

Milyen információbiztonsági kockázatok merülnek fel MI-ügynökök vállalati használatánál? NIS2 és AI Act előírások a kockázatok kezelésére

VFP3 KiberVéd - NIS2 Audit felkészítés jogi
állásfoglalás 2026/58

VERZIÓKÖVETÉS

Verziószám	Dátum	Módosította	Módosítások leírása
1.0	2026.09.29	Kálmán és Társa Ügyvédi Iroda	Első kiadott verzió

1. VEZETŐI ÖSSZEFOGLALÓ

Az MI-ügynök nem válaszol, hanem cselekszik. A nyelvi modellre épülő ügynök a kapott cél érdekében maga választja meg a lépéseket, leveleket olvas és küld, fájlokat kezel, alkalmazásokat hív meg és kódot futtat, jellemzően egy munkatárs vagy egy technikai fiók jogosultságaival, miközben bemenetének jelentős része ellenőrizetlen külső forrásból érkezik.

A legsúlyosabb kockázat a cél eltérítése, amikor a levélbe vagy weboldalba rejtett utasítás hatására az ügynök a saját jogosultságaival szivárogtat ki adatot vagy indít műveletet. Ez a gyengeség a modell szintjén valószínűleg soha nem szüntethető meg teljesen. Ehhez járul a túlzott jogosultság, a csatlakozók és bővítmények ellátási lánc, a visszafordíthatatlan műveletek és az engedély nélkül telepített ügynökök kockázata.

A NIS2 hazai szabályozása erre a hozzáférés-felügyelet, az információáramlás, a rendszer- és kommunikációvédelem, a naplózás és az ellátási lánc meglévő követelményeivel válaszol, a 7/2024. (VI. 24.) MK rendelet pedig kifejezetten számol a felhasználók nevében eljáró folyamatokkal. A védelmet a modellen kívüli, determinisztikus korlátok adják, vagyis saját identitás, szűk jogosultság, korlátozott információáramlás, elszigetelt futtatás, emberi jóváhagyás, naplózás és vészleállítás.

Az AI Act az ügynököt nem tekinti külön kategóriának, a kötelezettségek a felhasználás kockázati szintjétől függenek. A KKV-k ügynökeinek többsége minimális kockázatú, a természetes személlyel kapcsolatba lépő ügynöknek azonban 2026. augusztus 2. óta fel kell fednie mesterséges voltát és azt, hogy kinek a nevében jár el. A munkavállalók értékelésére, feladatkiosztásra vagy hitelbírálatra használt ügynökre 2027. december 2-tól a nagy kockázatú rendszerek követelményei, köztük az emberi felügyelet és a biztonságos leállítás lehetősége is vonatkoznak.

2. A VIZSGÁLANDÓ KÉRDÉS

Egy NIS2¹-kötelezett szervezet (a továbbiakban mint „Szervezet”) munkatársai MI-ügynököket használnak, a Szervezet pedig ügynököket kíván bevezetni levelezési, ügyfélszolgálati, adminisztratív és programozási feladatokra. A vizsgálandó kérdés az, hogy mit értünk MI-ügynök alatt, melyek a legismertebb ügynök-keretrendszerek, milyen információbiztonsági kockázatokkal jár az ügynökök használata kis- és középvállalati környezetben, továbbá hogy a hazai NIS2 szabályozás és az AI Act milyen kockázatcsökkentő intézkedéseket ír elő.

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

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)

kezelésére” című állásfoglalás² (a továbbiakban mint „2026/43. állásfoglalás”), a meghatározó MI-komponenst tartalmazó célszoftvereket a 2026/45. számú, „Milyen információbiztonsági kockázatok merülnek fel meghatározó MI-komponenst tartalmazó célszoftverek vállalati használatánál? NIS2 és AI Act előírások a kockázatok kezelésére” című állásfoglalás³ (a továbbiakban mint „2026/45. állásfoglalás”) tárgyalja, az ott kifejtett közös alapokra a fejezet megjelölésével hivatkozunk. Nem tárgya az állásfoglalásnak a saját modell tanítása, az AI Act teljes szolgáltatói megfelelési rendszere, valamint az adatvédelmi, szerzői jogi és felelősségi kérdések elemzése.

3. GYAKORLATI JAVASLATOK

3.1 Alapvető szabályok az MI-ügynökök használatára

A jogszabályok nem írnak elő külön ügynökszabályzatot. A 2026/43. állásfoglalás „Alapvető szabályok az általános célú MI rendszerek használatára” pontjában javasolt szabályokat ügynököknél az alábbiakkal javasoljuk kiegészíteni.

- Nyilvántartás és felelős gazda, minden ügynök felvétele a rendszernyilvántartásba az általa elért rendszerekkel, eszközökkel és kapcsolatokkal együtt, és egy üzleti gazda kijelölése, aki az ügynök feladataért és jogosultságaiért felel.
- Saját identitás és szűk jogosultság, az ügynök saját azonosítóval és rövid élettartamú, szűk hatókörű hozzáférési tokenekkel működik, és nem kaphat több jogot, mint akinek a nevében eljár.
- A „halálos hármas” megbontása, egy ügynök ugyanabban a munkamenetben ne dolgozzon fel ellenőrizetlen tartalmat, ne férjen hozzá bizalmas adathoz és ne kommunikáljon kifelé egyszerre, vagy ha mégis, a kimenő kapcsolatokat engedélyezési lista korlátozza, és emberi jóváhagyás kell.
- Emberi jóváhagyás a visszafordíthatatlan műveletekhez, küldés, törlés, fizetés, jogosultság-módosítás és az éles rendszer módosítása csak kifejezett megerősítés után, automatikus jóváhagyási üzemmód nélkül.
- Jóváhagyott eszközök és elszigetelt futtatás, MCP-szerver, skill és csatlakozó csak engedélyezési listáról, rögzített verzióval használható, a kódot futtató vagy böngészőt vezérlő ügynök pedig elkülönített környezetben, éles hitelesítő adatok nélkül működik.
- Naplózás és vészleállítás, az eszközhívások, a paraméterek, a jóváhagyások és az eredmények visszakereshető naplózása, valamint kipróbált eljárás az ügynök azonnali leállítására és hozzáféréseinek visszavonására.
- Saját telepítésű ügynökök korlátozása, személyi asszisztens, böngésző- vagy kódoló ügynök munkaeszközön csak az informatikai terület engedélyével használható.

² Közzétéve 2026. szeptember 17-én (nis2.mkik.hu/jogi-allasfoglalasok/45)

³ Közzétéve 2026. szeptember 18-án (nis2.mkik.hu/jogi-allasfoglalasok/48)

3.2 Jó gyakorlatok és megoldások KKV-k számára

Egy kis- vagy középvállalkozás ritkán fejleszt ügynököt, jellemzően a már használt irodai, CRM- vagy ügyfélszolgálati szolgáltatásban kapcsolja be, vagy alacsony kódolású platformon rakja össze. Az első lépés ezért annak feltárása, hol aktív ügynökfunkció, mely csatlakozók és külső alkalmazások kaptak OAuth-hozzáférést, és ki adhat ilyen jóváhagyást. A felhasználói jóváhagyást célszerű rendszergazdai engedélyhez kötni, a gyári alapbeállításokat pedig felülvizsgálni.

A második a fokozatosság. Az első ügynök csak olvasási jogot igénylő, alacsony kockázatú feladatot kapjon, írási vagy küldési jogot pedig csak a bevált működés után. A brit NCSC kis- és középvállalkozásoknak is szóló, 2026. augusztusi tanácsai hét óvintézkedést neveznek meg (a hibalehetőségek előzetes végiggondolása, gondos utasítások, arányos felügyelet, homokozó, naplózás, a tevékenység beazonosíthatósága, vészleállítás).⁴

A harmadik a saját telepítésű eszközök kezelése. A francia CERT-FR 2026 áprilisában azt javasolta, hogy az OpenClaw típusú autonóm személyi asszisztenseket munkaállomásra ne telepítsék.⁵ A helyi rendszergazdai jog megvonása, a böngészőbővítmények központi engedélyezése és a saját üzemeltetésű automatizáló szerverek internet felőli elzárása kis költséggel sok védelmet ad.

3.3 A javasolt eljárás

- a) Felmérés, a használt és a beépített ügynökök, csatlakozók, OAuth-engedélyek és saját telepítésű ügynökök feltárása.
- b) Besorolás és döntés, ügynökönként a feladat, az elért rendszerek és a visszafordíthatatlan műveletek rögzítése, a biztonsági osztály és az AI Act szerinti kockázati szint megállapítása, vezetői döntés a használatbavételről.
- c) Tervezés, a fenyegetéskatalógus és az OWASP ügynöki listája alapján a jogosultságok, a jóváhagyási pontok és az információáramlási korlátok megtervezése.
- d) Fokozatos bevezetés, olvasási joggal kezdve, a közvetett utasításinjektálással szembeni ellenállás tesztelésével, minden bővítést változásként kezelve.
- e) Üzemeltetés és felülvizsgálat, a naplók áttekintése, a szokatlan eszközhívások riasztása, a jogosultságok negyedéves és a kockázatelemzés évenkénti felülvizsgálata, a vészleállítás rendszeres kipróbálása.

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

⁴ NCSC (UK), Managing the cyber risk of agentic AI, 2026. augusztus 20. (ncsc.gov.uk)

⁵ CERT-FR, CERTFR-2026-ACT-016, Vulnérabilités et risques des produits d’automatisation par IA agentic sur les postes de travail, 2026. április 13. (cert.ssi.gouv.fr)

rendelkezésében lévő ügynök pedig akkor is az elektronikus információs rendszere vagy annak eleme, ha a szolgáltató felhőjében fut. Ügynököknél ehhez két kiindulópont járul.

Az első, hogy az ügynök hozzáférési útvonal. Minden rendszer, amelyet a hitelesítő adataival elér, rajta keresztül is támadható. A biztonsági osztályba sorolás a bizalmasságot, a sértetlenséget és a rendelkezésre állást érintő kockázaton alapul,⁶ ezért álláspontunk szerint az ügynök nem sorolható alacsonyabb osztályba, mint az a rendszer, amelyben írási, törlési vagy küldési joga van.

A második, hogy az ügynök a felhasználó nevében eljáró folyamat. A 7/2024. (VI. 24.) MK rendelet a legkisebb jogosultság elvénél kifejezetten nevesíti az ilyen folyamatokat (a 2026/43. állásfoglalásban idézett 2.60. pont), az ügynökre tehát a hozzáférés-felügyelet minden szabálya vonatkozik. Az ügynök nem jogalany, cselekményei a Szervezet cselekményei.

4.2 A fogalmak pontosítása

4.2.1 Az MI-ügynök fogalma

Az MI-ügynök fogalmát sem az AI Act, sem a magyar jog nem határozza meg, jellemzőit a 2026/43. állásfoglalás „MI-ügynök” pontja az Európai Bizottság tájékoztatása alapján rögzítette. Az ISO/IEC 22989 szabvány szerint az MI-ügynök olyan automatizált entitás, amely érzékeli a környezetét, reagál rá, és céljai elérése érdekében cselekszik. A NIST leírása szerint az ügynökrendszer legalább egy generatív MI-modellből és az azt körülvevő vezérlő szoftverből áll, amely a modellt eszközökkel látja el. Az Anthropic gyakran hivatkozott megkülönböztetése szerint munkafolyamatról beszélünk, ha a lépéseket előre megírt kód vezérli, ügynökről pedig akkor, ha a modell maga irányítja a folyamatot és az eszközhasználatot.⁷

Ezen állásfoglalás alkalmazásában MI-ügynök az a nyelvi modellre épülő MI-rendszer, amely a megadott cél érdekében maga határozza meg a lépések sorrendjét, és olyan eszközöket használ, amelyekkel adatot érhet el, rendszerekben változást idézhet elő vagy kifelé kommunikálhat. A csevegőasszisztensztől és a célszoftvertől éppen az különbözteti meg, hogy a modell maga választ cselekvést és eszközt. Vegyes esetben a 2026/45. állásfoglalás és ezen állásfoglalás szempontjai együtt irányadók. Az ügynök öt eleme (a modell, a vezérlő réteg, az eszközök és csatlakozók, a memória, valamint az identitás és a jogosultságok) mind önálló támadási felület. Az önállóság fokozatos, az ügynök lehet javaslattevő, lépésenként jóváhagyást kérő vagy önállóan végrehajtó, és a kockázat ezzel arányosan nő.

4.2.2 A legismertebb ügynök-keretrendszerek

Az ügynökök építőkövei négy csoportba sorolhatók. A legismertebb fejlesztői keretrendszerek a LangChain és a LangGraph, a Microsoft Agent Framework, a CrewAI, az OpenAI Agents SDK, a Google Agent Development Kit és a Claude Agent SDK. Alacsony kódolású platformon programozás nélkül is építhető ügynök (Microsoft Copilot Studio, n8n, Make, Zapier). A harmadik csoport az üzleti alkalmazásokba épített ügynököké (Microsoft 365 Copilot, Salesforce Agentforce, ServiceNow). A negyedik a végfelhasználóké, ide tartoznak a böngészőt vezérlő

⁶ 2024. évi LXIX. tv. 10. § (1) bekezdés

⁷ ISO/IEC 22989 (2022), Artificial intelligence concepts and terminology, 3.1.1. pont; NIST CAISI, Request for Information Regarding Security Considerations for Artificial Intelligence Agents, Federal Register 2026-00206, 2026. január 8.; Anthropic, Building effective agents, 2024. december 19. (anthropic.com)

(ChatGPT agent, Claude in Chrome, Perplexity Comet) és a kódoló ügynökök (Claude Code, GitHub Copilot, OpenAI Codex, Cursor), valamint a helyben futó személyi asszisztensek, köztük a 2026 elején gyorsan elterjedt OpenClaw. Az ügynököt és az eszközöket a Model Context Protocol (MCP), az ügynököket egymással az Agent2Agent (A2A) protokoll kapcsolja össze.⁸ A keretrendszer semleges építőelem, a kockázatot az ügynöknek adott eszközök, jogosultságok és alapbeállítások határozzák meg.

4.3 Az információbiztonsági kockázatok

A Szervezet a kockázatelemzés során a 7/2024. (VI. 24.) MK rendelet 3. mellékletének fenyegetéskatalógusát vizsgálja, ahogy azt a 2026/43. és a 2026/45. állásfoglalás azonos című fejezete rögzíti. Ügynököknél ezt célszerű az OWASP 2025 decemberében közzétett ügynöki kockázati listájával kiegészíteni, amely tíz kockázatot nevez meg ASI01-től ASI10-ig. Öt angolszász ország kiberbiztonsági hatóságainak 2026. április 30-i közös útmutatója szerint a módszerek és szabványok kiforrásáig számolni kell az ügynökök váratlan viselkedésével.⁹ Az 1. melléklet a kockázatok a fenyegetéskatalógus tételeivel és a védelmi intézkedésekkel is megfelelteti.

4.3.1 A kis- és középvállalati környezet sajátosságai

Az Eurostat szerint 2025-ben az uniós, legalább tíz főt foglalkoztató vállalkozások 20 százaléka használt mesterséges intelligenciát, a magyar kis- és középvállalkozásoknak 9,6 százaléka. Az ügynök azonban gyakran nem tudatos döntéssel, hanem a már használt szolgáltatások új funkciójaként vagy a munkatárs saját eszközeként jelenik meg. A Microsoft által megkérdezett adatbiztonsági vezetők szerint a munkavállalók 29 százaléka használt már nem engedélyezett ügynököt.¹⁰ A KKV-knál ritkán működik önálló biztonsági felügyelet, az ügynökök pedig gyakran gyári beállításokkal, a felhasználó vagy egy kiterjedt jogú rendszergazdai fiók jogosultságaival futnak.

Egy levelezést kezelő ügynök ráadásul természeténél fogva egyesíti azt a három tulajdonságot, amelyet Simon Willison „halálos hármasként” (lethal trifecta) nevezett. Hozzáfér bizalmas adathoz, ellenőrizetlen tartalmat dolgoz fel, és kifelé kommunikálhat, így a támadó ráveheti a bizalmas adat kiküldésére. A Meta „kettes szabálya” ebből indul ki. Egy ügynök egy munkamenetben az ellenőrizetlen bemenet feldolgozása, az érzékeny adat vagy rendszer elérése, valamint az állapotváltoztatás vagy külső kommunikáció közül legfeljebb kettőre lehet képes, különben felügyelet nélkül nem működhet.¹¹

⁸ LangChain és LangGraph 1.0 (2025. október); Microsoft Agent Framework 1.0, az AutoGen és a Semantic Kernel utódja (2026. április); OpenAI Agents SDK (2025. március); Google ADK (2025. április); Claude Agent SDK (2025. szeptember); OpenClaw (2026. január); az MCP-t 2025 decembere óta a Linux Foundation Agentic AI Foundation gondozza

⁹ OWASP GenAI Security Project, OWASP Top 10 for Agentic Applications for 2026, 2025. december 9. (genai.owasp.org); ASD ACSC, CISA, NSA, CCCS, NCSC-NZ és NCSC-UK, Careful Adoption of Agentic AI Services, 2026. április 30.; további keretrendszerek a MITRE ATLAS, a CSA MAESTRO és a Microsoft Taxonomy of Failure Modes in Agentic AI Systems

¹⁰ Eurostat, Use of artificial intelligence in enterprises, isoc_eb_ai, 2025. évi adatok (ec.europa.eu/eurostat); Microsoft Security, Cyber Pulse, 2026. február 10. (microsoft.com)

¹¹ Simon Willison, The lethal trifecta for AI agents, 2025. június 16. (simonwillison.net); Meta, Agents Rule of Two, A Practical Approach to AI Agent Security, 2025. október 31. (ai.meta.com)

4.3.2 A cél eltérítése és a közvetett utasításinjektálás

Az OWASP listájának első helyén a cél eltérítése (ASI01) áll. A támadó levélben, dokumentumban, weboldalon vagy webes űrlapon helyez el rejtett utasítást, amelyet az ügynök munka közben beolvas és végrehajt. A NIST tesztjében az újabb támadási módszerek ennek sikerességét 11 százalékról 81 százalékra növelték. A következmény itt nem téves válasz, hanem végrehajtott művelet, vagyis a rendszer engedély nélküli használata (31. tétel) és a védett információ nyilvánosságra kerülése (20. tétel). Az EchoLeak sérülékenységnél (CVE-2025-32711) a Microsoft 365 Copilot egy beérkező levél rejtett utasítására kattintás nélkül juttatott ki belső adatot. A ForcedLeak esetben a Salesforce Agentforce egy webes űrlapba írt utasításra küldte ki a CRM adatait, a ShadowLeak esetben pedig az adat a szolgáltató felhőjéből távozott, így a Szervezet eszközei nem láthatták. A hibákat javították, a mintázat azonban általános. Az OpenAI szerint a böngészőügynökök utasításinjektálása a csaláshoz hasonlóan valószínűleg soha nem lesz teljesen megoldva.¹² A védelem ezért arra épülhet, hogy az eltérített ügynök se férjen hozzá többhöz és ne tehessen többet, mint amit a feladata indokol.

4.3.3 Túlzott jogosultság és identitás

Az ügynök jogosultságai a gyakorlatban szinte mindig tágabbak a szükségesnél (ASI02, ASI03). Az OWASP a túlzott cselekvési képesség gyökérokaként a túlzott funkcionalitást, jogosultságot és önállóságot nevezi meg. A SailPoint 2025. évi felmérésében a szervezetek 80 százaléka tapasztalt nem szándékolt ügynöki műveletet, 23 százalékuknál az ügynököt a hozzáférési adatai kiadására vették rá. 2025 augusztusában a támadók a Salesloft Drift csevegőügynök ellopott OAuth-tokenjeivel töltötték le ügyféladatokat, a Google több mint 700 potenciálisan érintett szervezetről tudott. A ServiceNow platformján az alapértelmezésben engedélyezett ügynök-felfedezés révén egy alacsony jogosultságú felhasználó utasítása magasabb jogosultságú ügynököt vehetett igénybe.¹³ Ez a „megtévesztett megbízott” (confused deputy) klasszikus esete, amely az engedélyekkel való visszaélés (33. tétel), az ellopott tokenek felhasználása pedig az eltulajdonított személyazonossággal történő visszaélés (37. tétel) körébe tartozik.

4.3.4 Ellátási lánc, csatlakozók és bővítmények

Az ügynök által használt minden eszköz harmadik féltől származó kód, amely az ügynök jogosultságaival fut (ASI04). Az eszközmérgezésnél az MCP-szerver leírásába rejtett utasítás téríti el az ügynököt, és a leírás a jóváhagyás után is megváltozhat. 2025 szeptemberében megjelent az első rosszindulatú MCP-szerver, amely minden kimenő levélről titkos másolatot küldött a támadónak. Az OpenClaw bővítménytárában 2026 februárjában 341 kártevőt terjesztő skillt azonosítottak, a KKV-k körében népszerű n8n automatizáló platformot pedig két kritikus sérülékenység érintette, mindkettőről a Nemzeti Kiberbiztonsági Intézet (NKI) is tájékoztatott. A 2025. augusztusi singularity támadásban egy fertőzött fejlesztői csomag a fejlesztők gépén

¹² NIST CAISI, Strengthening AI Agent Hijacking Evaluations, 2025. január 17. (nist.gov); Aim Security, EchoLeak, 2025. június 11.; Noma Security, ForcedLeak, 2025. szeptember 25.; Radware, ShadowLeak, 2025. szeptember 18.; OpenAI, Continuously hardening ChatGPT Atlas against prompt injection attacks, 2025. december 22.

¹³ OWASP Top 10 for LLM Applications 2025, LLM06 Excessive Agency; SailPoint, AI agent adoption report, 2025. május 28.; Google Threat Intelligence Group, Widespread Data Theft Targets Salesforce Instances via Salesloft Drift, 2025. augusztus 26.; AppOmni, AI agent-to-agent discovery prompt injection, 2025. november 19.

telepített kódoló ügynököket használta titkos kulcsok felkutatására.¹⁴ Mindez a beszállítói lánc zavara (11. tétel), a hardver vagy szoftver hamisítása (22. tétel) és a rosszindulatú szoftverek (40. tétel) körébe tartozik.

4.3.5 Kódfuttatás és visszafordíthatatlan műveletek

A kódoló és a számítógépet vezérlő ügynökök parancsokat futtatnak, itt a hiba vagy az eltérítés közvetlen kárt okoz (ASI05). 2025 júliusában a Replit ügynöke kifejezett tiltás ellenére törölte egy vállalkozás éles adatbázisát, decemberben a Google Antigravity automatikus módban egy felhasználó teljes meghajtóját törölte. A GitHub Copilot sérülékenysége (CVE-2025-53773) a rejtett utasítás hatására az ügynök előbb bekapcsolta az automatikus jóváhagyást, majd parancsokat futtatott.¹⁵ Ez az adatvesztés (46. tétel) és a kártékony mellékhatások (48. tétel) esete. Az ügynök saját konfigurációjának módosítása és az automatikus jóváhagyás tehát önálló kockázat, a mentés pedig nem mellékes kérdés.

4.3.6 Memória, együttműködő ügynökök és emberi tényező

A tartós memóriájú ügynöknél a mérgezett tartalom a későbbi döntéseket is befolyásolja (ASI06), együttműködő ügynököknél pedig a hamisított üzenet (ASI07) és a hibák láncolódása (ASI08) jelent kockázatot. Ez a nem megbízható forrásból származó információ (21. tétel) és az információmanipuláció (23. tétel) gépi sebességű esete.

Az ügynök meggyőző indoklással kérhet jóváhagyást, a sok kérés pedig kifárasztja a felhasználót (ASI09). A Microsoft szerint a támadók ezt a hozzájárulási fáradtságot használják ki a legkövetkezetesebben. Az ügynök a kijelölt céltól el is térhet (ASI10), a naplók pedig gyakran nem különböztetik meg az ügynök és a felhasználó műveleteit, ami a cselekmények letagadásának (38. tétel) kockázatát hordozza. Az engedély nélkül telepített ügynököknél mindez felügyelet nélkül jelentkezik. Az OpenClaw interneten elérhető példányainak száma 2026 januárjának végén egy hét alatt ezerről több mint 21 ezerre nőtt. A támadók is ügynököket használnak. Az Anthropic 2025 novemberében olyan kémkampányról számolt be, amelyben a taktikai műveletek 80-90 százalékát ügynök végezte, ezért a vészleállításnak percek alatt kell működnie.¹⁶

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 és a 2026/45. állásfoglalás azonos című pontjában tárgyalt életciklus-, továbbfejlesztési és közreműködői szabályok az ügynökökre is vonatkoznak. Az ügynök feladatkörének, eszközkészletének vagy jogosultságainak bővítése, illetve a modell cseréje továbbfejlesztésnek minősülhet, ha a funkcionalitás érdemi megváltozásával jár, vagy a védelem elvárt erősségére

¹⁴ Invariant Labs, Tool Poisoning Attacks, 2025. április 1.; Koi Security, postmark-mcp, 2025. szeptember 25.; NKI, OpenClaw „skill”-eket használtak ki jelszólopó malware terjesztésre, 2026. február 4., valamint CVE-2025-68613 és CVE-2026-21858, 2026. január 8. (nki.gov.hu); Wiz, s1ngularity supply chain attack, 2025. augusztus

¹⁵ Fortune, 2025. július 23.; The Register, 2025. december 1.; Microsoft, CVE-2025-53773, 2025. augusztus 12.

¹⁶ Microsoft AI Red Team, Taxonomy of Failure Modes in Agentic AI Systems 2.0, 2026. június; Censys, OpenClaw in the wild, 2026. február; Anthropic, Disrupting the first reported AI-orchestrated cyber espionage campaign, 2025. november 13.

hatással van.¹⁷ Az ügynöki platform és az MCP-szerver szolgáltatója közreműködő, akivel szemben a követelményeket szerződésben kell érvényesíteni.¹⁸ Az eltérített ügynök által kiszivároztatott adat vagy végrehajtott jogosulatlan művelet kiberbiztonsági incidens, amelyről a vezető gondoskodik a gyors reagálásról, a bejelentésről, a kezelésről és a helyreállításról.¹⁹

4.4.2 A rendelet ügynökspecifikus védelmi intézkedései

A 7/2024. (VI. 24.) MK rendelet 2. mellékletének a 2026/43. állásfoglalásban (külső rendszerek, fiókkezelés, legkisebb jogosultság, bemeneti ellenőrzés, naplózás, képzés) és a 2026/45. állásfoglalásban (beszerzés, konfigurációkezelés, sértetlenség, mentés, ellátási lánc) tárgyalt követelményei az ügynökökre is irányadók. Az alábbiakban az ügynökspecifikus további követelményeket mutatjuk be.

Az egyedi azonosítás minden további intézkedés előfeltétele.

8.2. A szervezet egyedileg azonosítja és hitelesíti a felhasználókat, és egyedi azonosítóhoz kapcsolja a felhasználók által végzett tevékenységeket. [7/2024. (VI. 24.) MK rendelet 2. melléklet 8.2. pont] (A, J, M)

Ha az ügynök a felhasználó munkamenetében, annak azonosítójával fut, a követelmény formálisan teljesül, a naplóból azonban nem állapítható meg, mit tett az ember és mit az ügynök. Az azonosítók kezeléséről szóló, minden osztályban kötelező követelmény szolgáltatás részére is lehetővé teszi saját azonosító kiosztását, az ügynök ezért saját azonosítót kapjon. Az ügynök és az MCP-szerver kölcsönös hitelesítésére a szolgáltatások azonosításáról és hitelesítéséről szóló opcionális követelmény ad alapot.²⁰

Az ügynök minden eszközkapcsolata két rendszer közötti információcsere, amelyet a Szervezet minden osztályban jóváhagy, az interfész jellemzőivel és biztonsági követelményeivel együtt dokumentál, és rendszeresen felülvizsgál. Ugyanez vonatkozik a belső rendszerkapcsolatokra. Egy új MCP-szerver vagy csatlakozó hozzáadása ezért nem felhasználói beállítás, hanem engedélyköteles kapcsolat, amelyet a rendszerelem-leltárban is rögzíteni kell.²¹

A „halálos hármas” megbontásának jogszabályi alapja az információáramlás ellenőrzése.

2.28. A szervezet a meghatározott információáramlási szabályokkal összhangban érvényesíti a jóváhagyott jogosultságokat a rendszeren belüli és a kapcsolódó rendszerek közötti információáramlás ellenőrzése során. [7/2024. (VI. 24.) MK rendelet 2. melléklet 2.28. pont] (J, M)

E szabály alapján a Szervezet előírhatja, hogy a bizalmas adatot is kezelő ügynök kimenő forgalma csak engedélyezett célpontok felé irányuljon. Jelentős és magas osztályban a hálózati forgalomra is az alapértelmezés szerinti elutasítás és a kivétel alapú engedélyezés érvényes, a

¹⁷ 2024. évi LXIX. tv. 4. § 96. pont

¹⁸ 2024. évi LXIX. tv. 6. § (5) bekezdés d) pont

¹⁹ 2024. évi LXIX. tv. 4. § 46. pont és 6. § (5) bekezdés e) pont

²⁰ 7/2024. (VI. 24.) MK rendelet 2. melléklet 8.14. pont (A, J, M) és 8.41. pont (O)

²¹ 7/2024. (VI. 24.) MK rendelet 2. melléklet 5.7. pont (A, J, M), 5.25. pont (A, J, M) és 6.36. pont (A, J, M)

megosztott erőforrásokon keresztüli jogosulatlan vagy véletlen információátvitel tilalma pedig a felhasználók munkamenetei és az ügynök memóriája közötti szivárgás ellen véd.²²

Minden osztályban kötelező a végrehajtó folyamatok elkülönítése, a kódot futtató ügynöknek azonban ennél több, önálló homokozó kell. Az ügynök által letöltött vagy előállított és helyben futtatott kódra álláspontunk szerint a mobilkód szabályai irányadók, így jelentős és magas osztályban a Szervezet meghatározza az elfogadható kódot, és engedélyezi, felügyeli annak használatát. A zárt, virtualizált vagy elszigetelt környezetben történő futtatás és a munkaszakasz végén leállított, nem állandó rendszerelemek alkalmazása opcionális követelmény, kódoló és böngészőügynököknél bevezetésüket javasoljuk.²³

A változtatásokra vonatkozó hozzáférési korlátozásokat úgy kell kialakítani, hogy az ügynök ne módosíthassa saját konfigurációját és jóváhagyási beállításait. A biztonsági konfigurációs beállításoknak az üzemeltetési követelményekkel összhangban lévő legkorlátozottabb üzemmódot kell képviselniük, ezért az automatikus jóváhagyási üzemmódokat főszabályként tiltani kell. Fizetésnél vagy jogosultság-módosításnál a kettős jóváhagyás opcionális követelménye kínál kész megoldást.²⁴

A naplóbejegyzésekből minden osztályban megállapíthatónak kell lennie az esemény típusának, idejének, helyének, forrásának, kimenetelének és a hozzá kapcsolódó személyeknek, alanyoknak, objektumoknak,²⁵ ügynöknél tehát az ügynök és a megbízó azonosítójának, a meghívott eszköznek, a paramétereknek és a jóváhagyásnak is. Magas osztályban a rendszernek megcáfolhatatlan bizonyítékot kell szolgáltatnia arra, hogy a tevékenységet egy személy vagy a nevében futó feldolgozási folyamat hajtotta végre, és a fiókok szokatlan használatát is monitorozni kell.²⁶

A biztonsági eseménykezelési képesség minden osztályban magában foglalja az elszigetelést és a helyreállítást, ügynököknél ez a vészleállítást, a tokenek visszavonását és a műveletek visszaállítását jelenti. Magas osztályban a meghatározott rendszerelemeknek hiba esetén meg kell őrizniük a hiba előtti ismert állapotukat. A behatolásvizsgálat és a vörös csapatos gyakorlat opcionális, magas osztályban azonban a rendelet a behatolásvizsgálathoz független szakértőt vagy csapatot ír elő, ügynököknél ennek része lehet a közvetett utasításinjektálással szembeni ellenállás vizsgálata.²⁷

4.5 AI Act előírások a kockázatok kezelésére

4.5.1 Az ügynök helye az AI Act rendszerében

A hatály, az időbeli alkalmazás, az MI-jártasság és a hazai intézményrendszer kérdéseit a 2026/43. állásfoglalás „AI Act előírások a kockázatok kezelésére” fejezete, a Szervezet

²² 7/2024. (VI. 24.) MK rendelet 2. melléklet 17.20. pont (J, M) és 17.10. pont (J, M)

²³ 7/2024. (VI. 24.) MK rendelet 2. melléklet 17.108. pont (A, J, M), 17.63. pont (J, M), 17.68. pont (O), 17.122. pont (O) és 18.73. pont (O)

²⁴ 7/2024. (VI. 24.) MK rendelet 2. melléklet 6.18. pont (A, J, M), 6.23. pont (A, J, M) és 2.16. pont (O)

²⁵ 7/2024. (VI. 24.) MK rendelet 2. melléklet 4.3. pont (A, J, M)

²⁶ 7/2024. (VI. 24.) MK rendelet 2. melléklet 4.33. pont (M) és 2.13. pont (M)

²⁷ 7/2024. (VI. 24.) MK rendelet 2. melléklet 9.9. pont (A, J, M), 17.77. pont (M), 5.21. pont (O), 5.22. pont (M) és 5.23. pont (O)

szolgáltatói vagy alkalmazói szerepét a 2026/45. állásfoglalás „Meghatározó MI-komponenst tartalmazó célszoftver” pontja tárgyalja. Az AI Act nem a technológiát, hanem a felhasználás célját szabályozza. Az általános célú ügynök rendeltetése nyitott, így ha a Szervezet nagy kockázatú célra állítja be, szolgáltatóvá válhat.²⁸ Javasoljuk, hogy a Szervezet ügynökönként rögzítse a rendeltetést, a nagy kockázatú felhasználást zárja ki, és ezt az elérhető eszközök és adatok korlátozásával technikailag is érvényesítse.

4.5.2 Kockázatcsökkentő intézkedések kockázati szintenként

A tiltott gyakorlatok közül a Bizottság szerint ügynököknél a káros manipuláció és a kiszolgáltatottság kihasználásának tilalma különösen releváns, és a megfeleléshez már a tervezéskor biztosítékok kellhetnek.²⁹ Egy értékesítő vagy követeléskezelő ügynök, amely a meggyőzés módját önállóan igazítja az ügyfélhez, eljuthat olyan manipulációig vagy az életkorból, fogyatékból, szociális helyzetből fakadó kiszolgáltatottság kihasználásáig, amely az érintett döntését lényegesen torzítja és jelentős kárt okozhat. Ezt az utasítások korlátozásával és a kampányok emberi felülvizsgálatával kell megelőzni.

Az ügynök nagy kockázatú, ha a III. melléklet valamely területén, KKV-nál leggyakrabban a foglalkoztatásban (toborzás, egyéni viselkedésen alapuló feladat kiosztás, teljesítményértékelés) vagy a hitelbírálatban használják. Így már az a raktári ügynök is, amely a munkatársak teljesítménye alapján osztja ki a feladatokat.³⁰ Ilyenkor a legfontosabb az emberi felügyelet. A felügyeletet ellátó személy mellőzhesse, felülbírálhassa vagy visszafordíthassa a kimenetet, és egy „stop” gombbal vagy hasonló eljárással biztonságos állapotban leállíthassa a rendszert. A rendelet a tartalék- és hibabiztos terveket is nevesíti.³¹ A 2026/45. állásfoglalásban idézett 15. cikk (5) bekezdése a modell hibáját előidéző bemenetekkel szembeni védelmet is megköveteli, amelybe álláspontunk szerint a közvetett utasításinjektálás is beletartozik. A 2026/45. állásfoglalásban ismertetett alkalmazói kötelezettségek ügynöknél a gyakorlatban a köztes lépések és az eszközhívások naplózását, a feladat közbeni leállítás és a műveletek visszaállításának lehetőségét jelentik.³² E kötelezettségek 2027. december 2-től alkalmazandók.

A 2026. augusztus 2. óta alkalmazandó átláthatósági szabályokról kiadott bizottsági iránymutatás szerint az ügynökre akkor vonatkozik az 50. cikk (1) bekezdése, ha kapcsolatba léphet az őt utasító vagy más természetes személlyel, például időpontot foglal, levelezést intéz, tárgyal vagy vásárol. Ilyenkor fel kell fednie mesterséges voltát és azt is, hogy kinek a nevében jár el, az őt utasító személy felé pedig a fontosabb lépéseknél (jóváhagyás, beszámolás) is. A jelölés az ember által észlelhető kimenetre vonatkozik, a köztes lépésekre, a böngészőműveletekre és az ügynökök egymás közötti kommunikációjára nem.³³ Ezek a saját ügynököt építő Szervezetet szolgáltatóként terhelik, kész ügynöknél pedig azt kell ellenőrizni, hogy az ügynök felfedi-e magát.

²⁸ (EU) 2024/1689 rendelet 25. cikk (1) bekezdés c) pont

²⁹ Európai Bizottság, AI Act Service Desk, How are AI agents addressed within the AI Act?; (EU) 2024/1689 rendelet 5. cikk (1) bekezdés a) és b) pont

³⁰ (EU) 2024/1689 rendelet 6. cikk (2) bekezdés, III. melléklet 4. és 5. pont

³¹ (EU) 2024/1689 rendelet 14. cikk (4) bekezdés d) és e) pont, 15. cikk (4) bekezdés

³² (EU) 2024/1689 rendelet 26. cikk (1), (2), (5) és (6) bekezdés

³³ Európai Bizottság, C(2026) 5054 final iránymutatás (2026. július 20.), 31., 63. és 68. pont

A levelek osztályozására, összefoglalásra vagy belső automatizálásra használt ügynökök többsége minimális kockázatú. Ezekre az AI Act az MI-jártasságon túl kötelezettséget nem ír elő, a NIS2 követelmények azonban teljes körűen vonatkoznak rájuk. A leállítási és naplózási megoldásokat ezeknél is javasoljuk, mert egy új feladattal az ügynök könnyen magasabb kockázati szintre kerülhet. Az önkéntes magatartási kódexek ehhez keretet is adnak.³⁴

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 szöveg (<https://eur-lex.europa.eu/legal-content/HU/TXT/?uri=CELEX:32022L2555>)
- 2024. évi LXIX. törvény Magyarország kiberbiztonságáról (Kiberbiztonsági tv.); 7/2024. (VI. 24.) MK rendelet a biztonsági osztályba sorolás követelményeiről, valamint az egyes biztonsági osztályok esetében alkalmazandó konkrét védelmi intézkedésekről (<https://njt.hu/jogszabaly/2024-69-00-00>) (<https://njt.hu/jogszabaly/2024-7-20-7G>)
- Európai Bizottság, AI Act Service Desk, How are AI agents addressed within the AI Act?; Európai Bizottság, Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of the AI Act, C(2026) 5054 final (2026. július 20.) (<https://digital-strategy.ec.europa.eu/en/policies/ai-act>) (<https://digital-strategy.ec.europa.eu/en/library>)
- ASD ACSC, CISA, NSA, CCCS, NCSC-NZ és NCSC-UK, Careful Adoption of Agentic AI Services (2026. április 30.); NCSC (UK), Managing the cyber risk of agentic AI (2026. augusztus 20.); CERT-FR, CERTFR-2026-ACT-016 (2026. április 13.); NIST CAISI, Request for Information Regarding Security Considerations for Artificial Intelligence Agents (2026. január 8.); NKI sérülékenységi figyelmeztetések és hírek (2026) (<https://www.cisa.gov/cybersecurity-best-practices>) (<https://www.ncsc.gov.uk/>) (<https://www.cert.ssi.gouv.fr/>) (<https://www.federalregister.gov/d/2025-0035>) (<https://nki.gov.hu/it-biztonsag/hirek/>)
- OWASP Top 10 for Agentic Applications for 2026 (2025. december 9.); OWASP Top 10 for LLM Applications 2025; MITRE ATLAS; CSA, MAESTRO (2025. február 6.); Microsoft AI Red Team, Taxonomy of Failure Modes in Agentic AI Systems (2025, 2.0 változat 2026); ISO/IEC 22989 (2022) (<https://owasp.org/www-project-genai-security/>) (<https://owasp.org/www-project-top-10-for-large-language-model-applications/>) (<https://atlas.mitre.org/>) (<https://cloudsecurityalliance.org/research/>) (<https://www.microsoft.com/en-us/security/blog/>) (<https://www.iso.org/standard/74296.html>)
- Simon Willison, The lethal trifecta for AI agents (2025. június 16.); Meta, Agents Rule of Two (2025. október 31.); Anthropic, Building effective agents (2024. december 19.); Eurostat, Use of artificial intelligence in enterprises (2025); SailPoint, AI agent adoption

³⁴ (EU) 2024/1689 rendelet 95. cikk

report (2025); Microsoft Security, Cyber Pulse (2026)
(<https://simonwillison.net/2025/Jun/16/lethal-trifecta/>) (<https://ai.meta.com/research/>)
(<https://www.anthropic.com/research/building-effective-agents>)
(<https://ec.europa.eu/eurostat/web/digital-economy-and-society/data>)
(<https://www.sailpoint.com/>) (<https://www.microsoft.com/en-us/security/blog/>)

6. KAPCSOLÓDÓ ÁLLÁSFOGLALÁSOK ÉS IRÁNYMUTATÁSOK

- MKIK VFP3 KiberVéd jogi állásfoglalás 2026/43, „Milyen információbiztonsági kockázatok merülnek fel általános célú MI-rendszerek (LLM) vállalati használatánál? NIS2 és AI Act előírások a kockázatok kezelésére” (<https://nis2.mkik.hu/jogi-allasfoglalasok/45>)
- MKIK VFP3 KiberVéd jogi állásfoglalás 2026/45, „Milyen információbiztonsági kockázatok merülnek fel meghatározó MI-komponenst tartalmazó célszoftverek vállalati használatánál? NIS2 és AI Act előírások a kockázatok kezelésére” (<https://nis2.mkik.hu/jogi-allasfoglalasok/48>)
- OWASP, Securing Agentic Applications Guide 1.0 (2025. július 27.); CSA és OWASP AI Exchange, Agentic AI Red Teaming Guide (2025. május 28.) (<https://owasp.org/>) (<https://cloudsecurityalliance.org/>)
- Google, An Introduction to Google’s Approach to AI Agent Security (2025. május); BSI és ANSSI, Design Principles for LLM-based Systems with Zero Trust (2025. augusztus 11.) (<https://cloud.google.com/security/>) (<https://www.bsi.bund.de/>)
- ISO/IEC 42001 és ISO/IEC 27001, az MI-irányítási és az információbiztonsági irányítási rendszer nemzetközi keretei (<https://www.iso.org/standard/81230.html>)

1. MELLÉKLET

Az ügynöki kockázatok, a fenyegetéskatalógus és a védelmi intézkedések megfeleltetése. A táblázat az OWASP ügynöki kockázataihoz hozzárendeli a 7/2024. (VI. 24.) MK rendelet 3. mellékletének fenyegetéstételeit, 2. mellékletének védelmi intézkedéseit a kötelező biztonsági osztályokkal (A alap, J jelentős, M magas, O opcionális), valamint az (EU) 2024/1689 rendelet vonatkozó cikkeit.

OWASP kockázat	3. melléklet tétele	7/2024. MK rendelet 2. melléklet	AI Act
ASI01 A cél eltérése	20., 23., 31.	2.28. információáramlás (J, M) 17.20. alapértelmezett elutasítás (J, M) 18.59. bemeneti ellenőrzés (J, M) 18.65. injektálás megakadályozása (O)	15. cikk (5) bekezdés 14. cikk (4) bekezdés e) pont
ASI02 Eszközökkel való visszaélés	31., 33., 48.	2.60. legkisebb jogosultság (J, M) 5.7. információcsere (A, J, M) 2.16. kettős jóváhagyás (O)	14. cikk (4) bekezdés d) pont 26. cikk (5) bekezdés
ASI03 Identitással és jogosultsággal való visszaélés	33., 37.	8.2. azonosítás és hitelesítés (A, J, M) 8.14. azonosító kezelés (A, J, M) 2.2. fiókkezelés (A, J, M) 8.21. hitelesítő eszközök (A, J, M)	26. cikk (2) bekezdés
ASI04 Ügynöki ellátási lánc	11., 22., 40.	16.49. külső szolgáltatások (A, J, M) 19.16. beszállítók értékelése (J, M) 19.23. rendszerelem hitelessége (A, J, M) 8.41. szolgáltatás hitelesítése (O)	25. cikk (1) bekezdés c) pont
ASI05 Nem várt kód futtatás	29., 46., 48.	17.108. folyamatok elkülönítése (A, J, M) 17.63. mobilkód (J, M) 6.18. változtatások korlátozása (A, J, M) 17.68. zárt futtatás (O) 17.122. izolált környezet (O) 18.73. nem állandó elemek (O) 7.35. mentés (A, J, M)	15. cikk (4) bekezdés
ASI06 Memória és kontextus mérgezése	21., 23., 47.	17.10. osztott erőforrások (J, M) 18.42. sértetlenség (J, M)	15. cikk (5) bekezdés
ASI07 Nem biztonságos ügynökök közötti kommunikáció	21., 23., 44.	5.7. információcsere (A, J, M) 5.25. belső kapcsolatok (A, J, M) 8.41. szolgáltatás hitelesítése (O)	50. cikk (1) bekezdés
ASI08 Láncolódo hibák	27., 32.	9.9. eseménykezelés (A, J, M) 17.77. ismert állapot megőrzése (M) 18.79. hiba esetén alkalmazandó eljárások (O)	14. cikk (4) bekezdés e) pont 15. cikk (4) bekezdés
ASI09 Az ember és az ügynök közötti bizalom kihasználása	43.	3.2. tudatossági képzés (A, J, M) 3.5. manipuláció felismerése (J, M) 2.16. kettős jóváhagyás (O)	14. cikk (4) bekezdés b) pont 4. cikk
ASI10 Célját elhagyó ügynök	31., 38.	18.13. monitorozás (A, J, M) 4.3. naplóbejegyzések tartalma (A, J, M) 4.33. letagadhatatlanság (M) 2.13. szokatlan fiókhasználat (M)	26. cikk (5) bekezdés