

## Ukryty Internet – nowe podejście

### Ukryty Internet – słowo wstępu i stan badań

Podejście do zagadnienia Ukrytego Internetu zmieniało się wielokrotnie, od pierwszego użycia tego terminu<sup>1</sup> do chwili obecnej. Występowały różnice w terminologii, zarówno anglojęzycznej (w której funkcjonują takie terminy, jak: *Invisible Web*, *Deep Web*, *Hidden Internet*), jak i polskiej (Ukryty Internet nazywany jest także siecią głęboką). W warstwie terminologicznej pojawiają się także nieścisłości. Ukrytym Internetem (ang. *Deep Web*) nazywany jest błędnie Ciemny Internet (ang. *Dark Web*)<sup>2</sup>, anonimowa sieć oparta na protokole *The Onion Router* (TOR). Rozbieżności ujawniły się także w obszarze definicyjnym tego pojęcia. Michael K. Bergman uważa, że Ukrytym Internetem (*Deep Web*) są strony wykreowane dynamicznie w wyniku wyszukiwania szczegółowego, prowadzonego przy użyciu wewnętrznych interfejsów<sup>3</sup>. Inne spojrzenie prezentują Chris Sherman i Gary Price. Definiują oni Ukryty Internet (*Invisible Web*) jako „strony, pliki i inne, często wysokojakościowe informacje, dostępne poprzez sieć WWW, których wyszukiwarki internetowe nie mogą, z powodów ograniczeń technologicznych, albo nie chcą, z powodów uznaniowych, dodawać do swoich indeksów”<sup>4</sup>. Podobne stanowisko

<sup>1</sup> Michael K. Bergman podaje, że w dzisiejszym jego rozumieniu zrobiła to Jill H. Ellsworth w 1994 r. Zob.: M.K. Bergman, *The Deep Web: surfacing hidden value*, The Journal of Electronic Publishing 2001, 1(7) [online] [dostęp: 15.05.2014] Dostępny w World Wide Web: <http://hdl.handle.net/2027/spo.3336451.0007.104>.

<sup>2</sup> Grupa Tense, *Tor dla opornych – sekrety ukrytego Internetu*. Gazeta.pl 15.11.2014 [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: [http://technologie.gazeta.pl/internet/1,113840,16969824,Tor\\_dla\\_opornych\\_sekrety\\_ukrytego\\_Internetu.html](http://technologie.gazeta.pl/internet/1,113840,16969824,Tor_dla_opornych_sekrety_ukrytego_Internetu.html); Bright Planet, *The Deep Web: What was House of Cards Talking About?* [online] 18.02.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.brightplanet.com/2014/02/deep-web-house-of-cards/>; *Anonimowa strona głębokiego Internetu – dla ciekawych*, [w:] *Wyszukiwanie w Głębokim Internecie* (szkolenie e-learningowe na platformie EBIB.pl) [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://moodle.ebib.pl/mod/lesson/view.php?id=49>; Por. też: J. Devine, F. Egger-Sider, *Going Beyond Google again*, Chicago 2014, s. 15.

<sup>3</sup> M.K. Bergman, op. cit.

<sup>4</sup> Ch. Sherman, G. Price, *The Invisible Web: uncovering information sources search engine can't see*, Medford, New Jersey 2003, s. 57.

definicyjne reprezentują Jane Devine i Francine Egger-Sider<sup>5</sup>. W literaturze pojawiła się także koncepcja Akademickiego Ukrytego Internetu (ang. *Academic Invisible Web*, dalej AIW), przedstawiona przez Dirka Lewandowskiego i Philippa Mayra<sup>6</sup>. Zdefiniowali go jako zasób przeznaczony dla środowiska akademickiego, składający się ze wszystkich baz danych i zbiorów informacji opublikowanych lub zdeponowanych online, których nie można wyszukać przy użyciu ogólnodostępnych wyszukiwarek internetowych<sup>7</sup>. Lewandowski i Mayr zauważyli, że trudno mówić o sztywnych granicach pomiędzy AIW a naukową siecią powierzchniową<sup>8</sup>.

Dyskutowany był także rozmiar Ukrytego Internetu. Pierwsze badania metodą próbkowania i porównania są zasługą M.K. Bergmana. Oszacował on, że zasoby sieci głębokiej są od 400-550 razy większe od sieci powierzchniowej<sup>9</sup>. Dane te szybko zakwestionował Ch. Sherman, wykazując, że rozmiar został przeszacowany i proporcja pomiędzy zasobami sieci przeszukiwalnej a Ukrytym Internetem wynosi 1:50<sup>10</sup>. Wyliczenia Bergmana podważył także Wolfgang G. Stock, uważając je za dziesięciokrotnie zawyżone<sup>11</sup>. Dokładniejsze rozważania na temat rozmiaru zasobów głębokich zaprezentowali jednak dopiero Lewandowski i Mayr, podając przy tym także wyliczenia dotyczące rozmiaru AIW. Ich zdaniem dane Bergmana są przeszacowane, a Akademicki Ukryty Internet można przyrównać rozmiarowo do indeksu dużej wyszukiwarki internetowej ogólnego przeznaczenia<sup>12</sup>.

## Ukryty Internet i sieć powierzchniowa – zmiany i relacje

Na przestrzeni lat zmieniało się podejście badaczy do zjawiska Ukrytego Internetu, ale także zmienił się Ukryty Internet. Wiele zasobów opisywanych jako nieindeksowane przez wyszukiwarki stało się częścią sieci powierzchniowej. Nowe algorytmy indeksujące wyszukiwarek zaczęły uwzględniać obszary opisywane jako sieć głęboka. Ukrytym Internetem nadal pozostają zasoby indeksowane w bazach danych, które nie posiadają indeksu w formie hipertekstowej, a wyszukiwanie informacji w ich zasobach odbywa się wyłącznie przy użyciu dedykowanych interfejsów. Opisywana przez Shermana i Price'a sieć nieprzezroczysta (ang. *The*

<sup>5</sup> J. Devine, F. Egger-Sider, *Beyond Google: The Invisible Web in the academic library*, *The Journal of Academic Librarianship* 2004, vol. 4, s. 265.

<sup>6</sup> D. Lewandowski, P. Mayr, *Exploring the academic invisible web*, *Library Hi Tech*, 2006, vol. 4, s. 532.

<sup>7</sup> Ibidem, s. 532.

<sup>8</sup> Ibidem.

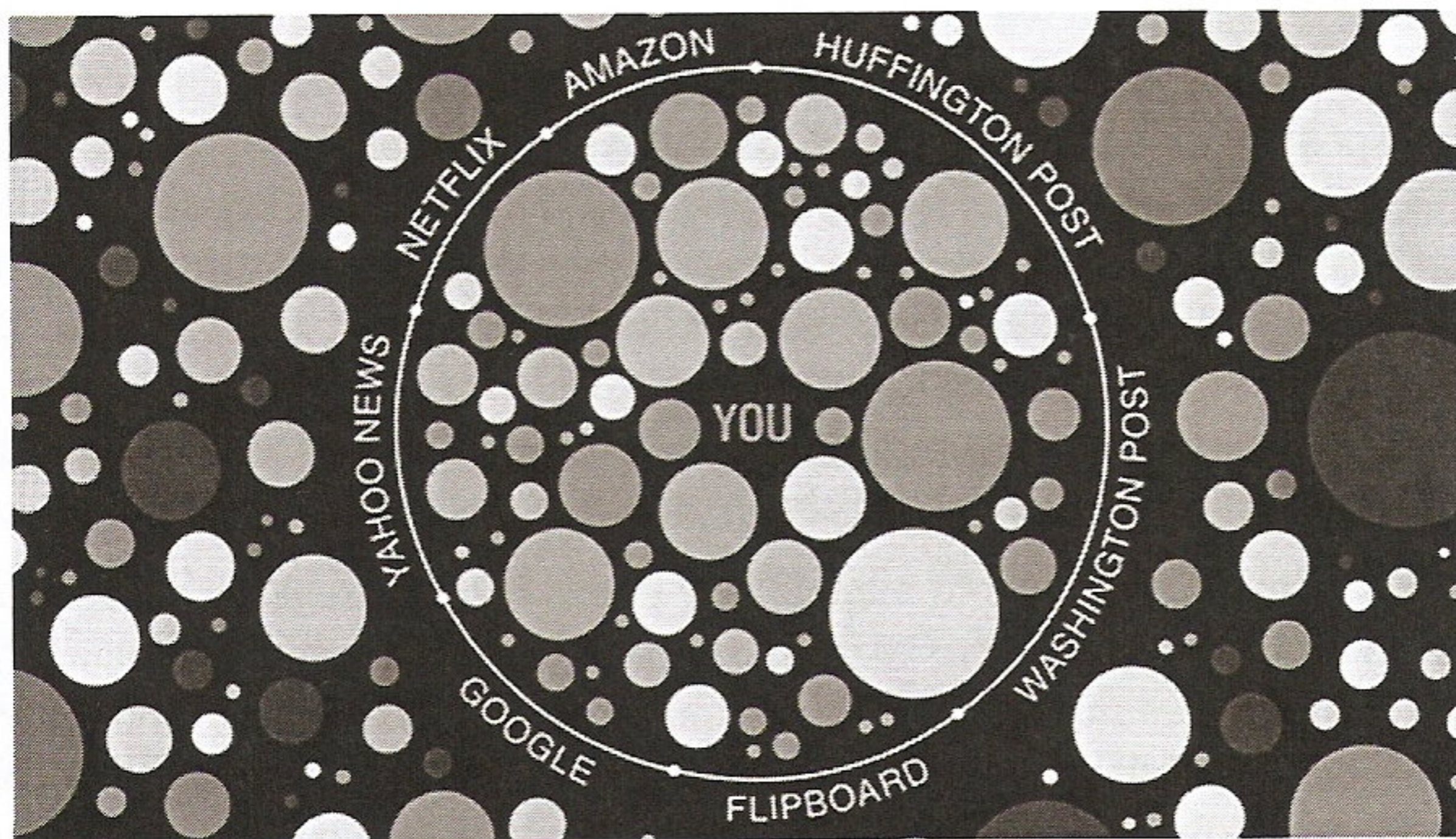
<sup>9</sup> M.K. Bergman, op. cit.

<sup>10</sup> Ch. Sherman, *Search for the Invisible Web*, *The Guardian* [online] 06.09.2001 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.theguardian.com/technology/2001/sep/06/internetnews.onlinesupplement>.

<sup>11</sup> W.G. Stock, *Weltregionen des Internet: Digitale Informationen im WWW und via WWW*, *Password* 2003, vol. 2, s. 27.

<sup>12</sup> D. Lewandowski, P. Mayr, op. cit., s. 533-536.

*Opaque Web*)<sup>13</sup> nadal pozostaje nieindeksowana ze względów uznaniowych. Limitowanie indeksowania w głąb witryny zależy od wielu różnych czynników, m.in. położenia podstrony w hierarchicznym układzie witryny oraz wartości PageRanku osiąganego przez stronę główną serwisu i jego podstrony<sup>14</sup>. Do kategorii zasobów opisywanych przez Shermana i Price'a jako *Opaque Web* należałoby jeszcze dołożyć zasoby spychane w wyniku wojen pozycjonerskich z kilku pierwszych stron wynikowych na strony dalsze, z których korzysta niewielki odsetek użytkowników wyszukiwarek. W tym kontekście zupełnie nową kategorią niewidzialności są zasoby pozostające poza osobistymi bańkami filtrującymi (ang. *The Filter Bubbles*) użytkowników. Algorytm wykorzystywany do powszechnej personalizacji wyników wyszukiwania w wyszukiwarce Google został udostępniony 4 grudnia 2009 roku<sup>15</sup>. W praktyce bańka filtrująca oznacza dobieranie wyników wyszukiwania do zgromadzonych przez system informacji o użytkowniku<sup>16</sup>. Istnienie mechanizmów personalizujących zaburza wielokontekstowy i obiektywny obraz rzeczywistości (por. rys. 1), stwarza też nową przestrzeń niewidzialności.



Rys. 1. Schemat bańki filtrującej

Źródło: E. Pariser, *Beware online "filter bubbles"*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.youtube.com/watch?v=B8ofWfx525s>

<sup>13</sup> Ch. Scherman, G. Price, op. cit., s. 70-72.

<sup>14</sup> E. Enge, *Matt Cutts Interviewed by Eric Enge*, Stone Temple Consulting [online] 14.03.2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.stonetemple.com/articles/interview-matt-cutts-012510.shtml>.

<sup>15</sup> E. Pariser, *The Filter Bubble: what the Internet is hiding from You*, London 2012, s. 3.

<sup>16</sup> Szczegółowy opis parametrów personalizacji zamieszczono w tekście: N. Pamuła-Cieślak, *Analiza czynników wpływających na jakościowe cechy wyników w wyszukiwarce Google*, Toruńskie Studia Bibliologiczne 2013, nr 1, s. 118-119.

Część powierzchniowa WWW zmienia się w Ukryty Internet, choć zasoby, o których mowa, są indeksowane i wyszukiwane. Wybiórczość ich wyświetlania użytkownikom sprawia, że stają się one prawdziwie ukryte, a niewidzialność ta jest zmienna, inna dla każdego użytkownika. Należy też zauważyć, że bańka filtrująca nie dotyczy wyłącznie wyszukiwarek takich jak Google, Bing czy Yahoo! Istnieje także w mediach społecznościowych (np. Facebook, Instagram, Google+, Flipboard), platformach medialnych (The Huffington Post, Yahoo! News, The Washington Post), portalach i serwisach usługowych (Amazon, Netflix)<sup>17</sup>.

## Ukryty Internet a media społecznościowe

Media społecznościowe to termin, pod którym kryją się różnorodne serwisy posiadające opcje dzielenia się i udostępniania treści przez użytkowników. David Nicholas i Ian Rowlands wyróżnili osiem typów serwisów społecznościowych<sup>18</sup>.

Są to:

- sieci społecznościowe (Facebook, LinkedIn);
- serwisy blogowe (Wordpress, Blogger);
- serwisy mikroblogowe (Twitter);
- współdzielone narzędzia do tworzenia treści (Google Docs);
- serwisy zakładkowe i wykorzystujące folksonomię (Delicious, CiteULike, Digg, Pearltrees);
- narzędzia do spotkań i konferencji online (Skype);
- społecznościowe planery czasu (Doodle, Kalendarz Google);
- serwisy umożliwiające publikowanie i dzielenie się plikami graficznymi i multimediami (Flickr, YouTube, Vimeo).

Media społecznościowe to niejednolite pod względem funkcji, oprogramowania czy architektury środowisko informacyjne. Większość z nich pozwala użytkownikom na dobranie ustawień prywatności według osobistych preferencji. Opcja ta występuje jako jedna z domyślnych funkcjonalności lub jest udostępniana wyłącznie użytkownikom posiadającym płatne konto w serwisie.

Do sieci powierzchniowej można zaliczyć bez wątpliwości publicznie jawne treści publikowane na blogach i w serwisach zakładkowych. Odróżniającym się od nich typem zasobów są grafiki i multimedia, udostępniane w stosownych serwisach o charakterze społecznościowym. W ich przypadku indeksowanie opiera się na innych mechanizmach, ale algorytmy agregujące różne typy treści do ogólnych wyników wyszukiwania (np. *Universal Search* w wyszukiwarce Google)

<sup>17</sup> Por. rys. 1.

<sup>18</sup> D. Nicholas, I. Rowlands, *Social Media use in the research workflow*, Information Services & Use 2011, vol. 1-2, s. 63.

także akustycznie je wyszukują i wyświetlają użytkownikom. Do sieci głębokiej należą zaś w pewnym procencie pozostałe serwisy społecznościowe. Społecznościowy Ukryty Internet można podzielić na dwie części. W jednej z nich znajdują się zasoby, których status wyświetlania został ograniczony przez użytkowników (np. prywatny, dostępny w określonym gronie odbiorców) – ma on charakter sieci prywatnej (*The Private Web*) lub zastrzeżonej (*The Proprietary Web*)<sup>19</sup>. Badania pokazują jednak wyraźnie, że nawet pośród postów i zasobów publicznie udostępnionych wskaźnik indeksowania w wyszukiwarce Google nie wynosi 100%. Indeksowanie treści pochodzących z serwisu Facebook zbadał Eric Enge dla serwisu Stone Temple Consulting<sup>20</sup>. Wykazał, że indeksowane są publiczne dane udostępniane w profilach. Najwyższy PageRank mają profile popularne, mające setki tysięcy polubień. Maksymalny wskaźnik indeksowania ilości postów dla tych najpopularniejszych kont wynosi ok. 85%, jeśli zawierają one odsyłacze do innych zasobów. Posty z profili popularnych niezawierające odnośników są indeksowane maksymalnie w 60%. Wskaźnik ten znacząco spada dla wpisów z profili posiadających mniej polubień. E. Enge przeprowadził także badania indeksowania treści z serwisu mikroblogowego Twitter<sup>21</sup>. Wskaźnik indeksowania wpisów (twittów) z tego serwisu okazał się bardzo niski. Najlepiej indeksowane są posty, które pochodzą z bardzo wpływowych profili, posiadających ponad 1 milion obserwujących użytkowników. W przypadku mikroblogów, których użytkownicy publikują wpisy „na gorąco”, poza wskaźnikiem indeksacji, liczy się także szybkość publikowania i upowszechnienia w sieci. Ten sam wskaźnik ważny jest przy ich wyszukiwaniu. Badanie szybkości indeksowania twittów przeprowadzono dla kont bardzo popularnych, posiadających ponad 5 mln śledzących je użytkowników. Niezbyt wysokie wskaźniki aktualizacji (6% wpisów zaindeksowanych w ciągu 24 h oraz 15% w ciągu 48 h) wskazują na to, że społecznościowe serwisy mikroblogowe mogą być uznane za fragment Ukrytego Internetu. Ma to szczególne znaczenie wobec faktu, że są one coraz częściej i w coraz większym stopniu wykorzystywane do celów zawodowych, edukacyjnych czy naukowych<sup>22</sup>. Wy-

<sup>19</sup> Oba te typy sieci głębokiej pochodzą z typologii zaproponowanej przez Ch. Shermana i G. Price’a. Zob. Ch. Sherman, G. Price, op. cit., s. 73-74.

<sup>20</sup> E. Enge, *Does Facebook Activity Impact SEO?*, Stone Temple Consulting [online] 12.10.2013 [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.stonetemple.com/does-facebook-activity-impact-seo/>.

<sup>21</sup> E. Enge, *How Does Google Index Tweets?*, Stone Temple Consulting [online] 07.09.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.stonetemple.com/how-does-google-index-tweets/>.

<sup>22</sup> Por. D. Nicholas, I. Rowlands, op. cit., s. 81-82; J. Devine, F. Egger-Sider, *Going Beyond Google again*, op. cit., s. 45-49; S.D. Smith, J.B. Caruso, *The ECAR Study of Undergraduate Students and Information Technology*, [online] 2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://anitacrawley.net/Resources/Reports/ECAR%20study%20highlights.pdf>.

szukiwarki typu ogólnego nie są natomiast w pełni przygotowane lub dopuszczone do indeksowania zamieszczanych tam treści.

Przestrzeń serwisów społecznościowych ma szczególne znaczenie dla rozważań nad zagadnieniem Ukrytego Internetu i wyszukiwania informacji. Już w 2010 roku Facebook znalazł się na pierwszym miejscu listy stron, od której internauci rozpoczynają swoją codzienną aktywność w przestrzeni internetowej<sup>23</sup>. Serwisy skupiające dużą społeczność użytkowników wyrastają na konkurencyjną, w stosunku do sieci WWW, przestrzeń informacyjno-komunikacyjną. Posiadają:

- własną architekturę i strukturę, sposoby komunikowania się pomiędzy użytkownikami (wiadomości w serwisie lub poprzez zintegrowany z nim komunikator);
- możliwość dzielenia się, tworzenia i prezentacji treści (profile osobiste, strony – tzw. fanpejdże);
- miejsca dyskusji osób o podobnych poglądach i zainteresowaniach (grupy);
- system wyszukiwania i powiadamiania o zmianach (wyszukiwarka – w tym Facebook Social Graph, powiadomienia o aktualizacjach);
- przestrzeń reklamową i promocyjną, rozrywkę (gry i aplikacje).

Ponadto zostały doskonale przystosowane do powszechnie używanych urządzeń mobilnych, co daje im przewagę nad systemem WWW, którego składowe są w różnej fazie dostosowania do dynamicznych zmian technologii dostępowej. W niedalekiej przyszłości możliwe jest zastąpienie przez serwisy społecznościowe dotychczasowego modelu funkcjonowania Internetu. Obecnie praktykowane wyszukiwanie informacji, oparte w dużym stopniu na narzędziach automatycznie indeksujących treści, straciłoby na znaczeniu. Wobec tak radykalnych zmian należy rozważyć pojawienie się nowego modelu wyszukiwania i nowych przestrzeni niewidzialności.

## Mobilny Internet a problematyka sieci głębokiej

Coraz więcej ludzi wykorzystuje urządzenia mobilne podczas codziennych aktywności o różnym charakterze, a szczególny wzrost ich użycia zauważalny jest w obszarze serwisów społecznościowych<sup>24</sup>. Środowisko mobilne i WWW to dwa odrębne, luźno powiązane ze sobą ekosystemy. Pełnią one podobne funkcje

<sup>23</sup> Y.Q. Mui, P. Whoriskey, *Facebook passes Google as most popular site on the Internet, two measures show*, [online]. 31.12.2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.washingtonpost.com/wp-dyn/content/article/2010/12/30/AR2010123004645.html>.

<sup>24</sup> *State of the media: the social media report 2012*, [online] 12.03.2012 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2012-Reports/The-Social-Media-Report-2012.pdf>.

i służą realizacji tych samych potrzeb, lecz sposób, w jaki to robią, diametralnie się różni. Użytkownicy przenoszą się z otwartej przestrzeni systemu hipertekstowego do bardziej hermetycznej przestrzeni informacyjno-komunikacyjnej skupionej w aplikacjach. Warto przyjrzeć się temu obszarowi pod kątem powstawania w nim nowych przestrzeni Ukrytego Internetu.

Środowisko mobilne to ogromne zmiany modelu korzystania z udostępnionych treści. Dotyczy to także wyszukiwania informacji (por. tab. 1).

Tab. 1. Porównanie różnych funkcji i zastosowań dwóch środowisk publikacji treści

	Wcześniej (WWW)	Teraz (Środowisko mobilne)
<b>Dostęp do treści</b>	Przeglądarka	Aplikacje
<b>Powiadomienia o publikacji</b>	Web syndication	Subskrypcje
<b>Wyszukiwanie</b>	Google!!!	Google??? Facebook???
<b>Model ekonomiczny dostępu do treści</b>	Darmowy dostęp	Dostęp freemium
<b>Język programowania</b>	Javascript	Objective-C
<b>Język formatowania treści</b>	HTML	XML

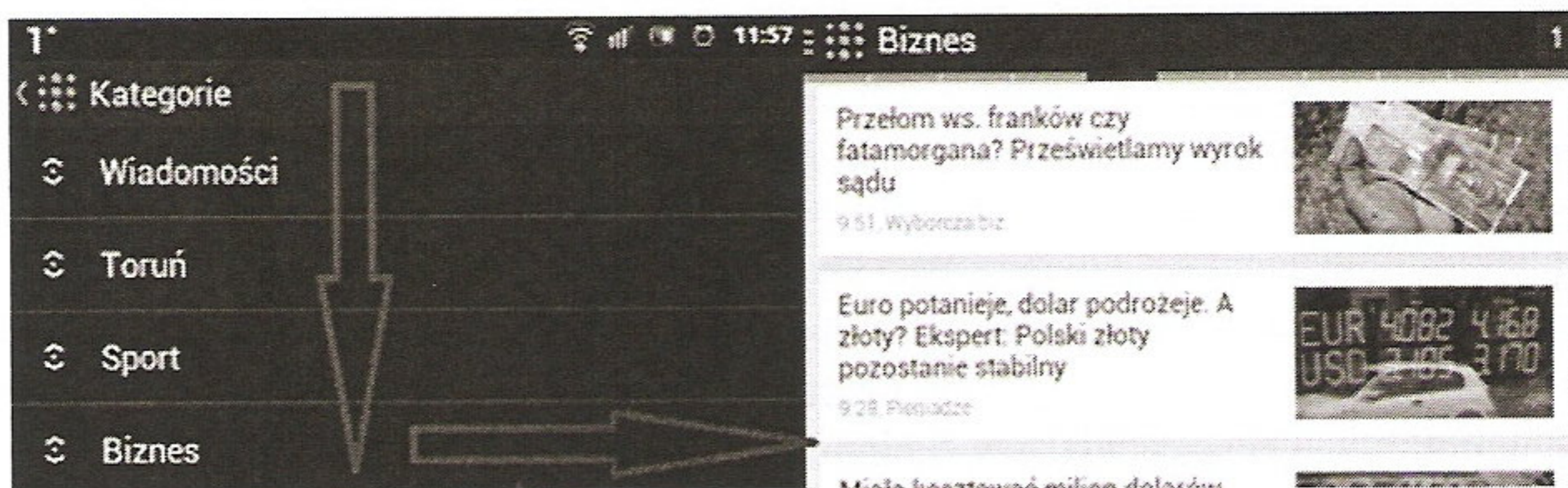
Źródło: opracowanie własne na podstawie: Ch. Anderson, M. Wolff, *The Web is dead. Long live the Internet*, Wired Magazine [online] 17.08.2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: [http://www.wired.com/2010/08/ff\\_webrip/](http://www.wired.com/2010/08/ff_webrip/)

Aplikacje mobilne to nowy sposób publikowania i korzystania z treści. Jak wynika z tab. 1, istnieją wątpliwości co do przyszłej, wiodącej marki i technologii wyszukiwawczej. Zasoby treściowe wewnątrz aplikacji sprawiają robotom wyszukiwarek podobne problemy, jak zasoby umieszczane wewnątrz baz danych. Dzieje się tak dlatego, że w aplikacjach umieszcza się odsyłacze do zasobów udostępnianych w sieci WWW, ale nie ma odsyłaczy prowadzących od witryn do aplikacji. Nie używa się także odnośników z aplikacji do aplikacji, choć nowe formaty programowania pozwalają je zagnieźdźać<sup>25</sup>. Przejście użytkowników i wydawców treści z witryn WWW do mobilnego Internetu wymusi zmiany w modelach i technologiach wyszukiwawczych. Obecnie trwają prace nad rozwiązaniami umożliwiającymi skanowanie treści udostępnianych w aplikacjach, zaprojektowanych

<sup>25</sup> J. Battelle, *What might a Facebook search engine look like?* [online] 16.01.2012 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://battellemedia.com/archives/2012/01/what-might-a-facebook-search-engine-look-like.php>.

dla urządzeń z systemem operacyjnym Android<sup>26</sup>. Istnieją jednak inne mobilne systemy operacyjne, których ta technologia nie obejmuje.

Mobilny Internet już wymusza zmiany, tak na rynku wyszukiwawczym, jak i w użytkownikach informacji. Wyszukiwanie odpowiedniej treści w aplikacjach różni się od przeglądania tych samych treści na stronie WWW. Przypomina bardziej szukanie przedmiotowe w obrębie dziedziny lub zagadnienia, ze względu na pionową (por. rys. 2) lub kafelkową architekturę informacji stosowaną w aplikacjach<sup>27</sup>.



Rys. 2. Pionowe wyszukiwanie kategorii i poziome jej rozwinięcie na przykładzie aplikacji mobilnej Gazeta.pl

Źródło: opracowanie własne

Zmianie ulega również wyszukiwanie informacji. Powstają wyszukiwarki dedykowane mobilnemu Internetowi (np. Mobile Content Networks<sup>28</sup>, Evi<sup>29</sup>), do sieci mobilnej wkraczają też marki znane z rynku wyszukiwawczego WWW. Już teraz dostosowują swoje narzędzia do specyfiki nowego środowiska informacyjnego, poprzez wprowadzanie opcji wyszukiwania głosowego czy limitowanie liczby rezultatów wyszukiwania w stosunku do pełnych wersji swoich wyszukiwarek WWW<sup>30</sup>. Takie ograniczenia powiększają zasięg sieci głębszej w strefie Internetu mobilnego.

Do narzędzi wyszukiwania informacji dla urządzeń mobilnych można zaliczyć także platformy udostępniające aplikacje. Najpopularniejszymi z nich, prze-

<sup>26</sup> G. Sterling, *Quixey Offering Deeper Search Results Inside Mobile Apps* [online] 25.02.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://searchengineland.com/quixey-offering-deeper-search-results-inside-mobile-apps-185160>; idem, *Google Sees Deeper App Content As Key To Mobile Usage* [online] 10.03.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://searchengineland.com/google-seeks-reassert-mobile-deeper-app-indexing-186102>.

<sup>27</sup> J. Devine, F. Egger-Sider, *Going Beyond Google again*, s. 49.

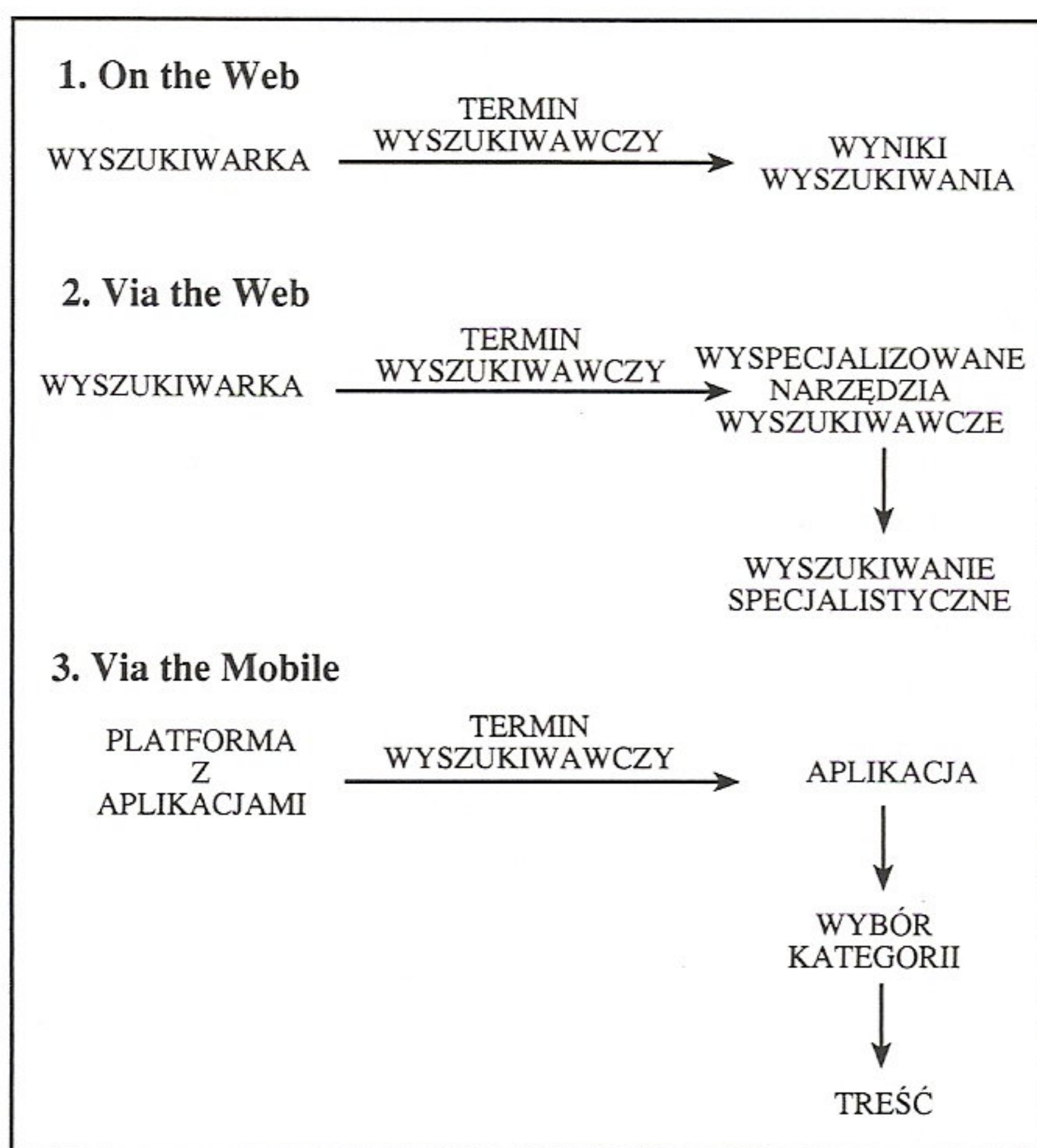
<sup>28</sup> *Mobile Content Networks*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://mcn-inc.com>.

<sup>29</sup> *Evi*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.evi.com>.

<sup>30</sup> A. Aldrich, *Universities and Libraries Move to the Mobile Web*, *Educase Quarterly* 2010, vol. 2 [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.educause.edu/ero/article/universities-and-libraries-move-mobile-web>.



znaczonymi dla trzech systemów operacyjnych, są: Google Play<sup>31</sup> (Android), iTunes<sup>32</sup> (iOS) oraz sklep z aplikacjami dla Windows Phone<sup>33</sup>. Ich funkcjonowanie w obszarze wyszukiwania mobilnego implikuje powstanie nowego modelu docierania do treści, co wiąże się także z powstaniem nowej strategii wyszukiwawczej. Do funkcjonującej w WWW strategii *On the Web*, polegającej na formułowaniu prostych zapytań w wyszukiwarce i przeciwstawionego wzorca wieloetapowego wyszukiwania poprzez sieć (ang. *Via the Web*), dołączyć można model wyszukiwania *Via the Apps* (patrz rys. 3). Chcąc porównać tę ostatnią strategię do funkcjonowania narzędzi wyszukiwawczych w systemie webowym, można użyć analogii do katalogów internetowych (ang. *directories*).



Rys. 3. Porównanie dwóch strategii wyszukiwawczych funkcjonujących w WWW z zaproponowanym modelem strategii *Via the Apps*, dedykowanym wyszukiwaniu w aplikacjach mobilnych

Źródło: opracowanie własne

<sup>31</sup> *Google Play*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://play.google.com/store>.

<sup>32</sup> *iTunes*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.apple.com/itunes/>.

<sup>33</sup> *Aplikacje dla systemu Windows Phone – Sklep Microsoft*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.windowsphone.com/store>.

Platforma z aplikacjami to katalog, w którym należy odszukać odpowiednie narzędzie (aplikację), aby dotrzeć do zawartej w niej treści. *Via the Apps*, podobnie jak *Via the Web*, jest wyszukiwaniem horyzontalno-wertykalnym, podczas gdy *On the Web* ma charakter liniowy.

W odniesieniu do wykorzystywania urządzeń mobilnych w procesie wyszukiwania i pozyskiwania informacji szczególnego znaczenia nabiera istnienie bańki filtrującej. Moduł GPS, lokalizacja urządzenia z wykorzystaniem technologii WiFi, konto użytkownika, pozwalające w pełni skorzystać z możliwości sprzętu czy dostęp do wiadomości e-mail, sms, rejestru rozmów telefonicznych przez aplikacje, predestynuje urządzenie mobilne do bycia jednym z najbardziej spersonalizowanych obszarów współczesnego ICT. Przekłada się to na wyszukiwanie informacji. Idea bańki filtrującej, przejawiająca się w doborze treści do użytkownika, może się w tym środowisku realizować w pełni, niosąc dla Ukrytego Internetu konsekwencje dużo dalej idące, niż to ma miejsce w środowisku WWW. To jeszcze jedna przestrzeń nowej niewidzialności.

## Semantyczny Web – wpływ konceptu na wyszukiwanie i Ukryty Internet

Jednym z motywatorów tworzenia struktury sieci semantycznej (zwanej też siecią trzeciej generacji) w obrębie World Wide Web było niedostosowanie standardów opisów dokumentów i obiektów internetowych, a zatem ograniczone możliwości ich skutecznego i trafnego wyszukiwania. Pomimo tego, że koncept semantycznej sieci znany jest już od dawna, a jego start zapowiadany był już 10 lat temu<sup>34</sup>, ciągle nie doczekał się on pełnej realizacji. Na kompletne wdrożenie sieci internetowej trzeciej generacji nie czekali jednak twórcy wyszukiwarek, którzy zastosowali w nich nowe technologie. Wyszukiwarki nazywane semantycznymi powinny posiadać co najmniej jedną z poniższych cech:

- odpowiadać na pytania zadawane w języku naturalnym (*Natural Language Search Engines*);
- grupować wyszukane zasoby w gniazda kontekstowe (*Context Search Engines*), czasem wizualizując je w różnych formach (*Visual Search Engines*);
- analizować znaczenie zapytania i wyszukanych dokumentów w oparciu o bazy wiedzy;
- przeszukiwać sieć semantyczną w oparciu o nagłówki plików RDF<sup>35</sup>.

<sup>34</sup> M. Adamczyk, *Człowiek cyfrowy*, Wprost 2002, nr 16, s. 57.

<sup>35</sup> Por. B. Gontar, J. Papińska-Kacperek, *Semantyczne wyszukiwarki internetowe*, Acta Universitatis Lodzensis: Folia Oeconomica 2011, nr 261, s. 172-175; A. Pulikowski, *Wyszukiwanie informacji w Internecie*, [w:] *Bibliotekarstwo*, red. A. Tokarska, Warszawa 2013, s. 457; N. Pamuła-Cieślak, *Wizualizacja i prezentacja informacji w wyszukiwarkach internetowych*, [w:] *Współczesne oblicza komunikacji i informacji*, red. E. Głowacka, M. Kowalska, P. Krysiński, Toruń 2014, s. 260-261.

Wyszukiwarki takie mają najczęściej charakter wertykalny, choć zdarzają się rozwiązania horyzontalne, wykorzystywane także w najpopularniejszych narzędziach (np. Graf Wiedzy Google).

Koncepcje semantyczne, zarówno w obszarze wyszukiwania, jak i całości organizacji danych online, byłyby doskonałym rozwiązaniem wielu problemów użytkowników, których dostarcza istnienie Ukrytego Internetu. Sieć trzeciej generacji i narzędzia rozumiejące zarówno zapytanie, jak i zgromadzone w systemach informacyjnych dane pozwoliłyby skrócić drogę wyszukującego od zapytania do odpowiedzi, poprzez pominięcie szeregu interfejsów. Zadaniem systemu wyszukiwawczego byłoby eksplorowanie niezbędnych baz w celu wydobycia z nich odpowiednich danych. Jest to jednak rozwiązanie przyszłościowe, na które użytkownikom Internetu przyjdzie jeszcze poczekać.

## Kognitywny Ukryty Internet – kolejna płaszczyzna niewidzialności

Poza obiektywnymi czynnikami natury technologicznej, ekonomicznej czy uznaniowej w obszarze narzędzi wyszukiwawczych, istnienie Ukrytego Internetu jest także warunkowane przez czynnik ludzki. Wydaje się, że pierwszymi badaczami sieci głębokiej, którzy zwrócili uwagę na „kognitywną niewidzialność” treści w Internecie, byli Nigel Ford i Yazdan Mansourian<sup>36</sup>. Problemem Ukrytego Internetu nie jest samo jego istnienie, lecz niewiedza użytkowników o kanałach dystrybucji informacji w sieci, metodach docierania do nich i narzędziach ułatwiających ich wyszukiwanie. Czynnikiem zniechęcającym do wieloetapowego wyszukiwania odpowiednich informacji w sieci głębokiej może być także czas i zaangażowanie, którego taki proces wymaga. Niewątpliwie przywiązanie do marki wyszukiwawczej<sup>37</sup>, zaufanie do algorytmów rankingowych wyszukiwarek<sup>38</sup>, a także swoiste lenistwo wyszukiwawcze, polegające na zaspokajaniu potrzeb informacyjnych przy użyciu jednego narzędzia, pomimo świadomości istnienia innych źródeł informacji<sup>39</sup>, tworzy obszar, określany mianem Kognitywnego Ukrytego Internetu (*Cognitive Invisible Web*)<sup>40</sup>. Winić za ten stan rzeczy można użytkowników, ale przede wszystkim systemy edukacyjne, które niedo-

<sup>36</sup> N. Ford, Y. Mansourian, *The invisible web: an empirical study of “cognitive invisibility*, *Journal of Documentation* 2006, vol. 5, s. 591.

<sup>37</sup> E. Hargittai i in., *Trust online: young adults’ evaluation of web content*, *International Journal of Communication* 2010, vol. 4, s. 479-482.

<sup>38</sup> Ibidem, s. 482-483; B. Pan i in., *In Google we trust: user’s decisions on rank, position and relevance*, *Journal of Computer – Mediated Communication* 2007, vol. 12, s. 803.

<sup>39</sup> S. Cisek, R. Sapa, *Komunikacja naukowa w Internecie – mity i rzeczywistość* [online] E-prints in Library and Information Science [dostęp: 15.11.2014] Dostępny w World Wide Web: [http://eprints.rclis.org/bitstream/10760/9148/1/cisek\\_sapa\\_2006.pdf](http://eprints.rclis.org/bitstream/10760/9148/1/cisek_sapa_2006.pdf).

<sup>40</sup> Pojęcia tego używają J. Devine i F. Egger-Sider w: *Going beyond Google again*, s. 12 i dalsze.

statecznie kształtują ich kompetencje informacyjne. Wpływ na kognitywną warstwę niewidzialności zasobów ma także z pewnością charakter Ukrytego Internetu. Rozproszenie, zalew informacyjny, mnogość i zmienność rozwiązań oraz nieintuicyjne modele wyszukiwawcze dedykowane temu obszarowi sieci sprawiają, że nie może być on postrzegany jako przyjazny użytkownikom.

## Podsumowanie

Kognitywny aspekt sieci głębokiej jest największym wyzwaniem w procesie edukacji informacyjnej, która musi nadążać za zmianami – tak w obszarze technologicznym, jak i w potrzebach i zachowaniach informacyjnych użytkowników. Należy wykorzystywać nowe technologie, zarówno w obszarze wyszukiwania, jak i kształcenia. Być tam, gdzie są użytkownicy – w aplikacjach społecznościowych na urządzeniach mobilnych, wykorzystując możliwości sieci semantycznej i nowych narzędzi wyszukiwawczych. Nie jest to zadanie łatwe. Zadając sobie pytanie o rolę bibliotekarzy i specjalistów informacji w tej zmieniającej się rzeczywistości technologiczno-internetowej, należy podkreślić, że to oni stanowią lub powinni stanowić główne ogniwo edukacyjne w nabywaniu wspomnianych kompetencji przez użytkowników. Ośrodki informacji czy biblioteki to nie tylko budynki i ludzie. Zostały one wyposażone w narzędzia, które przyniosła technologia. Już nie tylko statyczne witryny internetowe, ale także media społecznościowe, komunikatory, webinaria i narzędzia e-learningowe, aplikacje mobilne i serwisy dzielenia się wiedzą, pozwalają podążać za użytkownikami, szkolić i służyć pomocą w wirtualnej przestrzeni, w której są oni zanurzeni i w której szukają informacji. Projektowanie wzorców wyszukiwawczych, scenariuszy, wychodzenie naprzeciw problemom użytkowników, to klucz do zmierzenia się z nowymi płaszczyznami niewidzialności internetowych zasobów informacyjnych.

## Bibliografia i wykorzystane źródła internetowe

- Adamczyk M., *Człowiek cyfrowy*, Wprost 2002, nr 16.
- Aldrich A., *Universities and Libraries Move to the Mobile Web*, Educase Quarterly, 2/2010 [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.educase.edu/ero/article/universities-and-libraries-move-mobile-web>.
- Anderson Ch., Wolff M., *The Web is dead. Long live the Internet*, Wired Magazine [online] 17.08.2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: [http://www.wired.com/2010/08/ff\\_webrip/](http://www.wired.com/2010/08/ff_webrip/).
- Anonimowa strona głębokiego Internetu – dla ciekawych*, Wyszukiwanie w Głębokim Internecie [szkolenie e-learningowe na platformie EBIB.pl] [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://moodle.ebib.pl/mod/lesson/view.php?id=49>.
- Aplikacje dla systemu Windows Phone – Sklep Microsoft*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.windowsphone.com/store>.

- Battelle J., *What might a Facebook search engine look like?* [online] 16.01.2012 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://battellemedia.com/archives/2012/01/what-might-a-facebook-search-engine-look-like.php>.
- Bergman M.K., *The Deep Web: surfacing hidden value*, The Journal of Electronic Publishing 2001, vol. 1 [online] [dostęp: 15.05.2014] Dostępny w World Wide Web: <http://hdl.handle.net/2027/spo.3336451.0007.104>.
- Cisek S., Sapa R., *Komunikacja naukowa w Internecie – mity i rzeczywistość* [online] E-prints in Library and Information Science [dostęp: 15.11.2014] Dostępny w World Wide Web: [http://eprints.rclis.org/bitstream/10760/9148/1/cisek\\_sapa\\_2006.pdf](http://eprints.rclis.org/bitstream/10760/9148/1/cisek_sapa_2006.pdf).
- Devine J., Egger-Sider F., *Beyond Google: The Invisible Web in the academic library*, The Journal of Academic Librarianship 2004, vol. 4.
- Devine J., Egger-Sider F., *Going Beyond Google again*, Chicago 2014.
- Enge E., *Does Facebook Activity Impact SEO?*, Stone Temple Consulting [online] 12.10.2013 [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.stonetemple.com/does-facebook-activity-impact-seo/>.
- Enge E., *How Does Google Index Tweets?*, Stone Temple Consulting [online] 07.09.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.stonetemple.com/how-does-google-index-tweets/>.
- Enge E., *Matt Cutts Interviewed by Eric Enge*, Stone Temple Consulting [online] 14.03.2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.stonetemple.com/articles/interview-matt-cutts-012510.shtml>.
- Evi, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.evi.com>.
- Ford N., Mansourian Y., *The invisible web: an empirical study of "cognitive invisibility"*, Journal of Documentation 2006, vol. 5.
- Gontar B., Papińska-Kacperek J., *Semantyczne wyszukiwarki internetowe*, Acta Universitatis Lodzensis: Folia Oeconomica 2011, nr 261.
- Google Play, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://play.google.com/store>.
- Grupa Tense, *Tor dla opornych – sekrety ukrytego Internetu*. Gazeta.pl: Next, [online] 14.11.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: [http://technologie.gazeta.pl/internet/1,113840,16969824,Tor\\_dla\\_opornych\\_sekrety\\_ukrytego\\_Internetu.html](http://technologie.gazeta.pl/internet/1,113840,16969824,Tor_dla_opornych_sekrety_ukrytego_Internetu.html).
- Hargittai E. i in., *Trust online: young adults' evaluation of web content*, International Journal of Communication 2010, vol. 4.
- iTunes, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.apple.com/itunes/>.
- Lewandowski D., Mayr P., *Exploring the academic invisible web*, Library Hi Tech 2006, vol. 4.
- Mobile Content Networks, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://mcn-inc.com>.
- Mui Y.Q., Whoriskey P., *Facebook passes Google as most popular site on the Internet, two measures show*, [online]. 31.12.2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.washingtonpost.com/wp-dyn/content/article/2010/12/30/AR2010123004645.html>.
- Nicholas D., Rowlands I., *Social Media use in the research workflow*, Information Services & Use 2011, vol. 1-2.
- Pamuła-Cieślak N., *Analiza czynników wpływających na jakościowe cechy wyników w wyszukiwarce Google*, Toruńskie Studia Bibliologiczne 2013, vol. 1.

- Pamuła-Cieślak N., *Wizualizacja i prezentacja informacji w wyszukiwarkach internetowych*, [w:] *Współczesne oblicza komunikacji i informacji*, red. E. Głowacka, M. Kowalska, P. Krysiński, Toruń 2014.
- Pan B. i in., *In Google we trust: user's decisions on rank, position and relevance*, *Journal of Computer-Mediated Communication* 2007, vol. 12.
- Pariser E., *Beware online "filter bubbles"*, [online] [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.youtube.com/watch?v=B8ofWfx525s>.
- Pariser E., *The Filter Bubble: what the Internet is hiding from You*, London 2012.
- Pulikowski A., *Wyszukiwanie informacji w Internecie*, [w:] *Bibliotekarstwo*, red. A. Tokarska, Warszawa 2013.
- Sherman Ch., Price G., *The Invisible Web: uncovering information sources search engine can't see*, Medford, New Jersey 2003.
- Sherman Ch., *Search for the Invisible Web*, *The Guardian* [online] 06.09.2001 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.theguardian.com/technology/2001/sep/06/internetnews.onlinesupplement>.
- Smith D., Caruso J.B., *The ECAR Study of Undergraduate Students and Information Technology* [online] 2010 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://anitacrawley.net/Resources/Reports/ECAR%20study%20highlights.pdf>.
- State of the media: the social media report 2012*, [online] 12.03.2012 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2012-Reports/The-Social-Media-Report-2012.pdf>.
- Sterling G., *Google Sees Deeper App Content As Key To Mobile Usage* [online] 10.03.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://searchengineland.com/google-seeks-reassert-mobile-deeper-app-indexing-186102>.
- Sterling G., *Quixey Offering Deeper Search Results Inside Mobile Apps* [online] 25.02.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <http://searchengineland.com/quixey-offering-deeper-search-results-inside-mobile-apps-185160>.
- Stock W.G., *Weltregionen des Internet: Digitale Informationen im WWW und via WWW*, Password 2003, vol. 2.
- The Deep Web: What was House of Cards Talking About?* [w:] *BrightPlanet*, [online] 18.02.2014 [dostęp: 15.11.2014] Dostępny w World Wide Web: <https://www.brightplanet.com/2014/02/deep-web-house-of-cards/>.

Natalia Pamuła-Cieślak

## SUMMARY

This paper is an attempt to a new approach to the phenomenon of Deep Web. Despite of permanent development technology of crawling and searching, Invisible Web still exists. As a result of technological changes this phenomenon is also changing. Deep part of the Web described in the textbooks of the early twenty-first century now is a part of Surface Web. But there are a new areas of invisibility, associated with the increase of the amount of information online, semantic Web, new types of meta search engines, the filter bubble in search scores, social media or mobile technologies. Invisible Web should be define again with the new determinants of its transformation.