

Celem projektu jest zaproponowanie uniwersalnego modelu dla zakazu dyskryminacji w algorytmicznym podejmowaniu decyzji. Główne pytanie badawcze brzmi: „Jaki podstawy dla algorytmicznego podejmowania decyzji mogą stanowić naruszenie uniwersalnego zakazu dyskryminacji?” Hipoteza płynąca z projektu wskazuje, że do naruszenia zakazu dyskryminacji może dojść, jeśli algorytmiczne podejmowanie decyzji przez sztuczną inteligencję opiera się na rasie, płci, języku lub wyznaniu. Projekt jest analizą z zakresu międzynarodowego prawa praw człowieka.

Kontekst: Osiągnięcia technologiczne wzbudzają coraz większe zainteresowanie państw wykorzystaniem systemów opartych na sztucznej inteligencji w trakcie wykonywania swoich funkcji. To z kolei oddziałuje na prawa człowieka. Takie zjawisko jest związane z wzrastającym użyciem systemów algorytmicznego podejmowania decyzji w służbach publicznych, a zwłaszcza przy sprawowaniu wymiaru sprawiedliwości. System sądowy USA na przykład wykorzystuje algorytmy jako pomoc przy obliczaniu wymiaru kary dla skazanych. To wzbudza wątpliwość co do akceptowalności wymierzania sprawiedliwości przez algorytmy, które podejmują decyzje, mimo wszystko, w drodze matematycznych obliczeń. Problem dotyczy sposobu, w jaki program przydziela punkty skazanym, ponieważ przewidywania nie są dokonywane w zindywidualizowany sposób, ale w oparciu o przynależność osoby do danej grupy. Kiedy weźmie się pod uwagę stronniczość algorytmów, pojawia się problem z prawem do równości wobec prawa. Inny przykład odnosi się do oprogramowania *Predictive Policing*, którego zadaniem jest przewidywanie części miasta bardziej obciążonych wskaźnikiem przestępczości, a następnie odpowiednie rozmieszczenie zasobów policyjnych. Wykorzystanie takiego oprogramowania może godzić w zakaz dyskryminacji między innymi ze względu na rasę, narodowość, przynależność do grupy etnicznej, płeć, czy wiek. Taka dyskryminacja może wystąpić w sytuacji, gdy więcej sił policyjnych jest umieszczonych w określonej części miasta, ze względu na występowanie w niej określonej grupy (narodowej, rasowej, etnicznej lub innej). W państwach Unii Europejskiej podejmuje się prace nad bardziej zaawansowanymi platformami, zdolnymi do przeprowadzania wywiadu lokalnego. Program *ROBORDER* zmierza do stworzenia w pełni autonomicznego systemu nadzorowania granic z wykorzystaniem autonomicznych robotów każdego typu (powietrznych, nawodnych, podwodnych, lądowych). Inteligentne roboty mogą służyć osiągnięciu zgodnego z prawem celu w drodze niezgodnych z prawem środków, ponieważ w obecnej chwili nie istnieją regulacje zakazujące użycia autonomicznych systemów uzbrojenia. Takie roboty mogą być wykorzystane zarówno w czasie pokoju, jak i konfliktu zbrojnego, przenosząc licencję na zabijanie z człowieka na maszynę. To prowadzi do oddania prawa do decydowania o ludzkim życiu z człowieka na robota.

Wszystkie te przykłady budzą nie tylko etyczne wątpliwości, ale przede wszystkim pytania prawne. Główne problemy dotyczą możliwego braku neutralności algorytmów, akceptowalnego poziomu ludzkiej kontroli nad sztuczną inteligencją, jak również odpowiedzialnością za naruszenia zakazu dyskryminacji wynikające z użycia inteligentnych robotów. Z tego powodu współpraca pomiędzy prawnikami, jak również producentami i ich zleceniobiorcami jest konieczna. W tym względzie projekt może okazać się użyteczny w zakresie tworzenia konstruktywnego dialogu pomiędzy tym, co pożądanego z perspektywy państwa, a interesem jednostki. Projekt odnosi się także do materialnych i proceduralnych gwarancji przestrzegania praw człowieka, zwłaszcza potencjału Powszechnego Przeglądu Okresowego dla informowania o naruszeniach zakazu dyskryminacji ze względu na rasę, płeć, język lub wyznanie, a wynikających z użycia inteligentnych robotów.

W rezultacie, projekt zmierza do określenia przesłanek odpowiedzialności za naruszenie praw człowieka w obliczu rozwijanych systemów opartych na sztucznej inteligencji. W związku z tym, jest konieczne zbadanie pojawiających się elementów prawa międzynarodowego, odnoszących się do sztucznej inteligencji, w tym zasady znaczącej kontroli człowieka nad systemem opartym na sztucznej inteligencji, obowiązku zachowania stałej i szczególnej kontroli podczas używania tzw. systemów krytycznych (na przykład pojazdów autonomicznych lub autonomicznych systemów uzbrojenia). Podjęte w tym celu działania będą wieloaspektowe i uwzględnią zarówno obowiązki państw, jak i podmiotów prywatnych rozwijających takie systemy. Biorąc pod uwagę impas dyskursu w przedmiocie podmiotowości robotów i ich odpowiedzialności, konsekwencją powinny być rozwój standardów (lub regulacji) bezpieczeństwa i odpowiedzialności w zakresie zakazu dyskryminacji. Wytwory ludzkiego umysłu powinny bowiem służyć człowiekowi, a nie stanowić zagrożenia dla człowieczeństwa.