JAKUB MACIEJ ŁUBOCKI

STARE PRZESTRZENIE INFORMACJI. CO TRACIMY NA NIEWŁAŚCIWEJ AUTOMATYZACJI POCHODNYCH ŹRÓDEŁ INFORMACJI?*

Poglądy na temat relacji między terminami „automatyzacja”, „digitalizacja” i „cyfryzacja”. Struktura pochodnych źródeł informacji. Zjawiska towarzyszące automatyzacji pochodnych źródeł informacji. Straty powstałe przy niewłaściwej automatyzacji pochodnych źródeł informacji, ich skutki i sposoby zapobiegania im.

SŁOWA KLUCZOWE: wydawnictwa informacyjne, pochodne źródła informacji, automatyzacja bibliografii, automatyzacja słowników, automatyzacja encyklopedii

1. UŚCIŚLENIE DEFINICYJNE

Rozważania na temat zaprezentowany w tytule musi poprzedzić objaśnienie, dlaczego zdecydowano się na termin „automatyzacja”, mimo że przez niektórych jest on uważany za archaiczny. Analiza usytuowania jego znaczenia w kontekście innych wyrażen, znaczeniowo bardzo bliskich lub nawet sprawiających wrażenie synonimów, wykaże, że wciąż jest najbardziej adekwatny, a tym samym nadal zachowuje swoją użyteczność.

Pierwotnie tytuł miał zapowiadać tekst o stratach spowodowanych niewłaściwą digitalizacją źródeł. Jednakże, wedle *Słownika wyrazów obcych PWN*, jest to „nadawanie postaci cyfrowej danym pisany i drukowanym, zawartym na nośnikach magnetycznych lub innych”¹, a więc termin ten nie obejmuje źródeł o charak-

* Tekst jest poszerzoną wersją referatu wygłoszonego pod tym samym tytułem na III Ogólnopolskiej Konferencji Naukowej „Nowe przestrzenie informacji” (Wrocław, 21 kwietnia 2016) zorganizowanej przez Koło Naukowe Bibliotekoznawców Uniwersytetu Wrocławskiego. Pragnę tu podziękować dr Elżbiecie Herden, dr hab. Małgorzacie Góralskiej oraz drowi Witoldowi Ucherkowi za cenne wskazówki, jakich udzielili mi w trakcie jego opracowywania.

¹ *Słownik wyrazów obcych PWN z przykładami i poradami*. Oprac. Lidia Drabik, Warszawa: Wydawnictwo Naukowe PWN 2011, s. 201.

terze *born digital*. Ten sam słownik proponuje cyfryzację jako synonim digitalizacji², i była to kolejna z możliwości. Jednak wokół pojęcia cyfryzacji narosło nieco wątpliwości. Ponieważ nie jest to jeszcze termin, który doczekał się ugruntowanego znaczenia, bywa on stosowany w sposób dowolny, a przez to staje się nieostry. Stąd można spotkać głosy, że w istocie nie jest to synonim digitalizacji. Marek Zieliński, na podstawie analizy programu „Polska Cyfrowa”, wprowadzonego w 2014 roku przez ówczesne Ministerstwo Administracji i Cyfryzacji (*sic!*) oraz Ministerstwo Rozwoju Regionalnego, zauważa, że w programie tym cyfryzacja jest rozumiana „szeroko jako działania w celu zwiększenia dostępności Internetu i jego zasobów dla obywateli i wbudowanie mechanizmów elektronicznych w administrację państwową. Nie ma to więc nic wspólnego z digitalizacją”³ i w tym rozumieniu mamy właściwie do czynienia z informatyzacją⁴, która choć jest przedmiotem zainteresowania cyfryzacji, to jednak — jak zauważa Dominik Kazanowski — same komputery i sieci, a więc technologia, nie stanowią jeszcze cyfryzacji, podobnie jak digitalizacja nie ogranicza się do samego procesu skanowania. „Cyfryzacja, w odróżnieniu od pochodzącego od angielskiego »digital« terminu »digitalizacja«, ma wymiar holistyczny. Posiada także cechy doktryny. Cyfryzacja to łączne określenie na komputeryzację, informatyzację i usieciowienie”⁵. Zasadnicze rozróżnienie między digitalizacją a cyfryzacją znajduje się więc w podejściu do możliwości odbioru treści — digitalizacja to równoległość nośnika tradycyjnego i cyfrowego gotowego do odbioru: o ile nie wiecznie, to przynajmniej przez jakiś czas. „W tym sensie, [treści] są w dalszym ciągu dostępne dla człowieka. Człowieka nieuzbrojonego w narzędzia. Człowieka, który wyłącznie za sprawą swoich zmysłów jest w stanie odbierać rzeczywistość. Z cyfrowymi obrazami rzeczywistości jest już inaczej. Nie mamy do nich dostępu bez pośrednictwa narzędzi”⁶.

Z drugiej strony, Wanda Krzemińska swoją wnikliwą analizą językoznawczą⁷, badającą możliwości słowotwórcze oraz rzeczywiste stosowanie obu termi-

² *Ibidem*, s. 161.

³ Marek ZIELIŃSKI, *Digitalizacja a cyfryzacja*, 17 XII 2013, http://www.archiwistyka.pl/artykuly/artykuly_i_felietony/528 [dostęp: 17 IV 2016].

⁴ Taka też nazwa tego urzędu byłaby adekwatniejsza, co zresztą przyznał sam prezes Rady Ministrów, który we własnym rozporządzeniu zarządził: „Minister [Cyfryzacji — przyp. J.M.Ł.] kieruje działem administracji rządowej — informatyzacja”, rozporządzenie Prezesa Rady Ministrów z dnia 9 grudnia 2015 r. zmieniające rozporządzenie w sprawie szczegółowego zakresu działania Ministra Cyfryzacji, Dz.U. 2015 poz. 2090, § 1. Od czasu powołania pierwszego ministerstwa z „cyfryzacją” w nazwie (co zostało zarządzone rozporządzeniem Rady Ministrów z dnia 21 listopada 2011 r. w sprawie utworzenia Ministerstwa Administracji i Cyfryzacji, Dz.U. 2011 Nr 250, poz. 1501) dział ten znajdował się w jego gestii.

⁵ [Dominik KAZANOWSKI], *Definicja cyfryzacji*, 17 V 2014, <http://networkeddigital.com/2014/05/17/definicja-cyfryzacji/> [dostęp: 17 IV 2016].

⁶ *Ibidem*.

⁷ Wanda KRZEMIŃSKA, *Co z tą cyfryzacją?*, [w:] *Cyfryzacja w procesach komunikowania*. Pod red. Wandy Krzemińskiej i Piotra Nowaka, Poznań: Sorus Wydawnictwo i Drukarnia 2004, s. 11–23.

nów w sieci, dowiodła, że językowo są to terminy równoznaczne i w dziedzinie nowych technologii definiują dokładnie te same treści. W związku z tym oba terminy w naturalny sposób konkurują z sobą i analiza rzeczywistego materiału językowego dokonana przez Krzezińską wykazała pewną przewagę digitalizacji nad cyfryzacją: na korzyść digitalizacji przemawia jej regularność słowotwórcza, spolszczenie zapisu, przyswojenie wszystkich możliwych form, potwierdzone powszechnym użyciem w języku pisanym i mówionym oraz nieregularność, wariantowość i nieporęczność niektórych form tworzonych od cyfryzacji. Jednakże cyfryzacja, ponieważ pochodzi od zakorzenionego już od stuleci wyrazu „cyfra”, dla użytkownika języka polskiego jest bardziej naturalna i zintegrowana z zasobem nazewniczym, a tym samym rekonstrukcja jej znaczenia jest dla niego prostsza aniżeli w zapożyczeniu angielskim, jakim jest digitalizacja; ponadto od dawna słowa pochodne od wyrażenia „cyfra” służą do opisu nowych technologii⁸. Doprowadziło to do polaryzacji znaczeń i użycie obu terminów, usankcjonowanej praktyką językową użytkowników: jeśli sytuacja wymaga użycia formy rzeczownikowej, preferowana jest „digitalizacja” (a nie „cyfryzacja”) i stosunek ten wyraża się w proporcji 12 : 1; jeśli zaś formy przymiotnikowej — preferowany jest „cyfrowy” (a nie „digitalny”) w proporcji 1572 : 1⁹. Krzezińska w konkluzji pisze więc o „schizmie terminologicznej”, zauważa, że nie jest to sytuacja codzienna w języku polskim i może prowadzić do wielu utrudnień, jednakże choć jest to językowa niekonsekwencja, to w jakiś sposób funkcjonalna i ekonomiczna językowo.

Na zakończenie tego przeglądu stanowisk można jeszcze przytoczyć opinię Małgorzaty Góralskiej, która zauważa, że „digitalizacja odnosi się do zamiany tradycyjnych materiałów bibliotecznych (rękopiśmiennych i drukowanych) na postać cyfrową, natomiast zamiana analogowych przekazów audialnych i audiowizualnych na postać cyfrową częściej nazywana jest cyfryzacją — stąd ostatnio przechodziliśmy w Polsce proces cyfryzacji telewizji”¹⁰. Byłoby więc to rozróżnienie czynnościowe i w pewnych okolicznościach terminy te można uważać za synonimy, lecz nie w kontekście fachowego lub specjalistycznego użycia.

Niezależnie jednak od tego, do którego ze stanowisk się przychylimy, oba pojęcia wykluczają w swoim znaczeniu pewne formy dokumentów (w podejściu synonimicznym oba wykluczają wszystko to, co nie było zamianą sygnału analogowego na cyfrowy; w podejściu antonimicznym digitalizacja nie dotyczy dokumentów, których jedyna postać to oryginalna postać cyfrowa, a cyfryzacja — których pierwotna postać to postać tradycyjna), które miały zostać w niniejszych rozważaniach uwzględnione i stąd nie mogły znaleźć się w tytule. Ten problem nie

⁸ *Ibidem*, s. 14–16.

⁹ *Ibidem*, s. 20.

¹⁰ Z listu elektronicznego od M. Góralskiej z dnia 21 kwietnia 2016 roku (w archiwum autora).

występuje przy automatyzacji, albowiem — zgodnie z definicją — „mechanizacja i automatyzacja procesów informacyjnych” jest to

działalność mająca na celu odciążenie człowieka od wykonywania manualnie operacji procesu informacyjnego, zwiększenie jego możliwości w tym zakresie poprzez stosowanie maszyn i urządzeń wykonujących poszczególne polecenia człowieka (mechanizacja) lub zaprogramowanych w sposób pozwalający im na alternatywne wybieranie następnych czynności lub szeregu takich czynności (automatyzacja)¹¹.

A ponieważ w niniejszym tekście miały być uwzględnione wszelakie źródła informacji, w których człowiek nie jest jedynym czynnikiem aktywnym w procesie wyszukiwania¹², to właśnie automatyzacja okazała się najwłaściwszym określeniem. Pewną pomocą w zorientowaniu się w relacjach zachodzących między wszystkim podanymi tu pojęciami może być tabela 1.

Tab. 1. Zależności między pojęciami opisującymi procesy technologiczne i kulturowe zachodzące w trakcie rozwoju społeczeństwa informacyjnego według D. Kazanowskiego

społeczeństwo industrialne	działanie	umaszyno- wienie	rutynizacja	uprocесо- wienie	bezreflek- syjność	sztuczna re- fleksyjność	uniwersa- lizm
	technolo- gia	mechanizacja	komputery- zacja	informaty- zacja	automaty- zacja	algorytmi- zacja	społе- czeństwo cyfrowe
	kultura	globalizacja	digitalizacja	cyfryzacja	usiecio- wienie	wirtualiza- cja	
	podmiot	człowiek	człowiek	człowiek/ maszyna	maszyna	maszyna	?

Źródło: opracowanie własne na podstawie: [Dominik KAZANOWSKI], *Definicja cyfryzacji*. Adres dostępu do samej grafiki: http://networkeddigital.com/wp-content/uploads/2014/05/cyfryzacja_artificial_reflexivity.png [dostęp: 17 IV 2016].

¹¹ *Encyklopedia współczesnego bibliotekarstwa polskiego*. Pod red. Karola Głombiowskiego, Bolesława Świdorskiego i Heleny Więckowskiej, Wrocław: Zakład Narod. im. Ossolińskich — Wydawnictwo 1976, s. 201. Wyróżnienia własne.

¹² Co ma swoje dobre i złe strony, ponieważ — jak pisze Lev Manovich w swojej rozprawie o nowych mediach, tj. mediach, wszystkich które znamy, przełożonych na język danych numerycznych zrozumiałych przez maszyny — „postać liczbowa i modularność nowych mediów umożliwiają zautomatyzowanie wielu czynności [...]. W ten sposób — chociaż częściowo — można z procesu twórczego wyeliminować intencjonalność”, Lev MANOVICH, *Język nowych mediów*, Warszawa: Wydawnictwa Akademickie i Profesjonalne 2006, s. 97. A przecież właśnie intencjonalność jest jedną z podstawowych cech informacji, według Juliana Gałczyńskiego — immanentną: „Uznaje się za pewnik, że twórcą/emiterem i odbiorcą informacji jest człowiek i że ma ona charakter intencjonalny. [...] Jest to] cecha właściwa wszystkim rodzajom informacji — oddzielająca ją od oznak (adela) i informacji biologicznej”, Julian GAŁCZYŃSKI, *Holistyczne ujęcie informacji. Cz. 1. Założenia wstępne*, Praktyka i Teoria Informacji Naukowej i Technicznej (14) 2006, nr 4, s. 5, 7. Można więc zadać sobie pytanie: czy przypadkiem nie jest tak, że im bardziej źródło jest automatyczne, tym paradoksalnie mniejszy jest jego potencjał informacyjny? O tym, jak istotny jest aspekt człowieczeństwa w informacji, zob. też s. 342 (przypis 69) niniejszego tekstu.

Tabela przedstawia kolejne etapy przechodzenia od społeczeństwa industrialnego do społeczeństwa cyfrowego. Według Dominika Kazanowskiego pierwsze trzy etapy są elementami okresu rekonfiguracji starego ładu, a następne trzy — programowania społecznego, które w efekcie finalnym doprowadzi nas do „jednego, uniwersalnego społeczeństwa cyfrowego”, posługującego się jednym, wspólnym, maszynowym językiem, służącym do komunikacji w relacjach człowiek–człowiek oraz człowiek–maszyna.

2. PRZEDMIOT BADAŃ

W niniejszym artykule przedmiotem rozważań będą zatem zautomatyzowane (umożliwiające alternatywne drogi przeszukiwania) źródła informacji o charakterze pochodnym. Spośród różnych typów źródeł tego rodzaju wzięte pod uwagę będą słowniki ogólne językowe, bibliografie/katalogi oraz encyklopedie, a więc reprezentujące zarówno źródła skierowujące, jak i faktograficzne, i tylko w ten wybiórczy sposób będą rozumiane tu jako pochodne źródła informacji (PZI). Dokumenty tego rodzaju pełnią ważne zadanie w procesie naukowym, ponieważ „są źródłem informacji o dokumentach pierwotnych, o treści dokumentów pierwotnych lub źródłem informacji o zagadnieniu omówionym w jednym lub wielu dokumentach pierwotnych”¹³. W ten sposób w ogóle możliwa jest kumulacja wiedzy naukowej przez budowanie nowej myśli na podstawie wcześniejszych odkryć i doświadczeń, a także w dużej mierze wyeliminowany zostaje problem dochodzenia do już znanej wiedzy. Jednak aby było to możliwe, PZI muszą być funkcjonalne. W jaki sposób się tego dokonuje?

PZI podają informację w sposób przetworzony, to jest opracowaną według ustalonych zasad oraz w układzie dostosowanym do potrzeb użytkownika i najpełniej zaspokajającym jego potrzeby, w związku z czym niezwykle ważna jest właściwa segmentacja ich treści, przekładająca się na kolejne poziomy struktury źródła i ułatwiająca sprawne wyszukiwanie odpowiednich informacji. W ten sposób możemy zaobserwować, że źródła tego rodzaju mają mikrostrukturę¹⁴,

¹³ Wojciech PIRÓG, *Zagadnienia informacji i dokumentacji naukowej*, Warszawa: Państwowe Wydawnictwo Naukowe 1977, s. 24.

¹⁴ W leksykografii rozumianą jako: „Wewnętrzna konstrukcja artykułu hasłowego. W przeciwieństwie do ogólnej listy haseł (makrostruktury), mikrostruktura zawiera szczegółowe informacje na tematżądanego hasła, komentujące jego właściwości formalne i semantyczne (pisownia, wymowa, gramatyka, definicja, użycie, etymologia). Jeśli hasło ma więcej niż jedno znaczenie, informacja jest podana dla każdego z nich (podhasła). Słowniki różnią się w zależności od ilości informacji, którą podają i jak ją przedstawiają w tekście artykułu. Użytkownicy mogą nie mieć wystarczających umiejętności do zrozumienia zawartości mikrostruktury i mogą potrzebować wyraźnych wskazówek i/lub instrukcji, aby odnaleźć i wyodrębnić potrzebne elementy” (przeł. J.M.Ł., w oryg.: „The internal design of a reference unit. In contrast to the overall word-list (macrostructure) the microstructure provides detailed information about the headword, with comments on its

makrostrukturę¹⁵ i megastrukturę¹⁶. Obserwacje te pierwotnie zostały wypracowane na gruncie leksykografii (stąd specyficzne, słownikowe, nasączenie przytoczonych tu definicji), można jednak efektywnie przełożyć je na PŻI w ogólności. Mikrostruktura byłaby to wówczas wewnętrzna konstrukcja zasadniczej jednostki (artykułu hasłowego, który możemy rozumieć jako opis jednostki leksykalnej w słowniku językowym, opis desygnatów nazw w encyklopedii lub opis bibliograficzny dokumentu w bibliografii/katalogu) skomponowana według jednolitych, z góry przyjętych zasad, podająca według ścisłego i stałego porządku informacje skierowujące lub faktograficzne dotyczące słów, nazw lub dokumentów. Po skompletowaniu wszystkich jednostek przewidzianych dla źródła będzie ono podlegało nowemu procesowi komponowania na wyższym poziomie. W ten sposób uzyskamy makrostrukturę, czyli nadrzędny układ o charakterze płytkim lub wielostopniowym, porządkujący według z góry zdefiniowanego klucza wszystkie jednostki zasadnicze. Ten układ (właściwy słownikom, encyklopediom lub bibliografii) wraz z treścią jednostek (czyli tekstem głównym¹⁷) oraz materiałami uzupełniającymi (tablice, ilustracje, aneksy, przypisy) stanowiłby zrąb główny źródła. Po dopełnieniu go wszelkimi potrzebnymi tekstami pobocznymi, jakimi są karty tytułowe, materiały wprowadzające (przedmowy, wstępy, noty wydawnicze, podziękowania, dedykacje, motta, posłowania) oraz materiały infor-

formal and semantic properties (spelling, pronunciation, grammar, definition, usage, etymology). If the headword has more than one sense, the information is given for each of these (sub-lemma). Dictionaries vary according to the amount of information they provide, and how they present it in the text of the entry. Users may not have sufficient reference skills to follow the intricacies of the microstructure, and may need explicit guidance and/or instruction to find and extract the details required”), Reinhard Rudolf Karl HARTMANN, Gregory JAMES, *Dictionary of Lexicography*, London: Routledge 2001, s. 94.

¹⁵ W leksykografii rozumianą jako: „Ogólna struktura listy haseł, która pozwala kompilatorowi i użytkownikowi na zlokalizowanie informacji w kompilowanym źródle informacji. Najpopularniejszym układem struktury w słownikach Zachodu jest lista w układzie alfabetycznym (choć istnieją inne sposoby układu haseł, np. tematyczny, chronologiczny lub frekwencyjny); lista ta stanowi zasadniczą część źródła i może być uzupełniona przez aparat pomocniczy z przodu, w środku lub na końcu źródła” (przeł. J.M.Ł., w oryg.: „The overall list structure which allows the compiler and the user to locate information in a reference work. The most common format in Western dictionaries is the alphabetical word-list (although there are other ways of ordering the headwords, eg. thematically, chronologically or by frequency), which constitutes the central component. This can be supplemented by outside matter in the front, middle or back of the work”), *ibidem*, s. 91.

¹⁶ W leksykografii rozumianą jako: „Ogólna rama części składowych źródła informacji, zawierająca makrostrukturę i aparat pomocniczy” (przeł. J.M.Ł., oryg.: „The overall framework of the component parts of a reference work, including macrostructure and the outside matter”), *ibidem*, s. 93.

¹⁷ Pojęcie „tekst główny”, a także pojawiające się dalej „materiały uzupełniające”, „zrąb główny” (właśc. „podstawowy zrąb dzieła”), „karty tytułowe”, „materiały wprowadzające”, „materiały informacyjno-pomocnicze”, a co za tym idzie — poszczególne rodzaje tekstów nimi oznaczane, podaję tu w rozumieniu wszystkich arkuszy Polskiej Normy PN-N-01222:1978 *Kompozycja wydawnicza książki*.

macyjno-pomocnicze (słowniki, wykazy, indeksy, spisy treści, komentarze, bibliografie), składające się na aparat pomocniczy, otrzymamy megastrukturę. Tyle stwierdzają leksykografowie — bibliolodzy dopowiedzieliby, że ta troista w swojej budowie struktura treściowa (zobrazowana na il. 1) powinna zostać jeszcze dopełniona strukturą materialną źródła (czyli perytekstem edytorskim — okładką, formatem, materiałem typograficznym, nakładem, umieszczeniem w serii wydawniczej¹⁸), ponieważ dzieło zawsze istnieje w określonej formie¹⁹. Dla tej najbardziej zewnętrznej warstwy można zaproponować określenie ultrastruktura²⁰; wówczas owe cztery koncentryczne warstwy złożą się na całość samoistnego PZI.

Il. 1. Struktura pochodnego źródła informacji

Źródło: opracowanie własne. Schemat jest znacznie poszerzoną, przekomponowaną i dostosowaną do polskiej terminologii oraz treści artykułu wersją: Rudolf Karl HARTMANN, Gregory JAMES, *Dictionary of Lexicography*, London: Routledge 2001, s. 92.

¹⁸ Za: Anna ŁACH, *Perytekst edytorski. Definicja, funkcje i strategie perswazyjne*, Zeszyty Wydziału Humanistycznego [Kolegium Karkonoskiego w Jeleniej Górze] 2010, nr 5, s. 253.

¹⁹ Krzysztof MIGON, *Nauka o książce wśród innych nauk społecznych*, Wrocław: Zakład Narod. im. Ossolińskich 1976, s. 52.

²⁰ Podążając za znaczeniem przedrostka *ultra-*: „poza, po drugiej stronie, poza zasięgiem, poza granicami” (*Słownik wyrazów obcych PWN. Wydanie nowe*. Pod red. Elżbiety Sobol, Warszawa: Wydawnictwo Naukowe PWN 1995, s. 1143), które odpowiada miejscu tej warstwy, będącej dosłownie „poza granicami” treści źródła (przyjmując, że jej granice stanowi megastruktura) i mającej charakter przede wszystkim materialny i/lub symboliczny.

W związku z tym analiza będzie odnosić się do zagadnień funkcjonalności struktury i organizacji zgromadzonego w PŹI materiału, a nie jego konkretnej treści (a tym samym jej relewancji, prawdziwości czy słuszności). W związku z tym, czerpiąc tym razem z dorobku terminologicznego nauki o bibliografii, mowa będzie o sprawności informacyjnej²¹ źródła, a nie jego wartości informacyjnej²². Uogólniając te terminy na całość PŹI, można powiedzieć, że ocena wartości informacyjnej to ocena doboru materiału do źródła i stopnia jego kompletności w ramach przyjętych założeń, a także stopnia prawdziwości i prawidłowości jego opisu (badania, czy na przykład słowniki nie zawierają błędów logicznych w definicjach, bibliografie nie podają danych niedokładnych, niesprawdzonych lub błędnych, a encyklopedie nie zawierają błędów rzeczowych lub kłamstw), natomiast ocena sprawności informacyjnej to ocena wyboru metody opracowania i konstruowania źródła oraz jednolitości jej zastosowania (a zatem funkcjonalności: czy łatwo, szybko i niezawodnie zapewnia ona użytkownikowi dotarcie do odpowiednich — z punktu widzenia jego potrzeb informacyjnych — jednostek i niesionych przez nie informacji), a także adekwatności sposobu komunikacji do nośnika.

3. PODSTAWOWE OBSERWACJE

Niezależnie od nośnika, na którym jest zapisane PŹI — czy to tradycyjnego, czy też elektronicznego — wszystkie poziomy struktury powinny wystąpić²³ oraz być sprawnie skonstruowane. Niestety, w wielu przypadkach tak nie jest. Podstawowym zjawiskiem charakteryzującym PŹI w środowisku cyfrowym jest prze-

²¹ „Sprawność informacyjna zależy od formy danego spisu bibliograficznego, tj. można o niej mówić w odniesieniu do konkretnego zbioru materiałów w tym spisie zawartych. Pod pojęciem sprawności informacyjnej spisu bibliograficznego rozumiem taką metodę opracowania bibliografii, która łatwo, szybko i niezawodnie zapewnia użytkownikowi dotarcie do odpowiednich — z punktu widzenia jego potrzeb informacyjnych — pozycji w danej bibliografii. Sprawność informacyjna zależy więc od metody opisu, tj. od tego, czy opis wystarczająco informuje o zawartości i ujęciu danej pozycji, bo sam tytuł — jak wiadomo — nie zawsze jest adekwatny w stosunku do treści; zależy także od budowy bibliografii, tj. rodzaju układu, trafności przydziałów rzeczowych (m.in. zapewnionej także przez stosowanie opisów wielokrotnych i odsyłaczy), od istnienia i rodzaju indeksów względnie ich braku”, Henryk SAWONIAK, *Sprawność informacyjna bibliografii polskich w XIX wieku (do 1918 r.)*, Przegląd Biblioteczny (39) 1971, nr 1–4, s. 155.

²² „O wartości informacyjnej bibliografii decyduje przede wszystkim dobór materiałów. Jeśli autor bibliografii pominie ważne materiały, braku tego nie zrównoważy najstaranniejsza metoda opracowania. Tym większa jest wartość informacyjna, im bardziej zbliża się autor do kompletności w zebraniu potrzebnych i ważnych materiałów — w ramach przyjętych przez siebie założeń”, *ibidem*, s. 155.

²³ Wbrew pozorom nawet w PŹI na nośnikach elektronicznych mamy do czynienia z przynajmniej niektórymi elementami ultrastruktury. Jest to na przykład materiał typograficzny czy podobizna okładki.

sadne rozbudowanie mikrostruktury źródła przy jednoczesnym zaniedbaniu lub w ogóle pominięciu pozostałych poziomów struktury. Owo rozbudowanie przejawia się w skrupulatnym posegmentowaniu i oetykietowaniu elementów składowych treści jednostek zasadniczych skutkujących drobiazgową granulacją danych, zobrazowaną na ilustracjach 3 i 4. Należy podkreślić, że granulacja sama w sobie nie jest zjawiskiem niekorzystnym, natomiast jeśli odbywa się kosztem makro- i megastruktury, wówczas jest to sytuacja niepożądana. W niej użytkownik z poziomu źródła jako dokumentu przechodzi od razu do jednostkowych danych, bez wsparcia i przewodnictwa istniejących do tej pory takich elementów, jak układ zrębu głównego oraz aparat pomocniczy. Mówiąc potocznie, jest rzuwany od razu na głęboką wodę otchłani niezliczonych informacji, bez żadnego przewodnika prowadzącego od ogółu do szczegółu. Jeśli do tego dochodzi, istnieje wysokie

brak *m III, D. braku 1. bfm* «fakt nieistnienia czegoś, nieobecność, niedostatek» □ *B. czego (nie: na czym):* Brak czasu. Odczuwał brak sił. Mimo braku słońca opaliła się nieco. △ *Z braku kogoś, czegoś:* Z braku powietrza mało nie zemdlął. △ *W braku kogoś, czegoś:* W braku laku dobry i opłatek (przysłowie). △ *niepoprawne* W naszym osiedlu występuje brak (*zamiast:* brakuje) placówek handlowych. 2. «wada, skaza, błąd, niedociągnięcie; wyrób z usterkami» □ *Braki w czym:* Mieć braki w wykształceniu. Produkować braki.

HASŁO	coachka
WYMOWA	kojczka
DEFINICJA	«kobieta zajmująca się <i>coachingiem</i> »
PRZYKŁADY UŻYCIA	Z <i>coachką</i> spotykałam się raz w miesiącu przez pół roku (www.biblioteki.org).
	16 maja zostałam kobietą-matką. Z premedytacją piszę „kobietą-matką”, a nie „matką”. Zawdzięczam to określenie mojej koleżance <i>coachce</i> (celestyna.pl).
WIZUALIZACJE	Zobacz w grafice Google >
ODSYŁACZE	<od <i>coach</i> >
DATA	Rejestracja hasła: 01.12.2015, 14.23, ostatnia aktualizacja hasła: 03.12.2015, 23.47
AUTOR	Jerzy Buczek
REDAKTOR	Jan Burzyński

II. 2. a i b. Porównanie stopnia granulacji danych w słownikach ogólnych języka polskiego — forma drukowana vs. elektroniczna zautomatyzowana — na przykładzie artykułów hasłowych „brak” i „coachka”. Na ilustracji 2a. widać maksymalną kondensację treści artykułu hasłowego — dla niewprawnego użytkownika nie wszystkie informacje mogą być w nim zrozumiałe. Bardziej intuicyjny sposób prezentacji treści oferuje układ ukazany na ilustracji 2b. Warto zwrócić uwagę na pola, które nie występują w formie drukowanej (data rejestracji, wizualizacje).

Źródło: *Słownik wyrazów obcych PWN. Wydanie nowe*. Pod red. Elżbiety Sobol, Warszawa: Wydawnictwo Naukowe PWN 1995, s. 54; *Obserwatorium Językowe Uniwersytetu Warszawskiego. Najnowsze słownictwo polskie*. Pod red. Mirosława Bańki, Macieja Czeszewskiego i Jana Burzyńskiego, <http://nowewyrazy.uw.edu.pl/haslo/coachka.html> [dostęp: 23 VI 2016].

SIWECKA, Dorota

Światowy model informacji bibliograficznej : programy i projekty (1950-2010) / Dorota Siwecka ; Stowarzyszenie Bibliotekarzy Polskich. – Wrocław ; Warszawa : Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 2015. – 355, [1] s. : il. ; 24 cm. – (Nauka, Dydaktyka, Praktyka = Science, Didactics, Practice ; 162)

Tyt. równol.: A world model of bibliographic information : programmes and projects (1950-2010). – Spis treści także ang.

Bibliogr. s. 297-339. Indeks.

Współfinansowanie: Uniwersytet Wrocławski. Instytut Informacji Naukowej i Bibliotekoznawstwa

ISBN 978-83-64203-48-0 : zł 48

1. Bibliografia – teoria 2. Informacja naukowa – historia

01-027.21

001.102

b50986806

2016/2744

NR_BBN	PB 2016/2744
AUT:	Siwecka, Dorota
TYTUŁ/ODP:	Światowy model informacji bibliograficznej : programy i projekty (1950-2010) / Dorota Siwecka ; Stowarzyszenie Bibliotekarzy Polskich.
WYD:	Wrocław ; Warszawa : Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 2015.
OBJ:	355, [1] s. : il. ; 24 cm.
SER:	Nauka, Dydaktyka, Praktyka = Science, Didactics, Practice ; 162
UW1:	Tyt. równol.: A world model of bibliographic information : programmes and projects (1950-2010).
DOD:	Bibliogr. s. 297-339. Indeks.
PRZEDMIOT:	.4 1901-2000
PRZEDMIOT:	.4 1945-1989
PRZEDMIOT:	.4 1989-2000
PRZEDMIOT:	.4 2001-
PRZEDMIOT:	Bibliografia - teoria
PRZEDMIOT:	Informacja naukowa - historia
991	PB
UKD_BN	01-027.21
UKD_BN	001.102

II. 3. a i b. Porównanie stopnia granulacji danych w „Przewodniku Bibliograficznym” — forma elektroniczna typograficzna vs. elektroniczna zautomatyzowana — na przykładzie opisu bibliograficznego nr 2016/2744 (Dorota SIWECKA, *Światowy model informacji bibliograficznej*, Warszawa: Wydaw. SBP 2015). Wyraźnie widać, że na ilustracji 3b. ten sam opis bibliograficzny zajmuje więcej miejsca, lecz każdy z jego kolejnych elementów ma swoją etykietę. Niestety, prymitywne środowisko technologiczne tej bazy (skrótowość etykiet ociera się o absurd) sprawia, że — paradoksalnie — czytelniejszy jest opis przedstawiony na ilustracji 3a.

Źródło: Przewodnik Bibliograficzny. Urzędowy Wykaz Druków Wydanych w Rzeczypospolitej Polskiej 2016 (72(84)), nr 6, s. 2, <http://www.bn.org.pl/download/document/1455639531.pdf> [dostęp: 23 VI 2016]. *Przewodnik Bibliograficzny (MARC 21, 1973–)*, <http://mak.bn.org.pl/cgi-bin/makwww.exe?BM=23&IM=01&WI=SIWECKAbDOROTA&NU=01&DD=1> [dostęp: 23 VI 2016].

ryzyko, że wprawdzie jakieś informacje zostaną odnalezione, lecz nie wiadomo, w jakim stopniu relewantne do zapytania (czy są to wszystkie możliwe informacje do uzyskania?, czy są adekwatne?, czy są najlepsze jakościowo?). W ten sposób PŻI, rozbudowując możliwości wyszukiwawcze, zaniedbują walor informacyjny, a przecież jako elementy systemu informacyjno-wyszukiwawczego (SIW) powinny jednakowo zapewniać dostęp do obu tych aspektów.

Oprócz granulacji danych doświadczamy jeszcze innych problemów niewłaściwej automatyzacji pochodnych źródeł informacji. Zachłyśnięcie się nowoczesnymi technologiami sprawiło, że z jednej strony porzucamy ogromny walor informacyjny płynący z nadawania wyższych stopni struktury, do niedawna wymuszanych przez dotychczasowy nośnik tego typu źródeł (druk), z drugiej jednak — nie wykorzystujemy możliwości, jakie tkwią w tych technologiach, nie do urzeczywistnienia w drukowanej formie. Prowadzi to do dwóch zjawisk

obserwowanych w zautomatyzowanych PŻI. Pierwszym z nich jest googlizacja obsługi źródła²⁴. Jak pisze James Gleick,

w latach 90. młoda technologia wyszukiwarek popadła w ortodoksję. Myślano odruchowo w kategoriach tradycyjnych technik organizowania światowej informacji (encyklopedie, słowniki). Zdawano sobie sprawę, że porządek alfabetyczny zaczyna tracić znaczenie, ale za późno dostrzeżono, jak dynamicznym i nieuchwytnym fenomenem jest internet. [...] Google to zupełnie inna strategia — jego interface był pozbawiony reklam: nic, tylko okienko, gdzie użytkownik może wpisać pytanie, i przycisk szukaj w Google²⁵.

To właśnie obserwujemy dziś wśród automatyzowanych źródeł informacji, niezależnie od ich rodzaju: dążenie do upodobnienia strony startowej PŻI do strony startowej Google przez ascetyczny interfejs z dominującym paskiem wyszukiwania i ledwie sygnalizowanymi możliwościami korzystania, innymi niż wpisanie „czegokolwiek” do paska i obserwowanie, jakie odpowiedzi zostaną zwrócone. Wystarczy przyrzeć się stronom startowym bibliografii, katalogów, słowników oraz encyklopedii i porównać ich wygląd z wcześniejszymi wersjami, by doskonale zaobserwować ten proces (il. 4–5).

Multimedialny słownik angielskiego online

Wpisz angielskie lub polskie hasło

Il. 4. Przykład googlizacji w zautomatyzowanym słowniku ogólnym językowym

Źródło: Diki. Multimedialny słownik angielskiego online, www.diki.pl [dostęp: 17 V 2016].

²⁴ Którą tu rozumiem się technicznie jako upodobnianie się interfejsów różnych stron do interfejsu strony startowej wyszukiwarki Google (Google Polska, www.google.pl [dostęp: 16 V 2016]). Najczęściej jednak ten neologizm używany jest w aspekcie socjologicznym czy filozoficznym na określenie ogółu przemian w ekonomii, zarządzaniu i organizacji informacji oraz związanych z nimi obaw (etyczność szukania powiązań na podstawie historii wyszukiwań i perswazyjnego personalizowania odpowiedzi wyszukiwarki, dominacja w skali globalnej jednego ponadnarodowego gracza na rynku serwisów internetowych, bombardowanie jednostki informacjami). Więcej na temat tak ujętej googlizacji można znaleźć w książce o znamienym tytule, który Zbigniew Mach tłumaczy jako *W sercu Googlepleksu. Jak Google myśli, działa i kształtuje nasze życie*: Steven LEVY, *In the Plex. How Google thinks, works, and shapes our lives*, New York: Simon & Schuster 2011.

²⁵ James GLEICK, *Buszujący w głowach. Nie jesteśmy użytkownikami Google'a, jesteśmy jego produktem*, Książki. Magazyn do czytania 2011, nr 2, s. 28.

Il. 5. a i b. Ukazanie procesu postępującej googlizacji w automatyzowanym katalogu bibliotecznym na przykładzie przemiany interfejsu Głównego Katalogu Komputerowego Biblioteki Narodowej. Ilustracja 5a. przedstawia interfejs katalogu sprzed zmiany. Na zrzucie ekranu widać, że obok paska wyszukiwania oferowany jest wybór wyszukiwania pod kątem konkretnej cechy dokumentu i/lub w konkretnej bazie. Ilustracja 5b. przedstawia ten sam katalog po zmianie interfejsu, która nastąpiła przed 11 sierpnia 2015 roku (ustalono na podstawie kolejnych wersji strony zarchiwizowanych przez Internet Archive Wayback Machine, <https://archive.org/web/> [dostęp: 24 VI 2016]; dziś oba miejsca dostępu są w internecie, jednak komunikat „Zachęcamy do korzystania z nowszej wersji katalogu dostępnego pod adresem: katalogi.bn.org.pl” nie pozostawia wątpliwości, które miejsce preferują twórcy). Tu przeszukiwanie katalogu jest możliwe tylko przez wpisanie zapytania do paska wyszukiwania — pozostałe możliwości są ukryte w opcji „Wyszukiwanie zaawansowane”, anonowanej bardzo skromnym graficznie napisem.

Źródło: a. Katalogi Biblioteki Narodowej, <http://alpha.bn.org.pl/> [dostęp: 24 VI 2016]; b. Katalogi Biblioteki Narodowej, <http://katalogi.bn.org.pl/iii/encore/?lang=pol> [dostęp: 24 VI 2016].

Cytowana wcześniej ortodoksja wyszukiwarek była przykładem kolejnego zjawiska — skeumorfizmu. Określenie to oznacza „dokładne odtworzenie kształtu, barwy i funkcji przedmiotu oryginalnego w materiale zastępczym”²⁶ i w wypadku źródeł informacji oznacza dosłowne powielenie funkcji i możliwości przeniesionych z formy drukowanej, bez zaimplementowania nowych rozwiązań, możliwych do zastosowania dzięki środowisku cyfrowemu. Obecny w zautomatyzowanych PZI skeumorfizm może prowadzić do wypaczeń. Przykładem są te

²⁶ *Słownik wyrazów obcych PWN. Wydanie nowe*, s. 1017.

słowniki językowe, które w dalszym ciągu o wymowie informują przez alfabet fonetyczny, zamiast dołączenia nagrania dźwiękowego z rzeczywistą wymową słowa. Brak ilustracji także jest przejawem skeumorfizmu — słownik, jakkolwiek jego podstawową funkcją jest definiowanie, w niektórych przypadkach nigdy nie odda słowami tak dobrze tego, co mogłaby oddać ilustracja, która zawsze może pełnić funkcję dopełniającą definicję, a tym samym ułatwiać zrozumienie sensu słowa i wszystkich jego desygnatów. To samo można powiedzieć o stosowaniu skrótów, na przykład kwalifikator „wulg.” zamiast „wulgarne” — w zautomatyzowanych PZI nie ma w zasadzie żadnych ograniczeń co do objętości tekstu; można byłoby nawet pokusić się o stwierdzenie, że skrótowość w nich powinna być ograniczona do minimum. W przypadku bibliografii za skeumorfizm można uznać brak przekierowania do pełnotekstowej wersji dokumentu, mimo że jest ona dostępna w sieci, natomiast skeumorficzne encyklopedie będą unikać materiałów audiowizualnych: nagrań dźwiękowych, filmów, animacji, dynamicznych wykresów czy schematów obrazujących w sekwencjach pewne procesy, tabeli przestawnych itp. Bartek Papież tak tłumaczy sens skeumorfizmu:

ma ułatwiać użytkownikowi zrozumienie funkcji cyfrowego programu przez wykorzystanie odpowiedniego odniesienia do świata rzeczywistego. [...] Warto [...] zaznaczyć, że to nie to samo co mimetyzm. Mimetyzm jest jedynie stylem wizualnym, ledwie prześlizgującym się po najbardziej powierzchownej warstwie projektu — formie. Skeumorfizm sięga głębiej, starając się odwzorować również funkcję²⁷.

Dlatego tylko częściowo i nie wszystkie źródła informacji w formie plików w formacie .pdf mogą być uznane za skeumorficzne, albowiem choć wizualnie odtwarzają formę drukowaną w sposób zupełny, a często także przejmują jej funkcjonowanie (na przykład za pomocą zestawionych indeksów), to jednak oprócz tego jest możliwe pełnotekstowe ich przeszukiwanie, a niejednokrotnie zawierają też aktywne hiperłącza.

Oczywiście obecność wszystkich trzech rozwiązań nie jest z gruntu niewłaściwa. Granulacja porządkuje treść jednostki zasadniczej, nie pozostawia wątpliwości co do rodzaju podawanej informacji, znosi także obowiązek znajomości systemu zapisu tej jednostki (przykładowo nie trzeba znać języka opisu bibliograficznego, by wiedzieć, który fragment dotyczy umieszczenia dokumentu w serii). Googlizacja uprzyściplnia korzystanie ze źródła cyfrowym tubylcom, dla których wyszukiwanie informacji w inny sposób niż „googlowanie” jest często niewyobrażalne, jest symbolem intuicyjnego charakteru obsługi źródła, zachęca także do spontanicznego przeszukiwania zasobu, umożliwiając niejednokrotnie dotarcie do zupełnie nieprzewidywanych wyników. Skeumorfizm z kolei ułatwia cyfrowym imigrantom wejście w świat nowych technologii, zachęca do korzystania z nowych możliwości, jakie one oferują, i łagodzi lęk związany ze zmianą dotychczasowych nawyków wyszukiwawczych. Jednak ich nieprzemyślane stosowanie, a w szcze-

²⁷ Bartek PAPIEŻ, *Melodie, które znamy*, 2+3D. Grafika plus Produkt 2013, nr 2 (47), s. 26.

gólności hurraoptymistyczne przekonanie, że teraz wystarczy jedynie udostępnić dane, a reszta „zrobi się sama”, mogą prowadzić do bardzo poważnych nieporozumień, utrudniających maksymalne wykorzystanie potencjału źródła.

4. PODSTAWOWE STRATY

Z tych wstępnych obserwacji wywieść można trzy podstawowe straty, które ponosimy przy niewłaściwej automatyzacji tego typu źródeł.

4.1. ZANIK ADDYTYWNOŚCI

Makrostruktura była nie tylko sposobem na wprowadzenie porządku „pozwalającym kompilatorowi i użytkownikowi na zlokalizowanie informacji”, lecz także sama w sobie najczęściej stanowiła nową informację. Dowolny układ zrębu głównego (może poza najprostszym układem alfabetycznym) dawał nie tylko klucz dostępu do informacji (aspekt wyszukiwawczy SIW), lecz także, przez ich kumulację według cechy, dawał nowy obraz zgromadzonego materiału (aspekt informacyjny SIW). Układ tematyczny osadza jednostkę w szerszym kontekście, układ chronologiczny obrazuje pewien rozwój w czasie, układ frekwencyjny pokazuje powszechność lub rzadkość danego zjawiska itp. Brak troski o informacje niesione do tej pory przez makrostrukturę, które przecież mogą nadal być reprezentowane na przykład w postaci symboli klas schematu klasyfikacyjnego, sprawia, że tracimy pewien rodzaj informacji oraz dostępu do niej. Brak schematów klasyfikacji czy wykazów (a więc elementów megastruktury) zmusza nas do postawienia pytania: jak szukać czegoś, o czego istnieniu nie mamy pojęcia? A przecież najczęściej z PZI korzystamy właśnie w sytuacji niewiedzy; i nie chodzi tu o brak „twardej” wiedzy faktograficznej, lecz niewiedzy ogólnej, kiedy użytkownik nie jest pewien, czy to, czego poszukuje, w ogóle istnieje. Jest to istota heurystyki, jednego z najważniejszych procesów informacyjnych. Heurystyka to od greckiego *heurisko* — ‘znajduję’, czyli „umiejętność wykrywania nowych faktów i związków między faktami, dzięki którym dochodzi się do poznania nowych prawd”²⁸. W PZI z zaniedbaną makro- i megastrukturą zgromadzone małe części informacji przestają harmonijnie i łącznie współtworzyć większe części, informacja przestaje być addytywna²⁹, zaczyna się rozlewać. W tej sytuacji

²⁸ *Słownik wyrazów obcych PWN. Wydanie nowe*, s. 430. W informatologii heurystykę przeciwstawia się algorytmowi. Gdy heurystyka dotyczy wiedzy dopiero poszukiwanej i jest stosowana do rozwiązania zadania, uzupełniania luk i poszukiwania nowych danych — jest więc twórcza, ale i zawodna, to algorytm dotyczy wiedzy już poznanej i problemów już zbadanych, a jego zastosowanie zawsze kończy się sukcesem — jest więc rutynowy — Aneta FIRLEJ-BUZON, *Heurystyka — geneza oraz współczesne zastosowania*, *Zagadnienia Informatyki Naukowej* 2003, nr 1 (81), s. 30.

²⁹ Addytywność to jedna z cech informacji podawana przez Juliana Gałczyńskiego i określana przez niego jako „możliwość tworzenia z elementów informacji — całości znaczeniowej — tekstu skończonego logicznie”, J. GAŁCZYŃSKI, *op. cit.*, s. 7.

użytkownik doświadcza poczucia zalewu informacją oraz heurystycznej niemocy na skutek niemożliwości łączenia faktów.

4.2. POZORNA INTUICYJNOŚĆ

Przedmowy objaśniające zawartość i dobór materiałów, wstępy nakreślające genezę powstania i historię tworzenia źródła, a przez to ukazujące jego ostateczny kształt (w tym mocne i słabe strony), wykazy przyjętych skrótów czy kwalifikatorów, instrukcje postępowania... — to wszystko zapewniało zaplecze metodyczne niezbędne do właściwego i pełnego wykorzystywania źródła. Brak troski o te elementy, i co za tym idzie — o megastrukturę, pozostawia użytkownika bez wsparcia. I choć wydaje się, że korzystanie ze zautomatyzowanych źródeł jest tak intuicyjne, że już nie potrzeba tych wszystkich „zbędnych” dodatków, to w istocie ich użytkownik nierzadko porusza się w nich po omacku. Dobitnym przykładem zbytniego optymizmu i wiary pokładanej w intuicję użytkownika jest słowo wstępne do Elektronicznej Bazy Bibliografii Polskiej Estreichera, z którego dowiadujemy się, że „czerpanie z bazy jest na razie nieodpłatne i nie wymaga żadnych instrukcji poza podstawową umiejętnością obsługi komputera”³⁰, a przecież każdy, kto korzystał z tej bazy, wie, że już w swojej podstawowej strukturze odzwierciedla ona podział drukowanej wersji *Bibliografii Polskiej* Estreicherów na część staropolską, dziewiętnastowieczną i chronologiczną; o całym systemie doboru, selekcji i specyficznej metodzie opisu bibliograficznego nie mówiąc. Stąd, mimo że właściwa automatyzacja tej bibliografii byłaby kolosalnym krokiem w kierunku uprzyśpieszenia wyszukiwania w niej, to jej skeumorficzny charakter krok ten wyraźnie skrócił i nadal znajomość epokowej *Metody Bibliografii Polskiej* Karola Estreichera Marii Dembowskiej³¹ jest nieodzowna, jeśli chce się „Estreichera” wyzyskać w pełni.

4.3. BRAK ARCHIWIZACJI

Pierwsze z wymienionych strat biorą się głównie z automatyzacji nieprzemysłanej. Straty o charakterze trzecim wynikają z tej automatyzacji w ogóle. Niekwestionowana prostota użytkowania i inne walory zautomatyzowanych PZI (dostępność, elastyczność, otwartość itp.) doprowadziły do wypierania tradycyjnych form publikacji. Jednak te drugie wyróżniał bardzo istotny walor: mimowolnie, przez utrwalenie wydawnictw drukiem, utrwalały obraz informacji w danym miejscu i danym czasie. W źródłach zautomatyzowanych o ten walor trzeba intencjonalnie dbać — jeśli więc nie dokonuje się ich systematycznej archiwizacji, tracimy bardzo ważną możliwość badawczą, jaką jest analiza rozwoju i zmian (zarówno ilościowych, jak i jakościowych) informacji w danym obszarze. Źródła tradycyjne

³⁰ Elektroniczna Baza Bibliografii Polskiej Estreichera, www.estreicher.uj.edu.pl/ [dostęp: 19 IV 2016]. Wyróżnienie własne.

³¹ Maria DEMBOWSKA, *Metoda Bibliografii Polskiej Karola Estreichera*. Wyd. 3. uzup., Warszawa: Wydaw. SBP 2001.

oferują nam tę możliwość, ponieważ są dokumentami, które z definicji są „materialnie utrwalonym wyrazem myśli ludzkiej”³² — dokumenty elektroniczne już nie mają pierwotnej postaci fizycznej, jedynie fizyczną lokalizację i sposób organizacji informacji (struktura logiczna)³³ wraz narzędziami ich odczytu (kontekst technologiczny). A jak podkreśla Krzysztof Narojczyk

gdy źródło informacji elektronicznej nie ma waloru stałości, rozumianej tu jako posiadanie identycznej zawartości dla wszystkich użytkowników, nie możemy mówić o nim jako o dokumencie elektronicznym. Przekaz treści, jaki za jego pomocą uzyskujemy, tak długo, jak długo nie zostanie utrwalony, pozostaje jedynie informacją — tekstem, [...] [który] generowany jest dynamicznie z różnych elementów dostępnych w sieci i, z definicji, przechowywany jedynie czasowo w pamięci ulotnej komputera³⁴.

W związku z tym, przykładowo, baza danych jest dokumentem elektronicznym, lecz wyniki wyszukiwania — o ile nie zostaną zapisane — już nie, a zatem baza danych to źródło dokumentów jedynie możliwych/potencjalnych/*in spe*³⁵ i nieświadomość potrzeby archiwizacji jej treści może na dłuższą metę skutkować poważną luką źródłową o charakterze technologicznym. Lukę tę można porównać do luki powstającej na skutek stosowania kwaśnego papieru: niebawem do pewnych treści nie będziemy mogli w ogóle dotrzeć, ponieważ nośniki, na których były zapisane, rozsypią się w pył, a już dziś wiadomo, że nie wszystko uda się przenieść na nośniki alternatywne. Jeśli niearchiwizowanie będzie się utrzymywało, to w ten sposób stracimy kolejne obszary informacji, które niegdyś były utrwalone.

5. STRATY W POSZCZEGÓLNYCH TYPACH

Oprócz tych ogólnych zjawisk, jakich doświadczają zautomatyzowane PZI, można także zaobserwować trudności specyficzne dla konkretnych typów PZI, wymienionych na początku.

Słowniki językowe ogólne reprezentują uniwersum języka, a ich zasadnicze jednostki opisują znaczenie słów. Badanie tego uniwersum przez ich pryzmat może być utrudnione z powodu:

³² *Słownik terminologiczny informacji naukowej*. Pod red. Marii Dembowskiej, Wrocław: Zakład Narod. im. Ossolińskich 1979, s. 38. Wyróżnienie własne.

³³ Niekiedy postuluje się w ogóle wyłączenie dokumentów elektronicznych z pojęcia „dokument” i określanie ich nie jako grupy dokumentów, lecz jako zupełnie odrębne zjawisko: „zasoby elektroniczne”, zwracając uwagę, że bardziej chodzi w nich o treść, a nie (bardzo płynną) formę.

³⁴ Krzysztof NAROJCZYK, *Dokument elektroniczny i jego opis bibliograficzny w publikacjach humanistycznych*, Olsztyn: Wydaw. Uniwersytetu Warmińsko-Mazurskiego 2005, s. 19, 21.

³⁵ Ponieważ „immanentną cechą dokumentu, także elektronicznego, jest jego względna trwałość, wyrażająca się w tym, że jest on w określony sposób zapisany (utrwalony). Oznacza to, że można go wielokrotnie ekspozycjonować, mając gwarancję, że za każdym razem prezentowana jest ta sama treść. Na tym pewniku opiera się rzetelność obiegu informacji w nauce”, *ibidem*, s. 33–34. Tu także więcej o specyfice informacji generowanej na bieżąco i potrzebie jej utrwalania.

— nadmiaru mikrostruktury: współczesne programy służące do tworzenia automatycznego słowników są konstruowane dość sztywno i wprowadzają „terror schematu hasła”. Jean Pruvost tak opisuje współczesny proces powstawania słowników:

przed redakcją słownika opracowuje się tzw. opis typu dokumentu (OTD), mający na celu informacyjne zdefiniowanie wszystkich obiektów, uporządkowanych i zhierarchizowanych, które mogą się pojawić w artykule hasłowym. Taki opis umożliwia sporządzenie szablonu, wypełnianego następnie przez leksykografa. Dla przykładu, są w nim zaprogramowane wszystkie kategorie gramatyczne czy wszystkie kwalifikatory terminologiczne, łatwe do odnalezienia na rozwijanej liście³⁶.

Oczywiście Pruvost dostrzega zalety takiego rozwiązania (łatwość formatowania, korekty i zmiany nośnika ostatecznego produktu, a tym samym oszczędność czasu i pieniędzy), lecz, niestety, z drugiej strony leksykograf przed przystąpieniem do pracy nie jest w stanie przewidzieć wszystkich oboczności i osobliwości, jakie mogą wystąpić w języku, i które tym samym będą musiały być reprezentowane w jednostce zasadniczej. I choć doskonale zdaje sobie sprawę, że jeden schemat nie może być jednakowo wydajny wobec wszystkich typów, ponieważ jednostki, które opisuje, są bardzo różnorodne (przykładowo, w jednym artykule hasłowym wystarczy ilustracja kilkuwyrazową ekscerpcją, ale inny będzie wymagał już przykładu kilkudzaniowego) to bardzo często jest już wtedy za późno na zmiany:

opracowanie dobrego OTD zakłada, że zostały już przewidziane wszystkie mogące wystąpić problemy, co można sobie wyobrazić w przypadku reedycji albo słowników dawnych, a co jest wątpliwe w przypadku nowego słownika. Często bowiem zdarza się, że już po utworzeniu OTD, w trakcie analizy dotyczącej jakiegoś wyrazu hasłowego, pojawia się nowy obiekt, nieuwzględniony w OTD. Tymczasem, choć OTD powinien pozostać otwarty, w rzeczywistości często trudno go modyfikować w toku prac³⁷.

Tym samym uwolniono leksykografów od obowiązku maksymalnej zwięzłości i skrótowości (powodowanej miejscem ograniczonym fizycznymi możliwościami druku), lecz jednocześnie narzucono im rygor dostosowania się do sztywnej mikrostruktury. Pruvost zauważa też, że granulacja danych mikrostruktury

³⁶ Jean PRUVOST, *Les dictionnaires français outils d'une langue et d'une culture*, Paris: Ophrys 2006, s. 181. Przeł. Witold Ucherek, oryg.: „On élabore ainsi avant la rédaction du dictionnaire une DTD (Description de Type de Document) qui a pour objectif de définir informatiquement toutes les entités alors ordonnées et hiérarchisées pouvant intervenir. Une fois établie, cette DTD permet de concocter un masque de saisie que remplit le lexicographe. Y sont par exemple programmées toutes les catégories grammaticales, toutes les marques de domaine, aisément appelées sur un menu déroulant”.

³⁷ *Ibidem*. Przeł. Witold Ucherek, oryg.: „L'élaboration d'une bonne DTD suppose que soient déjà prévus tous les problèmes rencontrés, ce qui peut se concevoir pour les rééditions ou les dictionnaires anciens et ce qui reste aléatoire pour un nouveau dictionnaire. Ensuite, il est fréquent qu'une fois la DTD constituée, à la faveur de l'analyse conduite à propos d'un mot dont on bâtit l'article, surgisse une nouvelle entité qui n'a pas été prévue dans la DTD. Si une DTD doit rester ouverte, dans les faits, elle est souvent difficile à modifier en cours de route”.

w postaci numeracji kilku znaczeń jednego słowa we współczesnych słownikach przeczy nieco jego istocie i wspomina nienumerowane artykuły hasłowe w pierwszych wydaniach słownika Akademii Francuskiej czy słownika *Petit Larousse*:

czytając artykuł jednym ciągiem, lepiej się dostrzegało, w jakim stopniu słowo nie jest jednostką dającą się bez końca rozczłonkowywać w drodze analiz, choćby nie wiem jak naukowych. Któż może stwierdzić, że jakieś słowo polisemiczne posiada 8, 10, 12 czy 15 znaczeń? Literaci, a w szczególności poeci, dobrze wiedzą, że znaczenie słowa znajduje się częściej między znaczeniami, w osmozie, niż w z góry określonej przegródce. Lektura artykułu bez przesadnej liczby znaczników sprzyjała takiemu podejściu, co warto przypomnieć³⁸.

— braku makrostruktury: rezygnacja z uszeregowania w pewnym logicznym porządku jednostek zasadniczych może jawnie wprowadzać w błąd. Jeszcze do niedawna dwukierunkowy słownik przekładowy *Diki. Multimedialny słownik angielskiego online*³⁹ na zapytanie „Spacer” zwracał odpowiedź „s p a c e r, American English, rzeczownik, ‘przekładka’, technical”, mimo że poszukiwane było angielskie znaczenie polskiego słowa „spacer”: w większości przypadków homografy obu języków (na przykład smoking, wanna itp.) były tłumaczone w jednym oknie obok siebie, w tym jednak konkretnym, na skutek błędu, można było odnieść wrażenie, że w języku polskim nie ma słowa „spacer” i dlatego nie ma ono swojego tłumaczenia⁴⁰. Błąd ten nie mógłby w ogóle wystąpić w słowniku tradycyjnym, w którym szeregi wyrazów byłyby osobne dla kierunku „polski → angielski” oraz „angielski → polski” i użytkownik sam decydowałby, jakiego rodzaju informacja jest mu potrzebna. W środowisku automatycznym również można byłoby mu zapobiec, gdyby twórcy słownika zaoferowali użytkownikowi wybór, z którego kierunku tłumaczenia chce skorzystać. Taką funkcję ma na przykład słownik *Glosbe — wielojęzyczny słownik online*⁴¹, który najpierw pyta o wybór języka, a następnie o kierunek tłumaczenia. Jest to nieodzowne rozwiązanie, ponieważ jako słownik wielojęzyczny umożliwia przekłady w niezliczonych kombinacjach: na dzień 17 maja 2016 roku oferował zasoby leksykalne 7983 języków⁴², które następnie może ze sobą krzyżować, aby oferować tłumaczenia. Niestety, na skutek niekorzystnego wpływu googlizacji źródła twórcy *Diki* nie pomyśleli o takim rozwiązaniu.

³⁸ *Ibidem*, s. 181–182. Przeł. Witold Ucherek, oryg.: „En lisant l'article d'une traite, on y percevait mieux combien le mot n'est pas un élément dissécable à l'infini au gré des analyses, fussent-elles éminemment scientifiques. Qui peut dire qu'un mot polysémique comprend 8, 10, 12 ou 15 sens? Les gens de lettres et plus particulièrement les poètes savent bien que le sens du mot est souvent entre les sens, en osmose, plus que dans une case prédéterminée. La lecture d'un article sans balisage excessif favorisait cette approche qui mérite aussi d'être rappelée”.

³⁹ *Diki. Multimedialny słownik angielskiego online*, www.diki.pl [dostęp: 17 V 2016].

⁴⁰ Ostatnio błąd ten poprawiono, choć nadal w sekcji angielskiego wyrazu „spacer → pol. ‘przekładka’” istnieje link do porady *Przedimek przed rzeczownikiem „walk” — spacer?*, która przecież powinna się znaleźć pod „spacer → ang. ‘walk’” — *Diki. Multimedialny słownik angielskiego online* [szukaj: *spacer*], www.diki.pl/sownik-angielskiego?q=spacer [dostęp: 17 V 2016].

⁴¹ *Glosbe — wielojęzyczny słownik online*, pl.glosbe.com [dostęp: 17 V 2016].

⁴² *Wszystkie dostępne języki*, pl.glosbe.com/all-languages [dostęp: 17 V 2016]

— braku megastruktury: brak informacji o doborze słów, celu tworzenia słownika oraz do kogo jest on adresowany, a przede wszystkim o opisie metody definiowania i budowy wyrazu hasłowego także może prowadzić do nieporozumień. Brak właściwej relacji między mikrostrukturą a megastrukturą doprowadził słownik *Lingea Lexicon 5. Uniwersalny słownik francusko-polski, polsko-francuski*⁴³ do sprzeczności: oferuje możliwość wyszukiwania zaawansowanego według kwalifikatorów stylu lub dziedzin, lecz jednocześnie nie podaje listy tychże. Użytkownik musi więc się domyslać, jaką postać mają te kwalifikatory (na przykład szukać wyrazów wychodzących z użycia przez wpisywanie w odpowiednie miejsce „arch.” = archaiczny, „przest.” = przestarzały, „daw.” = dawniej, a może przez francuskie „vx.” = vieux?). Wpisanie dziedziny „biologia” zwróci brak wyników — dopiero zastosowanie „biol.” zaprowadzi nas do słownictwa z zakresu biologii. A przecież w automatycznych PŻI nie jesteśmy ograniczeni papierową objętością tradycyjnego nośnika: dołączenie bardzo obszernego i drobiazgowego wstępu objaśniającego wszystkie niuanse, w tym zastosowane rozwiązania oraz stosunek leksykografa do zebranego materiału (a także do tego, co celowo pominął), nie jest już żadnym problemem. Niestety, na skutek niekorzystnego wpływu pozornej intuicyjności twórcom *Lingei Lexicon 5* mikrostruktura jednostki zasadniczej zdała się oczywistością i wszelkie instrukcje oraz pomoce zawarte na dołączonej wkładce dotyczą jedynie obsługi technicznej (informatycznej) słownika. Zatem możemy się dowiedzieć, jak go zainstalować, a także jaka jest architektura jego interfejsu, lecz już, niestety, nie mamy wspomnianej wiedzy na temat zastosowanych kwalifikatorów — pozostaje nam ślepy traf. Brak informacji o doborze słów lub wykazu aneksów uniemożliwia ocenę, czy słownik ten uwzględnia chociażby nazwy własne (na przykład imiona czy nazwy państw), a przecież to może być podstawą wyboru słownika. Na marginesie tego akapitu można jeszcze rozważyć pytanie, dlaczego taki słownik nie miałby ukrywać elementów frankofońskich, jeśli zamierza się korzystać z opcji dla polonofonów.

— braku archiwizacji: jeśli ona nie występuje, uniemożliwi to w przyszłości prowadzenie badań nad pojawieniem się słowa (czy raczej jego leksykograficznej rejestracji) lub ewolucji jego definicji⁴⁴. Piotr Wierzchoń zwraca także uwagę na związane z tym zagadnienie aktualności słownika. Możliwość niemalże

⁴³ *Lingea Lexicon 5. Uniwersalny słownik francusko-polski i polsko-francuski* [dysk optyczny], Kraków: Lingea 2009.

⁴⁴ Piotr Żmigrodzki przestrzega, że słowniki nie do końca podają rzeczywiste znaczenie wyrazów, lecz jedynie gromadzą ich opisy leksykograficzne. Na tej podstawie nie powinniśmy mówić o tym, że słowniki ukazują rzeczywistość językową, a jedynie obraz leksykograficzny — podając autorskie definicje, przekazują jedynie ustalenia leksykografa na temat rzeczywistości, które chyba nigdy do końca nie mogą być obiektywne; rzeczywiste znaczenie słowa pokazują dopiero kolokacje, a więc umiejscowienie go w kontekście, co możliwe jest do zbadania jedynie w korpusach językowych. Tak samo z pojawieniem się słowa: słownik ukazuje tylko, kiedy leksykografowie dostrzegli je i uznali za właściwe zarejestrować — samo słowo, w rzeczywistości

natychmiastowej modyfikacji zasobu słownika w środowisku automatycznym (dodawanie, odejmowanie i zmienianie haseł lub znaczeń) jest fascynująca, jednakże dopiero upływ czasu zweryfikuje, czy osąd leksykograficzny był słuszny — wcześniej częściowo czas na odpowiedni namysł i wyważenie osądu zapewniało stosunkowo długi okres prac redakcyjnych oraz kolegialny sposób podejmowania decyzji⁴⁵. Poddaje także pod rozważę cenę, jaką zapłacimy za brak takiego przejawu dotychczasowej archiwizacji, jak suplementy:

Słowniki internetowe mogą być (teoretycznie) aktualizowane bezustannie w tym sensie, że efekty tej aktualizacji dostępne są użytkownikowi Internetu praktycznie natychmiastowo. Niwelowany więc jest okres, w którym po zamknięciu pewnego etapu prac nad słownikiem drukowanym rozpoczyna się pracę nad osobnym wydawnictwem — suplementem. Jednak czy brak jawnie oddzielnego zbioru leksykalnego nie będzie w niektórych przypadkach faktem niekorzystnym? Można na tak postawione pytanie odpowiedzieć twierdząco, jeśli za punkty analizy zamierza się przyjąć z pewnych względów tylko słownictwo zawarte w suplementie [...]. Słownik sieciowy problem taki rozwiązywałby tylko wtedy, gdyby przy każdym hasle podawać bądź datę oznaczającą jego wprowadzenie, bądź symbol przedziału czasowego (z założoną tolerancją, na przykład kwartalną, półroczną itp.), w którym zanotowano i opracowano to hasło⁴⁶.

Bibliografie i katalogi ogólne reprezentują uniwersum dokumentów lub jego faktycznie istniejący wycinek, a ich zasadnicze jednostki opisują dokumenty. Badanie tych uniwersów przez ich pryzmat może być utrudnione z powodu:

— źle zorganizowanej mikrostruktury: jedną z możliwości wyszukiwania w Narodowym Uniwersalnym Katalogu Centralnym⁴⁷ jest wyszukiwanie zaawansowane z użyciem hasła przedmiotowego. Niestety, algorytm wyszukiwania nie różni jednostek leksykalnych tego języka i oferując ogólne pole wyszukiwawcze „Temat”, nie umożliwia zaznaczenia formy znaczenia, które się tam wpisuje. Z tego powodu, po wpisaniu zapytania „Kolekcjonerstwo”, otrzymujemy wyniki, w których kolekcjonerstwo jest tematem, określnikiem lub dopowiedzeniem. Skutkuje to niską relewancją wyników, gdyż obok piśmiennictwa na temat zjawiska kolekcjonerstwa⁴⁸ (pożądanego) otrzymujemy także ujęcie kolekcjonerskie jakiegoś innego tematu⁴⁹

językowej, mogło się pojawić znacznie wcześniej. Por. Piotr ŻMIGRODZKI, *Leksykografia jako przedmiot uniwersyteckiej edukacji polonistycznej*, Poradnik Językowy 2001, z. 9, s. 42 (punkt 3).

⁴⁵ Piotr WIERZCHOŃ, *Internet a leksykografia*, [w:] *Dokumenty elektroniczne*. Pod red. Mirosława Górno i Piotra Nowaka, Poznań: Wydaw. Sorus 1999, s. 12.

⁴⁶ *Ibidem*, s. 13–14.

⁴⁷ NUKAT. *Prosto do informacji. Katalog zbiorów polskich bibliotek naukowych*, <http://katalog.nukat.edu.pl/> [dostęp: 17 V 2016].

⁴⁸ Przykład: Kolekcjonerstwo — Polska — 18–20 w. [Tomasz Feliks de ROSSET, *Polskie kolekcje i zbiory artystyczne we Francji w latach 1795–1919. Między „skarbnicą narodową” a galerią sztuki*, Toruń: Wydaw. Uniwersytetu Mikołaja Kopernika 2005].

⁴⁹ Przykład: Florencja (Włochy) — kolekcjonerstwo — zbiory [Francesca FIORANI, *The Marvel of Maps. Art, cartography and politics in Renaissance Italy*, New Haven: Yale University Press 2005].

(mniej istotne) oraz opisy konkretnych kolekcji⁵⁰ (najmniej istotne). Powoduje to ogromny nadmiar zwróconych wyników oraz pewne rozmycie prezentowanej literatury, co niewprawnemu użytkownikowi może bardzo utrudnić dotarcie do właściwego piśmiennictwa.

— braku makrostruktury: makrostrukturą w bibliografii lub katalogu jest układ bibliograficzny. Jeśli przyjmuje on kształt rzeczowy — umożliwia bardzo cenne dla użytkownika grupowanie informacji według tematyki dokumentu. Brak tej informacji skutkuje nie tylko problemem z dotarciem do odpowiednich dokumentów, o czym świadczą wypowiedzi bibliografów opracowujących bibliografie dziedzin i zagadnień⁵¹, lecz w dłuższym okresie uniemożliwi prowadzenie badań nad przyrostem piśmiennictwa w danych dziedzinach, a to spowoduje brak informacji naukowawczej o kondycji (rozwoju, stagnacji, upadku, kierunkach zmian) dyscypliny. Marta Skalska-Zlat pisze: „tytuły publikacji naukowych i fachowych łączy się przeważnie z symbolami klasyfikacyjnymi [...] w celu wykazania struktury tematycznej danej dyscypliny. Im bardziej szczegółowa klasyfikacja, tym precyzyjniejsza ocena kierunków rozwoju dyscypliny, jej dokonań i braków”⁵². Jednak dalej przytacza spostrzeżenie innej uznanej badaczki bibliografii:

takie spojrzenie na bibliografię [tj. wykorzystanie spisów bibliograficznych w badaniach nad stanem nauki] postulowała A[nn]a Sitarska na ogólnopolskiej naradzie bibliografów w 1978 r., zwracając uwagę, że ten sposób wykorzystania bibliografii jest szeroko praktykowany w świecie, przyczyniając się do lepszego poznania i „zinterpretowania formalnych i pozaformalnych cech systemu nauki”. W Polsce natomiast twórcy solidnych nieraz baz materiałowych zatrzymują się w pół drogi oddając je użytkownikom, którzy wykorzystują dane jednostkowe. Natomiast „potencjał zawartej

⁵⁰ Przykład: Jankowski, Andrzej (kolekcjonerstwo) — kolekcje kart pocztowych — katalogi. [Andrzej JANKOWSKI, Tomasz URBANIAK, Andrzej SZKUDNIEWSKI, *Koszalin na starych pocztówkach*, Słupsk: Stolpbud 2008].

⁵¹ „Jako autorka bibliografii z zakresu językoznawstwa [...] prace nad bibliografiami rozpoczynam od czytania »BZCz« [Bibliografii Zawartości Czasopism] i od »Przewodnika Bibliograficznego«, następnie, po wstępnej selekcji, przeglądam konkretne tytuły i artykuły. Brak wyszukiwania według działów nie tylko utrudnia pracę, ale ją zuboża. Być może ja powinnam zmienić zasady wyszukiwania, bo język hasel przedmiotowych nie jest doskonały i nie daje takiej precyzji jak wszelkie bibliografie dziedzinowe. Przez hasła w »BZCz« nie wyszuka się potrzebnych dziedzin, zwłaszcza, jeżeli są one interdyscyplinarne. Redagując bibliografię semantyczną, czytam rozdziały dotyczące filozofii, logiki, nauk społecznych, psychologii i inne, bo to wszystko się ze sobą łączy. Drukowana forma »BZCz« zapewnia kompletność materiału” — wypowiedź Aliny Grochowskiej w drugiej części dyskusji po V Ogólnokrajowej Naradzie Bibliografów — *Piąta Ogólnokrajowa Narada Bibliografów, Warszawa 11–13 czerwca 2003. Referaty i dyskusja*. Pod red. Joanny Gregorczyk, Warszawa: Biblioteka Narodowa 2004, s. 279–280. Cytat pochodzi z czasów, kiedy ważyły się losy formy publikacji tych bibliografii. Szczęśliwie obie, po przejściu na pełną automatyzację, zachowały reprezentację klasyfikacji (według UKD) oraz stabilne formy archiwalne (w postaci plików w formacie .pdf).

⁵² Marta SKALSKA-ZLAT, *Bibliografia w Polsce 1945–1996. Naukowawcza analiza dyscypliny*, Wrocław: Wydaw. Uniwersytetu Wrocławskiego 2002, s. 34.

w tych zbiorach wiedzy o całym zespole danych, o powiązaniach i współzależnościach formalnych i treściowych pozostaje niewykorzystany⁵³.

Można byłoby rzec, że w automatycznych PZI jest on niewykorzystany jeszcze bardziej, gdyż generowanie w nich zestawień ilościowych jest banalnie proste dzięki oprogramowaniu komputerowemu.

— braku megastruktury: brak informacji o doborze materiałów do bibliografii (czyli o jej zakresie i zasięgach) czy szerzej — o jej metodzie — skutecznie obniża wartość takiej bibliografii i stawia pod znakiem zapytania ogólny sens takiego źródła. Przykładem może być Bibliografia Historii Polskiej⁵⁴, która informuje jedynie o zasięgu chronologicznym wydawniczym bazy oraz że „zawiera te same dane, co kolejne roczniki wydawane w formie książkowej przez Instytut Historii PAN”⁵⁵. To implikuje konieczność sięgnięcia do wersji drukowanej bibliografii (lub jakichś zewnętrznych opracowań: opisu metodycznego lub recenzji), aby dowiedzieć się, co tak naprawdę można w niej znaleźć, ponieważ sformułowanie „historia polska” może być tak szeroko interpretowane (Polski w jakich granicach? Historii za jaki okres? Historia polska, czyli ogół dokumentów dot. historii wytworzonych przez polskich naukowców czy też ogół dokumentów dot. historii Polski?), że jest absolutnie niewystarczające. Oczywiście specjaliści z tego zakresu z pewnością wiedzą, czego mogą się spodziewać po tej bibliografii (mającej bezdyskusyjne uznanie w środowisku zarówno historyków, jak i bibliografów), jednakże upublicznienie bibliografii w Internecie przyczynia się do jej upowszechnienia, w związku z tym mogą natknąć się na nią również osoby bez specjalnego doświadczenia. Wówczas bez elementów megastruktury będą bezradne (abstrahując od problemu bariery językowej, spowodowanej wysyceniem specjalistyczną terminologią typowej przedmowy do bibliografii, czyniącą ją niezrozumiałą dla niewprawnego użytkownika).

— braku archiwizacji: brak stabilnej archiwizacji bibliografii może powodować komplikacje w poszukiwaniach o charakterze systematycznym — odbywających się nie doraźnie, lecz w regularnych interwałach czasowych, które bardzo często prowadzą bibliografowie tworzący bibliografie regionalne i lokalne. Dowodem na to są rzeczywiste wypowiedzi tego typu specjalistów⁵⁶, którzy wskazują,

⁵³ *Ibidem*, s. 167, gdzie cytuje: Anna SITARSKA, *Bibliografia a nauka*, [w:] *Trzecia Ogólnokrajowa Narada Bibliografów, Warszawa 5–7 VI 1978 r. Referaty i dyskusja*. Pod red. Barbary Karamać i Elżbiety Czermarmazowicz, Warszawa: Biblioteka Narodowa 1980, s. 64, 66–67.

⁵⁴ Bibliografia Historii Polskiej, <http://www.bibliografia.ipn.gov.pl/> [dostęp: 17 V 2016].

⁵⁵ *O bibliografii. Informacja o zawartości bazy Bibliografii Historii Polskiej*, <http://www.bibliografia.ipn.gov.pl/bhp/o-bibliografii/8346,Informacja-o-zawartosci-bazy-Bibliografii-Historii-Polskiej.html> [dostęp: 17 V 2016].

⁵⁶ „Wersja drukowana »BZCz«, a mówię to jako bibliograf regionalny, jest dla nas najważniejszym narzędziem pracy. Ja po prostu czytam »BZCz«, »Przewodnik Bibliograficzny« i nie wyobrażam sobie innego systemu pracy. [...] Kryterium doboru do bibliografii regionalnej, podczas przeglądania »BZCz«, jest dla mnie hasło geograficzne. Przeglądając indeks przedmiotowy,

że jakkolwiek dynamiczne bazy danych mają swoje zalety przy poszukiwaniach jednorazowych, to jednak nie dają pewności co do kompletności wyszukiwań, jeśli te mają za zadanie stale monitorować pojawianie się nowego piśmiennictwa; dodatkowo każdorazowe przeglądanie całej bazy (czy raczej każdorazowe kalibrowanie jej pod właściwym kątem) jest zdecydowanie bardziej uciążliwe niż lektura kolejnych zeszytów.

Encyklopedie i leksykony reprezentują uniwersum przedmiotów, osób, miejsc i zdarzeń. Badanie tego uniwersum przez ich pryzmat może być utrudnione z powodu:

— braku archiwizacji: niedostępność utrwalonego źródła sprawia, że nie jesteśmy w stanie odtworzyć obrazu wiedzy z danego okresu, gdyż aktualizowane na bieżąco encyklopedie zawierają jedynie jej stan aktualny. Na marginesie tego spostrzeżenia należy poczynić jeszcze jedną uwagę, być może bardziej dotkliwą aniżeli problem archiwizacji. Archiwizacja utrwali obraz, jednak czy jest to obraz reprezentatywny? Wśród tak zwanych wolnych encyklopedii, przeważających wśród zautomatyzowanych encyklopedii, jeszcze większym problemem jest niemożność stwierdzenia kanonu wiedzy. Tworzone są one w dużej mierze przez hobbystów, a więc z dużą dozą prawdopodobieństwa można powiedzieć, że w encyklopediach tych pisze się o tym, co ich interesuje, a nie o tym, co powinno być wiadome⁵⁷. Przykładowo, w kategorii „Informacja naukowa” Wikipedia podaje 64 strony (właśc. artykuły hasłowe; wraz ze stronami dwóch podkategorii)⁵⁸, gdy

przyjmuję lub odrzucam potrzebne mi materiały i informacje. Zasadniczą różnicą jest wyszukiwanie informacji w bazie komputerowej, która zawiera ponad 300 000 rekordów, a pozyskiwanie materiałów w trakcie czytania określonego zeszytu”, wypowiedź Alicji Browko w drugiej części dyskusji po V Ogólnokrajowej Naradzie Bibliografów — *Piąta Ogólnokrajowa Narada Bibliografów*, s. 299. To też cytat z czasów, kiedy ważyły się losy formy publikacji tych bibliografii, zob. uwaga w przyp. 53.

⁵⁷ Problem profesjonalizmu pojawia się także przy automatycznych słownikach, na co zwraca uwagę Piotr Wierchoń, dostrzegając jednak pozytywny aspekt tej sytuacji: „redaktorzy słowników, a zwłaszcza dzieł o znaczeniu fundamentalnym dla kultury i nauki polskiej [...], to przede wszystkim doświadczona kadra polonistów, językoznawców. Przy takim panteonie naukowców większość nazwisk sieciowych leksykografów [...] po prostu nigdy nie znajdzie swojego miejsca, ale warto w tym miejscu spytać, czy aby nie jest to zbyt pochopna i przedwczesna kwalifikacja młodych entuzjastów w obszar niebytu naukowego? [...] Często więc zdarza się, że sieciowi leksykografowie (z racji tego, iż w ich kręgu zainteresowania pozostają stosunkowo wąskie dziedziny rzeczywistości) dysponują faktycznie wysoce odpowiednią kompetencją we właściwym zakresie. [...] Czy więc słownik »licealisty-internauty« będzie słownikiem informatyka czy hobbisty? Z drugiej strony można zapytać, czy słownik piętnastoletniego ircowca [...] uważać można za mało wartościowe odpryski wolnego czasu nastolatka, czy też — ufając jego kompetencji — przyjąć opracowany przez niego zbiór za obiektywnie wystarczający zadowalający, a wreszcie — nie bójmy się tego powiedzieć — leksykograficznie korzystny i wartościowy”, P. WIERZCHOŃ, *op. cit.*, s. 16 (tu obszerna egzemplifikacja).

⁵⁸ Kategoria: Informacja naukowa, ostatnia modyfikacja 31 stycznia 2015, pl.wikipedia.org/wiki/Kategoria:Informacja_naukowa [dostęp: 17 V 2016]. Nawet, gdybyśmy powiększyli tę pulę

sam tylko *Słownik terminologiczny informacji naukowej* podaje około 1660 haseł głównych⁵⁹, a *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych* — aż 1804 artykułów hasłowych⁶⁰. Wprawdzie można wysunąć zarzut pod adresem tych wartości, wskazując na niewątpliwy fakt, że są to, po pierwsze, źródła dziedzinowe, a nie ogólne (a zatem będą bardziej drobiazgowo w rejestracji), a po drugie — są słownikami a nie encyklopediami (choć jeden z nich ma „charakter encyklopedyczny”), jednakże nie może to w całości wytłumaczyć rozbieżności między tymi liczbami. Niepodważalnych w porównywaniu danych dostarczyłaby liczba haseł z zakresu informacji naukowej z uznanej polskiej encyklopedii ogólnej, na przykład trzydziestotomowej *Wielkiej encyklopedii PWN*⁶¹, lecz bezwzględne dane są w tym momencie dla autora nieosiągalne⁶². Wystarczy powiedzieć, że wprowadzenie zapytania „bibliograf” do automatycznego indeksu encyklopedii zwróciło 91 wyników, w tym 85 relewantnych⁶³, co pokazuje nam liczbę biogramów samych tylko bibliografów, które przecież jak najbardziej należą do obszaru informacji naukowej właściwego do zarejestrowania w encyklopedii ogólnej — a co z desygnatami terminów, hasłami geograficznymi, instytucji, tytułów czasopism itp.? Stąd z pewnością można stwierdzić, że liczba haseł informatologicznych w Wikipedii jest zbyt skromna, aby można było uznać ją za reprezentatywną dla tej dziedziny. Zagrożeniem obrazowania stanu wiedzy może być także dokonywanie bezwolnych tłumaczeń treści z innych wersji języko-

o strony kategorii „Bibliografie” (Wikipedia nie wyróżnia samodzielnego działu dla obszaru nauki o bibliografii, najwyraźniej jej twórcy nie interesują się tym obszarem, nie wiedzą o nim lub zajmują anachroniczne stanowisko, że bibliografia to jedynie produkt, spis konkretnej literatury, a nie cały obszar wiedzy) w liczbie 23, to i tak otrzymamy ostateczny wynik 87 stron — Kategoria: Bibliografie, ostatnia modyfikacja 17 marca 2013, pl.wikipedia.org/wiki/Kategoria:Bibliografie [dostęp: 18 V 2016].

⁵⁹ *Słownik terminologiczny informacji naukowej*, s. 5.

⁶⁰ Bożenna BOJAR, „*Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*”. *Problemy metodologiczne*, Przegląd Biblioteczny (71) 2003, z. 1/2, s. 121.

⁶¹ *Wielka encyklopedia PWN*. Pod red. Jana Wojnowskiego. T. 1–31, Warszawa: Wydawnictwo Naukowe PWN 2001–2005.

⁶² Wprawdzie *Indeks tomów 1–31 na płycie CD-ROM* dołączony do *Suplementu* oferuje indeks podstawowych dziedzin wiedzy, lecz tylko jako kryterium zawężania liczby wyników w odpowiedzi na pytanie, natomiast nie ma możliwości wygenerowania listy wszystkich haseł z danej dziedziny; w dodatku informacja naukowa nie jest w nim wyodrębniona jako samodzielny dział, a artykuły hasłowe z tego zakresu są rozproszone po działach „Bibliotekoznawstwo i nauka o książce” oraz „Naukoznawstwo” (do którego z resztą przypisano hasło „Informacja naukowa”). Pewnym rozwiązaniem tego problemu mogłaby być aproksymacja liczby haseł na podstawie liczby w obrębie kilku liter lub wyrywkowa kontrola haseł porównywana z cytowanymi słownikami, jednakże w dalszym ciągu dane te mogłyby być obciążone błędem niewłaściwego przybliżenia; nie jest też to głównym przedmiotem tego artykułu.

⁶³ *Indeks tomów 1–31 na płycie CD-ROM* [dysk optyczny; w:] *Wielka encyklopedia PWN*. Pod red. Jana Wojnowskiego. T. 31 (Suplement), Warszawa: Wydawnictwo Naukowe PWN 2005 [szukaj: *bibliograf*]. Owe sześć odrzuconych wyników to trzy odsyłacze, dwa uzupełnienia haseł w suplementie i jedno hasło nerelewantne.

wych tego samego hasła, które mogą nie odzwierciedlać specyfiki danego przedmiotu w kraju, jego odmiennego postrzegania lub niewłaściwie tłumaczyć terminy specjalistyczne. Jest to jednak już zagadnienie wkraczające na pole wartości, a nie sprawności informacyjnej.

6. EFEKTY I JAK IM ZAPOBIEGAĆ

Jak łatwo zauważyć, każdy z tych typów PŹI obrazuje pewną rzeczywistość, dotyczy jakiegoś uniwersum — uniwersum języka, dokumentów, nazw (przedmiotów, osób, miejsc i zdarzeń) lub jego fizycznie osiągalnego fragmentu. Niewłaściwa automatyzacja PŹI prowadzi do nieutrwalenia obrazu tego uniwersum. Zanik kategoryzacji i porządkowania skutkuje rozlewaniem się informacji, zalewem informacyjnym potęgującym i tak już nadmierny szum informacyjny. Może to prowadzić do powstawania informacji nieprawdziwej lub nierelevantnej: z powodu braku kontroli specjalisty, który do tej pory czuwał nad właściwym opracowaniem źródła lub pozbawiania go narzędzi, którymi mógł to źródło brać w ryzy. Przy czym nie chodzi o postawę „policjanta drogowki” decydującego, kto „wjeżdża”, a kto nie do zasobu źródła, lecz „rybaka”, który zarzuca sieć, wyławia różne okazy z otchłani i je opisuje — jak trafnie określa charakter pracy współczesnego leksykografa (ale też opracowujących inne typy PŹI) Erin McKean⁶⁴, przy okazji

⁶⁴ „Ludzie myślą, że moja praca polega na wpuszczaniu dobrych słów do słownika i zatrzymywaniu złych. Ale ja nie chcę być gliniarzem z drogowki. [...] Decydowanie, które słowa są dobre, a które złe, prawdę mówiąc nie jest proste; ani przyjemne. A kiedy twoja praca nie jest prosta lub przyjemna, szukasz wymówki, aby jej nie robić. Więc jeśli miałabym pomyśleć o jakimś zajęciu jako metaforze mojej pracy, zdecydowanie bardziej byłby to rybak. Chcę zarzucić moją olbrzymią sieć do głębokiego, niebieskiego oceanu języka angielskiego i zobaczyć, jakie cudowne stworzenia mogą podnieść z jego dna. [...] Słowniki online powielają prawie wszystkie problemy słowników drukowanych, poza wyszukiwaniem. A kiedy poprawisz wyszukiwanie haseł, w zasadzie odbierzesz jedyną zaletę druku, którą jest szczęśliwy traf. Szczęśliwy traf jest wówczas, kiedy znajdujesz rzeczy, których nie szukałeś, ponieważ znajdowanie tego, czego się szuka, jest ogromnie trudne” (tłum. własne, w oryg.: „People think that my job is to let the good words make that difficult left-hand turn into the dictionary, and keep the bad words out. But the thing is, I don't want to be a traffic cop. [...] Deciding what words are good and what words are bad is actually not very easy. And it's not very fun. And when parts of your job are not easy or fun, you kind of look for an excuse not to do them. So if I had to think of some kind of occupation as a metaphor for my work, I would much rather be a fisherman. I want to throw my big net into the deep, blue ocean of English and see what marvelous creatures I can drag up from the bottom. [...] Online dictionaries replicate almost all the problems of print, except for searchability. And when you improve searchability, you actually take away the one advantage of print, which is serendipity. Serendipity is when you find things you weren't looking for, because finding what you are looking for is so damned difficult”), Erin McKean, *The joy of lexicography*, www.ted.com/talks/erin_mckean_redefines_the_dictionary [dostęp: 18 V 2016]. Referat wygłoszony na konferencji „TED2007. Icons. Geniuses. Mavericks” (Monterey, Kalifornia, 7–10 marca 2007). Treść tego referatu jest godna polecenia w całości, ponieważ zwięzłe, a jednocześnie lekko opisuje ogół problemów, jakich do-

dawanie dat kolejnych zmian i ich obrazu⁶⁷ wraz z możliwością odtwarzania stanu źródła w konkretnym dniu. W dłuższej perspektywie czasu na pewno warto byłoby rozważyć jednak trwałe archiwizowanie dynamicznego zautomatyzowanego PZI w wersji statycznej (na przykład w postaci plików w formacie .pdf) w jakimś interwale czasowym, co w efekcie uczyniłoby go pełnoprawnym dokumentem elektronicznym.

7. WNIOSKI KOŃCOWE

Opisane procedury naprawcze powinny zostać wdrożone jak najszybciej ze względu na poważne skutki, jakie niesie zaniechanie podjęcia takich działań, albowiem przytoczone w tym tekście zjawiska prowadzą wprost do utraty wiedzy pozanarzędziowej — wiedzy płynącej nie z twardych danych faktograficznych, lecz z doświadczenia i mądrości specjalisty przygotowującego źródło, a odzwierciedlonej w nadanej przez niego makro- i megastrukturze dokumentu. Wprawdzie tworzenie tych warstw lub ich reprezentantów wymaga dużego nakładu pracy i sporych kompetencji przygotowującego, ale w zamian skutkuje nie tylko niebagatelnym wzrostem waloru informacyjnego PZI, lecz także jego użyteczności w ogóle. Koncentracja wyłącznie na walorze wyszukiwawczym jest krótkowzroczna: nie od dziś wiadomo, że niektórzy użytkownicy PZI mają problem z rozróżnieniem kryteriów formalnych i rzeczowych stosowanych podczas procesu wyszukiwawczego. W ich właściwym stosowaniu pomagały właśnie te, zaniedbywane dziś w zautomatyzowanych PZI, elementy. W efekcie użytkownik, pozostawiony sam sobie i nieprowadzony niewidzialną ręką specjalisty, nie mogąc skorzystać z podszeptów płynących z makro- i megastruktury, myli miejsce wydania dokumentu z miejscem przedmiotu dokumentu (Wrocław jako miejsce wydania i jako fragment rzeczywistości, mający na przykład swoją historię, ludność, kulturę), osobę jako podmiot (autor dzieła) z osobą jako przedmiot (temat dzieła) czy datę wydania z datą przedmiotu (na przykład rok 1683 jako data wydania i jako czas wydarzenia się bitwy pod Wiedniem). W tej sytuacji intuicyjność w grupowaniu materiałów może prowadzić do wadliwych wyników. Oprócz tego że współczesne wyszukiwanie ma być intuicyjne, ma być także dynamiczne i bez wątplenia prowadzi ono do szybszych wyników niż wyszukiwanie statyczne. Tracimy jednak wówczas ów „statyczny obraz” w danym miejscu i danym czasie.

Obecnie wielkie nadzieje pokłada się w sieciach semantycznych, które mają być rozwiązaniem przynajmniej części bolączek trapiących użytkowników SIW. O ile tradycyjna sieć (a w niej między innymi automatyczne PZI) potrafi jedynie

⁶⁷ Tu jednak należałoby rozważyć, czy jest sens archiwizacji wszystkich kolejnych wersji (nawet tych, w których poprawki ograniczały się jedynie do usunięcia błędu ortograficznego lub literówki), czy tylko tych znaczniejszych, gdzie modyfikacja dotyczyła sensu treści.

dostarczać informację, bez rozróżniania jej znaczenia i niuansów oraz oceny jej przydatności, a zatem nie potrafi jej interpretować, o tyle sieć semantyczna ma umożliwiać maszynom przetwarzanie informacji w sposób odpowiedni do ich znaczenia. Jednak wiara w sieci semantyczne jest naiwna: nawet najbardziej inteligentna sieć będzie wciąż sztuczną inteligencją (ang. *artificial intelligence*, AI) i nigdy, wykorzystując tylko metody automatyczne, nie udzieli całkowicie relewantnej odpowiedzi na zapytanie zadane SIW. Cytowany już Lev Manovich, wypowiadając się o konieczności rozumienia przez komputer znaczenia zawartego w generowanych obiektach, pisze wprost:

trudno mówić o znaczących sukcesach w tej dziedzinie. [...] W latach 90. XX wieku przeciętny użytkownik komputera mógł zetknąć się ze sztuczną inteligencją nie dzięki interfejsowi człowiek-komputer, ale raczej w grach komputerowych. Niemal każda komercyjna produkcja zawiera „moduł sztucznej inteligencji”, czyli fragment kodu odpowiedzialny za zachowanie postaci. [...] Postaci w grach komputerowych — tak jak eksperckie systemy AI — dysponują wiedzą w kilku ściśle określonych, bardzo wąskich dziedzinach. Ale ponieważ gry komputerowe są w bardzo wysokim stopniu skodyfikowane i oparte na regułach, postaci te funkcjonują bardzo efektywnie, to znaczy bardzo efektywnie reagują na dozwolone polecenia użytkownika, takie jak: „Idź”, „Strzelaj”, „Podnieś przedmiot”. Postaci nie potrafią nic więcej, ale gra skonstruowana jest tak, że użytkownik nie ma możliwości, by się o tym przekonać. Na przykład, komputerowe gry walki i bijatyki nie przewidują zadawania pytań przeciwnikom ani — tym bardziej — prowadzenia rozmów. Można tylko „atakować” przeciwnika, naciskając przyciski, a komputer z pewnością odpowie „atakiem” i w tej — rządzącej się prostymi regułami — symulacji robi to bardzo efektywnie. Podsumowując, komputerowe postaci mogą się wydawać inteligentne i obdarzone różnymi umiejętnościami tylko dlatego, że programy bardzo ograniczają nasze kontakty z nimi. Mówiąc inaczej, komputery mogą udawać istoty inteligentne tylko dlatego, że komunikując się z nimi, używamy bardzo małej części naszego człowieczeństwa⁶⁸.

I to jest zasadniczy problem sieci semantycznych, a być może automatyzacji w ogóle.

Co może wydawać się zaskakujące, już Karol Estreicher senior, nie mając o tym pojęcia, poległ na problemie sieci semantycznych. Otóż z analizy Marii Dembowskiej dotyczącej odsyłaczy przedmiotowych w *Bibliografii Polskiej* wynika, że — przykładowo — na trzynastcie pozycji zarejestrowanych pod hasłem „Kamień”⁶⁹ tylko dwie dotyczą geologicznego znaczenia tego wyrazu — pozostałe traktują o kamieniach żłociowych, kamieniach węgielnych oraz stosują go w znaczeniu przenośnym⁷⁰. Wprawdzie do tego błędu doprowadziło Estreichera stosowanie opisu treści dokumentu tylko na podstawie wyrazu głównego tytułu, jednak doskonale obrazuje problem przekładania jednostek języka naturalnego na język sztuczny, a jeszcze pamiętajmy o różnicach występujących między językiem

⁶⁸ L. MANOVICH, *op. cit.*, s. 99–100. Wyróżnienie własne.

⁶⁹ Karol ESTREICHER, *Bibliografia Polska*. T. 19 (K–Kom), Kraków: Drukarnia Uniwersytetu Jagiellońskiego 1903, s. 73.

⁷⁰ M. DEMBOWSKA, *op. cit.*, s. 77–78.

naturalnym a językiem specjalistycznym, i związanym z nim zagadnieniem terminów i terminologii. Na zakończenie tych rozważań trzeba jeszcze przypomnieć o problemie trwałości nowych nośników oraz odczytu starzejących się technologii, którego nie trzeba tu szerzej przedstawiać, gdyż jest on powszechnie znany.

8. PODSUMOWANIE

Na — podkreślmy to — niewłaściwej automatyzacji PZI tracimy porządek, wiedzę, powiązania logiczne, a finalnie — informacje. Doświadczył tego nawet Umberto Eco, który upatruje w semiotyce możliwości zapobiegania zgubnym skutkom płynącym z niesprawnego wykorzystywania sieci, a w tym zauto-matyzowanych PZI:

Gdy pewnego dnia szukałem w sieci jakiejś informacji, wpadłem w pułapkę klikania w kolejne odesłania i linki. Poznałem mnóstwo ciekawych stron, ale przy okazji straciłem z pola widzenia cel i ideę moich poszukiwań. I właśnie jednym z obowiązków semiotyki będzie nie tylko odnajdywanie znaków, ale też edukowanie, jak wrócić do źródeł ich szukania⁷¹.

Właśnie przed taką pułapką powinni chronić użytkowników twórcy zauto-matyzowanych źródeł informacji o charakterze pochodnym, a oprócz tego, przez właściwie skonstruowaną i przyjazną makro- i megastруктурę źródła, umożliwiać im heurystyczne samokształcenie.

JAKUB MACIEJ LUBOCKI

OLD INFORMATION SPACES. WHAT DO WE LOSE DUE TO INAPPROPRIATE AUTOMATISATION OF DERIVATIVE SOURCES OF INFORMATION?

Summary

The belief that the very fact that information has been made available in the dynamic internet environment is enough for the users to find it may lead to very serious misunderstandings. Hence the importance of the reflection on what we lose because of careless and ill-considered automatisa-tion of derivative sources of information (bibliographies/catalogues, dictionaries, encyclopaedias). Admiration for modern technologies has, on the one hand, made us abandon the huge informational value of the higher level of data organisation (Googlisation and granulation) required by the printed form of such sources and on the other it has led to a situation in which we fail to use the potential of these technologies, impossible to realise in the printed form (skeuomorphism). This leads to the disappearance of the additive nature of the information contained in them and its spread, heuristic impotence stemming from the seemingly intuitive nature of the handling of the source as well as the loss of the archival image of the state of affairs at a given time and place, hitherto guaranteed

⁷¹ Karolina TATARZYŃSKA, *Umberto Eco doktorem UŁ*, Gazeta Wyborcza Łódź (120) 25 V 2015, s. 3.

by the durable printed form. All these processes cause a loss of non-tool-related knowledge. The author uses examples to demonstrate the losses — perhaps irrevocable — to the specialist information toolkit and ways of preventing them. The paper is preceded by a short overview of positions concerning the relations between the terms “automatisation” and “digitisation”.

KEY WORDS: informational publications, derivative sources of information, automatisation of bibliography, automatisation of dictionaries, automatisation of encyclopaedias