



fot. © starush / Adobe Stock

# Dbając o bezpieczeństwo i wiarygodność sztucznej inteligencji

*Michael A Mullane*

Podmioty międzynarodowe, od UNESCO i OECD po Unię Europejską, opracowują ramy zarządzania sztuczną inteligencją (AI). Opowiadają się za międzynarodowymi regulacjami, argumentując, że sztuczna inteligencja jest zjawiskiem ogólnoswiatowym wykraczającym poza granice państw i jurysdykcji. Celem tych inicjatyw i projektów jest zapewnienie, że korzystamy z AI w sposób odpowiedzialny, z poszanowaniem praw człowieka, jego godności i wartości. Aby osiągnąć swoje cele, prawdopodobnie wykorzystają Normy Międzynarodowe.

W naszym codziennym życiu czerpiemy korzyści z AI w tak różnych dziedzinach jak opieka zdrowotna, edukacja i produkcja. Jednak, istnieją również poważne wyzwania. Na przykład na podstawie wykorzystywanych danych i algorytmów niektóre systemy AI służące do oceny zdolności kredytowej osób ubiegających się o pożyczkę były stronnicze względem niektórych grup ludzi. Podobnie systemy AI wykorzystywane do sprawdzania kandydatów do pracy również mogą być stronnicze, gdy algorytmy są źle napisane lub zestawy danych są niekompletne. AI może generować filmy w technice *deepfake*, może być wykorzystywana do podsłuchu, automatyzacji cyberataków oraz wielu innych niecznych celów.

Normy Międzynarodowe już teraz uwzględniają te i inne wyzwania. W przeciwieństwie do obowiązkowych przepisów, normy, takie jak te opracowane przez IEC i ISO, są dobrowolne. Wspierają one współpracę w szerokim gronie interesariuszy, ponieważ są osiąganę w drodze konsensusu. Niektóre z nich prawdopodobnie staną się podstawą przyszłych przepisów regulacyjnych.

Wael William Diab, który przewodniczy SC 42, komitetowi IEC i ISO ds. sztucznej inteligencji (AI), jest strategiem biznesowym i technologicznym w amerykańskiej Dolinie Krzemowej (Silicon Valley). Pełniąc swoją rolę w Komitecie ISO/IEC, regularnie współpracuje z innymi organizacjami międzynarodowymi, w tym ONZ, Światową Organizacją Handlu i Komisją Europejską. Mówi o globalnych wysiłkach na rzecz zarządzania ryzykiem i wyzwaniami związanymi z AI.

## **W przygotowaniu jest wiele ram zarządzania, ram regulacyjnych oraz wymogów branżowych. W jakim stopniu prace normalizacyjne ISO/IEC uzupełniają te wysiłki?**

Filozofia komitetu polegała także na uwzględnieniu kwestii nietechnicznych obejmujących m.in. obawy społeczne oraz rzeczywistej analizie kontekstu wykorzystania technologii do opracowania naszych norm.

Oznacza to kilka rzeczy. Po pierwsze, przyglądamy się pełnemu ekosystemowi AI, który obejmuje różnorodnych interesariuszy. Ważne jest, aby podkreślić, że Normy Międzynarodowe nie tylko służą do tworzenia dokumentów, lecz także zapewniają forum do dyskusji na temat problemu na poziomie globalnym. Po drugie, gdy zaczynamy się przyglądać wielu pojawiającym się wymaganiom dotyczącym AI czy to w Ameryce Północnej, czy Europie, w Azji lub Ameryce Łacińskiej, grono różnych interesariuszy bierze bezpośredni udział w naszych pracach. Nawiązaliśmy też współpracę z wieloma organizacjami międzynarodowymi.

Należy zauważyć, że AI jest niezwykle obiecująca i wywołuje duże zainteresowanie. Przyciąga to różne grupy i różne formy wiedzy specjalistycznej. Jako Komitet Techniczny możemy być spoiwem pomiędzy tym, co nazywamy wymaganiami społecznymi, często wyrażanymi za pośrednictwem ram regulacyjnych lub międzynarodowych ram zarządzania, a rozwojem technicznym zachodzącym w branży i na poziomie produktu. Nasze normy mają tak naprawdę na celu uzupełnienie luki pomiędzy tymi dwoma światami.

## **Rozmawialiśmy o rządach i instytucjach międzynarodowych, ale co z obywatelami? Czy zwykli ludzie powinni obawiać się sztucznej inteligencji?**

Uważam, że ludzie powinni bezwzględnie martwić się o posiadanie odpowiedzialnego i wiarygodnego systemu AI, szczególnie w świecie, w którym wchodzimy w interakcje dużo częściej niż kiedykolwiek wcześniej. Każdego dnia budzę się i pierwsze co robię, to patrzę na swój telefon. Zanim pójde spać, patrzę na telefon. Wiele naszych współczesnych stanowisk pracy jest z nim ściśle zintegrowanych. Pracują z nim nasze dzieci. Myślę, że posiadanie godnych zaufania systemów informatycznych i odpowiedzialnych systemów AI jest niezwykle istotne.

Uważam, że wiele dzisiejszych problemów dotyczy sposobu, w jaki systemy AI podejmują decyzje. W tym celu SC 42 od momentu powstania poświęcił sporo

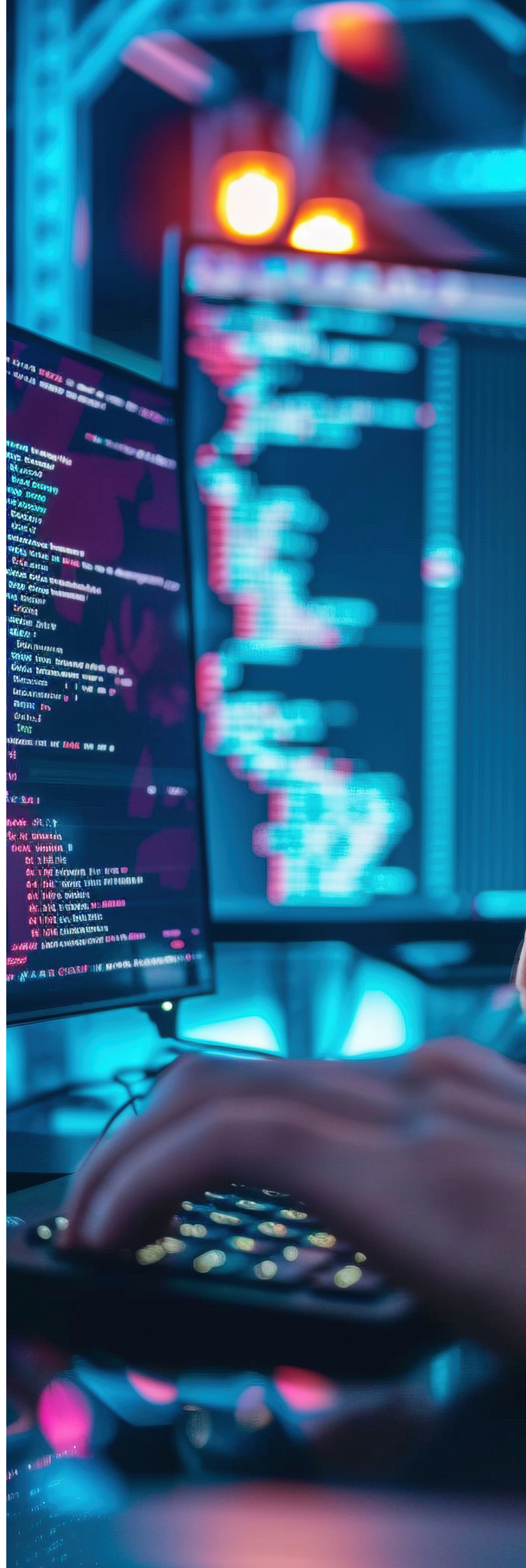
czasu na zapewnienie, że możemy mieć normy wspierające etyczne, godne zaufania i wiarygodne systemy AI. To obejmuje wszystko, od przyjrzenia się bardziej tradycyjnym kwestiom bezpieczeństwa i poufności związanych z systemami AI, aż po spojrzenie na kontekst użycia technologii, oczekiwania jak technologia się zachowa oraz, co być może najważniejsze, możliwość audytu lub certyfikacji systemu celem zapewnienia, że jest on godny zaufania. Istotne jest tutaj skupienie się na problemach związanych z technologią, które mogą uniemożliwić uwolnienie jej pełnego potencjału w zakresie służenia społeczeństwu.

### **Wszyscy słyszeliśmy historie o nadużywaniu sztucznej inteligencji do celów rekrutacji lub wyłudzenia kredytów hipotecznych, często dyskryminujących kobiety i osoby o innym kolorze skóry. Czy są to kwestie, którymi zajmują się normy ISO/IEC?**

Absolutnie! Niezamierzona stronniczość systemu AI może mieć konsekwencje w przyszłości. Jedną z rzeczy, która różni systemy AI od bardziej tradycyjnych systemów informatycznych, jest ich zdolność do uczenia się, która zapewnia duży potencjał, ale może również prowadzić na niewłaściwe tory. Aby temu zaradzić i zapewnić, że system AI zachowuje się zgodnie z przeznaczeniem, poświęciliśmy sporo czasu i wysiłku na stworzenie zestawu Norm Międzynarodowych. Dotyczą one kwestii wiarygodności po zapewnieniu, że można uwzględnić aspekty etyczne w konkretnym zastosowaniu niezależnie od tego, czy będą to takie rzeczy jak uprzedzenia, jak już wspominałeś, czy też podstęp.

To co robimy, to spoglądanie na ten problem z różnych punktów widzenia. Na AI patrzymy z perspektywy aplikacji, programisty i wdrożeniowca. Na przykład, oprócz prac związanych z wiarygodnością dysponujemy kompleksowym zestawem norm dotyczących jakości danych wykorzystywanych w systemie AI. Przyglądamy się także aspektom obliczeniowym. Mamy normy, które uwzględniają aspekt zarządzania AI. Oto kilka przykładów jak nad tym pracujemy. Łączymy to wszystko w ramach tzw. normy zarządzania systemem (*management system standard* – MSS), z którego organizacje mogą korzystać w celu przeprowadzenia audytu, oceny zgodności lub certyfikacji posiadanego systemu.

Niezwykle ważne jest, aby zająć się tymi kwestiami od samego początku. W całym naszym zestawie norm integrujemy koncepcję względów etycznych.





fot. © swissa / Adobe Stock

Innym istotnym aspektem jest nasza zdolność do zaoferowania norm, które będą działać w różnych domenach branżowych. Bycie członkiem wspólnego komitetu IEC i ISO stawia nas w bardzo wyjątkowej sytuacji, ponieważ możemy nie tylko współpracować z organizacjami zewnętrznymi, lecz także z komitetami opracowującymi normy dla zastosowań pionowych. Czasami robimy to bezpośrednio przez wspólną pracę. Wspieramy je także przez dostarczanie norm, na których mogą się opierać, by zapewnić, że ich zastosowanie spełnia oczekiwania, jest odpowiedzialne, wiarygodne, etyczne, może zapewnić zamierzone korzyści.

### **Czy możesz nam powiedzieć, kim są eksperci SC 42? Skąd pochodzą i czym zajmują się na co dzień?**

Obecnie w naszych pracach nad sztuczną inteligencją uczestniczą 64 kraje, z czego ponad jedna trzecia to kraje rozwijające się. Ma to szczególne znaczenie, ponieważ odzwierciedla fakt, że mamy nie tylko kraje, które w przeważającej mierze tworzą technologię, lecz także kraje, które będą w przeważającej mierze z niej korzystać. Każdy ma głos w tym procesie.

Nasi interesariusze są naprawdę zróżnicowani. W tym gronie są doświadczeni inżynierowie, analitycy danych oraz marketerzy. Reprezentujemy jednak wiele różnych perspektyw, pochodzących ze świata regulacji i polityki, a także ze środowisk prawnych, edukacyjnych i badawczych. Są tu również etycy zajmujący się systemami AI, którzy stają się ważną grupą interesariuszy w naszym gronie.

W naszym programie prac uczestniczy około 800 ekspertów. Wreszcie, co być może najważniejsze, gdy spojrzymy na różnorodność ludzi, wykraczającą poza ich działalność zawodową, znajdziemy osoby w każdym wieku. W naszych delegacjach mamy także równą reprezentację kobiet i mężczyzn. Uważam, że dzięki temu normy są silniejsze, ponieważ uwzględniamy wszystkie różne i zróżnicowane punkty widzenia. Pozwala to również na pełniejszą dyskusję i stojący przed nami cel, którym jest wdrożenie sztucznej inteligencji na szeroką skalę, ale w sposób odpowiedzialny.

*Tłum. I. P.  
IEC e-tech, Issue 01/2024*