

Тетяна Симоненко,
мол. наук. співробітник НБУВ

Репозитарій «Наукова періодика України»: шлях оброблення електронних версій

У статті розглянуто реалізацію конвеєрної інформаційної технології оброблення нових надходжень до загальнодержавного репозитарію «Наукова періодика України», що дає змогу проводити оперативне опрацювання та онлайнне представлення 2000 наукових журналів та збірників наукових праць.

К л ю ч о в і с л о в а: електронна копія, конвеєрна технологія, наукова періодика, репозитарій.

The article is devoted to an implementation of the technological conveyor for processing of new arrivals in the state repository «Scientific Periodicals of Ukraine». This conveyor allows rapid processing and online publication of 2000 academic journals and scientific bulletins.

K e y w o r d s: electronic copy, conveyor technology, scientific periodicals, repository.

Конвеєрні технології оброблення нових надходжень нині широко використовуються потужними бібліотеками, вони вважаються найефективнішим способом організації бібліотечної роботи в умовах інтенсивного приросту нових надходжень. Суть конвеєрного оброблення полягає у розподілі функціональних обов'язків на споріднені групи (етапи) і виділенні для виконання кожного з етапів окремого бібліотечного працівника або групи працівників. Завдяки такому підходу стає можливим оптимізація процесів опрацювання партій нових надходжень та суттєве піднесення продуктивності праці бібліотекарів.

На початку 2000-х рр., з інтенсифікацією електронного комплектування, у НБУВ розпочалися пошуки оптимальної конвеєрної технології оброблення електронних документів [1, 2, 3]. Проблема набула ще більшої актуальності у 2008–2009 рр., коли Бібліотека розпочала реалізацію загальнодержавного Інтернет-проекту «Наукова періодика України». В рамках цього проекту НБУВ взяла на себе зобов'язання представляти на єдиній технологічній платформі всю номенклатуру наукових періодичних видань України з можливостями вільного доступу до повних текстів випусків кожного з видань.

Нормативною базою для реалізації цього проекту став Закон України «Про Основні засади розвитку інформаційного суспільства в Україні на 2007–2015 роки» [5], у якому, зокрема, наголошується на необхідності створення «електронних інформаційних ресурсів в архівах, бібліотеках та музеях, науково-дослідних установах з визначенням вимоги щодо обов'язкового зберігання в єдиному електронному форматі результатів наукової діяльності» та забезпе-

чення вільного доступу «до результатів наукових досліджень, створених за рахунок коштів Державного бюджету України». З наведеного випливає вимога оприлюднення знань, одержаних у процесі науково-дослідної, дослідно-конструкторської, проектно-технологічної та виробничої діяльності вітчизняних фахівців, у загальнодоступних електронних журналах і збірниках наукових праць. Підзаконний нормативний акт ВАК України і НАН України «Про затвердження Порядку передавання електронних копій періодичних друкованих наукових видань на зберігання до Національної бібліотеки України імені В. І. Вернадського» передбачає, що засновники журналів, збірників наукових праць, де можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора та кандидата наук і на які можна посилатися у наукових статтях та дисертаціях, передають до НБУВ електронні копії видань, а бібліотека безоплатно розміщує ці копії на своєму сайті і забезпечує до них безкоштовний доступ [4].

Отже, виникла необхідність реалізації дієвої технології, яка б дала змогу проводити оперативне оброблення та онлайнне представлення 2000 наукових журналів і збірників наукових праць. Для реалізації цього проекту в Центрі бібліотечно-інформаційних технологій НБУВ було створено групу у складі 10 працівників. Таким чином, в умовах серйозного дефіциту кадрового ресурсу, виконання проекту вимагало високого ступеня технологічності організації праці.

Наповнення репозитарію «Наукова періодика України» вдалося здійснити засобами конвеєрної технології оброблення нових надходжень «Шлях електронної періодики», розробленої з урахуван-

ням специфіки функціонування даного проекту.

Апаратно-програмний комплекс проекту «Наукова періодика України» був реалізований з використанням семи автоматизованих робочих місць операторів конвеєрної технології та трьох серверів (поштовий сервер, файл-сервер, веб-сервер).

До складу групи увійшли два фахівці з загально-координаційними функціями (керівник проекту та адміністратор репозитарію), системний адміністратор та сім операторів, які, власне, і реалізують технологічний конвеєр. Розглянемо детально посадові обов'язки кожного з них.

Керівник проекту відповідає за загальну координацію роботи групи, вироблення стратегічних цілей, завдань та напрямів модернізації проекту, приймає остаточне рішення зі спірних питань, регулює проблеми, що виникають у взаємовідносинах бібліотеки з видавцями, презентує можливості проекту потенційним технологічним партнерам, реалізує співпрацю з ВАК України, НАН України та іншими заінтересованими інституціями.

Адміністратор репозитарію керує технологічним конвеєром, забезпечує безперебійність його роботи, приймає результати праці у операторів і формує з них цілісний пакет веб-документів для публікації у репозитарії, а також спільно з керівником проекту забезпечує науково-методичний супровід діяльності групи.

Системний адміністратор забезпечує супровід апаратно-програмного комплексу проекту, функціонування серверів та робочих комп'ютерів співробітників.

Безпосередньо з редакціями наукових часописів спілкуються інспектори з прийому електронних копій. У складі групи два інспектори: один приймає електронні копії на компакт-дисках або флеш-носіях (принесені особисто, надіслані Укрпоштою, службою кур'єрської доставки тощо), другий приймає інформацію від видавців електронною поштою. Обов'язки обох інспекторів однакові: вони перевіряють комплекtnість та відповідність встановленим вимогам матеріалів, що надаються видавцями, проводять первинну перевірку всіх наданих файлів та ухвалюють рішення щодо прийняття комплекту документів або повернення його видавцю на доопрацювання.

Літературний редактор редагує анкети з вихідними відомостями періодичних видань, наукові біографії вчених надані видавцями, відомості про наукові організації. Статті часописів публікуються в оригінальному форматі, без додаткових правок.

Інспектор верстки відповідає за роботу з файлами повних текстів статей у форматах pdf та html.

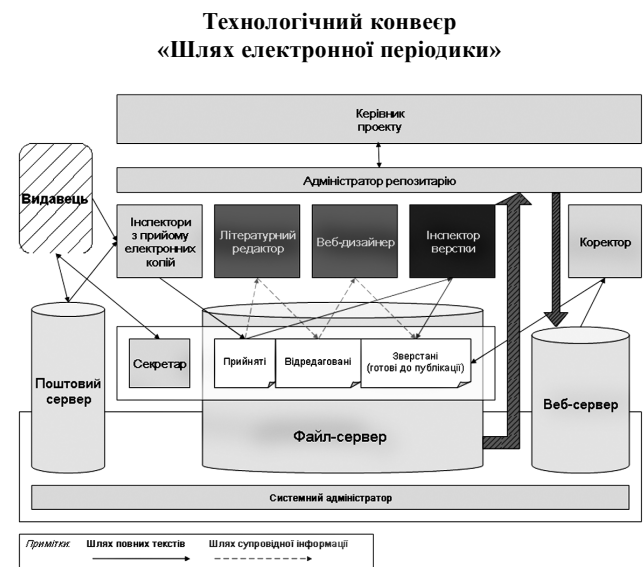
Він інспектує працездатність кожного окремого файлу, у разі потреби, виправляє незначні дефекти або переконвертовує файли з використанням більш оптимальних алгоритмів представлення.

Веб-дизайнер робить веб-верстку анкет з вихідними відомостями періодичних видань, наукових біографій вчених та відомостей про наукові організації.

Коректор регулярно інспектує зміст репозитарію, він відповідає за пошук опечаток, оновлення застарілих даних тощо.

Секретар веде всю документацію проекту, забезпечує прийом відвідувачів. До його обов'язків входить ведення картотеки обліку проходження електронних копій, виготовлення договорів, довідок для ВАК України, інших діловодних документів, надання телефонних консультацій видавцям, прийом і передавання документів факсом, виконання службових доручень керівника проекту та адміністратора репозитарію.

Технологічний конвеєр «Шлях електронної періодики» реалізується таким чином (див. схему):



При першому зверненні до НБУВ редакція видання одноразово формує пакет супровідних документів (відповідно до наказу «Про затвердження Порядку передавання електронних копій періодичних друкованих наукових видань на зберігання до Національної бібліотеки України імені В. І. Вернадського»). До цього пакета входять такі документи:

1. Офіційний лист-клопотання від установи-видавця, засвідчений підписом керівника та печаткою установи. Лист готується у довільній формі, без зазначення конкретних випусків, томів, які видавець просить взяти на зберігання.

2. Сканована копія «Свідоцтва про державну реєстрацію засобу масової інформації» оформлена у вигляді файлу формату jpeg розміром не менше 500x700 пікс.

3. Сканована копія обкладинки одного з нових випусків видання, оформлена у вигляді файлу формату jpeg розміром не менше 140x200 пікс.

4. Тека, що містить усі файли окремого випуску видання (у випадку передавання кількох випусків видання, кожний з них оформляється окремою текою файлів):

- файл у форматі html, де повністю розкритий зміст окремого випуску видання і кожна стаття у змісті супроводжується окремим гіперпосиланням на файл повного тексту. У файлі не допускається реклама або посилання за зовнішні ресурси Інтернет;
- тека, що містить усі файли повних текстів статей, названих у змісті випуску. Кожна окрема стаття повинна оформлятися у вигляді окремого файлу формату pdf або html; в іменах файлів допускаються лише цифри та символи латиниці;
- файл з інформацією для авторів та вимогами до оформлення статей (за бажанням видавця);
- файл(и) змісту видання альтернативною мовою (за бажанням видавця).

5. Наукові біографії головного редактора та/або членів редколегії, довідково-енциклопедична стаття про установу-видавця та/або про видання, його історію та специфіку (за бажанням видавця):

- тексти біографій/довідкових статей, оформлені у вигляді файлів стандартних текстових форматів (doc, odt, rtf, txt);
- фото в форматі jpeg.

6. Анкета – «Інформація для титульного файлу», оформлена у вигляді файлу стандартного текстового формату (doc, odt, rtf, txt). Поля анкети:

| | |
|------------------------------------|--|
| Рік заснування: | |
| Проблематика: | |
| ISSN: | |
| Свідоцтво про державну реєстрацію: | |
| Фахова реєстрація у ВАК України: | |
| Галузь науки: | |
| Періодичність: | |
| Мова видання: | |
| Засновник: | |
| Головний редактор: | |
| Заступники головного редактора: | |
| Відповідальний секретар: | |
| Члени редколегії: | |
| Адреса редакції: | |

Цей пакет супровідних документів представник редакції видання привозить особисто, надсилає Укрпоштою або іншою службою доставки документів. Спочатку його опрацьовує інспектор з прийому електронних копій, відповідальний за прийом матеріалів на компакт-дисках та флеш-носіях. У подальшому, після розміщення видання в репозитарії, передавання електронних копій проводиться, в основному, засобами e-mail, тобто матеріали потрапляють на опрацювання до інспектора з прийому електронних копій, відповідального за роботу з електронною поштою. В цьому випадку редакція видання надсилає електронною поштою не весь комплект документів, а лише матеріали, які необхідно додати/оновити (теки файлів нових випусків видання та, за необхідності, оновлення до анкети видання, однієї з наукових біографій тощо).

Інспектор з прийому електронних копій проводить первинну експертизу комплектності та відповідності отриманого пакета документів встановленим вимогам. Інформація про отримання нових матеріалів реєструється у картотеці обліку проходження електронних копій. У разі виявлення відсутніх, пошкоджених або неналежно оформлених документів інспектор з прийому електронних копій повідомляє про це редакцію видання і робить відповідні помітки у картотеці обліку.

Картотека обліку проходження електронних копій є базою даних, розміщеною на файл-сервері репозитарію. В ній фіксується вся службова інформація щодо проходження електронних копій технологічним конвеєром та стану готовності матеріалів до публікації. Доступ до картотеки (з повними правами читання та вибірково правами запису) мають всі члени технологічної групи, за ведення картотеки відповідає секретар.

Якщо пакет документів оформлений належним чином і редакція видання потребує довідки про внесення електронної копії видання на зберігання для перереєстрації часопису у ВАК України, інспектор з прийому електронних копій дає доручення секретарю підготувати необхідну довідку. Крім того, на вимогу редакції видання може бути встановлений період затримки представлення нових випусків у відкритому доступі. У разі, якщо видавець бажає встановити термін такої затримки (до двох років), секретар, крім довідки, готує також договір про передачу електронної копії на зберігання, у якому вписуються умови та вказуються терміни публікації нових випусків видання у репозитарії.

Інспектор з прийому електронних копій вивантажує файли, отримані від редакції видання, на файл-

сервер репозитарію у теку «Прийняті». Файл-сервер репозитарію сконфігурований таким чином, щоб забезпечити технологічній групі захищений доступ до матеріалів, які проходять конвеєрне оброблення в процесі підготовки до публікації на веб-сервері, та журналізацію дій з файлами. В основному, весь вхідний потік файлів потрапляє до теки «Прийняті». Але з деякими видавцями, з якими репозитарій вже тривалий час успішно співпрацює, практикується так звана «швидка публікація». У цьому випадку інспектор з прийому електронних копій вивантажує нові файли безпосередньо у теку «Готові до публікації». Це скорочує час опрацювання електронних копій і забезпечує оперативну появу нових випусків видання у репозитарії.

У теці «Прийняті» окремо зберігаються файли, що потребують опрацювання літературним редактором та випуски видань, які публікуються у репозитарії без додаткового редагування. Відповідно, випуски видань і супровідні файли проходять різні шляхи на технологічному конвеєрі (див. схему 1).

З теки «Прийняті» літературний редактор завантажує на власний комп'ютер нові файли, зміст яких потребує перегляду та коригування. До таких належать анкети з вихідними відомостями періодичних видань, надані видавцями наукові біографії вчених та відомості про наукові організації. Згідно з існуючими вимогами, ці матеріали видавці подають у вигляді файлів стандартних текстових форматів (doc, odt, rtf, txt). Літературний редактор редагує ці файли за допомогою власного текстового редактора, реєструє свої дії у картотеці обліку проходження електронних копій і у тому ж текстовому форматі вивантажує відредаговані версії на файл-сервер репозитарію до теки «Відредаговані».

Всі файли, опрацьовані літературним редактором, готує до публікації у репозитарії веб-дизайнер. З теки «Відредаговані» він завантажує на власний комп'ютер файли у текстовому форматі і верстає на їх основі html-сторінки згідно зі стилем, прийнятим у репозитарії. До обов'язків веб-дизайнера входить також робота з фото та іншими ілюстративними матеріалами, наданими у комплекті документів. Він відповідає за те, щоб привести всі матеріали до формату, оптимального для представлення у веб-середовищі. За результатами роботи веб-дизайнер вивантажує на файл-сервер репозитарію до теки «Готові до публікації» цілісний комплект веб-сторінок, які мають бути додані на веб-сервер.

Комплекти файлів нових випусків видань проходять інший технологічний шлях. З теки «Прийняті» інспектор верстки завантажує на власний

комп'ютер цілісні комплекти файлів, які стосуються єдиного випуску видання, перевіряє працездатність усіх навігаційних посилань у межах випуску та бездефектність всіх повних текстів у форматах html та/або pdf. При потребі інспектор верстки виправляє незначні дефекти або переконвертовує файли з використанням більш оптимальних алгоритмів представлення. Політика репозитарію передбачає публікацію матеріалів у тому вигляді, у якому їх надали видавці. Тому інспектору верстки суворо заборонено вносити у файли будь-які зміни, пов'язані з коригуванням їх змісту. Головний принцип, якого дотримується інспектор верстки: в усіх файлах, наданих редакцією видання, графічне відображення всієї інформації повинно залишитися незмінним, але у файли можуть вноситися технічні правки, пов'язані з оптимізацією механізмів завантаження та зменшенням фізичного розміру файлів або виправленням синтаксичних помилок мови html. Проінспектовані комплекти файлів вивантажуються на файл-сервер до теки «Готові до публікації».

З текою «Готові до публікації» працює адміністратор репозитарію. У технологічному процесі лише він має право оновлювати репозитарій, додаючи нові файли на веб-сервер. Після остаточної передпублікаційної інспекції та перевірки режиму публікації конкретного видання (оперативно або із затримкою), адміністратор репозитарію переносить усі необхідні файли з файл-серверу на веб-сервер. На цьому процес публікації нового випуску видання у репозитарії завершується.

Післяпублікаційну інспекцію проводить коректор. Він, за особливим алгоритмом, інспектує всі веб-сторінки репозитарію, виявляє опечатки в опублікованій супровідній інформації, відслідковує актуальність даних про видання, установи-засновники та членів редколегій. Крім того, коректор відповідає за синхронізацію посилань між виданнями. Типовим прикладом такої синхронізації є оновлення посилань у списках членів редколегій. Відомі вчені часто є членами редколегій одразу кількох видань. Якщо редакція одного з видань надала для публікації у репозитарії наукову біографію ученого, потрібно відслідкувати всі видання, де цей учений є членом редколегії та оформити відповідні гіперпосилання. Для виконання своїх обов'язків коректор безпосередньо з репозитарію завантажує на власний комп'ютер веб-сторінки, які потребують коригування, локально редагує їх та вивантажує виправлені версії сторінок на файл-сервер до теки «Готові до публікації».

Завдяки конвеєрній технології «Шлях електронної періодики» вдалося досягти високого рівня організації праці. Пропускна здатність конвеєра дає змогу щоденно приймати від видавців, опрацьовувати та готувати до публікації понад 50 випусків періодичних видань (у середньому по 20 статей у кожному). В результаті проект «Наукова періодика України» за короткий час став найбільшим у Європі сховищем безкоштовних повних текстів наукових статей.

Роботи з подальшого удосконалення реалізованої технології ведуться кількома напрямками. По-перше, дедалі більше уваги приділяється питанням автоматизації взаємодії видавців з файл-сервером репозитарію. Практика показала, що серед вітчизняних видавців наукової періодики багато тих, хто, виробляючи стратегію розвитку власного часопису, покладає серйозні сподівання на співпрацю з НБУВ. Такі видавці готові виділяти додаткові кадрові ресурси задля створення в межах репозитарію повноцінних веб-сайтів власних видань з високотехнологічним пошуково-навігаційним апаратом. НБУВ має серйозний технологічний базис, але, на жаль, не має необхідного кадрового ресурсу, щоб власними силами створювати і оновлювати високотехнологічні веб-сайти для тисяч українських наукових видань. У даній ситуації цілком оптимальною є модель централізованого загальнодержавного науково-видавничого середовища, побудованого на базі інфраструктури «хмарних обчислень» та інноваційних засобів веб-комунікації. У такому середовищі НБУВ відводиться роль технологічної корпорації – партнера українських видавців.

По-друге, важливим напрямом технологічного вдосконалення репозитарію є організація каналів його взаємодії з провідними світовими науково-інформаційними порталами. За обсягами представленої інформації «Наукова періодика України» входить до числа найбільших науково-інформаційних сховищ світу. Тому розвиток проекту перебуває у полі зору багатьох потужних інформаційних агрегаторів. Від Інтернет-компаній НБУВ отримала декілька пропозицій щодо співпраці. Наразі стратегічним технологічним партнером проекту обрана корпорація Google. Інструментарій, який постачається Google, забезпечує функціонування базового пошукового апарату репозитарію. Але на сьогодні проект досяг тих масштабів, коли наявних базових можливостей для повноцінного пошуку наукової інформації

користувачам вже замало. В ході консультацій адміністрації репозитарію з фахівцями Google вдалося виробити нову модель даних, яка сприятиме запровадженню розширеного пошуково-навігаційного апарату. Перехід до нової моделі передбачає запровадження кількох нових ланок у технології формування репозитарію, проте базовий принцип функціонування технологічного конвеєра залишиться незмінним.

Третій напрям удосконалення технології функціонування «Наукової періодики України» пов'язаний із запровадженням каналів автоматизованої взаємодії проекту з його замовниками – ВАК України та НАН України. Аналізуючи завдання, поставлені перед проектом, варто зазначити, що він розглядається управлінськими структурами як джерело отримання формалізованих оцінок щодо ефективності діяльності суб'єктів наукового процесу. Репозитарій повинен стати прикладним інструментом управління, зручним для отримання довідкової, статистичної, наукометричної інформації, пошуку плагіату тощо.

Список використаних джерел

1. Баркова О. В. Застосування технології конвеєрного формування електронних ресурсів у процесі співпраці виробників наукової інформації / О. В. Баркова // Документознавство та інформаційна діяльність. Наука. Освіта. Практика : зб. матеріалів наук. конф., 18 грудня 2002 р., м. Київ. – К. : ДАКККіМ, 2003. – С. 43–46.
2. Баркова О. В. Мережева конвеєрна технологія формування електронних ресурсів наукової інформації / О. В. Баркова // Науково-технічна інформація. – 2002. – № 1. – С. 44–49.
3. Костенко Л. И. Конвеєрна інформаційна технологія обробки текущих поступлений в Национальной библиотеке Украины имени В. И. Вернадского / Л. И. Костенко, Ю. В. Яковлева // Бібліотечні будівлі в контексті сучасної моделі інформаційно-комунікативної діяльності бібліотек України: матеріали Всеукраїнського симпозиуму, Одеса, 3 вересня 2007 р. – Одеса, 2007. – С. 40–47.
4. Про затвердження Порядку передавання електронних копій періодичних друкованих наукових видань на зберігання до Національної бібліотеки України імені В. І. Вернадського [Електронний ресурс] : наказ ВАК України та НАН України від 07.07.2008 № 436/311; зареєстровано в Міністерстві юстиції України 15.01.2009 № 20/16036. – Режим доступу : http://www.nbu.gov.ua/law/08_elkop.html. – Назва з екрана.
5. Про Основні засади розвитку інформаційного суспільства в Україні на 2007–2015 роки : Закон України від 9 січ. 2007 № 537-V // Відомості Верховної Ради України. – 2007. – № 12. – С. 102.