Negative Citing: Scenario Analysis of Inaccurate Evaluation of Scientific Content

Alina Petrushka and Maria Komova

Lviv Polytechnic National University, S. Bandery st. 12, Lviv, 79013, Ukraine

Abstract

The article investigates the influence of variability in the implementation of scientific communication on the quality of application of traditional metrics based on accounting the number of publications and their citations, consider the context of the citation. The purpose of the study is to conduct a scenario analysis of the realization of methodological risk of negative citing when using traditional metrics for evaluating scientific content. To achieve this goal the following tasks were performed: conducting a scenario analysis of the basic citation analysis to form the conceptual basis of the study; formation of principles of realization variative component of citation analysis, taking into account the risk of negative citing; conducting a scenario analysis of the citation scheme taking into account the risk of negative citing. The main research methodology is the method of scenario analysis, which was used to analyze the realization of methodological risk of negative citing in the evaluation of scientific content by traditional metrics. To increase the objectivity of evaluation of scientific content it is necessary to move to a variative level of citation analysis, which involves the identification and consideration of possible methodological risks of traditional metrics. Implementation of a variative component of evaluation of scientific content to account for the risk of negative citing provides identifying a marker of the scientist's belonging to the risk group of negative citing, verification of negative citation and analysis of the reasons for criticism and indication negative citation to establish the weight of the marker in the certain scientific subject area. Combining the results of expert evaluation and parameterization will allow reducing the subjectivity of the negative citing indication.

Keywords 1

Scientific content, scientific information, researcher impact, evaluation, quantitative assessment, citation analysis

1. Introduction

The development of new alternative metrics or the improvement of existing ones is an important methodological problem of modern bibliometric research. The search for new solutions makes it obvious that the existing methodological tools fail to adequately and comprehensively create an evaluative picture of the scientific information space. Despite the imperfections of traditional bibliometric tools today, it determines the basis for the functioning of undisputed monopolies in the field of assessing the quality of scientific knowledge Scopus and Web of Science. Aware of the possible shortcomings of traditional metrics, the use of citation analysis remains the only source of data that allows mapping the scientific information space. Today traditional metrics are the basic means of navigation for researchers under conditions of the rapid accumulation of scientific content.

The purpose of the study is to conduct a scenario analysis of the realization of the methodological risk of negative citing using traditional metrics for evaluating scientific content. To achieve this goal, the following tasks were performed: conducting a scenario analysis of the basic citation analysis to

COLINS-2021: 5th International Conference on Computational Linguistics and Intelligent Systems, April 22–23, 2021, Kharkiv, Ukraine
EMAIL: alina.i.petrushka@lpnu.ua (A. Petrushka); mariia.v.komova@lpnu.ua (M. Komova)
ORCID: 0000-0002-8769-4599 (A. Petrushka); 0000-0002-4115-3690 (M. Komova)


form the conceptual basis of the study; formation of conceptual principles of realization of a variative component of the citation analysis taking into account risk of negative citing; conducting a scenario analysis of the citation scheme, taking into account the risk of a negative citation.

The main research methodology is the method of scenario analysis, which was used to analyze the realization of the methodological risk of negative citing in the process of evaluating scientific content by traditional metrics. Scenario analysis involves the use of tools of scientific methods of analysis, synthesis, formalization, visualization. Analysis and synthesis were used to form the conceptual principles for the implementation of basic and variative citation analysis as components of the system of evaluation of scientific content. The methods of formalization and visualization were used to represent and illustrate the process of citing as a reflection of the communication process between authors of scientific publications with the help of the conceptual and mathematical apparatus of the set theory.

2. Recent research and publications

The study of scientific value and relevance of the achievements of the individual researcher or institution, publication's impact on the development of scientific creativity is a relevant and popular area of research in academic, information-analytical, library structures.

2.1. Potential and limitations of traditional metrics and altmetrics for evaluating the researcher

Cabezas-Clavijo A. and Delgado-López-Cózar E. study the features, advantages and limitations of scientific evaluation methods based on Google Scholar: method of calculating the h-index, which is a standard bibliometric indicator. The h-index of the most relevant journals and researchers based on data was obtained from Web of Science, Scopus and Google Scholar. The results show that the average values of the h-index in Google Scholar are almost 30% higher than the values obtained in Web of Science, and approximately 15% higher than those collected by Scopus. However, there are no significant differences in the ratings [1].

Zhang C.-T. believes that the h-index is a simple and effective indicator for evaluating the educational activities of researchers, but it has critical shortcomings. It leads to bias and inaccuracy of assessments. A method for calculating the h'-index is proposed, which takes into account the values of h-square, excess and h-tail. Compared to the h-index, the h'-index has a better correlation with the indices of the total number of citations and citations from a single publication, has the ability to distinguish between types of citations. The h'-index improves the h-index, as well as the indices of the total number of citations and references to the publication. It can distinguish between types of citations and can be used to make a more accurate assessment [7].

Liu Y. and Sun Z. also state that one of the most popular indicators for assessing the scientific impact of the authors is the h-index. The h-index is efficient and easy to use, but it has disadvantages. It does not take into account significant and important segments of information. To address these shortcomings, scientists have proposed the g-index and EM-index, but they also operate only on the number of articles with a high level of citation, and the total number of cited articles is not fully taken into account. However, any cited article is important for scientific evaluation and contains unique information about scientists. Therefore, the concept of S-index is proposed, which justifies the use of bilateral h-index. The S-index takes into account all citations of publications and forms a single number for human scientific influence. To illustrate the advantage of the S-index, an experimental analysis of 100 cited articles was performed. The analysis showed that the S-index is a more balanced and accurate method of quantifying the scientific impact of the authors [24].

Maqsood S., Islam M. A., Afzal M.T. and Masood N. state that the authors of research papers are given a rating, which is taken into account when shortlisted for scientific awards, prestigious positions, appointments, invitations to the keynote speaker or distribution of grants for research. The rating of scientists is established based on traditional bibliometric indicators, which include: the number of publications, the number of citations and the h-index. However, clear citations appear only

over time. Therefore, traditional bibliometric indicators for ranking authors are not favourable for young researchers. They reduce their competitiveness over senior colleagues. Measures for the creation of a centralized network of co-authors are proposed. The purpose of the experiment: identification and ranking of authors who have been awarded by various scientific societies. The rating is additionally influenced by bibliographic indicators and network indicators. The experiment showed that network indices and bibliometric h-index have the same potential to identify influential authors, especially in the case of networks with a large value of the cluster coefficient [19].

D'Agostino M., Dardanoni V. and Ricci R.G. investigate methods of assessing scientific productivity and scientific impact, in particular, to establish relative estimates using percentages. This is a purely quantitative method that establishes a formal assessment. In the case of such a method, citation ratings in terms of relative values coincide with absolute [14].

Tam C. H., Tan S. B. and Soo K. C. study methods of assessing the scientific quality of individual research articles and individual journals as factors of their influence. The factor of influence is taken into account when considering grant applications, resolving career issues and forms of material incentives. Determining the impact factor of the journal as an indicator of research quality, establishing an index of scientific citations and the use of bibliometric methods are the topical areas of research. The journal's impact factor qualifies as a quantitative indicator based on the ratio between annual citations in a particular journal and the total number of citations in that journal for the previous 2 years. However, its use to measure the quality of research should be within the statement, tolerance, trend. The quality of research cannot be measured by the influence of the journal alone as dominant or unique [5].

Sedighi M. studies the impact of scientific work, calculated using traditional methods of scientometrics and using the methods of altmetry (metrics of social networks). The study covers cited papers published in journals (Scientometrics and Journal of Informetrics) over a five-year basis. The study used Springer and ScienceDirect databases, an Altmetric Explorer service provided by Altmetric.com. 830 scientific papers with altmetric scores were identified, which reflect the quantitative and qualitative indicators of the authors' attention to social media. The study found a correlation between altmetric scores and citation scores using correlation tests. The results showed a significant, positive and weak statistical relationship between the number of citations of works calculated using traditional scientometric methods and the altmetric scores of these works. One aspect of the study was to establish the number of readers of the researched works on social networks (Mendeley and Citeulike) and compare them with the number of their citations. The results of the study show that some social networks and their indicators can reproduce the impact of scientific work. However, altmetric indicators can be used as additional scientometric indicators in assessing the impact of research [15].

Lazarev V. S. and Nazarovets S. A. point to the limitations of traditional bibliometric research in evaluating non-English references in the context of linguistic expansion in the field of scientific publication. Replacing non-English references with their English-language counterparts results in a distorted citation pattern [23].

Thus, recent research in the field of bibliometrics shows that traditional metrics, such as the h-index, serve as a basis for evaluating the impact of the individual researcher. At the same time, the obvious shortcomings of quantitative evaluation based on citation analysis and the limitations of providing a qualitative component of scientific impact assessment require the improvement of existing metrics or the formation of new approaches. Another important trend in the development of bibliometrics at the present stage is the active involvement in the traditional bibliometrics methodology of altmetrics tools as a qualitative supplement to quantitative indicators.

2.2. Use of traditional metrics and altmetrics for evaluating scientific editions

Lazarev V. S. investigates the discipline impact factor (DIF), which is used to evaluate serial publications. For the study, an analysis of the scientific literature, evaluative judgments was conducted. The information base of the research is important, thorough bibliometric publications on the influence of bibliometrics and its perception. Examples of the application of DIF and its

modifications are presented. The use of DIF to evaluate serial publications is relevant to university research libraries that have limited funding for subscription. Evaluation of scientific serials is an important area of library activity, even in the conditions of using database subscriptions and open access to a significant number of publications. Open access to publications is not able to compensate for the gaps. DIF is important when choosing the best journal to submit an article. DIF is offered for wide practical application. The original indicator - "symmetrical" DIF ("discipline susceptibility factor") and its differences from DIF in terms of content and purpose of evaluation are presented. Symmetric DIF is designed to search for serial publications on intersectoral research that are of potential value for widespread application of research results. Both methods are of interest to scientific libraries of universities in the field of scientific information services [22].

Guerrero-Sosa J. D. T., Menendez-Dominguez V. H. and Castellanos-Bolaños M. E. study quantitative statistics to measure the scientific value of publications concentrated in high-performance open content digital repositories. The proposed methodology of action research (AR), which provides an algorithm of connections: research is related to practice; research informs practice; practice informs research. AR covers five phases: development of a problem scenario; analysis of modern technical support; testing methodology; publication of results. The method is used to characterize group and individual research results at the State University of Mexico. Data from Scopus and the National Open Access Repository of Mexico were studied. It is established that the AR system does not depend on a particular technological tool and is based on universal standards for the description and exchange of metadata, which creates conditions for the integration of new elements for evaluation. The proposed method involves the use of indicators that assess the scientific value of the publication based on the importance of its type and degree of cooperation [12].

De Filippo D., Aleixandre-Benavent R. and Sanz-Casado E. study the characteristics of scientific activity in the field of social sciences and humanities, which are less than other fields represented in international databases. Authors state that it is difficult to assess the quality of social and humanitarian journals, ranking researchers and institutions working in these fields. A methodology for ranking Spanish journals in the social sciences and humanities, developed under the auspices of the Spanish Science and Technology Foundation (FECYT), is proposed. Model of rating journals by two dimensions (impact and visibility) based on quantitative criteria. To obtain indicators in each of these dimensions, a set of indicators is set: Index of citation of science, Index of citation of social sciences, Index of citation of arts and humanities, Index of citation of new sources [8].

Thus, an important segment in the field of bibliometric research is to assess the quality of scientific editions. Establishing a correlation between the level of authority of an individual researcher and the level of a scientific journal, where he publishes, can be used in solving personnel problems. Evaluation of scientific content in terms of subscription can be used for rational distribution of the budget of institutions under conditions of limited funding.

2.3. Alternative approaches to the implementation of citation analysis

Petersen A. M., Pan R. K., Pammolli F., Fortunato S. indicate that quantitative evaluation of research requires clear, relatively simple, interdisciplinary methods that do not limit the time of publication of the article. The decision for levelling of restrictions in quantitative estimation of researches in the aspect of time of publication of the article is offered. The authors note that the real value of citation depends on the time of its creation. The nominal number of citations does not correspond to the real scientific influence of the scientist. Therefore, it is necessary to apply a systematic revision of methods for evaluating citations, especially citations of researchers, whose activities cover several decades [2].

Kobayashi N. and Toyoda T. conduct statistical analysis of citations in the semantic network to establish a quantitative assessment of research results. Authors propose to rank researchers in a particular subject area by presenting the k-index, which is the number of papers that contain specific keywords and that are published by the researcher in a certain period [16].

Schoonbaert D. and Roelants G. explore the basic concepts of citation analysis and the factors influencing the journal, which are considered in the context of assessing the quality of individual scientific publications, individual researchers and research structures. Relevant aspects of the study

are technical capabilities, selectivity of the database; restrictions on time, language, subject, type of publication; numerous benefits of authorship and motivation for citation. Citation analysis based on the application of journal influencing factors can be used to assess the scientific weight of publications of individual researchers or scientific institutions, but it should not be considered as the only evaluation criterion [9].

Zhou Q. and Zhang C. study citations to measure the academic impact of books. The only quantitative metric for evaluating books is citation frequency. Quantitative evaluation does not take into account the value of the content of the book. It indicates the lack of completeness of the evaluation results. Therefore, to obtain reliable book evaluation results, it is necessary to integrate frequency metrics and content metrics. Cited books reflect the frequency of citation of books, motives and context of citation, current topics, authoritative authors. The citation does not indicate transparently the context of the citation, the hidden information. Multidimensional citation metrics were used to study the value of books. The next stage is the comparison of the obtained results with the evaluation of books using traditional metrics [17].

Vicencio-Ríos G., Araya-Castillo L., Millán-Toledo C. and Rivera Flores Y. conduct a descriptive and comparative bibliometric study. To identify detailed and systematic information about the brands of scientific activity in 1995-2018, 481 articles published in the WoS database were analyzed, with a total of 11,840 citations. The most influential article was written by Aaker Aaker, the most relevant author was Aaker Aaker, the most famous journal was Psychology & Marketing, the most productive institution was the University of Texas at Austin, and the most influential was the United States. A bibliometric map was compiled, which highlights the results of using the method [10].

Thus, the citation itself, as a basis of bibliometric research, requires a rethinking and revision of approaches to its identification and evaluation.

2.4. Relationship between altmetrics and features of scientific activity

Delgado-Carreón C. C., Machin-Mastromatteo J. D., Romo-González J. R. and Pacheco-Mendoza J. study the influence of the creative component on the scientific activity of university teachers. A survey of 120 university teachers was conducted to evaluate articles on specialized topics. Articles were evaluated on five criteria: novelty; flexibility, fluidity; achievements, dedication; confidence; problem-solving. After the survey, the data were compared with three other data sets: bibliometric data (Scopus), Altmetrics (Dimensions) and reviews (Publons) for the period from 2013 to 2018. A description, statistical analysis, correlation of collected data. It is established that there is a small correlation between the indicators of creativity of researchers and their bibliometric and altmetric indicators. The highest marks for measuring creativity were flexibility and smoothness, devotion. During the analyzed period, researchers published 379 documents, but only 61 of them published articles in journals indexed in Skopje. The analysis shows that all researchers did not find significant differences in their own assessments of their work, despite the level of indicators of scientific activity [4].

Castanha R. G. and Cabrini Grácio M. C. study the inheritance of theoretical and methodological achievements of the researcher by his academic descendants (formation of scientific schools) using the method of bibliographic communication. The method of bibliographic communication contributed to the analysis of the impact of the transfer of scientific identity of a scientist in the academic genealogical network of research researchers [18].

Keshavarz H. and Esmaeili Givi M. study methods for assessing the accuracy of the information available on scientific websites. A scale for measuring the reliability of information was developed based on a questionnaire. It was conducted based on a stratified random sample among 672 students of Iranian universities. The results show that the reliability of the information on scientific websites can be measured by evaluating 8 criteria: ethics, writing style, website appearance, website identity, professional information, accuracy, ease of use and interaction. These are the most important criteria for studying the assessment of the reliability of scientific web information by students [11].

Davidavičiene V., Al Majzoub K. and Meidute-Kavaliauskiene I. study the development and use of information and communication technologies (ICT), which have enabled the development of new types of work organizations - virtual organizations, virtual teams, distance work and learning. Virtual

team (VT) technology allows you to bring together creative, talented professionals who do not have direct access to the office but can ensure the competitiveness of the firm. However, the problem area is the exchange of knowledge in virtual teams. Studies of this problem have been conducted in the United States and Europe. The results of a study of knowledge exchange in virtual teams, which was conducted in the United Arab Emirates (UAE) in the IT industry and takes into account the specifics of the region. A model for assessing the impact on knowledge sharing in virtual teams has been created. It was found that the factors that have a direct impact on the exchange of knowledge in the field of VT include: culture, motivation, conflict, ICT, trust and leadership. The results show that language does not affect the exchange of knowledge [21].

Smolinsky L. and Lercher A. J. investigate the influence of bibliometric methodology on scientific activity, as well as the dependence of the nature of the impact on the method of evaluating citations of publications of an individual researcher. The study uses the printed journals of the American Physical Society to classify articles by subject and to gather information about citations, publications, and authors. Researchers determine the total number of citations, the number of publications of the author on a particular topic. To summarize the results of the application of the bibliometric methodology, the method of lending to individual scientists is used [13].

Makris C., Pispirigos G. and Simos M. A. study the process of text annotation, which provides the ability to identify the meanings of the text segment within a given context. There are limitations to the word paradigm in annotations. Therefore, the use of information retrieval technologies and artificial intelligence necessitates overcoming the semantic difficulties that arise due to the presence of the phenomena of polysemy and homonymy in natural language. The overactive search for information is inherent in such scientific fields as digital marketing, bioinformatics, chemical engineering, neurology and social sciences, and community discovery. A new distributed controlled community discovery methodology based on the use of a community discovery algorithm to annotate Wikipedia text is presented. This is the concept of a coherent community as a metric for establishing compatible local community coherence. The affinity and coherence of author communities in Wikipedia bring significant improvement and accuracy of new or incomplete annotations [6].

Thus, in recent years there has been a tendency to combine the tools of traditional and alternative metrics to provide quantitative and qualitative components of the evaluation of scientific content. However, the methodologies for implementing citation analysis, both within traditional and alternative approaches, do not provide for the identification and consideration of the context of citations, namely the identification of negative evaluative judgments of the author of the citation.

3. Results

The traditional methodology for evaluating the authority and demand for scientific content is developed on basic citation analysis. It involves the identification and accounting of citations to the publication – the primary source, based on content analysis of the list of references given in publications – secondary sources [3].

Consider the example of the scheme of basic citation analysis, which will form the conceptual basis for the research.

3.1. Scenario analysis of the basic citing scheme

Some primary sources were cited in certain secondary sources and make up the set

$$PS = \{PS_1, PS_2, \dots, PS_j, \dots\} = \{PS_j\}_{j=1}^{N^{(PS)}}, \quad (1)$$

where $N^{(PS)}$ is the number of primary sources in the set PS with powerset

$$|PS| = N^{(PS)}. \quad (2)$$

The fact of citing is considered as the implementation of the communication process between the author of primary sources and the authors of secondary sources. As a result of this scientific communication, links-citation $Cit(PS)$ appears between the certain primary source and the

corresponding secondary source. Citability of primary sources can be represented by countable sets of citations

$$Cit_{(PS_1)} = \{Cit_1(PS_1), Cit_2(PS_1), \dots, Cit_i(PS_1), \dots\} = \{Cit_i(PS_1)\}_{i=1}^{N^{(Cit,PS_1)}}, \quad (3)$$

$$Cit_{(PS_2)} = \{Cit_1(PS_2), Cit_2(PS_2), \dots, Cit_i(PS_2), \dots\} = \{Cit_i(PS_2)\}_{i=1}^{N^{(Cit,PS_2)}}, \quad (4)$$

……

$$Cit_{(PS_j)} = \{Cit_1(PS_j), Cit_2(PS_j), \dots, Cit_i(PS_j), \dots\} = \{Cit_i(PS_j)\}_{i=1\dots N^{(Cit,PS_j)};\, j=1\dots N^{(PS)}}, \quad (5)$$

where $N^{(Cit,PS_1)}, N^{(Cit,PS_2)}, \dots, N^{(Cit,PS_j)}$ are the numbers of citations to primary sources $PS_1, PS_2, \dots, PS_j$ accordingly in the sets $Cit_{(PS_1)}, Cit_{(PS_2)}, \dots, Cit_{(PS_j)}$ with powersets

$$\left|Cit_{(PS_1)}\right| = N^{(Cit,PS_1)}, \quad (6)$$

$$\left|Cit_{(PS_1)}\right| = N^{(Cit,PS_2)}, \quad (7)$$

……

$$\left|Cit_{(PS_j)}\right| = N^{(Cit,PS_j)}. \quad (8)$$

Moreover, the sets of citations $Cit_{(PS_1)}, Cit_{(PS_2)}, \dots, Cit_{(PS_j)}$ do not intersect. The set of citations to all primary sources forms a cumulative set of citations $Cit_{(Cum)}$, which subsets are sets of citations to each primary source

$$Cit_{(Cum)} = Cit_{(PS_1)} \bigcup Cit_{(PS_2)} \bigcup \dots \bigcup Cit_{(PS_j)}. \quad (9)$$

Powerset of cumulative set is determined by the sum of the subsets' powersets (Fig. 1)

$$\left|Cit_{(Cum)}\right| = \sum_{j=1}^{N^{(PS)}} \left|Cit_{(PS_j)}\right|. \quad (10)$$

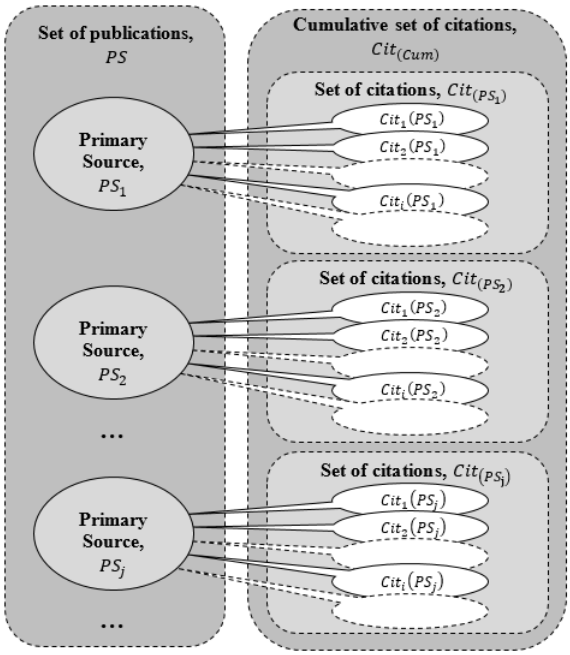

Figure 1: Basic citing scheme

Quantitative assessment of the importance of primary source is determined by the corresponding powersets of the sets of citations $\left|Cit_{(PS_1)}\right|, \left|Cit_{(PS_2)}\right|, \dots, \left|Cit_{(PS_j)}\right|$: the importance of a scientific publication is the greater, the greater is the powerset of the set of its citations.

Quantitative assessment of the influence of the researcher is determined by the powerset of the set of publications $|PS|$, powerset of the cumulative set of citations $\left|Cit_{(Cum)}\right|$ and the scatter of the powerset values of the sets of citations $\left|Cit_{(PS_1)}\right|, \left|Cit_{(PS_2)}\right|, \dots, \left|Cit_{(PS_j)}\right|$:

- level of publishing activity is the higher, the greater is the powerset of the set of publications | PS |;
- cumulative citation index of the researcher is the greater, the greater is the powerset of the cumulative set of citations $\left|\mathrm{Cit}_{(\mathrm{Cum})}\right|$;
- researcher's h-index is the greater, the greater is the powerset of the citation sets $\left|Cit_{(PS_1)}\right|, \left|Cit_{(PS_2)}\right|, \dots, \left|Cit_{(PS_j)}\right|$ with minimal scatter of their values and high publishing activity.

The considered basic citation scheme illustrates the idealized process of scientific communication. It involves reviewing the results of scientific activities without taking into account the qualitative assessment of the citation. It is obvious that the basic citation analysis, which forms the methodological apparatus of traditional metrics, does not cover the possible variability of the results of scientific communication. That is why the importance of a scientific publication and the impact of the researcher can't be adequately assessed only by the powerset of the set of publications and citations.

To increase the objectivity of the evaluation of scientific content, the concept of "variative citation analysis" was proposed and the possible methodological risks of traditional metrics were identified [3]. Consider the scheme of realization of the risk of negative citing, which arises due to the neglect of the context of citing.

3.2. Variative citation analysis for evaluating negative citing

A variative component of the evaluation of scientific content with taking into account the risk of negative citing involves identifying the context of citing. Within the research of this subject area the following stages are distinguished:

- identification and accounting of negative citing, which makes it possible to establish a marker of the scientist's belonging to the risk group of negative citing;
- indication of citation with negative context to establish the weight of the marker of negative citing;
- verification of negative citing. At this stage, an unambiguous verdict is made on the presence or absence of negative citing of the publication based on the analysis of the reasons for criticism.

Identification and accounting of negative citing involve the transition from the identification of citations as bibliographic records in the list of references to the analysis of the contextual group of physical attributes of citations directly in full-text secondary sources.

Identification and accounting of negative citing involve identifying contextual groups of physical attributes of citations. The methodological tool for this process can be a method of sentiment analysis. It allows in an automated mode using computational linguistics to detect the presence of emotionally coloured and evaluative statements of the author about a particular object.

The fact of detecting a marker of negative citing necessitates the indication and verification of a negative citation for further in-depth study of a particular subject area.

Indication of negative citing implies its parameterization. Negative citing parameterization is the quantification of identified citations using an indicator that correlates with the number of citations with a negative response and compares with the weight of the negative citing marker. For ease of evaluation, this indicator should be a dimensionless quantity (coefficient) that can vary within certain limits. This will allow forming unambiguous judgments about the importance of negative citing.

To indicate citation with a negative response, we propose to use the coefficient of negative citing $K^{(-)}$, which is equal to the ratio of the number of citations with a negative context to the total number of citations:

$$K^{(-)} = \frac{N^{(Cit^-)}}{N^{(Cit)}} = \frac{N^{(Cit^-)}}{N^{(Cit^+)} + N^{(Cit^-)}}, \tag{11}$$

where $N^{(Cit^-)}$ is the number of citations with a negative context; $N^{(Cit^+)}$ is the number of citations with a positive (neutral) context; $N^{(Cit)}$ is the total number of citations to publication.

Depending on the number of identified negative citations, the coefficient of negative citations can vary within $0 < K^{(-)} \leq 1$. The greater is the number of identified negative citations, the greater is the value of $K^{(-)}$. It can be used to indicate a negative citing of a single publication and set of publications.

To determine the significance of the negative citing marker in a particular subject area, the expert must also take into account the social impact of the scientific publication. Under the social influence, we understand the consequences of the practical implementation of scientific research on the living conditions (life quality, life expectancy and health) of a person and society as a whole. For example, it is obvious that from the point of view of social impact, negative citing in the field of medicine is more important than in the humanities. Expert discipline differentiation of the indication of negative citing allows grading the weight of the marker of negative citing to establish the maximum allowable value of the number of citations with a negative response. So, experts should determine the limit values of the coefficient of negative citing $K^{(-)}$ for each subject area. Verification of a negative citing is a complex cognitive process that involves expert assessment of the contextual field of the citation. This process requires a high level of competence, objectivity, professionalism, scientific intuition and adherence to moral and ethical standards.

We believe that the key aspects of expert assessment of negative citing are:

- motivational, within which the expert analyzes the motivation of the negative response;
- discipline, within which the expert determines the importance of the marker of negative citing in a particular subject area.

When analyzing the motivation of a negative response, the expert must take into account the scientific, ideological, gender, moral and ethical basis of its appearance. The presence of citations with negative responses may indicate:

- the presence of significant shortcomings in the scientific research;
- non-compliance with the opponent's principles of academic integrity, professional and moral and ethical standards for their benefit;
- controversial results of scientific research, which at the same time find both their supporters and opponents;
- innovativeness/resonance of research results, which ahead of time causes criticism of the dominant opinion of the scientific community.

To nominate the ahead of time phenomenon, we propose the term "Copernicus Effect". This term was already used to describe the latest scientific thought that causes us to look at the universe differently [20]. The historical aspect of the introduction of this term is the precedent of harsh criticism and prohibition of the spread of the heliocentric theory by the Polish medieval scientist Nicolaus Copernicus by his contemporaries. Copernicus' revolutionary theory became the antithesis of the dominant geocentric theory of the universe among theologians and scholars of the time. The absolute influence of religion and the church in all spheres of public life led to the limitation of the development of science by ideological boundaries. The teachings of Copernicus was not without flaws but became a cornerstone for the formation of astronomy as an exact science. There have been many similar precedents in the history of science. Drawing parallels with modernity, an example of the realization of the Copernican effect can be considered a brilliant Ukrainian cosmologist Yuri Kondratyuk (Olexander Shargei). He is the author of fundamental theoretical works in the field of interplanetary flights, which were significantly ahead of the technical progress of his time and therefore remained unnoticed for about half a century. It was Kondratyuk who in 1914 developed the scheme, later called the Kondratyuk Route, which was used in the first flight to the moon under the Apollo program in 1969. The neoteric scientific idea usually acquires the status of true with difficulty. This is primarily due to the unwillingness of society to accept it through mental, worldview, ideological, technical barriers.

Therefore, the Copernicus Effect is characterized by the postponement of its fixation from the moment of publication of innovative research results to the moment of disappearance of barriers to its perception.

Expert assessment of the motivation of negative citing is heuristic, as it is largely determined by experience, creative intuition, worldview and scientific position of the expert. That is, expert judgment a priori contains a component of subjectivism.

Thus, if the identification and accounting of negative citing is a process of automated content analysis, the indication of negative citing is a process of expert content analysis of secondary sources. Combining the results of expert evaluation and parameterization will reduce the level of subjectivity of the indication of negative citing.

Consider an example of the risk of negative citation in assessing the importance of a scientific publication.

3.3. Scenario analysis of the negative citing scheme

By the basic citation analysis, it was found that the scientific publication, the primary source, was cited $N^{(Cit,PS)}$ times. According to the methodology of traditional metrics, the importance of a publication is determined by its citation. That is, a cumulative set of citations $Cit_{(Cum)}$, which is formed by all references to the publication outside the context of citation. The measure of the publication's citability without taking into account the risk of negative citing is the powerset of the cumulative set of citations

$$\left|Cit_{(Cum)}\right| = N^{(Cit,PS)}. \quad (12)$$

To take into account the risk of negative citation, we move on to the variative component of the evaluation. As a result of the analysis of the tonality of the text of the context fields of citations, $N^{(Cit^{+},PS)}$ citations with a positive or neutral response and $N^{(Cit^{-},PS)}$citations with a negative response to the primary source were identified (Fig. 2).

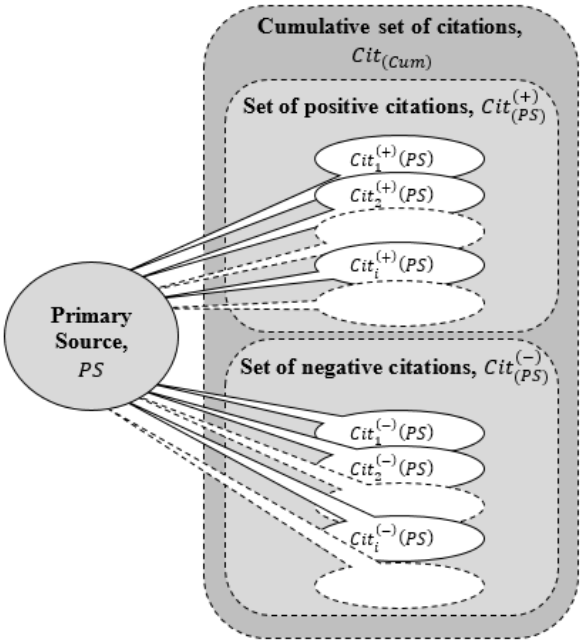

Figure 2: The scheme of realization of negative citing

In this case, the citability of primary source is represented by a set of positive citations

$$Cit^{(+)}_{(PS)} = \left\{Cit^{(+)}_1(PS), Cit^{(+)}_2(PS), \dots, Cit^{(+)}_i(PS), \dots\right\} = \left\{Cit^{(+)}_i(PS)\right\}_{i=1}^{N^{(Cit^{+},PS)}} \quad (13)$$

and a set of negative citations

$$Cit^{(-)}_{(PS)} = \left\{Cit^{(-)}_1(PS), Cit^{(-)}_2(PS), \dots, Cit^{(-)}_i(PS), \dots\right\} = \left\{Cit^{(-)}_i(PS)\right\}_{i=1}^{N^{(Cit^{-},PS)}}, \quad (14)$$

the union of which forms a cumulative set of all citations

$$Cit_{(Cum)} = Cit_{(PS)}^{(+)} \bigcup Cit_{(PS)}^{(-)}, \quad (15)$$

$$\left|Cit_{(Cum)}\right| = N^{(Cit)} = N^{(Cit^{+},PS)} + N^{(Cit^{-},PS)}. \quad (16)$$

The fact of identification of citations with a negative context is defined as a marker of the scientist's belonging to the risk group of negative citing. This necessitates indication using the coefficient of negative citing and expert assessment. In general, the process of implementing a variative citation analysis taking into account the negative citation can be described by the following algorithm (Fig. 3).

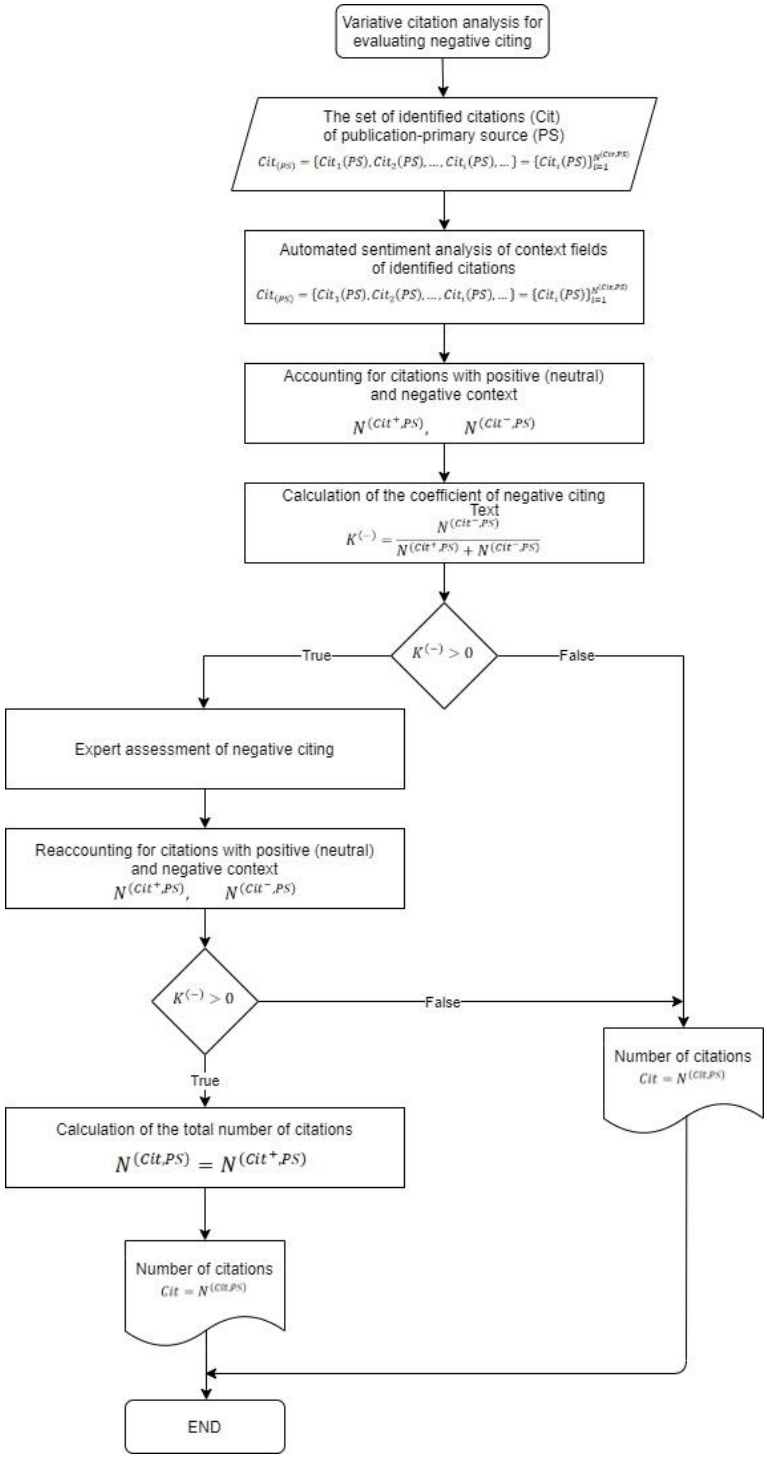

Figure 3: Algorithm for the implementation of variative citation analysis

4. Discussion

The proposed method of variative citation analysis is developed to identify the risk of negative citation when evaluating the quality of scientific content. The use of variative citation analysis will allow:

- to increase the objectivity of evaluating the authority and impact of the researcher and the quality of scientific content at the quantitative and qualitative levels. Identifying a marker of negative citing will make it impossible to "wind up", "overestimate" the quantitative indicators of scientific activity (total number of citations, h-index), which occurs during the summation of all citations without regard to their context. Expert verification of negative citing is proposed to strengthen the automated citation analysis. It will provide a final verdict on the identification of the fact of a negative citing. That is, a variative citation analysis will make it possible to calculate quantitative indicators for evaluating the authority of the scientist only based on citations with a positive (neutral) context;
- to return the scientific discourse in the direction of an objective, critical assessment of scientific achievements. After all, a critical assessment puts the publications of a certain author in the focus of the scientific community. This will lead to a sharp increase in the citability of his works and, accordingly, the growth of his quantitative scientometric indicators and the level of authority. Therefore, researchers are interested in maintaining and increasing their positions, and therefore prefer neutral, uncritical assessments of the results of research by colleagues, opponents. Such an approach can have negative consequences both in the business sphere and in the social sphere of public life. For example, uncritical attitudes toward medical research can reduce the quality of health care. In all scientific and technical fields, this leads to a decrease in the competitiveness and efficiency of scientific developments, in particular those introduced into production. In sum, the lack of free critical discussion on scientific issues will hamper scientific and technological progress.

Based on the results of the study, an algorithm for the implementation of variative citation analysis was developed, which can be implemented in the tools of modern bibliometric systems for the identification and accounting of citations.

The method of variative citation analysis requires further development at the empirical level.

5. Conclusions

A slice of modern bibliometric research indicates the urgent need to develop new alternative metrics or improve existing ones. The methodological basis for quantitative assessment of the authority and demand for scientific content is a basic citation analysis, which involves the identification and accounting of citations of the primary source based on the content analysis of bibliographic references in secondary sources.

The considered basic citation scheme illustrates the idealized process of scientific communication. It does not take into account the contextual field of the citation and does not cover the possible variability of the results of scientific communication.

To increase the objectivity of the evaluation of scientific content, it is necessary to move to a variative level of citation analysis. It involves the identification and consideration of possible methodological risks of traditional metrics.

Implementation of a variative component of scientific content evaluation to take into account the risk of negative citing involves identifying a marker of a scientist's risk group of negative citing, verifying negative citing and analyzing the reasons for criticism, and indicating negative citing to establish the weight of the marker in the certain subject area. Combining the results of expert evaluation and parameterization will reduce the level of subjectivity of the indication of negative citation.

6. References

[1] A. Cabezas-Clavijo, E. Delgado-López-Cózar, Google Scholar e índice h en biomedicina: la popularización de la evaluación bibliométricaGoogle Scholar and the h-index in biomedicine: The popularization of bibliometric assessment, Medicina Intensiva 37(5) (2013) 343–354. doi: ttps://doi.org/10.1016/j.medin.2013.01.008.

[2] A. M. Petersen, R. K. Pan, F. Pammolli, S. Fortunato, Methods to account for citation inflation in research evaluation, Research Policy 48(7) (2019) 1855–1865. doi: https://doi.org/10.1016/j.respol.2019.04.009.

[3] A. Petrushka, Methodological Risks of Evaluating the Quality of Scientific Content, Theoretical and Practical Aspects of the Development of the European Research Area, 3rd ed., Baltija Publishing, Riga, Latvia, 2020, pp. 92–111. doi: https://doi.org/10.30525/978-9934-588-53-2-34.

[4] C. C. Delgado-Carreón, J. D. Machin-Mastromatteo, J. R. Romo-González, J. Pacheco-Mendoza, Creativity-related traits and the scientific production of professors from the autonomous university of Chihuahua, Digital Library Perspectives, ahead-of-print (2021). doi: 10.1108/DLP-08-2020-0077.

[5] C. H. Tam, S. B. Tan, K. C. Soo, The journal impact factor: Too much of an impact?, Annals of the Academy of Medicine Singapore 35(12) (2006) 911–916.

[6] C. Makris, G. Pispirigos, M. A. Simos, Text Semantic Annotation: A Distributed Methodology Based on Community Coherence, Algorithms 13(7) (2020). doi: https://doi.org/10.3390/a13070160.

[7] C.-T. Zhang, The h'-Index, Effectively Improving the h-Index Based on the Citation Distribution, PLoS ONE 8(4) (2013). doi:10.1371/journal.pone.0059912.

[8] D. De Filippo, R. Aleixandre-Benavent, E. Sanz-Casado, Toward a classification of Spanish scholarly journals in social sciences and humanities considering their impact and visibility, Scientometrics 125 (2020) 1709–1732. doi: https://doi.org/10.1007/s11192-020-03665-5.

[9] D. Schoonbaert, G. Roelants, Citation analysis for measuring the value of scientific publications: quality assessment tool or comedy of errors?, Tropical Medicine and International Health 1(6) (1996) 739–752. doi: https://doi.org/10.1111/j.1365-3156.1996.tb00106.x.

[10] G. Vicencio-Ríos, L. Araya-Castillo, C. Millán-Toledo, Y. Rivera Flores, Development of research in brand personality, Revista Venezolana de Gerencia 25(92) (2020) 1583–1599.

[11] H. Keshavarz, M. Esmaeili Givi, A scale for credibility evaluation of scientific websites: findings from a cross-contextual approach, Online Information Review 44(7) (2020) 1369–1386. doi: https://doi.org/10.1108/OIR-04-2020-0127.

[12] J. D. T. Guerrero-Sosa, V. H. Menéndez-Domínguez, M. E. Castellanos-Bolaños, An indexing system for the relevance of academic production and research from digital repositories and metadata, The Electronic Library, ahead-of-print (2021). doi: https://doi.org/10.1108/EL-06-2020-0160.

[13] L. Smolinsky, A. J. Lercher, Co-author Weighting in Bibliometric Methodology and Subfields of a Scientific Discipline, Journal of Data and Information Science 5(3) (2020) 84–96. doi: https://doi.org/10.2478/jdis-2020-0021.

[14] M. D'Agostino, V. Dardanoni, R. G. Ricci, How to standardize (if you must), Scientometrics 113 (2017) 825–843. doi: https://doi.org/10.1007/s11192-017-2495-7.

[15] M. Sedighi, Evaluating the impact of research using the altmetrics approach (case study: the field of scientometrics), Global Knowledge, Memory and Communication 69(4/5) (2020) 241–252. doi: https://doi.org/10.1108/GKMC-02-2019-0013.

[16] N. Kobayashi, T. Toyoda, Statistical search on the Semantic Web, Bioinformatics 24(7) (2008) 1002–1010. doi: https://doi.org/10.1093/bioinformatics/btn054.

[17] Q. Zhou, C. Zhang, Evaluating wider impacts of books via fine-grained mining on citation literatures, Scientometrics 125 (2020) 1923–1948. doi: https://doi.org/10.1007/s11192-020-03676-2.

[18] R. G. Castanha, M. C. Cabrini Grácio, Bibliographic Coupling Indicators for the evaluation of theoretical-methodological proximity in academic genealogy networks a study applied to the PQ

fellowship researchers descending from Aldo Barreto, Revista Digital de Biblioteconomia e Ciencia da Informacao 18 (2020). doi: 10.20396/RDBCI.V18I00.8661393.
[19] S. Maqsood, M. A. Islam, M. T. Afzal, N. Masood, A comprehensive author ranking evaluation of network and bibliographic indices, Malaysian Journal of Library and Information Science 25(1) (2020) 31–45. doi: 10.22452/mjlis.vol25no1.2.
[20] T., Davis, The Copernicus Effect, Susan Davis International (2016). URL: https://www.susandavis.com/the-copernicus-effect/.
[21] V. Davidavičiene, K. Al Majzoub, I. Meidute-Kavaliauskiene, Factors Affecting Knowledge Sharing in Virtual Teams, Sustainability 12(17) (2020). doi: https://doi.org/10.3390/su12176917.
[22] V. S. Lazarev, Discipline Impact Factor: Some of Its History, Some of the Author's Experience of Its Application, the Continuing Reasons for Its Use and... Next beyond, Journal of Data and Information Science 5(3) (2020) 197–209. doi: https://doi.org/10.2478/jdis-2020-0015.
[23] V. S., Lazarev, S.A., Nazarovets, Don't dismiss non- English citation, Nature 556 (2018) 174. doi: https://doi.org/10.1038/d41586-018-04169-2.
[24] Y. Liu, Z. Sun, S-Index: A New Measure to Evaluate the Scientific Impact of Scholars, in: Proceedings of the 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering, AEMCSE 2020, Shenzhen, China, 2020, pp. 529–535. doi: 10.1109/AEMCSE50948.2020.00119.