



- narzędzie do mierzenia zrozumiałości polskich tekstów

Włodzimierz Gruszczyński
SWPS Uniwersytet Humanistycznospołeczny

Czym jest?

- **Jasnopis** jest programem komputerowym, z którego można bezpłatnie korzystać *online* pod adresem: www.jasnopis.pl/aplikacja.
- Za pomocą **Jasnopisu** można mierzyć trudność (zrozumiałość) tekstu polskiego.
- Trudność jest wyrażona liczbowo w skali 1-7. Sprawdziliśmy ją za pomocą badań ankietowych.
- **Jasnopis** nie wymaga dodatkowego oprogramowania ani wiedzy informatycznej lub lingwistycznej.
- **Jasnopis** podaje oceny wyliczane innymi, wcześniej znanymi metodami.
- **Jasnopis** podaje także informacje statystyczne, które mogą się przydać bardziej zaawansowanym użytkownikom.
- **Jasnopis** zaznacza niektóre trudniejsze fragmenty tekstu – akapity, zdania i wyrazy.
- Planujemy udostępnienie tzw. wtyczek (plug-in), dzięki którym będzie można korzystać z Jasnopisu w czasie pracy za pomocą programów MS Word i LibreOffice.

Czym nie jest?

- **Jasnopis** nie poprawia tekstu ani nie zmienia go pod jakimkolwiek względem; w szczególności go **nie upraszcza!**
- **Jasnopis** nie jest przeznaczony do oceny:
 - tekstów artystycznych, zwłaszcza poetyckich;
 - tekstów bardzo krótkich, kilkuwyrazowych (sprawdza się w odniesieniu do tekstów minimum 100-wyrazowych, a najlepiej 300-wyrazowych),
 - tekstów zawierających bardzo wiele wzorów matematycznych, chemicznych, wyrażeń symbolicznych itp.
 - tekstów zawierających wiele cytatów obcojęzycznych.
- **Jasnopis** nie jest programem, który „rozumie” tekst, więc nie odróżnia ciągu słów mniej lub bardziej przypadkowych od prawdziwego tekstu polskiego – zakładamy, że użytkownik nie będzie wobec Jasnopisu „złośliwy”.
- **Jasnopis** nie ocenia jakości tekstu, nie odróżnia tekstów „dobrych” od „złych”, cokolwiek by to znaczyło.

Kto za to zapłacił?

Jasnopis powstał w ramach projektu badawczego pod tytułem
***Mierzenie stopnia zrozumiałości polskich tekstów
użytkowych (pozaliterackich).***

- Projekt sfinansowało **Narodowe Centrum Nauki** na podstawie DEC-2011/03/B/HS2/05799 z dn. 2012-06-20.
- Planowany okres realizacji projektu: 24 mies. (08.2012 – 08.2014).
- Przedłużony okres realizacji: do 15.12.2015.
- Środki finansowe: **390215,00 PLN.**
- Jednostka koordynująca: Wydział Nauk Humanistycznych i Społecznych Uniwersytetu SWPS.
- Jednostka współpracująca: Instytut Podstaw Informatyki PAN.

Kto tworzył?



Interdyscyplinarny zespół:

- Włodzimierz Gruszczyński (kierownik, USWPS) – lingwistyka,
- Bartosz Broda (Samsung) – informatyka, inżynieria lingwistyczna,
- Edyta Charzyńska (UŚ) – psychologia, filozofia,
- Łukasz Dębowski (IPI PAN) – statystyka, informatyka, lingwistyka,
- Milena Hadryan (UAM) – legilingwistka, skandynawistyka,
- Bartłomiej Nitoń (IPI PAN) – informatyka (programowanie),
- Maciej Ogrodniczuk (IPI PAN) – informatyka, lingwistyka,
- Grażyna Oblas (UW) – ekonomia, socjologia,

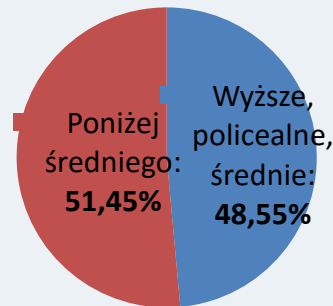
Okresowo: Laura Polkowska (UKSW) – lingwistyka.

Wsparcie:

- organizacyjne: Renata Łągowska (SWPS),
 - graficzne: Bartosz Szymański (SWPS),
 - marketingowe: Dział Marketingu SWPS,
 - kontakty z mediami, PR: Marta Komor (SWPS).
- Studentki, studenci, koleżanki, koledzy, którzy przeprowadzali ankiety.

Dlaczego chcieliśmy to zrobić?

- W Polsce istnieje zbyt dużo zbyt trudnych tekstów w przestrzeni publicznej.
- Wielu ludzi sobie z nimi nie radzi, bo są za mało wykształceni.



- Nie radzą sobie z nimi osoby z różnymi rodzajami dysfunkcji utrudniającymi rozumienie tekstu, tzn. ok. 10 proc. dorosłej populacji! Są to m.in.:
 - głusi (tzw. „milczący cudzoziemcy”),
 - niewidomi,
 - osoby z zaburzeniami poznawczymi i intelektualnymi (niski iloraz inteligencji, ADHD, autyzm, dysleksja i in.),
- Kłopoty z trudnymi tekstami mają też imigranci, cudzoziemcy z kartami stałego pobytu.

Przykład tekstu dla prostego człowieka

- Działający z upoważnienia wnioskodawcy – Dyrektor Zarządzający – Prokurent T. P. zaproponował aby ustalenie przebiegu granicy nastąpiło po rozstrzygnięciu sprawy prowadzonej w Sądzie Okręgowym do sygn. [...] z wniosku skarżącego o wydanie nieruchomości.
- Podniesiono, że w odpowiedzi na zapytania skarżącego o zaawansowanie prac, informowano skarżącego pismami, które są w aktach, że zakończenie postępowania rozgraniczeniowego warunkowane jest dostarczeniem przez Geodetę dokumentów, które w pierwszej kolejności posłużą do wydania przez Wójta Gminy opinii w trybie art. 33 ust. 2 pkt 1 ustawy z dnia 17 maja 1989 r. Prawo Geodezyjne i Kartograficzne. Po uzyskaniu tej opinii i przyjęciu dokumentów do powiatowego zasobu dokumentacji geodezyjnej i kartograficznej prowadzonego przez Starostę Powiatowego, dokumentacja będzie zdatna do wydania decyzji w przedmiocie rozgraniczenia.

Dlaczego chcieliśmy to zrobić?

- Widzieliśmy potrzebę narzędzia umożliwiającego sprawdzenie zrozumiałości tekstu metodą, która:
 - jest obiektywna,
 - jest (względnie) szybka,
 - **została zweryfikowana za pomocą badań empirycznych,**
 - jest w pełni dostosowana do cech języka polskiego,
 - daje dla porównania również wyniki obliczane według innych znanych sposobów,
 - daje możliwości doskonalenia i dostosowania do zastosowań specjalistycznych.

Jakie metody stosuje się za granicą?

- Na świecie powstało i powstaje wiele metod sprawdzających zrozumiałość tekstów, istnieją też narzędzia informatyczne służące do tego celu (głównie dla języka angielskiego).
- W niektórych krajach sprawdza się stopień zrozumiałości tekstów, które mają funkcjonować w przestrzeni publicznej, np.:
 - w USA w wielu urzędach istnieje wymóg sprawdzania tekstów pod względem czytelności przed ich przekazaniem odbiorcom;
 - w Szwecji od kilkunastu lat wdrażany jest program *klarspråk*, czyli „jasny język”, dzięki czemu udało się wyprzeć z urzędów typowy ciężki styl kancelaryjny.
- Trwają badania nad metodami upraszczania tekstów, w tym także upraszczania automatycznego.

Co było wcześniej w Polsce?

- W latach 60. XX wieku Walery Pisarek stworzył i opublikował wzór, pozwalający obliczyć współczynnik trudności tekstu polskiego; wzór znany był przede wszystkim w środowiskach dziennikarzy:

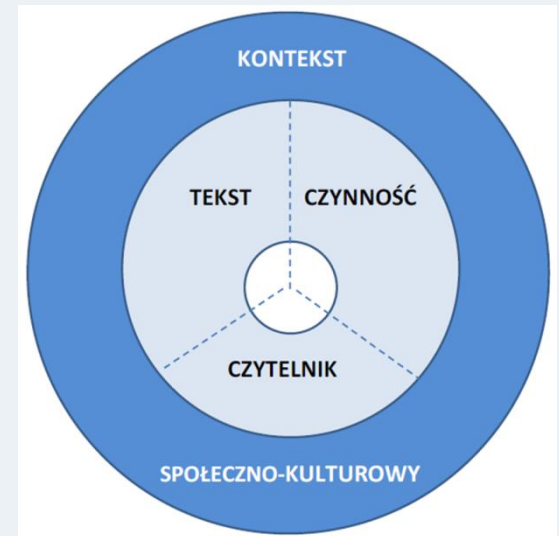
$$T = \frac{\sqrt{T_s^2 + T_w^2}}{2}$$

- Stanisław Gajda podjął próby obliczania trudności tekstów naukowych (1990) a następnie administracyjno-prawnych (2004); tą samą metodą posłużył się Andrzej Malinowski w odniesieniu do tekstów prawnych (2006).
- W ostatnich latach w Pracowni Prostej Polszczyzny na UW. opracowano i udostępniono w internecie narzędzie LOGIOS obliczające indeks mglistości dla polskiego tekstu : <http://www.logios.pl>.
- .W internecie można znaleźć inne programy, które obliczają indeks mglistości, por. np.: <http://www.trzyw.pl/gunning-fog-index/>.

Od czego zależy rozumienie tekstu?

- Od cech czytelnika: jego uwagi, pamięci, umiejętności krytycznego myślenia, wnioskowania, wizualizacji; jego motywacji i wiedzy (znajomości tematu, kompetencji językowej); osobistych doświadczeń.
- Od cech i okoliczności czytania (cel, zaangażowanie, konsekwencje).
- Od cech tekstu – zewnętrznych (graficznych) i językowych.
- Od kontekstu społeczno-kulturowego.

Autor tekstu ma wpływ tylko na cechy językowe i dlatego zajmujemy się głównie nimi.



Komu w Polsce zależy na upraszczaniu tekstów?

- Rzecznik Praw Obywatelskich (cyt.: *20—30% skarg nie trafiłoby do nas, gdyby ludzie naprawdę zrozumieli, o co chodziło*).
- Rada Języka Polskiego (m.in. raport o języku dokumentów konsumenckich).
- Niektóre ministerstwa: Ministerstwo Rozwoju Regionalnego, Ministerstwo Finansów, Ministerstwo Administracji i Cyfryzacji, Ministerstwo Zdrowia, Ministerstwo Edukacji Narodowej (ocena trudności podręczników szkolnych).
- Niektóre instytucje centralne: Korpus Służby Cywilnej, Rządowe Centrum Legislacji, Zakład Ubezpieczeń Społecznych, Agencja Rozwoju Przedsiębiorczości.
- Niektóre duże instytucje prywatne, zwłaszcza banki, towarzystwa ubezpieczeniowe.
- Szczególne znaczenie ma inicjatywa „Obywatel”:
www.obywatel.gov.pl.

Mamy nadzieję, że się przyda...

Dziękuję za uwagę!