

MONIKA ŁASZKIEWICZ

STRESZCZANIE RĘCZNE A STRESZCZENIE AUTOMATYCZNE TEKSTU NAUKOWEGO – STUDIUM PRZYPADKU

WPROWADZENIE

Dążenie do automatyzacji etapów pracy naukowej jest nie tylko ważne, ale wręcz konieczne w czasach, w których musimy sobie szybko i skutecznie poradzić z napływającymi zewsząd informacjami¹. Od wielu lat są prowadzone badania oraz wdrażane nowe rozwiązania, które mają ułatwić pracę zwłaszcza z dużymi zbiorami dokumentów lub nagrań. Automatyczna korekta tekstów, transkrypcja, anonimizacja, analiza morfologiczna, automatyczne tłumaczenia na różne języki obce, parafrazowanie, wydobywanie słów kluczowych, terminów lub obcojęzycznych wtrąceń, tworzenie własnych przeszukiwalnych korpusów, ocena zrozumiałości stworzonego tekstu i wiele innych, to możliwości, które współcześnie daje nam inżynieria lingwistyczna i humanistyka cyfrowa. Bardzo często oferowane narzędzia są ogólnodostępne, bezpłatne, a dodatkowo w ramach szkoleń i warsztatów można poznać ich funkcje i sposoby zastosowania².

Przedmiotem rozważań w tym artykule jest jakość streszczenia tekstu naukowego – chcę sprawdzić, w jaki sposób z zadaniem streszczenia tekstu poradzi

Dr MONIKA ŁASZKIEWICZ – Uniwersytet Marii Curie-Skłodowskiej w Lublinie, Wydział Filologiczny, Instytut Językoznawstwa i Literaturoznawstwa; adres do korespondencji: pl. Marii Curie-Skłodowskiej 4A, 20-031 Lublin; e-mail: monika.laszkiwicz@mail.umcs.pl; ORCID: <https://orcid.org/0000-0003-1825-444X>.

¹ Stanisława Niebrzegowska-Bartmińska pisała: „Żyjemy w świecie tekstów, obcujemy z tekstami, z tekstami stykamy się na każdym kroku, jesteśmy uwikłani w teksty, działamy tekstami – to obiegowe slogany, które potwierdzają rolę tekstów jako rzeczywistych narzędzi komunikacji i określają wagę praktycznych umiejętności związanych z tworzeniem tekstu, jego przekształcaniem i operowaniem na tekście jako gotowym wytworze” (2014, ss. 19–20).

² Od lat tego typu wykłady, warsztaty i szkolenia są oferowane przez konsorcjum Clarin-pl (zob. <https://clarin-pl.eu/index.php/o-nas/>).

sobie człowiek, a w jaki maszyna oraz jakie podobieństwa i jakie różnice można zaobserwować między tymi rodzajami streszczeń.

STAN BADAŃ

Próby stworzenia programu do automatycznego streszczania mają na świecie znacznie dłuższą tradycję niż w Polsce. Pierwsza praca *The Automatic Creation of Literature Abstracts* została opublikowana pod koniec lat pięćdziesiątych przez Hansa Petera Luhna, który wykorzystał metody statystyczne do stworzenia automatycznych streszczeń tekstów specjalistycznych (technicznych). Badał on częstość kluczowych dla publikacji słów, przyjmując założenie, że autor tekstu ważne dla siebie terminy powtarza wielokrotnie (Luhn, 1958, s. 160). Na tej podstawie oraz na podstawie odległości tych słów od siebie w zdaniu algorytm stworzony przez Luhna obliczał współczynnik istotności. Następnie wybierał zdania szczególnie ważne, dokonywał ich ekstrakcji i łączył je ze sobą w tekst streszczenia.

Od tego czasu na świecie trwają intensywne prace nad maszynowym streszczaniem tekstów w różnych językach. Piotr Glenc w 2021 roku przeanalizował liczbę publikacji na ten temat indeksowanych w bazie Scopus – do 2000 roku było ich 8, natomiast między 2016 a 2020 już 334. Łącznie od 1995 roku opublikowano 616 prac, w których poruszano tematykę automatycznego streszczania (Glenc, 2021). Stan na luty 2024 roku pokazuje skokowy wręcz wzrost opracowań – obecnie indeksowanych jest 2813³, z czego 281 prac powstało w 2021 roku, 322 – w 2022, 289 – w 2023 i już 26 – w 2024 roku.

W Polsce⁴ badania na szerszą skalę zaczęto prowadzić po 2000 roku⁵, co doprowadziło w efekcie do powstania różnych narzędzi do ekstrakcyjnego streszczania tekstów (zob. np. PolSum2, narzędzie Branny i Gajęckiego⁶, Summa-

³ Źródło: <https://www.scopus.com/search/form.uri?display=basic#basic> [dostęp: 18.02.2024], wyszukiwana fraza: "automatic AND text AND summarization".

⁴ Piotr Glenc (2021) zwraca uwagę na dwa podejścia do streszczania: „podejście ekstrakcyjne (ang. *extractive*) – wybór najważniejszych fragmentów z oryginalnego tekstu (zgodnie z przyjętymi kryteriami ważności) i tworzenie streszczenia (ekstraktu) z ich wykorzystaniem; podejście abstrakcyjne/abstraktowe (ang. *abstractive*) – tworzenie gramatycznie spójnych streszczeń (abstraktów) opisujących zawartość streszczanego dokumentu z wykorzystaniem zaawansowanych technik generowania języka naturalnego, bez wykorzystywania oryginalnych fragmentów tekstu”.

⁵ W 2002 roku Barbara Gawronska opisała swoją propozycję systemu, który umożliwiał grupowanie anglojęzycznych tekstów prasowych zgodnie z tematem oraz zestawiał kilkudzaniową informację o nich w językach: angielskim, duńskim i polskim. W badaniach wykorzystano m.in. Wordnet (Gawronska, 2002, s. 110).

⁶ Rozwiązanie to miało pewne znaczące ograniczenia, które w kolejnych projektach starano się przezwyciężyć. Po pierwsze – testowano je na tekstach prasowych i dostosowywano je wyłącznie

rizer⁷, Nicolas⁸). Streszczenia tego typu są łatwiejsze do wygenerowania, bo stanowią wyciąg kluczowych zdań z tekstu głównego. Należy jednak zaznaczyć, że dużym wyzwaniem dla algorytmów było wytypowanie słów i zdań istotnych, a w kolejności powiązanie ich w spójną i logiczną całość.

Kolejnym etapem w pracach nad skutecznym programem do streszczania tekstów w języku polskim stało się opisane w 2020 roku przez Wojciecha Ozimka narzędzie NLPer, służące do streszczania abstrakcyjnego (Ozimek, 2020, ss. 23–43) – pierwsze na naszym gruncie. Prace nad automatycznym streszczaniem trwają. Mimo najczęściej pozytywnej autoewaluacji rozwiązań, wskazywania wyższości własnego programu/algorytmu/systemu, efektywność (z punktu widzenia użytkownika streszczeń) nie jest jeszcze wysoka. Wynika to z trudności i skomplikowania tego działania na tekście – streszczanie wymaga zrozumienia tekstu, umiejętności selekcji informacji, uogólniania i kondensacji, a więc głębokiego przetworzenia tekstu, nieograniczonego jedynie do selekcji właściwych zdań.

Koniec 2022 roku przyniósł użytkownikom Internetu nowe możliwości – dynamiczny rozwój badań nad sztuczną inteligencją (AI) doprowadził do opublikowania nowego bota. Jest to zaawansowany model językowy, który został wytrenowany przez OpenAI na 45 terabajtach danych⁹. *ChatGPT (generative pre-trained transformer)* rozwiązuje wiele z dotychczasowych problemów, buduje spójne teksty, potrafi przetworzyć informacje, wykorzystać przekazane mu informacje, ma pamięć jednej sesji, więc można odwoływać się do tego, co już zostało powiedziane. Można go wykorzystać do różnych celów i zadań, w tym także do streszczania (zob. przypis 14). Mimo pojawiającej się krytyki i obaw co do bezpieczeństwa tego narzędzia, bot od początku cieszy się ogromnym zainteresowaniem użytkowników¹⁰.

do nich, po drugie – zbiór tekstów był bardzo ograniczony (tylko jedna długość streszczenia tylko trzech tekstów) oraz po trzecie – algorytm mógł być zastosowany do tekstów nie dłuższych niż 50 zdań (Branny, Gajęcki, 2005, s. 36).

⁷ Program opierał się zarówno na płytkich cechach tekstu, danych statystycznych, jak i na analizie wykorzystującej wiedzę językoznawczą (tokenizacja, stemming, lematyzacja, analiza morfologiczna, parsowanie struktury retorycznej tekstu, obliczanie statystyk, przetwarzanie i selekcja, centralność zdania) (Świetlicka, 2010, ss. 15–23).

⁸ Oparty na mechanizmach uczenia maszynowego (Kopeć, 2018, ss. 145–156).

⁹ 1 TB to 1024 GB, a dane, które wykorzystano do trenowania, objęły m.in. Wikipedię, tematyczne serwisy internetowe, publiczne bazy tekstów i książek (także naukowych), blogi oraz serwisy informacyjne, jak zauważa Przemysław Kazienko (zob. Górski, 2023), <https://www.chip.pl/2023/03/chatgpt-politechnika-wroclawska-badanie>.

¹⁰ W Wikipedii można znaleźć informację na temat niespotykanego na taką skalę zainteresowania programem ChatGPT: „zyskał [on] ponad milion użytkowników w ciągu zaledwie 5 dni (dla porównania Netflixowi zajęło to 41 miesięcy, Facebookowi 10 miesięcy, a Instagramowi 2,5 miesiąca). Zdobył 100 milionów użytkowników zajęło ChatGPT dwa miesiące, podczas gdy TikTok dokonał tego w dziewięć” (<https://pl.wikipedia.org/wiki/ChatGPT> [dostęp: 14.05.2023]).

PODSTAWA MATERIAŁOWA I METODA PRACY

Korpus tekstów, będący podstawą obserwacji i analiz, stanowi 40 streszczeń (35 streszczeń to streszczenia wykonane przez ludzi, a 5 to streszczenia maszynowe) tego samego tekstu – artykułu naukowego Jerzego Bartmińskiego pt. „Styl potoczny” z tomu *Współczesny język polski*. Wszystkie zgromadzone teksty zostały zdeponowane w *korpusomacie*¹¹, co ułatwiło mi przeszukiwanie i porównywanie zbioru.

34 streszczenia zostały sporządzone przez studentów 1 roku e-edytorskiego i technik redakcyjnych UMCS w 2017 roku i powstały w ramach zajęć z teksteologii edytorskiej¹²; jedno – streszczenie eksperckie – pozyskałam z *Bibliografii adnotowanej dorobku naukowego Profesora Jerzego Bartmińskiego* (Łaszkiewicz i in., 2024, ss. 108–109). W tym miejscu zamieszczam je oraz jedno przykładowe streszczenie studenckie, choć w analizach odwołuję się do całego zbioru.

Przykłady – streszczenie ręczne

Styl potoczny, zdaniem autora, jest pierwszym i naturalnym językiem człowieka. Stanowi centrum systemu stylowego języka i bazę derywacyjną dla innych stylów. Potoczność występuje zarówno w odmianie ustnej, jak i pisanej. Dotyczy nie tylko form wyrazu, lecz tkwi w głębszej warstwie języka, w sferze znaczeń i postrzegania rzeczywistości, sięga sfery zachowań człowieka. Styl potoczny utrwała „naiwny” obraz świata, budowany z postawy zdroworozsądkowej i punktu widzenia „prostego człowieka”. Wszystko to znajduje swój wyraz w strukturze słownictwa, jego semantyce, frazeologii, derywatach, wzorcach budowy zdań i sposobach konstruowania tekstów. Słownik podstawowy w tym stylu tworzą wyrazy nazywające części ciała, rozwój i funkcjonowanie człowieka, odżywianie, dom i wyposażenie, rodzinę itp. Styl potoczny ma rejestry – neutralny i emocjonalny oraz swobodny i staranny. Podstawowe wartości stylu potocznego to: antropocentryzm i konkretność. Mechanizmy porządkujące wypowiedź potoczną bazują na kolekcji, scenariuszu i typizacji. Potoczność wchodzi w trzy podstawowe relacje, które wyznaczają jej pozycję w systemie wariantów współczesnego języka polskiego: potoczność vs. poetyckość, potoczność vs. naukowość, potoczność vs. oficjalność. W refleksji nad potocznością autor wyróżnił dwie koncepcje: opolską, w której jest ona traktowana jako podstawowy wariant języka narodowego (co łączy się z antropologiczną koncepcją stylu), oraz warszawską, rozumianą stylistycznie, sytuacyjnie, pragmatycznie, a nie semantycznie. Sam opowiada się za koncepcją pierwszą. [SNB]

¹¹ <https://korpusomat.pl>.

¹² Studenci przed zajęciami zapoznali się z tekstem artykułu Jerzego Bartmińskiego, następnie na pierwszym spotkaniu otrzymali informacje na temat tego, czym jest streszczenie oraz jaka długość w tym wypadku jest preferowana (podano ogólną informację, że oczekiwana objętość tekstu to maksymalnie pół strony znormalizowanego maszynopisu). Studenci sami decydowali, czy przygotowane streszczenie będzie streszczeniem ekstrakcyjnym czy abstrakcyjnym – nie narzucono tego z góry. Łącznie studenci mogli na streszczenie poświęcić około 3h. Analizowany zbiór tekstów to wyniki ich pracy.

Jerzy Bartmiński w swoim tekście porusza temat stylu potocznego. Już na wstępie wspomina o tym, że język potoczny jest pierwszym wariantem języka jaki przyswajamy jako dzieci oraz jest używany przez największą liczbę osób. Pełni również rolę tła, na którym funkcjonują style wyspecjalizowane. Bez niego nie udaje się wyodrębnić i opisać innych stylów językowych. Autor zauważa, że język potoczny jest widoczny w odmianie pisanej (listy, dzienniki, pamiętniki) jak i mówionej (rozmowa telefoniczna, na ulicy, przy stole) oraz, że istnieje podział w stylu na rejestry neutralny i emocjonalny oraz swobodny i staranny. Styl potoczny ma swoje wartości takie jak: antropocentryzm, konkretność, kolektywność i typizacja. Wchodzi on również w trzy podstawowe relacje wyznaczające mu pozycję. Opozycja: potoczność a poetyckość, potoczność a naukowość oraz potoczność a oficjalność. Bartmiński porusza kwestię podejścia do stylu potocznego – koncepcji „warszawskiej”, w której mamy do czynienia z leksykalno-stylistycznym, pragmatycznym, a nie semantycznym rozumieniem potoczności oraz „opolskiej” gdzie przyznaje się potoczności status pełnowartościowego, a nawet podstawowego i głównego wariantu języka narodowego. [SKR]

Do wygenerowania streszczeń automatycznych wytypowałam różne programy dostępne dla języka polskiego. O ostatecznym wyborze zdecydowała: 1) dostępność narzędzia w sieci oraz 2) możliwość zadania oczekiwanej długości streszczenia¹³, gdyż zależało mi na tym, żeby była ona porównywalna z długością streszczeń zgromadzonych w 2017 roku¹⁴. Na potrzeby analiz wzięłam pod uwagę: program *Resoomer* w wersji płatnej PRO¹⁵, bezpłatny program, który został oparty (jak zadeklarowano na stronie) na złożonych algorytmach i sztucznej inteligencji – *Smodin*¹⁶, oraz bezpłatną wersję linguabota *ChatGPT*¹⁷. Następnie podejmowałam trzy próby wygenerowania streszczenia artykułu. *Resoomer* i *Smodin* za każdym

¹³ Nie wykorzystałam takich programów, jak NLPer (nie udało mi się go zlokalizować) oraz Summarize (wygenerowane streszczenie nie spełnia założeń – ma ponad 15 300 znaków).

¹⁴ Przewaga tego bota nad innymi, z którymi spotykałam się wcześniej, polega na tym, że prowadzona rozmowa nie wydaje się sztuczna. To skłoniło mnie do kilkukrotnego przetestowania go w tak wymagającym zadaniu, jakim jest streszczenie tekstu naukowego. Przykładowa wymiana zdań: „M: Czy umiesz pisać streszczenia? ChatGPT: Oczywiście! Mogę pomóc w napisaniu streszczenia. Proszę podaj mi tekst, który chciałbyś, abym streszczył [sic!], oraz informacje na temat długości i szczegółowości, które preferujesz w streszczeniu.”; „M: Napisz streszczenie artykułu Jerzego Bartmińskiego o stylu potocznym. ChatGPT: Niestety, nie posiadam dostępu do konkretnych artykułów Jerzego Bartmińskiego o stylu potocznym. Jestem ogólnym modelem [sic!] AI, a moja wiedza jest ograniczona do danych dostępnych do września 2021 roku. Mogę jednak podać ogólne informacje na temat stylu potocznego, jeśli Cię to interesuje. M: Artykuł Jerzego Bartmińskiego jest wcześniejszy. Pochodzi z tomu *Współczesny język polski*. ChatGPT: Przepraszam za zamieszanie. Niestety, nie mam dostępu do konkretnego artykułu Jerzego Bartmińskiego z tomu *Współczesny język polski*. Jednakże, mogę przedstawić ogólne informacje na temat stylu potocznego w języku polskim” [dostęp: 7.05.2023].

¹⁵ Źródło: <https://resoomer.com/pl>.

¹⁶ Źródło: <https://smodin.io/pl>.

¹⁷ Źródło: <https://chat.openai.com>.

razem dawał ten sam wynik, w przypadku bota *ChatGPT* otrzymałam 3 różne streszczenia.

W tym miejscu zamieszczam przykładowe streszczenie automatyczne.

Przykład – streszczenie maszynowe

Język potoczny to wyróżniony wariant języka narodowego, używany do codziennej komunikacji. Jest prosty, konkretny i bliski, umożliwiając nam zrozumienie nowych wyrażeń i kontrolę nad wieloznacznymi słowami. Styl potoczny jest podstawą dla innych stylów językowych i jest obecny zarówno w mowie, jak i piśmie. Składa się z dwóch rejestrów: neutralnego i emocjonalnego. Rejestr neutralny jest najczęściej używany i przekazuje obiektywną informację, podczas gdy rejestr emocjonalny zawiera dodatkowo emocje i postawę mówiącego. Styl potoczny ma również różne poziomy formalności: swobodny i staranny. Swobodny rejestr jest luźniejszy i zawiera obrazowe frazeologizmy, podczas gdy rejestr staranny jest bardziej formalny. Potoczność obejmuje również wartościowanie, orientację przestrzenną opartą na ciele człowieka oraz typizację i kolektywność. Istnieją trzy główne opozycje, którym podlega styl potoczny: potoczność kontra poetyckość, potoczność kontra naukowość i potoczność kontra oficjalność. Relacje między tymi stylami językowymi polegają na tym, że język potoczny jest bardziej wszechstronny i emocjonalny, podczas gdy język naukowy jest bardziej specjalistyczny. W kulturze języka potoczność jest analizowana z różnych perspektyw, a jej wartość może być oceniana zarówno przez aspekty leksykalno-stylistyczne, jak i pragmatyczne. [ChatGPT1]

W kolejności dokonałam analizy jakościowej zgromadzonego materiału. Sprawdziłam obecność kluczowych elementów treści (słów, fraz i zdań) w tekstach, aby ocenić ich walor poznawczy, oraz przyjrzałam się różnym nieścisłościom i błędom, by sprawdzić, czy mają one związek z tym, w jaki sposób streszczenie powstało (kto je wykonał).

STRESZCZENIE RĘCZNE A STRESZCZENIE MASZYNOWE ARTYKUŁU JERZEGO BARTMIŃSKIEGO „STYL POTOCZNY”

Przygotowanie streszczenia składa się z szeregu złożonych operacji myślowych: od oswojenia się z tekstem wyjściowym (poprzez czytanie), przez jego zrozumienie, określanie głównej myśli, wyszukanie najważniejszych informacji, po zredagowanie ostatecznej wersji tekstu docelowego (por. Dąbrowska, 2014, ss. 30–31)¹⁸. Także systemy generujące streszczenia automatyczne muszą wykonać

¹⁸ Na temat etapów tworzenia streszczenia lub elementów, które powinny zostać w nim zawarte, powstało wiele prac, por. np. Bartmiński (1992, ss. 7–14); Zaśko-Zielińska i in. (2008, ss. 66–72); Wróbel (2014, s. 6); Niebrzegowska-Bartmińska (2014, ss. 24–29).

pewne kroki (wykorzystując do tego zadania algorytmy) – od wstępnego przetwarzania tekstu wyjściowego (analiza na różnych poziomach języka), jego głębszą transformację, przez selekcję istotnych informacji, a następnie łączenie wszystkich informacji na dany temat w blok (agregacja), uogólnienie zebranych informacji, po stworzenie tekstu streszczenia (Świetlicka, 2010, ss. 7–8). *ChatGPT* referuje proces tworzenia streszczenia przez siebie w następujący sposób: „Streszczenie tekstu opiera się na analizie i zrozumieniu treści, której dokonuje model językowy, taki jak ja. Proces ten obejmuje: 1) Zrozumienie tekstu: Model analizuje tekst, rozpoznaje znaczenie słów, zdań i kontekst. 2) Identyfikacja kluczowych informacji: Model wyodrębnia główne pomysły, wydarzenia, postacie i inne istotne elementy. 3) Odrzucenie szczegółów [...]. 4) Tworzenie skrótu [...]. 5) Sprawdzanie spójności [...]. 6) Prezentacja skrótu: Model prezentuje skrót tekstu w sposób, który zawiera najważniejsze elementy oryginalnego tekstu, ale jest bardziej zwięzły”¹⁹.

Streszczenie, będąc tekstem „drugiego stopnia”, może zostać uznane za metatekst, „tekst o tekście” (por. Bartmiński, Niebrzegowska-Bartmińska, 2009, ss. 186–191; Łaskiewicz, 2017, s. 153). Jednak, by tak się stało, muszą zostać spełnione podstawowe warunki. Porównując ze sobą streszczenia ręczne i maszynowe, a następnie oceniając ich jakość, warto wziąć pod uwagę selekcję informacji – właściwe wytypowanie najważniejszych informacji, przy jednoczesnym zidentyfikowaniu i pominięciu informacji drugorzędnych oraz unikaniu redundancji – są to umiejętności szczególnie potrzebne w przypadku tworzenia krótkich tekstów, bo prowadzą do precyzji i pozwalają uniknąć rozmycia treści.

W zgromadzonym materiale widać wyodrębnienie przez autorów streszczeń ręcznych tych samych informacji – kluczowych słów, fraz lub zdań. Przeprowadzona analiza tekstów pozwala zauważyć, które z informacji pojawiają się najczęściej. Niekiedy wyekscerpowano je z artykułu w niezmienionej formie, innym razem zostały one przez autorów sparafrazowane. Żadne ze streszczeń ręcznych nie jest wyłącznie ekstrakcyjne, a stopień parafrazowania, uogólniania, łączenia ze sobą informacji i pisanie „własnymi słowami” rośnie wraz ze wzrostem kompetencji autora.

W streszczeniach przygotowywanych ręcznie pojawiły się informacje przede wszystkim o tym, że styl potoczny ma swoiste wartości (35 streszczeń), wśród nich najczęściej wymieniano antropocentryzm (33) oraz konkretność (31 razy). W 33 streszczeniach jest mowa o *trzech podstawowych relacjach/opozycjach, w jakie wchodzi styl potoczny*. W 32 uwzględniono, że potoczność *obejmuje szeroką gamę wypowiedzi ustnych (oralnych) i pisemnych/występuje w odmianie ustnej*

¹⁹ Etapy przygotowania streszczenia przez *ChatGPT* [dostęp: 11.02.2024].

i w odmianie pisanej. W 27 streszczeniach znalazła się informacja, że jest centrum systemu stylowego języka lub że styl potoczny zajmuje centralną pozycję pośród innych stylów. W 26 była mowa o tym, że styl potoczny to baza derywacyjna/punkt wyjścia dla innych stylów. Kluczowe okazuje się, zdaniem anotatorów, także stwierdzenie, że język potoczny jest przyswajany jako pierwszy w kolejności/jest pierwszym językiem człowieka – informacja ta została uwzględniona w 21 streszczeniach. Jedenaście osób zwróciło uwagę na eksplikacyjną funkcję stylu i stwierdziło, że pozwala [on] na uproszczenie wypowiedzi trudnej/jest najbardziej zrozumiały/dzięki niemu jesteśmy w stanie zrozumieć wszystkie inne style²⁰.

Tylko w 9 streszczeniach wymieniono *mechanizmy porządkujące: kolekcje, scenariusz i typizację*; jedynie w 8 jest mowa o tym, że styl potoczny utrwała naiwny obraz świata. Rzadziej pojawiały się też informacje o tym, że język potoczny jest używany najczęściej / przez największą liczbę osób lub że jest powszechny (7); że jest on *naturalnym językiem człowieka* (7).

Mimo różnego stopnia nasycenia tekstów słowami czy frazami kluczowymi, można uznać, że streszczenia przygotowane ręcznie spełniają wymóg informacyjności. Stopień informacyjności jest jednak zindywidualizowany – uzależniony od kompetencji autora.

W zbiorze streszczeń automatycznych również można zaobserwować różną ich jakość. W przypadku wyniku wygenerowanego przez *Smodin* trudno nawet mówić o streszczeniu, bo jest to raczej zbiór przypadkowych fraz, zdań lub fragmentów tekstu wyjściowego, który nie spełnia podstawowej funkcji informacyjnej:

por.: Buttlera, 1978a; Buttlar, Markowski, 1988; Markowskiego, 1990). Wśród wariantów języka narodowego język potoczny wyróżnia miejsce, wręcz wyjątkowe. Polega ona m.in. Co to jest demokracja? moje – J.B.] doświadczenia językowego” (Bogusławski, Język w błędzie, 1988, s. 73). Odpowiedź: „To, co postrzegamy wtedy, kiedy patrzymy, że stanie się nam coś złego, i pragniemy, żeby się nie stało” [Smodin]²¹.

Nieco lepiej z zadaniem poradził sobie płatny *Resoomer*, choć otrzymany wynik wymaga od człowieka uzupełnień i redakcji, a informacje ważne (podkreślone przeze mnie) funkcjonują w tekście obok informacji drugorzędnych lub nieistotnych:

forma rządów, w której wszyscy obywatele są równi wobec prawa i mają prawo brać udział w procesie podejmowania i wdrażania decyzji. Język i komunikacja są integralną częścią życia człowieka, a pierwszym językiem, którego się uczy my, jest język

²⁰ Dodatkowo w kolejnych czterech streszczeniach jest mowa o tym, że jest to język *łatwy do przyswojenia*, a nawet *najprostszy*.

²¹ Zamieszczony fragment jest początkiem wygenerowanego przez narzędzie streszczenia.

potoczny. Jest to język najprostszy i najbardziej konkretny, pozwalający nam porozumiewać się w codziennych sytuacjach. Gdy się starzejemy i musimy omawiać bardziej złożone tematy, pierwszy język służy jako podstawa do zrozumienia i wyjaśnienia nowych słów i pojęć, a także do kontrolowania niejednoznacznych i niejasnych słów. Pozwala nam również wyjaśnić bardziej złożone pojęcia, takie jak strach czy demokracja [Resoomer].

W przypadku streszczeń wygenerowanych przez *ChatGPT* można mówić o nagromadzeniu słów i fraz kluczowych, które jednakże nie zawsze pojawiają się we właściwym kontekście: *Styl potoczny wyraża wiele wartości, takich jak antropocentryzm, konkretność, kolektywność i typizacja. Styl potoczny jest praktyczny, zdroworoządkowy i oparty na konkretnych doświadczeniach człowieka [ChatGPT 2]* – wartości stylu i mechanizmy porządkujące rzeczywistość zostały ze sobą zrównane. Niekiedy wygenerowane teksty brzmią sztucznie: *W mediach o dużej liczbie czytelników, takich jak „Gazeta Wyborcza” stosuje się prosty, codzienny język, aby czytelnik czuł się osobiście zwrócony [ChatGPT 2]; Często prosimy rozmówcę, aby przetłumaczył swoją wypowiedź na nasze słowa i mówił zrozumiałym językiem [ChatGPT 2].*

Zdarzają się sformułowania mało precyzyjne: *W kulturze języka potoczność jest analizowana z różnych perspektyw, a jej wartość może być oceniana zarówno przez aspekty leksykalno-stylistyczne, jak i pragmatyczne [ChatGPT1]* – jest tu co prawda mowa o podejściu do potoczności z różnych perspektyw, ale pominięte zostają szczegółowe informacje, które dla zrozumienia artykułu Bartmińskiego są kluczowe (na czym polega wskazane patrzeć na potoczność z różnych perspektyw).

W tekstach generowanych automatycznie, mimo ich porównywalnej długości do tekstów sporządzonych ręcznie, nagromadzenie słów i fraz kluczowych jest zróżnicowane. Najlepiej wypada *ChatGPT*, wygenerowane (wszystkie trzy) wyniki są zbliżone do tekstów przygotowanych przez człowieka, gdyż uwzględniają ważne informacje z artykułu (np. o tym, że styl potoczny jest prosty i bliski, że cechują go swoiste wartości, jak antropocentryzm i konkretność, że jest pierwszym językiem, że stanowi centrum systemu stylowego i że jest podstawą dla innych stylów). Mniej kluczowych informacji zostało zawartych w streszczeniu wykonanym programem *Resoomer*, gorzej także to streszczenie wypada pod względem selekcji materiału oraz pod względem językowym. Streszczenie wygenerowane przez *Smodin* jest całkowicie błędne.

Problemem pojawiającym się przy okazji selekcji materiału jest umiejętność pomijania informacji drugorzędnych. Bez straty dla zrozumiałości, a z zyskiem dla kondensacji treści można byłoby zrezygnować w streszczeniach przygotowywanych ręcznie chociażby z przykładów: *Autor zauważa, że język potoczny*

jest widoczny w odmianie pisanej (*listy, dzienniki, pamiętniki*) jak i mówionej (*rozmowa telefoniczna, na ulicy, przy stole*) [SKR]. Zdarzają się informacje błędnie wytypowane przez autora jako ważne i wprowadzone do tekstu: *Istnieje współczesny ekspansywny rejestr wulgaryzmów, który jest określany jako „niski”* [SAK]. Należy jednak zaznaczyć, że informacje te, jeśli pojawiają się w streszczeniach sporadycznie, to nie zakłócają jego informacyjności.

W przypadku streszczeń automatycznych, w których wydobycie właściwych słów kluczowych nie zawsze zostało wykonane w sposób zadowalający, włączanie fragmentów drugorzędnych prowadzi do jeszcze większego rozmycia treści, np. obok kluczowego: *Istnieją trzy główne opozycje, którym podlega styl potoczny: potoczność kontra poetyckość, potoczność kontra naukowość i potoczność kontra oficjalność*, pojawia się zdanie: *Relacje między tymi stylami językowymi polegają na tym, że język potoczny jest bardziej wszechstronny i emocjonalny, podczas gdy język naukowy jest bardziej specjalistyczny* [ChatGPT1], które z jednej strony jest zbyt wąskie (nie zostały opisane pozostałe opozycje przedstawiane w tekście, tj. potoczność a poetyckość i potoczność a oficjalność), z drugiej zaś – niepotrzebnie włączane do tekstu streszczenia. W analizowanym streszczeniu dużo uwagi poświęcono także rejestrom stylu potocznego, co przy nieuwzględnieniu głównych wartości tego stylu budzi zastrzeżenia. Trzeba jednak przyznać, że ChatGPT poradził sobie z zadaniem znacznie lepiej niż programy Resoomer czy Smodin, a wygenerowane streszczenia automatyczne trudniej jest odróżnić od streszczeń ręcznych. Ma to związek z ilością danych, na których ChatGPT był trenowany, bo dzięki temu język przygotowanych streszczeń jest bardziej naturalny, bogaty, zrozumiały dla użytkownika, a wygenerowany tekst wydaje się także bardziej spójny i profesjonalny.

PODSUMOWANIE

Streszczenia ręczne i streszczenia maszynowe są efektem wieloetapowej pracy człowieka lub algorytmów. Analiza wielu streszczeń tego samego tekstu naukowego pokazuje mocne i słabe strony obu sposobów streszczania. Widać, że nadal streszczenia przygotowane ręcznie są dokładniejsze i przekazują więcej informacji, ale tylko i wyłącznie wtedy, kiedy są przygotowywane przez kompetentne osoby. W streszczeniach ręcznych trafniej identyfikowano kluczowe elementy treści i skuteczniej unikano umieszczania informacji drugorzędnych. Botowi ChatGPT udało się tego dokonać w mniejszym stopniu, ale otrzymane wyniki dają wgląd w treść artykułu. Płatny program Resoomer nie poradził

sobie z tym zadaniem w zadowalającym stopniu. Wynik wymagałby poprawek ręcznych – usunięcia błędów, dopisania brakujących fragmentów, redakcji (a więc dodatkowego czasu i nakładu pracy). Program *Smodin* w ogóle sobie nie poradził z zadaniem.

Każde z 35 streszczeń ręcznych jest inne mimo podobnie wytypowanych słów i fraz kluczowych. Streszczenie wygenerowane przez *Resoomer* i *Smodin* jest zawsze takie samo, natomiast streszczenia generowane przez *ChatGPT* są za każdym razem inne. To zbliża nowoczesne streszczanie automatyczne do tradycyjnego sposobu streszczania tekstu przez człowieka (a nawet różnych ludzi), którzy są w stanie wygenerować różne, poprawne streszczenia tego samego tekstu²².

Przygotowanie ręcznego streszczenia tekstu jest czasochłonne. Przewaga wszelkich automatycznych narzędzi w tym punkcie jest wyraźna – automatyczne streszczenie zajęło programom mniej niż minutę.

Z narzędzi do automatycznego streszczania mogą korzystać i autorzy (którzy potrzebują streszczeń do celów publikacyjnych), i czytelnicy artykułów naukowych (którzy chcą szybko zorientować się w treści danego artykułu i ocenić na tej podstawie jego przydatność do własnych badań). Jednym i drugim pozwala zaoszczędzić czas poświęcony na – w pierwszym przypadku – redakcję tekstu streszczenia, w drugim przypadku – przeczytanie całego tekstu wyjściowego w sytuacji, gdyby miał okazać się nieprzydatny dla przyjętych celów poznawczych; nawet jeśli jakość takiego streszczenia nie jest jeszcze w pełni satysfakcjonująca.

*

Dla wzbogacenia analizy warto byłoby uwzględnić także inne aspekty (pominięte w tym artykule z powodu obowiązujących ograniczeń): (1) błędy rzeczowe – ponieważ nawet właściwa selekcja materiału nie zagwarantuje, że powstanie dobre streszczenie, jeśli pojawią się w tekście deformacje treści, fałszujące ustalenia i wnioski zawarte w streszczanym artykule; (2) budowa streszczenia – wzięcie pod uwagę poprawności jego struktury, która powinna oddawać strukturę artykułu i w której informacje powinny się pojawiać w kolejności omawiania ich w tekście wyjściowym; (3) spójność i integralność, które za autorami *Tekstologii* można potraktować jako najważniejsze cechy tekstu

²² Na tę potencjalną możliwość generowania różnych streszczeń tego samego tekstu zwracał uwagę już O. Wojtasiewicz. W swojej teorii streszczenia zakładał, że t_j jest streszczeniem t_i (jednym z wielu możliwych) wtedy i tylko wtedy, gdy a) t_j nie jest tekstem pustym, b) t_j nie składa się z tautologii logicznych, c) zbiór konsekwencji tekstu t_j nie równa się zbiorowi wszystkich zdań, d) długość t_j jest mniejsza niż długość t_i , e) zbiór konsekwencji tekstu t_j zawarty jest w zbiorze konsekwencji t_i . (Wojtasiewicz, 1977, s. 188).

(Bartmiński, Niebrzegowska-Bartmińska, 2009, s. 292); (4) błędy językowe – kryterium świadczące o kompetencjach językowych i dbałości o kulturę wypowiedzi (ważne zwłaszcza w przypadku streszczeń ręcznych, ale pokazujące również stopień zaawansowania narzędzi do automatycznego streszczania). Pozatekstowym, ale także ważnym kryterium z punktu widzenia osób tworzących streszczenia jest także czas i nakład pracy potrzebny do jego przygotowania, a z punktu widzenia osób poszukujących informacji – czas i nakład pracy potrzebny do ich zdobycia (5). Kwestie te rozwinę w kolejnych artykułach.

BIBLIOGRAFIA

- Bartmiński, J. (1992). Streszczenie w aspekcie typologii tekstów. W: T. Dobrzyńska (red.). *Typy tekstów. Zbiór studiów* (ss. 7–14). IBL PAN.
- Bartmiński, J. (1993). Styl potoczny. W: J. Bartmiński (red.). *Encyklopedia kultury polskiej XX wieku. T. 2: Współczesny język polski* (ss. 115–134). Wiedza o Kulturze.
- Bartmiński, J. (2015). O działaniach na tekście z perspektywy tekstologii integralnej. W: S. Niebrzegowska-Bartmińska, M. Nowosad-Bakalarczyk, T. Piekot (red.). *Działania na tekście w edukacji szkolnej i uniwersyteckiej* (ss. 9–18). Wydawnictwo UMCS.
- Bartmiński, J., Niebrzegowska-Bartmińska, S. (2009). *Tekstologia*. Wydawnictwo Naukowe PWN.
- Branny, E., Gajęcki, M. (2005). Text summarizing in Polish. *Computer Science*, 7, 31–48.
- Dąbrowska, D. (2014). Jak ćwiczyć umiejętność streszczania. *Polonistyka*, 1, 30–34.
- Dudczak, A. (2007). Zastosowanie wybranych metod eksploracji danych do tworzenia streszczeń tekstów prasowych dla języka polskiego [praca magisterska]. Politechnika Poznańska, <http://www.cs.put.poznan.pl/dweiss/research/lakon/publications/thesis.pdf> [dostęp: 20.04.2023].
- Gajda, S. (1982). *Podstawy badań stylistycznych nad językiem naukowym*. Wydawnictwo Naukowe PWN.
- Gawronska, B. (2002). Employing cognitive notions in multilingual summarization of news reports. W: S. Winter (red.). *Proceedings of NLULP-02* (ss. 103–119), <http://www.cs.haifa.ac.il/~shuly/nlulp02/papers/gawronska.Pdf> [dostęp: 10.04.2023].
- Glenc, P. (2021). Narzędzia do automatycznego streszczania tekstów w języku polskim. Stan badań naukowych i prac wdrożeniowych. *E-mentor*, 2 (89), <https://www.e-mentor.edu.pl/mobi/artukul/index/numer/89/id/1513> [dostęp: 6.04.2023].
- Górski, S. (2023). Polscy naukowcy prześwietlili ChatGPT. Jak to mówią, jak coś jest do wszystkiego.... *Chip*, <https://www.chip.pl/2023/03/chatgpt-politechnika-wroclawska-badanie> [dostęp: 20.05.2023].
- Kopec, M. (2018). Summarization of Polish press articles using coreference [praca doktorska]. Instytut Podstaw Informatyki Polskiej Akademii Nauk, <http://zil.ipipan.waw.pl/MateuszKopec?action=AttachFile&do=view&target=m.kopec-phd-thesis.pdf> [dostęp: 10.04.2023].
- Luhn, H.P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2 (2), 159-165, <https://doi.org/10.1147/RD.22.0159>.
- Łaskiewicz, M. (2017). Streszczenie, abstrakt, adnotacja i ich funkcja w tekście naukowym. W: U. Sokółska (red.). *Socjolekt – idiolekt – idiolekt: historia i współczesność* (ss. 149-166), https://repozytorium.uwb.edu.pl/jspui/bitstream/11320/9711/1/M_Laskiewicz_Streszczenie_abstrakt_adnotacja_i_ich_funkcja_w_tekscie_naukowym.pdf [dostęp: 30.05.2023].

- Łaskiewicz, M., Niebrzegowska-Bartmińska, S., Nowosad-Bakalarczyk, M. (2024). *Bibliografia adnotowana dorobku naukowego Profesora Jerzego Bartmińskiego*. Wydawnictwo UMCS.
- Niebrzegowska-Bartmińska, S. (2014). Działania na tekście. *Polonistyka*, 1, 24-29.
- Ozimek, W. (2020). Automatic summary of texts in Polish [praca magisterska]. Uniwersytet Jagielloński w Krakowie, <https://thesiscommons.org/a72g3/> [dostęp: 1.05.2023].
- Świetlicka, J. (2010). Metody maszynowego uczenia w automatycznym streszczeniu tekstów [praca magisterska]. Uniwersytet Warszawski, <http://nlp.ipipan.waw.pl/~adamp/msc/swietlicka.joanna/TekstPracy.pdf.gz> [dostęp: 1.04.2023].
- Wojtasiewicz, O.A. (1977). Próba formalnej definicji pojęcia streszczenia. *Studia Semiotyczne*, 7, 185–191.
- Wróbel, I. (2014). *Jak napisać dobre streszczenie? Poradnik dla ucznia (i nauczyciela)*. Instytut Badań Edukacyjnych, <http://www.ibe.edu.pl/images/prasa/Materia%20pomocniczy%20do%20matury%20-%20streszczenie%20.pdf> [dostęp: 4.05.2016].
- Zaśko-Zielińska, M., Majewska-Tworek, A., Piekot, T. (2008). *Sztuka pisania. Przewodnik po tekstach użytkowych*. Wydawnictwo Naukowe PWN.

NARZĘDZIA

- ChatGPT, <https://chat.openai.com>
- Korpusomat, <https://korpusomat.pl>
- Resoomer, <https://resoomer.com/pl>
- Smodin, <https://smodin.io/pl>

STRESZCZANIE RĘCZNE A STRESZCZENIE AUTOMATYCZNE
TEKSTU NAUKOWEGO – STUDIUM PRZYPADKU

Streszczenie

Celem artykułu było porównanie streszczeń tego samego tekstu – artykułu naukowego Jerzego Bartmińskiego pt. „Styl potoczny”. 35 streszczeń zostało wykonanych ręcznie, 5 to rezultat streszczenia maszynowego (wykorzystano: *Resoomer*, *Smodin*, *ChatGPT*). Analiza wykazała, że streszczenia przygotowane ręcznie spełniają wymóg informacyjności, jednak jej stopień jest zindywidualizowany – uzależniony od kompetencji autora. W streszczeniach tradycyjnych trafniej identyfikowano kluczowe elementy treści i skuteczniej unikano umieszczania informacji drugorzędnych. Botowi *ChatGPT* udało się tego dokonać w mniejszym stopniu niż człowiekowi, ale otrzymane wyniki dają wgląd w treść artykułu. Streszczenie otrzymane przy pomocy programu *Resoomer* wymaga dodatkowej pracy (uzupełnień, redakcji). Efekt pracy narzędziem *Smodin* trudno uznać za tekst. Każde ze streszczeń ręcznych jest inne, podobnie jak streszczenia generowane przez *ChatGPT*. Streszczenia wygenerowane przez *Resoomer* i *Smodin* są zawsze takie same. Widać więc pewne podobieństwo nowoczesnego streszczenia automatycznego do sposobu streszczenia przez człowieka (a nawet różnych ludzi), którzy są w stanie wygenerować różne, poprawne streszczenia tego samego tekstu.

Słowa kluczowe: streszczenie; streszczenie ręczne; streszczenie automatyczne; streszczenie maszynowe

MANUAL SUMMARIZATION *VERSUS* AUTOMATIC SUMMARIZATION
OF A SCIENTIFIC TEXT: A CASE STUDY

Summary

The purpose of the article was to compare summaries of the same text – a scientific article by Jerzy Bartmiński titled “Colloquial Style”. Thirty-five abstracts were made by hand, five were the result of machine summarization (*Resoomer*, *Smodin*, *ChatGPT* were used). During the analysis, it was noted that the abstracts prepared by hand met the requirement of informativeness, but its degree is individualized – depending on the competence of the author. These abstracts more accurately identified key elements of the content and more effectively avoided the inclusion of secondary information. *ChatGPT* managed to do this to a lesser extent, but the results obtained provide insight into the content of the article. The summary obtained with *Resoomer* requires additional human work (additions, editing). The result of the *Smodin* tool can hardly be considered a text. Each of the manual summaries is different, as are the summaries generated by *ChatGPT*. The abstracts generated by *Resoomer* and *Smodin* are always the same. Thus, one can see some similarity in modern automatic summarization to the traditional way of summarizing by humans (or even different humans), who are able to generate different, correct summaries of the same text.

Keywords: summary; human summarization; automatic summarization; machine summarization