Czy mamy się bać sztucznej inteligencji?

Wywiad Vadima Makarenko z **Katarzyną Szymielewicz**, szefową Fundacji Panoptykon

27 lutego 2021 | 08:55

https://wyborcza.biz/biznes/7,177150,26819148,czy-mamy-sie-bac-sztucznej-inteligencji.html

O tym, czy pandemia zwiększa nasze lęki przed sztuczną inteligencją i czy nasza wizja życia z nią nie jest błędna, mówi Katarzyna Szymielewicz, szefowa Fundacji Panoptykon**, która opublikowała opracowanie "Sztuczna inteligencja non-fiction".*

Vadim Makarenko: Zacznijmy od artykułu, który napisała sztuczna inteligencja w brytyjskim dzienniku „The Guardian". Zostałem zasypany pytaniami, czy się boję o moją przyszłość. Nie boję się, ale chciałbym poznać twoją opinię – czy mam się bać?

Katarzyna Szymielewicz: Myślę, że jest to świetne pytanie dla piszących. Oboje piszemy. Ja też się nie boję, ponieważ mam kontakty z wytworami sztucznej inteligencji: i tekstowymi, i z bootami, które rozmawiają z nami na infoliniach.

Nie ma wątpliwości, że ta inteligencja potrzebuje jeszcze wielu lat treningu, żeby napisać coś, co ma sens, zrobić jakąś podstawową kompilację wiadomości czy opracować bardzo prostą depeszę.

Podejrzewam, że możemy się spodziewać w ciągu pięciu do dziesięciu lat takiego postępu techniki analizy tekstu, że takie depesze czy skróty z konferencji, wywiadów pisane przez AI (artificial intelligence) będą miały sens.

Natomiast napisanie prawdziwego tekstu – który zawiera opinię, analizuje świat – przez AI długo nam jeszcze nie grozi.

Jest taka popularna opinia, że sztuczna inteligencja zabierze nam pracę. Jednak śmiem twierdzić, że sztuczna inteligencja – aby nam zagrozić – musi się nauczyć jeszcze wielu rzeczy, np. zbierania informacji, ich przetwarzania.

– To jest świetny przykład pokazujący ograniczenia tej technologii. Myślenie o AI w kontekście pracy – czyli zestawu rozmaitych zadań, które wykonuje człowiek – jest nieporozumieniem. AI może być użyta do przetwarzania informacji, analizy informacji – wszędzie tam, gdzie występują jakieś wzorce. Wzorce, które się nie zmieniają w czasie, które da się zapisać w postaci statystycznych reguł i kodu. Więc np. analiza tekstu spisanego z długaśnej konferencji pod kątem słów kluczowych – czy padł tam COVID, bo szukamy informacji o COVID-zie – to świetne zadanie dla AI.

Natomiast analiza sytuacji, którą wykonuje polityk, lekarz, dziennikarz – każdy, kto próbuje zrozumieć świat i coś z tego wyciągnąć – nie jest zadaniem dla AI, ponieważ takich funkcji nie można wpisać w kod, którym posługują się pracujące systemy. Jeśli coś nas czeka, to myślę, że zastąpienie nas w pewnego typu zadaniach. I mogę sobie wyobrazić, że dziennikarz korzysta z analizy tekstu, którą wykonuje dla niego taka technologia, ale już nie korzysta z zastępstwa. Takiego zastępstwa po prostu nie będzie.

Zabranie nam pracy to popularna narracja. Kolejna – to narracja o wspaniałym pomocniku. Wokół AI jest sporo zachwytów, a oto jedne z nich: „Sztuczna inteligencja pomaga w wynalezieniu np. szczepionki na koronawirusa". Czy zachwyty są bardziej uzasadnione niż lęki?

– Po stronie zarówno lęków, jak i zachwytów pokutują pewne mity o sprawczości AI, o jej niezależności od człowieka, o jej omnipotencji w generowaniu wiedzy.

AI – tak jak funkcjonuje dziś – to po prostu wyrafinowana forma analizy statystycznej. Mówimy o matematyce i statystyce, a nie o humanistyce. Nie mówimy o jakieś nauce i rezultatach badań, jakie uzyskuje człowiek, który krytycznie myśli, ma świadomość, ma wolną wolę i dokonuje ocen moralnych. W tej sferze AI z definicji nie może funkcjonować.

Jeżeli odróżnimy mity od prawdy i skupimy się na tym, co jest możliwe dzięki analizie statystycznej i dzięki przetwarzaniu ogromnych ilości danych, to zaczniemy rozumieć, gdzie ta AI może nam pomóc, a gdzie nie bardzo.

Jeśli chodzi o szczepionkę na COVID, to AI nie wymyśli dla nas sposobu postępowania z pandemicznym wyzwaniem, nie zdefiniuje problemu, nie wymyśli nowych badań. Coś takiego może zrobić tylko człowiek – a wręcz grupa ludzi – z odpowiednim doświadczeniem i olbrzymią wiedzą kontekstową. Ale kiedy już cel jest postawiony, załóżmy, że w postaci badania konkretnych próbek genetycznych – więc mamy takich próbek miliony – i coś musi je dla nas przetworzyć, to to jest świetne zadanie dla AI. Myślę, że w takim zakresie AI może być użyteczna i może przyspieszyć tempo prac.

Jednak AI nie zastąpi ekipy naukowców, która wymyśliła, że będziemy pracować nad szczepionkami taką, a nie inną metodą. Ta ekipa jest odpowiedzialna za etyczną analizę, za testy na ludziach i za ich zaprojektowanie.

Tak jak dziennikarstwo nie polega wyłącznie na pisaniu tekstów, tak samo prace nad szczepionkami mają dużo odcinków. Załóżmy, że sztuczna inteligencja pojawi się kiedyś na wszystkich tych odcinkach, czyli – w przypadku dziennikarstwa – będzie potrafiła ocenić wiarygodność rozmówcy, wiarygodność dokumentu, wybrać kluczowe cytaty z tego dokumentu do tekstu, który potem ma napisać. To możliwe?

– Arvind Narayanan z Uniwersytetu Princeton rozróżnia trzy sposoby użycia AI.

Po pierwsze, AI może wspierać percepcję. I to jest najłatwiejsze do uzyskania – jesteś dziennikarzem, który prowadzi śledztwo, i chciałbyś sprawdzić, czy jakaś broń, którą wykorzysta mocarstwo X, przejeżdżała przez teren Y. Masz tysiące albo miliony zdjęć z internetu. Świetne zadanie dla AI, ona będzie wspierała twoją percepcję – podobnie jak wspiera percepcję lekarza, który szuka śladów raka na wynikach skanu ludzkiego ciała. A w kontekście szczepionek AI pewnie może wspierać genetyków w analizie kodów itd. Tutaj nie ma kontrowersji – powinniśmy inwestować w AI, żeby stawała się coraz lepsza.

Drugi przypadek to jest automatyzowanie oceny. Chcemy ocenić, jaka jest sytuacja, np. chciałbyś ocenić, czy rozmówca kłamie albo czy ulica jest niebezpieczna. Co się na niej dzieje – bójka czy taniec? Ludzie biegną dla sportu czy uciekają? Ocena jest trudna, bo wymaga znajomości kontekstu. Człowiek, który wykonuje zawód sędziego, prawnika, dziennikarza, etyka powinien się czuć dość bezpiecznie, bo liczba kontekstów, które musiałaby przetworzyć maszyna, jest trudna do opisania.

Internet podsuwa nam mnóstwo przykładów, w których łatwo zacierają się granice między mową nienawiści a krytyką. AI jest w tym kontekście wykorzystywana i myślę, że niestety będzie. Nie do końca ufam w jej zdolności, ponieważ poruszamy się na polu, którego nie da się zakląć we wzorce matematyczne i w reguły statystyczne.

I trzeci poziom to przewidywanie tego, co się wydarzy. To najtrudniejsze i – zdaniem Arvinda Narayanana z Princeton – w zasadzie niemożliwe. Chyba że posługujemy się stereotypami i mamy na to zgodę. Uważam, że społeczeństwo powinno się naradzić samo ze sobą, jak chce funkcjonować. Czy chcemy działać wedle stereotypów, które sobie opracowaliśmy, i zakodować AI, żeby pomogła nam uruchomić system szybkiego reagowania na ulicy, w sklepie, w szpitalu, gdziekolwiek? To jest potężny dylemat etyczny i polityczny, dlatego w tej sferze przewidywania zachowań jest chyba najtrudniej.

Jest polski start-up, który – w niebezpieczny dla ludzi sposób – eksperymentuje z rozpoznawaniem twarzy. Jak rozmawiać o tym z ludźmi?

– Przypominają mi się dwa przykłady. Pewnie wszyscy pamiętamy modę na aplikację, która postarzała twarze ludzi – FaceApp. Ludzie wchodzili w to masowo – z ciekawości, z chęci zagrania w coś. Przekazywali swoje zdjęcia do tej obróbki, być może bez refleksji nad tym, komu je przekazują i co może dalej się z nimi wydarzyć. Nasze zdjęcia służyły trenowaniu silnika sztucznej inteligencji, który potem można zastosować do dowolnej innej aktywności.

Firma, która wytrenowała swój algorytm na dziesiątkach milionów twarzy z całego świata, może sprzedać go policji, żeby wspierał działania śledcze, może sprzedać go armii albo Facebookowi, który wykorzysta go do tego, żeby tagować nas na zdjęciach.

W pandemii rządy również sięgnęły po rozmaite techniki analizy danych. Na przykład rząd Polski zastosował prosty algorytm do porównywania zdjęć, które ludzie wrzucają do obowiązkowej aplikacji Kwarantanna domowa. Chodzi o osoby, które są poddane kwarantannie i muszą wrzucać parę razy dziennie swoje zdjęcie, które system porównuje ze zdjęciem wgranym do rządowej bazy danych.

I odbiór takiego działania jest zupełnie inny niż w przypadku korzystania z aplikacji postarzającej twarze. Jako Panoptykon dostaliśmy sporo sygnałów od zaniepokojonych ludzi: co to w zasadzie jest? Czy te dane są bezpieczne? Sama stawiałam takie pytania, czy czasami rząd nie będzie chciał wykorzystać kwarantanny, żeby potrenować jakieś algorytmy, a potem użyć ich np. do analizy zdjęć demonstrantów albo analizy twarzy obywateli w jakimś innym kontekście. Byliśmy uspokajani, że nie, że są tam rozmaite zabezpieczenia. Pokazuję różnicę, kontekst zmienia wszystko – kontekst zabawy, jakiegoś żartu internetowego a kontekst władzy publicznej, która sięga po podobną technologię, to zupełnie coś innego. To, że ktoś wytrenuje fragment kodu na zadaniu pozornie banalnym, nie znaczy, że ten sam kod nie zaistnieje w sytuacji dla nas niebanalnej, a nawet groźnej.

Do tego, żeby ocenić przydatność danej zabawki, technologii, aplikacji do innych „tajemniczych" celów, potrzeba nieraz doświadczenia i dużej wyobraźni.

– Dlatego zdecydowaliśmy się na publikację obszernego przewodnika „Sztuczna inteligencja non-fiction" dla ludzi, których to interesuje i chcieliby przełamać swoje ograniczenia wynikające z braku wiedzy czy braku rozumienia pewnych zależności. To nie jest coś, czego uczono w szkole, to nie jest coś, czego możemy się uczyć z mediów, to nie jest coś, czego uczą nas firmy, które te systemy wykorzystują.

W marketingu dostrzegam dwie tendencje. Pierwsza polega na wprowadzaniu w błąd: „coś" – lodówka, lalka, szminka – nie ma nic wspólnego z AI, ale twórcy takiego rozwiązania mówią o swoim produkcie „smart", „inteligentne", „automatyczne". Druga tendencja jest odwrotna: „coś" ma w środku AI. To coś jest np. Facebookiem, który filtruje treści w sposób, który ma wywołać pewien efekt w naszym mózgu. Ma nas zaciekawić, utrzymać naszą uwagę dłużej, niż może byśmy chcieli, pobudzić nasze emocje. Ma w środku AI, ale nie chce o tym mówić, żeby nas nie stresować i nie stracić naszego zaufania, ujawniając gry, które z nami prowadzi. Takich przypadków wykorzystania AI jest sporo.

W tym momencie faktycznie – jako odbiorcy – mamy prawo być zagubieni. I dlatego takie próby jak nasza, żeby pewne pojęcia porządkować, pewne prawidła w tej dziedzinie wiedzy tłumaczyć. Jednak nadal brakuje nam wiedzy o tym, co się dzieje na rynku. Możemy tylko spekulować na temat tego, co kto ma i czy na pewno Izrael, Chiny i USA mają „tego" najwięcej, a polski rząd „tego" jeszcze nie kupił, bo nie ma pieniędzy.

*Katarzyna Szymielewicz – prawniczka, założycielka i szefowa Fundacji Panoptykon, która opublikowała opracowanie „Sztuczna inteligencja non-fiction”.

**Fundacja Panoptykon – polska organizacja pozarządowa, której celem jest ochrona podstawowych wolności wobec zagrożeń związanych z rozwojem współczesnych technik nadzoru nad społeczeństwem. Fundacja została założona 17 kwietnia 2009 roku.