

**Lovas Róbert**

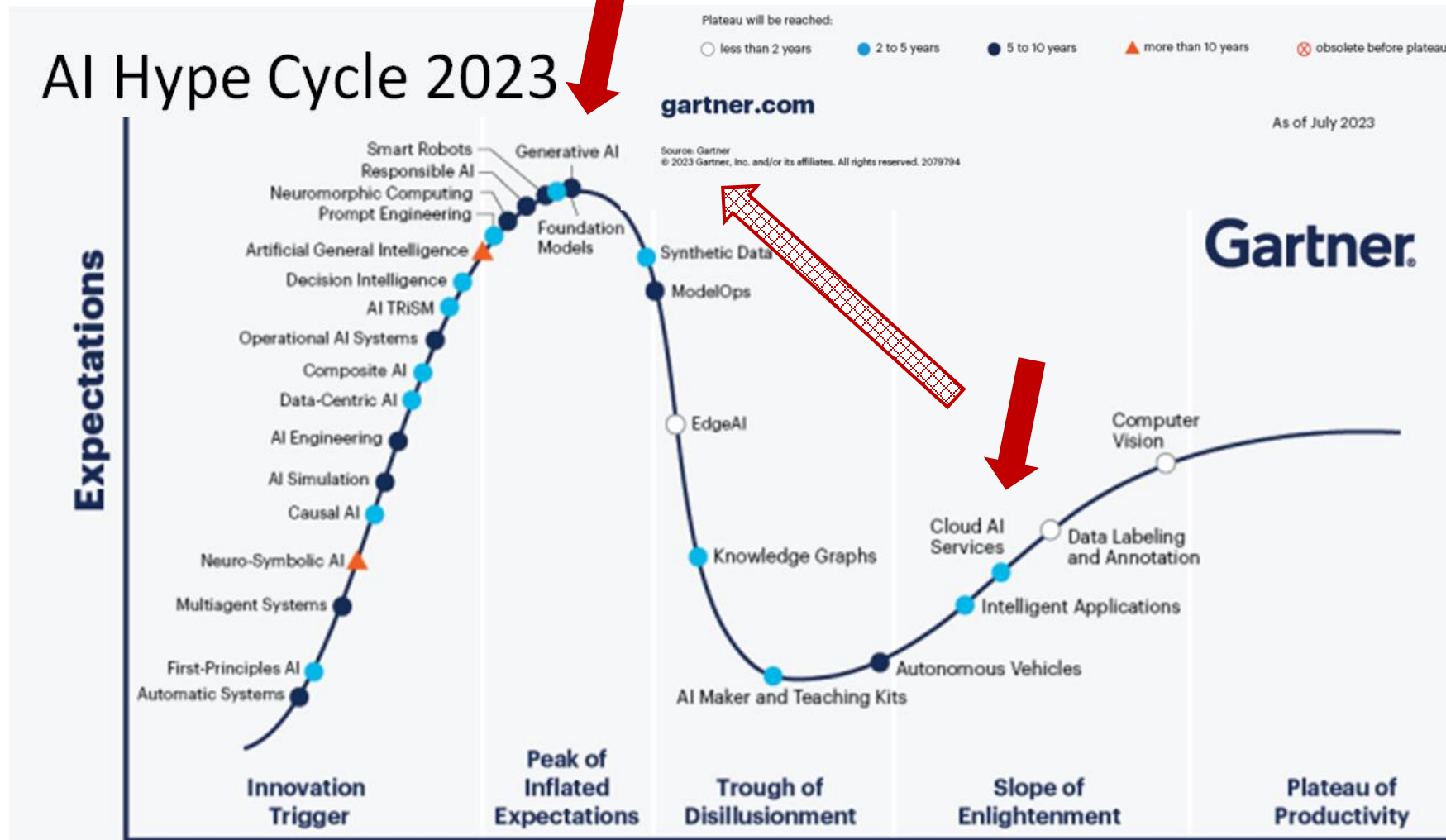
HUN-REN SZTAKI  
igazgatóhelyettes



A HUN-REN Cloud szerepe a kutatási  
hálózat AI4Science programjában

<https://hun-ren.hu/ai-4-science>

## AI Hype Cycle 2023



	KLASSZIKUS AI	GENERATÍV AI
<b>TÁMOGATÓ</b> jobb kutatási kéességek	Könnyebb, mélyebb, több modalitású adatelemzési kéesség	Könnyebb, jobb minőségű írás, kódolás, irodalomelemzés
<b>TRANSZFORMATÍV</b> új kutatási lehetőségek	Új szimulációs, adatgyűjtési, metodológiai lehetőségek (pl. Alphafold)	Új modellezési lehetőségek (pl. Computational social science)

Mennyire tartja akadálynak az alábbi tényezőket az MI adta lehetőségek jobb kihasználásakor?

1. egyáltalán nem akadály    2. kisebb mértékben akadály    3. közepes mértékben akadály  
4. jelentős akadály    5. kritikus    6. Nem tudom megítélni

MI alapú, licenszelhető szoftverek megvásárlása

MI alapú, nyílt forráskódú szoftverek telepítése, futtatása

MI alapú szoftverekhez szükséges számítási infrastruktúra elérhetősége

MI alapú szoftverekhez szükséges számítási infrastruktúra könnyű használhatósága, vag...

MI alapú szoftverek használatához szükséges saját ismeretek



Ön szerint a kutatásai során használt MI megoldásokhoz szükséges informatikai erőforrások mennyisége a közeljövőben (2-3 éven belül) hogyan fog változni?



## 1. Infrastrukturális fejlesztések és elérhetőség

**Szuperszámítógépes infrastruktúra hozzáféréseinek egyszerűsítése:** A kutatók szerint a hozzáférési folyamatok túlzottan bonyolultak, és az egyszerűsítés sokat segítené, különösen az új felhasználóknak.

**Számítási kapacitás és hardver frissítések:** Több kutató említette, hogy a szerverek és a lokális munkaállomások hardvere elavult, és gyakori frissítésre van szükség, hogy lépést tartsanak az AI technológiák gyors fejlődésével.

**Jobb hozzáférés a szerverparkokhoz:** A HUN-REN Cloud és hasonló szerverparkok használata jelenleg nem elég intuitív. Ezen felül, a szerverek számítási teljesítménye nem mindig kielégítő.

## 2. Képzés és tudásbázis bővítése

**AI és gépi tanulás oktatásának fejlesztése:** A felsőoktatásban az AI alapú technológiák oktatásának modernizálására és bővítésére van szükség, hogy a kutatók jobban felkészüljenek a jövő technológiáira.

**Szakértői támogatás és konzultáció:** Szükség van olyan platformokra, ahol a kutatók konkrét segítséget kaphatnak AI projektek megvalósításához, valamint az AI eszközök használatának megismertetésére.

**Átfogó képzési programok:** Több kutató jelezte, hogy szükség van átfogó képzési programokra, amelyek segítik őket az AI technológiák hatékony használatában.

**KUTATÁSOK  
HATÉKONYSÁGÁNAK  
NÖVELÉSE**

*Innovációs értéklánc -  
növekedjen a megvalósuló  
kutatások, innovációk  
társadalmi hasznosulása*

**HAZAI ÉS NEMZETKÖZI  
EGYÜTTMŰKÖDÉS  
FEJLESZTÉSE,  
BEÁGYAZÓDÁS A  
NEMZETKÖZI  
HÁLÓZATBA**

*Cél: a tudásmegosztás, a  
multidiszciplináris  
együttműködések serkentése*

**LEGMODERNEBB  
TECHNOLÓGIÁK  
HASZNÁLATÁNAK  
TÁMOGATÁSA**

*Cél: a rendelkezésre álló  
technológiák folyamatos  
fejlesztése*

**A szolgáltatási funkció fejlesztésének zászlóshajója: az AI4Science program.  
Indítás: 2024. szeptember 23.**

## INTÉZMÉNYI AI NAGYKÖVETI HÁLÓZAT



*Milyen témákkal lehet a kutatóhelyi AI Nagykövethez fordulni?*

**Minden kutatóhelyen dedikált AI Nagykövet** segíti a kutatókat, hogy a mesterséges intelligencia használatát beépítsék a kutatásaikba. A nagykövetek inspirációt adnak az AI használatához, elősegítik az új módszerek megismerését és támogatják, hogy minden kutató megtalálja a számára legnagyobb hozzáadott értéket, amelyet az AI hozhat.

## AI OKTATÁS ÉS INSPIRÁCIÓ



*Az AI milyen területein mélyíthetik el a kutatók a tudásukat?*

### Alapozó képzések

A cél az általános alapok elsajátítása, amelynek nyomán a kutatók tovább tudnak lépni személyes témák felé.

### Specialista képzések

A hálózaton belül fejlesztett saját képzések, amelyek egy-egy tudományterület egyedi kérdéseit járják körül

## SZEMÉLYES AI TÁMOGATÁS



*Hová lehet fordulni az AI-t igénylő kutatási ötlettel?*

### HUN-REN Központ AI szakértői csapat

1. kutatási ötletek validálási és kutatástervezési támogatása
2. infrastruktúra-támogatás
3. alkalmazástámogatás

### Külső szakértői csapat

Neuron Solutions külső szerződéses partnerként

### AI kutatási partnerkereső

AI kompetenciákkal rendelkező, más kutatási területek iránt nyitott kutatók + AI alkalmazási ötletekkel rendelkező kutatók.

## ELÉRHETŐ AI TECHNOLÓGIÁK



*Milyen biztonságos AI technológiák használhatók érzékeny kutatási területekhez?*

### Számítási infrastruktúra

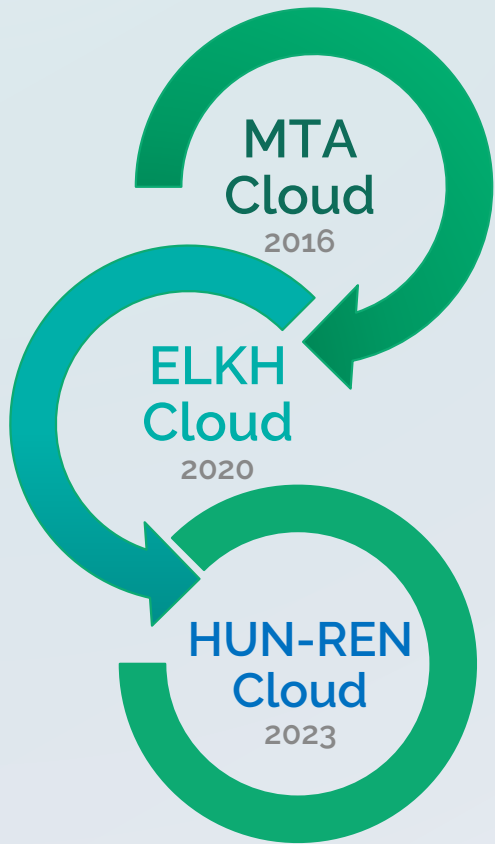
1. Biztonságos GenAi futtatása, akár saját környezetben
2. Gépi tanulás elősegítése akár elosztott, több GPU kártya támogatásával

### Saját AI keretrendszer

1. HUN-REN Cloud
2. Komondor



# A HUN-REN Cloud projekt céljai



- Európai színvonalú és kapacitású számítási infrastruktúra biztosítása a **HUN-REN SZTAKI** és **HUN-REN Wigner Adatközpont** felhőinek föderációjával
- Nem csak az HUN-REN tagintézmények számára, hanem a teljes magyar kutatói közösség felé nyitni
- Kiemelten támogatni a **mesterséges intelligencia** kutatásokat
- Terjeszteni a felhő alkalmazásának **kultúráját** a hazai kutatók körében
- Támogatni a kutatókat alkalmazásaik felhőre történő **adaptálásában**
- Bekapcsolódni az **európai informatikai infrastruktúra**-fejlesztések ökoszisztémájába

**Összesen igényelt projektek száma**  
Aggregált adatok évenkénti lépésekben  
(db)



# A HUN-REN Cloud kapacitásának bővülése (2021-2022)

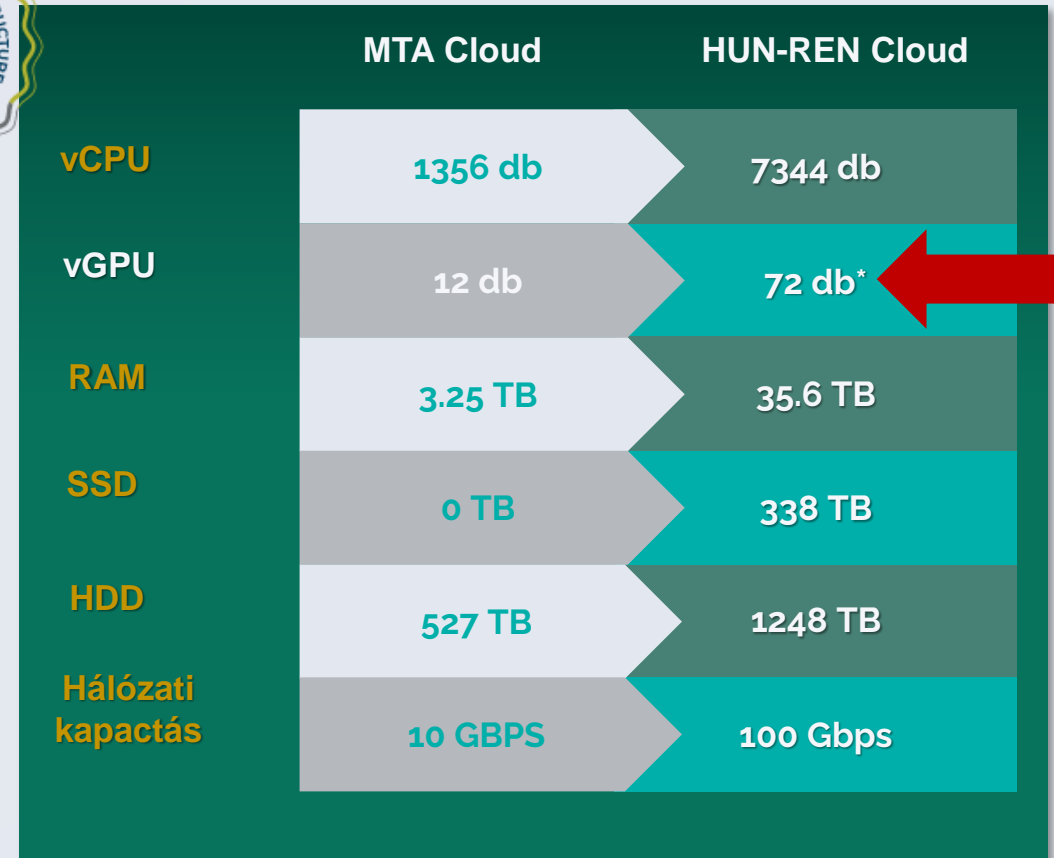


Hazai és nemzetközi eredmények:

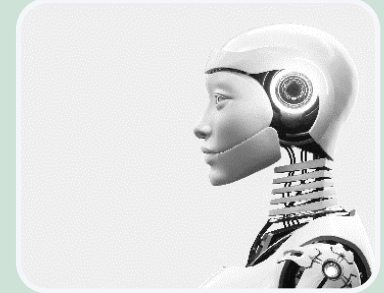
- Elnyerte el az NKFIH „**Kiváló Kutatási Infrastruktúra**” címét
- Illesztés az **Európai Nyílt Tudományos Felhőhöz (EOSC)**, az EU egyik legjelentősebb digitális infrastruktúrájához
- A **Kutatási Infrastruktúrák Európai Stratégiai Fóruma (ESFRI)** legújabb, 20-25 éves időtávot átölelő roadmap-jén



*További MI erőforrások a hazai kutatók számára is*



# Disszemináció és hasznosítás



**232**

tudományos publikáció

2021-23: **80+** db

**21**

új online szakmai rendezvény

 **1300+** résztvevő

**29**

igénybe vevő intézmény

 **353** projekt

**9**

újonnan támogatott egyetem

BME ELTE SZTE PTE  
SE ÓE PE DE SZE

**3**

támogatott nemzeti laboratórium

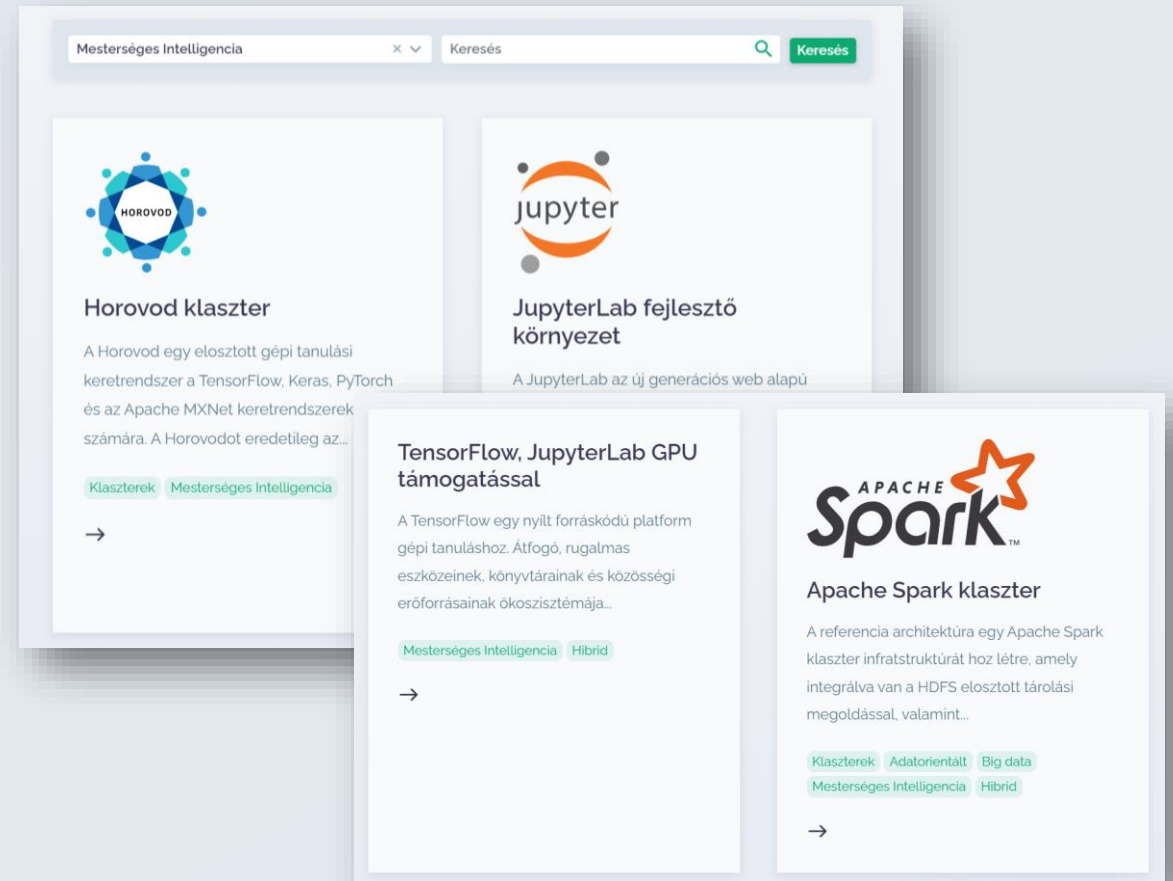
 **MESTERSÉGES INTELLIGENCIA**  
Nemzeti Laboratórium

 **AUTONOMOUS SYSTEMS**  
National Laboratory

 **DH** National Laboratory for Digital Heritage

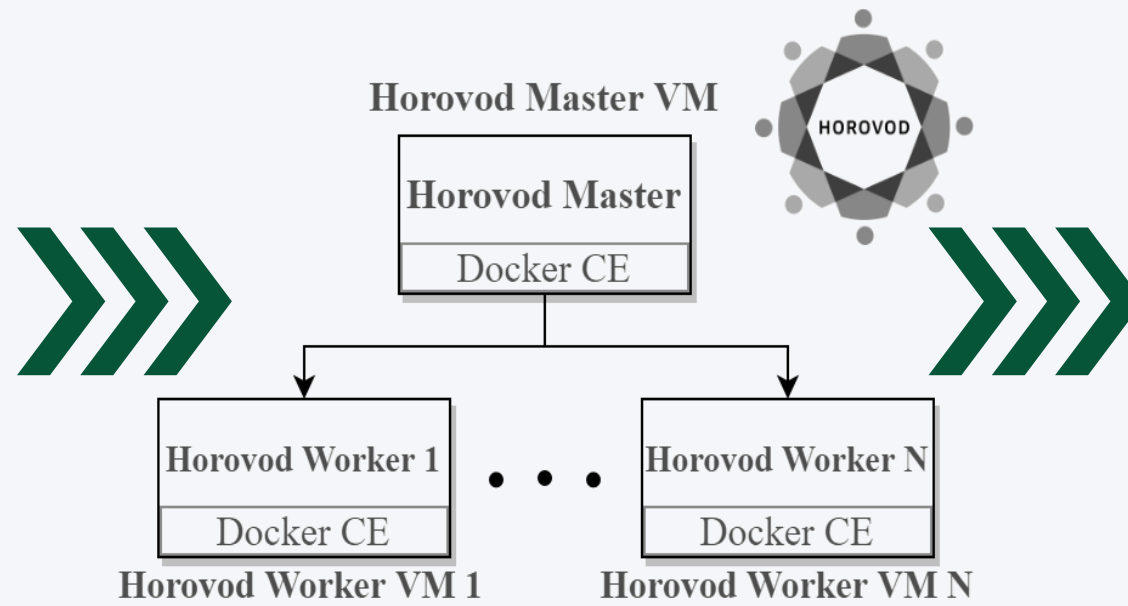
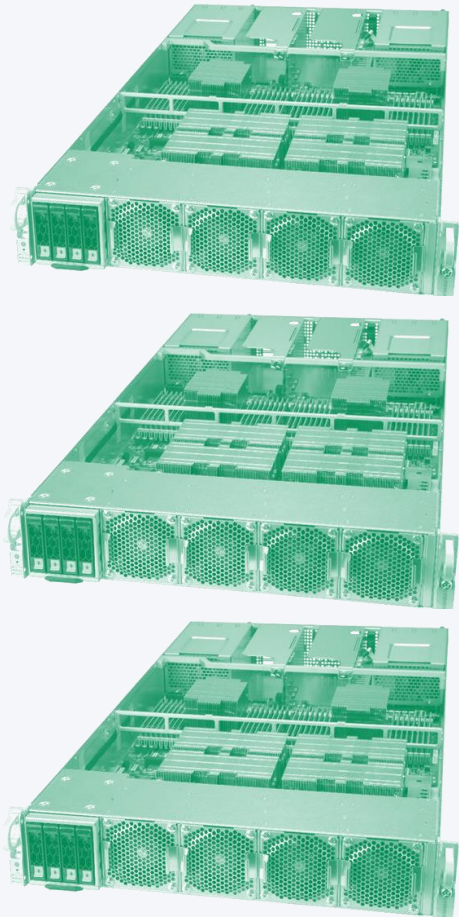
# Referencia architektúrák, mint kulcsrakész platformok

- A kutatásokhoz gyakran összetett, nagy léptékű platformot kell kialakítani, ami számos szoftver és szolgáltatás összehangolt működésén alapul, ezért átszabható, megbízható és skálázható **referencia architektúrákat** biztosítunk a felhasználóknak a legjobb gyakorlatok alapján
- A különböző, gyakran igényelt 16 referencia architektúra részletes dokumentációval elérhető a felhő webportálján – köztük **4 MI célú**



➔ <https://science-cloud.hu/referencia-architekturak>

# Mély gépi tanulás elosztott GPU platformon



# Tudásbázis építése és tudástranszfer

← Címlap | Előadások → Mesterséges Intelligencia és gépi tanulás

## Rekurrens Neurális Hálózatok

- ▶ Idősoros adatokon végzett feldolgozás
- ▶ Visszacsatolás a hálózatban a korábbi állapot tárolására

Input: a Word → Stateful Model (Recurrent Neural Network) → Output: Most likely next word

Memory of previous words influence next prediction

Output to for: Machine

ELKH Cloud

Gábor Kertész

- Mesterséges Intelligencia és gépi tanulás 00:01
- Jelen 01:41
- Terminológia 06:37
- Adatok 08:38

← Címlap | Előadások → ELKH Cloud általános használata

## Bejelentkező képernyő

openstack.

Log in

Authenticate using

SAML Federation login (eduID, eduGAIN, Security Assertion Mark)

If you are not sure which authentication method to use, contact your administrator

Sign In

ELKH Cloud

SZTAI

ELKH

00:36 / 41:36

← Címlap | Előadások → ELKH Cloud referencia architektúra koncepció

## Horovod Performance Evaluation

Sample	Ideal (%)	Measured (%)
1	100	100
2	100	98
3	100	95
4	100	92
5	100	88
6	100	82
7	100	78
8	100	89

Sample	Ideal (%)	Measured (%)
1	100	100
2	100	97
3	100	94
4	100	92
5	100	89
6	100	79
7	100	77
8	100	73

2023.10.26. Reference architecture concept on HUN-REN Cloud

HUN-REN SZTAI

15:08 / 22:49

← Címlap | Előadások → Az ELKH Cloud előnyei a POLTEXT projekt példáján

## Coding MN articles

Coded NS articles → Uncoded MN articles

Features: stopwords removed, stemmed, TF-IDF

Loop for rounds of coding

Training set → Virgin test set

Classification SVM

- training one vs all
- for all codes in training set
- proportional training sets
- T samples \* T SVM models

Sample 1 SVM 1, SVM 2, SVM 7

Sample 2 SVM 1, SVM 2, SVM 7

Sample 7 SVM 1, SVM 2, SVM 7

Move newly accepted coded articles from test set to training set

Human validation of coded MN articles → Uncoded MN articles

- If all SVM results for all samples concur we have a verdict for that code for given article
- If there is only one verdict for an article, it is classified accordingly
- Human supervision of results (not validation)

01:07:43 / 01:36:11

08:25 / 10:31

- Az ELKH Cloud előnyei a POLTEXT projekt példáján 00:01
- Research context 00:31
- Research problem 00:57

➔ <https://science-cloud.hu/eloadasok>

# Új AI szolgáltatás: GenAI4Science

The screenshot displays the GenAI4Science web interface. At the top, there are navigation tabs: Chat, Default, Notebook, Parameters, Model, Training, and Session. The main chat area shows a conversation:

- AI:** How can I help you today?
- You:** Generate a hello-world code in python
- AI:** Here's your code:  

```
print('Hello world!')
```
- You:** Generate me a tensorflow code for cnn network
- AI:** Sure! Here's an example of how you could create a simple CNN using TensorFlow:  

```
import tensorflow as tf
from tensorflow import keras
from skimage import io, filters
import numpy as np

# Define data
X = np.random.rand(1000, 28, 28)
y = np.random.randint(0, 10, size=1000)

# Create model
model = Sequential()
model.add(Conv2D(32, (3, 3), activation='relu', input_shape=(28, 28, 1)))
model.add(MaxPooling2D((2, 2)))
model.add(Flatten())
model.add(Dense(64, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(10, activation='softmax'))
```

On the right side, there are control panels for 'Start reply with' (with a text input 'Sure thing!'), 'Mode' (with radio buttons for 'chat', 'chat-instruct', and 'instruct'), and 'Chat style' (with a dropdown menu set to 'cai-chat'). At the bottom, there is a 'Send a message' input field and a 'Generate' button.

On the left side, there is a terminal window showing the output of 'nvidia-smi' and 'ps' commands:

```
Every 2.0s: nvidia-smi
Tue Feb 27 15:55:01 2024
+-----+
| NVIDIA-SMI 525.125.06   Driver Version: 525.125.06   CUDA Version: 12.0   |
+-----+-----+
| GPU   Name               Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf    Pwr:Usage/Cap|  Memory-Usage | GPU-Util  Compute M. |
|-----+-----+-----+
| 0   GRID V100DX-32C     On         | 00000000:06:00:0 | Off         | 0          |
| N/A   N/A    P0     N/A /  N/A | 18273MiB / 32768MiB | 57%      Default  |
|-----+-----+-----+
| Processes:
| GPU   GI   CI          PID    Type   Process name          GPU Memory |
|-----+-----+-----+
| 0   N/A  N/A         18752    C   python                 18273MiB |
+-----+-----+-----+
```

Szolgáltató:



# PULI demonstrációs rendszer

## Elkezdődött a migráció a HUN-REN Cloud infrastruktúrára

HUN-REN NYELVTUDOMÁNYI KUTATÓKÖZPONT

### PULI nyelvmodellcsalád

#### Utasításkövető (*instruct*) modellek

Ezeket a modelleket úgy finomhangoltuk, hogy képesek legyenek követni az utasításokat. Így vagy egy kérdést várnak, amire válaszolnak ("Ki volt Dugovics Titusz?", "Mi Spanyolország fővárosa?"), vagy egy utasítást, hogy tegyenek meg valamit ("Írj egy fogalmazást arról, hogy ...", "Készíts programtervet egy háromnapos budapesti városnézéshez...").

**ParancsPULI**  
Magyar nyelvű utasításkövető finomhangolt PULI-GPTrio

**PULI Llumix 32K Instruct**  
32K kontextus ablakú, magyar nyelvű utasításkövető Llama-2 modell

#### Nyers (*pretrained*) modellek

Ezeket a modelleket nagy mennyiségű szövegen tanítottuk. Amikor kapnak egy promptot, a nyers modellek kiszámolják, melyek azok a szavak, amelyek statisztikailag valószínűsíthetően követik a kapott szöveget; egyszerűen tehát megfelelő folytatást generálnak az általunk beírt szöveghez ("Írtam egy tündérmesét. Egyszer volt, hol nem volt, volt egyszer egy ..."). Nem feladatuk utasítást végrehajtani, vagy kérdésre válaszolni.

**PULI 35X**  
Az első csak magyar nyelvű GPT-3 alapmodell

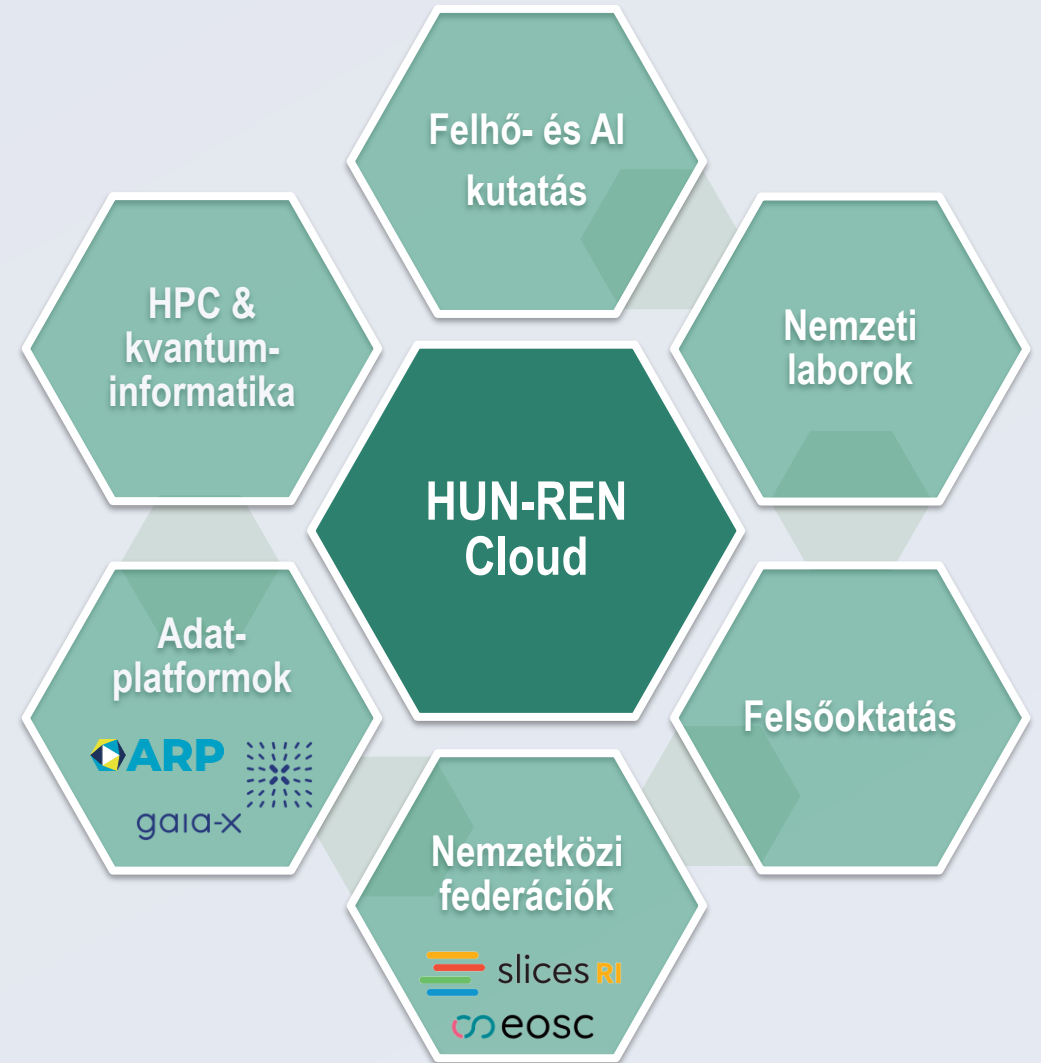
**PULI Trio**  
Háromnyelvű (magyar, angol, kínai) GPT-3 modell

**PULI Llumix 32K**  
32K kontextus ablakú, magyar nyelvű továbbtanított Llama-2 modell



# Jövőkép

- HUN-REN Cloud szolgáltatásainak folyamatos fejlesztése **AI irányába**
- Hazai partnerségek erősítése
- Nemzetközi beágyazódás növelése
- Átjárhatóság és szinergiák
  
- További **erőforrásokhoz hozzáférés** biztosítása
- **Felhasználói bázis** szélesítése
  
- *Hatékonyabb K+F folyamatok*
- *Még versenyképesebb hazai kutatóműhelyek*



***Köszönöm a figyelmet!***

