INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Computational Innovation Studies:

Understanding innovation studies through novel scientometric approaches

Ana Teresa Martins Sousa Santos

PhD in Complexity Sciences

Supervisor:
Professor Doutor Sandro Mendonça
Assistant Professor
Instituto Universitário de Lisboa (ISCTE-IUL) -
Business Research Unit (BRU-IUL)

December, 2021

TECNOLOGIAS
E ARQUITETURA

Computational Innovation Studies:

Understanding innovation studies through novel scientometric approaches

Ana Teresa Martins Sousa Santos

PhD in Complexity Sciences

Jury:

Doctor Jorge Louçã, Full Professor, Instituto Universitário de Lisboa (ISCTE)
Doctor João Caraça, Full Professor, Instituto Superior de Economia e Gestão da Universidade de Lisboa (ISEG)
Doctor Aurora Teixeira, Full Professor, Faculdade de Economia da Universidade do Porto
Doctor Rui Lopes, Associate Professor, Instituto Universitário de Lisboa (ISCTE)
Doctor Sandro Mendonça, Assistant Professor, Instituto Universitário de Lisboa (ISCTE)

December, 2021

“Quando não há transparência, não há confiança”
– Dalai Lama

Acknowledgements

The following personal acknowledgements will be addressed in Portuguese:

Começo por agradecer ao meu Orientador. Ao Professor Doutor Sandro Mendonça por todo o apoio que me deu ao longo dos últimos anos, por toda a confiança, por se ter tornado um amigo. Agradeço-lhe por me ter ensinado tanto sobre inovação e sobre as especificidades do conhecimento contemporâneo sobre o fenómeno da inovação, com uma abordagem que só é possível a quem se dedica de corpo e alma, há tantos anos, às causas da economia da mudança e do progresso técnico. Muito, muito obrigada!

Agradeço à minha Instituição de acolhimento, o Instituto Universitário de Lisboa (ISCTE) da Universidade de Lisboa e também a Faculdade de Ciências que no decorrer do curso do meu doutoramento contribuíram extraordinariamente para superar minhas limitações. Também ao atual Diretor do Programa Doutoral em Ciências da Complexidade, Professor Jorge Louçã que me deu todas as condições para que pudesse executar os meus trabalhos de doutoramento. Agradeçxs4o ao meu comité de tese e em particular ao Professor Doutor Rui Jorge Lopes pelos seus sábios conselhos e tudo o que ensinou sobre teorias de redes.

Não posso deixar de agradecer ainda a todos os meus Professores ao longo dos anos, em especial aos Professores de Mestrado na Universidade de Aveiro e no ISEG/ULisboa que fizeram nascer em mim o gosto pela investigação e pela ciência que me conduziu até aqui. Um agradecimento especial aos colegas de todos os cantos do planeta, com quem me cruzei, no programa Doutoral Ciências da Complexidade, por me terem aconselhado, ouvido e ajudado em tantos momentos, cada um a sua maneira. Outro agradecimento especial vai para o Professor Doutor Manuel Mira Godinho, o Professor Doutor Vitor Corado Simões, o Doutor Hugo Confraria, o Professor Doutor Bram Timmermans e o Professor Doutor Jon Mikel Zabala-Iturriagagoitia pelos seus comentários e sugestões particularmente brilhantes.

Não posso deixar de agradecer a todos os revisores anónimos que pacientemente responderam com *insights* - muitas vezes instigantes - e comentários em todos os artigos que submeti a conferências e revistas. Isso definitivamente tornou o meu trabalho mais robusto, mais detalhado, e ajudou-me a transformar comentários negativos em forças positivas para melhorar a qualidade do meu projeto de investigação. Também a todos os participantes que nas conferencias *24th Nordic Workshop on Bibliometrics and Research Polic*y, Reykjavik 27-29 November 2019; 18th International Conference of

the International Society for Scientometrics and Informetrics, Leuven 12-15 July 2021; and *II Simpósio Latinoamericano sobre Estudios Metricos en Ciencia y Tecnologia*, Bolgotá 13-15 September 2021 contribuíram com questões pertinentes e geradoras de novas ideias.

Agradeço especialmente à minha amiga Raquel Banheiro, que tanto pensou comigo, fez comigo, fez por mim, e por ser, na sua calma, uma força da natureza. Em particular à Maricarmen Rodriguez por tanta dedicação e ajuda nas apresentações em espanhol que se tornou tão conhecedora dos meus trabalhos quanto eu.

Não menos importantes, quero também agradecer aos meus amigos e família de sangue e de coração, que sempre me apoiaram, que compreenderam as ausências e os cansaços, e que tantas vezes me ouviram e me deram o braço. Um agradecimento muito especial à minha Mãe e ao meu Pai que me fizeram chegar até aqui, que me levam ao colo nos momentos difíceis, e se alegram mais do que eu com as minhas conquistas. Ao meu avô... que não deixa a neta dele atirar a toalha ao chão!

Resumo

A cientometria é uma importante área de investigação dedicada ao estudo quantitativo da ciência e está a expandir-se a um ritmo sem precedentes. Surgiu como um paradigma de avaliação e espera-se que ajude na resolução de problemas sociais complexos. Apesar da sua importância, pouco se sabe sobre os guardiões da ciência e os mecanismos de governação editorial mais amplos que ajudam a orientar os esforços científicos.

Neste projeto, seguimos uma perspetiva pouco explorada (assumimos os conselhos editoriais e as revistas como veículo institucional), numa área específica de investigação científica (os Estudos de Inovação). Abordamos diferentes aspetos em três etapas: em primeiro lugar, produzimos um retrato abrangente do fenómeno editorial, sondando as características estruturais heterogéneas dos conselhos editoriais, que são dominados por editores masculinos, anglo-americanos que exibem uma concentração de 85% das posições editoriais em 20% dos países; em segundo lugar, comparamos os materiais publicitários das revistas (*blurbs*) com uma medida de semelhança do cosseno identificando seis revistas com mais de 80% de semelhança semântica com a *Research Policy* (a revista principal) e descobrimos que as revistas podem ser classificadas em quatro grupos; e em terceiro lugar, combinamos os resumos (*abstracts*) das revistas realmente publicados com a descrição publicitária, revelando que o conteúdo selecionado em cinco revistas teria tido maior interesse para outras. Por fim, desenvolvemos uma ferramenta interativa que permite comparar a semelhança dos conteúdos publicados pelas revistas. Estas estratégias de investigação apresentadas juntam-se ao portfólio de metodologias que os analistas de política científica podem usar para compreender sistematicamente as agendas de revistas, a fim de refletir sobre o que foi realizado e o que ainda está por fazer.

Palavras-chave: conselhos editoriais; estudos de inovação; revistas científicas, *text mining*, classificação de texto, *machine learning*, estratégias de submissão.

Abstract

Scientometrics is an important research field that is dedicated to the quantitative study of science and is expanding at an unprecedented rate. It emerged as an evaluation paradigm and is expected to assist in the resolution of complex societal problems. For years, the impact of research has been at the top of the agenda for policymakers, however little is known about the gatekeeping processes and the broader editorial governance mechanisms that helps steer scientific efforts.

In this project, we will pursue an under-explored perspective (we take on editorial boards and the journals as an institutional vehicle) and apply to a specific field of academic research (*Innovation Studies*). We address different aspects in three steps: first, we provide a comprehensive portrait of the editorship phenomenon by probing the heterogeneous structural features of boards, which dominated by men and angloamerican editors displaying a concentration of 85% of editorial positions in 20% of the countries; second, we compare journals' advertising materials (blurbs) with a cosine similarity measure identifying six journals with more than 80% semantic similarity with *Research Policy* (the leading journal) and find out that the journals can be classified into four groups; and third, we match journal blurbs with the abstracts of papers actually published disclosing that the contents from five journals would have greater interest to other outlets. Finally, an interactive tool was developed so that researchers are better empowered to compare the similarity of journals contents in the future. These research strategies presented add to the portfolio of methodologies that science policy analysts can use to systematically understand journal agendas in order to reflect on what has been accomplished and what remains to be done.

Keywords: Editorial boards; innovation studies, interlocking editorships; journals, text mining, text classification, machine learning, submission strategies.

List of outputs: posters, publications, presentations and products

This work generated 14 tangible outputs:

- #2 Posters, #7 papers, #5 presentations and #1 product.

Posters

Judging journals by their "blurbs": a text mining approach to innovation journals, 24th Nordic Workshop on Bibliometrics and Research Policy, Reykjavik 27-29 November 2019;

Watching over innovation studies: Profiling the gatekeepers, ISSI 2021 Conference, 12-15 July 2021.

Papers published in proceedings

Journals' agendas versus actual publications: A first look at article dynamics in innovation journals, Proceedings of the 18th International Conference of the International Society for Scientometrics and Informetrics, 997–1008, 12-15 July 2021.

The small world of editorships: A network on innovation studies, Proceedings of the 18th International Conference of the International Society for Scientometrics and Informetrics, 985-996, 12-15 July 2021.

Watching over innovation studies: Profiling the gatekeepers, Proceedings of the 18th International Conference of the International Society for Scientometrics and Informetrics, 1537–1538, 12-15 July 2021.

Papers in journals (Q1)

The small world of Innovation Studies: An "editormetrics" perspective, Scientometrics, 2022.

Do papers (really) match journals' "Aims & Scope"? A computational study of Innovation Studies, Scientometrics, 2022.

Papers in journals (under review)

Keeping a close watch on innovation studies: Opening the black box of journal editorships, Quantitative Science Studies, 2021.

Presentations in conferences

Judging journals by their "blurbs": a text mining approach to innovation journals, 24th Nordic Workshop on Bibliometrics and Research Policy, Reykjavik 27-29 November 2019 (poster).

Watching over innovation studies: Profiling the gatekeepers, ISSI 2021 Conference, 12-15 July 2021 (poster).

Journals' agendas versus actual publications: A first look at article dynamics in innovation journals, 18th International Conference of the International Society for Scientometrics and Informetrics, Leuven, 12-15 July 2021.

The small world of editorships: A network on innovation studies, 18th International Conference of the International Society for Scientometrics and Informetrics, Leuven, 12-15 July 2021.

Os guardiões dos Estudos de Inovação: O lugar dos editores latino-americanos nas revistas de referência, Latmetrics 2021 Conference, Bogotá, 13-15 September 2021.

Product

"ScienceSubmit", an interactive app designed to function as science management tool for prospective authors willing to submit their work to indexed journals, available online at: https://sciencesubmit.shinyapps.io/ShinyApp/

Index

List of Tables

List of Figures

Acronyms

AMJ Academy of Management Journal

AMR Academy of Management Review

ASQ Administrative Science Quarterly

CJE Cambridge Journal of Economics

EB Editorial Board

HR Human Relations

ICC Industrial and Corporate Change

IJTM International Journal of Technology Management

IDF Inverse Document Frequency

JIBS Journal of International Business Studies

JIF Journal impact factor

JMS Journal of Management Studies

LSA Latent Semantic Analysis

MS Management Science

NLP Natural language processing

OSc Organization Science

OSt Organization Studies

RDM R&D Management

RS Regional Studies

RP Research Policy

SBE Small Business Economics

SSS Shiny Server Side

SUI Shiny User Interface

SNA Social Network Analysis

SMJ Strategic Management Journal

SPRU Science Policy Research Unit

SVD Singular Value Decomposition

TASM Technology Analysis & Strategic Management

TF Term Frequency

TFSC Technological Forecasting and Social Change

Tec Technovation

WoS Web of Science

CHAPTER 1

Introduction

This Thesis advances new approaches to analyse, in automated but theory-based ways, the dynamics behind the contents of academic journals. We aim to study the background of editorial processes and how it might impact the scientific production. By doing so, we situate this Thesis in the broader agenda of scientometric research. Underexploited techniques are presented here not only as useful but valuable in the assessment of board editors, the journals themselves and contents, both advertised and published ones.

We set out to contribute to a better understanding of the current mechanisms of governance of scientific areas as revealed by the contents, blurbs, and editorial bodies of journals. Several questions about this problem are addressed in this Thesis, ranging from knowing the demographic characteristics of the editors to the interactions between editors and journals; whether it is possible to semantically compare journals based on the contents published and their self-statement of interest, and whether it is possible to derive methods that automatically compare journal contents. These problems present some unique challenges, such as the volume of data produced by journals, gathering that data, and processing it in a timely and manageable manner. To tackle these challenges, this Thesis proposes a data-driven/computationally-intensive approach to assess published research: it applies recently developed qualitative (text content) and quantitative (paper counts) methods to understand how one particular field of work is organized, *Innovation Studies*.

In this work, we also make inferences about editorial governance in a different perspective: as a complex system. Complexity is a type of world framework that envisions systems through the analysis of agents and their interactions (Watts, 2004). Individual verification of the parties is left out and a vision with a focus on the whole is chosen (McKelvey, 2001). Castellani and Hafferty (2009) created a map that depicts the history and evolution of complexity sciences. It covers specific topics, key scholarly themes, and notable researchers[1]. The map is seen as an evolving introduction to the multiplicity of actors and relations. With his work, the editorial governance is seen as a complex system. Science is now viewed as an open, multi-actor, multi-dimensional and multi-output phenomenon that needs novel approaches so that its processes could be fully understood.

[1] An updated version of the map can be found at https://www.art-sciencefactory.com/complexity-map_feb09.html

In this project, the methods proposed are applied to real data, this means facing the difficulties of actual empirical work and attempting to generate results that may be useful to advance the very research program of *Innovation Studies*. The field of *Innovation Studies* is taken here as the source ground for being dynamic (evolving over time) and encompassing journals with high classification suggesting high standards of quality. This area of study is committed to the understanding of the factors, processes and effects of science, technology and innovation (Caraça et al., 2007). It is also a multidisciplinary enterprise and an intellectual project geared toward drawing practical insights that can be applied to management and policy (Fagerberg et al., 2013; Martin, 2019). This Thesis is an account of the study of networks in the context of complex systems, with a focus on the journals that emerge and evolve over time, as well as the editors who interact with one another (Louçã et al., 2004; Pereira et al., 2021).

Research questions

The work conducted in this Thesis revolves around the following overarching research question: "What under-exploited data (blurbs) and under-utilised techniques reveal about the scientific organization of 'Innovation Studies'?". Indeed, we are driven by the opportunity of addressing scientometrics from a computational perspective and believe to have found an apt body of research to deploy it.

Several operative subquestions branch out from this main research question. These start by exploring the editorial structure of innovation journals, by converting journals promotional materials into an empirical source, and by relating actual published paper outputs to what journals' state are their agendas. How to probe the structural features of journal boards?

- How editorial board composition and characteristics illuminate the organisation of the field of "Innovation studies"?
- How editorial blurbs provide a source to map and measure the features of "Innovation Studies"?
- How the published content/self-advertised agendas compare and contrast in "Innovation Studies"?

Following a critical review of some of the current knowledge on these and related topics, I attempt to answer these questions in the three chapters that follow.

Thesis outline

The Thesis is divided into five chapters followed by a conclusion. After the bibliographic references there is a section about posters and published and relevant to this work.

Chapter 1 discusses the role of journals in the academic ecosystem as well as their historical significance in research. For being such an important piece of the puzzle, journals are taken as a basic institutional background in this Thesis. Innovation Studies are presented as the scientific field providing the data for the analysis and allowing the methodology to be tested.

Chapter 2 analyses the identity of editorial board members. For emerging as decisive power-brokers and scrutinisers in the contemporary science "ecosystem", but seldom subject of systematic study, a comprehensive portrait of the editorship phenomenon in *Innovation Studies* is presented by probing the structural features of the boards of 20 leading innovation-oriented journals as of 2019.

Chapter 3 compares journals' advertising materials, the blurbs, for enhancing a scientometric approach to journals. A latent semantic analysis is applied to mine for themes and context in journals to organise outlets into clusters and rank them based on their similarity with the main reference in the field, *Research Policy*.

Chapter 4 matches journals' blurbs and abstracts already published to understand how past published papers reveal the actual publication profile interests. It is assumed that the interests described in introductory sections are consistent with the content of regular issues and that "blurbs" can guide authors, readers, and policymakers. Because of being an informative set, the chapter does not make that assumption but tests it with a machine learning approach.

Chapter 5 presents an interactive tool so that researchers are better empowered to compare the similarity of journals contents in the future. The algorithm developed in chapter 4 is applied here to generate an app made available online.

Finally, in chapter 6, the main results found are reviewed. The limitations of this methodology are discussed, new avenues for future research are presented, and the importance of discussing and researching editorial governance is emphasized. The appendices of this Thesis provide the posters presented during this thesis.

CHAPTER 2

The journals in the innovation studies field

2.1. The journals and their boards

2.1.1. The academic journal as a unit of analysis

As mainstream communication channels of scientific production, journals are subject to evaluation in multiple perspectives, such as dissemination (Bryce et al., 2020; Chavarro et al., 2017). Hence, numerous indexes emerged to provide an answer for the multiple interests explored. However, to measure quality, journal impact factor (JIF) gained traction and eventually became an "industry standard" (Garfield, 1972; Pinski and Narin, 1976). Despite the science indicators research program's accomplishments, much has been written regarding the limitations and even noxious consequences of single indexes (Drivas and Kremmydas, 2020; Martin, 2019; Moed and van Leeuwen, 1996; Wilhite et al., 2019). Indeed, no other metric in bibliometrics has attracted as much discussion as the JIF (Archambault and Larivière, 2009; Callaway, 2016; Larivière and Sugimoto, 2019; Moed and van Leeuwen, 1996; Seglen, 1997). The pioneer of the JIF concept, (Garfield, 1983), nonetheless argued that an article-based journal score system would help editors in deciding when to cease its print or to librarians when selecting subscriptions. Researchers themselves came to look at these prestige pointers when choosing the journal to submit their works for publication (Harzing, 2012; Welch, 2012). In turn, this structure created pressure on journals which, having higher scores, receive many more publications than they can publish, leading to lower acceptance rates, and a race in tables of ranks (Hall and Martin, 2019a; Sugimoto et al., 2013; Vogel et al., 2017).

Thus, professional academic journals have continued to hold fast to their role as *switching boards* of the modern process of exchanging scientific information (vis-à-vis monographs or proceedings, for example; see González-Albo and Bordons (2011) and Juzych et al. (1991)). Journals carry out a dual role: as a locus of readership/followership (demand-side) and a target for contributions/leadership (supply-side). They are intermediary institutions, connecting producers and consumers of scientific work. As a third party with a degree of strategic autonomy, journals could be said as benefiting the status of a *platform* in the science sectoral system (including in an economic sense; see Mansell and Steinmueller (2020). That is, rather than being a simple functional interface in the frictionless world of science, journals have become central institutions in the structured ecosystem of contemporary research, performing a roundabouts function for academic attention and being active modulators of scholarly behaviour.

Hence the relevance of "journalytics", a term elected here to refer to data science and statistical techniques devoted to assessing journals' profile and status as key institutions of modern organised science. This could be implemented by turning the tools of scientometrics more explicitly on journals, i.e., setting up an empirical inquiry of journals as a phenomenon. Besides feeding science studies (sociology of science, economics of science, etc.), increasing focus on journals can reorient evaluative scientometrics by redefining the topics addressed. Whereas displacing the JIF from its dominant role has not yet happened, at least some effort can be invested into enriching the pluralism of journal indicators and ranking tools (for alternatives to the JIF see Larivière and Sugimoto (2019); and Rafols et al. (2012)), namely through content-based and algorithmic subject categorisation and comparisons (Rafols and Leydesdorff, 2009; Rousseau et al., 2019).

Advances in scientometrics have always been partially contingent on developing of new sources of (meta)data. We claim that some of the data could be better appreciated as *internal* to the journal and *qualitative* in nature. We start from an assumption: journals are not just carriers of contents; they *are* content in themselves. Journals, however, are constructs too, i.e., intellectual, and organisational projects and show a capacity to assert themselves in their own terms and evolve, either by the entrepreneurial initiative of editors or through the instituted decision-making processes of holding scientific societies and commercial publishers, and in so doing they leave a data trail (Braun, 2005; Mendonça et al., 2018).

If science is to be conceived as a "complex, self-organizing, and evolving network of scholars, projects, papers, and ideas" (Fortunato et al., 2018, p. 359), we argue that the "serial literature" should be present in the definition to be included in the analysis. Since they are essential nodes in the process of knowledge circulation and validation (Hicks, 2004), journals remain the key locus where research results are formally shown, validated and exchanged (Van Raan, 2005). Hence, the case is made that journals can be taken *de per se* as another building block of the science system and be seen to constitute an actual source of new data in scientometrics.

2.2. Innovation studies

Innovation is a variegated and vicarious process. It is simultaneously patterned and unpredictable, both across space and time. Innovation is recognised as the essential factor for economic growth either at the firm, regional or national levels (Fagerberg and Godinho, 2004; Grillitsch et al., 2019). It is also fundamental to the wholesome development of societies in a way consistent with long-run existential challenges such as climate change and human dignity (Fagerberg et al., 2018; Martin, 2019; Schot and

Steinmueller, 2019). There is a private side to this process, but also a public one; that is to say, entrepreneurs are engaged in it, regulators frame it, and government agencies push it, while a wide array of not-for-profit actors in between, like universities and professional associations, enable, empower and modulate it. An influential contemporary view is that innovation is a composite and evolving phenomenon, better understood as a system or a network (see, e.g., Freeman, 2008; Jarrahi and Sawyer, 2019).

Although firmly rooted in a tradition starting in Adam Smith, going through Marx, and culminating in Schumpeter (Freeman and Soete, 1997), it is in the second half of the twentieth century that innovation economics emerges and gains coherence (Fagerberg, 2003). The field that is currently best known as *innovation studies* carries the hallmark of Schumpeter's definition of innovation, giving it a recognisable epistemic signature (Fagerberg et al., 2013). A relatively precise definition of innovation studies is found in Fagerberg et al. (2012):

"the scholarly study of how innovation takes place and what the important explanatory factors and economic and social consequences are."

Despite its origins in classic economic thinking, this is an example "area studies", hence a strongly multidisciplinary field, i.e., *object*-driven (innovation) rather than *discipline*-driven. As a many-splintered phenomenon, innovation is akin to treatment from economics and organisational sciences perspectives and is approachable from both policy and managerial outlooks; it also lends itself to be appreciated from geographical and historical perspectives; it includes issues so contradictory such as competitiveness objectives and societal problems like unemployment or inequality. Another feature of the field is the pursuit of an empirically grounded set of research approaches (from case study methodologies to the creation of an increasingly rich toolbox of indicators, in which patents are still paramount) that are nonetheless predicated on the goal of building generalisable knowledge (stylised facts, success factors, etc.) that are, notwithstanding, able to underpin policy (Cimoli et al., 2020; Santiago, 2020). The study of innovation is constantly confronted by new issues; consequently, it is an evolving body of (actionable) knowledge in which it mirrors its subject (Fagerberg, 2018; Fagerberg and Verspagen, 2020; Louçã, 2020; Martin, 2019).

The upsurge of the interest around innovation coincided with a growing pluralism of contributions (Fagerberg and Verspagen, 2009). The innovation studies community is a "broad church" and started as an intellectual movement (a critique to mainstream economics, considered too deductivist and removed from the actual business life) but flourished in many different academic quarters (including hybrid institutional environments) (Williams, 2019). Innvation journals started to appear in the 1960s, and several achieved a very high standing (Singh et al., 2020). Surely enough, new journals were founded and were able to carve sub-specialism orientations, some more affirmative of their theoretical stance (say, evolutionary economic theory), others more business school oriented (emphasising a capability-view apt for real-world strategising), and yet others more of a critical persuasion (elaborating on constructive socio-technical appraisals of transformative change) (Martin, 2012).

Broadly capitalising on the early insights of Joseph Schumpeter, who moved across subjects such as entrepreneurship and structural change in historical time, the field expanded into a plural research programme and a global research community in the coming decades. By the 1970s and 1980s, the field had moved beyond economics as a dominant perspective (it had acquired a definite interdisciplinary hallmark), beyond its pioneering UK and US teaching and research institutions (to encompass continental Europe, Latin America, and the emerging Asian Tigers), beyond strictly academic circles (to impact the OECD, for instance) (Fagerberg, 2018; Nelson, 2012). By the 1990s, a number of journals had grown around the field (*Research Policy*, *Industrial and Corporate Change*, etc.) and the first whole textbooks became available (Freeman and Soete, 1997; Tidd et al., 1999). By the 2000s, a number of initiatives began taking stock of the lessons learned so far, this was the case with new reference handbooks being published (Fagerberg and Godinho, 2004; Hanusch and Pyka, 2007) and the first collections of articles by founding contributors with personal notes about the origins and evolution of the field (Dosi, 2000; Freeman, 2008; Pavitt, 1999). Also, by the 2000s, the first anniversaries of seminal papers appeared in the form of journal special issues/sections (Chesbrough et al., 2006; Dosi et al., 2006), which in the 2010s was compounded with books celebrating earlier path-breaking classics in the area of the economics of innovation and evolutionary change (Lerner and Stern, 2013; Nelson et al., 2018).

As with any other field, consolidation work in Innovation Studies partially takes the form of review papers. Generalisation-seeking formal reviews of Innovation Studies have appeared every now often, with at least the periodicity of once in a decade, and such works have been mainly of an interpretative kind (Castellacci et al., 2005; Dosi, 1988; Gopalakrishnan and Damanpour, 1997; Martin, 2016a; Østergaard et al., 2011). However, in the turn to the 2000s, as academic work on Innovation Studies was becoming more reflexive, i.e., providing integrative accounts of research progress, review articles became more quantitative. Leading examples were two works evaluating the impact of Keith Pavitt and reconstituting his intellectual environment: Meyer et al. (2004) , who followed his career through a citation analysis, and Verspagen and Werker (2004), who collected evidence concerning relations among innovation scholars through an online questionnaire. Later in the decade, a first integrated picture of this community of scholars emerged: Fagerberg and Verspagen (2009) combined a thorough literature-based and web-based survey approach to show the vital scientific references of the field (inspiring figures, foundational volumes, etc.) and the structure of collaboration (influential journals, conferences, academic centres, professional associations, etc.). In the 2010s, the trend toward scientometric syntheses of this expansive research area gained momentum. A special issue in *Research Policy* Fagerberg et al. (2012) adopted a particular methodology to identify the core literature in science, technology, entrepreneurship, and Innovation Studies: an analysis of the references of authoritative handbooks.

The nature of the field suggests it as a testing site for the perspective and methods advanced in this research. It is both heterogeneous and characterised by key shared attributes. Its contributors are distributed across geographies (research groups are not disproportionally concentrated in the North-Atlantic region); still, their work is continuously synchronised by a core set of journals (Fagerberg and Verspagen, 2009).

2.2.1. What are the key titles in innovation studies?

Fagerberg and Verspagen (2009) approached practitioners to identify their organisations and institutions and concluded innovation studies have grown with contributions from several thousand academics worldwide. In the absence of an overarching academic society or professional association, innovation-oriented journals emerge as the field's main communication infrastructure.

As the authors point out, the most central channel and the most established platform for scholarly work on innovation is *Research Policy*, launched in 1972 by Christopher Freeman, also the first director of Science Policy Research Unit (SPRU), established in 1966 in the University of Sussex. In their survey, this journal was found to be the preferred outlet and acknowledged as the leading journal in several other sources (Andersen and Alkærsig, 2016; Lundvall, 2014; Sáez-Martínez et al., 2019). The dominance of this journal gives a particular structure to the landscape of innovation studies, making the field akin to a "gold standard" approach in terms of analysis, that is, holding this journal as a "ground truth measure" (the reference object) against which others can be judged (Banerjee et al., 2018; Versi, 1992).

Adopting an exploratory design, we will not limit our analysis to a *convenience sample*, but will include the total population of validated and influential innovation-oriented journals. Recognising RP as an important publishing journal from the outset, Fagerberg et al. (2012) went further exploring other (19) publishing journals on innovation studies[2]. Table 1 summarises the features of the twenty top journals found in their paper. For each journal, additional descriptors were collected: categories based on subject-area(s) covered, journal's quartile, year of journal launch, and the journal publisher. Descriptors are gleaned from the publishers' official webpages and Scimago, the freely available portal that computes journal indicators derived from the Scopus database (González-Pereira et al., 2010).

Table 1: List of journals.

Journal	**Subject area[3]**	**Quartile (* if FT50)**	**Highest rank**	**Launch**	**Publisher**
Academy of Management Journal *(AMJ)*	Business	Q1*	1st Business	1958	Academy of Management
Academy of Management Review *(AMR)*	Business	Q1*	2nd Business	1976	Academy of Management
Administrative Science Quarterly *(ASQ)*	Arts and Humanities;	Q1*	1st Arts and Humanities	1956	SAGE Publications

[2] Authoritative handbooks supplied a source to explore the core literature of the field. Supporting references were recognised and used to build this list; these journals collectively accounted for almost a quarter of all citations in these large survey books.

	Social Sciences				
Cambridge Journal of Economics *(CJE)*	Economics	Q1	143rd Economics	1977	Oxford University Press
Human Relations *(HR)*	Arts and Humanities; Business; Social Sciences	Q1*	7th Arts and Humanities	1947	SAGE Publications
Industrial and Corporate Change *(ICC)*	Economics	Q1	108th Economics	1992	Oxford University Press
International Journal of Technology Management *(IJTM)*	Business; Computer Science; Engineering; Social Sciences	Q1	83rd Engineering	1986	Interscience Enterprises
Journal of International Business Studies *(JIBS)*	Business; Economics.	Q1*	3rd Business	1970	Palgrave Macmillan
Journal of Management Studies *(JMS)*	Business	Q1*	16th Business	1964	John Wiley & Sons
Management Science (*MS)*	Business; Decision Sciences	Q1*	2nd Decision Sciences	1954	Institute for Operations Research and the Management Sciences
Organization Science *(OSc)*	Business	Q1*	6th Business	1990	Institute for Operations Research and the Management Sciences
Organization Studies *(OSt)*	Business	Q1*	14th Business	1980	SAGE Publications
R&D Management *(RDM)*	Business	Q1	29th Business	1970	John Wiley & Sons

***Regional Studies** (RS)*	Environmental Science; Social Sciences	Q1	11th Environmental Science	1967	Taylor & Francis
***Research Policy** (RP)*	Business; Decision Sciences; Engineering	Q1*	4th Decision Sciences/ Management Science	1971	Elsevier
***Small Business Economics** (SBE)*	Business; Economics.	Q1	74th Economics Business	1989	Springer
***Strategic Management Journal** (SMJ)*	Business	Q1*	4th Business	1980	John Wiley & Sons
***Technology Analysis & Strategic Management** (TASM)*	Business; Decision Sciences	Q2	55th Decision Sciences	1989	Taylor & Francis
***Technological Forecasting and Social Change** (TFSC)*	Business; Psychology	Q1	28th Business	1969	Elsevier
***Technovation** (Tec)*	Business; Engineering	Q1	11th Engineering	1981	Elsevier

As "innovation" is neither one of the 27 major thematic categories nor one of 313 specific subject categories of Scopus, we analysed the areas and categories these set of journals are assigned to. As a testimony of the multidisciplinary of the field, we find the set distributed along ten different thematic categories (Arts and Humanities, Computer Science, Decision Sciences, etc.), with half of the journals allocated to more than one area (three of them even figuring in more than two disciplinary areas), which serves. Of the twenty journals, ten belong to the Business category, while the two other more common categories are Economics and Social Sciences.

Most of the outlets could be considered top ones in the sense of being located among the 25% top journals in their subject categories: 19 belong to the first quartile of Scimago's subject lists (only TASM was found in a Q2 position in the competitive area of Business). Nine journals have their best performance in the Business area. On average, journals from Business ranked the 11th position, the 125th position in Economics and the 47th position on Social Sciences. Furthermore, this set has no less than 11 titles in the Financial Times top 50 journals list[4].

Along with the increasing number of innovation studies articles published as noted by Fagerberg and Verspagen (2009), scientific journals have also experimented an exponential growth (Tenopir and King, 2014). In this sample, while most of the journals (eleven) were launched in the 1970s and 1980s some important ones were founded in the 1990s (like OSc and ICC). Behind the journals there are ten different publishing houses; all located in the West, between the US and the European bloc. Although balanced, three publishers were found more frequently, each associated with three journals: Elsevier, John Wiley & Sons, and SAGE.

2.2.2. A review of innovation studies surveys

To provide a sensitivity test of the journal listing provided in Fagerberg et al. (2012) we carried out a further exercise, which comprised two main stages. The first step is illustrated in Figure 1:.

[4] Not available at the time of Fagerberg et al. (2012), the widely used FT50 is now seen as a salient distinction of performance for business schools and researchers in general (Vidgen et al., 2019).

To start with, and to generate a snow-ball (see Cottrill et al., 1989), we looked at those articles citing Fagerberg et al. (2012) and Fagerberg and Verspagen (2009). Articles in Web of Science (WoS) covering the years spanning 2009-2019 were searched for: of 221 hits, 193 papers were retained after the removal of repeats. We searched within the title and abstract to infer whether the article was a survey and identified journals according to their "innovation studies" orientation: it was possible to determine that 171 items were not addressing innovation-oriented issues from a survey perspective. We thus considered 22 articles and read this in full. In the end, four articles were identified in which innovation-related journals were found to be discussed[5]. For these final articles Table 2 details methodologies, timeframe covered, and the key journals highlighted in the papers[6].

[5] Cancino et al. (2017) and Merigó et al. (2016) are considered as the same reference as the set of journals studied were the same.

[6] Shafique (2013) presented a global view of the field by combining longitudinal and structural perspectives based on bibliometric data from four major social science disciplines - economics, sociology, psychology and management. Also using bibliometric analysis Cancino et al. (2015, 2017) pointed out the main authors as well as journals driving forward innovation studies. In Teixeira (2013) and Rakas and Hain (2019) other journal rankings can be obtained.

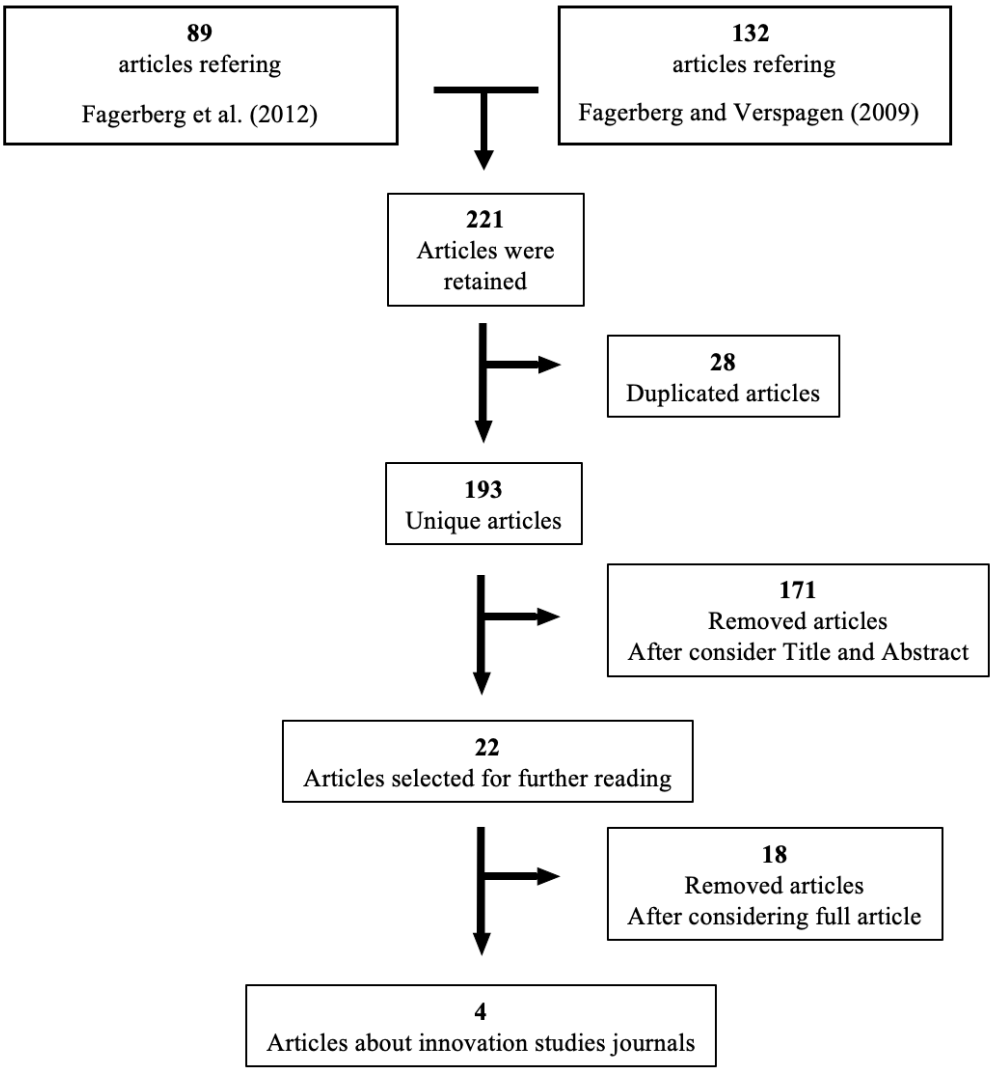


Figure 1: Reviewing surveys in "innovation studies"-oriented journals.

A further step was taken, aimed to expand the literature coverage and provide an additional sensitivity check. To ascertain if the top twenty baseline journals are the key platforms for innovation research, we considered two additional articles published before and after Fagerberg et al. (2012) and proceeded to examine their reference lists as shown in Table 2 (the last two papers). As an older record, Castellacci et al. (2005) was selected for it provides a survey of the established knowledge and emerging agenda on innovation phenomena. Kotsemir and Abroskin (2013), who did not make a reference to Fagerberg et al. (2012), addressed different aspects regarding the development of innovation concepts.

Table 2: Further studies identifying innovation-oriented journals.

Article	Methodology	Timeframe	Common journals with Fagerberg (2012)	*RP* as the reference journal Innovation studies?
Fagerberg and Verspagen (2009)	Most important journals for survey respondents.	2004 – 2005	#2: *ICC, RP.*	Yes
Fagerberg et al. (2012)	Handbook analysis (described above).	Up to 2008	Baseline	Yes
Shafique (2013)	Co-citation analysis of 50 most cited innovation papers from top 20 journals (Source WoS).	1988 – 2008	#7: *AMJ, AMR, ASQ, MS, OSc, RP, SMJ.*	Yes
Teixeira (2013)	Journals publishing more articles about National Systems of Innovation on EconLit and Scopus.	1991 – 2010	#5: *CJE, ICC, RP, TASM, TFSC*	Yes
Cancino and Merigó (2015, 2017) and Merigo et al. (2016)	Most cited journals about innovation research in WoS.	1989 – 2013	#15: *AMJ, AMR, ASQ, ICC, JMS, MS, OSc, OSt, RDM, RP, RS, SBE, SMJ, Tec, TFSC*	Yes
Rakas and Hain (2019)	Journals publishing more articles related to innovation systems in WoS.	1980 – 2018	# 8: ICC, IJTM, RP, RS, SMJ, TASM, Tec, TFSC,	Yes
Castellacci (2005)	Journals most referenced within the article.	Not applicable	#9: AMJ, ASQ, CJE, ICC, MS, OSc, RP, RS, SMJ.	Yes
Kotsemir (2013)	Journals most referenced within the article.	Not applicable	#10: ASQ, CJE, MS, RDM, RS, RP, SBE, SMJ, Tec, TFSC.	Yes

The procedure yields the patterns visible in Table 3. Overall, the summary is: the top 20 journals highlighted by Fagerberg et al. (2012) appear at least once[7]; RP is the only journal that appears prominently in all of these six sources, something that greatly buttresses its claim to the "gold standard" status in the field[8]. For the purposes of our study, these findings are significant and allow to proceed further.

Table 3: Common journals cited in different review articles.

Journal	Fagerberg and Verspagen (2009)	Shafique (2013)	Teixeira (2013)	Cancino et al. (2017)	Rakas and Hain (2019)	Castellaci et al. (2005)	Kotsemir et al. (2013)
AMJ		X		X			
AMR		X		X			
ASQ		X		X		X	X
CJE			X			X	X
HR							
ICC	X		X	X	X	X	
IJTM					X		
JIBS					X		
JMS				X			
MS		X		X		X	X
OSc		X		X		X	
OSt				X			
RDM				X			X
RS				X	X	X	X
RP	X	X	X	X	X	X	X
SBE				X			X
SMJ		X		X	X		X
TASM			X		X		

[7] In Castellacci et al. (2005) the journals in Fagerberg et al. (2012) list account for 28,6% (16 out of 56) of the total references; for Kotsemir and Abroskin (2013) that figure is 12,6% (33 out of extensive list of 262 references). Castellacci et al. (2005) refer seven journals out of the top 20 baseline, three of which many times, while RP is, by far, the most used journal, followed at a distance by ICC and OSc. Kotsemir and Abroskin (2013) refers ten journals: RP is, again, the most resourced journal, followed by MS, RS and Tec.

[8] After RP the most frequent journal was ICC.

TFSC	X	X	X	X
Tec		X	X	X

CHAPTER 3

Journal editorships

The section that follows was published in Scientometrics. Two prior versions were also accepted in the 18th International Conference of the International Society for Scientometrics and Informetrics as research article published in the conference proceedings and a poster. In addition, another version limited to South American editors was also accepted as a research article for the Latmetrics 2021.

3.1. Introduction

Academic serials (especially peer-reviewed journals) play an essential role in the research system as the main channel for spreading new knowledge (Braun, 2005; Hames, 2001; Liwei and Chunlin, 2015; Willett, 2013). As an autonomous institution, a journal is contingent on good editorial governance, usually defined by the integrity and independence of stewardship (Rynes, 2006). Elite journal board members, as "gatekeepers of knowledge" (Goyanes and de-Marcos, 2020, p. 791), "occupy key roles as opinion formers, gatekeepers and arbiters of disciplinary values" (Burgess and Shaw, 2010, p. 629). They act as mediators and influencers of academic discourse (Wang, 2018). While directing the policies and positioning of journals and choosing reviewers and having the last word on what gets published, editorship holders shape the discipline and the community (Faria, 2005; Nästesjö, 2020; Zsindely et al., 1982). This paper approaches journals from an editorial board (EB) angle and brings this perspective to the field of Innovation Studies.

Although some debate about journal governance has taken place in a few research fields, little is still known about board composition and the environment surrounding the editorial process (Bedeian et al., 2009; Horan et al., 1993; Miner, 2003; Ozbilgin, 2004; van Fleet et al., 2006). For instance, Harzing and Metz (2012) suggested editorial membership clusters around the founding nationalities, highlighting the dominance of EB members from US, with the UK coming a distant second, like many others noticed (Braun and Dióspatonyi, 2005; García-Carpintero et al., 2010). Furthermore, female under-representation on journal EBs has also been reported for fields like information systems (Cabanac, 2012), medicine (Amrein et al., 2011) and management (Burgess and Shaw, 2010). Besides, many of these and other socio-demographic characteristics have been related to the success journals are able to achieve (Harzing and Metz, 2011; Metz and Harzing, 2009). Moreover, institutional and geographic correlates of journals' scholarly impact have also been explored (Wu et al., 2020).

Despite the large literature on Innovation Studies in recent years, to the best of our knowledge, little has been written on the community of scholars acting as gatekeepers of the field. Most of the extant research falls into a couple of major categories: qualitative work of a historical orientation (Fagerberg et al., 2013) or ethnographic bent (Sharif, 2006), and quantitative ventures into the bibliometric profile of founding researchers (Fagerberg et al., 2011; Meyer et al., 2004), existing groups and networks (Fagerberg et al., 2012; Fagerberg and Verspagen, 2009), and prevailing concepts and research streams (Rakas and Hain, 2019; Rossetto et al., 2018). In this work, we stretch this reflexive work on the knowledge-base of Innovation Studies by leveraging the available, but limited, methodological understanding concerning journal EBs.

We do this by investigating the EB composition of the twenty foremost innovation-oriented outlets identified by Fagerberg et al. (2012). For this sample, we compiled the names of a total population of 2,440 individual scholars and built a complete editorship dataset comprising 3,005 editorial memberships from original, exhaustive, and triangulated search work. The contribution of this paper is twofold: first, to produce a collection of new descriptive statistics on previously unacknowledged editorial stylised facts, to establish links between editorship characteristics and journal performance metrics, and to unveil the social structures of board memberships in Innovation Studies. Second, to extend the analysis of Innovations Studies to include an editormetrics perspective and discuss the governance of journals as academic institutions.

The paper is organised as follows. In section 2, argues for considering the editorship approach as part of the scientometric enterprise. Section 3 describes the evidence, and the data collection, cleaning and consolidation procedures Section 4 presents the basic stylised facts of the editorial crowd steering the top innovation journals. Section 5 explores key patterns linking editorial features to journals' metrics. Section 6 show how networks of journal and countries' gatekeepers are structured. The paper closes with concluding notes and some prospects for further research.

3.2. An editormetrics approach

Journals are integral to the landscape of contemporary research and do frame academic activities. Only recently, however, have they been incorporated in the systematic study of science. In scientometrics, this agenda has been implemented in the analysis of the governing structures of journals: EBs. The quantitative approach gatekeeping patterns can be referred to as "editormetrics" (Mendonça et al., 2018, 2016; Pacher et al., 2021). The main purpose of EBs is to foster the quality of published articles (Cabanac, 2012), but in turn it can be shown that the scientific impact of EB members influence the perceived worth of a journal (Krishnan and Bricker, 2004). In principle, this approach

could be used as a partial, complementary indicator of the knowledge base and organisational set-up of a given academic area.

In their pioneering quantitative assessment, Zsindely et al. (1982) noted the stratified organisation of research communities and considered participation on such structures as representing "a distinctive recognition of the scientific work of the invitee." Indeed, publishers and scholars have an interest in having talented and prestigious scholars in these strategic places (Baccini and Barabesi, 2010; Bakker and Rigter, 1985). Miniaci and Pezzoni (2015) found that the quantity and quality of research, a doctorate from a reputed university, career length, but also social connections, are factors at play in editor selection. The prime role of individual editors is to determine what gets approved for publication (Bedeian et al., 2009) and in steering the research priorities of their field by, for instance, approving or promoting certain special issues or special sections (Bedeian et al., 2009). Moreover, they have the responsibility of authenticating the knowledge that enters the archives of science while keeping author and reviewer unethical behaviour at bay (Hall and Martin, 2019a). Thus, editors indirectly impact the scientific labour market through the mediation of the peer-review process (Petersen et al., 2017). However, the burgeoning diversity of topics and the increasing specialisation within fields places a cognitive burden on editors, suggesting they may be under strain to remain at the forefront and to be acquainted with the key experts so as to ensure the appropriate reviewers are selected (Biagioli et al., 2019). As such, empirical information on editorial capabilities and context has substantive potential for tackling several issues of relevance to scientometrics and science policy (Chan et al., 2005; Mazov and Gureyev, 2016; Pacher et al., 2021).

The A number of contributions have increasingly explored scientific eminence, institutional representation, international diversity, socio-demographic factors, and network relationships using these empirical sources (Morton and Sonnad, 2007; Wilkes and Kravitz, 1995). However, it remains a fact that editorial practices and journal governance are not directly observable through editorship characteristics (Baccini and Barabesi, 2014; Mendonça et al., 2018; Resnik and Elmore, 2016), neither is publisher influence on the board composition (for a rare account see (Waltman et al., 2020)[9]. Editors as journal managers also feel crushing pressure to improve publication metrics and, thus, are not immune to risks of academic misconduct (Hall and Martin, 2019a). Instances have been reported of EBs developing stratagems of pumping their titles up the rankings, from coercive citations to cross-journal citation cartels, Martin (2016b) from journal capture by an editorial click to the entrepreneurial establishment of pseudo or predatory journals (Biagioli et al., 2019). Editors have a degree of organisational power since the standards, policy and agenda of the outlet are primarily regulated by them. This work will uncover the socio-demographic and geographical-institutional patterns in this field. Editormetrics can contribute to unveil journal strategy and shed some light on what otherwise remains a rather opaque segment of the academic ecosystem.

3.3. Editors of Innovation Studies journals

In order to understand how the journals are organized, some descriptive analysis are provided below to allow some comparisons and the inference of the *modus operandis*.

3.3.1. The size of Boards in Innovations Studies

The mean size is 150 editors. Seven journals that are clearly above this level: they are the US-based, management-oriented, high-impact journals (Figure 2). There is also what seems to be a mid-sized group (where core innovation journals are located such as RP, ICC and TFSC) and a smaller group (where the innovation management specialism seems to be concentrated, e.g., TASM, IJTM, RDM). EB size variance seems to be a common stylised fact (Dhanani and Jones, 2017; Harzing and Metz, 2013; Kocher and Sutter, 2001).

[9] The limited autonomy in managing the journal's daily business along with editor appointments by Elsevier, motivated the collective resignation of the editorial community of the Journal of Informetrics who moved to launch Quantitative Science Studies as an open access journal published by MIT Press.

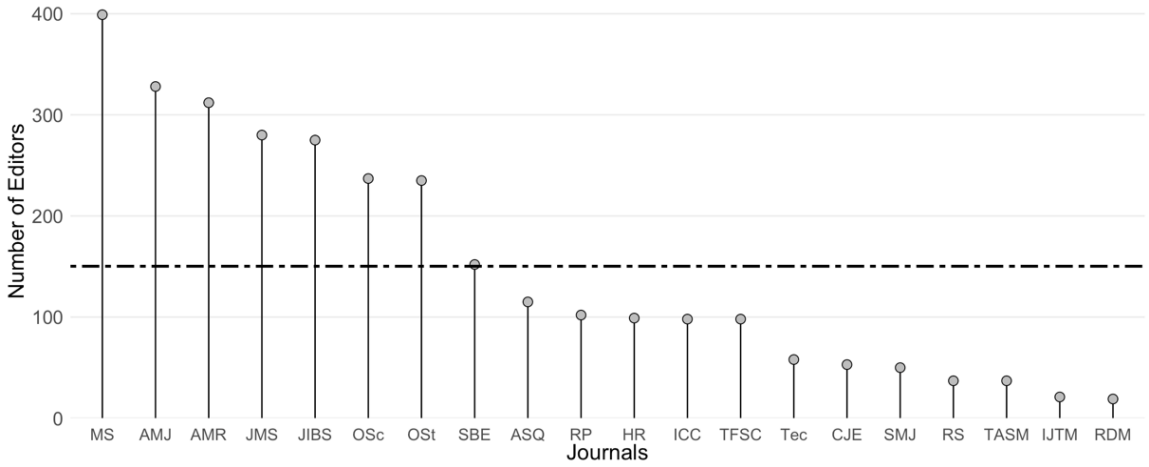


Figure 2: Frequency of editors per journal board.

Note: dashed line represents the mean.

3.3.2. Geographies of editorship

Greater international work experience is thought to provide contextual understanding, an asset essential to appraise and integrate knowledge flows (Belderbos et al., 2020). Although this attribute is difficult to proxy through the data available, all journals in the sample are nonetheless plural in terms of countries represented in their EB. Surely enough, large journals tend to encompass more proveniences while smaller journals less so. But differences can be observed that show that editorial cosmopolitanism varies, for instance, some journals do have a great internationalisation performance in relation to the size of their EBs, e.g., IJTM, RDM and Tec.

Figure 3: shows country and continent distribution of editorships. Taken together, the Global North takes most seats (North America and Europe, 88%) and the same can be said about the Anglosphere (English-speaking countries and territories, 73%), while the least represented are Latin America and Africa. In terms of individuals' countries, the US is far ahead (47.7%), followed by the UK (14.9%), with Canada (5.9%), Australia (3.1%), Italy and The Netherlands (both with 3.0%) coming next. Outside the "West" (Anglosphere and Europe), the most represented places are Hong-Kong, Mainland China, Japan, Israel and South Korea (taken together, however, these five players reach little over 4%). Scholars' location could be also described as approximating a Pareto distribution, since that 20% of nations host editors taking 85% of the editorial seats.

Figure 3: Affiliation of editorial members.

This uneven landscape is in line with other areas. For fifteen scientific disciplines García-Carpintero et al. (2010) found that 53% of EB members were located in the US with 32% coming from Europe as a whole. For a total of 246 economics journals Wu et al. (2020) confirm the overwhelming primacy of the US (48.8% of all editors), followed by the UK (11.2%). Studies of the 30 top economics journals and 16 major finance journals show EB membership to be even more skewed towards the US (Chan and Fok, 2003; Hodgson and Rothman, 1999). A clear dominance of the US, UK, Canada and Australia occurs also for 39 journals in communications and media studies (Ozbilgin, 2004). Saunders (2019) who studied information systems, found 48.5% of EB members to be affiliated to the US and 13.7% to the UK, followed by Canada and Australia, suggesting dominance by Anglophone countries. The same happens for 22 human resource management journals, with China and Japan as the only bright spots beyond the North-Western corner of the word (Ozbilgin, 2004). Our data indicate that Innovation Studies journals are not distinguished by their openness or inclusivity norms.

3.3.3. The presence abroad in the institutions

Table 4: lists the top 20 institutions by counting the number of different editorial seats occupied by their staff (a measure of "scale") and by the number of different journals in which they have a role ("scope"). The 14 highest ranked institutions in terms of sheer numbers of editorial positions occupied belong to the US. However, the dominance of the US universities is less pronounced in terms of the portfolio of different journals reached. It seems, therefore, that US editorial leadership is more characterised by depth (that is, heavily concentrated in some journals and themes) rather than breath (i.e., spread across a wide array of journals and specialisms).

Table 4: Frequency of editorships in top 20 institutions.

Region	Country	Institution	Quantity of editorships	Number of journals
Northern America	US	University of California	50	9
Northern America	US	University of Pennsylvania	49	7
Northern America	US	New York University	37	9
Northern America	US	Harvard University	36	11
Northern America	US	Cornell University	34	7
Northern America	US	Stanford University	34	8
Northern America	US	University of Michigan	33	8
Europe	DK	Copenhagen Business School	31	9
--	WO	INSEAD	30	7
Northern America	US	University of Southern California	30	8
Northern America	CA	University of Toronto	29	9
Europe	IT	Bocconi University	28	9
Europe	UK	London Business School	28	9
Northern America	US	Arizona State University	26	8
Northern America	US	Massachusetts Institute of Technology	26	5
Northern America	US	University of Texas at Dallas	24	8
Northern America	US	University of Washington	24	6
Northern America	US	Duke University	23	6
Europe	NL	Erasmus University Rotterdam	23	7
Northern America	US	University of Maryland	22	8

Not all affiliations are universities or research institutions. In addition, a small number of other entities (31) were found, such as firms (like Uber, DaimlerChrysler Corp, IGS Group, Hong Kong international Management Consulting Company, Public Policy Forecasting Inc, Saint-Gobain Group, Growth Dynamics, Infosys Technologies, and AT&T Information Systems), international organisations and comparable institutions (the European Commission, European Management Association, Center for Global R&D and Innovation (GLORAD), European Management Association, OECD, SMS Executive Office in Chicago, Economic Commission for Latin America and the Caribbean - ECLAC, African Development Bank, and the United Nations Conference on Trade and Development - UNCTAD and the Inter-American Development Bank), Foundations/Think Tanks (Fundação Getúlio Vargas, Fundação de Amparo à Pesquisa do Estado de São Paulo, Kuwait Foundation for the Advancement of Sciences) and others from the civil society and government (Council of Great Lakes Industries, Swedish Entrepreneurship Forum, Federal Government of Brazil, Embassy of Mexico in South Africa, and European Commission Joint Research Centre).

3.3.4. The gender balance

Women are a minority in editorial responsibilities (26.9%) and a considerable number of countries do not have any female editorship (15 out of 53, or 28.3% of the countries). A total of 25 countries have more male than female editors; the countries for which men are not the majority tend to have a smaller number of editors; only six countries have more women than men, and only marginally so. Figure 4: shows a breakdown along gender lines for the largest 15 countries and illustrates what is a high a significant correlation. This pattern means that countries like the US and the UK, with the largest numbers of editorships, have the largest female presence. In what innovation-oriented journals are concerned, the median for the proportion of editorial seats held by women is 24% (journal like RS and SMJ situated here), while the US-based OSC is the most balanced with 40% with the UK-based TASM on the other end of the spectrum with 8%.

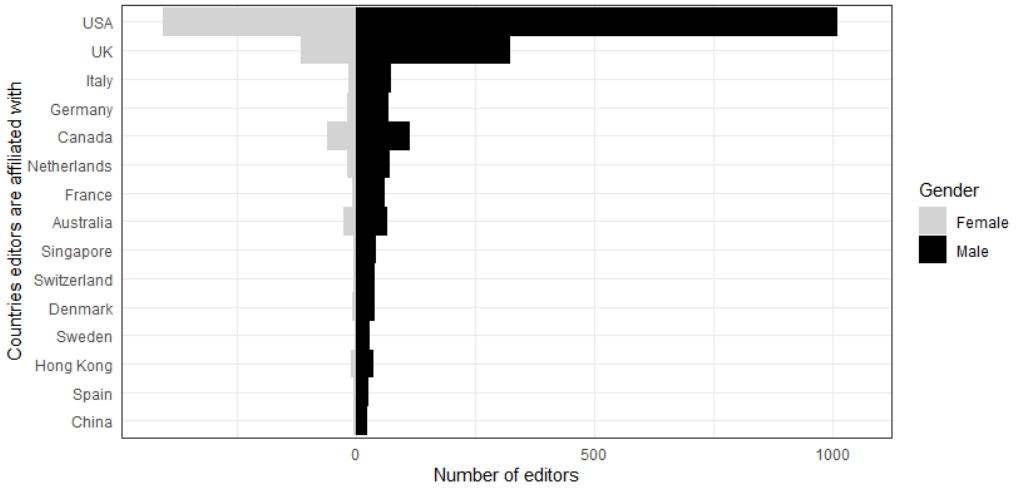


Figure 4: Number of editors affiliated to the 15 countries hosting more editors, by gender.

Gender balance in EBs attracts interest from numerous authors and has been studied in different areas (Harris et al., 2019; Sarna et al., 2020). Topaz and Sen (2016) found 8,9% of women editorships on boards from Mathematical Sciences. Mendonça et al. (2018) identified 28,2% of women editors among five leading journals on Sub-Saharan Africa. In educational psychology, Robinson et al. (1998) found only 11% of female editors and 43% female editorial board members. Later, Greenbaum et al. (2018) updated these figures and contrasted the 60% of female editors and 47% of EBs members with those registered in the past. Thus, there is a significant gap in female representation and, although women are entering boards, Metz and Harzing (2009) proved the imbalance is decreasing for the area of management and business but only slowly.

Thus, the overall picture is that EBs do not have a gender parity outlook, something that reflects the broader patterns observed in what science, technology, engineering, and mathematics (STEM) careers are concerned (Lerchenmueller and Sorenson, 2018).

3.3.5. Editorial duties

The internal modus operandi of an EB is not commonly described in general, which is what happens in this case. Brinn and Jones (2008) explained that editors typically serve in one of three prominent roles: Editor-in-chief who is responsible for overseeing the editorial process and policies across the journal; the managing editor who monitors the workflow of manuscript submission and final publication details; and the associate editors who are responsible for receiving manuscripts, allocating reviewers, evaluating the reviews, and making recommendations for or against publication.

Among different journals a substantial variability of role labelling was found. For instance, there are journals with only one title for all editors, such as HR where everyone is "editor" while MS exhibited 35 different titles. According to the titles assigned, different internal organisations can notionally be inferred: some journals can have more sophisticated organisational set-ups and more defined internal structures, sometimes including regional specialisation, than others. Table 5: provides the summary of the editorial duties, organised by us in broad descriptors according to their presentation on official websites.

Table 5: Editorial labelling in the journals.

Role family	Roles found
Associate Editor	Associate Editor, Associate Editors.
Book Review Editor	Book Review Editor, Book Review Editors.
Copy Editor	Copy Editor, Copy/Production Editor.
Editor-in-Chief	Editors-in-Chief, Editor-in-Chief, Editor in Chief, Editor-In-Chief, Editor, Managing Editor.
Editor	Editor, Editors, Co-Editors, Members, Editorial Board, Editorial Board Members, Editorial Review Board, Editorial Advisory Board.
Advisory Editors	Advisory Board, Advisory Editors.
Senior Editor	Senior Managing Editor/Licensing Services Manager, Senior Managing Editor.
Regional Editor	Regional Editor, Continental Europe Editor, North America Editor, United Kingdom And Scandinavia Editor, Regional Editor China, Urban and Regional Horizons Editor, Regional Editor Far and Middle East.
Region Editorial Manager	Region Editorial Manager, Continental Europe Editorial Manager, United Kingdom And Scandinavia Editorial Manager.
Area Editor	Area Editor, Big Data Analytics Editor, Accounting Editor, Business Strategy Editor, Decision Analysis Editor, Entrepreneurship and Innovation Editor, Finance Editor, Healthcare Management Editor, Information Systems Editor, Marketing Editor, Operations Management Editor, Optimization Editor, Organizations Editor, Revenue Management and Market Analytics Editor, Stochastic Models and Simulation Editor.
Area Associate Editor	Area Associate Editor, Accounting Associate Editor, Business Strategy Associate Editor, Decision Analysis Associate Editor, Entrepreneurship and Innovation Associate Editor, Finance Associate Editor, Healthcare Management Associate Editor, Information Systems Associate Editor, Marketing Associate Editor, Operations Management Associate Editor, Optimization Associate Editor, Organizations Associate Editor, Big Data Analytics Associate Editor, Revenue Management and Market Analytics Associate Editor, Stochastic Models and Simulation Associate Editor.

Special Topics	Commercializations, Emerging Economies, IT and National Policy, Marketing and New Product Development, Strategy and Innovation, Sustainable Development Innovation, Technological Entrepreneurship, Technology Management in Industry.
Patrons	Founding Editor, Founding Editor and Editor-in-Chief 1969-2010, Patrons, Former Editors-in-Chief.

We found 44 scholars taking multiple duties inside the same journal, remarkably something that seems not to have been pointed out in the literature so far. In addition, we observed 41 editors having two simultaneous appointments and three with three. This phenomenon happens on 8 journals (AMR, ASQ, CJE, JIBS, MS, OSt, SBE and TFSC) and no rationale for it is offered. Figure 5: displays the journal distributions of editors with one, two and three editorship duties within the same journal.

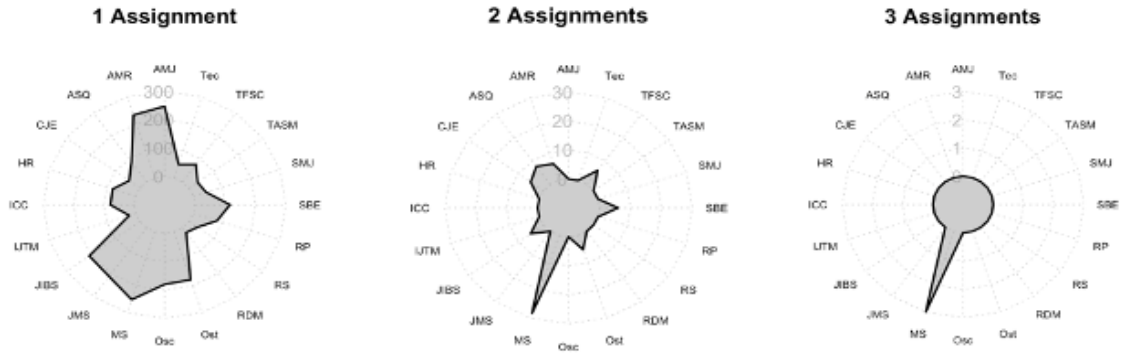


Figure 5: Number of editors with one or multiple duties in the same journal.

3.3.6. The inter-journals shared editors

The database comprises 3,005 memberships and includes 43 seats occupied in the same journal by the same people; this we will disregard for the remainder of this paper. Hence, the individuals in our database occupy 2958 editorial positions in one or more journals. Of the total 2,440 unique editors 448 (18.4%) are "serial" editors, i.e., they appear repeatedly across journals). These serial editors are spread over all the 20 journals, i.e., every journal has at least one editor of this type. Most of the serial editors (324, or 72.3%) are present in two EBs while the others in more (we call them "super-editors" for simplicity): 73 are in three, 12 in four, and 3 are in five journals. Amid serial and super editors, we found 120 women serial and 21 women super editors. Serial editors come mostly from the US (212) and the UK (68), meaning that these countries exhibit a great proportion of serial editors in their total stock of editors: 18.7% and 18.8%, respectively. It should also be noted that their reach is extensive: US and UK serial editors are spread across 16 and 18 different journals, respectively. In what super-editors are concerned, most of them come also from the US (49) and the UK (11). The extreme situation

of a single individual to occupy five editorships is an exclusive of these countries: one is American, two are British. The US and UK-based editors are thus serious accumulators of editorial powers.

Table 6: shows (left-hand side) that the proportion of serial editors in the journals ranges from 2% (in CJE, the most isolated journal) to 56% (ASQ and SMJ, the most inclusive of networked editors). It also shows (right-hand side) the 30 countries with serial editors and the sub-set of 18 countries with super-editors.

Table 6: Frequency of editors and shared ones per journal.

Journal	Total individual editors	Shared editors	Shared editors (%)
AMJ	328	145	44%
AMR	310	149	48%
ASQ	111	62	56%
CJE	52	1	2%
HR	99	26	26%
ICC	98	38	39%
IJTM	21	2	10%
JIBS	274	66	24%
JMS	280	99	35%
MS	365	21	6%
OSc	237	117	49%
OSt	234	82	35%
RDM	19	3	16%
RS	37	3	8%
RP	102	38	37%
SBE	150	14	9%
SMJ	50	28	56%
TASM	37	12	32%
TFSC	96	12	13%
Tec	58	12	21%
Total	**2958**	**930**	

Country	Serial editors*	Super editors**
US	212	49
UK	69	11
Canada	33	9
Italy	18	5
Denmark	12	1
France	11	2
Australia	11	1
Germany	7	3
Netherlands	7	1
Singapore	7	0
Switzerland	6	3
Hong Kong	6	2
China	5	0
Finland	4	0
Sweden, Japan	3	1
South Korea	3	0
Spain	2	2
Russia, Israel	2	1
Taiwan, Austria	2	0
Peru, New Zealand	1	1
Belgium, Brazil, Cyprus, India, Ireland, Norway	1	0

Evidence for multiple memberships and serial editors has been reported in other fields. For instance, in the accounting field, Brinn and Jones, 2008 identified two individuals assuming editorial duties for six journals simultaneously. Chan and Fok, 2003 found three scholars with eight, a maximum number of memberships in the field of international business.

3.4. Further patterns on the editorial boards

3.4.1. The correlates of EB diversity

The journals from our sample were founded in the 30 years that span between 1964 and 1996. In relating the founding years of journals to their EB diversity we generate the patterns displayed in Figure 5. No link is found between journal cohorts and geographical diversity (Figure 6A); that is to say, younger journals are no more cosmopolitan than older journals. This pattern seems to corroborate what Shaw et al. (2016) witnessed for the information systems research field; using a dynamic perspective, geographic diversity remained unchanged over time. Likewise, no clear connections emerge between journal cohort (the pattern is only less linear, Figure 6B); notwithstanding, while the two older journals display little above 20% of female presence, the two younger journals have little below 40% of female editors in the EB. Dickersin et al. (1998) concluded for epidemiology that newer journals had a higher proportion of women editors on EBs.

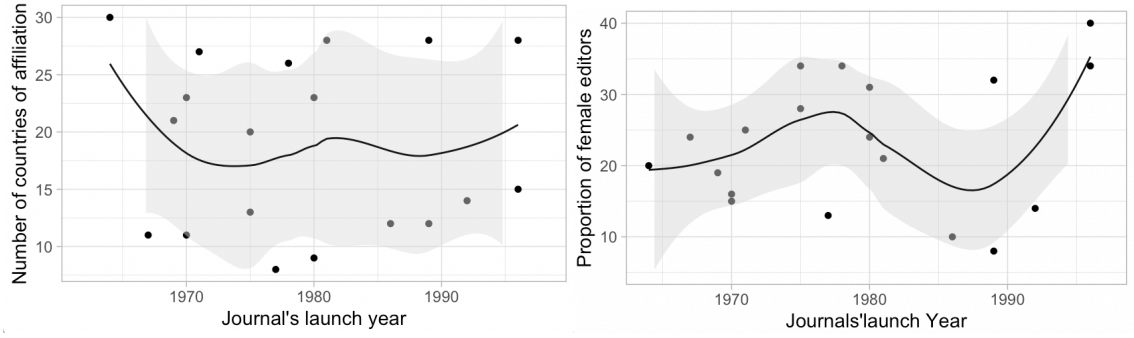


Figure 6: Journals' cohorts and diversity: A) geographical; and B) gender.

We also explored the association between the board size and diversity, both in terms of geographical dispersion (Figure 7A) and gender parity (Figure 7B). Figure 7A shows that journals with large numbers of editors do not necessarily embrace global compositions; there seems to be a plateau, especially for big journals (EB size above 100), with possibly even a negative relationship for very large journals (EB size above 300). Similar evidence was also compiled by Tutarel (2004) for management journals. (Wu et al., 2020) also noticed that the number of editors from different continents was not higher for the largest boards in economics. Figure 7B depicts what seems a pronounced "inverted U" relationship between EB size and gender representation. This result is consistent with the idea of a "glass ceiling" limiting women presence on EBs (Hymowitz and Schelhardt, 1986). This phenomenon is found in different disciplines, such as medicine (Amrein et al., 2011; Hafeez et al., 2019) and maths (Topaz and Sen, 2016).

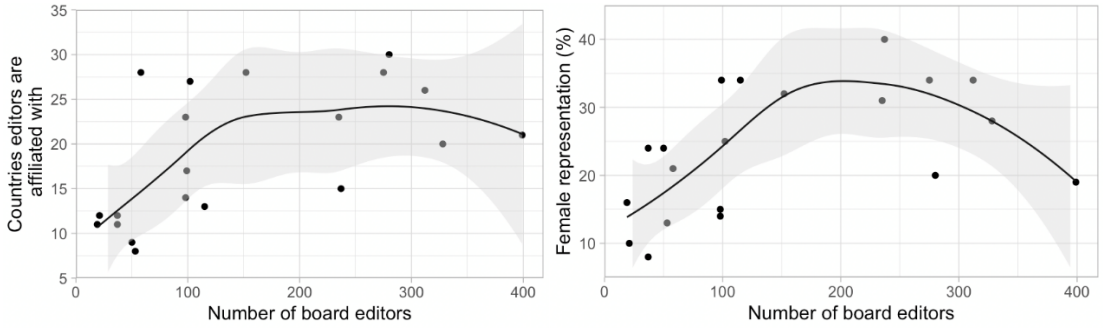


Figure 7: Board size and diversity A) geographical; and B) gender.

3.4.2. Drivers of journal impact

In order to understand socio-demographic characteristics of EBs (geography and gender) might have on journal achievement, we studied the association of both variables with the journals' H-index. Figure 8A shows a positive association with geographic diversification. There is evidence on the importance of diverse backgrounds among editors, but not all kinds of diversity are alike in terms of impact (Østergaard et al., 2011). Nisonger (2002) also found an overall positive association in a number of areas including business and political science and argues that a heterogeneous EB's international composition is a marker of an higher impact journal. In the economics field, Wu et al. (2020) also found a significant correlation between affiliation diversity and the five-year impact. Figure 8B finds a mostly positive, although wobbling, relationship between gender representation and impact. For EBs with

over 30% of female editors the journal quality is conspicuously not affected. A general positive association was noticed by Metz and Harzing (2009) in the field of management.

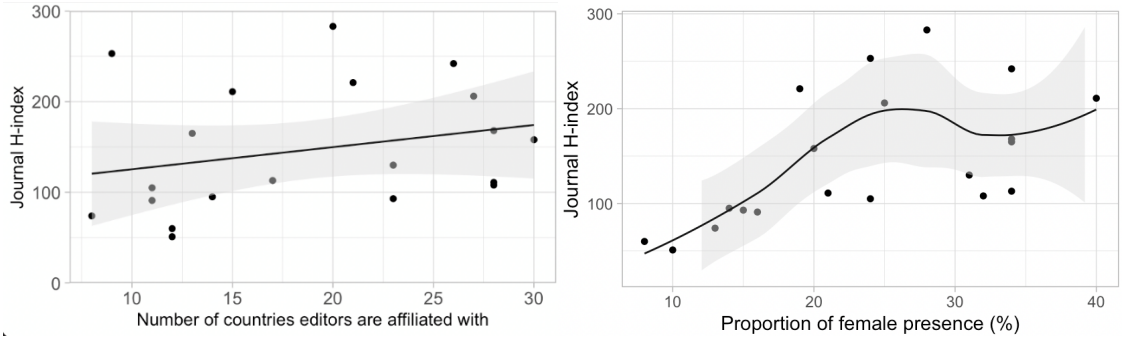


Figure 8: Country representation (A), female presence (B) and journal performance.

A final variable set against journal performance is the proportion of networked editors: Figure 9 finds an overall positive association. This much was observed by Aguinis et al. (2012) in the field of management. These authors argued that editors with multiple memberships tend to be more embedded in the science system, its norms, and know-how. For these reasons, high impact journals compete for experienced scholars and try to hoard reputed editors for their EBs. As Baccini and Barabesi (2010) did, it is also worth noting that when scholars become members of the EBs of more than one journal there are connections created between the journals in a way that is reminiscent of "interlocking directorates" in the corporate world. These inter-organisational ties constitute potentially important conduits of communication and epistemic proximity between the journals. Rather than being a spurious or trivial pattern, economic sociologists see here a central characteristic of modern social structures and explain this phenomenon as the result of strategic decisions and ameneable to analysis through network theory (Mizruchi, 1996). The analysis of social networks generated by the presence of shared editors in different journals is one important, albeit still relatively underexplored, way to study the contemporary informal dynamics of science as an institution (Baccini et al., 2020).

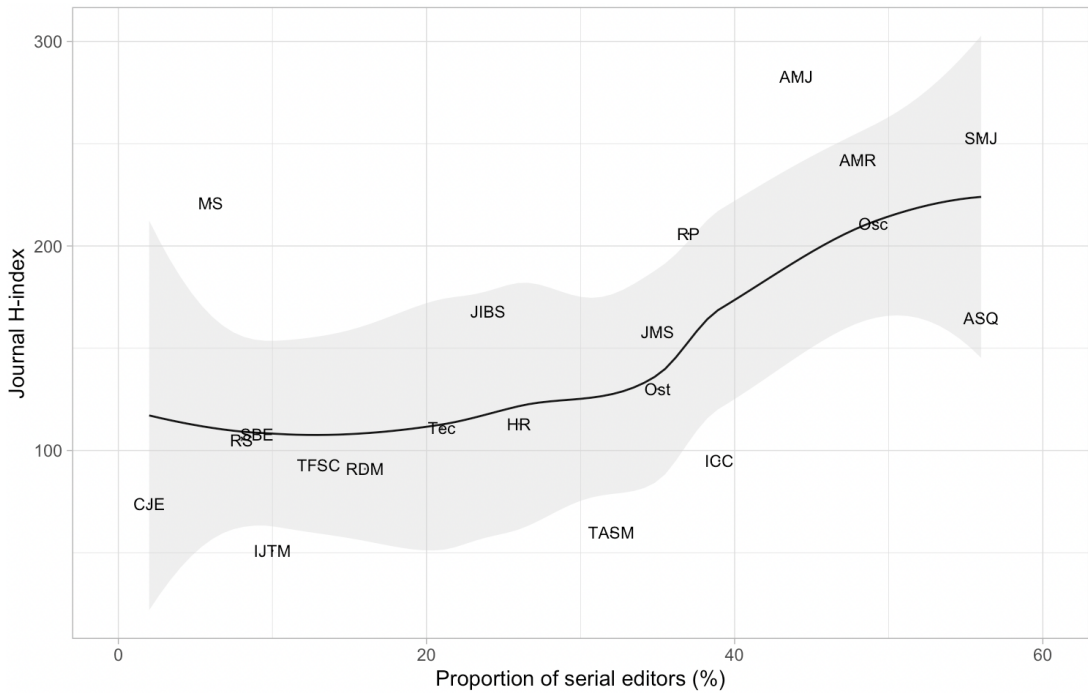


Figure 9: Serial editors and the journal performance.

3.5. Innovation Studies editorial networks

3.5.1. The editors: Analysing the connections

In modern science governance literature, editors demonstrate role legitimacy through high academic standing (publication profile and performance) and further signal their intellectual and social capital through EB affiliations (Vandome, 2013; Wright, 1985). As Powell and Grodal (2006) pointed out in relation to innovative activities, the heterogeneity of team members profiles, provides a wide stock of cognitive resources. In this sense, the editorial process is a particular type of knowledge work and bridging activity for sustaining professional communities. Relational evidence can be captured in a number of ways; for instance, editors not only belong to formal groups (EBs) but as they interact with fellow editors in other journals so that their cross-cutting connections reveal latent informal network structures that can be studied (Vespignani, 2018). A social network was defined by Wasserman and Faust (1994a) “as a finite set or sets of actors and the relationship or relationships between them.” Social Network Analysis (SNA) unpacks structures in terms of nodes (or vertices) and ties (edges).

To address the boards’ structure and dynamics, SNA can be applied. A network is defined as a "collection of vertices joined by edges" that can represent the pattern of connections between various objects (Wasserman and Faust, 1994a). In EB settings, we can consider the actors as the journals, and

the ties are the editors shared between journals. This relationship is characterized by a common presence of same scholars in multiple boards without directionality (i.e., for each interaction both are senders and receivers), which make an undirected network (i.e., the links do not have a direction).

Figure 10: is a bipartite graph that displays two different groups: the scholars in the database (small dark grey dots for women, light grey for men) and the journals they are attached too (the larger, tagged circles); connections to the journal represent EB memberships. Sharing editors brings journals together and to the centre of the picture. By contrast some journals appear more isolated, i.e., their editors tend not to occupy EB seats elsewhere: RS, IJTM, CJE, RDM. These journals share few members with few other journals, for instance, CJE shares only one scholar with a single journal (JIBS).

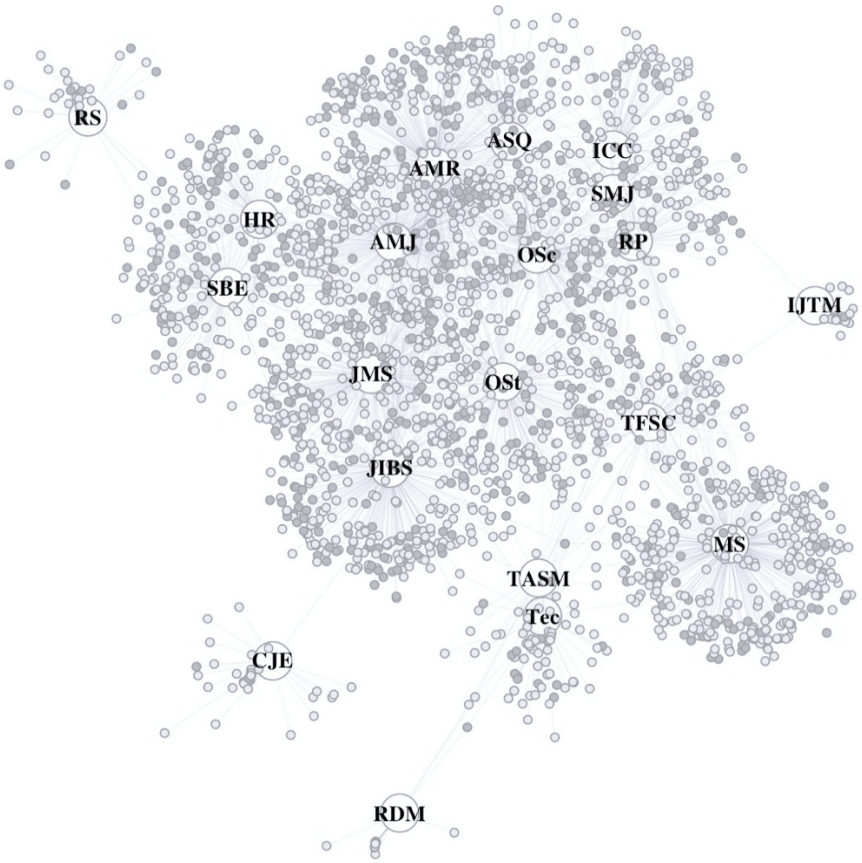


Figure 10: Bipartite graph of the editorial network in Innovation Studies.

Considering only serial editors, the ones with multiple editorials jobs within our sample, F emerges. Each little dot is a serial editor, and always has more than one link connecting to journals. It is possible to see pairs of journals sharing more editors, depicting a larger number of linkages to the same editors. Pairs of journals like AMR and AMJ, JMS and OSt, RP and ICC have many editors in common.

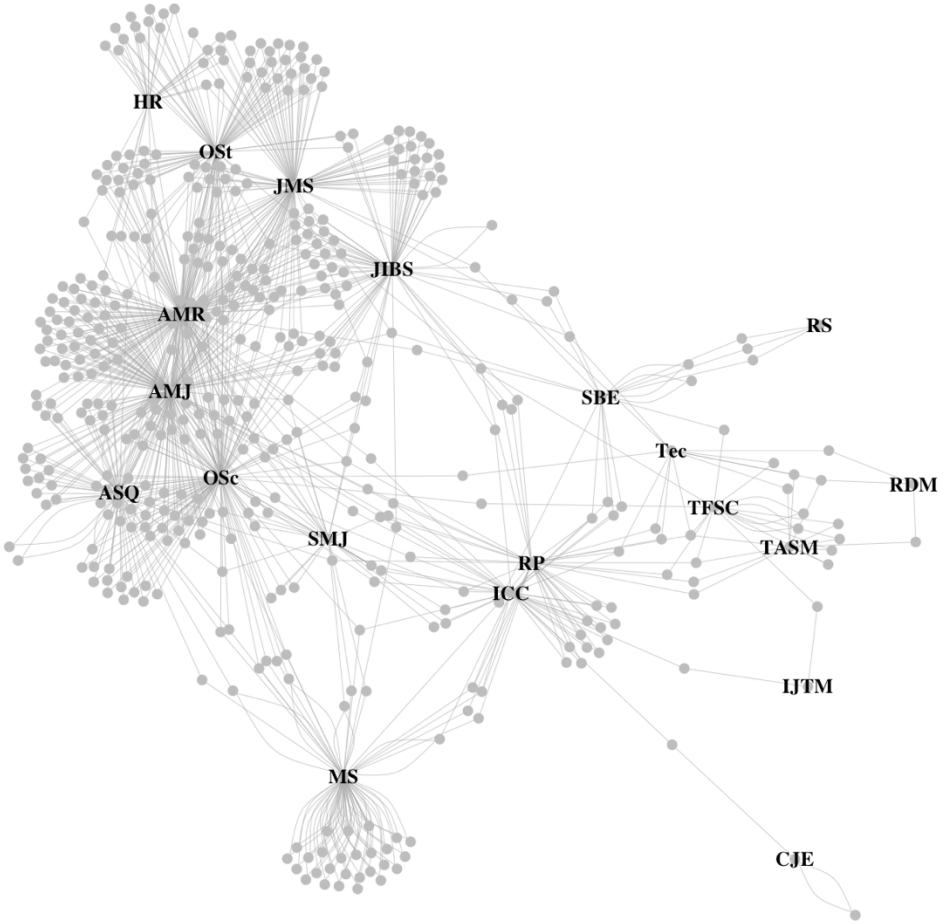


Figure 11: Bipartite graph of the editorial network in Innovation Studies, serial editors.

A giant cluster is evident in the top left corner comprising American editorial houses like ASQ, AMJ, AMR, HR, OSc, OSt and JMS. This dense cluster suggests a high number of editors are shared between them, which could be geographically explained by the common location of the journals, the US.

The full network journals include 2.440 editors who are linked to the 20 journals via 3.005 links. The density of the interlocking editorship network (the ratio of the actual number of connections to the maximum number of connections in the network) is 6%. This means that each editor could be linked to all 20 journals at the same time, resulting in a network with 48.800 connections, but only 3005 links were discovered, representing a density of 6%. This figure is significantly lower than the 14% found by Baccini et al. (2009) for statistical journals.

If only journals and their connections are considered, the density of such a network is 41,6%. This means that each of the 20 journals could be linked to the other 19, allowing 380 connections to be established and 157 links were discovered between journals. The following section will go over the relationship between journals in greater depth.

3.5.2. The interlocking editorship network: Journal centrality

A key purpose in SNA is to distinguish between the most central and the peripheral components of a network. In our case, the goal is to perceive the journals in a central position, with greater power or influence. For this reason, the original network was projected into a bipartite network. EB networks can be conceptually understood as web of journals (the large vertices) linked (edges) to the editors (small vertices) they share. Centrality reveals the status of an entity in a network and, as suggested by Wasserman and Faust (1994a), can be unpacked in three main measures: Degree, Closeness and Betweenness. A complementary measure, such as PageRank, can be computed with algorithm used by Google Search to rank web pages measuring their importance (Page et al., 1999).

The simplest measure for the centrality of a journal in an editorial cobweb is represented by the overlap among EBs. Degree centrality, hence, is the number of direct edges a journal has to other journals measured by the editors they have in common (Bonacich, 1987; Nieminen, 1973; Shaw, 1954). It measures journals' ability to access, share knowledge or other resources, and, ultimately, influence the wider network. Thus, the more ties a journal has to other journals, the more connected it will be and the more central its position in the network (Borgatti and Foster, 2003; Wasserman and Faust, 1994b).

Closeness is proxied by the distance between a journal and all the other journals. This is computed by the shortest paths between all nodes (BEAUCHAMP, 1965; Sabidussi, 1966). A journal is central in this sense if members in its EBs can quickly interact with all the other EBs. The more direct and indirect connections a given journal has with others the stronger the pull will have in the network. A journal informational potential passes through multiple other journals to reach any end journal take longer and is slower than those that interact directly. In other words, the number of editors the message must pass through before reaching its destination represents a cost that rises as the number of editors in the path grows. Thus, journals occupying a central location are best placed to reach the entire network most efficiently.

The idea behind Betweenness is that communication between two nonadjacent journals might depend on other journals in the network, especially on those outlets lying on the paths between the two. The number of times a given node lies on the shortest path between other nodes highlights how it acts as a "bridge" in the network (Freeman, 1979; Hansen et al., 2020; McKnight, 2014). This proxies for an ability to intermediate information, govern the resource flow and bring together otherwise disparate parts of the system.

Finally, PageRank can be computed with algorithm used by Google Search to rank web pages measuring their importance (Page et al., 1999). Each node in a network is assigned a score based on its number of incoming links (its 'indegree'). These links are also weighted depending on the relative score of its originating node. In Table 7:, centrality measures calculated using igraph package from R are provided for all the outlets in the network to identify the most powerful top-tiers in innovation studies.

Table 7: Journals' centrality measures.

Journal	Degree	Closeness	Betweenness	PageRank
AMJ	12	0.037	11.35	0.069
AMR	7	0.037	11.35	0.069
ASQ	9	0.030	2.65	0.053
CJE	1	0.022	0	0.013
HR	6	0.027	0	0.037
ICC	11	0.037	61.02	0.073
IJTM	6	0.024	0.80	0.019
JIBS	11	0.037	9.58	0.064
JMS	10	0.033	12.85	0.059
MS	9	0.033	3.01	0.054
OSc	13	0.038	21.71	0.075
OSt	9	0.031	6.09	0.053
RDM	2	0.021	0	0.020
RS	1	0.021	0	0.013
RP	13	0.040	36.71	0.077
SBE	9	0.033	40.81	0.060
SMJ	9	0.032	2.40	0.053
TASM	4	0.027	6.28	0.033
TFSC	7	0.031	20.72	0.049
Tec	8	0.032	36.69	0.055

Note: the three highest scores for each metric are highlighted.

Within the network, RP and OSc come ahead as the outlets with the highest Degree, i.e., with more connections of individual editors to other journals. RP showed also to be the journal with the highest Closeness to other journals, with OSc as the second again. As with Betweenness, ICC proved to be the most mediating of outlets, with SBE a far second. Regarding RP and ICC: it should be noted that RP comes third in Betweenness, and also that ICC exhibits the third highest values for Degree and Closeness (at par with management journals in each metric). Overall, and as PageRank also shows, these findings consistently highlight RP and ICC as the most central outlets in Innovation Studies.

Figure 12: provides an integrated picture of the network centrality evidence. Edges illustrate the common editors shared between each pair of journals, so wider ones mean higher number of common scholars. The size of nodes illustrates journals' Closeness (i.e., the average shortest distance) to all other outlets. Larger journals are those with the shortest average distance to other outlets, while the smaller ones are those with largest distance to the others. A grey scale is applied to illustrate Betweenness centrality measure, where darker colours represent journals linking higher numbers of other non-adjacent outlet pairs.

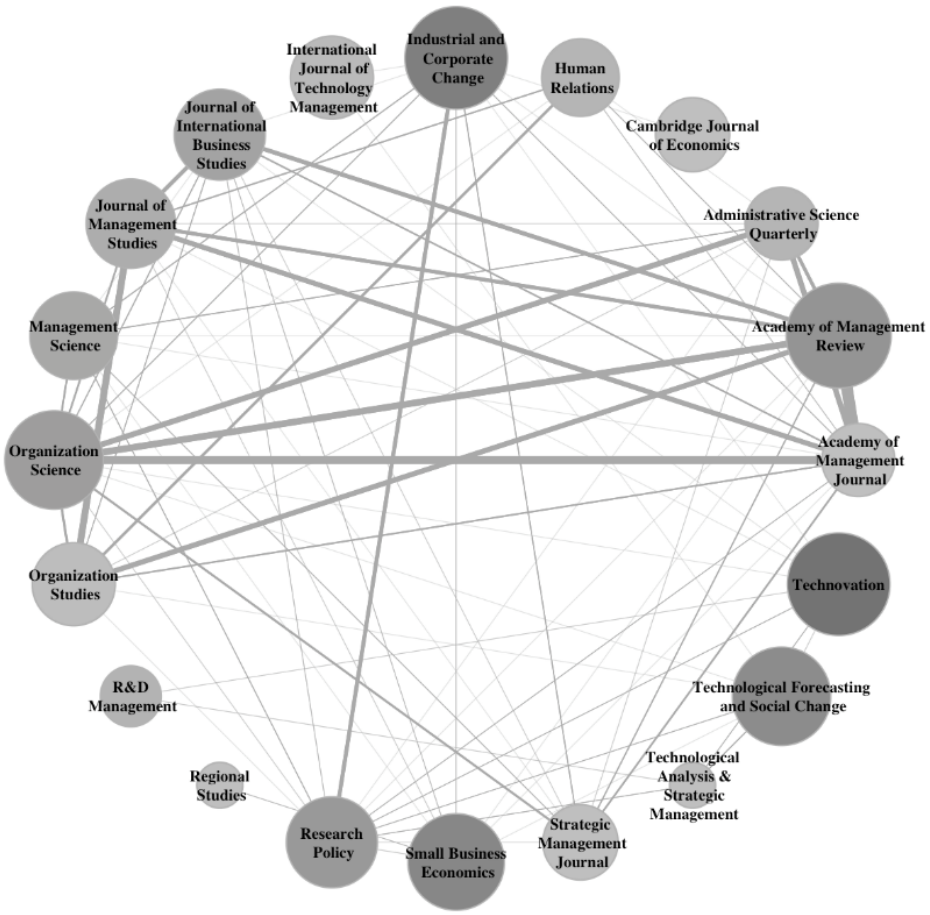


Figure 12: Projection of journals' centrality measures.

The largest nodes are RP and RDM. with the highest closeness scores and shortest average distance to other nodes. Occupying such a position may suggest they are a reference for other outlets. RS and RDM are the smallest nodes connected to the whole network by only 1 and 2 shared editors, respectively. With the highest Betweenness score, ICC is the darkest node responsible for facilitating the communication between innovation-oriented journals. The large number of editors ICC shares with other outlets and its interdisciplinary nature makes communication exchange possible between non-adjacent journals. As no journals presented a zero degree, all of them have at least an edge connecting to at least another journal.

All these centrality measures were also applied to the network structure of EBs by Andrikopoulos and Economou (2015) to identify the sources of major influence in Finance. Likewise, Goyanes and de-Marcos (2020) identified the editors, geographic areas and institutions with the highest centrality in Communication Sciences. Teixeira and Oliveira (2018) mapped the most influential journals in Knowledge Management. Finally, Baccini and Barabesi (2011) also ranked journals in Information and Library Sciences based using centrality measures.

3.5.3. Editorial communities: Country perspective

Editors share their knowledge on the EBs but also in the countries they live. Barabási and Albert (1999) have shown real networks are not random graphs because they have great heterogeneities, revealing a degree of organisation. Considering the editors' countries of affiliation, we applied a multi-level modularity optimisation algorithm for finding community structure. Based on modularity, a measure of community presence, we aimed to identify the community structures in the network with dense connections between the nodes within communities but sparse connections between nodes in different communities. Fortunato (2010) explained communities as "groups of vertices that probably share common properties and/or play similar roles within the graph". In our study, communities are those countries hosting editors.

To reach an optimised modularity the "Louvain algorithm" was applied (Blondel et al., 2008). With it, each vertex is moved to the community with which it achieves the highest contribution to modularity. When no vertices can be reassigned, each community is considered a vertex on its own, and the process starts again with the merged communities. The process stops when there is only a single vertex left or when the modularity cannot be increased any more in a step. Even though, the modularity achieved was 0.05329, a small value alerting for the number of connections between the nodes within a community would be a few higher than the connections between nodes in different communities.

F depicts the three communities of countries. The network consists of countries that are connected the co-presence of editors on the same EB. The 53 countries are found to be placed in three communities: 1) there is what appears a central group with 18 countries, in yellow, where countries like the US and UK are situated; 2) a large group, in blue, encompassing 24 countries, in which many European countries are found; 3) and a smaller group of 11 countries, in green, where many Asian countries appear.

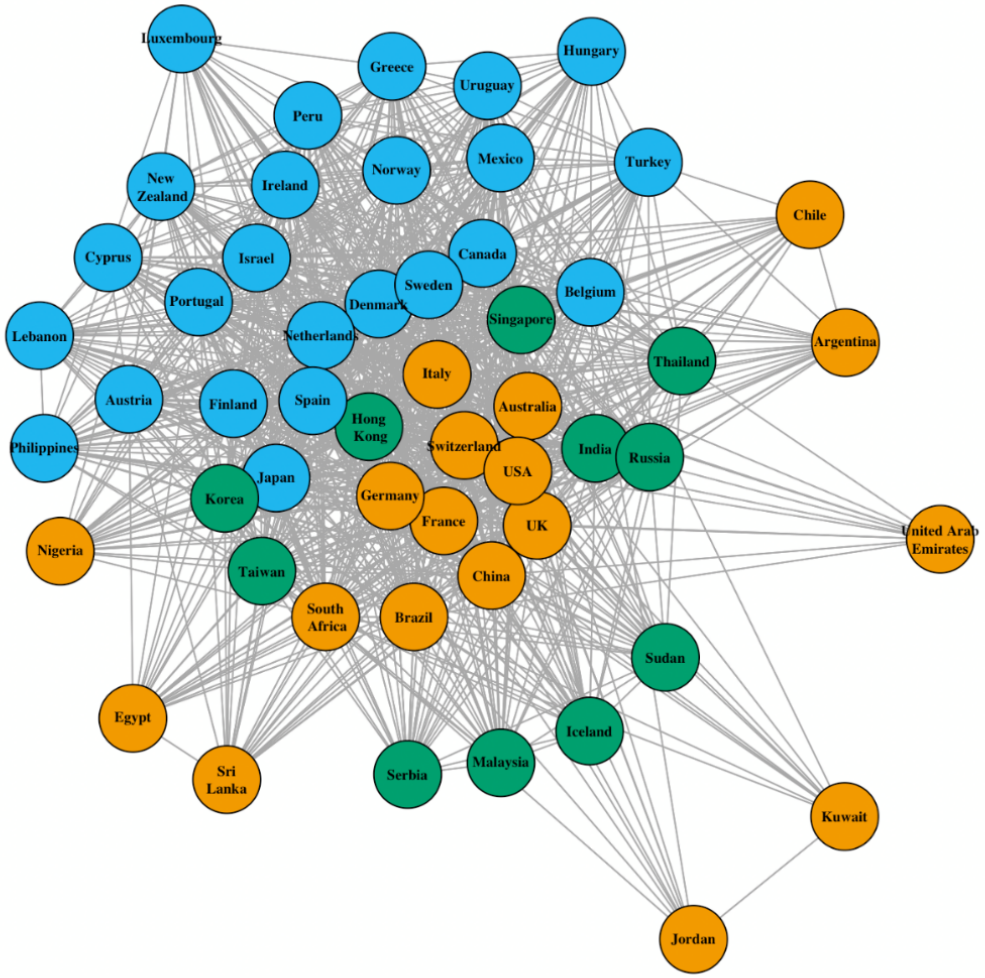


Figure 13: Communities of countries found among the Innovation Studies editors.

Running the usual centrality metrics, it emerges that the countries scoring the highest values (for Degree, Betweenness and Closeness) are indeed in the "yellow" community: US, UK, Australia, France and Switzerland. However, this group includes a few countries with a limited number of editors and low centrality scores, such as Kuwait and Jordan, which have few editors and seem only to appear on EBs whenever the most central nationalities are there.

3.6. Conclusions

Editorships are a central feature in modern, journal-based, professional research communication. This paper offers a snapshot of the editorial characteristics and structures in 20 leading Innovation Studies journals. Three empirical goals can justify such a project: to uncover the socio-demographic and geographical-institutional patterns in this field; to highlight stylised facts linking editorial features to journal performance; and, finally, to provide an outline of the social organisation of the editorship communities in this area of research. There are two additional goals: a methodological one, which is stretching the efforts regarding the mapping and measurement of Innovation Studies as a field, and a science policy one, which relates to the implications concerning the governance of journals as academic institutions.

First, editorships in Innovation Studies are heterogeneous. The editorial corps in the journal sample ranges from a few dozen to hundreds. Using a new dataset of 2,440 individuals and 3,005 different EB memberships, a few gaps are revealed, namely in terms of gender and geography: this is a male-dominated and AngloAmerican-based editorial elite. It is also a concentrated phenomenon, as 20% of the countries hold 85% of the editorial positions. A lopsided distribution is also apparent from an institutional perspective, as the top-20 universities with more say at the journal table are all located in seven countries from the northern hemisphere: US, UK, Italy, Netherlands, Canada, Denmark and Singapore. This outlook is compounded in what editors-in-chief are concerned: Only two countries hold more than two of these positions, the US and the UK.

Second, linking the editorship characteristics to the outlets' broader record also yields an image on the correlates of scientific standing. The inclusion of editors seating on multiple boards, more diverse national representation, and a greater proportion of women on EBs seems to be positively associated with journal performance. The importance of identity factors is likely to be non-linear, in particular, there is only such an extent to which gender balance is linked to journal impact metrics. Journal organizational indicators show that neither younger nor larger journals display greater diversity.

Third, the editorial environment is seemingly made of people that interact with each other across journal boundaries and this matters for developing a fuller picture of the forces shaping the course of Innovation Studies. Serial editors are almost 20% of the editors, thanks to them all the journals are connected, and their co-location on different boards builds up a journal network with properties that are worth analysing. The common centrality measures (Degree, Closeness, Betweenness) reaffirm Research Policy as the prime reference in Innovation Studies, with Industrial and Corporate Change coming second.

Our study has a few limitations and leaves open the possibility that some patterns can be better described, revised, and explained. To get the most up-to-date and accurate data, future research may consider conducting a questionnaire survey to editors highlighting questions about recruiting, actual work, and the level of perceived impact of their role in the field. The journal list in Fagerberg et al. (2012) can be made more customised to innovation-oriented journals[10] as well as updated[11].

[10] Management, organizational and strategy journals could be removed while others from sociology and innovation critical studies could be added (see Santos and Mendonça, 2021a, 2021b)).

[11] Journals that are on topic but absent for some reason could be included. Titles could be harvested from the 40 titles in the field of Innovation Studies ranked on the ABS list (the last one published by the Chartered Association of Business Scholars, Academic Journal Guide 2021) and derived from systematic surveys (e.g. Rakas and Hain, 2019; Kashani and Roshani, 2019). This effort could see the inclusion of outlets like *Industry and Innovation*, the *Journal of Technology Transfer*, *Science and Public Policy*, *Foresight and STI Governance* and the *Journal of Innovation Management*. The last one, for instance, was launched in 2013 (that is, after Fagerberg et al. (2012)).

Innovation Studies is a reflexive field (Morlacchi and Martin, 2009) and proved to be a fruitful object for studying the usefulness of Editormetrics as a methodology that can add to the stock of knowledge. The interest on the Editormetrics as a complement to bibliometrics has resulted in certain changes so far. Elsevier journals have already reported editors' gender information[12], and Cell Press has established an "Inclusion & Diversity" statement as a path to parity[13]. Regional Studies employed an open method for hiring editors with a statement regarding their search for a new editor[14].

The business of editorship is indeed displaying more entrepreneurial traits recently. This editorial activism is related to new strategies in developing novel "product lines" and in "promoting" their content. Research Policy occasionally publishes "Discussion Papers" on themes that are deemed relevant and the editors commission responses to a leading article in the form of commentaries with the different views. Industrial and Corporate Change provide a list "editor's choice" articles and make them free to read. Technological Forecasting & Social Change allows for occasional opinion essays called "From my perspective" which are not standard research papers. Management Studies created an "Insights Blog" to give attract attention of particular pieces of published research, highlighting the practical and societal consequences. Human Relations gives the "Paper of the Year Award" to the article that the EB considers that best encapsulates broad readership appeal and sound methods. All these editors' initiatives highlight the importance of understanding the "who's who" in journal gatekeeping processes as a source of information about the structure and strategy of the contemporary research ecosystem.

[12] The announcement was made on February 11th, 2021 and it accessible through: https://www.elsevier.com/about/press-releases/corporate/elseviers-journals-now-displaying-editors-gender-in-support-of-diversity

[13] Available on https://www.cell.com/cell/fulltext/S0092-8674(20)31689-5?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0092867420316895%3Fshowall%3Dtrue

[14] The announcement is made available on journal's website: https://www.regionalstudies.org/news/editor-rs/ . An outlet like the *Journal of Hazardous Materials* also intendes to reward exceptional early-career scholars by encouraging them to pursue careers as editors.

Given that scholarly journals remain a prime plank of the modern scientific infrastructure it is important to address the need for metrics and evaluation. However, the scope for a wider understanding is also warranted since trust in academic institutions as it a premium in an increasingly complex ecosystem of knowledge claims (see Biagioli et al., 2019; Biagioli and Lippman, 2020; Hall and Martin, 2019). This paper came across, and documented, the proliferation of editorial duty labelling within the same epistemic community: a phenomenon that is underexplained in the literature as far as we are aware. But it also points out to something that is rarely described in journals: the relative degree of autonomy among different editors and, in particular, the binding powers yielded by editors-in-chief. This paper only scratched the surface of phenomena of serial or super-editorships whereas future probes could go deeper into issues of editor mobility, the "market for editorial talent", or the effects of overlapping editorships on scholarship over the long-run. In spite of the shortcomings, this contribution found structural and behavioural accountability to be lacking: the selection of editors is mostly unclear, data on their processing of papers is absent, as well as procedures for appealing decisions. If information pertaining to editors, the editorial process and journal governance appears to be less developed in journals' frontmatter it is nonetheless possible to make more available. For instance, regarding individual editors: their background and expertise, year of enlisting, competing interests, papers managed per year, etc. Regarding the editorial process: publishing a code of editorial conduct, establishing oversight mechanisms for re-appreciating cases or issues by another authority in the journal, etc. Regarding journal management provisions: showing the bylaws governing the workings of the journal by pointing out, namely, its internal structure, reporting practices, editor selection and evaluation methodology, etc. Surely adaptative editorial guidelines and indicators could be specified to assist editors in their roles while helping the prospective author and readers to better navigate the journal content and metadata. It is believed that a better balance between flexibility and governance is possible, that is, enhancing transparency without bringing in unnecessary bureaucracy and excessive audit culture (Heidari et al., 2016).

With this work, we make a modest proposal for integrating the editormetric perspective into the efforts to map and measure Innovation Studies. The insights emerging from our study invite a reflection on the outlook of the gatekeeping crowd, as well as their impact on the evolution of research. Future studies may lead to a better understanding of the editorial black box.

CHAPTER 4

Comparing journals' interests with the gold standard in the field

The following section was submitted to *Quantitative Science Studies* journal. A prior version was also accepted in the 24th Nordic Workshop on Bibliometrics and Research Policy, Reykjavik, 27-29 November 2019, as a poster.

4.1. Introduction

Understanding the aims and scope of academic journals is important for both substantive and instrumental reasons. For analysts of the science system, the mapping and measurement of journals' standing is an instrumental perspective to shed light into the actual institutions channelling the outputs of the research enterprise (Altbach and Wit, 2018; Evans, 2013). For authors, the rapid growth of scientific literature and research topics makes the task of determining where to submit their contributions harder and harder (Freeman, 1969; Mackey et al., 2019; Moravcsik, 1975). For policymakers, who have to look for findings and to care for the visibility of research, the challenges are likewise daunting (May, 1997; Reale et al., 2018; Sooryamoorthy, 2020). We specifically advance (and implement) a methodology to compare (and order) journals by making use of their meta-materials through the deployment of text mining techniques.

International peer-reviewed journals typically offer information about themselves. The promotional briefs published as the front-end material present the goals and topic coverage to readers, who can also be authors or become part of the reviewer pool. Such general statements both summarise the key coordinates of the research field to which the journal caters to (subject matter, academic discipline, etc.) but also announce the differentiating attributes of the individual journal (its specific purposes or belonging to a community, its claim to a particular angle of analysis or specialised topic coverage, etc.). For purposes of expediency, we will call "blurbs" to these self-referential materials, a shorthand we take from the promotional briefs that are used in manuscripts to entice book agents to go through the whole work or on the back of finished books to catch the attention of potential readers (Hall, 2016).

We believe that blurbs can add something to the research policy field, particularly to the scientometric toolbox broadly understood here as the study of quantitative aspects of the scientific enterprise as a communication system (Hood and Wilson, 2001). Throughout the years, bibliometrics-based scoring systems, like the JIF, the H-index[15], and cheer citation volume, have been used to provide public accountability, guidance to researchers choosing where to submit their works, and signal where readers should invest their time and attention. However, no metrics are immune to criticism especially as they become performative, that is, benchmarks become targets in themselves for the private optimising behaviours of actors like journal editors and publishers (Freeman and Soete, 2009; Hall and Martin, 2019a; Martin, 1996). Focusing on journals as a unit of analysis, the novelty of our approach lies in the usage of the journals' own (internal, qualitative) contents to relate and rank journals rather than on more conventional (external, quantitative) statistics. While the exploding availability of digital scholarly data helped to nurture the area of "science of science" (Fortunato et al., 2018; Wang and Barabási, 2021), we contend that the science of journals is an adjunct to such a work programme being aimed at a fuller understanding of the research process.

In this paper, we experiment with underutilised journal narrative raw material, which can be processed and converted into metrics. Natural language processing (NLP) is a computation-driven approach towards managing content and, in going further than counting language pieces like terms or sentences, it provides leverage on the context around the words to capture hidden structures and saliences. Taking advantage of the increasing number of open-source tools (such as R) for transforming textual data into packets of numbers encoding words (Kotu and Deshpande, 2019), we propose a two-step approach to examine and evaluate the blurb sections of journals. For the whole of this exercise, blurbs will be represented as vector objects, which makes it possible to measure the cosine of the angle between vectors and determine the ones which are closer or further away (Han et al., 2012). First, we perform a cluster analysis to categorise innovation-oriented journals into a number of "like-minded" groups. Second, we compare journals' blurbs to find the ones with the most similar interests with RP, the one which Fagerberg and Verspagen (2009) showed to be "only channel of communication that reaches the entire field" and one of the world's most prestigious journals (Zhang, 2020); in other words, inspired by the concept of "gold standard", which in medicine refers to the best treatment or diagnostic tool available against which new ones should be compared, in this second step we study which journals are closer to the key innovation studies platform rather than simply stay with the number of clusters derived from the first step (Cardoso et al., 2014; Versi, 1992).

[15] The H-index is an authors' performance indicator which became also latter applied to the journals (Braun et al., 2013).

The field of innovation studies is taken here as the source ground for our empirical analysis. This area of study is committed to the understanding of the factors, processes and effects of science, technology and innovation phenomena. It is also a multidisciplinary enterprise and an intellectual project geared toward drawing practical insights that can be applied to management and policy (Fagerberg et al., 2013; Martin, 2019). For being an emergent field with a strong firm-centric focus in the early years and for nurturing a broader perspective later on with a greater emphasis on the environment in which knowledge actors operate, it offers itself both as a cohesive and diverse field to determine how journals differentiate and position when presenting themselves as vehicles for publishing.

In the end, we are interested in addressing the following questions: Can journals be comprehended, clustered and classified according to the words they use to describe themselves? Can an alternative journal ranking procedure be developed, i.e., one that considers the substance of their editorial agenda? Answers for innovation studies are shown to be positive. We learn that our approach detects four groups of journals, which we labelled as “Innovation and technology”, “Sectoral studies”, “Organisational theory”, and “Strategic management”. We are also able to confirm RP as the “gold standard” in the field and go on to uncover an ordinal list of journals in terms of blurb proximity, six of which presenting above 80% similarity to RP. While the work reported here would appear to further our understanding regarding how the field of innovation studies is structured, it is also submitted that the methodology is of broader application. Indeed, the knowledge-base and linkages underpinning the contemporary research journal ecosystem constitute a pivotal dimension to unveil in any holistic overview of a field.

This paper is organised as follows. In section 2, we introduce how the scientometric project has contributed to the study of journal phenomena. In section 3, we briefly describe how innovation studies have matured. Then, in section 4, we detail how data was acquired while, in section 5, we go through the content analysis and appraise the results. Finally, section 6 presents a critical synthesis and offers some ideas for the future.

4.2. “Journalytics”: The place of journals in scientometrics

4.2.1 Indicators of science

The rapid progress and development of science created a demand for methods to measure and explain its growth. This demand came from the need to determine and assess the allocation of resources to

"big science" projects (Bush, 1945; de Solla Price, 1965). A further need to transcend simple evaluation was soon felt by those out and around the science silo who felt increasingly anxious to define priorities, direct commitments, and demonstrate effects having in mind outputs and the broader outcomes that could only be appreciated effectively in relative terms vis-à-vis a range of research groups and institutions, other firms and industries, and other countries and geographical blocks (Archambault and Larivière, 2009; Freeman, 1982; Godin, 2019; Smith, 2004).

While by the early 1960s the "science of science" research programme was being articulated (de Solla Price, 1962, p. vii), by the late 1960s it was developing into an "industry" with an increasingly solid foundation in quantitative measurement and statistical theory: Bibliometrics was becoming the prime way to detail and analyse output indicators (Freeman, 1969; Pritchard, 1969). Within years, bibliometrics expanded to incorporate numerous different kinds of research items and interlinkages between them (Broadus, 1987; Garfield, 1982; Patel and Pavitt, 1995) and became a key component of the wider agenda that focuses on the quantitative study of scientific research activities, not just outputs (Egghe, 2005; Glänzel et al., 2019). Indeed, intra-scientific issues, societal concerns, policy-driven and technical demands far exceed today the traditional bibliometric boundaries (Glänzel and Chi, 2020).

The novelty of our contribution, which does not rely on "classic" records (say, papers or references) or web 2.0 sources (like social media or other digital trails) is two-fold. First, we shift the focus from the research piece (the publication) to the research-vehicle (the journal), that is, instead of the product we study the *platform* that distributes the research outputs. Second, we make use not of the materials inherent to scientific practice (say, the academic content) but instead of auxiliary information (i.e., the marketing and advertising corpus deployed for institutional promotion).

4.2.3 Toward a blurb-based research design

Exploring journals' wordage regarding the field which they are placed provides the means for an evidence-based account of how the area itself is defined and a way to investigate how they delineate their intellectual propositions inside the field. It is possible to argue that journal *self-descriptions* (e.g., "aims and scope") offer evidence regarding the state-of-the-art of research since this background text can evolve over time. These small statements present the journals' agenda and goals to the public and can yield powerful indicators of what scholarly platforms are offering and which advances they prefer to harbour.

By tapping into "meta-knowledge" (i.e., the information about information that comes from "the critical scrutiny of what is known, how, and by whom") (Evans and Foster, 2011), increased possibilities come forward so as to identify patterns in the scientific system. For our purposes, the synopsis that accompanies journals is taken as the data source. This material is particularly rich since this text has both descriptive (information, i.e., what is the journal about) and persuasive (advertising, i.e., why stakeholders should be attracted to it) purposes.

Our approach brings forward the (quantitative) indicators that can be derived from such textual evidence (blurbs). The methodology adds to the toolbox of relevant materials and metrics which are used to understand journals. To illustrate and operationalise our contribution, we deploy a particular set of text mining techniques (latent semantic analysis) to the top journals communication "pitches". For this purpose, we apply a "journalytics" approach to the field of innovation studies. In what follows we proceed to establish a perimeter of innovation-oriented outlets; and identify the blurbs from which a corpus can be built. The steps are shown in Figure 14: and are comprehensively described in the following sections of the paper.

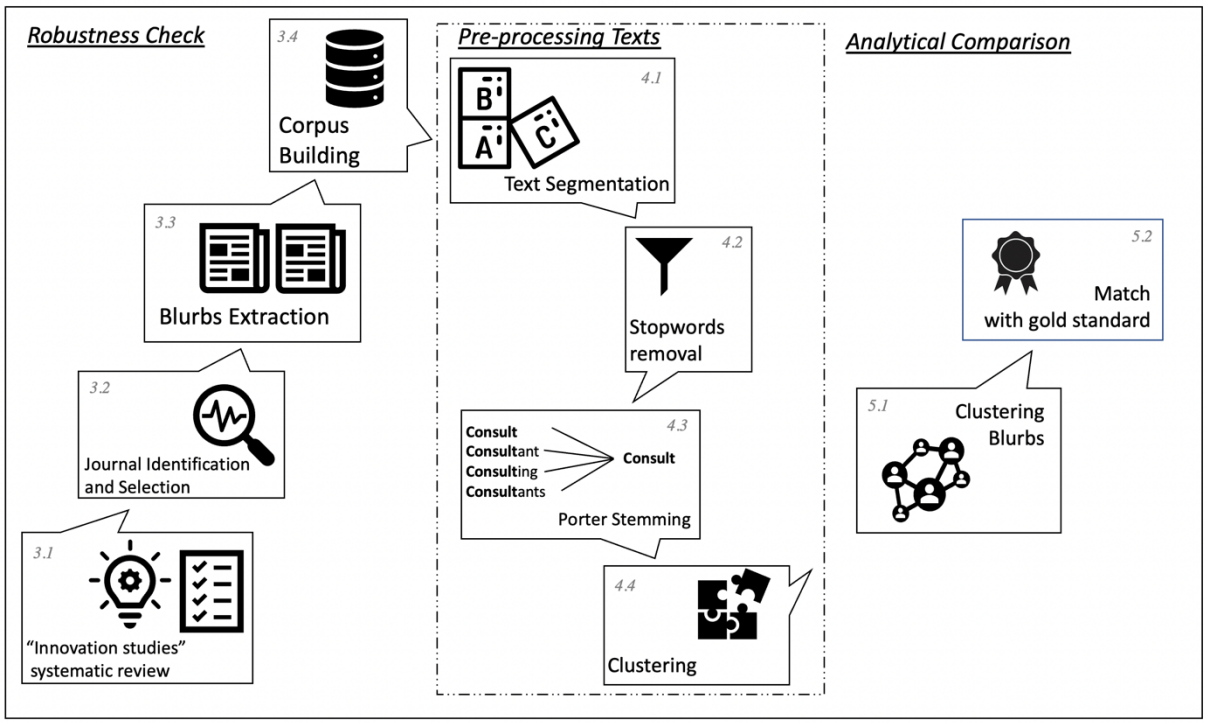


Figure 14: Deriving scientometric value from blurbs: A methodology.

Note: Paper sub-sections addressing each step are identified in top corners, in light grey. Steps identified in 3.1 and 3.2 are described in chapter 1, section 1.2.

4.3. Blurbs text retrieval and corpus building

Journals' homepages display a number of special features that can instruct us about journals' thematic focus. By considering and analysing these contents, we arrive at a novel method of comparing, clustering, and ranking journals based on a systematic treatment of the journals' own material. By performing the sequential steps shown in Figure 14:, a corpus of self-description journal texts was processed to investigate the existence of clusters and find those more innovation-focused (when compared to RP, the main title of the field).

Although journals' self-description is present in all journal's webpages, their location, title and size differ and may be found easily in some journals but not for all. For some outlets a short description may be found in the home page, as the main text, while in others the self-description in found under sections like "Aims and Scope" or "About". The main characteristics such as dimension and location are similar between outlets from the same publisher[16]. Blurbs were extracted from journals' webpages in January 2019. Through a text mining technique, we harness this reflexive data of journals about themselves. In applying such a method, multiple steps were required in order to convert a document collection into a document-by-term matrix.

4.4. Preparing journals blurbs for analytical assessment

To the best of our knowledge, there is as yet no systematic content account of the introductory materials of journals connected to innovation studies. If this (meta)material is indeed there to differentiate journals and to entice scholarly attention and stakeholder recognition, then we will be able to represent a journal's semantic contours and to uncover ways in which individual outlets position in relation to one another. Our "journalytics" methodology, already sketched in Figure 14:, is fleshed out in what follows.

[16] In journals published by Elsevier the blurb section was found immediately before the "Aims and Scope" section on the homepage. SAGE Publications and Oxford University Press present journals' blurbs as the main text on journal's homepage entitled "About this journal" and "About the journal", respectively. Wiley-published journals present their "Aims and Scope" section inside the "About" topic. For those published by Taylor & Francis and Springer, journals' self-description is found in the "Aims and Scope" section, clearly identified in the frontpage. The Academy of Management does something different by defining their journals with a "Mission statement" on the homepage and those from Institute for Operations Research and the Management Sciences which present an "Editorial Statement".

4.4.1. Text segmentation

With text segmentation, the original text is divided into words (Hinterberger et al., 2009). It is also known as “tokenization”, which is the process of setting the sequence of characters into words by locating words boundaries (e.g., white spaces or punctuation).

4.4.2. Removing stopwords

In order to focus on the words with real meaning on our blurbs, we removed word pieces that are trivial or merely auxiliary. This set of words is called “stopwords”. As blurbs had words like “e.g.”, “journal” and “publish” which were considered not relevant to assess journals interests, they were removed from the analysis. This improves the computational capacity (Schofield et al., 2017).

4.4.3. Stemming

We applied a stemming algorithm to reduce morphological variants to a root/base word. This step includes the removal of words prefixes and suffixes applying the Porter (1980) algorithm. Stemming improves the ability to retrieve more relevant information by solving the problem of spurious vocabulary variation and reducing the vocabulary size. With this strategy the number of repeated words increases, and the most frequent ideas are highlighted.

4.4.4. Clustering journal blurbs

Different approaches have been offered to measure the similarity between texts. The term-based similarity models treat each string (or document) as a set of tokens (stem words). The main characteristic of term-based similarity is the use of the overlap of two documents as likeness quantification. In this study, blurbs contents will be matched with cosine similarity because it is calculated using only the dot product and magnitude of each vector. Therefore, it is affected only by the terms both vectors have in common. The angle between two vectors defines cosine similarity. As cos θ may range between [-1,1], 0 means that the two vectors, say x and y, are orthogonal (uncorrelated) and, thus, there is not semantic relation between the two texts but if cos θ is near to 1, their vectors are closer, and therefore their contents are more similar. The cosine measure of the angle between two non-zero vectors of an inner product space may be defined as:

$$\cos\theta = \frac{x \cdot y}{\|x\|\|y\|} \qquad (1)$$

With cosine, document similarity is measured regardless of texts size. Although vectors could have lots of attributes with frequency zero, it does not over-penalise content's overlap for non-common terms (Landauer et al., 1998). It has been proposed as a robust metric for scoring the similarity between two strings by Tata and Patel (2007). Similar documents can be grouped as clusters. The hierarchical clustering will make it possible to generate more clusters with more similar documents within or fewer clusters with more heterogenous units.

4.5. Results and discussion

4.5.1. Analysing the blurbs: The similarity of journal agendas

The application of the methodology to the journal text data allows for the study innovation-oriented research communities. A total of 2,111 words were retrieved from the 20 top journals. Blurb size varies considerably across journals (TFSC being the smallest with 31 words and OSc the largest with 281) and appears to be heavily influenced by publisher policy (space awarded to this information on the websites).

Although we may discuss about the most frequent words and the ones distinctive for each blurb, we aim to understand the relationships between blubs to determine similar contents. Within the semantic space, the cosine between blurbs vectors was computed and the similarity obtained (Günther et al., 2014). The higher the score, greater the semantic similarity between blurbs. Results are illustrated in Figure 15:, where darker colours indicate closer blurbs. No white spaces were found suggesting no blurbs were semantic opposites supporting the journal sample coherence as a whole.

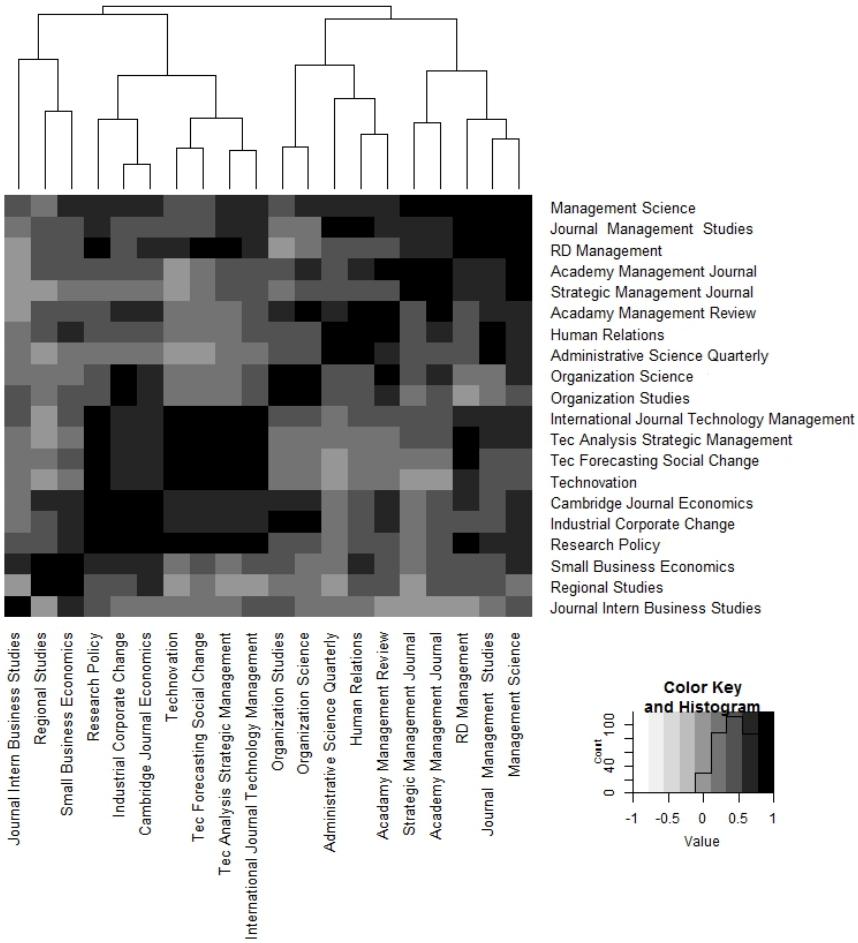


Figure 15: Heatmap and dendrogram of the cosine similarity between journals.

Note: Nearest neighbour (or single linkage clustering) method.

Clear clusters emerge based on the cosine similarity semantic proximity between blurbs. Around the heatmap, the dendrogram illustrates the arrangement of the clusters produced by the corresponding analyses (i.e., hierarchical clustering). Each cluster was furthermore labelled based on the common topics of the journals it encompasses; a content-explicit representation is displayed in Figure 16:. To suggest labelling solutions, the journals' blurbs from each group were analysed and a wordcloud computed [17].

[17] Wordclouds are a visual method of text data, which can display the corresponding size of words with the frequencies they occurred: the more frequent a word appears, the bigger it will be exhibited in the wordcloud (Bateman et al., 2008).

Figure 16: Wordcloud of each journal cluster.

Note: The larger the stem-word, the higher its frequency.

The clusters display communalities among them. Stems like innov*, research* and techn* are all frequently found in all clusters, which is reassuring since we depart from the assumption that the twenty journals contribute to the innovation studies (see, again, Fagerberg et al., 2012). Other common terms are also present, such as develop*, econom*, organ*, public*, social*, strateg* and theor*.

For the "Innovation and technology" group, the largest one and where RP is located, the terms most used were tech* and then innov* which supports the idea of a strong innovation orientation cluster and the central concept used to label it. Also found frequently were manag* and polici*. Encompassing only three outlets, the "Sectoral studies" group shows fewer frequent words. However, busi* is the stem happening most often which lead us to label this group as industry-oriented, whereas the region* term is also indicative of the empirical bent underpinning of this cluster. For the "Organisation theory" group, the most frequent stem is organ* which suggests an organisation-oriented group. Finally, in the "Strategic management" the stem manag* occurs the most.

The key characteristics of each cluster are detailed in Table 8: along with the average number of words and the most frequent stems revealed among the journals' blurbs.

Table 8: Overview of clusters' main characteristics.

Cluster	Journals	Average Year of Foundation	Scimago subject Area	Editorial region	Average JIF	Average no. words (and range)	Most frequent stem-words
Innovation and technology (#7)	CJE, ICC, IJTM, RP*, TASM, Tec, TFSC	1981	Business; Economics	European	3.406	83 (31-148)	techn (#20), innov (#13), econom (#12), manag (#11), organ (#10), polici (#10)
Sectoral studies (#3)	JIBS*, RS, SBE	1975	Business; Economics. Environmental Science and Social Sciences	European	5.758	135 (104-197)	busi (#13), intern (#9), econom (#7), region (#7), research (#6), studi (#6)

Organisational theory (#5)	AMR*, ASQ*, HR* OSc*, OSt*	1970	Business; Arts and Humanities	American	5.395	115 (59-281)	organ (#23), theor (#16), research (#11), manag (#9), scienc (#8)
Strategic management (#5)	AMJ*, JMS*, MS* RDM, SMJ*	1965	Business	European and American	4.943	110 (48-180)	manag (#27), research (#11), strateg (#9) practic (#6), scienc (#6), theor (#6)

Note: * referring to presence in the *Financial Times* Top 50 journal ranking.

The “Innovation and technology” cluster, which has European roots and the smallest Blurb wordage, is the youngest and has the lowest JIF. Sectoral studies”, the second youngest cluster, also has all the journals located in Europe but has the highest JIF. The oldest clusters, “Organisational theory” and “Strategic management”, are more American-based.

It also stands out that the identified clusters reflect the journals’ subjects as collected in Table 1. The “Innovation and technology” group includes two journals from Economics, Econometrics and Finance subjects, five journals focused on Business, Management and Accounting, and also in other fields such as Decision Sciences and Engineering. Thus, we may conclude that Economics and Business provide a robust disciplinary background to those journals that address the “Innovation and technology” agenda more fully. Conversely, “Sectoral studies” seem to be more inter-disciplinary whereas “Organisational theory” and “Strategic management” talk more to business school communities.

4.5.2. A substance-sensitive rank of journals’: capturing their innovation interest

To realise which journal blurbs are the most similar to the “gold standard” the cosine similarity between all blurbs vectors and RP’s were computed. Believing the blurb is stating the desired topics to be published, we may say the more semantically similar a blurb is to RP, the more focused on the core innovation subject-matter a given journal is. In Figure 17., and independently of the clusters assigned before, each journal is ranked according to their blurb cosine similarity to RP.

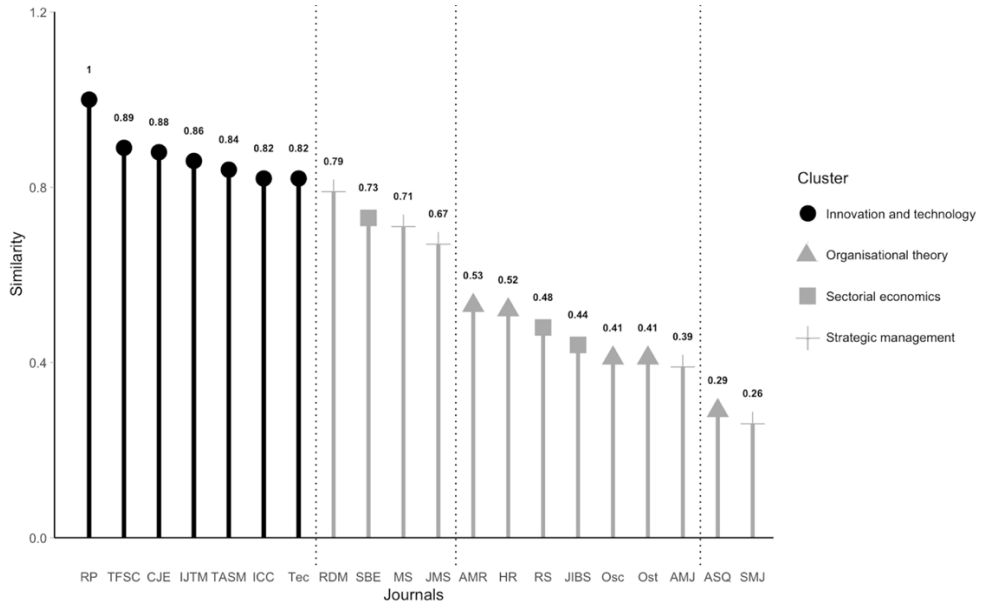


Figure 17. Blurbs’ cosine similarity between Research Policy and all other journals.

In the first positions, it is possible to identify the six journals previously identified in “Innovation and technology” cluster (marked in black). All these journals have a high degree of similarity, above 0.8. In this kind of analysis, the value of the cosine that can be interpreted as defining a pattern is arbitrary, however, “a threshold has to be set” (Leydesdorff, 2007). Other tiers were delineated following a criterion of discrepancy: the existence of two large drops in the cosine were found. As a result, and besides the core group, we posit the presence of three tiers of journals. Three notes suggest themselves.

First, the second-tier journals still seem to exhibit a substantive affinity to innovation concepts and phenomenon. This set of journals is well above 0.5, a threshold that is believed to separate strong from weak similarities (Oghbaie and Mohammadi Zanjireh, 2018).

Second, the below-core tiers are not as neatly patterned as the first tier since each has journals that belong to several clusters previously identified (symbols are used to make this visually evident). For instance, the “Strategic management” journals are found spread across all the three lower tiers.

Third, aligning the ranking obtained from a conventional impact metric like the H-index scores achieved by journals with the one achieved through cosine similarities, it possible to notice the rearranged positions. Figure 18 illustrates how journals change their ranking position according to the method of comparison (H-index vs cosine similarity). The Spearman correlation shows to be of -0.6516, with a p-value of 0.0018, suggesting a significant and negative association between both ranking approaches. That is, the outcome of this analysis does not simply replicate some well-known stylized facts of the field of innovation studies; the comparison of our semantic-sensitive results (showing a hierarchy of journals explicitly oriented to innovation) to a conventional, bibliometric-driven ranking of journals (in which strong preferences for the innovation agenda are not picked up).

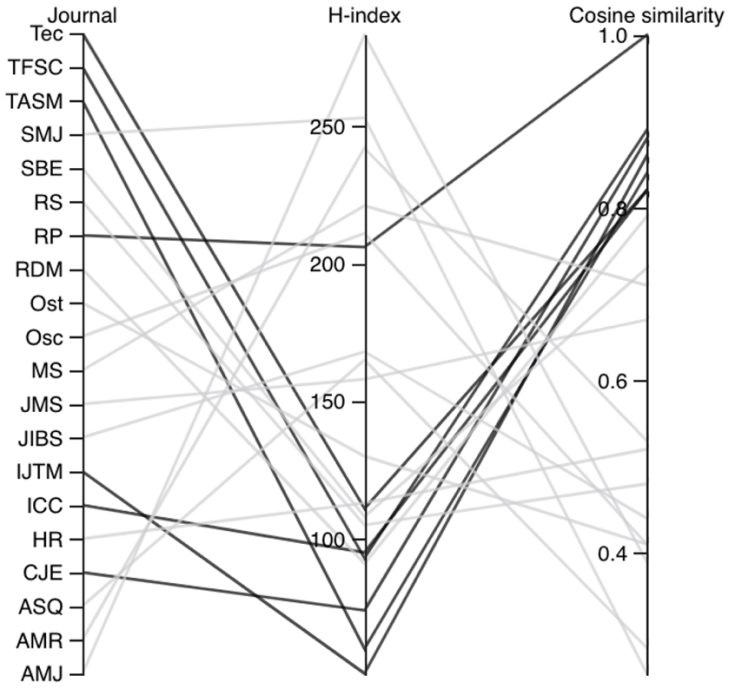


Figure 18. Journals' position for H-index and cosine similarity.

Note: The journals identified in the "Innovation and technology" cluster are in black.

Since all journals are ordered strictly in terms of innovation-oriented positioning (defined as proximity to RP)[18], we believe that this discourse-dependent approach yields an alternative way to systematically order journals (Ana Teresa Santos and Mendonça, 2021)[19]. The innovation-inclined scholarly community may find usefulness in a new type of ranking: a hierarchical listing of journals formed on the basis of thematic organising principles. Witnessing the outcome in terms of the ranking rearrangement makes clear the prime importance of selecting the proper feature based on which to compare outlets and inform stakeholder decisions.

4.5.3. Discussion

Our analysis of the journals' blurb evidence has generated novel insights about the outlets more focused on innovation studies and pointers that suggest this textmetric approach is of potential use for future research and in other areas. In particular, practitioners, including information scientists,

[18] Clustering methods relate all items simultaneously as it were, while here the reference point for all is just one of the items (the "gold standard").
[19] In this work, (Ana Teresa Santos and Mendonça, 2021) map journals self-description with the abstracts from the articles published in the previous ten years.

administrators that monitor productivity and progress, and policy-makers who are concerned with the rhythm and direction of the science system, may benefit from efforts to open the academic journal black-box.

Taking advantage of short self-promotional contents, referred here as blurbs, four clusters of journals emerged that have not been recognised before: “Innovation and technology”, “Sectoral economics”, “Organisational theory”, and “Strategic management”. Along the way, our research was able to validate the (heterogeneous) innovation-friendly journal sample we departed from. While RP is recognised as exceptionally dedicated to innovation studies, six other journals are shown to belong to the core group of innovation studies (the “Innovation and technology” cluster).

When selecting a journal for submission, one of the criteria is its prestige (Salinas and Munch, 2015). However, manuscript rejection rates are a powerful reminder that one of the main reasons for publication denial is still wrong journal selection (Pierson, 2004), mismatch with journal’s scope (Griffiths and Norman, 2016; Khadilkar, 2018; Menon et al., 2020), and irrelevance to the field (Fischer et al., 2017). Unfit papers represent an enormous expenditure in terms of time and resources (Powell, 2016). In this work, we propose journals’ promotional speech as an additional element to be considered when selecting a journal (Ana Teresa Santos and Mendonça, 2021). In judging our top-20 journal by looking to their current JIF or H-scores, a researcher may conclude the best journal to submit an innovation work is JIBS; however, our semantic-sensitive ranking would suggest submitting an innovation-oriented paper to a target where such an article would be more valued, like RP itself or other the journals more focused on innovation studies such as TFSC or ICC. This hierarchy of journals may present a ladder for authors targeting innovation topics.

The evidence we obtained suggests it is possible to adopt a "journalytics" approach to scientometrics; this can be done by considering small and easy-to-read narrative meta-information that usually frames academic media. Conventional scientometric studies centred in metrics such as citations are working at the product level (the paper) while here we are drawing attention to the platform level (the journal). With such an approach, in which analysed contents are taken from websites, the data is assumed to reflect the updated editorial board's preferences and current focus (Glänzel et al., 2003; Van Raan, 2004), making it relevant to quantitative discourse analysis (Salkind, 2010). In such a moment when Big Data is transforming the way research is conducted, we submit that there is still enough mileage (substantive knowledge) in "small data". Such small data (say, "blurbs") is not so affected by the noise that blunts the analytical value of automated informational processing (Banbura and Modugno, 2014; Nate, 2013). In other area studies where a journal can be found to be a scholarly cornerstone, substance-sensitive journal rankings may also be derived as a correction or counterbalance to the conventional citation-based prestige tables. There is a need for broader and more diverse set of approaches to measure the quality of journals, and to give authors guidance to improve paper-journal fit. Journal promotional briefs are up to now a largely overlooked source that can potentially lead to different or more nuanced results when compared to available bibliometric techniques.

The approach is not without drawbacks. For instance, little is known about blurb stability. By checking the blurbs two years after their extraction, in January 2021, we noticed four of them changed: JMS and MS from the "Strategic management" cluster and JIBS and SBE from the "Sectoral studies". Some increased the blurb (such as SBE), others shortened it (like MS). Some re-engineer sentences so as to make them more reader-centric (MS, when moving from "we encourage cross-functional, multidisciplinary research..." to "readers will also find cross-functional, multidisciplinary research"), others enhance their marketing appeal (JIBS, in highlighting its FT50 kudus), while others highlight topic evolution (SMJ, in including "entrepreneurial process"). In our sample we find a 20% change in journals statements in two years, illustrating a rather poorly understood phenomenon: there is dynamics in journal discourse, it might not be evenly distributed among different groups of journals, and blurb dynamics is probably worth monitoring in future "journalytics" studies.

4.6. Conclusions

In this article, we articulate an argument to expand the realm of scientometrics and bring in journals more explicitly to the fore. We apply an approach that may address concerns from various

stakeholders, from article authors to science users, from grant evaluators to policymakers, from editors to publishers. By applying cosine similarity to journal blurbs in the field of innovation studies, where apparently no such prior exercise was implemented before, we are able to distinguish unacknowledged journal profiles and positionings: “Innovation and technology”; “Sectoral studies”; “Organisational theory”; and “Strategic management”.

We confirm Research Policy to be the leading outlet in the field and to belong firmly to the “Innovation and technology” cluster. By taking this journal as the gold standard in innovation-related research, we can also propose a semantic alternative to conventional metrics, i.e., a novel “soft” ordering of journals which can be used to guide journal-paper matching. This discourse-driven or substance-sensitive mapping of journals is quite different from the usual, citation-dependent rankings methods. We find that the “Innovation and technology” journals are the closest ones to RP, followed by a hierarchy of other journals not connected to a particular journal cluster or correlated with impact factor.

With this exploratory analysis, we suggest that self-characterisation wordage (the small promotional materials we dubbed blurbs) may enhance our understanding of journals (Ana Teresa Santos and Mendonça, 2021). By developing an inductive, “bottom-up”, approach to identifying meaningful clusters of journals by relying on their declared editorial preferences instead of predefined classifications, we provided a first attempt at considering journal “front-office” matter for addressing a number of issues of interest for research policy analysts, namely: what do (exactly) journals stand for?; how do academic outlets resemble each other?; how to provide content-based representations of scholarly agendas to navigate the challenges of research identification, monitoring, evaluation?

The answers found for the innovation studies field suggest that the methodology can also be useful to comprehend, compare and classify serial literature channels, i.e., the journals, in other academic fields. Moreover, in innovation studies the “science of journals” agenda could be extended by including a larger number of outlets and going beyond the “top” ones identified in Fagerberg et al. (2012). There are, however, limits in judging journals by their blurbs: there is incomplete knowledge on how these marketing pronouncements come to be formulated; the blurb corpus can be often too synthetic and hence an imperfect mirror of journals’ purposes; they are static pronouncements and can fall behind the pace of the research that gets published, and represent what the journal stands for (regardless the editorial pronouncements). Nonetheless, unveiling embedded themes and revealing the inter-journal structures in sets of academic journals is a worthwhile venture. More should, and can, be done in the future about journals as objects of analysis and modern platforms of research distribution.

CHAPTER 5

Matching journals' self-descriptions and contents published

The section that follows was published in Scientometrics. A prior version was also accepted in the 18th International Conference of the International Society for Scientometrics and Informetrics, 2021.

5.1. Introduction

5.1.1. A content-oriented paper-journal matching approach

The exponential growth in the number of scientific publications makes it difficult for readers to find their interests, for researchers to choose the right journal to publish their work, and for science managers to keep track of the journals that are emerging and their aims (Bornmann and Mutz, 2015; Evans, 2013; Gu and Blackmore, 2016; Jinha, 2010; Shiffrin et al., 2018; Ware and Mabe, 2015). Based on citations assessment (Garfield, 2006, 1972), many other metrics have been developed to compare journals, measure their influence and importance, and make science policy assessments of the content produced and published by journals (Bergstrom, 2007; Bollen et al., 2009a, 2009b, 2006; Bornmann and Marx, 2015; Chen et al., 2007).

The H-index (Ball, 2005; Bornmann and Daniel, 2007; Hirsch, 2005), proposed by Jorge E. Hirsch in 2005 to measure both researcher productivity and citation impact, is also used to evaluate journals (Braun et al., 2013; Mingers et al., 2012; Vanclay, 2008). In 2016, the CiteScore was developed to count the citations of articles, reviews, conference papers, book chapters, and data papers published within four years divided by the number of units published (James et al., 2018; Teixeira da Silva, 2020; Teixeira Da Silva et al., 2017). Social web metrics, altmetrics, for academic publications have also been proposed to assess more than traditional citation metrics such as tweets, blog mentions, news media, social bookmarking, article views, and downloads (Galligan and Dyas-Correia, 2013; Priem et al., 2012; Sud and Thelwall, 2013).

The lack of content sensitivity is one of the major limitations of these metrics. Recommendation algorithms, such as content-based analysis, have been proposed in various contexts to address the problem of information overload and heterogeneity (Bergamaschi et al., 2010; Park et al., 2012; Villegas et al., 2018). It has been used in a variety of contexts, from tweets and other personal opinions to sports news and movie reviews. Pazzani and Billsus (2007) examined the content of articles that a user had previously rated or selected to discover articles that had not yet been considered but were similar to those that had been selected. Computer-aided screening and analysis can contribute to more efficient content review and classification (Lewis et al., 1996). Machine learning algorithms allow for efficient classification of texts that include subjective content of sentences, source identification, and sentiment classification (Cohen and Singer, 1996).

5.1.2. Supervised learning tool as a machine learning technique

Inspired by the seminal work "As we may think" by Vannevar Bush's (1945) in which he proposed a system that would expand human knowledge and understanding, Alan Turing developed a work that addressed the simple question "Can machines think? (Turing, 1950). He proposed a method to evaluate whether machines can do things that humans call thinking, paving the way for the development of what we now call Artificial Intelligence (McCarthy et al., 1955). It involves the simulation of human intelligence processes by machines such as learning (acquiring information and rules for using the information), reasoning (using the rules to reach approximate or final conclusions), and self-correction (Dick, 2019; Ghatak, 2017). Machine Learning emerged as a branch of artificial intelligence, defined as "the study of computer algorithms that enable computer programs to improve automatically through experience" (Chartrand et al., 2017; Liu et al., 2021; Mitchell, 1997).

For the construction of this predictive model, a supervised learning algorithm was applied (Caruana and Niculescu-Mizil, 2006). This is a computer science technique that aims to teach machines to predict unknown cases based on patterns identified from previously observed examples, usually a training set (Breiman, 2001). The learning algorithm attempts to discover and model the relationship between the target feature (the predicted feature, i.e., journals) and other features (i.e., words and their frequencies) (Jain et al., 1999). This type of model predicts a value based on other values in the dataset, so we can use the published texts to predict the journals in which they were published.

Rather than relying on a single model, we hoped to improve overall performance by combining multiple models. This is called an "ensemble" and is based on the idea that combining several weaker learning algorithms creates a stronger learner. The ensemble-based method we chose for our model was Random Forest, as it can process categorical or continuous features and selects only the most important features (Zaklouta and Stanciulescu, 2012; Zhang and Haghani, 2015). Random Forest can process extremely large datasets by using only a small, random portion of the entire feature set.

5.1.3. The conception of the model

Several studies based on content similarity have been conducted for different purposes (Beel et al., 2016). Lewis et al. (1996) proposed a computer-based technique to help teachers group their students for collaborative activities based on their recent work so that their collaboration leads to meaningful interactions. Mooney and Roy (2000) developed a system for generating personalized book recommendations. Protasiewicz et al. (2016) proposed a content-based recommendation system for selecting reviewers (experts) to evaluate research proposals. In this work, our goal is to develop a model that can confirm whether the interest statements provided by journals match the actual published content (Sebastiani, 2002; Wang et al., 2018).

Scientific journals develop specific content to attract researchers and announce the topic they intend to publish. In recent years, an additional area has been added to attract researchers' attention and influence their decision: the blurb[20]. The concept encompasses both an informative and a persuasive purpose. Originally referring to the text on the back of a book (Basturkman, 1999; Gesuato, 2007; Kathpalia, 1997) the blurb is increasingly used on websites to describe journal interests (Gea Valor, 2005; Leo, 2016; Price, 2020). Considering the ultimate goal of blurbs on journals is to persuade authors to select the outlet described to submit their works, we will use these short pieces of information to match with the journals abstracts.

[20] The term *blurb* seems to have appeared in 1907 in a comic book jacket decorated with a drawing of the beautiful 'Miss Blurb' with subsequent definitions of the blurb as a 'flamboyant advertisement', a brief description functioning as a 'commendatory advertisement' and a 'puff piece' (Bhatia, 2004).

We use the abstracts published by the journals to collect them and train the model. F illustrates the steps planned for the development. Since it is possible to know which journal the abstracts belong to, the model is trained with 70% of the extracted abstracts associated with the titles of the journals that have published. The abstracts are converted into a usable format by extracting their features i.e., words. The labelled training data is used to develop the model by training the algorithm to recognize the content patterns in each journal. The remaining 30% of the abstracts are used to test the algorithm and check how accurate it is. In the end, the algorithm will be able to predict the journals to which the blurb texts belong by analysing their content and checking the patterns that fit best.

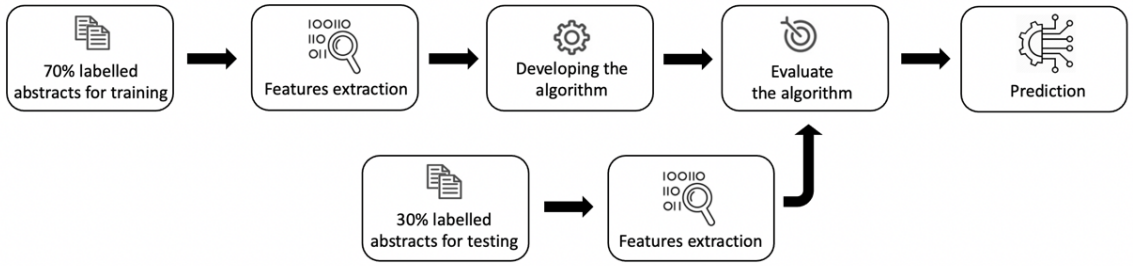


Figure 19: Supervised learning process.

We followed these steps to build a prediction model and test the functionality of our classification algorithm. With this method, a training dataset of abstracts (previously classified in journals) is used to identify unique patterns representing each top-tier group, and these identified patterns are then used to correctly predict the blurb that best matches the published abstracts.

5.2. The journal abstracts

5.2.1. Abstracts retrieval

Data on journal publications are publicly available on the Web of Science, where we searched for all titles published in the twenty journals identified by Fagerberg et al. (2012) between 2010 and 2019. We extracted the title, abstract, and year of publication for each article. Only articles with all these details were included. A total of 16,803 abstracts were considered. We also found that each journal published a different number of articles per year, which changed over time. Figure 20. shows the evolution of the number of articles published in each journal over ten years.

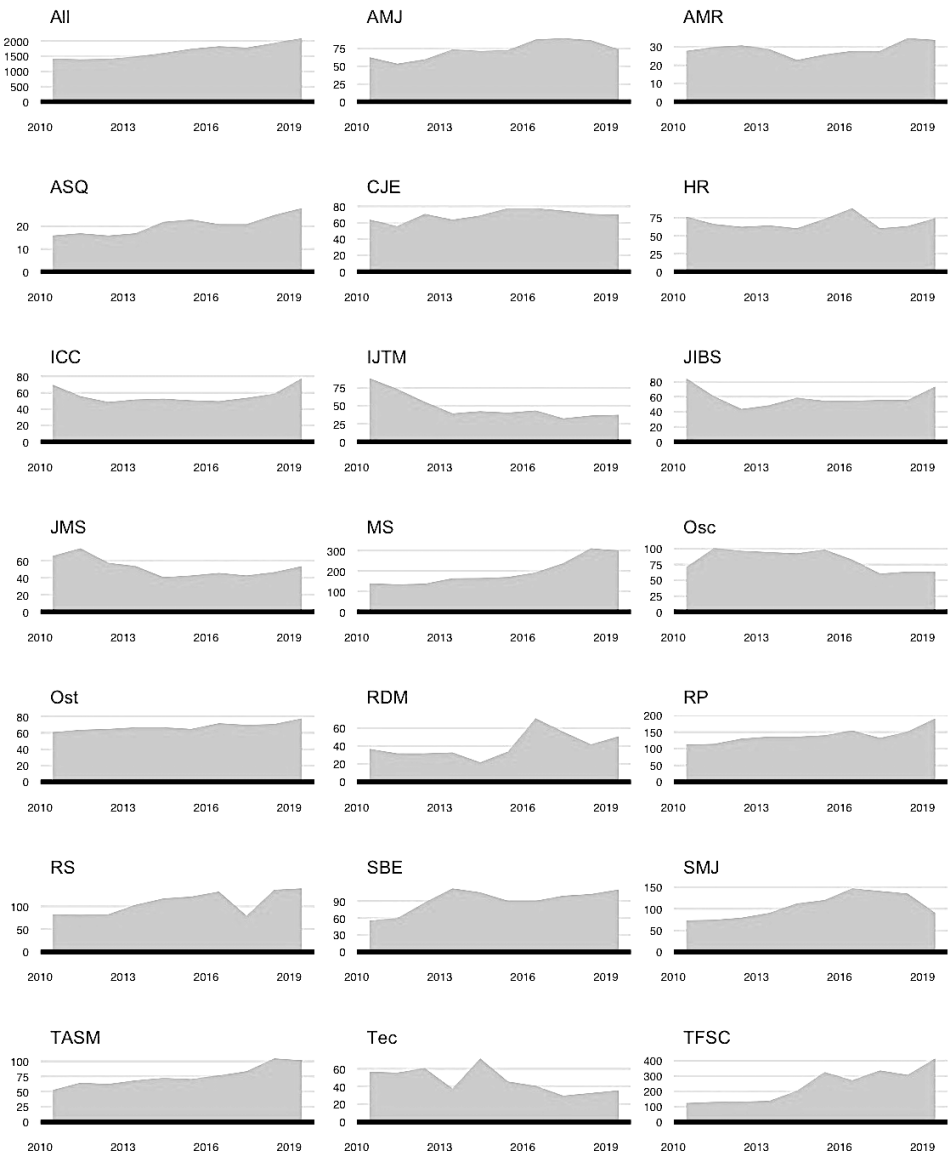


Figure 20. Number of articles published in each top-tier from 2010 to 2019.

The number of articles published in each journal varies, and the overall trend has been upward over the years, with the exception of three journals: JMS, IJTM, and Tec. Both TFSC and MS saw a significant increase in the number of manuscripts accepted for publication, with fewer than 150 articles published in 2010 and more than 300 articles published in 2019. Some titles experienced slower growth, such as AMR and CJE, which published 28 and 64 articles, respectively, in 2010 and 34 and 70 articles, respectively, in 2019.

Such a growth trend was predicted by de Solla Price (1961) who first published quantitative data on the growth of science from 1650 to 1950 at a growth rate of 5.6% per year and predicted that the number of journals would reach one million by the year 2000. Björk and Annikki (2008) also predicted that 1.350.000 articles would be published in peer-reviewed journals by 2006, with lower coverage in the Science Citation Index than in other databases and a decreasing trend over time (Landhuis, 2016; White, 2019).

5.2.2. The abstracts' contents

The manuscript abstract is a stand-alone summary of the paper and provides a quick overview of the purpose of the paper, the research conducted, and the possible conclusions (Hartley and Cabanac, 2017; Smet et al., 1994). These are generally small pieces of text, but their size and internal composition vary. To conduct further analyses with these short statements, the raw text documents (the originals) were converted into numerical feature vectors. This was done in four steps: 1) tokenization of words; 2) conversion to lowercase; 3) removal of stop words; and 4) stemming.

Tokenization of words is the process of separating words into tokens, units of speech or writing. Tokenization can be done simply by adopting spaces and punctuation as separators between words. Each individual unit is called a token. In case of capitalization, all upper-case letters are converted to lower case letters. The goal of removing stop words is to remove meaningless words from the English language, such as the, and, a, is... These common words do not add meaning to the text and therefore can be omitted without having a negative impact (Schofield et al., 2017). Finally, stemming involves reducing inflected words to their root form. In this study, we used the Porter Stemmer algorithm, which is one of the most commonly used stemming algorithms (Porter, 1980).

This pre-processing stage is considered to be beneficial as it can reduce computational complexity and improve algorithm performance without significantly affecting text categorization accuracy (HaCohen-Kerner et al., 2020; Song et al., 2005). With the obtained dataset, the contents of the abstracts could be analysed. Figure 21. shows the distribution of the master data of the abstracts of each journal. The length of the violin represents the distribution of word stems in the abstracts (the longer the violin, the more unique word stems in the journal blurb), and the width represents the frequency of occurrence of the word stems. Large parts of the violin represent the stems that occur with greater frequency.

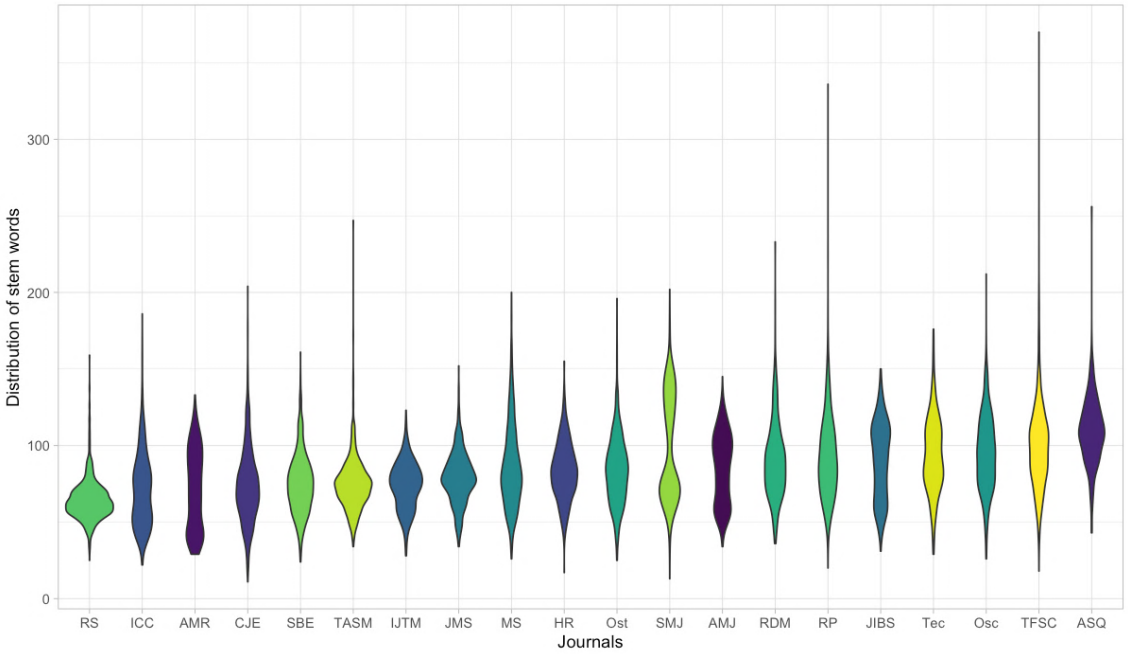


Figure 21. Distribution of stem-words per journal.

Both RP and TFSC are the outlets with the greatest diversity of stems, i.e., greater number of distinct stems. These journals also show a low frequency of the stems' occurrence. RS, TASM and IJTM, on the other hand, have a greater number of stem repetitions. We may infer these three last journals may be more concentrated in the topics they select for publishing, while it is suggested RP and TFSC may publish a large variety of topics studied.

5.2.3. The trends on topics preference

With these 16,803 abstracts from 10 years of publications, we aim to understand whether there are any topics more frequent in each journal. We developed Figure 22. where the lollipops illustrate the five words which occurred the most in each top tier as well as their frequency. Some of the most used words were common to more than one journal.

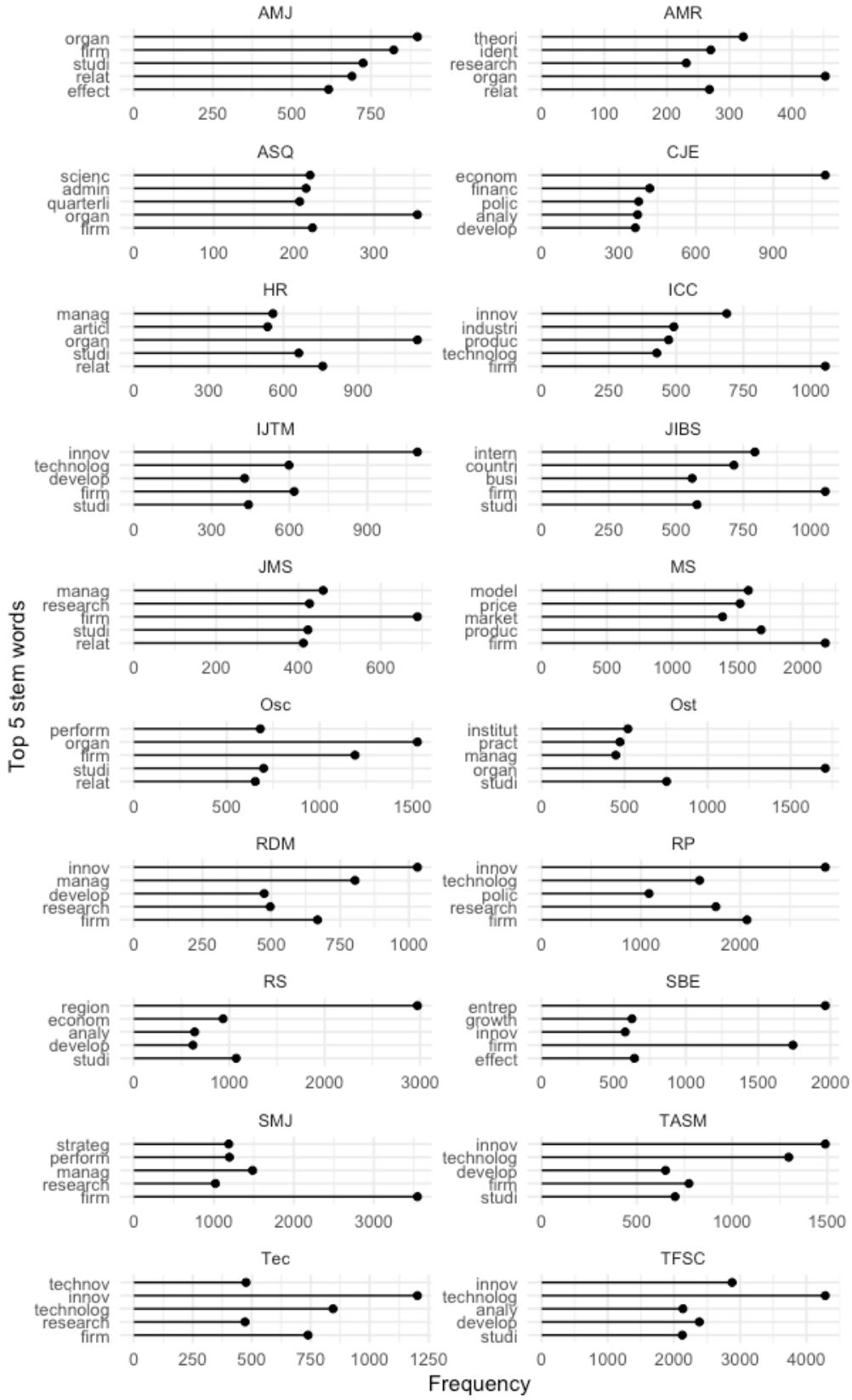


Figure 22. Top 5 stem words occurring the most between 2010 and 2019, per journal.

As expected in a journal set devoted to innovation studies, the word “innov” appears as one of the most frequently used in seven of them: Tec, TASM, RP, ICC, TFSC, RDM and IJTM. Both “technolog” and “organ” were among the most frequent word for six other journals: ICC, IJTM, RP, TASM, Tec and TFSC, and AMJ, AMR, ASQ, HR, OSc and OSt, respectively. It appears that “firm” is also important, as it is the most common stem for SMJ, JMS and JIBS. RS is the most specific journal, with a unique most frequent stem “region”. This distinction could be explained by the fact that the focus is on “Environmental Science,” as revealed in Table 1.

5.2.4. One-word pattern over the years

Even though the top five words (and stems) in each channel have already been identified, this set of words could have been changed or, at least, their proportion, over time. Relevant topics are typically defined by the interests of multiple stakeholders, including commercial agents, research groups, government policymakers, and, more recently, consumer needs (Martin et al., 2009; Phillips, n.d.). We believe that changing editorial directions can be accomplished by encouraging article submissions or selecting for publication topics of critical importance at the time, thereby adjusting the balance of editorial interest. These new intellectual directions in the editorial pipeline are frequently described in the outlets' editorial sections (Caruana and Niculescu-Mizil, 2006). To see the evolution of some key concepts' usage, we plotted their frequency for each year over a ten-year period of data collection. In Figure 23:, it is illustrated how the stem “innov” changed over time in each scientific journal.

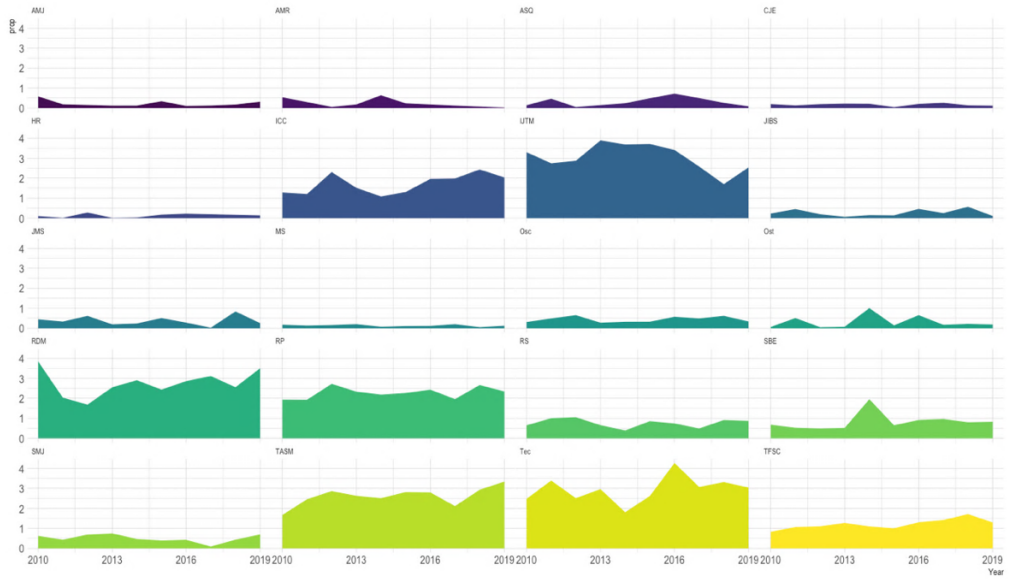


Figure 23: Usage of the stem "innov" over the 10-years period by each of the journals.

The pattern of “innov” usage generates three identifiable clusters encompassing journals with similar proportions of this stem occurrence over the years. The first one with the ones with higher frequencies include IJTM, RDM, RP, TASM and Tec. The second cluster noticed was made up of ICC and TFSC with moderated use of the “innov” expression during the 10 years period. The third one includes all the others, using “innov” more rarely, although present almost every year. This large group includes 12 journals: AMJ, AMR, ASQ, CJE, HR, JIBS, JMS, MS, OSC, OST, RS, SBE and SMJ.

RP seems to present a more constant usage of the word while the other outlets exhibit some changes over the years. The increasing occurrence of the usage of “innov” in RDM, ICC, TASM and Tec contrasts with decreasing noticed for IJTM. Another word noticed as being among those most frequent was “technolog”. Figure 24: shows the proportion of its usage in each journal for the 10 years period studied. As expected, based on the Figure 22., IJTM, RP, TASM, Tec and TFSC use this stem in the abstracts published in a great proportion.

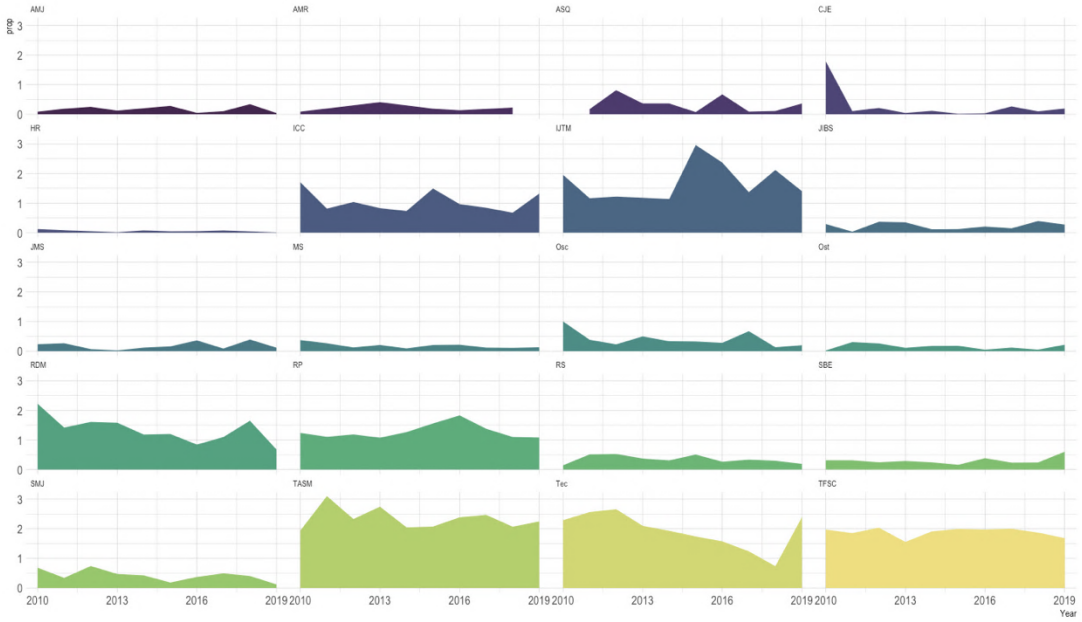


Figure 24: Usage of the stem "technolog" over the 10-years period by each of the journals.

TASM and Tec, a pair of journals with greater proportions of “technolog” usage, showed also a decreasing trend with a significant reduction in 2018 for Tec. Larger tendencies were also noticed in 2015 for IJTM and ICC, followed by RP and ASQ in 2016, TASM and OSc in 2017 and RDM and IJTM again in 2019. For studding the pattern of usage of “polic”, Figure 25: was also plotted. Although only reported among the top 5 most frequent words for RP in Figure 22., other journals revealed to be using it in large proportions.

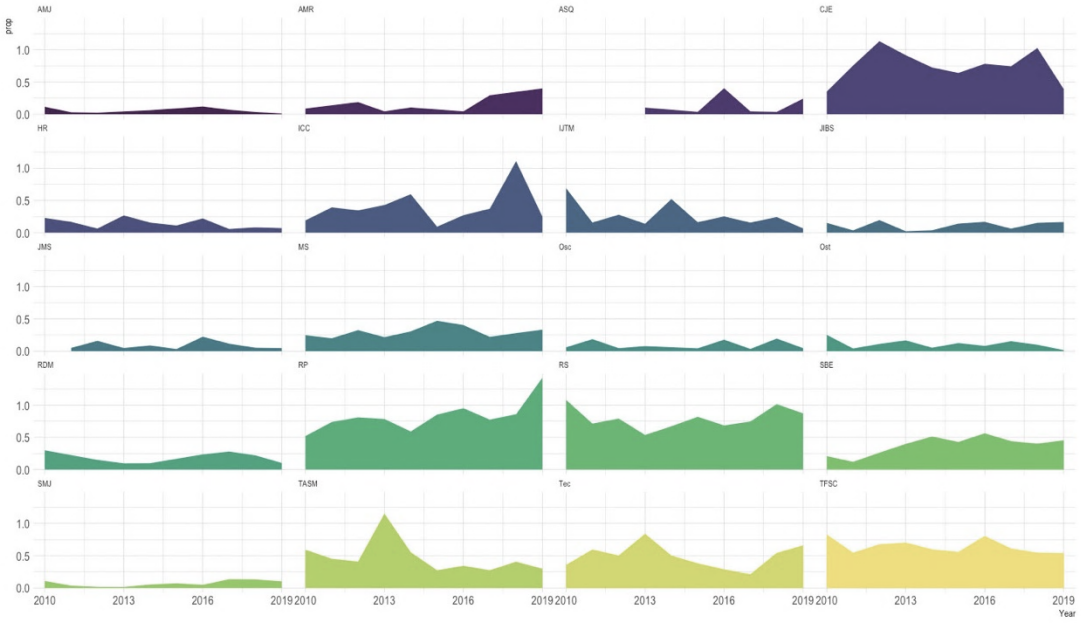


Figure 25: Usage of the stem "polic" over the 10-years period by each of the journals.

CJE, RS and RP revealed to be the journals with larger proportion of usage of this word. In 2012, it was the year when "polic" proportion of usage achieved a maximum, followed by TASM and Tec in 2013. While ICC was using in a great proportion this stem in 2014, RP registered a low usage although increasing in the following years. ICC and RS achieved maximum proportions in 2018 and RP followed the trend in the year after.

5.2.5. Ten-words pattern of journals' interests over time

In addition to the topic of innovation, which was covered in all journals over a ten-year period and which we consider to be the most important topic for this group of top candidates, many other topics attracted the interest of this group of top candidates. We compared the frequency with which the topics appeared in the abstracts of each journal each year to examine the evolution of the topics over time. Some additional word-stems were removed as they related to publications rather than research topics[21]. From Figure 26: and Figure 27: show the top ten of word stems with the largest changes in

[21] The excluded terms are: "studi", "aim", "neg", "set", "paper", "author", "approach", "discuss", "suggest", "right", "reserve", "view", "analysi", "term", "active", "data", "region", "find", "found", "question", "research", "result", "u.", "articl", "literatur", "draw", "base", "form", "moder", "cluster", "journal", "find", "elsevi", "analyz", "analys", "offer", "respons" and "publish".

frequency for ICC, IJTM, RP and TASM. For each journal, the word stems and their proportion of use over time can be viewed.

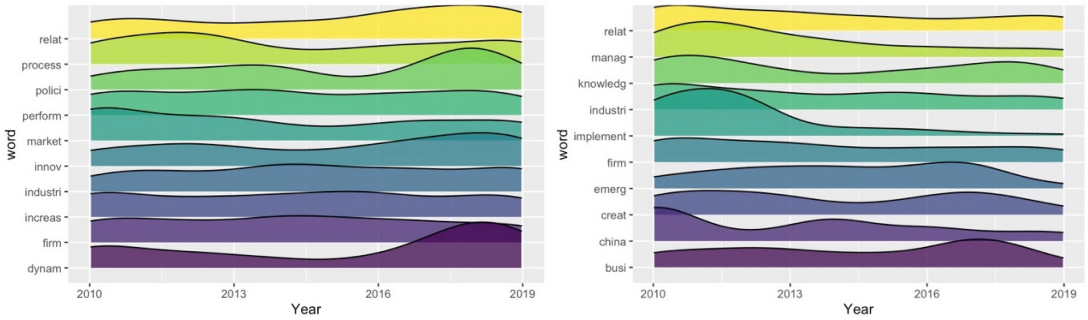


Figure 26: Topic trends with largest variation in A) ICC; and B) IJTM.

According to Figure 26:, ICC's interest in "firm," "industry," and "innovation" has remained consistent over the years. In 2012 and 2014, “process” drew the most attention, followed by “dynamics” in 2015 and 2016, “relations” in 2014, and “policies” in 2015. For, IJTM, the term "firm" was a recurring theme over the years. Topics such as "firm" and "industry" declined, while "management" increased. Furthermore, “China” increased its occurrence while “emerging” decreased. Between 2012 and 2016, “implementation” became more common, and “creation” had a strong showing in 2014 and 2015.

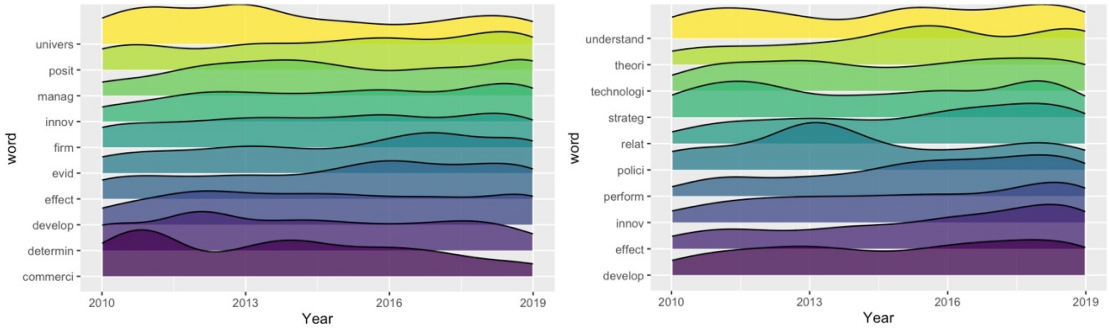


Figure 27: Topic trends with largest variation in A) RP; and B) TASM.

For RP, "innovation," "development," "effects," and "firm" were the topics identified as the recurring ones over time. On the other hand, “management” had outlived its usefulness by 2011, and “determin” was irregular, with a focus on specific years such as 2011, 2013, 2016, and 2018. In TASM record, the terms “develop,” “innov,” and “technologi” were the most important during this time period. In 2011, “strateg” was a major topic, followed by “polici” in 2013 and “understand” in 2014. In 2015, the topic of "perform" became significant, and in 2016, the topic of "effect". In 2017, there was a lot of talk about “theori”.

5.3. Journal blurbs vs abstracts published

5.3.1. Similarity between blurbs and abstracts

Considering all the words that appear in the abstracts published during the ten-year period studied, a text clustering algorithm was implemented to find related blurb texts and groups of abstracts. A similarity measure is used to represent the similarity between texts in a single numerical value that depends on two factors: the properties of the two objects and the measure itself. We used cosine similarity because it is a measure that is applied to text documents, such as in numerous applications for information retrieval (Yates and Neto, 1999) and also clustering (Larsen and Aone, 1999), and because it is independent of the length of the document, which is important when the length is so heterogeneous in the sample under consideration.

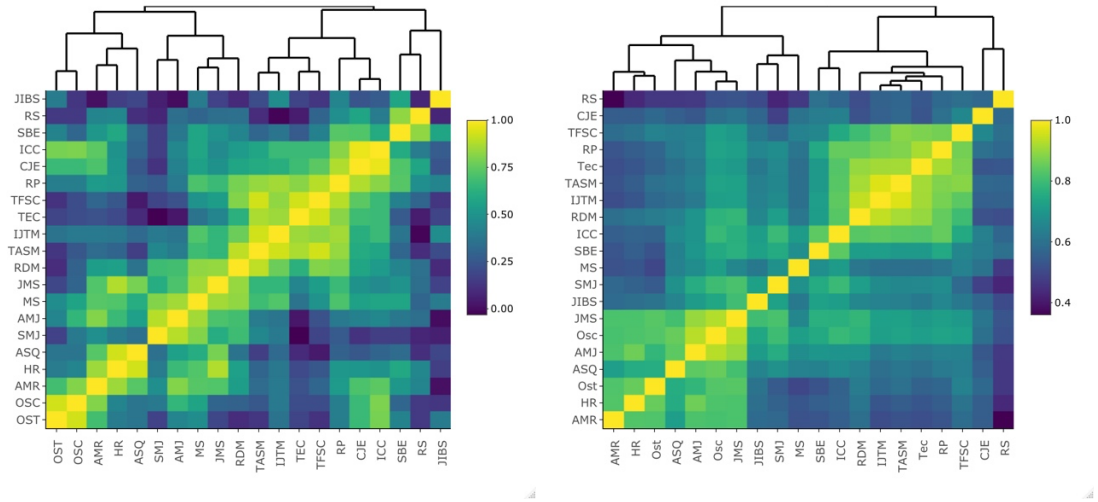


Figure 28: Cosine similarity between journals based on: A) blurbs; B) abstracts published.

The similarity shown in Figure 28: indicates that the journals would be grouped differently when abstracts or blurbs are considered. Considering four clusters based on the similarity of the blurb texts, we find 1) a sectoral cluster with OST, OSC, AMR, HR and ASQ; 2) a management cluster with SMJ, AMJ, MS, JMS and RDM; an innovation and technology cluster of TASM, IJTM, TEC, TFSC, RP, CJE and ICC; and finally, 4) a business cluster with SBE, RS and JIBS. However, if the clusters are formed based on the abstracts published between 2010 and 2019, the four clusters would be: 1) a sectoral cluster with AMR, HR, OST, ASQ, AMJ, OSC and JMS; 2) management cluster with JIBS, SMJ and MS; 3) an innovation cluster with SBE, ICC, RDM, IJTM, TASM, TEC, RP and TFSC; and 4) a small group of two business journals, the CJE and RS.

These clusters may indicate that some journals have different interests than the published abstracts. For example, the content published by SBE in its abstracts is closer to other innovation-oriented journals, while the stated interests are more similar to those of business journals such as JIBS and RS. Both SMJ and MS state objectives in a blurb that are closer to other management journals, while the abstracts published during this ten-year period are closer to a business journal, JIBS.

5.3.2. The matching model & the learning algorithm

Among the supervised learning algorithms known, those combining multiple learning models using either boosting or bagging techniques have shown to achieve better results than simple learning models (Caruana and Niculescu-Mizil, 2006). We studied three different classifiers: linear Support Vector Machine (Cristianini and Shawe-Taylor, 2000; Osuna, 1997), Decision Tree (Murthy, 1998; Swain and Hauska, 1977) and Random Forest (Breiman, 2001) as their performance does not depend on the input's dimensionality and they are robust to high dimensions. In other works, Random Forests outperformed the other learning algorithms (Breiman, 2001). Actually, also Fernández-Delgado et al., (2014) evaluated the Random Forest classification performance versus support vector machines, decision trees and many others on 121 public datasets and the Random Forest algorithm achieved the best classification results.

Random Forest is a specific application of ensemble learning (Rokach, 2010). The Random Forest is built on classification and regression trees. To increase the diversity of each tree in the forest, Random Forest is trained using a bootstrap aggregating (bagging) technique (Breiman, 1996). The prediction ability for a single tree is limited and given a large dataset, overfitting is common for a single decision tree. Random Forest was operated by constructing a multitude of decision trees during training, then outputting the class that was the mean prediction of the individual decision trees. Like any other supervised learning techniques, the goal of Random Forest model is to identify patterns in a training set and then use these identified patterns to predict future unseen cases.

The main assumption of machine learning is that the distribution of training data is identical to the distribution of test data and future examples. If the learning algorithm accurately classifies the set used for testing, then the machine learning assumption suggests that it will perform as well for future unseen cases (Russell and Norvig, 2002). For the training purposes, our classification dataset made of abstracts was split into two disjoint sets so 80% of the observations were used for training and 20% for testing and check accuracy.

5.3.3. Cross-validation & accuracy check

Cross-validation is a technique to evaluate predictive models by partitioning the original sample into a training set to train the model, and a test set to evaluate it. Training accuracy was determined by a ten-fold cross validation (Arlot and Celisse, 2010; Kohavi, 1995). In a ten-fold cross validation, the input data is divided into ten equal disjoint subsets. Each subset is used as the test set while all the others are used as the training set. For each validation, accuracy is measured by the ratio between the "number of correctly classified cases" and the "total number of cases". The final accuracy will be the average accuracy of the ten different subsets (Zhang, 1993).

Predictability is usually defined as the possible maximum accuracy of a recommendation algorithm. Depending on the application the learning model is designed for, the minimum acceptable accuracy level may be different as well as maximum degree of predictability (Song et al., 2010). Figure 29: shows the accuracy (in percentage) of our Random Forest model classifying the abstracts to the journals which published them. The overall accuracy of our model was 80% which means our model is able to correctly predict 4 out of 5 abstracts. Previous articles have described and used accuracy measures that are widely recognized as the standard way to determine the test and training accuracies of prediction models (Mullen and Collier, 2004).

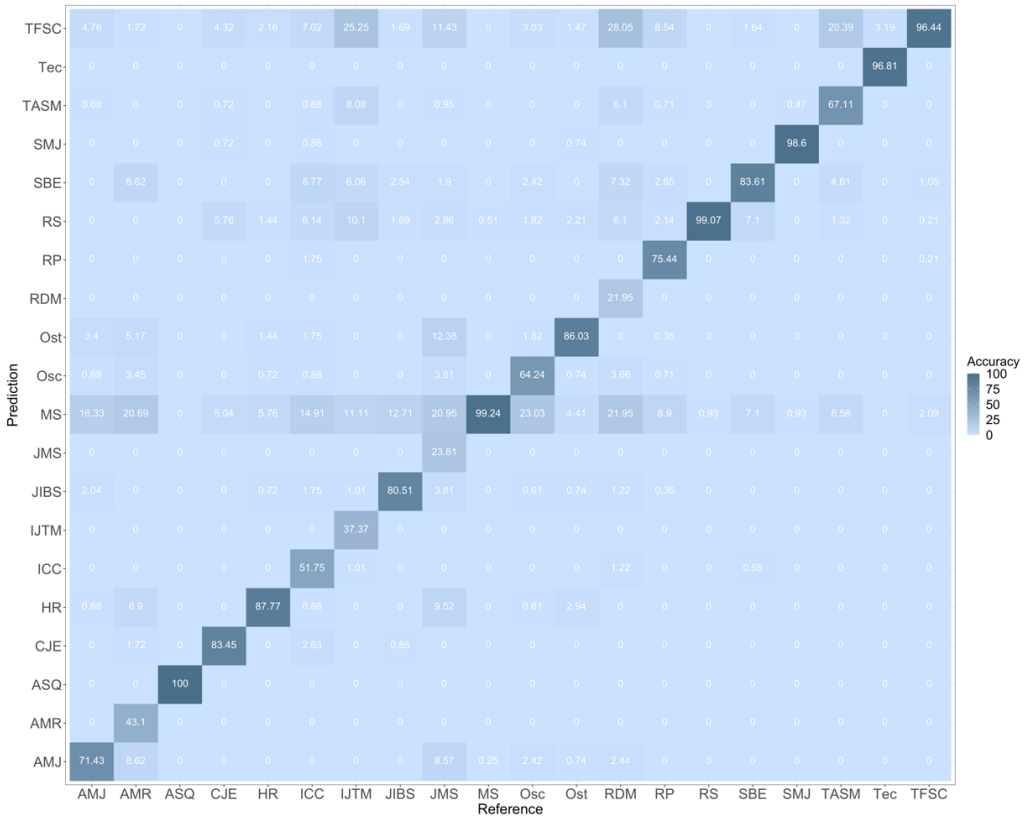


Figure 29: Prediction accuracy between abstracts and the journal which published them.

Although the overall accuracy was 80%, the accuracy achieved by model classifying the abstracts belonging to each specific journal varies greatly. Figure 29: shows that all of the ASQ abstracts were correctly classified. Closer to this figure are the abstracts published on SMJ (98,6%), RS (99,07%) and MS (99,24%). Nonetheless, the model's performance in predicting the abstracts published in RDM, JMS, IJTM, and AMR was subpar. The model's accuracy differences can be explained by the type of content published. It is suggested that the more specialized the journals, the greater the predictability of the model. The opposite is also true for journals that publish a greater variety of topics. The model assigns to these journals abstracts that have previously been published in other outlets, usually with low levels of predictability. This is the case with MS and TFSC. Both are predicted to publish many abstracts from other journals. When stating their eclectic and diverse interests, their editorial announcement defines themselves not only as generalists, but also as direct competitors of others.

5.3.4. Matching blurbs with the abstracts

Using the model developed, we used it to predict which blurb matches the set of articles previously published. The goal is to realize whether the journals self-description matches with the current articles being selected for publishing. For this purpose, we applied the model trained with the abstracts published for ten years to see whether the stated topics of interest in blurbs match the journals which published them. All pre-process steps were applied to the set of blurbs in order to obtain a dataset to be used by the model and match the journals. Figure 30: presents the classifications suggested by the model. On the left, we see set of abstracts published in each journal during the 10 years period considered and used to train the model while in the right side, the blurb text is considered. The lines show how the model classified the blurbs.

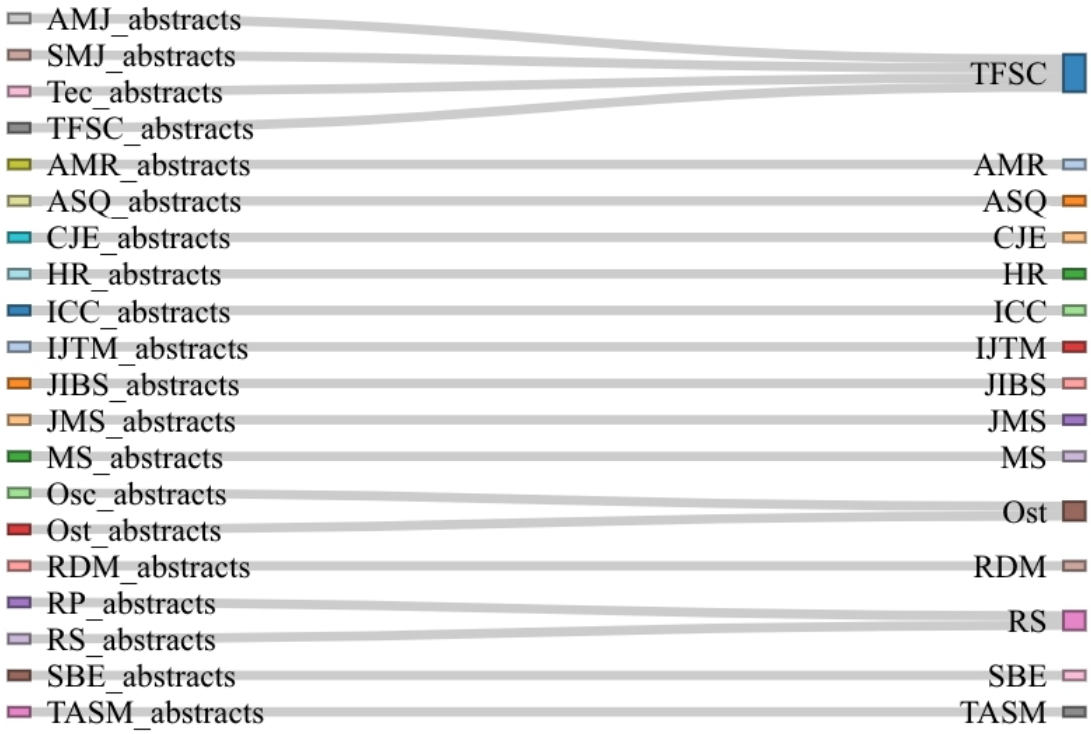


Figure 30: Blurbs' classification.

In the sample, fifteen blurb texts were correctly linked to the journals to which they belonged. This finding suggests that the blurb texts in these journals accurately describe the content of the article abstracts selected for publication. Five blurb texts, however, were linked to different journals: RP, Osc, AMJ, SMJ, and Tec. The abstracts from AMJ, SMJ, TEC, OSc, and RP were more semantically similar to other journals' blurbs. The blurb published by OSt was considered to be similar to the contents published in the abstracts from OSc and OSt. The model also discovered that AMJ, SMJ and Tec abstracts more related to the TFSC's description and RP and RS abstracts published over a ten-year period were semantically more similar to the RS blurb.

These findings suggest that for 25% of the journals contents published could have been selected for elsewhere. Their contents are more in line with other journals' blurb descriptions than their own. Although RS is seen as a more generalist journal and RP as more specialized, the model appears to indicate that RP is looking beyond its specific topics of interest expressed in the blurb and is interested in publishing topics addressed by the RS. The same can be said for OSt's blurb. The contents published in both OSt and OSc are semantically more similar to the interests stated in OSt's blurb. Even more journals match the introductory statement of TFSC's interests. Its content corresponds to articles published in AMJ, SMJ, Tec, and TFSC itself. This could imply that the diversity of journals is lower than if we looked at blurbs. When reading the introductory sections of journals, one might imagine a broader range of topics being addressed and published, even though the actual publications suggest otherwise.

5.4. Discussion

In this study, we presented a methodology for authors, science managers and policymakers to select journals based on the approved content. Although some of the utilised methods have been known for decades, to best of our knowledge, they have not been applied in the context of recommending scientific journals for research manuscripts submission, or contents providers. Our goal in this work was to create an algorithm capable of providing such valuable guidance to researchers based on contents published. We discovered that using modern text analysis methods, it is possible to text mine narratives from existing units of knowledge (abstracts) and yield advice about the outlets where a new manuscript best fits or is more focused on a specific field.

Because the aims and scopes of outlets change over time as well as the contents received and accepted for publishing, journals fair recommendations would benefit from regular state of the art analysis so that their actual interests could be identified and made available to the research community. In this work, we limited our sample to the top innovation oriented scientific outlets identified by Fagerberg et al. (2012) and 16,803 articles were retrieved from the twenty journals sampled. The frequency of articles reassured the global trend of increasing the number of papers published each year.

When it comes to Machine Learning classification, the key question is not whether one learning algorithm is better than another, but rather under what conditions a specific method can significantly outperform others on a given application problem. Because each approach has its own set of drawbacks, the combination of all approaches is critical. They can be overcome, however, by combining them. Although an ensemble procedure was used to achieve the highest classification accuracy possible, some flaws remain. The first is increased storage as a result of the requirement that all component classifiers be stored after training rather than a single classifier.

With the application of a Random Forest algorithm, we achieved an overall 80% of accuracy in determining the journal's contents using machine learning techniques. For the journals ASQ, SMJ, RS and MS, almost 100% of their abstracts were correctly recognized. However, the model misclassifies larger proportion of articles from the RDM and JMS. Our hypothesis is that the automatic classifier can do better with specialized journals whereas the ones publishing a broader range of scientific topics, such as RDM and JMS, mislead the algorithm into identifying a characteristic pattern of interesting topics.

This trained algorithm was able to check whether the contents published semantically matched the blurb statements in the official webpages. For 25% of journals, the published contents proved to be not corresponding exactly to their statements of interest and ultimately more similar to other journal blurbs than their owns. This suggests that journals published contents are not so diverse as one might think when analysing the self-descriptions. An example was RP, a specialized journal, which proved to be publishing contents beyond its stated interests, more similar to the ones described by RS.

5.5. Concluding Remarks

In this paper, we proposed a text classification approach for determining the best venue for submitting research articles or locating relevant content. Based on the content of accepted manuscripts for the top twenty innovation-studies oriented journals identified by Fagerberg et al. (2012), between 2010 and 2019, we developed, trained, and validated a model which proved to have 80% of accuracy matching the abstracts with the journals that published them, i.e., the model is correctly assigning supervised learning model 4/5 abstracts to the journals which actually published them. This model also allowed us to determine that 25% of these journals were selecting for publication contents matching other top-tiers' preferences according to the front matter of scientific outlets (which is gaining popularity among the research community).

Text classification, or computer-aided analysis of textual data, offers a great opportunity to advance journal selection strategies, as has been recognized in many other research fields. In this proof-of-concept study, we also distinguished between blurbs that described the exact outlet and those that expressed common interests to others. We believed it could be used to address the real problem of selecting the most appropriate venue to submit an article based on the content of previously accepted manuscripts because of its high potential for efficient and effective classification. Such methodology may also benefit researchers from areas of science knowledge other than innovation studies, or it may support other critical tasks such as peer-review (Kovanis et al., 2016).

Some limitations should eventually be reported and addressed in future works. This project encompasses the content published during a period of ten years (between 2010 and 2019), which was primarily motivated by the need to balance data size and computational capacity. A larger set of training abstracts, blurbs, and journals would be desirable for future exploration of the potential of machine learning algorithms to compare journal contents. A validation study could be useful to demonstrate the necessary adjustments when considering broader scientific fields.

CHAPTER 6

An interactive tool

6.1. The purpose of the app

A key achievement to any scientist is to turn a scientific manuscript into a published article and reach a large number of readers and authors. However, choosing an appropriate outlet to submit a research work is a difficult decision. In the modern world where the pressure to publish on academics worldwide is growing every day, having an article being rejected because of not fitting the aim of the journal can have a negative impact on a yearly academic review and prevent a pioneering idea from entering the desired field (Alberts et al., 2008).

On the other hand, those with top-level responsibilities such as science managers and policymakers are facing some challenges to map the scientific knowledge. Human-assigned metadata, such as subject category classification of articles or journals, has been the dominant source to map science (Börner, 2010). Fitting new knowledge into pre-existing and finite categories appear to be, at best, limited and poor (Small, 2004). Suominen and Toivanen (2016) argued that makes no sense to use "a rearview mirror to predict where a fast-moving car is heading."

In the last years, several metrics emerged for journals comparison and multiple web services have been developed to assist authors in selecting a publishing venue. A variety of performance metrics that can be used as filters when selecting journals form the basis of these rankings and comparison services. Cofactor Journal Selector, developed by the London-based firm Cofactor, guides users through a detailed set of filters to match an author's publishing needs ("Journal Selector," n.d.). Others, such as Elsevier Journal Finder and EndNote Match, developed by Elsevier and Thomson Reuters, respectively, require users to enter key information about the article to be published (e.g., title, abstract, keywords/phrases) and then use that information to find the best matching journals (Kang et al., 2015). However, these services provided by publishers proved to be limited to their own pool of publications, assuming authors begin their decision-making process by first choosing a publisher (Forrester et al., 2017) and narrowing the available options. Furthermore, no tool was shown to provide content-specific guidance.

For this reason, that compared journals based on the content that was actually published. The user could choose the journals to compare, as well as the articles to consider and the years to compare. With such latitude, each researcher could compare the outlets of his or her choice, regardless of publisher. To ensure that everyone can deal with it intuitively, a user-friendly environment was also planned.

6.2. Functionalities and user interface

The name "Science Submit" was chosen for instrumental reasons because it could provide some guidance to researchers about the journals for publishing a work and its purpose could be easily understood by those seeing the app for the first time. It was also intended to be generalist in nature, as it could be applied to the map out of any scientific field. Despite the fact that other tools for bibliometrics and scientometrics research have been developed, such as Bibliometrix (Aria and Cuccurullo, 2017), to the best of our knowledge, content analysis has not yet been included and articles selected for publishing have not been considered for journal comparison.

This one-of-a-kind tool is implemented using R/Shiny library and modular structure. R/Shiny package easily allows developing advanced and practical interfaces in a web-based approach combined with the power of the R statistical instrument. Shiny apps were originally designed for small applications consisting of two main entities: the Shiny User Interface (SUI) that provides all the aesthetic components the user interacts with and the Shiny Server Side (SSS) that performs the required computations. A module is conceived as a shiny independent app, with its SSS and SUI. R programming language allows for a wide range of statistical and graphical techniques. Thus, Science Submit is flexible, can be quickly upgraded, and can be integrated with other statistical R-packages because it is written in R. As a result, it is useful in a field that is constantly changing, such as bibliometrics.

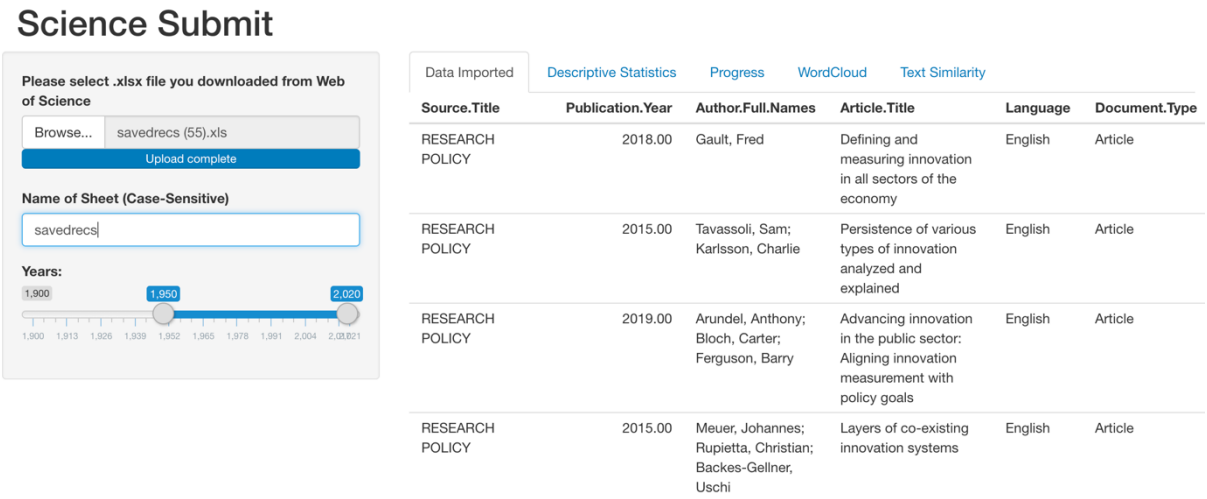


Figure 31: depicts the main layout of the tool, including a screenshot of the app with input boxes. Users can choose which contents to compare by using three interactive fields. The first box allows the researcher to upload a file that was downloaded from WoS. A "Browser" button is also provided so that users can search their machines for the downloaded files to be used in the Science Submit app. The second box allows the user to specify the name of the excel file sheet containing the data. WoS sets the excel sheet name to "savedrecs" by default. However, this option is provided so that the user can have different datasets in multiple excel sheets from the same file and choose which sheet to use by the algorithm. A range slider is also provided for selecting the years. The goal is for the researcher to be able to experiment with and change the years of publication without having to upload another subset of the previously selected data.

Science Submit
Please select .xlsx file you downloaded from Web of Science
Browse...
savedrecs (55).xls
Upload complete
Name of Sheet (Case-Sensitive)
savedrecs
Years:
1,900
1,950
2,020
Data Imported
Descriptive Statistics
Progress
WordCloud
Text Similarity
Source.Title
Publication.Year
Author.Full.Names
Article.Title
Language
Document.Type
RESEARCH POLICY
2018.00
Gault, Fred
Defining and measuring innovation in all sectors of the economy
English
Article
RESEARCH POLICY
2015.00
Tavassoli, Sam; Karlsson, Charlie
Persistence of various types of innovation analyzed and explained
English
Article
RESEARCH POLICY
2019.00
Arundel, Anthony; Bloch, Carter; Ferguson, Barry
Advancing innovation in the public sector: Aligning innovation measurement with policy goals
English
Article
RESEARCH POLICY
2015.00
Meuer, Johannes; Rupietta, Christian; Backes-Gellner, Uschi
Layers of co-existing innovation systems
English
Article

Figure 31: The layout of the developed tool.

Besides the interactive menu on the left, there are also four main tabs which provide some outputs related to the data imported. In the first tab about the “Data Imported”, a table is provided with some variables provided by the WoS. T describes the variables and their attributes. Some examples are also provided so that the user can get a sense of the information displayed.

Table 9: Variables made available in tab "Source.Title".

Name of the variable	**Meaning of the data provided**	**Type of variable**	**Examples**
Source.Title	Title of the journal	Text	Nature, Science…
Publication.Year	Year of publication of the article	Numeric	1998, 2021…
Author.Full.Names	Full names of authors of the article	Text	Smith, J; Brown, W; …
Article.Title	Title of the article	Text	Attention Is All You Need; Rebuilding marine life; …
Language	Language used the articles	Text	English, Russian...
Document.Type	Type of article published	Text	Article, Review…
Volume	Number of journal’s volume where the articles was published.	Numeric	8, 12…
Issue	Number of journal’s issue where the articles was published.	Numeric	1, 4…
Start.Page	Page where the article starts.	Numeric	22, 118...
End.Page	Page where the article ends.	Numeric	35, 257…
DOI	Digital Object Identifier is a string	Text	http://dx.doi.org/10.1002/asi.43002;

	of numbers, letters and symbols used to uniquely identify an article or document, and provided with a permanent web address.		http://dx.doi.org/10.1007/s1192-012-0896-1Uddin;
WoS.Categories	The subject category of its source publication in the *Web of Science.*	Text	Management, accounting…

The frequency of articles selected for each journal per year is provided in the second tab, titled "Descriptive statistics" (see

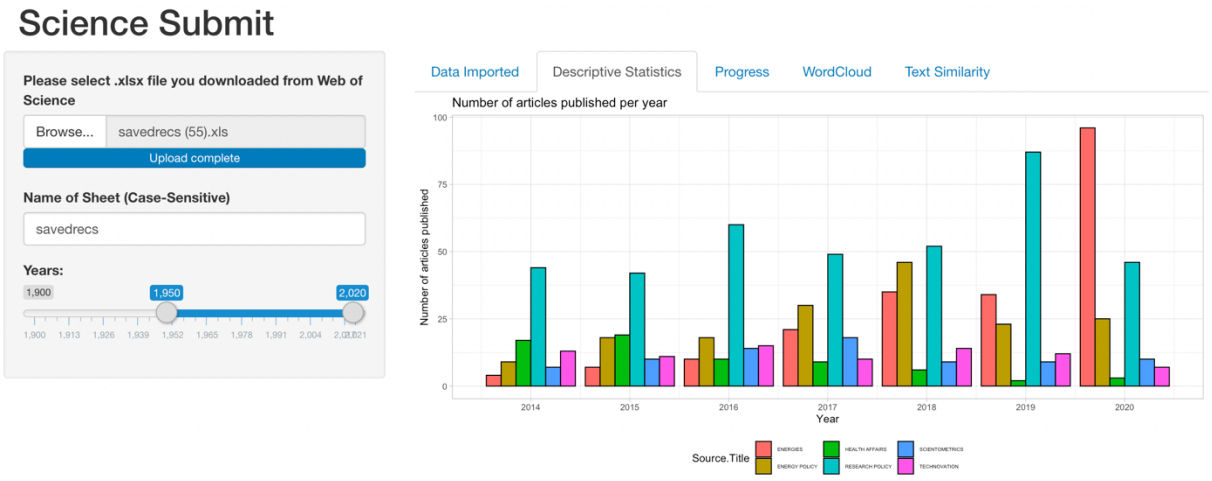


Figure 32: A printscreen of the tab "Descriptive statistics"). In the grouped barplot, the years are provided in x-axis and the frequency of the articles found in the dataset are represented in y-axis. The journals are identified in the legend below the plot with different colours.

Figure 32: A printscreen of the tab "Descriptive statistics".

Aside from absolute figures, the proportion is also available in the tab "Progress," where the bar plot displays the stacked sum of articles from each journal in the x-axis and the year in the y-axis (see

Figure 33:). This was intended to make researchers aware of the dataset composition through an easier-to-interpret graphical output. The proportions provided by journal allow user to confirm their sample uploaded.

Figure 33: A printscreen of the tab "Progress".

Based on the abstracts, which are seen as short versions of the articles (Hartley and Cabanac, 2017), a wordcloud is computed and provided in the tab “WordCloud” (see Figure 34:). The algorithm considers all the abstracts from the samples, pre-processes the text to extract words, remove punctuations, transform capital letter to small ones and reduce words to their root form so plurals and different tense verb form are considered the same. The word cloud depicts in different sizes the words found. The bigger and bolder a word appears, the more often it’s mentioned in the abstracts.

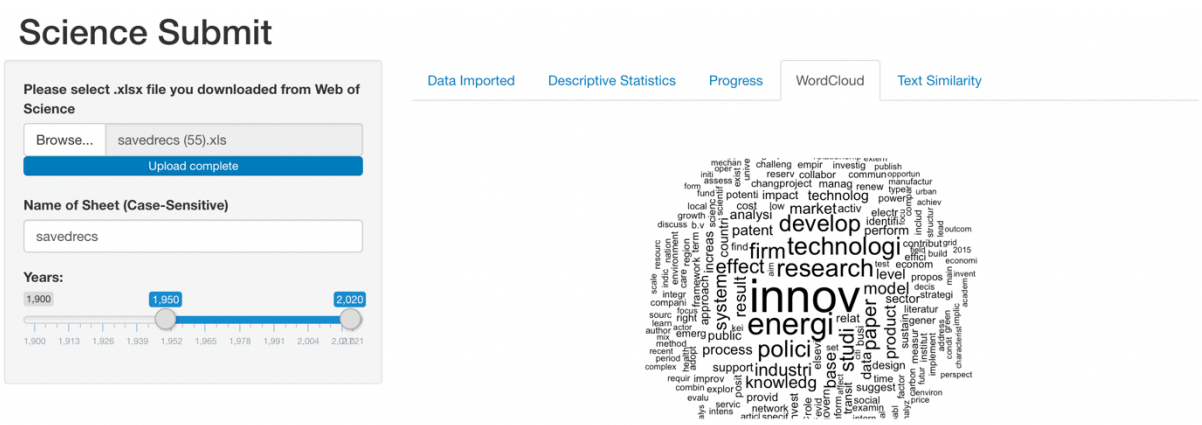


Figure 34: A printscreen of the tab "WordCloud".

The final tab, "Text Similarity," mal, allowing researchers to determine which ones are semantically more similar (see

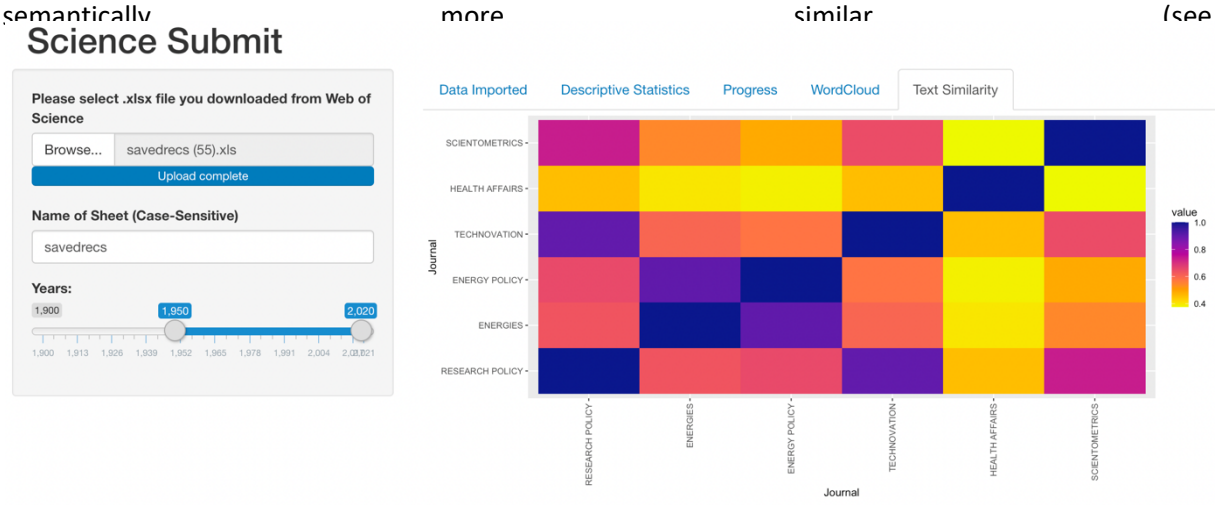


Figure 35:). Darker the colour, greatest the similarity between the abstracts from the journals being compared. For this comparison, the abstracts are processed and transformed in data frames which the algorithm uses to compare. The measure of similarity used is the cosine for not being influenced by the different sized of abstracts.

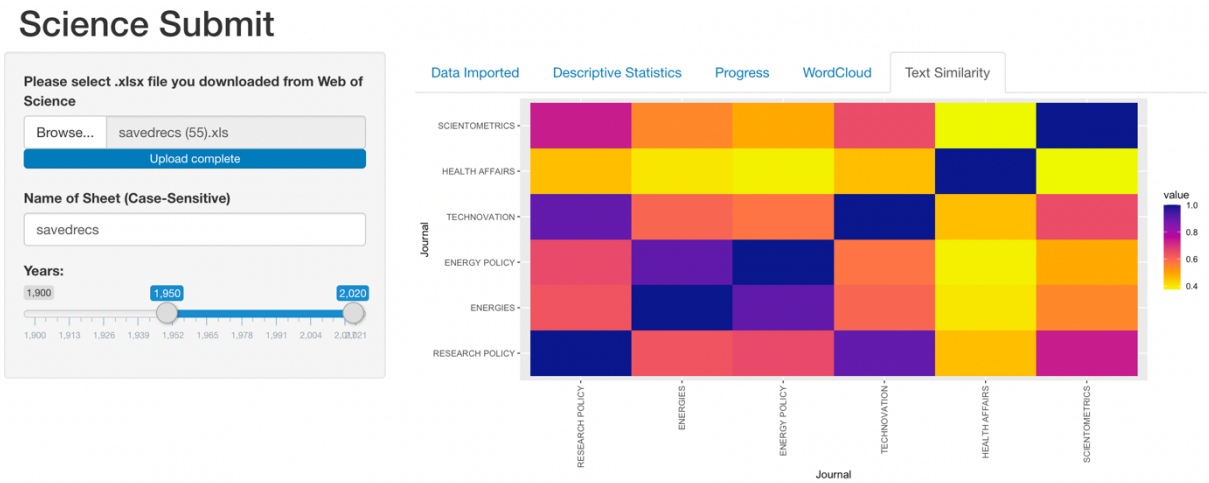


Figure 35: A printscreen of "Text Similarity" tab.

For all the tabs considered, users are allowed to change the sheet selected as well as the years range considered. The graphical content will change based on the choices made in a few seconds. If a new document is uploaded, the outputs provided will also change to display the data available in the document uploaded.

6.3. Future developments

Science Submit tool is made available to researchers all over the world online when accessing the following link: https://sciencesubmit.shinyapps.io/ShinyApp/. A dataset downloaded from the WoS is required containing the journals for comparison, the abstracts to be considered and the years of abstracts publication. This first step of data retrieval is left to the researchers so they could select the journals of their interest. The standard excel output fits the regular work of the app.

Besides WoS, many other online bibliographic databases that store metadata about scientific works, such as Scopus (http://www.scopus.com), Google Scholar (http://scholar.google.com), and Science Direct (http://www.sciencedirect.com/) can be sources of bibliographic information. Other similar databases exist for specific fields, such as Medline, Astrophysics Data System, arXiv, DBPL, or CiteSeerXPatent. They do not cover the same scientific fields and journals, so the choice is not neutral (Waltman, 2016). Because of this, we plan to add review the app presented previously so that it could also read and run data from other sources than WoS. Ultimately, we aim the app to run with data from multiple datasets and consider them all for the generation of suggestions.

CHAPTER 7

Conclusion and future work

This final chapter closes the Thesis by briefly reviewing and discussing its findings, limitations, and future work prospects.

Main conclusions

In this thesis, we conducted three interconnected works that provide different but complementary perspectives on the editorial governance debate. For each essay, distinct data was produced and analysed using unrelated techniques to address distinct goals. Original evidence, extracted and developed for this project, on journal editorial boards, journals self-advertising-materials and abstracts of published were independently scrutinized through descriptive statistics, network analysis, text mining techniques and machine learning approaches. The whole Thesis thus tries to provide a thorough and balanced perspective on lesser-known dimensions of the contemporary scientific enterprise.

In chapter 2, the identity of editorial board members was found to be a heterogeneous, male-dominated, Angloamerican-dominated editorial elite. Journal decision-makers are geographically concentrated: 20% of the countries hold 85% of the editorial positions. The presence of editors on multiple boards, more diverse national representation, and a higher proportion of women on boards appear to be associated with improved journal performance. The editorial environment appears to be made up of people who interact with one another across journal boundaries, with nearly 20% of editors serving on more than one board. The conventional network centrality measures (Degree, Closeness, and Betweenness) reaffirm *Research Policy* as the primary reference in Innovation Studies, with *Industrial and Corporate Change* coming in second. This essay also discovered a proliferation of editorial duty labelling within the same epistemic community: a phenomenon that, to the best of our knowledge, has gone unnoticed in the literature. There is also a lack of evidence regarding structural and behavioural accountability: the process for selecting editors, as well as the procedures for appealing decisions, is largely unclear. If information about editors, the editorial process, and journal governance appears to be lacking in the journal's frontmatter, more can be done to make these matters more accessible to the community.

Chapter 3 confirmed *Research Policy* as the leading outlet in the field and when considered as the "gold standard" in innovation-related research, a semantic alternative to conventional metrics, i.e., a novel "soft" ranking of journals, was proposed. This discourse-driven or substance-sensitive ordering of journals is quite different from the usual, citation-dependent rankings methods. It yields the identification of six other journals (*Technological Forecasting and Social Change*, C*ambridge Journal of Economics*, *International Journal of Technology Management*, *Technology Analysis & Strategic Management*, *Industrial and Corporate Change* and *Technovation*) that follow a pattern not connected to the impact factor ranking. The findings for the case of *Innovation Studies* suggest that the methodology could be applied to other fields so that researchers can better comprehend, compare, and classify serial literature channels in their fields of interest.

Chapter 4 analysed the semantic similarity between journals through the clusters formed which showed to be different when considering blurbs or abstracts. To understand which journals had a self-presentation material different from the contents published, we developed, trained, and validated a supervised learning model which proved to have 80% of accuracy to match blurbs with abstracts selected during the last ten years. Five journal blurbs were found to be more related to the topics chosen for publication in other venues than their owns. This meant 25% of the journals were describing interests matching other journals contents rather than their own. Although it is commonly assumed that the interests described in the blurb section are consistent with the content of regular issues, this machine learning approach proved to be useful in testing that assumption and may provide help in guiding authors, readers, and policymakers about journals' actual publishing interests.

The results obtained with the development of this new algorithm aimed to compare contents published between journals were also applied in Chapter 5 to the construction of a new tool named as *Science Submit*. It was also developed and made public online so that researchers all over the world could semantically compare the articles published by the journals of their choice and get advice on the outlets to submit their work. This tool can be accessed on: https://sciencesubmit.shinyapps.io/ShinyApp/. This discourse-dependent approach yields an alternative way to systematically compare journals and could also be useful to science managers and policymakers in assessing and monitoring journal publishing patterns over time.

With this work, we make a modest proposal for incorporating 'editormetrics' and 'journalytics' perspectives in scientometrics at large and into the ongoing efforts to map and measure *Innovation Studies*. The findings of our study persuade us to develop further analyse the origins of the gatekeeping crowd, as well as the utility of previously published articles and journal self-presentation material as a guide for content assessment and comparison. We understand that the debate over editorial policies may appear to be a big challenge. However, in line with the quotation from Heraclitus at the start of this Thesis, researchers change day by day with the contents they read and the experiences they gain, as do the journals to which we submit our work. Only further research could help us better understand the editorial processes and develop more effective policies that could yield great benefits to the scientific community.

Limitations, opportunities, and implications

This Thesis has a number of limitations and leaves open the possibility of better describing, revising, and explaining some patterns. The journal list in Fagerberg et al. (2012) can be made more customised (expanded or trimmed) to innovation-oriented journals as well as updated. Titles could be harvested from the 40 titles in the field of *Innovation Studies* ranked on the ABS list (the last one published by the Chartered Association of Business Scholars, Academic Journal Guide 2021) and derived from systematic surveys (e.g., Rakas and Hain, 2019; Kashani and Roshani, 2019). This effort could see the inclusion of outlets like *Industry and Innovation*, the *Journal of Technology Transfer*, *Science and Public Policy*, *Foresight and STI Governance* and the *Journal of Innovation Management*. The last one, for instance, was launched in 2013 (that is, after Fagerberg et al., 2012, the source of our work).

Future research may consider contacting journals for updated board members list to obtain the most up-to-date and accurate data about the editorial board members, as there is a common gap between new members being welcomed on board and the official website being updated accordingly. Furthermore, conducting a questionnaire survey with editors, focusing on questions about recruiting, actual work, and the level of perceived impact of their role on the field, could provide us with additional data to study other aspects of those gatekeepers responsible for editorial decisions and the process of their duties at journal boards. So far, the interest in Editormetrics as a supplement to bibliometrics has resulted in a few changes. Elsevier journals have already published gender information for editors, and Cell Press has established an "Inclusion & Diversity" statement as a step toward parity. *Regional Studies* used an innovative method for hiring editors by publishing a statement about their search for a new editor, and the *Journal of Hazardous Materials* intended to reward exceptional early-career scholars in order to encourage them to pursue careers as editors.

In recent years, advertising sections, the *blurbs*, have been added to journals with the purpose of attracting researchers' attention and influence their decision, but little is known about their stability. When we checked the blurbs two years later, in January 2021, we discovered that four out of twenty had changed: *Journal of Management Studies*, *Management Science*, *Journal of International Business Studies*, and *Small Business Economics*. Some lengthened the blurb (for example, *Small Business Economics*), while others shortened it (like *Management Science*). Some re-engineered sentences to make them more reader-centric (*Management Science*, changing "we encourage cross-functional, multidisciplinary research..." to "readers will also find cross-functional, multidisciplinary research"), while others spiced-up their marketing appeal (*Journal of International Business Studies*, promoting its FT50 kudus), and still others highlight topic evolution (*Strategic Management Journal*, including "entrepreneurial process"). Over two years, a 20% churn was observed, illustrating a rather poorly understood phenomenon: there is dynamics in journal discourse, it may not be evenly distributed across different groups of journals, and blurb dynamics is likely worth monitoring in future "journalytics" studies.

The content-sensitive approaches used in this exploratory analysis to compare journals and find those with similar interests based on marketing pronouncements or abstracts already approved for publication revealed that there is underutilized data available that could be used to improve our understanding of journals. A methodology similar to the one proposed here could be useful in a variety of other research areas. Those that can identify a set of specialized journals could also benefit from a similar approach. Some validation studies would be desirable in the future to demonstrate the degree of accuracy and the adjustments required to make similar tools come true in other fields.

Finally, debates on contemporary science must explicitly consider journals as institutions and, within journals, appraise their governance mechanisms. Boards and their internal rules are still unknown. An open policy would be a desirable safeguard against potential bias and would ensure that both readers and science managers have enough information considering the studies and articles published in the pages of journals. We see three goals for the better practice of editorial governance: 1) understanding how journals actually work and are managed by editors; 2) making sense and keeping track of journal mission statements; 3) monitoring and reflexively examine what journals actually publish.

References

Aguinis, H., Gottfredson, R.K., Culpepper, S.A., Dalton, D.R., de Bruin, G.P., 2012. Doing Good and Doing Well: On the Multiple Contributions of Journal Editors. Academy of Management Learning & Education 12, 564–578.

Alberts, B., Hanson, B., Kelner, K.L., 2008. Reviewing peer review. Science (1979).

Altbach, P.G., Wit, H. de, 2018. Too much academic research is being published. International Higher Education 96, 2–3.

Amrein, K., Langmann, A., Fahrleitner-Pammer, A., Pieber, T.R., Zollner-Schwetz, I., 2011. Women underrepresented on editorial boards of 60 major medical journals. Gender Medicine 8, 378–387.

Andersen, P.D., Alkærsig, L., 2016. Profile and Trends of FTA and Foresight. In: Anticipating Future Innovation Pathways Through Large Data Analysis. Springer, pp. 45–58.

Andrikopoulos, A., Economou, L., 2015. Editorial board interlocks in financial economics. International Review of Financial Analysis 37, 51–62.

Archambault, É., Larivière, V., 2009. History of the journal impact factor: Contingencies and consequences. Scientometrics 79, 635–649.

Aria, M., Cuccurullo, C., 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics 11, 959–975.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection *. Statistics Surveys 4, 40–79.

Baccini, A., Barabesi, L., 2010. Interlocking editorship. A network analysis of the links between economic journals. Scientometrics 82, 365–389.

Baccini, A., Barabesi, L., 2011. Seats at the table: The network of the editorial boards in information and library science. Journal of Informetrics 5, 382–391.

Baccini, A., Barabesi, L., 2014. Gatekeepers of economics: the network of editorial boards in economic journals. In: The Economics of Economists: Institutional Setting, Individual Incentives, and Future Prospects. Cambridge University Press.

Baccini, A., Barabesi, L., Khelfaoui, M., Gingras, Y., 2020. Intellectual and social similarity among scholarly journals: an exploratory comparison of the networks of editors, authors and co-citations. Quantitative Science Studies.

Baccini, A., Barabesi, L., Marcheselli, M., 2009. How are statistical journals linked? A network analysis. Chance 22, 35–45.

Bakker, P., Rigter, H., 1985. Editors of medical journals: Who and from where. Scientometrics 7, 11–22.

Ball, P., 2005. Citation-based league table determines scientists' pecking order. Nature 436, 900.

Banbura, M., Modugno, M., 2014. Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data. Journal of Applied Econometrics 29, 133–160.

Banerjee, T., Peterson, M., Oliver, Q., Froehle, A., Lawhorne, L., 2018. Validating a commercial device for continuous activity measurement in the older adult population for dementia management. Smart Health 5–6, 51–62.

Barabási, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science (1979) 286, 509–512.

Basturkman, H., 1999. A Content Analysis of Elt Textbook Blurbs: Reflections of Theory-in-Use. RELC Journal 30, 18–38.
Bateman, S., Gutwin, C., Nacenta, M., 2008. Seeing things in the clouds: The effect of visual features on tag cloud selections. HYPERTEXT'08: Proceedings of the 19th ACM Conference on Hypertext and Hypermedia, HT'08 with Creating'08 and WebScience'08 193–202.
BEAUCHAMP, M.A., 1965. AN IMPROVED INDEX OF CENTRALITY. Behavioral science 10, 161–163.
Bedeian, A.G., van Fleet, D.D., Hyman, H.H., 2009. Scientific achievement and editorial board membership. Organizational Research Methods 12, 211–238.
Beel, J., Gipp, B., Langer, S., Breitinger, C., 2016. Research-paper recommender systems: a literature survey. International Journal on Digital Libraries 17, 305–338.
Belderbos, R., Lokshin, B., Boone, C., Jacob, J., 2020. Top management team international diversity and the performance of international R&D. Global Strategy Journal.
Bergamaschi, S., Guerra, F., Leiba, B., 2010. Guest editors' introduction: Information overload. IEEE Internet Computing 14, 10–13.
Bergstrom, C., 2007. Eigenfactor. College and Research Libraries News.
Bhatia, V., 2004. Worlds of Written Discourse : A Genre-Based View, 1st ed, Worlds of Written Discourse : A Genre-Based View. Bloomsbury Academic.
Biagioli, M., Kenney, M., Martin, B.R., Walsh, J.P., 2019. Academic misconduct, misrepresentation and gaming: A reassessment. Research Policy 48, 401–413.
Biagioli, M., Lippman, A., 2020. Gaming the Metrics: Misconduct and Manipulation in Academic Research. MIT Press.
Björk, B.-C., Annikki, R., 2008. Global annual volume of peer reviewed scholarly articles and the share available via different Open Access options. In: Open Scholarship: Authority, Community, and Sustainability in the Age of Web 2.0 - Proceedings of the 12th International Conference on Electronic Publishing. Toronto, Canada.
Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008, P10008.
Bollen, J., Rodriquez, M.A., van de Sompel, H., 2006. Journal status. Scientometrics 69, 669–687.
Bollen, J., van de Sompel, H., Hagberg, A., Bettencourt, L., Chute, R., Rodriguez, M.A., Balakireva, L., 2009a. Clickstream data yields high-resolution Maps of science. PLoS ONE 4, e4803.
Bollen, J., van de Sompel, H., Hagberg, A., Chute, R., 2009b. A Principal Component Analysis of 39 Scientific Impact Measures. PLoS ONE 4, e6022.
Bonacich, P., 1987. Power and Centrality: A Family of Measures. American Journal of Sociology 92, 1170–1182.
Borgatti, S.P., Foster, P.C., 2003. The network paradigm in organizational research: A review and typology. Journal of Management 29, 991–1013.
Börner, K., 2010. Atlas of Science. The MIT Press 47405.
Bornmann, L., Daniel, H.-D., 2007. What do we know about the h index? Journal of the American Society for Information Science and Technology 58, 1381–1385.
Bornmann, L., Marx, W., 2015. Methods for the generation of normalized citation impact scores in bibliometrics: Which method best reflects the judgements of experts? Journal of Informetrics 9, 408–418.

Bornmann, L., Mutz, R., 2015. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. J Assoc Inf Sci Technol 66, 2215–2222.
Braun, T., 2005. Keeping the Gates of Science Journals. In: Handbook of Quantitative Science and Technology Research. Springer Netherlands, pp. 95–114.
Braun, T., Dióspatonyi, I., 2005. The counting of core journal gatekeepers as science indicators really counts. The scientific scope of action and strength of nations. Scientometrics 62, 297–319.
Braun, T., Glänzel, W., Schubert, A., 2013. A Hirsch-type index for journals. Scientometrics 2006 69:1 69, 169–173.
Breiman, L., 1996. Bagging predictors. Machine Learning 24, 123–140.
Breiman, L., 2001. Random forests. Machine Learning 45, 5–32.
Brinn, T., Jones, M.J., 2008. The composition of editorial boards in accounting: A UK perspective. Accounting, Auditing & Accountability Journal 21, 5–35.
Broadus, R.N., 1987. Toward a definition of "bibliometrics." Scientometrics 12, 373–379.
Bryce, C., Dowling, M., Lucey, B., 2020. The journal quality perception gap. Research Policy 49, 103957.
Burgess, T.F., Shaw, N.E., 2010. Editorial Board Membership of Management and Business Journals: A Social Network Analysis Study of the Financial Times 40. British Journal of Management 21, 627–648.
Bush, V., 1945. Science: the Endless Frontier. Washington.
Cabanac, G., 2012. Shaping the landscape of research in information systems from the perspective of editorial boards: A scientometric study of 77 leading journals. Journal of the American Society for Information Science and Technology 63, 977–996.
Callaway, E., 2016. Beat it, impact factor! Publishing elite turns against controversial metric. Nature.
Caraça, J., Ferreira, J.L., Mendonça, S., 2007. A chain-interactive innovation model for the learning economy: Prelude for a proposal (No. 12-2007/DE), DE Working papers.
Cardoso, J.R., Pereira, L.M., Iversen, M.D., Ramos, A.L., 2014. What is gold standard and what is ground truth? Dental Press Journal of Orthodontics 19, 27–30.
Caruana, R., Niculescu-Mizil, A., 2006. An empirical comparison of supervised learning algorithms. In: ACM International Conference Proceeding Series. pp. 161–168.
Castellacci, F., Grodal, S., Mendonca, S., Wibe, M., 2005. Advances and challenges in innovation studies. Journal of Economic Issues 39, 91–121.
Castellani, B., Hafferty, F.W., 2009. Mapping complexity. Understanding Complex Systems 2009, 241–261.
Chan, K.C., Fok, R.C.W., 2003. Membership on editorial boards and finance department rankings. Journal of Financial Research 26, 405–420.
Chan, K.C., Fung, H.-G., Lai, P., 2005. Membership of editorial boards and rankings of schools with international business orientation. Journal of International Business Studies 2005 36:4 36, 452–469.
Chartrand, G., Cheng, P.M., Vorontsov, E., Drozdzal, M., Turcotte, S., Pal, C.J., Kadoury, S., Tang, A., 2017. Deep learning: A primer for radiologists. Radiographics.

Chavarro, D., Tang, P., Ràfols, I., 2017. Why researchers publish in non-mainstream journals: Training, knowledge bridging, and gap filling. Research Policy 46, 1666–1680.
Chen, P., Xie, H., Maslov, S., Redner, S., 2007. Finding scientific gems with Google's PageRank algorithm. Journal of Informetrics 1, 8–15.
Chesbrough, H., Birkinshaw, J., Teubal, M., 2006. Introduction to the research policy 20th anniversary special issue of the publication of "Profiting from Innovation" by David J. Teece. Research Policy 35, 1091–1099.
Cimoli, M., Dosi, G., Yu, X., 2020. Industrial policies, patterns of learning, and development - An evolution perspective. In: The Oxford Handbook of Industrial Policy. pp. 93–124.
Cohen, W.W., Singer, Y., 1996. Context-sensitive learning methods for text categorization. In: SIGIR Forum (ACM Special Interest Group on Information Retrieval). pp. 307–316.
Cristianini, N., Shawe-Taylor, J., 2000. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press.
de Solla Price, D.J., 1961. Science since Babylon. Yale U.P.
de Solla Price, D.J., 1965. Littel Science, Big Science. Columbia University Press.
Dhanani, A., Jones, M.J., 2017. Editorial boards of accounting journals: gender diversity and internationalisation. Accounting, Auditing and Accountability Journal 30, 1008–1040.
Dick, S., 2019. Artificial Intelligence. Harvard Data Science Review 1.
Dickersin, K., Fredman, L., Flegal, K.M., Scott, J.D., Crawley, B., 1998. Is there a sex bias in choosing editors? Epidemiology journals as an example. J Am Med Assoc 280, 260–264.
Dosi, G., 1988. Sources, Procedures, and Microeconomic Effects of Innovation. Journal of Economic Literature 26, 1120–1988.
Dosi, G., 2000. Innovation, Organization and Economic Dynamics: Selected Essays. Edward Elgar.
Dosi, G., Malerba, F., Ramello, G.B., Silva, F., 2006. Information, appropriability, and the generation of innovative knowledge four decades after Arrow and Nelson: an introduction. Industrial and Corporate Change 15, 891–901.
Drivas, K., Kremmydas, D., 2020. The Matthew effect of a journal's ranking. Research Policy 49, 103951.
Egghe, L., 2005. Expansion of the field of informetrics: Origins and consequences. Information Processing and Management 41, 1311–1316.
Evans, J.A., 2013. Future science. Science (1979).
Evans, J.A., Foster, J.G., 2011. Metaknowledge. Science (1979) 331, 721–725.
Fagerberg, J., 2018. Innovation, Economic Development and Policy: Selected Essays. Edward Elgar.
Fagerberg, J., Fosaas, M., Bell, M., Martin, B.R., 2011. Christopher Freeman: social scienc entrepreneur. Research Policy 40, 897–916.
Fagerberg, J., Fosaas, M., Sapprasert, K., 2012. Innovation: Exploring the knowledge base. Research Policy 41, 1132–1153.
Fagerberg, J., Godinho, M.M., 2004. Innovation and Catching-Up. In: The Oxford Handbook of Innovation.

Fagerberg, J., Laestadius, S., Martin, B.R., 2018. The Three Great Issues Confronting Europe – Economic, Environmental and Political. In: Innovation Systems, Policy and Management. Cambridge University Press, pp. 464–491.
Fagerberg, J., Verspagen, B., 2009. Innovation studies—The emerging structure of a new scientific field. Research Policy 38, 218–233.
Fagerberg, J., Verspagen, B., 2020. Innovation–diffusion, the economy and contemporary challenges: a comment. Industrial and Corporate Change 29, 1067–1073.
Fagerberg, Jan., Martin, B.R., Andersen, E.S., 2013. Innovation studies: evolution and future challenges. Oxford University Press.
Faria, J.R., 2005. The Game Academics Play: Editors versus Authors. Bulletin of Economic Research 57, 1–12.
Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D., Fernández-Delgado, A., 2014. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?, Journal of Machine Learning Research.
Fischer, E., Gopaldas, A., Scaraboto, D., 2017. Why papers are rejected and how to get yours accepted: Advice on the construction of interpretive consumer research articles. Qualitative Market Research 20, 60–67.
Forrester, A., Björk, B.-C., Tenopir, C., 2017. New web services that help authors choose journals. Learned Publishing 30, 281–287.
Fortunato, S., 2010. Community detection in graphs. Physics Reports.
Fortunato, S., Bergstrom, C.T., Börner, K., Evans, J.A., Helbing, D., Milojević, S., Petersen, A.M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., Barabási, A.L., 2018. Science of science. Science (1979) 359.
Freeman, C., 1969. Measurement of output of research and experimental development: A review paper, OECD, Statistical Reports and Studies. Paris.
Freeman, C., 1982. The Economics of Industrial Innovation, 2nd ed. Francis Pinter, London.
Freeman, C., 2008. Systems of Innovation: Selected Essays in Evolutionary Economics. Edward Elgar.
Freeman, Chris, Soete, L., 1997. The Economics of Industrial Innovation, 1st Edition. ed. The MIT Press.
Freeman, C., Soete, L., 2009. Developing science, technology and innovation indicators: What we can learn from the past. Research Policy 38, 583–589.
Freeman, Christopher, Soete, Luc., 1997. The Economics of Industrial Innovation. MIT Press.
Freeman, L.C., 1979. Centrality in social networks conceptual clarification. Social Networks 1.
Galligan, F., Dyas-Correia, S., 2013. Altmetrics: Rethinking the Way We Measure. Serials Review 39, 56–61.
García-Carpintero, E., Granadino, B., Plaza, L., 2010. The representation of nationalities on the editorial boards of international journals and the promotion of the scientific output of the same countries. Scientometrics 84, 799–811.
Garfield, E., 1972. Citation Analysis as a Tool in Journal Evaluation. Science (1979) 178, 471–479.
Garfield, E., 1982. In Tribute to V.V. Nalimov: Renaissance Scholar and Scientometrician Par Excellence. Essays of an Information Scientist 5, 417–427.
Garfield, E., 1983. How to Use Journal Citation Reports. Essays of an Information Scientist 6, 131–134.

Garfield, E., 2006. The history and meaning of the journal impact factor. JAMA 295, 90–93.
Gea Valor, M.L., 2005. Advertising books: a linguistic analysis of blurbs. Ibérica 10, 41–62.
Gesuato, S., 2007. Evaluation in Back-Cover Blurbs. Textus XX, 83–102.
Ghatak, A., 2017. Machine Learning with R, Machine Learning with R. Springer Singapore.
Glänzel, W., Chi, P.S., 2020. The big challenge of Scientometrics 2.0: exploring the broader impact of scientific research in public health. Scientometrics 1–21.
Glänzel, W., Moed, H.F., Schmoch, U., Thelwall, M., 2019. Springer Handbook of Science and Technology Indicators. Springer International Publishing.
Glänzel, W., Schlemmer, B., Thijs, B., 2003. Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon. Scientometrics 58, 571–586.
Godin, B., 2019. The Invention of Technological Innovation: Languages, Discourses and Ideology in Historical Perspective. Edward Elgar, Cheltenham, UK.
González-Albo, B., Bordons, M., 2011. Articles vs. proceedings papers: Do they differ in research relevance and impact? A case study in the Library and Information Science field. Journal of Informetrics 5, 369–381.
González-Pereira, B., Guerrero-Bote, V.P., Moya-Anegón, F., 2010. A new approach to the metric of journals' scientific prestige: The SJR indicator. Journal of Informetrics 4, 379–391.
Gopalakrishnan, S., Damanpour, F., 1997. A review of innovation research in economics, sociology and technology management. Omega (Westport) 25, 15–28.
Goyanes, M., de-Marcos, L., 2020. Academic influence and invisible colleges through editorial board interlocking in communication sciences: a social network analysis of leading journals. Scientometrics 123.
Greenbaum, H.K., Goodsir, H.L., Smith, M.C., Robinson, D.H., 2018. Female Participation as Top-Producing Authors, Editors, and Editorial Board Members in Educational Psychology Journals from 2009 to 2016. Educational Psychology Review 30, 1283–1289.
Griffiths, P., Norman, I., 2016. Why was my paper rejected? Editors' reflections on common issues which influence decisions to reject papers submitted for publication in academic nursing journals. International Journal of Nursing Studies 57, A1–A4.
Grillitsch, M., Schubert, T., Srholec, M., 2019. Knowledge base combinations and firm growth. Research Policy 48, 234–247.
Gu, X., Blackmore, K.L., 2016. Recent trends in academic journal growth. Scientometrics 108, 693–716.
Günther, F., Dudschig, C., Kaup, B., 2014. LSAfun - An R package for computations based on Latent Semantic Analysis. Behavior Research Methods 47, 930–944.
HaCohen-Kerner, Y., Miller, D., Yigal, Y., 2020. The influence of preprocessing on text classification using a bag-of-words representation. PLoS ONE 15.
Hafeez, D.M., Waqas, A., Majeed, S., Naveed, S., Afzal, K.I., Aftab, Z., Zeshan, M., Khosa, F., 2019. Gender distribution in psychiatry journals' editorial boards worldwide. Comprehensive Psychiatry 94, 152119.

Hall, J., Martin, B.R., 2019a. Towards a taxonomy of research misconduct: The case of business school research. Research Policy 48, 414–427.
Hall, J., Martin, B.R., 2019b. Towards a taxonomy of research misconduct: The case of business school research. Research Policy 48, 414–427.
Hall, R., 2016. Writing Book Blurbs and Synopses: Professional Techniques for Fiction Authors. CreateSpace Independent Publishing Platform.
Hames, I., 2001. Editorial boards: realizing their potential. Learned Publishing 14, 247–256.
Han, J., Kamber, M., Pei, J., 2012. Data Mining: Concepts and Techniques, Data Mining: Concepts and Techniques. Elsevier Inc.
Hansen, D.L., Shneiderman, B., Smith, M.A., Himelboim, I., 2020. Calculating and visualizing network metrics. In: Analyzing Social Media Networks with NodeXL. Elsevier.
Hanusch, H., Pyka, A., 2007. Elgar companion to neo-Schumpeterian economics . Edward Elgar, Cheltenham.
Harris, C.A., Banerjee, T., Cramer, M., Manz, S., Ward, S.T., Dimick, J., Telem, D.A., 2019. Editorial (Spring) Board? Gender Composition in High-impact General Surgery Journals Over 20 Years. Annals of Surgery 269, 582–588.
Hartley, J., Cabanac, G., 2017. Thirteen Ways to Write an Abstract. Publications 2017, Vol. 5, Page 11 5, 11.
Harzing, A.-W., 2012. How to become an author of ESI Highly Cited Papers? University of Melbourne.
Harzing, A.W., Metz, I., 2011. Gender and geographical diversity in the editorial board of the Journal of International Business Studies. AIB Insights 11, 3–7.
Harzing, A.W., Metz, I., 2012. Explaining geographic diversity of editorial boards: The role of conference participation and English-language skills. European Journal of International Management 6, 697–715.
Harzing, A.W., Metz, I., 2013. Practicing what We Preach: The Geographic Diversity of Editorial Boards. Management International Review 53, 169–187.
Heidari, S., Babor, T.F., de Castro, P., Tort, S., Curno, M., 2016. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. Research Integrity and Peer Review 1.
Hicks, D., 2004. The four literatures of social science. In: Handbook of Quantitative Science and Technology Research. pp. 473–496.
Hinterberger, H., Domingo-Ferrer, J., Kashyap, V., Khatri, V., Snodgrass, R.T., Terenziani, P., Koubarakis, M., Zhang, Y., Joshi, J.B.D., Gamper, J., Böhlen, M., Jensen, C.S., Tansel, A.U., Böhlen, M., Böhlen, M.H., Jensen, C.S., Snodgrass, R.T., Khatri, V., Revesz, P., Mamoulis, N., Jensen, C.S., Snodgrass, R.T., Jensen, C.S., Snodgrass, R.T., Wijsen, J., Jensen, C.S., Snodgrass, R., Jensen, C.S., Snodgrass, R.T., Jensen, C.S., Snodgrass, R.T., Bettini, C., Wang, X.S., Jajodia, S., Jensen, C.S., Snodgrass, R.T., Dyreson, C., Wijsen, J., Gao, D., Chomicki, J., Toman, D., Shoshani, A., Combi, C., Terenziani, P., Jensen, C.S., T. Snodgrass, R., Jensen, C.S., Snodgrass, R.T., Böhlen, M., Chomicki, J., Toman, D., Jensen, C.S., Snodgrass, R.T., Torp, K., F. Roddick, J., Toman, D., Schiel, U., Silva, S.F., Dyreson, C., Grandi, F., Plachouras, V., Lalmas, M., El-Khair, I.A., Carterette, B., Shen, D., Li, H., Ferragina, P.I., Nitto, I., Zhang, L., Sun, J.-T., Navarro, G., Huang, H., Zhang, B., De Moura, E.S., Cai, Y., Sun, J.-T., Srinivasan, P., Yan, J., Huang, H., Zhang, B., Yan, J., Hu, J., Liu, N., Shen, D., Huang, H., Zhang, B., El-Khair, I.A., Hinterberger, H., Arenas, M., Breunig, M., Al-Houmaily, Y.J., Samaras,

G., Mankovskii, S., George, B., Shekhar, S., Alonso, O., Gertz, M., Montanari, A., Chomicki, J., Øhrstrøm, P., Hasle, P.F. V., Jensen, C.S., Snodgrass, R.T., Jensen, C.S., T. Snodgrass, R., Lorentzos, N.A., Gao, L., Wang, X.S., Jensen, C.S., Snodgrass, R.T., Dyreson, C., Jensen, C.S., T. Snodgrass, R., Liu, N., Caverlee, J., Jacobsen, H.-A., Marian, A., Hoel, E., Felice, P. Di, Clementini, E., Kemme, B., Güting, R.H., Vossen, G., Shasha, D., Vossen, G., Reuter, A., Vossen, G., Jensen, C.S., Snodgrass, R.T., Alonso, G., Schuldt, H., Moro, M.M., Tsotras, V.J., Manolopoulos, Y., Theodoridis, Y., Tsotras, V.J., Fekete, J.-D., Novák, V., Floriani, L. De, Magillo, P., Crochemore, M., Lecroq, T., Despotovic, Z., Agarwal, N., Liu, H., Sion, R., Snodgrass, R.T., Bonnet, P., Shasha, D., Fagin, R., Chen, L., Al-Houmaily, Y.J., Samaras, G., Lechtenbörger, J., Lausen, G., Amati, G., Jacobsen, H.-A., 2009. Text Segmentation. In: Encyclopedia of Database Systems. Springer US, Boston, MA, pp. 3072–3075.

Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 102, 16569–16572.

Hodgson, G.M., Rothman, H., 1999. The Editors and Authors of Economics Journals: a Case of Institutional Oligopoly? The Economic Journal 109, 165–186.

Hood, W.W., Wilson, C.S., 2001. The Literature of Bibliometrics, Scientometrics, and Informetrics. Scientometrics 52, 291–314.

Horan, J.J., Weber, W.L., Fitzsimmons, P., Maglio, C.J., Hanish, C., 1993. Further Manifestations of the MOMM Phenomenon. The Counseling Psychologist 21, 278–287.

Hymowitz, C., Schelhardt, T.D., 1986. The Glass-Ceiling Why Women Can't Seem to Break the Invisible Barrier that Blocks Them from Top Jobs. The Wall Street Journal 57, D4–D5.

Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: A review. In: ACM Computing Surveys. ACM PUB27 New York, NY, USA , pp. 264–323.

James, C., Colledge, L., Meester, W., Azoulay, N., Plume, A., 2018. CiteScore metrics: Creating journal metrics from the Scopus citation index. Learned Publishing 32, 367–374.

Jarrahi, M.H., Sawyer, S., 2019. Networks of innovation: the sociotechnical assemblage of tabletop computing. Research Policy: X 1, 100001.

Jinha, A.E., 2010. Article 50 million: an estimate of the number of scholarly articles in existence. Learned Publishing 23, 258–263.

Journal Selector [WWW Document], n.d. URL https://cofactorscience.com/journal-selector (accessed 9.15.20).

Juzych, M.S., Shin, D.H., Coffey, J.B., Parrow, K.A., Tsai, C.S., Briggs, K.S., 1991. Pattern of Publication of Ophthalmic Abstracts in Peer-reviewed Journals. Ophthalmology 98, 553–556.

Kang, N., Doornenbal, M., Schijvenaars, B., 2015. Elsevier journal finder: Recommending journals for your paper. In: RecSys 2015 - Proceedings of the 9th ACM Conference on Recommender Systems. Association for Computing Machinery, Inc, New York, New York, USA, pp. 261–264.

Kathpalia, S.S., 1997. Cross-cultural Variation in Professional Genres: A Comparative Study of Book Blurbs. World Englishes 16, 417–426.

Khadilkar, S.S., 2018. Rejection Blues: Why Do Research Papers Get Rejected? Journal of Obstetrics and Gynecology of India.

Kocher, M.G., Sutter, M., 2001. The Institutional Concentration of Authors in Top Journals of Economics During the Last Two Decades. The Economic Journal 111, 405–421.
Kohavi, R., 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. International Joint Conference on Artificial Intelligence 14, 1137–1143.
Kotsemir, M.N., Abroskin, A., 2013. Innovation Concepts and Typology – An Evolutionary Discussion. SSRN Electronic Journal.
Kotu, V., Deshpande, B., 2019. Data science : concepts and practice, 2nd ed.
Kovanis, M., Porcher, R., Ravaud, P., Trinquart, L., 2016. The Global Burden of Journal Peer Review in the Biomedical Literature: Strong Imbalance in the Collective Enterprise. PLOS ONE 11, e0166387.
Krishnan, C.N.V., Bricker, R., 2004. Top finance journals: Do they add value? Journal of Economics and Finance 28, 361–378.
Landauer, T.K., Laham, D., Foltz, P., 1998. Learning Human-like Knowledge by Singular Value Decomposition: A Progress Report.
Landhuis, E., 2016. Scientific literature: Information overload. Nature.
Larivière, V., Sugimoto, C.R., 2019. The Journal I 3 Part A | 1.1 1. The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects.
Larsen, B., Aone, C., 1999. Fast and effective text mining using linear-time document clustering 16–22.
Leo, J.R. di, 2016. Blurbs in the Post-Truth Era. American Book Review 38, 2–10.
Lerchenmueller, M.J., Sorenson, O., 2018. The gender gap in early career transitions in the life sciences. Research Policy 47, 1007–1017.
Lerner, J., Stern, S., 2013. The Rate and Direction of Inventive Activity Revisited, The Rate and Direction of Inventive Activity Revisited. University of Chicago Press.
Lewis, D.D., Schapire, R.E., Callan, J.P., Papka, R., 1996. Training algorithms for linear text classifiers. In: SIGIR Forum (ACM Special Interest Group on Information Retrieval). pp. 298–306.
Leydesdorff, L., 2007. Betweenness centrality as an indicator of the interdisciplinary of scientific journals. Journal of the American Society for Information Science and Technology 58, 1303–1319.
Liu, N., Shapira, P., Yue, X., 2021. Tracking developments in artificial intelligence research: constructing and applying a new search strategy. Scientometrics 1–40.
Liwei, Z., Chunlin, J., 2015. Social Network Analysis and Academic Performance of the Editorial Board Members for Journals of Library and Information Science. Collnet Journal of Scientometrics and Information Management 9, 131–143.
Louçã, F., 2020. Chris Freeman forging the evolution of evolutionary economics. Industrial and Corporate Change 29, 1037–1046.
Louçã, J., Rosa, C., Guimarães, F., Meneses, V., 2004. MULTI-AGENT PROPOSITIONS TO MANAGE ORGANIZATIONAL KNOWLEDGE - Position paper concerning a three-dimensional research project. In: Proceedings of the 4th International Workshop on Pattern Recognition in Information Systems. SciTePress - Science and and Technology Publications, pp. 436–441.
Lundvall, B.-Å., 2014. Innovation Studies: A Personal Interpretation of 'The State of the Art.' In: Innovation Studies. Oxford University Press, pp. 20–70.

Mackey, T.K., Shah, N., Miyachi, K., Short, J., Clauson, K., 2019. A Framework Proposal for Blockchain-Based Scientific Publishing Using Shared Governance. Frontiers in Blockchain 2, 19.
Mansell, R., Steinmueller, W.E., 2020. Advanced Introduction to Platform Economics. Edward Elgar Publishing.
Martin, B., 2012. The evolution of science policy and innovation studies. Research Policy 41, 1219–1239.
Martin, B., 2019. The future of science policy and innovation studies: some challenges and the factors underlying them. In: Handbook on Science and Public Policy. pp. 523–542.
Martin, B.R., 1996. The use of multiple indicators in the assessment of basic research. Scientometrics 36, 343–362.
Martin, B.R., 2016a. Twenty challenges for innovation studies. Science and Public Policy 43, 432–450.
Martin, B.R., 2016b. Editors' JIF-boosting stratagems - Which are appropriate and which not? Research Policy.
Martin, B.R., Bell, M., Callon, M., Grupp, H., Kodama, F., Kuhlmann, S., Fleming, L., von Tunzelmann, N., Powell, W., 2009. EES and the continuing evolution of Research Policy. Research Policy.
May, R.M., 1997. The Scientific Wealth of Nations. Science (1979) 275.
Mazov, N.A., Gureyev, V.N., 2016. The editorial boards of scientific journals as a subject of scientometric research: A literature Review. Scientific and Technical Information Processing 43, 144–153.
McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E., 1955. A proposal for the Dartmouth summer conference on artificial intelligence.
McKelvey, B., 2001. What Is Complexity Science? It Is Really Order-Creation Science. Emergence: Complexity and Organization 3, 137–157.
McKnight, W., 2014. Graph Databases. In: Information Management. Elsevier.
Mendonça, S., Pereira, J., Ferreira, M.E., 2016. Gatekeeping African studies: A preliminary insight on what do editorial boards indicate about the nature and structure of research brokerage. STI2016 - Proceedings of the 21ST international conference on science and technology indicators.
Mendonça, S., Pereira, J., Ferreira, M.E., 2018. Gatekeeping African studies: what does "editormetrics" indicate about journal governance? Scientometrics.
Menon, V., Varadharajan, N., Praharaj, S.K., Ameen, S., 2020. Why Do Manuscripts Get Rejected? A Content Analysis of Rejection Reports from the Indian Journal of Psychological Medicine. Indian Journal of Psychological Medicine 025371762096584.
Metz, I., Harzing, A.-W., 2009. Gender Diversity in Editorial Boards of Management Journals. Academy of Management Learning & Education 8, 540–557.
Meyer, M., Pereira, T.S., Persson, O., Granstrand, O., 2004. The scientometric world of Keith Pavitt: A tribute to his contributions to research policy and patent analysis. Research Policy 33, 1405–1417.
Miner, J.B., 2003. Commentary on Arthur Bedeian's "The Manuscript Review Process: The Proper Roles of Authors, Referees, and Editors." Journal of Management Inquiry 12, 339–343.
Mingers, J., MacRi, F., Petrovici, D., 2012. Using the h-index to measure the quality of journals in the field of business and management. Information Processing & Management 48, 234–241.

Miniaci, R., Pezzoni, M., 2015. Is Publication in the Hands of Outstanding Scientists? A Study on the Determinants of Editorial Boards Membership in Economics. GREDEG Working Papers.
Mitchell, T., 1997. Introduction. In: Machine Learning. McGraw Hill.
Mizruchi, M.S., 1996. What do interlocks do? An Analysis, Critique, and Assessment of Research on Interlocking Directorates. Annual Review of Sociology 22, 271–298.
Moed, H.F., van Leeuwen, T.N., 1996. Impact factors can mislead. Nature 381, 186.
Mooney, R.J., Roy, L., 2000. Content-based book recommending using learning for text categorization. Proceedings of the ACM International Conference on Digital Libraries 195–204.
Moravcsik, M.J., 1975. Phenomenology and models of the growth of science. Research Policy 4, 80–86.
Morlacchi, P., Martin, B.R., 2009. Emerging challenges for science, technology and innovation policy research: A reflexive overview. Research Policy.
Morton, M.J., Sonnad, S.S., 2007. Women on professional society and journal editorial boards. J Natl Med Assoc 99, 764–771.
Mullen, T., Collier, N., 2004. Incorporating topic information into sentiment analysis models. In: Empirical Methods in Natural Language Processing. pp. 412–418.
Murthy, S.K., 1998. Automatic construction of decision trees from data: A multi-disciplinary survey. Data Mining and Knowledge Discovery 2, 345–389.
Nästesjö, J., 2020. Navigating Uncertainty: Early Career Academics and Practices of Appraisal Devices. Minerva 2020 59:2 59, 237–259.
Nate, S., 2013. The Signal and the Noise: The Art and Science of Prediction. Penguin Books.
Nelson, R., 2012. Some Features of Research by Economists on Technological Change Foreshadowed by The Rate and Direction of Inventive Activity. In: The Rate and Direction of Inventive Activity Revisited. University of Chicago Press.
Nelson, R.R., Dosi, G., Helfat, C.E., Pyka, A., Saviotti, P.P., Lee, K., Dopfer, K., Malerba, F., Winter, S.G., 2018. Modern Evolutionary Economics, Modern Evolutionary Economics. Cambridge University Press.
Nieminen, U.J., 1973. On the centrality in a directed graph. Social Science Research 2, 371–378.
Nisonger, T.E., 2002. The relationship between international editorial board composition and citation measures in political science, business, and genetics journals. Scientometrics 54, 257–268.
Oghbaie, M., Mohammadi Zanjireh, M., 2018. Pairwise document similarity measure based on present term set. Journal of Big Data 5, 52.
Oldham, S., Fulcher, B., Parkes, L., Arnatkevičiūtė, A., Suo, C., Fornito, A., 2019. Consistency and differences between centrality measures across distinct classes of networks. PLOS ONE 14, e0220061.
Østergaard, C.R., Timmermans, B., Kristinsson, K., 2011. Does a different view create something new? The effect of employee diversity on innovation. Research Policy 40, 500–509.
Osuna, E., 1997. Support Vector Machines: Training and Applications.
Ozbilgin, M., 2004. "International" human resource management: Academic parochialism in editorial boards of the "top" 22 journals on international

human resource management. In: Personnel Review. Emerald Group Publishing Ltd.
Pacher, A., Heck, T., Schoch, K., 2021. Open Editors: A Dataset of Scholarly Journals' Editorial Board Positions.
Page, L., Brin, S., Motwani, R., Winograd, T., 1999. The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab.
Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K., 2012. A literature review and classification of recommender systems research. Expert Systems with Applications 39, 10059–10072.
Patel, P., Pavitt, K.L.R., 1995. Patterns of technological activity: their measurement and interpretation. In: Handbook of the Economics of Innovation and Technological Change. pp. 14–51.
Pavitt, K., 1999. Technology, Management and Systems of Innovation, undefined. Edward Elgar Publishing.
Pazzani, M.J., Billsus, D., 2007. Content-based recommendation systems. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, pp. 325–341.
Pereira, L.R., Lopes, R.J., Louçã, J., 2021. Community identity in a temporal network: A taxonomy proposal. Ecological Complexity 45, 100904.
Petersen, J., Hattke, F., Vogel, R., 2017. Editorial governance and journal impact: a study of management and business journals. Scientometrics 112, 1593–1614.
Phillips, F., n.d. Editorial: A Time of Transition for TFSC. Technological Forecasting and Social Change 78, 899–901.
Pierson, D.J., 2004. The Top 10 Reasons Why Manuscripts Are Not Accepted for Publication. Respiratory Care 49.
Pinski, G., Narin, F., 1976. Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. Information Processing and Management 12, 297–312.
Porter, M.F., 1980. An algorithm for suffix stripping. Program 14, 130– 137.
Powell, K., 2016. Does it take too long to publish research? : Nature News & Comment. Nature 530.
Powell, W.W., Grodal, S., 2006. Networks of Innovators. In: The Oxford Handbook of Innovation. Oxford University Press.
Price, R., 2020. Editorial: Blurbs, a new tool for AJP readers. Citation: American Journal of Physics 88, 341.
Priem, J., Groth, P., Taraborelli, D., 2012. The Altmetrics Collection. PLOS ONE 7, e48753.
Pritchard, A., 1969. Statistical bibliography or bibliometrics. Journal of Documentation.
Protasiewicz, J., Pedrycz, W., Kozłowski, M., Dadas, S., Stanisławek, T., Kopacz, A., Gałęzewska, M., 2016. A recommender system of reviewers and experts in reviewing problems. Knowledge-Based Systems 106, 164–178.
Rafols, I., Leydesdorff, L., 2009. Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and indexer effects. Journal of the American Society for Information Science and Technology 60, 1823–1835.
Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A., 2012. How journal rankings can suppress interdisciplinary research: A comparison

between Innovation Studies and Business & Management. Research Policy 41, 1262–1282.
Rakas, M., Hain, D.S., 2019. The state of innovation system research: What happens beneath the surface? Research Policy 48, 103787.
Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P., Larkin, C., Lepori, B., Mosoni-Fried, J., Oliver, E., Primeri, E., Puigvert, L., Scharnhorst, A., Schubert, A., Soler, M., Soòs, S., Sordé, T., Travis, C., Van Horik, R., 2018. A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. Res Eval 27, 298–308.
Resnik, D.B., Elmore, S.A., 2016. Ensuring the Quality, Fairness, and Integrity of Journal Peer Review: A Possible Role of Editors. Science and Engineering Ethics 22, 169–188.
Robinson, D.H., McKay, D.W., Katayama, A.D., Fan, A.C., 1998. Are Women Underrepresented as Authors and Editors of Educational Psychology Journals? Contemporary Educational Psychology 23, 331–343.
Rokach, L., 2010. Ensemble-based classifiers. Artificial Intelligence Review 33, 1–39.
Rossetto, D.E., Bernardes, R.C., Borini, F.M., Gattaz, C.C., 2018. Structure and evolution of innovation research in the last 60 years: review and future trends in the field of business through the citations and co-citations analysis. Scientometrics.
Rousseau, R., Zhang, L., Hu, X., 2019. Knowledge integration: Its meaning and measurement. In: Springer Handbooks. Springer, pp. 69–94.
Russell, S., Norvig, P., 2002. Artificial Intelligence: A Modern Approach, 2nd Edition. ed. Prentice Hall.
Rynes, S.L., 2006. “Getting on board” with AMJ: Balancing quality and innovation in the review process. Academy of Management Journal 49, 1097–1102.
Sabidussi, G., 1966. The centrality index of a graph. Psychometrika 31, 581–603.
Sáez-Martínez, F.J., Triguero, Á., González-Moreno, Á., 2019. A review of Open-innovation and Eco-innovation strategies in SMEs. In: Research on Open-Innovation Strategies and Eco-Innovation in Agro-Food Industries. Chartridge Books Oxford, pp. 9–24.
Salinas, S., Munch, S.B., 2015. Where Should I Send It? Optimizing the Submission Decision Process. PLOS ONE 10, e0115451.
Santiago, F., 2020. Industrial Policies in the BRICS. In: The Oxford Handbook of Industrial Policy. Oxford University Press, pp. 748–778.
Santos, Ana Teresa, Mendonça, S., 2021. Journals’ agendas versus actual publications: A first look at article dynamics in innovation journals. In: 18th International Conference of the International Society for Scientometrics and Informetrics. pp. 985–996.
Santos AT, Mendonça S, 2021. The small world of editorships: A network on innovation studies. In: 18th International Conference on Scientometrics & Informetrics. pp. 985–1028.
Santos, A.T., Mendonça, S., 2021. Exploring innovation studies in a blurb lenses: a textmetric approach to journal positioning .
Sarna, K. v., Griffin, T., Tarlov, E., Gerber, B.S., Gabay, M.P., Suda, K.J., 2020. Trends in gender composition on editorial boards in leading medicine, nursing, and pharmacy journals. In: Journal of the American Pharmacists Association. Elsevier B.V., pp. 565–570.

Saunders, M., 2019. My first 12 months as a journal editor [WWW Document]. URL https://ecologyisnotadirtyword.com/2019/01/20/my-first-12-months-as-a-journal-editor/ (accessed 11.8.20).
Schofield, A., Magnusson, M., Mimno, D., 2017. Pulling Out the Stops: Rethinking Stopword Removal for Topic Models, the Association for Computational Linguistics.
Schot, J., Steinmueller, W.E., 2019. Transformative change: What role for science, technology and innovation policy?: An introduction to the 50th Anniversary of the Science Policy Research Unit (SPRU) Special Issue. Research Policy.
Sebastiani, F., 2002. Machine Learning in Automated Text Categorization. ACM Computing Surveys.
Seglen, P.O., 1997. Why the impact factor of journals should not be used for evaluating research. BMJ 314, 498–502.
Sharif, N., 2006. Emergence and development of the National Innovation Systems concept. Research Policy 35, 745–766.
Shaw, M.E., 1954. Group Structure and the Behavior of Individuals in Small Groups. Journal of Psychology: Interdisciplinary and Applied 38, 139–149.
Shaw, N., Grimshaw, P., Burgess, T., 2016. Diversity of the Information Systems Research Field: A Journal Governance Perspective. Information Systems Research.
Shiffrin, R.M., Börner, K., Stigler, S.M., 2018. Scientific progress despite irreproducibility: A seeming paradox. Proc Natl Acad Sci U S A 115, 2632–2639.
Singh, S., Dhir, S., Das, V.M., Sharma, A., 2020. Bibliometric overview of the Technological Forecasting and Social Change journal: Analysis from 1970 to 2018. Technological Forecasting and Social Change 154, 119963.
Small, H., 2004. On the shoulders of Robert Merton: Towards a normative theory of citation. Scientometrics 2004 60:1 60, 71–79.
Smet, A.A. de, Manaster, B.J., W A Murphy, J., 1994. How to write a successful abstract. https://doi.org/10.1148/radiology.190.2.8284418 190, 571–572.
Smith, K., 2004. Measuring Innovation. In: Fagerberg, J., Mowery, D.C. (Eds.), The Oxford Handbook of Innovation. Oxford University Press.
Song, C., Qu, Z., Blumm, N., Barabási, A.L., 2010. Limits of predictability in human mobility. Science (1979) 327, 1018–1021.
Song, F., Liu, S., Yang, J., 2005. A comparative study on text representation schemes in text categorization. Pattern Analysis and Applications 8, 199–209.
Sooryamoorthy, R., 2020. Scientometrics for the study of sociology. International Sociology 35, 461–479.
Sud, P., Thelwall, M., 2013. Evaluating altmetrics. Scientometrics 2013 98:2 98, 1131–1143.
Sugimoto, C.R., Larivière, V., Ni, C., Cronin, B., 2013. Journal acceptance rates: A cross-disciplinary analysis of variability and relationships with journal measures. Journal of Informetrics 7, 897–906.
Suominen, A., Toivanen, H., 2016. Map of science with topic modeling: Comparison of unsupervised learning and human-assigned subject classification. J Assoc Inf Sci Technol 67, 2464–2476.
Swain, P.H., Hauska, H., 1977. The decision tree classifier: Design and potential. IEEE Trans Geosci Electron GE-15, 142–147.
Tata, S., Patel, J.M., 2007. Estimating the selectivity of tf-idf based cosine similarity predicates. SIGMOD Record 36, 7–12.

Teixeira da Silva, J.A., 2020. CiteScore: Advances, Evolution, Applications, and Limitations. Publishing Research Quarterly 2020 36:3 36, 459–468.
Teixeira Da Silva, J.A., Aamir, •, Memon, R., Memon, A.R., 2017. CiteScore: A cite for sore eyes, or a valuable, transparent metric? Scientometrics 111, 553–556.
Teixeira, E.K., Oliveira, M., 2018. Editorial board interlocking in knowledge management and intellectual capital research field. Scientometrics 117, 1853–1869.
Tenopir, C., King, D.W., 2014. The growth of journals publishing. In: The Future of the Academic Journal: Second Edition. Elsevier Inc., pp. 159–178.
Tidd, J., Bessant, J., Pavitt, K., 1999. Managing Innovation: Integrating Technological, Market, and Organizational Change. Wiley.
Topaz, C.M., Sen, S., 2016. Gender Representation on Journal Editorial Boards in the Mathematical Sciences. PLoS ONE 11.
Turing, A.M., 1950. Computing Machinery and Intelligence . Mind LIX, 433–460.
Tutarel, O., 2004. Composition of the editorial boards of leading medical education journals. BMC Medical Research Methodology 4, 1–4.
van Fleet, D.D., Ray, D.F., Bedeian, A.G., Downey, H.K., Hunt, J.G., Griffin, R.W., Dalton, D., Vecchio, R.P., Kacmar, K.M., Feldman, D.C., 2006. The Journal of Management's First 30 Years. Journal of Management 32, 477–506.
Van Raan, A.F.J., 2004. Sleeping Beauties in science. Scientometrics 59, 467–472.
Van Raan, A.F.J., 2005. Measuring Science. In: Handbook of Quantitative Science and Technology Research. Springer Netherlands, pp. 19–50.
Vanclay, J.K., 2008. Ranking forestry journals using the h-index. Journal of Informetrics 2, 326–334.
Vandome, R., 2013. The advancement of science: James McKeen cattell and the networks of prestige and authority, 1894-1915. American Periodicals 23, 172–187.
Versi, E., 1992. "Gold standard" is an appropriate term. British Medical Journal 305, 187.
Verspagen, B., Werker, C., 2004. Keith Pavitt and the Invisible College of the Economics of Technology and Innovation. Research Policy 33, 1419–1431.
Vespignani, A., 2018. Twenty years of network science. Nature 558, 528–529.
Villegas, N.M., Sánchez, C., Díaz-Cely, J., Tamura, G., 2018. Characterizing context-aware recommender systems: A systematic literature review. Knowledge-Based Systems 140, 173–200.
Vogel, R., Hattke, F., Petersen, J., 2017. Journal rankings in management and business studies: What rules do we play by? Research Policy 46, 1707–1722.
Waltman, L., 2016. A review of the literature on citation impact indicators. Journal of Informetrics 10, 365–391.
Waltman, L., Larivière, V., Milojević, S., Sugimoto, C.R., 2020. Opening science: The rebirth of a scholarly journal. Quantitative Science Studies 1, 1–3.
Wang, D., Barabási, A.-L., 2021. The Science of Science. Cambridge University Press.
Wang, X., 2018. The relationship between SCI editorial board representation and university research output in the field of computer science: A quantile regression approach. Malaysian Journal of Library and Information Science 23, 67–84.

Wang, Y., Wang, L., Rastegar-Mojarad, M., Moon, S., Shen, F., Afzal, N., Liu, S., Zeng, Y., Mehrabi, S., Sohn, S., Liu, H., 2018. Clinical information extraction applications: A literature review. Journal of Biomedical Informatics.
Ware, M., Mabe, M., 2015. The STM Report: An overview of scientific and scholarly journal publishing.
Wasserman, S., Faust, K., 1994a. Social Network Data: Collection and Applications. In: Social Network Analysis: Methods and Applications. Cambridge University Press.
Wasserman, S., Faust, K., 1994b. Social Network Analysis. Cambridge University Press.
Watts, D.J., 2004. The "New" Science of Networks. https://doi.org/10.1146/annurev.soc.30.020404.104342 30, 243–270.
Welch, S.J., 2012. Selecting the right journal for your submission. J Thorac Dis 4, 336–8.
White, K., 2019. Publications Output: U.S. Trends and International Comparisons.
Wilhite, A., Fong, E.A., Wilhite, S., 2019. The influence of editorial decisions and the academic network on self-citations and journal impact factors. Research Policy 48, 1513–1522.
Wilkes, M.S., Kravitz, R.L., 1995. Policies, practices, and attitudes of north american medical journal editors. Journal of General Internal Medicine 10, 443–450.
Willett, P., 2013. The Characteristics of Journal Editorial Boards in Library and Information Science. International Journal of Knowledge Content Development & Technology 3, 5–17.
Williams, R., 2019. Why science and innovation policy needs Science and Technology Studies? In: Handbook on Science and Public Policy. pp. 503–522.
Wright, P., 1985. Editing: Policies and Processes. In: Duffy, T.M., Robert, W. (Eds.), Designing Usable Texts.
Wu, D., Lu, X., Li, Jianping, Li, Jing, 2020. Does the institutional diversity of editorial boards increase journal quality? The case economics field. Scientometrics 124, 1579–1597.
Yates, R.B., Neto, B.R., 1999. Text and multimedia languages and properties. Modern information retrieval 141–162.
Zaklouta, F., Stanciulescu, B., 2012. Real-time traffic-sign recognition using tree classifiers. IEEE Transactions on Intelligent Transportation Systems 13, 1507–1514.
Zhang, P., 1993. Model Selection Via Multifold Cross Validation. The Annals of Statistics 21, 299–313.
Zhang, T., 2020. Will the increase in publication volumes "dilute" prestigious journals' impact factors? A trend analysis of the FT50 journals. Scientometrics 1–7.
Zhang, Y., Haghani, A., 2015. A gradient boosting method to improve travel time prediction. Transportation Research Part C: Emerging Technologies 58, 308–324.
Zsindely, S., Schubert, A., Braun, T., 1982. Editorial gatekeeping patterns in international science journals. A new science indicator. Scientometrics 4, 57–68.

Annexes: Posters presented in international conferences

Judging Journals by their "Blurbs": A Text Mining Approach to Innovation Journals

Ana Teresa Santos[1]; Sandro Mendonça[1,2,3]
[1]ISCTE-IUL, [2]ISEG/ University of Lisbon, Lisbon, Portugal, [3]SPRU, University of Sussex, UK

1 Abstract

Design: We take "blurbs", the meta-information that is published on the journal front-matter or official websites describing the aims and purposes of academic outlets, as an under-research resource for enhancing the agenda of the scientometrics of journals (or "journalytics", as we short-handed it) in innovation.

Methods: Topic Modelling, Latent Semantic Analysis and Social Networking.

Contribution: Our experiment contributes to a portfolio of ways by which researchers and science policy analysts can judge journals finding unambiguously the agenda of interest.

Results: We find six journals that have more than 80% of (cosine) similarity to *Research Policy*, arguably the leading scholarly vehicle for the innovation studies. Also *Industrial Corporate Change* is an important link between journals.

Originality: In this work, we present a latent semantic analysis that mines for unseen themes behind the journal "blurbs" to examine the positioning and standing of journals.

2 Examining Journals Place

This study is intended to serve both intellectual and strategic needs of stakeholders who rely on journals to make sense of a variety of knowledge agendas.

In today's information-saturated world, **text analysis** has been used to set the stage for a data-driven approach towards managing content.

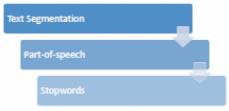
In the present research, we propose a new method based on cosine similarity to be applied to blurb sections from a list of journals previously ranked by Fagerberg and colleagues (2012) as the most dedicated to innovation.

The novelty of our research lies on the usage of the journals content to compare journals rather than previous publication and authors' profile. By changing the focus to journals instead of people, we proposed a content analysis to order the journals interest on innovation subject.

We also analyze the **social structure** of editorial board membership in innovation research assuming the contribution of a certain editor as the sum of boards he works for.

3 Blurb Coding & Networking

With **Topic Modelling**, a text mining technique, hidden topics in the document are discovered with methods for identifying co-occurring keywords to summarize large collections of textual information.

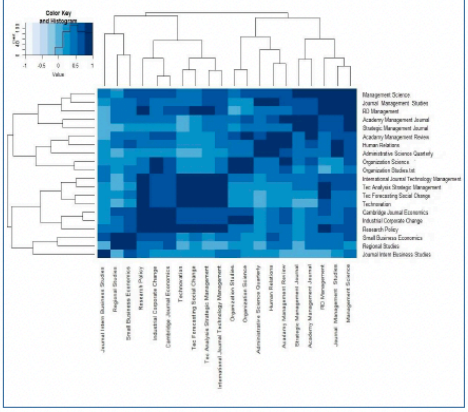


Latent Semantic Analysis is a language processing procedure that allows recognition of textual associative patterns and permits statistical extraction of common textual themes that characterize an entire set of documents, as well as tracking the relative prevalence of each theme over time and across entities.

Social Network Analysis is a technique aimed to explores structural properties of the network generated by the editorial activities in innovation studies. Out of 3006 editors from the 20 journals recorded, 560 are repeated names.

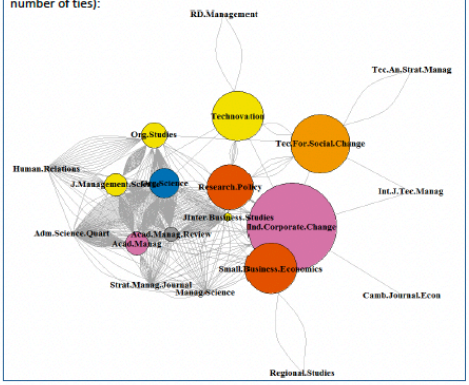
4 Cosine Similarity of Journal Blurbs

Semantic Space by Similarities Calculation



5 Innovation Journals Networks

Innovation Journals Networks according to Journals Betweenness (bigger size nodes for more central ones) and **Degree** (different colors mean different number of ties):



6 Discussion

Dendogram ranks the closest blurbs to the *Research Policy* one, which is recognized as a reference journal for innovation studies (Martin, 2012). Believing the blurb is stating the desired topics to be published, we may say the more similar a blurb is to the *Research Policy* blurb, the more focused on innovation the journal is. (Dziallas, 2019).

Cosine	Journal
1.00	Research Policy
0.89	Technological Forecasting and Social Change
0.88	Cambridge Journal Economics
0.86	International Journal of Technology Management
0.84	Technology Analysis & Strategic Management
0.82	Industrial Corporate Change
0.82	Technovation
0.79	R&D Management

There are journals more independent than others although all the 20 journals are integrated in the network.

Degree: the more ties a journal has to other journals, the more central is its position in the network. Both *Academy Management Journal* and *Academy Management Review* have more than 200 connections to other journals.

Closeness: A journal is central if its board can quickly interact with all the other boards. Journals occupying a central location are best placed to influence the entire network most quickly. With exception of *Cambridge Journal of Economics*, *Human Relations*, *R&D Management*, *International Journal of Technology Management* and *Technology Analysis & Strategic Management*, all other journals have a similar score.

Betweenness: journal is more central in this respect if it is an important intermediary in links between other journals. *Industrial Corporate Change* is a "brokerage" for around 33 short paths linking other journals.

7 Conclusions

In this research, we mapped the interest about publishing on innovation. We identified consistent self-characterisation wordage data, thereby enhancing our understanding of journals impetus for contributing to this field's collective advances. Further, we established a conceptual and analytical link between the research itself and journals.

We provided what is, to the best of our knowledge, the first attempt considering the "front-office" promotional information for addressing a number of matters of interest for scientometricians, namely: what do (exactly) journals stand for; how do academic outlets contrast and each other; how to provide a semantic maps to help navigate the challenges of research identification, monitoring, evaluation, and the like. We also develop a new method to combine the contribution of the editor itself and of his status in the co-editor network. We use the times of co-editorship to calculate the distance between each journal.

Our experimental results indicate that the proposed techniques provide a reasonable domain-dependent ranking, also closer to the rankings already published through other methods. The network analysis also shows that there are journals more independent than others although all the 20 journals are integrated in this network.

INNOVATION | BIBLIOMETRICS| BLURBS | TEXT MINING | |TOPIC MODELING | JOURNAL CLASSIFICATION | EDITORIAL BOARDS

Ana Teresa Santos
aa.teresa.ss@gmail.com

References

Dziallas, M., & Blind, K. (2019). Innovation indicators throughout the innovation process: An extensive literature analysis. *Technovation*, *80–81*, 3–29. https://doi.org/10.1016/J.TECHNOVATION.2018.05.005
Fagerberg, J.; Fosaas, M.; Sapprasert, K. (2012). Innovation: Exploring the knowledge base. *Research Policy*, *41*(7), 1132–1153. https://doi.org/10.1016/J.RESPOL.2012.03.008
Fagerberg, J., & Verspagen, B. (2009). Innovation studies—The emerging structure of a new scientific field. *Research Policy*, *38*(2), 218–233. https://doi.org/10.1016/j.respol.2008.12.006
Martin, B. (2012). The evolution of science policy and innovation studies. *Research Policy*, *41*(7), 1219–1239. https://doi.org/10.1016/J.RESPOL.2012.03.012

Watching over innovation studies: Profiling the gatekeepers

Ana Teresa Santos[1] and Sandro Mendonça[1]
[1]Instituto Universitário de Lisboa (ISCTE-IUL) – Business Research Unit (BRU-IUL), Lisbon, Portugal.

Introduction

Academic serials play a very critical role in the scientific ecosystem. Responsible for articles selection (Bedeian et al., 2009; Feldman, 2008), the editorial board (EB) members ensure the scientific quality of publications. So far, board elites have not been subject to a scrutiny proportional to their decision power (Burgess & Shaw, 2010). In this work, we shed some light on demographical characteristics of editorial members.

The Boards of "Innovation Studies"

Fagerberg et al. (2012) studied both the most productive journals in innovation as well as those using the most findings already published. For the 20 top-tier journals identified, we aim to present the demographic features of scholars behind. Scholars' names, affiliation country and gender were collected from official journal's editorial page. A total of 3,005 available seats were recorded occupied by 2,440 distinct persons.

The size of EBs

Boards mean size is 150 editors. Although *R&D Manag* Board was found with only 19 editors, others have over 300 scholars: the likes of North-American outlets as *Acad Manage J* and *Acad Manage Rev* while *Manage Sci* stands out as the one with the largest team, 399 editors as presented in Figure 1.

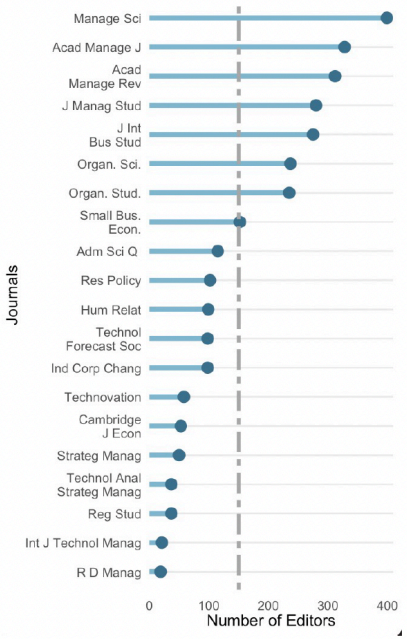


Figure 1: Editorial Board size per journal intercepted by the mean number of editors.

Gender balance

In Figure 2, we matched gender proportion in 10 countries with higher numbers of editors. All countries show a higher frequency of memberships held by men editors than women.

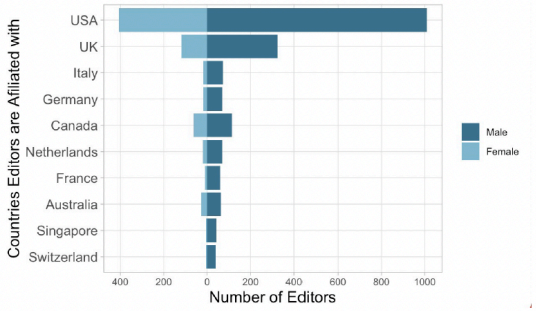


Figure 2: Number of editors affiliated to the top 10 countries, by gender.

Geographies of editorship

Figure 3 shows a world representation of editors' frequency found in each country. Darker the colour, higher the frequency of editors. The US overwhelming dominance of membership is clear. Europe and India host a significant number of editors. With exception of Australia, nations from the South represent a negligible role in this editorial process.

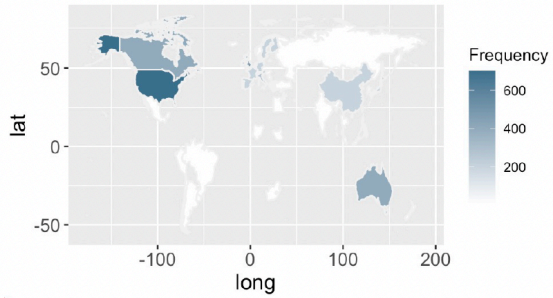


Figure 3. Geographical location of editorial members.

Conclusion

Innovation EB members showed to be diverse quantitative and qualitatively. Despite of being an eclectic topic, a low diversity for gender (male predominance) and country affiliation (high representation of US and UK editors) is perceived. A greater diversity is important for EB meet their missions (Jagsi et al., 2008) and a key driver in academia for knowledge development by applying different methodologies and paradigms (Robinson & Dechant, 1997).

References

Bedeian, A. G., van Fleet, D. D., & Hyman, H. H. (2009). Scientific Achievement and Editorial Board Membership. Organizational Research Methods, 12(2), 211–238.

Feldman, D. C. (2008). Building and Maintaining a Strong Editorial Board and Cadre of Ad Hoc Reviewers. In Opening the Black Box of Editorship (pp. 68–74). Palgrave Macmillan UK.

Burgess, T. F., & Shaw, N. E. (2010). Editorial Board Membership of Management and Business Journals: A Social Network Analysis Study of the Financial Times 40. British Journal of Management, 21(3), 627–648.

Fagerberg, J., Fosaas, M., & Sapprasert, K. (2012). Innovation: Exploring the knowledge base. Research Policy, 41(7), 1132–1153.

Jagsi, R., Tarbell, N. J., Henault, L. E., Chang, Y., & Hylek, E. M. (2008). The representation of women on the editorial boards of major medical journals: A 35-year perspective. Archives of Internal Medicine, 168(5), 544–548.

Robinson, G., & Dechant, K. (1997). Building a Business Case for Diversity. The Academy of Management Executive, 11(3), 21–31.

Share with us your thoughts to: atmss@iscte-iul.pt