

Zaawansowane procedury NLP jako przesłanka rekonstrukcji idei wiedzy

Rafał Maciąg  <https://orcid.org/0000-0003-1812-3538>

Uniwersytet Jagielloński
e-mail: rafal.maciag@uj.edu.pl

ORYGINALNY ARTYKUŁ NAUKOWY

Źródła finansowania publikacji: Artykuł powstał w wyniku realizacji projektu badawczego o numerze 2018/29/B/HS1/01882, finansowanego ze środków Narodowego Centrum Nauki

Polityka open access: artykuł dostępny na podstawie licencji: CC BY 4.0

Informacja o konflikcie interesów: brak konfliktu interesów

Sugerowane cytowanie artykułu: Maciąg Rafał (2022). Zaawansowane procedury NLP jako przesłanka rekonstrukcji idei wiedzy. *Zarządzenie w Kulturze*, 23(1), 37–53.

Abstract

Advanced NLP Procedures as Premises for the Reconstruction of the Idea of Knowledge

The article presents the current state of development of the Natural Language Processing (NLP) technology, in particular the GPT-3 language model, and presents its consequences for understanding the phenomenon of knowledge. The NLP technology has been experiencing remarkable development recently. The GPT-3 language model presents a level of advancement that allows it to generate texts as answers to general questions, as summaries of the presented text, etc., which reach the level surpassing the analogous level of human texts. These algorithmic operations lead to the determination of the probability distribution of its components. Texts generated by such a model should be considered as autonomous texts, using immanent, implicit knowledge embedded in language. This conclusion raises questions about the status of such knowledge. Help in the analysis is provided also by the theory of discourse, as well as the theory of discursive space based on it, that proposes the interpretation of knowledge as a trajectory of discourses in a dynamical space. Recognizing that knowledge may also be autonomous, and in particular not be at the exclusive disposal of humans, leads to the question of the status of artificial cognitive agents, such as the GPT-3 language model.

Keywords: knowledge, Natural Language Processing, GPT-3, discourse, hermeneutics

Wprowadzenie

Przetwarzanie języka naturalnego (*Natural Language Processing*, NLP) należy do zespołu technologii określanych jako sztuczna inteligencja (AI). Jest to dziedzina bardzo obszerna, która w ostatnich latach przeżywa bezprecedensowy rozwój także w zakresie zastosowań praktycznych. W niniejszym tekście zostaną przedstawione przesłanki, które pozwalają zinterpretować obecne, najbardziej rozwinięte algorytmy NLP, w tym przede wszystkim model językowy GPT-3, jako istotne zmienne w interpretacji funkcji tekstu oraz, szerzej, problemu wiedzy. Ta ostatnia przechodzi istotną ewolucję rozumienia w wieku XX, tracąc swój czysto podmiotowy, tj. związany wyłącznie z człowiekiem, charakter, a także przestaje być traktowana jedynie jako przedmiot filozofii (epistemologii). Jest natomiast interpretowana jako istotny zasób oraz czynnik zmiany społecznej. Tę problematykę omawiam szerzej gdzie indziej (Maciąg 2022). Także rozwój technologii cyfrowej stanowi proces rekonfiguracji idei wiedzy, pozostając w ścisłym związku z jej nowymi interpretacjami.

Przetwarzanie języka naturalnego jest jednym z obszarów tej technologii o szczególnej dynamice rozwoju. Według ostatniego, bardzo obszernego i dobrze udokumentowanego, raportu na temat sztucznej inteligencji, sygnowanego przez Uniwersytet Stanforda, jedno z najnowszych rozwiązań w tym obszarze, jakim jest model językowy GPT-3, którym zajmujemy się tutaj bliżej, należy do najważniejszych przełomów technicznych w obrębie sztucznej inteligencji w okresie objętym raportem, tj. w roku 2020 (Zhang et al. 2021: 79). Model ten cechuje się także istotnym wpływem społecznym (Tamkin et al. 2021). Rozwojowi technicznemu towarzyszy wzrost nakładów rynkowych na tę technologię, które w roku 2020 w USA wynosiły 16,53 miliarda \$, a prognoza na rok 2028 szacuje je na 127,26 miliarda \$, co oznacza średni wzrost rok do roku o blisko 30% (CAGR). Rynek amerykański jest istotny, ponieważ jego udział w rynku światowym tej technologii szacuje się na 43% (Fortune Business Insights 2021).

Imponujące są także cytowane w raporcie stanfordzkim osiągnięcia nowych technologii NLP. Raport powołuje się na dwa badania porównawcze skuteczności tych technologii w porównaniu z analogiczną skutecznością człowieka w wykonywaniu pewnych zadań, związanych z używaniem języka do konkretnych celów. Badanie pierwsze, o nazwie SuperGLUE, jest testem oceniającym wykonanie szeregu zadań związanych ze zrozumieniem tekstu. Tych zadań jest osiem i stanowią zbiór składający się głównie z wcześniej opublikowanych propozycji. Znajdują się w nim głównie zadania polegające na udzielaniu odpowiedzi na rozmaicie formułowane w kontekście zadanego tekstu pytania, wymagające jednak pewnej wiedzy uprzedniej i procedury wnioskowania (Wang et al. 2018). W tym teście oceniano model języka pn. DeBERT, dostarczony przez Microsoft, o mniej zaawansowanej konstrukcji niż GPT-3. Osiągnął on uśrednioną jakość oszacowaną na 90,3 punktu, która przewyższyła wynik osiągnięty przez człowieka dla analogicznych zadań

oszacowany na 89,8 punktu. Badanie drugie opierało się na teście pn. Stanford Question Answering Dataset (SQuAD). Polega on na udzielaniu krótkich odpowiedzi na pytania dotyczące testowego tekstu. Wersja pierwsza testu zawierała 100 tysięcy pytań dotyczących Wikipedii (Rajpurkar et al. 2016), druga dodawała jeszcze 50 tysięcy pytań, nie posiadających odpowiedzi w ogóle, co badany system miał także umieć rozpoznać (Rajpurkar et al. 2018). Również tutaj algorytm okazał się skuteczniejszy od człowieka w dwu kolejnych badaniach, osiągając poziom jakości szacowany na 95,4 oraz 93 punkty (wyniki człowieka to odpowiednio 91,2 i 89,5 punktu) (Zhang et al. 2021: 62–63).

Zarówno wyniki uzyskane w opisanych badaniach, jak i bezpośrednie obcowanie z tekstem powstałym wskutek interakcji z maszyną są spektakularne i zdecydowanie wypełniają warunki tzw. testu Turinga, który został wymyślony do oszacowania pojawienia się sztucznej inteligencji dorównującej ludzkiej (Turing 1950). Podobną opinię wygłosił ostatnio Ilya Sutskever, będący założycielem i szefem organizacji OpenAI, w której laboratoriach powstał GPT-3¹. Choć podobnie daleko idący wniosek wydaje się przedwczesny, pojawia się jednak problem inteligibilnego tekstu będącego tworem sztucznego systemu kognitywnego. Trzeba też dodać, że badania opisane w cytowanym raporcie poprzedziły publikację najbardziej obecnie zaawansowanego algorytmu NLP, modelu językowego GPT-3 (Brown et al. 2020), który jest przedmiotem opisu w tym tekście. Model ten jest uważany za najbardziej zaawansowany, choć jego parametry techniczne zostały przekroczone już w innych modelach, choć opartych na podobnych rozwiązaniach. Do tych modeli należą: Megatron-Turing NLG 530B, stworzony przez Nvidię i Microsoft, opublikowany 11 października 2021 roku, Wu Dao 2.0, stworzony przez Beijing Academy of Artificial Intelligence, opublikowany 31 maja 2021 roku, oraz M6 – Multi-Modality to Multi-Modality Multitask Mega-transformer, stworzony przez Alibaba DAMO Academy, opublikowany 8–10 października 2021 roku. GPT-3 zachowuje jednak pozycję rozwiązania najbardziej zaawansowanego, poddawanego jedynie udoskonaleniom. OpenAI zaproponowało także własne rozwiązania tego rodzaju, dotyczące różnych aspektów jego funkcjonowania, takich jak poprawiony mechanizm tzw. osadzania słów (*embedding*) (Neelakantan et al. 2022) lub rekonstrukcja algorytmu pod kątem jego łatwiejszej i bardziej zbliżonej do umiejętności użytkowników obsługi, nazwana InstructGPT (Ouyang et al. 2022).

Badania z zakresu NLP mają długą historię. Li Deng i Yang Liu uznają za ich inaugurację artykuł Alana Turinga z roku 1950, w którym proponuje on sławny test, nazwany od jego imienia testem Turinga (Deng, Liu 2018). Ten długi okres rozwoju dzielą oni na trzy fale, z których najważniejsza z naszego punktu widzenia jest trwająca obecnie fala trzecia, która rozpoczęła się około roku 2010 i jest związana z tzw. głębokim uczeniem (*deep learning*), czyli wykorzystaniem sztucznych

¹ <https://twitter.com/ilyasut/status/1491554478243258368> [odczyt: 13.02.2022].

sieci neuronowych. W ramach tej fali doszło do spektakularnego rozwoju technologii NLP, którego jak na razie najbardziej zaawansowanym rozwiązaniem (tzw. *the state-of-art*) jest model językowy pn. GPT-3 (Generative Pre-trained Transformer), oparty w przeważającej większości na poprzednim modelu, pn. GPT-2, i różniący się od poprzednika przede wszystkim wielkością danych wykorzystanych do wytrenowania tego modelu (Brown et al. 2020; Radford et al. 2019).

Model językowy to według definicji podanej przez Dana Jurafsky'ego i Jamesa H. Martina model, „który przypisuje prawdopodobieństwa sekwencjom słów” (Jurafsky, Martin 2020: 30). Charu C. Aggarwal tłumaczy prościej, iż „model językowy służy do stworzenia probabilistycznej reprezentacji tekstu” (Aggarwal 2018: 4). Eugene Charniak jeszcze ogólniej pisze, że „model języka to rozkład prawdopodobieństwa wszystkich łańcuchów językowych” (Charniak 2019: 71). Podobną definicję podaje Ian Goodfellow et al.: „Model językowy definiuje rozkład prawdopodobieństwa sekwencji tokenów w języku naturalnym” (Goodfellow et al. 2016: 461). Tokeny można opisać jako wypreparowane, sensowne części słów. W technologii NLP model językowy jest strukturą, która powstaje jako skutek analizy zgromadzonych w tym celu tekstów, nazywanych korpusem językowym. Dzięki tej analizie każdemu słowu (tokenowi) zostaje przypisany zestaw liczb, definiujących jego położenie w wielowymiarowej przestrzeni. Tych wymiarów jest zwykle 512 lub 1024, ale w przypadku GPT-3 także znacznie więcej. Położenie to ma znaczenie semantyczne w tym sensie, że odzwierciedla zależności semantyczne między wyrazami (tokenami), a co więcej – można wykonywać działania matematyczne, które zachowują te zależności. Taka przestrzeń powstaje dzięki technologii zwanej osadzeniem (*embedding*), zaproponowanej przez Tomáša Mikolova et al. (Mikolov et al. 2013). W modelu GPT-3 podstawowa technika osadzania jest uzupełniona przez procedury z obszaru tzw. rekurencyjnych sieci neuronowych, w szczególności ich wariant pn. LSTM (*Long Short-Term Memory*) (Hochreiter, Schmidhuber 1997) oraz tzw. procedurę uwagi (*attention*) (Vaswani et al. 2017).

Jednak najważniejszy z punktu widzenia niniejszego tekstu jest podstawowy korpus językowy, na którym opiera się cały model i służy do jego wytrenowania. Jest to zbiór tekstów, o którym autorzy modelu piszą, iż ma prezentować „tak duży i zróżnicowany zbiór danych, jak to możliwe, aby zebrać reprezentację zadań realizowanych przez język naturalny w możliwie różnych dziedzinach i kontekstach” (Radford et al. 2019). W wypadku obydwu modeli te zbiory pochodzą z Internetu, w szczególności są wynikiem pracy tzw. pająka sieciowego (*web crawler*), czyli algorytmu przeszukiwania sieci pn. Common Crawl. Można powiedzieć, że w ten sposób zostaje odnowiona wczesna idea Pierre'a Lévy'ego pn. inteligencji zbiorowej (*collective intelligence*) (Lévy 1999), która jest rezultatem połączenia wiedzy i aktywności pojedynczych ludzi za pomocą sieci.

The Common Crawl Foundation jest zarejestrowaną w Kalifornii inicjatywą internetową non profit, której celem jest „zbudowanie i utrzymanie otwartego

repozytorium danych powstających dzięki indeksowaniu sieci, do których każdy może uzyskać dostęp i je analizować” (Common Crawl Foundation 2021). Gromadzi ona co miesiąc całość tekstów obecnych w Internecie od roku 2013. Zbiory danych, które je zawierają, mają ogromną objętość. Na przykład komplet plików z października 2021 roku ma objętość ok. 116 tebibajtów (TiB, *tebabyte*; 1 tebibajt równa się 1 099 511 627 776 bajtom). Te liczby pozwalają właściwie ocenić ilość współcześnie dostępnych danych pochodzących z Internetu i pozostających do dyspozycji różnych projektów, w tym modeli językowych.

Korpus będący podstawą GPT-3 jest także oparty na zbiorze tekstów z Internetu, dostarczonych w większości przez Common Crawl, przy czym ten zasób, aby „poprawić średnią jakość”, został przefiltrowany i poddany tzw. rozmytej deduplikacji (*fuzzily deduplication*), tzn. „pozbawiony dokumentów w wysokim stopniu nakładających się na inne dokumenty”. Filtrowanie odbyło się za pomocą klasyfikatora wytrenowanego wcześniej na zbiorach dokumentów, takich jak WebText, Wikipedia, oraz kolekcjach książek dostępnych w sieci. Uzyskany zbiór tekstów został uzupełniony przez „kilka wyselekcjonowanych zbiorów danych wysokiej jakości”, takich jak rozszerzona wersja WebText, dwie dalsze kolekcje książek z sieci i anglojęzyczna Wikipedia. Objętość nieprzefiltrowanego zbioru Common Crawl z lat 2016–2019 wynosiła 45 TB, a przefiltrowanego – 570 GB, co stanowi ok. 78% ostatecznego zbioru tokenów przeznaczonych do trenowania. Pozostała część przypada na resztę zbiorów. Jednocześnie, w czasie trenowania modelu, zbiory tekstów nie były traktowane jednakowo. Zbiór powstały na podstawie Common Crawl oraz jednego ze zbiorów książek był używany w czasie trenowania istotnie rzadziej. Najczęściej używana była Wikipedia. Było to wynikiem faktu, iż autorzy algorytmu „postrzegali pewne zbiory danych jako wyższej jakości” (Brown et al. 2020).

Efektom działania modelu GPT-3 są teksty skonstruowane poprawnie pod względem składniowym i semantycznym, a także wnoszące swoisty wkład poznawczy, tj. oparte na pewnej wiedzy, która ma charakter immanentny, wewnętrzny. Ta ostatnia kwestia pozostaje także w zgodzie z założeniami teoretycznymi dotyczącymi tzw. rozproszonych modeli konekcyjnych sztucznej inteligencji (*distributed connectionist models of artificial intelligence*), tj. sztucznych sieci neuronowych (Flasiński 2016). Powstaje jednak pytanie o źródło tej wiedzy.

Istnieje szereg testów, jakim poddaje się rozwiązania informatyczne w obszarze sztucznych sieci neuronowych, zastosowanych także w przypadku modelu GPT-3 do takich zadań, jak modelowanie języka (*language modeling*), kwestie domykania i dokończania zdań (*cloze and completion*), tłumaczenia (*translation*), zdroworozsądkowego rozumowania (*common sense reasoning*), czytania ze zrozumieniem (*reading comprehension*) itd. Część z nich zawierają opisane wcześniej procedury z raportu stanfordzkiego. Nie istnieje jednak żaden test pozwalający potwierdzić

równoważność tekstów sztucznych i ludzkich lub zaprzeczyć tej równoważności, a w szczególności potwierdzić ich inteligibilność lub jej zaprzeczyć na podstawie jawnych i enumeratywnych kryteriów. Jediną taką procedurą, która raczej potwierdza fakt tego braku, jest tzw. test Turinga (Turing 1950). W wypadku inteligibilności kluczowe znaczenie może mieć nie tyle tzw. inteligencja, ile raczej wiedza, choć trzeba zaznaczyć, że obydwie kategorie nie posiadają precyzyjnych denotatów, lecz raczej spełniają definicję Ludwiga Wittgensteina, tj. są rezultatem swoistej i ciągłej gry interpretacyjnej (Wittgenstein 1958). Nie ulega jednak wątpliwości, że wiedza jako przedmiot refleksji zyskuje nową płaszczyznę badawczą w wypadku tekstów generowanych sztucznie. Tego rodzaju badania są już prowadzone, ale mają z reguły techniczny charakter, a pojęcie wiedzy ma ograniczony charakter (Bender, Koller 2020; Petroni et al. 2019; Porada et al. 2021; Safavi, Koutra 2021; Zhou et al. 2020).

Zgodnie z przyjętym tutaj założeniem inteligibilny tekst jest fenomenem autonomicznym, a tym samym jest równy tekstom wytworzonym przez człowieka. Tego rodzaju podejście zostało ukształtowane w naukach humanistycznych, a ściślej na polu tzw. hermeneutyki. Opiera się ono na założeniu, że tekst samodzielnie niesie pewną złożoną wartość semantyczną. Dyskusja na ten temat toczy się w kulturze europejskiej co najmniej od Platona, który zajął się nim w *Fajdrosie*, jednak hermeneutyka przeżyła niezwykle rozkwit w wieku XX. Opierając się na jej rozumowaniu, można przyjąć, że inteligibilny tekst – w tym przypadku sztuczny – jest wydarzeniem autonomicznym w znaczeniu posiadania własnej, immanentnej wartości semantycznej, a także szerzej – pewnego sensu. Ponieważ jest on produktem pewnego autonomicznego procesu, należy zadać pytanie, jaki jest status tego procesu i jak się on lokuje wobec analogicznego procesu tworzenia tekstu przez człowieka.

Pewnej pomocy może udzielić w tym względzie fakt, iż tekst jest jednym z najstarszych obiektów refleksji filologicznej i filozoficznej. Można bez przesady powiedzieć, że jest fundamentem cywilizacji zachodniej, tak samo zresztą jak i wielu innych. Jednocześnie tę rolę należy rozpatrywać w kontekście innych językowych materializacji, takich jak mowa (ekspresja oralna) czy druk. Każda z nich wprowadza własną specyfikę i posiada własną literaturę przedmiotu. Należy przypomnieć, że choć w wypadku modeli GPT pojawia się jedynie tekst zapisany w postaci elektronicznej, to jednak nie powinno to uniemożliwiać jego analizy analogicznej do wytworów ludzkich.

Tego rodzaju analiza zwykle jest łączona z tzw. postępowaniem hermeneutycznym. Stanley E. Porter i Jason C. Robinson przywołują ponadto legendarne źródło hermeneutyki i podkreślają relację „pomiędzy” (*in-between*), odwołującą się do postaci Hermesa, który pośredniczył między bogami a ludźmi (Porter, Robinson 2011). Hermeneutyka może zostać zestawiona z filozofią jako poszukiwanie i wytworzenie sensu za pomocą interpretacji. W tym zestawieniu tekst pojawia się jako ekwiwalent świata, a procedury analityczne mogą być odpowiednio rozszerzone (Malpas, Gander 2015). Jak pisze Josef Bleicher, „hermeneutykę można luźno

zdefiniować jako teorię lub filozofię interpretacji znaczenia” (Bleicher 1980: 1), po czym dodaje, że „hermeneutyka jest konsekwentnie zaangażowana w dwa zadania: pierwsze to ustalenie dokładnego znaczenia, treści słowa, zdania, tekstu itp.; i drugie – odkrycie instrukcji zawartych w formach symbolicznych” (Bleicher 1980: 11). Aczkolwiek hermeneutyka pojawia się co najmniej od czasów antycznych, jej samoświadoma (filozoficzna) wersja rozwija się intensywnie od połowy XIX wieku, a wśród jej licznych przedstawicieli znajdują się tacy filozofowie jak Fryderyk Nietzsche czy później Martin Heidegger, Hans-Georg Gadamer, a także przedstawiciele tzw. poststrukturalizmu: Michel Foucault, Jacques Derrida, Gilles Deleuze, Richard Rorty. Była ona także przedmiotem namysłu innych szkół filozoficznych, takich jak angielska filozofia analityczna (e.g. Bertrand Russell, Gilbert Ryle, John Langshaw Austin) czy tzw. koło wiedeńskie z jego najbardziej znanym przedstawicielem Ludwigiem Wittgensteinem. Wydaje się, że hermeneutyka jest prawdopodobnie jedną ze skutecznych strategii, jakie można zastosować w celu niebezpośredniej ewaluacji inteligibilności tekstów generowanych sztucznie.

Należy dodać, że umiejętność generowania inteligibilnych tekstów była uważana za wyłącznie ludzką kompetencję. Tego rodzaju założenie formułuje w traktacie pt. *Rozprawa o metodzie* Descartes, który zbudował europejskie filozoficzne podstawy myślenia o człowieku (Descartes 1637). W tym traktacie formułuje on także swoje najślawniejsze twierdzenie, fundujące owo myślenie: *cogito ergo sum*. Przedstawia w nim dwa argumenty, które dowodzą według niego całkowitej niemożliwości stworzenia automatu podobnego do człowieka. Podstawą pierwszego z nich jest zdolność człowieka do używania języka w sposób nieosiągalny dla maszyny, tzn. zgodnie z kontekstem znaczeniowym, a więc odnosząc język do pewnych praktycznych okoliczności wytwarzania się sensu: „niemożliwe jest, aby [maszyna – przyp. R.M.] składała w różny sposób słowa, aby odpowiadała z sensem na wszystko, co się powie w jej obecności” (Descartes 2013, 48). Choć ta okoliczność wymaga szerszej analizy, można powiedzieć, że teksty stworzone przez algorytmy NLP wydają się spełniać ten warunek.

Problem kompetencji aktualizuje pytanie dotyczące statusu ontologicznego autora, którym w wypadku tekstów generowanych sztucznie jest sztuczny system kognitywny. Ta konstatacja z kolei rodzi pytanie o status tego systemu wobec człowieka w zakresie tej kompetencji. Należy w tym miejscu podkreślić, że nie mamy do czynienia ze swobodną fantazją na ten temat, mającą charakter prognozy, która proponuje mniej lub bardziej dowolne interpretacje (e.g. Yuval Noah Harari, Nick Bostrom, Ray Kurzweil), ale ze zjawiskiem czysto doświadczalnym.

Uznanie tekstów sztucznych za równoprawne ludzkim powoduje, że istotna staje się kwestia kontekstu języka, jaki jest przyjmowany *expressis verbis* w różnych koncepcjach i modelach pojawiających się w obszarze NLP (e.g. Deng, Liu 2018; Goodfellow et al. 2016; Jurafsky, Martin 2020). Kwestia języka pojawia się w kontekście generalnie semantycznego i w tym sensie jakościowego nastawienia.

To nastawienie z konieczności aktywizuje także wyższy niż sama instancja tekstu poziom opisu, tj. poziom lingwistyczny, a także filozoficzny. Tego rodzaju podejście posiada uzasadnienie w obszernej, wspomnianej już refleksji hermeneutycznej (por. Foucault, Derrida, Austin itd.). Bardzo bogaty, przeprowadzony z punktu widzenia technologii i jednocześnie lapidarny opis tej problematyki podaje John. F. Sowa (Sowa 2010).

Język jako fenomen pozwala także na skuteczną interpretację fenomenu wiedzy, co staje się możliwe przy użyciu takiego narzędzia jak dyskurs. Zostało ono wprowadzone w roku 1952 przez Zelliga Harrisa i od tamtego czasu jego zastosowanie rozwinęło się w bardzo wiele podejść, które „rozciągają się między analizą dyskursu zorientowaną bezpośrednio na tekst i koncentrującą się głównie na jego właściwościach lingwistycznych a podejściem zorientowanym bardziej społecznie, które uwzględnia funkcje tekstu w kontekście środowiska społecznego i kulturowego, w którym występuje” (Paltridge 2006: 1). Dyskurs jest uważany powszechnie za emanację wiedzy przez różnych badaczy: „teoria dyskursu często obraca się wokół związku władzy, wiedzy i podmiotowości” (Angermuller et al. 2014: 6); „dyskurs jest społecznie twórczy i społecznie ukształtowany: tworzy sytuacje, przedmioty wiedzy oraz społeczne tożsamości i relacje między ludźmi i grupami ludzi” (Hyland, Paltridge 2011: 39); „walkę między różnymi roszczeniami wiedzy można rozumieć i badać empirycznie jako walkę między różnymi dyskursami, które reprezentują różne sposoby rozumienia aspektów świata i konstruują różne tożsamości mówiących” (Jørgensen, Phillips 2002: 2). Jeden z najważniejszych badaczy dyskursu, Teun van Dijk, pisze: „rola kontekstu w tworzeniu i rozumieniu dyskursu jest fundamentalna. Ponieważ wiedza jest częścią kontekstu, każdy poziom struktury dyskursu zależy od wiedzy uczestników” (Dijk 2013: 592).

Dyskurs w kontekście wiedzy przyjmuje co najmniej dwa, różniące się skalą, podejścia. Pierwsze dotyczy pragmatycznego zastosowania dyskursu jako narzędzia rekonstrukcji wiedzy w sensie pragmatycznym, w tym także organizacyjnej, co oznacza wykorzystanie go w praktyce zarządzania: „teza «wiedzy jako działania społecznego» przywołuje rozumienie «działania społecznego» jako dyskursu (rozmowy i tekstu) pojawiającego się w procesie interakcji” (Crane 2016: 77). Z drugiej strony pojęcie dyskursu umożliwia także podejście o charakterze ogólnym, tj. filozoficznym, pozwalając opisać i zinterpretować procesy historyczne o wielkiej skali. Ten nurt reprezentują Jean-François Lyotard (Lyotard 1979) i Michel Foucault (Foucault 1966, 1969). Foucault jest także autorem, za którego refleksją podąża większość współczesnych badań dyskursu (Fairclough 2003: 2; Jørgensen, Phillips 2002: 13). Idea wiedzy magazynowanej w dyskursie odwołuje się do refleksji Nietzschego, Wittgensteina oraz Marksa i opiera się na istotnej roli języka w procesach komunikacyjnych. Prowadzi ona także do swoistej formy władzy, jaką zdobywa dyskurs w procesach społecznych (Foucault 1981). Koncepcja Foucaulta interpretuje dyskurs jako językową instancję wiedzy, będącej rezultatem kolektywnej, społecznej

i historycznej aktywności, dzięki której pojawia się swoisty i autonomiczny byt, nazywany przez Foucaulta właśnie dyskursem (Foucault 1966, 1969, 1971).

Zaproponowana przez autora teoria przestrzeni dyskursywnej opiera się na koncepcji Foucaulta, proponując rozumienie dyskursu jako fenomenu doświadczalnego i możliwego do interpretacji fizycznej. Ponieważ dyskurs jest rozumiany w tej teorii jako zjawisko retencji i artykulacji wiedzy, zgodnie z ideą wprowadzoną przez Foucaulta, przestrzeń dyskursywna jest więc także opisem istnienia i funkcjonowania wiedzy. Przestrzeń dyskursywna to n -wymiarowa przestrzeń dynamiczna, w której dyskursy, jako autonomiczne instancje wiedzy, przebiegają w czasie trajektorie opisujące rzeczywisty stan wiedzy w przedmiocie, którego dotyczą. Kategorie przestrzeni można ekstrapolować na rozmaitość, co pozwala wprowadzać liczne przestrzenie o swoistych wymiarach, stanowiące różne instancje wiedzy (Maciąg 2018, 2020). Korpus tekstów, na których opiera się tworzenie przestrzeni dyskursywnej, nie jest bynajmniej jednorodną i stabilną masą, lecz dynamiczną strukturą. Model przestrzeni dyskursywnej odsłania tę złożoną strukturę, która wydaje się włóknista (w czasie), gdzie poszczególne włókna powstają z indywidualnych trajektorii dyskursów.

Dyskurs nie jest prostą sumą wypowiedzi, ale rezultatem ich dynamicznych, współzależnych interakcji, zależnych także od lokalnych uwarunkowań społecznych, kulturowych, etnicznych itp. Wypowiedzi te mogą przyjmować różną formę materialną (symboliczną), choć najłatwiej dostępna jest forma tekstowa. Wspomniane interakcje w momencie artykulacji w postaci wypowiedzi ujawniają się jako chwilowo stabilne relacje różnego rzędu i rodzaju, w których jest zawarta wiedza na temat treści tej wypowiedzi, przy czym owa chwilowość może mieć bardzo różny czas trwania. Przestrzeń dyskursywna pozwala zilustrować wspomnianą dynamikę interakcji. Jest to istotne, ponieważ pozwala dostrzec brak tej właściwości w wypadku modeli GPT, które korzystają ze względnie stabilnych korpusów o maksymalnie dużej objętości.

Spostrzeżenia

Przy braku jakichkolwiek przekonywających argumentów wykluczających osiągnięcie wystarczającego poziomu jakości poznawczej oraz poprawności językowej sztucznych tekstów – te, które są tworzone przez model GPT-3 należy uznać za teksty pełnoprawne. Do takiego wniosku prowadzi ich analiza hermeneutyczna, która stawia je na równi z tekstami tworzonymi przez człowieka.

Sztuczne teksty otwierają nową perspektywę w dyskusji na temat wiedzy. Jeżeli założymy, że teksty te pokazują dostęp do ich immanentnej (wewnętrznej, implicytnej) wiedzy, pytanie o rodzaj i źródło tej wiedzy staje się uprawnione. Ponieważ jedynym zasobem semantycznym, na którym opierają się algorytmy tworzące te

teksty, jest język, reprezentowany przez zbiór realnych, spontanicznych tekstów, należy uznać, że jest to jedyne źródło wiedzy dla tych algorytmów.

Ta konkluzja oznacza wsparcie dla idei wiedzy, dla której kontenerem jest język. Tego typu koncepcja jest już obecna i jej rozwinięty wariant reprezentuje koncepcja dyskursu, który pełni funkcję kontenera wiedzy. Kwestia umieszczenia zjawiska wiedzy na poziomie języka jest osobnym problemem, bo w rzeczywistości modele opierają się na zbiorach tekstów, które stanowią jeden z rodzajów realizacji języka. Jednak tego rodzaju uogólnienie, które pojawia się na poziomie samych definicji modelu językowego, jest uprawnione, co wynika z uzasadnień dostarczonych przez hermeneutykę.

Należy także dodać, że pomimo odwoływania się do kategorii języka obydwie modele nie korzystają praktycznie z narzędzi lingwistycznych, ale proponują własne techniki, takie jak tokenizacja, LSTM czy idea tzw. uwagi (*attention*). Są to swoiste i skuteczne techniki manipulowania tekstem, który został poddany dycytywacji, tj. którego struktura semantyczna została zinterpretowana numerycznie.

Sztuczne teksty, traktowane na równi z tekstami ludzkimi, można uznać za jedne z bezpośrednich dowodów na istnienie wiedzy o charakterze poza-ludzkim, tj. niezwiązanej bezpośrednio z kognitywnymi bądź mentalnymi kompetencjami człowieka.

Tego rodzaju wniosek prowadzi do pytania o instancję, która realizuje artykulację wiedzy, przy czym, analogicznie do procedury dyskursywnej, istnieje wyraźny związek z kontekstem wytwarzanym przez człowieka, który ma charakter społeczny w szerokim tego słowa znaczeniu. Język jest bowiem tworem społecznym. Tę problematykę rozwinął Michel Foucault, którego teksty stały się inspiracją dla dalszych, bardzo obszernych i kontynuowanych nadal badań (Foucault 1966, 1971). Związek ten ujawnia się w dyskursie, który jest fenomenem językowym stającym na granicy kontekstu społecznego i pojedynczych instancji aktów językowych, realizowanych przez wypowiedzi człowieka.

Sytuacja modeli językowych, a w szczególności modelu GPT-3, jako sztucznych agentów kognitywnych w kontekście generowania tekstu cechuje się daleko idącą analogią do sytuacji i kompetencji człowieka. W szczególności dotyczy to ontologicznej podstawy wiedzy, jaką w obydwu sytuacjach jest język, który w przypadku modeli językowych materializuje się w postaci ustalonego korpusu językowego. Chociaż w wypadku człowieka tego rodzaju podstawa nie jest podobnie skonkretyzowana, jednak, na co wskazują badania dyskursu, jest nią społeczna praktyka użycia języka, będąca jedynym źródłem wiedzy ludzkiej, a materializująca się jako teksty czy wypowiedzi. Język jest bowiem jedynym nośnikiem wiedzy ludzkiej i materializuje się jedynie w taki sposób.

Wskazane cechy instancji artykułującej wiedzę nadają jej status ontologiczny do pewnego stopnia podobny do ludzkiego. Jest to sytuacja bezprecedensowa i wymagająca przemyślenia od nowa statusu maszyny cyfrowej.

Dyskusja

Przedstawiony model GPT-3 pozwala zinterpretować bardzo istotny związek, jaki istnieje między wiedzą a językiem, który pojawia się bezpośrednio jako masywne zbiory tekstów. Istnieje rozbudowana teoria tego związku, oparta na idei dyskursu. Z drugiej strony istnieje także rozbudowana teoria tekstu jako przedmiotu interpretacji (znaczeniowej), tj. hermeneutyka, uzasadniająca pełną wartość tekstu generowanego sztucznie. Jednak należy wskazać na pewne, utrudniające postępowanie badawcze, okoliczności.

1. Hermeneutyka, jako zespół refleksji filozoficznych, ma charakter całkowicie jakościowy. Jej procedury nie zostały także opracowane z myślą o ewaluacji tekstów, ale ich interpretacji. Jednak z drugiej strony istotne jest założenie leżące u podstaw hermeneutyki, które nadaje tekstowi całkowitą autonomiczność i odrywa go od instancji autora – człowieka. W pierwotnej wersji ta autonomiczność oznaczała boskie pochodzenie tekstu poddawanego egzegezie, co w wypadku tekstów tworzonych sztucznie stanowi korzystną okoliczność.
2. Zbiórów danych tekstowych używanych jako korpus językowy przy tworzeniu modelu języka GPT-3 i podobnych nie można uznać za ostateczną reprezentację zasobów językowych. Podlegają one bowiem procesom ciągłego konstruowania. Procesy te w znacznym stopniu są spowodowane przez wysoką zmienność tego rodzaju zasobów internetowych, na których w większości są oparte, choć wspomniany korpus obejmuje także zasoby o charakterze względnie stałym, np. różnego rodzaju publikacje.
3. Internetowe zasoby tekstowe nie są ani stałe, ani jednorodne. Są natomiast skutkiem aktywności różnego typu społeczności o różnym charakterze kulturowym i społecznym, a także wspólnot narodowych, co różnicuje je także językowo. Nie stanowią one także spójnej całości. Mają natomiast charakter rozproszony, heteronomiczny i zmienny w czasie. Ich wysoki stopień wewnętrznych powiązań i relacji sprawia, iż stanowią one bez wątpienia zbiór o wysokim stopniu złożoności. W tej sytuacji deterministyczne uporządkowanie tego zbioru, np. tworzenie precyzyjnych kryteriów dla jego porządkowania i filtrowania, jest trudne lub wręcz niemożliwe.
4. Zbiory tekstów pełniące funkcję materiału bazowego w procesie tworzenia korpusu dla modelu językowego mają arbitralny, a nawet przypadkowy charakter. Autorzy modeli często posługują się nieostrych kryteriami wyboru, takimi jak jakość, w żaden sposób tej jakości nie definiując i przyjmując ją intuicyjnie. Ich ogólne założenie – maksymalne bogactwo – jest jedynie luźnym postulatem. Każda ingerencja w korpus powiększa stopień jego przypadkowości, która wcale nie musi oznaczać osiągnięcia poziomu statystycznej reprezentatywności wiedzy, języka itp.

5. Każda z procedur (algorytmów) stosowanych w obrębie modeli GPT ma wpływ na strukturę semantyczną początkowego korpusu językowego, począwszy od procesu tokenizacji, która w wypadku modelu GPT-3 korzysta z techniki tzw. kodowania przez pary bajtów (*Byte Pair Encoding*, BPE). Polega ona na poddawaniu słów zakodowanych w postaci bajtów procedurze ujednolicenia i uproszczenia. Pozwala to na stworzenie matematycznych parametrów semantyki tekstów za pomocą techniki osadzania (*embedding*). Proces digityzacji umożliwia dalsze procedury przetwarzania tekstu, który przyjmuje charakter liczbowy w postaci wektorów, tj. list wartości numerycznych. Takie procedury, bardzo liczne i wyrafinowane, muszą mieć istotny wpływ na formowanie się ostatecznej struktury semantycznej (Helwe et al. 2021; Kozłowski et al. 2019; Merrill et al. 2021; Qiu, Xu 2022; Traylor et al. 2021; Yoshida, Oseki 2022).

Wnioski

Model językowy GPT-3 można traktować jak generator autonomicznego tekstu równego ludzkiemu, co powoduje dalsze konsekwencje w postaci zmiany rozumienia fenomenu wiedzy ujawnianej w tym tekście. Zmiana ta polega na dopuszczeniu istnienia wiedzy, której dysponentem/właścicielem nie jest bezpośrednio człowiek. Takie stwierdzenie oznacza zmianę dotychczasowego myślenia o wiedzy.

Uzasadnienia dla autonomicznej wartości tekstu generowanego sztucznie dostarcza podejście hermeneutyczne, które reprezentuje fundamentalne podejście do tekstu jako takiego. To podejście, które rozwinęło się w zaawansowaną refleksję filozoficzną, pozwala traktować tekst jako swoisty i samodzielny fenomen o wysokim stopniu złożoności. Jednocześnie nie przesądza ona o źródle tekstu – w swoich początkach hermeneutyka zajmowała się egzegezą tekstów boskich. Teksty generowane sztucznie, a więc w pewnym sensie nie-ludzkie, zrównują się na tym tle z tekstami pochodzącymi od człowieka, a więc mogą i powinny podlegać także analizie dyskursywnej.

Opisany model GPT-3 potwierdza istnienie wiedzy zawartej w języku, operacyjnie zrealizowanym jako korpus tekstów. Nie pozwala on jednak zrozumieć natury tego istnienia, mechanizmów, jakim podlega itd., ze względu na implicytny charakter wiedzy, jaka pojawia się w wypadku tego modelu. Do takiej wiedzy mogą mieć zastosowanie procedury oparte na językowym charakterze wiedzy, takie jak oparte na idei dyskursu, np. teoria przestrzeni dyskursywnej.

Granica między podejściem ilościowym i jakościowym a odpowiednio formalnym (matematycznym, syntaktycznym) i treściowym (semantycznym) w obrębie technologii cyfrowej wydaje się praktycznie pokonana, ale zastosowane założenia i wybory, których dokonywali twórcy, wywołują następne problemy i pytania.

Model GPT-3 i jego następcy jako zrealizowany doświadczalnie autonomiczny model generowania inteligibilnego tekstu otwiera także pytania dotyczące sposobu istnienia jego algorytmów i realizującej go maszyny obliczeniowej w kontekście analogicznych kompetencji ludzkich. Takie pytania mogą prowadzić do konieczności zrekonstruowania istniejących poglądów na temat wyłączności i niepowtarzalności pewnych kognitywnych kompetencji ludzkich oraz przemyślenia statusu tych maszyn i algorytmów. Ta problematyka znajduje się jeszcze w bardzo początkowej fazie rozwoju, ale nie można jej zlekceważyć ze względu na istniejące przesłanki doświadczalne.

Bibliografia

- Aggarwal Charu C. (2018). *Machine Learning for Text*. Cham, Switzerland: Springer International Publishing.
- Angermüller Johannes, Maingueneau Dominique, Wodak Ruth (2014). An Introduction. W: Johannes Angermüller, Dominique Maingueneau, Ruth Wodak (red.), *The Discourse Studies Reader: Main Currents in Theory and Analysis*. Amsterdam–Philadelphia: John Benjamins Publishing Company, 1–14.
- Bender Emily M., Koller Alexander (2020). Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data. W: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185–5198, <https://doi.org/10.18653/v1/2020.acl-main.463> [odczyt: 21.06.2022].
- Bleicher Joseph (1980). *Contemporary Hermeneutics: Hermeneutics as Method, Philosophy and Critique* (reprint, 1982 edition). London–Boston: Routledge & Kegan Paul.
- Brown Tom B., Mann Benjamin, Ryder Nick, Subbiah Melanie, Kaplan Jared, Dhariwal Prafulla, Neelakantan Arvind, Shyam Pranav, Sastry Girish, Askell Amanda, Agarwal Sandhini, Herbert-Voss Ariel, Krueger Gretchen, Henighan Tom, Child Rewon, Ramesh Aditya, Ziegler Daniel M., Wu Jeffrey, Winter Clemens, Hesse Christopher, Chen Mark, Sigler Eric, Litwin Mateusz, Gray Scott, Chess Benjamin, Clark Jack, Berner Christopher, McCandlish Sam, Radford Alec, Sutskever Ilya, Amodei Dario (2020). *Language Models are Few-Shot Learners*, ArXiv:2005.14165 [Cs], <https://arxiv.org/abs/2005.14165> [odczyt: 21.06.2022].
- Charniak Eugene (2019). *Introduction to Deep Learning*. Cambridge, Massachusetts: The MIT Press.
- Common Crawl Foundation (2021). *Common Crawl*, <https://commoncrawl.org/> [odczyt: 21.06.2022].
- Crane Lesley (2016). *Knowledge and Discourse Matters: Relocating Knowledge Management's Sphere of Interest onto Language*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Deng Li, Liu Yang (red.) (2018). *Deep Learning in Natural Language Processing*. Singapore: Springer Singapore, <https://doi.org/10.1007/978-981-10-5209-5> [odczyt: 21.06.2022].

- Descartes Rene (1637). *Discours de la methode pour bien conduire sa raison, & chercher la verité dans les sciences. Plus La dioptrique. Les meteores. Et La geometrie. Qui sont des essais de cete methode*. Leyde: De l'imprimerie de Ian Maire.
- Descartes Rene (2013). *Rozprawa o metodzie. Właściwego kierowania rozumem i poszukiwania prawdy w naukach*, tłum. Tadeusz Boy-Żeleński. Kęty: Wydawnictwo ANTYK.
- Dijk Teun Adrianus van (2013). Discourse and Knowledge. W: James Paul Gee, Michael Handford (red.), *The Routledge Handbook of Discourse Analysis*. London: Routledge, 587–603.
- Fairclough Norman (2003). *Analysing Discourse: Textual Analysis for Social Research*. London: Routledge.
- Flasiński Mariusz (2016). *Introduction to Artificial Intelligence*. Cham, Switzerland: Springer.
- Fortune Business Insights (2021). *Natural Language Processing (NLP) Market Size & Growth, 2028*, <https://www.fortunebusinessinsights.com/industry-reports/natural-language-processing-nlp-market-101933> [odczyt: 21.06.2022].
- Foucault Michel (1966). *Les mots et les choses: Une archéologie des sciences humaines*. Paris: Gallimard.
- Foucault Michel (1969). *L'archéologie du savoir*. Paris: Gallimard.
- Foucault Michel (1971). *L'ordre du discours: Leçon inaugurale au Collège de France prononcée le 2 décembre 1970*. Paris: Gallimard.
- Foucault Michel (1981). *The Order of Discourse*. W: Robert E. Young (red.), *Untying the Text: A Post-Structuralist Reader*. Boston: Routledge & Kegan Paul, 51–78.
- Goodfellow Ian, Bengio Yoshua, Courville Aaron (2016). *Deep Learning*. Cambridge, Massachusetts: MIT Press, <https://www.deeplearningbook.org> [odczyt: 21.06.2022].
- Helwe Chadi, Clavel Chloé, Suchanek Fabian M. (2021). *Reasoning with Transformer-based Models: Deep Learning, but Shallow Reasoning*, 3rd Conference on Automated Knowledge Base Construction, June 22, https://openreview.net/forum?id=Ozp1WrgtF5_ [odczyt: 21.06.2022].
- Hochreiter Sepp, Schmidhuber Jürgen (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780, <https://doi.org/10.1162/neco.1997.9.8.1735> [odczyt: 21.06.2022].
- Hyland Ken, Paltridge Brian (red.) (2011). *Bloomsbury Companion to Discourse Analysis*. London–New York: Bloomsbury.
- Jørgensen Marianne, Phillips Louise (2002). *Discourse Analysis as Theory and Method*. London–Thousand Oaks, California: Sage Publications.
- Jurafsky Dan, Martin James H. (2020). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 3rd ed. Upper Saddle River, New Jersey: Prentice Hall.
- Kozłowski Austin C., Taddy Matt, Evans James A. (2019). The Geometry of Culture: Analyzing Meaning through Word Embeddings. *American Sociological Review*, 84(5), 905–949, <https://doi.org/10.1177/0003122419877135> [odczyt: 21.06.2022].
- Lévy Pierre (1999). *Collective Intelligence, Mankind's Emerging World in Cyberspace*. Cambridge, Massachusetts: Perseus Books.
- Liotard Jean-François (1979). *La condition postmoderne: Rapport sur le savoir*. Paris: Éd. de Minuit.

- Maciąg Rafał (2018). Discursive Space and Its Consequences for Understanding Knowledge and Information. *Philosophies*, 3(4), <https://doi.org/10.3390/philosophies3040034> [odczyt: 21.06.2022].
- Maciąg Rafał (2020). Ontological Basis of Knowledge in the Theory of Discursive Space and Its Consequences. *Proceedings*, 47(1), 11, <https://doi.org/10.3390/proceedings47010011> [odczyt: 21.06.2022].
- Maciąg Rafał (2022). *Wiedza jako opowieść. Przestrzeń dyskursywna*. Kraków: TAIWPN Universitas.
- Malpas Jeff, Gander Hans-Helmut (red.) (2015). *The Routledge Companion to Hermeneutics*. Abingdon, Oxon: Routledge.
- Merrill William, Goldberg Yoav, Schwartz Roy, Smith Noah A. (2021). Provable Limitations of Acquiring Meaning from Ungrounded Form: What Will Future Language Models Understand? *Transactions of the Association for Computational Linguistics*, 9, 1047–1060, https://doi.org/10.1162/tacl_a_00412 [odczyt: 21.06.2022].
- Mikolov Tomáš, Sutskever Ilya, Chen Kai, Corrado Greg, Dean Jeffrey (2013). *Distributed Representations of Words and Phrases and their Compositionality*, ArXiv:1310.4546 [Cs, Stat], <https://arxiv.org/abs/1310.4546> [odczyt: 21.06.2022].
- Neelakantan Arvind, Xu Tao, Puri Raul, Radford Alec, Han Jesse Michael, Tworek Jerry, Yuan Qiming, Tezak Nikolas, Kim Jong Wook, Hallacy Chris, Heidecke Johannes, Shyam Pranav, Power Boris, Nekoul Tyna Eloundou, Sastry Girish, Krueger Gretchen, Schnurr David, Such Felipe Petroski, Hsu Kenny, Thompson Madeleine, Khan Tabarak, Sherbakov Toki, Jang Joanne, Welinder Peter, Weng Lilian (2022). *Text and Code Embeddings by Contrastive Pre-Training*, ArXiv:2201.10005 [Cs], <https://arxiv.org/abs/2201.10005> [odczyt: 21.06.2022].
- Ouyang Long, Wu Jeff, Jiang Xu, Almeida Diogo, Wainwright Carroll L., Mishkin Pamela, Zhang Chong, Agarwal Sandhini, Slama Katarina, Ray Alex, Schulman John, Hilton Jacob, Kelton Fraser, Miller Luke, Simens Maddie, Askill Amanda, Welinder Peter, Christiano Paul, Leike Jan, Lowe Ryan (2022). *Training Language Models to Follow Instructions with Human Feedback*, <https://arxiv.org/abs/2203.02155> [odczyt: 21.06.2022].
- Paltridge Brian (2006). *Discourse Analysis: An Introduction*. London–New York: Continuum.
- Petroni Fabio, Rocktäschel Tim, Lewis Patrick, Bakhtin Anton, Wu Yuxiang, Miller Alexander H., Riedel Sebastian (2019). Language Models as Knowledge Bases? W: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2463–2473, <https://doi.org/10.18653/v1/D19-1250> [odczyt: 21.06.2022].
- Porada Ian, Sordani Alessandro, Cheung Jackie C.K. (2021). *Does Pre-training Induce Systematic Inference? How Masked Language Models Acquire Commonsense Knowledge*, ArXiv:2112.08583 [Cs], <https://arxiv.org/abs/2112.08583> [odczyt: 21.06.2022].
- Porter Stanley E., Robinson Jason C. (2011). *Hermeneutics: An Introduction to Interpretative Theory*. Grand Rapids, Michigan–Cambridge, U.K.: William B. Eerdmans Publishing Company.

- Qiu Wenjun, Xu Yang (2022). *HistBERT: A Pre-trained Language Model for Diachronic Lexical Semantic Analysis*, ArXiv:2202.03612 [Cs], <https://doi.org/10.13140/RG.2.2.14905.44649> [odczyt: 21.06.2022].
- Radford Alec, Wu Jeffrey, Child Rewon, Luan David, Amodei Dario, Sutskever Ilya (2019). *Language Models are Unsupervised Multitask Learners*, https://d4mucfpksywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf [odczyt: 21.06.2022].
- Rajpurkar Pranav, Jia Robin, Liang Percy (2018). *Know What You Don't Know: Unanswerable Questions for SQuAD*, ArXiv:1806.03822 [Cs], <https://arxiv.org/abs/1806.03822> [odczyt: 21.06.2022].
- Rajpurkar Pranav, Zhang Jian, Lopyrev Konstantin, Liang Percy (2016). *SQuAD: 100, 000+ Questions for Machine Comprehension of Text*, CoRR, ArXiv:1606.05250, <https://arxiv.org/abs/1606.05250> [odczyt: 21.06.2022].
- Safavi Tara, Koutra Danai (2021). *Relational World Knowledge Representation in Contextual Language Models: A Review*, ArXiv:2104.05837 [Cs], <https://arxiv.org/abs/2104.05837> [odczyt: 21.06.2022].
- Sowa John F. (2010). The Role of Logic and Ontology in Language and Reasoning. W: Roberto Poli, Johanna Seibt (red.), *Theory and Applications of Ontology: Philosophical Perspectives*. Dordrecht: Springer Netherlands, 231–263, https://doi.org/10.1007/978-90-481-8845-1_11 [odczyt: 21.06.2022].
- Tamkin Alex, Brundage Miles, Clark Jack, Ganguli Deep (2021). *Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models*, ArXiv:2102.02503, <https://arxiv.org/abs/2102.02503v1> [odczyt: 21.06.2022].
- Traylor Aaron, Feiman Roman, Pavlick Ellie (2021). AND does not Mean OR: Using Formal Languages to Study Language Models' Representations. W: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*. Vol. 2: *Short papers*, 158–167, <https://doi.org/10.18653/v1/2021.acl-short.21> [odczyt: 21.06.2022].
- Turing Alan M. (1950). Computing Machinery and Intelligence. *Mind*, 236(14), 433–460.
- Vaswani Ashish, Shazeer Noam, Parmar Niki, Uszkoreit Jakob, Jones Llion, Gomez Aidan N., Kaiser Lukasz, Polosukhin Illia (2017). *Attention Is All You Need*, ArXiv:1706.03762 [Cs], <https://arxiv.org/abs/1706.03762> [odczyt: 21.06.2022].
- Wang Alex, Singh Amanpreet, Michael Julian, Hill Felix, Levy Omer, Bowman Samuel R. (2018). *GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding*, ArXiv:1804.07461, <https://arxiv.org/abs/1804.07461> [odczyt: 21.06.2022].
- Wittgenstein Ludwig (1958). *Philosophical Investigations* (second edition). Oxford: Basil Blackwell.
- Yoshida Ryo, Oseki Yohei (2022). Learning Argument Structures with Recurrent Neural Network Grammars. W: *Proceedings of the Society for Computation in Linguistics*, 5(1), 101–111, <https://doi.org/10.7275/kne0-hc86> [odczyt: 21.06.2022].

- Zhang Daniel, Maslej Nestor, Brynjolfsson Erik, Etchemendy John, Lyons Terah, Manyika James, Ngo Helen, Niebles Juan C., Sellitto Michael, Sakhaee Ellie, Shoham Yoav, Clark Jack, Perrault Raymond (2021). *The AI Index 2021 Annual Report*, Stanford, California: AI Index Steering Committee, Human-Centered AI Institute, Stanford University.
- Zhou Xuhui, Zhang Yue, Cui Leyang, Huang Dandan (2020). Evaluating Commonsense in Pre-Trained Language Models. W: *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05), 9733–9740, <https://doi.org/10.1609/aaai.v34i05.6523> [odczyt: 21.06.2022].