

Cyfrowy policjant nas nie uratuje

15 września 2024

„To samo, co zrobiono z kasjerami w supermarketach, dzieje się powoli z policjantami i sędziami. Obiecano nam, że technologia rozwiąże wszystkie problemy, o ile my zgodzimy się wszystkie swoje problemy i troski powierzyć w ręce technologii” – mówi Griff Ferris, badacz wpływu technologii na prawa człowieka w wymiarze sprawiedliwości.

– Gdy próbowałem zgłębić temat wpływu sztucznej inteligencji na wymiar sprawiedliwości, co krok napotykałem się na optymistyczne prognozy i scenariusze. „Sztuczna inteligencja przyniesie rewolucję w podejściu do bezpieczeństwa i uwolni wielkie zasoby po stronie policji” – to tylko jeden z wielu podobnych cytatów. Czytam, że policja dzięki zastosowaniu algorytmów będzie miała więcej czasu, biurokracja zostanie ograniczona, a automatyzacja i cyfryzacja przyspieszą działanie wymiaru sprawiedliwości. Gdzie zatem problem?

– Cóż, technologiczny solucjonizm jest problemem sam w sobie. Widzimy to w kolejnych dziedzinach życia: w transporcie, ochronie zdrowia, świadczeniach społecznych, a obecnie w bezpieczeństwie publicznym i wymiarze sprawiedliwości. Za każdym razem słyszymy, że nowe technologie nas uratują, że sztuczna inteligencja zajmie się za nas różnymi patologiami. Teraz – że AI wyłapie przestępców, oczyści nasze ulice z chuligaństwa, wniesie cyfrową precyzję i neutralność do sądów... A nikt nie zadaje podstawowych pytań.

– Na przykład jakich?

– Co AI, czyli „sztuczna inteligencja”, miałaby w tym kontekście oznaczać. Kiedy ktoś przekonuje nas o zbawiennych skutkach jakiejś technologii, a my nawet nie wiemy, jak ona działa – to czy nie powinno nas to zastanawiać? A dziś jest

tak, że właściwie wszystko, co opiera się na jakiejś bazie danych, oprogramowaniu i nawet bardzo luźno pojętej automatyzacji, od razu nazywane jest sztuczną inteligencją. Gdy więc na przykład policja wprowadza rozwiązania bazujące na kartotekach czy informacjach, które już posiada, także nazywa to „wdrażaniem AI”. Można mieć wrażenie, że rzeczywiście wchodzimy na inny poziom. To jednak złudne.

– Dlaczego?

– Z każdym narzędziem bazującym na przetwarzaniu danych przez algorytmy, a tym bardziej na uczeniu maszynowym, wiąże się zasadnicze ryzyko. Takie, że algorytm będzie powielał prawidłowości i uprzedzenia zawarte w tych danych i „uczył się” ich. Bazy danych odzwierciedlają bowiem strukturę władzy i mechanizmy działania tego społeczeństwa, które je stworzyło. W policyjnych kartotekach w państwach europejskich znajdziemy nieproporcjonalnie dużo ludzi biednych, o innym niż biały kolorze skóry czy Romów, osób marginalizowanych – historycznie zawsze tak było. To osoby o mniej pewnej pozycji w społeczeństwie były najczęściej obiektem nadzoru, przeszukań, aresztów i śledztw. Zatem organy ścigania zawsze dysponowały większą bazą danych na ich temat. W oczywisty sposób wprowadzenie tutaj rachunku algorytmów i „sztucznej inteligencji” będzie skutkowało brakiem neutralności systemów cyfrowych.

– Rozumiem: wymiar sprawiedliwości historycznie nie był wolny od uprzedzeń. Ale co zmienia w tym AI?

– Jeśli jesteśmy świadomi, że istnieje dyskryminacja w jakiegokolwiek formie, to chyba nie powinniśmy się zgadzać na jej dalsze trwanie, prawda? Dziś godzimy się nie tylko z istnieniem tych uprzedzeń – rasizmu, seksizmu, homofobii – ale i na przeniesienie ich na poziom kolejnych algorytmów i technologicznych narzędzi w naszym wymiarze sprawiedliwości. Dopuszczamy myśl, że systemy, które będą wspomagać (a może nawet wyręczać) człowieka w fundamentalnych decyzjach, same

będą dalekie od neutralności i obiektywizmu.

– Ktoś może łatwo odpowiedzieć, że ryzyko błędu jest zawsze, ale wymiar sprawiedliwości w państwie demokratycznym gwarantuje prawo do obrony, odwołania się do wyższej instancji, domniemanie niewinności i wiele innych zabezpieczeń. Czy ryzyko, że padniemy ofiarą rasistowskiego czy seksistowskiego algorytmu, rzeczywiście jest aż tak duże?

– Spójrzmy na to inaczej. Istnieje większe ryzyko, że nigdy nie dowiemy się, czy w naszej sprawie sięgnięto po jakieś narzędzia cyfrowe czy algorytmy – ani też w jaki sposób wpłynęły one na decyzję policji, prokuratury czy sądu. Dlaczego? Po pierwsze: władza nie ma ochoty ani obowiązku się z tego tłumaczyć. Po drugie: nawet jeśli wiemy, że nasze państwo korzysta z narzędzi sztucznej inteligencji, to firmy, które dostarczają te technologie, nie ujawniają, jak one działają. To ich pilnie strzeżony sekret: prawo do algorytmu i wiedza o jego funkcjonowaniu to często przecież najdroższe, co mają. Jeśli więc nie wiemy, co i jak jest wdrażane do działania organów ścigania i policji, to nie mamy się na co skarżyć i od czego odwoływać, prawda?

– Ale to chyba najlepszy moment, żeby w końcu powiedzieć, jakich konkretnie systemów czy narzędzi powinniśmy się twoim zdaniem obawiać.

– Jeśli chodzi o policję i wykorzystanie sztucznej inteligencji, mówimy dziś tak naprawdę o dwóch typach narzędzi: o algorytmach predykcyjnych i o ocenie ryzyka w odniesieniu do konkretnych osób. Te pierwsze przeczesują bazy danych policji i innych agencji, aby na tej podstawie (z uwzględnieniem aspektów geograficznych, urbanistycznych, demograficznych) wytypować miejsca, gdzie może dojść do łamania prawa. Te drugie zaś mają za zadanie oszacować, czy istnieje prawdopodobieństwo, że wytypowana osoba popełni w przyszłości przestępstwo.

– I ktoś tego naprawdę używa?

– Jest wiele różnych przykładów używania tego typu systemów. W Amsterdamie w Holandii działa narzędzie opracowane przez władze miejskie w porozumieniu z policją i prokuraturą, które bazuje na profilowaniu młodzieży poniżej 16 roku życia. W ramach tego systemu powstaje lista obejmująca „top 400” osób wysokiego ryzyka. Kiedy w okolicy dochodzi do przestępstwa, to właśnie nazwiska z tej listy mogą znaleźć się jako pierwsze w kolejce do przesłuchań, przeszukań i aresztowań. A rodzicom dzieci z takiej listy częściej grozi odebranie praw rodzicielskich. To tylko jeden z przykładów.

– Takich rzeczy jest więcej?

– Oczywiście. W Wielkiej Brytanii używano systemu indywidualnej oceny ryzyka, którego wskazanie wpływa na prawdopodobieństwo postawienia zarzutów. To znaczy: jeśli ktoś mieścił się w profilu, który wskazywał na ryzyko złamania prawa w przyszłości, było bardziej prawdopodobne, że usłyszy zarzuty. Ale i na odwrót: osoby, które rzekomo popełniły jakieś wykroczenie, ale ich profil wskazywał na niskie ryzyko, mogły w ogóle nie usłyszeć zarzutów. Uznawano, że i bez tego się zresocjalizują i poprawią swoje postępowanie. To jest przykład bardzo ważnych decyzji, które mogą wpłynąć na całe życie człowieka, a które w coraz większym stopniu powierzamy wyrokom AI.

– Wybacz, że to powiem, ale bardzo to wszystko „angielskie”. Mocno w duchu tradycji, która nie tylko stworzyła Benthamowski Panoptykon, ale i dalej szuka „klas niebezpiecznych” i domaga się narzędzi wyłuskiwania podejrzanych jednostek na podstawie cech charakterystycznych.

– Nie mam odruchu obrony swojej ojczyzny i nie będę zaprzeczał. Chciałbym móc powiedzieć, że to wyłącznie nasz problem. Ale tak nie jest. Mówiłem już chociażby o Holandii, ale do listy krajów stosujących podobne narzędzia można

dopisać też Hiszpanię. Tam działa między innymi system do oceny ryzyka ekstremizmu. Szybko przekształcił się w narzędzie do śledzenia muzułmanów, bo tak go zaprojektowano. Problem w tym, że niektóre zjawiska, którymi miałyby się zająć sztuczna inteligencja, wcale nie są łatwe do zdefiniowania. A od tego, jakie kryteria „ekstremizmu” czy „dżihadyzmu” przyjmie rząd, zależy, co stanie się potem. Może się stać tak, że masę młodych mężczyzn wrzucimy do bazy danych o ekstremistach, choć nie popełnili żadnego przestępstwa. Nie robią nic zakazanego, lecz zostali zaklasyfikowani tak przez algorytm z powodu swojego wieku, narodowości i religii. Źle zaprojektowany algorytm z miejsca uzna całą kategorię ludzi za podejrzanych – a co potem robi z tym państwo, nie wiemy. Oprócz profilowania pod kątem przestępczości i ekstremizmu warto wspomnieć jeszcze o tym, że w kolejnych krajach Europy sprzedaje się na rynku komercyjnym narzędzia do profilowania osadzonych w zakładach karnych, które obiecują „rzetelną i opartą na danych” ocenę ryzyka recydywizmu. To kolejny ważny obszar.

– Ale skoro Unia Europejska ma zakazać profilowania rasowego w algorytmach czy predykcyjnego postępowania policyjnego, to może cała nasza rozmowa nie ma sensu, bo nieetyczne praktyki i tak zostaną objęte zakazami zawartymi choćby w największej ustawie regulującej użycie sztucznej inteligencji, czyli AI Act?

– Bardzo bym się ucieszył, gdyby to była prawda. Oczywiście, Unia Europejska podjęła ważny i odpowiedzialny krok, rozważając zakazanie najbardziej niebezpiecznych technologii wykorzystujących AI. I zgadzam się, że właśnie policyjne algorytmy predykcyjne należy dopisać do tej listy. Co więcej, na pochwałę zasługuje to, że europarlamentarzyści zagłosowali za zakazem. Problem w tym, że zapisy, które ostatecznie trafiły do europejskiego prawodawstwa w AI Act, nie odpowiadają temu, o co wnioskowały organizacje broniące praw człowieka, gdy domagały się zakazu. Mówiąc w największym skrócie: zakaz obejmie pewne formy profilowania i zastosowania

algorytmów predykcyjnych. Pozostaje jednak szerokie pole do interpretacji, a zarazem otwarta droga do ewentualnego kwestionowania nadużyć w sądach. W rzeczy samej, trochę na to właśnie liczę – że władze będą musiały dowodzić w sądach, że mają prawo do sięgania po tego typu narzędzia. Teraz niestety, w formie rozwodnionej przez Komisję i Radę Europejską, nie możemy mówić o jednoznacznym zakazie. Dopiero się przekonamy, co to będzie oznaczać w praktyce.

– I co się wydarzy? Pytam zarówno w krótkim, jak i w szerszym horyzoncie czasowym.

– W XXI wieku, a w ostatniej dekadzie w szczególności, poznaliśmy doskonale tę logikę, która mówi, że trzeba oszczędzać pieniądze i zastępować ludzi maszynami tam, gdzie to tylko możliwe. Technologia będzie tańsza, bardziej efektywna i mniej wymagająca niż ludzie – a przynajmniej tak nam powtarzano. W jakimś sensie teraz to samo, co zrobiono z kasjerami w supermarketach, dzieje się powoli z policjantami i sędziami. Obiecano nam, że technologia rozwiąże wszystkie problemy, o ile my zgodzimy się wszystkie swoje problemy i troski powierzyć w ręce technologii. Są takie momenty jak protesty Black Lives Matter w 2020 roku, gdy cały świat mówi, że nasze instytucje są wadliwe. Drogi wyjścia są dwie. Można naprawić – czy choćby próbować naprawić – złe procedury, porzucić pewne założenia, odrzucić logikę, która normalizuje rasizm. Ale jest też pokusa przeciwna: nałożyć na to wszystko jeszcze narzędzia cyfrowe, algorytmy, sztuczną inteligencję. Wtedy zawsze będzie można powiedzieć, że człowiek nie zawiódł, że w konkretnym rozstrzygnięciu nie ma ludzkiej winy czy rasizmu – takie po prostu są wyroki systemu komputerowego. Jeśli nawet ktoś się pomylił, to na pewno nie policjant, sędzia czy urzędnik – ale komputer. Łatwo zrozumieć, dlaczego z pewnych względów niektórym taka wizja zautomatyzowania policji i wymiaru sprawiedliwości się podoba. Pozwala rozwodnić odpowiedzialność i w razie niepowodzeń przerzucić winę na technologię. Ale nie musimy się na to godzić.

– Jednak czy argument, że zdjęcie obowiązków z barków policji i uwolnienie zasobów wymiaru sprawiedliwości pomoże w skuteczniejszym ściganiu przestępstw, jest bezzasadny?

– Generalnie policja nie jest zbyt dobra w ściganiu przestępstw. To, czy w jakiejś dzielnicy pracuje 10 czy 100 policjantów, nie ma przełożenia na to, jaka jest wykrywalność przestępstw. Dużo większe znaczenie ma dostępność miejsc pracy, jakość usług publicznych, poziom zaufania do państwa czy skuteczność walki z biedą i wykluczeniem. Zasoby wydane na poprawę jakości życia mieszkańców, zabezpieczenie społeczne czy edukację mogą się bardziej przyczynić do poprawy bezpieczeństwa. To mój punkt pierwszy. Punkt drugi dotyczy tego, że systemy bazujące na AI też nie przyczyniają się do poprawy bezpieczeństwa. Skuteczność niektórych z nich w ocenie ryzyka nie przekracza 50%, więc nie są lepsze od rzutu monetą. Zazwyczaj modele, na których się opierają, są skrzywione i wskazują, że obcokrajowcy są bardziej skłonni do popełniania przestępstw. Co nijak nie pomaga walczyć z przemocą domową czy przemocą w najbliższej okolicy. Sądzę więc, że przejdziemy przez podobny cykl jak w przypadku monitoringu w miejscach publicznych w latach 1990. i w pierwszym dziesięcioleciu XXI wieku. Kiedyś wydawało się, że kamery w miastach uchronią nas przed przestępczością. Dziś jest ich pełno, ale trudno powiedzieć, żeby przyczyniły się do zwiększenia bezpieczeństwa i skuteczności ścigania przestępstw. Mamy więc olbrzymi wzrost nadzoru, ale niekoniecznie jakościową zmianę w sferze bezpieczeństwa.

– I tego powinniśmy się obawiać?

– Najgorszy scenariusz, który może się zrealizować, jest taki, że te kamery, systemy rozpoznawania twarzy, algorytmy predykcyjne i masowe profilowanie się po prostu przyjmą. I wtedy wszyscy – nie tylko groźni przestępcy – ale wszyscy, którzy podpadli władzom, będą obiektem takiego nadzoru. Uczestnicy protestów ulicznych, członkowie mniejszości narodowych, przeciwnicy władzy... Narzędzia służące rzekomo do

walki z przestępczością będą mogły zostać użyte przeciwko nim z wielką łatwością. Nietrudno sobie wyobrazić dzielnice biedy czy imigranckie osiedla zmieniające się w swego rodzaju państwo policyjne napędzane technologiami nadzoru i wolne od niego enklawy zamkniętych osiedli dla bogatych. Podobnie – łatwo sobie wyobrazić skutki trafienia do bazy ekstremistów czy innych „osób wysokiego ryzyka” dla niewinnego człowieka. Gdy pewnego dnia zacznasz zadawać sobie pytanie „Dlaczego nie mogę dostać pracy, stypendium, miejsca na uczelni czy umówić się do lekarza?” – być może jesteś podejrzany, a nawet o tym nie wiesz. Wszystko to brzmi jak koszmar, ale tak naprawdę jeśli nie zatrzymamy wdrażania pewnych technologii już teraz, to ta wizja nie jest wcale odległa.

Z Griffim Ferrisem rozmawiał Jakub Dymek

Źródło: Panoptikon.org