

Milyen információbiztonsági kockázatok merülnek fel meghatározó MI-komponenst tartalmazó célszoftverek vállalati használatánál? NIS2 és AI Act előírások a kockázatok kezelésére

VFP3 KiberVéd - NIS2 Audit felkészítés jogi
állásfoglalás 2026/45

VERZIÓKÖVETÉS

Verziószám	Dátum	Módosította	Módosítások leírása
1.0	2026.09.18	Kálmán és Társa Ügyvédi Iroda	Első kiadott verzió

1. VEZETŐI ÖSSZEFOGLALÓ

A meghatározó MI-komponenst tartalmazó célszoftver más természetű kockázatot hoz a szervezetbe, mint egy MI chatbot. Itt nem a munkatárs visz ki adatot egy promptban, hanem a szervezet egy üzleti folyamat egy részét (csalásfigyelés, számlafeldolgozás, önéletrajz-szűrés) bízta egy tanított modellre, amelynek működése a szervezet számára gyakran egyáltalán nem átlátható. A modell nem előre megírt szabályok, hanem a tanítóadatokból levont minták alapján dönt, ezért a tanítóadat, a modellfájl és a modellt kiszolgáló beszállítói lánc olyan támadási felület, amellyel a hagyományos szoftverbeszerzés nem számol.

A NIS2 hazai szabályozása erre a szoftverbeszerzés és a fejlesztés meglévő követelményeivel válaszol. A célszoftver a szervezet elektronikus információs rendszere vagy annak eleme, ezért biztonsági osztályba kell sorolni, a szerződésben rögzíteni kell a biztonsági követelményeket, a modell cseréje pedig változás, amelynek hatását előre meg kell vizsgálni. A 7/2024. (VI. 24.) MK rendelet beszerzési, konfigurációkezelési, sértetlenségi és ellátási lánc követelményei közvetlenül alkalmazhatók, ha a modellt, a tanítóadatot és a kiszolgáló környezetet rendszerelemként kezeljük.

Az AI Act a célszoftvereket a rendeltetés szerint különbözteti meg. A szervezet a legtöbb célszoftvernek alkalmazója, kötelezettségei a kockázati szinttől függenek. A munkahelyi érzelmfelismerés tiltott, a toborzás, a hitelbírálat, egészségügyben alkalmazott MI vagy a kritikus infrastruktúra biztonsági alkotórésze nagy kockázatú, a legtöbb csalásfigyelő, előrejelző vagy dokumentumfeldolgozó rendszer minimális kockázatú. A nagy kockázatú rendszerek alkalmazói kötelezettségei az (EU) 2026/1744 rendelettel módosított határidők szerint 2027. december 2-től, a termékekbe épített rendszereknél 2028. augusztus 2-től alkalmazandók, a szolgáltatói kiberbiztonsági követelmények azonban már ma a beszerzési átvilágítás mércéjéül szolgálnak. A javasolt megközelítés nem új rendszer, hanem a meglévő beszerzési és változáskezelési folyamat kiegészítése néhány MI-specifikus kérdéssel.

2. A VIZSGÁLANDÓ KÉRDÉS

Egy NIS2¹-kötelezett szervezet (a továbbiakban mint „Szervezet”) olyan szoftvereket vezet be vagy fejleszt, amelyek alapfunkcióját gépi tanulással tanított modell adja, például csalásfigyelő, számlafeldolgozó, karbantartás-előrejelző vagy önéletrajz-szűrő rendszert. A vizsgálandó kérdés az, hogy mit értünk meghatározó MI-komponenst tartalmazó célszoftver alatt, milyen információbiztonsági kockázatokat hordoz ezek vállalati használata, továbbá hogy a hazai NIS2 szabályozás és az AI Act milyen kockázatcsökkentő intézkedéseket ír elő.

Az állásfoglalás a 2026/43. számú, „Milyen információbiztonsági kockázatok merülnek fel általános célú MI-rendszerek (LLM, MI-ügynökök) vállalati használatánál? NIS2 és AI Act

¹ Az Európai Parlament és a Tanács (EU) 2022/2555 irányelve (2022. december 14.) az Unió egész területén egységesen magas szintű kiberbiztonságot biztosító intézkedésekről (NIS 2 irányelv)

előírások a kockázatok kezelésére” című állásfoglalás (a továbbiakban mint „2026/43. állásfoglalás”) párja. Nem tárgya ezért az általános célú MI-rendszerek és az MI-ügynökök használata, a prompt injection, az MI-jártassági képzés és a közös jogszabályi alap, ezekre a 2026/43. állásfoglalás fejezetének megjelölésével hivatkozunk. Nem tárgya továbbá a saját alapmodell tanításának kockázata, az AI Act teljes szolgáltatói megfelelési rendszere, valamint az adatvédelmi és szerzői jogi kérdések elemzése.

3. GYAKORLATI JAVASLATOK

3.1 Alapvető szabályok a meghatározó MI-komponenst tartalmazó célszoftverekre

A jogszabályok nem írnak elő önálló MI-beszerzési szabályzatot, a követelmények a beszerzési szabályzatba, a fejlesztési eljárásrendbe és a változáskezelésbe beépíthetők. Javasoljuk, hogy a Szervezet legalább a következőket rögzítse.

- A szerep és a besorolás tisztázása a beszerzés előtt, vagyis annak eldöntése, hogy a Szervezet alkalmazó vagy szolgáltató lesz-e, a célszoftver felvétele a rendszernyilvántartásba, biztonsági osztályba sorolása a kezelt adatok és a támogatott folyamat alapján, valamint az AI Act szerinti kockázati szint dokumentált megállapítása, a tiltott mesterséges intelligencia-gyakorlatok kizárásával.
- A beszerzési követelmények, a modell rendeltetésének, korlátainak, tanítóadat-forrásainak és pontossági mérőszámainak dokumentálása, a biztonsági tesztelés eredményeinek átadása, a modellváltozásokról szóló előzetes értesítés, a Szervezet adatainak tanításra való felhasználásának kizárása, a mögöttes modellszolgáltató megnevezése.
- A modell mint rendszerelem, a modellfájl és a tanítóadat felvétele a rendszerelem-leltárba, mentése, eredetének és sértetlenségének ellenőrzése, továbbá minden újratanítás, modellcseré és szállítói modellfrissítés kezelése változásként, mért teljesítménnyel és biztonsági hatásvizsgálattal.
- A kimenet és a bemenet felügyelete, a modell pontosságának és eltolódásának mérhető mutatói, a kimenet emberi ellenőrzése ott, ahol a döntés személyt érint vagy visszafordíthatatlan, a döntések visszakereshető naplózása, a bemeneti adatok érvényességének ellenőrzése, a modell manipulációjának gyanúja kiberbiztonsági eseményként kezelendő.

3.2 Jó gyakorlatok és megoldások KKV-k számára

Egy kis- vagy középvállalkozás jellemzően kész terméket vesz igénybe, ezért a legnagyobb kockázatcsökkentést a beszerzési kérdőív hozza. Érdemes a szállítótól írásban megkérdezni, mire tanították a modellt, milyen pontosságot mér és hogyan, kinek az infrastruktúráján fut a következtetés, felhasználja-e a Szervezet adatait további tanításra, hogyan és mikor cseréli a modellt, milyen naplót ad át, és ki a mögöttes modellszolgáltató. Aki ezekre nem tud válaszolni, annak a termékét a Szervezet nem tudja biztonsági osztályba sorolt rendszerként üzemeltetni.

A második lépés annak eldöntése, mely döntéseket hozhat meg a modell önállóan (például egy számla mezőinek kitöltése), és melyeket csak emberi jóváhagyással (például egy ügyfél elutasítása). A harmadik a szállítói modellfrissítések kezelése. A Szervezet előzetes értesítést kér,

a frissítés után pedig rövid ellenőrző tesztet futtat egy megőrzött, ismert eredményű mintakészleten. Ez a legolcsóbb eszköz a csendes teljesítményromlás észrevételére.

3.3 A javasolt eljárás

- a) Felmérés és besorolás, a meglévő szoftverek átvizsgálása annak megállapítására, hol működik már ma tanított modell, ezek felvétele a nyilvántartásba, a Szervezet AI Act szerinti szerepének, a rendszer kockázati szintjének és NIS2 biztonsági osztályának dokumentált megállapítása a vezető jóváhagyásával.
- b) Kockázatelemzés, a fenyegetéskatalógus alapján a modellre, a tanítóadatra és az ellátási láncra vonatkozó kockázatok dokumentálása, a biztonsági hatásvizsgálat elvégzése a bevezetés előtt.
- c) Beszerzés és szerződés, a biztonsági, dokumentációs, tesztelési és értesítési követelmények rögzítése a szerződésben, egyedi fejlesztésnél a fejlesztői konfigurációkezelés és biztonsági tesztelés előírása.
- d) Bevezetés, átvételi teszt ismert eredményű adatkészleten, a modell és a tanítóadat felvétele a leltárba és a mentésbe, a naplózás és a monitorozási mutatók beállítása, a használatbavételről szóló vezetői döntés.
- e) Üzemeltetés és felülvizsgálat, a teljesítménymutatók rendszeres áttekintése, minden modellváltozás változáskezelése, a kockázatelemzés legalább évenkénti, jelentős változás esetén soron kívüli felülvizsgálata.

4. VÁLASZ ÉS INDOKOLÁS

4.1 A vizsgálat kiindulópontjai

Az állásfoglalás a 2026/43. állásfoglalás „A vizsgálat kiindulópontjai” pontjában rögzített alapokra épül. A Szervezet a Kiberbiztonsági tv. 1. § (1) bekezdés d) pontja alá tartozik. A célszoftvernél ehhez két kiindulópont járul.

Az első, hogy a célszoftver a Szervezet önálló elektronikus információs rendszere vagy egy meglévő rendszer eleme, ezért biztonsági osztályba kell sorolni. A törvény a besorolást a rendszer sértetlenségének és rendelkezésre állásának, valamint a kezelt adat bizalmasságának, sértetlenségének és rendelkezésre állásának kockázata alapján írja elő.² A besorolást tehát nem az dönti el, hogy a szoftverben MI működik, hanem az, hogy milyen adatot kezel és milyen folyamatot támogat. Mivel a „jelentős” osztályhoz tartozik a nagy mennyiségű személyes adat, a jogszabállyal védett adat és az érzékeny üzleti folyamat sérülésének lehetősége,³ egy hitelbírálati, csalásfigyelő vagy HR-szűrő modell jellemzően legalább „jelentős” osztályú rendszer eleme.

A második, hogy a rendelet 2. mellékletének követelményei a rendszerelemekre is vonatkoznak, és ha a Szervezet rendelkezési joga csak egyes elemekre terjed ki, a követelményeket ezekre kell

² 2024. évi LXIX. tv. 10. § (1) bekezdés

³ 7/2024. (VI. 24.) MK rendelet 1. melléklet 2.2.3. pont

teljesíteni.⁴ A tanított modell, a tanítóadat és a következtetést végző környezet rendszerelem, a modellre így a szoftverre, a tanítóadatra pedig az információra írt szabályok vonatkoznak.

4.2 A fogalmak pontosítása

4.2.1 MI-rendszer és MI-komponens

A hazai kiberbiztonsági jogszabályok az MI fogalmait nem határozzák meg, ezért itt is az AI Act fogalmait vesszük alapul. A rendelet az MI-rendszert a következtetési képességgel határozza meg.

1. „MI-rendszer”: gépi alapú rendszer, amelyet különböző autonómiaszinteken történő működésre terveztek, és amely a bevezetését követően alkalmazkodóképességet tanúsíthat, és amely a kapott bemenetből – explicit vagy implicit célok érdekében – kikövetkezteti, miként generáljon olyan kimeneteket, mint például előrejelzéseket, tartalmakat, ajánlásokat vagy döntéseket, amelyek befolyásolhatják a fizikai vagy a virtuális környezetet; [(EU) 2024/1689 rendelet 3. cikk 1. pont]

A hagyományos szoftver előre megírt szabályokat hajt végre, az MI-rendszer a tanítóadatokból tanult összefüggések alapján állít elő kimenetet. Az Európai Bizottság iránymutatása szerint a kizárólag ember által megírt szabályokat vagy egyszerű statisztikai módszereket alkalmazó rendszerek nem tartoznak a fogalom alá.⁵ Az MI-rendszer működhet önállóan vagy egy termék alkotóelemeként, az MI-komponens tehát nem külön jogi kategória, hanem az MI-rendszer azon formája, amikor egy tágabb alkalmazásba épül. Ha a modell biztonsági funkciót lát el, biztonsági alkotórésznek minősülhet,⁶ ennek a nagy kockázatú besorolásnál van jelentősége.

4.2.2 Meghatározó MI-komponenst tartalmazó célszoftver

Ezen állásfoglalás alkalmazásában meghatározó MI-komponenst tartalmazó célszoftver az a meghatározott üzleti feladatra készült alkalmazás, amelynek alapfunkcióját egy vagy több tanított modell adja, a modell rögzített rendeltetéssel működik, a felhasználó nem ad neki szabadon megfogalmazott utasítást, és a modell nélkül a szoftver a rendeltetését nem tudná betölteni. A „meghatározó” jelző különbözteti meg az olyan szoftvertől, amelyben az MI csak kiegészítő kényelmi funkció, például helyesírás-javaslat. A vállalati gyakorlatban ide tartozik a csalásfigyelő, a hitelbírálati, a dokumentumfeldolgozó, a karbantartási igényt előrejelző, a gépi látáson alapuló minőségellenőrző, az önéletrajz-szűrő, a hálózati anomáliákat felismerő biztonsági eszköz és a rögzített tudásbázisra korlátozott ügyfélszolgálati válaszadó. Ez utóbbi nyelvi modellre épül, mégis célszoftver, ha rendeltetése rögzített, a nyelvi modell sajátos kockázataira ilyenkor a 2026/43. állásfoglalás „Az információbiztonsági kockázatok” fejezete is irányadó.

⁴ 7/2024. (VI. 24.) MK rendelet 2. § (3) bekezdés

⁵ Európai Bizottság, Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689, C(2025) 924, 2025. február 6. (digital-strategy.ec.europa.eu)

⁶ (EU) 2024/1689 rendelet (12) preambulumbekkezdés, 6. cikk (1) bekezdés a) pont és 3. cikk 14. pont

A Szervezet szerepe háromféle lehet, és ez az AI Act szerinti kötelezettségeket alapvetően meghatározza. Ha kész terméket vesz igénybe, amelyben a szállító tanította és üzemelteti a modellt, a Szervezet alkalmazó. Ha a modellt maga fejleszti vagy fejleszteti, és saját neve alatt helyezi üzembe, szolgáltatónak minősül akkor is, ha a rendszert kizárólag saját használatra szánja, mert az üzembe helyezés a saját használatra történő rendelkezésre bocsátást is magában foglalja.⁷ Ha külső, letöltött vagy API-n elérhető modellt épít saját alkalmazásba, és azt saját neve alatt helyezi üzembe, az alkalmazás szolgáltatója szintén a Szervezet. Nagy kockázatú rendszernél a szerep utólag is megváltozhat, ha a Szervezet a rendszert jelentősen módosítja, vagy rendeltetését úgy változtatja meg, hogy az nagy kockázatúvá válik.⁸

4.3 Az információbiztonsági kockázatok

A Szervezet a kockázatelemzés során a 3. melléklet fenyegetéskatalógusának elemeit vizsgálja,⁹ ahogyan azt a 2026/43. állásfoglalás azonos című fejezete is rögzíti. A célszoftver kockázatai abban különböznek az általános célú rendszerekétől, hogy a támadási felület nem a felhasználói prompt, hanem a tanítóadat, a modellfájl, a bemeneti csatorna és a beszállítói lánc.

4.3.1 Sértetlenség, a modell manipulálása és tévedése

A tanított modell viselkedését nem a forráskód, hanem a tanítóadat határozza meg. Aki a tanítóadatot befolyásolja, a modell döntéseit is befolyásolja. Ez az adatmérgezés, amely a hardver vagy szoftver hamisítása (22. tétel) és az információmanipuláció (23. tétel) körébe tartozik. Különösen kitett az a rendszer, amely működés közben folyamatosan tanul, mert a támadó a mindennapi bemenettel tudja elmozdítani a döntési határt. Az AI Act ezért külön nevesíti a visszacsatolási hurkok kockázatát.¹⁰ A modellmérgezés ugyanezt az előtanított összetevő manipulálásával éri el. A NIST 2025. évi taxonómiája a mérgezést, a kijátszást és a bizalmassági támadásokat tekinti a gépi tanulás elleni támadások fő típusainak.¹¹

A második sértetlenségi kockázat a kijátszás. A támadó olyan bemenetet állít elő, amelyet az ember helyesen ítél meg, a modell viszont tévesen, például apró módosításokkal jogosnak láttatott tranzakciót (29. és 35. tétel), és ezt a hagyományos sérülékenységvizsgálat nem tárja fel, mert a tanult összefüggésekben rejlik. A harmadik kockázat nem támadás, hanem a modell természetes tévedése. A pontosság az idő múlásával romlik, mert a valóság eltávolodik a tanítóadatoktól, és ez az eltolódás csendes, a rendszer nem jelez hibát, csak egyre többször dönt rosszul (27. és 47. tétel). Ha a döntést emberi ellenőrzés nélkül hajtják végre, a tévedés automatikusan kárrá válik, ezért az AI Act az automatizálási torzítás tudatosítását a nagy kockázatú rendszerekénél külön követelménnyé teszi.¹²

4.3.2 Bizalmasság, a tanítóadat és a modell mint vagyontárgy

⁷ (EU) 2024/1689 rendelet 3. cikk 3., 4. és 11. pont

⁸ (EU) 2024/1689 rendelet 25. cikk (1) bekezdés b) és c) pont

⁹ 7/2024. (VI. 24.) MK rendelet 2. § (4) bekezdés és 3. melléklet

¹⁰ (EU) 2024/1689 rendelet 15. cikk (4) bekezdés

¹¹ NIST AI 100-2 E2025, Adversarial Machine Learning, A Taxonomy and Terminology of Attacks and Mitigations, 2025. március (nvlpubs.nist.gov)

¹² (EU) 2024/1689 rendelet 14. cikk (4) bekezdés b) pont

A tanított modell részben megőrzi a tanítóadatot. Megfelelően megválasztott lekérdezésekkel megállapítható, hogy egy adott személy adatai szerepeltek-e a tanításban, egyes esetekben a tanítóadat részben rekonstruálható. Ha a Szervezet személyes adaton vagy üzleti titkon tanított modellt tesz kívülről elérhetővé, a védett információ nyilvánosságra kerülése (20. tétel) és a személyes adatokkal történő visszaélés (39. tétel) kockázata a modellen keresztül áll fenn, tömeges lekérdezéssel pedig a modell maga is lemásolható. Kész termék igénybevételekor a kockázat iránya fordított. A szállító a Szervezet adatait kapja, és a szerződési feltételek szerint ezeket saját modelljének továbbtanítására is felhasználhatja, így a Szervezet adatai egy másik ügyfélnél működő modellbe kerülhetnek, a tanítási célú felhasználás kikapcsolása gyakran külön nyilatkozatot igényel.

4.3.3 Rendelkezésre állás és ellátási lánc

A felhőben futó következtetés a külső szolgáltató zavara (12. tétel) esetén kiesik, és a modell helyben nem pótolható. A szállító a modellt bejelentés nélkül cserélheti vagy kivezetheti, ami a rendszer viselkedését megváltoztatja anélkül, hogy a Szervezet bármit módosított volna (19. tétel). A gépi tanulás ellátási láncja emellett hosszabb, mint a hagyományos szoftveré. A modellt előtanított modellekből, nyilvános adatkészletekből, nyílt forráskódú keretrendszerekből és jellemzően egy nagy modellszolgáltató szolgáltatásából állítják össze, és ezek bármelyike a beszállítói lánc zavarának (11. tétel) vagy a szoftver hamisításának (22. tétel) forrása lehet. Az előtanított modellek nyilvános tárhelyein rendszeresen találunk olyan modellfájlokat, amelyek betöltéskor kódot futtatnak, mert az elterjedt sorosítási formátumok ezt lehetővé teszik.¹³ A brit NCSC és a CISA közös iránymutatása ezért az ellátási lánc biztonságát és a modellek, adatok eredetének nyomon követését a biztonságos MI-fejlesztés alapelvei közé sorolja.¹⁴ A szervezeti kockázat mindezek mögött az, hogy gyakran senki sem felel a modellért, a szoftvert az üzleti terület veszi, az informatika üzemelteti, a modell viselkedését pedig egyik sem érti.

4.4 NIS2 előírások a kockázatok kezelésére

4.4.1 A törvényi keret

A szervezet vezetőjének a 2026/43. állásfoglalás „A törvényi keret” pontjában bemutatott feladatai (felmérés, nyilvántartás, kockázatmenedzsment, védelmi intézkedések, szabályzat, használatbavételi döntés, képzés, a közreműködők szerződéses kötelezése) a célszoftverre is vonatkoznak. Az alábbi rendelkezések különös súlyt kapnak:

A Kiberbiztonsági törvény a védelmet a rendszer teljes életciklusában írja elő, és az életciklust a tervezéstől a megszüntetésig határozza meg,¹⁵ a tanított modellenél ez a tanítóadat gyűjtésével kezdődik.

A törvény továbbfejlesztésnek tekinti a működő rendszer olyan fejlesztését, amely funkcionalitásának érdemi megváltozásával jár, vagy védelmének elvárt erősségére hatással

¹³ JFrog Security Research, Examining Malicious Hugging Face ML Models with Silent Backdoor, 2024. február (jfrog.com); CISA, NSA, FBI és partnerügynökségek, AI Data Security, 2025. május 22. (cisa.gov)

¹⁴ NCSC (UK), CISA és partnerügynökségek, Guidelines for Secure AI System Development, 2023. november 27. (ncsc.gov.uk)

¹⁵ 2024. évi LXIX. tv. 5. § (1) bekezdés és 4. § 26. pont

van,¹⁶ a modell újratanítása vagy cseréje ilyen lehet. Végül, ha a Szervezet közreműködőt vesz igénybe a rendszer létrehozásához, üzemeltetéséhez vagy karbantartásához, a vezető gondoskodik arról, hogy a kiberbiztonsági követelmények szerződéses kötelemként teljesüljenek,¹⁷ a modellt fejlesztő vagy üzemeltető szállító pedig közreműködő.

4.4.2 A rendelet védelmi intézkedései

A 7/2024. (VI. 24.) MK rendelet 2. mellékletének a 2026/43. állásfoglalás „A rendelet védelmi intézkedései” pontjában tárgyalt követelményei (külső rendszerek használata, szerződéses követelmények a külső szolgáltatóval szemben, felhasználói szoftvertelepítés, legszűkebb funkcionalitás, hozzáférés-ellenőrzés, naplózás, képzés) a célszoftverre is irányadók, ezeket nem ismételjük. Az alábbiakban azokat a követelményeket emeljük ki, amelyek a tanított modell sajátosságaira adnak választ. A kiindulópont a beszerzés, ahol a rendelet minden osztályban egységes biztonsági követelmények szerződéses rögzítését várja el.

16.7. A szervezet a beszerzési folyamat során - beleértve a fejlesztést, az adaptálást, a rendszerkövetést és a karbantartást is - a szerződéseiben egységes nyelvezetet alkalmaz, továbbá követelményként rögzíti az alábbiakat:

16.7.1. A funkcionális biztonsági követelményeket.

[...]

16.7.4. Az érintett EIR biztonsági osztályát és az ahhoz tartozó, illetve a szervezet által meghatározott további biztonsági követelmények teljesítéséhez szükséges védelmi intézkedéseket.

16.7.5. A biztonsággal kapcsolatos dokumentációs követelményeket.

[...]

16.7.9. A teljesítési kritériumokat. [7/2024. (VI. 24.) MK rendelet 2. melléklet 16.7. pont] (A, J, M)

A tanított modellnél a funkcionális biztonsági követelmény a modell pontosságának és stabilitásának mérhető elvárása, a dokumentációs követelmény a modell rendeltetésének, tanítóadat-forrásainak, korlátainak és ismert gyengeségeinek leírása, a teljesítési kritérium pedig az átvételi teszt ismert eredményű adatkészleten. A Szervezet a fejlesztőtől minden osztályban megköveteli a védelmi intézkedések funkcionális leírását és a rendszer dokumentációját, jelentős és magas osztályban a tervezési és megvalósítási információkat is.¹⁸

Egyedi fejlesztésnél a rendelet jelentős és magas osztályban a fejlesztőtől biztonságértékelési tervet, egység-, integrációs-, rendszer- és regressziós tesztelést, az eredmények dokumentálását és ellenőrizhető hibajavítási folyamatot követel meg.¹⁹ Gépi tanulásnál a regressziós teszt a modell új verziójának összevetése a korábbival ugyanazon adatkészleten, a biztonsági tesztelés pedig a kijátszási és mérgezési ellenálló képesség vizsgálata, amelyhez a NIST taxonómiája és a MITRE ATLAS keretrendszere kész fenyegetéslisát ad. A fejlesztői változáskövetés ugyanezen

¹⁶ 2024. évi LXIX. tv. 4. § 96. pont

¹⁷ 2024. évi LXIX. tv. 6. § (5) bekezdés d) pont

¹⁸ 7/2024. (VI. 24.) MK rendelet 2. melléklet 16.8. pont (A, J, M), 16.9. pont (J, M) és 16.15. pont (A, J, M)

¹⁹ 7/2024. (VI. 24.) MK rendelet 2. melléklet 16.66. pont (J, M)

osztályokban megköveteli, hogy a fejlesztő csak a Szervezet által jóváhagyott változtatásokat hajtsa végre, és dokumentálja azok biztonsági hatását,²⁰ ez a modell újratanítására is kiterjed.

A modell és a tanítóadat rendszerelem, ezért a konfigurációkezelés követelményei vonatkoznak rájuk. A rendszerelem-leltár minden osztályban kötelező,²¹ a modellfájl verziója, a tanítóadat-készlet azonosítója és a következő környezet ide tartozik. A változáskezelés jelentős és magas osztályban megköveteli a módosítások biztonsági szempontú jóváhagyását és dokumentálását, a biztonsági hatásvizsgálat pedig minden osztályban kötelező a változtatás bevezetése előtt, magas osztályban elkülönített tesztkörnyezetben.²² A szállítói modellfrissítésnél a hatásvizsgálat a frissítés utáni ellenőrző tesztet jelenti. Az éles adatok tesztkörnyezeti használatának jóváhagyása opcionális,²³ a tanítóadat azonban jellemzően éles adat, ezért a tanítókönyvet védelme indokolt.

A sértetlenség védelmére a rendelet jelentős és magas osztályban sértetlenségellenőrző eszközöket ír elő a meghatározott szoftverek és információk jogosulatlan változtatásának észlelésére, és ugyanezen osztályokban kötelező a bemeneti információ ellenőrzése.²⁴ A modellfájl és a tanítóadat lenyomatának ellenőrzése betöltéskor az előbbit teljesíti, amely megakadályozza a fájlok jogosulatlan utólagos módosítását (így a fájlok hiteles forrásának igazolásával hozzájárul a modellszolgáltatás elleni védelemhez), a bemeneti adatok formátumának és tartományának ellenőrzése pedig a kijátszási kísérletek egy részét kiszűri. A kimeneti információ ellenőrzése opcionális,²⁵ a modell kimenetének elvárt tartományhoz mérése azonban az egyetlen technikai eszköz a csendes teljesítménycsökkenés felismerésére, ezért bevezetését javasoljuk. Az ellátási lánc körében a rendszerelemek hitelességének ellenőrzése és a hamisított elemek kiszűrése minden osztályban kötelező, az eredet és a módosíthatatlanság hitelesítése opcionális.²⁶ Az előtanított modell aláírásának, lenyomatának és forrásának ellenőrzése ezeket tölti meg tartalommal. A beszállítók rendszeres értékelése jelentős és magas osztályban kötelező,²⁷ és a mögöttes modellszolgáltatóra is kiterjed.

A naplózás körében a Szervezet meghatározza a naplózandó eseményeket, és indokolja, miért alkalmasak a biztonsági események utólagos kivizsgálására.²⁸ A modell döntéseinek (bemenet, modellverzió, kimenet, emberi felülbírálat) naplózása nélkül a manipuláció nem vizsgálható ki. A mentés minden osztályban kiterjed a rendszerszintű információkra, a modellfájl és a tanítóadat mentése nélkül a rendszer nem állítható helyre, a szállító távoli hozzáférést a modell frissítéséhez pedig jóvá kell hagyni és naplózni kell.²⁹ A folyamatos felügyelet minden osztályban kötelező,³⁰ a rendszerszintű metrikák közé a modell teljesítménymutatói tartoznak, ezek adják meg a vezetőnek a használat folytatásáról szóló döntés alapját.

²⁰ 7/2024. (VI. 24.) MK rendelet 2. melléklet 16.58. pont (J, M)

²¹ 7/2024. (VI. 24.) MK rendelet 2. melléklet 6.36. pont (A, J, M)

²² 7/2024. (VI. 24.) MK rendelet 2. melléklet 6.7. pont (J, M), 6.15. pont (A, J, M) és 6.16. pont (M)

²³ 7/2024. (VI. 24.) MK rendelet 2. melléklet 16.5. pont (O)

²⁴ 7/2024. (VI. 24.) MK rendelet 2. melléklet 18.42. pont (J, M) és 18.59. pont (J, M)

²⁵ 7/2024. (VI. 24.) MK rendelet 2. melléklet 18.77. pont (O)

²⁶ 7/2024. (VI. 24.) MK rendelet 2. melléklet 19.23. pont (A, J, M), 19.11. pont (O) és 19.12. pont (O)

²⁷ 7/2024. (VI. 24.) MK rendelet 2. melléklet 19.16. pont (J, M)

²⁸ 7/2024. (VI. 24.) MK rendelet 2. melléklet 4.2. pont (A, J, M)

²⁹ 7/2024. (VI. 24.) MK rendelet 2. melléklet 7.35. pont (A, J, M) és 10.11. pont (A, J, M)

³⁰ 7/2024. (VI. 24.) MK rendelet 2. melléklet 5.15. pont (A, J, M)

4.5 AI Act előírások a kockázatok kezelésére

A hatály, az időbeli alkalmazás és a hazai intézményrendszer kérdéseit a 2026/43. állásfoglalás „Hatály és időbeli alkalmazás” pontja tárgyalja, az MI-jártassági kötelezettséget és az átláthatósági szabályokat ugyanott az „MI-jártasság” és az „Átláthatóság” pont ismerteti, ezek a célszoftverekre is vonatkoznak. A célszoftvereknél a kockázati szint szerinti különbségtétel a lényeges, mert az AI Act kötelezettségei ettől függenek.

4.5.1 A kockázati szintek és a célszoftverek

A rendelet négy kockázati szintet ismer. A tiltott gyakorlatok közül a célszoftverek szempontjából a munkahelyi érzelmefelismerés a legjelentősebb. A rendelet 2025. február 2-től tiltja a természetes személyek érzelmeiből következtetést levonó MI-rendszerek használatát a munkahelyek és az oktatási intézmények területén, az orvosi vagy biztonsági okból történő használat kivételével.³¹ A tilalom a hívások hangulatelemzésétől a videointerjú-értékelő szoftverig több kész termékre kiterjedhet.

Nagy kockázatú a rendszer, ha az I. melléklet szerinti termék biztonsági alkotórésze, vagy a III. melléklet valamely területére tartozik.³² A vállalati célszoftverek közül ide tartozik a biometrikus azonosítás, a kritikus digitális infrastruktúra és a közműszolgáltatások irányításában biztonsági alkotórészként használt rendszer, a toborzásra, előléptetésre, feladat kiosztásra és teljesítményértékelésre használt rendszer, a hitelképesség értékelése, valamint az élet- és egészségbiztosítási kockázatértékelés. A csalásfigyelő rendszert a rendelet kifejezetten kiveszi a hitelképességi körből.³³ A III. melléklet szerinti rendszer kivételesen nem nagy kockázatú, ha jól körülhatárolt eljárási feladatot lát el, korábbi emberi tevékenység eredményét javítja, döntési mintáktól való eltérést észlel emberi felülvizsgálat mellett, vagy előkészítő feladatot végez. A természetes személyekre vonatkozó profilalkotás azonban mindig nagy kockázatú, az értékelést pedig a szolgáltatónak dokumentálnia kell,³⁴ alkalmazóként a Szervezet ezt a szállítótól kérje el. A harmadik szintet az átláthatósági kötelezettség alá eső rendszerek adják, a célszoftverek közül az érzelmefelismerő és a biometrikus kategorizáló rendszer, amelynek alkalmazója tájékoztatja az érintett személyeket.³⁵ Minden más célszoftver, így a csalásfigyelő, az előrejelző és a dokumentumfeldolgozó rendszer minimális kockázatú, amelyre az AI Act az MI-jártasságon túl kötelezettséget nem ír elő, a NIS2 követelmények azonban teljes körűen vonatkoznak rá.

4.5.2 Kockázatcsökkentő intézkedések nagy kockázatú rendszernél

A nagy kockázatú rendszer szolgáltatója számára a rendelet a kiberbiztonságot tervezési követelménnyé teszi.

(5) A nagy kockázatú MI-rendszereknek reziliensnek kell lenniük a jogosulatlan harmadik felek arra irányuló kísérleteivel szemben, hogy a

³¹ (EU) 2024/1689 rendelet 5. cikk (1) bekezdés f) pont és 113. cikk a) pont, az (EU) 2026/1744 rendelettel megállapított szöveg

³² (EU) 2024/1689 rendelet 6. cikk (1) és (2) bekezdés

³³ (EU) 2024/1689 rendelet III. melléklet 1., 2., 4. és 5. pont

³⁴ (EU) 2024/1689 rendelet 6. cikk (3) és (4) bekezdés

³⁵ (EU) 2024/1689 rendelet 50. cikk (3) bekezdés

rendszer sebezhetőségeinek kiaknázása révén megváltoztassák a rendszer használatát, kimeneteit vagy teljesítményét.

A nagy kockázatú MI-rendszerek kiberbiztonságának biztosítását célzó műszaki megoldásoknak megfelelőnek kell lenniük a releváns körülmények és a kockázatok szempontjából.

*Az MI-specifikus sebezhetőségek kezelésére szolgáló műszaki megoldásoknak adott esetben magukban kell foglalniuk azon intézkedéseket, amelyek a tanítóadat-készlet manipulálását megkísérlő támadásoknak (adatmérgezés) vagy a tanítás során használt előtanított összetevők manipulálását megkísérlő támadásoknak („modellmérgezés”), a modell hibájának előidézésére szolgáló bemeneteknek (támadó szempontú példák vagy modellkijátszás), a bizalmasság elleni támadásoknak vagy a modellhibáknak a megelőzésére, felderítésére, az azokra való reagálásra, azok megoldására és ellenőrzésére irányulnak.
[(EU) 2024/1689 rendelet 15. cikk (5) bekezdés]*

Ez a rendelkezés pontosan azokat a kockázatokat nevesíti, amelyeket a fenti fejezet leírt. A szolgáltatót emellett kockázatkezelési rendszer, a tanítóadatok minőségi követelményei, automatikus naplózás, a használati utasításban a pontosság, a stabilitás és a kiberbiztonság várható szintjének közlése, valamint az emberi felügyeletet lehetővé tevő tervezés terheli.³⁶ Alkalmazóként a Szervezet e követelményeket a beszerzés során mérceként használja, a szállítótól az EU-megfelelőségi nyilatkozatot és a használati utasítást kéri. Ha a nagy kockázatú rendszer egyben a kiberrezilienciáról szóló rendelet hatálya alá tartozó, digitális elemekkel rendelkező termék, a gyártó a 15. cikk kiberbiztonsági követelményeit e rendelet alapvető követelményeinek teljesítésével is igazolhatja.³⁷

Az alkalmazó saját kötelezettségei a rendszer üzemeltetésére vonatkoznak. Biztosítja, hogy a rendszert a használati utasítás szerint használják, az emberi felügyeletet a szükséges szakértelemmel, képzéssel és hatáskörrel rendelkező személyekre bízva, gondoskodik a bemeneti adatok relevanciájáról, nyomon követi a működést, kockázat gyanúja esetén tájékoztatja a szolgáltatót és a piacfelügyeleti hatóságot, és felfüggeszti a használatot, a súlyos váratlan eseményt pedig haladéktalanul jelenti. Az automatikusan generált naplókat legalább hat hónapig megőrzi, munkáltatóként a bevezetés előtt tájékoztatja a munkavállalók képviselőit és az érintett munkavállalókat, a személyekre vonatkozó döntéseknél pedig az érintett természetes személyeket.³⁸ A hitelképességet és a biztosítási kockázatot értékelő rendszer alkalmazója a bevezetés előtt alapvető jogi hatásvizsgálatot végez.³⁹ Ezek a kötelezettségek a III. melléklet szerinti rendszerekre 2027. december 2-től, az I. melléklet szerinti rendszerekre 2028. augusztus 2-től alkalmazandók. Az ezen időpontok előtt üzembe helyezett rendszerekre a rendelet csak

³⁶ (EU) 2024/1689 rendelet 9. cikk (1) bekezdés, 10. cikk (1) bekezdés, 12. cikk (1) bekezdés, 13. cikk (3) bekezdés b) pont ii. alpont és 14. cikk (1) bekezdés

³⁷ Az Európai Parlament és a Tanács (EU) 2024/2847 rendelete (kiberrezilienciáról szóló rendelet) 12. cikk (1) bekezdés

³⁸ (EU) 2024/1689 rendelet 26. cikk (1), (2), (4), (5), (6), (7) és (11) bekezdés

³⁹ (EU) 2024/1689 rendelet 27. cikk (1) bekezdés

akkor alkalmazandó, ha ezt követően a tervezésükben jelentős változás következik be, a hatóságok által használt rendszereknél pedig a megfelelést 2030. augusztus 2-ig kell biztosítani.⁴⁰ Az alkalmazói kötelezettségek megsértése miatt Magyarországon az MI piacfelügyeleti hatóság, a pénzügyi szolgáltatásokhoz közvetlenül kapcsolódó rendszereknél a Magyar Nemzeti Bank járhat el.⁴¹ Ha a Szervezet nagy kockázatú célszoftvert saját neve alatt fejleszt és helyez üzembe, a szolgáltatói kötelezettségek teljes köre terheli, a 15. cikk (5) bekezdése szerinti műszaki intézkedéseket ilyenkor magának kell megvalósítania.

5. FELHASZNÁLT FORRÁSOK

- Az Európai Parlament és a Tanács (EU) 2022/2555 irányelve (NIS 2 irányelv); az (EU) 2024/1689 rendelet (AI Act), az (EU) 2026/1744 rendelettel módosított, 2026. július 27-től hatályos egységes szerkezetű szöveg; az (EU) 2024/2847 rendelet (CRA)
- 2024. évi LXIX. törvény Magyarország kiberbiztonságáról (Kiberbiztonsági tv.); 7/2024. (VI. 24.) MK rendelet a biztonsági osztályba sorolás követelményeiről és a védelmi intézkedésekről; 418/2024. (XII. 23.) Korm. rendelet
- 2025. évi LXXV. törvény az Európai Unió mesterséges intelligenciáról szóló rendeletének magyarországi végrehajtásáról; 344/2025. (X. 31.) Korm. rendelet; VFP2 KiberVéd jogi állásfoglalás 2026/43
- Európai Bizottság, Guidelines on the definition of an artificial intelligence system (2025. február 6.); NIST AI 100-2 E2025 (2025. március); NCSC (UK) és CISA, Guidelines for Secure AI System Development (2023. november); CISA, NSA, FBI, AI Data Security (2025. május); MITRE ATLAS

6. KAPCSOLÓDÓ ÁLLÁSFOGLALÁSOK ÉS IRÁNYMUTATÁSOK

- MKIK VFP3 KiberVéd jogi állásfoglalás 2026/43 (általános célú MI-rendszerek, MI-ügynökök, MI-jártasság, átláthatóság);
- ENISA, Multilayer Framework for Good Cybersecurity Practices for AI (2023);
- NIST AI 100-2 E2025;
- MITRE ATLAS;
- ISO/IEC 42001, 23894 és 27090, a gépi tanulás kiberbiztonságának nemzetközi keretei.

⁴⁰ (EU) 2024/1689 rendelet 113. cikk c) pont és 111. cikk (2) bekezdés, az (EU) 2026/1744 rendelettel megállapított szöveg

⁴¹ 2025. évi LXXV. törvény 1. § c) pont cb) alpont, 5. § (1) bekezdés és 7. § (1) bekezdés; 344/2025. (X. 31.) Korm. rendelet 1. § (1) bekezdés b) pont