

A SZÜKSÉG MEGOLDÁST SZÜL – AKÁR TÖBBET IS a „NAMIEZ”- projekt és az anonimizálás

2026.10.19.

Tóth Péter, elnöki főtanácsadó



Nemzeti Adó-
és Vámhivatal

Mit várunk az MI-megoldásoktól?

Három korlát gyorsan leépül

- „Túl sok” adat
 - Nagy tömegű tranzakció, számla, kapcsolat és ügylet mintázatainak elemzése.
- Strukturálatlan információ
 - Dokumentum, jogeset, fénykép, röntgenkép, hang vagy kézírás feldolgozható adattá válik.
- Technikai belépési küszöb
 - Természetes nyelven kérhetünk elemzést, kódot, összefoglalót vagy adatkinyerést.

Kísérletezésből tömeges alkalmazás

- 29 / 38 OECD-tag adóhatóság AI-használatot jelzett
- >90% FTA-tagok: AI-használat vagy bevezetés alatt álló megoldás
- 3 fő irány:
 - Compliance
 - Döntéstámogatás
 - Szolgáltatások



Nem az a kérdés, hogy „szabad-e AI-t használni?”

Mire? Milyen adaton? Milyen döntési hatással? Milyen kontrollokkal?

A technikai lehetőség ≠ jogszerű használat

- Célhoz kötött adathasználat
- Adat- és titokvédelem
- AI Act és ágazati szabályok
- Auditálhatóság és jogorvoslat

Egy jó modell ≠ üzemszerűen működő AI-rendszer

- Demo-t készíteni könnyű.
- Biztonságos, auditálható, integrált és fenntartható szolgáltatást építeni nehéz.

Az AI felnagyítja a szervezeti hiányosságokat, nem megoldja

- Adatminőség
- Adatszótár- és metaadat-hiány
- Inkonzisztens tudásbázis
- Jogosultsági hézag
- Kompetenciahiány

AI-képességet kell építeni, nem csak AI-eszközt venni.

- Adat és módszertan
- Hardver- és szoftverinfrastruktúra
- Fejlesztési, üzemeltetési és felhasználói tudás
- Jog és governance

Nem zsenialitás kell, hanem fáradhatatlan precizitás

Az AI dolga: elvégezni a feladatot – nem lenyűgözni



Nagyon okos frontier modellek

Már többet tudnak, mint amennyi a napi munkához kell.



Ingyenes, nyílt súlyú modellek

Fél-egy éves lemaradással, legálisan, ingyen futtathatók.

Amihez nem kell frontier modell

- szöveg összefoglalása
- anonimizálás
- két dokumentum összevetése
- jegyzetek és kérdéslisták
- kulcsszavak és nevek kinyerése
- átírás más stílusban
- gyors fordítás

A tudásmunka nagyjából fele ilyen feladatokból áll.

Miért érdemes kisebbet választani?



Adatvédelem

A bizalmas anyag nem hagyja el a házat.



Költség

Sok felhasználóval gyorsan gyűlik a tokenszámla.



Illeszkedés

A feladat egyszerűen nem igényel többet.

Nem egy AI-eszközt építünk, hanem kontrollált eszközrendszert

A feladat, az adat érzékenysége és a szükséges kontroll szintje dönti el, melyik környezet a megfelelő.

EGY FELADAT → MEGFELEŐ MODELL → MEGFELEŐ KÖRNYEZET → MEGFELEŐ KONTROLL

1. ÁG

Prediktív AI és gépi tanulás

- Kockázati minták és anomáliák
- Kép- és számlaadat-elemzés
- Validált modellek üzemi folyamatokban

Céltzott, mérhető, reprodukálható modellek

2. ÁG

Generatív AI és nyelvi modellek

Három egymást kiegészítő használati réteg

01

NAMIEZ

Zárt, on-prem

Érzékeny NAV-adatok, teljes kontroll, célfeladatok.



02

Copilot

Vállalati környezet

Általános irodai és tudásmunka-támogatás, adatkontroll mellett.



03

Külső AI-eszközök

Szabályozott hozzáférés

NAV-adat nélkül, szoros jogosultság- és adatkezeléssel.



Nem minden feladatot ugyanazzal a modellel kell megoldani – és nem minden adat kerülhet ugyanabba a környezetbe.

NAMIEZ – biztonságos AI-eszköztár a NAV belső feladataihoz

Olyan rutinfeladatokat támogat, amelyeknél az adat nem hagyhatja el a NAV infrastruktúráját.

MIÉRT?

1

Időt nyer

Ismétlődő dokumentum- és adatfeldolgozási feladatokat gyorsít.

2

Adatot véd

Érzékeny dokumentumok nem kerülnek külső AI-szolgáltatásba.

3

Képességet épít

A modellek, működés és fejlesztési tudás NAV-kontroll alatt marad.

NAMIEZ

NAV MI Eszköztár

„digitális svájcbicska”



**Első „quick win” funkciók: anonimizálás ·
adatkinyerés · dokumentum-összevetés ·
határozat-összefoglalás**

MI LESZ?



- ✓ Zárt NAV-infrastruktúrán fut
- ✓ Előre kialakított célfunkciók
- ✓ Egyszerű feltöltés → feldolgozás → letöltés
- ✓ Fokozatosan bővíthető technológiai alap

MI NEM LESZ?



- Nem „NAV ChatGPT”
- Nem autonóm döntéshozó
- Nem a Copilot kiváltása
- Nincs szabad promptolás

Kontrollált pilotból skálázható belső AI-képesség

A szakmai és technikai előkészítés párhuzamosan halad; a projekt a meglévő CPU-környezetben indul, de GPU-s környezetben skálázható.

Projektállapot

-  **Use case-ek kijelölve** **KÉSZ**
A négy első fázisú „quick win” funkció megerősítve.
-  **Szakértői szerződés** **KÉSZ**
A szerződés szakmai tartalma előkészítve, augusztus végi indulás.
-  **Belső CPU-pilot** **OKTÓBERTŐL**
Modellértékelés és funkciófejlesztés kontrollált környezetben.
-  **GPU-kapacitás** **NOVEMBERTŐL**
Beszerzés folyamatban; későbbi skálázás és nehezebb use case-ek.

 **NAMIEZ – NAV MI Eszköztár**
Zárt, belső AI-eszköztár

 Strukturált adatkinyerés

 Dokumentum-összevetés

 Személyes adatok anonimizálása

 Beérkező levél összefoglalása

 Szkennelt dokumentum feldolgozása

 Hanganyag leiratozása

 Programkód dokumentálása

 SQL-lekérdezés magyarázata

 Határozat összefoglalója

MOCK-UP

Belső használatra

NAV munkatárs

 **Határozat összefoglalója – rövid**
Töltsön fel egy legfeljebb kb. 20 oldalas határozatot, a rendszer tömör, strukturált összefoglalót készít a fő megállapításokról, jogalapról és rendelkező részről.

1

Bemenet
Határozat feltöltése

 Húzza ide a fájlt, vagy kattintson a tallózáshoz.

 **határozat_2026_04_18.pdf**
3,42 MB



2

Feldolgozás indítása
A rendszer automatikusan tömör, strukturált összefoglalót készít a feltöltött határozatból.

3

Kimenet
Automatikusan generált, strukturált összefoglaló (nem chatbot-válasz).

 **Megállapítások**

- A határozat tárgya hatósági ellenőrzés megállapításai.
- Jogsértés megállapításra került az Áfa tv. 120. § (1) bekezdés d) pontja alapján.
- Az ellenőrzés időszaka: 2024.01.01. – 2024.12.31.
- A megállapított adókülönbözlet: 5 240 000 Ft.

 **Jogalap**

- Áfa tv. 120. § (1) bekezdés d) pont
- Áfa tv. 169. §
- Art. 50. § (1) bekezdés
- Art. 115. § (1) bekezdés

 **Rendelkező rész**

- A hatóság 5 240 000 Ft adókülönbözetet állapít meg.
- Késedelmi pótlék fizetési kötelezettség kerül előírásra.
- A kötelezettség teljesítésének határideje: 2026.05.18.
- A határozat ellen fellebbezésnek helye nincs.

 **határozat_osszefoglalo.pdf**
412 KB • PDF dokumentum



 Előnézet

 Letöltés

 Az összes feldolgozás a NAV belső környezetében történik. Az adatok nem hagyják el a NAV belső környezetét.

Következő mérföldkő: az első funkciók user tesztje és mérési eredményekkel alátámasztott skálázási döntés.

Gyakorlati működés (előzetes!)

Anonimizálás

The screenshot shows a web browser window displaying the 'NAMIEZ - NAV MI Eszköztár' application. The browser's address bar shows 'localhost:4200'. The application has a dark sidebar on the left with a file explorer and a main content area. The main content area has a header with the title 'NAMIEZ - NAV MI Eszköztár' and a sub-header 'Eszköztár'. Below the sub-header, there is a description: 'Válassza ki a feladathoz illő eszközt. A futtatás a háttérben történik, az eredményt a Feladataim oldalon találja.' The main content area is divided into three sections: 'Eszköztár', 'Feladataim', and 'Admin'. The 'Eszköztár' section is active and contains a sub-section 'Anonimizálás és adatvédelem'. This section has a title 'Személyes adatok anonimizálása' and a button 'ANO-01 Dokumentum vagy szöveg'. Below this, there is a paragraph of text: 'Személyes és azonosító adatok (nevek, adószámok, TAJ-számok, lakcímek, telefonszámok, iktatószámok) cseréje egységes szögletes zárójeles jelölőkre, a szöveg többi részének változatlan megőrzésével. Az iktatószám és az ügyszám alapértelmezésben szintén jelölőt kap; a Kiegészítő utasításban kérhető a megtartásuk.' Below this, there is a section 'Adatkinyerés' with two sub-sections: 'Strukturált adatok kinyerése szövegből' and 'Nyomtatvány-adattáblázat kinyerése'. The first sub-section has a button 'ADK-01 Dokumentum vagy szöveg' and the second has a button 'ADK-06 Dokumentum vagy szöveg'. The bottom of the browser window shows the Windows taskbar with various icons and the system clock.

NAMIEZ - NAV MI Eszköztár

Eszköztár

Válassza ki a feladathoz illő eszközt. A futtatás a háttérben történik, az eredményt a Feladataim oldalon találja.

Anonimizálás és adatvédelem

Személyes adatok anonimizálása

ANO-01 Dokumentum vagy szöveg

Személyes és azonosító adatok (nevek, adószámok, TAJ-számok, lakcímek, telefonszámok, iktatószámok) cseréje egységes szögletes zárójeles jelölőkre, a szöveg többi részének változatlan megőrzésével. Az iktatószám és az ügyszám alapértelmezésben szintén jelölőt kap; a Kiegészítő utasításban kérhető a megtartásuk.

Adatkinyerés

Strukturált adatok kinyerése szövegből

ADK-01 Dokumentum vagy szöveg

Nyomtatvány-adattáblázat kinyerése

ADK-06 Dokumentum vagy szöveg

Egy use-case példa – több megoldási irány ... de baj ez?

Feladatspecifikus NER / transformer

Erős direkt azonosítók, domainen belül

Stabilabb ismételhetőbb recall/precision

Gyengébb quasi-identifier és kontextus

Korlát domainváltás, ritka kategóriák

Magyar evidencia

Klinikai szövegen F1 97,7–98,7%; ismételt futtatás:
residual PII 1,18% → 0,33%.

Generatív LLM

Erős lehet zajos, változatos szabad szöveg

Előny kontextus, implicit attribútumok

Kockázat prompt- és modellfüggőség

Költség inferencia és validálás drágább

Vegyes eredmény

GPT-4: P 99,25%, de R 83,18%; más korpuszon frontier
LLM recall 96% körül.

Megoldáskombinációs lehetőségek

Regex /
formális
azonosítók

Finomhangolt
NER /
transformer

LLM privacy
ellenőrzés

LLM felfedési
kísérlet

Humán
felülvizsgálat /
kockázati
értékelés

NEM! Sokszor nincs egyetlen, mindenre jó válasz!

Párhuzamos belső fejlesztések

Mellette szól

- Más igények jönnek a felszínre
- Eltérő felhasználási helyzetekre optimalizálják
- Többféle folyamatintegrációs lehetőség
- Képességfejlődés
- Gyorsabb lehet a valódi jó megoldás(ok) megtalálása

Ellene / kockázatok

- Mérés nélkül ismeretlen működési minőség, felhasználhatóság kérdéses
- Nagyobb erőforrás-felhasználás kapacitáskorlát mellett
- Párhuzamosság, pazarlásveszély

Csak

Konzisztens belső mérés
Folyamatos belső tudásmegosztás

mellett!

Köszönöm a figyelmet!



Nemzeti Adó-
és Vámhivatal