

Ez a dokumentum kifejezetten a pályázati tanácsadó számára készült.A dokumentum hivatkozik a **DIMOP Plusz-1.2.3/A-24** (digitális infrastruktúra) és a **GINOP Plusz-2.1.3-24** (innováció) pályázatokban elszámolható költségkategóriákra.

MŰSZAKI SPECIFIKÁCIÓ ÉS IGÉNYBEJELENTŐ

Projekt megnevezése: AI Ügynökrendszer (BAS) és Logisztikai Platform (Pohi AI Pro)
Infrastrukturális Hátterének Megteremtése

Kedvezményezett: Pohánka és Társa Kft.

Dátum: 2026. január 7.

1. SZOFTVER ÉS FELHŐSZOLGÁLTATÁSOK (Cloud Infrastructure)

Cél: A Brunella rendszer skálázható, biztonságos futtatási környezetének és a nagy nyelvi modellek (LLM) API hozzáféréseinek biztosítása.

Szolgáltatás Megnevezése	Specifikáció / Csomag Részletei	Szakmai Indoklás (A DFK-hoz és Pályázathoz)



<p>Google Cloud Platform (GCP)</p>	<p>Compute & Serverless:</p> <ul style="list-style-type: none"> • Cloud Run (konténerizált backend futtatása) • Compute Engine (GPU-gyorsított VM-ek) <p>AI & Machine Learning:</p> <ul style="list-style-type: none"> • Vertex AI: Hozzáférés a Gemini 1.5 Pro és Flash modellekhez • AutoML: Saját modellek finomhangolása (Fine-tuning) 	<p>A Brunella Agent System mikroszolgáltatás alapú architektúrájának (Python/FastAPI) kiszolgálása. A Vertex AI biztosítja az üzleti logikát vezérlő nagy nyelvi modelleket. A Cloud SQL tárolja a LangGraph ágensek állapotát (memory) és a vállalati tudásbázist. Ez a szolgáltatás a "Szoftver bérlet / Felhő szolgáltatás" kategóriában számolható el.</p>
---	---	---



	<p>Adattárolás:</p> <ul style="list-style-type: none"> • Cloud SQL for PostgreSQL (vektoros kiegészítéssel) • Cloud Storage (dokumentumok tárolása) 	
<p>Fejlesztői Eszközök & DevOps</p>	<p>CI/CD és Container Registry:</p> <ul style="list-style-type: none"> • Google Artifact Registry (Docker image-ek) • Cloud Build (automatizált tesztelés és deploy) 	<p>A folyamatos fejlesztés (CI/CD) és a stabil működés biztosítása. A fejlesztői környezet és az éles rendszer közötti szinkronizációhoz elengedhetetlen, különösen a Docker konténerek kezelésénél.</p>



	Monitoring: <ul style="list-style-type: none"> • Google Cloud Operations Suite (Logging, Monitoring) 	
Kiegészítő AI Szolgáltatások	Egyéb API Előfizetések: <ul style="list-style-type: none"> • OpenAI API (GPT-4o fallback) • Anthropic API (Claude 3.5 Sonnet kódoláshoz) • Browserbase / Puppeteer (Headless böngészéshez) 	A "Piacfigyelő" és "Őrszem" modulok működéséhez szükséges redundáns AI modellek és webes adatgyűjtő (scraping) technológiák biztosítása.

2. ESZKÖZBESZERZÉS: AI FEJLESZTŐI MUNKAÁLLOMÁS (Hardware)

Cél: A fejlesztési ciklus gyorsítása, lokális modellfuttatás (Ollama/Llama 3), adatbiztonság és a jelenlegi technikai korlátok (C: meghajtó telítődés, lassú inferencia) felszámolása.



Eszköz Típusa	Ajánlott Konfiguráció (Minimum Specifikáció)	Szakmai Indoklás (A DFK-hoz és Pályázathoz)
Processzor (CPU)	<p>High-End Workstation CPU</p> <ul style="list-style-type: none"> • Típus: AMD Ryzen Threadripper 7000 széria (vagy ekvivalens Intel Xeon W) • Magok száma: Min. 24 mag / 48 szál 	<p>A párhuzamos adatfeldolgozás (Data Preprocessing) és a Docker konténerek (backend, adatbázis, cache) egyidejű futtatásához szükséges számítási kapacitás.</p>
Grafikus Gyorsító (GPU)	<p>Dual-GPU Setup (AI Training fókusz)</p> <ul style="list-style-type: none"> • Opció A (Ár/Érték): 2 db NVIDIA GeForce RTX 4090 (24GB VRAM / kártya) • Opció B (Enterprise): 1 db NVIDIA RTX 6000 Ada Generation (48GB VRAM) 	<p>A projekt legkritikusabb eleme. A 48GB VRAM szükséges a Llama 3 70B és Qwen 2.5 Coder modellek lokális futtatásához és finomhangolásához (fine-tuning). A jelenlegi hardverek memóriakorlátja akadályozza a fejlesztést.</p>



	<ul style="list-style-type: none"> • <i>Összes VRAM igény: Min. 48 GB</i> 	
Memória (RAM)	128 GB - 256 GB DDR5 ECC (Error Correction Code támogatással)	A nagy méretű adathalmazok és a memóriában tartott AI modellek kezeléséhez. Az ECC (hibajavító) memória kritikus a hosszú futású tanítási folyamatok stabilitásához.
Háttértár (Storage)	Osztott SSD Rendszer: <ul style="list-style-type: none"> • Rendszer: 2 TB NVMe Gen5 SSD (>10,000 MB/s) • Adat/Cache: 4 TB NVMe Gen4 SSD 	A Docker image-ek és a "Layer Caching" okozta tárhelyproblémák végleges megoldása. A Gen5 sebesség szükséges a vektoradatbázisok gyors betöltéséhez.
Hűtés és Ház	Egyedi Workstation Ház + Folyadékhűtés <ul style="list-style-type: none"> • Kiváló légáramlás (Airflow) 	A 24/7-es üzemmódban futó gépi tanulási folyamatok (training runs) hőtermelésének kezelése és az üzembiztonság garantálása.



	<ul style="list-style-type: none"> • 1200W+ Tápegység (Platinum minősítés) 	
--	---	--

3. ÖSSZEGZÉS A TANÁCSADÓNAK (Copy-Paste javaslat)

*"Tisztelt Tanácsadó Partnerünk!

A **Digitális Fejlesztési Konceptió (DFK)** összeállításához mellékeljük a cégünk innovációs céljaihoz (AI Ügynökrendszer fejlesztése és piacra lépése) elengedhetetlenül szükséges technológiai specifikációt.

Kérjük, hogy a pályázati anyagban (DIMOP Plusz / GINOP Plusz) a fenti tételeket az alábbi tevékenységekhez rendelje hozzá:

1. **K+F Tevékenység:** Saját MI modellek tanítása és tesztelése (Hardver igény).
2. **Szolgáltatásfejlesztés:** A Brunella rendszer felhőalapú kiszolgálása és API integrációk (Google Cloud igény).
3. **Digitalizáció:** A vállalati folyamatok automatizálása és adatok biztonságos tárolása.

A megjelölt hardver konfiguráció és felhőszolgáltatás nem általános irodai célokat szolgál, hanem speciális, **nagy számítási igényű mérnöki fejlesztőeszköz**, amely a projekt megvalósíthatóságának alapfeltétele."*



