


## Sztuczna inteligencja (SI) w badaniach lingwistycznych

### Artificial Intelligence (AI) in Linguistic Research

**Alicja SZTUK**

Uniwersytet Warszawski/ University of Warsaw

E-mail: [alicja.sztuk@uw.edu.pl](mailto:alicja.sztuk@uw.edu.pl), 

**Abstract:** The main purpose of the paper is both to present and to highlight the wide range of artificial intelligence appliance in linguistic research. I intend to define the so called ‘linguistic intelligence’ in the sense of machine learning, based mainly on artificial neural networks. Linguistic intelligent solutions seem to be not only up-to-date but also very promising in the area of developing and improving any intelligent linguistic tools, such as intelligent tutoring systems that are able to interact with human being, or the voice (speech) recognition systems that are able to receive, interpret (understand) and sometimes even carry out spoken commands. Finally, I intend to present the area of so called ‘terminotics’. The term refers to the meeting point of three interrelated disciplines: terminology, computational linguistics and linguistic engineering. This branch is also assisted by computer tools and new technologies based on artificial intelligence and machine learning. These (tools) are mainly designed for term extraction and corpora development but lately there are also some new possibilities to use these tools to increase the quality of terminology infrastructure as well.

**Keywords:** artificial intelligence, machine learning, linguistic intelligence, linguistic research, intelligent tutoring system, linguistic smart software system for glottodidactics and translation, intelligent voice recognition, chatbot, terminotics

### 1. Wstęp

Sztuczna inteligencja (SI) jest dziś bezsprzecznie popularna i ten trend nie zmienia się najprawdopodobniej w ciągu najbliższych lat. Może on jedynie, jak przypuszczam, przybrać na sile. Wszystko za sprawą możliwości jej (SI) praktycznego zastosowania, jak i bardzo gwałtownego rozwoju w tym zakresie (więcej na ten temat zob. A. Sztuk 2018). Wiele dziedzin ludzkiej działalności postawiło na tzw. inteligentne rozwiązania, aby przede wszystkim zwiększyć swoją efektywność, konkurencyjność, wydajność, bezawaryjność, a także, by zautomatyzować pewne procesy i w efekcie końcowym móc zredukować koszty do niezbędnego minimum.

Konsekwencją rozwoju idei sztucznej inteligencji jest uczenie maszynowe, czyli inaczej samouczenie się maszyn (ang. *machine learning*). Termin *uczenie maszynowe* został wprowadzony, by podkreślić możliwości praktycznego zastosowania dokonanych w zakresie SI. Mowa tu o tworzeniu automatycznych systemów potrafiących (samo)doskonalić się przy pomocy zgromadzonego doświadczenia (danych) oraz nabywania na tej podstawie nowej wiedzy (praktycznych umiejętności do wykona-

nia określonych działań). Zastosowanie odpowiednich algorytmów<sup>1</sup> ma za zadanie pozwolić oprogramowaniu na zautomatyzowanie procesu pozyskiwania i analizy danych, w celu ulepszenia własnego systemu. Proces uczenia się maszyn, chociaż porównywany do uczenia ludzkiego, jest od niego odmienny, bowiem samouczenie się maszyn polega przede wszystkim na analizowaniu ogromnych ilości danych, wyciąganiu wniosków na podstawie tych danych oraz wykrywaniu nawet najmniejszych zależności między tymi danymi, tzw. danymi wejściowymi (ang. *input data*) (por. A. Sztuk 2018a). H. Simon (1983: 76) definiuje *uczenie maszynowe* jako „zmiany w systemie, które mają charakter adaptacyjny w tym sensie, że pozwalają systemowi wykonać za następnym razem takie samo zadanie lub zadania podobne bardziej efektywnie”.

Początków uczenia maszynowego można doszukiwać się na przełomie lat pięćdziesiątych i sześćdziesiątych<sup>2</sup>, jednak za zdecydowany przełom w tym zakresie uznaję, za P. Domingosem (2015), początek lat dziewięćdziesiątych, kiedy to Gerald Tesauro stworzył program *TD-Gammon*, potrafiący konkurować w grze *Backgammon* z mistrzami świata<sup>3</sup>.

W świetle faktu, iż SI ma tak szerokie spektrum zastosowania w wielu dziedzinach ludzkiej działalności, rodzi się dla piszącej te słowa jako lingwisty – terminologa – pytanie, jak rozwój SI, a wraz z nim rozwój uczenia maszynowego, wpływa/wpłynął na lingwistykę. Mam tu na myśli zarówno rozwój badań w tym zakresie, jak i rozwiązania praktyczne implementowane jako tzw. inteligentne rozwiązania i narzędzia lingwistyczne.

W niniejszym artykule podejmę próbę odpowiedzi na pytanie, w jakim stopniu fenomen SI, a wraz z nim uczenie maszynowe, dają się wykorzystać (i są rzeczywiście aktualnie wykorzystywane) w badaniach i/lub praktycznych rozwiązaniach lingwistycznych. W tym celu po pierwsze podejmę próbę zdefiniowania wyrażenia *lingwistyczna inteligencja* w kontekście dokonań w zakresie SI. Ponadto dokonam przekrojowego przeglądu aktualnych tendencji w praktycznym zastosowaniu inteligentnych rozwiązań lingwistycznych oraz przedstawię pewne konkretne przykłady zastosowania tzw. inteligentnych narzędzi lingwistycznych.

---

<sup>1</sup> Algorytm – to sekwencja instrukcji mówiących komputerowi, co ma zrobić. Komputery składają się z miliardów maleńkich przełączników zwanych *tranzystorami*, a algorytmy sterują przełączaniem się tranzystorów miliardy razy na sekundę (P. Domingos 2015: 21).

<sup>2</sup> Pierwszym przykładem maszynowego uczenia się może być projekt Arthura Samuela z firmy IBM, który w latach 1952–1962 rozwijał program do szkolenia zawodników szachowych. Przełomem w dziedzinie sztucznej inteligencji i maszynowego uczenia się było powstanie systemu eksperckiego *Dendral* na Uniwersytecie Stanforda w 1965 roku. System ten powstał w celu zautomatyzowania analizy i identyfikacji molekuł związków organicznych, które dotychczas nie były znane chemikom. Wyniki badań, otrzymane dzięki systemowi *Dendral*, opublikowano w prasie specjalistycznej. Było to pierwszym w historii odkryciem dokonany przez komputer.

<sup>3</sup> Aby dojść do takiej perfekcji program ten uczył się swojej strategii grając jako przeciwnik w ponad milionie gier i w ten sposób zdobywał doświadczenie (dane).

Na zakończenie podejmę próbę ogólnego przedstawienia stosunkowo nowej<sup>4</sup>, jednakże pręźnie rozwijającej się, poddziedziny<sup>5</sup> – tzw. *terminotyki*, która wykorzystuje osiągnięcia SI w badaniach i rozwiązaniach terminologicznych oraz terminograficznych.

## 2. Lingwistyczna inteligencja w świetle sztucznej i ludzkiej inteligencji

Wspomniane wcześniej uczenie maszynowe jest często mylone lub błędnie utożsamiane z SI. Z czysto technicznego punktu widzenia, jak zostało to zaznaczone we wstępie, uczenie maszynowe to dziedzina należąca do SI lub po prostu jej poddziedzina. Celem SI jest nauczenie komputerów tego, co ludzie obecnie robią lepiej, a uczenie jest zapewne najważniejszą z tych umiejętności. Bez niej żaden komputer nie dotrzyma kroku człowiekowi na dłuższą metę; bo tylko dzięki zdolności uczenia się można pozyskiwać kolejne informacje, które zaś mogą przełożyć się na wiedzę, także tę aplikatywną oraz umiejętności i sprawności (por. P. Domingos 2016: 27). Reasumując, przyjmuję, że zdolność uczenia się i zdobywania nowej wiedzy oraz umiejętności jest najwyższym przejawem inteligencji, zarówno człowieka, jak i maszyny (jeśli założymy, że poziom inteligencji maszyn będzie w przyszłości w pewnych aspektach porównywalny z poziomem inteligencji ludzkiej, a wiele dzisiejszych opracowań teoretycznych i rozwiązań praktycznych na ten fakt wskazuje). Kwestią otwartą pozostawiam natomiast, w jakim stopniu jest/może być inteligentny człowiek, a w jakim stopniu przypisać można tę cechę maszynie oraz, czy w ogóle możliwe jest lub będzie w przyszłości porównywanie<sup>6</sup> tych dwóch zasobów inteligencji, a więc ludzkiej z komputerową.

Zanim jednak zagłębię się w uczenie maszyn, pragnę wrócić do inteligencji jako takiej. Stawiam w tym miejscu pytanie, czy istnieje zależność między inteligencją maszyny (komputera), a inteligencją człowieka. Z. Vetulani (2004) twierdzi, że zależność taka jest i to znacząca, gdyż ta pierwsza powinna być jak najwierniejszym odwzorowaniem tej drugiej. Zdaniem badacza (2004: 23 i nast.) zadaniem sztucznej inteligencji jest komputerowe modelowanie tych funkcji czy przejawów aktywności myślowej człowieka, których realizacja angażuje wyższe czynności centralnego układu nerwowego. A więc sztuczna inteligencja ma, w idealnym założeniu, naśladować i powielać funkcje kognitywne mózgu ludzkiego.

---

<sup>4</sup> Terminotyka jest, w moim przekonaniu, nowym kierunkiem badawczym w Polsce. Jak dotąd opublikowano niewiele prac poświęconych tej tematyce. Za prekursorów w tej dziedzinie (zarówno w jej rozwoju, jak i opisie teoretycznym) uznaje się Kanadę. To właśnie uniwersytety kanadyjskie (prym w tej dziedzinie wiedzy Uniwersytet w Ottawie) (por. M. Tryuk 2017: 133) propagują wykorzystanie narzędzi elektronicznych na wszystkich etapach pracy terminograficznej.

<sup>5</sup> Na aktualnym etapie uznaję terminotykę za poddziedzinę zarówno terminologii, jak i lingwistyki komputerowej oraz inżynierii lingwistycznej.

<sup>6</sup> Pisząc o możliwości porównywania obu zasobów inteligencji mam na myśli w tym miejscu osiągnięcie przez komputer takiego poziomu inteligencji, który nie będzie klasyfikowany przez nieświadomego (niepoinformowanego) człowieka jako zachowanie nieprawdopodobne dla istoty ludzkiej.

Mając na myśli naśladowanie ludzkiej inteligencji rozumieć przede wszystkim tworzenie przez człowieka sztucznych, zewnętrznych sieci neuronowych (ang. *artificial neural networks*), które są inspirowane strukturą budowy naturalnych (ludzkich) neuronów i łączących je synaps. Takie podejście określa się w literaturze przedmiotu *podejściem koneksjonistycznym* (por. m.in.: P. Domingos 2016: 67; M. Lister i in. 2009: 626; P. Szerszeń 2014: 50). Ponadto P. Szerszeń (tamże) zwraca uwagę, w odniesieniu do uczenia się języka, iż w modelach koneksjonistycznych<sup>7</sup>, w przeciwieństwie do teorii Chomsky'ego, zwraca się uwagę na to, że nabywanie języka nie wymaga posiadania zdeterminowanej biologicznie wiedzy językowej lub typowych dla przetwarzania językowego struktur, lecz polega na interakcji pomiędzy mechanizmami uczenia i językowymi bodźcami.

Należy tu jednakże od razu zaznaczyć, że żadna z powstałych (jak dotąd!) sieci neuronowych nie była w stanie dorównać i konkurować z doskonałością mózgowych sieci neuronowych. Tak więc podzielam pogląd P. Szerszenia (2018: 63) w tej sprawie, że sztuczna inteligencja (a zatem także i uczenie maszynowe – przyp. A. Szuk) jest jedynie dość niedoskonałą próbą odzwierciedlenia i naśladowania genialnej ludzkiej inteligencji. Pisząc o niedoskonałości w kontekście sztucznej inteligencji mam na myśli jej wymiar jakościowy, bo o ile człowieka dzięki jego wyposażeniu można nazwać istotą w pełni inteligentną, o tyle sztuczna inteligencja jest na tyle „inteligentna”, na ile pozwala na to aktualny poziom rozwoju badań nad SI (por. M. Lister i in. 2009: 80).

W tym miejscu nasuwa się pytanie, jaki status wobec powyższego ma tzw. inteligencja lingwistyczna, i jak należy ją rozumieć, do czego odnosić. W pierwszej kolejności należy podkreślić, że z całą pewnością tzw. inteligencja lingwistyczna bardziej jest zbliżona do sztucznej, niż do ludzkiej – w pewnym sensie doskonałej – inteligencji. Jak pisze bowiem Z. Vetulani (2004) sztuczna inteligencja odnosi się do komputerowego modelowania takich przejawów aktywności myślowej człowieka, które wymagają aktywacji najgłębszych pokładów centralnego układu nerwowego. Do aktywności tych należy z pewnością również wszelka aktywność związana z modelowaniem kompetencji językowej. Kompetencja językowa to szczególnie rodzaj kompetencji komunikacyjnej, którą można rozumieć jako zbiór sprawności i umiejętności oraz wiedzy, dzięki którym człowiek jest w stanie produkować, jak i odbierać (interpretować) poprawne wypowiedzi (zdania czy pełne teksty), których nigdy dotąd nie wypowiadał i nie odbierał. Oznacza to zatem, że człowiek jest w stanie generować oraz odbierać nieskończoną ilość struktur za pomocą skończonej ilości środków językowych. To zatem sprawia, że człowiek jest istotą w pewnym sensie modelową – doskonałą, a co za tym idzie jego inteligencja jest na najwyższym, godnym naśladowania, poziomie.

---

<sup>7</sup> Istnieją oczywiście inne podejścia do uczenia się maszyn oprócz modelu koneksjonistycznego, m.in. model ewolucyjny, konektywistyczny czy bayesowski (od jego twórcy T. Bayesa). Nie będę ich jednak tu omawiała, gdyż nie jest to konieczne z punktu widzenia tematu niniejszego artykułu, a także ze względu na ograniczenia objętościowe nakładane na tego rodzaju teksty naukowe.

Próby rekonstrukcji modelu ludzkiej kompetencji językowej podejmowane były wielokrotnie, zazwyczaj z niezadowolającym skutkiem. Prób takich podejmował się m.in. N. Chomsky, opracowując swego rodzaju aparat matematyczny w postaci tzw. gramatyk generatywnych, które to zapoczątkowały z kolei teorię gramatyk formalnych (por. Z. Vetulani 2004: 22). W ten oto sposób prace Chomsky'ego umożliwiły stworzenie teoretycznych i metodologicznych ram do dalszych działań w zakresie tworzenia sztucznych, zautomatyzowanych systemów, imitujących ludzkie kompetencje językowe. To właśnie owe opracowania<sup>8</sup> dotyczące gramatyk generatywnych nadały kierunek poszukiwań i badań w zakresie rozwijania i ulepszania inteligentnych systemów, wyposażonych w sztuczną, aczkolwiek pseudo-inteligentną kompetencję komunikacyjną, tj. systemów, które potrafią odbierać polecenia człowieka, interpretować je, a także są wyposażone w funkcję produktywną i są w stanie wygenerować poprawną wypowiedź skierowaną do adresata ludzkiego. Opracowanie i stworzenie tego rodzaju systemów sprawiło, iż traktowane do tej pory jako absurdalne wizje o robocie antropoidalnym, potrafiącym komunikować się z człowiekiem w jego języku, stały się dość realistyczne.

Rozumiana w ten sposób lingwistyczna inteligencja, a więc zdolność sztucznego systemu do tworzenia i odbierania oraz interpretowania komunikatów jak najbardziej zbliżonych do tych, które stworzyłby człowiek, jest więc niewątpliwie częścią działań w ramach, omawianej w tej artykule, sztucznej inteligencji. Podsumowując powyższe rozważania można wysunąć wniosek, iż głównym zadaniem, czy inaczej założeniem wyróżniania i definiowania lingwistycznej inteligencji i tzw. sztucznej (a więc nie tej w odniesieniu do człowieka lecz do maszyny) kompetencji językowej, jest tworzenie systemów z szeroko rozumianą (bo obejmującą różne sprawności językowe począwszy od słuchania i interpretowania, po generowanie niepowtarzalnych i unikatowych wypowiedzi), emulowaną<sup>9</sup> kompetencją językową (por. Z. Vetulani 2004: 23).

Odwoływanie się do sztucznej inteligencji i uczenia maszynowego w kontekście lingwistycznym, to zatem w gruncie rzeczy odwoływanie się do sztucznej inteligencji lingwistycznej, która to z kolei obejmuje, najogólniej rzecz ujmując, wszystkie procesy związane ze sztucznym (maszynowym, komputerowym) modelowaniem kompetencji językowej (a więc wszystkich jej komponentów, tj. tzw. sprawności językowych, a więc słuchania, mówienia, wymawiania poszczególnych wyrazów etc.).

---

<sup>8</sup> Poza gramatyką generatywną Chomsky'ego ważną funkcję odegrała tu również zasada kompozycyjności języka opracowana przez G. Fregego opisana w książce pt.: *Über Sinn und Bedeutung* z 1892 roku. Zasada ta opiera się na przekonaniu, iż analizę semantyczną jednostki tekstowej można sprowadzić do analizy jej poszczególnych składników, gdyż znaczenie tekstu jest sumą znaczeń jej poszczególnych części oraz sposobu ich łączenia.

<sup>9</sup> Wyrażenie *emulowana kompetencja językowa* odnosi się do takiej kompetencji językowej, która powstała w wyniku naśladowania innej kompetencji w ramach np. innego systemu. Tu: mowa o tzw. sztucznej kompetencji językowej wzorowanej na idealnym modelu tej kompetencji, tj. ludzkiej kompetencji językowej.

### 3. Sztuczna inteligencja lingwistyczna – praktyczne zastosowanie

W dostępnej dziś literaturze lingwistycznej znaleźć można wiele wzmianek o kolejnych zastosowaniach nowych narzędzi stworzonych na fundamencie inteligencji lingwistycznej – tzw. inteligentnych narzędzi lingwistycznych. Chciałabym w tym artykule przedstawić kilka, moim zdaniem najbardziej interesujących propozycji, w których zastosowano inteligentne rozwiązania odnoszące się m.in. do udoskonalania różnych sprawności językowych lub do usprawniania różnego rodzaju prac lingwistycznych. Inteligencja narzędzi lingwistycznych często odnosi się do ich możliwości dostosowywania się do potrzeb poszczególnych użytkowników, bądź konkretnej sytuacji, czyli tzw. adaptacyjności (ang. *Adaptivity*; niem. *Adaption/ Adaptation*) (por. P. Szerszeń 2018: 64).

Inteligentne rozwiązania (narzędzia) lingwistyczne znalazły dość szerokie zastosowanie, przede wszystkim w dziedzinie glottodydaktyki. Systemy te mogą być nazywane po prostu systemami glottodydaktycznymi albo glottodydaktycznymi systemami adaptującymi, czyli dostosowującymi się do indywidualnych potrzeb ucznia. Tego typu system adaptujący D. Michie (1991: 56) definiuje jako system, który: „wykorzystuje zewnętrzne dane empiryczne w celu tworzenia i aktualizacji podstaw dla udoskonalonego działania na podobnych danych w przyszłości oraz wyrażania tych podstaw w zrozumiałej i symbolicznej postaci”.

Przykładem takiego systemu adaptującego się do potrzeb konkretnej osoby są systemy do nauki słownictwa, które w mniejszym lub większym stopniu wykorzystują możliwości dopasowywania się do konkretnej osoby i jej potrzeb językowych, a także glottodydaktycznych. Jako przykład jednego z tego typu inteligentnych narzędzi do nauki słownictwa, P. Szerszeń (2018) wskazuje system *iVocabulary*<sup>10</sup> Ch. Beera (por. ibidem: 68). System umożliwia uczącemu organizowanie niemalże całego procesu nauki nowego słownictwa, w tym terminów, poprzez swobodne gromadzenie, zarządzanie i grupowanie nowej leksyki oraz pozostałe funkcje, które mają ułatwić pracę z systemem. Ponadto *iVocabulary* daje możliwość synchronizowania różnych środowisk pracy osoby uczącej się i zarządzanie słownictwem z poziomu zarówno komputera, smartfona czy tableta, dzięki pracy w chmurze (*via iCloud*) lub korzystając z Dropboxa<sup>11</sup>. Inteligentny system *iVocabulary* umożliwia zarówno tworzenie własnych glosariuszy i słowniczków poprzez możliwość dowolnej organizacji słownictwa, np. według kategorii lub daty dodania, jak również zarządzanie leksyką już włączoną do systemu, np. przez innych użytkowników. Ciekawą funkcją dodatkową jest możliwość skorzystania z rozwiązań wypracowanych przez innych użytkowników systemu. Inteligencja systemu *iVocabulary* polega jednak przede wszystkim na udostępnianiu dopasowanych do potrzeb uczących się ćwiczeń i

---

<sup>10</sup> Więcej przykładów inteligentnych systemów do nauki słownictwa i zarządzania nim por. P. Szerszeń 2018: 68.

<sup>11</sup> Dropbox – usługa świadczona przez Dropbox Inc., polegająca na udostępnieniu przestrzeni dyskowej na serwerach tej firmy. Wysyłanie, przeglądanie i pobieranie danych jest możliwe poprzez zwykłą przeglądarkę lub poprzez dedykowaną aplikację zainstalowaną na komputerze.

tekstów leksykalnych oraz automatyczną ewaluację wyników wraz z informacją zwrotną i interpretacją popełnionych błędów. Dodatkowo system ma zintegrowaną funkcję rozpoznawania mowy, aby osoba ucząca się mogła dodatkowo ćwiczyć poprawną wymowę poszczególnych słów.

Kolejnym inteligentnym lingwistycznie narzędziem są inteligentne systemy tutorialne (IST). Są to systemy, które z reguły nie tylko inteligentnie reagują na podany *input* językowy, ale w dużej mierze są w stanie wygenerować także inteligentny lingwistycznie *output* w postaci wszelkiego rodzaju komunikatów zwrotnych przekazywanych osobie uczącej się przez tzw. pedagogicznych agentów (PA), którzy niejednokrotnie zwiualizowani są w systemie w formie np. postaci awatara (por. P. Szerszeń 2018: 69). Można powiedzieć, że tego typu inteligentnym systemom tutorialnym przypisuje się dość wysoki poziom lingwistycznej inteligencji, gdyż systemy te „opanowały” zdolność uczenia się poprzez gromadzenie, klasyfikowanie i analizowanie danych. Wiedza ta jest przez nie wykorzystywana, by m.in. trafnie reagować na zapytania czy komunikaty osób uczących się, a także, by inteligentnie dostosowywać się do potrzeb i poziomu konkretnej osoby uczącej się. Ponadto te bardziej zaawansowane systemy zostały wyposażone w funkcje rozpoznawania mowy, dzięki czemu oferta, chociażby ćwiczeń dydaktycznych, jest jeszcze szersza i jeszcze bardziej atrakcyjna. System tym samym ma więcej możliwości do zbierania danych o osobie uczącej się i może jeszcze efektywniej wykorzystywać tę wiedzę, m.in. po to, by jeszcze trafniej proponować uczącemu się pomoc w nauce lub, by móc zaoferować mu takie ćwiczenia, które będą najefektywniejsze w obszarach, w których uczeń wykazał do tej pory najwięcej deficytów. Przykładami, na które powołują się dość często badacze zajmujący się tematyką IST są dwa flagowe, bo zarazem najbardziej zaawansowane technologicznie narzędzia: *Tell Me More Campus* oraz *ALELO*. W obu przypadkach mamy do czynienia z dość zaawansowanymi technologiami, przyjaznym interfejsem użytkownika, bogatą ofertą ćwiczeniową, a także zdolnością do dopasowania systemu do potrzeb i poziomu indywidualnego użytkownika. Ponadto program *Tell Me More Campus* został dodatkowo wyposażony w funkcję rozpoznawania mowy w celu umożliwienia wirtualnej rozmowy (a przez to także ćwiczeń fonetycznych) użytkownika z inteligentnym agentem. Jak wykazuje w swojej pracy P. Szerszeń (2018) funkcjonalność ta pozostawia jednak wiele do życzenia, bo rozmowy mogą się odbywać na jedynie określone i dość okrojone tematy. Niemniej jednak, zastosowanie przez autorów tego rodzaju funkcjonalności należy uznać za swoiste *novum*, które warto jest starać się rozwijać i ulepszać, chociażby z uwagi na bardzo dużą atrakcyjność tego rodzaju ćwiczeń oraz możliwość wypracowywania wręcz spektakularnych efektów (glotto)dydaktycznych u uczących się (w porównaniu z tradycyjnymi metodami)<sup>12</sup>.

---

<sup>12</sup> Swoją tezę motywuję przede wszystkim niewątpliwie współwystępującą dużą motywacją uczących się w obliczu tego rodzaju ćwiczeń, które angażują nowoczesne technologie i zarazem są czymś nowym i nieznanym, wymagającym większego wysiłku intelektualnego. Uczniowie są bowiem nierzadko znudzeni tradycyjnymi formami (glotto)dydaktycznymi, a wszelkiego rodzaju nowinki technologiczne rozbudzają ich ciekawość i poprzez to chęć do nauki.

Kolejnym narzędziem, dość interesującym w świetle rozważań nad inteligentnymi narzędziami lingwistycznymi, jest e-learningowy system translo- i glottodydaktyczny LISTiG<sup>13</sup>/LISST<sup>14</sup>. W rozwój tego inteligentnego narzędzia lingwistycznego zaangażowane były na poszczególnych etapach różne podmioty, m.in. uniwersyteckie jednostki badawcze<sup>15</sup>, jak również pozauczelniane jednostki<sup>16</sup> o uznanej międzynarodowej renomie w branży IT oraz w zakresie rozwijania różnego rodzaju narzędzi lingwistycznych. Złożone narzędzie, jakim jest bez wątpienia system LISST/LISTiG, jest e-learningowym systemem translo- i glottodydaktycznym, który dzięki metodom informatyczno-lingwistycznym przekazuje uczącym się automatyczne komunikaty zwrotne w odpowiedzi na swobodnie wprowadzone przez nich teksty. Po wykonaniu ćwiczeń translacyjnych w systemie uczący się otrzymują automatycznie szczegółowe komunikaty zwrotne w odpowiedzi na wprowadzone przez siebie, przetłumaczone zdania. Program LISST/LISTiG (również dzięki swojej strukturze) jako pierwszy umożliwił w pełni zautomatyzowane rozpoznawanie rodzaju popełnionego błędu translacyjnego oraz wysyłanie na ten temat informacji zwrotnej do osoby uczącej się. Uczący się otrzymuje więc bezpośrednią i konkretną odpowiedź na temat wprowadzonej przez siebie propozycji tłumaczenia i może w ten sposób poprawić konkretny, wskazany przez system błąd. Interfejs prowadzących zajęcia jest nieco bardziej zróżnicowany i umożliwia nie tylko wprowadzenie do systemu wariantów tłumaczeniowych, lecz także zintegrowanie indywidualnych komunikatów o błędach z konkretnymi zjawiskami językowymi oraz samodzielne wprowadzanie opisu zadania. Dodatkowo system dokonuje automatycznej oceny wprowadzanych tekstów pod kątem gramatyki i ortografii. Na podstawie morfosyntaktycznych i semantycznych analiz system porównuje wprowadzone przez prowadzących zajęcia poprawne i błędne odpowiedzi z wariantami tłumaczeń wprowadzonymi przez osoby uczące się. Osoby, które wykonały ćwiczenia otrzymują automatyczne komunikaty zwrotne, w których wskazane są błędy w ich tłumaczeniach w stosunku do uprzednio wprowadzonych do systemu przez prowadzących zajęcia wzorcowych odpowiedzi. System porównuje wprowadzone przez uczących się poszczególne części zdania, jak i ich kolejność, z odpowiednimi informacjami, które wcześniej do systemu wprowadził prowadzący.

Ostatnim typem narzędzi, których tematykę chciałabym podjąć w niniejszym artykule są wielozadaniowe, inteligentne lingwistycznie narzędzia wspomagające

---

<sup>13</sup> LISTiG (PL) – Lingwistycznie Inteligentne Systemy Translo- i Glottodydaktyczne; LISST (DE) – Linguistisch Intelligente Softwaresysteme für die Sprach- und Translationsdidaktik.

<sup>14</sup> System powstał w oparciu o międzynarodową współpracę wiodących przedstawicieli środowisk akademickich i pozaakademickich z Polski i Niemiec (więcej informacji pod adresem: <http://www.lisst.de:8090>).

<sup>15</sup> Mowa tu m.in. o Uniwersytecie Warszawskim, Uniwersytecie J. Gutenberga w Moguncji oraz Uniwersytecie Kraju Saary (Niemcy). Więcej szczegółów znaleźć można na stronie [www.lisst.de:8090](http://www.lisst.de:8090).

<sup>16</sup> Mowa tu m.in. o takich jednostkach jak Institut der Gesellschaft zur Förderung der Angewandten Informationsforschung e.V. – IAI. Więcej szczegółów znaleźć można na stronie [www.lisst.de:8090](http://www.lisst.de:8090).



tłumaczenie. Jak słusznie konstatuje M. Nikishina (2018a: 172 i nast.) narzędzia te oferują nie tylko możliwość optymalizacji pracy tłumacza poprzez korzystanie z pamięci tłumaczeniowej, ale jednocześnie tłumaczenie maszynowe (oparte na sieciach neuronowych). Zdaniem badaczki takie hybrydowe rozwiązania (jednoczesne wykorzystanie TM – *Translation Memory* oraz MT – *Machine Learning*) cechują się większą trafnością i dokładnością, w porównaniu z zastosowaniem wyłącznie tłumaczenia maszynowego lub wyłącznie wsparcia w postaci pamięci tłumaczeniowych i poprzez ten fakt zyskują na popularności. M. Nikishina (2018b) wskazuje na fakt, iż, jak dowodzą przeprowadzone badania, w przypadku łączenia obu metod wsparcia procesu tłumaczenia (TM+MT) odnotowuje się znaczący przyrost poprawnych dopasowań translacyjnych (tzw. *perfect matches*). Jednym z przykładów takiego interaktywnego i lingwistycznie inteligentnego programu wspierającego proces translacji jest program *Lilt*. Program działa na bazie tzw. *Predictive Translation Memory*, opracowanej na Uniwersytecie w Stanford i służy *stricto* do edycji tłumaczenia maszynowego. Pamięć tłumaczeniowa wykorzystywana w tym oprogramowaniu służy nie tylko do przechowania przetłumaczonych segmentów, ale również do samouczenia się systemu; propozycje tłumacza zmieniają się w zależności od poprawek wprowadzonych przez tłumacza, w efekcie czego jakość tłumaczenia maszynowego się poprawia (por. *ibidem*: 175). Można więc stwierdzić, że również w tym zakresie (translacji) samouczenie się systemów staje się coraz popularniejsze i coraz bardziej efektywne.

Przytoczone przykłady oscylowały głównie wokół języka pisanego, chociaż przy okazji IST wspomniałam o inteligentnych systemach rozpoznawania mowy. W kwestii języka mówionego większość dostępnych narzędzi bazuje na koncepcji tzw. botów językowych (ang. *chatterbot*<sup>17</sup> lub *chatbot*). Najlepsze tego typu systemy są w stanie prowadzić rozmowę z żywym człowiekiem w czasie rzeczywistym i potrafią niejednokrotnie dorównywać już człowiekowi poziomem konwersacji. Do stosowanych przez nie metod do prowadzenia inteligentnej, ludzkiej konwersacji należy m.in. częsta zmiana tematu rozmowy czy udzielanie wymijających wypowiedzi oraz zarzucanie rozmówcy pytaniami w celu unikania udzielania odpowiedzi na zadawane im pytania. Inteligencja linguabotów (czy inaczej chatbotów) jest efektem opracowania tzw. bazy wiedzy dla tego typu systemów, która powstaje przy użyciu języka znaczników opartego na języku XML, a mianowicie AIML (*Artificial Intelligence Markup Language*).

#### **4. Terminotyka – czyli lingwistycznie inteligentne rozwiązania w terminologii i terminografii**

Możliwość zastosowania sztucznej inteligencji w rozwiązaniach terminologicznych i terminograficznych odkryto stosunkowo niedawno (w porównaniu chociażby z rozwojem innych, wspomnianych w poprzedniej części artykułu, inteligentnych lingwistycznie narzędzi). Fascynacja możliwościami, jakie otworzyły się w tym

---

<sup>17</sup> Nazwa *chatterbot* została po raz pierwszy użyta przez Michaela Mauldina, twórcę *Verbota* w 1994 roku. Źródło: <https://pl.wikipedia.org/wiki/Chatbot#AIML> [pobrano 20.10.2018].

zakresie była i jest nadal tak duża, że postanowiono nadać temu zjawisku osobną nazwę – terminotyka.

Wyrażenie *terminotyka* jest stosunkowo nowe w języku polskim i na próżno właściwie szukać go w słownikach. Biorąc pod uwagę czysto językowy aspekt *terminotyka* to zrost językowy powstały ze zrośnięcia dwóch wyrazów: *terminologia* i *informatyka*. W spektrum zainteresowań tej nowej dziedziny jest tworzenie i udoskonalanie zautomatyzowanych (częściowo lub w pełni), zaawansowanych technologicznie narzędzi, wykorzystywanych w pracach terminologicznych i terminograficznych.

Jak już wspominałam wcześniej za prekursora w rozwoju tej dziedziny uznaje się Kanadę, która boryka się z problemami terminologicznymi i terminograficznymi podwójnie, z racji dwóch oficjalnych języków urzędowych. W związku z tym kraj ten od dawna szukał rozwiązań, ułatwiających zarządzanie terminologią. Ja jednak w niniejszym artykule nie skupię się na działaniach kanadyjskich, lecz chciałabym przedstawić lingwistycznie inteligentne rozwiązania terminotyczne, które zostały opracowane w innym kraju, tj. w Niemczech. Będą to rzecz jasna tylko wybrane przykłady z zakresu terminologii i terminografii, jednak obrazujące ogólne tendencje rozwoju w tym zakresie.

Inteligentne narzędzia terminologiczne rozwijane są dziś głównie w celu automatycznej ekstrakcji terminów z coraz obszerniejszych korpusów tekstowych, zautomatyzowanej kontroli<sup>18</sup> nad coraz większymi zasobami terminologicznymi, wykorzystującej zaawansowane technologie przetwarzania języka oraz wreszcie, w celu tworzenia coraz bardziej zaawansowanych pamięci tłumaczeniowych (ang. *Translation Memory Systems*).

Punktem wyjścia do rozważań nad terminologią w kontekście inteligentnych rozwiązań terminotycznych jest stwierdzenie, jak ją „wyłowić” z tekstów, aby można było wykonywać na niej jakiegokolwiek działania. Logicznym jest, że im dany korpus tekstowy z danej dziedziny jest obszerniejszy, tym po pierwsze trudniej wydobyc z niego terminy (przeszukać go i wybrać odpowiednie jednostki wyrażeniowe), jednak z drugiej strony ten sam fakt (wielkość korpusu) wpływa na ilość zawartych w nim terminów, a tym samym można powiedzieć, że taki zbiór terminologiczny będzie dokładniejszy (w odniesieniu do konkretnej dziedziny), bardziej przekrojowy. Łatwiej będzie więc zdecydować, że pojawienie się tego terminu w tekście nie jest np. kwestią przypadku czy błędu tłumaczeniowego. Jeśli bowiem termin pojawia się często, świadczy to o tym, że po pierwsze nie pojawił się on tam przez przypadek, a po drugie jest relewantny dla danego obszaru specjalistycznego. Jak więc sprawić, by ekstrakowanie terminologii było prostsze, dokładniejsze i efektywniejsze?

A. Theofilidis (2018) zauważa, że kluczem do sukcesu w tworzeniu niezawodnego zasobu terminologicznego jest nie tylko poprawne wykrycie (wyekstrahowanie) nowych terminów z korpusu, lecz również zweryfikowanie pod kątem aktualno-

---

<sup>18</sup> Pisząc „kontrola” mam tu na myśli (za)planowane zarządzanie terminologią (ang. *terminology management*).

ści tej części terminologii, która została włączona do zasobu wcześniej: „Dabei geht es nicht allein darum, die spezifisch neuen Terme zu identifizieren. Es gilt auch festzustellen, welcher Anteil, der für den erweiterten Fachbereich bereits etablierten Terminologie in der neu aufkommenden Fachdomäne ebenfalls gültig ist” (A. Theofilidis 2018: 22). Ponadto, co jest godne podkreślenia i stosunkowo innowacyjne, autor zauważa, że oprócz wspomnianych kwestii, równie istotne, a może nawet bardziej relewantne, jest stwierdzenie istniejących lub mogących się pojawić zależności (relacji) między poszczególnymi jednostkami terminologicznymi. Zarówno między już uwzględnionymi w bazie terminami, jak i między nowymi, a także pomiędzy tymi płaszczyznami, a więc między zaakceptowanymi terminami i kandydatami (por. *ibidem*: 23).

Nowoczesne inteligentne lingwistycznie rozwiązania terminotyczne mogą dużo więcej niż człowiek (mam tu na myśli zarówno jednostkę, jak i grupę ludzi). Dzieje się tak, dlatego że komputer jest dziś po prostu bardziej wydajny i efektywny. Ponadto komputer nie jest tak podatny na popełnianie błędów jak człowiek. Ale co jeszcze, poza opisanymi polami działania potrafią maszyny, aby usprawnić pracę terminologów i terminografów? Przeszukując korpusy specjalnie zaprogramowane systemy są w stanie „wyłowić” z tekstu także tzw. *unerlaubte Terme*, a więc terminy niedozwolone. To oczywiście nie wszystko – specjalnie „przeszkolone”<sup>19</sup> programy wyróżniają w tekstach także wyrażenia, które wprawdzie terminami nie są, bo np. nie spełniają kryteriów formalnych, ale w języku potocznym pokrywają podobny zakres znaczeniowy, co sklasyfikowane i zaakceptowane już terminy. Kolejną inteligentną funkcją narzędzi terminotycznych jest umiejętność rozumienia znaczenia i przypisania go konkretnej dziedzinie specjalistycznej. To z kolei przekłada się na zdolność systemów do typowania tzw. *kandydatów terminologicznych*, a więc potencjalnych nowych terminów, o które można/ trzeba rozszerzyć bazę.

W świetle powyższego nasuwa się pytanie, na jakiej podstawie inteligentne narzędzia terminotyczne uczą się i zdobywają nową wiedzę. Z racji faktu, iż większość tego typu narzędzi powstaje w ośrodkach komercyjnych, metodologia rozwijania oprogramowania nie jest upubliczniana, co więcej, jest dość dobrze chroniona przed konkurencją. Niemniej jednak powstają, jak dotąd stosunkowo nieliczne pozycje naukowe, które skoncentrowane są wokół tej tematyki. Do osób, które w ostatnich latach zainteresowały się tą tematyką zaliczyć można takie osoby jak m.in.: U. Reuther i A. Zerfaß. Autorzy ci zgodnie wskazują, że u podstaw tak rozwiniętych rozwiązań terminotycznych musi zawsze leżeć solidnie przeprowadzona, głęboka analiza morfologiczna i gramatyczna. Podczas tych procesów terminy (i/lub kandydaci terminologiczni) są rozpoznawane pod kątem struktur gramatycznych oraz cech lingwistycznych, takich jak część mowy, etymologia, przynależność do klasy semantycznej etc. (por. U. Reuther 2017: 532 oraz A. Zerfaß 2016: 98).

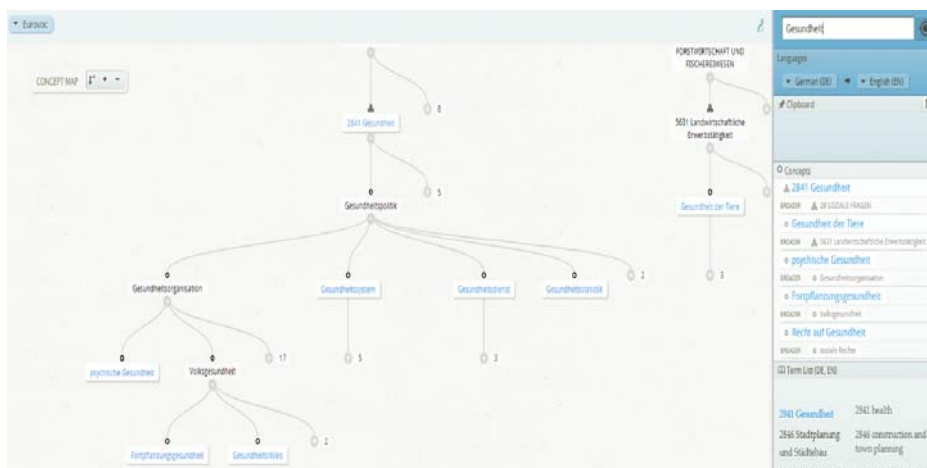
Wskazane pozycje odnoszą się do rozwiązań wypracowanych na potrzeby terminologii niemieckojęzycznej, a więc nie wszystkie propozycje w nich przedstawione mają odniesienie do innych języków. Przykładem może tu być np. rozklada-

---

<sup>19</sup> Zaprogramowane przez człowieka.

nie złożzeń (niem. *Komposita*) na czynniki pierwsze, aż do momentu, gdy wyróżnić będzie można wyraz podstawowy (niem. *Kopfwort*) oraz wyrazy go określające<sup>20</sup>. Rozłożenie wyrazu na części pierwsze umożliwia z kolei uwzględnienie w całym procesie także tych terminów, które np. zostały zapisane w inny sposób lub, gdy użyto synonimów względem wybranych części składowych całego terminu, np.: *Emissions-Reduktion*, *Emissionsreduzierung*, *Emissionsminderung*, *Reduktion der Emission*, *Minderung der Emission* (por. A. Theofilidis 2018: 28).

Kolejną, wartą dostrzeżenia funkcjonalnością jest automatyczna hierarchizacja terminów i ewentualnych kandydatów terminologicznych. Jest to opcja dość innowacyjna. Na podstawie m.in. rozpoznawania struktury oraz umiejętności przypisania przez oprogramowanie konkretnego znaczenia terminowi, możliwe jest odtworzenie tzw. siatki relacyjnej, zarówno wertykalnej, jak i horyzontalnej. W ten sposób możliwe jest zobrazowanie zależności między poszczególnymi terminami, które jak powszechnie wiadomo, nigdy nie występują jako pojedyncze, wyrwane z kontekstu jednostki, lecz jako części większych rodzin terminologicznych. W języku niemieckim wyekstrahowanie terminów nadrzędnych i podrzędnych jest o tyle proste, iż w przypadku złożzeń można tego dokonać poprzez odrzucanie kolejno części określających wyraz nadrzędny (a więc tych występujących po lewej stronie względem wyrazu głównego), aż dotrzemy do wyrazu najistotniejszego – określanego. Przykładem rzeczywistego zastosowania tej funkcjonalności jest narzędzie opracowane przez niemiecką firmę COREON, które obrazuje mapy koncepcyjne poszczególnych terminów. Widać tu wyraźnie relacje konkretnego wyrażenia do wyrażen sąsiednich w linii poziomej (sąsiedztwo horyzontalne), jak i zależność typu nadrzędny – podrzędny (sąsiedztwo wertykalne).



Rys. 1 COREON – mapa conceptualna niemieckiego terminu *Gesundheit*.

<sup>20</sup> Język niemiecki charakteryzuje się dużą ilością różnego rodzaju złożzeń. Główna zasada ich tworzenia mówi, iż wyrazy określające stawiane są zawsze w prepozycji względem wyrazu określanego. Tak więc główna część terminu znajduje się na końcu danego wyrażenia np. *Brennstoffzelle*.

Ponadto istotne jest, z punktu widzenia inteligentnych narzędzi, weryfikowanie częstotliwości (niem. *Vorkommenshäufigkeit*; ang. *Tf – term frequency*) występowania danego terminu lub kandydata, zarówno w danym tekście, w danym typie tekstów, jak i w całym korpusie. Pozwala to określić, jak ważny jest dany termin lub kandydat dla danej dziedziny. Im częstsze występowanie, tym większa „waga” (ważność, relewancja) terminu lub kandydata. Na podstawie m.in. tych informacji (oraz innych dodatkowych parametrów) można określić tzw. relewancję informacyjną terminu. Aby móc określić wartość relewancji informacyjnej danego terminu należy znać (oprócz częstotliwości występowania) również inny wskaźnik określany w literaturze skrótowo jako IDF (niem. *inverse Dokumentfrequenz*). Wskaźnik ten służy wykazaniu, iż termin występujący często, jednak w niewielkiej ilości tekstów (dokumentów) jest ważniejszy (informacyjnie), niż ten, który występuje rzadko, za to w dużej ilości dokumentów. Na podstawie iloczynu TF oraz IDF wyliczana jest tzw. waga terminu, która przekłada się bezpośrednio na jego relewancję informacyjną oraz ewentualną konieczność uwzględnienia w bazie terminologicznej. Innym, nieco mniej istotnym wskaźnikiem opcjonalnym, może być również pozycja terminu w tekście, a więc w której części zazwyczaj on występuje, czy da mu się przypisać stałą pozycję w tekstach specjalistycznych (np. wyłącznie wprowadzenie lub wyłącznie część opisująca wnioski etc.) (por. A. Theofilidis 2018: 30).

Automatyzacja prac terminologicznych i terminograficznych umożliwia dodatkowo sklasyfikowanie danego terminu lub kandydata do konkretnej klasy semantycznej, m.in. na podstawie umiejętności rozpoznawania dziedziny przez system na podstawie najczęściej występujących w danej dziedzinie klas. I tak np. w chemii najczęściej występującymi terminami są terminy sklasyfikowane ogólnie jako materiał (substancja, ciecz, gaz itp.).

Ostatnią nowością opracowaną przez niemieckie firmy zajmujące się opracowywaniem narzędzi do automatycznego przetwarzania języka, jest opracowywanie tzw. chmur terminologicznych (niem. *Termwolken*). Wyrażenia (nie tylko terminy) wykazujące m.in. duże podobieństwo semantyczne, przypisywane są tym samym chmurom. To z kolei pozwala na rozróżnienie przez system synonimów. W innej chmurze semantycznej wystąpi wirus wywołujący chorobę, a w zupełnie innej wirus powodujący zakłócenia pracy komputera. Natomiast nakładanie się na siebie chmur jest zjawiskiem normalnym i przewidzianym, wskazuje bowiem na coraz większą interdyscyplinarność nauki.

## 5. Wnioski

Nadrzędnym celem SI jest, ogólnie rzecz ujmując, symulacja ludzkiej inteligencji przez maszynę – komputer. Nierzadko spotkać się można ze stwierdzeniem, że rozwijanie kompetencji językowych jest najbardziej zaawansowanym przejawem i zarazem osiągnięciem intelektualnym człowieka. W związku z tym ewentualna możliwość rozwijania w maszynie ludzkich (a przynajmniej zbliżonych do ludzkich) kompetencji językowych jest wizją bardzo kuszącą i pożądaną, aczkolwiek od razu dostrzec należy poziom skomplikowania i trudności tego przedsięwzięcia. Należy mieć jednak na uwadze, że implementowanie szeroko pojętej SI w coraz to nowe

„zakamarki” lingwistyki sprawia, iż możliwa jest coraz sprawniejsza i efektywniejsza automatyzacja percepcji oraz generowania języka ludzkiego (naturalnego) przez komputer. To z kolei sprawia, iż NLP (*Natural Language Processing*) staje się podstawą wielu (a wkrótce może nawet wszystkich) sprawnie działających narzędzi lingwistycznych, począwszy od tzw. systemów glottodydaktycznych przez inteligentne, samouczące się oprogramowanie do wykonywania tłumaczeń, kończąc na narzędziach rozpoznających i zarządzających terminologią, nawet jej ogromnymi, wielojęzycznymi zasobami.

Niniejszy artykuł miał na celu zaprezentowanie wybranych możliwości, jakie daje zastosowanie SI i uczenia maszynowego w badaniach i rozwiązaniach lingwistycznych. Jak widać na podstawie mnogości zaprezentowanych przykładów, które są jedynie, jak podkreślam, wybranymi spośród wielu, ten trend przynosi duże korzyści i możliwości. Na podstawie zastosowanych w kilku poddziedzinach lingwistycznych rozwiązań (w glottodydaktyce translatoryce czy terminologii) widać, jakie perspektywy otwierają się przed lingwistyką wraz z zastosowaniem inteligentnych technologii. I chociaż, jak widać gołym okiem, zaprezentowane narzędzia te nie są (jeszcze!) idealne, to wszystko wskazuje na to, że ich ciągle udoskonalanie oraz tempo, w jakim poprawia się ich jakość, już wkrótce umożliwi wyeliminowanie wszelkich defektów i dostosowanie poszczególnych rozwiązań do potrzeb poszczególnych użytkowników lub grupy użytkowników.

Biorąc pod uwagę powyższe rozważania, a także szerzej poczynione obserwacje ludzi oraz przejawów ich inteligencji, oraz porównując je z możliwościami komputerów, jestem zmuszona przyznać, że jak dotąd komputery nie dorównały pod względem inteligencji człowiekowi. W świetle poczynionych rozważań wydaje się jednak, że to kwestia czasu, gdyż inteligentne narzędzia i rozwiązania są coraz doskonalsze. Mając na uwadze, z jaką szybkością uczą się dziś maszyny, jak szybko wysuwają wnioski oraz z jaką prędkością i zarazem dokładnością są w stanie opracować duże zasoby danych, a także jak szybki obserwujemy postęp w innych dziedzinach, w których od lat stosuje się inteligentne rozwiązania, należałoby założyć, że w nieodległej, jednak bliżej nieokreślonej przyszłości, komputery będą w stanie dorównać człowiekowi, również pod względem lingwistycznym, a więc na najbardziej wysublimowanej płaszczyźnie jego działalności kognitywnej. Co więcej tempo działań i dokładność komputerów pozwala na wyciągnięcie dalej idących wniosków, iż niedługo po tym, jak maszyny dorównają człowiekowi, mogą zostawić jego (ludzką) wydajność daleko w tyle.

W wielu krajach prowadzone są badania nad tzw. „naczelnym algorytmem”, który byłby na tyle samowystarczalny i inteligentny, iż człowiek w ogóle nie byłby mu potrzebny (wyluczając z tego fazę początkową, gdy człowiek musiałby go zaprogramować i ustawić w taki sposób, aby sam gromadził dane, przetwarzał je i wyciągał wnioski). Wynalezienie go byłoby niewątpliwie jednym z największych odkryć wszech czasów, ponieważ taki jeden algorytm byłby w stanie nauczyć się tylu różnych rzeczy, że byłby w zasadzie samowystarczalny. Badacze zgodnie twierdzą, że istnieje wiele dowodów na jego istnienie, lecz jak dotąd żaden człowiek nie był na tyle inteligentny, by go wyprowadzić. W obliczu takich hipotez, które

zresztą zdają się nie być wcale tak irracjonalne, jak jeszcze kilka lat temu, można założyć że opracowanie dużo prostszego algorytmu, który skupiałby się wyłącznie na kwestiach lingwistycznych i ułatwiałby pracę lingwistów, umożliwiając np. szybkie i proste tworzenie i aktualizację baz danych terminologicznych<sup>21</sup>, to założenia wysoce prawdopodobne. Tak więc można założyć, że w najbliższych latach będą powstawały nowe/ ewentualnie będą udoskonalane już istniejące narzędzia do automatyzacji wielu procesów związanych z percepcją i generowaniem wypowiedzi językowych, które będą stworzone nie przez żywą osobę, lecz przez uczącą się maszynę.

## Bibliografia

- Frege, G. (1892), *Über Sinn und Bedeutung*, (w:) *Zeitschrift für Philosophie und philosophische Kritik*. Konstanz.
- Domingos, P. (2015), *Naczelny algorytm – jak jego odkrycie zmieni nasz świat*. Gliwice.
- Lister, M./ J. Dover/ S. Giddings et al. (2009), *Nowe media. Wprowadzenie*, tłum. M. Lorek/ A. Sadza/ K. Sawicka. Kraków.
- Michie, D. (1991), *Machine Intelligence*. Londyn.
- Nikishina, M. (2018a), *Programy wspomagające tłumaczenie „w chmurze” – co i jak w trawie piszczy?*, (w:) „Applied Linguistics Papers” 25/1, 171–179.
- Nikishina, M. (2018b), *Tłumaczenie w programie CAT z podglądem tłumacza: chwilowa fascynacja czy przyszłość tłumaczeń pisemnych?*, (w:) „Applied Linguistics Papers” 25/4 (w druku).
- Reuther, U. (2017), *Darf’s ein bisschen weniger sein? – Termextraktion mit linguistischen Verfahren*, *tekom-Jahrestagung/tcworld conference 2017 – Tagungsband*, 531–543.
- Simon, H. (1983), *Machine Learning. An Artificial Intelligence Approach*. Trossingen.
- Szerszeń, P. (2014), *Platformy (glotto)dydaktyczne. Ich implementacja w uczeniu specjalistycznych języków obcych*, (Studi@ Naukowe 15). Warszawa.
- Szerszeń, P. (2018), *Na czym polega „lingwistyczna inteligencja” współczesnych systemów glottodydaktycznych? Adaptacyjność w skali mikro- i makro w obecnych rozwiązaniach e-learningowych*, (w:) „Applied Linguistics Papers” 25/1, 61–72.
- Sztuk, A. (2018a), *Nowoczesne metody aktualizacji baz danych terminologicznych. Statystyka za cenę prywatności?*, (w:) „Komunikacja specjalistyczna” 14/2018 (w druku).

---

<sup>21</sup> Wskazanie na bazy terminologiczne w tym kontekście nie jest przypadkowe. Jak pokazały przeprowadzone przeze mnie badania (por. A. Sztuk 2018b), dotyczące jakości infrastruktury terminologicznej, jakość wielu baz danych terminologicznych pozostawia wiele do życzenia i jest niewątpliwie jednym z najsłabszych punktów dzisiejszej infrastruktury terminologicznej.

- Sztuk, A. (2018b), *Bazy danych terminologicznych – aktualne tendencje w zarządzaniu terminologią*. Orbis Linguarum (w druku).
- Theofilidis, A. (2018), *Methoden der inhaltlichen Erschließung neuer Fachdomänen auf der Grundlage von Terextraktionsverfahren*. DDT-Symposion 2018 (w druku).
- Tryuk, M. (2017), *Frankońska szkoła terminologii*, (w:) S. Grucza/ M. Małachowicz (red.), *Polskie i europejskie nurty terminologiczne*, (Studi@ Naukowe 38). Warszawa, 121–140.
- Vetulani, Z. (2004), *Komunikacja człowieka z maszyną. Komputerowe modelowanie kompetencji językowej*. Warszawa.
- Zerfaß, A. (2016), *Terminologieextraktion*, (w:) J. Hennig/ M. Tjarks-Sobhani (red.), *Terminologiewerk für technische Dokumentation, tekom Schriften zur Technischen Kommunikation, Band 21*, 93–111.

### **Źródła internetowe**

- <http://apps.chbeer.de/ivocabulary/>. [Pobrano 12.10.2018].
- <https://www.tellmemorecampus.com/>. [Pobrano 10.10.2018].
- <https://www.tellmemorecampus.com/>. [Pobrano 4.10.2018].
- <https://lilt.com/>. [Pobrano 11.10.2018].
- <https://lilt.com/>. [Pobrano 7.10.2018].
- <https://pl.wikipedia.org/wiki/Chatbot#AIML>. [Pobrano 20.10.2018].