

ЛІХОНОСОВА ГАННА

Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут»
<https://orcid.org/0000-0001-6552-8920>
e-mail: a.likhonosova@gmail.com

ІНФОРМАТИЗАЦІЯ ТА УПРАВЛІННЯ НАУКОВИМИ ДОСЛІДЖЕННЯМИ: АЛГОРИТМ ПОБУДОВИ РЕЙТИНГУ ПУБЛІКАЦІЙНОЇ АКТИВНОСТІ

У статті актуалізовано проблему доцільності розробки програмної системи, покликаної допомогти керівникам дослідницьких груп, а також індивідуальним дослідникам приймати рішення щодо покращення публікаційної активності, наукометричної результативності, визначати слабкі та сильні сторони, напрямки та конкретні рішення у вигляді автоматизованих рекомендацій щодо підвищення величини та репрезентативності наукометричних показників. Дослідження базується на використанні теорії системного аналізу, яка дозволяє комплексно розглянути всі аспекти, пов'язані з процесами пошуку шляхів між статичними і динамічними аспектами публікаційної активності. Охарактеризовано методіку побудови публікаційного рейтингу на підставі профілів вчених в Scopus. У ході дослідження показано, що в даний час доцільною є розробка програмної системи, покликаної допомогти керівникам дослідницьких груп, а також індивідуальним дослідникам приймати рішення щодо поліпшення публікаційної активності, наукометричної результативності, визначати слабкі та сильні сторони, напрями та конкретні рішення у вигляді автоматизованих рекомендацій щодо підвищення величини та репрезентативності наукометричних показників. Запропонована методика, включаючи алгоритм збору, передобробки та ранжирування наукометричних даних є основою методичного інструментарію аналізу та оцінки публікаційної активності.

Ключові слова: інформатизація, наукові дослідження, рейтинг, алгоритм побудови системи підтримання рішення, індивідуальні наукометричні показники.

LIKHNOSOVA GANNA

National Aerospace University "Kharkiv Aviation Institute"

INFORMATIZATION AND MANAGEMENT OF SCIENTIFIC RESEARCH: PUBLICATION ACTIVITY RATING ALGORITHM

The article updates the problem of the feasibility of developing a software system designed to help heads of research groups, as well as individual researchers, make decisions about improving publication activity, scientometric performance, identify weaknesses and strengths, directions, and specific solutions in the form of automated recommendations for increasing the value and representativeness of scientometric indicators. The research is based on the use of the theory of system analysis, which allows a comprehensive consideration of all aspects related to the processes of finding ways between static and dynamic aspects of publishing activity. The method of building a publication rating based on profiles of scientists in Scopus is characterized. In the course of the study, it was shown that it is currently appropriate to develop a software system designed to help heads of research groups, as well as individual researchers, make decisions about improving publication activity, scientometric performance, identify weaknesses and strengths, directions, and specific solutions in the form of automated recommendations for increasing the value and representativeness of scientometric indicators. The proposed technique, including the algorithm for collecting, refining and ranking scientometric data is the basis of the methodological toolkit for the analysis and assessment of publication activity. The practical significance of the obtained scientific results lies in the proposal of approaches and algorithms, on the basis of which it is possible to develop a software system designed to help heads of research groups, as well as researchers to make decisions on increasing publication activity, scientometric performance, identify weaknesses and strengths, directions and specific solutions in the form of automated recommendations for increasing the value and representativeness of scientometric indicators.

Keywords: informatization, scientific research, rating, decision support system construction algorithm, individual scientometric indicators.

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Інформаційні технології надають можливість будь-якому суб'єкту громадської діяльності поширювати та обробляти інформацію в Інтернеті шляхом створення власної сторінки. Структура наукових видань зазнала безліч змін з впровадженням сучасних інформаційних технологій. Кожне наукове видання апріорі повинно мати власний сайт, з метою популяризації стейкхолдерам власної діяльності.

Незважаючи на те, що немає об'єктивних підстав вважати, що науко-метричні показники, засновані на вимірі публікацій та цитат, здатні безпосередньо відображати рівень і якість наукових досліджень, як таких, багато дослідників відзначають, що розробка наукометричних вимірів та їх зіставлення відкривають нові межі наукової діяльності, дають можливість скоригувати її напрями, структуру та зміст, дати їй оцінку, порівняти її результати з результатами інших досліджень.

Очевидно, що не кожен дослідник має вільний час і компетенції, щоб відповідати на ці питання, при цьому відповідей на них багато в чому залежить його репутація та доступність результатів дослідження для світової наукової спільноти. Особливо дана проблема загострюється на рівні керівництва науковими колективами, лабораторіями та дослідницькими організаціями в цілому, коли керівнику необхідно приймати управлінські рішення щодо фінансування тих чи інших досліджень, забезпечення кадровими та іншими ресурсами. Керівник потребує аналітичної інформації щодо перспективності тих чи інших досліджень, дослідницького та публікаційного потенціалу тих чи інших дослідників.

Деяку інформацію для керівника надають міжнародні реферативні та бібліографічні бази даних та їх аналітичні програми. Індивідуальні наукометричні показники, що розраховуються в них, відіграють ключову роль при вивченні динаміки наукових досліджень, побудові міжнародних рейтингів дослідницьких та освітніх організацій, проходженні атестації дослідниками та прийнятті рішення про виділення дослідницьких грантів.

Аналіз останніх досліджень і публікацій

У галузі наукометрії та інфометрії загалом напрацьовано великий базис методів для дослідження динаміки наукових досліджень та публікаційної активності, зокрема значний внесок сформовано В.В. Налімовим, З.М. Мульченко [1], L. Blackert, S. Siegel [2], O. Nacke [3] та ін. Відомі у міжнародній практиці аналітичні методи та показники, що найбільш повно відповідають задачі індивідуальної наукометричної оцінки автора (J. Hirsch), статті та журналу (E. Garfield) тощо. Удосконалюється аналітичний інструментарій (InCites, SciVal etc.) у програмних системах міжнародних реферативно-бібліографічних баз даних Scopus, Web of Science.

Також розвитком технологій в області поширення інформатизації наукових досліджень займаються такі автори: Спирін О. М., Яцишин А. В. [4], Іванов С. М., Кільченко А. В. [5], Урядові та неурядові компанії й організації, що займаються цифровізацією результатів наукових досліджень є Springer, E-SCIENCE, Scopus, Web of Science, Periodikos, Національна бібліотека ім. Вернадського ті багато інших. Проте сучасні тенденції біфуркаційного розвитку соціально-економічної системи актуалізують необхідність аналізу та узагальнення наукових напрацювань, що присвячені досліджуваній темі, що й зумовило мету цієї роботи.

Виділення невирішених раніше частин загальної проблеми, котрим присвячується стаття.

Незважаючи на значну кількість досліджень у галузі інформатизації наукових досліджень та теорії рейтингування, вагомий внесок учених у дослідження видів та ефективності алгоритмів рейтингування публікаційної активності, сьогодні з'являються нові проблеми у сфері управління науковими дослідженнями та автоматизації прийняття рішення щодо побудови міжнародних рейтингів дослідницьких та освітніх організацій, проходженні атестації дослідниками та прийнятті рішення про виділення дослідницьких грантів.

Формулювання цілей статті

Метою статті є застосування теорії інформатизації та управління науковими дослідженнями до розробки алгоритм побудови рейтингу публікаційної активності як елементу система підтримки прийняття рішення в управленні науковими дослідженнями на підставі аналізу наукометричних даних.

Виклад основного матеріалу.

На сучасному етапі розвитку інформаційних технологій важливу роль набувають інтелектуальні рекомендаційні системи, як програми, які намагаються передбачити, які об'єкти (фільми, музика, книги, публікації, веб-сайти) будуть цікавими для користувача, маючи певну інформацію про його профіль. Зрозуміло, і наукометричний профіль дослідника чи дослідницької організації загалом у тому чи іншого наукометричної бази даних цілком відповідає необхідним умовам розробки рекомендацій.

Отже, в даний час доцільна розробка програмної системи, покликаної допомогти керівникам дослідницьких груп, а також індивідуальним дослідникам приймати рішення щодо покращення публікаційної активності, наукометричної результативності, визначати слабкі та сильні сторони, напрямки та конкретні рішення у вигляді автоматизованих рекомендацій щодо підвищення величини та репрезентативності наукометричних показників.

Найважче завдання аналізу публікаційної активності та оцінки індивідуального вкладу з метою управління науковою діяльністю в цілому вимагає врахування базових засад у рамках науково-технічної політики наукової організації чи інституту.

Насамперед, слід забезпечити диференційованість оцінки. Вочевидь, що з вирішення цього завдання у межах різних наукових наук напрямів чи афілійованих організацій досить робити вибірку (наприклад, за ключовими словами, приналежності до організації та інших.).

Важливим елементом диференційованого підходу, з погляду, є облік тривалості публікаційної активності. Справді, зіставлення наукометричних показників дослідника, працюючого у своїй галузі протягом багатьох років і дослідника, який нещодавно захистив дисертацію, є скрутним. Облік цього зауваження зазвичай виявляється у введенні поправочного коефіцієнта пропорційного тривалості публікаційної активності або віку дослідника. Вважаємо достатнім зважування за допомогою розподілу на тривалість публікаційної активності. У разі може вирішуватися й інше важливе завдання – стимулювання молодих вчених. Для наукового колективу, у разі потреби, як оцінку тривалості публікаційної активності слід приймати максимальну індивідуальну тривалість.

Важливим завданням є розробка підходу щодо забезпечення консервативності оцінки. У ряді робіт у зв'язку з цим пропонується використовувати функції з спадною похідною. Забезпечення спадного вкладу обумовлено необхідністю обмежити спекуляцію наукометричними показниками та врахувати деяку специфіку їх динаміки (наприклад, кількість цитувань статті лінійно залежить від числа співавторів [6], а динаміка індексу Хірша апроксимується сигмоїдальною функцією [7]).

Зокрема, вибір подібної функції (1) дозволяє зробити облік цитування більш консервативним, а вклад кожної наступної цитати помітно менший за вклад попередньої. Іншим варіантом є використання логарифмічного перетворення, при цьому логарифмування наукометричних даних вирішує задачу наближення розподілу до нормального.

$$\sqrt[4]{\frac{N^2}{n}}, \quad (1)$$

де N – число цитат, n – число публікацій.

Принципи комплексності та обґрунтованості оцінки можуть бути реалізовані використанням повних наборів даних за показниками, наукометричних баз - Web of Science, Scopus та «Наука України». Залежно від стратегічних завдань із стимулювання реферування у тій чи іншій базі даних, наборам показників з метою оцінки необхідно привласнити відповідні ваги (у разі ранжирування – пріоритетні ранги).

Важливим критерієм вибору тих чи інших наборів показників є можливість їхнього автоматизованого вилучення з доступних джерел та інтеграція наукометричної оцінки в інформаційну систему організації (університету).

Спираючись на описані вище підстави, пропонуємо оцінку публікаційної активності на основі формування агрегованого рейтингу. Основні етапи методики полягають у наступному (на прикладі Scopus).

1. Відбір наукометричних показників, що у побудові рейтингу, з принципів оцінки та управлінських завдань.

2. Формування масиву та вивантаження показників, загальнодоступних для перевірки у системі Scopus, у розрізі областей дослідження.

3. Коригування значень показників (зважування) (2):

$$x'_i = \frac{x_i}{y_i + 1}, \quad (2)$$

де x_i - вихідний показник по i -му автору; y_i – коригування по i -му автору (наприклад, кількість років від першої публікації).

Коригування дозволяє отримати питомі, порівняні за авторами, величини, а також дозволяє формувати певні стимули публікаційної активності, зокрема: віднесення кількості публікацій та індексу Хірша до років від першої публікації з одного боку зумовлює необхідність підтримки публікаційної активності на належному рівні, а з іншого – дозволяє стимулювати молодших авторів; віднесення загального числа цитат до публікацій ставить пріоритет на публікації малої кількості великих цитованих робіт за порівнянням з численними маловпливовими роботами (попереджає дроблення наукових результатів на кілька публікацій); віднесення числа публікацій до співавторів з одного боку орієнтує на залучення продуктивних співавторів та участь у значущих роботах, а з іншого – попереджає номінальне співавторство.

4. Перетворення значень показників у ранги в результаті відображення на рангову множину (сортування за спаданням):

$\langle x'_i \rangle \rightarrow \text{sort} \downarrow \langle 1, 2, \dots, N \rangle$, ранг i -го автора визначено, як:

$$r_i = \text{sort} \downarrow (x'_i), r_i \in 1, 2, \dots, N \quad (3)$$

5. Ранжування показників за значимістю для оцінки індивідуальної публікаційної активності, потім граничне агрегування [8], яка, як раніше зазначено, забезпечує некомпенсаторність підсумкової оцінки.

На даному етапі слід розділити наукометричні показники на дві групи: до першої ставляться показники, які априорі можна ранжувати по значимості (наприклад, входження в top-10 % більш значиме, аніж в top-25 %), а до другої ставляться показники умовно рівної значимості. Загальне рішення може бути також знайдене на основі аналізу експертних оцінок.

Для цього має бути проведене анкетування експертів, у число яких увійшли керівники науковою працею різного рівня, що дозволить визначити погоджену порівняльну значимість наукометричних показників, доступних в Scopus.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

У ході дослідження показано, що в даний час доцільною є розробка програмної системи, покликаної допомогти керівникам дослідницьких груп, а також індивідуальним дослідникам приймати рішення щодо поліпшення публікаційної активності, наукометричної результативності, визначати слабкі та сильні сторони, напрями та конкретні рішення у вигляді автоматизованих рекомендацій щодо підвищення величини та репрезентативності наукометричних показників. Запропонована методика, включаючи алгоритм збору, передобробки й ранжирування наукометричних даних являє собою основу методичного інструментарію аналізу й оцінки публікаційної активності.

Література

1. Nalimov V.V., Mulchenko Z.M. *Scientometrics*. Studying the development of science as an information process. Riga: Nauka. 1969, 192 p.
2. Blackert L. Siegel S. *Ist in ger wissenschaftlichen information Platz fur die Informetrie?* Wiss. Zeitschrift Tech. Hochschule. 1979. Bd 25, Hf 6. P. 187-199.
3. Nacke O. *Informetria*: Name fur eine neue Disziplin. Nach. Dokument. 1979. Bd 30, Hf 6. P. 219-226.
4. Спірін О. М., Яцишин А. В., Іванова С. М., Кільченко А. В., Лупаренко Л. А. Модель інформаційно-аналітичної підтримки педагогічних досліджень на основі електронних систем відкритого доступу.

Інформаційні технології і засоби навчання. 2017. Т. 59, вип. 3. С. 134-154.

5. Спирін О. М., Яцишин А. В., Іванова С. М., Кільченко А. В., Лупаренко Л. А. Використання електронних систем відкритого доступу для інформаційно-аналітичної підтримки педагогічних досліджень. *Інформаційні технології і засоби навчання*. 2016. Т. 55, вип. 5. С. 136-174.

6. Vieira E.S, Gomes J.A.N.F. Citations to Scientific Articles: Its Distribution and Dependence on the Article Features. *Journal of Informetrics*. 2010. Vol. 4. № 1. P. 1-13.

7. Тарасевич Ю.Ю., Шиняєва Т.С. Часова динаміка індексу Хірша. Математичне моделювання та програмування. 2016. № 1 С. 32-45.

8. Ван Вухт Ф., Вестерхайден Д. Багатомірне ранжування: новий інструмент прозорості у галузі вищої освіти. *Вісник міжнародних організацій: освіта, наука, нова економіка*. 2012. №1. С. 9-33.

References

1. Nalimov V.V., Mulchenko Z.M. *Scientometrics. Studying the development of science as an information process*. Riga: Nauka. 1969, 192 p.

2. Blackert L. Siegel S. Ist in ger wissenschaftlichen information Platz fur die Informetrie? *Wiss. Zeitschrift Tech. Hochschule*. 1979. Bd 25, Hf 6. P. 187-199.

3. Nacke O. Informetria: Name fur eine neue Diszipline. *Nach. Dokument*. 1979. Bd 30, Hf 6. P. 219-226.

4. Spirin O. M., Yatsyshyn A. V., Ivanova S. M., Kilchenko A. V., Luparenko L. A. Model informatsiino-analitychnoi pidtrymky pedahohichnykh doslidzhen na osnovi elektronnykh system vidkrytoho dostupu. *Informatsiini tekhnolohii i zasoby navchannia*. 2017. Т. 59, вып. 3. С. 134-154.

5. Spirin O. M., Yatsyshyn A. V., Ivanova S. M., Kilchenko A. V., Luparenko L. A. Vykorystannia elektronnykh system vidkrytoho dostupu dlia informatsiino-analitychnoi pidtrymky pedahohichnykh doslidzhen. *Informatsiini tekhnolohii i zasoby navchannia*. 2016. Т. 55, вып. 5. С. 136-174.

6. Vieira E.S, Gomes J.A.N.F. Citations to Scientific Articles: Its Distribution and Dependence on the Article Features. *Journal of Informetrics*. 2010. Vol. 4. № 1. P. 1-13.

7. Tarasevych Yu.Iu., Shyniaieva T.S. Chasova dynamika indeksu Khirsha. *Matematychni modeliuvannia ta prohramuvannia*. 2016. № 1 С. 32-45.

8. Van Vukht F., Vesterkhaiden D. Bahatomirne ranzhuvannia: novyi instrument prozorosti u haluzi vyshchoi osvity. *Visnyk mizhnarodnykh orhanizatsii: osvita, nauka, nova ekonomika*. 2012. №1. С. 9-33.