

Teun Lucassen
Roald Dijkstra
Jan Maarten Schraagen

Czytelność Wikipedii

Abstrakt: *Wikipedia jest powszechnie uznawana za wiarygodne źródło informacji encyklopedycznej. Pewnym problemem jest jednak jej czytelność. Artykuły Wikipedii bywają napisane w sposób niezrozumiały dla przeciętnego użytkownika. Do zbadania tego aspektu w odniesieniu do wszystkich dostępnych artykułów angielskiej wersji Wikipedii wykorzystano indeks czytelności Flescha. W rezultacie badań okazało się, że czytelność Wikipedii jest na niskim poziomie: 75% z wszystkich artykułów plasuje się poniżej spodziewanego poziomu. W teście lepiej wypada Wikipedia w wersji uproszczonej (Simple English Wikipedia), ale i jej czytelność okazuje się nieodpowiednia dla docelowej grupy użytkowników. Wersja demonstracyjna przyjętej metody badań dostępna jest na stronie: www.readabilityofwikipedia.com.*

Słowa kluczowe: czytelność, Wikipedia, Simple English Wikipedia, indeks Flescha

1. Wprowadzenie

Wikipedia, od momentu powstania w roku 2001, stale się rozwija, by stać się obszernym zbiorem informacji encyklopedycznej. W wywiadzie udzielonym dla Slashdot w 2004 r. [1] Jimmy Wales, współzałożyciel Wikipedii, wprost określił cele *dla wolnej encyklopedii, której redaktorem może być każdy: Wyobraźmy sobie świat, w którym wszyscy ludzie na Ziemi mają wolny dostęp do sumy całej ludzkiej wiedzy. Ku temu właśnie zmierzamy*. Wydaje się, że Wales w dużym stopniu zrealizował te zamierzenia. W marcu 2011 r. sama angielskojęzyczna wersja Wikipedii zawierała już 3,5 mln artykułów [2]. Ponadto dostępna była w 279 innych językach.

Przeprowadzono już badania zarówno na temat jakości informacji zawartych w Wikipedii (Giles, 2005; Wilkinson, Huberman, 2007), jak i postrzegania tej jakości przez użytkowników (Kubiszewski i in., 2011; Lucassen, Schraagen, 2011). Obecnie podnoszony jest inny problem, a dotyczy on czytelności Wikipedii.

Badanie ścisłości informacji dostępnych w Wikipedii na temat raka (Rajagopalan et al., 2010) pokazało, że informacje były porównywalne pod względem dokładności z wiadomościami dostępnymi w źródle fachowym, natomiast ich czytelność była dużo gorsza. Sprawdzono to za pomocą indeksu czytelności Flescha-Kincaida (Kincaid et al., 1975), który zastosowano do obu serwisów internetowych, uzyskując wynik według skali indeksu. Dzięki temu możliwe było porównanie ze spodziewanymi poziomami czytelniczymi na różnych etapach amerykańskiego systemu szkolnictwa. Niska czytelność artykułów Wikipedii, którą wykazano na przykładzie tematyki związanej z rakiem, nie daje podstaw do tego, by sądzić, że opracowania na inne tematy będą się plasować wyżej.

Wikipedia nie ma określonego odbiorcy, poza tym wskazanym przez Jimmy'ego Walesa w cytowanej wyżej wypowiedzi. Stwierdził on, że cała ludzka wiedza winna być dostępna dla każdego. Lecz, jak przypuszczamy, informacje umieszczone w Wikipedii są często dostępne nielicznym, gdyż są zbyt trudne w odbiorze. Walraven i inni badacze (2009) stwierdzili częste wykorzystanie Wikipedii przez 14-letnich licealistów. Nie mają oni jeszcze w pełni rozwiniętych umiejętności czytelniczych i z tego względu korzystanie z Wikipedii podlega u nich oczywistym ograniczeniom.

1.1. Indeks czytelności Flescha

Czytelność można zdefiniować jako łatwość rozumienia przekazu skierowanego do czytelnika przez twórcę tekstu. Podjęto wiele prób zautomatyzowanej oceny czytelności tekstów. (DuBay, 2004). Jednym z pierwszych był indeks czytelności Flescha (Flesch, 1948), który połączył w prostym algorytmie dwie, dające się łatwo zmierzyć, cechy, istotne dla czytelności. Po pierwsze przyjęto, że długie zdania są trudniejsze w czytaniu. Mierzy się to jako średnią liczbę słów na zdanie (SL — *sentence length*). Podobnie w przypadku słów — im dłuższe, tym trudniejsze. Przy czym długość słowa (WL — *word length*) określa się średnią liczbą sylab na słowo.

Wzór służący do obliczania czytelności według indeksu Flescha (RE — *reading ease*) jest następujący:

$$RE = 206.835 - (1.015 \times SL) - (84.6 \times WL)$$

Wynik obliczenia przyjmuje najczęściej wartość w przedziale 0–100 (choć wyższe i niższe wyniki też są teoretycznie możliwe) i interpretuje się zgodnie z poniższą tabelą.

Tab. 1. Indeks Flescha — przedziały wyników i ich interpretacja.

Wynik	Interpretacja (poziom czytelności)
90–100	bardzo łatwy
80–90	łatwy
70–80	średnio łatwy
60–70	normalny
50–60	średnio trudny
30–50	trudny
0–30	bardzo trudny

Kierując się zapisami przedstawionymi w tabeli 1, można przyjąć, że teksty kierowane do szerokiego grona odbiorców (jak angielska wersja Wikipedii) powinny celować na poziomie normalnej czytelności, czyli w zakresie między 60 i 70 (normalny). Dla odbiorców nie-

władających biegle językiem angielskim należałoby przyjąć znacznie wyższą czytelność, zapewne na poziomie co najmniej 80 (łatwy).

Indeksy czytelności, takie jak indeks czytelności Flescha lub indeks Flescha-Kincaida (stosującego te same atrybuty tekstu), wykorzystuje się w różnych celach. W amerykańskim Ministerstwie Obrony używa się indeksu Flescha-Kincaida do określania czytelności dokumentów (Kincaid i in., 1975). Co więcej, agencje ubezpieczeniowe na Florydzie są prawnie zobowiązane do przestrzegania czytelności na poziomie przynajmniej 45 [3].

Zasadność stosowania indeksu Flescha do przewidywania rozumienia tekstów zawsze była przedmiotem dyskusji. Z pewnością dwie proste miary, jak długość słowa i długość zdania, nie mogą oddać w pełni pojęcia czytelności. Zaobserwowano jednak silne współzależności występujące pomiędzy indeksem Flescha a sprawdzianami z rozumienia tekstów (DuBay, 2004). Tym niemniej indeksy czytelności (budowane na dowolnych wzorach czytelności) powinno się przyjmować z pewną dozą krytycyzmu.

1.2 Uproszczona angielska wersja Wikipedii — Simple English Wikipedia

Hipotetyczna słaba czytelność Wikipedii została zauważona przez Wikimedia Foundation. U uruchomiono więc uproszczoną wersję Wikipedii (Simple English), która w marcu 2011 r. zawierała 68 000 artykułów [4]. Serwis ten stworzono po to, by [...] *dostarczał informacji encyklopedycznych ludziom o zróżnicowanych potrzebach: studentom, dzieciom, dorosłym mającym trudności w uczeniu się oraz tym, którzy uczą się języka angielskiego* [5]. W tym celu zastosowano prostszą składnię i podstawowe słownictwo. Współtwórcom artykułów zaleca się korzystanie jedynie z poziomu podstawowego „Basic English” (Ogden, 1968), obejmującego 850 słów. W razie konieczności można sięgać do wersji poszerzonej obejmującej 1500 słów.

Przeprowadzono wiele badań sprawdzających efektywność uproszczonej wersji Wikipedii. Pokazano w nich, że przez pierwsze trzy lata po jej uruchomieniu w 2003 r. czytelność, zgodnie z indeksem Flescha, obniżała się wraz z przyrostem artykułów (Besten, Dalle, 2008). O ile w drugiej połowie roku 2003 średnia czytelność była na poziomie 80 (łatwy), o tyle w 2006 spadła do nieco ponad 70 (średnio łatwy). Interesujące byłoby ponowne przeanalizowanie aktualnego poziomu czytelności uproszczonej wersji Wikipedii i stwierdzenie, czy tendencja ta się utrzymała. Jeśliby się to potwierdziło, to uproszczona wersja Wikipedii stawałaby się coraz mniej czytelna dla użytkowników o specjalnych potrzebach, a wartość dodana w stosunku do zwykłej Wikipedii — zanikłaby.

W niniejszym opracowaniu zbadano czytelność obu angielskojęzycznych wersji Wikipedii: zwykłej i uproszczonej, za pomocą indeksu Flescha (Flesch, 1948). Postawiono hipotezę, że dla wielu osób zwykła wersja Wikipedii jest zbyt trudna w odbiorze. Ponadto przyjęto, że wersja uproszczona powinna rozwiązać ten problem dzięki lepszej czytelności. Jeśli jednak tendencja spadkowa, którą zauważono we wcześniejszych badaniach (Besten, Dalle, 2008), potwierdzi się, to wartość dodana uproszczonej Wikipedii spadnie, gdyż osoby niewładające biegle językiem angielskim nie będą w stanie zrozumieć jej treści.

2. Metodologia

2.1 Zestaw danych

Wybór grupy reprezentatywnej artykułów pochodzących z obu wersji Wikipedii zastąpiono próbą zindeksowania wszystkich artykułów. Dzięki temu uzyskano pewność, że wyniki rzeczywiście odzwierciedlają poziom czytelności Wikipedii na dany czas. Artykuły ze zwykłej wersji Wikipedii uzyskano z bazy danych pobranej w formacie XML w dniu 17 czerwca 2010 r. [6] Teksty z wersji uproszczonej ściągnięto z bazy danych 2 września 2010 r.

2.2. Przetwarzanie

Pliki w formacie XML zaimportowano do lokalnej MediaWiki za pomocą MWDumper — narzędzia do importu danych MediaWiki [7]. Następnie zawartość każdego artykułu była wyszukana z bazy danych MediaWiki i przefiltrowana do postaci wyłącznie pełnych zdań (patrz następna sekcja). Indeks Flescha policzono za pomocą *php-text statistics*, gdzie zastosowano wzór, o którym była mowa w sekcji 1.1. Uzyskany rezultat był przenoszony do lokalnej bazy MySQL. Po przetworzeniu wszystkich artykułów wyniki eksportowano do plików w formacie CSV (*comma-separated values*) a następnie importowano do SPSS do dalszej analizy.

2.3. Filtrowanie

Zgodnie z metryką stosowaną w indeksie czytelności Flescha (długość zdania i długość słowa) powinno się poddawać analizie jedynie pełne zdania. To oznacza, że każda inna zawartość powinna być usunięta przed testowaniem. Najpierw usunięto wszystkie artykuły oznaczone jako strony ujednoznaczniające, przekierowania i wykazy oraz strony zawierające wyłącznie inne nośniki treści, jak grafiki, dokumenty dźwiękowe czy pliki PDF. Należało je pominąć, gdyż nie stanowią one zawartości w postaci pełnych zdań.

Następnie odfiltrowano wszystkie pozostałe elementy artykułów, które nie tworzyły (lub nie stanowiły części) pełnych zdań. Przykładem są nagłówki, tablice, linki URL. Włącza się je do treści artykułów Wikipedii przy użyciu języka znaczników MediaWiki. Język ten umożliwia proste znakowanie układu tekstu. Na przykład nagłówek tworzy się, obejmując tytuł znakami równości (np. „===Wprowadzenie===”).

Aby usunąć z tekstu wszystkie nieodpowiednie elementy, wprowadzono do PHP standardowe wyrażenia, dzięki którym namierzano określające je znaczniki treści. Niektóre elementy usuwano z tekstu w całości, jak nagłówki, tablice i obrazki. Inne przekształcano, usuwając wszystkie informacje kontekstowe, które nie stanowiły części samego tekstu. Dotyczyło to linków, powołań i cytowań. Nie usuwano pozostałości tych elementów, stanowiących części pełnych zdań.

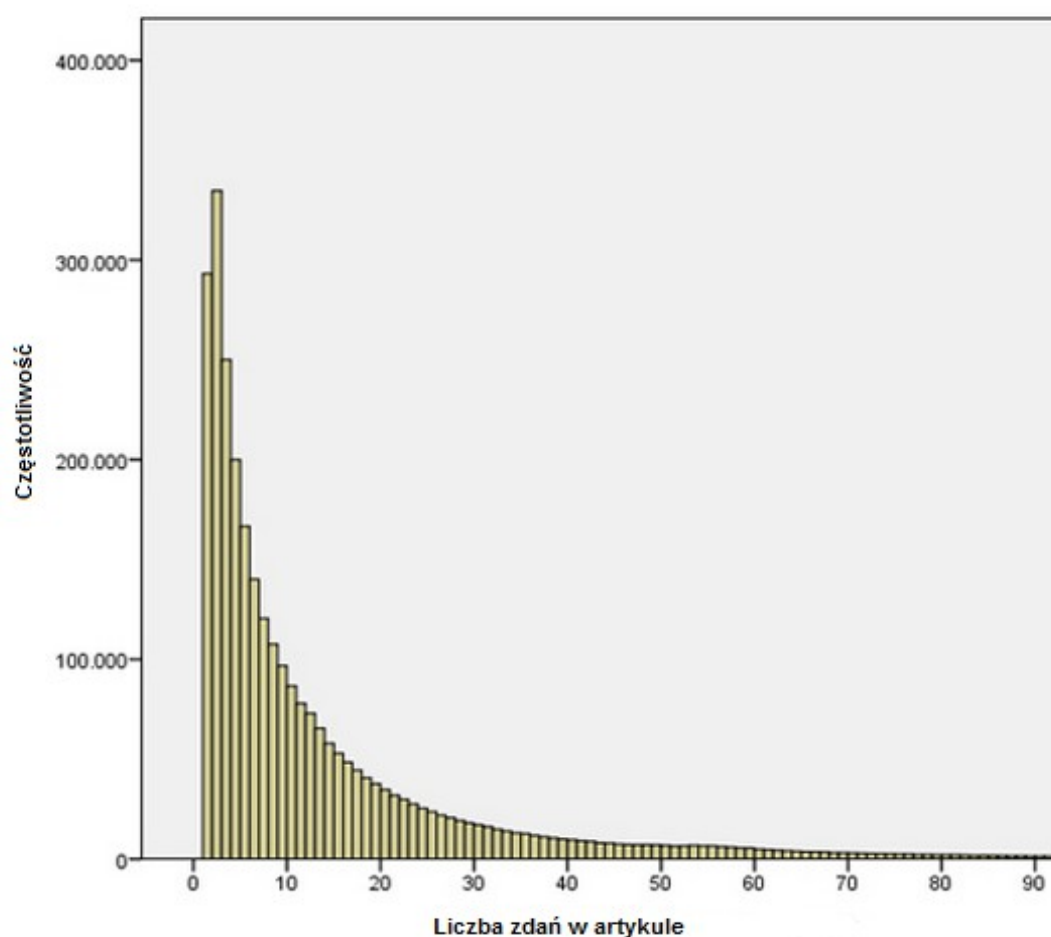
Nie wszystkie artykuły Wikipedii nadawały się do przetestowania. Niektóre na przykład po przefiltrowaniu nie zawierały ani jednego pełnego zdania. Dotyczyło to w dużej mierze wykazów zawierających przykładowo linki do pokrewnych tematów. Ponadto w niektórych przypadkach autorzy tekstów zastosowali inny język znaczników, np. HTML. Powodowało to problemy w mechanizmie filtrowania, który zaprogramowano dla języka MediaWiki.

W takich sytuacjach artykuły generujące błędy były automatycznie wyłączane z dalszej analizy. Z obu serwisów: zwykłej Wikipedii i jej wersji uproszczonej, udało się wygenerować i poddać analizie odpowiednio 88 i 85% artykułów zawierających pełne zdania.

3. Wyniki

3.1. Angielskojęzyczna wersja Wikipedii (English Wikipedia)

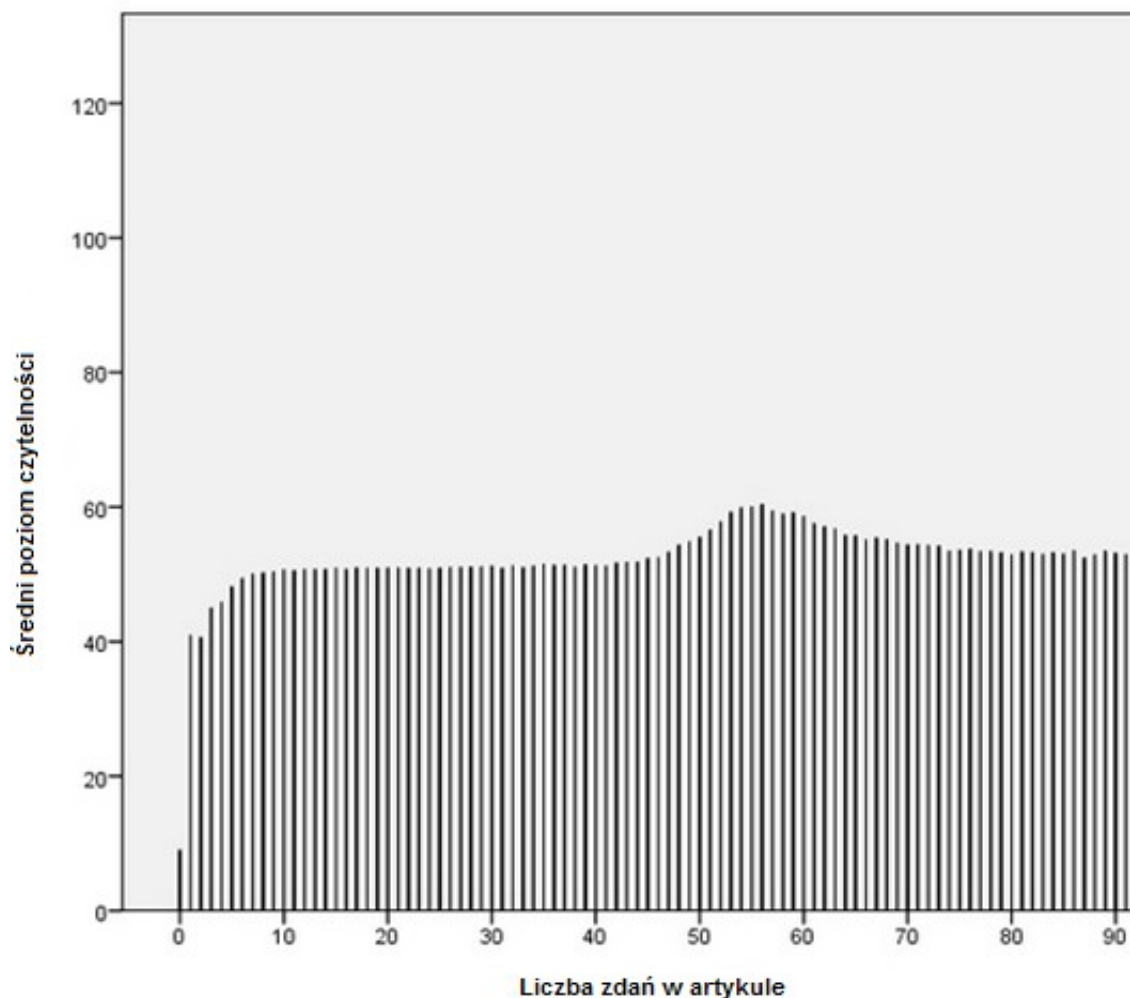
Analizie poddano w sumie 2 955 210 artykułów, po uprzednim ich przefiltrowaniu do pełnych zdań. Szybko jednak okazało się, że wiele z nich zawiera zaledwie po kilka zdań. Jak uwidoczniło na wykresie 1 około 40% artykułów składa się z co najwyżej pięciu zdań.



Wykr. 1. Rozdział liczby zdań w artykułach angielskojęzycznej Wikipedii

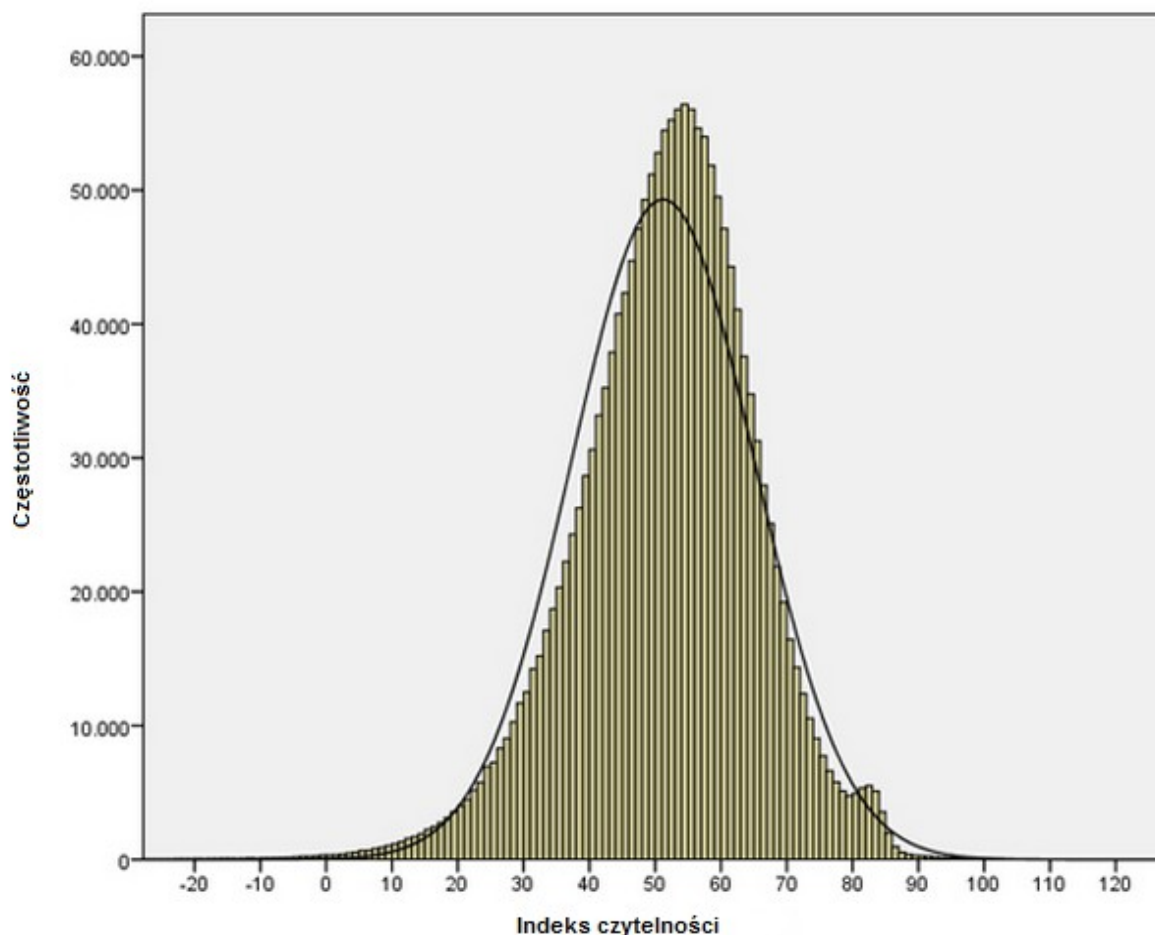
Wykres 2 pokazuje, że poziom czytelności wyrównuje się przy artykułach zawierających co najmniej pięć zdań. Ponadto większa zmienność w czytelności występuje w przypadku krótszych artykułów (SD = 24.32 dla artykułów zawierających pięć lub mniej zdań, SD = 13.84 dla artykułów powyżej pięciu zdań). Niewielki skok w średniej czytelności zaobser-

wowano na poziomie 55 zdań, co wynikało z dodania wielu haseł na temat amerykańskich miast (Ortega et al., 2009). Artykuły te cechuje podobna czytelność i niemal ta sama liczba zdań.



Wykr. 2. Poziomy czytelności artykułów Wikipedii z uwzględnieniem liczby zdań

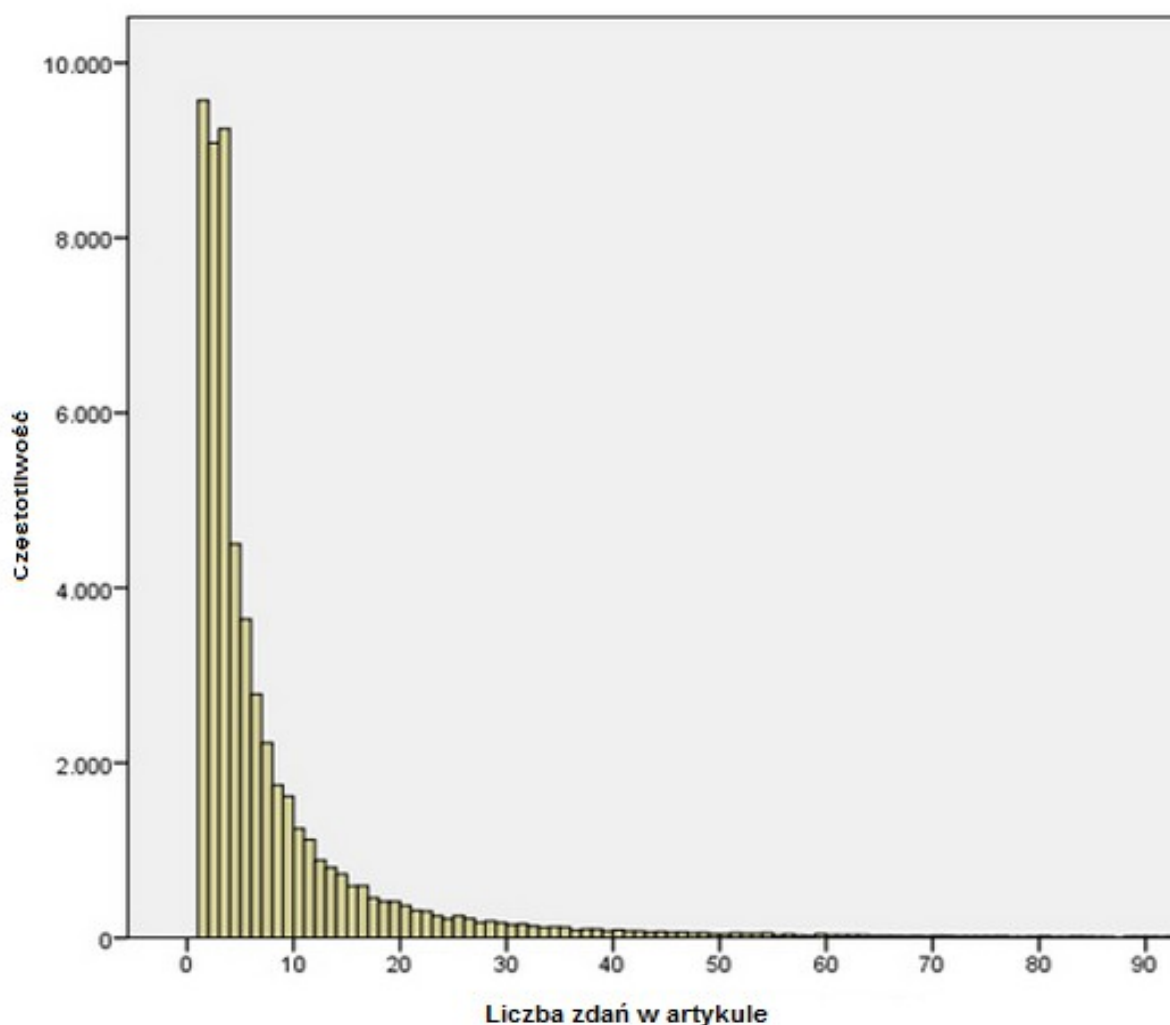
Uważamy, że tak krótkie artykuły nie stanowią grupy reprezentatywnej dla artykułów Wikipedii. Dlatego z dalszych badań wykluczaliśmy artykuły składające się z pięciu lub mniejszej liczby zdań. Wykres 3 przedstawia rozdział poziomów czytelności dla pozostałych 1 710 752 artykułów anglojęzycznej Wikipedii, zawierających więcej niż pięć zdań. Średni poziom czytelności wynosił 51,18 (SD = 13.84). Okazało się, że 73,5% wszystkich artykułów wypadło poniżej zakładanego poziomu 60 (zwykły), przy czym 45% można było zakwalifikować jako trudne, a nawet jeszcze gorzej (< 50).



Wykr. 3. Histogram poziomów czytelności angielskojęzycznej Wikipedii

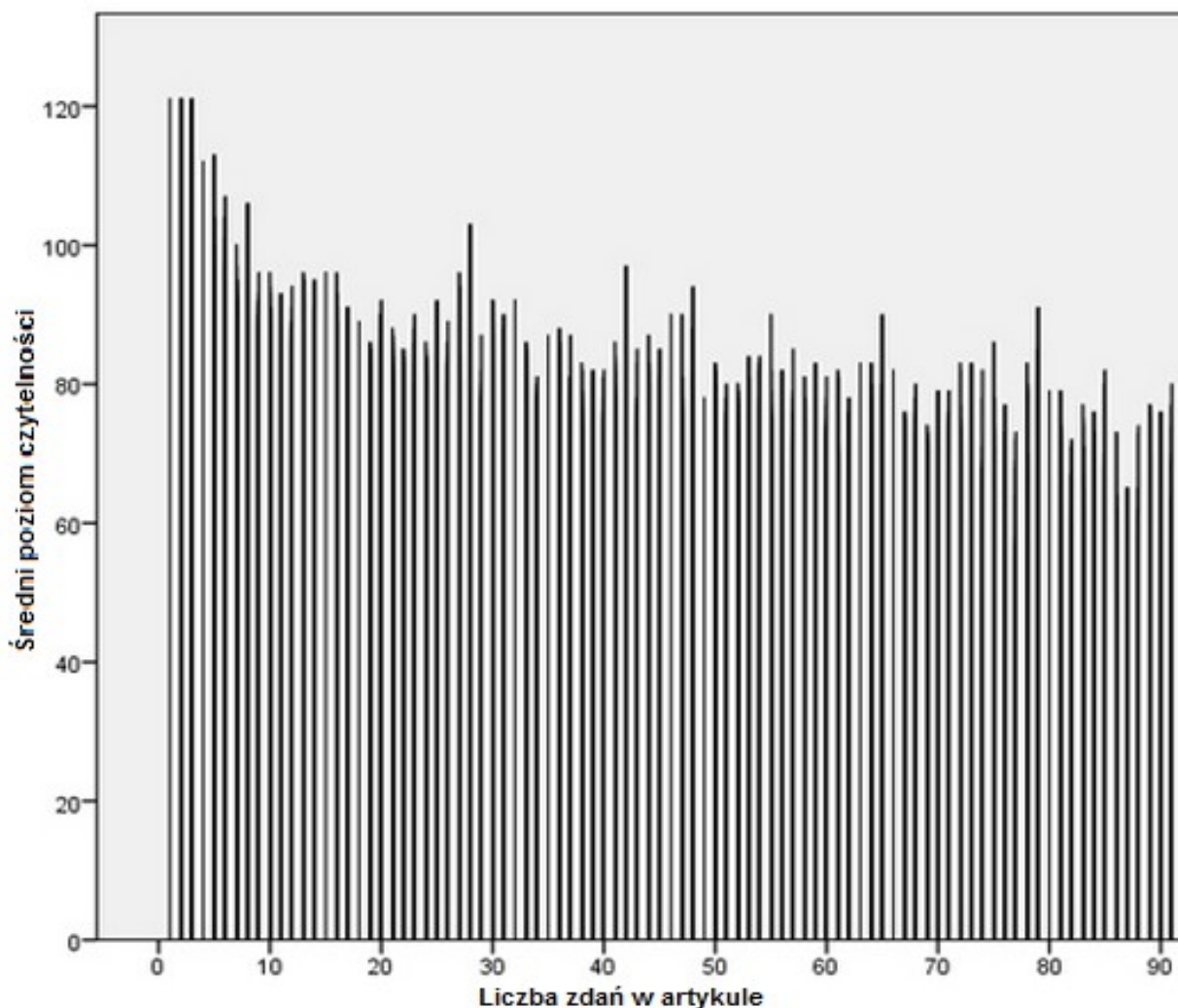
3.2. Uproszczona wersja angielskojęzycznej Wikipedii (Simple English Wikipedia)

W porównaniu ze zwykłą Wikipedią z 57 422 artykułów pochodzących z uproszczonej angielskojęzycznej wersji Wikipedii (Simple English Wikipedia) jeszcze większa ich część składała się z pięciu lub mniejszej liczby zdań (patrz wykres 4). Udział takich artykułów w uproszczonej wersji Wikipedii wyniósł ponad 60%.



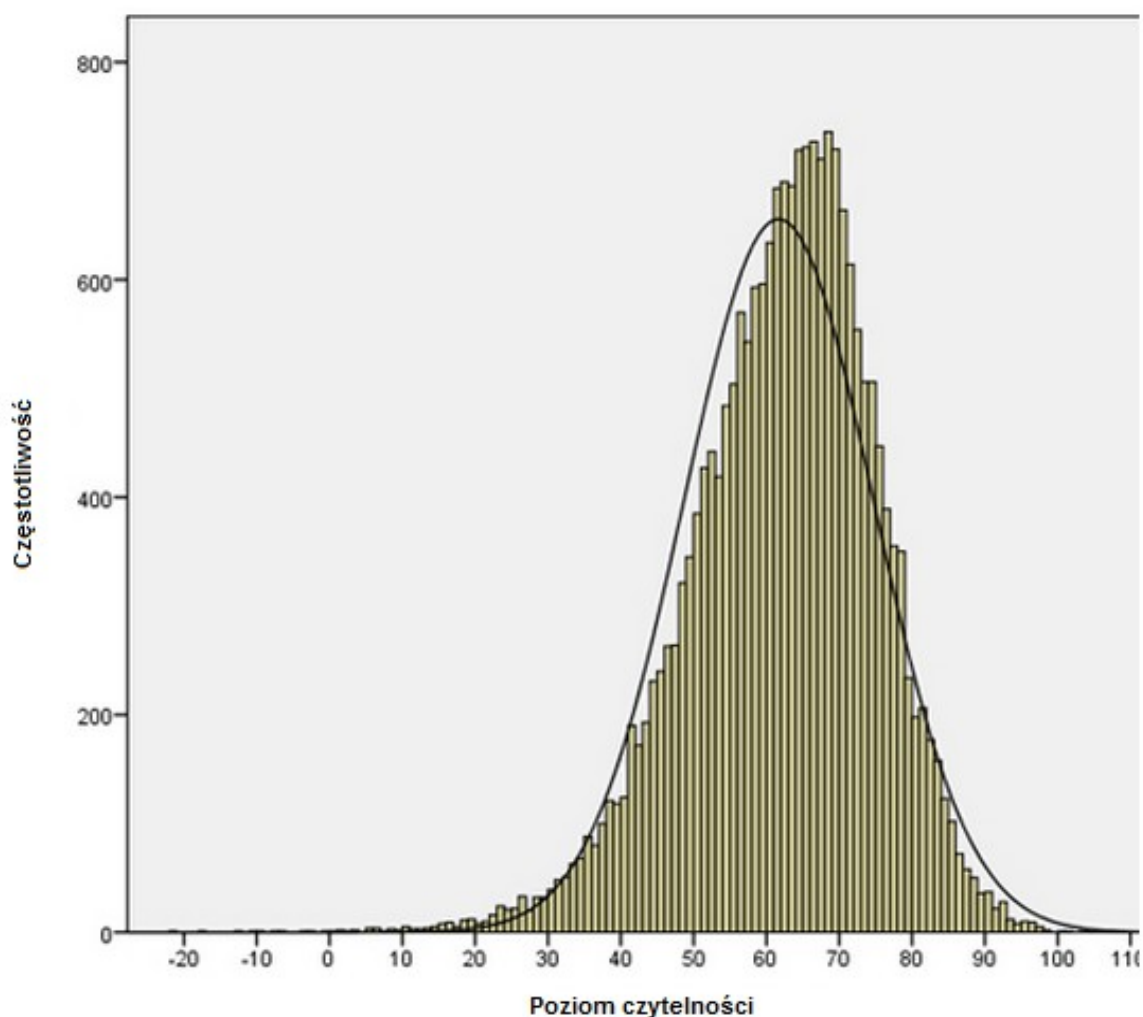
Wykr. 4. Rozdział liczby zdań w artykułach z uproszczonej wersji Wikipedii

Podobnie jak w przypadku zwykłej Wikipedii poziom czytelności artykułów z wersji uproszczonej stabilizował się przy artykułach składających się z co najmniej pięciu zdań (patrz wykres 5). Krótsze artykuły cechowała większa zmienność w poziomie czytelności ($SD = 19,52$) niż w przypadku artykułów składających się z więcej niż pięciu zdań ($SD = 13,00$).



Wykr. 5. Poziomy czytelności artykułów z uproszczonej angielskojęzycznej wersji Wikipedii z uwzględnieniem liczby zdań

Na wykresie 6 przedstawiono rozdział poziomów czytelności artykułów uproszczonej wersji Wikipedii, zawierających więcej niż pięć zdań. Zbadano wszystkie z 21 366 artykułów, z rezultatem średniego poziomu czytelności wynoszącego 61,69 (SD = 13.00). Okazało się, że 94,7% artykułów plasowało się poniżej oczekiwanego poziomu 80 (łatwy), podczas gdy 42,3% nie osiągnęło nawet poziomu 60 (zwykły), co odpowiadałoby bardziej odbiorcy ogólnemu, ale nie temu o szczególnych wymaganiach językowych.



Wykr. 6. Histogram poziomów czytelności uproszczonej wersji Wikipedii

3.3. Porównanie

Bezpośrednie porównywanie poziomów czytelności obu wersji Wikipedii nie byłoby rzetelne. Różni je przede wszystkim liczba dostępnych artykułów. Dlatego do porównania wybraliśmy wyłącznie artykuły o takiej samej stronie tytułowej występującej w obu wersjach Wikipedii ($N = 9603$). Uwzględniono również, o czym już wspomniano, wyłącznie artykuły składające się z więcej niż pięciu zdań. Biorąc pod uwagę zredukowaną liczbę artykułów, średnia czytelność zwykłej angielskojęzycznej wersji Wikipedii plasowała się na poziomie 49,27 ($SD = 11.33$), przy średniej czytelności uproszczonej wersji Wikipedii będącej na poziomie 61,46 ($SD = 12.95$). W przypadku takiego zestawu artykułów wersja uproszczona charakteryzuje się znacząco wyższymi poziomami czytelności niż zwykła Wikipedia ($t(9602) = 99.63$, $p < .001$).

4. Dyskusja i wnioski

W niniejszym opracowaniu zbadano obecny stan czytelności angielskojęzycznej wersji Wikipedii. Wykazano, że znaczna liczba artykułów plasuje się poniżej oczekiwanego poziomu. Rozwiązanie w postaci uproszczonej Wikipedii cechuje lepsza czytelność, jednakże powinno się w niej dążyć do uzyskania jeszcze wyższego poziomu czytelności, gdyż kierowana jest do osób o ograniczonych umiejętnościach czytelniczych. Ponadto wskazano, że spadek czytelności, który obserwowano od czasu uruchomienia uproszczonej wersji Wikipedii w roku 2003, nadal się utrzymuje.

Wyniki tego badania pokazują, że czytelność angielskojęzycznej Wikipedii jest zdecydowanie poniżej oczekiwanego poziomu. Mimo że średni wynik 51,18 nie wydaje się daleki od oczekiwanego, niemal 75% artykułów plasowało się poniżej poziomu 60 pkt według indeksu Flescha. Co więcej, połowę z nich można zaklasyfikować do trudnych lub jeszcze niżej. To wszystko potwierdza przypuszczenie, że znaczna liczba artykułów Wikipedii jest zbyt trudna dla wielu czytelników.

Problem czytelności Wikipedii był już podnoszony w odniesieniu do artykułów objaśniających tematykę raka (Rajagopalan i in., 2010). Artykuły te plasowały się na poziomie 14 według indeksu Flescha-Kincaida, co odpowiada czytelności na poziomie 30. Na szczęście średnia czytelność angielskojęzycznej Wikipedii jest stosunkowo wyższa niż ten wynik, aczkolwiek wiele artykułów (ok. 7%, tj. ok. 12 500 artykułów) plasuje się na poziomie 30 lub niżej. To oznacza, że czytelność może znacząco się wahać, w zależności od tematyki. Analiza tematyczna (podobna do tej przeprowadzonej przez Halavaisa i Lackaffa, 2008), mogłaby ujawnić, które z artykułów pilnie wymagają poprawienia ich czytelności. Hasła o tematyce technicznej mogą być mniej czytelne niż artykuły z innego zakresu, np. kultury popularnej.

Konieczność modyfikacji artykułów zależy także od statystyki czytelniczej. Artykuły z wysoką liczbą odwiedzin powinno się poprawiać w pierwszej kolejności, przed tekstami o niszowej tematyce, którą czyta niewielu. Można to połączyć z potrzebami czytelników o ograniczonej znajomości języka angielskiego, którzy mogą interesować się określoną tematyką. Wprowadzenie otwartych danych dotyczących odwiedzin na stronach każdego artykułu Wikipedii pomogłoby w ustaleniu hierarchii ważności tematów wymagających poprawy.

Fundacja Wikimedia zaproponowała rozwiązanie problemu czytelności w postaci uproszczonej wersji angielskojęzycznej Wikipedii (Simple English Wikipedia). Wypada ona lepiej w indeksie czytelności niż zwykła Wikipedia. Jednakże wyniki pokazują, że jej czytelność spadła o 10 punktów w ciągu ostatnich czterech lat (Besten, Dalle, 2008), i obecnie utrzymuje się na poziomie standardowym. Wynik ten jest niepokojący, bo oznacza, że uproszczona wersja Wikipedii zatraciła swój sens. Stała się odpowiednia dla przeciętnego użytkownika zamiast dla osób o ograniczonych umiejętnościach językowych. Jednym ze sposobów sprawdzania uproszczonej Wikipedii mogłoby być testowanie liczby użytych słów, które faktycznie występują na liście słownictwa podstawowego (*Basic English lists*) (Ogden, 1968).

Źródeł niskiej czytelności obu wersji Wikipedii można doszukiwać się w jej współtwórcach. Są to przede wszystkim ludzie dobrze wykształceni, gdyż tworzenie zawartości Wikipedii wymaga gruntownej wiedzy dziedzinowej, której jeszcze w Wikipedii nie opisano. Możliwe, że piszą dla swoich rówieśników, nie biorąc pod uwagę ogółu odbiorców Wikipedii. A nawet jeśli zwracają szczególną uwagę na odbiorców, jak w przypadku uproszczonej Wikipedii, to i tak przeceniają ich umiejętności językowe.

Interesujące byłoby znalezienie sposobu na zachęcenie i wsparcie współtwórców Wikipedii w zakresie poprawy czytelności ich artykułów. Niewykluczone, że można by to osiągnąć za pomocą dodatkowego narzędzia w opcjach edytorskich Wikipedii, które by sygnalizowało autorowi zbyt długie zdania lub trudne słowa. W trakcie edycji pojawiałyby się podpowiedzi z podziałem zdania lub użyciem prostszego słownictwa. Umożliwia to indeks czytelności Flescha, który zastosowany w środowisku on-line jako miernik czytelności byłby na podorzędziu twórców artykułów Wikipedii.

Należy podkreślić, że uproszczona Wikipedia jest wciąż na etapie rozwoju. Świadczy o tym nie tylko liczba artykułów, lecz także ich średnia objętość. W trakcie badań zauważono, że mniej więcej połowa artykułów składa się z co najwyżej trzech zdań. To dodatkowo motywuje do zwrócenia uwagi współtwórcom Wikipedii na polepszenie czytelności już na tym etapie. Wydaje się, że pisanie tekstów od razu prostszym językiem będzie efektywniejsze niż późniejsze poprawianie ich czytelności.

W badaniu zastosowano tylko jeden (zautomatyzowany) miernik czytelności. Mimo że użyteczność indeksu Flescha jest powszechnie uznana, dobrze byłoby poddać go — jak i inne mierniki czytelności — weryfikacji, aby uzyskać pewność, że wyniki odpowiadają rozumieniu i postrzeganiu czytelności przez użytkownika.

5. Wersja demonstracyjna

Wykorzystanie indeksu czytelności Flescha do testowania artykułów Wikipedii wymagało opracowania narzędzia dostępnego on-line. Dzięki temu czytelność artykułów Wikipedii można sprawdzić „na żywo”, w momencie klikania na hasło. Testować można obie angielskojęzyczne wersje Wikipedii: zwykłą i uproszczoną. Oprócz wartości bezwzględnej, określającej poziom czytelności konkretnego artykułu, podano również czytelność testowanych w tym badaniu artykułów w postaci procentowej. Dodatkowo można wprowadzić własny tekst i sprawdzić jego czytelność. Wersja demonstracyjna dostępna jest na stronie: <http://www.readabilityofwikipedia.com>.

Przekład Agnieszka Brachfogel

Noty o autorach:

Teun Lucassen jest doktorantem na Wydziale Psychologii Poznawczej i Ergonomii na Uniwersytecie Twente w Enschede w Holandii. Tytuł magistra uzyskał na Wydziale Mediów Społecznościowych na tymże uniwersytecie. Obecnie jako doktorant zajmuje się badaniami poziomu wiarygodności repozytoriów społecznych, ze szczególnym uwzględnieniem Wikipedii.

Roald Dijkstra jest właścicielem Babbletics, firmy zajmującej się rozwojem intuicyjnych aplikacji internetowych. Pracuje jako wolny strzelec przy projektach interakcyjnych, równolegle studiuje na Wydziale Mediów Społecznościowych na Uniwersytecie Twente.

Jan Maarten Schraagen jest naukowcem w holenderskiej Organizacji Stosowanych Badań Naukowych (Netherlands Organization for Applied Scientific Research TNO) i profesorem poznawczej psychologii stosowanej na Uniwersytecie Twente. Jego naukowe zainteresowania obejmują analizę zadań, zespołowego podejmowania decyzji, wiarygodności repozytoriów społecznych oraz adaptacyjnych interakcji między człowiekiem a komputerem. Jest współredaktorem książek *Cognitive task analysis* (Mahwah, N.J.: L. Erlbaum Associates, 2000) oraz *Naturalistic decision making and macrocognition* (Aldershot, England: Ashgate, 2008). Dr Schraagen jest doktorem nauk w dziedzinie psychologii poznawczej na Uniwersytecie Amsterdamskim w Holandii.

Adresy stron WWW:

1. <http://interviews.slashdot.org/article.pl?sid=04/07/28/1351230>.
2. http://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia.
3. <http://law.onecle.com/florida/insurance/627.4145.html>.
4. <http://simple.wikipedia.org/wiki/Special:Statistics>.
5. http://en.wikipedia.org/wiki/Simple_English_Wikipedia.
6. <http://download.wikimedia.org>.
7. <http://www.mediawiki.org/wiki/Mwdumper>.
8. <http://code.google.com/p/php-text-statistics/>.

Bibliografia:

1. Matthijs L. den Besten and Jean-Michel Dalle, 2008. Keep it simple: a companion for Simple Wikipedia? *Industry and Innovation*, volume 15, number 2, pp. 169–178
<http://dx.doi.org/10.1080/13662710801970126>.
2. William H. DuBay, 2004. The principles of readability at <http://www.nald.ca/library/research/readab/readab.pdf>, accessed 20 August 2012.
3. Rudolph Flesch, 1948. "A new readability yardstick, *Journal of Applied Psychology*, volume 32, number 3, pp. 221–233 <http://dx.doi.org/10.1037/h0057532>.
4. Jim Giles, 2005. Internet encyclopaedias go head to head, *Nature*, volume 438, number 7070 (15 December), pp. 900–901, and at <http://www.nature.com/nature/journal/v438/n7070/full/438900a.html>, accessed 20 August 2012.
5. Alexander Halavais, Derek Lackaff, 2008. An analysis of topical coverage of Wikipedia, *Journal of Computer-Mediated Communication*, volume 13, number 2, pp. 429–440
<http://dx.doi.org/10.1111/j.1083-6101.2008.00403.x>.
6. J. Peter Kincaid, Robert P. Fishburne, Richard L. Rogers, Brad S. Chissom, 1975. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel, Millington, Tenn.: Chief of Naval Technical Training, Naval Air Station Memphis; Springfield, Va.: distributed by National Technical Information Service (NTIS).
7. Ida Kubiszewski, Thomas Noordewier, Robert Costanza, 2011. Perceived credibility of Internet encyclopedias, *Computers & Education*, volume 56, number 3, pp. 659–667
<http://dx.doi.org/10.1016/j.compedu.2010.10.008>.
8. Teun Lucassen, Jan Maarten Schraagen, 2011. Factual accuracy and trust in information: the role of expertise, *Journal of the American Society for Information Science and Technology*, volume 62, number 7, pp. 1,232–1,242.
9. Charles K. Ogden, 1968. *Basic English: international second language*. New York: Harcourt, Brace & World.

10. Felipe Ortega, Jesus M. Gonzalez–Barahona, Gregorio Robles, 2009. Quantitative analysis of the top ten Wikipedias, In: Joaquim Filipe, Boris Shishkov, Markus Helfert, Leszek Maciaszek (editors). *Software and data technologies: Second International Conference, ICSoft/ENASE 2007, Barcelona, Spain, July 22–25, 2007, revised selected papers. Communications in Computer and Information Science*, volume 22. Berlin: Springer–Verlag, pp. 257–268.
11. M.S. Rajagopalan, V. Khanna, M. Stott, Y. Leiter, T.N. Showalter, A. Dicker, Y.R. Lawrence, 2010. Accuracy of cancer information on the Internet: a comparison of a Wiki with a professionally maintained database, *Bodine Journal*, volume 3, number 1, article 8, at <http://jdc.jefferson.edu/bodinejournal/>, accessed 20 August 2012.
12. Amber Walraven, Saskia Brand–Gruwel, Henny P.A. Boshuizen, 2009. How students evaluate information and sources when searching the World Wide Web for information, *Computers & Education*, volume 52, number 1, pp. 234–246 <http://dx.doi.org/10.1016/j.compedu.2008.08.003>.
13. Dennis M. Wilkinson, Bernardo A. Huberman, 2007. Cooperation and quality in Wikipedia, *WikiSym '07: Proceedings of the 2007 International Symposium on Wikis*, pp. 157–164, and at <http://www.hpl.hp.com/research/scl/papers/wikipedia/wikipedia07.pdf>, accessed 20 August 2012.

Nota redakcyjna: Oryginalny tekst: LUCASSEN T., DIJKSTRA R., SCHRAAGEN J.M. Readability of Wikipedia. *First Monday* [on-line] 2012, vol. 17, nr 9 [dostęp 9.09.2015]. Dostępny w: <http://firstmonday.org/ojs/index.php/fm/article/view/3916>. DOI: <http://dx.doi.org/10.5210/fm.v0i0.3916>. Tekst został opublikowany w Domenie Publicznej. Autorzy zostali poinformowani o przekładzie i publikacji.