

УДК 681.3.06

ЗАСОБИ ПІДВИЩЕННЯ INTERNET СТАТУСУ ДРУКОВАНИХ ЖУРНАЛІВ

© О. П. Цурін, к.т.н., с.н.с., НТУУ «КПІ», Київ, Україна

InterNet статус печатных журналов связан с их вхождением в мировые поисковые системы, внесением своего вклада в Webometrics рейтинг и индексацию в ведущих базах данных отслеживания цитирования (SCOPUS и Google Sholar). Рассмотрены особенности и проблемы использования баз цитирования. Подробное внимание сосредоточено на роли и значении Web-сайтов. Анализируется регистрация Web-сайта на 3 известных поисковиках, учитываемых Webometrics — Google, Bing, Yahoo. В статье рассмотрены результаты работы программной системы мониторинга Web-сайтов печатных журналов, издаваемых в НТУУ «КПИ», и особенности сравнения и рекомендаций по Web-сайтах журналов.

InterNet status of print journals due to their incorporation into the global search engines contribute to the Webometrics ranking and indexing in major databases to track quotations (SCOPUS and Google Sholar) The peculiarities and problems of using citation databases. Detailed attention is focused on the role and value of Web-sites. Analyzes Web-site registration on 3 known search engines that included Webometrics — Google, Bing, Yahoo. The article reviews the results of the monitoring program Web-site print magazines published in NTUU «KPI», and feature comparisons and recommendations on Web-site logs.

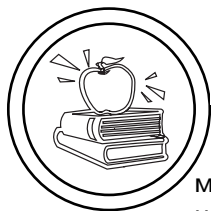
Постановка проблеми

Підвищення статусу друкованих журналів пов'язано з їх входженням у світові пошукові системи, внесення свого вкладу у Webometrics рейтинг та індексацию в провідних базах даних відстеження цитування. Найбільш популярними з баз відстеження цитування є SCOPUS та Google Sholar. Згідно з новими вимогами ВАК РФ, достатньою умовою для включення наукового видання в «Перелік» є його індексация у світових провідних базах даних відстеження цитуван-

ня. Такий відомий Санкт-Петербурзький державний інститут ввів у дію указ, по якому при заміщенні посад викладачів пропонується представляти дані про цитування праць, опублікованих ними за останні п'ять років. На Заході цитування часто використовують не тільки як оцінку науковця, але і як оцінку колективу, наприклад, редакційної колегії журналу.

Основні бази відстеження цитування

SCOPUS — бібліографічна і реферативна база даних та інстру-



мент для відстеження цитованості статей, що публікувалися у наукових виданнях. Індексують 18000 назв наукових видань з технічних, медичних та гуманітарних наук 5000 видавців. База даних індексує наукові журнали, матеріали конференцій та книжкові видання. Дані Scopus використовуються в рейтингу провідних університетів світу Times Higher Education Supplement: World University Rankings (QS TopUniversities), що представлено на сайті <http://www.topuniversities.com/>. Українських університетів у цьому рейтингу немає. Критерії відбору видань для включення в базу даних SCOPUS достатньо детально розглянуто в роботі [1]. До основних з них слід віднести:

1. Наявність сайту журналу зі стандартним набором розділів (про журнал, редакційна колегія, інформація для авторів та інше). Є попередження, що не рахується як сайт журналу розділ іншого сайту (і бібліотеки Вернадського також);

2. Видання повинне мати англomовне назву і публікувати англomовні версії рефератів всіх наукових статей (і це не анотація, а реферат статті, оцінюється також якість англomовних рефератів);

3. Періодичне видання повинне публікувати нові випуски з регулярністю не рідше, ніж 1 раз на рік;

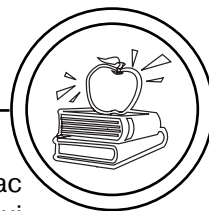
4. Присутність на сторінках Web-сайту повних текстів статей не є обов'язковою умовою, але є бажаним для забезпечення можливості переходу зі сторінок у Scopus до сторінок повних текстів статей.

Проблемою використання бази Scopus є необхідність оплачування. Тому більшою популярністю користується безплатна база Google Scholar

Google Scholar-Пошукова система наукової літератури. Включає статті великих наукових видавництв, архіви препринтів, публікації на сайтах університетів, наукових суспільств і інших наукових організацій. Шукає статті у тому числі і російською мовою. Вона розраховує індекс цитування публікацій і дозволяє знаходити статті, на які були посилання з тих, що є в базі. Ця база проста і доступна у використанні.

Однією з суттєвих проблем пошуку цитованості є «однофамільці» та транслітерація. Ці дві проблеми обумовлюють необхідність безпосередньої участі автора у відборі робіт при оцінці цитованості.

Першою великою пошуковою системою, яка почала активно використовувати індекс цитування, стала Google (використання алгоритму PageRank). PageRank — це метод Google для вимірювання «важливості» сторінки. Коли всі інші фактори, такі як тег Title та ключові слова враховані, Google використовує PageRank, щоб відкоригувати результати так, щоб більш «важливі» сайти піднімуться відповідно вгору на сторінці результатів пошуку користувача. Тобто, порядок ранжирування в Google працює наступним чином: знаходить всі сторінки, що відповідають ключовим словам пошуку; робить ранжування відповідно «сторінковим факторам», таким, як ключові слова;



враховує текст посилань на сторінки; коригує результати даними PageRank.

PageRank — це число, що характеризує голосуючу здатність всіх вхідних посилань на сторінку і те, як сильно вони рекомендують цю сторінку.

У PageRank є одна важлива властивість. На нього більш важко впливати, ніж на будь-який інший фактор ранжування. Це означає, що у PageRank є можливість дати вам перевагу перед конкурентами, якщо він використаний у комбінації з іншими прийомами оптимізації для пошукових машин.

Роль та значення Web-сайту для друкованого журналу

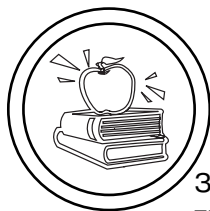
Існують два види подання наукових журналів: 1. Традиційна друкована форма. 2. Електронні журнали.

Ці дві форми мають як позитивні, так і негативні властивості. Так, друкована форма дозволяє ознайомитися з текстом, не використовуючи додаткових електронних засобів. У серйозних журналах «Редакційна колегія» — це об'єднання висококваліфікованих фахівців в науковому напрямку відповідного журналу. Кожна редакція має зв'язки з відповідними рецензентами за певними напрямками. Це одні з основних переваг друкованих журналів. Електронна форма журналу є порівняно новим напрямком і її можна вважати елементом сучасного інформаційного суспільства. Необхідно відзначити, що основним її недоліком є відсутність або незначна кількість людей, які виконують завдання ре-

дакційної колегії. У той же час електронна форма має наступні переваги порівняно з друкованою формою: здійснювати пошук по тексту, перехід за гіперпосиланнями; може міститися багато журналів за різний період; розмір шрифту тексту можна динамічно змінювати; вартість електронного журналу є набагато нижче, ніж друкована версія, або можливий відкритий доступ; поширення і пошук електронних журналів вимагає істотно менших витрат, ніж друкованих версій; наявність Web-сторінки, на якій би обговорювалися думки читачів з приводу статті; наявність підписки на нові статті. Побудова Web-сайту на CMS [2] дає додаткові можливості без додаткових витрат: можливість спілкування редакції з авторами (і в online, і в offline режимах); відсутність проблем у зміні контенту; просте поповнення бази даних журналу при публікації нового видання; можливість обговорення нового видання з метою покращення наступних видань. Сайт журналу як, і будь-який Web-сайт, після створення повинен супроводжуватися і розкручуватися. У більшості випадків при розкрученні виконується реєстрація Web-сайтів журналів на відомих пошукових системах. Для отримання пристойного місця у Webometrics рейтингу доцільно реєструвати Web-сайт на трьох популярних пошуковиках: Google, Bing та Yahoo.

Мета роботи

Створення засобів моніторингу Web-сайтів друкованих журналів Вищого Навчального



Закладу, що дозволяють оцінити InterNet статус відповідних журналів. Чим вище статус, тим більше внесок журналу як у Webometrics рейтинг [3], так і у цитованість не тільки журналу, але й його авторів.

Результати проведених досліджень

Друковані журнали НТУУ «КПІ»

Статус Web-сайтів журналів можна оцінити на відомих пошуковиках по його адресі на сайті <http://push2check.com>. Програму, яку розроблено при виконанні магістерської роботи Р. Іванченко, дозволяє робити моніторинг сайтів із визначеними IP адресами по бажанню чи через певний інтервал часу і при цьому надається + чи – зміни параметру. Важливим для ВНЗ є

статус по Webometrics рейтингу [3], який рекомендує три основних пошуковика — Google, Bing, Yahoo.

При цьому видаються:

Ind (Index) — Індексованість у відповідній пошуковій системі;

BL (Back Link) — Посилання на сайт з пошукової системи;

PR (Page Rank) — Складний параметр пов’язаний з цитованістю (видає Google);

RF (Rich Files) — Цінні файли, які враховуються у Webometrics [3];

M — місце.

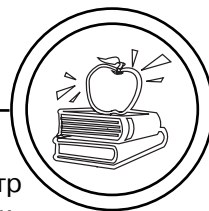
На жаль в НТУУ «КПІ» є тільки 4 Web-сайти друкованих журналів, які існують більше року, і для них було зібрано параметри, які надано у таблиці 1.

Щоб підрахувати місце відповідного журналу треба ввести значимість відповідних коефі-

Таблиця 1

Web-сайти друкованих журналів НТУУ «КПІ»

Пошукові системи	Google					Bing				Yahoo			
	Ind.	BL	PR	RF	M	Ind.	BL	RF	M	Ind.	BL	RF	M
Ізвестия Вузов «Радио-електроника»	279	28	4	23	2	<u>48</u>	<u>13</u>	0	2	<u>159</u>	<u>155</u>	1	3
Технологія і техніка друкарства	531	45	5	163	1	<u>18</u>	<u>8</u>	0	3	<u>121</u>	<u>852</u>	11	1
Електроніка і зв'язь	<u>58</u>	<u>3</u>	4	40	3	<u>66</u>	<u>8</u>	55	1	<u>34</u>	<u>27</u>	1	4
Системні дослідження та Інформаційні Технології	<u>847</u>	<u>17</u>	5	5	4	<u>0</u>	<u>2</u>	0	4	<u>0</u>	<u>602</u>	0	2
Наукові вісті Вісник НТУУ «КПІ» Філософія, Психологія, Педагогіка	Новий Web-сайт (менше 3 місяців)												
	Новий Web-сайт (менше 3 місяців)												



цієнтів. Для Google була прийнята наступна значимість:

Індексованість (Ind.) — 0.01, Зворотні посилання (B L) — 0.2, P R — 0.39, Цінні файли R F — 0.4.

Можна обґрунтувати ці значення. Так «Індексованість» характеризує те, що Google сам, чи з Вашою участю проіндексував сторінки Вашого сайту. Цей коефіцієнт каже про те, що Ваш сайт видно у Google і його значимість мала. Більшу значимість мають зворотні посилання. І найбільшу значимість мають PR та RF, по яким рахується рейтинг Webometrics.

Для Bing та Yahoo була прийнята наступна значимість: Індексованість (Ind.) — 0.01, Зворотні посилання (BL) — 0.49, Цінні файли RF — 0.5.

Аналіз таблиці дозволяє не тільки і не стільки порівнювати журнали по їх представленню в InterNet, але й показує, на що треба звернути увагу. Наприклад, журналу «Системні дослідження та Інформаційні Технології» на реєстрацію у пошуковиках Bing та Yahoo і ще на наявність на сайті PDF файлів, що

підвищить R F, цей же параметр треба підвищити і журналу «Електроника и связь».

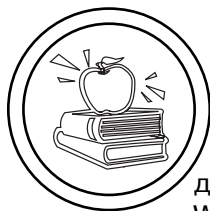
Аналізуючи параметри, які впливають на Webometrics рейтинг, можна виділити деякі кількісні параметри — об'єм Web-сайту, об'єм документів та об'єм і кількість PDF файлів. Ці параметри для існуючих Web-сайтів взяти з InterNet не можливо і тому магістрами кафедри СП-САПР Самойлік, Ушеренко та Шафрай під керівництвом доцента Цуріна О. П. було розроблено програмну систему мовою PHP, яка виконує ці функції. Програмна система достатньо складна і складається з двох підсистем — підсистеми «Робот», який обходить сайти та збирає інформацію і підсистеми «Статистика», яка обробляє зібрану інформацію і видає її користувачу. Результати роботи системи надано у таблиці 2.

Ця таблиця дає можливість, по-перше, порівняти сайти по об'єму, по-друге, сказати, що у «Известия Вузів Радиоелектроника» документи не викладаються, «Системні дослідження та Інформаційні Технології» —

Таблиця 2

Характеристики програмної системи

Назва журналу	Об'єм сайту	М	Об'єм документ.	М	Об'єм PDF	М
Технологія і техніка друкарства	129189	3	103783	2	102991	1
Електроника и связь	21263	5	21105	4	21105	3
Известия Вузів Радиоелектроника	126306	2	2073	5	2073	5
Наукові вісті	66452	4	58261	3	58261	2
Системні дослідження та Інформаційні Технології	194846	1	114307	1	10587	4



документи викладаються у WORD, а документи у PDF становлять 10 %.

Ці таблиці по журналам НТУУ «КПІ», як і таблиці по інститутам, факультетам та кафедрам планується виставляти за адресою <http://webometr.kpi.ua>.

Висновки

Сформовані таблиці фактично характеризують статус Web-сайтів журналів в InterNet просторі. Значення коефіцієнтів у таблицях дозволяють вибрати стратегію підвищення цього статусу.

1. Журнали НАН України і Scopus: актуальні питання : матеріали семінару «Електронні продукти для науки та освіти : повнотекстова база даних ScienceDirect та реферативна база даних Scopus». — 2010 [Електронний ресурс] — Режим доступу : http://www.publications.nas.gov.ua/books/serii/academy/1102010/Documents/2009_02/a2.pdf. 2. Цурін О. П. Web-портал поліграфічних видань журналів Університету / О. П. Цурін, Є. А. Антонович // Технологія і техніка друкарства. — 2009. — № 1-2. — С. 35—39. 3. Цурін О. П. Автоматизовані засоби підняття світового Webometrics-рейтингу Web-сайтів наукових установ / О. П. Цурін, Р. М. Іванченко // Системний аналіз та інформаційні технології: матеріали 11-ї міжнародної науково-технічної конференції SAIT 2009. — Київ. — 2009. — С. 230.

Надійшла до редакції 30.06.10