

[Akceptuje](#)

W ramach naszej witryny stosujemy pliki cookies w celu świadczenia państwu usług na najwyższym poziomie, w tym w sposób dostosowany do indywidualnych potrzeb. Korzystanie z witryny bez zmiany ustawień dotyczących cookies oznacza, że będą one zamieszczone w Państwa urządzeniu końcowym. Możecie Państwo dokonać w każdym czasie zmiany ustawień dotyczących cookies. Więcej szczegółów w naszej [Polityce Prywatności](#)

[Portal](#) [Informacje](#) [Katalog firm](#) [Praca](#) [Szkolenia](#) [Wydarzenia](#) [Porównania międzylaboratoryjne](#)
[Kontakt](#)



[Laboratoria](#)
[.net](#)
[Innowacje](#)
[Nauka](#)
[Technologie](#)



[Logowanie](#) [Rejestracja](#) [pl](#)

Newsletter

zapisz się

Naukowy styl życia

Nauka i biznes

- [Nowe technologie](#)
- [Felieton](#)
- [Tygodnik "Nature"](#)
- [Edukacja](#)
- [Artykuły](#)
- [Przemysł](#)

[Strona główna](#) > [Informacje](#)

Polskie okno na ChatGPT

Wrocławscy badacze pracują nad polskim odpowiednikiem ChatuGPT. Aby go jednak opracować, potrzebują jak najwięcej danych o tym, jak wyglądają rozmowy Polaków ze sztuczną inteligencją. Dlatego apelują o to, by w ramach rozmów z ChatemGPT korzystać z przygotowanego przez nich polskiego okienka.

ChatGPT udostępniony został w listopadzie ub.r. przez amerykańską firmę OpenAI. To bazujący na sztucznej inteligencji generator treści - bot, z którym porozumiewać się można językiem naturalnym. Narzędzie - posługuje się ono również językiem polskim - jest w stanie odpowiadać na pytania, tłumaczyć dokumenty na różne języki, przeprowadzić korektę i redakcję językową tekstów, streszczać i analizować prace naukowe, sugerować rozwiązania przeróżnych problemów, pisać eseje, scenariusze, poprawiać błędy w kodach programistycznych, przeszukiwać bazy danych. A jego zastosowania dopiero są odkrywane.

"Szacujemy, że nawet do 70 proc. osób w Polsce nie miało do czynienia z tym czatem. Dla wielu osób trudnością nie do pokonania jest choćby to, że ChatGPT nie ma polskiego interfejsu. W dodatku, aby móc z czatu skorzystać, trzeba się zalogować - używając logowania przez konto Google, czy podając numer telefonu. To bariera, której wiele osób nie jest w stanie przekroczyć. My wychodzimy tym problemom naprzeciw" - tłumaczy w rozmowie z PAP dr Jan Kocoń z projektu CLARIN-PL Politechniki Wrocławskiej.

Zespół z Wrocławia przygotował polskie okienko dialogowe dla ChatuGPT. Pomysł jest bardzo prosty: za pośrednictwem polskiej strony rozmawia się z ChatemGPT, ale wgląd do tej rozmowy mają również polscy badacze. Dzięki temu polscy użytkownicy mają łatwiejszy dostęp do amerykańskiej aplikacji, a badacze zdobywają bazę zapytań do czatu i informacje, czego w tych odpowiedziach brakuje.

Strona zespołu CLARIN-PL przygotowana jest w języku polskim. A kilka pierwszych pytań można zadać natychmiast, bez logowania. Dla użytkowników, którzy się zalogują (rejestracja jest darmowa) przewidziane są większe limity niż w przypadku bezpłatnego dostępu do ChatuGPT. W ten sposób badacze chcą zachęcić polskich internautów do uchylenia rąbka tajemnicy na temat tego, jak wyglądają ich konwersacje ze sztuczną inteligencją.

"Pracujemy nad polskim odpowiednikiem ChatuGPT. Żeby to rozwiązanie miało szansę zaistnieć, musimy zebrać jak najwięcej informacji o metodach używania takich czatów przez polskich odbiorców. ChatGPT powstał za granicą i niekoniecznie przy jego opracowywaniu pochyłono się tam nad problemami, które są istotne dla polskich użytkowników" - mówi dr Kocoń.

Ocenia, że ChatGPT znacznie słabiej posługuje się językiem polskim niż np. angielskim. Popelnia błędy językowe i nie najlepiej rozumie zapytania sformułowane w naszym języku. To widać na przykład wtedy, kiedy prosimy chat o napisanie wiersza czy piosenki - zwraca uwagę badacz. Po angielsku wychodzi to całkiem nieźle, ale po polsku tekst zwykle się nawet nie rymuje.

"Nie mamy informacji, jak powstawał model OpenAI, ale naszym głównym podejrzeniem jest to, że on polskiego 'widział' relatywnie niewiele w stosunku do innych języków. Najprawdopodobniej model stosuje międzyjęzykowy transfer wiedzy na podstawie bazy tłumaczeń" - ocenia naukowiec.

Badacz opisuje, że tworzenie sztucznej inteligencji składa się z dwóch głównych etapów: trzeba mieć dużą bazę danych - w tym przypadku tekstów, na podstawie których model uczy się języka. A potem potrzebna jest jeszcze baza zapytań i odpowiedzi, na podstawie których sztuczna inteligencja uczy się generować pożądane treści.

Problemem jest nie tylko to, że ChatGPT widział mało tekstów w języku polskim na etapie tworzenia modeli języka, ale i na etapie szkolenia widział mało polskich instrukcji, zapytań.

Dlatego naukowcy z Wrocławia chcą opracować model, w którego sercu będzie od samego początku język polski. "Nie jesteśmy w stanie konkurować z OpenAI w języku takim jak angielski, ale jeśli chodzi o języki słowiańskie - mamy dużo do zaoferowania. Mamy zebraną bardzo dużą bazę tzw.

tekstów korpusowych (służących do badań lingwistycznych) - w języku polskim. Na ich podstawie jesteśmy w stanie zrobić duży model językowy. A potem chcemy go stroić na instrukcjach, które dostaniemy od użytkowników” - tłumaczy naukowiec.

“Najważniejsze dla nas jest to, aby za pośrednictwem naszego okienka użytkownicy zgłaszali różnego rodzaju nieprawidłowości wynikające z używania czata” - tłumaczy Jan Kocoń.

Jeśli czat poda odpowiedź niezgodną z oczekiwaniami, można wcisnąć smutną buźkę pod okienkiem dialogowym. Już to wystarczy badaczom jako sygnał, żeby przyjrzeć się odpowiedzi bota i sprawdzili, co jest nie tak.

Po każdej ocenie odpowiedzi otwiera się okienko, w którym można naukowcom wpisać swój komentarz - np. by zwrócić uwagę na błędy językowe; na to, że ChatGPT zmyślił jakieś informacje albo że jego odpowiedź była obraźliwa. A nawet wpisać, jaka odpowiedź by nas satysfakcjonowała. W ten sposób nie tylko pomagamy polskim badaczom, ale również trenujemy swoje krytyczne myślenie i ćwiczymy ograniczone zaufanie do sztucznej inteligencji.

Komentarze i reakcje internautów będą przeglądane przez polski zespół badaczy. A zapytania i wnioski z tych rozmów będą służyły do prac nad polskim botem. “Żeby powstał dobry czat, musi mieć do dyspozycji dużo instrukcji. Tak zrobiło OpenAI - zatrudniło mnóstwo ludzi, którzy rozmawiali z botem i korygowali odpowiedzi czatu”- mówi rozmówca PAP. W ten sposób czat uczył się, jakie treści są pożądane, a jakich treści nie powinien generować.

Chodzi m.in. o to, by nauczyć sztuczną inteligencję, że jest pewna klasa pytań, na które czat nie może odpowiedzieć wprost (np. treści mogące ułatwić popełnienie przestępstwa, naruszające prywatność lub obrażające uczucia religijne). Ktoś musiał więc kiedyś ręcznie dla takiej klasy zapytań przygotować wzorcowe odpowiedzi, a model stroił się do tych instrukcji.

Polscy badacze nie mają takich nakładów jak amerykańska firma. Nie są w stanie sami przewidzieć wszystkich możliwych zastosowań czata i sprawdzić, czy dobrze się on spisuje. Skorzystają więc z interakcji polskich użytkowników z ChatemGPT, żeby trenować swój model.

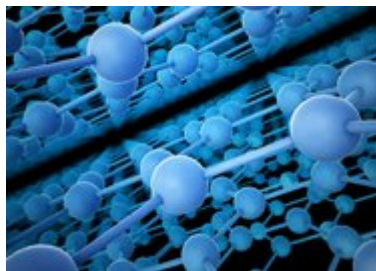
Badacz informuje, że jego uczelnia - we współpracy z Wrocławskim Centrum Sieciowo-Superkomputerowym - jest w trakcie zakupów sprzętu obliczeniowego, który ma przydać się w rozwijaniu badań nad polską sztuczną inteligencją. Budżet jest duży - to ok. 80 mln zł, ale sprzęt trafi do naukowców dopiero w przyszłym roku. Już teraz jednak badacze przygotowali dodatkowy budżet na dostęp do mocy obliczeniowych. “Nie chcemy czekać na sprzęt z założonymi rękami, już teraz przygotowujemy dane, na których będziemy trenować model” - relacjonuje.

Naukowiec apeluje też do badaczy i specjalistów z różnych dziedzin o pomoc w pracach nad polską sztuczną inteligencją. “Jeśli chcemy mieć polskie wysokomarżowe technologie, musimy je rozwijać i badać. A bez danych nie ruszymy z miejsca” - podkreśla.

W skład zespołu wchodzi: Bartosz Walkowiak, Dawid Banach, Tomasz Walkowiak, Magdalena Drewniak, Jan Wiczorek, Paweł Kazienko, Tomasz Naskręt, Jan Kocoń, Maciej Piasecki.(PAP)

Źródło: pap.pl

<http://laboratoria.net/aktualnosci/31836.html>



28-05-2024

[Drżące nanorurki](#)

Właściwości zależą m.in. od tego, w jaki sposób struktury te wibrują.



28-05-2024

[Naukowcy znaleźli sposób na recykling betonu](#)

Informuje "Nature".



28-05-2024

[ADHD zdiagnozowano u co dziewiątego dziecka w USA](#)

W roku 2022 dzieci z diagnozą ADHD było o milion więcej niż w roku 2016.



28-05-2024

Testy na obecność HPV

Co osiem lat równie skuteczne, co regularna cytologia.



28-05-2024

Do środowiska trafiło ponad 1 mld komarów GMO

Przeznaczonych do walki z malarią.



28-05-2024

Może to owady uratują nas przed zwałami plastiku

Niektóre gatunki owadów są w stanie zjadać plastik.



28-05-2024

[Terapia daremna przedłuża cierpienie, przedłuża agonię](#)

Terapia daremna nie jest w stanie pomóc pacjentowi.



28-05-2024

[Widzimy eskalację zaburzeń związanych ze stresem](#)

Szeroko rozumianych lękowo-depresyjnych.

Informacje dnia: [Drżące nanorurki](#) [Naukowcy znaleźli sposób na recykling betonu](#) [ADHD zdiagnozowano u co dziewiątego dziecka w USA](#) [Testy na obecność HPV](#) [Do środowiska trafiło ponad 1 mld komarów](#) [GMO](#) [Może to owady uratują nas przed zwałami plastiku](#) [Drżące nanorurki](#) [Naukowcy znaleźli sposób na recykling betonu](#) [ADHD zdiagnozowano u co dziewiątego dziecka w USA](#) [Testy na obecność HPV](#) [Do środowiska trafiło ponad 1 mld komarów](#) [GMO](#) [Może to owady uratują nas przed zwałami plastiku](#)

Partnerzy