REGULAR PAPER

Framework for scoring the scientific reputation of researchers

Isaac Martín de Diego[1] · Juan Carlos Prieto[1] · Alberto Fernández-Isabel[1] · Javier Gomez[1] · César Alfaro[1]

Abstract

In the scientific community, there is no single, objective, and precise metric for ranking the work of researchers based on their scientific merit. Most existing metrics are based on the number of publications associated with an author along with the number of citations received by those publications. However, there is no standard metric officially used to evaluate the researchers' careers. In this paper, the *Framework for Reputation Estimation of Scientific Authors* (*FRESA*) to address this issue is depicted. It is a system able to estimate the reputation of a researcher focusing on the achieved publications. It calculates two indexes making use of the relevance and the novelty concepts in the scientific domain. The system can depict the scientific trajectories of the researchers through the proposed indexes to illustrate their evolution over time. *FRESA* uses web information sources and applies similarity measures, text mining techniques, and clustering algorithms to also rank and group the researchers. The presented work is experimental, rendering promising results.

Keywords Reputation index · Relevance index · Novelty index · Scientific quality measures · Author scientific trajectory

1 Introduction

Nowadays, there is no automatic and unique mechanism to qualify researchers to assign seats in public universities or to award them with scientific prizes. This is large since the scientific

---

✉ Alberto Fernández-Isabel
alberto.fernandez.isabel@urjc.es

Isaac Martín de Diego
isaac.martin@urjc.es

Juan Carlos Prieto
juancarlos.prieto@urjc.es

Javier Gomez
javier.gomez@urjc.es

César Alfaro
cesar.alfaro@urjc.es

1   Data Science Laboratory, Rey Juan Carlos University, C/ Tulipán, s/n, 28933 Móstoles, Spain

community has not agreed to define a standard that allows ranking researchers objectively and uniquely [33]. Several attempts to define reputation metrics based on publication efforts to produce trust and recognition of researchers in the scientific area have been achieved.

However, no metric has reached a special relevance to become a standard [45]. It is also a great issue that there is no unique repository where all the publications are collected. Nonetheless, there are well-known repositories that try to provide an acceptable solution to the matter. Famous instances of them are Google Scholar [31] and Semantic Scholar [13]. The main problem in these cases is to gather the information from authors (i.e., the publications made by them or the associated citations). Therefore, elaborating a standard metric to score researchers is a demanding task that could influence the scientific community to feed these repositories transforming them into a more complete knowledge base.

In this sense, well-known metrics such as h-index [21], g-index [9] or Google's i10 index [8] have been elaborated to mitigate the issue. However, they are too simple to be fair to all existing authors' profiles because they award productivity over the quality or novelty of the approaches.

This paper presents the *Framework for Reputation Estimation of Scientific Authors (FRESA)* framework. It is focused on estimating the reputation of the authors' profiles considering not only the relevance of the publications but also their novelty. This fact is a novel issue as historically the reputation of an author in science has been estimated only with the quality and the number of publications [27]. In this case, the proposed reputation makes use of two original indexes: the relevance index and the novelty index. The first one uses well-known features like the number of citations and information provided by the JCR index [27]. The second provides new insights into evaluating the innovation of the approaches according to its research topics discussed in three-year cycles [2].

Three different experiments have been addressed to show the viability of the proposal. In the first one, the reputation trajectory of a recognized author is graphically represented by the system and the results are validated. In the second experiment, some comparisons with well-known indexes are shown highlighting their limitations. In the third experiment, clustering tasks are performed to group different authors by their estimated reputation. Thus, it illustrates the ability of the system to classify authors into representative groups according to the calculated reputation.

The remainder of the paper is organized as follows. Section 2 establishes the context and describes similar approaches. Section 3 introduces the proposed framework. Section 4 presents the experiments used to illustrate the viability of the proposal. Finally, Sect. 5 concludes and provides future lines of work.

2 Background

This section introduces the foundations of the *FRESA* framework. First, the three different perspectives included in the system are detailed: reputation estimation, novelty detection and evaluation, and relevance calculation. The first one addresses the reputation concept and its estimation in different contexts (see Sect. 2.1). The second one details the novelty concept, and how it is detected and evaluated (see Sect. 2.2). The third one focuses on the relevance concept and how it is estimated in multiple domains (see Sect. 2.3). Finally, the most typical bibliometric methods and techniques are tackled (see Sect. 2.4). This allows establishing the indexes and the reputation calculation algorithm defined by *FRESA* in the domain.

2.1 Reputation estimation

Reputation as a concept has been widely used in commercial issues. For instance, online collaboration, e-commerce websites, and transactions are some of the most common domains. There, reputation helps to reduce the uncertainty between peers, as a highly reputed member will be as a consequence more reliable and safe [39].

Reputation systems provide essential input for computational trust as predictions on future behavior based on the past actions of a peer [14]. In this increasingly interconnected world, the need for reputation is becoming more important as larger numbers of people and services interact online. Reputation is a tool to facilitate trust between entities, as it increases the efficiency and effectiveness of online services and communities. As most entities will not have any direct experience with other entities, they must increasingly come to rely on reputation systems. Such systems allow estimating who is likely to be trustworthy based on the feedback from past transactions [20].

In Online Rating Systems [29], users' reputation is critical to detecting spammers and eliminating their ratings. The effect of such spammers can be mitigated by a reputation system that can estimate the credibility of users based on their behavior. Some reputation systems detect spammers based on the deviation between their ratings and the true rating of the item. This deviation can be usually estimated using a penalty function [30]

On the other hand, measuring the influence and reputation of academic researchers has been a challenging issue that has attracted much attention from the scientific community.

Scientific reputation can be defined as the trust and recognition given to a scientific researcher for his or her work. This recognition should be reduced to a certain number so that it can be compared with others and build a valid metric to evaluate researchers. This scientific recognition should consider the quality and novelty of the work, not the productivity of the author.

In this regard, many bibliometric indicators have been and continue to be proposed, most of them based on citations of papers and their variations, such as h-index [21], g-index [9], hg-index [1], or ca-index [3]. Many other methods for the assessment of the researchers' reputation use citation weighting [46] and other similar features [12], or network-based [7, 32] models.

Multi-criteria metrics based on a profile model of the researchers have been also developed [5]. They usually involve other elements of the scientific career production apart from bibliometry as participation in research projects, conduction of application-oriented research, and dissemination of research results via conferences, among others [24].

2.2 Novelty detection and evaluation

Novelty detection is the task of identifying novel information given a set of already accumulated background information. Potential applications of novelty detection systems are abundant, given the "information overload" in all domains [16].

Several methods have been proposed to determine novel information from raw text documents. To settle whether information within a sentence has been seen in the previously read material, some approaches integrate the information about the context of the sentence with novel words and named entities within the sentence and use a specialized learning algorithm to tune the system parameters [40]. Named entity recognition and part-of-speech tagging have achieved great results comparing various types of entities within sentences [35].

In the case of the scientific community, scientists are under increasing pressure to produce novel research. Review studies from the sociology of science have been made to anticipate how emphasizing novelty might impact the structure and function of the scientific community [6]. Thus, it is necessary to find a balance between novelty and reproducibility [22]. Besides, the technological impact is not always defined by the novelty of the research [43].

Different indicators have been proposed in the literature to detect, analyze, and assess the novelty of scientific research. Instances of these types of approaches are [44] and [4]. The first one developed a novelty indicator for assessing scientific publications by examining whether published research makes first-time-ever combinations of referenced journals, taking into account the difficulty of making such combinations. Therefore, the novelty of a paper is measured as the number of new journal pairs in its references weighted by the Cosine Similarity [28] between the newly paired journals. The second one proposed a novelty indicator based on the frequency of pairwise combinations of author keywords applied to all research publications during different years. Thus, given a year and a research field, papers can be ranked based on their novelty score where the proposal with the most infrequent pairwise combination of keywords obtains high scores capturing their originality.

2.3 Relevance evaluation

The concept of relevance was recognized for the first time at the end of the 1950s for providing a solution to a concept that had been dealt with since the 17th century with the publication of the first scientific journals [34].

In the case of raw textual content, one of the most important methods to measure relevance is the PageRank algorithm [36]. It can measure the importance of a text generating a conceptual network from it and later by counting the number and quality of the links (relationships) of a node (usually a concept).

In the scientific context, the relevance is evaluated using quantitative measures. There are several indicators to measure the scientific relevance of an author, a paper, or a journal. The basic idea is that highly relevant papers will be cited more often than less relevant papers. Bibliography Coupling [26] is a well-known instance to measure the similarity between documents and the recent citation number, considering this recent citation number of a scientific paper in a certain time.

A recommendation system for scientific relevance for researchers using an ontological model previously described [42] is proposed in [19]. Moreover, to measure the relevance of the researchers, several indicators have been proposed. Thus, traditional metrics, such as the Scimago Journal Ranking [10] and the publication citations number, are usually combined with other context-dependent features such as the type of publication or the city and country of the researcher's institution.

2.4 Scientific quality measures

Over the years, different quality metrics have been used in the scientific community to assess scientific researchers. These metrics usually present the same problem, they reward productivity over the quality of the work. Instances of these measures are the total number of citations, impact factor, h-index, i10 index, g-index, UNIKO reputation index, and altmetrics.

The total number of citations is a bibliometric indicator that reflects the relevance of an article. It penalizes young authors, since an author with few highly relevant articles may have a much lower number of citations than an author with many irrelevant articles.

The impact factor [17] measures the importance of a scientific publication. It is calculated every year by the Institute for Scientific Information (ISI) for those publications of the Journal Citation Reports (JCR). The one-year impact factor is calculated as the ratio of A to B, where A is the number of citations in the current year to any publications in a certain journal in the previous two years; and B denotes the number of articles published in the journal in the previous two years. It presents two problems: its calculation is based on the number of citations so it does not measure the quality of the publication, and the base calculation period is very short since classic articles are frequently cited even after decades.

The h-index [21] measures the professional quality of researchers, considering both the number of publications of the researcher and the total number of citations to those works. Given a set of scientific publications from a researcher ranked in decreasing order of the number of citations, the h-index of the researcher is the number in which the ordinal number coincides with the number of citations. The main drawback of the h-index is that it penalizes researchers with a short but brilliant career since being limited by the number of publications they will not have a high h-index.

The i10 index [8] is a Google metric available since 2011 that indicates the number of academic papers with at least 10 citations that an author has published. It presents the same problem as the h-index and the impact factor, it rewards productivity over the quality of work.

The g-index [9] is a metric to improve the h-index. It is calculated as follows: given a set of articles sorted in decreasing order of the number of citations, the g-index is the highest number that fulfills that these g articles have received in total at least g^2 citations. This metric relieves rewarding productivity since articles with more citations are given more value.

The UNIKO reputation index [11] aims to compensate for the problems introduced by metrics such as the h-index, g-index, or i10 index, mitigating the influence of the author's productivity. This metric is calculated as a weighted sum of the number of influential citations, the number of total citations, the seniority, and the number of articles published. However, this metric does not prevent highly relevant authors with a short career from being penalized.

Finally, in recent years, new measures that consider the impact and dissemination of the work on the Internet have been suggested. They are generically called altmetrics [37]. A well-known measure of this type is a novel bibliometric indicator called citing authors index [3].

3 Framework architecture

The *FRESA* framework is an innovative platform focused on providing a brand new metric to rank authors by their reputation. It estimates this reputation according to the relevance (i.e., the number of publications and the quality of the journal) and the novelty (i.e., the different branches of innovation) during the scientific careers of authors. The system calculates and also depicts a reputation trajectory based on the relevance and novelty of the authors for each year to illustrate their evolution. This process is completely novel because of the new insights provided by the novelty and the depiction of the trajectories, which also allows producing visual comparisons between scientific careers. To achieve these tasks, *FRESA* gathers information from external web sources, obtains relevant knowledge, and achieves estimations in two-dimensional graphics.

Regarding the architecture of the system, it presents four main modules which are detailed in the next sections (see Fig. 1): the *Information Updating* module, the *Relevance Calculator* module, the *Novelty Calculator* module, and the *Trajectories Analysis* module. These four
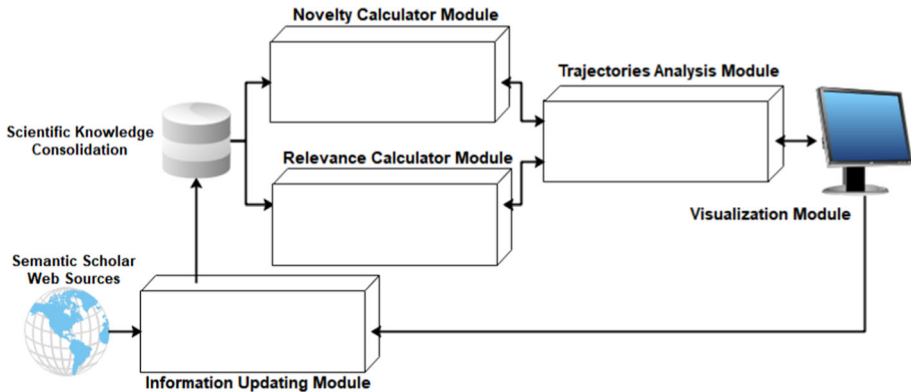
Fig. 1 Overview of the architecture of the *FRESA* framework

modules collect and update the information from web sources using the Semantic Scholar API. Then, they estimate the relevance and novelty indexes for each author and year and calculate the final reputation score. Finally, they build the reputation trajectories for each author representing the entire scientific work of the author. Besides, the system also holds a *Visualization* module and the *Scientific Knowledge Consolidation* module that acts as a knowledge base. The last one contains the information regarding the JCR journals and Computing Research and Education (CORE) [23] conferences such as the journal name, impact factor, quartile, CORE index, journal topic, or conference acronym. It is also updated with the authors' and papers' information. Finally, the Semantic Scholar API is used to collect information from the papers. It is an AI-backed search engine for academic publications [18] designed to highlight the most important and influential papers.

3.1 Information updating module

This module collects and updates the information from the web sources and processes these data to enrich the *Scientific Knowledge Consolidation* module. It presents two components (see Fig. 2): the *Author's Information Collector* and the *Author's Information Processor*. The first one uses the Semantic Scholar API to collect information relative to the authors and their publications. The information collected from each published paper is the following: title, abstract, authors, citations, fields of study, doi, topics, journal name, and publication year. Due to the lack of quartiles and CORE indexes information in Semantic Scholar, the *Author's Information Processor* is needed. It receives the information from the *Author's Information Collector* and the information related to JCR journals and CORE conferences from the *Scientific Knowledge Consolidation* module: journal name, impact factor, quartile and category for JCR journals; and conference name, acronym, and CORE index for CORE conferences. This component consolidates the desegregated data based on a TF-IDF similarity estimation [38] between the name of the journal extracted from Semantic Scholar and the name of the journal from the *Scientific Knowledge Consolidation* module. Notice that journal names provided by Semantic Scholar do not always match with the official names provided by JCR.
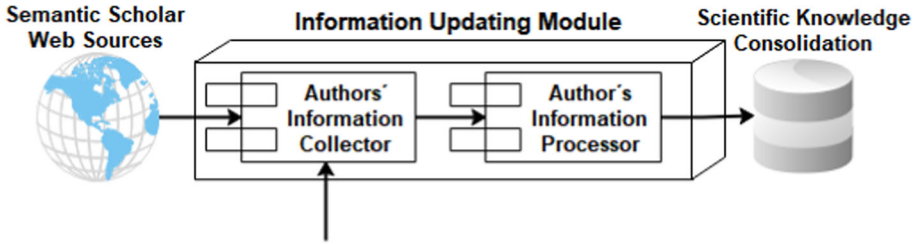
Fig. 2 Excerpt of the *information updating* module and its components
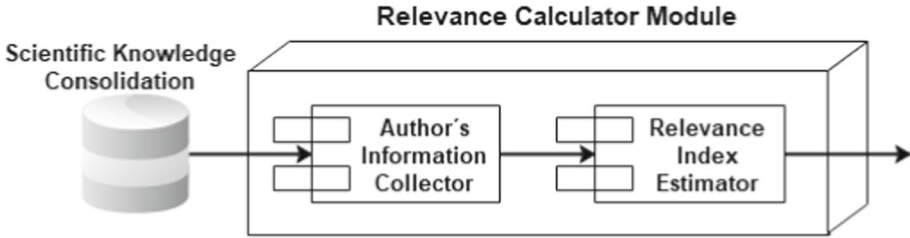

Fig. 3 Excerpt of the architecture of the *relevance calculator* module

3.2 Relevance calculator module

This module calculates the relevance index for each author and year. It presents two components (see Fig. 3): the *Author's Information Collector* and the *Relevance Index Estimator*. The first one extracts information from the *Scientific Knowledge Consolidation* module to feed the *Relevance Index Estimator*. This latter calculates the author relevance index for each year. It uses a three-year moving average to soften the impact of years with few or low-quality publications. This is applied to avoid the relevance of a reputed author being 0 in a year where no publications are made. In the same way, a five-year moving average is not used because it would flatten the trajectory too much, penalizing the good years excessively and rewarding the bad ones.

In the following, the metric used to estimate the relevance index is specified. First, the basic definitions are presented to clarify the proposed metric explanation:

- *Citations*: is the total number of citations of an author's articles in that year.
- J_P: is the number of articles published in JCR journals with a P quartile, where P is Q1, Q2, Q3 or Q4.
- C_P: is the number of articles published in CORE conferences with a P index, where P is A^*, A, B, or C. Notice that the equivalence in importance between Q3 and Q4 with A^* and A, respectively, has been considered. Thus, CORE B and CORE C articles are considered similar in importance.
- A_P: represents the number of authors who have published articles with a P quartile JCR journals or P index CORE conferences, where P is Q1, Q2, Q3, Q4, A^*, A, B or C.
- w_j: refers to the weight element, with $w_j \geq 0$ and $\sum_{j=1}^{6} w_j = 1$.
- *Seniority*: refers to the time elapsed, in years, between the last published paper of the author and the first one.
- *Papers*: is the total number of published papers of authors during their career.

– *Current Seniority*: refers to the time elapsed between the year for which the relevance is being calculated and the publication year of the first article of the author.
– *Current Number of Years*: refers to the number of the year for which the relevance is being calculated. For instance, for an author whose first publishing year is 2020, the *Current Number of Years* for 2020 will be 1, for 2021 will be 2, and so on

The relevance index is calculated through Eq. (1):

$$\text{Relevance Index} = \frac{\text{Citations} * \text{Weighted Citations Value}}{\text{Years Penalty}}. \tag{1}$$

Thus, the relevance index is based on the weighted number of citations over a penalty regarding the number of years. The value for the weight is calculated as follows:

$$\text{Weighted Citations Value} = \frac{(S1 + S2 + S3 + S4 + S5)}{\text{Productivity Penalty}}, \tag{2}$$

where S_i for i in $[1, 5]$ are related to the number of articles in $i-$quartile and core conferences over a penalty related to the number of authors. The Productivity Penalty is about the total number of published papers over the seniority of the author. That is, authors with many papers in $Q1$, and a low number of co-authors during their careers are associated with high Weighted Citations Value.

$$S1 = \frac{J_{Q1}}{A_{Q1}} * w_1, \tag{3}$$

is related to the number of articles published in $Q1$ over the number of co-authors.

$$S2 = \frac{J_{Q2}}{A_{Q2}} * w_2, \tag{4}$$

is related to the number of articles published in $Q2$ over the number of co-authors.

$$S3 = \frac{J_{Q3} + C_{A*}}{A_{Q3} + A_{A*}} * w_3, \tag{5}$$

is related to the number of articles published in $Q3$ and CORE A^* over the number of co-authors.

$$S4 = \frac{J_{Q4} + C_A}{A_{Q4} + A_A} * w_4, \tag{6}$$

is related to the number of articles published in $Q4$ and CORE A over the number of co-authors.

$$S5 = \frac{C_B + C_C}{A_B + A_C} * w_5, \tag{7}$$

is related to the number of articles published in CORE B and CORE C over the number of co-authors.

$$\text{Productivity Penalty} = \frac{\text{Papers}}{(\text{Current Seniority} * w_6)}, \tag{8}$$

is the total number of published papers of authors during their career over their seniority. Thus, it is high when there is a high number of papers in a low number of years.
Finally:

$$\text{Years Penalty} = ((\text{Seniority} - \text{Current Number of Years}) * w_6) + 1 \tag{9}$$
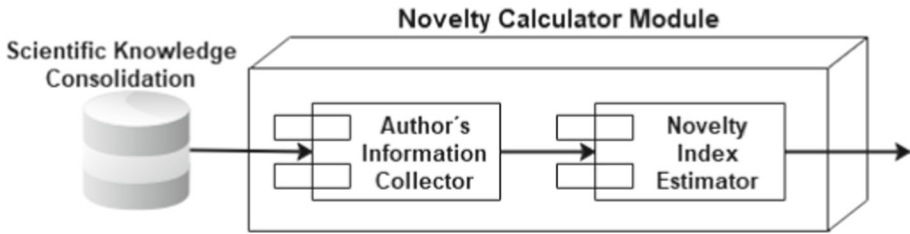
is a penalty related to seniority.

Fig. 4 Excerpt of the architecture of the *novelty calculator* module

As a summary, the *RelevanceIndex* is based on the value of the citations corrected for a weighted based on the number of articles, the number of co-authors, and the seniority. In this way, for instance, it is avoided that an author with 1, 000 citations in an article published in a Q4 quartile journal from having greater relevance than an author with 500 citations in an article published in a Q1 quartile journal.

In addition, to avoid rewarding productivity over the quality of the articles, the weighted value of the citations is divided by the number of articles published. Besides, the time elapsed until the year in which the relevance is calculated is taken into account to prevent authors with very short but highly influential scientific lives from being penalized.

3.3 Novelty calculator module

This module calculates the novelty index for each author and year. It presents two components (see Fig. 4): the *Author's Information Collector* and the *Novelty Index Estimator*. The first one extracts the information needed from the *Scientific Knowledge Consolidation* module to feed the *Novelty Index Estimator*. This latter calculates the author's novelty index for every year. The novelty index offers an annual score to measure the originality of the published articles of each author in that year. It is calculated by comparing the topics of the articles using similarity techniques. For this, the title, keywords, and abstract of the scientific articles are used. First, three natural language processing (NLP) techniques are applied to gather the relevant types of words: part of the speech, lemmatization, and named entity recognition (NER). Part of the speech technique detects common nouns, which are lemmatized to produce the linguistic root of the word (mainly eliminating plural forms). On the other hand, NER technique detects proper noun phrases (mainly names of people in this context). These tasks improve the results obtained through the similarity measure. Then, TF-IDF [38] and Cosine Similarity [28] techniques are applied to calculate the global similarity measure among the articles.

The author's novelty index in a year is estimated through the average of the similarity indexes calculated for each paper in that year. The similarity of the papers is calculated as follows: the first paper published in the year has a similarity equal to 1 since it is original, the second paper published in the year has a similarity equal to 1 minus the similarity between the first and second paper, the third paper published in the year has similarity equal to 1 minus the maximum of the similarities between paper 1 and 3, and 2 and 3, and so on. The novelty index is limited to the originality of each author taking into account his work.
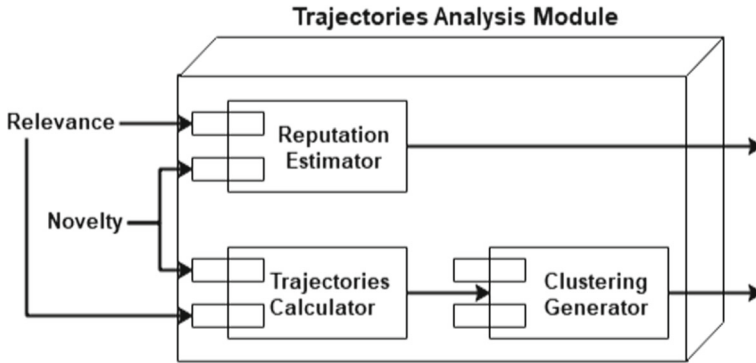
Trajectories Analysis Module

Fig. 5 Excerpt of the architecture of the *trajectories analysis* module

3.4 Trajectories analysis module

This module accomplishes three different tasks. The first one consists of estimating the final reputation score. The second one builds the reputation trajectories for each author representing the entire scientific work of the author. The third one groups the authors into three clusters based on their reputations. It presents three components associated with these tasks (see Fig. 5): the *Reputation Estimator*, the *Trajectories Calculator*, and the *Clustering Generator*, respectively.

The module receives the relevance and novelty indexes for each author from the *Relevance Calculator* and the *Novelty Calculator* modules.

Delving into the calculation of the tasks, the reputation score is calculated as the geometric mean between the *relevance* and the *novelty*. That is:

$$\text{Reputation Score} = (\text{Relevance Index} * \text{Novelty Index})^{1/2}. \tag{10}$$

Notice that the novelty index is bounded by 1, but the relevance index is not bounded at all. Thus, the geometric mean has been chosen since it gives more weight to small items than the arithmetic mean.

In the case of the reputation trajectories, they consist of a sequence of points. Each point has associated spatial information given by the relevance and novelty indexes and temporal information given by the year of those indexes. These points are estimated as the combination of the relevance index and the novelty index in that year. The reputation trajectories are plotted on a graph whose X-axis is the relevance index and the Y-axis is the novelty index.

Finally, for the groups of authors based on the reputation trajectories, a grid-based analysis is performed. The grid is divided into 25 cells to convert the trajectories into one-dimensional arrays. It uses 25 cells due to the compromise between the complexity introduced by inserting more cells and the granularity necessary to obtain good results. Once all the trajectories are converted, the dynamic time warping (DTW) algorithm [41] is applied to calculate the distances. This algorithm allows measuring the similarity between time sequences, obtaining a good fit even with a time lag. This is achieved by looking for the closest point between each point of the two trajectories allowing discerning similar shapes that may be out of phase. To do this, a similarity matrix is generated that compares the two trajectories and calculates the distances between all the points of each trajectory; and, based on this matrix, the most optimal trajectory is selected. Finally, the hierarchical clustering algorithm [25] is used to cluster the different authors.

4 Experiments

This section addresses a set of experiments that evaluate the metrics proposed and validate the overall performance of the system. For the setup of the system, the value of the citations is weighted with the following weights: $w_1 = 0.4$, $w_2 = 0.25$, $w_3 = 0.15$, $w_4 = 0.1$, $w_5 = 0.05$, and $w_6 = 0.05$. These values have been provided according to the knowledge acquired from previous well-established approaches in the domain (e.g., the UNIKO framework [11] and the Webelance system [12]). Thus, the relative importance given to the quartiles of the journals in which an article is published decreases as the quartile number increases. In addition, the seniority is considered with a moderate relevance following the previous literature in the domain [11].

Delving into the experiments, the first one, described in Sect. 4.1, shares a complete explanation of the reputation trajectory calculated for three recognized authors. These reputation trajectories are heuristically interpreted based on the author's publications to confirm that the databases used and the metrics proposed are valid.

The second experiment, described in Sect. 4.2, addresses three different issues. In the first one, the relevance index, the novelty index, and the reputation score are calculated and compared to the h-index and the i10 indexes. In the second one, a comparison is made to prove the limitations presented by the h-index and the i10 when the author presents a short scientific career but with a great impact. Finally, it introduces a hypothetical scenario to compare the performance using the metric proposed, the h-index, and the i10 when a fraud is committed.

The third experiment, described in Sect. 4.3, is focused on elaborating profiles according to the scientific performance of the researchers based on their reputation trajectories. It runs a clustering algorithm to achieve this task for 45 authors.

4.1 Reputation trajectory comparison of recognized authors

This experiment is used to validate the purpose of this paper heuristically. For this, the reputation trajectory of three particular authors is calculated and interpreted based on their publications. These authors are David Tax, Terrance E. Boult, and Isaac M. de Diego (see Figs. 6, 7, and 8). Table 1 shows the reputation scores of David Tax for each year from 1997 to 2020 as well as the relevance and novelty indexes for this author. Table 2 shows the reputation scores of Terrance E. Boult for each year from 1985 to 2020 as well as the relevance and novelty indexes for this author. Table 3 shows the reputation scores of Isaac M. de Diego for each year from 1997 to 2022 as well as the relevance and novelty indexes for this author.

In the first year of scientific publishing of David Tax, the relevance index is 0.19 and the novelty index is 1. These results are logical, as the work made has not yet had an impact on the community. In addition, the articles are novel since they are the firsts. In 1998, the author publishes 5 articles. Two of these articles are published in high-quartile journals and are cited more than 130 times, so it seems natural that his relevance index increases to 0.72. Since the topics covered this year are similar to those of the previous year, it is also normal for the novelty index to decrease to 0.92. In 1999, the author published 7 articles. Several of those get numerous citations, reaching 1, 803 citations in an article published in a $Q1$ journal. Therefore, the relevance index increased to 1.05. As this index is calculated in cycles of three years (and the first year penalizes), the rise in the relevance index is not very pronounced. The novelty index drops for the same reason to 0.54. In 2000, the author published 5 new papers.
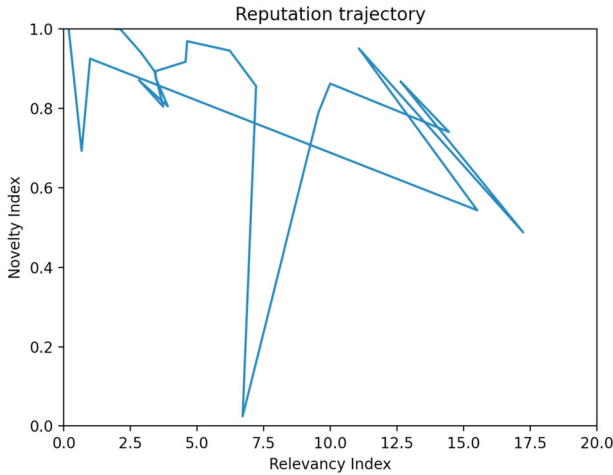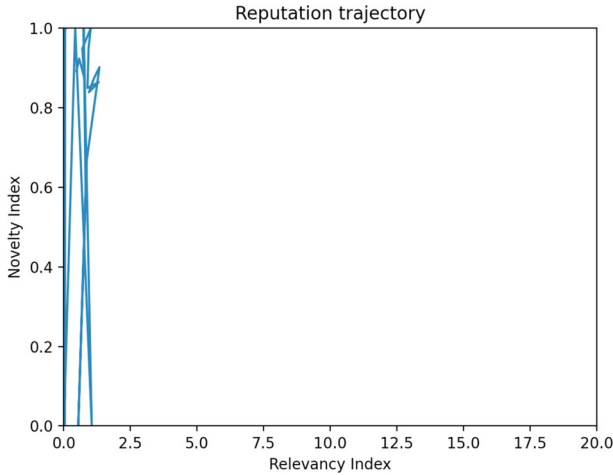
Fig. 6 Reputation trajectory of David Tax
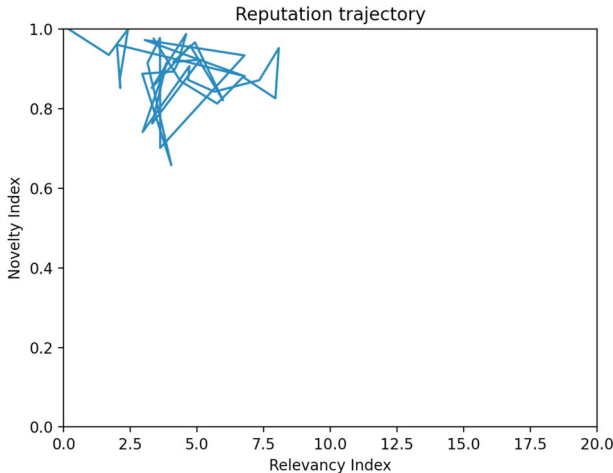
Fig. 7 Reputation trajectory of Terrance M. Boult

One of them reached 8, 136 citations. Besides, the first year (1997) is no longer included in the calculation of the relevance index for his fourth year which contributes to a very steep increase in the relevance index, reaching 16.3. Instead, the novelty index increases again to 0.95, since the topics are novel compared to the publications in previous years. In 2001, the author published 4 articles without too many citations. Furthermore, three of them are published at conferences. In this way, the relevance index has a slight drop to 11.76, although it is still high due to the trajectory of recent years. The novelty index this year decreases to 0.48 because the topics of the articles are similar. In 2002, the author published 6 articles. Three of them are published as the first author, one being the sole author and the other two sharing them only with another author. In addition, it reached a high number of citations. The relevance index of this year is the highest of his scientific career, 18.03. The novelty index increases to 0.86. In 2003, the author published 5 articles with very little impact. The journals

Fig. 8 Reputation trajectory of Isaac M. de Diego

where these articles were published have a low quartile and the articles barely get citations. The relevance index decreases to 13.26, although it is still high due to the previous two years. In 2004, the author published 14 articles. One of them is published in a $Q1$ magazine. This article is written by only two authors and David Tax is the first author. Also, it gets 2, 838 citations. However, the rest of the articles do not have much impact. Therefore, the relevance index remains almost constant, at 13.21. However, due to the lack of new topics in these articles, the novelty index decreases to 0.74. In the following years, the trend is the same and the calculated indexes explain the relevance and novelty of the publications (see Table 1). In 2007, there is a significant drop in the novelty index to 0.02. This is produced due to the publication of only two very similar articles. In 2008, the relevance index also decreased to 6.92. In subsequent years, the relevance index began to decline gradually. The publications were having less and less impact, not exceeding 45 citations in any of them. In addition, David Tax stops publishing as the first author and the articles have a greater number of authors.

There is a pattern and a trend in the author's trajectory. The pattern shows that the author intersperses years in which articles with novel topics and less impact are published, with years in which articles with high impact on subjects in which the author is an expert are published. In the former, the author obtains high novelty indexes and low relevance indexes. In the latter, the author gets low novelty indexes and high relevance indexes. The trend shows that the impact of scientific work is ascending until 2005 and descending from 2006 onward.

In the early years of Terrance E. Boult, the relevance was low because the articles did not get many citations. However, his relevance score is higher than Tax's relevance score in the first three years. Terrance publishes articles alone or with very few authors in higher quartile journals, which explains this fact. The novelty of his articles is also greater than Tax's because he does not focus on the same subjects. In 1991 and 1993, Bault obtains a relevance higher than 6.00 because he exceeds 200 and 500 citations, respectively. However, Boult does not reach a relevance as high as Tax because he published many articles during these years in journals of the lower quartile that do not have citations. Novelty remains high because the topics he deals with remain varied. During the 1990s and 2000s, the relevance index remained constant at around 4.00, far from the relevance indexes reached by Tax in the 2000s because Bault continues to publish in lower quality journals and does not get as many

Table 1 Relevance index, novelty index, and reputation score of David Tax

Year	Relevance index	Novelty index	Reputation score
1997	0.19	1.00	0.44
1998	0.72	0.92	0.81
1999	1.05	0.54	0.75
2000	16.32	0.95	3.94
2001	11.76	0.48	2.38
2002	18.03	0.86	3.94
2003	13.26	0.86	3.38
2004	13.21	0.74	3.13
2005	14.86	0.86	3.57
2006	10.35	0.78	2.84
2007	9.82	0.02	0.44
2008	6.92	0.85	2.43
2009	7.37	0.94	2.63
2010	6.37	0.96	2.47
2011	4.73	0.91	2.07
2012	4.65	0.89	2.03
2013	3.46	0.80	1.66
2014	3.78	0.87	1.81
2015	2.86	0.80	1.51
2016	3.94	0.87	1.85
2017	3.50	0.88	1.75
2018	3.48	0.94	1.81
2019	2.96	1.00	1.72
2020	2.15	1.00	1.47

citations as Tax. In addition, Boult shares articles with several authors and Tax publishes alone or with only another author. In 2014, 2015, and 2016, an increase in Bault's relevance index can be seen. In 2012 and 2016, his two articles with the highest number of citations are published. Note that the relevance index is calculated in three-year cycles; therefore, the year 2014 benefited from the 2012 articles. In addition, in 2013, 2014, and 2015 most of the articles are written by himself with another author. The novelty has remained fairly constant over the years. None of Bault's articles reach the citations generated by the three Tax articles with the most citations.

Finally, Isaac has a much lower relevance than the two previous authors. In his first years, an extremely low relevance index is achieved, with an average of 0.04, although 1997 is the year in which the article with the highest number of citations was published by him. During the first years, his publications were shared with many authors in low-quartile journals. From 2005, his relevance index rises, although it is far from the relevance indexes of the other two authors in the same years, provoked by the publishing with fewer authors and the achievement of more citations. In 2012 and 2013, the second and third highest relevance indexes of his career are obtained thanks to articles published in higher quartile journals and with fewer authors. In addition, his second article with more citations was published in 2010. Note that the calculation of relevance is in three-year cycles. Regarding novelty, Isaac obtained highs and lows during the first 10 years. There are years in which the same topic is treated and

Table 2 Relevance index, novelty index, and reputation score of Terrance E. Boult

Year	Relevance index	Novelty index	Reputation score
1985	0.19	1.00	0.44
1986	1.69	0.93	1.44
1987	2.41	1.00	1.80
1988	2.11	0.87	1.50
1989	2.12	0.85	2.05
1990	2.00	0.96	3.31
1991	6.79	0.88	2.44
1992	3.04	0.97	1.48
1993	6.78	0.93	2.39
1994	3.62	0.70	1.25
1995	3.61	0.97	2.49
1996	3.15	0.91	1.40
1997	4.04	0.65	2.22
1998	2.95	0.88	4.48
1999	4.13	0.89	0.44
2000	4.30	0.91	1.44
2001	4.59	0.98	1.80
2002	3.30	0.85	1.50
2003	4.74	0.95	2.05
2004	5.98	0.82	3.31
2005	4.92	0.96	2.44
2006	4.10	0.91	1.48
2007	5.12	0.92	2.39
2008	3.32	0.76	1.25
2009	3.86	0.91	2.49
2010	2.96	0.74	1.40
2011	4.71	0.90	2.22
2012	4.66	0.87	4.48
2013	5.66	0.84	2.36
2014	7.34	0.87	1.27
2015	8.07	0.95	3.61
2016	7.94	0.82	1.84
2017	6.72	0.87	3.49
2018	5.75	0.81	3.61
2019	4.37	0.86	1.84
2020	3.37	0.97	3.49

years in which the topic is radically changed. This explains why Isaac goes from 1.00 to 0.00 during the first ten years. The novelty has become more constant and high in the last ten years.

The interpretation of the reputation trajectory of these three authors shows that the proposed metrics provide interesting results, taking into account factors ignored by other metrics. Therefore, the reputation of an author should not be measured solely based on the relevance of

Table 3 Relevance index, novelty index, and reputation score of Isaac M. de Diego

Year	Relevance index	Novelty index	Reputation score
1997	0.09	1.00	0.27
1998	0.05	1.00	0.24
1999	0.04	0.00	0.03
2000	0.03	0.00	0.02
2001	0.03	1.00	0.22
2002	0.02	0.00	0.02
2003	0.02	0.00	0.01
2004	0.01	0.00	0.01
2005	0.43	1.00	0.54
2006	1.05	0.00	0.84
2007	0.75	1.00	0.80
2008	0.82	0.59	0.78
2009	0.82	0.59	0.77
2010	0.55	0.00	0.44
2011	0.87	0.66	0.83
2012	1.34	0.90	1.25
2013	1.14	0.87	1.09
2014	0.95	0.83	0.93
2015	1.35	0.86	1.22
2016	0.89	0.84	0.88
2017	0.03	0.95	0.93
2018	1.01	1.00	1.01
2019	0.70	0.94	0.75
2020	0.78	0.87	0.80
2021	0.58	0.92	0.65
2022	0.48	0.89	0.56

the produced work. Besides, relevance should capture details such as the number of authors participating in the published article and should be penalized if the author's work begins to lose quality over the years. Metrics such as h-index or i10 do not decrease even if the researcher starts to publish very low-quality articles. That explains why Bault has a higher h-index despite never having managed to write an article that equals the number of citations of the three articles with the highest citations in Tax, for example.

4.2 Comparison between the metrics proposed and alternatives

The purpose of the second experiment is to validate the performance of the proposed metrics. For this intent, three different validations have been carried out to compare the proposed metrics with the h-index and i10 index (both provided by Google Scholar). The first one calculates the relevance index, the novelty index, and the reputation score for 45 different authors. The second one compares the h-index and i10 to the reputation score for an author with a very short scientific career but a great impact. The third one introduces a hypothetical scenario to detect possible fraud.

4.2.1 Correlation between the reputation score and h-index and i10 index

This first validation has been carried out with 45 different authors. These authors are selected based on their public reputation due to the h-index: 25 medium reputation authors, 18 high reputation authors, and 2 top reputation authors. The authors of each group have a similar number of articles, citations, and seniority, to avoid the h-index and i10 introduce errors due to the limitations they present. The objective of this experiment is to compare the reputation scores with the h-index and i10 for each author. For this, the reputation score has been calculated for all authors. Table 4 shows the relevant information for each author.

The Pearson correlation between the reputation score and h-index is 0.75, between the reputation score and i10 index is 0.54, and between the h-index and i10 index is 0.93. The high correlation between the h-index and i10 index is because both metrics are calculated similarly. The reputation score considers the quality and the novelty of the works; hence, it is normal that the correlation between the reputation score and the other metrics is lower.

As the correlation with the h-index is higher than with the i10 index, the Pearson correlation between the h-index, and relevance and novelty indexes are considered. This way it can be tested the weight of the quality and the novelty of works in the h-index. The former is 0.72 and the latter is 0.49. Notice that the correlation between the relevance index and the h-index is not as high as between the h-index and the i10 index. The variables included in the reputation score that consider the quality of the works provoke this drop. Besides, the correlation between novelty and the h-index is even lower. These results are because the h-index does not consider the novelty of the works of the authors. Therefore, the reputation score proposed in this paper seems more complete than the h-index and i10 index.

4.2.2 H-index and i10 limitations when an author has a short scientific career but a great impact

The lack of information about the quality and novelty of the works when calculating the h-index and i10 index makes these indexes useless to profile the reputation of an author with a short scientific career but a great impact. There are enough authors with non-prolific careers that have had a huge influence on the evolution of important science matters. One of them is Evariste Galois [15], whose case is well known in the scientific community and is always used as an example of one of the limitations of those metrics such as h-index or i10 present. The Galois reputation score cannot be calculated because the values of the variables required to calculate the metric are not known. However, this experiment, it is calculated that the reputation score of a hypothetical author could be Galois himself. Hence, it is proposed a scientific researcher published only two articles in his entire life and did so in the same year. Both articles were published in Q1 journals as sole author and received 15,550 citations and 33,450 citations, respectively. The reputation score for this author is 490. Instead, this author has an h-index and i10 of 2. The reputation of this author is high due to the influence of his two articles. Thus, it can be concluded that it is important to consider the quality and novelty of the works when calculating a reputation metric.

4.2.3 Fraud detection

The last issue addressed in the second experiment is fraud detection. A hypothetical scenario is introduced to prove that metrics such as h-index and i10 do not detect possible fraud, while the proposed reputation measure can detect it.

Table 4 Reputation score, h-index, and i10 for a set of 45 authors

Author name	Reputation score	H-index	i10
Andrew N. Meltzoff	4.08	123	296
D.E. Walling	2.31	120	477
Xavier Marie	0.74	69	226
Sefaattin Tongay	1.46	69	181
Sebastien Marcel	1.85	66	167
Prat N	1.04	62	200
Terrance E. Boult	1.90	62	177
Salvador García	2.03	61	125
Douglas Irwin	1.81	56	113
Timothy J. Hatton	1.60	56	128
Jan Luiten Van Zanden	0.87	56	157
Tom Heskes	2.11	55	151
Cindy Morris	1.54	55	107
David Studholme	2.29	54	136
Víctor Izquierdo-Roca	1.11	53	138
Joao Carlos Setubal	1.91	53	123
Francesc Gallart	1.03	51	122
Giorgio Giacinto	1.94	51	104
David Tax	2.12	50	114
Jean Greenberg	2.81	49	66
Matthieu Boisgontier	0.86	31	64
Nerantzis Kazakis	1.12	31	50
Ioannis Kougias	1.05	30	36
Marco Petitta	0.56	27	78
Teodoro Estrela	0.69	27	51
Polemio M	0.55	26	68
Olga Petrucci	0.63	26	52
Katja Fiehler	1.28	26	50
Antonio Alonso Ayuso	0.93	26	38
David S. Jacks	1.14	25	33
Zanoni Dias	0.83	24	53
Tine de Moor	0.98	24	42
Bert de Vries	1.20	24	39
Bas Van Leeuwen	0.40	23	47
Rainer Fremdling	0.78	23	40
Giovanni Pardini	0.69	23	36
Ángela Aurora Pasqua	0.57	23	36
Laura Ding	0.54	22	34
Tjeerd Dijkstra	0.56	22	34
Eduardo M. García Roger	1.04	22	31
Thorsten Rissom	0.36	21	33
Isaac M. de Diego	0.61	20	36
Syantan Ghosal	2.11	20	31
Bishnupriya Gupta	0.76	19	31
Kelvin S. Oie	1.00	17	18
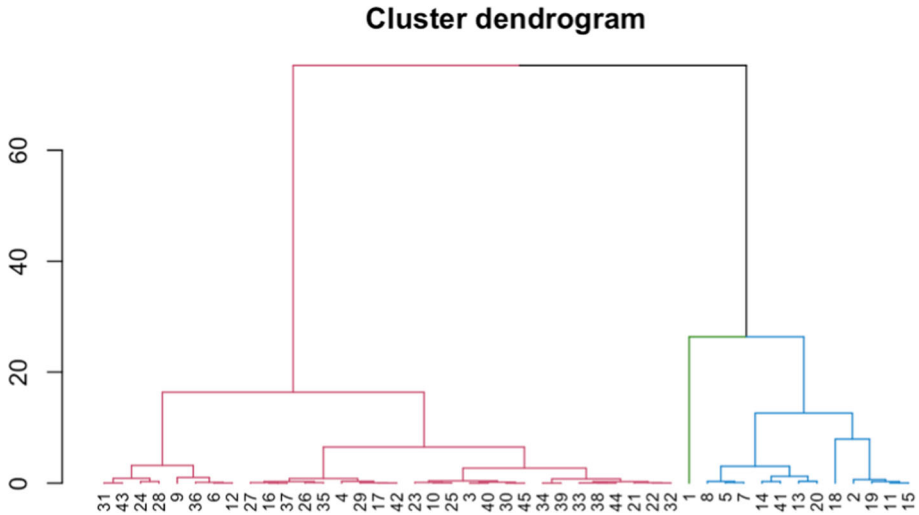
Fig. 9 Hierarchical clustering based on the reputation trajectories of 45 authors

The hypothetical fraudulent scenario is as follows: a group of 10 authors achieves an h-index of 30 in two years using fraudulent techniques. To do this, 30 articles are published among the 10 authors. Each author writes only 3 articles in these two years and all the authors are included in all the papers. To facilitate the publishing work, all articles are submitted to CORE C conference publications indexed by Google Scholar. To achieve the 30 citations per article necessary to reach the h-index equal to 30, all articles are cited in subsequent articles, as well as in articles by friend scientists. This way, these 10 authors achieve an h-index of 30 in 2 years, despite having published in low-reputation publications, sharing these publications with another 9 authors, and getting non-legitimate citations. The h-index does not consider important factors such as the number of authors per article or the reputation of the journal where the articles have been published, among other things. Therefore, it does not penalize the possible fraud in these 10 authors. Instead, the reputation score takes into account these factors and others, penalizing the overall reputation score of these authors.

4.3 Author's clustering by reputation score and h-index

The third experiment studies the clustering of the set of 45 authors previously presented in Sect. 4.2. This clustering is based on their reputation trajectories and results are compared with the one based on the h-indexes.

Notice that the clustering is not based on the reputation scores but on the reputation trajectories. This way the information of the whole scientific trajectory enriches the results. For this, grid-based analysis and the well-known DTW algorithm is performed to calculate distances from trajectories. These distances are used to feed a hierarchical clustering algorithm. The resulting dendrogram with cluster information is shown in Fig. 9.

It can be observed that 3 different clusters have been generated. Table 5 shows the comparison of the clusters generated with the proposed metric and the clusters based on the h-index. When profiling these clusters, it can be indicated that the first author of the table is a top researcher, 12 authors labeled as B have a high reputation and, finally, the rest of the authors have a medium reputation. The clustering based on the reputation trajectories is quite similar

Table 5 Trajectories clusters and h-index clusters

Author name	Trajectories cluster	H-index cluster
Andrew N. Meltzoff	A	A
D.E. Walling	B	A
Xavier Marie	C	B
Sefaattin Tongay	B	B
Sebastien Marcel	B	B
Prat N	C	B
Terrance E. Boult	B	B
Salvador García	C	B
Douglas Irwin	C	B
Timothy J. Hatton	C	B
Jan Luiten Van Zanden	B	B
Tom Heskes	B	B
Cindy Morris	C	B
David Studholme	B	B
Víctor Izquierdo-Roca	C	B
Joao Carlos Setubal	B	B
Francesc Gallart	C	B
Giorgio Giacinto	B	B
David Tax	B	B
Jean Greenberg	B	B
Matthieu Boisgontier	C	C
Nerantzis Kazakis	C	C
Ioannis Kougias	C	C
Marco Petitta	C	C
Teodoro Estrela	C	C
Polemio M	C	C
Olga Petrucci	C	C
Katja Fiehler	C	C
Antonio Alonso Ayuso	C	C
David S. Jacks	C	C
Zanoni Dias	C	C
Tine de Moor	C	C
Bert de Vries	C	C
Bas Van Leeuwen	C	C
Rainer Fremdling	C	C
Giovanni Pardini	C	C
Ángela Aurora Pasqua	C	C
Laura Ding	C	C
Tjeerd Dijkstra	C	C
Eduardo M. García Roger	C	C
Thorsten Rissom	C	C
Isaac M. de Diego	C	C
Syantan Ghosal	B	C
Bishnupriya Gupta	C	C
Kelvin S. Oie	C	C

to the clustering based on the h-indexes with some differences. Regarding these differences, only one author labeled as C in the h-index clusters has been labeled as B in the trajectories clusters. This author is Syantan Ghosal, and the reason is that the novelty of his works is high. Regarding the authors labeled as B in the h-index clusters, 8 out of 18 have been labeled as C in the trajectories clusters. Most of them have been downgraded due to the quality and novelty of their work. However, Xavier Marie and Prat N. have been downgraded because Semantic Scholar has not indexed their most influential articles. Regarding the authors labeled as A in the h-index clusters, D.E. Walling has been labeled as B in the trajectories clusters. This is due to two main reasons, the novelty of his works is not high and the quality of his latest works has decreased abruptly.

In conclusion, this clustering has shown that it provides acceptable results to classify authors according to their reputations. The proposed metrics act as a filter discarding edge and fraudulent cases. Thus, it can be said that this framework is a functional prototype to analyze and cluster authors by their scientific trajectories.

5 Conclusions

This paper has presented the *FRESA* framework focused on estimating the reputation of researchers. Semantic Scholar and other web information sources have been used to collect specific knowledge from authors and their associated publications. This functionality is completed with the estimation of two indexes: the relevance index and the novelty index. They have been defined according to different features. For the relevance index, seven features have been used, and text-based similarity measures were applied to calculate them. In the case of the novelty index, simple NLP techniques are used to gather the textual information used to estimate it. Both indexes are used to calculate the reputation score.

By calculating the proposed indexes over the years for authors, it is possible to analyze their scientific trajectory. This eases to study in greater detail than other indexes of the scientific career of researchers. For instance, it could be detected when authors have published very relevant and novel works during a time. In the same way, this feature could identify when those authors reduce the publishing workload or their research has stalled on similar topics over the years.

Promising results have been accomplished in the experiments carried out to prove the viability of the proposal comparing it with alternative indexes. Thus, the proposal has shown that it is a more complete approach than the existing ones. It has been able to detect similar features through the relevance index, while it has provided more information using the novelty index. For this reason, it can be said that the framework abilities and the reputation index successfully overcome the literature of the domain.

FRESA has the potential to be used as a relevant tool for research agencies, expert evaluation panels, and the human resources sections. Thanks to its capability to depict trajectories and make visual comparisons, it is relatively easy to discriminate between profiles of researchers to find out the best ones who adapt to certain established prerequisites.

Even though the *FRESA* framework represents a complete system, it is only an initial prototype. Thus, future enhancements can be applied to increase the overall performance. For instance, the novelty index can be improved if information about the papers of other authors is entered. Notice that in this paper the novelty index is calculated considering only the publications of the author. Besides, the relevance index could be redefined taking into account the order of the authors within the paper. This fact would update the component that

has already been introduced in the metric by considering the number of authors per paper. Also, information about the organization where researchers have developed the different scientific approaches could be included. This would consider the role of institutions and how they can be associated with the author's research focus. However, it is an open issue that requires additional study. Finally, further research with other clustering algorithms would be interesting to select several heterogeneous profiles of authors to detect future improvements to include.

References

1. Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F (2010) hg-index: a new index to characterize the scientific output of researchers based on the h-and g-indices. Scientometrics 82(2):391–400
2. Ballester O, Penner O (2019) Evolution of topics and novelty in science. In: ISSI, pp 1606–1611
3. Cappelletti-Montano B, Columbu S, Montaldo S, Musio M (2021) New perspectives in bibliometric indicators: moving from citations to citing authors. J Informet 15(3):101164
4. Carayol N, Agenor L, Oscar L (2019) The right job and the job right: novelty, impact and journal stratification in science. Impact J Stratific Sci (March 5, 2019)
5. Cervi CR, Galante R, de Oliveira JPM (2013) Comparing the reputation of researchers using a profile model and scientific metrics. In: 2013 IEEE 16th international conference on computational science and engineering. IEEE, pp 353–359
6. Cohen BA (2017) Point of view: How should novelty be valued in science? Elife 6:e28699
7. de Freitas Veneroso JM, Dias M, Ueda A, Ribas S, Ribeiro-Neto B, Ziviani N, e Silva ES (2019) P-score: a reputation bibliographic index that complements citation counts. Scientometrics 121(3):1269–1291
8. Dhamdhere SN (2018) Cumulative citations index, h-index and i10-index (research metrics) of an educational institute: a case study. Int J Libr Inf Sci 10(1):1–9
9. Egghe L (2006) An improvement of the h-index: the g-index. ISSI Newsl 2(1):8–9
10. Falagas ME, Kouranos VD, Arencibia-Jorge R, Karageorgopoulos DE (2008) Comparison of scimago journal rank indicator with journal impact factor. FASEB J 22(8):2623–2628
11. Fernández-Isabel A, Prieto JC, Ortega F, de Diego IM, Moguerza JM, Mena J, Galindo S, Napalkova L (2018) A unified knowledge compiler to provide support the scientific community. Knowl Based Syst 161:157–171
12. Fernández-Isabel A, Barriuso AA, Cabezas J, de Diego IM, Pinheiro JV (2020) Knowledge-based framework for estimating the relevance of scientific articles. Expert Syst Appl 161:113692
13. Fricke S (2018) Semantic scholar. J Med Library Assoc JMLA 106(1):145
14. Fujimura K, Nishihara T (2003) Reputation rating system based on past behavior of evaluators. In: Proceedings of the 4th ACM conference on electronic commerce, pp 246–247
15. Galois É, Neumann PM (2011) The mathematical writings of Évariste Galois, vol 6. European Mathematical Society
16. Gamon M (2006) Graph-based text representation for novelty detection. In: Proceedings of TextGraphs: the first workshop on graph based methods for natural language processing, pp 17–24
17. Garfield E et al (1994) The impact factor. Curr Contents 25(20):3–7

18. Gormley C, Tong Z (2015) Elasticsearch: the definitive guide: a distributed real-time search and analytics engine. O'Reilly Media Inc, New York
19. Guerrero-Sosas JD, Chicharro FPR, Serrano-Guerrero J, Menendez-Dominguez V, Castellanos-Bolaños ME (2019) A proposal for a recommender system of scientific relevance. Procedia Comput Sci 162:199–206
20. Hendrikx F, Bubendorfer K, Chard R (2015) Reputation systems: a survey and taxonomy. J Parallel Distrib Comput 75:184–197
21. Hirsch JE, Buela-Casal G (2014) The meaning of the h-index. Int J Clin Health Psychol 14(2):161–164
22. Holding AN (2019) Novelty in science should not come at the cost of reproducibility. FEBS J 286(20):3975–3979
23. Inc C (2022) The computing research and education association of Australasia. https://www.core.edu.au/. Accessed: 2022-09-29
24. Jamali HR, Nicholas D, Herman E (2016) Scholarly reputation in the digital age and the role of emerging platforms and mechanisms. Res Eval 25(1):37–49
25. Johnson SC (1967) Hierarchical clustering schemes. Psychometrika 32(3):241–254
26. Kessler MM (1963) Bibliographic coupling between scientific papers. Am Doc 14(1):10–25
27. Krampl A (2019) Journal citation reports. J Med Libr Assoc JMLA 107(2):280
28. Li B, Han L (2013) Distance weighted cosine similarity measure for text classification. In: Intelligent data engineering and automated learning–IDEAL 2013: 14th international conference, IDEAL 2013, Hefei, China, October 20–23, 2013. Proceedings 14. Springer, pp 611–618
29. Liao H, Zeng A, Xiao R, Ren ZM, Chen DB, Zhang YC (2014) Ranking reputation and quality in online rating systems. PLoS ONE 9(5):e97146
30. Ling G, King I, Lyu MR (2013) A unified framework for reputation estimation in online rating systems. In: Twenty-third international joint conference on artificial intelligence, pp 2670–2676
31. López-Cózar ED, Orduña-Malea E, Martín-Martín A (2019) Google scholar as a data source for research assessment. In: Springer handbook of science and technology indicators. Springer, pp 95–127
32. Massucci FA, Docampo D (2019) Measuring the academic reputation through citation networks via pagerank. J Informet 13(1):185–201
33. Meho LI, Rogers Y (2008) Citation counting, citation ranking, and h-index of human-computer interaction researchers: a comparison of scopus and web of science. J Am Soc Inform Sci Technol 59(11):1711–1726
34. Mizzaro S (1997) Relevance: the whole history. J Am Soc Inf Sci 48(9):810–832
35. Ng KW, Tsai FS, Chen L, Goh KC (2007) Novelty detection for text documents using named entity recognition. In: 2007 6th international conference on information, communications and signal processing. IEEE, pp 1–5
36. Page L, Brin S, Motwani R, Winograd T (1999) The pagerank citation ranking: bringing order to the web. Technical report, Stanford InfoLab
37. Priem J, Taraborelli D, Groth P, Neylon C (2011) Altmetrics: a manifesto. http://altmetrics.org/manifesto. Accessed: 2022-09-29
38. Ramos J et al (2003) Using tf-idf to determine word relevance in document queries. In: Proceedings of the first instructional conference on machine learning, Citeseer 242, pp 29–48
39. Ruohomaa S, Kutvonen L, Koutrouli E (2007) Reputation management survey. In: The second international conference on availability, reliability and security (ARES'07). IEEE, pp 103–111
40. Schiffman B, McKeown K (2005) Context and learning in novelty detection. In: Proceedings of human language technology conference and conference on empirical methods in natural language processing, pp 716–723
41. Senin P (2008) Dynamic time warping algorithm review. Inf Comput Sci Dept Univ Hawaii Manoa Honolulu, USA 855(1–23):40
42. Sosa JDTG, Domínguez VHM, Bolaños MEC, Montalvo JRG (2019) Use of an ontological model to assess the relevance of scientific production. IEEE Lat Am Trans 17(09):1424–1431
43. Veugelers R, Wang J (2019) Scientific novelty and technological impact. Res Policy 48(6):1362–1372
44. Wang J, Veugelers R, Stephan P (2017) Bias against novelty in science: a cautionary tale for users of bibliometric indicators. Res Policy 46(8):1416–1436
45. Yu MC, Wu YCJ, Alhalabi W, Kao HY, Wu WH (2016) Researchgate: an effective altmetric indicator for active researchers? Comput Hum Behav 55:1001–1006
46. Zhu X, Turney P, Lemire D, Vellino A (2015) Measuring academic influence: not all citations are equal. J Am Soc Inf Sci 66(2):408–427