



**WYDZIAŁ  
ZARZĄDZANIA**

Uniwersytet Łódzki

## **Czy sztuczna inteligencja może nam pomóc w walce z terrorem fake newsów?**

**Komentuje dr Dominika Kaczorowska-Spychalska z Wydziału Zarządzania UŁ**

**O infodemii, czyli epidemii fałszywych informacji na temat koronawirusa, od kilku tygodni mówi Światowa Organizacja Zdrowia. Rozprzestrzeniająca się jak wirus plaga *fake newsów* prowadzi do dezinformacji, zakłamując rzeczywistość.**

*Fake newsy* nie są jednak efektem ubocznym obecnej pandemii, a zjawiskiem, które od dłuższego czasu rozwija się w Internecie. Setki informacji pojawiających się codziennie w formie artykułów, podcastów czy materiałów wideo niosą za sobą ogrom informacji, wśród których coraz trudniej odróżnić te prawdziwe. Według ekspertów Gartnera do 2022 roku będziemy konsumować więcej *fake newsów* niż informacji prawdziwych, głównie za pośrednictwem mediów społecznościowych. Tego typu komunikaty wzmacniają nasze subiektywne poglądy, wzbudzają nasze zainteresowanie i budzą niejednokrotnie skrajne emocje. Powtarzane przez użytkowników, w krótkim czasie wciągają nas w świat półprawd i kłamstw - w efekcie nie tylko nie potrafimy ich zidentyfikować, ale także coraz częściej odrzucamy informacje prawdziwe. Czy sztuczna inteligencja może nam pomóc w ich identyfikacji? A może sama potrafi je stworzyć?

Generatory tekstów, jak chociażby OpenAI, mogą być skuteczne w zwalczaniu *fake newsów*, ale jednocześnie mogą stać ich wprawnymi twórcami. Potrafią wykrywać teksty noszące znamiona nieprawdziwych, jak i samodzielnie pisać artykuły jedynie w oparciu o zadany im temat. Ryzyko, że tego typu narzędzia staną się skutecznym instrumentem w tworzeniu *fake newsów* sprawiło, że w przypadku OpenAI, zwlekano z ujawnieniem jego pełnej wersji. I chociaż obecnie jest ona dostępna na rynku w formie otwartej licencji, eksperci są zdania, że prezentuje ona zaledwie fragment swoich faktycznych możliwości.

Technologia *deep fake* (będąca połączeniem angielskich słów *deep learning* i *fake*), oprócz swojego potencjału rozrywkowego sprawdza się także przy tworzeniu *fake newsów*. Pozwala, przerobić dostępne, nawet w czasie rzeczywistym, informacje tekstowe, dźwiękowe czy wideo, które dla przeciętnego użytkownika są w zasadzie nie do odróżnienia od tych prawdziwych. To na nich osoby znane i rozpoznawalne mówią lub robią rzeczy, których nigdy nie powiedziały i nie zrobiły. Przekonali się o tym chociażby Keanu Reeves, Tom Cruise, Barack Obama, Joe Biden czy Nancy Pelosi. Ciekawym przykładem mogą być także wygenerowani, dzięki tej

Agnieszka Wołowicz

**Wydział Zarządzania UŁ**

ul. Matejki 22/26, 90-237 Łódź

tel.: 601 082 770, e-mail: [agnieszka.wolowicz@uni.lodz.pl](mailto:agnieszka.wolowicz@uni.lodz.pl)

[www.wz.uni.lodz.pl](http://www.wz.uni.lodz.pl)

technologii, prezenterzy wiadomości w Chinach. Wydają się być oni niezwykle autentyczni, bowiem ich ruchy, mimika i mowa ciała są całkowitym odwzorowaniem ich ludzkich odpowiedników. W efekcie *deep fake* może okazać się wyjątkowo skutecznym, narzędziem w procederze dezinformacji.

Okazuje się również, że nie tylko ludzie odpowiadają za rozprzestrzenianie się fałszywych wiadomości. Aktywne w tym procesie są także boty, które według raportu Distil Networks „Bad Bot Report 2019”, generują blisko 40% całego ruchu w sieci. Przypominają prawdziwych internautów, co oznacza, że na ich profilach znajdziemy np. szereg danych osobowych, zdjęcia, a także wpisy urealniaszące ich codzienne życie. W rękach twórców *fake newsów*, stosujących coraz bardziej zaawansowane metody dotarcia do odbiorców, boty są coraz bardziej atrakcyjnym narzędziem oddziaływania w mediach społecznościowych. Potwierdzają to badania IAB Polska „Dezinformacja w sieci”, które wskazują na *social media* jako główne źródło nieprawdziwych informacji. Według badań Państwowego Instytutu Badawczego NASK co piąty polski internauta twierdzi również, że nie sprawdza wiarygodności internetowych informacji ani źródeł ich pochodzenia. Może to sprzyjać, już wkrótce, rozwojowi zjawiska, nazwanego przez ekspertów Gartnera tzw. *fake of everything*, czyli sytuacji, gdy wszystko wokół nas staje się elementem podróbek, fałszerstw i wyrafinowanego kłamstwa, trudnych czy wręcz niemożliwych do odróżnienia.

Na szczęście ta sama sztuczna inteligencja może okazać się pomocna przy identyfikowaniu nieprawdziwych informacji, co oznacza, że już wkrótce nie będziemy musieli jedynie ograniczać się w tym zakresie do korzystania z metod konwencjonalnych, jak chociażby weryfikacja autora, daty publikacji czy medium, które upowszechniło dany tekst. Prace nad stworzeniem sztucznej inteligencji umożliwiającej nam wykrywanie *fake newsów* są reakcją na jej wykorzystanie w procesie ich kreacji. Im bardziej sztuczna inteligencja jest zaawansowana w procesie tworzenia nieprawdziwych materiałów, tym bardziej potrzebne są oparte na niej algorytmy i systemy pozwalające demaskować efekt jej pracy twórczej.

I chociaż obecnie wykrywaniem *fake newsów* wciąż zajmują się głównie ludzie, tzw. *fact checkerzy* (ang. *fact-checkers*), to sztuczna inteligencja w ich rękach mogłaby okazać się nie tylko wsparciem zwiększającym skuteczność tych działań, ale szansą na to, by przenieść się na nowy poziom interakcji w sieci – bardziej rzeczywisty i prawdziwy. Pozwoli ona bowiem w krótkim czasie odnajdywać nieprawdziwe informacje i blokować ich zasięg. Wymaga to jednak znacznych badań, z jednej strony, nad dynamiką ich rozprzestrzeniania, jak w przypadku wirusów, z drugiej, nad ich wpływem na osoby „zakażone” ze szczególnym uwzględnieniem wymiaru etycznego i społeczno-kulturowego tego zjawiska.

Pomocne w tym procesie okazują się algorytmy sztucznej inteligencji wykrywające cechy *fake newsów* czy określające poziom przewidywalności tekstów, jak chociażby GLTR (ang. *Giant Language Model Test Room*), wielowymiarowe repozytoria danych jak np. FakeNewsNet, stanowiące podstawę wielu badań czy technologie uwierzytelniania instalowane na urządzeniach rejestrujących obraz. Pojawiają się także coraz liczniej platformy, jak chociażby Deep Fake Detection Challenge czy Grover (ang. *Generating aRticles by Only Viewing Metadata Records*) pracujące nad stworzeniem narzędzi pozwalających na skuteczne rozpoznawanie *fake newsów* i ich zwalczanie. W tego typu działania angażują się również naukowcy czy agencje prasowe, które już dzisiaj wykorzystują sztuczną inteligencję jako wsparcie przy opracowaniu profesjonalnych tekstów.

Pandemia koronawirusa stworzyła przyjazne warunki do gwałtownego rozwoju i rozprzestrzeniania się *fake newsów*, w tym w szczególności tych, które go dotyczą. Bazujące

Agnieszka Wołowicz

Wydział Zarządzania UŁ

ul. Matejki 22/26, 90-237 Łódź

tel.: 601 082 770, e-mail: [agnieszka.wolowicz@uni.lodz.pl](mailto:agnieszka.wolowicz@uni.lodz.pl)

na towarzyszącym nam lęku przed nieznanym, strachu przed śmiercią czy irytacji i bezsilności wynikających z samoizolacji, rozprzestrzeniają się równie szybko jak sam wirus, dając nam zarówno nadzieję, jak i ją odbierając. Można by pokusić się o stwierdzenie, że infodemia stała się swoistą formą terroryzmu precyzyjnie wymierzonego w pojedynczego człowieka. Dlatego też w walkę z nią zaangażowały się już największe portale społecznościowe, które z jednej strony blokują podejrzane konta i ograniczają zasięgi rozprzestrzenianych przez nie treści, z drugiej włączają się w kampanię informacyjną, wskazując rzetelne i wiarygodne źródła informacji o chorobie. Platformy *fact checkerów*, jak chociażby Konkret24, na bieżąco weryfikują wiadomości, które mogą nosić znamiona *fake newsów* i publikują swoje rzetelne opinie na ich temat, a projekty takie jak #FakeHunters, zainicjowane przez PAP i GovTech Polska, zwiększają rolę społecznej weryfikacji tego typu treści związanych z koronawirusem. Również i w tym przypadku w walce z infodemią sięgnięto po sztuczną inteligencję, czego przykładem jest europejski projekt badawczy SocialTruth, w którym uczestniczy między innymi Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy. I chociaż pierwotnie projekt dotyczył rozpoznawania *fake newsów* jako takich, to w obliczu zalewu nieprawdziwych informacji dotyczących obecnej pandemii, część prowadzonych działań jest ukierunkowana właśnie na treści dotyczące koronawirusa.

Sztuczna inteligencja nie wyręczy nas w krytycznym myśleniu, analizowaniu i filtrowaniu konsumowanych treści, może nam jednak pomóc w walce z dezinformacją. Według futurologów technologia ta może być lepsza od nas w wielu obszarach aktywności już w ciągu najbliższych 20 lat. Może zatem okazać się bardziej skuteczna również w procesie rozpoznawania *fake newsów* czy ich eliminacji. Niewątpliwie, patrząc na jej możliwości, już dzisiaj wydaje się być ona jednym z najbardziej intrygujących efektów ewolucji człowieka.

#### Więcej o autorce:

**Dominika Kaczorowska-Spychalska** jest Dyrektorem Centrum Mikser Inteligentnych Technologii CMIT na Wydziale Zarządzania Uniwersytetu Łódzkiego. Jako ekspert uczelni w obszarze technologii cyfrowych, zajmuje się w szczególności sztuczną inteligencją (AI) i Internetem Rzeczy (IoT) oraz ich implikacją w biznesie. Pasjonuje ją wpływ tych technologii na zachowania człowieka (Homo Cyber versus Homo Roboticus) oraz zakres i charakter interakcji między nim a technologią (Human-to-Machine, Machine-to-Human), w tym problematyka Digital Ethics. Jest autorką (w części współautorką) prawie 60 publikacji wydanych zarówno w wydawnictwach polskich, w tym czasopismach branżowych skierowanych do praktyków (np. „Marketing w Praktyce”), jak i zagranicznych (np. Springer). Reprezentowała Polskę podczas Konferencji Państw Grupy Wyszehradzkiej "V4 Conference on Artificial Intelligence" (Workshop Session: Societal challenges and labour market impacts by AI), a także była w grupie ekspertów zewnętrznych Ministerstwa Cyfryzacji zaangażowanych w prace nad przygotowaniem „Założeń do strategii AI w Polsce”. Jest pomysłodawcą i współautorką programu „Inteligentne technologie i człowiek”.