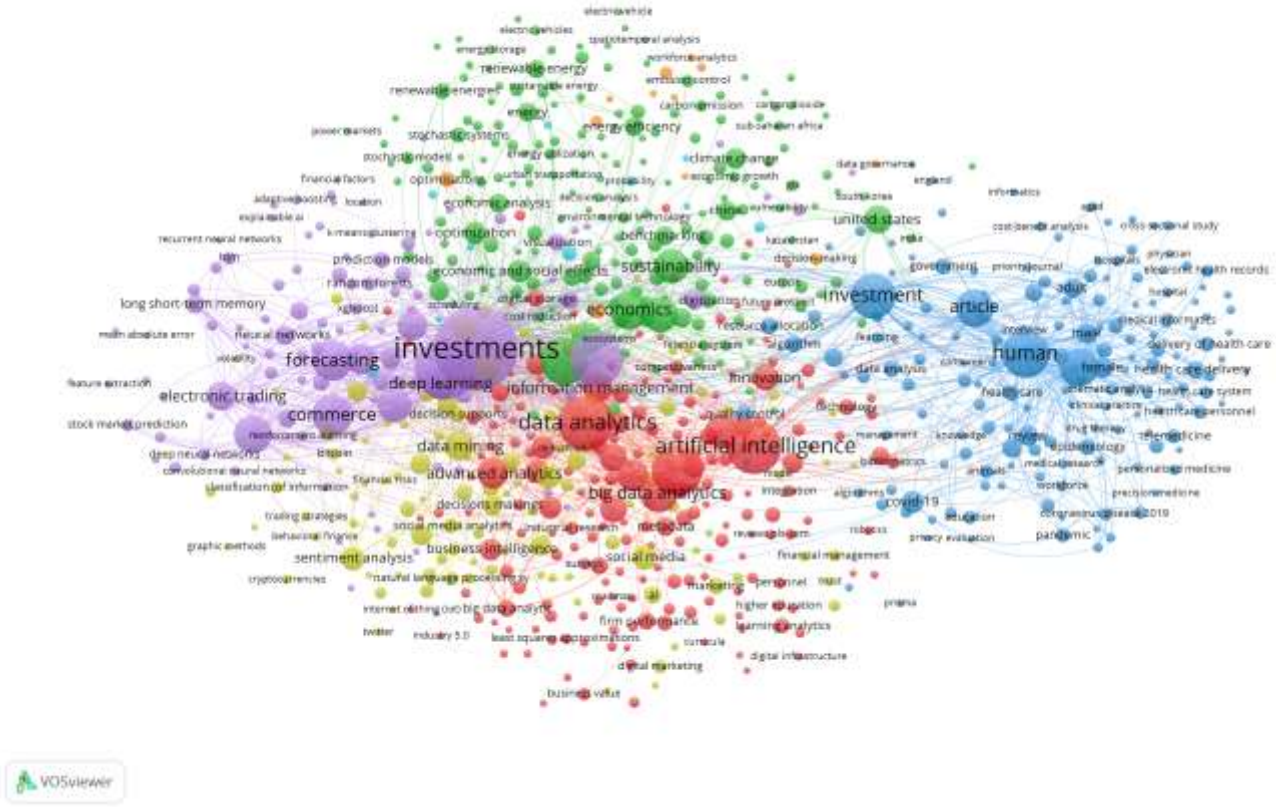


Fig -5: Keyword co-occurrence network at a minimum of 5 occurrences. Seven clusters form, among them analytics and artificial intelligence, economics and sustainability, investments and machine learning, and two peripheral groups.

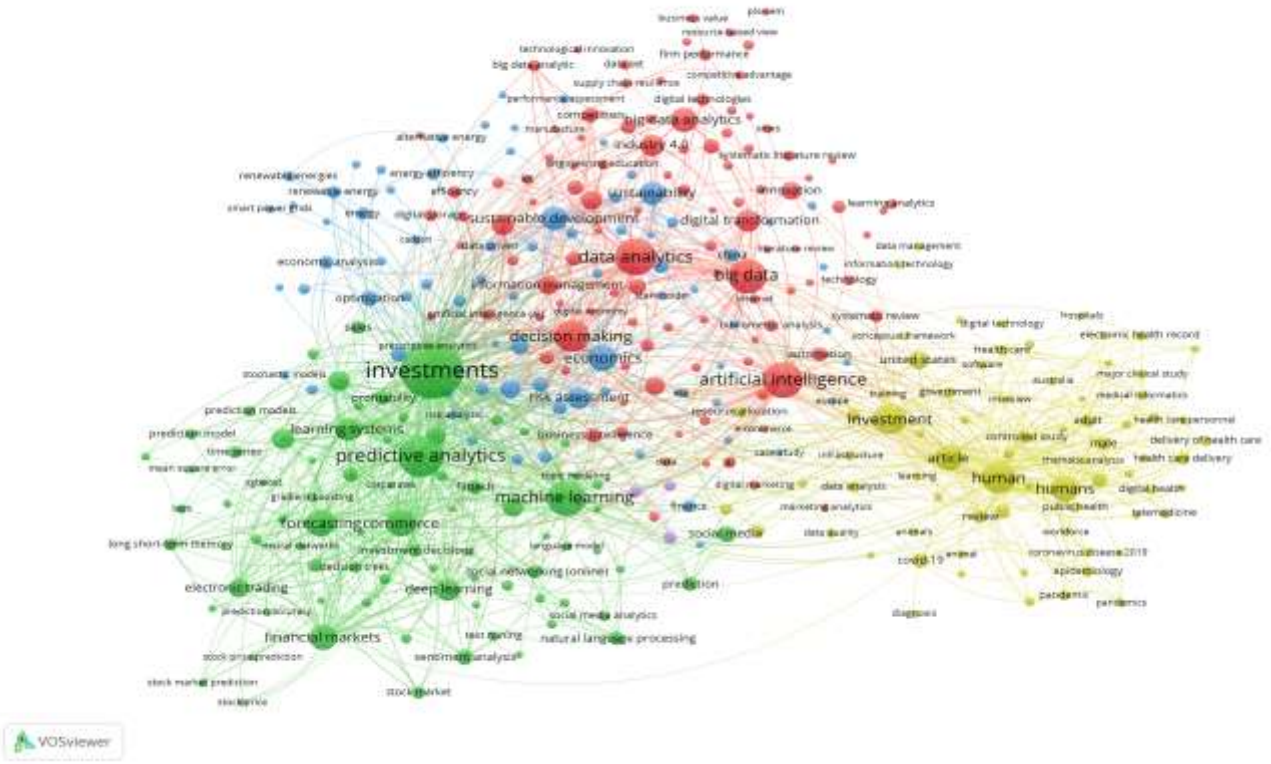


Fig -6: Keyword co-occurrence network at a minimum of 10 occurrences. Five clusters form; economics and risk assessment merges with analytics, while predictive modelling separates from the largest cluster.

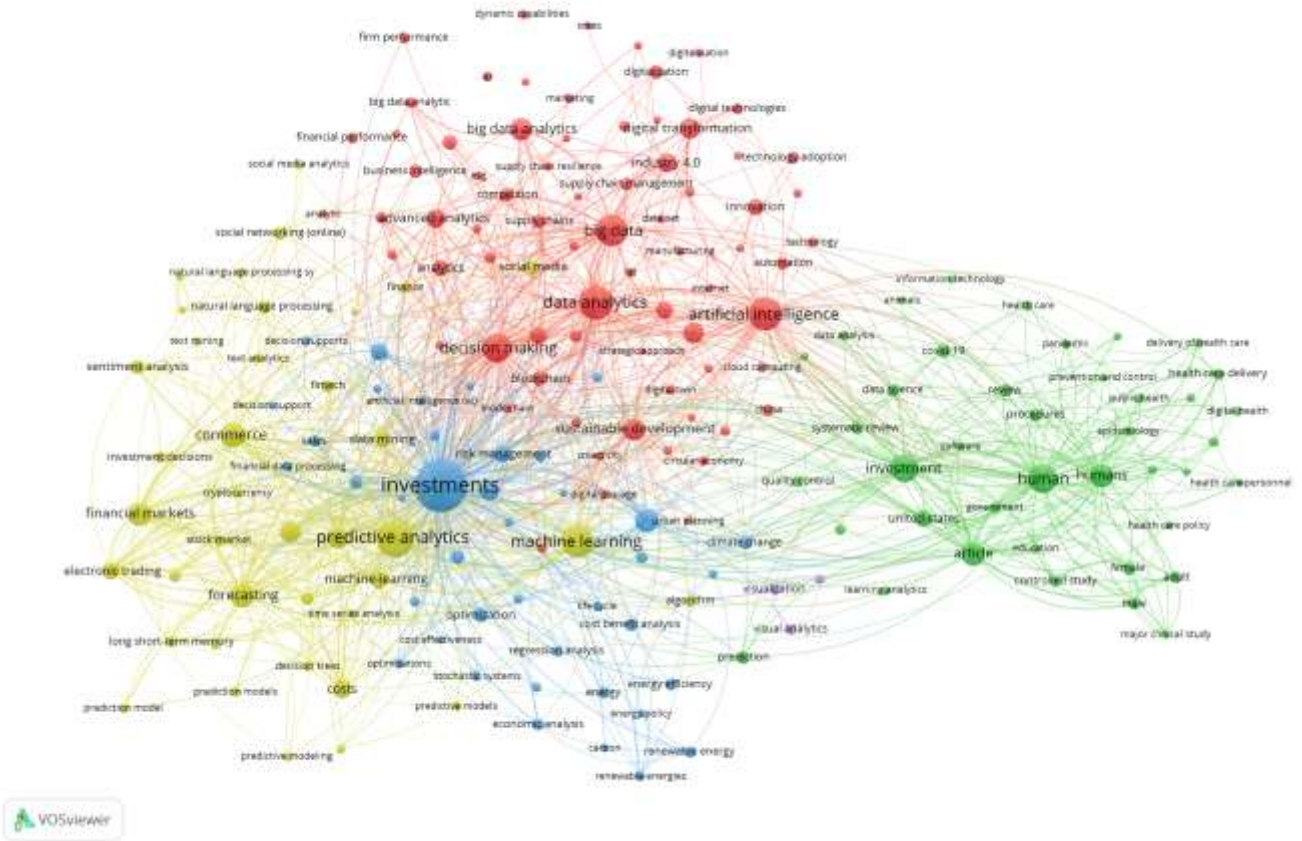


Fig -7: Keyword co-occurrence network at a minimum of 15 occurrences. Four clusters remain, the smallest a visual-analytics cluster of three circles.

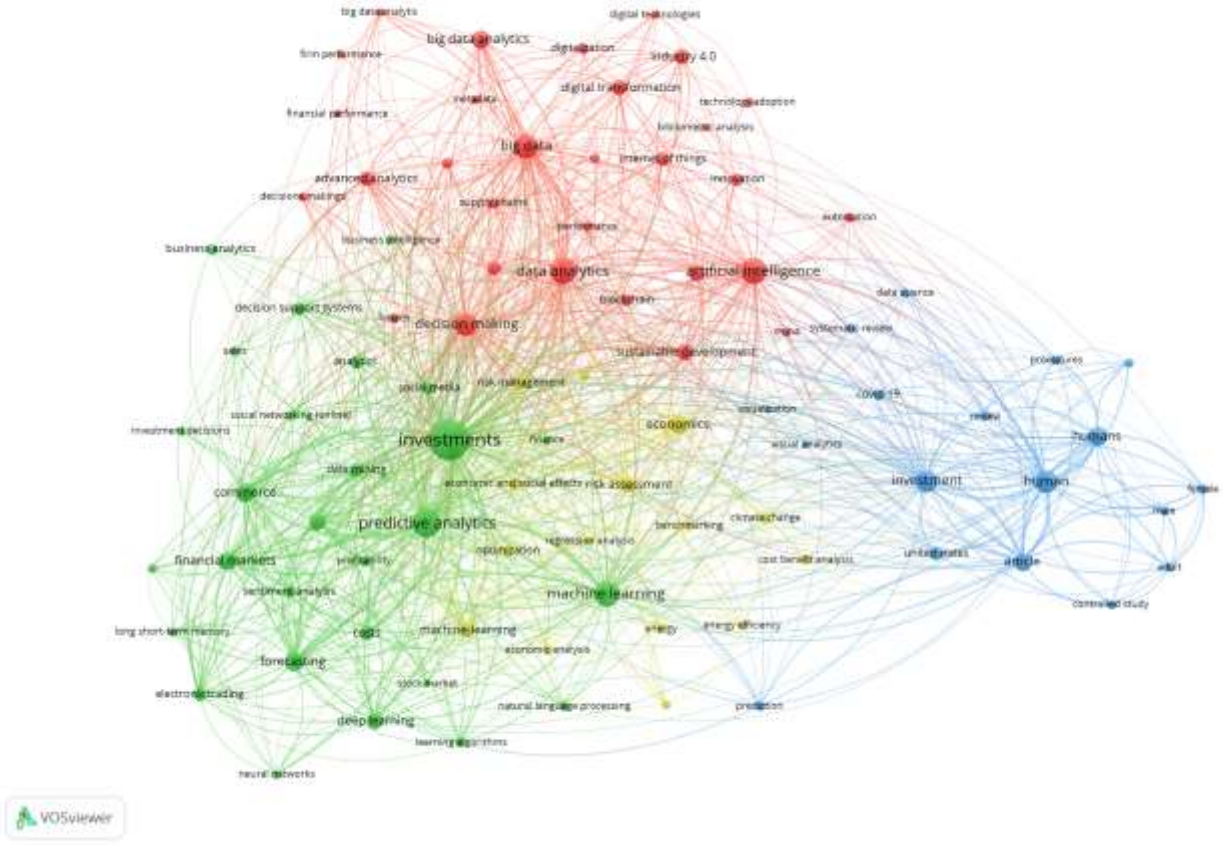


Fig -8: Keyword co-occurrence network at a minimum of 25 occurrences. Four clusters remain, 91 circles in all.

Table -5: Cluster composition across the four co-occurrence thresholds

Threshold	Circles	Clusters	Document vocabulary	Investment and prediction	Economics, risk, evaluation	Analytics and AI
5	791	7	158	106	192	199
10	339	5	73	79	merged with analytics	108
15	196	4	41	39	40 (joined by investment)	73
25	91	4	18	28	15	30

As Table 5 shows, the lowest threshold is crowded, 791 keywords across seven clusters, and the document-vocabulary terms form a cluster of their own, 158 circles strong: human, humans, article, female and review. These are indexer artefacts rather than subject terms, surviving into all four maps. Cluster count should not be read as a count of sub-domains.

The investment and prediction material are remarkably stable. 'Investments' is the most used keyword, appearing 554 times across all the maps, while predictive analytics and machine learning hold second and third place, respectively. The group only changes its cluster number at each threshold, but retains all of its members, showing persistence like no other group. It only diverges at fifteen occurrences, with 'investments' with economics and risk terms, and prediction-words forming a distinct cluster of 39. The analytics group decreases from 199 circles over to 30, its boundary with the most variation: economics and risk vocabulary at 10 occurrences, and sustainability at 15. On these maps, the two groupings could either easily pass for a single sub-domain, or remain separate, and again, whichever way they are, showing which one to combine or separate is a decision more for analysts rather than the data.

The supporting papers for the stable groups are the analytics capacity studies that established the instrument vocabulary (Mikalef et al., 2019, 2020), the econometric analysis that related that capacity to firm outcomes (Müller et al., 2018), the data-driven supply chain review (Kamble et al., 2020), the financial sentiment study of the prediction keywords (Sohangir et al., 2018), the network/systemic risk work (Galaz et al., 2021), and the energy applications review (Ahmad et al., 2022). For practitioners the low-threshold map remains more useful, such as keeping the peripheral vocabulary that is missing for the 25-occurrence map.

5. FUTURE RESEARCH AGENDA

In the literature, ‘investment’ is divided between two intellectual camps, neither of which maps onto corporate capital budgeting. The first, investment is something to forecast: stock market trends (Nabipour et al., 2020) or returns for portfolio optimisation (Ma et al., 2021). In the second, investment is an expenditure that has to be justified. Sivarajah et al. (2017) propose that big data analysis can facilitate better investment decisions, and later studies connect analytics capability with firm performance (Mikalef et al., 2019; Müller et al., 2018). This division is reflected in keyword maps. For every threshold the most frequent keyword is ‘investments’ and the group splits into clusters only at 15. There is also no mention of methods such as net present value (NPV), internal rate of return (IRR) or real options. This agenda deals with predictive analytics and corporate finance and suggests some actionable recommendations .

5.1 From Forecast Accuracy To Realised Value

A model can be predictive and still lead an investment committee to make a bad capital decision. Empirical research must therefore distinguish three constructs:

- Accuracy of forecasts – errors relative to actual operating results;
- Ex ante decision quality: If the investment was rational at the time of approval, given what was known.
- Ex post realised value: lifecycle costs, operational returns and durability of the asset in use. Yalcin et al. (2022) demonstrated that multicriteria approaches assist firms in choosing analytics platforms; however, no such evaluation has been carried out in the literature for AI-assisted project selection. A comprehensive study should compare algorithmic valuations with NPV, IRR, optimisation techniques and Monte Carlo simulation. Does algorithmic assistance help with selection under capital constraints, or reduce cost overruns? The question is best answered with longitudinal designs: archive forecasts and recommendations at approval, and follow assets through to completion. Staggered corporate adoption enables quasi-experimental difference-in-differences evaluation.

5.2 Adaptive Appraisal Under Uncertainty

In the traditional view of corporate finance, investment appraisal is a one-off gatekeeping exercise: a static appraisal that is locked in once the project is approved. We propose an adaptive framework of capital budgeting in which valuation is a continuously updated path. Machine learning systems would consume telemetry, commodity prices and demand, and update cash flow distributions in real time. Then real options, scenario modelling, robust optimisation and Monte Carlo techniques would be used to decide whether to continue, delay, expand or abandon an investment. The closest evidence concerns supply chain resilience (Spieske & Birkel, 2021), not capital allocation, so applying these methods to capital budgeting is an untested conjecture. Do adaptive models outperform static hurdle rates under rapidly changing conditions? Is it their failure when the training data is sparse or nonstationary, or when frequent updating results in expensive errors? Field case studies should be preceded by backtesting over turbulent periods with assumptions and uncertainties stated in advance.

5.3 Explainability, Judgement And Accountability

Ethics is present in 106 of the publications we reviewed. 83 of those are from 2025 and early 2026. Generative AI also experiences a similar spike, with 47 of 54 mentions over the same two years. systemic risks (Galaz et al., 2021). But more keywords in the literature does not mean capital recommendations can be audited. One governance model combining humans and AI could look this way:

- Delegated autonomy: the AI autonomously handles routine, low-stakes replacement expenditures.
- Human review: a human steps in for high-risk commitments, under-sourced, or shifting data, and inputs out of the tested range.
- Institutional auditability: record the model settings, initial assumptions, and the reasons for any managerial override. Test this approach on real-life executives. Do the explanations help them make better decisions, or just give them more confidence when they differ from the model? Ifas an investment should fail, you know whom to blame. Analysts and model builders, of course. And the executives who approved it.

Consumer data is the single most common topic. It appears in 84 papers. Capital allocation, cybersecurity, and governance round out the others. How do you deal with governance and keep generative AI analysis with a company's audited numbers.

5.4 Dual-Value Appraisal: Financial Return, Sustainability And Resilience

And then there are the sustainability clusters, the investments and the big data analysis. There are only 15 co-mentions in the analytics cluster. Recentness is a better measure. 2025 and 2026: 55 ESG publications issued.Only 40 have been published so far. Mixing appraisal into this work is something new. Machine learning can predict ESG ratings and environmental risks (Seow, 2025). Digital transformation is linked to better corporate sustainability scores (Bindeeba et al., 2025). But the gaps in appraisal make the reporting unreliable, and this affects the discount rate used in capital budgeting. Instead of mixing all the sustainability indicators into one black-box number, we suggest an appraisal framework with two dimensions. Returns and Emissions Climate Risks and Resilience.

- Long term performance – do projects with lower initial returns but lower environmental risk provide superior returns over multi-decade horizons?
- Forensic testing: Can ML flag false sustainability claims in audited carbon ledgers?

5.5 Context: Firm Size, Economies And Sectors

The authors are affiliated with organisations in the United States, China, and India in the areas of finance, logistics, energy, and agriculture (Kamble et al., 2020): Affiliations tell you where the authors work, not where the methods work . Comparative studies should therefore examine AI-assisted capital appraisal by sector, country and firm size. Practical questions include: What is the minimal data infrastructure that enables algorithmic tools to give reliable recommendations? But the capital-intensive nature of hospital projects makes them an ideal testing ground.

The corporate use of these tools is only plausible when studies compare AI-supported appraisal with established valuation methods on the same projects, report assumptions, data quality and uncertainty openly and follow realised multiyear cash flows. . proof ≠ number of citations and keyword growth.

6. CONCLUSION

This study aimed to conduct a mapping of the research field of investment and analytics from 2017 to 2026. The study's premise is the notion that analytics is used to inform investment decisions and is an area that organisations must invest in. The authors intended to explore how research on these two concepts has developed, which topics it encompasses, who leads the field and how its language has evolved. To achieve this goal, the authors performed a Scopus database search on 11 September 2026 with the terms "Investment AND Analytics" following the PRISMA approach, resulting in a final set of 2,071 documents, including 1,874 articles and 197 reviews. The data were analysed using bibliometric and VOS viewer. The present review has achieved its intended purpose as a consolidated overview of a fragmented body of work that spans a range of sectors and uses a variety of terms. The findings are summarised below according to the three research questions asked in the Introduction.

The first research question asked about the trends in publications and citations. The findings suggest that there has

been a significant rise in the research on investment and analytics. The number of annual publications rose from 59 in 2017 to 483 in 2026, an annual growth rate of 26.32% with an R2 value of 0.898231. Remarkably, 2025 and 2026 account for 938 publications, or 45.3% of the dataset, while 2026's figure is expected to continue to climb as the year is incomplete. The literature has been cited 44,409 times, with an average of 21.44 citations per document, with (Sivarajah et al., 2017) being cited 1,736 times. Thus, while the literature is well-cited, early influential works have had a significant impact on the field while more recent writings have had less time to accumulate citations.

The second research question examined the main themes, sources, authors and countries. The field is globally diverse, with 7,755 authors, 1,190 sources and 116 countries, an average of 3.97 authors per article and 31.39% international collaboration. Output is concentrated, as the top ten countries account for 63.4% of mentions, led by the US (1,024), China (591), India (542), the UK (368) and Australia (332). Keyword maps at thresholds of 5, 10, 15 and 25 show "investments" as the top keyword (554 occurrences). An investment, prediction and machine learning cluster appears at every level, while the analytics and AI group is less stable, with seven clusters at the lowest level. Finally, general terms like "human," "article," and "female" are consistently grouped together across all levels and should not be interpreted as a theme of the field.

The third research question analysed how terms and themes evolved over time and identified what gaps remain. "Big data" dominated author keywords from 2017 to 2020, "machine learning" and "data analytics" led from 2021 to 2024, and "artificial intelligence" became prominent in 2025-2026. Ethics (106 publications) and ESG (55 publications) surged in the last two years reflecting a shift toward AI-driven applications and their societal and environmental impacts. Gaps still remain including inconsistent analytics definitions, a disconnect between investment appraisal and financial machine learning research, and uneven geographical output.

This review is useful to a number of different groups. For academics it provides a coherent research map that brings together a disparate body of literature, helping both established and new researchers to see the major streams, sources and countries of contribution. The core of the research is clearly investment, prediction and machine learning, with AI, ethics and ESG all growing significantly in 2025-2026. It also has some clear methodological implications for future bibliometric work. For managers, investors and decision makers the consensus is that the value of analytics is highly context dependent. Müller et al. (2018) found productivity gains only in IT-intensive or highly competitive industries and Mikalef et al. (2020) found firm skills and resources to be important. Managers should evaluate projects in stages using multi-criteria methods (Yalcin et al., 2022) and investment managers using machine learning should factor in trading costs as Ma et al. (2021) show that high turnover can effectively eliminate much of the return. For society Bindeeba et al. (2025) combine 59 studies on digital change and ESG performance and Giordino et al. (2026) find a greater AI focus is associated with better ESG performance (which may be of benefit in areas such as energy use, health care and food). There are however some significant caveats. Galaz et al. (2021) point to the dangers of algorithmic bias and unequal access to AI and the IEA (2026) report that data centre electricity use is rising rapidly so the benefits of AI need to be weighed against these costs. For policymakers this is all evidence based guidance. Narayan et al. (2025) find that poor ESG data leads to higher financing costs for SMEs in emerging markets so standardised ESG data and reporting could improve transparency. Support for adoption should be sectoring specific and responsible AI requirements, risk assessment mechanisms and investment support policies should be developed in parallel. Given that only 31.39% of the publications involve authors from more than one country, there is a case for funding to encourage cross-country collaboration.

This review has some limitations and each of them points to a chance for future research. The study was done within the time and resources of an academic project, so the authors could not add more databases, more checks or more advanced methods. The authors kept only English articles and reviews from journals. Other records such as conference papers and book chapters were left out hence the data set may be smaller than the full field. The search words were broad, and the method does not report a check of each record for relevance. The general words found in the keyword maps show that some records may not be closely linked to the topic. The data were taken on 11 September 2026, so 2026 is not a full year. The output of 483 documents in 2026 covers only a part of the year. In the same way the new papers have had less time to gain citations that is why the citation counts favour older papers. Finally, this review is descriptive. It did not use co-citation analysis, bibliographic coupling or thematic evolution analysis, and it did not judge the quality of the studies.

REFERENCES

1. Ahmad, T., Zhu, H., Zhang, D., Tariq, R., Bassam, A., Ullah, F., AlGhamdi, A. S., & Alshamrani, S. S. (2022). Energetics systems and artificial intelligence: applications of industry 4.0. *Energy Reports*, *8*, 334–361. https://doi.org/10.1016/j.egyr.2021.11.256
2. Aria, M., & Cuccurullo, C. (2017). Bibliometrix: an r-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007
3. Baker, H. K., Kumar, S., & Pattnaik, D. (2020). Twenty-five years of review of financial economics: a bibliometric overview. *Review of Financial Economics*, *38*(1), 3–23. https://doi.org/10.1002/rfe.1095
4. Bindeeba, D. S., Tukamushaba, E. K., & Bakashaba, R. (2025). Digital levers for sustainability: a meta-analytic review of digital transformation's influence on ESG performance. *Cogent Business and Management*, *12*(1). https://doi.org/10.1080/23311975.2025.2564919
5. Chen, C. (2017). Science mapping: a systematic review of the literature. *Journal of Data and Information Science*, *2*(2), 1–40. https://doi.org/10.1515/jdis-2017-0006
6. De Luca, L. M., Herhausen, D., Troilo, G., & Rossi, A. (2021). How and when do big data investments pay off? the role of marketing affordances and service innovation. *Journal of the Academy of Marketing Science*, *49*(4), 790–810. https://doi.org/10.1007/s11747-020-00739-x
7. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: an overview and guidelines. *Journal of Business Research*, *133*(May), 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070
8. Fosso Wamba, S., & Akter, S. (2019). Understanding supply chain analytics capabilities and agility for data-rich environments. *International Journal of Operations and Production Management*, *39*, 887–912. https://doi.org/10.1108/IJOPM-01-2019-0025
9. Galaz, V., Centeno, M. A., Callahan, P. W., Causevic, A., Patterson, T., Brass, I., Baum, S., Farber, D., Fischer, J., Garcia, D., McPhearson, T., Jimenez, D., King, B., Larcey, P., & Levy, K. (2021). Artificial intelligence, systemic risks, and sustainability. *Technology in Society*, *67*. https://doi.org/10.1016/j.techsoc.2021.101741
10. Garad, A., Riyadh, H. A., Al-Ansi, A. M., & Beshr, B. A. H. (2024). Unlocking financial innovation through strategic investments in information management: a systematic review.

Discover Sustainability, *5*(1). https://doi.org/10.1007/s43621-024-00542-6

11. Giordino, D., Ballesio, E., Alshaghdali, N., & Galgotia, D. (2026). The relationship between organizational focus on AI, financial growth and sustainable development: evidence from europe. *Technological Forecasting and Social Change*, *224*. https://doi.org/10.1016/j.techfore.2025.124499
12. Islam, M. A., Fakir, S. I., Masud, S. B., Hossen, M. D., Islam, M. T., & Siddiky, M. R. (2024). Artificial intelligence in digital marketing automation: enhancing personalization, predictive analytics, and ethical integration. *Edelweiss Applied Science and Technology*, *8*(6), 6498–6516. https://doi.org/10.55214/25768484.v8i6.3404
13. Kamble, S. S., Gunasekaran, A., & Gawankar, S. A. (2020). Achieving sustainable performance in a data-driven agriculture supply chain: a review for research and applications. *International Journal of Production Economics*, *219*, 179–194. https://doi.org/10.1016/j.ijpe.2019.05.022
14. Kamis, A. (2026). A review of business analytics, machine learning, and generative artificial intelligence research 2020–2025: toward responsible artificial intelligence. *Algorithms*, *19*(6). https://doi.org/10.3390/a19060491
15. Kappelman, L., Maurer, C., Mclean, E. R., Kim, K., Johnson, V. L., Guerra, K., Torres, R., & Snyder, M. (2021). The 2020 SIM it issues and trends study. *MIS Quarterly Executive*, *20*(1), 69–107. https://www.scopus.com/pages/publications/85110798908?origin=resultslist
16. Kappelman, L., Torres, R., McLean, E., Maurer, C., Johnson, V., & Kim, K. (2018). The 2018 SIM it issues and trends study. *MIS Quarterly Executive*, *18*(1), 51–84. https://www.scopus.com/pages/publications/85063062495?origin=resultslist
17. Khalifa, M., & Albadawy, M. (2024). AI in diagnostic imaging: revolutionising accuracy and efficiency. *Computer Methods and Programs in Biomedicine Update*, *5*. https://doi.org/10.1016/j.cmpbup.2024.100146
18. Ma, Y., Han, R., & Wang, W. (2021). Portfolio optimization with return prediction using deep learning and machine learning. *Expert Systems with Applications*, *165*. https://doi.org/10.1016/j.eswa.2020.113973
19. Mariani, M., & Borghi, M. (2021). Customers' evaluation of mechanical artificial intelligence in hospitality services: a study using online reviews analytics. *International Journal of Contemporary Hospitality Management*, *33*(11), 3956–3976. https://doi.org/10.1108/IJCHM-06-2020-0622
20. Mauro, M., Noto, G., Prenestini, A., & Sarto, F. (2024). Digital transformation in healthcare: assessing the role of digital technologies for managerial support processes. *Technological Forecasting and Social Change*, *209*. https://doi.org/10.1016/j.techfore.2024.123781
21. Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: findings from a mixed-method approach. *Journal of Business Research*, *98*, 261–276. https://doi.org/10.1016/j.jbusres.2019.01.044
22. Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information and Management*, *57*(2). https://doi.org/10.1016/j.im.2019.05.004
23. Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007
24. Müller, O., Fay, M., & vom Brocke, J. (2018). The effect of big data and analytics on firm performance: an econometric analysis considering industry characteristics. *Journal of Management Information Systems*, *35*(2), 488–509. https://doi.org/10.1080/07421222.2018.1451955
25. Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. *IEEE Access*, *8*, 150199–150212. https://doi.org/10.1109/ACCESS.2020.3015966
26. Narayan, M., Kumar, N., Parida, V. K., & Kumari, P. (2025). Sustainable finance in emerging markets: adaptive governance and environmental, social, and governance innovation for equitable climate resilience. *Development and Sustainability in Economics and Finance*, *8*. https://doi.org/10.1016/j.dsef.2025.100101
27. Nobanee, H., Dilshad, M. N., Al Dhanhani, M., Al Neyadi, M., Al Qubaisi, S., & Al Shamsi, S. (2021). Big data applications the banking sector: a bibliometric analysis approach. *SAGE Open*, *11*(4). https://doi.org/10.1177/21582440211067234
28. Paul, J., & Criado, A. R. (2020). The art of writing literature review: what do we know and what do we need to know?. *International Business Review*, *29*(4). https://doi.org/10.1016/j.ibusrev.2020.101717
29. Raguseo, E., & Vitari, C. (2018). Investments in big data analytics and firm performance: an empirical investigation of direct and mediating effects. *International Journal of Production Research*, *56*(15), 5206–5221. https://doi.org/10.1080/00207543.2018.1427900
30. Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions. *Electronics (Switzerland)*, *10*(21). https://doi.org/10.3390/electronics10212717
31. Seow, R. Y. C. (2025). Transforming ESG analytics with machine learning: a systematic literature review using TCCM framework. *Corporate Social Responsibility and Environmental Management*, *32*(6), 7358–7389. https://doi.org/10.1002/csr.70089
32. Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, *70*, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001
33. Sohangir, S., Wang, D., Pomeranets, A., & Khoshgoftaar, T. M. (2018). Big data: deep learning for financial sentiment analysis. *Journal of Big Data*, *5*(1). https://doi.org/10.1186/s40537-017-0111-6
34. Spieske, A., & Birkel, H. (2021). Improving supply chain resilience through industry 4.0: a systematic literature review under the impressions of the COVID-19 pandemic. *Computers and Industrial Engineering*, *158*. https://doi.org/10.1016/j.cie.2021.107452
35. Trieu, V. H. (2017). Getting value from business intelligence systems: a review and research agenda. *Decision Support Systems*, *93*, 111–124. https://doi.org/10.1016/j.dss.2016.09.019
36. Vo, N. N. Y., He, X., Liu, S., & Xu, G. (2019). Deep learning for decision making and the optimization of socially responsible investments and portfolio. *Decision Support Systems*, *124*. https://doi.org/10.1016/j.dss.2019.113097
37. Yalcin, A. S., Kilic, H. S., & Delen, D. (2022). The use of multi-criteria decision-making methods in business analytics: a comprehensive literature review. *Technological Forecasting and Social Change*, *174*. https://doi.org/10.1016/j.techfore.2021.121193
38. Yue, X., Bai, J., Liu, Q., Tang, Y., Puri, A., Li, K., & Qu, H. (2020). SPortfolio: stratified visual analysis of stock portfolios. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 601–610. https://doi.org/10.1109/TVCG.2019.2934660
39. *Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. Organizational Research Methods, 18(3), 429–472. https://doi.org/10.1177/1094428114562629*