Kto pyta nie błądzi

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

PPNT
Poznański Park Naukowo-Technologiczny

Transkrypcja wystąpienia

SZTUCZNA INTELIGENCJA

Bartosz Naskręcki

Transkrypcja wystąpienia, które odbyło się podczas II cyklu spotkań organizowanych w ramach projektu Kto pyta nie błądzi – Nauka dostarczy odpowiedzi. Wystąpienie miało miejsce 11 czerwca 2024 roku.

Bartosz Naskręcki: Dzień dobry, witam was serdecznie. Nazywam się Bartek, możecie się do mnie tak zwracać, myślę, że będzie łatwiej. Dzisiaj otrzymałem dosyć trudne zadanie – porozmawiać z wami o haśle, które nazywa się sztuczna inteligencja. Swoją drogą, jeśli będziecie mieli pytania, widzę, że tutaj zrobiła się dosyć potężna przerwa między nami. Wcześniej cały pierwszy rząd był wypełniony szóstoklasistami, ale myślę, że będziecie chcieli zadawać pytania. Jest kilka rzeczy, które musimy sobie wyjaśnić, więc może rozpocznijmy od prostego pytania. Jak myślicie, czym jest inteligencja? Z czym wam się to kojarzy? Macie jakiś pomysł? Możecie do mnie po prostu krzyknąć, powiedzieć, co wam przychodzi do głowy. Nie ma głupich odpowiedzi. No co, co to może być inteligencja? Czy na przykład, jeśli siedzę na matematyce i rozwiązuję zadanie, i

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

uda mi się je rozwiązać, to jest to objaw inteligencji? A na przykład, jeśli idę ulicą i widzę fajnego chłopaka, i mi się podoba, to czy to też jest forma inteligencji? Czy uczucia i takie rzeczy mają coś wspólnego z inteligencją, czy niekoniecznie? A słuchajcie, jak już jesteśmy w tym temacie, i ktoś się w kimś zakocha, to zakochujemy się mózgiem czy sercem? Mózgiem, bije nam szybko serce, ale to wszystko dzieje się w naszej głowie. Teraz ten temat, który będziemy zgłębiać, dotyczy różnych zjawisk związanych z myśleniem. Czyli inteligencja to jest myślenie. Co to jest myślenie? To jest forma interakcji z otoczeniem, która przynosi różnego rodzaju efekty, tak? Czy coś zmieniamy w naszym otoczeniu. Tak, w pewnym sensie możemy myśleć o inteligencji. A słowo sztuczna? Jak pomyślicie o rzeczach, które spotkaliście w swoim życiu i nazwaliście inteligentnymi, to czy na przykład krzesło można nazwać inteligentnym? Dlaczego? No właśnie, co robi krzesło? W sensie, siedzimy na nim i tyle, prawda? Ale na przykład, jeśli postawimy krzesło w sytuacji, w której mówimy: okej, tutaj mamy krzesło, i teraz chcemy, żeby to krzesło wykonało jakąś czynność, na przykład pojechało samochodem – to niespecjalnie. Ale z drugiej strony, jeśli weźmiecie coś, co jest w sumie trochę bardziej skomplikowane od krzesła, no dużo bardziej, ale nie aż tak – mrówka – to czy mrówka jest inteligentna? Ale czy można na przykład usiąść z mrówką, wypić kawę i porozmawiać z nią na przykład o Mickiewiczu? No może mrówka nie lubi Mickiewicza, prawda? Ale czy można z mrówką porozmawiać o Mickiewiczu? Można, tylko że ta rozmowa będzie bardzo jednostronna, prawda? Więc jakby w drugą stronę nie będziemy mieli możliwości odzyskania informacji zwrotnej. A można porozmawiać z własnym psem? Można. I co ciekawe, pies nawet zareaguje, prawda? Czyli jest tu jakaś forma inteligencji. A teraz popatrzcie na to w ten sposób: mam ze sobą telefon i teraz w tym telefonie mam taką aplikację, która się nazywa Chat GPT. I teraz ja z tą aplikacją mogę zrobić taką rzecz, uwaga, spróbuję to zrobić. Zawsze się udaje. Włączę do mikrofonu część. Cześć, jak mogę ci dzisiaj pomóc? Słuchaj, jestem na wykładzie i mam opowiadać o sztucznej inteligencji, powiesz mi, co to jest? Oczywiście, sztuczna inteligencja, czyli AI, to dziedzina informatyki zajmująca się tworzeniem systemów zdolnych do wykonywania zadań, które normalnie wymagają ludzkiej inteligencji. Obejmuje to takie czynności, jak rozpoznawanie mowy, podejmowanie decyzji, rozumienie języka naturalnego, uczenie się i adaptacja. AI może być używana w różnych dziedzinach – od medycyny, przez motoryzację. Czy to jest forma inteligencji? W sensie, przed chwilą mówiliśmy, że z krzesłem nie da się porozmawiać, z mrówką nie da się porozmawiać, da się porozmawiać z psem – ta rozmowa jest mniej lub bardziej interaktywna. Ale ta rozmowa ewidentnie była interaktywna. Rozmawiałem z czymś, co odpowiadało. Czy to było naturalne? W sensie, czy to, co tutaj mam, to jest element naszej natury, powstało w toku ewolucji? Nie bardzo. Więc to jest przykład czegoś, co się dosyć niedawno pojawiło. Mamy algorytm, który jest sztuczny, stworzony

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

przez człowieka, i to jest przykład interakcji. My możemy coś realnego zrobić. To jest przykład inteligencji. I o tym będziemy dzisiaj rozmawiać.

MIT 1

Czy sztuczna inteligencja przejmie władzę nad światem?

Bartosz Naskręcki: Okej, przejdźmy do mitów związanych z tymi różnymi algorytmami, o których będziemy rozmawiać. Mit numer 1, jeszcze ładna animacja. Czy sztuczna inteligencja przejmie władzę nad światem? Dobra, to jest ciekawe pytanie. Teraz pomyślmy sobie w ten sposób: to, co ja mam w tym telefonie, rozmawiało ze mną, tak? Czy jesteście sobie w stanie wyobrazić scenariusz, gdzie mój telefon wychodzi z mojej kieszeni, idzie w świat i zaczyna w jakiś sposób wpływać na ludzi? Czy to jest sensowny scenariusz, czy niespecjalnie? No właśnie, telefon nie ma nóg, chyba że ja go, nie wiem, wyposażę w robotyczne nogi. No więc w oczywisty sposób nie może przejąć władzy nad światem. Ale z drugiej strony, zastanówmy się nad tym. Co to znaczy przejąć władzę nad światem? Czy ktoś z was oglądał wczoraj może film na jakiejś platformie streamingowej? Okej, super. A czy zastanawialiście się nad tym, jak to jest, że jak uruchamiasz telewizor i masz listę filmów, to te filmy wam się zazwyczaj podobają, te które są najczęściej podpowiadane? W sensie, co powoduje, że jak włączasz ten telewizor, to widzisz propozycje filmów, które wam się podobają? Algorytm. Ale skąd się wziął ten algorytm? No właśnie, ktoś napisał jakiś program. Czyli sztuczna inteligencja, jakkolwiek by to postrzegać, to co przed chwilą zrobiłem, to nie jest mrówka, to nie jest nasz pies, to jest algorytm, który wykonuje pewne jasno określone zadania. No z tym jasno to jeszcze do tego wrócimy, ale ewidentnie jest to algorytm. Ale teraz z drugiej strony pomyślmy sobie tak: czy w pewnym sensie to, co robi mrówka, pojedyncza mrówka, też nie jest formą algorytmu? Jak myślicie? Trochę jest. No a teraz przeskalujmy to dalej. To, jeżeli mrówka wykonuje algorytm, to jeżeli wezmę mojego psa na łąkę i go puszczę luzem, to czy pies wykonuje jakiś algorytm? A jaki konkretnie algorytm wykonuje pies? W sensie, co ten pies robi na tej łące? Różne rzeczy, prawda? Zobaczcie, to jest dosyć istotna rzecz w tych wszystkich rozważaniach, które się tu pojawią, że z inteligencją to jest trochę tak jak z tą mrówką. Im mniejsza skala, tym łatwiej jest nam wskazać, że coś jest inteligentne, nie jest inteligentne, w sensie czynności, które wykonuje jakiś dany agent. Natomiast jak się skala powiększa, dochodzimy do psa albo do roju mrówek i tak dalej, to nagle okazuje się, że odpowiedź na pytanie, na przykład czy człowiek, czy ktokolwiek z was, myśli o sobie w ten sposób, że jesteście pewną formą algorytmu. A kto was

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

zaprogramował? Życie. No właśnie, to jest ciekawa sprawa. Życie. Można pomyśleć o tym, że życie jest formą programowania algorytmów, my jesteśmy emanacjami tych algorytmów, ale jak widać, to jest bardziej skomplikowane. Działanie tych algorytmów, przynajmniej tych, o których mówimy, jest na pewno ściśle uzależnione od człowieka. Nie ma na razie takich algorytmów, które moglibyśmy puścić całkowicie luzem i one byłyby zupełnie poza ludzką kontrolą. To się jeszcze nie wydarzyło, ale niewątpliwie takie rzeczy powoli zaczynają się dziać. Natomiast to, co jest ważne, to to, że algorytm, jakkolwiek by o nim nie myśleć, nie ma żadnych swoich imperatywów, myśli, pomysłów, przynajmniej niczego takiego nie zaobserwowaliśmy do tej pory. On po prostu jest i wykonuje ściśle określone zadanie. Ale teraz pojawia się pewien problem, bo jeżeli zaczniemy budować coraz bardziej skomplikowane zadania, typu przeczytaj książkę, wyciągnij z niej streszczenie – no to jeszcze wydaje się całkiem rozsądne zadanie – ale jeżeli poprosimy na przykład Chata GPT o stworzenie wiersza i ten wiersz okaże się całkiem fajny, no to skąd tak naprawdę ten wiersz się wziął? Przecież tego wiersza nikt wcześniej nie napisał. Ktoś napisał ewidentnie ten algorytm, który działa, został on wytrenowany na danych, które pochodzą od ludzi, ale wiersza wcześniej nie było. No więc jakaś forma inteligencji się tutaj pojawia. To, co na pewno możemy powiedzieć z perspektywy tego, co dzisiaj wiemy, to to, że algorytmy rządzą naszym życiem, a właściwie kompletnie przenikają nasze życie. I teraz pytanie do was, takie dosyć niepokojące: wyobraźcie sobie sytuację w ostatnim roku waszego życia, kiedy robiliście coś całkowicie z własnej woli. Czy coś wam przychodzi do głowy? W sensie, pójście do szkoły to jest pewna forma przymusu, jest nawet taki obowiązek ustawowy. Co to jest ustawa? Ustawa to pewien zestaw procedur, w których organizuje się społeczeństwo, to też jest forma algorytmu. Czy zrobiliście coś w ostatnim roku swojego życia, co nie byłoby podporządkowane albo jakiejś potrzebie przymusowej, po prostu z własnej woli? Coś wam przychodzi do głowy? No śmiało, czy ktoś z was na przykład poczuł potrzebę chwili i poszedł na łąkę, położył się na łące i obejrzał sobie niebo? Okej, super. A teraz pomyślcie dalej, to może być trochę przerażające, co spowodowało, że poszliście na łąkę? Na przykład to, że obejrzeliście dzień wcześniej film w telewizji o tym, że ktoś poszedł na łąkę. To jest pewna nadzieja, ale warto pomyśleć o tym w ten sposób, że żyjemy w takich dziwnych czasach, że jak się głęboko zastanowimy nad tym, co na co dzień robimy, to często dojdziemy do konkluzji, że to nie jest wcale takie oczywiste, czy to, co robimy, nie jest już pewnego rodzaju formą podporządkowania się algorytmowi. Więc moja rada dla was: jeżeli coś macie wynieść z tego wykładu, to czasami wyjdźcie na tę łąkę i zacznijcie myśleć po swojemu, w sensie próbujcie wyjść poza proste algorytmy, które ktoś wam zewnątrz narzuca. Ktoś wam każe oglądać film, nie oglądajcie tego filmu, spróbujcie pomyśleć o tym, co byście zrobili innego. Okej? Jeżeli rozwiązujecie

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

jakiś problem, nie czytajcie podręcznika, spróbujcie ten problem zanalizować, spróbujcie coś zrobić po swojemu. Ale niestety problem polega na tym, że nasze życie stało się przyjemne między innymi dlatego, że żyjemy w cywilizacji, która w zasadzie jest formą algorytmu. Więc jeżeli my patrzymy na mrówkę z taką wyższością i mówimy: o, zobacz, ty jesteś taka mało rozgarnięta, ale my to jesteśmy inteligentni – no to przeskalujmy nas do całego społeczeństwa. To my jesteśmy tymi mrówkami, nie? I to, co robimy na co dzień, często nie jest wcale tylko podporządkowane naszej woli, ale różnego rodzaju algorytmom. I te algorytmy one często są niewidoczne dla nas.

MIT 2

Czy sztuczna inteligencja jest całkowicie bezstronna?

Bartosz Naskręcki: Mit numer 2. Czy sztuczna inteligencja jest całkowicie bezstronna? No to jest dziwne pytanie, bo jak właściwie o tym myśleć? Jeżeli ja zapytam algorytm, czy woli banany, czy gruszki, tak zapytam Chata GPT, to jak myślicie, czego się możemy spodziewać jako odpowiedzi? No ale algorytm coś nam odpowie, nie? I teraz pytanie jest takie: skąd się jego odpowiedź wzięła? Jak myślicie, że one zostały wytrenowane na obrazach, które już te cywilizacyjne skrzywienia w sobie prezentują, czyli jeżeli chcemy mieć całkowicie bezstronny algorytm, gdybyśmy teraz pomyśleli, czy możemy stworzyć algorytm, który będzie rozstrzygał w sprawach sądowych, to sytuacja jest dość poważna. Bo w jaki sposób zagwarantować, że ten algorytm, który będzie się uczył na podstawie różnego rodzaju spraw karnych, które się wcześniej wydarzyły, będzie obiektywny? On musiałby być tak stworzony, tak zaprojektowany, żeby potrafił oddzielić nasze ludzkie emocje i preferencje i różnego rodzaju uprzedzenia od tego, w jaki sposób interpretować sprawiedliwość. To jest kwestia ludzka, a nie samego algorytmu. Czyli algorytm jest zasadniczo bezstronny. Ale ponieważ karmi się danymi pochodzącymi od nas, bo ma z nami wchodzić w interakcje, więc zawsze w pewnym sensie będzie stronniczy, ale tylko z tego powodu, że to my taki algorytm stworzymy, a nie on sam z siebie taki powstanie. A swoją drogą, jak pomyślicie o ludziach i o ich preferencjach, to skąd się biorą różnego rodzaju preferencje u ludzi, że na przykład jedni wolą banany, a nie gruszki? Czy coś wam przychodzi do głowy, co może powodować, że jedni wolą to, a inni tamto? W sensie, najprostsza odpowiedź byłaby taka, że nasz mózg nam to podpowiada. Ale właściwie skąd się biorą te myśli w naszym mózgu? Czy na przykład to jest tak, że my wszyscy myślimy w podobny sposób, tylko czasami jedni myślą trochę inaczej, to się modyfikuje, zmienia. Warto nad tym się zastanowić. Okej, czyli obiektywność sztucznej inteligencji to jest tak, jakby pytać, czy nasz komputer

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

potrafi być bezstronny. Komputer tak, ale dane, którymi będziemy ten komputer karmić i obsługiwać, już niekoniecznie. A dane to jesteśmy my. No i dalej: inteligencja ona nie jest po naszej stronie czy przeciwko nam. Czyli jeżeli na przykład rozmawiacie z Chatem GPT, to to nie jest tak, że on jest waszym psychoterapeutą i on was dobrze rozumie i czuje wasze emocje. Ten algorytm w zasadzie tylko przetwarza pewne dane, i te dane, które my wprowadzamy i na których trenujemy ten algorytm, one wpływają na to, w jaki sposób to działa. Jak porównacie sobie różne wersje tych czatów, które w ciągu ostatnich dwóch lat powstały, to część z nich nie potrafi rozróżniać wielu aspektów etycznych, tylko dlatego, że nie zostały na takich danych w ogóle wytrenowane. Ale teraz z drugiej strony to jest trochę przerażające, bo jeżeli stworzymy program, który wydawałoby się, że będzie w stanie rozumieć sens przetwarzania danych i pojęcia etyczne, to czy możemy powiedzieć, że ta sztuczna inteligencja ma pewnego rodzaju formę moralności? Czy wyobrażacie sobie taki scenariusz, że siadacie przed komputerem i ten komputer decyduje o tym, czy wasze zachowanie było etyczne? Zazwyczaj robi to człowiek, prawda? Ale człowiek ma pewnego rodzaju doświadczenie, które jest oparte na wiedzy, na pewnych relacjach z innym człowiekiem. Nagle okazuje się, że wydaje się, że możemy powoli stworzyć program, który może emulować formę moralności. Więc pytanie jest takie: czy to już jest moralność? Czy to jest tylko nadal przetwarzanie danych? Mit jest obalony.

MIT 3

Czy sztuczna inteligencja może myśleć i czuć jak ludzie?

Bartosz Naskręcki: Zobaczmy na następny mit numer 3. Czy sztuczna inteligencja może myśleć i czuć jak ludzie? Teraz do was mam znowu pytanie: jeżeli mówiliśmy już wcześniej o myśleniu, a teraz spróbujmy porozmawiać o odczuciach. Jeżeli rozmawiacie z jakąś inną osobą i ta osoba mówi, że ją boli palec, to czy to, co wy odczuwacie, to jest też ból palca, czy jakieś współczucie? Co właściwie możemy poczuć w takim przypadku? Czy na przykład będziecie podejrzewali, że ta osoba zmyśla? Można pomyśleć. A teraz, gdybyście byli postawieni przed sytuacją, że siedzicie przed komputerem i komputer wam mówi, że go coś boli, to uwierzycie temu komputerowi? Jeżeli zapytam algorytm i on mi powie, że go coś boli, to jak myślicie, czy to będzie miało sens? Czy komputer może odczuwać ból? No właśnie, ale to jest kwestia trudna do rozstrzygnięcia, bo jeżeli ktoś mówi wam, inny człowiek, że go boli głowa, to co jesteście w stanie pomyśleć, to to, że ta osoba wygląda jak ja, mnie też kiedyś

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

bolała głowa, więc wiem, co to znaczy ból głowy. Więc zaczynam empatyzować z tą osobą, bo wiem, co to znaczy, że boli mnie głowa, prawda? Ale to nie zawsze musi być tak samo, bo jeżeli na przykład ktoś wam opowiada o tym, jak był na Mount Everest i mówi, jakie to jest niesamowite uczucie wejść na tak wysoką górę, a wy nigdy nie byliście na Mount Everest, albo w ogóle nie byliście w górach, to trudno jest wam sobie wyobrazić uczucia, które w tym momencie ma ta osoba. Czyli często jest tak, że to, co my mówimy i to, co wydaje nam się, że czujemy, albo to, co osoby nam opowiadają, że czują, to jest bardziej projekcja naszych własnych myśli niż jakieś obiektywne przekazanie informacji. I teraz, postawieni przed tym pytaniem, powinniśmy się zastanowić, czy w takim razie, jeżeli mamy algorytm i algorytm w jakimś sensie nam powie, że go coś boli, to w pewnym sensie to może być prawdą. To znaczy wewnętrzny stan tego algorytmu na coś wskazuje, ale to w żaden sposób nie przypomina ludzkiego uczucia. Czyli jeżeli my powiemy algorytmowi: okej, boli mnie głowa, to on nie zrozumie, o co nam chodzi, bo to nie jest jego inteligencja. W sensie, my jesteśmy zupełnie inni, ale to nie znaczy, że ten algorytm nie może czuć. Czyli problem polega na tym, że sztuczne sieci neuronowe powstają w totalnie różny sposób niż powstaje nasza inteligencja. Za chwileczkę do tego wrócimy. Czy to jest forma jakiejś cyfrowej reprezentacji, przesyłu informacji? Ale z drugiej strony, nasz mózg to też jest przetwarzanie informacji, tylko w trochę inny sposób. Czyli teraz możemy sobie naprawdę zadać pytanie: okej, czy algorytm może odczuwać? No i teraz pojawia się z tego pewien kłopot, bo czy możemy w ogóle porównać nasze odczuwanie z odczuwaniem komputera? Prawdopodobnie nie. Te sieci neuronowe nie mogą odczuwać w takim sensie jak my. Zresztą, przed chwilą próbowałem was przekonać, że nawet trudno jest sobie dobrze wyobrazić, co czują inni ludzie, więc tym bardziej trudno sobie wyobrazić, że algorytm odczuwa w takim sensie jak my. Ale to wcale nie jest takie oczywiste, że algorytm nie może mieć pewnego rodzaju myśli, przetwarzania informacji w ogólnym sensie, które coś powodują, na przykład, że ten algorytm będzie się dziwnie zachowywał. Jeszcze jedna rzecz, która jest dosyć istotna, mianowicie kwestia rozwoju naszego intelektu. Każdy z was, który tutaj siedzi na sali, miał swoją mamę, tak? Urodziliście się, mama was urodziła. Jak myślicie, ile lat ma ludzkość? Szacunkowo. Ludzkość jako ludzie, w sensie tak jak my wyglądamy, będzie więcej niż 1000 lat. Ile? 200 000 lat. Czyli ja mam mamę, moja mama miała mamę i tak dalej, i 200 000 lat temu, koniec. No właśnie, ale to teraz, jak to jest? Czyli tak, każda mama ma taki mózg, prawda? Dość podobny. Te mózgi są takie same? Zmieniają się jakoś? Zmieniają się. A teraz, które mózgi przetrwały do dzisiaj? Jak myślicie? Te, które się najlepiej adaptowały do otoczenia, nie? Gdyby się nie adaptowały... Czy to się nazywa ewolucją? Jak popatrzycie na ten obrazek, to te wszystkie rodzaje mózgów, które tu widzicie, pochodzą od różnych człekokształtnych, od różnego rodzaju gatunków małp. I one się ewidentnie

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

różnią, one się zmieniały przez setki tysięcy lat. I teraz ta inteligencja, która z nich się pojawiła, ona powstawała w toku ewolucji, prawda? To nie było tak, że ktoś jednego dnia usiadł i stworzył sobie mózg, po prostu usiadł i jest mózg i wszystko działa. To się pojawiało. A teraz z drugiej strony mamy sytuację taką, że my tworzymy algorytmy, które mają nam emulować czy przypominać inteligencję. I te algorytmy przynajmniej na razie nie są poddane żadnego rodzaju ewolucji takiej klasycznej, w sensie, że nie ma tak, że Chat GPT, żeby dotrwał do jutra, to musi przetrwać jakąś próbę. Jak przetrwa tę próbę, to się zaadaptuje i przekaże jakąś swoją mądrość do kolejnego pokolenia i tak dalej. W ogóle, co to znaczy przekazanie mądrości? W naszym przypadku to są geny, prawda? Geny to jest to, co przekazujemy kolejnemu pokoleniu i to kolejne pokolenie odtwarza z tych genów całe ciało, cały organizm, łącznie z mózgiem. Więc ten mózg się pojawia jako nowa kopia, troszeczkę zmodyfikowana, za każdym razem trochę inna. I zawsze przetrwają te mózgi, które najlepiej sobie radzą. Czyli ewidentnie to, co różni nas od sztucznej inteligencji przede wszystkim, to jest to, że nasz mózg jest świetnie zaadaptowany do środowiska. I to jest rzecz, którą już pewnie zauważyliście, że jak zaczniecie gadać z tą sztuczną inteligencją, to często jest tak, że ona w wielu sytuacjach sobie już poradzi. Ale jeżeli podacie jej dostatecznie dziwną sytuację, która w ogóle nie jest w żadnym sensie opisana w żadnych danych, na których była trenowana, to ta sztuczna inteligencja często może sobie nie poradzić. Ale z drugiej strony, mamy algorytmy, na przykład takie, które grają w gry, na przykład w Go albo w szachy, które świetnie poradziły sobie już z taką genetyczną ewolucją. Były modyfikowane kolejne wersje tych algorytmów i one coraz lepiej grały. Okazuje się, że można to doprowadzić już do takiej sytuacji, że ten algorytm gra lepiej tak naprawdę niż człowiek. No i teraz, jeżeli pomyślimy o tych rzeczach, które już się pojawiają w naszym otoczeniu, to może jest tak, że jesteśmy w stanie w jakimś sensie puścić te algorytmy w taką ewolucję, tak jak myśmy ewoluowali, i one staną się coraz lepsze. To jest w zasadzie pytanie otwarte, więc na razie jeszcze nie znamy na nie pełnej odpowiedzi.

MIT 4

Czy sztuczna inteligencja jest stronnicza i uczciwa?

Bartosz Naskręcki: Mit numer 4. Czy sztuczna inteligencja jest stronnicza lub uczciwa? Już częściowo się to pojawiło, więc zastanówmy się nad tym troszeczkę głębiej. Po pierwsze, sieci neuronowe nie mają w sobie wbudowanego pojęcia prawdy. Nie wiem, czy zdajecie sobie sprawę, ale jak rozmawiacie z Chatem GPT

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

i zapytacie go, ile jest 2+2, i zazwyczaj da wam odpowiedź 4. Tam nie zostało przeprowadzone w żadnym klasycznym sensie rozumowanie takie matematyczne, które właściwie stanowiłoby uzasadnienie tego faktu, że 2+2=4, tylko to jest na zasadzie trochę takiej papugi, która jak już zobaczyła wiele razy, że gdzieś było napisane, że 2+2=4, to najprawdopodobniejsza odpowiedź to jest 4. Ale z drugiej strony, jak zaczniecie to przeskalowywać, to okazuje się, że w pewnym sensie to, co robią algorytmy tego typu, to jest trochę manipulowanie prawdą. Ale to nie jest tak do końca jak w naszym życiu. Czyli jeżeli my próbujemy wykonywać pewne rozumowanie, to my też często mylimy się, ale to nasze mylenie jest zupełnie inne. I tutaj na to trzeba zwrócić uwagę, że w przypadku zwłaszcza dużych modeli językowych, czyli tych algorytmów, które nazywamy teraz głównie sztuczną inteligencją, takich właśnie czatów, tam nie ma czegoś takiego jak pojęcie prawdy. Czyli jeżeli zapytacie ten algorytm, czy w danej lekturze coś się pojawiło, to będzie zanalizowany z punktu widzenia częstości kontekstu pojawiania się tego pojęcia w tym tekście, ale to nie jest żadne rozumowanie w takim sensie klasycznym matematycznym. Jedyna rzecz, która może nam pomóc w tym przypadku, to jest to, że jeżeli poprosimy algorytm o napisanie programu, który ma coś wykonać, to jeżeli ten algorytm będzie poprawnie zaimplementowany, to ten algorytm możemy uruchomić, wykonać na komputerze w taki klasyczny sposób i on da nam odpowiedź. Ale sam Chat GPT tak naprawdę – mówimy, że halucynuje – czyli on, to co nam odpowiada, to jest tylko pewna kwestia przetwarzania języka, czyli tego, co już usłyszał. Czyli jeżeli poprosicie Chat GPT, żeby dał wam receptę na wszystkie problemy ludzkości, to jedyne, co on wam poda, to opisy czy jakieś wskazówki na podstawie tego, na czym został wytrenowany. Czyli to nie jest tak, że jeżeli poprosicie go o świetną rozprawkę na temat "Pana Tadeusza", to dostaniecie tekst, który jest całkowicie oryginalny. Nie, ten tekst będzie tylko przetworzeniem informacji, które już tam były. Czy to jest prawdziwe? To trudno powiedzieć, bo to nie ma nawet sensu. To pytanie, czy to, co sztuczna inteligencja nam daje, często nie ma nic wspólnego z taką klasyczną prawdą rozumianą w sensie nawet Arystotelesa, czyli że rzeczywiście możemy coś weryfikować w rzeczywistości, to jest tylko kwestia pewnego rodzaju kontekstu. I to jest dość niebezpieczne, ponieważ coraz częściej, zwłaszcza jak rozmawiam z różnymi osobami, spotykam się z taką sytuacją, że podczas rozmowy z takim czatem, utwierdzamy się w przekonaniu, że rozmawiamy z osobą, że zapominamy o tym, że to jest dosyć prosty algorytm, który nie miał takiego szczęścia jak my, że ewoluował przez wiele tysięcy lat i został poddany wielu miliardom prób. No, w pewnym sensie został, ale jeszcze nie aż tak wielu i tam nie ma czegoś takiego jak klasyczne pojęcie prawdy, takiego sprawdzania, weryfikowania w rzeczywistości, te algorytmy często tego w ogóle nie mają. No i teraz problemy ze stronniczością, czyli to, co dostajemy jako odpowiedzi, to jest znowu kwestia danych. Jeżeli chcecie, żeby algorytm

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

działał w sposób, którego się spodziewacie, dobry, jakościowo dobry, to wszystko jest ukryte w tych danych. I teraz oczywiście dane, to nie znaczy, że musimy coś podać algorytmowi książkę i tak dalej. Możemy stworzyć osobny algorytm, który będzie tworzył sztuczne dane. Dokładnie w ten sposób działa algorytm, który gra w grę Go, który został wytrenowany na całkowicie sztucznych partiach. On eksplorował, w jaki sposób grać w grę, żeby nauczyć się optymalnych ruchów. No i w tym sensie, to były dla niego dane. Jeden z najnowszych algorytmów, który rozwiązuje zadania matematyczne z geometrii europejskiej, bardzo trudne zadania zresztą, on działa w ten sposób, że nauczył się, został wytrenowany za pomocą sztucznych danych do tego, żeby rozwiązywać problemy matematyczne. I teraz, czy on jest matematykiem? Wydawałoby się, że w pewnym sensie jest, ale to nie jest taki matematyk jak człowiek, który potrafi odczuwać, ma pewne intuicje, ma pewne uprzedzenia, ma cały bogaty kontekst kulturowy. To jest tylko nadal dość prosty algorytm, który wykonuje dość skomplikowane, ale bardzo precyzyjnie postawione pytania. No i teraz kwestia prawdy, czyli dochodzenia do jakichś obiektywnych faktów, udowodnienia czegoś, pokazania czegoś. To znaczy, nasze emocje, nasze intuicje, nasze wyobrażenia, intelekt, to może być całkowicie oderwane od człowieka. Maszyna, algorytm, który wykonuje jasno określone zadania, ale nie ma w tym takiego bogactwa kulturowego, które mamy my. I tutaj to jest jeden z problemów, z którymi się często spotykam, że jak ktoś myśli właśnie o inteligencjach, to za każdym razem próbujemy przykładać ludzkie miary. Na przykład kwestia świadomości jest taką dosyć istotną sprawą. Czy algorytm może być czegoś świadomy? A czy to znaczy, że my jesteśmy świadomi? Świadomość w pewnym sensie jest nadal niepoznana, my nie wiemy do końca, czym jest świadomość, więc w jaki sposób możemy spróbować policzyć, sprawdzić, że na przykład Chat GPT jest świadomy? Jak zapytamy go, czy jest świadomy, to coś odpowie, ale czy to cokolwiek oznacza? Trudno powiedzieć.

MIT 5

Czy dążenie do sztucznej inteligencji ogólnej (AGI) i superinteligencji doprowadzi do zagłady ludzkości?

Bartosz Naskręcki: No i ostatni chyba temat już, wydaje mi się, dosyć nośne pytanie. Mit numer 5. Czy dążenie do sztucznej inteligencji ogólnej i superinteligencji doprowadzi do zagłady ludzkości? Po pierwsze, rozbierzmy to pytanie na części: sztuczna inteligencja ogólna – co to jest? W zasadzie nikt nie wie, bo nikt tak naprawdę do końca nie wie, czym jest taka uniwersalna inteligencja, na przykład jak my. Swoją drogą, jak pomyślimy o rodzajach

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

inteligencji, z którymi się spotkaliśmy, to nie mamy zbyt dużo przykładów inteligencji naturalnych, które gdzieś możemy zaobserwować we wszechświecie. Generalnie większość inteligencji, o których wiemy cokolwiek, sprowadza się tylko do naszej planety, więc bardzo trudne jest powiedzenie, co to w ogóle może być sztuczna inteligencja ogólna. Czy to jest algorytm, który potrafi sobie ze wszystkim poradzić? No a jeżeli taki algorytm, powiedzmy, będzie miał zdecydować na przykład o życiu wszystkich ludzi, to co zrobi w takim przypadku? Jak to w ogóle zaprogramować? To jest bardzo trudne pytanie. Superinteligencja to jest w ogóle takie pojęcie, które jeszcze wykracza poza to. Czy dojdziemy za chwilę do sytuacji, że będziemy mieli algorytmy, które będą w stanie nie tylko rozumować tak jak ludzie, ale robić dużo więcej niż ludzie? I teraz, czy one będą na przykład obdarzone moralnością, czy będą w stanie na przykład wchodzić z nami w interakcje w takim sensie, jak my wchodzimy ze sobą, czy będą nas bardziej traktować jako mrówki? Stąd się bierze to pytanie, bo łatwo sobie wyobrazić teraz taki scenariusz, że jeżeli rzeczywiście coś takiego stworzymy, to może to rzeczywiście doprowadzić do zagłady ludzkości, w sensie, że staniemy się dla tych algorytmów jacyś bardzo słabi i one przejmą nad naszym życiem władzę i na przykład postanowią nas zlikwidować. Ja powiem tak: po pierwsze, tak jak już wspomniałem, my nadal w zasadzie nie wiemy, czym jest inteligencja, więc w sumie to wszystko, co robimy związane ze sztuczną inteligencją, w zasadzie powinno nam pomóc, ponieważ próbujemy różnego rodzaju metodami trochę lepiej zrozumieć, czym jest inteligencja. Ale tutaj wydaje mi się też istotna kwestia, że ludzkość tak naprawdę nie potrzebuje sztucznej inteligencji, aby doprowadzić się do krawędzi zagłady. Jak tutaj widzimy jakieś susze, te problemy, to trochę jest jak z taką bajką, którą mój syn ogląda, nazywa się Psi Patrol. Najpierw robi się niezłą grandę, a potem próbuje się to naprawić. To w pewnym sensie ludzkość jest na tym etapie właśnie, że doprowadziliśmy naszą cywilizację do takiego etapu, że potrafimy popsuć środowisko. I teraz zastanawiamy się, czy możemy to naprawić. Sztucznej inteligencji nie potrzebujemy do tego, żeby popsuć środowisko, ale być może będziemy potrzebowali sztucznej inteligencji do tego, żeby środowisko naprawić. Tutaj pojawia się dodatkowa dziwna myśl, mianowicie: czy to nie jest tak, że jak te inteligencje powstają, to też nie jest forma ewolucji? W pewnym sensie może to jest kolejny etap ewolucji, że teraz, jeżeli powstanie jakaś kolejna inteligencja, to coś się wydarzy, powstanie osobny gatunek. Trudno powiedzieć, to jest na tyle nowe, że właściwie nie mamy żadnych przykładów w przyrodzie, żeby coś takiego się działo, że jakaś naturalna inteligencja stworzyła odrębne byty. No i teraz pytanie, co się dalej stanie? Myślę, że więcej jest tutaj pozytywów niż negatywów na razie, z tego powodu, że mamy mnóstwo problemów, gdzie te algorytmy, nawet w ograniczonym sensie, takie które nie do końca są w stanie rozumować w pełni tak jak my, ale mogą nam na pewno pomóc, zwłaszcza ze

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska

ul. Rubież 46, 61-612 Poznań
Sekretariat: +48 61 827 97 00
e-mail: ppnt@ppnt.poznan.pl
www.ppnt.poznan.pl

względu na to, że są w stanie przetwarzać olbrzymie ilości danych. Czy sztuczna inteligencja przejmie władzę nad światem? Raczej bym się tym nie przejmował. Bardziej bym się przejmował tym, czy my potrafimy wykorzystać algorytmy i te różne sztuczne inteligencje do tego, żeby nasze życie było lepsze. I tutaj wydaje mi się, że jest zdecydowanie więcej pozytywów niż negatywów. Oczywiście, takie wizje katastroficzne, że jakiś algorytm zapanuje nad nami, można roztaczać, ale tak jak już wcześniej powiedziałem, w zasadzie algorytmy rządzą naszym życiem do pewnego stopnia, więc my nawet nie potrzebujemy wyrafinowanej sztucznej inteligencji, bo my sami jako cywilizacja programujemy się do pewnego rodzaju takich sytuacji. Ale na pewno to, co możemy stwierdzić, to to, że jeżeli nie będziemy eksplorować tego tematu, jeżeli wy nie będziecie zadawać sobie kolejnych pytań, czy ja jestem w stanie poprawić swoje życie za pomocą algorytmów, czy ja jestem w stanie rozwiązać jakieś realne problemy ludzkości za pomocą algorytmów, to na pewno ludzkość nie będzie się dalej rozwijała. Jak widzimy, największe osiągnięcia mamy tam, gdzie ludzie zadają sobie trudne pytania i próbują na nie odpowiedzieć. I sztuczna inteligencja to jest kolejna dziedzina, gdzie jesteśmy w stanie zadawać trudne pytania i próbować na nie odpowiadać. I co ciekawe, uzyskujemy coraz lepsze efekty, więc wydaje mi się, że raczej nie zagraża nam przynajmniej w najbliższym czasie żadna forma opresji ze strony sztucznej inteligencji. Natomiast możemy sobie w życiu pomóc i wydaje mi się, że tym się trzeba kierować. Ja myślę, że to wszystko, dziękuję wam serdecznie za uwagę.

---KONIEC---

Ministerstwo Nauki i Szkolnictwa Wyższego
Społeczna Odpowiedzialność Nauki
Projekt dofinansowany ze środków budżetu państwa, przyznanych przez Ministra Nauki i Szkolnictwa Wyższego w ramach Programu Społeczna odpowiedzialność nauki II – moduł: popularyzacja nauki. Dofinansowanie z MNiSW: 404 800,00 PLN. Całkowita wartość projektu: 452 350,00 PLN
Rzeczpospolita Polska