Pató Viktória Lilla

Emberközpontúság kontra vállalati érdekek az EU mesterségesintelligencia-szabályozásában

Ha digitális technológiákról van szó, az Európai Unió progresszív és emberközpontú mesterségesintelligencia-politikájával határozza meg a globális kodifikációs teret. Az európai technológiai szabályozás a centralista Kína és a decentralista Amerikai Egyesült Államok között helyezkedik el, úgy igyekszik kialakítani mesterségesintelligencia-stratégiáját, hogy ösztönözze a tudományos és vállalati innovációt, ugyanakkor megvédje a társadalmat a túlkapásoktól. Mindezt teszi úgy, hogy a digitális térben ugyanazon jogokat és kötelezettségeket kívánja garantálni, amelyek az offline térben megilletik a polgárokat.

Az Európai Unió intézményeiben 2018 óta aktív diskurzus folyik az MI-technológiák fejlesztését és alkalmazását szabályozó folyamatok kialakításáról,[1] ahogy az UNESCO és más nemzetközi szervezetek is figyelemmel kísérik a technológia evolúcióját.[2] Az EU mesterséges intelligenciával kapcsolatos politikai vitáinak alapjai az etikai keretek, az irányelvek, az értékelés, valamint az ágazati ajánlások. A politikai diskurzust a megbízhatóság, a kiszámíthatóság, az átláthatóság, az emberközpontúság, fenntarthatóság és a diverzitás értékei vezetik.

A megbízható mesterséges intelligencia etikai irányelvei hét olyan kulcsfontosságú követelményt tartalmaznak, amelyeknek meg kell felelniük ahhoz, hogy megbízhatónak minősüljenek.

Az emberközpontúság elvének több definíciója is lehet, társadalomtudományos szempontból vizsgálva a technológiai fejlesztés célja az emberi evolúció előmozdítása, vagyis a gépnek az ember szolgálatába állítása. A biztonságtudományok területéről vizsgálva az emberközpontúság olyan technológiai megoldást is jelent, amely a végső irányítást meghagyja az embernél, a biztonságos mesterséges intelligencia addig garantálható, ameddig végső soron az ember felelős az irányításért olyan döntéseknél, amelyek káros hatásúak is lehetnek. Jogtudományi szempontból a technológiai fejlesztéskor az egyetemes emberi jogoknak érvényesülniük kell.

[1] Európai Bizottság [é. n.].
[2] Srinivan 2018.

https://doi.org/10.36250/01145_24

A mesterséges intelligencia tanításához adatokra van szükség, viszont az adatokat létrehozó személynek vagy csoportnak engedélyeznie kell adatuk felhasználásának célját. A magánélet és az adatvédelem teljes körű tiszteletben tartása mellett megfelelő adatkezelési mechanizmusokat is biztosítani kell, figyelembe véve az adatok minőségét és integritását, valamint az adatokhoz való legitim hozzáférést. Erre a Mesterséges Intelligencia Jogszabálycsomagban *(AI Act)* a szabványosítás jelenthet némiképp megoldást. A jelenleg gyártott MI-modellek többsége feketedoboz-szerű, felépítésükből fakadóan működési mechanizmusuk nem elemezhető az ember által, hiszen több milliárd információból állít elő egy újat. Az adatkezelési hozzájárulásnál az átláthatóság elve azt is jelölheti, hogy az MI-rendszerek működését és azok döntéseit az érintett felekhez igazodó módon kell elmagyarázni, illetve a rendszerek képességéről és korlátairól tájékoztatást adni. Erre a GDPR rendelettel kötelezővé tett adatkezelési tájékoztatók jóváhagyása nem jelent teljes körű megoldást, hiszen az adatok szolgáltatói legtöbb esetben nem értelmezik a hozzájáruláshoz szükséges rendelkezést. Fontos lenne az is, hogy egy kijelölt ellenőrző hatóság felelősségre vonható legyen a mesterségesintelligencia-rendszerek tevékenységeiért.

A sokszínűség előmozdítása érdekében a mesterségesintelligencia-rendszereknek mindenki számára hozzáférhetőnek kell lenniük, ehhez azonban olyan képzési rendszerre van szükség, amelyre még nem készült fel a tömegoktatás. A sokszínűség elvének érvényesülése azt is jelenti, hogy törekedni kell arra, hogy egyik népcsoport se marginalizálódhasson, így a mesterséges intelligencia fejlesztéséhez használt adatokra a lehető legszínesebb adatbázist kell építeni. Az MI-rendszereknek minden ember javát kell szolgálniuk, beleértve a jövő generációit is. Ezért biztosítani kell, hogy ezek fenntarthatók és környezetbarátak legyenek, és alaposan meg kell fontolni társadalmi hatásukat.

A felelősség és elszámoltathatóság érvényesülésében az auditálhatóságnak kulcsszerepe van. Utóbbi lehetővé teszi az algoritmusok, adatok és tervezési folyamatok értékelését, különösen a kritikus alkalmazásokban a megfelelő jogorvoslat lehetőségének megteremtésével. Jelenleg nincs politikai konszenzus abban, hogy ki a felelős az MI hibáiért, mert a fejlesztők nem okolhatók miután modelljeiket legyártották, azok az alkalmazottak pedig, akik a rendszerek működését biztosítják, egy szabályozott protokollt követnek, amitől eltérő hibáért nem vonhatók felelősségre.

A jogalkotás során arra törekednek Európa-szerte, hogy mindig legyen emberi felelős és jogorvoslati szint is a mesterségesintelligencia-rendszer mögött. Ezzel szemben az EU jelenlegi politikája és stratégiája az „algoritmus

hibáztatása" megközelítést ösztönzi, amely a vállalati érdekeket jobban szolgálja, mint az polgári jogokat.

Ha az EU versenyezni szeretne a Szilícium-völgy innovációs piacával a mesterséges intelligencia szektorban, az emberközpontú technológiai megközelítést háttérbe kellene szorítani a vállalati érdek érvényesüléséhez. Az Európai Unió Tanácsának cseh elnökségi prioritásai között szerepel a mesterséges intelligencia jogszabálytervezet körüli vita sürgetése. Jelenleg a mesterséges intelligencia szűkebb definíciójának kidolgozása zajlik, amely az európai alapokmányokban foglalt értékek érvényesítésére törekszik, azonban a politikai viták erősen átitatottak a vállalati lobbi befolyásával. A cseh elnökség júliusban a definíció pontosítása mellett a magas kockázatú MI-rendszerek listájának felülvizsgálatát és lerövidítését, az MI igazgatótanács szerepének megerősítését és a nemzetbiztonsági mentesség átfogalmazását javasolta.[3] 2024-ben, a magyar elnökség idején várható a jogszabály elfogadása.

2022. augusztus 25.

Felhasznált irodalom

Bertuzzi, Luca (2022): AI Act: Czech Presidency Pushes Narrower AI Definition, Shorter High-risk List. *Euractiv,* 2022. július 18. Online: www.euractiv.com/section/digital/news/ai-act-czech-presidency-pushes-narrower-ai-definition-shorter-high-risk-list

Európai Bizottság [é. n.]: *Ethics Guidelines for Trustworthy AI.* Online: https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

Srinivan, Rajeev (2018): The Ethical Dilemmas of Artificial Intelligence. In Srinivasan, Rajeev: *Human Decisions: Thoughts on AI.* 104–107. Online: https://unesdoc.unesco.org/ark:/48223/pf0000367514?posInSet=19&queryId=76d829de-c50c-4220-877c-4dd6ee379305

[3] Bertuzzi 2022.