Journal of Library and Information Studies 23:1 (June 2025) pp.1-77
https://doi.org/10.6182/jlis.202506_23(1).001

Mapping the Generations of Research Impact Science: A Scoping Review of Metrics, Frameworks, and Predictive Approaches

Mudassar Arsalan[1], Omar Mubin[2], Abdullah Al Mahmud[3]

Abstract

This scoping review systematically examines the evolution of research impact assessment frameworks, categorising their development into 4 generations. The study follows Arksey and O'Malley's (2005) scoping review methodology, enhanced by Levac et al. (2010) and aligned with PRISMA-ScR guidelines. A comprehensive search was conducted across Scopus, Web of Science, PubMed, IEEE Xplore, and Google Scholar, as well as grey literature, to identify relevant studies on research impact frameworks, metrics, and methodologies. A total of 139 studies were selected based on predefined inclusion criteria, encompassing bibliometric indices, multidimensional frameworks, predictive analytics, and alternative metrics.

The findings highlight the transition from traditional citation-based measures to sophisticated, data-driven methodologies. The first generation (bibliometric indices) included 26 key metrics focusing on publication productivity. The second generation introduced 41 multidimensional frameworks incorporating societal, economic, and policy indicators. The third-generation integrated machine learning and predictive analytics to assess impact across 17 aspects and 9 data-driven factors. The fourth generation utilised 16 alternative metrics, including Altmetric Attention Scores and PlumX Metrics, to capture real-time digital engagement.

Key challenges identified include limited standardisation, regional biases, and the under-exploitation of emerging technologies, such as large language models. The study underscores the need for predictive multidimensional frameworks and standardised taxonomies to enhance scalability and foresight in impact assessment. By structuring the evolution of research impact science, this review provides actionable insights to refine assessment methodologies, ensuring their relevance for addressing societal challenges and guiding strategic research investments.

Keywords: Research Impact Assessment; Generational Frameworks; Bibliometric Metrics; Multidimensional Models; Predictive Analytics

1. Introduction

Research plays a critical role in addressing societal challenges, driving innovation, and informing policies across diverse academic, social, economic, and cultural contexts (Ward et al., 2023). Its impacts are as diverse as the methodologies and contexts in which it operates, emphasising the need for robust frameworks to capture and measure its influence comprehensively. Historically, research evaluation prioritised academic metrics, such as citations and journal impact factors (Kostoff, 1994b; Moed et al., 1985). However, growing demands from policymakers, funding bodies,

[1,2] School of Computer, Data and Mathematical Sciences, Western Sydney University, Penrith, Australia
[3] Centre for Design Innovation, Swinburne University of Technology, Melbourne, Australia
* Corresponding Author: Mudassar Arsalan, E-mail: mharsalan@gmail.com

and the public have shifted the focus toward a broader assessment of research's societal and economic contributions (Martinuzzi et al., 2023; Sørensen et al., 2022).

This paradigm shift has led to the emergence of "research impact science," a multidisciplinary and empirical field that systematically examines the diverse impacts of research across academic, societal, economic, and policy domains. Anchored in the theoretical and methodological traditions of scientometrics, evaluation studies, and data science, this field employs rigorous, standardised, and replicable approaches. Retrospective methods evaluate historical contributions, while prospective approaches predict future impacts, providing policymakers and institutions with actionable insights. The evolution from traditional bibliometric measures focused on academic productivity to multidimensional frameworks, such as the UK's Research Excellence Framework (REF), demonstrates the field's increasing complexity and scientific rigour. These frameworks integrate qualitative and quantitative metrics, capturing aspects such as public engagement, policy influence, economic contributions, and societal change (Giménez Toledo, 2018; Research Excellence Framework [REF], 2014). By uniting evidence-based methodologies with predictive capabilities, research impact science builds a comprehensive and testable foundation for understanding and maximising the value of research in addressing global challenges.

The growing complexity of global challenges, including climate change, health crises, and economic inequality, has further underscored the importance of aligning research outputs with societal priorities. This alignment requires methodologies that assess historical contributions and employ predictive models to anticipate future impacts. Advanced tools like machine learning (ML) and real-time analytics are pivotal in evaluating research trajectories and ensuring strategic resource allocation (Thelwall, 2021; Yazdizadeh et al., 2024).

Despite significant advancements, research impact science still needs to be more cohesive, with diverse methodologies, indicators, and frameworks needing a unified structure. Early bibliometric indices, while foundational, are limited in scope, focusing primarily on academic productivity. Later generations of frameworks address broader aspects but often struggle with standardisation, complexity, and scalability issues. The emergence of alternative metrics and data science-driven models introduces innovative approaches but raises challenges related to reliability, comparability, and implementation (Boshoff & de Jong, 2020).

This study addresses these gaps by systematically consolidating the evolution of research impact aspects, indicators, models, and frameworks. Using a generational lens, it categorises the development of impact assessment methodologies, identifying patterns and trends while proposing a structured foundation for future research. The novelty of this work lies in its integrative approach, which combines historical analysis with advanced predictive modelling to establish research impact science as a cohesive and forward-looking discipline. By providing a comprehensive review and addressing existing challenges, this study aims to enhance the understanding and practice of research impact evaluation, aligning academic rigour with societal relevance.

The remainder of the study is organised as follows: Section 2 reviews the literature, tracing the origins and development of research impact science, emerging trends, and global policy contexts. Section 3 details the scoping review methodology employed in this study. Section 4 presents the findings. Section 5 discusses aligning these findings with the study's objectives, explores implications, and charts future directions for the field. Section 6 concludes with a summary of key findings and recommendations for advancing research impact assessment practices.

2. Literature Review

2.1 Origins and development of research impact science

The evolution of research impact science is deeply rooted in the broader field of scientometrics, which systematically analyses scholarly communication and the intellectual progression of disciplines. The earliest formal attempts to measure academic output quantitatively can be traced back to E.W. Hulme's pioneering work, *Statistical Bibliography about the Growth of Modern Civilization*, which introduced foundational methods for assessing the dissemination and societal influence of academic research (Hulme, 1923). Hulme's contributions set the stage for later advancements, including Derek de Solla Price's seminal work, *Little Science, Big Science* (1963), which highlighted the exponential growth of scientific research and emphasised the need for methodologies capable of systematically tracking its evolution and impact (Price, 1963).

A significant leap in this field came with Eugene Garfield's introduction of the Science Citation Index (SCI) in the mid-20th century. Initially conceived as a tool to navigate scientific literature, the SCI quickly became a cornerstone for evaluating the influence of research through citation networks (Garfield, 1955). Garfield's subsequent innovation, the Impact Factor, introduced in 1963, provided a standardised metric for assessing journal-level influence, marking a transformative moment in research evaluation (Garfield & Sher, 1963). These developments established the quantitative foundations of scientometrics and laid the groundwork for future advancements in assessing research productivity and impact (Garfield, 1979).

Scientometrics was formally defined by Nalimov and Mul'chenko (1971) as "quantitative methods of the research on the development of science as an informational process." (p. 2) This definition underscored the growing recognition of science as a system of interconnected outputs, including publications and citations, that could be systematically measured and analysed. At the same time, societal and policy-driven imperatives catalysed the field's evolution. Increasing demands for accountability in publicly funded research fuelled the integration of scientometric tools into policymaking, as exemplified by the National Science Board's Science Indicators Reports in 1972, which used citation data to align research outputs with societal and policy objectives (Wyatt et al., 2017).

The institutionalisation of scientometrics further accelerated its application and relevance. The establishment of the journal *Scientometrics* in 1978 by Tibor Braun provided a dedicated platform for exploring the quantitative aspects of scientific productivity and its societal implications (Hérubel, 1999). This was complemented by the creation of the International Society for Scientometrics and Informetrics in 1994, which fostered global collaboration and solidified the discipline's status. Methodological innovations, such as co-citation analysis, introduced by Small in 1973 and co-word mapping, developed by Callon in 1983, significantly enhanced the ability to visualise and analyse complex scholarly networks (Callon et al., 1983; Small, 1973).

As scientometrics matured, its scope expanded to encompass broader interdisciplinary and socio-political aspects. For instance, Linda Smith's "Citation Analysis" (1981) applied citation metrics to evaluate scholarly influence across disciplines, reflecting the growing complexity of academic ecosystems and the need for frameworks capable of capturing this diversity. This expansion laid the foundation for research impact science, which became a specialised focus within scientometrics. At the same time, retaining the quantitative tools of its parent field, research impact science emphasises assessing research's societal, economic, and policy outcomes.

Structured frameworks played a pivotal role in formalising research impact science. The Payback Framework, introduced in the 1990s, systematically linked research outputs with tangible societal outcomes such as health improvements and economic benefits, establishing a structured approach to impact evaluation (Penfield et al., 2014). Similarly, the productive interaction framework of Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society (SIAMPI) emphasised the importance of stakeholder collaboration in translating research into societal benefits (Spaapen et al., 2011). These frameworks expanded the scope of research evaluation and highlighted the need to align academic research with public priorities.

Integrating impact assessment into national and institutional strategies further solidified its role within the field of scientometrics. Initiatives such as the UK's REF and Australia's Research Quality Framework (RQF) have incorporated research impact into funding and policy decisions, reflecting a growing emphasis on accountability and societal relevance (Penfield et al., 2014). The emergence of altmetrics and other real-time digital tools further broadened the scope of impact evaluation by capturing online engagement and immediate societal interactions with research outputs.

Together, these developments highlight the dual role of scientometrics as both a discipline and a foundation for the evolution of research impact science. While scientometrics focuses on the quantitative analysis of scholarly productivity, research impact science extends this focus to assess the broader societal implications of research. This interplay drives innovation in methodologies and frameworks, bridging the academic, societal, and policy domains.

2.2 Emerging trends, challenges, and global policy context in research impact science

Research impact science continues to evolve, shaped by technological advancements, policy demands, and societal expectations. Emerging trends reveal a move towards integrative frameworks, real-time data analysis, and innovative indicators for assessing research outcomes. These developments align with the global shift towards accountability and evidence-based policymaking, underscoring the importance of demonstrating research benefits in diverse domains.

One of the prominent trends in research impact assessment is the adoption of interdisciplinary approaches and the integration of novel data streams. Frameworks like SIAMPI's productive interactions emphasise the importance of stakeholder engagement, recognising that impactful research often emerges from collaborative networks rather than isolated efforts (Spaapen et al., 2011). Additionally, altmetrics, which capture digital engagement, such as social media mentions and online discussions, have gained traction as complementary tools to traditional citation metrics. These metrics offer insights into the societal impact of research, providing real-time feedback loops for researchers and policymakers (Penfield et al., 2014).

Adopting mixed-method approaches, which combine quantitative and qualitative indicators, has become essential for capturing the multifaceted nature of research impact. Empirical data enriches this approach, and narrative case studies allow for a comprehensive evaluation of research contributions, ranging from academic advancements to societal transformations. The UK's REF exemplifies this trend, employing case studies to assess the significance and reach of research outcomes while incorporating metrics for accountability (REF, 2014).

Despite these advancements, significant challenges persist. Attribution remains one of the most complex issues in impact assessment, as research outcomes often result from cumulative and collaborative efforts over extended periods. The intricate networks of knowledge exchange and the non-linear nature of impact pathways complicate efforts to isolate specific contributions from individual projects or institutions (Hughes & Kitson, 2012). Additionally, the time lag between research activities and measurable impacts poses further difficulties, particularly for disciplines where outcomes may take decades to materialise.

The dynamic nature of impact introduces variability, as the significance of research outcomes can change over time and across contexts. Temporary impacts may dissipate, while initially overlooked findings can later gain relevance in unforeseen ways. This temporal variability necessitates adaptable evaluation frameworks that accommodate the evolving nature of research contributions (Börner et al., 2011).

Globally, the push for research impact assessment reflects broader policy trends emphasising transparency, accountability, and societal relevance. National and international initiatives, such as Australia's RQF and the European Research Council's efforts to integrate impact evaluations, highlight the growing importance of linking research funding to demonstrable societal benefits (Penfield et al., 2014). These frameworks often serve as policy tools to allocate resources strategically, ensuring that public investments in research yield tangible benefits for communities.

The intersection of research impact science with global challenges, such as climate change, health crises, and economic inequality, underscores the necessity of aligning academic endeavours with societal priorities. Policymakers increasingly demand evidence of research contributions to addressing these challenges, driving the development of targeted evaluation tools and collaborative platforms (Donovan & Hanney, 2011).

The need for a systematic, global approach to organising aspects, indicators, frameworks, and models of research impact has never been more apparent. Efforts to standardise methodologies while retaining flexibility for disciplinary and contextual variations can enhance the reliability and comparability of impact assessments. Emerging technologies, such as data mining and artificial intelligence (AI), hold promise for automating the collection and analysis of impact evidence, reducing administrative burdens on researchers and institutions (Moed & Halevi, 2015). While research impact science has made significant strides, addressing its challenges requires a concerted effort to develop innovative tools, foster interdisciplinary collaboration, and align academic research with societal needs.

The review of the historical development and key advancements in research impact science highlights the remarkable diversity of aspects, indicators, models, and frameworks that have emerged over time. While each generation of research impact methodologies has contributed significantly to the field, they also reveal notable limitations. Early bibliometric indices, although foundational, still need to be broadened to encompass a broader focus on academic productivity. Comprehensive frameworks and multidimensional models developed in later generations have advanced accountability and inclusivity but often require greater complexity, increased resource intensity, and limited comparability across disciplines and contexts. The emergence of alternative metrics and data science-driven approaches offers new possibilities but introduces challenges related to data reliability, standardisation, and implementation scalability.

A critical gap remains in the systematic organisation of these diverse methodologies and their integration into a coherent framework. The absence of standardisation and comparability across aspects, indicators, and models leads to fragmented practices that hinder their broader applicability and scalability. Furthermore, the evolving demands of modern research ecosystems—including interdisciplinarity, sustainability, and digital transformation—underscore the need for innovative and adaptable approaches. Despite decades of development, research impact science lacks a unified structure to guide stakeholders in effectively selecting, applying, and refining tools for evaluating the influence of research.

These challenges highlight the necessity of a systematic scoping review to consolidate and organise the research impact aspects, indicators, models, and frameworks with the evolutionary approach. By identifying patterns and trends across generations, such an effort can provide a structured foundation for research impact assessment. This organisation is essential for addressing current limitations and charting a clear path forward, ensuring that research impact evaluation aligns with academic rigour and societal relevance. Consequently, this review aims to fulfil this critical need, offering a comprehensive and forward-looking perspective to support the evolving demands of the research community.

3. Methodology

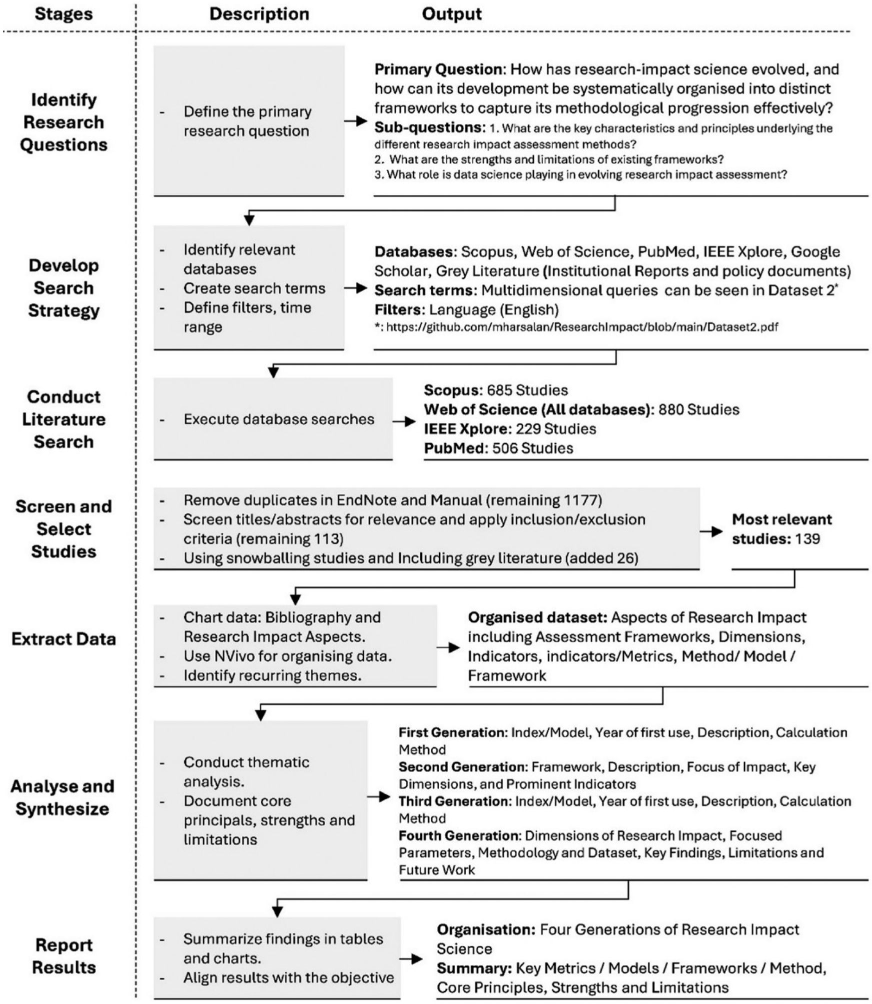
This study employs a scoping review methodology to systematically explore the evolution of research impact science, focusing on the trio framework of factors, indicators, and methods. The framework informs the approach of Arksey and O'Malley (2005), enhanced by Levac et al. (2010), and aligned with the PRISMA-ScR checklist for scoping reviews (Tricco et al., 2018). Scoping reviews are particularly suited for examining broad and complex research areas where diverse methodologies, disciplines, and types of evidence converge (Peters et al., 2021). This approach was chosen to comprehensively map the existing literature, identify gaps, and synthesise emerging trends in research impact science while capturing the role of data science and ML. The review was conducted in distinct stages to ensure methodological rigour and transparency (see Figure 1).

The first stage focused on defining the research questions that would guide the study. The primary research question was framed as: How has research-impact science evolved, and how can its development be systematically organised into distinct frameworks to capture its methodological progression effectively? This overarching question was supported by 3 sub-questions: What are the key characteristics and principles underlying different research impact assessment methods? What are the strengths and limitations of existing frameworks? Moreover, what role is data science playing in the evolution of research impact assessment? These questions established a clear focus and structured the subsequent stages of the review.

A comprehensive search strategy was developed in the second stage to identify relevant literature. Key databases, including Scopus, Web of Science, PubMed, IEEE Xplore, and Google Scholar, as well as grey literature, such as institutional reports and policy documents, were selected to ensure diverse coverage. Multidimensional search queries, aligned with the research objectives, were employed, with filters applied for language (English). Detailed search terms were documented in Dataset 1 (https://github.com/mharsalan/ResearchImpact/blob/main/Dataset1.pdf) to maintain transparency and reproducibility. The third stage involved conducting the literature search across these databases, yielding a substantial pool of studies: 685 from Scopus, 880 from Web of Science, 229 from IEEE Xplore, and 506 from PubMed.

The inclusion criteria for this scoping review were designed to ensure a comprehensive and relevant selection of studies addressing the evolution of research impact science. Eligible studies focused on frameworks, methodologies, metrics, or tools for assessing research impact, as well as literature exploring the development and progression of research impact science across various disciplines. The review encompassed multiple sources, including peer-reviewed journal articles, conference proceedings, institutional reports, grey literature (such as policy documents and white papers), and books. Studies from key domains such as health, environmental sciences, social sciences, higher education, and technology were included to ensure a multidisciplinary perspective, provided they applied or discussed research impact assessment frameworks. Only studies offering sufficient details on frameworks, metrics, methodologies, or indicators used to evaluate research impact were considered,

Figure 1. Methodological Framework for the Evolution of Research Impact Science



Note. PRISMA-ScR Scoping Review Methodology after Tricco et al. (2018).

ensuring their substantive contribution to the analysis. Both theoretical and empirical contributions were prioritised, including articles that present conceptual models or empirical data analysing the strengths, limitations, or evolution of research impact assessment. This approach enabled a holistic and nuanced understanding of the methodologies and frameworks shaping the field.

Following the literature search, the fourth stage focused on screening and selecting studies. Duplicate records were removed using EndNote and manual review, and titles and abstracts were screened for relevance based on predefined inclusion and exclusion criteria. Grey literature and institutional reports were included to capture non-academic perspectives, and a snowballing approach was employed to identify additional relevant studies. This rigorous process resulted in the selection of 139 studies that were most applicable to the research questions. In the fifth stage, data extraction focused on aspects of research impact. NVivo software was employed to organise the data, identifying recurring themes and ensuring systematic categorisation. The organised dataset included assessment frameworks, aspects, indicators/metrics, and methodologies.

The sixth stage involved analysing and synthesising the extracted data to document the thematic progression of research impact science across 4 distinct generations. Key features for each generation were categorised, including the models and indices of the first generation, the multidimensional frameworks of the second, the alternative metrics and digital tools of the third, and the advanced data-driven methodologies of the fourth. The analysis captured the core principles, strengths, and limitations of each generational framework. Finally, in the seventh stage, the findings were summarised in tables and charts, aligning the results with the study's objectives. This stage systematically organised the evolution of research impact science into four distinct generations, highlighting key metrics, models, methodologies, core principles, strengths, and limitations. This structured methodological approach ensured that the review offered a robust and comprehensive analysis of the progression of the research impact assessment.

4. Results

The evolution of research impact aspects, indicators, assessment methods, and frameworks reflects the growing recognition of the need to measure and demonstrate the value of scientific research systematically. Over time, assessment practices have evolved from simple quantitative metrics focused on scholarly productivity to more sophisticated frameworks and predictive models that capture a broader range of impacts. This evolution can be categorised into 4 distinct generations.

4.1 First generation: Bibliometric indices and productivity metrics

4.1.1 Historical foundations of bibliometric approaches

The first generation of research impact assessment evolved from its early focus on bibliometric indices, such as raw citation counts and publication numbers, into a more comprehensive suite of metrics addressing individual and journal-level evaluation needs (see Appendix, Table 1). These methods provided the foundational tools for quantifying academic productivity and impact, enabling researchers, institutions, and funding agencies to evaluate the systematic outputs of scientific endeavours. Over time, the first generation expanded its scope by developing innovative metrics that sought to address the limitations of earlier approaches while maintaining the principles of objectivity, simplicity, and reproducibility.

4.1.2 Key indicators and indices

The first generation began with simple publication counts and citation analyses, which measured the frequency and patterns of citations as proxies for research influence (Hulme, 1923). Eugene Garfield's introduction of the Impact Factor (IF) in the 1960s marked a significant advancement, providing a journal-level metric based on the average number of citations received per article within a specific time frame. The IF quickly became a standard tool for evaluating the prestige and influence of journals (Garfield, 1979). However, as research evaluation demands grew, new metrics were introduced to enhance individual-level assessments. The h-index, proposed by Hirsch in 2005, combined publication output and citation impact into a single measure, quickly gaining popularity for its simplicity and applicability across disciplines (Hirsch, 2005).

To address the limitations of the h-index, derivatives such as the g-index and A-index emerged. The g-index placed greater weight on highly cited articles, reflecting the cumulative impact of the most significant works of an author (Egghe, 2006). Similarly, the A-index calculated the average citations per article within the h-core, offering a refined view of citation impact (Jin et al., 2007). More recently, the GFsa index introduced a novel approach by incorporating a researcher's "scientific age" into its calculations. It provides fairer assessments by normalising citations based on the length of a researcher's career (Fernandes & Fernandes, 2024). These advancements illustrate the adaptability of first-generation metrics to evolving evaluation needs.

At the journal level, indices such as the SCImago Journal Rank (SJR) and Source Normalized Impact per Paper (SNIP) offered field-weighted measures that accounted for citation behaviours specific to different disciplines. For instance, SJR weighted citations based on the prestige of the citing journal, while SNIP normalised citation counts to account for variations in citation practices across fields (Falagas et al., 2008; Oosthuizen & Fenton, 2014). These innovations enabled more accurate cross-disciplinary comparisons, addressing a significant limitation of earlier bibliometric methods.

4.1.3 Core principles

The first generation adhered to 3 core principles – objectivity, simplicity, and replicability. Metrics like the IF, h-index, and g-index were designed to provide quantifiable and standardised measures that could be easily calculated and compared across contexts. Their widespread adoption was facilitated by their integration into major bibliometric databases, including Scopus, Web of Science, and Google Scholar, which ensured accessibility and consistency in their application (Elsevier, 2024). Field-normalized metrics, such as SNIP and Field-Weighted Citation Impact (FWCI), further aligned these principles by addressing disciplinary differences, enabling fairer evaluations across research domains (Purkayastha et al., 2019).

4.1.4 Strengths and limitations of first-generation indices

The enduring relevance of first-generation metrics is attributed mainly to their strengths. These metrics provide transparent and reproducible measures of research productivity and influence. They are simple to calculate, widely available, and applicable across diverse academic contexts. However, they also face significant limitations. One major criticism is their inherent disciplinary bias, as metrics like the IF and h-index tend to favour fields with higher publication and citation rates, thereby disadvantaging disciplines such as the humanities and social sciences (Hirsch, 2005). Similarly, journal-level metrics, such as the IF, are susceptible to skewed citation distributions, where a few highly cited articles can disproportionately elevate a journal's overall score (University of Bradford, 2020).

Another key area for improvement is the inability of these metrics to capture qualitative dimensions of research impact, such as societal or policy contributions. The focus on citations as the primary indicator of influence overlooks the broader effects of research outside the academic community (Bornmann & Marx, 2013). Moreover, first-generation metrics are vulnerable to manipulation, with practices like excessive self-citation and strategic publication timing artificially inflating scores.

Despite their limitations, first-generation metrics remain central to research evaluation due to their adaptability and continuous refinement. The development of indices, such as the GFsa and field-weighted measures, reflects ongoing efforts to address biases and enhance the robustness of these methods (Fernandes & Fernandes, 2024). Additionally, their integration into widely used databases and evaluation frameworks ensures accessibility and relevance. As new tools and methodologies emerge, first-generation metrics provide a foundational framework upon which more sophisticated approaches can be built.

4.2 Second generation: Multidimensional frameworks for research impact assessment

4.2.1 Rationale for expanding beyond citations

The development of comprehensive frameworks characterised the second generation of research impact assessment, which evaluates the diverse and multidimensional benefits of research. These frameworks emerged in response to the growing demand for accountability in publicly funded research and the need to demonstrate tangible outcomes, such as improved health systems, economic growth, and evidence-based policymaking, reflecting societal priorities. Unlike first-generation methods that relied primarily on publication and citation metrics, these frameworks aimed to capture the broader societal implications of research, providing a holistic understanding of its value. Table 2 in the Appendix summarises the key frameworks developed during this generation, highlighting their focus, aspects, and indicators.

The need for these frameworks arose from recognising the limitations of first-generation metrics in capturing the full spectrum of research impacts. Policymakers, funding agencies, and other stakeholders require tools to evaluate research's translational and implementation aspects, ensuring that scientific advancements translate into meaningful societal benefits (Bernstein et al., 2007; Donovan & Hanney, 2011). Frameworks like the Payback Model and the Canadian Academy of Health Sciences (CAHS) Framework addressed these demands by systematically linking research activities to outcomes across multiple domains (Canadian Academy of Health Sciences [CAHS], 2009; Donovan & Hanney, 2011). These methodologies enhanced research accountability and provided insights into its alignment with national priorities and strategic goals.

4.2.2 Key frameworks across domains

Health and Biomedical Research Impact Frameworks. Health and biomedical research have been a leading domain in the development and application of second-generation research impact assessment frameworks. The critical nature of health-related challenges and their profound societal and economic implications made this domain an early focus for structured frameworks. These tools were developed to evaluate the tangible benefits of health research, ensuring accountability to funders and alignment with public and policy priorities. Frameworks for health and biomedical research provided the foundation for evaluating the impacts of research across other disciplines, underscoring the central role of this domain in the second generation of research assessment.

The importance of health and biomedical research in driving the development of this generation lies in its unique position to influence human well-being, healthcare systems, and economic productivity. The urgency to translate research findings into improved healthcare practices and policy changes created a strong demand for comprehensive frameworks that could capture these multifaceted outcomes. The domain also receives substantial public funding globally, necessitating transparent and systematic methods to demonstrate societal returns on investment. Furthermore, the tangible nature of health outcomes—such as reduced disease burden, improved quality of life, and cost savings in healthcare systems—made health and biomedical research well-suited for structured impact assessment. These factors collectively positioned health research as a pioneer in developing second-generation frameworks.

Several frameworks emerged to assess the impacts of health and biomedical research, each addressing the complexities and unique requirements of the domain. The Payback Framework, introduced in 1996, was one of the most influential models. It categorised research outcomes into 5 aspects: knowledge production, research targeting, policy impact, health benefits, and economic benefits. Initially designed for health services research, it provided a structured approach to linking research activities to societal outcomes and became widely applicable across various disciplines. The Payback Framework remains foundational, influencing the design of subsequent frameworks (Donovan & Hanney, 2011).

The Canadian Institutes of Health Research (CIHR) Impact Framework is built on the Payback model, incorporating capacity building, partnerships, and knowledge dissemination. It emphasised collaborative approaches and practical applications, aligning health research with Canada's national healthcare priorities. This framework highlighted the importance of partnerships in translating research into actionable outcomes and policies (Bernstein et al., 2007). Similarly, the CAHS Framework broadened the focus to evaluate societal returns on investment, emphasising improved health outcomes, enhanced health systems, and economic benefits. The CAHS Framework addressed national priorities and demonstrated the societal relevance of health research, offering a comprehensive model for assessing the impacts (CAHS, 2009).

The Becker Model, designed for translational research, was another significant contribution in this domain. It captured the progression from basic science to practical applications, illustrating the pathways from knowledge generation to tangible societal benefits. By incorporating indicators such as collaborations, patents, publications, and health outcomes, it provided a systematic approach to understanding the translational impact of research (Sarli et al., 2010). Similarly, the Wellcome Trust's Assessment Framework emphasised the real-world applications of research findings, particularly in addressing global health challenges in low-resource settings. This framework underscored the importance of adapting research outputs to local contexts, emphasising translational and implementation science (Wellcome Trust, 2009).

Health and biomedical research frameworks led the second generation due to the immediate societal relevance of health challenges, the substantial public investment in the domain, and the interdisciplinary nature of health research. The direct implications of health research for population well-being made it an ideal focus for impact assessment. Moreover, the strong links between health research and policy development underscored the need for frameworks that capture policy influence and legislative changes. The dominance of health research frameworks also reflects the clarity and measurability of outcomes in this field, such as reduced morbidity and mortality rates, which align well with structured evaluation models.

The continued prominence of health and biomedical research frameworks in the second generation can be attributed to the evolving complexity and scale of health challenges. Issues such as aging populations, emerging diseases, and healthcare inequalities have increased the need for robust evaluation frameworks to ensure that research investments translate into meaningful societal benefits. The emphasis on translational science and implementation research has further driven the development of models, such as the Becker Model, which maps the journey from bench to bedside (Sarli et al., 2010). Additionally, frameworks in this domain remain central due to their adaptability and ability to address national and global health priorities.

Higher Education and Academic Research Impact Frameworks. Higher education and academic research have been another pivotal domain in the evolution of second-generation research impact assessment frameworks. These frameworks were developed to address the dual roles of higher education institutions (HEIs): advancing scientific knowledge and addressing societal challenges. They aim to measure research quality and broader societal contributions, aligning institutional priorities with national and global goals. HEIs are particularly significant in research evaluation due to their critical role in knowledge production and accountability for public investments. While early bibliometric analyses provided valuable insights into research productivity and informed initial efforts to evaluate research outputs in HEIs, they were often limited in scope, focusing narrowly on quantitative indicators such as citations and publication counts (Moed et al., 1985). These methods underscore the need for more comprehensive and multidimensional frameworks to capture the societal and economic impacts of research and ensure alignment with public priorities and policy objectives.

The Research Assessment Exercise (RAE), introduced in the UK in 1992, was one of the earliest frameworks in this domain. It utilised peer-review panels to assess the quality of research outputs across disciplines and allocate public funding accordingly. The RAE had a significant influence on funding decisions and shaped institutional strategies for achieving research excellence. It laid the foundation for the REF, launched in 2014, which added societal and economic impact as core evaluation aspects. The REF employed bibliometrics, case studies, and expert reviews to assess the quality of outputs, societal benefits, and research environment, making it one of the most comprehensive tools for research impact evaluation (Boaden & Cilliers, 2001; REF, 2014).

In Australia, the Excellence in Research for Australia (ERA) framework, established in 2010, assesses the quality of research outputs from HEIs against international benchmarks. It incorporates a mix of quantitative and qualitative metrics, including traditional academic outputs, such as publications, and non-traditional outputs, like creative works and patents. By comparing research performance across disciplines, the ERA ensures accountability and encourages continuous improvement in research quality (Australian Research Council [ARC], 2018). Complementing the ERA, the Engagement and Impact Framework (EIF) was used to evaluate how Australian universities engage with industry and address societal challenges. The EIF emphasises collaborative efforts, public policy contributions, and socio-economic benefits, highlighting the broader impacts of academic research (ARC, 2017).

The World University Rankings, first introduced by Times Higher Education (THE) in 2004, offer an additional framework for evaluating HEIs. These rankings use diverse performance indicators, including teaching quality, research output, international collaboration, and industry income. By emphasising global engagement and knowledge transfer, this framework aligns research outputs with the strategic priorities of universities and nations. It has become an essential tool for assessing institutional performance and competitiveness on a global scale (Times Higher Education [THE], 2018). Similarly, the UNESCO Framework for Knowledge Translation aligns research impacts with the United Nations Sustainable Development Goals (SDGs). This framework systematically introduces the SDGs into university programs, addressing institutional, thematic, structural, and personal aspects. It evaluates universities based on their contributions to the SDGs. First introduced in 2019, this framework is used to assess the impact of universities on sustainable development (Leal Filho et al., 2021). Additionally, this framework is utilised to determine the university's impact ranking by Times Higher Education (THE, 2018).

Multidisciplinary Research Impact Frameworks. Multidisciplinary research frameworks have gained prominence due to the increasing need to address complex global challenges, including climate change, public health crises, and sustainable development. These issues require collaborative approaches that integrate knowledge and methods across disciplines. Multidisciplinary frameworks aim to evaluate the processes and outcomes of such research, ensuring that the efforts lead to meaningful societal benefits.

The VINNOVA Impact Logic Framework, developed by the Swedish Governmental Agency for Innovation Systems, evaluates the translation of research into practical innovations. It emphasises collaboration between academia, industry, and public entities, with indicators such as patents, public finance, and societal contributions to debates. This framework highlights the role of multidisciplinary research in driving innovation and addressing societal challenges (Kolbenstvedt et al., 2007). Another key model is the Global Challenges Research

Fund (GCRF), which supports interdisciplinary research tackling global issues. The GCRF prioritises societal impact, equity, and sustainable solutions, ensuring alignment with international development goals (Carden et al., 2023).

Policy and Social Science Research Impact Frameworks. Frameworks for policy and social science research focus on assessing the societal, cultural, and policy impacts of research. These frameworks bridge the gap between academic research and decision-making processes, ensuring that research effectively informs policies, practices, and public discourse.

The Decision-Making Impact Model, introduced in 2003, evaluates the impact of applied health research on diverse target audiences, including policymakers, clinicians, and the general public. This Canadian model tailors its evaluation to the needs of specific user groups, avoiding a one-size-fits-all approach. Drawing from the Payback Framework, it emphasises interactive measures such as user-pull and producer-push processes, which assess the uptake and utilisation of research in decision-making. The model supports evidence-based policy by fostering a culture of continuous research application and knowledge exchange (Lavis et al., 2003).

The Research Impact Framework (RIF), developed in the UK in 2006, categorises research impact into 4 domains: research-related, policy, service, and societal impacts. This framework builds upon earlier models, such as the Payback Framework, and highlights the value of research beyond academia. The RIF enhances accountability by focusing on real-world applications and demonstrates how research initiatives align with national and organisational goals. Key indicators include publications, patents, policy networks, and societal benefits such as empowerment and sustainable development (Kuruvilla et al., 2006).

The Flows of Knowledge, Expertise, and Influence Model, introduced in 2008, assesses the non-linear pathways of knowledge transfer in social science research. Developed by Meagher, Lyall, and Nutley, this UK model identifies researchers, policymakers, and intermediaries as critical actors in effective knowledge dissemination. It evaluates the impacts of research on policy formation, cultural attitudes, and professional practices, emphasising engagement processes and networks. The framework addresses the UK's priorities for demonstrating societal value from research investments, particularly in public health and quality of life (Meagher et al., 2008).

Environmental Health and Safety Research Impact Frameworks. Environmental health and safety research frameworks emerged to address the intricate relationships between environmental factors and public health outcomes. The increasing global emphasis on sustainability, public health, and environmental policy has propelled the development of evaluation models in this domain. These frameworks aim to assess the effectiveness of interventions, policies, and research programs in mitigating environmental hazards and improving societal well-being.

The National Institute of Environmental Health Sciences (NIEHS) Logic Model, developed in 2008, provides a structured approach for evaluating programs related to environmental health. This model outlines the relationships between inputs, activities, outputs, and outcomes, providing a comprehensive understanding of the short-term and long-term impacts of environmental health interventions. By aligning with national health priorities, the NIEHS Logic Model enables evidence-based decision-making and fosters public health improvements through targeted research and community engagement. It focuses on key aspects, such as environmental hazard awareness, policy changes, emission reductions, and behavioural shifts in the public, emphasising its societal relevance (Engel-Cox et al., 2008).

The prominence of environmental health and safety research frameworks reflects the global urgency to address environmental challenges, including climate change, pollution, and ecosystem degradation. These frameworks connect scientific research with actionable policy measures, ensuring that research outcomes contribute to tangible benefits for society and the environment. By facilitating interdisciplinary collaboration and emphasising real-world applications, these models address the growing need for sustainable solutions to complex environmental issues.

4.2.3 Core principles

Core principles guided the development of these frameworks, emphasising multidimensionality, accountability, and stakeholder engagement. Multidimensionality evaluated impacts across academic, economic, societal, and policy domains. At the same time, accountability highlighted the need for research to demonstrate returns on investment to funders and the public (Bernstein et al., 2007; REF, 2014). Stakeholder engagement played a crucial role, as frameworks were designed to be relevant to diverse audiences, including policymakers, healthcare providers, and industry partners (Ward et al., 2023). These principles enabled the development of flexible and adaptable tools that could be applied across various disciplines and sectors.

4.2.4 Strengths and limitations of second-generation frameworks

These frameworks offered significant strengths, including their ability to evaluate various impacts and their relevance to diverse stakeholders. Addressing societal and policy implications provided a more comprehensive view of research outcomes, ensuring that funding decisions aligned with public priorities (CAHS, 2009; Donovan & Hanney, 2011). Additionally, their flexibility allowed for customisation to suit specific disciplines or national contexts (Bernstein et al., 2007).

However, they also faced challenges. Many frameworks relied heavily on self-reported data, raising concerns about subjectivity and potential bias in evaluations (Wellcome Trust, 2009). Their implementation was often resource-intensive, requiring significant time and expertise, which posed challenges for institutions with limited resources (Sarli et al., 2010). Furthermore, the diversity of research impacts made it challenging to develop standardised metrics and methodologies that could be universally applied (Buykx et al., 2012).

Despite these challenges, the second generation remains highly relevant in research impact assessment. Integrating these frameworks into funding policies and institutional strategies has ensured their continued use in evaluating research outcomes. Their alignment with global priorities, such as the United Nations SDGs, highlights the enduring significance of addressing complex societal challenges (Bernstein et al., 2007; CAHS, 2009; Leal Filho et al., 2021). By providing a structured and multidimensional approach to assessment, these frameworks have set a benchmark for evaluating the broader value of research in a rapidly changing academic and policy environment.

4.3 Third generation: Data-driven and predictive research impact models

4.3.1 The shift towards computational intelligence

The third generation of research impact assessment marks a transformative shift toward leveraging advanced data science techniques to evaluate, predict, and understand the influence of research. This generation builds on the foundations of prior generations by integrating methodologies such as ML, natural language processing (NLP), and graph-based network analysis to address the growing complexity and scale of modern research ecosystems. Unlike earlier generations, which emphasised bibliometric indicators or multidimensional frameworks, the third generation introduces predictive capabilities, enabling stakeholders to anticipate future trends, evaluate emerging fields, and optimise resource allocation (Ji et al., 2024; Porwal & Devare, 2024).

Unlike the retrospective approaches of the first and second generations, which primarily analysed past research outputs and impacts, the third generation adopts a forward-looking, prospective perspective. It leverages advancements in data science, particularly the proliferation of ML algorithms, big data analytics, and semantic analysis tools, to forecast research impact aspects such as citation trajectories, interdisciplinary growth, and policy influence. These developments enable stakeholders to identify emerging trends, strategically allocate resources, and enhance decision-making processes. By harnessing the power of big data and computational methods, the third generation represents a significant step forward in addressing the dynamic and predictive requirements of modern research ecosystems (Vital & Amancio, 2022; Wu et al., 2022).

4.3.2 Key models, advanced analytical techniques and data-driven contributors

Citation Context, Trends, and Forecasting. Third-generation frameworks have significantly advanced research impact assessment by leveraging sophisticated methodologies in citation context, trends, prediction, and forecasting (see Appendix, Table 3). Techniques such as ML, deep learning, and NLP have enabled researchers to develop predictive models that analyse long-term citation trajectories and semantic content while identifying temporal dynamics in scholarly communication (Du et al., 2024; Ji et al., 2024; Porwal & Devare, 2024; Zhou et al., 2022). Recent innovations have introduced dynamic and context-aware models, including Dynamic Multi-Context Attention Networks (DMA-Nets), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs), which utilise temporal and network features for improved accuracy. These models rely on datasets such as the Microsoft Academic Graph and arXiv to identify trends in citation dynamics, offering valuable insights for recognising high-impact research outputs. Despite their potential, challenges such as computational demands and limited applicability across diverse domains persist as significant barriers (Abbas et al., 2023; Abrishami & Aliakbary, 2019; Ji et al., 2024; Li et al., 2019; Zhu & Ban, 2018).

Semantic metadata analysis has emerged as a transformative approach for enhancing citation prediction. By examining abstracts and technical terms using methodologies such as Document-to-Vector (Doc2Vec) and Bidirectional Long Short-Term Memory (Bi-LSTM) networks, researchers have uncovered the critical role of semantic features in influencing citation dynamics. These methods, validated across datasets in AI, life sciences, and computer science, demonstrate the capacity of semantic analysis to refine citation forecasting. Advanced tools, such as Bidirectional Encoder Representations from Transformers (BERT), are expected to further expand the scope of these models and enhance their predictive accuracy in broader disciplinary contexts (Baba & Baba, 2018; Baba et al., 2019; Ma et al., 2021; Porwal & Devare, 2024). Similarly, studies focusing on the Physics and Astronomy Classification Scheme (PACS) codes have employed decision trees and statistical models to analyse over 399,000 articles. These studies revealed strong correlations between classification codes and citation counts, emphasising the impact of domain-specific features such as "Condensed Matter" research (PACS 60). Nonetheless, the limited generalizability of these findings to other fields underscores the need to integrate additional features, including author networks and semantic content, into these frameworks (Enduri et al., 2022).

The temporal dynamics of citation bursts also play a crucial role in understanding the evolution of research influence. Enhanced models incorporating regression analyses and clustering techniques have improved the prediction of highly cited articles. Researchers have developed methodologies that capture long-tail citation distributions by utilising datasets such as AMiner (a research insight database) and the Open Academic Graph, offering more profound insights into the drivers of citation behaviours. Future efforts in this area aim to apply these models to more diverse datasets and incorporate dynamic features to refine predictions further (Amjad et al., 2022; Du et al., 2024; Pradhan et al., 2019). Deep learning models, such as sequence-to-sequence RNNs and Hierarchical Extreme Learning Machines (H-DELM), have also demonstrated their ability to handle high-dimensional data and predict citation trends. However, challenges related to model interpretability and scalability remain critical areas of concern, particularly in ensuring the applicability of these models across various academic disciplines (Li et al., 2019; Pobiedina & Ichise, 2016; Srinivasa, 2019; Zhou et al., 2022).

Academic Collaboration for Broader Research Impact. Academic collaboration is pivotal in enhancing research impact by uniting diverse expertise and resources. Data science methodologies have been crucial in analysing and facilitating these collaborations. Techniques such as network dynamics modelling, ML, and deep learning have been applied to predict future collaborations, refine link prediction in author networks, and detect anomalies within citation networks. Multi-network representation learning, integrating textual, structural, and temporal features, has proven particularly effective in fostering impactful academic partnerships. These methodologies, supported by Web of Science, COVID-19 Open Research Dataset (CORD-19), AMiner, and IEEE Xplore datasets, have provided actionable insights into collaborative structures' influence on research dissemination and citation metrics. Frameworks like Graph Learning for Anomaly Detection (GLAD), for example, help maintain the integrity of scholarly communication by detecting anomalous citations (Hou et al., 2023; Y. Liu et al., 2022; Vital & Amancio, 2022; Yang et al., 2023).

Research predicting future collaborations has demonstrated the efficacy of multi-network representation learning frameworks in improving link prediction accuracy. Studies employing knowledge graphs, node embeddings, and spatial-temporal factors have highlighted the importance of integrating diverse features to enhance prediction reliability. However, computational intensity and scalability issues remain critical challenges, necessitating future efforts to optimise models and extend them to other domains (Kanakaris et al., 2021; Makarov & Gerasimova, 2019; Yang et al., 2023; Zhang et al., 2019; Zhou et al., 2019). Similarly, efforts to enhance link prediction in author networks have utilised nodal-attribute-based predictors and ML models to achieve high accuracy. While effective, these methods' reliance on topological features underscores the need for incorporating temporal and content similarity metrics to improve scalability and practical applications (Roopashree & Umadevi, 2014; Song et al., 2022; Vital & Amancio, 2022).

Studies analysing co-authorship networks have revealed their significant influence on citation metrics and research impact. Researchers have identified collaboration trends across domains such as AI, ML, and medical research by examining network properties and employing methods like Structural Variation Analysis (SVA) and Propensity Score Matching (PSM). However, the need for more data on informal collaborations and discipline-specific focus presents limitations. Future research aims to expand these analyses to subfields, explore causal relationships, and refine computational methods for broader applicability (Grodzinski et al., 2021; Hou et al., 2023; Vinayak et al., 2023). Additionally, frameworks like GLAD have demonstrated their effectiveness in detecting anomalous citations using synthetic data. Yet, their reliance on artificial datasets and scalability challenges necessitates further refinement to apply them to real-world scenarios (J. Liu et al., 2022).

Specialised Research Impact. The application of data science to specialised facets of research impact has enabled more nuanced evaluations of journals, researchers, institutions, and academic domains. Techniques such as ML and deep learning have facilitated predictive modelling for journal metrics, researcher evaluation, and institutional impact. For instance, studies have employed k-Nearest Neighbors (kNN) imputation and regression models with Long Short-Term Memory (LSTM) networks to address missing journal impact factor (JIF) values, while eXtreme Gradient Boosting (XGBoost) models have been used to predict institutional impact (Bai et al., 2017; Croft & Sack, 2022; Hua & Huynh, 2024). However, the restricted scope of datasets and computational demands still need to be addressed. Expanding the scope of data fields and integrating advanced techniques such as altmetrics could provide valuable insights into these specialised areas (Croft & Sack, 2022; Hua & Huynh, 2024).

In researcher evaluation, models leveraging network features and proximities have demonstrated their ability to rank researchers and predict future impact. Researchers have highlighted critical parameters that influence research impact by employing CNNs, decision trees, and deep neural networks (DNNs). Limitations such as manual dataset selection and high computational costs suggest future directions for expanding domains, integrating additional factors, and ensuring scalability (Alshdadi et al., 2023; de Abreu Batista et al., 2021; He et al., 2022). Additionally, domain-specific studies have employed LSTM models and deep learning methods to develop classification schemes and analyse cultural impacts. However, these studies have demonstrated their utility, class imbalance and limited scope, highlighting opportunities for further refinement (Huang et al., 2022; Yaniasih & Budi, 2021).

Data-driven Contributors. Data-driven contributors are grounded in the quantitative analysis of empirical data associated with research activities and outputs and play a pivotal role in evaluating, predicting, and enhancing the impact of research (see Appendix, Table 4). These factors leverage various metrics and advanced analytics, such as citation metrics, predictive modelling, and collaboration network analysis, providing a robust framework for evidence-based decision-making. By quantifying research influence, these elements illuminate current impacts and enable the development of sophisticated models that project future trajectories of research influence.

Central to these metrics are citation-based indicators, which provide a detailed view of the scholarly impact of research outputs. Temporal dynamics such as inter-citation durations—the time intervals between consecutive citations—reveal patterns of influence that fluctuate over time, providing insights into the persistence and fading of scholarly impact (Ji et al., 2024). Additionally, citation arrival times, mainly when they occur shortly after publication in widely read or influential journals, can signal imminent spikes in citation frequency. This phenomenon is particularly significant in rapidly evolving fields such as AI, biotechnology, and renewable energy technologies (Abbas et al., 2023). Moreover, the aging effect is a critical factor, where older publications tend to see a plateau or decline in citation frequency, emphasising the need for research topics to remain relevant and timely (Abbas et al., 2023).

The context of citations also holds substantial value. Citation sentiment assesses the tone—positive, negative, or neutral—of citations, providing deeper insights into the academic reception of the cited work. Positive citations, especially those in prestigious journals, not only enhance the visibility of a paper but also indicate recognition of its quality and scholarly influence. Network centrality within citation networks also provides crucial data, as papers occupying central positions within these networks are often regarded as foundational or pivotal, resulting in frequent citations and high scholarly esteem (Ke et al., 2024; Shi et al., 2019).

Content features further influence the visibility and impact of research. The relevance of keywords and trending topics within a paper can significantly amplify its visibility and citation potential (Shi et al., 2019; Vinayak et al., 2023). For example, emerging terms like "blockchain" or "cybersecurity" in computer science can rapidly elevate a paper's prominence within the scholarly community. Linguistic patterns, including specific word sequences in research titles and abstracts, have been shown to correlate with higher citation counts. Advanced predictive models, such as LSTM networks, utilise these linguistic and historical citation patterns to forecast future citation performance, providing a forward-looking assessment of research impact (Du et al., 2024).

Collaboration and co-authorship networks are also critical in enhancing the impact of research. The structure of these networks, as quantified by metrics such as node degree and betweenness centrality, suggests potential for collaboration and the dissemination of ideas (Aljohani et al., 2021). Highly connected individuals or papers within these networks often achieve greater visibility and play a pivotal role in the diffusion of new research findings. The dynamics of such networks, including the roles of familiar neighbours and shared collaborators, often forecast ongoing and future collaborative efforts, which are vital for sustaining the influence and development of research domains (Makarov & Gerasimova, 2019; Song et al., 2022; Zhou et al., 2019).

Institutional and individual metrics, such as h-indices and productivity scores, quantitatively assess research impact and quality. High h-indices, which measure both the quantity and impact of an author's publications, indicate significant scholarly influence (de Abreu Batista et al., 2021; Heo et al., 2023). At the institutional level, collective h-indices, which aggregate the h-indices of all affiliated researchers, reflect an institution's overall research strength and impact, often correlating with leadership in advancing specific fields of study (Bai et al., 2017).

Journal metrics such as IF, SNIP, and SJR are crucial in assessing the relative influence of research publications (Croft & Sack, 2022). These metrics help researchers identify which journals offer the best potential for enhancing the visibility and impact of their work. Whether open-access or subscription-based, the publication type also significantly affects a paper's reach and impact. Open-access publications, due to their accessibility, tend to attract a wider audience, resulting in higher citation rates and broader dissemination of research findings.

Finally, economic and geographical factors significantly influence the development and impact of research. Geographic proximity to major academic conferences, research hubs, and technological centres enhances opportunities for networking, collaboration, and access to innovative resources (Zhang et al., 2019). Similarly, economic conditions, as measured by indicators such as GDP per capita, influence the level of research funding and infrastructure available, thereby shaping the volume and quality of research outputs from different regions. These factors create a complex ecosystem that fosters research activities, generating impact and providing a multifaceted picture of the dynamics in the global research landscape (Grodzinski et al., 2021).

4.3.3 Core principles

The third generation of research impact assessment is underpinned by 4 key principles: precision, scalability, adaptability, and predictiveness. Precision is achieved through advanced ML techniques that extract nuanced patterns from large and complex datasets, significantly enhancing the accuracy of predictions (Zhang et al., 2021). This level of detail enables researchers and institutions to anticipate future trends and understand the multifaceted aspects of research influence. Scalability ensures that these computationally intensive models can accommodate the analysis of global research datasets, making them applicable across diverse disciplines and geographic regions. This scalability is critical for addressing the increasing volume and diversity of research outputs in a globalised academic ecosystem (Vital & Amancio, 2022).

Adaptability is another cornerstone of third-generation frameworks, facilitated by integrating semantic, temporal, and social features into analytical models. This adaptability allows these frameworks to be customised for specific research domains or institutional priorities, enhancing their relevance and utility in varied contexts (Porwal & Devare, 2024). Finally, predictiveness defines the ability of these models to forecast future research trajectories, including citation patterns, collaboration trends, and overall research influence. By enabling proactive decision-making, these predictive capabilities empower institutions, policymakers, and funding agencies to allocate resources effectively and strategically (Li et al., 2019; Wu et al., 2019).

4.3.4 Strengths and limitations of third-generation models and analytical techniques

The third generation introduces significant strengths that enhance the impact of research assessment. One of its most notable advantages is its enhanced predictive power, which enables these models to forecast future trends, identify high-impact research, and highlight emerging disciplines with remarkable accuracy (Vital & Amancio, 2022). Integrating citation trends, semantic content, and network dynamics enables multidimensional insights, providing a holistic understanding of research influence that surpasses the capabilities of earlier generations (Porwal & Devare, 2024).

Additionally, advanced computational techniques that facilitate the analysis of datasets across diverse regions and disciplines ensure the global applicability of third-generation frameworks. This scalability ensures their relevance in an increasingly interconnected research landscape (Zhang et al., 2021). Furthermore, these frameworks provide strategic decision-making tools that guide institutions, funding agencies, and policymakers in resource allocation, strategic planning, and prioritising impactful research areas, enabling more effective and data-driven decision-making processes (Wu et al., 2022).

Despite their strengths, third-generation frameworks face challenges that limit their broader adoption and utility. These models' computational demands are substantial, requiring significant resources and expertise that are often unavailable to smaller or underfunded institutions, thereby creating disparities in access (Vital & Amancio, 2022). Data bias and quality present another limitation, as the dependence on diverse datasets introduces biases based on geographic and disciplinary disparities, which can skew predictions and reduce accuracy (Zhang et al., 2021).

Moreover, the interpretability issues inherent in complex ML models make their results difficult for non-technical stakeholders to understand and apply, hindering broader adoption in policymaking and institutional planning (Xiao et al., 2019). Finally, many frameworks exhibit domain-specific constraints, requiring significant customisation to be applicable across different fields, which increases implementation complexity and limits generalizability (Porwal & Devare, 2024).

4.4 Fourth generation: Alternative metrics

4.4.1 Emergence of altmetrics

The fourth generation of research impact assessment represents a shift from traditional bibliometric methods and structured frameworks to alternative metrics, known as altmetrics (Priem et al., 2011). These metrics capture the online influence of research. Unlike first-generation metrics, which emphasised academic citations and publication outputs, or second-generation frameworks, which focused on structured and multidimensional evaluation across specific domains, the fourth generation emerged as a response to the digitalisation of research dissemination.

The rise of social media, digital repositories, and online academic platforms has created new avenues for engaging with research, necessitating metrics that can evaluate non-traditional research outputs and diverse forms of impact in real-time. Altmetrics measure interactions such as mentions, shares, and downloads, enabling real-time assessment of public, academic, and policymaker engagement. This generation has broadened the scope of impact evaluation by addressing aspects often overlooked by prior approaches, such as societal discourse, public outreach, and online academic collaboration (Sugimoto et al., 2017). Table 5 in the Appendix provides an overview of key metrics and platforms introduced in this generation, highlighting their role in advancing research evaluation practices and broadening the understanding of impact.

4.4.2 Key metrics and platforms

Aggregated Metrics Platforms. Metrics like the Altmetric Attention Score, introduced in 2011, aggregate data from multiple sources, including social media platforms, news outlets, policy documents, and blogs, to provide a weighted and comprehensive snapshot of a research output's online visibility. Represented visually through the "Altmetric donut," this score reflects the varying importance of different types of engagement (Priem, 2010; Priem et al., 2011; Roemer & Borchardt, 2015). Similarly, PlumX Metrics, launched in 2016, categorises online interactions into 5 dimensions—Usage, Captures, Mentions, Social Media, and Citations—offering researchers and institutions a detailed breakdown of how their work is consumed and shared digitally (Lindsay, 2016). These tools provide real-time analytics that can guide researchers in optimising dissemination strategies.

The PLOS Article-Level Metrics, introduced in 2009, were among the first to deliver granular insights at the article level, tracking views, downloads, and citations alongside social media mentions. This approach enabled researchers to evaluate the impact of individual articles, bypassing traditional journal-level assessments (Yan & Gerstein, 2011). Expanding on these ideas, Dimensions Badges, launched in 2018, integrate citation counts from conventional sources with altmetric data, creating a more holistic framework for evaluating research influence (Jamwal & Kumar, 2022). These platforms bridge the gap between traditional bibliometric measures and alternative metrics, aligning with the principles of multidimensionality introduced in the second generation.

Specialised Platforms for Alternative Metrics. Several platforms focus on specific dimensions of research dissemination. For example, Impactstory Metrics, introduced in 2012, enable researchers to track online engagement across platforms such as Twitter, GitHub, and Mendeley, providing comprehensive overviews of their work's societal and academic reach (Konkiel, 2014); similarly, Kudos Metrics, established in 2014, focuses on enabling researchers to explain and share their work effectively, tracking metrics that correlate visibility-enhancing activities with engagement outcomes (Erdt et al., 2017).

Platform-specific Metrics. Beyond aggregated platforms, specific platforms, such as ResearchGate Score and Mendeley Readers, provide unique metrics tailored to their respective user communities. For instance, the ResearchGate Score, developed in 2016, considers factors like publications, reads, and citations, reflecting a blend of academic and social engagement, albeit through a proprietary algorithm (Hoffmann et al., 2016). Similarly, Mendeley in 2009 measures readership by counting how often a research output is added to users' libraries, serving as a proxy for interest and accessibility (Sugimoto et al., 2017; Zahedi & Costas, 2020). These metrics offer insights into the impact of researchers and facilitate peer-to-peer connections.

Platforms like GitHub Repository Stars/Forks reflect the impact of research outputs such as software and code repositories, which are increasingly critical in fields like computer science and bioinformatics. These metrics measure user interaction, reuse, and adaptation, demonstrating the tangible application of research in practice (Dozmorov, 2018). Tools like Crossref Event Data, introduced in 2016, expand this by collecting mentions and interactions associated with research DOIs across various online platforms (Rittman, 2020).

Digital and Social Media Engagement. Altmetrics also emphasises the role of social media and online platforms in amplifying research influence. Metrics such as Social Media Mentions, Wikipedia Citations, and YouTube Video Views track engagement across platforms like Twitter, LinkedIn, and Wikipedia, capturing the broader public and societal interaction with research outputs. Wikipedia citations, for example, measure the frequency with which scholarly works are referenced in articles, highlighting their role in disseminating public knowledge (Stalder & Hirsh, 2002). Similarly, YouTube video views quantify the reach of research-related presentations and discussions, showcasing the value of visual and multimedia content in engaging diverse audiences (Dai & Wang, 2023).

4.4.3 Core principles

The fourth generation builds upon the principles of multidimensionality established in the second generation but emphasises immediacy and inclusivity to address the limitations of earlier approaches. While second-generation frameworks, such as the Payback Model and the CAHS Framework, systematically linked research activities to societal and policy outcomes, they often relied on extensive data collection and long-term evaluations. Altmetrics, by contrast, utilises real-time data from digital platforms to provide immediate feedback on research influence. This immediacy enables researchers and stakeholders to assess engagement with newly published work and emerging fields promptly. Additionally, the inclusivity of altmetrics enables the evaluation of diverse research outputs, including datasets, software, multimedia, and policy briefs, thereby expanding the scope of what constitutes impactful research (Priem et al., 2011; Roemer & Borchardt, 2015).

4.4.4 Strengths and limitations of fourth-generation metrics

The fourth generation brought about a fundamental shift in understanding the influence of research by recognising the value of public engagement, including policy discussions, community dialogues, and the application of research in non-academic settings. By capturing real-time mentions, shares, and conversations across platforms such as Twitter, Facebook, and ResearchGate, altmetrics provide insights into how research impacts diverse audiences beyond academia. This broader perspective aligns with the objectives of second-generation frameworks, such as the REF and the EIF, which emphasise societal relevance and public accountability (ARC, 2017; REF, 2014). However, altmetrics differ in their ability to capture these impacts dynamically and across digital spaces, addressing the growing demand for immediacy in impact evaluation.

Despite their strengths, the fourth generation faces notable limitations. Altmetrics relies on data from digital platforms, which can introduce biases based on geographic and cultural disparities in platform usage. For instance, regions with lower social media penetration may be underrepresented, leading to incomplete evaluations (Sugimoto et al., 2017). Additionally, altmetrics focus on attention rather than sentiment, failing to distinguish between positive and negative engagement. Practices such as coordinated sharing and gaming of metrics can also artificially inflate scores, thereby undermining the reliability of these tools (Roemer & Borchardt, 2015; Stalder & Hirsh, 2002). These challenges necessitate cautious interpretation and the development of more robust methodologies to enhance the accuracy and credibility of altmetrics.

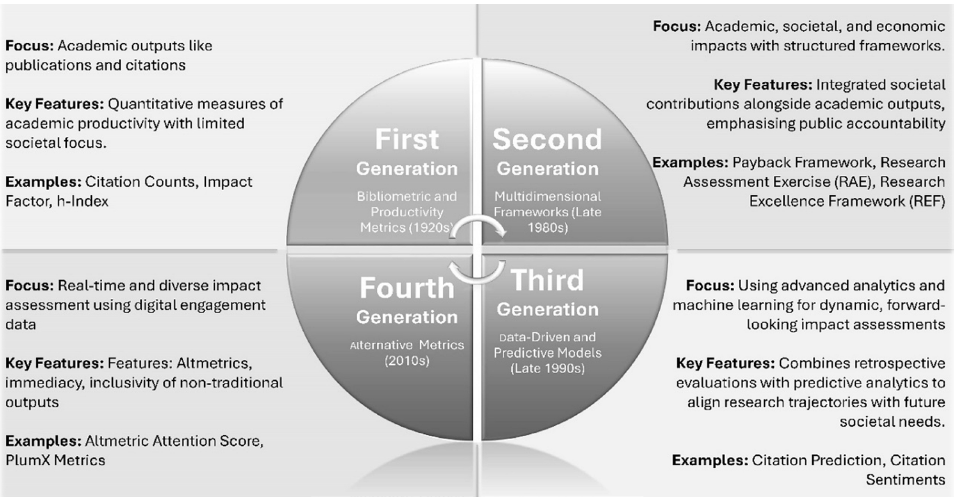
Nevertheless, integrating altmetrics into widely used platforms such as ResearchGate, Mendeley, and PLOS underscores their continued relevance in the research ecosystem. These metrics align with the principles of open science, fostering greater accessibility and accountability by emphasising public and policy engagement. Their ability to evaluate diverse outputs and provide timely feedback ensures their utility for stakeholders across academia, industry, and government (Hoffmann et al., 2016; Jamwal & Kumar, 2022). Altmetrics complements the comprehensive evaluation frameworks of the second generation by providing an agile and dynamic approach to assessing research influence, reflecting the evolving needs of a digitally connected world.

5. Discussion

5.1 Comparative analysis of research impact generations

The progressive development of research impact assessment reflects significant transformations in methodologies, principles, and tools, each adapting to the evolving demands of academia and society (see Figure 2). The first generation laid the foundation by focusing on quantitative metrics, emphasising citation

Figure 2. Four Generations of Research Impact Science



counts, publication numbers, and journal-level evaluations. Metrics such as the IF and h-index were pivotal in providing standardised tools for systematically assessing research output, establishing the groundwork for bibliometric databases like Scopus and Web of Science (Garfield, 1955; Garfield & Sher, 1963; Hirsch, 2005). The second generation expanded the scope to multidimensional frameworks, addressing broader societal, economic, health, and policy impacts. Frameworks like the Payback Model and the CAHS Framework redefined evaluation by integrating societal goals, demonstrating how research outcomes could inform public priorities and policymaking (CAHS, 2009; Donovan & Hanney, 2011).

The third generation transitioned from retrospective analyses to forward-looking evaluations by incorporating data-driven and predictive models. It leveraged advancements in data science, such as ML, big data analytics, and NLP, to analyse large-scale, multidimensional datasets. These tools enabled stakeholders to anticipate citation trajectories, identify emerging fields, and optimise resource allocation, emphasising the predictive and strategic potential of research impact assessment (Ji et al., 2024; Porwal & Devare, 2024). The fourth generation embraced alternative metrics (altmetrics) as a response to the digitalisation of research dissemination. Altmetrics introduced real-time and diverse influence indicators, capturing the broader societal, policy, and online presence of research. By measuring interactions such as social media mentions, shares, downloads, and multimedia engagement, this generation broadened the understanding of research impact in the digital ecosystem, providing timely feedback and inclusivity in evaluating non-traditional research outputs (Priem et al., 2011; Sugimoto et al., 2017).

The underlying principles guiding these generations evolved to reflect increasing complexity in research evaluation. Objectivity, simplicity, and replicability defined the first generation, enabling standardised and transparent metrics that became the basis for bibliometric analyses (Garfield & Sher, 1963; Hirsch, 2005). These principles laid the foundation for robust bibliometric databases, facilitating consistency and accessibility. The second generation introduced multidimensionality, accountability, and stakeholder engagement, extending evaluation to encompass societal, economic, and policy impacts. This era emphasised aligning research priorities with national and global objectives, fostering relevance and collaboration (Bernstein et al., 2007; REF, 2014).

The third generation shifted toward predictiveness, scalability, and interdisciplinarity, leveraging advancements in data science, including ML, NLP, and big data analytics. These tools enabled forward-looking evaluations by identifying emerging trends, predicting citation trajectories, optimising resource allocation, and addressing the dynamic needs of modern research ecosystems (Ji et al., 2024; Porwal & Devare, 2024). The fourth generation prioritised immediacy and inclusivity, leveraging digital data to evaluate non-traditional outputs, such as datasets, software, and policy briefs. Alternative metrics (altmetrics) enabled real-time feedback on research influence, capturing broader societal and online engagement and addressing diverse research outputs within the digital ecosystem (Priem et al., 2011; Roemer & Borchardt, 2015).

The metrics and indicators of these generations illustrate a growing sophistication in evaluating research impact. The first generation relied on bibliometric indices such as citation counts, publication numbers, and journal-level metrics like the IF, with author-level metrics such as the h-index offering insights into individual productivity and influence (Garfield & Sher, 1963; Hirsch, 2005). The second generation expanded evaluation to include health, policy, and economic metrics such as quality-adjusted life years (QALYs), return on investment (ROI), and legislative influence, reflecting the integration of societal and policy considerations (CAHS, 2009; Sarli et al., 2010). The third generation significantly advanced by integrating predictive models and advanced analytics. It employed citation dynamics, semantic analysis, and network metrics to enable multidimensional and forward-looking assessments. Methodologies such as BERT and Doc2Vec have facilitated predictions about citation trajectories, interdisciplinary collaboration, and broader research trends, reflecting the rise of data-driven approaches (Ji et al., 2024; Porwal & Devare, 2024). The fourth generation embraced digital engagement, introducing tools such as Altmetric Attention Scores and PlumX Metrics to measure influence across online platforms. Platform-specific metrics, such as the ResearchGate Score, highlighted the digital ecosystem's growing importance in evaluating research impact, capturing interactions like social media engagement and public discourse in real-time (Roemer & Borchardt, 2015; Sugimoto et al., 2017).

The frameworks supporting these metrics demonstrate the progression from traditional bibliometric methods to more multidimensional and predictive approaches. Using standardised metrics, the first generation relied on platforms like Web of Science and Google Scholar to measure academic output and influence within defined disciplinary boundaries (Falagas et al., 2008). Second-generation frameworks, such as the CAHS Framework and REF, have tailored their approaches to specific domains, including health, higher education, and policy, thereby highlighting broader societal impacts. These frameworks emphasised multidimensionality and accountability, aligning research evaluation with societal goals and public priorities (Engel-Cox et al., 2008; Lavis et al., 2003). The third generation incorporated advanced predictive models and data science-driven frameworks. It utilised methodologies like DMA-Nets and LSTM to enable predictive analysis, offering insights into citation patterns, collaboration trends, and research trajectories. These tools reflected the growing importance of forward-looking evaluations in modern research ecosystems (Du et al., 2024). The fourth generation leveraged digital platforms, such as Altmetric and ResearchGate, to capture social media engagement and digital dissemination. These frameworks expanded the scope of research evaluation by integrating real-time metrics to assess diverse forms of societal and online engagement, emphasising inclusivity and immediacy in capturing research impact (Priem et al., 2011; Yan & Gerstein, 2011).

Each generation's strengths reflect its alignment with the demands of its time. The first generation's metrics were simple, objective, and widely applicable, ensuring accessibility and reproducibility (Garfield & Sher, 1963; Hirsch, 2005). The second generation provided a holistic perspective, addressing diverse impacts and demonstrating accountability through evidence of tangible returns on investment (CAHS, 2009). The third generation excelled in predictive capabilities and scalability, leveraging data science to analyse complex, multidimensional datasets and providing actionable insights for strategic decision-making (Vital & Amancio, 2022; Zhang et al., 2021). Finally, the fourth generation stood out for its real-time feedback and inclusivity, capturing diverse outputs and societal engagement through digital metrics such as social media and online interactions (Priem et al., 2011; Sugimoto et al., 2017).

Despite these strengths, each generation faced limitations. The first generation's reliance on quantitative metrics led to disciplinary bias and vulnerability to manipulation, such as self-citation practices (Bornmann & Marx, 2013). The second generation required significant resources, introduced subjectivity through self-reported data, and struggled with standardisation due to its complexity (Sarli et al., 2010). The third generation faced computational demands and interpretability challenges, which limited accessibility and broader adoption, particularly for underfunded institutions (Porwal & Devare, 2024; Xiao et al., 2019). The fourth generation encountered data bias and variability across platforms, which reduced metric reliability and hindered the differentiation between positive and negative engagement (Sugimoto et al., 2017).

The platforms and tools associated with each generation illustrate the progressive methodological advancements in research impact assessment. The first generation introduced foundational citation analysis and tracking tools, including Scopus and Web of Science platforms. This generation was characterised by indices and metrics such as citation counts, publication numbers, journal-level evaluations like the IF, and author-level metrics like the h-index, laying the groundwork for systematic bibliometric analyses. In this study, 26 key indices and metrics were collected, tabulated, and examined as representative examples of this foundational generation (Garfield, 1955; Hirsch, 2005). The second generation expanded the evaluation scope to societal and policy impacts through multidimensional frameworks, such as the REF and the NIEHS Logic Model. While this generation encompasses numerous frameworks, this study focused on 41 representative research impact frameworks, highlighting their integration of diverse indicators to address broader societal and economic contributions (CAHS, 2009; REF, 2014). The third generation marked a transition to predictive modelling, leveraging data-driven methodologies such as DMA-Nets, RNNs, and CNNs. In this study, 17 aspects of research impact—including citation prediction, sentiment analysis, and emerging research trends—were explored alongside 9 key groups of data-driven factors, such as temporal dynamics parameters, network analysis, and big data integration (Ji et al., 2024; Porwal & Devare, 2024). Finally, the fourth generation leveraged digital and societal engagement through tools like Altmetric and Mendeley, which captured real-time metrics to assess diverse outputs, including social media mentions and multimedia interactions. This study reviewed and tabulated 16 prominent alternative metrics, including the Altmetric Attention Score, PlumX Metrics, and Kudos Metrics, emphasising their role in broadening the inclusivity and scope of research impact (Priem et al., 2011; Roemer & Borchardt, 2015).

This comparative analysis addresses the overarching research question by systematically organising the evolution of research impact science into 4 distinct generational frameworks. Each generation reflects a response to the changing demands of academia and society, illustrating a clear progression in methodologies, principles, and tools. From the foundational bibliometric metrics of the first generation to the multidimensionality of the second, the predictive capabilities of the third, and the inclusivity of the fourth, these frameworks reveal how research impact assessment has evolved to encompass broader academic, societal, and technological priorities. The findings demonstrate the methodological advancements and underlying principles that guide these frameworks by synthesising these transitions. This analysis lays the groundwork for addressing the sub-questions by highlighting each generation's key characteristics, strengths, and limitations while emphasising the transformative role of data science in shaping modern research impact assessment.

5.2 Implications of the study

This review highlights the evolution of research impact assessment across 4 generational frameworks, mapping their underlying principles, metrics, and methodologies. Synthesising existing knowledge underscores the theoretical progression from simple bibliometric measures to advanced, predictive, and digital-driven approaches. The findings demonstrate how each generational shift reflects broader changes in academic and societal priorities, offering insights into the theoretical underpinnings of impact assessment methodologies. This synthesis identifies opportunities for further theoretical refinement, particularly in integrating data-driven analytics and interdisciplinary approaches into future frameworks. By mapping trends and gaps, this review provides a foundation for scholars to build more cohesive and theoretically informed research impact assessment models.

This review identifies actionable insights from existing research impact frameworks, emphasising their relevance to stakeholders such as policymakers, funding agencies, and academic institutions. By mapping the evolution of research impact science into 4 distinct generational frameworks, the findings address the overarching research question by offering a systematic organisation of methodologies and their progression over time. Furthermore, this synthesis aligns with the sub-questions by highlighting the key characteristics, strengths, limitations, and the transformative role of data science in shaping the modern frameworks of each generation. For example, the fourth generation's focus on real-time altmetrics offers lessons on leveraging digital platforms to enhance research visibility and engagement. In contrast, the third generation's predictive capabilities highlight the potential for resource optimisation and proactive planning. These insights can guide decision-makers in adopting and adapting impact assessment tools that align with their priorities, ensuring that research activities remain responsive to evolving societal and policy demands. Furthermore, the review highlights the importance of striking a balance between computational sophistication and accessibility, particularly for institutions with limited resources.

5.3 Limitations and future directions

This scoping review provides a comprehensive synthesis of research impact assessment frameworks; however, several limitations warrant attention. First, the review relies exclusively on secondary data and existing literature, which inherently introduce selection biases tied to the availability and accessibility of sources. While the breadth of frameworks covered offers valuable insights into their evolution across disciplines, the lack of empirical data—such as stakeholder interviews, case studies, or longitudinal analyses—limits the review's ability to address the contextual nuances influencing research impact within specific fields. Addressing this gap through empirical methodologies would enhance understanding of the principles and characteristics underpinning different frameworks, directly contributing to the systematic organisation of research impact science into distinct generational models, as highlighted in Sub-question 1.

Another area for improvement lies in the geographic focus of many frameworks, which are predominantly drawn from research conducted in high-income countries, primarily the United States, the United Kingdom, and Australia. While many frameworks have been developed in high-income countries, addressing their regional biases is critical to ensuring global applicability and incorporating the unique challenges and contributions of low- and middle-income regions. Such limitations restrict the ability to comprehensively identify the strengths and limitations of frameworks across diverse contexts, as explored in Sub-question 2. Developing a standardised taxonomy or nomenclature for research impact indicators would address this gap, facilitating cross-disciplinary and international comparisons. Collaborative efforts among researchers, policymakers, and funding agencies will ensure the practical relevance and applicability of a taxonomy.

Emerging trends in data science, AI, and large language models (LLMs) also pose challenges to the timeliness of this review. The rapid evolution of these technologies makes it challenging to capture their transformative potential for research impact assessment fully. For instance, integrating predictive analytics, ML models, and LLMs into impact frameworks remains an underexplored area that could significantly enhance forecasting capabilities, streamline knowledge transfer, and improve the visibility of impactful research. Investigating the role of these tools in modern research ecosystems is crucial for addressing Sub-question 3, particularly in understanding how data science can complement traditional bibliometric tools and inform interdisciplinary and emerging research trajectories.

Future research should prioritise empirical approaches to address these limitations. Stakeholder interviews, case studies, and longitudinal analyses can provide deeper insights into the contextual factors that shape research impact across different fields and regions. This would enable a more comprehensive evaluation of the characteristics and principles underlying different frameworks, contributing to the systematic refinement of research impact science. Methodologically, integrating advanced analytics, AI-driven models, and predictive frameworks into assessments could offer a more dynamic and scalable understanding of research trajectories. Dedicated evaluations of how ML and LLMs can automate impact identification, map citation networks, and forecast interdisciplinary collaborations hold significant promise for advancing evidence-based decision-making.

Furthermore, addressing gaps in emerging fields such as AI, biotechnology, and blockchain will be essential to ensure frameworks remain relevant and adaptive to diverse research impact pathways. Collaborative efforts across sectors will align future frameworks with societal and technological advancements, ensuring accessibility and utility for institutions with varying resources. Developing methodologies tailored to these rapidly evolving fields will refine existing frameworks, enhancing their scalability and precision to foster actionable insights for institutional, national, and global stakeholders.

6. Conclusion

This scoping review systematically examined the evolution of research impact assessment, categorising its progression into 4 generations and addressing the overarching research question: How has research-impact science evolved, and how can its progression be systematically captured? The study analysed 26 bibliometric indices and metrics in the first generation, 41 multidimensional frameworks in the second generation, 17 aspects and 9 groups of data-driven factors in the third generation, and 16 prominent alternative metrics in the fourth generation. This generational approach underscores the field’s shift from traditional bibliometric tools to multidimensional, predictive, and data-driven methodologies.

The first generation established foundational bibliometric tools for measuring academic productivity, while the second generation introduced multidimensional frameworks aligned with societal and policy goals. The third generation incorporated data science and predictive analytics for forward-looking evaluations, and the fourth generation leveraged alternative metrics to capture real-time digital engagement and non-traditional research outputs.

This review highlighted the strengths and limitations of these frameworks, underscoring the transformative role of data science in research impact assessment. Stakeholder engagement emerged as a critical factor, enhancing relevance and alignment with societal needs. Future advancements should integrate predictive analytics, AI-driven models, and standardised taxonomies to improve consistency and foresight. Empirical studies involving stakeholder collaboration and longitudinal analyses will further contextualise impact pathways.

In conclusion, research impact assessment has evolved to capture the benefits that research provides to society. By embracing innovation and fostering collaboration, the field is well-positioned to address contemporary challenges and drive meaningful societal change. Continued refinement of these frameworks is essential for their relevance and effectiveness in an increasingly complex world.

References

Abbas, K., Hasan, M. K., Abbasi, A., Mokhtar, U. A., Khan, A., Abdullah, S. N. H. S., Dong, S., Islam, S., Alboaneen, D., & Ahmed, F. R. A. (2023). Predicting the future popularity of academic publications using deep learning by considering it as temporal citation networks. *IEEE Access, 11*, 83052-83068. https://doi.org/10.1109/ACCESS.2023.3290906

Abrishami, A., & Aliakbary, S. (2019). Predicting citation counts based on deep neural network learning techniques. *Journal of Informetrics, 13*(2), 485-499. https://doi.org/10.1016/j.joi.2019.02.011

Aljohani, N. R., Fayoumi, A., & Hassan, S.-U. (2021). A novel deep neural network-based approach to measure scholarly research dissemination using citations network. *Applied Sciences, 11*(22), Article 10970. https://doi.org/10.3390/app112210970

Alnowaiser, K. (2024). Scientific text citation analysis using CNN features and ensemble learning model. *PLOS ONE, 19*(5), Article e0302304. https://doi.org/10.1371/journal.pone.0302304

Alshdadi, A. A., Usman, M., Alassafi, M. O., Afzal, M. T., & AlGhamdi, R. (2023). Formulation of rules for the scientific community using deep learning. *Scientometrics, 128*(3), 1825-1852. https://doi.org/10.1007/s11192-023-04633-5

Amjad, T., Shahid, N., Daud, A., & Khatoon, A. (2022). Citation burst prediction in a bibliometric network. *Scientometrics, 127*(5), 2773-2790. https://doi.org/10.1007/s11192-022-04344-3

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social*

Research Methodology, 8(1), 19-32. https://doi.org/10.1080/1364557032000119616

Australian Research Council. (2017). *EI 2018 submission guidelines*. https://www.arc.gov.au/sites/default/files/2022-07/era_ei_2018_submission_guidelines_0.pdf

Australian Research Council. (2018). *ERA 2018 submission guidelines*. https://web.archive.org.au/awa/20220302235108mp_/https://www.arc.gov.au/file/3781/download?token=Wq9o-CbM

Baba, T., & Baba, K. (2018). Citation count prediction using non-technical terms in abstracts. In O. Gervasi, B. Murgante, S. Misra, E. Stankova, C. M. Torre, A. M. A. C. Rocha, D. Taniar, B. O. Apduhan, E. Tarantino, & Y. Ryu (Eds.), *Computational science and its applications–ICCSA 2018* (pp. 366-375). Springer. https://doi.org/10.1007/978-3-319-95162-1_25

Baba, T., Baba, K., & Ikeda, D. (2019). Citation count prediction using abstracts. *Journal of Web Engineering, 18*(1-3), 207-228. https://journals.riverpublishers.com/index.php/JWE/article/view/3181

Bai, X., Zhang, F., Hou, J., Xia, F., Tolba, A., & Elashkar, E. (2017). Implicit multi-feature learning for dynamic time series prediction of the impact of institutions. *IEEE Access, 5*, 16372-16382. https://doi.org/10.1109/ACCESS.2017.2739179

Banzi, R., Moja, L., Pistotti, V., Facchini, A., & Liberati, A. (2011). Conceptual frameworks and empirical approaches used to assess the impact of health research: An overview of reviews. *Health Research Policy & Systems, 9*, Article 26. https://doi.org/10.1186/1478-4505-9-26

Bernstein, A., Hicks, V., Borbey, P., Campbell, T., McAuley, L., & Graham, I. D. (2007). A framework to measure the impact of investments in health research. In *Science, technology and innovation indicators in a changing world: Responding to policy needs* (pp. 231-250). OECD Publishing. https://doi.org/10.1787/9789264039667-en

Boaden, R. J., & Cilliers, J. J. (2001). Quality and the research assessment exercise: Just one aspect of performance? *Quality Assurance in Education, 9*(1), 5-13. https://doi.org/10.1108/09684880110381283

Boell, S. K., & Wilson, C. S. (2010). Journal Impact Factors for evaluating scientific performance: Use of h-like indicators. *Scientometrics, 82*(3), 613-626. https://doi.org/10.1007/s11192-010-0175-y

Börner, K., Glänzel, W., Scharnhorst, A., & van den Besselaar, P. (2011). Modeling science: Studying the structure and dynamics of science. *Scientometrics, 89*(1), 347-348. https://doi.org/10.1007/s11192-011-0429-3

Bornmann, L., & Marx, W. (2013). How good is research really? Measuring the citation impact of publications with percentiles increases correct assessments and fair comparisons. *EMBO reports, 14*(3), 226-230. https://doi.org/10.1038/embor.2013.9

Boshoff, N., & de Jong, S. P. L. (2020). Conceptualizing the societal impact of research in terms of elements of logic models:

A survey of researchers in sub-Saharan Africa. *Research Evaluation, 29*(1), 48-65. https://doi.org/10.1093/RESEVAL/RVZ020

BRIF workshop group. (2011). The role of a bioresource research impact factor as an incentive to share human bioresources. *Nature Genetics, 43*(6), 503-504. https://doi.org/10.1038/ng.831

Buykx, P., Humphreys, J., Wakerman, J., Perkins, D., Lyle, D., McGrail, M., & Kinsman, L. (2012). 'Making evidence count': A framework to monitor the impact of health services research. *The Australian Journal of Rural Health, 20*(2), 51-58. https://doi.org/10.1111/j.1440-1584.2012.01256.x

Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information, 22*(2), 191-235. https://doi.org/10.1177/053901883022002003

Canadian Academy of Health Sciences. (2009). *Making an impact: A preferred framework and indicators to measure returns on investment in health research*. https://www.cahs-acss.ca/wp-content/uploads/2011/09/ROI_FullReport.pdf

Carden, F., Vogel, I., Hepworth, C., & Stevenson, C. (2023). *Evaluation of the Global Challenges Research Fund: Midpoint synthesis report: Assessing quality, impact positioning and early outcomes against GCRF's theory of change*. Department for Business, Energy & Industrial Strategy. https://assets.publishing.service.gov.uk/media/65b39720a0ae1b00125260b2/evaluation_of_the_gcrf__assessment_of_research_quality__positioning_for_use_and_results.pdf

Commonwealth Scientific Industrial Research Organisation. (2024). *Impact evaluation guide*. https://www.csiro.au/~/media/About/Files/Our-impact-framework/CSIROImpactEvaluationGuide_WEB.pdf

Connor, J. (2011, November 16). *Google Scholar citations open to all*. Google Scholar Blog. https://scholar.googleblog.com/2011/11/google-scholar-citations-open-to-all.html

Croft, W. L., & Sack, J.-R. (2022). Predicting the citation count and CiteScore of journals one year in advance. *Journal of Informetrics, 16*(4), Article 101349. https://doi.org/10.1016/j.joi.2022.101349

Dai, X., & Wang, J. (2023). Effect of online video infotainment on audience attention. *Humanities & Social Sciences Communications, 10*, Article 421. https://doi.org/10.1057/s41599-023-01921-6

de Abreu Batista, A., Jr., Gouveia, F. C., & Mena-Chalco, J. P. (2021). Predicting the Q of junior researchers using data from the first years of publication. *Journal of Informetrics, 15*(2), Article 101130. https://doi.org/10.1016/j.joi.2021.101130

Deloitte Access Economics. (2011). *Returns on NHMRC funded research and development*. Australian Society for Medical Research. https://asmr.org.au/wp-content/uploads/library/NHMRCReturns.pdf

Dembe, A. E., Lynch, M. S., Gugiu, P. C., & Jackson, R. D. (2014). The translational research impact scale: Development, construct validity, and reliability testing. *Evaluation & the Health Professions, 37*(1), 50-70. https://doi.org/10.1177/0163278713506112

Donovan, C., & Hanney, S. (2011). The 'Payback Framework' explained. *Research Evaluation, 20*(3), 181-183. https://doi.org/10.3152/095820211X13118583635756

Dozmorov, M. G. (2018). GitHub statistics as a measure of the impact of open-source bioinformatics software. *Frontiers in Bioengineering & Biotechnology, 6*, Article 198. https://doi.org/10.3389/fbioe.2018.00198

Du, W., Li, Z., & Xie, Z. (2024). A modified LSTM network to predict the citation counts of papers. *Journal of Information Science, 50*(4), 894-909. https://doi.org/10.1177/01655515221111000

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics, 69*(1), 131-152. https://doi.org/10.1007/s11192-006-0144-7

Elsevier. (2024, August 14). *Metrics in SciVal – What are they and what are their strengths and weaknesses?* Journal Article Publishing Support Center. https://service.elsevier.com/app/answers/detail/a_id/13936/supporthub/publishing/

Emamy, K., & Cameron, R. (2007). Citeulike: A researcher's social bookmarking service. *Ariadne, 2007*(51). http://www.ariadne.ac.uk/issue/51/emamy-cameron/

Enduri, M. K., Sankar, V. U., & Hajarathaiah, K. (2022). Empirical study on citation count prediction of research articles. *Journal of Scientometric Research, 11*(2), 155-163. https://doi.org/10.5530/jscires.11.2.17

Engel-Cox, J. A., van Houten, B., Phelps, J., & Rose, S. W. (2008). Conceptual model of comprehensive research metrics for improved human health and environment. *Environmental Health Perspectives, 116*(5), 583-592. https://doi.org/10.1289/ehp.10925

Erdt, M., Aung, H. H., Aw, A. S., Rapple, C., & Theng, Y.-L. (2017). Analysing researchers' outreach efforts and the association with publication metrics: A case study of Kudos. *PLOS ONE, 12*(8), Article e0183217. https://doi.org/10.1371/journal.pone.0183217

Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008). Comparison of SCImago journal rank indicator with journal impact factor. *The FASEB Journal, 22*(8), 2623-2628. https://doi.org/10.1096/fj.08-107938

Fernandes, G. V. O., & Fernandes, J. C. H. (2024). GFsa (GF "Scientific Age") Index application for assessment of 1020 highly cited researchers in dentistry: A pilot study comparing GFsa Index and H-Index. *Publications, 12*(2), Article 18. https://doi.org/10.3390/publications12020018

Fu, H.-Z., & Ho, Y.-S. (2014). Top cited articles in adsorption research using Y-index. *Research Evaluation, 23*(1), 12-20. https://doi.org/10.1093/reseval/rvt018

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science, 122*(3159), 108-111. https://doi.org/10.1126/science.122.3159.108

Garfield, E. (1979). *Citation indexing: Its theory and application in science, technology, and humanities*. Wiley.

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation, 14*(3), 195-201. https://doi.org/10.1002/asi.5090140304

Giménez Toledo, E. (2018). La evaluación de las humanidades y de las ciencias sociales en revision [Research assessment in humanities and social sciences in review]. *Revista Española de Documentación Científica, 41*(3), Article e208. https://doi.org/10.3989/redc.2018.3.1552 (in Spanish)

Glänzel, W., & Schubert, A. (1995). Predictive aspects of a stochastic model for citation processes. *Information Processing & Management, 31*(1), 69-80. https://doi.org/10.1016/0306-4573(94)E0012-Q

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health, 89*(9), 1322-1327. https://doi.org/10.2105/ajph.89.9.1322

Grodzinski, N., Grodzinski, B., & Davies, B. M. (2021). Can co-authorship networks be used to predict author research impact? A machine-learning based analysis within the field of degenerative cervical myelopathy research. *PLOS ONE, 16*(9), Article e0256997. https://doi.org/10.1371/journal.pone.0256997

He, C., Wu, J., & Zhang, Q. (2022). Proximity-aware research leadership recommendation in research collaboration via deep neural networks. *Journal of the Association for Information Science & Technology, 73*(1), 70-89. https://doi.org/10.1002/asi.24546

Heo, G. E., Ko, Y. S., Xie, Q., & Song, M. (2023). High acknowledgement index: Characterizing research supporters with factors of acknowledgement affecting paper citation counts. *Journal of Informetrics, 17*(4), Article 101447. https://doi.org/10.1016/j.joi.2023.101447

Hérubel, J.-P. V. M. (1999). Historical bibliometrics: Its purpose and significance to the history of disciplines. *Libraries & Culture, 34*(4), 380-388.

Hinrichs, S., Montague, E., & Grant, J. (2015). *Researchfish: A forward look: Challenges and opportunities for using Researchfish to support research assessment*. Policy Institute at King's College London.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569-16572. https://doi.org/10.1073/pnas.0507655102

Hoffmann, C. P., Lutz, C., & Meckel, M. (2016). A relational altmetric? Network centrality on ResearchGate as an indicator of scientific impact. *Journal of the Association for Information Science & Technology, 67*(4), 765-775. https://doi.org/10.1002/asi.23423

Hou, J., Zheng, B., Wang, D., Zhang, Y., & Chen, C. (2023). How boundary-spanning paper sparkles citation: From citation count to citation network. *Journal of Informetrics, 17*(3), Article 101434. https://doi.org/10.1016/j.joi.2023.101434

Howard, H. C., Mascalzoni, D., Mabile, L., Houeland, G., Rial-Sebbag, E., & Cambon-Thomsen, A. (2018). How to responsibly acknowledge research work in the era of big data and biobanks: Ethical aspects of the Bioresource Research Impact Factor (BRIF). *Journal of Community Genetics, 9*(2), 169-176. https://doi.org/10.1007/s12687-017-0332-6

Hua, V., & Huynh, B. (2024). Forecasting a Journal Impact Factor under missing values based on machine learning. *IEEE Access, 12*, 85745-85760. https://doi.org/10.1109/ACCESS.2024.3416345

Huang, W., Song, F., Zhang, S., & Xia, T. (2022). Influence of deep learning-based journal reading guidance system on students' national cognition and cultural acceptance. *Frontiers in Psychology, 13*, Article 950412. https://doi.org/10.3389/fpsyg.2022.950412

Hughes, A., & Kitson, M. (2012). Pathways to impact and the strategic role of universities: New evidence on the breadth and depth of university knowledge exchange in the UK and the factors constraining its development. *Cambridge Journal of Economics, 36*(3), 723-750. https://doi.org/10.1093/cje/bes017

Hulme, E. W. (1923). *Statistical bibliography in relation to the growth of modern civilization: Two lectures delivered in the University of Cambridge in May, 1922*. Butler & Tanner Grafton.

Jacob, R., & Mcgregor, M. (1997). Assessing the impact of health technology assessment. *International Journal of Technology Assessment in Health Care, 13*(1), 68-80. https://doi.org/10.1017/S0266462300010242

Jamwal, V., & Kumar, H. (2022). An overview of dimensions and dimensions badge. *Library Hi Tech News, 39*(6), 8-13. https://doi.org/10.1108/LHTN-01-2022-0010

Ji, T., Self, N., Fu, K., Chen, Z., Ramakrishnan, N., & Lu, C.-T. (2024). Citation forecasting with multi-context attention-aided dependency modeling. *ACM Transactions on Knowledge Discovery from Data, 18*(6), Article 144. https://doi.org/10.1145/3649140

Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The R- and AR-indices: Complementing the h-index. *Chinese Science Bulletin, 52*(6), 855-863. https://doi.org/10.1007/s11434-007-0145-9

Joshi, K. P., Joshi, A., Yesha, Y., & Krishnapuram, R. (1999). Warehousing and mining web logs. In C. Shahabi (Ed.), *WIDM '99: Proceedings of the 2nd international workshop on web information and data management* (pp. 63-68). Association for Computing Machinery. https://doi.org/10.1145/319759.319792

Kakushadze, Z. (2016). An index for SSRN downloads. *Journal of Informetrics,* *10*(1), 9-28. https://doi.org/10.1016/j.joi.2015.11.005

Kanakaris, N., Giarelis, N., Siachos, I., & Karacapilidis, N. (2021). Shall I work with them? A knowledge graph-based approach for predicting future research collaborations. *Entropy, 23*(6), Article 664. https://doi.org/10.3390/e23060664

Ke, Z. T., Ji, P., Jin, J., & Li, W. (2024). Recent advances in text analysis. *Annual Review of Statistics & Its Application, 11*, 347-372. https://doi.org/10.1146/annurev-statistics-040522-022138

Kolbenstvedt, M., Elvik, R., Elvebakk, B., Hervik, A., & Braein, L. (2007). *Effects of Swedish traffic safety research 1971 – 2004: Main report*. VINNOVA. https://www.vinnova.se/contentassets/508b2679042e43888dc3cb1cb7f1289c/va-07-10.pdf

Konkiel, S. (2014, December 3). *Impact challenge day 30: Create a comprehensive impact profile at Impactstory.org*. OurResearch blog. https://blog.ourresearch.org/impact-challenge-impactstory-profile/

Kostoff, R. N. (1994a). Federal research impact assessment: State-of-the-art. *Journal of the American Society for Information Science, 45*(6), 428-440. https://doi.org/dhxnmf

Kostoff, R. N. (1994b). Research impact assessment: Principles and applications to proposed, ongoing, and completed projects. *Investigative Radiology, 29*(9), 864-869. https://doi.org/10.1097/00004424-199409000-00013

Kuruvilla, S., Mays, N., Pleasant, A., & Walt, G. (2006). Describing the impact of health research: A research impact framework. *BMC Health Services Research, 6*, Article 134. https://doi.org/10.1186/1472-6963-6-134

Landry, R., Amara, N., & Lamari, M. (2001). Climbing the ladder of research utilization: Evidence from social science research. *Science Communication, 22*(4), 396-422. https://doi.org/10.1177/1075547001022004003

Lavis, J., Ross, S., McLeod, C., & Gildiner, A. (2003). Measuring the impact of health research. *Journal of Health Services Research & Policy, 8*(3), 165-170. https://doi.org/10.1258/135581903322029520

Leal Filho, W., Frankenberger, F., Salvia, A. L., Azeiteiro, U., Alves, F., Castro, P., Will, M., Platje, J., Lovren, V. O., Brandli, L., Price, E., Doni, F., Mifsud, M., & Ávila, L. V. (2021). A framework for the implementation of the Sustainable Development Goals in university programmes. *Journal of Cleaner Production, 299*, Article 126915. https://doi.org/10.1016/j.jclepro.2021.126915

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science, 5*, Article 69. https://doi.org/10.1186/1748-5908-5-69

Li, M., Xu, J., Ge, B., Liu, J., Jiang, J., & Zhao, Q. (2019). A deep learning methodology for citation count prediction with large-scale biblio-features. In *2019 IEEE international conference on systems, man and cybernetics*

(SMC) (pp. 1172-1176). IEEE. https://doi.org/10.1109/SMC.2019.8913961

Liao, S., Lavender, C., & Zhai, H. (2024). Factors influencing the research impact in cancer research: A collaboration and knowledge network analysis. *Health Research Policy & Systems, 22*, Article 96. https://doi.org/10.1186/s12961-024-01205-8

Lindsay, J. M. (2016). PlumX from Plum Analytics: Not just altmetrics. *Journal of Electronic Resources in Medical Libraries, 13*(1), 8-17. https://doi.org/10.1080/15424065.2016.1142836

Liu, J., Xia, F., Feng, X., Ren, J., & Liu, H. (2022). Deep graph learning for anomalous citation detection. *IEEE Transactions on Neural Networks & Learning Systems, 33*(6), 2543-2557. https://doi.org/10.1109/TNNLS.2022.3145092

Liu, Y., Liu, S., Shao, X., & He, Y. (2022). Policy spillover effect and action mechanism for environmental rights trading on green innovation: Evidence from China's carbon emissions trading policy. *Renewable & Sustainable Energy Reviews, 153*, Article 111779. https://doi.org/10.1016/j.rser.2021.111779

Ma, A., Liu, Y., Xu, X., & Dong, T. (2021). A deep-learning based citation count prediction model with paper metadata semantic features. *Scientometrics, 126*(8), 6803-6823. https://doi.org/10.1007/s11192-021-04033-7

Makarov, I., & Gerasimova, O. (2019). Predicting collaborations in co-authorship network. In *2019 14th international workshop on semantic and social media adaptation and personalization (SMAP)*. IEEE. https://doi.org/10.1109/SMAP.2019.8864887

Martinuzzi, A., Hametner, M., Windsperger, A., & Brunnhuber, N. (2023). A framework for assessing the climate impacts of research and innovation projects and programmes. *Sustainability, 15*(24), Article 16600. https://doi.org/10.3390/su152416600

Meagher, L., Lyall, C., & Nutley, S. (2008). Flows of knowledge, expertise and influence: A method for assessing policy and practice impacts from social science research. *Research Evaluation, 17*(3), 163-173. https://doi.org/10.3152/095820208X331720

Mester, G. (2015). New trends in scientometrics. In P. Möhringer (Ed.), *Papers of 33rd international scientific conference "Science in Practice"* (pp. 22-27). Hochschule für Angewandte Wissenschaften Würzburg-Schweinfurt. https://nbn-resolving.org/urn:nbn:de:bvb:863-opus-2692

Moed, H. F., Burger, W. J. M., Frankfort, J. G., & van Raan, A. F. J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy, 14*(3), 131-149. https://doi.org/10.1016/0048-7333(85)90012-5

Moed, H. F., & Halevi, G. (2015). Multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science & Technology, 66*(10), 1988-2002. https://doi.org/10.1002/asi.23314

Mostert, S. P., Ellenbroek, S. P. H., Meijer, I., van Ark, G., & Klasen, E. C. (2010). Societal output and use of research performed by health research groups. *Health Research Policy & Systems, 8*, Article 30. https://doi.org/10.1186/1478-4505-8-30

Nalimov, V. V., & Mul'chenko, Z. M. (1971). *Measurement of science. Study of the development of science as an information process*. Foreign Technology Division.

Ochi, M., Shiro, M., Mori, J., & Sakata, I. (2021). Which is more helpful in finding scientific papers to be top-cited in the future: Content or citations? Case analysis in the field of solar cells 2009. In F. D. Mayo, M. Marchiori, & J. Filipe (Eds.), *Proceedings of the 17th international conference on web information systems and technologies (Webist)* (Vol. 1, pp. 360-364). https://doi.org/10.5220/0010689100003058

Oosthuizen, J. C., & Fenton, J. E. (2014). Alternatives to the impact factor. *The Surgeon, 12*(5), 239-243. https://doi.org/10.1016/j.surge.2013.08.002

Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). Assessment, evaluations, and definitions of research impact: A review. *Research Evaluation, 23*(1), 21-32. https://doi.org/10.1093/reseval/rvt021

Peters, M. D. J., Marnie, C., Colquhoun, H., Garritty, C. M., Hempel, S., Horsley, T., Langlois, E. V., Lillie, E., O'Brien, K. K., Tunçalp, Ö., Wilson, M. G., Zarin, W., & Tricco, A. C. (2021). Scoping reviews: Reinforcing and advancing the methodology and application. *Systematic Reviews, 10*, Article 263. https://doi.org/10.1186/s13643-021-01821-3

Pobiedina, N., & Ichise, R. (2016). Citation count prediction as a link prediction problem. *Applied Intelligence, 44*(2), 252-268. https://doi.org/10.1007/s10489-015-0657-y

Porwal, P., & Devare, M. H. (2024). Citation count prediction using weighted latent semantic analysis (wlsa) and three-layer-deep-learning paradigm: A meta-heuristic approach. *Multimedia Tools & Applications, 83*(11), 32245-32276. https://doi.org/10.1007/s11042-023-16957-8

Pozen, R., & Kline, H. (2011). Defining success for translational research organizations. *Science Translational Medicine, 3*(94), Article 94cm20. https://doi.org/10.1126/scitranslmed.3002085

Pradhan, D. K., Chakraborty, J., Nandi, S. (2019). Applications of machine learning in analysis of citation network. In *CODS-COMAD '19: Proceedings of the ACM India joint international conference on data science and management of data* (pp. 330-333). Association for Computing Machinery. https://doi.org/10.1145/3297001.3297053

Price, D. J. d. S. (1963). *Little science, big science*. Columbia University Press.

Priem, J. [@jasonpriem]. (2010, September 29). *I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics* [Post]. X. https://twitter.com/jasonpriem/status/25844968813

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2011). *Altmetrics: A manifesto*. https://doi.org/10.5281/zenodo.12684249

Purkayastha, A., Palmaro, E., Falk-Krzesinski, H. J., & Baas, J. (2019). Comparison of two article-level, field-independent citation metrics: Field-Weighted Citation Impact (FWCI) and Relative Citation Ratio (RCR). *Journal of Informetrics, 13*(2), 635-642. https://doi.org/10.1016/j.joi.2019.03.012

Research Excellence Framework. (2014). *REF 2014*. Higher Education Funding Council for England. https://2014.ref.ac.uk/

Rittman, M. (2020, October 6). *Event data*. Crossref. https://www.crossref.org/documentation/event-data/

Rizzo, F., Hamilton, B. B., & Keagy, R. D. (1975). Orthotics research evaluation framework. *Archives of Physical Medicine & Rehabilitation, 56*(7), 304-308.

Roemer, R. C., & Borchardt, R. (2015). Introduction to altmetrics. In *Altmetrics* (Library Technology Reports, Vol. 51, No. 5, pp. 5-10). ALA TechSource. https://doi.org/10.5860/ltr.51n5

Roopashree, N., & Umadevi, V. (2014). Future collaboration prediction in co-authorship network. In *2014 3rd international conference on eco-friendly computing and communication systems* (pp. 183-188). IEEE. https://doi.org/10.1109/Eco-friendly.2014.45

Royal Netherlands Academy of Arts and Sciences, Association of Universities in the Netherlands, & Netherlands Organisation for Scientific Research. (2009). *Standard evaluation protocol 2009-2015: Protocol for research assessment in the Netherlands*. https://www.uu.nl/sites/default/files/sep2009-2015.pdf

Rycroft-Malone, J., Burton, C., Wilkinson, J., Harvey, G., McCormack, B., Baker, R., Dopson, S., Graham, I., Staniszewska, S., Thompson, C., Ariss, S., Melville-Richards, L., & Williams, L. (2015). Collective action for knowledge moblisation: A realist evaluation of the collaborations for leadership in applied health research and care. *Health Services & Delivery Research, 3*(44). https://doi.org/10.3310/hsdr03440

Sadaf, F., Shahid, M. H., & Islam, M. A. (2021). Predicting most influential paper award using citation count. In *2021 international conference on digital futures and transformative technologies (ICoDT2)*. IEEE. https://doi.org/10.1109/ICoDT252288.2021.9441487

Sarli, C. C., Dubinsky, E. K., & Holmes, K. L. (2010). Beyond citation analysis: A model for assessment of research impact. *Journal of the Medical Library Association, 98*(1), 17-23. https://doi.org/10.3163/1536-5050.98.1.008

Schapper, C. C., Dwyer, T., Tregear, G. W., Aitken, M., & Clay, M. A. (2012). Research performance evaluation: The experience of an independent medical research institute. *Australian Health Review, 36*(2), 218-223. https://doi.org/10.1071/AH11057

Scott, C. S., Nagasawa, P. R., Abernethy, N. F., Ramsey, B. W., Martin, P. J., Hacker,

B. M., Schwartz, H. D., Brock, D. M., Robins, L. S., Wolf, F. M., Carter-Dubois, M., & Disis, M. L. (2014). Expanding assessments of translational research programs: Supplementing metrics with value judgments. *Evaluation & the Health Professions, 37*(1), 83-97. https://doi.org/10.1177/0163278713500984

Searles, A., Doran, C., Attia, J., Knight, D., Wiggers, J., Deeming, S., Mattes, J., Webb, B., Hannan, S., Ling, R., Edmunds, K., Reeves, P., & Nilsson, M. (2016). An approach to measuring and encouraging research translation and research impact. *Health Research Policy & Systems, 14*, Article 60. https://doi.org/10.1186/s12961-016-0131-2

Shi, C., Wang, H., Chen, B., Liu, Y., & Zhou, Z. (2019). Visual analysis of citation context-based article influence ranking. *IEEE Access, 7*, 113853-113866. https://doi.org/10.1109/ACCESS.2019.2932051

Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics, 72*(2), 253-280. https://doi.org/10.1007/s11192-007-1722-z

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*(4), 265-269. https://doi.org/10.1002/asi.4630240406

Smith, L. C. (1981). Citation analysis. *Library Trends, 30*(1), 83-106.

Song, X., Zhang, Y., Pan, R., & Wang, H. (2022). Link prediction for statistical collaboration networks incorporating institutes and research interests. *IEEE Access, 10*, 104954-104965. https://doi.org/10.1109/ACCESS.2022.3210129

Sørensen, O. H., Bjørner, J., Holtermann, A., Dyreborg, J., Sørli, J. B., Kristiansen, J., & Nielsen, S. B. (2022). Measuring societal impact of research—Developing and validating an impact instrument for occupational health and safety. *Research Evaluation, 31*(1), 118-131. https://doi.org/10.1093/reseval/rvab036

Spaapen, J., van Drooge, L., Propp, T., van den Meulen, B., Shinn, T., Marcovich, A., van den Besselaar, P., de Jong, S., Barker, K., Cox, D., Morrison, K., Sveinsdottir, T., Pearson, D., D'Ippolito, B., Prins, A., Molas-Gallart, J., Tang, P., & Castro-Martínez, E. (2011). *SIAMPI final report: Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society*. http://www.siampi.eu/Content/SIAMPI_Final%20report.pdf

Srinivasa, G. (2019). Relevance of innovations in machine learning to scientometrics. *Journal of Scientometric Research, 8*(2s), s39-s43. https://doi.org/10.5530/jscires.8.2.23

Stalder, F., & Hirsh, J. (2002). Open source intelligence. *First Monday, 7*(6). https://doi.org/10.5210/fm.v7i6.961

Stančiauskas, V., Brozaitis, H., Notten, A., Hollanders, H., Papageorgiou, H., Manola, N., Zagame, P., & Le-Mouel, P. (Eds.). (2022). *Study to support the monitoring and evaluation of the Framework Programme for research and innovation along Key Impact Pathways: Indicator methodology and metadata handbook*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/44653

Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science & Technology, 68*(9), 2037-2062. https://doi.org/10.1002/asi.23833

Teixeira da Silva, J. A., & Memon, A. R. (2017). CiteScore: A cite for sore eyes, or a valuable, transparent metric? *Scientometrics, 111*(1), 553-556. https://doi.org/10.1007/s11192-017-2250-0

Thelwall, M. (2021). Measuring societal impacts of research with altmetrics? Common problems and mistakes. *Journal of Economic Surveys, 35*(5), 1302-1314. https://doi.org/10.1111/joes.12381

Times Higher Education. (2018). *World University Rankings 2019: Methodology Times Higher Education (THE)*. https://www.timeshighereducation.com/world-university-rankings/world-university-rankings-2019-methodology

Torres-Salinas, D., Moreno-Torres, J. G., Delgado-López-Cózar, E., & Herrera, F. (2011). A methodology for Institution-Field ranking based on a bidimensional analysis: The IFQ ^{2} A index. *Scientometrics, 88*(3), 771-786. https://doi.org/10.1007/s11192-011-0418-6

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467-473. https://doi.org/10.7326/M18-0850

University of Bradford. (2020). *The pros and cons of key metrics*. https://www.bradford.ac.uk/library/research-support/bibliometrics-and-altmetrics/The_pros_and_cons_of_key_metrics.pdf

Vanclay, J. K., & Bornmann, L. (2012). Metrics to evaluate research performance in academic institutions: A critique of ERA 2010 as applied in forestry and the indirect H_2 index as a possible alternative. *Scientometrics, 91*(3), 751-771. https://doi.org/10.1007/s11192-012-0618-8

Vinayak, Raghuvanshi, A., & Kshitij, A. (2023). Signatures of capacity development through research collaborations in artificial intelligence and machine learning. *Journal of Informetrics, 17*(1), Article 101358. https://doi.org/10.1016/j.joi.2022.101358

Vinkler, P. (2009). The π-index: A new indicator for assessing scientific impact. *Journal of Information Science, 35*(5), 602-612. https://doi.org/10.1177/0165551509103601

Vinkler, P. (2014). The use of the Percentage Rank Position index for comparative evaluation of journals. *Journal of Informetrics,* *8*(2), 340-348. https://doi.org/10.1016/j.joi.2014.01.001

Vital, A., & Amancio, D. R. (2022). A comparative analysis of local similarity metrics and machine learning approaches: Application to link prediction in author citation networks. *Scientometrics, 127*(10), 6011-6028. https://doi.org/10.1007/s11192-022-04484-6

Wang, L., Wen, H., & Liu, Y. (2016). AHP based quantitative evaluation index system of teacher's research performance in the university. *International Journal of Multimedia & Ubiquitous Engineering, 11*(7), 391-402. https://doi.org/10.14257/ijmue.2016.11.7.39

Ward, R. L., Nutbeam, D., Mijnhardt, W., Nelson, P., Todd, A., Rees, M. I., Richards, J., Khan, N. N., Ho, I., & Chung, S. (2023). Development of a novel and more holistic approach for assessing impact in health and medical research: The Research Impact Assessment Framework. *Australian Health Review, 47*(5), 589-595. https://doi.org/10.1071/AH23152

Weiss, A. P. (2007). Measuring the impact of medical research: Moving from outputs to outcomes. *American Journal of Psychiatry, 164*(2), 206-214. https://doi.org/10.1176/ajp.2007.164.2.206

Wellcome Trust. (2009). *How we are making a difference: Assessment framework report summary 2008/09*. https://cms.wellcome.org/sites/default/files/wtx059577_0.pdf

Wiegers, S. E., Houser, S. R., Pearson, H. E., Untalan, A., Cheung, J. Y., Fisher, S. G., Kaiser, L. R., & Feldman, A. M. (2015). A metric-based system for evaluating the productivity of preclinical faculty at an academic medical center in the era of clinical and translational science. *Clinical & Translational Science, 8*(4), 357-361. https://doi.org/10.1111/cts.12269

Williams, V. L., Eiseman, E., Landree, E., & Adamson, D. M. (2009). *Demonstrating and communicating research impact: Preparing NIOSH programs for external review*. RAND.

Wu, M., Zhang, Y., & Li, X. (2022). Exploring associations within disease-gene pairs: Bibliometrics, word embedding, and network analytics. In D. F. Kocaoglu, T. R. Anderson, D. C. Kozanoglu, K. Niwa, & H. J. Steenhuis (Eds.), *PICMET'22: Portland international conference on management of engineering and technology (PICMET) proceedings: Technology management and leadership in digital transformation looking ahead to post-COVID era*. https://doi.org/10.23919/PICMET53225.2022.9882884

Wu, Z., Lin, W., Liu, P., Chen, J., & Mao, L. (2019). Predicting long-term scientific impact based on multi-field feature extraction. *IEEE Access, 7*, 51759-51770. https://doi.org/10.1109/ACCESS.2019.2910239

Wyatt, S., Milojević, S., Park, H. W., & Leydesdorff, L. (2017). The intellectual and practical contributions of scientometrics to STS. In U. Felt, R. Fouché, C. A. Miller, & L. Smith-Doerr (Eds.), *The handbook of science and technology studies* (pp. 87-112). MIT Press.

Xiao, C., Han, J., Fan, W., Wang, S., Huang, R., & Zhang, Y. (2019). Predicting scientific impact via heterogeneous academic network embedding. In A. C. Nayak & A. Sharma (Eds.), *PRICAI 2019: Trends in artificial intelligence* (pp. 555-568). Springer. https://doi.org/10.1007/978-3-030-29911-8_43

Yan, K.-K., & Gerstein, M. (2011). The spread of scientific information: Insights from the web usage statistics in PLoS article-level metrics. *PLoS ONE, 6*(5), Article e19917. https://doi.org/10.1371/journal.pone.0019917

Yang, C., Wang, C., Zheng, R., & Geng, S. (2023). Link prediction in research collaboration: A multi-network representation learning framework with joint training. *Multimedia Tools & Application, 82*(30), 47215-47233. https://doi.org/10.1007/s11042-023-15720-3

Yaniasih, Y., & Budi, I. (2021). Systematic design and evaluation of a citation function classification scheme in Indonesian journals. *Publications, 9*(3), Article 27. https://doi.org/10.3390/publications9030027

Yazdizadeh, B., Ahmadi, A., Najafi, F., Mohammad, K., Fariden, M., Khalili, D., Mahdavi, M., Rahimpour, E., Jouyban, A., Kelishadi, R., Monazzam, M. R., Eftekhari, M. B., Falahat, K., Nikooee, S., & Majdzadeh, R. (2024). Establishing research impact assessment in Iran: The first report from a non-high-income country. *Journal of Global Health, 14*, Article 04050. https://doi.org/10.7189/JOGH.14.04050

Zahedi, Z., & Costas, R. (2020). Do online readerships offer useful assessment tools? Discussion around the practical applications of Mendeley readership for scholarly assessment. *Scholarly Assessment Reports, 2*(1), Article 14. https://doi.org/10.29024/sar.20

Zerva, C., Nghiem, M.-Q., Nguyen, N. T. H., & Ananiadou, S. (2020). Cited text span identification for scientific summarisation using pre-trained encoders. *Scientometrics, 125*(3), 3109-3137. https://doi.org/10.1007/s11192-020-03455-z

Zhai, L., Yan, X., & Zhu, B. (2013). The h_1 -index: Improvement of *h*-index based on quality of citing papers. *Scientometrics, 98*(2), 1021-1031. https://doi.org/10.1007/s11192-013-1039-z

Zhang, Q., Mao, R., & Li, R. (2019). Spatial-temporal restricted supervised learning for collaboration recommendation. *Scientometrics, 119*(3), 1497-1517. https://doi.org/10.1007/s11192-019-03100-4

Zhang, X., Xie, Q., & Song, M. (2021). Measuring the impact of novelty, bibliometric, and academic-network factors on citation count using a neural network. *Journal of Informetrics, 15*(2), Article 101140. https://doi.org/10.1016/j.joi.2021.101140

Zhou, H., Sun, J., Zhao, Z., Yang, Y., Xie, A., & Chiclana, F. (2019). Attention-based deep learning model for predicting collaborations between different research affiliations. *IEEE Access, 7*, 118068-118076. https://doi.org/10.1109/ACCESS.2019.2936745

Zhou, Y., Wang, R., & Zeng, A. (2022). Predicting the impact and publication date of individual scientists' future papers. *Scientometrics,* *127*(4), 1867-1882. https://doi.org/10.1007/s11192-022-04286-w

Zhu, X., & Ban, Z. (2018). Citation count prediction based on academic network features. In L. Barolli, M. Takizawa, T. Enokida, M. R. Ogleia, L. Ogleia, & N. Javavid (Eds.), *IEEE AINA 2018: 32nd IEEE international conference on advanced information networking and applications*. https://doi.org/10.1109/AINA.2018.00084.

(Received: 2024/10/13; Accepted: 2025/2/21)

Appendix: Generational Advancements in Research Impact Metrics, Frameworks, and Models

Table 1. Key Metrics of Research Impact Assessment in the First Generation

Index/Model	Year	Description	Calculation method	Reference
Early bibliometric indices and productivity metrics (Pre-1990s)				
Statistical Bibliography	1920s	Early attempts at quantitatively analysing published literature to understand the growth and trends in scientific disciplines.	Quantitative analysis of publication counts over time.	Hulme (1923)
Academic Productivity Metrics	1960s–1980s	Focused on evaluating academic productivity through the number of publications and citations received.	Counting total publications and citations per researcher or institution.	Garfield (1955, 1979) Kostoff (1994a) Rizzo et al. (1975)
Citation Analysis	1950s–1960s	The study of the frequency, patterns, and graphs of citations in articles and books laid the foundation for later bibliometric indicators.	Analysing citation counts and networks among scholarly works.	Garfield (1955)
Impact Factor (IF)	1960s	Measures the frequency with which the average article in a journal has been cited in a particular year.	Total citations in a year to articles published in the previous two years are divided by the total number of articles published in those two years.	Garfield (1979) Garfield & Sher (1963)
Development of author-level indices (2005–2016)				
h-Index	2005	Measures both the productivity and citation impact of a researcher's publications.	A researcher has an h-index of h if h of their papers have at least h citations each.	Hirsch (2005)
g-Index	2006	Gives more weight to highly cited articles, improving upon the h-index.	The highest number g such that the top g articles received collectively at least g^2 citations.	Egghe (2006)
A-Index and R-Index	2007	The A-index measures average citations per article in the h-core; the R-index combines the h-index and the A-index.	A-Index: Total citations of h-core divided by h. R-Index: Square root of the sum of h-core citations.	Jin et al. (2007)

(continued)

Table 1. Key Metrics of Research Impact Assessment in the First Generation (contiune)

Index/Model	Year	Description	Calculation method	Reference
Contemporary h-Index (hc-Index)	2007	Adjusts the h-index by giving more weight to recent publications.	Citations are weighted by a decaying function based on the publication age.	Sidiropoulos et al. (2007)
i10-Index	2011	Measures the number of publications with at least 10 citations, indicating influential works.	Counts the total number of a researcher's publications that have been cited at least 10 times.	Connor (2011)
π-Index	2009	Reflects both the productivity and impact of a researcher's most cited publications.	Based on the ratio of citations received to the total number of citations possible for top publications.	Vinkler (2009)
H_2-Index	2012	Squares the h-index to give more weight to highly productive researchers.	H_2 = h-index squared (H^2).	Vanclay & Bornmann (2012)
H_I-Index	2013	Considers co-authorship order, emphasising contributions of first or corresponding authors.	Adjusts citations based on the position of authorship in publications.	Zhai et al. (2013)
Y-Index	2014	Reflects the number of first-author and corresponding-author articles, indicating leadership in research.	Combines the number of first-author papers (F) and corresponding-author papers (C) into a vector or angle measure.	Fu & Ho (2014)
PRP-Index	2014	Provides a nuanced view of individual contributions by considering the rank percentile of publications.	Based on the percentile ranks of publications within a set, emphasising higher-ranked works.	Vinkler (2014)
AHP Index	2016	Applies the Analytic Hierarchy Process to bibliometric evaluation, integrating multiple criteria for assessment.	Combines various bibliometric indicators weighted according to their importance in a hierarchical structure.	Wang et al. (2016)
GFsa Index	2024	Considers total citations and "scientific age" to provide a fairer evaluation among researchers.	GFsa = (Total number of citations) ÷ (Scientific Age)^2, where Scientific Age is the time since a researcher's first publication.	Fernandes & Fernandes (2024)

(continued)

Table 1. Key Metrics of Research Impact Assessment in the First Generation (contiune)

Index/Model	Year	Description	Calculation method	Reference
Journal-level metrics and indices				
SCImago Journal Rank (SJR)	2008	Accounts for the number of citations received and the importance of the journals from which such citations come.	Weighted citations per document, with weights depending on the SJR of the citing journals.	Falagas et al. (2008)
Eigenfactor Score	2008	Measures the influence of a journal based on citation networks, considering the origin of incoming citations.	Like Google's PageRank algorithm, citations are weighted based on the journal's influence.	Oosthuizen & Fenton (2014)
Source Normalized Impact per Paper (SNIP)	2010	Weights citations based on the total number of citations in a subject field to measure contextual impact.	Citations are normalised by the citation potential in the subject field.	Oosthuizen & Fenton (2014)
IF^2-Index	2010	Considers the impact of journals where an author's articles are published.	Combines individual article citations with the impact factors of the journals in which they appear.	Boell & Wilson (2010)
JIF Percentile	2010	Provides a percentile rank for a journal within its subject category based on the Journal Impact Factor.	Ranks journals by impact factor percentile within their subject categories.	Boell & Wilson (2010)
IFQ^2A Index	2011	Assesses research performance by combining journal impact factor with article quality and quantity.	Integrates impact factors, article counts, and quality indicators into a composite score.	Torres-Salinas et al. (2011)
h5-Index and h5-Median	2015	Google's metrics evaluating journal impact based on articles from the last five complete years.	h5-Index: The highest number of articles received at least h citations in the last five years. h5-Median: Median number of citations for articles in the h5-core.	Mester (2015)
CiteScore	2016	Elsevier's journal metric covers a wider range of documents than traditional impact factors.	Citations received in one year to documents published in the previous three years are divided by the number of documents published in those years.	Teixeira da Silva & Memon (2017)

(continued)

Table 1. Key Metrics of Research Impact Assessment in the First Generation (contiune)

Index/Model	Year	Description	Calculation method	Reference
Field-normalized citation metrics				
Field-Weighted Citation Impact (FWCI)	2010	Compares the actual number of citations received by a publication to the expected number of citations for similar publications.	Calculate by dividing the number of citations received by the expected number for the subject field, publication type, and publication year.	Purkayastha et al. (2019)
Relative Citation Ratio (RCR)	2016	Measures the citation performance of a paper relative to other papers in its field.	Calculated by comparing a paper's citation rate to the average citation rate of NIH-funded papers in the same field.	Purkayastha et al. (2019)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Health and biomedical research impact frameworks				
CETS	Health policy, cost management	Impact on Health Policy, Impact on Health Costs	Ministry decisions on health services, Hospital rules, Cost minimisation, Optimisation of healthcare system	Jacob & Mcgregor (1997)
RE-AIM Model	Public health interventions, long-term sustainability	Reach, Efficacy, Adoption, Implementation, Maintenance	Proportion of target population, Success rate, Adoption rate, Implementation, Program sustainability	Glasgow et al. (1999)
The Matrix	Health, economy, policy influence, knowledge dissemination	Knowledge Transfer, Economic Impact, Health and Social Outcomes, Policy and Practice Impact, Capacity Building	Publications, Patents, Training programs, Policy changes, Health outcomes, Social engagement	Wiegers et al. (2015)
Canadian Institutes of Health Research (CIHR) Impact Framework	Knowledge dissemination, health outcomes, economy	Knowledge Production, Research Targeting, Policy Impact, Health Benefits, Economic Impacts	Publications, Citation impact, Public health improvements (PYLL), Patents, Commercialization, Cost savings	Bernstein et al. (2007)
Medical Research Logic Model	Healthcare improvements, patient outcomes	Initial, Intermediate, Long-Term Impacts	Awareness of scientific evidence, Change in clinical practice, Improvement in patient well-being	Weiss (2007)
Canadian Academy of Health Sciences Framework (CAHS)	Health outcomes, knowledge generation, policy	Advancing Knowledge, Capacity Building, Decision Making, Health Impacts, Health System Indicators, Economic and Social Impacts	Relative citation impact, Funding levels, Research use in healthcare, Morbidity/mortality data, QALYs, PROMs, Commercialisation, Social benefits	CAHS (2009)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
The Wellcome Trust's Assessment Framework	Knowledge generation, health outcomes, policy	Knowledge Generation, Researcher Development, Health Impact, Technology Development, Policy Development	Publications, Citations, RBI, Awards, Fellowships, Patents, Policy impact, Capacity-building, Media coverage	Wellcome Trust (2009)
The Becker Model	Biomedical research, clinical implementation	Research Output, Knowledge Transfer, Clinical Implementation, Community Benefit	Biological materials, Patents, Medical devices, Clinical guidelines, Health care outcomes, Quality of life	Sarli et al. (2010)
NIOSH Logic Models	Workplace safety, health hazard reduction	Construction Program, Mining Program, Health Hazard Evaluation Program, Personal Protective Technology Program, Overall NIOSH Program	Peer-reviewed articles, Safety guidelines, Workplace policies, Best practices, Technology adoption, Standards & Regulations	Williams et al. (2009)
Societal Quality Score (Leiden University Medical Centre)	Knowledge dissemination, stakeholder engagement, economic gains	Knowledge Production, Knowledge Exchange, Knowledge Use, Earning Capacity	Contributions to media, Patents, Speeches for companies, Use of medical protocols, Charity funding, Indirect funding, Contract funding	Mostert et al. (2010)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Institute for Translational Health Sciences (ITHS) Kellogg Logic Model WHO Health Services Assessment Model	Health service improvements, capacity building, translational research	Inputs, Activities, Outputs, Outcomes/Impacts	Relevance: The degree to which services meet translational research needs. Efficiency: How well the research processes are streamlined. Adequacy: The extent to which services are suitable for achieving positive health impacts. Effectiveness: The ability of the services to improve research outcomes. Equity: Access to resources for diverse cultural groups. Impact: The influence of research on population health. Sustainability: Long-term viability of resources and services.	Scott et al. (2014)
Payback Framework	Knowledge dissemination, health, economy	Knowledge, Future Research, Policy/Product Development, Health/Health Sector Benefits, Economic Benefits	Journal articles, Development of research skills, Political decisions, Pharmaceutical products, Cost reduction, Commercial exploitation	Donovan & Hanney (2011)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Monetary Value Approach	Economic impact, health system efficiency	Well-being Gains, Avoided Health Costs, Productivity, Commercialization of R&D	DALYs, Value of a statistical life (VSL), Avoided health costs, Productivity gains, Commercialisation benefit/Cost ratio, ROI	Deloitte Access Economics (2011)
Banzi's Research Impact Model	Knowledge, health outcomes, economic, social	Advancing Knowledge, Capacity Building, Policy/Product Development, Health/Sector Benefits, Economic Benefits	Peer-reviewed publications, Patents, Epidemiologic data, QALYs, Social benefits, Product sales, Spin-off companies	Banzi et al. (2011)
Research Performance Evaluation Framework	Knowledge creation, commercialisation, public health	Knowledge Creation, Research Inputs, Commercial, Clinical & Public Health Outcomes	Peer-reviewed publications, Funding, Research students, Patents, Commercialisation, Adoption and implementation of research findings	Schapper et al. (2012)
Translational Research Organizations Performance Model	Research funding, validation, collaboration	Funding KPIs, Talent KPIs, Creation KPIs, Validation KPIs, Dissemination KPIs, Uptake KPIs, Collaboration KPIs	Grant dollars secured, Peer-reviewed publications, Citations, Licensing agreements, Patents, Spinoffs, Partnerships, Co-authorship	Pozen & Kline (2011)
Health Services Research Impact Framework	Knowledge, policy, services, societal impact	Advancing Knowledge, Policy Impact, Service Impact, Societal Impact	Peer-reviewed publications, Clinical guidelines, Capacity building, Policy briefs, Validated research adoption, Health outcomes	Buykx et al. (2012)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Translational Research Impact Scale	Research quality, clinical practice, societal impact	Research Impacts, Translational Impacts, Societal Impacts	Research networks, IRB processes, Grant submissions, Patents, Clinical guidelines, Health care improvement, Job growth, Policy changes	Dembe et al. (2014)
Hunter Medical Research Institute Framework	Health outcomes, policy influence, economic impact	Advance Knowledge, Clinical Implementation, Community Benefit, Legislation, Economic Impact, SROI, Case Studies	PhD completions, Clinical guidelines, Quality of life (QoL) improvement, Policy citations, Cost avoided, SROI ratio	Searles et al. (2016)
Research Impact Assessment Framework	Health outcomes, Policy and practice change, Economic benefits, Improved health systems, Enhanced research capability	Research Quality, Research Relevance, Research Engagement, Research Translation, Research Sustainability	Publications and citations, Research funding, Collaboration with stakeholders, Translation of research into practice, Impact on health outcomes, Policy influence, Economic return on investment	Ward et al. (2023)
Higher education and academic research				
Research Assessment Exercise (RAE)	Academic excellence, policy influence, societal benefit	Research Quality, Research Output, Research Environment, Research Impact	Publications, Research income, Collaborations, Impact on policy, Societal benefits	Boaden & Cilliers (2001)
VINNOVA (Swedish Governmental Agency for Innovation Systems)	Academic impact, public safety, economic impacts	Academic Results, Effects for Users, Diffusion of Research, Economic Impacts	Publications, PhDs, Patents, Traffic safety measures, Public finance, Public debate, Workforce migration	Kolbenstvedt et al. (2007)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Matrix Scoring System	Academic productivity, research performance	Research, Education, Authorship, Administration/ Service	Direct & indirect costs, Peer-reviewed publications, Patents, Contact hours, Grant submissions, Committee service	Wiegers et al. (2015)
Excellence in Research for Australia (ERA)	Academic output, creative works, funding	Traditional Research Outputs, Non-Traditional Research Outputs, Research Income	Books, Journal articles, Creative works, Curated exhibitions, Competitive grants, Industry research income	ARC (2018)
Research Excellence Framework (REF)	Health outcomes, economic growth, policy impact	Quality of Outputs, Health and Welfare, Society and Culture, Economy, Commerce, Public Policy, Production, Environment, Practitioners	Bibliometrics, Impact statement, Case studies	REF (2014)
Engagement and Impact Framework	Academic engagement, socioeconomic impact	Research-related, Socio-economic, Healthcare	Mobility of researchers, Patents, Licencing agreements, Co-authorship, Spin-outs, Public lectures and seminars	ARC (2017)
World University Ranking Model by Times Higher Education	Academic reputation, research influence	Teaching, Research, Citations, International Outlook, Industry Income	Teaching environment, Research volume, Citation impact, International collaboration, Knowledge transfer	THE (2018)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
SDG Implementation Framework for Universities	Sustainable development in higher education.	Institutional Commitment, Work Plan, Budget Agreement, Progress Mapping, Operations Integration, Cross-Disciplinary Integration, Research Consideration, Staff and Student Involvement, External Stakeholder Communication, Monitoring and Reporting	Institutional policies, Strategic plans, Budget allocations, Progress reports, Operational practices, Interdisciplinary teaching, Research projects, Stakeholder engagement, Communication strategies, Monitoring outcomes	Leal Filho et al. (2021)
Multidisciplinary research				
Research Utilization Ladder	Research utilisation, knowledge transfer	Technology, Economic, Institutional, Social Interaction	Transmission, Cognition, Reference, Effort, Influence, Application	Landry et al. (2001)
Royal Netherlands Academy of Arts and Sciences (KNAW) -Standard Evaluation Protocol (SEP)	Academic reputation, societal relevance	Research Output, Earning Capacity, Academic Reputation, Societal Relevance, Viability	Number of publications, PhDs, Project funding, Prizes, Socio-cultural/ Economic quality, Resource management, Innovative capacity	KNAW et al. (2009)
SIAMPI	Stakeholder engagement, societal, economic	Knowledge Dissemination, Stakeholder Interest, Impact and Use of Results	PhDs in industry, Joint roadmaps, Patents, Market launch of new products, Spin-offs, Staff exchanges	Spaapen et al. (2011)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Researchfish	Regular monitoring of research outputs and impacts	Publications, Collaboration, Further Funding, Engagement, Policy Influence, Clinical Trials, Awards	Publications, Collaboration data, Further funding, Policy influence, Clinical trials data, IP data, Awards and recognition	Hinrichs et al. (2015)
Multidimensional Assessment Model	Knowledge growth, societal and economic impacts	Knowledge Growth, Research Networks, Publication Outlets, Social Impact, Technological Impact, Economic Impact, Cultural Impact	Publications and citations, Co-authorship, Patents, Revenues from commercialisation, Employability, Social media mentions	Moed & Halevi (2015)
Research Quality Plus Plus (RQ++)	Research quality and impact	Scientific Rigour, Research Legitimacy, Research Importance, Positioning for Use, Research and Innovation Results	Methodological integrity, Partnership mutuality, Originality, Knowledge accessibility, High-quality research outputs	Carden et al. (2023)
Realist Evaluation	Knowledge mobilisation, collaboration, cultural change	Direct, Processual & Conceptual Impacts, Personal Impacts, Collaboration for Leadership Impact	Tool/project impacts, Publications, Grant capture, Cultural change	Rycroft-Malone et al. (2015)
Commonwealth Scientific Industrial Research Organisation (CSIRO) Impact Evaluation Framework	Knowledge dissemination, societal well-being, environmental sustainability, policy influence, economic growth	Knowledge Creation, Societal Well-being, Economic Impact, Environmental Impact, Policy Impact, Capacity Building	Publications, Patents, Collaborations, Policy changes, Environmental improvements, Economic returns	CSIRO (2024)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Horizon Europe Key Impact Pathways	Innovation, science-policy interface, social and environmental impact, economic competitiveness	Creating High-Quality New Knowledge, Strengthening Human Capital, Fostering Diffusion of Knowledge, Scientific Impact, Addressing EU Policy Priorities, Delivering Benefits through Research Missions, Strengthening Research Uptake in Society, Generating Innovation-Based Growth, Creating Better Jobs	Number of publications and citations, Collaborations with industry and academia, Skills development and training metrics, Societal engagement, Economic indicators, Policy alignment, Innovation metrics	Stančiauskas et al. (2022)
Policy and social science research				
Decision-Making Impact Model	Knowledge transfer, policy impact	Producer-push Process, User-pull Process, Exchange Process	Publications, Policy briefs, Decision-makers awareness, Commissioned research projects, Involvement in decision-making processes	Lavis et al. (2003)
Research Impact Framework (RIF)	Knowledge, policy, services, societal impact	Research-related, Policy, Service, Societal Impacts	Publications, Patents, Leadership awards, Policy networks, Cost-effectiveness, Health literacy, Empowerment, Sustainable development	Kuruvilla et al. (2006)

(continued)

Table 2. Second Generation's Research Impact Assessment Frameworks – Focus on Impact, Key Aspects and Prominent Indicators (contiune)

Framework name	Focus of impact	Key aspects	Prominent indicators	Reference
Flows of Knowledge, Expertise and Influence	Knowledge transfer, social impact on policy and culture	Policy and Practices, Culture and Attitudes, Influences on Processes Leading toward Impacts	Policy formation, Conceptual impacts, Professional practice changes, Knowledge dissemination, Networks, Engagement processes	Meagher et al. (2008)
Environmental Health and Safety Research				
NIEHS Logic Model	Environmental health, societal impacts	Awareness, Policy Assessment, Knowledge Accumulation, Environmental/ Health Impact, Societal Change	Environmental hazard awareness, Policy changes, Reduction in emissions, New grant programs, Public behaviour change	Engel-Cox et al. (2008)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
Aspects of research impact: citation context, trends, prediction and forecasting						
1	Long-term Citation Prediction Models	Objective: • Propose robust models for forecasting citation counts over long durations using deep learning techniques. Motivation: • Long-term citation prediction is crucial for understanding academic impact. • Existing methods lack generalisation and accuracy.	Methodology: • Developed models like DMA-Nets, temporal network frameworks, RNNs, and CNNs. • Techniques included attention mechanisms, sequence-to-sequence models, and incorporation of temporal features. Datasets: • USPTO patents, Microsoft Academic Graph, arXiv High-Energy Physics, Web of Science, AMiner.	Improved prediction accuracy over state-of-the-art methods. Demonstrated effectiveness in capturing long-term citation trends. Highlighted the importance of temporal and network features.	Limitations: • Computationally intensive models. • Generalization to other domains untested in some studies. Future Work: • Reduce computational complexity. • Extend models to other domains. • Integrate additional features like author metrics.	Abbas et al. (2023) Abrishami & Aliakbary (2019) Glänzel & Schubert (1995) Ji et al. (2024) Li et al. (2019) Zhu & Ban (2018)
2	Impact of PACS Codes on Citations	Objective: • Investigate the impact of PACS codes on research article citations. • Develop a universal approximation curve to predict citations based on keywords.	Methodology: • Analysed citation impact at three hierarchical PACS levels. • Applied decision trees and statistical methods.	"Condensed Matter" (PACS 60) received the highest citations. Developed a universal approximation curve for third-level PACS codes.	Limitations: • Limited to APS journals and PACS codes. • Prediction accuracy capped at 88%.	Enduri et al. (2022)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
		Motivation: • Keywords influence citations, but exact correlations are underexplored. • Provides actionable insights for authors and editors.	Dataset: • 399,713 articles from APS Physical Review Journals (1985–2012).	Strong correlations were found at second-level PACS codes.	Future Work: • Extend methodology to higher-level PACS codes and other disciplines. • Integrate author networks and article semantics.	
3	Semantic and Content-Based Citation Prediction	Objective: • Develop citation prediction models leveraging semantic features and abstracts. Motivation: • Existing methods often neglect semantic metadata. • Abstracts are rich in information but underutilised.	Methodology: • Combined Doc2Vec for metadata encoding with Bi-LSTM and attention mechanisms. • Used SVM classifiers on abstracts with technical/non-technical terms. Datasets: • 83,331 Computer Science papers (Research.com). • 9,117 AI journal papers (Scopus). • PNAS abstracts (49,171 and 52,425 papers).	Semantic features significantly enhanced prediction accuracy. Abstracts alone were effective for citation prediction. Technical terms in abstracts had a significant impact.	Limitations: • Limited to specific domains (AI, life sciences). • Dependency on specific encoding methods. Future Work: • Extend to other disciplines. • Incorporate author metrics and altmetrics. • Apply advanced models like BERT.	Baba & Baba (2018) Baba et al. (2019) Ma et al. (2021) Porwal & Devare (2024)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
4	Citation Burst and Temporal Dynamics	Objective: • Predict citation bursts and study temporal dynamics in citation patterns. • Enhance models to handle long-tail citation distributions. Motivation: • Understanding citation bursts aids in assessing research impact. • Traditional models fail to predict long-tail distributions effectively.	Methodology: • Analysed features like author productivity and h-index using regression models. • Applied k-means clustering on citation time-series data. • Modified LSTM networks with power-law adjustments. Datasets: • AMiner (617,740 articles). • Open Academic Graph. • 11,209 articles from information science journals.	Identified key factors influencing citation counts. Classified citation profiles into distinct clusters. Improved prediction of highly cited articles.	Limitations: • Limited to specific domains and time periods. • Challenges with data sparsity and computational demands. Future Work: • Include dynamic modelling of citation profiles. • Apply frameworks to other datasets. • Incorporate author and semantic features.	Amjad et al. (2022) Du et al. (2024) Pradhan et al. (2019)
5	Advanced Deep Learning Models in Scientometrics	Objective: • Propose innovative deep learning methods for scientometric tasks, including impact prediction and representation learning.	Methodology: • Developed sequence-to-sequence RNNs for individual impact prediction. • Introduced GERscore for citation network analysis. • Proposed Hierarchical Extreme Learning Machines (H-DELM).	Improved prediction accuracy and representation learning capabilities. Demonstrated the potential of models like H-DELM.	Limitations: • Challenges with model interpretability and computational demands. • Generalization across domains requires testing.	Li et al. (2019) Pobiedina & Ichise (2016) Srinivasa (2019) Zhou et al. (2022)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
		Motivation: • Deep learning offers new opportunities for complex scientometric tasks. • Traditional models may not handle high-dimensional data effectively.	• Surveyed deep learning applications in scientometrics. Datasets: • APS dataset (678,916 papers). • Hep-Th and ArnetMiner datasets. • Benchmark datasets (MNIST, UCI).	Highlighted benefits of deep learning in high-dimensional data analysis.	Future Work: • Explore transfer learning in scientometrics. • Enhance the interpretability of deep learning models. • Validate across diverse datasets.	
6	Topic-SCORE Analysis based on Citations	Objective: • Propose and apply Topic-SCORE, a statistically robust scientific research method for analysing text data. Motivation: • Existing models like neural networks are resource-intensive. • Need accessible methods for social scientists with smaller datasets.	Methodology: • Combined Singular Value Decomposition (SVD) and anchor word analysis for text clustering. • Extended analysis with TR-SCORE for ranking topics based on citations. Dataset: • MADStat: 83,331 papers across 36 statistics journals (1975–2015).	Identified 11 representative topics in statistics. Developed cross-topic citation graphs. Effectively ranked journals, highlighting influential ones.	Limitations: • Limited to statistics; may not generalise to other fields. • Assumes anchor words, which may not be valid for all datasets. Future Work: • Apply Topic-SCORE to other scientific fields. • Improve robustness of anchor-word condition. • Explore applications in diverse languages.	Ke et al. (2024)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
7	Citation Context Analysis	Objective: • Perform sentiment analysis and classification of citation contexts. • Analyze research dissemination patterns and evaluate article impact. Motivation: • Citation metrics often lack qualitative aspects. • Understanding citation context enhances the evaluation of scientific works.	Methodology: • Utilized CNN features and voting classifiers with SMOTE for sentiment analysis. • Designed CNN-based classification models with FastText embeddings. • Developed citation context-based influence ranking models (S-SPEAR). Datasets: • ACL Anthology Network citation corpus. • ACL Anthology Reference Corpus (ARC). • Web of Science dataset (IEEE TVCG articles).	Achieved high accuracy and F1 scores in classification tasks. Demonstrated improvement in bibliometric measures by incorporating sentiment analysis. Visual analysis tools allowed intuitive exploration of citation relationships.	Limitations: • Potential overfitting with SMOTE-generated data. • Limited to specific datasets; generalisation may be constrained. Future Work: • Combine handcrafted and word-embedding features. • Improve scalability for large citation networks. • Incorporate advanced semantic features.	Aljohani et al. (2021) Alnowaiser (2024) Shi et al. (2019)
8	Feature-Based Prediction Models	Objective: • Develop models leveraging unique features like acknowledgment indices and user ratings.	Methodology: • Employed NER for acknowledgment entity classification. • Used Gradient Boosting Regressor for predicting influential papers. • Developed UT-CDAE incorporating user rating trends.	Acknowledgment index correlated with citation impacts. High recall in predicting influential papers.	Limitations: • Biomedical focus may limit generalizability. • Results are dataset-specific.	Heo et al. (2023) Ochi et al. (2021) Sadaf et al. (2021)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
		Motivation: • Acknowledgments and user ratings provide insights into research impact and recommendation systems. • Traditional models often overlook these features.	Datasets: • PubMed Central (2,265,937 articles). • 19,651 papers from five conferences with citations. • MovieLens datasets (ML-100K and ML-1M).	UT-CDAE outperformed CDAE in recommendation accuracy.	Future Work: • Extend analysis to other fields. • Combine citation trends with content features. • Expand UT-CDAE to diverse datasets.	
9	Future Top-Cited Paper Identification	Objective: • Compare content-based and citation-based representations to identify future top-cited papers. Motivation: • Early identification of impactful papers is crucial. • Effectiveness of using content vs. citation data needs exploration.	Methodology: • Implemented Sentence-BERT for content embeddings. • Used SEAL (Graph Neural Network) for citation embeddings. • Applied clustering and entropy analysis. Dataset: • 57,935 papers on solar cells (2006–2009) from Scopus.	SEAL outperformed Sentence-BERT in clustering top-cited papers. Citation data was more effective than content data for prediction. Achieved lower entropy, indicating better concentration of top-cited papers.	Limitations: • Limited to a single domain and time period. • Sentence-BERT not pre-trained on scientific literature. Future Work: • Extend analysis to interdisciplinary fields. • Combine content and citation data models. • Pre-train models on scientific literature.	Zerva et al. (2020)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
Aspects of research impact: Academic collaboration						
10	Collaboration Prediction Models	Objective: • Develop models to predict future research collaborations using network dynamics and deep learning. Motivation: • Predicting collaborations aids in fostering academic partnerships. • Existing methods often lack integration of multi-dimensional factors.	Methodology: • Constructed multi-network representation learning frameworks. • Used knowledge graphs, node embeddings, LSTM-CNN models, and spatial-temporal factors. Datasets: • Web of Science (Information Science and Library Science). • CORD-19 dataset. • AMiner and HSE datasets. • IEEE Xplore (Guangdong-Hong Kong-Macao Greater Bay Area).	Improved link prediction accuracy over baseline methods. Demonstrated the importance of integrating textual, structural, and temporal features. Achieved higher precision and recall in collaboration predictions.	Limitations: • Computational intensity and scalability issues. • Generalization to other regions or disciplines requires validation. Future Work: • Extend models to other domains. • Incorporate additional features like institutional influence. • Optimize computational efficiency.	Kanakaris et al. (2021) Makarov & Gerasimova (2019) Yang et al. (2023) Zhang et al. (2019) Zhou et al. (2019)
11	Link Prediction in Author Networks	Objective: • Improve link prediction techniques in co-authorship and citation networks using various features and methods.	Methodology: • Evaluated similarity metrics and machine learning models (ANN, SVM, XGBoost).	Achieved high prediction accuracy and AUC values. Incorporating nodal attributes improved predictions.	Limitations: • Limited to specific datasets; scalability untested. • Relied on topological features, excluding semantic information.	Roopashree & Umadevi (2014) Song et al. (2022) Vital & Amancio (2022)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
		Motivation: • Accurate link prediction is crucial for understanding research collaborations. • Traditional models often neglect temporal dynamics and nodal attributes.	• Developed nodal-attribute-based predictors like SIN and KMC. • Employed supervised learning with minimal feature sets. Datasets: • APS journals (450,000 articles). • Collaboration data from statistical journals. • NetScience dataset.	Minimal features were effective for accurate link prediction.	Future Work: • Extend to more extensive networks. • Include temporal and content similarity features. • Optimize models for real-world applications.	
12	Co-authorship and Collaboration Networks	Objective: • Analyze co-authorship networks and their implications for citation and research impact. Motivation: • Research collaborations are critical for capacity building. • Understanding network structures influences policy formulation and impact prediction.	Methodology: • Developed co-authorship networks from bibliometric data. • Analyzed network properties and metrics (degree distribution, centrality). • Used Structural Variation Analysis (SVA) and Propensity Score Matching (PSM).	Identified growth patterns and collaboration trends across countries. Found that network characteristics predict research impact. Boundary-spanning papers had higher citation counts.	Limitations: • Data may miss informal collaborations. • Limited to specific disciplines. Future Work: • Extend to subfields and integrate new metrics. • Investigate causal relationships between network changes and policies.	Grodzinski et al. (2021) Hou et al. (2023) Vinayak et al. (2023)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
			Datasets: • Web of Science for AI and ML research (USA, China, India). • Scopus data on Degenerative Cervical Myelopathy. • Microsoft Academic Graph (MAG).		• Refine similarity computation methods.	
13	Anomaly Detection in Citation Networks	Objective: • Develop a deep graph learning framework (GLAD) for detecting anomalous citations. Motivation: • Anomalous citation practices undermine academic integrity. • Detecting anomalies is challenging due to network complexity.	Methodology: • Proposed GLAD using graph neural networks. • Incorporated semantic mining via citation purpose classification (CPU algorithm). • Used autoencoders for edge feature representation. Dataset: • Microsoft Academic Graph with synthetic anomalous citation data.	GLAD achieved a high F1-score in detecting anomalous citations. CPU algorithm showed high accuracy and robustness. Effectively identified anomalous citation relationships.	Limitations: • Reliance on synthetic data may limit real-world applicability. • Scalability to large datasets is challenging. Future Work: • Apply GLAD to real-world datasets. • Optimize efficiency for more extensive networks. • Explore hybrid models for better semantic integration.	Liu, Xia et al. (2022)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
Aspects of research impact: Specialised research impact						
14	Journal Metrics Prediction	Objective: • Predict journal impact factors and related metrics using time series and bibliometric data. Motivation: • Missing JIF values and lack of predictive models hinder journal evaluation. • Predicting future performance aids stakeholders in decision-making.	Methodology: • Developed kNN Imputation for missing data handling. • Constructed regression models using ML and deep learning (LSTM). Datasets: • 145 AI journals (1997–2021) with various metrics. • 24,000+ journals from Scopus (2000–2020).	kNNI outperformed conventional methods in filling missing JIF values. LSTM models achieved low MAPE for citations and CiteScore predictions. Demonstrated efficacy of deep learning in predictive bibliometrics.	Limitations: • Limited dataset sizes restricted model usage. • LSTM complexity increases computational demands. Future Work: • Collect data across diverse fields. • Integrate state-of-the-art techniques for larger datasets. • Incorporate additional features like altmetrics.	Croft & Sack (2022) Hua & Huynh (2024)
15	Researcher Impact and Leadership	Objective: • Rank researchers and predict research potential using network features and proximities.	Methodology: • Developed Scientific Quantitative Rules (SQR) using CNNs and decision trees. • Implemented supervised learning with LR and DNN models. • Proposed PRLR model incorporating cognitive, geographical, and institutional proximities.	Identified impactful parameters for award recipients. Senior co-author characteristics predicted junior researchers' future impact.	Limitations: • Dataset limitations due to manual selection. • Models may be computationally expensive.	Alshdadi et al. (2023) de Abreu Batista et al. (2021) He et al. (2022)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
		Motivation: • Accurate assessment of researchers aids in academic evaluations. • Existing measures lack standardisation and integration of proximity metrics.	Datasets: • Records from 1,500 researchers across three domains. • ACM citation network data. • Web of Science for various disciplines.	PRLR outperformed collaborator recommendation models.	Future Work: • Extend frameworks to additional domains. • Integrate factors like funding and collaboration networks. • Investigate scalability for broader datasets.	
16	Institutional Impact Prediction	Objective: • Develop a methodology for predicting the future impact of institutions using various features. Motivation: • Predicting institutional impact aids in resource allocation and funding decisions. • Existing methods often rely solely on historical relevance scores.	Methodology: • Proposed an XGBoost-based predictive model. • Leveraged feature selection techniques to identify relevant features. Dataset: • MAG dataset: 33,953 authors, 19,343 papers, 4,524 institutions (2000–2015).	Author-based features were significant predictors. XGBoost outperformed GBDT in prediction accuracy. Geographic and economic factors improved model performance.	Limitations: • Limited to eight conferences and specific datasets. • Computational cost is high for large datasets. Future Work: • Include more diverse conferences and datasets. • Incorporate features like institutional funding. • Explore dynamic predictive modelling.	Bai et al. (2017)

(continued)

Table 3. Third Generation Research Impact Assessment Models – Key Objectives, Methodologies, and Findings (contiune)

No.	Focused parameter	Objective & motivation	Methodology & dataset	Key findings	Limitations & future work	References
17	Domain-Specific Studies	Objective: • Study the cultural impact of art journal reading and develop classification schemes for specific fields. Motivation: • Enhancing national cognition and cultural acceptance among students. • Existing citation classifications lack generalisation across disciplines.	Methodology: • Utilized deep learning algorithms for journal recommendation. • Designed surveys and applied statistical evaluation. • Developed a five-category citation classification scheme. Datasets: • Survey data from 200 college students. • 2,153 citation sentences from Indonesian food science journals.	Art journal reading improved national cognition among students. LSTM outperformed other classifiers in citation function classification. Categories aligned with journal article sections.	Limitations: • Limited scope to specific universities and journals. • Class imbalance impacted performance. Future Work: • Extend frameworks to larger datasets. • Incorporate semantic features for classification. • Refine guidance systems for broader applicability.	Huang et al. (2022) Yaniasih & Budi (2021)

Table 4. Data-Driven Factors Explaining and Predicting Research Impact

Factor/Category	Key insights for research prediction	Analytical challenges	Utility in research impact prediction	Approaches to leverage the factor
Citation Metrics	Citation counts measure scholarly influence and visibility (Ji et al., 2024).	May not capture qualitative contributions or broader societal impact.	Serves as a benchmark for evaluating the influence and reach of research outputs.	Combine citation metrics with qualitative assessments to gain a nuanced understanding of research impact.
Temporal Dynamics Parameters	Patterns like citation bursts and aging effects reveal trends in research influence over time (Abbas et al., 2023).	Requires continuous monitoring of citations to capture changing trends.	Helps identify the longevity and evolving relevance of research topics or publications.	Conduct longitudinal studies to track and predict long-term research trajectories.
Predictive Models/ Methods	Models like LSTM networks forecast future citation trajectories and research impact (Du et al., 2024).	Requires robust datasets and computational resources for accurate predictions.	Provides forward-looking insights into the potential influence of research works.	Develop institutional expertise in predictive modelling and validate models using empirical data.
Network Analysis Parameters	Collaboration and citation networks highlight influential researchers and work (Aljohani et al., 2021; Liao et al., 2024).	Managing and analysing large, complex networks effectively.	Identifies key nodes for collaboration and pathways for the diffusion of knowledge.	Utilise advanced network visualisation tools and conduct periodic analyses of collaboration structures.
Content Relevance Parameters	Trending keywords and topics enhance visibility and citation potential (Shi et al., 2019; Vinayak et al., 2023).	Rapidly changing trends require constant adaptation in research focus.	Helps align research with emerging academic and societal interests.	Apply text mining and topic modelling to identify emerging trends and align research focus accordingly.

(continued)

Table 4. Data-Driven Factors Explaining and Predicting Research Impact (contiune)

Factor/Category	Key insights for research prediction	Analytical challenges	Utility in research impact prediction	Approaches to leverage the factor
Institutional Metrics	Metrics like h-indices provide insights into individuals' or institutions' overall research productivity and quality (Bai et al., 2017; Heo et al., 2023).	Overemphasis on metrics may overshadow qualitative contributions.	Reflects productivity trends and leadership in advancing research within specific fields.	Use h-indices in combination with qualitative assessments to ensure balanced evaluations.
Journal Metrics	Impact factors and journal rankings indicate the potential visibility of published research (Croft & Sack, 2022).	Metric reliability varies across journals, and open-access models pose additional challenges.	Guides researchers in selecting journals for maximising visibility and potential impact.	Analyse journal trends and encourage publication in high-impact, relevant journals.
Big Data Integration Parameters	Big data and biobanks facilitate large-scale analyses and predictions in healthcare and biomedicine (BRIF workshop group, 2011; Howard et al., 2018).	Ethical concerns and challenges in integrating heterogeneous data sources.	Supports predictive models and evidence-based decision-making for future research directions.	Develop robust data governance frameworks and invest in infrastructure for handling large datasets.
Economic and Geographic Factors	Geographic proximity and economic resources shape research capacity and outcomes (Grodzinski et al., 2021; Zhang et al., 2019).	Regional disparities in resource availability and access to research infrastructure.	Explains variations in research productivity and potential across regions.	Use geographic data to identify underserved areas and foster collaborations for equitable research development.

Table 5. Key Metrics and Platforms of the Fourth Generation – Focus on Alternative and Digital Measures of Research Impact

Index/Model	Year	Description	Calculation method	Reference
Aggregated alternate matrix platforms				
Altmetric Attention Score	2011	A weighted count of all the online attention an individual research output receives, represented by the Altmetric "donut" badge.	Combines mentions from social media, news outlets, policy documents, blogs, and other platforms, each weighted differently to reflect relative importance.	Priem (2010) Priem et al. (2011) Roemer & Borchardt (2015)
PlumX Metrics	2016	Provides insights into how people interact with individual pieces of research output online.	Tracks five categories: Usage, Captures, Mentions, Social Media, and Citations, aggregating data from various online sources.	Lindsay (2016)
PLOS Article-Level Metrics	2009	Offers detailed metrics at the article level, including views, downloads, citations, and social media interactions for articles published in PLOS journals.	Collects data on article views, downloads, citations (from Scopus, CrossRef), social media mentions, and other usage metrics directly from the PLOS platform and third-party providers.	Yan & Gerstein (2011)
Dimensions Badges	2018	Provides citation counts and Altmetric data for individual research outputs, integrating traditional and alternative metrics.	Aggregates citations from scholarly sources and altmetric data, displaying them through badges on publisher or institutional repositories.	Jamwal & Kumar (2022)
Impactstory Metrics	2012	Allows researchers to track the online impact of their work through various alternative metrics.	Aggregates data from sources like Twitter, Mendeley, GitHub, and others to display metrics for individual research outputs.	Konkiel (2014) Sugimoto et al. (2017)
Kudos Metrics	2014	Helps researchers explain and share their work to enhance its visibility and impact, tracking subsequent metrics.	Tracks metrics such as views, downloads, citations, and altmetrics, correlating them with researchers' dissemination efforts through the Kudos platform.	Erdt et al. (2017)

(continued)

Table 5. Key Metrics and Platforms of the Fourth Generation – Focus on Alternative and Digital Measures of Research Impact (contiune)

Index/Model	Year	Description	Calculation method	Reference
Platform-specific metrics				
ResearchGate Score	2016	A metric from the ResearchGate platform, combining publications, interactions, and reputation.	Based on an undisclosed algorithm that considers publications, reads, citations, and engagement on the platform.	Hoffmann et al. (2016)
Mendeley Readers	2009	Measures the number of times a research output has been added to users' libraries on Mendeley, indicating readership and interest.	Counts the number of unique Mendeley users who have added the document to their personal libraries.	Sugimoto et al. (2017) Zahedi & Costas (2020)
SSRN Downloads	1994	Counts the number of times a paper has been downloaded from SSRN, indicating interest in the work.	Counts total downloads for each paper uploaded to SSRN, including abstract views and full-text downloads.	Kakushadze (2016)
CiteULike Bookmarks	2004	Indicates how many users have saved a research output to their CiteULike libraries.	Counts the number of unique users who have added the document to their CiteULike libraries.	Emamy & Cameron (2007)
GitHub Repository Stars/Forks	2008	Measures engagement and reuse of code or software associated with research outputs.	Counts the number of stars (likes) and forks (copies) of repositories on GitHub.	Dozmorov (2018)
Crossref Event Data	2016	Collects data on online mentions and interactions using Crossref DOIs.	Tracks events like social media mentions and links associated with DOIs from multiple sources.	Rittman (2020)
Social media and online mentions				
Social Media Mentions	2004	Tracks mentions of a research output on platforms like Twitter, Facebook, and LinkedIn.	Aggregates mentions, shares, likes, and comments related to a research output across various social media platforms.	Sugimoto et al. (2017)
Scholarly Blog Mentions	1999	Tracks mentions of scholarly works in academic and research blogs.	Aggregates data from blog aggregators to count mentions of research outputs in blog posts.	Joshi et al. (1999)
Wikipedia Citations	2001	Measures how often a scholarly work is cited within Wikipedia articles.	Counts the number of times a research output is referenced in Wikipedia entries across different languages.	Stalder & Hirsh (2002)
YouTube Video Views	2005	Measures views for videos related to research outputs, such as presentations.	Counts the total number of views for videos associated with a research output on YouTube.	Dai & Wang (2023)

圖書資訊學刊　第23卷第1期　（2025.6）　頁1-77
https://doi.org/10.6182/jlis.202506_23(1).001

構繪研究影響力科學的不同世代：對衡量指標、框架與預測方法的範域文獻回顧

Mapping the Generations of Research Impact Science: A Scoping Review of Metrics, Frameworks, and Predictive Approaches

Mudassar Arsalan[1], Omar Mubin[2], Abdullah Al Mahmud[3]

摘　要

本文對研究影響力評估框架進行系統性範域文獻回顧，採Arksey與O’Malley（2005）提出、Levac等人（2010）改良、並對應PRISMA-ScR準則的方法，查找Scopus、Web of Science、PubMed、IEEE Xplore、Google Scholar以及灰色文獻後，篩選出有關評估框架、標準和方法之139篇研究論文。研究影響力科學的發展從引文測量轉變至複雜的資料驅動，分為四個世代：第一代書目計量著重於出版生產力；第二代多維框架納入社會、經濟和政策指標；第三代資料驅動整合機器學習和預測分析；第四代另類計量掌握即時數位參與。主要挑戰包括標準化受限、地區性偏誤，以及新興技術運用不足，故本研究指出需建立預測性多維框架和標準化分類，提升影響力評估的擴展性和前瞻性，有助精進評估方法，以面對社會挑戰並指引策略研究之經費補助。

關鍵字：研究影響力評估、世代框架、書目計量指標、多維度模型、預測分析

[1,2] 澳洲西雪梨大學資訊與數據科學學院
School of Computer, Data and Mathematical Sciences, Western Sydney University, Penrith, Australia
[3] 澳洲斯威本理工大學設計創新中心
Centre for Design Innovation, Swinburne University of Technology, Melbourne, Australia
* 通訊作者Corresponding Author: Mudassar Arsalan, E-mail: mharsalan@gmail.com

註：本中文摘要由圖書資訊學刊編輯提供。
以APA格式引用本文：Arsalan, M., Mubin, O., & Al Mahmud, A. (2025). Mapping the generations of research impact science: A scoping review of metrics, frameworks, and predictive approaches. *Journal of Library and Information Studies, 23*(1), 1-77. https://doi.org/10.6182/jlis.202506_23(1).001
以Chicago格式引用本文：Arsalan, Mudassar, Omar Mubin, and Abdullah Al Mahmud. “Mapping the Generations of Research Impact Science: A Scoping Review of Metrics, Frameworks, and Predictive Approaches.” *Journal of Library and Information Studies* 23, no. 1 (2025): 1-77. https://doi.org/10.6182/jlis.202506_23(1).001.