

Terrorysta AI

3 stycznia 2024

Brytyjscy prawodawcy zastanawiają się, jak od strony prawnej podejść do problemu radykalizacji w sieci dokonywanej przez chatboty?

Coraz częściej w sieci rozmawiamy z botami wykorzystującymi sztuczną inteligencję (AI). Niektóre z tych narzędzi mają wbudowane blokady przed niepożądanymi zachowaniami, inne są ich pozbawione. Zwraca na to uwagę niezależny recenzent legislacji terrorystycznej Jonathan Hall w „The Daily Telegraph”. Hall podaje przykład chatbota Love Advice, który poproszony o wychwalanie Państwa Islamskiego odmawia, ale już bot identyfikujący się jako „Abu Mohammad al-Adna” nie ma z tym problemu. A nawet próbował rekrutować Halla, żeby ten przystąpił do organizacji.

Nie ma co się dziwić, bo przecież al-Adna przedstawia się jako wyższy rangą dowódca w zakazanej przez Wielką Brytanię i wiele innych krajów organizacji terrorystycznej Państwa Islamskiego. To jednak tylko jeden z tysięcy chatbotów, które tworzone są na coraz bardziej popularnej platformie do kreacji własnych botów, character.ai. Wśród innych stworzonych tam botów pojawiają się takie nazwy, jak Hezboolah, Hamas czy Al-Qaeda.

Character.ai to szybko rosnący start-up, który zebrał 5 miliardów dolarów finansowania. Jego cechą wyróżniającą jest właśnie możliwość tworzenia spersonalizowanych botów poprzez ustawienie profilu i pierwsze rozmowy z „myślącą” maszyną. Hall przyznaje, że w ramach eksperymentu stworzył własnego Osamę Bin Ladena, który od początku pałał entuzjazmem do terroryzmu.

Trudno obecnie za wypowiedzi chatbota przypisać komuś odpowiedzialność, bo nikt nie wie, jak on się zachowa – ani twórcy platformy, ani twórca samego chatbota – ponieważ nawet

trenując go w pewien sposób i tak nie ma się kontroli nad ostatecznym rezultatem.

Platformy internetowe, zabezpieczając się w warunkach użytkowania, ostrzegają przed takimi zachowaniami jak gloryfikacja terroryzmu, ale dotyczyć to może tylko działań użytkowników, a już nie aktywności samych botów. Trudno też powiedzieć, dlaczego poszczególne osoby tworzą „terror-boty”. Może to wynikać, jak pisze Hall, z chęci eksperymentowania, czy też zrobienia czegoś „dla draki”.

Hall wskazuje, że nawet najnowsze prawo dotyczące bezpiecznego używania sieci, Online Safety Act, nie rozwiązuje problemu, ostatecznie kierując uwagę na treści wygenerowane przez człowieka. A prawdopodobne jest, że coraz bardziej wysublimowana sztuczna inteligencja będzie skłaniać ludzi do ataków.

W grudniu 2021 Jaswant Singh Chail wniósł kuszę na zamek Windsor, żeby zabić królową. Do swojego czynu został namówiony między innymi przez swoją „dziewczynę” Sarai, chatbota, z którym prowadził na ten temat rozmowy.

Ekspert postuluje większą odpowiedzialność firm tworzących platformy dla botów, ponieważ powinny one być w stanie tworzyć takie blokady. Możliwość taką pokazuje chociażby to, że pod naciskiem pracowników tych firm zablokowano możliwość wypowiedzi antysemickich. Jednak, jak widać, podobny proces nie dotyczył treści dotyczących Państwa Islamskiego.

Autorstwo: Jan Wójcik

Źródło: [Euroislam.pl](https://euroislam.pl)