УДК 026.06:004.91+930.25:004

Ульріх Йоганнес Шнайдер
доктор філософії, професор
філософії, директор
Ляйпцігської університетської
бібліотеки (м. Ляйпціг, ФРН)

ШАНСИ ТА РИЗИКИ ОЦИФРОВУВАННЯ КУЛЬТУРНОЇ СПАДЩИНИ

Завдяки оцифровуванню книжок, здається, втілюється європейська мрія, мрія нескладного поєднання знання в мережах та мрія повного доступу до нього. В казковій країні сканованого світу книжок я можу дуже просто мати доступ до всіх текстів, котрі тільки забажаю. Таким чином, у мене є будь-який предметний текст, чи архівний документ, або музичний твір, такої ж якості, незалежно від сторіччя його походження. Всі зацікавлені дослідники можуть працювати в Інтернеті, як тільки-но ця візія (бачення) технічно втілюється, не дивлячись на місцеві бібліотечні фонди та умови доступу до них. Багато будуть робити це, бо вони вже дано сподіваються на нескладний доступ до ресурсів. Завдяки оцифровуванню збуваються дослідницькі мрії [1].

Саме у сфері культурної спадщини, котрою керують бібліотеки, на даний час відбулося найбільш масштабніше оцифровування ретроспективних ресурсів. Бібліотеки вже давно зрозуміли, яким чином можна модернізувати свої послуги. Спочатку вони виставили каталоги в режимі он-лайн, що означало великі зміни у поданні інформації, хоча до сих пір поєднання каталогів у мережі ще не завершено до кінця. Далі були організовані електронні служби обслуговування користувачів. Замовлення, видачу, повернення книг, чи продовження терміну користування книгами можна здійснювати з будь-якого терміналу. Потім настала черга фондів, й тут саме бібліотеки стали інноваційними у технічному відношенні, ще задовго до того, як видавництва розпочали пропонувати свої монографії в електронному вигляді.

Оцифровані фото книг роблять, головним чином, з двох причин: з одного боку, заради захисту фондів особливо цінних документів, з іншого боку – заради більш кращого користування. Оцифрована картинка, - й це можуть засвідчити співробітники, котрі

займаються опрацюванням рукописів, - у відношенні оптичної якості переважає оригінал, її можна рівень за рівнем збільшувати, навіть до того рівня, якого не сягає людське око. Крім цього, ці дані (файл) можна в любий момент копіювати, й при цьому не завдавати шкоди оригіналу. Саме тому, немає нічого дивного в тому, що бібліотеки з великими давніми фондами дуже охоче використовують таку новітню техніку.

Оцифровування книжок відбувається повсюди у світі. Зокрема, є комерційні та ліцензовані банки даних, як наприклад, «Ранні американські друки» (EAI, Early American Imprints), «Ранні англійські книжки он-лайн» (EEBO, Early English Books Online), котрі разом мають біля 200000 оцифрованих творів. Багато бібліотек в Німеччині пропонують цифрові копії давніх друків, наприклад Державна бібліотека в Мюнхені (Staatsbibliothek in München) з більш ніж півмільйона цифрових назв. Але інтерес представляють також невеликі зібрання, наприклад, часописи 18-го сторіччя в Білєфельдській університетській бібліотеці, емблематичні книжки в Бібліотеці Герцога Августа у Вольфенбюттелі, або арабські рукописи в Ляйпцігській університетській бібліотеці. Деякі спеціальні каталоги допомагають при виявленні нових творів, зокрема Центральний опис оцифрованих друків (ZVDD) з даними про майже 600 000 назв.

Майже скрізь ці оцифровані друки можна побачити після натискання миші комп'ютера як файли-зображення, наприклад, більше мільйона назв можна побачити на платформі «Gallica» Французької національної бібліотеки. Тут можна шукати лише мега-дані, інколи – також структурні дані, але поки що – не самі тексти всередині. Для того щоб користувачам надати можливість повнотекстового пошуку, необхідні дуже великі кошти, саме це призвело до того, що на даний час здебільшого в бібліотечних фондах є лише оцифровані фото текстів. Це стосується також Німецької цифрової бібліотеки (Deutsche Digitale Bibliothek), котру складають разом багато різних учасників, та котра стала платформою он-лайн в 2012 році.

Хоча Німецька цифрова бібліотека (НЦБ) – це бібліотека, однак тут зібрані насправді не лише книжки, але й об'єкти (з музеїв) та документи (з архівів). Поєднання у мережах – це те, над чим працюють у новоствореній «компетентній мережі», що відбувається заради певного бачення, або візії: «Візія Німецької цифрової

бібліотеки звучить так, що всі зацікавлені громадянки та громадяни можуть завдяки центральному доступу в Інтернеті запросити та подивитися необхідні об'єкти» (www.deutsche-digitale-bibliothek.de). Тобто, це дуже велика пошукова машина творів німецької культури, величезний каталог 30000 німецьких культурних закладів, котрі стосуються політичних цільових намірів. Те, як все це поєднати між собою, оплатити та зробити технічно привабливим – представляє собою цілий конгломерат проблем, котрий повинна вирішити ціла мережа органів та інституцій.

Візія (бачення) знань, поєднаних однією мережею, втілюється в НЦБ завдяки технічній взаємодії різних культурних закладів з очевидним підпорядкуванням, що культура завжди й однаково цікавить всіх. На веб-сторінці Німецької цифрової бібліотеки (НЦБ) повідомляється: «Той, хто задає мишкою пошук терміну «Бетховен», отримує в результаті дослідницьку літературу, ноти, портрети, йому необхідно лише ще раз натиснути на мишку, щоб узнати більше». Й на це витрачаються мільйони грошей окремих федеральних земель та всієї країни? Якщо той, хто цікавиться Бетховеном, є журналістом, то він знає свої джерела. Якщо ж це дослідник, то само собою зрозуміло.

Коли ж у цьому випадку звичайний громадянин немає ніякого уявлення, то йому допоможе стаття «Бетховен» у Вікіпедії, котра подає посилання (лінки) на ноти та літературу, звукові записи, зображення та багато інше. Що ж більше стосовно цього НЦБ може надати, так це мабуть відбитки оригінальних творів, може ще пару бюстів Бетховена у форматі 3Д, архівні документи, - побачимо в майбутньому. Потрібен ще трішки час, може реалізований проект НЦБ в майбутньому ще більше зацікавить заінтересованого в цьому громадянина. Бо саме він став причиною того, що він не зовсім повністю був задоволений сучасним станом з інформацією про Бетховена.

Оцифровуванням називають не лише цифрове фотографування, котре дозволяє циркуляцію в Інтернеті таким чином отриманих культурних творів спадщини та їх поєднання в мережі. Оцифровуванням можна назвати також спосіб перетворення надрукованих текстів в документи, котрі можна читати. OCR – це чарівне слово, за допомогою якого сьогодні для гуманітарних та культурологічних досліджень утворюються зовсім нові простори: OCR (Optical Character Recognition), тобто «оптичне розпізнавання

тексту», іншими словами – перетворення друкованого тексту по буквам у файл, котрий читається машинально. Сьогодні кожен знає службу «Google books» / «Гуггл книги», завдяки якій в цифровій формі, подібно до текстового файлу, можна проводити пошуки в багатьох текстах (численні мільйони томів), котрі вільно представлені в Інтернеті. Можна вже зараз передбачити, що наше письмове культурне минуле на екрані комп'ютера та інших технічних засобів може перетворюватися в світ текстів, котрі можна читати.

Й не одна лише фірма Google / Гуггл надає змогу читати повні тексти надрукованих творів. Є німецькі некомерційні проекти, наприклад проект «Гутенберг» / «Gutenberg», на котрому зберігаються понад 5500 романів 1200 авторів, також банк даних www.zeno.org, котрий вихваляє себе як «найбільша німецькомовна повнотекстова бібліотека» та пропонує поряд з текстами 7000 авторів, ще й 40000 зображень. Американська платформа загального користування www.archive.org пропонує 3 мільйони текстів у вільному доступі. Багато інших платформ працюють над перетворенням цієї написаної інформації в тексти, котрі будуть читатися за допомогою машин.

Цифрові повні тексти подібні до E-Books, котрі випускаються сьогодні, вони технічно вписуються напряму в наукову текстову культуру, котра на сьогодні є вже реальністю в сфері природничих наук, техніки та медицини. Дослідник, котрий представляє природничі, технічні чи медичні дисципліни, сьогодні вже не може працювати без електронної передачі текстів. Всі актуальні для нього часописи – це електронні публікації. Крім цього, обмін результатами досліджень відбувається сьогодні електронними шляхами. Така культура наукової праці на основі банків даних зараз також переходить в сферу гуманітарних та соціальних наук, котрі поступово (в залежності від національних мов та культурного середовища), але безперервно використовують нові можливості більш широкої комунікації.

Презентації Німецької цифрової бібліотеки (НЦБ) в мережі та під час багатьох заходів, наприклад, під час з'їздів бібліотекарів, відмовляються від порівнянь успіхів. Пропозиції «Google books» / «Гуггл книги», котрі частково конкурують з НЦБ, ніде не тематизуться, також не обговорюється те, в чому саме полягає різниця між «Google books» / «Гуггл книги» та НЦБ. Наприклад,

здається, що НЦБ поєднує оцифровані матеріали, тобто сфотографовані зображення сторінок текстів, не через свій сервер (що було б гігантським викликом, як у технічному, так і у фінансовому відношенні), а подає посилання (лінки) на ці оцифровані матеріали. Тобто тут НЦБ не робить крок до національної платформи, а лише оперує як координаційний та каталоговий центр, котрий не бажає відбирати можливості у володарів оригіналів оцифрованих матеріалів. Конкретним наслідком цього є те, що повнотекстовий пошук неможливо організувати уніфіковано, а лише згідно з технічними параметрами, котрі установлені інституціями, які надають вказані матеріали.

Можна собі легко уявити, що Баварська державна бібліотека (Bayerische Staatsbibliothek), котра співпрацює з Google, свої оцифровані книжки в рамках проекту Німецької цифрової бібліотеки (НЦБ) виставить у форматі «оптичного розпізнання тексту» (OCR), однак можливості багатьох інших бібліотек не будуть дозволяти це бачити. Тоді НЦБ стала б мега-каталогом для даних назв книг, й таким чином вона б принципово нічим не відрізнялася від інших мега-каталогів, наприклад «Карлсруеського віртуального каталогу» (KVK, Karlsruher Virtueller Katalog), або «Вьорлдкат» (Worldcut), завдяки яким можна шукати електронні ресурси (але не в рамках їх самих). НЦБ може запропонувати на відміну від цих бібліотечних інструментів, розширення за рахунок зібрань архівів та музеїв, однак необхідну кількість таких фондів можна буде досягти лише завдяки сумісній великій підтримці відповідних федеральних земель, що на даний момент є малоймовірним.

Дилема Німецької цифрової бібліотеки (НЦБ) пов’язана зі словом «оцифровування», оскільки це слово, здається, має магічну силу щодо тих, хто приймає політичні рішення. Однак однозначно поки ще не визначено, які саме виграші для культури несе в собі Німецька цифрова бібліотека (НЦБ) та на які дійсно важливі питання може надати відповіді цей інструмент. НЦБ виступає як нова інституція поряд з іншими, вже існуючими інституціями бібліотек, об’єднаннями певних бібліотек, міжнародними мега-каталогами. При цьому ще не видно, щоб НЦБ виконувала б певну роботу, котра б зменшувала навантаження уже існуючих активних інституцій, які виконують оцифровування матеріалів. Звичайно, в рамках НЦБ стане можливим широке покриття оцифровуванням навіть невеликих інституцій, котрі до сих пір з причин недостатніх

технічних компетенцій не могли приймати в цьому участь. Ще пройдуть десятиріччя, поки НЦБ стане центральною національною інстанцією для німецьких бібліотек, архівів та музеїв, котрі зберігають культурну спадщину [2].

Як здається, перш за все, в даний момент відсутня програма розвитку, тобто планування послідовних, недешевих робочих кроків. Як закликав президент Гьоте-Інститутів Клаус-Дітер Лєманн, Німецька цифрова бібліотека (НЦБ) спочатку повинна представити свої досягнення у бібліотечній сфері. Однак такі прагматичні підходи очевидно ще ніде не були прийняті на програмному рівні. Недолік сучасного бібліотечного ландшафту в Німеччині Наукова Рада (Wissenschaftsrat) нещодавно вбачала в тому, що підтверджуючі інструменти знаходяться в руках різних об'єднань, й не є уніфікованими. З приблизно 200 наукових бібліотек Німеччини лише 16 отримують гроші від федеральних земель, поза межами цього розшарування допомагає лише Німецьке дослідницьке товариство (Deutsche Forschungsgemeinschaft). Отже збоку НЦБ бажаний був би ефект, котрий хоча б в сфері оцифровування культурної спадщини призвів до об'єднання речей, котрі взаємопов'язані між собою. Однак стан речей залишається незадовільним, оскільки поки ще немає нікого, хто б надав довідку для НЦБ окрім технічних та організаторських можливостей. Німецька цифрова бібліотека (НЦБ) – це інституційно-організаційний монстр й зовсім не візія, котрої б бажали багато людей, які працюють з культурною спадщиною [3].

Література:

1. Ulrich Johannes Schneider, Die Erfindung des allgemeinen Wissens. Enzyklopädisches Schreiben im Zeitalter der Aufklärung. – Berlin: Akademie Verlag, 2013. – 259 S.

2. Ulrich Johannes Schneider (Hg.), Seine Welt wissen. Enzyklopädien in der Frühen Neuzeit. – Darmstadt: Primus Verlag, 2006. – 256 S.

3. Ulrich Johannes Schneider (Hg.), Textkünste. Buchrevolution um 1500. –Darmstadt: Wissenschaftliche Buchgesellschaft, 2016. – 224 S.

Переклав з німецької мови Володимир Абашнік