

**Миронов Олексій Вадимович**

Аспірант кафедри інформаційних технологій, ORCID: 0000-0001-7119-2415

Київський національний університет будівництва і архітектури, Київ

**Білощицький Андрій Олександрович**

Доктор технічних наук, завідувач кафедри інформаційних технологій, ORCID: 0000-0001-9548-1959

Київський національний університет будівництва і архітектури, Київ

**РОЗРОБКА МАТЕМАТИЧНОЇ МОДЕЛІ ПРЕДСТАВЛЕННЯ,  
УПРАВЛІННЯ ТА ОЦІНКИ НАУКОМЕТРИЧНИХ СУБ'ЄКТІВ**

***Анотація.** Однією зі складових оцінок діяльності ВНЗ у всьому світі є визначення узагальненого показника якості та результатів наукових досліджень окремого вченого, кафедри університету і вищих навчальних закладів в цілому. У сучасному світі інформаційних технологій саме множина публікацій, які доступні у web-просторі, дозволяє оцінити науковий рівень досліджень. Проте, відсутність єдиних вимог та стандартів до розміщення та управління науковими працями створює реальні перешкоди на шляху до якісної оцінки результатів діяльності наукометричних суб'єктів. Вирішення даної проблеми потребує: по-перше, визначення основних сутностей суб'єктів наукометрії та зв'язків між ними; по-друге, створення належного ступеня формалізації процесів управління науковими публікаціями на різних етапах їх обробки; по-третє, кінцевим етапом розробки моделі є створення глобальної бази даних вчених, наукових публікацій, наукових журналів, закладів для визначення рейтингу цитованості та популярності вищезгаданих суб'єктів. В роботі було розроблено модель представлення науково-публіцистичного контенту в мережі Інтернет, на основі якої запропоновано систему стандартизації вимог до науково-публіцистичних робіт. Окрім того, проведено аналіз та математичний опис моделі механізму оцінки наукової діяльності наукометричних суб'єктів, процесу їх оцінювання та середовища, в якому зазначений процес відбувається.*

**Ключові слова:** : наукометричні суб'єкти; публікації; цитування; оцінка рейтингу

**Вступ**

Відомо, що підґрунтям для створення наукових праць є ретельний аналіз наявних проєктів. Однак, якщо донедавна основним інструментом пошуку необхідних матеріалів були бібліотеки, то в умовах стрімкої глобальної інформатизації предметно-практичної діяльності людини та суспільства в цілому, одним з основних факторів, що каталізує розвиток сучасної науки, стає пошук відповідних напрацювань у web-просторі [1].

Сучасний розвиток інформаційних Інтернет-технологій в області організації сховищ даних, репозитаріїв та електронних бібліотек для зберігання наукових публікацій і надання (у разі потреби) доступу до баз даних наукових публікацій породжує нові можливості і задачі в частині супроводження інформаційних процесів, а також узагальненої оцінки якості і результатів наукової діяльності [2].

**Мета статті**

Мета – визначення основних сутностей суб'єктів наукометрії та зв'язків між ними. Створення належного ступеня формалізації процесів управління науковими публікаціями на різних етапах їх обробки. Кінцевим етапом розробки моделі є створення глобальної бази даних вчених, наукових публікацій, наукових журналів, закладів для визначення рейтингу цитованості та популярності наукометричних суб'єктів. Задача даної роботи – розробка моделі представлення науково-публіцистичного контенту в мережі інтернет, а також аналіз та математичний опис моделі механізму оцінки наукової діяльності наукометричних суб'єктів, процесу їх оцінювання та середовища, в якому зазначений процес відбувається.

**Виклад основного матеріалу****Представлення наукового контенту**

Початок наукометрії як самостійної дисципліни поклав американський вчений Юджин Гарфілд.

У 1955 р. він створив інститут наукової інформації, в 1961 р. запровадив індекс наукового цитування та оцінки наукових журналів. Згодом, 1995 – рік заснування платформи *ISI Web of Knowledge* (нині має назву: *Web of Science*). У 2005 р. заснована нині популярна наукометрична база даних *Scopus*, а в 2010 р. у *Scopus* з'являються нові журнальні метрики *SJR* і *SNIP*.

Розробка будь-якої моделі передбачає виконання цілого ряду вимог, які задовольнятимуть такі критерії оцінки даних [1; 3]:

1. Структурна достовірність, яка полягає у відповідності способу визначення та організації інформації;

2. Простота, що передбачає зручність вивчення моделі як професіоналами в області розробки інформаційних систем, так і звичайними користувачами;

3. Відсутність надмірності, яка полягає у вилученні надмірної інформації, тобто будь-яку частину даних можна представити лише один раз;

4. Сумісне використання – це відсутність приналежності до особливої прикладної програми або технології та відповідно можливість використання у багатьох прикладних програмах і технологіях;

5. Розширюваність – це властивість розвиватись та вмішувати нові вимоги з мінімальним впливом на роботу наявних складових;

6. Схематичне представлення – це можливість представлення моделі за допомогою наочних схематичних позначень.

Для побудови моделі представлення наукового контенту спершу визначимо основні сутності та зв'язки наукометричних суб'єктів. Подамо концептуальну схему представлення контенту наукових публікацій у вигляді повного графу. Ядром даної моделі наукового інформаційного простору є п'ять базових сутностей (рис. 1):



Рисунок 1 – Базові сутності концептуальної моделі наукових даних

Розглянемо математичну модель представлення наукових публікацій в мережі інтернет як систему

з впорядкованим набором елементів, і подамо її у вигляді кортежу [4]:

$$N = \langle U, S, R, M, f_{u,s} \rangle,$$

де  $U$  – множина користувачів системи (автори, редактори і т.д.);  $S$  – множина науково-публіцистичних об'єктів системи;  $R$  – заявки користувачів системи (вхідний параметр);  $M$  – повідомлення користувачам системи (вихідний параметр);  $f_{u,s}$  – функція підрахунку кількості зовнішніх посилань, отриманих публікацією  $S$  від користувачів:  $U' (U' \in U)$ .

Множину науково-публіцистичних об'єктів системи  $S$  визначають наявні в ній наукові публікації, кожна з яких  $S_i$  детермінована сукупністю атрибутів та метаданих:

$$S_i = \{v, f, m, D\},$$

де  $v$  – фізичний об'єм наукової публікації (МБ);  $f$  – формат публікації (\*.rtf, \*.pdf, і т.д.);  $m$  – додаткові супровідні матеріали до публікацій (графіки, діаграми і т.д.);  $D$  – множина метаданих наукової публікації.

Заявкою  $R$  вважатимемо запит від користувача  $U$  на публікацію наукової статті  $S_i$ , і визначимо його як кортеж:

$$R = \langle U^r, S_i^r \rangle,$$

де  $U^r$  – користувач, який є джерелом запиту;  $S_i^r$  – наукова стаття, що подана користувачем на розгляд.

Повідомленням вважатимемо сукупність рекомендацій щодо виправлення/редагування або погодження надісланої публікації, які визначаються як набір:

$$M = \langle R, T \rangle,$$

де  $T$  – сукупність повідомлень та матеріалів, рекомендацій до виправлення, або погодження наукової публікації.

Припустимо, що деякий журнал  $J$  складається з  $n$  науково-публіцистичних об'єктів  $O$  (випусків):

$$J = \{O_1, O_2, \dots, O_i, \dots, O_n\}, i = \overline{1, n}.$$

Кожному об'єкту  $O$  шляхом деякої композиції можна зіставити  $m_o$  відображень (наукових публікацій), кожне з яких є науково-публіцистичним об'єктом  $S$ , тоді отримуємо:

$$\gamma_o : S \rightarrow \{S_1^o, S_2^o, \dots, S_i^o, \dots, S_{m_o}^o\}, i = \overline{1, m_o}.$$

Окрім того, необхідно забезпечити виконання наступних умов для коректного формату представлення публікацій:

$$\forall S_i \in O \subset J, \left\{ \begin{array}{l} v_{s_i} \leq v_{\max}, v_{s_i} \in S_i^o \\ p_{i \min} \leq p_i \leq p_{i \max}, p_i \in D \subset S_i^o \end{array} \right\},$$

де  $p_i$  – атрибут метаданих статті, який відображає кількість внутрішніх посилань у публікації.

У загальному вигляді весь ланцюг процесів управління науковими публікаціями в інформаційній системі (від подання до опублікування) можна зобразити у вигляді каскадної моделі (рис. 2).



Рисунок 2 – Каскадна модель процесів управління науковими публікаціями в інформаційній системі

Останнім етапом опублікування наукового подання є його затвердження, яке має місце тоді, коли публікація пройшла корегування та редактування. Даний етап включає в себе розміщення надісланих статей  $S$  у відповідних випусках  $O$ , і, аналогічно, розміщення опублікованих випусків, що вже містять в собі необхідні публікації, у журнал  $J$  (рис. 3):

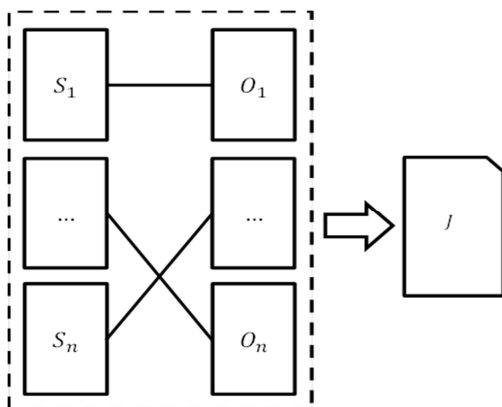


Рисунок 3 – Представлення науково-публіцистичних об'єктів

### Оцінка наукової діяльності

Щодо механізму роботи функції оцінки діяльності наукометричних суб'єктів  $f_{u,s}$ , то він виглядає дещо складніше ніж просто підрахунок кількості зовнішніх посилань. Це пов'язано з тим, що до розрахунку залучаються не зовнішні посилання як такі, а співвідношення загального об'єму (ваги) конкретної цитати, запозиченої з відповідного джерела, до загального об'єму тексту наукової публікації. Формально, ресурсами оцінки якості і результатів наукової діяльності наукометричних суб'єктів  $E$  виступають самі ж суб'єкти:

$$E = \langle S, J, I \rangle,$$

де  $S$  – статті, що оцінюються;  $J$  – журнали, що оцінюються;  $I$  – заклад, що оцінюється.

Оскільки маємо справу з базою даних наукових публікацій, доцільно було б використовувати функцію оцінки всередині відповідного тригера. Розглянемо БД компонентів наукового контенту  $DB$  як об'єднання множин даних, метаданих і документів:

$$DB = \bigcup_{i=1}^n d_i(C_i),$$

де  $d_i(C_i)$  – множина даних, метаданих і документів науково-публіцистичного компонента  $C_i$ . Науково-публіцистичний компонент відображає поточний рівень ієрархії наукометричного суб'єкта (статті, журнали, заклади):

$$f_{u,s} : E \rightarrow C_i,$$

$$C_i = \{(S_1, S_2, \dots, S_l) \vee (J_1, J_2, \dots, J_m) \vee (I_1, I_2, \dots, I_n)\}.$$

Так, функція оцінки наукової діяльності, наприклад журналу, матиме вигляд:

$$f_{u,s} = \frac{\sum_{i=1}^m \left( \prod_{j=1}^l N_w^{S_j w_k^{S_j}} \right)^{\frac{1}{\sum_{j=1}^l w_k^{S_j}}}}{V_{S_j} * f(V_{S_j})} = \frac{\sum_{i=1}^m \exp \left( \frac{1}{\sum_{j=1}^l w_k^{S_j}} \sum_{j=1}^l w_k^{S_j} \ln(N_w^{S_j}) \right)}{V_{S_j} * f(V_{S_j})},$$

де  $w_k^{S_j}$  – вага цитати  $k$ -го джерела статті  $S_j$ ;  $N_w^{S_j}$  – поточне (на визначеній ітерації) внутрішнє посилання в статті  $S_j$ ;  $V_{S_j}$  – об'єм статті  $S_j$ ;

$f(V_{S_j})$  – функція нормалізації, яка базується на об'ємі статті, що оцінюється.

Виходячи із зазначеного, математична модель БД, що є основою для оцінки наукової діяльності суб'єктів набуде вигляду:

$$DB = \left\langle \bigcup_{i=1}^{n_S} d_i^S(C_i), \bigcup_{i=1}^{n_J} d_i^J(C_i), \bigcup_{i=1}^{n_I} d_i^I(C_i) \right\rangle.$$

Якщо розглядається БД, то мова йде про динамічну систему, тобто систему, яка змінюється у часі. Змінним у даній системі є кількість наукових публікацій, тому модифікуємо модель таким чином:

$$DB = \left\langle \bigcup_{t_0}^{t_1} \left( \bigcup_{i=1}^{n_S} d_i^S(C_i) \right)_{t_1}, \left( \bigcup_{i=1}^{n_J} d_i^J(C_i) \right)_{t_1}, \left( \bigcup_{i=1}^{n_I} d_i^I(C_i) \right)_{t_1} \right\rangle,$$

$$\left\{ \begin{array}{l} t \in [t_0; t_1]; \\ J = \text{const}; \\ I = \text{const}; \end{array} \right\},$$

де  $t_0$  – момент впровадження БД в експлуатацію;  $t_1$  – поточний момент часу.

Тоді, запит  $Z$  до БД для оцінки, наприклад наукових публікацій, матиме вигляд:

$$Z(DB) = \pi \left\{ \sigma \left\{ \begin{array}{l} v_{s_i} \leq v_{\max}, v_{s_i} \in S_1^o \\ p_{i \min} \leq p_i \leq p_{i \max}, p_i \in D \subset S_i^o \end{array} \right\} (S_1 \times \dots \times S_n) \right\},$$

де  $\pi$  – оператор селекції для отримання необхідної перестановки, проєкції або значення атрибутів результуючого набору даних;

$S$   $\left\{ \begin{array}{l} v_{s_i} \in J, v_{\max}, v_{s_i} \in OS_i^o \\ p_{i \min} \in J, p_i \in J, p_{i \max}, p_i \in ODMS_i^o \end{array} \right\}$  – умова вибору

результуючого набору даних;  $S_j$  – наукова публікація, до якої сформовано запит.

## Висновки

Висвітлені фундаментальні концепції представлення наукових публікацій у web-просторі, запропонована на їх основі модель процесів управління на кожному з етапів, починаючи з подання і закінчуючи публікацією, надають підстави для виділення основних відношень між атрибутами наукових праць, і, власне, елементами наукометричної інформаційної системи.

Введені терміни та поняття сприяють формуванню науково-методологічної концепції представлення наукових публікацій. Розроблена модель представлення науково-публіцистичного контенту в мережі інтернет включає в себе як інформаційну складову, так і математичне обґрунтування. Саме на її основі запропонована система стандартизації вимог до науково-публіцистичних робіт. Окрім того, в роботі наведено науково-обґрунтований математичний опис моделі механізму оцінки наукової діяльності наукометричних суб'єктів, процесу їх оцінювання та середовища, в якому зазначений процес відбувається.

Наведена модель є поштовхом до подальших науково-дослідницьких робіт у сфері оптимізації моделі контенту наукових праць.

## Список літератури

1. Миронов О.В. Управління інформаційною системою розміщення наукових публікацій в мережі інтернет [Текст] // XII міжнародна наукова конференція «Управління проектами у розвитку суспільства». 2015. – 179 с.
2. Бурков, В.Н., Белоцицкий, А.А., Гогунский, В.Д. Параметры цитируемости научных публикаций в наукометрических базах данных [Текст] // Управління розвитком складних систем. – 2013. – № 15. – 134 с.
3. Філіпова Л.Я., Шелестова А.М. Інформаційне моделювання контенту університетського веб-сайту: загальні принципи та методика [Текст] // Вісник книжкової палати. – № 3. – 2014. – 3 с.
4. Миронов О.В. Розробка моделі представлення контенту наукових публікацій у web-просторі [Текст] // II міжнародна науково-практична конференція «Управління розвитком технологій». – 2015. – 66 с.
5. Бушуев, С.Д., Білоцицкий, А.О., Гогунський, В.Д. Наукометричні бази: характеристика, можливості і завдання [Текст] // Управління розвитком складних систем. – 2014. – № 18. – 145 с.
6. Шеннон К.Е. Работы по теории информации и кибернетике / К.Е. Шеннон: под ред. Р.Л. Добрушина, О.Б. Лупанова. – М.: Иностранная литература, 1963. – 832 с.
7. Керносов М.А. Математическая модель контента информационно-аналитической системы [Текст] // Бионика интеллекта. – 2013. – №1(80). – С. 54-61.
8. Codd E.F. Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate / E.F. Codd, S.B. Codd, C.T. Salley. – Hyperion Solutions Corporation. – 1993. – 24 p.
9. Белоцицкий А.А. Управление проблемами в методологии проектно-векторного управления образовательными средами [Текст] / А. А. Белоцицкий // Управління розвитком складних систем. – 2012. - № 9. – С. 104 – 107.

10. Оборський Г.О. Стандартизація і сертифікація процесів управління якістю освіти у вищому навчальному закладі [Текст] / Г.О. Оборський, В.Д. Гогунський, О.С. Савельєва // Тр. Одес. политехн. ун-та. – Вып. 1(35). – 2011. – С. 251 – 255.

Стаття надійшла до редколегії 18.06.2015

Рецензент: д-р техн. наук, проф. С.Д. Бушуєв, Київський національний університет будівництва і архітектури, Київ.

**Мионов Алексей Вадимович**

Аспирант кафедры информационных технологий, ORCID: 0000-0001-7119-2415

Киевский национальный университет строительства и архитектуры, Киев

**Белошицкий Андрей Александрович**

Доктор технических наук, профессор кафедры информационных технологий, ORCID: 0000-0001-9548-1959

Киевский национальный университет строительства и архитектуры, Киев

**РАЗРАБОТКА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПРЕДСТАВЛЕНИЯ,  
УПРАВЛЕНИЯ И ОЦЕНКИ НАУКОМЕТРИЧЕСКИХ СУБЪЕКТОВ**

*Аннотация.* Одной из составляющих оценки деятельности вузов во всем мире является определение обобщенного показателя качества и результатов научных исследований отдельного ученого, кафедры университета и высших учебных заведений в целом. В современном мире информационных технологий именно множество публикаций, которые доступны в web-пространстве, позволяет оценить научный уровень исследований. Однако, отсутствие единых требований и стандартов к размещению и управлению научными трудами создает реальные препятствия на пути к качественной оценке результатов деятельности наукометрических субъектов. Решение данной проблемы требует во-первых: определение основных сущностей субъектов наукометрии и связей между ними, во-вторых: создание надлежащей степени формализации процессов управления научными публикациями на разных этапах их обработки, в-третьих: конечным этапом разработки модели является создание глобальной базы данных ученых, научных публикаций, научных журналов, заведений для определения рейтинга цитируемости и популярности вышеупомянутых субъектов. Разработана модель представления научно-публицистического контента в сети интернет, на основе которой предложена система стандартизации требований к научно-публицистическим работам. Кроме того, проведен анализ и математическое описание модели механизма оценки научной деятельности наукометрических субъектов, процесса их оценки и среды, в которой указанный процесс происходит.

**Ключевые слова:** наукометрические субъекты; публикации; цитирования; оценка рейтинга

**Myronov Oleksii**

Postgraduate student of information technology, ORCID: 0000-0001-7119-2415

Kyiv National University of Construction and Architecture, Kyiv

**Biloshchytskyi Andrii**

Doctor of Technical Sciences, Head of information technology, ORCID: 0000-0001-9548-1959

Kyiv National University of Construction and Architecture, Kiev

**DEVELOPMENT OF MATHEMATICAL MODEL OF SCIENTOMETRIC SUBJECTS REPRESENTATION,  
MANAGEMENT AND EVALUATION**

*Abstract.* One of the universities' rating components in the world is the definition of a generalized indicator of quality and results of individual research scientists, university departments and higher education institutions as a whole. In today's world of information technology there is a set of publications that are available in the web-space, to evaluate the level of scientific research. However, the lack of uniform standards and requirements for placement and management of scientific work creates real obstacles to the qualitative assessment of scientometric subjects' performance. Solving this problem requires firstly, identifying key actors of scientometric entities and relations between them, secondly, creating a proper degree of formalization of scientific publications management in various stages of processing, thirdly, the final step is to create global database of scientists, scientific publications, scientific journals, institutions for citation ranking and popularity of the aforementioned entities. The paper was developed a model of presenting scientific and journalistic content on the Internet, which includes the system of requirements standardization for scientific and publication work. In addition, the analysis and mathematical description of the scientometric evaluation mechanism model of scientific subjects, the process of assessment and the environment in which this process occurs.

**Keywords:** Scientometric subjects; publications; citation; rating assessment

**References**

1. Myronov, Oleksii (2015). *Information system management of allocation of scientific publications on the Internet. XII International Conference "Project Management in the development of society"*, 179.
2. Burkov, Vladimir, Biloshchytskyi, Andrii, & Gogunsky, Victor (2013). *Parameters citation of scientific publications in scientometric databases. Management of Development of Complex Systems*, (15), 134-139.
3. Philipova, L.Ya., & Shelestova, A.M. (2014). *Information Modeling of University' website content: General Principles and Methods. Bulletin of the Book Chamber*. (3), 3.
4. Myronov Oleksii (2015). *Developing a representation model of the content of scientific publications in the web-space. II International scientific-practical conference "Management of technologies development"*, 66.
5. Bushuyev Sergii, Biloshchytskyi Andrii, & Gogunsky Victor (2014). *Scien tometric database: characteristics, opportunities and challenges. Management of Development of Complex Systems*, (18), 145.
6. Shannon, K.E. (1963). *Work on the theory of information and cybernetics. Moscow: Foreign Literature [in Russian]*.
7. Kernosov, M.A. (2013). *Mathematical model of the content of information-analytical system. Intelligence Bionics*, 1(80), 54-61.
8. Codd, E.F., (1993). *Providing OLAP (On-Line Analytical Processing) to User-Analysts: IT Mandate Hyperion Solutions Corporation [in English]*.
9. Beloshchitskii, Andrei (2012). *Management problems in the methodology of design and vector control education environments. Management of Development of Complex Systems*, (9), 104–107.
10. Oborskii, G.O. (2011). *Standardization and certification processes of education quality management in higher education institution. Odessa polytechnic institute*, 1(35), 251–255.

**Посилання на публікацію**

- APA Myronov O. & Biloshchytskyi, A., (2015). *Development of mathematical model of scientometric subjects representation, management and evaluation. Management of Development of Complex Systems*, 23 (1), P. 147 – 152.
- ГОСТ Миронов, О.В. Розробка математичної моделі представлення, управління та оцінки наукометричних суб'єктів [Текст] / О.В. Миронов, А.О. Білощичський // *Управління розвитком складних систем*. – 2015. - № 23(1). – С. 147- 152.