



Sztuczna inteligencja – wyścig o naszą wolność



JACEK KAWALEC

Wiceprezes Zarządu, VoiceLab

Była już era elektryczności, była era komputerów i internetu, a dziś wchodzimy w nową – erę sztucznej inteligencji. Stawką globalnego wyścigu o prym w tym obszarze jest tak naprawdę nasza wolność. To, czy będziemy widzieli świat naszymi oczami, czy też oczami marketingowców największych koncernów technologicznych. To, czy nasze wypowiedzi w sieci nie będą cenzurowane lub usuwane. To, czy algorytmy należące do potentatów pokroju Google'a nie będą wiedziały o nas więcej niż my sami. Dlatego też, chcąc walczyć o naszą niezależność – ale też i korzyści biznesowe – powinniśmy dbać o to, by rozwijać nasze własne rozwiązania oraz bazy danych w obszarze AI. Czy stać nas na to na Pomorzu?

Rozmowę prowadzi Marcin Wandałowski, redaktor prowadzący Pomorskiego Przeglądu Gospodarczego.

VoiceLab jest jednym z pomorskich pionierów w zakresie rozwijania projektów w obszarze sztucznej inteligencji. Czym konkretnie się zajmujecie?

Firma powstała około 10 lat temu i od początku działalności zajmuje się przetwarzaniem mowy. Stworzyliśmy system, który przy wykorzystaniu sztucznej inteligencji zamienia dźwięk na tekst. Nie chodzi tu jednak wyłącznie o zamianę pliku audio na ciąg znaków tekstowych. Nasze rozwiązanie zostało zaprojektowane w taki sposób, że rozumie sens wypowiedzi – co dana osoba miała na myśli, czy wypowiedź miała charakter pozytywny czy negatywny itd.

W efekcie uzyskujemy rezultat zbliżony do tego, jak dane słowa zrozumiałby i zinterpretował człowiek. Oczywiście – maszyna nie ma ludzkiego mózgu i jej „rozumienie” też jest inne. Wszystko bazuje na nauczaniu systemu – „nakarmieniu” go bardzo dużą ilością danych, dzięki którym sam potrafi wysnuwać własne wnioski.

Kluczowym elementem koniecznym do rozwijania projektów AI jest więc dostępność danych?

Tak – sztuczna inteligencja, a właściwie część tego obszaru zwana uczeniem maszynowym, faktycznie opiera się na „uczeniu” systemu za pomocą bardzo dużych ilości danych. Dajemy mu pewne dane wyjściowe – w naszym wypadku: surowe pliki audio, opisane przez transkrypcję, gdzie każdy dźwięk ma przypisaną odpowiadającą mu literę. Mówimy komputerowi: „jeśli słyszysz taki dźwięk, to on ma taką reprezentację, a tamten ma inną”. Żeby to elastycznie działało, potrzebnych jest bardzo wiele przykładów. My mamy do dyspozycji ponad 10 tys. godzin nagrań, a w wypadku Google mogą to być setki tysięcy, a nawet miliony godzin.

Często jednak – przynajmniej w Polsce – niewystarczającą ilość danych wskazuje się jako jedną z barier rozwoju AI...

Wydaje mi się, że w tym temacie panuje pewne ogólne niezrozumienie – wiele osób mówi przecież, że mamy dziś w sieci bardzo dużo darmowych danych z otwartych zbiorów, które moglibyśmy wykorzystać do uczenia systemów. Danych faktycznie jest sporo, natomiast problem w tym, że większość z nich nie jest użytecznych: nie są to dane uczące – czyli odpowiednio etykietowane, opisane. Nawiązując do naszej branży – brak jest, chociażby, otwartych zbiorów z plikami audio i ich transkrypcją. Takie zbiory można oczywiście zakupić – głównie w języku angielskim. Odmowną barierą jest jednak cena. Chcąc kupić godzinną, przetranskrybowaną rozmowę muszę się liczyć z wydatkiem rzędu 250 dolarów – precyzyjna transkrypcja takiej rozmowy zajmuje bowiem 5-7 godzin pracy człowieka. W dodatku za ceną niekoniecznie idzie jakość – często są to dane „symulowane”, przedstawiające rozmowę opartą na gotowym skrypcie, a nie „naturalną”. System nauczy się z nich mniej – najlepiej, gdy plik audio dotyczy normalnej, autentycznej interakcji międzyludzkiej, a nie dwóch osób, które czytają tekst z kartek.

“ **Wiele osób mówi, że mamy dziś w sieci bardzo dużo darmowych danych z otwartych zbiorów, które moglibyśmy wykorzystać do uczenia systemów. Danych faktycznie jest sporo, natomiast sęk w tym, że większość z nich nie jest użytecznych: nie są to dane uczące – czyli odpowiednio etykietowane, opisane.**

To jednak niejedynie bariery. Chcąc dobrze zaprojektować system, potrzebujemy reprezentatywnej próbki głosów ludzkich, rozłożonych według płci, wieku itd. Gdyby to nie było konieczne, moglibyśmy przecież nagrać i przetranskrybować dziesiątki godzin audycji radiowych czy telewizyjnych – występują w nich często te same głosy: prezenterów, ekspertów, redaktorów prowadzących. Jest to jednak zbyt mała różnorodność.

Warto w tym wszystkim mieć jednak na uwadze, że dane audio są chyba dość szczególnym, „wrażliwym” typem danych przez sam fakt tego, że nie można ot, tak nagrywać sobie prywatnych rozmów, a następnie transkrybować ich i przekazywać komputerowi.

Systemy AI mogą się też przecież uczyć z innego typu danych, np. tekstowych czy obrazowych. Tu również pojawia się problem?

Niestety tak. A przecież wiele danych jest na wyciągnięcie ręki. Moim zdaniem najbardziej potrzebne, a zarazem relatywnie proste w zdobyciu są dane medyczne – np. zdjęcia z tomografu czy rezonansu magnetycznego. Większość państw w Europie zbiera dziś takie właśnie, zanonimizowane dane, które są odpowiednio opisane: „tu jest serce, tu są płuca, tu jest wątroba. Na tym zdjęciu jest zdrowy organ, a na tym organ zaatakowany przez nowotwór”. Kiedy komputer otrzyma kilkanaście tysięcy takich zdjęć, nauczy się samemu wskazywać, czy pacjent jest zdrowy, czy chory. Nie chodzi o to, by technologia zastąpiła radiologów, lecz by była dla nich wsparciem, szczególnie w sytuacji nagłej, dynamicznej, gdy np. do szpitala jest przywożony pacjent wypadkowy z licznymi obrażeniami, który potrzebuje szybkiej pomocy medycznej. Takie rozwiązanie znacznie zmniejszyłoby ryzyko pominięcia przez lekarzy jakiegoś złamania czy innego typu urazu.

Rozumiem, że jest to jeden z przykładów pokazujący generalny problem ze zbieraniem i udostępnianiem danych. Jest on tym silniejszy, gdyż każdy projekt AI wymaga tak naprawdę określonego, dedykowanego mu zbioru informacji. Czy Pana zdaniem przedsiębiorstwa powinny uzyskiwać wsparcie w ich gromadzeniu?

VoiceLab jest dobrym przykładem firmy gromadzącej i kupującej wiele danych. Bierzymy udział w różnego typu grantach i projektach, udostępniających nam finansowanie na ten cel. Niemniej jednak uważam, że nie jest to dobry kierunek – jak wspominałem, dane są bardzo drogie. Moim zdaniem potrzebne byłyby swego rodzaju publiczne banki danych, za którymi stałyby organizacje niekomercyjne. W Stanach Zjednoczonych taką funkcję pełnią uczelnie wyższe. Pozostając w warunkach rynkowych, będziemy skazani na zaporowe ceny danych, co będzie nadal stanowiło wysoki próg wejścia do AI.

Bardziej jednak niż tworzenie instytucji przydałaby nam się ogólna zmiana filozofii, *mindsetu*, świadomości w zakresie zdobywania danych uczących. Najlepszy przykład? Na wielu stronach internetowych administracji państwowej zainstalowane są boty, sprawdzające, czy użytkownik jest „realny”, czy nie. Zna to każdy z nas – to obrazki, na których jesteśmy proszeni o zaznaczenie roweru, sygnalizacji świetlnej czy przejścia dla pieszych. Od strony systemu: są to dokładnie dane uczące, za pomocą których komputer jest uczony lokalizowania danego obiektu na obrazku. Zmierzam do tego, że – przynajmniej na stronach państwowych – zamiast korzystać z usług firmy Google, powinniśmy używać własnego narzędzia pozwalającego na zbieranie danych i budować swoje własne „magazyny danych”. Podobnego typu inspiracji do zmian mógłbym oczywiście przytoczyć wiele więcej.

“ **Przydałaby nam się dziś w Polsce ogólna zmiana filozofii, *mindsetu*, świadomości w zakresie zdobywania danych uczących dla potrzeb AI.**

Wróćmy do działalności VoiceLab. Jakie produkty stworzone na bazie systemu rozpoznania mowy oferujecie swoim klientom?

Zaprojektowaliśmy szereg aplikacji skierowanych głównie do biznesu. Najważniejsze z nich dotyczą wsparcia interakcji głosowej konsultantów firmy z klientami – czy sprzedawcy nie mówią za dużo, za szybko, co powinni powiedzieć w danym momencie itd. – oraz automatyzacji pewnych stałych procesów. Dla przykładu: do niektórych firm wielu klientów dzwoni w tej samej sprawie i przebieg tych rozmów jest mniej więcej taki sam. Jeśli struktura rozmowy jest powtarzalna, możemy ją zautomatyzować przy wykorzystaniu voicebota lub chatbota. Jesteśmy jedną z pierwszych firm w Polsce, która oferuje klientom wykorzystanie obydwu tych technologii. Naszym rozwiązaniem jest też m.in. automatyczny scoring rozmów telefonicznych – system VoiceLab zastępuje człowieka w ocenie, w jaki sposób sprzedawca rozmawiał z klientem, o co się pytał, czy uzyskał zgody marketingowe, jaka była ogólna atmosfera rozmowy. Dla naszych klientów to potężna oszczędność czasu i pieniędzy.

Transkrybujemy też wszelkie medialne programy informacyjne, także archiwalne. Mamy nagrania radiowe z ostatnich lat, które dzięki naszemu rozwiązaniu możemy przetworzyć i odszukać w nich kluczowe słowa: nazwiska, tematy, miejsca. Opracowaliśmy także tzw. biometrię głosową, dzięki której system jest w stanie wyszukać w archiwum nagrań wypowiedzi danej osoby, odnajdując ją po barwie głosu, a nie po nazwisku.

Jakiego typu przedsiębiorstwa korzystają z Waszych usług?

W kontekście analizy rozmów mamy wielu klientów głównie z branży ubezpieczeniowej oraz bankowej. Wykorzystując możliwości automatycznej transkrypcji audio, bierzemy też udział w dużym projekcie transkrypcji rozpraw sądowych. Mamy klientów zarówno w Polsce, jak i za granicą – w zasadzie w ciągu kilku miesięcy jesteśmy wdroyć w nasze rozwiązanie każdy język na świecie. Wszystko w zależności od dostępności danych. W tym momencie system wykorzystuje język polski, angielski, niemiecki, rosyjski i włoski.

W jaki sposób rywalizujecie z konkurencją? Przecież podobnego typu rozwiązania, pozwalające przetwarzać mowę na tekst, oferuje nawet Google...

Staramy się wyszukiwać nisze – tak technologiczne, jak również rynkowe. Odnośnie tych pierwszych – Google wyspecjalizowało się w rozwiązaniu działającym w częstotliwości 16 kHz, czyli w jakości np. dyktowania tekstu na telefon czy obsługiwanie inteligentnych asystentów głosowych typu Siri. My natomiast skupiamy się na jakości 8 kHz, odpowiadającej rozmowie telefonicznej, gdzie rozmówca słyszy przez słuchawkę skompresowany głos drugiej osoby.

Znacznie ważniejsze od nisz technologicznych są dla nas nisze rynkowe, związane np. z językami nieopłacalnymi dla gigantów – a do takich należą np. polski czy włoski. Naszym atutem jako firmy o nieporównanie mniejszej skali niż globalne koncerny jest też elastyczność – bez problemu dostarczymy dedykowany system pod specyficzne wymagania danego klienta.

“ Staramy się wyszukiwać nisze – tak technologiczne, jak również rynkowe. Zdecydowanie ważniejsze są dla nas te drugie, związane np. z językami nieoptycznymi dla gigantów – a do takich należą chociażby polski czy włoski.

VoiceLab jest inicjatorem pierwszego na Pomorzu, a jednego z pierwszych w Polsce Hubu Innowacji Cyfrowych (DIH – *Digital Innovation Hub*) – di4.ai. Na czym polega ta inicjatywa?

Koncept DIH jest inicjatywą Ministerstwa Rozwoju, które zaplanowało w 2019 r. stworzenie lokalnych hubów pomagających przejść przedsiębiorstwom przez proces cyfryzacji. Jako di4.ai stanowimy miejsce, w którym przedsiębiorcy – w szczególności mali i średni – uzyskują wsparcie w zakresie m.in. zrozumienia samej idei cyfryzacji, narzędzi potrzebnych do jej wdrożenia, opracowywania strategii cyfryzacji, poszukiwania odpowiednich dostawców rozwiązań, szkoleń dla pracowników itd. Wszystko to ma na celu, by finalnie przedsiębiorcy byli w stanie wykorzystywać w swojej działalności najnowsze technologie.

“ Di4.ai stanowi miejsce, w którym przedsiębiorcy – w szczególności mali i średni – uzyskują wsparcie w zakresie m.in. zrozumienia samej idei cyfryzacji, narzędzi potrzebnych do jej wdrożenia czy opracowywania strategii cyfryzacyjnych. Wszystko to ma na celu, by finalnie byli w stanie wykorzystywać w swojej działalności najnowsze technologie.

W ramach di4.ai współpracujemy blisko m.in. z Politechniką Gdańską oraz Gdańskim Parkiem Naukowo-Technologicznym. Dzięki naszym wspólnym wysiłkom otworzyliśmy w Parku robotyczne centrum prototypowania – Space4Makers. Centrum jest wyposażone m.in. w specjalne komputery z projektorami, urządzenia do pomiaru, drukarki 3D, etc. Umożliwia to zbudowanie od początku do końca prototypu danego urządzenia. To przykład działań, które chcemy realizować. Owszem – organizujemy też szkolenia, doradztwo dla firm czy hackathony, ale zależy nam, by nasze wsparcie miało jak najbardziej praktyczny wymiar. Nie chcemy być ośrodkiem skupiającym się wyłącznie na doradzaniu, lecz być *learning factory* – miejscem, w którym „na żywym organizmie” pokazujemy przedsiębiorcom, ale też np. domenę publicznej, w jaki sposób automatyzować dany proces.

Nazwa DIH-u wskazuje na to, że wspieracie firmy także w zakresie AI.

Sztuczna inteligencja jest naszą silną stroną ze względu na kompetencje VoiceLab, jak również środowiska, w jakim funkcjonujemy. Organizujemy meet-upy dotyczące AI, współpracują z nami osoby wyspecjalizowane w *machine learningu*. Oprócz kompetencji własnych wiemy więc też gdzie szukać partnerów z zewnątrz. Korzystamy w szczególności z zasobów Politechniki Gdańskiej, z prof. Jackiem Rumińskim na czele.

Obecnie startujemy w naborze na europejskie DIH-y. Stawka jest duża – wybrane inicjatywy uzyskają 7-letnie wsparcie o dużo większej skali niż obecnie. Chcemy, by taki DIH był także w Gdańsku, by służyły tu środki na rozwój nowoczesnych technologii, byśmy mogli cały czas animować to środowisko. Jeśli Pomorze chce zaistnieć na technologicznej mapie Europy, musi ściągać tu takie inicjatywy – dają one zarówno finansowanie, jak również wyspecjalizowane miejsca pracy, a także możliwość współpracy z DIH-ami z innych państw. Na tej bazie może powstać prawdziwy ekosystem, środowisko dla rozwijania i wdrażania najnowocześniejszych technologii w naszym regionie.

No właśnie – jakie są kluczowe elementy konieczne do tego, by Pomorze miało szansę na międzynarodowe zaistnienie, w szczególności w warstwie AI?

O pierwszym z tych elementów już mówiłem – to kwestia dostępu do danych uczących. W tym kontekście postawiłbym zdecydowanie na medycynę, przy wykorzystaniu współpracy z naszymi świetnymi lokalnymi uczelniami – Gdańskim Uniwersytetem Medycznym oraz Politechniką Gdańską. Dostępność tych danych może być nie tylko warunkiem, ale wręcz lokomotywą do pociągnięcia rozwoju branży na Pomorzu.

Drugim elementem, znacznie prostszym do spełnienia, jest dostępność odpowiednio dużej mocy obliczeniowej, czyli *de facto* komputerów wyposażonych w akceleratory, pozwalające na równoległe przetwarzanie bardzo dużych ilości danych. W ten sposób w praktyce „uczą się” maszyny. W tym momencie polskie uczelnie czy instytuty badawcze mają raczej skromne zasoby. Bez mocy obliczeniowych nie ma jednak rozwoju – na szczęście rozwiązanie tego problemu jest banalne, bo skupia się na konieczności poniesieniu kosztów, które nie są też znowu niebotyczne.

Trzeci i ostatni element, czyli dostępność kapitału intelektualnego, jest już na Pomorzu w zasadzie spełniony. Mamy tu dobre uczelnie, dobrze wykształconych specjalistów, a w dodatku jesteśmy miejscem atrakcyjnym do życia, co jest istotne w kontekście przyciągania tu talentów z zewnątrz. Teraz tylko ten kapitał intelektualny należałoby odpowiednio wykorzystać.

Dlaczego powinno nam tak bardzo zależeć na wzięciu udziału w tym wyścigu?

O sztucznej inteligencji mówi się czasem, że jest „nową elektrycznością”. Tak jak kiedyś świat został odmieniony przez prąd, a następnie przez komputery i internet, tak teraz takim *game changerem* jest właśnie AI. Nie ma dziedziny życia, której ona nie dotknie, będzie z nią związana cała gama zastosowań oraz rozwiązań.

Obecnie w obszarze AI nie tylko firmy konkurują z firmami – tutaj ścigają się wszyscy, także państwa z państwami. To, że „wszyscy” to robią, samo w sobie jest oczywiście słabym argumentem na to, że my też powinniśmy się zaangażować – wszak obszarów globalnego konkurowania jest dużo. Świetnie byłoby też przecież, gdybyśmy byli potęgą atomową czy militarną – tyle tylko, że w tych obszarach niezbędne byłyby potężne inwestycje. Infrastrukturalny próg wejścia w tego typu obszarach jest bardzo wysoki.

A w wypadku AI? Mamy ludzi z potencjałem, trzeba jeszcze dokupić komputery oraz ułatwić dostęp do danych. Zaczyna się coraz więcej u nas dziać w tym obszarze – chociażby na Politechnice Gdańskiej. Choć Polska nie ma wielkich osiągnięć światowych w tym zakresie, to jest wielu Polaków, którzy uczestniczyli w opracowywaniu bardzo ważnych, przełomowych nieraz rozwiązań i którzy mogą stanowić pewną inspirację, a część z nich – zapewne także i pomoc. Uważam, że 90% sukcesu mamy tu na wyciągnięcie ręki.

“ W obszarze AI 90% sukcesu Pomorze ma na wyciągnięcie ręki – mamy ludzi z potencjałem, trzeba byłoby jeszcze dokupić komputery oraz ułatwić dostęp do danych. Choć Polska nie ma wielkich osiągnięć światowych w AI, to jest wielu Polaków, którzy uczestniczyli w opracowywaniu bardzo ważnych rozwiązań i którzy mogą stanowić pewną inspirację, a nawet pomoc.

Rozumiem, że w czarnym scenariuszu nadal będziemy skazani na rozwiązania opracowywane przez globalne koncerny technologiczne. Czy ta wizja powinna nas aż tak martwić?

Uważam, że oddanie nas w ręce globalnych koncernów jest scenariuszem nie tyle czarnym, co bardzo ryzykownym. Byliśmy ostatnio świadkami tego, że wiodące sieci społecznościowe zablokowały Donaldowi Trumpowi dostęp do swoich kont. Co będzie dalej? Firmy te będą mogły decydować o tym, jakie informacje udostępniać na swoich łamach użytkownikom, a jakie nie? Coraz częściej już teraz jesteśmy tego świadkami. Takie praktyki mogą zostać przeniesione właściwie na każdą dziedzinę życia.

Jestem przekonany o tym, że stawką obecnego wyścigu jest tak naprawdę nasza wolność. W coraz większym stopniu żyjemy dziś, będąc pod kontrolą wielkich globalnych firm – nosimy ich urządzenia w kieszeni, pracujemy dla nich, korzystamy z ich aplikacji, udostępniamy im swoje dane, nieraz również te wrażliwe. Czy myśląc o naszym bezpieczeństwie, w obecnych realiach bardziej od innych państw nie zagrażają nam wielkie firmy, które mogą – obecnie lub w przyszłości – mieć inne interesy, niż my?

Najlepszym sposobem na to, by czuć się dziś wolnymi, jest przeciwdziałanie nadmiernemu uzależnieniu od koncernów i budowanie własnych zdolności i zasobów w obszarze AI. W innym wypadku już za kilka lat każdy z nas może widzieć świat oczami marketingowców dużych firm. Moim zdaniem – i mówię to z pełną odpowiedzialnością – byłaby to dla nas cywilizacyjna katastrofa.

“ Najlepszym sposobem na to, by czuć się dziś wolnymi, jest przeciwdziałanie nadmiernemu uzależnieniu od koncernów i budowanie własnych zdolności i zasobów w obszarze AI. W innym wypadku już za kilka lat każdy z nas może widzieć świat oczami marketingowców dużych firm.

O rozmówcy



JACEK KAWALEC

Wiceprezes Zarządu, VoiceLab

Jacek Kawalec – Wiceprezes Zarządu firmy VoiceLab, z którą jest związany od 2012 r. Doświadczenie zawodowe zdobywał w Oracle w Stanach Zjednoczonych, od 1995 r. był współzałożycielem i udziałowcem portalu Wirtualna Polska (wp.pl), gdzie odpowiadał za rozwój technologii. Po sprzedaży Wirtualnej Polski w 2005 r. zaangażował się w inwestowanie w start-upy. Absolwent Wydziału Psychologii Katolickiego Uniwersytetu Lubelskiego.

Partnerzy „PPG”



SAMORZĄD
WOJEWÓDZTWA POMORSKIEGO



GDAŃSK

