

Intensyfikacja eksploatacji publikacji cyfrowych - abstrakt

Biblioteki cyfrowe stają się coraz bardziej integralną częścią instytucji kultury oraz instytucji edukacyjnych, których celem działania jest gromadzenie oraz udostępnianie i popularyzowanie różnego rodzaju treści: książek, dokumentów, fotografii itp. Bardzo często dla tego typu instytucji możliwość zaistnienia w internecie jest dodatkowym elementem promocyjnym, ale też elementem wspierającym efektywność realizacji ich misji. Trudno się jednak oprzeć wrażeniu, że mimo możliwości jakie daje dzisiaj rozwój technologii informacyjnych i komunikacyjnych znaczna część instytucji posiadających własne biblioteki cyfrowe, albo współpracujących w konsorcjach tworzących regionalne biblioteki cyfrowe, funkcjonuje w modelu komunikacyjno-organizacyjnym, który w analogii do Web 1.0¹ można określić jako model instytucji 1.0. W erze Twittera, Facebooka i Google+, w której coraz większego znaczenia nabiera globalna eksploatacja pojedynczych myśli, użytkownicy oraz gospodarka sieciowa stawiają przed bibliotekami coraz większe wymagania.

Można wysnuć tezę, że sama technologia nie wystarczy aby wypracowane latami tradycyjne relacje pomiędzy instytucjami kultury a ich użytkownikami zmieniły się nagle z dnia na dzień. Technologia może jednak być dobrym impulsem do zainicjowania nowych form współdziałania oraz komunikowania się, bazujących na modelu 2.0², w którym użytkownik staje się podmiotem dostarczającym instytucji wartości dodanej w postaci własnej twórczości oraz własnej pracy na rzecz rozwoju danej instytucji. Taka interakcja może być korzystna zarówno dla instytucji jak i dla użytkownika ponieważ stymuluje rozwój autentycznego partnerstwa umożliwiającego realizację własnych celów przy jednoczesnym wspieraniu i potęgowaniu realizacji celów innych podmiotów.

Celem prezentacji jest zaproponowanie rozwiązań, które w opinii autora mogą przyczynić się do zwiększenia zainteresowania bibliotekami cyfrowymi wśród czytelników oraz podmiotów, które mogłyby być zainteresowane współfinansowaniem rozwoju kolekcji cyfrowych. Prezentowane rozwiązania są pochodną doświadczeń zdobytych podczas pracy przy tworzeniu Biblioteki Cyfrowej Uniwersytetu Wrocławskiego (www.BibliotekaCyfrowa.pl) oraz wydawnictwa publikacji elektronicznych.

¹ Wikipedia [Dostęp 13.06.2011]. Dostępny w World Wide Web: http://en.wikipedia.org/wiki/Web_1.0

² Analogia do Web 2.0. Wikipedia [Dostęp 13.06.2011]. Dostępny w World Wide Web: http://pl.wikipedia.org/wiki/Web_2.0

Intensyfikacja eksploatacji publikacji cyfrowych

Tomasz Kalota



Biblioteka Uniwersytecka we Wrocławiu
Digitalizacja.pl

Proces budowania Bibliotek Cyfrowych w Polsce zaczyna przybierać charakter masowy. Ich konstrukcja przypomina rodzaj cyfrowego magazynu z książkami, czy czasopismami, głównie ze względu na to, iż przedmiotem opisu i prezentacji staje się jednostka inwentarzowa. W obliczu tak dużego wyzwania jak zdigitalizowanie wszystkich zbiorów jest to praktyka zrozumiała, ale w erze Twittera, Facebooka i Google+, w której coraz większego znaczenia nabiera globalna eksploatacja pojedynczych myśli, użytkownicy oraz gospodarka sieciowa stawiają przed bibliotekami coraz większe wymagania.

Jak to zmienić?

Małe kroczki

**Cukierki
zamiast
cukierni**



Wykorzystano slajd z prezentacji
Emanuela Kulczyckiego
udostępnionej na licencji CC BY 3.0
<http://www.slideshare.net/emanuelkulczycki/>

Plan prezentacji

- Granulacja treści metodą na zwiększenie mobilności myśli
- Formaty mobilne
- Wieloformatowość publikacji
- Dobry OCR warunkiem dobrego udostępniania
- Linkowanie do obiektów cyfrowych

"Granulacja" informacji czyni ją bardziej mobilną

Steve Jobs - prezes firmy Apple przekonywał muzyków do tworzenia utworów niezależnych od "opakowania" w postaci albumów, twierdząc, że w ten sposób łatwiej będzie je dystrybuować.

Granulacja treści metodą na zwiększenie mobilności myśli - przykłady

"Nowe horyzonty" - książka Rafała Dutkiewicza - Prezydenta Wrocławia:

<http://www.bibliotekacyfrowa.pl/publication/22322>

Każdy rozdział jest osobnym plikiem PDF, dzięki czemu takie funkcje BC jak: statystyki odwiedzin, tagowanie, ocenianie, opisywanie, można wykorzystać dla każdego rozdziału niezależnie, a publikacja jest mimo to spójna i prezentowana jako całość.

Zobaczmy co Europeana wie na temat hasła **Nowa Europa**

<http://www.europeana.eu/>

Granulacja treści metodą na zwiększenie mobilności myśli - przykłady

"Nowe idee początku XXI wieku" pod red. Marcina Winiarskiego:

<http://www.bibliotekacyfrowa.pl/publication/20501>

Jest to praca zbiorowa, w której każdy artykuł ma swojego autora. Dzięki granulacji treści każdy artykuł można osobno opisać i przy pomocy schematu Dubli Core można nadać każdemu tekstowi właściwy tytuł, opisać go właściwymi słowami kluczowymi i co najważniejsze każdy ma właściwego autora, który jest indeksowany w bibliotece cyfrowej jako autor. Rozwiązuje się więc problem kolejności autorów przy zbiorowej publikacji. Mimo tego rozdrobnienia cała praca zbiorowa zachowuje pierwotną spójność.

Zobaczmy co Europeana wie na temat hasła **Rynek wiedzy**

<http://www.europeana.eu/>

Formaty mobilne



EPUB – (skrót od electronic publication) otwarty standard, oparty na języku XML, służący do publikowania elektronicznych książek (e-booków). Jest to bezpośredni następca rozwijanego wcześniej formatu Open eBook. Tworzone w nim książki nie mają podziału na strony, choć istnieje możliwość wyświetlania na marginesie numeru strony pochodzącego z książki drukowanej. Format ten jest coraz szerzej stosowany w czytnikach książek elektronicznych.

MOBI, PRC – format Mobipocket – jest to format książek na Kindle

PDF (ang. Portable Document Format, przenośny format dokumentu) – format plików służący do prezentacji, przenoszenia i drukowania treści tekstowo-graficznych, stworzony i promowany przez firmę Adobe Systems. Język opisu pliku PDF jest okrojoną wersją języka programowania PostScript wzbogaconą o elementy hipertekstowe.

Wieloformatowość publikacji

Przykłady publikacji udostępnionych w wielu formatach:

Wolne lektury:

<http://wolnelektury.pl/katalog/lektura/krolowa-sniegu/>

Złote Myśli:

http://www.zlotemysli.pl/kategorie/0/katalog-publicacji.html?sb=publish_date&fb=epub

eBooki.com.pl:

<http://www.ebooki.com.pl/Postepowanie-nieprocesowe-w-sprawach-osobowych-oraz-rodzinnych>

Dobry OCR warunkiem dobrego udostępniania

OCR (ang. Optical Character Recognition) – zestaw technik lub oprogramowanie służące do rozpoznawania znaków i całych tekstów w pliku graficznym o postaci rastrowej. Zadaniem OCR jest zwykle rozpoznanie tekstu w zeskanowanym dokumencie (np. papierowym formularzu lub stronie książki).

, Nip. może być opiekunem, ani mieć
iaki.;
kiegożkol.", iek wpływu do opieki.
' .
ie roo.te b) i\Jż 8wiad
ie
w a k c i e u r 0-
czystym, albo urzędowym, ani przypu..
szczanym (10 ś. A11adczenia w sądzie.
Nie n, oze !la. Ć w sądzie, 8m iako P o \V Ó
<l,
ani tako P n z w 8 n y, tylko p
d i. n. ieniem i
prz, ez zastęp. ftwo szcz
gólnego K, u r a t o

Terminem „literatura staroskandynawska” lub „nordycka” tradycja naukowa określa wspólnotę literacką w początkowym okresie piśmiennictwa na obszarze Skandynawii, tzn. w Danii, Szwecji i Norwegii oraz na północnych wyspach Atlantyku. W okresie wędrówek ludów osiedlili się tam Normanowie, spośród których w IX wieku wyodrębnili się językowo Danowie (Duńczycy), zaś w X wieku — Szwedzi i Norwegowie.

Breslau den 18. Julii.

Ehnerachtet die Feinde nach ihrem eignen Beschlusse den Obersten Dardowicz mit seinem Paul Lattara u. den Oberstl. Deißlitz mit 600 Husaren ausgeschickt, welche nach überdem durch 1000 Mann teilsche Cavallerie unter dem Commando des Oberstl. Grafen von Kollowratz, unterstützt worden, um den Hn. Obrist Lieutn. v. Schüg, welcher mit seinen unterhabenden Husaren, das ganze Land durchschiffte und in Contribution setzte, zu verfolgen und aufzuheben; so hat es denselben doch nicht gelingen wollen, ihn zu erreichen u. zu cooperiren, sondern er ist glücklich mit einer großen Anzahl erbeuteten Pferden und einer guten Summa Geldes, welche er jenzeit der Elbe an Contribution d. Quart. den den 14. Jul. im Königl. Lager vor Königsmarck einsetzte. Aus der Schlachten wird berichtet, daß Sr. Exell. der Hr. General-Lieutn. v. Nassau sein Quartier zu Neustadt genommen habe. Die ungarische Wälg hat auf Annäherung unserer Truppen Dapeln als einen unbalstbaren Ort verlassen u. sich weiter zurück gezogen.

Wien d. n. 30. Junii.

In vorangem. Semestri haben Ihre Kaiserl. Majest. eine Militair-Promotion vorgenommen: Der Fürst Wengel v. Sichtenheim ist Feldmarschall. Bey der Infanterie sind General-Feldzeugmeister, die Herren Generals Wallis, Wotta, (der nicht gestorben, wie man falsch berichtet.) Dammig, Chancios, Daun, Fürst Salza, Browne, Schulenburg und Pallavicini. Feldmarschall-Lieutenant: Grafen v. Dungenroth, Karl, General-Major, die Obersten Wengel, Maroull, Vinolpbus, Stägerberg, S. Andre, Wurmbrand, Elberfeld, Lujas, Cold, Sinter, Caproni, Erös, Lujan. Bey der Cavallerie, General der Cavallerie, Graf Solaburg. Feldmarschall-Lieutenant

37'

/3»»o y ^ f . Montags den 19. M . M. 54.

Loc 1

0%

W ilustracji wykorzystano czasopismo Schlesische Privilegirte Staats- Kriegs- und Friedens-Zeitung, które jest jednym z najcenniejszych i najważniejszych periodyków śląskich, ukazującym się nieprzerwanie przez ponad 200 lat we Wrocławiu.

<http://www.bibliotekacyfrowa.pl/publication/28062>

Breßlau brn 18. Julii« sein Quartier zu Neustadt
genommen habe.

Ohnnachtet die Feinde nach ihrem eignen Vi«
Unqari.che Militz ha« auf Annäherung Geständniß den
Obersten Bartoszewitz mit sei« unsrer Truppen Oppeln
als einen unhaltbaren Pulck Tartarn u den Oberst-L.
Desoffl mit ren Ort verlassen, u. sich weiter zurück
gezogen.

6aa Husaren ausgeschicket, welche noch über, W i m
h»n zc?. I u n i i .

tzem durch lc»c»Mann teutsche Caoallme un- An
oergangenm sonntag haben I h r o Kö-

ter dem ComMando des Obersten, Grafen von nigl.
Majest. etne Müitair Promotion porge-Kollowrath,
unterstützet worden, um dm H n . nomlnen: Der Fürst
Wentzel v. Lichtenstei»

Obltst Lieutn. v. Schütz, welcher m,t seinen ist
Feldmarschall. Bey der Infanterie find unterhabenden
Husaren, das gantze Land General-Feld Zeugmeisters,
die Henen Gene-durchst.elft« uno in C^ntribuon
setzte, u ver. rals Wallis. Pott«, (der nicht gestorben,
wie folgmond aufzuheben; so hat es denenselben man
falsch berichtet,) Dammtz, Chanclos, hoch nicht
gelingen wollm, ihn zu erreichen u. Saun, Fürst Salm,
Proune, Schulenburg zu coupiren, sondern ei «st

W przypadku czasopism istotne jest właściwe oznaczenie bloków tekstu aby zachować prawidłowy przepływ tekstu w szpaltach. Automatyczna analiza układu strony nie zawsze umożliwia uzyskanie oczekiwanego efektu, co jest szczególnie widoczne w przypadku starszych czasopism.

Ilustracja prezentuje warstwę tekstową, otrzymaną w wyniku zastosowania automatycznej analizy układu strony. OCR w takiej postaci może być przydatny w formatach hybrydowych (np. PDF, DjVu), w których warstwa graficzna wymusza rozkład warstwy tekstowej, ale w formatach mobilnych (np. ePUB, MOBI) jest bezużyteczna.

glücklich ma eine« und Pallapicini.
Feldmarschall, Lieutnants; grossen Anzahl etdeu, «ten
P, erben und emer Helfreich, Dungern,
Keul. Gendrab Majors, guten Summa Gtlo«s welche er
jenseit der die Obersten, Wetzet, Maroulll, Vmulphus,
Elde an Co!tt«ounonlyg«trl. ben, de», »4huj.
Vtahremberg, S. Andre, Wutmbrand Elbe«-

im Königl. L'gervor Komgizingratz eingetrof« selb,
Cujas/Cosa, Sine«r, Caplianl. EroK?u-fen. Aus Odir
Schlesien wird deichttet, daß zan. Bey der Capalleri.,
General der Cavql-

ße. H n t l. d«, Hr. General-Lieutn. v. Naffa« fttie, Glsf
Gallburg. Feldmalich«ll, Lleu«-

nsh

W celu uniknięcia efektu prezentowanego na ilustracji konieczne jest ręczne oznaczanie układu treści przed wykonaniem OCR. Ta czynność wydłuża cały proces kilkadziesiąt krotnie. Fakt ten zmusza instytucje zajmujące się udostępnianiem zdigitalizowanych treści do postawienia sobie pytania:

Czy produkować publikacje cyfrowe szybciej ale na krótszą metę, czy wolniej, ale z uwzględnieniem oczekiwań czytelnika, który za kilka lat coraz częściej będzie korzystał z treści cyfrowych przy pomocy urządzeń mobilnych?

Dlaczego możliwość linkowania do obiektów jest ważna?

1. Ułatwia “komunikację cytowań”.
2. Ma znaczenie w wymiarze promocyjnym.
3. Zewnętrzne linki kierujące do systemu BC mają pozytywny wpływ na pozycję w wynikach wyszukiwań.

W Internecie jest tak, że na wyniki wyszukiwań powaga państwowej instytucji kultury nie ma żadnego znaczenia. Znaczenie ma wiarygodność źródła, dobrze przygotowana publikacja cyfrowa, stałe i nie zmienne miejsce dostępu do obiektu z przejrzystą konstrukcją adresów internetowych. To te elementy powinny być gwarantowane powagą instytucji kultury.

Linkowanie do obiektów cyfrowych w BC

Treści graficzne (nie rozpoznawalne w formie tekstu - OCR) nie są wyeksponowane, jeśli nie staną się przedmiotem unikalnego opisu. Można spotkać przykłady pewnego rozdrobnienia publikacji na mniejsze części w celu wyeksponowania jej istotnych stałych elementów np. widoki:

<http://www.bibliotekacyfrowa.pl/publication/37873>

wówczas przytaczanie ich w sieci jest proste.

Przykłady adresów wiodących do konkretnych podstron publikacji

Format PDF:

http://www.bibliotekacyfrowa.pl/Content/37367/Odra_1961_05.pdf#page=39



- link do rekordu publikacji



- publikacja



- podstrona

Przykłady adresów wiodących do konkretnych podstron publikacji

Format DjVu:

<http://www.bibliotekacyfrowa.pl/Content/14302/directory.djvu?djvuoops&page=102>



- link do rekordu publikacji



- publikacja



- podstrona

Przykłady adresów wiodących do konkretnych podstron publikacji

Format Zoomify +



[http://www.bibliotekacyfrowa.pl/Content/29259/2449-IV.B.htm?
z=100&x=5208&y=4337&p=2449-IV.B_01](http://www.bibliotekacyfrowa.pl/Content/29259/2449-IV.B.htm?z=100&x=5208&y=4337&p=2449-IV.B_01)



- link do rekordu publikacji



- publikacja



- podstrona wraz z koordynatami wybranego miejsca obrazu

Dziękuję za uwagę

Tomasz Kalota



Biblioteka Uniwersytecka we Wrocławiu
Digitalizacja.pl