

Megszületett a világ első átfogó MI-szabályozása

Történelmi pillanatot éltek át a brüsszeli tárgyalásokon: a Tanács és a Parlament megegyezett a mesterséges intelligencia szabályozásáról (AI Act). Európa a legelső kontinens, amely egyértelmű szabályokat határoz meg a mesterséges intelligencia (MI) használatára vonatkozóan.

Maratoni tárgyalássorozat után az Európai Parlament és az Európai Unió Tanácsa megállapodott az AI Act szövegében, noha ennek néhány hete még nem sok esélyt jósoltak¹ a szakértők, hozzáteszem, megalapozottan, hiszen november elején megtorpanni látszottak a tárgyalások.

Végül a spanyol elnökségnek sikerült az utolsó hónapjában letennie azt a mérföldkövet, amelyre biztosan később is hivatkozni fognak, nevezetesen, hogy a világ első átfogó mesterséges intelligencia szabályozásának elfogadásához járultak hozzá. Az, hogy mennyire átfogó ez a szabályozás, az EU globális digitális jogalkotói énképe körüli örömtől és e történelmi pillanat mámorában nem biztos, hogy kommunikációs szinten reálisan értékelt, de elmondható, hogy a háromnapos tárgyalássorozat, amely a trilógus záróakkordja volt, 2023. december 8-án éjfélkor eredményesen zárult, így az AI Act tervezetének² előterjesztése (2021. április 21.) után két és fél évvel megszületett – globális viszonylatban is elsőként – a mesterséges intelligencia szabályozása az Európai Unióban. Az AI Act megalkotói a gyorsan változó technológiai közegre igyekeztek általánosan alkalmazható, a versenyt nem túlzóan korlátozó, a fejlődést, ugyanakkor a polgárok biztonságát szem előtt tartó, a transzparenciának, az átláthatóságnak és az etikai elveknek (emberközpontú AI) megfelelő szabályozást létrehozni. A tárgyalás megtorpanását az okozta javarészt, hogy vita övezte a ChatGPT-hez hasonló alapmodellek szabályozásának beilleszthetőségét, amire végül egy puha eszközt vezettek be, nevezetesen azt, hogy az általános célú mesterségesintelligencia-rendszereknek (General Purpose AI, GPAI) a piacra lépésük előtt meg kell felelniük az átláthatósági követelményeknek, úgymint a műszaki dokumentáció megléte, az EU szerzői jogi

¹ ZÖDI 2023.

² Európai Bizottság 2021.

törvényeinek való megfelelés, valamint a képzéshez felhasznált tartalmakról szóló részletes összefoglalók terjesztése. Ezzel egy kicsit hozzáérttek ugyan a GPAI-szabályozáshoz, de átfogónak nem tekinthető.

A Bianchini–Ancona-szerzőpáros szerint³ az EU vitája az alapmodellek szabályozása körül polarizálódott, és az Európai Parlament a három legnagyobb gazdaságtól megkövetelt magatartási kódexeknél jóval szigorúbb rezsimet szorgalmazott. Az uniós intézményeknek a 21 nyitott kérdésből álló lista pontjaiban kellett megállapodniuk. A tárgyalások első körben 22 óra után szüneteltek, mert nem sikerült megegyezni a mesterségesintelligencia-eszközök bűnüldözésben való használatának kérdésében, amely a spanyol elnökség javaslata⁴ volt. A vitatott javaslatnál az uniós tagállamok abban egyetértettek, hogy a katonai és védelmi célú mesterséges intelligenciák az AI Act hatálya alól kivételt képezzenek.

Több okból volt nehezebb megállapodni, amihez hozzátartozik az is, hogy a tiltott kategóriába több szereplőt sorolt volna az Európai Parlament, mint a Tanács. Az is vitatott téma volt,⁵ hogy az unióban tiltott alkalmazásokat az európai cégek harmadik államban értékesíthetik-e, valamint hogy a szabályozás vonatkozik-e azokra a mesterségesintelligencia-rendszerekre, amelyek az alkalmazás megkezdése előtt is forgalomban voltak, ha jelentős változáson mennek keresztül.

Az előbbieket szerint az AI Act tiltott kategóriái közé tartozó rendszerek korlátozásait, mint például a biometrikus arcfelismerés, nemcsak az európai piacon szerették volna biztosítani, hanem az uniós székhelyű vállalatokat is szerették volna megakadályozni abban, hogy az EU határain kívül értékesítsék azokat. Ez a kísérlet azonban jogalap híján kudarcba fulladt. A parlamenti képviselők a digitális szolgáltatásokról szóló törvény hatálya alá tartozó „rendszerüzemeltető” közösségimédia-ajánlórendszereket is a tiltott kategóriába sorolták volna, ám ezt ismét megakadályozták a tagállamok. A migrációs trendeket előrejelző és határőrző rendszerek terén azonban a Parlamentnek sikerült a szigorítást képviselnie. A magas kockázatú rendszerek tekintetében az EP-képviselők elérték, hogy az állami szervek vagy közszolgáltatást nyújtó magánszervek magas kockázatú mesterséges intelligencia használatakor alapjogi hatásvizsgálatot végezzenek. A Tanács a vitás pontok közül e tárgyban mentességet vívott ki a bűnüldöző szervek javára, sőt a közbiztonsággal összefüggő

³ BIANCHINI–ANCONA 2023.

⁴ Az Európai Unió Tanácsa 2023.

⁵ BERTUZZI 2023b.

kivételes körülmények között a rendvédelmi hatóságok olyan magas kockázatú rendszert is használhatnak, amelyek még nem estek át a bírósági engedélykérő eljáráson.

A rendszerkockázattal járó GPAI-modellek esetében a Parlament tárgyalóinak sikerült szigorúbb kötelezettségeket biztosítaniuk. Ha ezek a modellek megfelelnek bizonyos kritériumoknak, modellértékeléseket és kontradiktórius tesztelést kell végezniük, fel kell mérniük és mérsékelniük a rendszerszintű kockázatokat, továbbá jelentéstételi kötelezettségük is van a Bizottság felé energiahatékonyságukról, valamint súlyos kiberbiztonsági incidensek esetén. A képviselők ragaszkodtak ahhoz is, hogy a harmonizált uniós szabványok közzétételéig a rendszerszintű kockázatot jelentő GPAI-k magatartási kódexekre hagyatkozzanak a rendeletnek való megfelelés érdekében.

Az AI Act a tiltott, a magas kockázatú, az alacsony kockázatú és a kockázatmentes kategóriák szerint osztályozza a mesterségesintelligencia-rendszereket, és a következő pontokban állapodtak meg a trilógus alatt a társjogalkotók:

- A tiltott alkalmazások listája, mint például a Clearview,⁶ amelyek arcfelismerésen alapuló technológiák és a manipuláció eszközévé válhatnak a sebezhetőségek kihasználásával.
- A munkahelyeken és iskolákban érzelemkifejezést felismerő technológiák is tiltólistára kerültek, ez alól csak a biztonsági szempontok kivételek.
- A prediktív rendészeti szoftvereket, amelyek felmérlik az egyén jövőbeni bűncselekmény-elkövetésének kockázatát, szintén korlátozták.
- Tiltottak azok az MI-rendszerek is, amelyek érzékeny személyes tulajdonságok alapján csoportosítanak egyéneket.
- A valós idejű távoli biometrikus azonosítás használatának tilalma területén az Európai Parlament engedett a tagállamoknak, ugyanis terror-támadások vagy súlyos bűncselekmények megelőzése érdekében a gyanúsítottak és az áldozatok felkutatására használhatók. Az indokoltság fennállását egy független hatóság hagyja jóvá előzetesen, és a Bizottság felügyeli.
- A magas kockázati kategóriába azokat a rendszereket sorolták be, amelyek az emberek biztonságát, illetve alapvető jogait sérthetik. Érzékeny területek közé tartozik az oktatás, a foglalkoztatás, a kritikus infrastruktúra, a közszolgáltatások, a bűnüldözés, a határellenőrzés

⁶ Clearview AI [é. n.].

és az igazságszolgáltatás. A magas kockázatú rendszerek használatát egy közös európai adatbázisba szükséges bejelenteni a jövőben.

- Az általános célú mesterségesintelligencia-rendszerek szolgáltatóinak (mint a ChatGPT) információt kell nyújtaniuk a törvényi megfelelésről azokkal a szolgáltatókkal szemben, amelyek magas kockázatú kategóriába tartozó alkalmazást hoznak létre.
- A tiltott rendszerekkel kapcsolatos intézkedések megsértése esetén a pénzbírság elérheti a vállalat éves globális forgalmának 6,5%-át vagy 35 millió eurót, a rendszer- és modellszolgáltatók kötelezettségszegésért 3%-át vagy 15 millió eurót, a szolgáltatás elmulasztásáért pedig 1,5%-át vagy félmillió eurót.
- Az AI Act a hatálybalépése után két évvel alkalmazható, kivéve a tiltott kategóriák esetében, mert ott a közzétételt követő hat hónappal számolnak.
- A magas kockázatú mesterségesintelligencia-rendszerekre, a megfelelésértékelő testületekre és az irányítási fejezetre vonatkozó követelmények a hatálybalépést követő egy évvel alkalmazandók.

Mindebből látszik, hogy az AI Act több szeretett volna lenni, mint egy szabálykönyv, sokkal inkább az EU kutatóinak és startupjainak kilövőállomása a globális mesterségesintelligencia-versenyben. Éppen ezért biztosítani akarták, hogy a vállalkozások, különösen a kkv-k, az értékláncot irányító iparági óriások indokolatlan elnyomása nélkül fejleszthessenek mesterségesintelligencia-megoldásokat. Ennek érdekében a megállapodás támogatja az úgynevezett szabályozási homokozókat (*regulatory sandbox*) és a valós világban végzett tesztelést (*real-world-testing*), amelyeket a nemzeti hatóságoknak kell létrehozniuk az innovatív mesterséges intelligencia kifejlesztésére és betanítására a forgalomba hozatal előtt.

A taxonómia összehangolása

A Gazdasági Együttműködési és Fejlesztési Szervezet (OECD) Tanácsa 2023. november 8-án frissítette⁷ a mesterségesintelligencia-rendszerek 2019-es definícióját,⁸ amelyet az AI Act tárgyalásaiba is becsatornáztak.

⁷ Gazdasági Együttműködési és Fejlesztési Szervezet 2023.

⁸ Gazdasági Együttműködési és Fejlesztési Szervezet 2019.

A korábbi ajánlás felülvizsgálata az MI-rendszer meghatározásának aktualizálása érdekében történt, biztosítva, hogy az ajánlás továbbra is műszakilag pontos legyen, és tükrözze a technológiai fejlődést, beleértve a generatív mesterséges intelligenciát is. A meghatározást október közepén vitatta meg az OECD Digitális Gazdaságpolitikai Bizottsága és a Mesterséges Intelligencia Kormányzási Munkacsoportja. „A mesterségesintelligencia-rendszer egy gépi alapú rendszer, amely explicit vagy implicit célok esetén a kapott bemenetből arra következtet, hogyan hozhat létre kimeneteket, például előrejelzéseket, tartalmat, ajánlásokat vagy döntéseket, amelyek befolyásolhatják a fizikai vagy virtuális környezetet. A különböző mesterségesintelligencia-rendszerek autonómiájukat és alkalmazkodóképességüket tekintve eltérőek a telepítésük után.” Fontos egyértelműsíteni az utolsó sort, az alkalmazkodóképességre való hivatkozás azt tükrözi, hogy egyes mesterségesintelligencia-rendszerek a tervezési és telepítési szakasz után is fejlődhetnek, különösen azok, amelyek gépi tanulási technikákon alapulnak.

Az OECD 2019-ben fogalmazta meg először a mesterséges intelligencia fogalmát, azonban a változtatásokra szükség volt, hogy egy jövőt állóbb, az elmúlt négy év változásait nyomon követő fogalom alakuljon ki. Az egyik fő változtatás az volt, hogy eltávolították a hivatkozást arra a tényre, hogy a céloknak ember által meghatározottnak kell lenniük, hogy a fogalom részévé váljanak azok az esetek, amikor az MI-rendszer új célokat tud megtanulni.

Az AI Act a következőképp fogalmaz: a mesterséges intelligencia rendszereit úgy tervezték, hogy különböző szintű autonómiával működjenek, ami azt jelenti, hogy legalább bizonyos mértékben függetlenek az emberi irányítástól. A gépalapú kifejezés arra utal, hogy az AI-rendszerek gépeken futnak. Az explicit vagy implicit célokra való hivatkozás kiemeli, hogy a mesterségesintelligencia-rendszerek explicit, ember által meghatározott vagy implicit célok szerint működhetnek. Az AI-rendszer céljai eltérhetnek az AI-rendszer tervezett céljától egy adott környezetben. Az előrejelzésekre való hivatkozás magában foglalja a tartalmat is, mint a mesterségesintelligencia-rendszer egyik lehetséges kimenetét. A környezet az a kontextus, amelyben az MI-rendszerek működnek, míg az MI-rendszer által generált kimenetek, azaz előrejelzések, ajánlások vagy döntések, a rendszer céljaira reagálnak a bemeneti adatok alapján, miközben a kimenet tovább befolyásolja a környezetet, akár csak új információk bevezetésével is.

Az OECD lépése összehangoltnak bizonyult a Parlamenttel, hiszen utóbbi az AI Act fogalomtárába beépítette az OECD definícióinak főbb elemeit, ahogy abban 2023 márciusában már megállapodtak az EP-képviselők nemzetközi

szemantikai összhang fenntartása érdekében. Az OECD is egyértelműsítést eszközölt 2019-es definícióján, a parlamenti képviselők ebből kifolyólag úgy fejlesztették ki⁹ az AI Act definícióját, hogy nem tudták, az OECD milyen módosításokat végez. Végül összhang keletkezett¹⁰ az OECD és az AI Act „mesterségesintelligencia-rendszer” definíciója között.

A közös taxonómiára nemcsak az EU és az OECD törekszik, hanem az EU–USA Kereskedelmi és Technológiai Tanács is, hiszen a kiszámíthatósághoz és az egyértelműséghez elengedhetetlen, hogy közös fogalmak¹¹ jöjjenek létre a technológiai szférában is világszerte.

Egy mondatban összefoglalva elmondható, hogy az AI Act célja a mesterségesintelligencia-rendszerek szabályozása annak alapján, hogy mennyire kockázatosak. Az AI Act végrehajtására a Bizottságon belül létrehoznak egy Mesterséges Intelligencia Irodát az alapmodelleket érintő rendelkezések érvényesítésére, amelynek működtetésére külön költségvetése lesz a függetlenség garantálására. A mesterségesintelligencia-rendszereket az illetékes nemzeti hatóságok felügyelik, amelyek az Európai Mesterséges Intelligencia Testületbe tömörülnek a törvény következetes alkalmazásának biztosítása érdekében. Tanácsadó fórumok és szakértői, illetve tudományos testületek követik nyomon az AI Act végrehajtását.

2023. december 11.

Felhasznált irodalom

- Az Európai Unió Tanácsa (2023): *Preparation of the Council (Competitiveness (Internal Market, Industry, Research and Space))*. 2023. november 17. Online: <https://data.consilium.europa.eu/doc/document/ST-15118-2023-INIT/en/pdf>
- BERTUZZI, Luca (2023a): OECD Updates Definition of Artificial Intelligence 'to Inform EU's AI Act'. *Euractiv.com*, 2023. november 9. Online: www.euractiv.com/section/artificial-intelligence/news/oecd-updates-definition-of-artificial-intelligence-to-inform-eus-ai-act/
- BERTUZZI, Luca (2023b): AI Act: EU Policymakers Nail Down Rules on AI Models, Butt Heads on Law Enforcement. *Euractiv.com*, 2023. december 7. Online: www.euractiv.com/section/artificial-intelligence/news/ai-act-eu-policymakers-nail-down-rules-on-ai-models-butt-heads-on-law-enforcement/

⁹ Európai Parlament 2023.

¹⁰ BERTUZZI 2023a.

¹¹ Európai Bizottság 2023.

Megszületett a világ első átfogó MI-szabályozása

- BIANCHINI, Niccolò – ANCONA, Lorenzo (2023): Artificial Intelligence: Europe Needs to Start Dreaming Again. *Robert-schuman.eu*, 2023. november 28. Online: www.robert-schuman.eu/en/european-issues/728-artificial-intelligence-europe-needs-to-start-dreaming-again
- Clearview AI [é. n.]. Online: www.clearview.ai/
- Európai Bizottság (2021): *Az Európai Parlament és a Tanács rendelete a mesterséges intelligenciára vonatkozó harmonizált szabályok (a mesterséges intelligenciáról szóló jogszabály) megállapításáról és egyes uniós jogalkotási aktusok módosításáról*. 2021. április 21. Online: <https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:52021PC0206>
- Európai Bizottság (2023): *EU–U.S. TTC: Call for Input on First Edition of WG1 Terminology and Taxonomy for Artificial Intelligence*. 2023. november 3. Online: <https://digital-strategy.ec.europa.eu/en/news/eu-us-ttc-call-input-first-edition-wg1-terminology-and-taxonomy-artificial-intelligence>
- Európai Parlament (2023): *Report on the Proposal for a Regulation of the European Parliament and of the Council on Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. 2023. május 22. Online: www.europarl.europa.eu/doceo/document/A-9-2023-0188_EN.html
- Gazdasági Együttműködési és Fejlesztési Szervezet (2019): *Artificial Intelligence & Responsible Business Conduct*. Online: <https://mneguidelines.oecd.org/RBC-and-artificial-intelligence.pdf>
- Gazdasági Együttműködési és Fejlesztési Szervezet (2023): *Recommendation of the Council on Artificial Intelligence*. 2023. november 8. Online: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- ZŐDI Zsolt (2023): Kisiklik az EU MI-szabályozása? *Ludovika.hu*, 2023. november 23. Online: www.ludovika.hu/blogok/itkiblog/2023/11/23/kisiklik-az-eu-mi-szabalyozasa/