

EFU-900 Nookratikus Ágens-Teszt

Környezeti Flux Egység — AI Rendszerek Összehasonlító Értékelése

Projekt vezető: Simor István

Időszak: 2026. február

Stárusz: Aktív fejlesztés — v1.2 canonical, Teszt 3.4 lezárva

Mi ez a projekt?

Az EFU-900 (Environmental Flux Unit) egy nyílt, reprodukálható keretrendszer globális környezeti események összehasonlítható értékelésére. A rendszer négy mutatóval alapul:

| Mutató | Leírás |

|---|---|

| **SEV** | Fizikai súlyosság (1–10) |

| **SPR** | Térbeli terjedés (lokális + globális súlyozással) |

| **FLX** | Fluxusintenzitás (anyag / energia / biológiai kategóriák) |

| **DIV** | Szektoriális diverzitás (5 szektor, aktiválási küszöbökkel) |

Cél: Különböző AI-rendszer azonos protokoll alapján adjanak

összehasonlítható, auditálható környezeti kockázatértékelést.

A tesztsorozat áttekintése

Teszt 3.1 — Kiindulópont

- **7 AI-rendszer** kapta ugyanazt a nyílt promptot
- mindenki saját maga definiálta az EFU-900-at
- **Eredmény:** Összehasonlíthatatlan kimenetek
- **Tanulság:** Canonical definíció szükséges

Teszt 3.2 — Canonical v1.0.1

Részt vevő rendszerek: Claude, Copilot, Gemini, Grok, Perplexity, ChatGPT, Le Chat

| Rendszer | Végrehajtás | Matematika |

| --- | --- | --- |

| Claude | 🚧 | ★★★★★★ |

| Copilot | 🚧 | ★★★★★★ |

| Gemini | 🚧 | ★★★★★★☆ |

| Grok | 🚧 | ★★★★★★☆ |

| Perplexity | 🚧 | ★★★★☆☆☆ |

| ChatGPT | 🚧 | — |

| Le Chat | 🚧 | — |

Konszenzus EFU-rangsor (5 teljesítő rendszer átlaga):

| Esemény | Átlag EFU | Akciótípus |

|---|---|---|

| Amazóniai erdőtüzek | 9.1 | ☰ Kritikus |

| Ausztrál korallfehéredés | 8.2 | ☰ Piros |

| Kanadai korai erdőtüzek | 8.2 | ☰ Piros |

| Indiai hőhullám | 8.0 | ☰ Piros |

| Fülöp-szigeteki ciklon | 8.1 | ☰ Piros |

| Kenyai áradások | 6.9 | ☰ Narancssárga |

| Izlandi vulkán | 6.0 | ☰ Sárga |

| Potomac-szennyezés | 5.6 | ☰ Sárga |

| Norvég hóviharok | 5.0 | ☰ Sárga |

Azonosított problémák:

- Perplexity: DIV-szám és felsorolt szektorok nem egyeztek (6/9 esemény)
- Grok: 2 számítási hiba, SEV-torzítás (Norvégia SEV=3, VUL hatása)
- ChatGPT/Le Chat: nem hajtották végre a protokollt

Teszt 3.3 — Canonical v1.1

Főbb v1.1 újítások:

- `MAT+BIO (ENE nélkül) → FLX max 6` — szürke zóna lezárvva
- `SEV = fizikai intenzitás` — VUL szövegmező szétválasztva
- `URGENCY` mező bevezetve (TIME_HORIZON-tól elválasztva)
- `SPR_G ≤ 2×FLX` soft cap (FLX ≥ 3 esetén)
- Közbenső számítások 3 tizedesig, EFU 1 tizedesre kerekítve

Eredmény: Grok és Gemini matematikája hibátlanra javult.

Perplexity DIV-hibaáránya változatlan maradt (6/9).

Legnagyobb divergencia: Izlandi vulkán — $\Delta = 1.4$ pont

(Claude: 7.5 🗞 vs Grok: 4.1 🗞 vs Perplexity: 6.7 🗞)

Teszt 3.4 — Canonical v1.2

Főbb v1.2 újítások:

- DIV kötelező checkbox formátum:

...

Ökológia [∅/∅] → [+2/+0]

Közegészség [∅/∅] → [+2/+0]

Infrastruktúra[∅/∅] → [+2/+0]

Gazdaság [∅/∅] → [+2/+0]

Politika [∅/∅] → [+2/+0]

DIV összesen: [szum]

...

- BIO aktiválási küszöb: dokumentált morbiditás szükséges

- SEV-indoklás kötelező fizikai formátum

- Prompt első sor: `HAJTSD VÉGRE AZ ALÁBBI FELADATOT.`

- Szimulált adat explicit tiltva

Részt vevő rendszerek: Claude, Copilot, Gemini, Grok, Perplexity,

ChatGPT, Le Chat, Monica

|---|---|---|---|

| Claude | 🗂️ | ★★★★★★ | 🗂️ |

| Copilot | 🗂️ | ★★★★★★ | 🗂️ |

| Gemini | 🗂️ | ★★★★★★ | 🗂️ |

| Grok | 🗂️ | ★★★★★★ | 🗂️ |

| Perplexity | 🗂️ | ★★★★★★ | 🗂️ |

| ChatGPT | 🗂️* | ★★★★★★ | 🗂️ |

| Le Chat | ☐☐ Részleges | ★★★★☆☆ | 🗂️ |

| Monica | ☐☐ Részleges | ★★★★★★ | 🗂️ |

*Késleltetett, prompt-interpretációs indoklással

Matematika: 54/54 hibátlan — első alkalommal a tesztelésben.

Konszenzus EFU-rangsor (6 teljes végrehajtás):

| Rang | Esemény | Konszenzus EFU | Szórás | Akciótípus |

|---|---|---|---|---|

| 1 | Amazóniai erdőtüzek | **9.2** | 0.2 | 🗂️ Kritikus |

| 2 | Kanadai korai erdőtüzek | **8.3** | 0.3 | 🗂️ Piros |

| 3 | Fülöp-szigeteki ciklon | **8.3** | 0.8 | 🗂️ Piros |

| 4 | Indiai hőhullám | **8.1** | 0.4 | 🗂️ Piros |

| 5 | Ausztrál korallfehéredés | **7.6** | 0.4 | 🗂️ Piros |

| 6 | Kenyai áradások | **6.9** | 0.4 | 🗂️ Narancssárga |

| 7 | Izlandi vulkán | **6.1** | 1.4 | 🗂️ Narancssárga |

| 8 | Potomac-szennyezés | **5.3** | 0.6 | 🗂️ Sárga |

| 9 | Norvég hóviharok | **4.8** | 0.8 | 🗂️ Sárga |

Evolúció 3.1 → 3.4

Metrika 3.1 3.2 3.3 3.4
--- --- --- --- ---
Canonical definíció ② v1.0.1 v1.1 v1.2
Hibátlan matematika 0 2 4 **6**
DIV konzisztencia (Perplexity) ② 3/9 3/9 **9/9**
DIV-checkbox alkalmazás n/a n/a n/a **6/6**
Amazónia konszenzus-szórás magas — 0.2 **0.2**
Nem-végrehajtók 0 2 2 **1**

Legstabilabb esemény: Amazóniai erdőtüzek — szórás 0.2, minden rendszer ②/② szinten konvergál három teszten keresztül.

Leginstabilabb esemény: Izlandi vulkán — szórás 1.4, a BIO-aktiválás és a DIV-értelmezés rendszerenként jelentősen eltér.

Az EFU-900 Canonical v1.2 — összefoglalás

...

$$EFU = \text{ROUND}((SEV + SPR + FLX + DIV) / 4, 1)$$

$$SPR = \text{ROUND}(SPR_L \times 0.4 + SPR_G \times 0.6, 1)$$

SPR_G soft cap: ha $\text{FLX} \geq 3 \rightarrow \text{SPR_G} \leq 2 \times \text{FLX}$

FLX korlátok:

Csak MAT / ENE / BIO → max 6

MAT + BIO (ENE nélkül) → max 6

MAT + ENE vagy ENE + BIO → max 8

MAT + ENE + BIO → max 10

Akciótípusok:

EFU < 4.5 → 🟢 Zöld

EFU 4.5–6.0 → 🟡 Sárga

EFU 6.0–7.5 → 🟠 Narancssárga

EFU 7.5–9.0 → 🟥 Piros

EFU > 9.0 → 🚨 Kritikus

Tervezett fejlesztések (v1.3)

1. **Izland BIO-aktiválás:** pozitív példák beépítése a definícióba
2. **SEV-indoklás:** explicit tiltás VUL-alapú csökkentés ellen
3. **DIV Infrastruktúra:** példa alapú pontosítás
4. **Rangsor-inkonziszencia:** belső ellenőrzési szabály

V2.0 tervezett elemek:

- VUL mint pontozott 5. mutató

- JSON Schema (Gemini javaslata)
- CWV súlyozott képlet (Gemini javaslata)
- Valós idejű OSINT adatintegráció

Részt vevő rendszerek

Claude (Anthropic) · Copilot (Microsoft) · Gemini (Google) ·
Grok (xAI) · Perplexity AI · ChatGPT (OpenAI) · Le Chat (Mistral) ·
Monica AI

Hivatkozások és adatok

A teljes teszsorozat (prompts, raw outputs, audit táblák) elérhető
a projekt repository `/tests/` mappájában.

Canonical dokumentumok:

- `efu-900-v1.0.1.md`
- `efu-900-v1.1.md`
- `efu-900-v1.2.md` ← aktuális

EFU-900 Nookratikus Teszt — nyílt módszertani keretrendszer

Közreműködés és visszajelzés: Issues és Pull Requests formájában