

AI TRAINING APP

LETNÍ ŠKOLA CZECH AI FACTORY

TOMÁŠ MARTINOVIČ
9.9.2026



EuroHPC
Joint Undertaking



Spolufinancováno
Evropskou unií

MSMT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

AI TRÉNINK NA SUPERPOČÍTAČI

LUMI AI Factory



- Vybrat si superpočítač
- Vyplnit žádost o zdroje
- Čekat na potvrzení
- Vytvořit účet
- Nastavit SSH klíč
- Znalost práce s Linux terminálem
- Přihlásit se na superpočítač
- Naučit se používat dávkový plánovač SLURM
- Naučit se používat moduly
- Naučit se pracovat s daty na superpočítači
- ... konečně něco spočítat

OD PROBLÉMŮ K INOVACÍM

- Spuštění výpočtů jedním tlačítkem
- Optimalizovaná běhová prostředí
- Automatizace a reprodukovatelnost experimentů
- Centralizace (týmová spolupráce)
- Bezpečná správa dat
- Monitorování logů
- Sledování hodnotících metrik modelů



Založená na LEXIS Platformě:

docs.lexis.tech

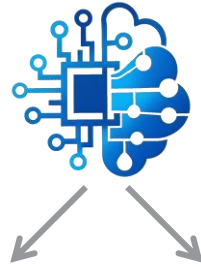
opencode.it4i.eu/lexis-platform



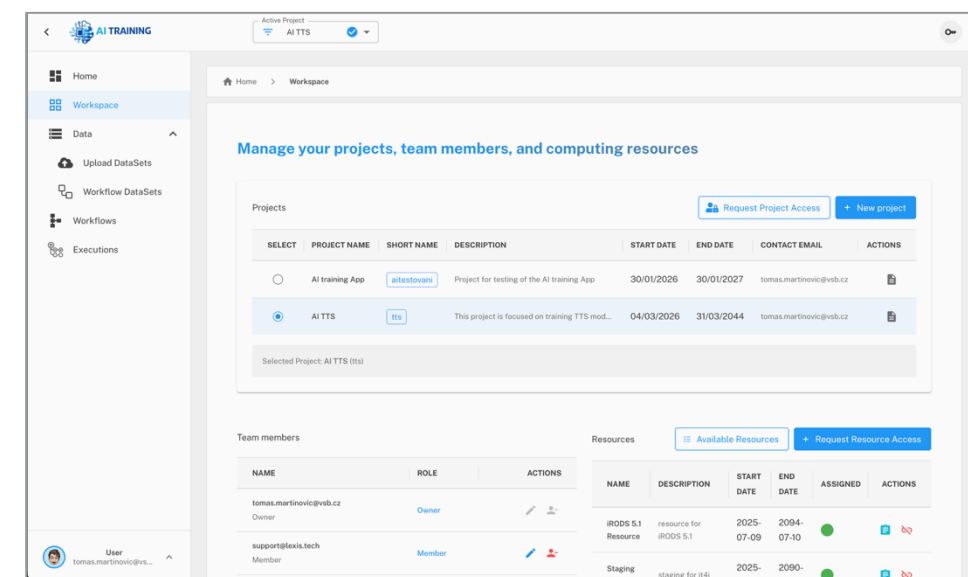
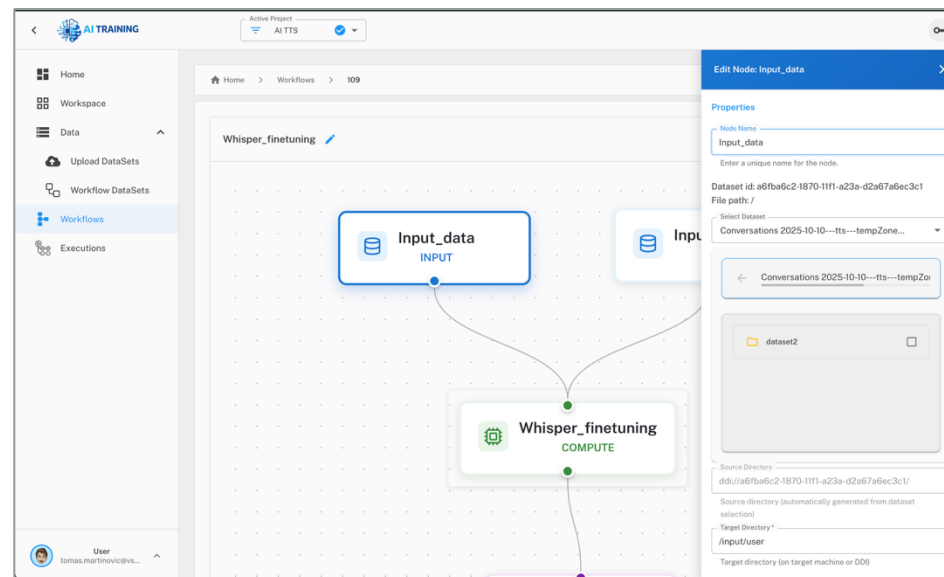
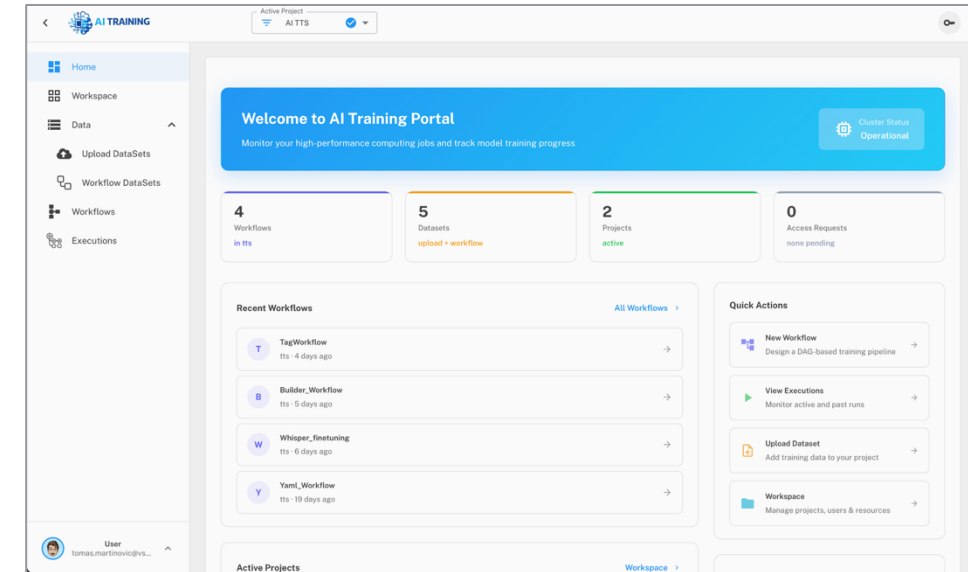
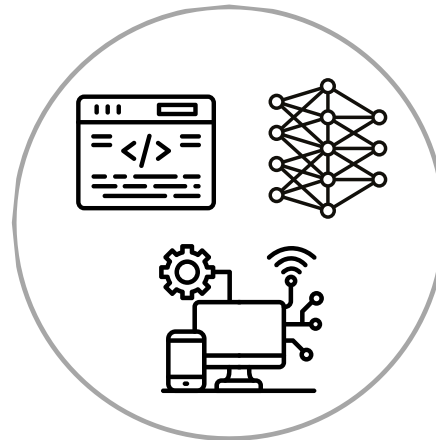
The screenshot shows the AI TRAINING platform landing page. At the top, it says 'AI TRAINING' and 'POWERED BY LEXIS PLATFORM 2'. The main heading is 'AI Training at Supercomputer Scale'. Below this, a subheading reads: 'Train your deep learning models with enterprise-grade high-performance computing resources. Experience unprecedented speed and efficiency in every training cycle.' The section 'Why Choose Our Platform?' features four benefits: 'High Performance Computing' (Leverage distributed computing power for faster AI model training), 'Accelerated Training' (Reduce training time from days to hours with optimized infrastructure), 'Scalability' (Scale your training from a single GPU to thousands of GPUs), and 'Flexibility' (Supports a wide range of deep learning frameworks and models). On the right side, there is a 'Ready to Get Started?' section with the text 'Join thousands of researchers using LEXIS for cutting-edge AI training' and a 'Get Started' button.

PRO KOHO JE NAŠE APLIKACE URČENA?

ZAČÁTEČNÍCI



EXPERTI



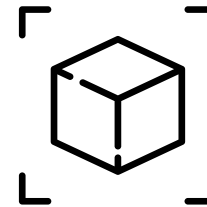
TRÉNINK AI MODELŮ

- Úloha
- Volba vstupních dat a AI modelu
- Předzpracování vstupních dat
- Volba hyperparametrů AI modelu
- Trénink AI modelu
- Hodnocení natrénovaného AI modelu pomocí hodnotících metrik
- Dávková inference
- Optimalizace hyperparametrů AI modelu
- Finální řešení



<https://aitraining.lexis.tech/>

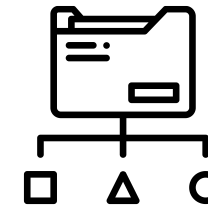
Detekce a
segmentace objektů
v obrazových datech



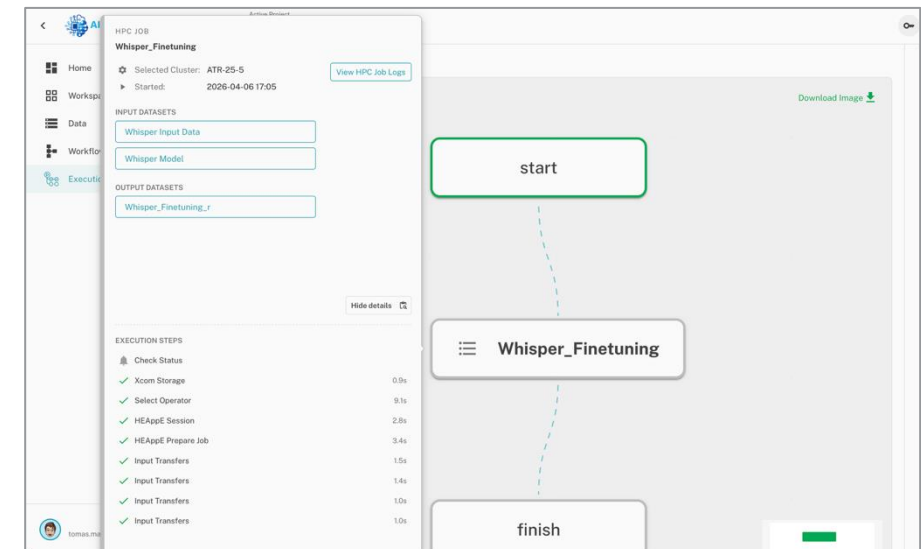
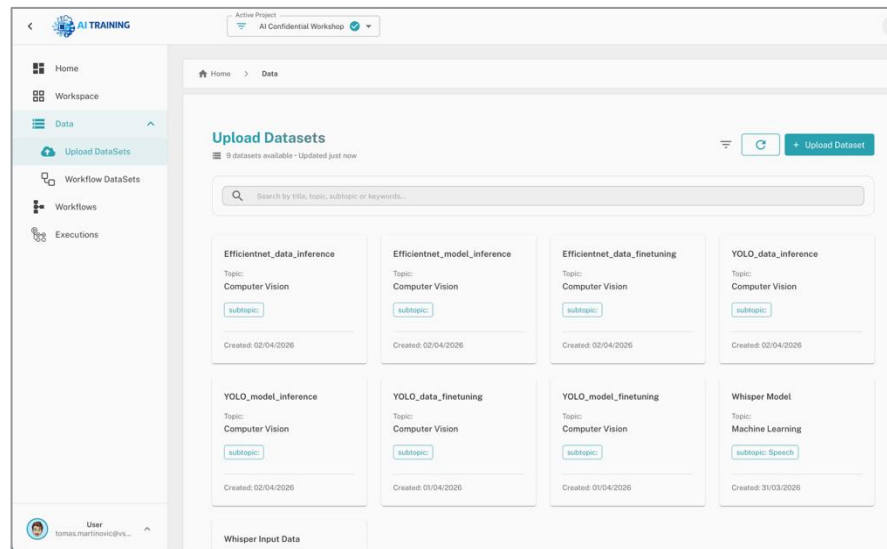
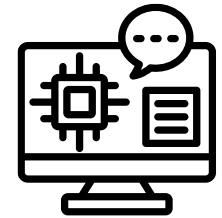
Převod audio
záznamů do
textové podoby



Klasifikační a
regresní úlohy



Fine-tuning LLMs



INFERENČNÍ SLUŽBA



LUM | AI Factory

- OpenAI API kompatibilita
- Podpora
 - vLLM
 - SGLang
 - NVIDIA Triton
- Využívá HPC batch job a uživatel zdroje sdílí jenom s členy projektu - nativní HPC přístup
- Schopnost přejít na nový job a job pre-scheduling
- Import modelů z HuggingFace

<https://aitraining.lexis.tech/>
<https://inference.exa4mind.eu>

Inference Service
Deploy and serve Hugging Face models on GPU resources

SERVICE STATUS ●

The service is active. Use the access token below as the OpenAI API key in your AI tools or clients.

Authentication Token

eyJhbGciOi...

OpenAI Base URL

https://chat.exa4mind.eu/p/atr-26-1/openai/v1

Available Models

• Qwen/Qwen3-4B-Instruct-2507	Not loaded Engine: vllm Uses 1 GPU 30 minutes waittime 30s idle timeout	Load	Unregister
• google/gemma-4-31B-it	Not loaded Engine: vllm Uses 4 GPU 30 minutes waittime 30s idle timeout	Load	Unregister
• google/gemma-4-E2B-it	Not loaded Engine: vllm Uses 1 GPU 30 minutes waittime 30s idle timeout	Load	Unregister
• google/gemma-4-31B-it-qat-w4a16-ct	Not loaded Engine: vllm Uses 2 GPU 60 minutes waittime 60s idle timeout	Load	Unregister

Manage Model Registry

Hugging Face Model ID:

Inference Engine:

GPU Count:

Waittime (seconds):

Waittime Ratio (0-1):

Idle Timeout (seconds):

☐ Autoscale

☒ Is Refreshing

[Register Model](#)



Nové kontejnery

- Šablony



MLflow integrace

- Metriky modelu
- Monitorování využití systému



Vylepšená správa

- Možnost importu vlastního modelu
- Podpora RAG

LUMI AI Factory

KONTAKTUJTE NÁS

EXPERTI LUMI AI FACTORY
ai-factory@lumi-supercomputer.eu

ČESKÝ KONTAKT LUMI AI FACTORY
ai-factory@it4i.cz



EuroHPC
Joint Undertaking



Spolufinancováno
Evropskou unií

MŠMT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY