


ORGANIZAČNÍ INFORMACE

//
*Proč je přihlášení do AI training app
tak komplikované?*
//

Co očekávat v odpolední části?

- Představení IT4I a co jsou AI Factories?
- Jak trénovat model (nejen) na superpočítači Karolina
- Praktická ukázka trénování modelu
- Hands-on práce s natrénovaným modelem se zadáním
- Coffee break
- Benchmarking a efektivita tréninkového procesu
- Technické okénko- jak nastavit tréninkový proces?
- Q&A + networking



LUMI
AI Factory

Kde trénovat vlastní
model?

Kde najít podporu pro
vývoj AI?

ai-factory@it4i.cz

AI Confidential Workshop

Praha
27.5.2026



EuroHPC
Joint Undertaking



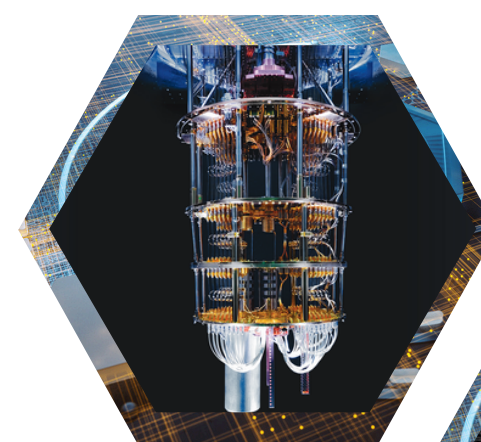
Spolufinancováno
Evropskou unií



Ministerstvo
školství, mládeže
a tělovýchovy

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA

IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM

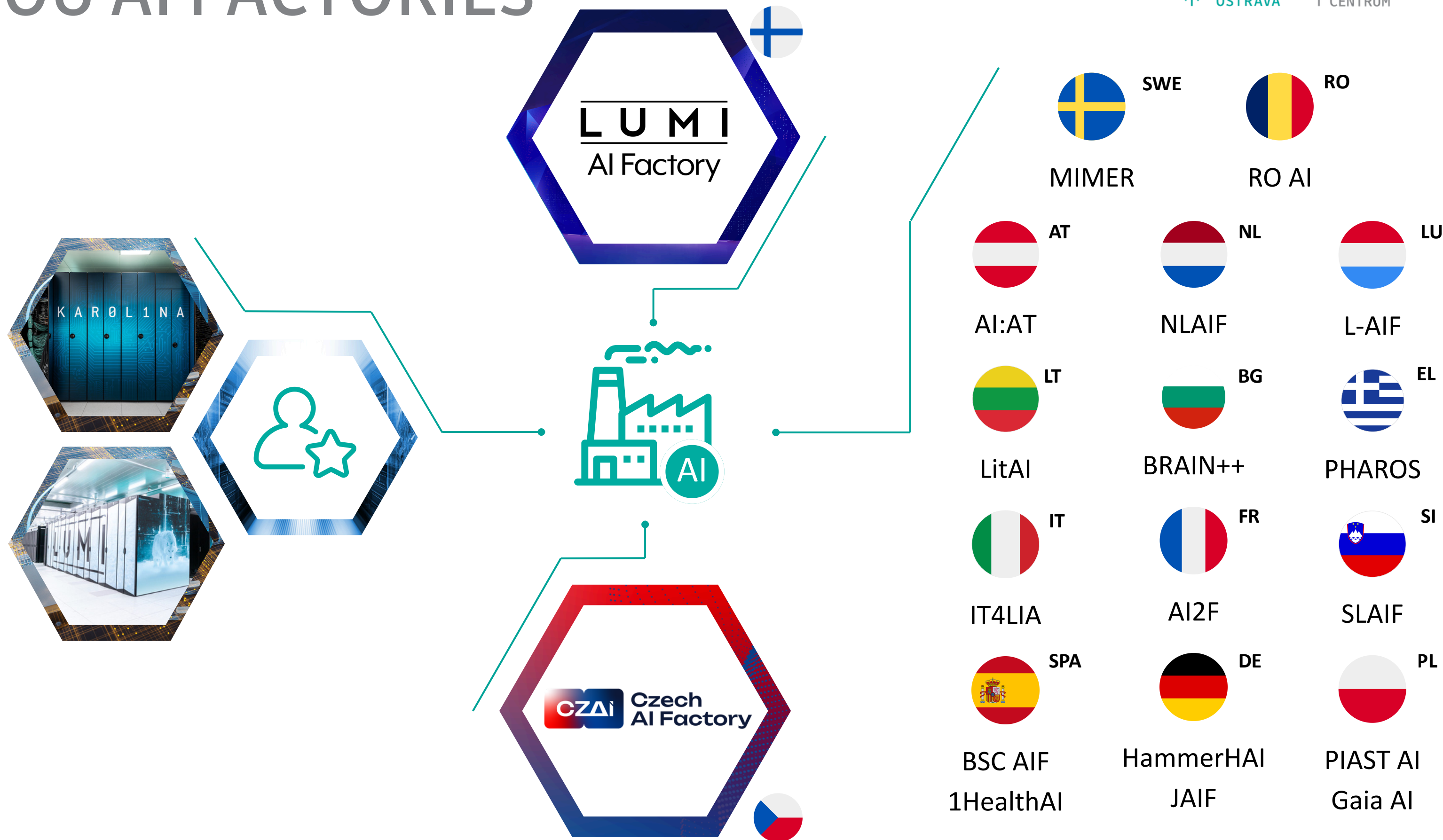


VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA

IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM



CO JSOU AI FACTORIES



PILÍŘE AI FACTORIES



6 310 × 3 570 ×

KAROLINA

A teal-bordered box containing two columns of icons: a laptop and a smartphone. Below each icon is a count. To the right is a hexagonal frame with a blue circuit pattern and the word 'KAROLINA' in white capital letters.

120 800 × 213 600 ×

A teal-bordered box containing two columns of icons: a laptop and a smartphone. Below each icon is a count. To the right is a hexagonal frame with a dark border and an image of a server rack.

Konzultace

Školení

PŘÍSTUP K VÝPOČETNÍ SÍLE

Playground
access

Fast lane
access

Large scale
access

AI for Science
and Collaborative
EU Projects

- Pro začátečníky a nové uživatele
- Do 5 000 GPU hodin
- Přístup do 2 pracovních dnů
- Doba trvání alokace až 3 měsíce

- Pro pokročilé uživatele
- 10 000 až 50 000 GPU hodin
- Rozhodnutí o přijetí do 4 dnů
- Doba trvání alokace až 3 měsíce

- Pro náročné AI projekty
- Více než 50 000 GPU hodin
- Rozhodnutí o přijetí do 10 dnů
- Doba trvání alokace až 12 měsíců

- Pro výzkumné AI projekty
- 20 000 až 90 000 GPU hodin
- Rozhodnutí o přijetí do 1 měsíce
- Doba trvání alokace až 6 měsíců

VLASTNÍ PROJEKT

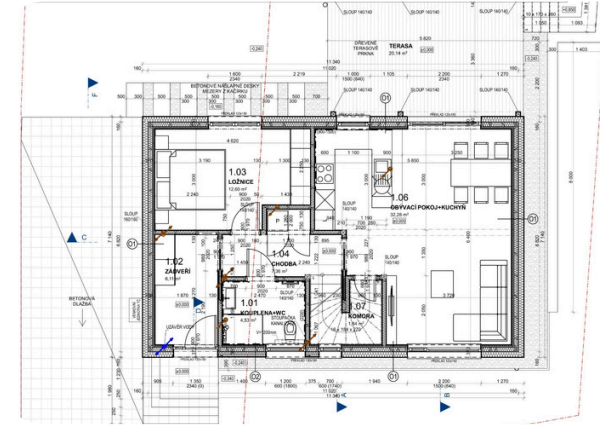
LUMI AI Factory

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA | IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM

Newcomers



Explorers



Adopters



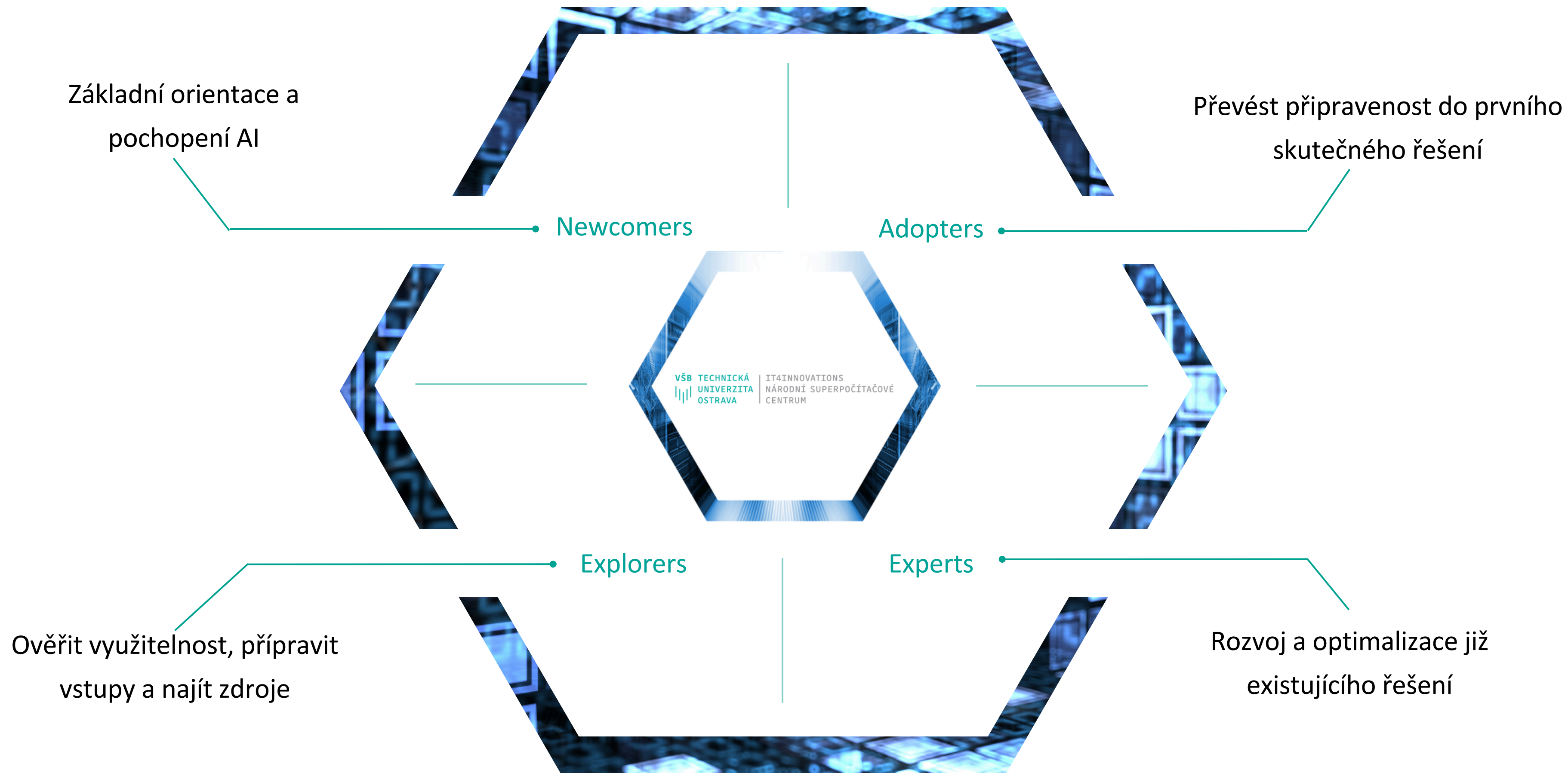
Experts



V ČEM MŮŽEME POMOCI

LUMI AI Factory

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA | IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM



AI SLUŽBY

LUMI AI Factory

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | IT4INNOVATIONS NÁRODNÍ SUPERPOČÍTAČOVÉ CENTRUM



PŘÍKLADY FIREM

LUMI AI Factory

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA | IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM

 my
paperwork

BANDITS

BottleCap AI

 **ValkaAI**

vctrm

AI FACTORIES

LUMI AI Factory

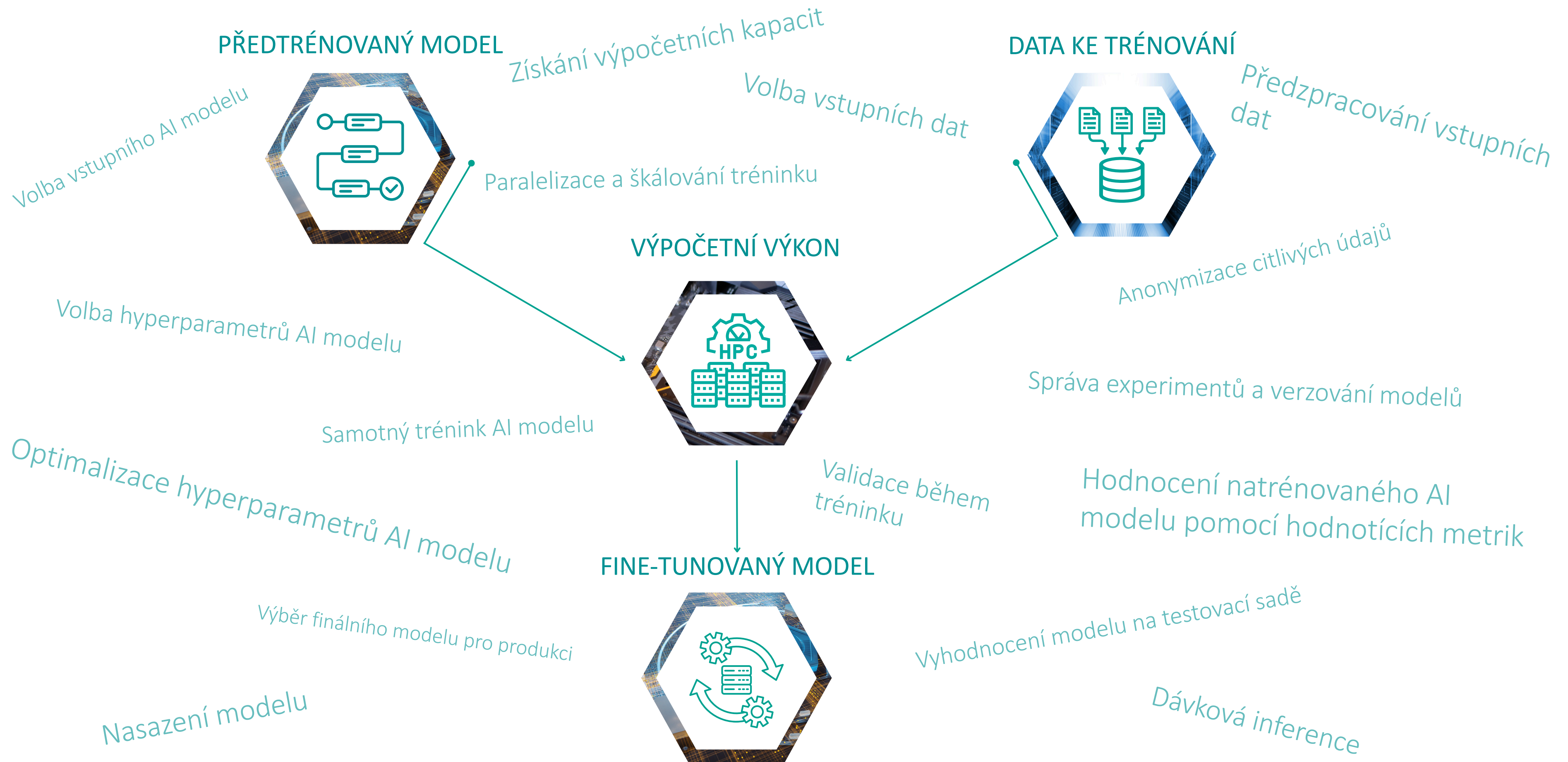
VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA | IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM

LUMI AI Factory



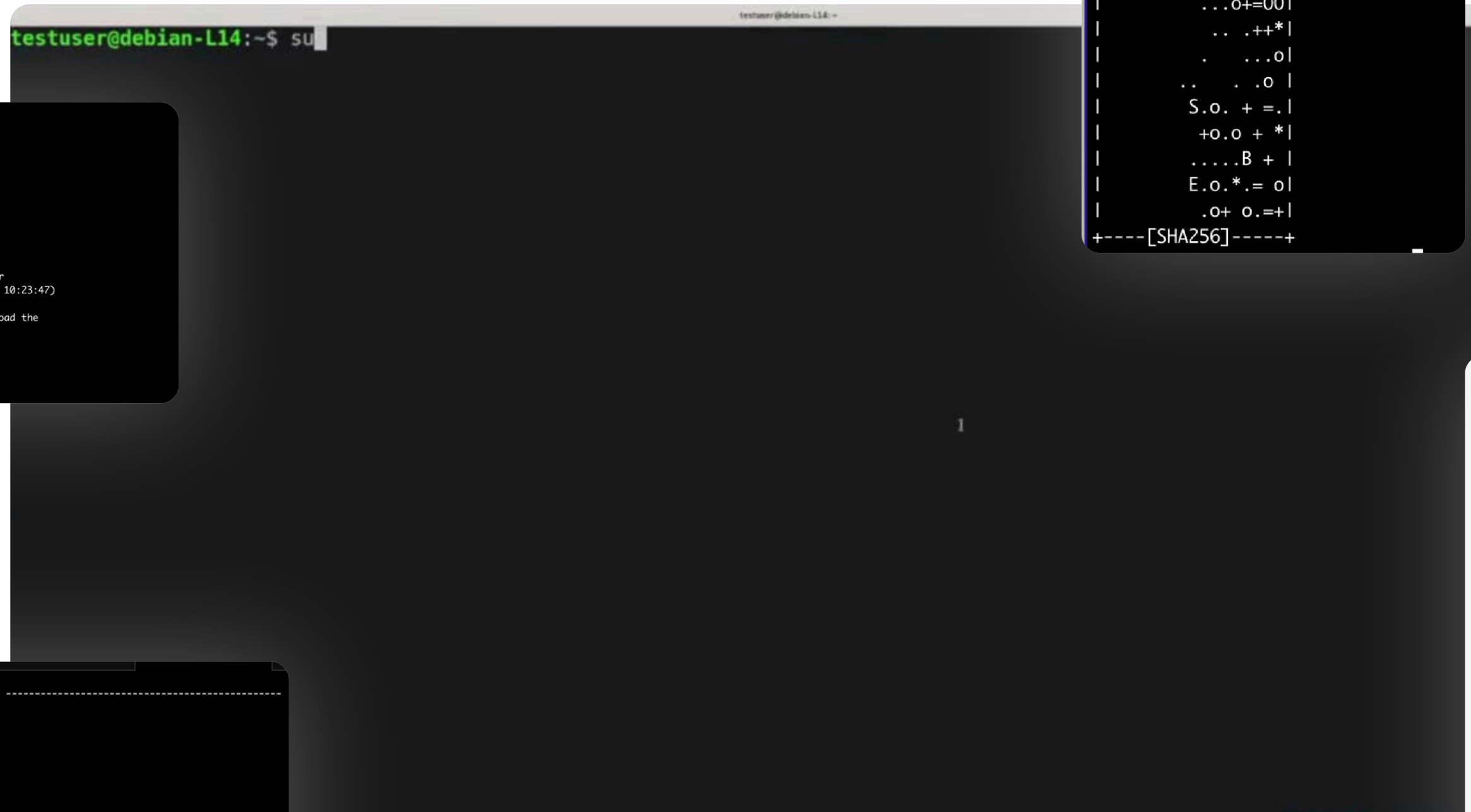
ai-factory@it4i.cz

TRÉNOVÁNÍ MODELU



KLASICKÝ PŘÍSTUP K HPC

```
Your public key has been saved in  
The key fingerprint is:  
SHA256:fFp7L5cfGT0jmWRJ1TvVwHRJCzUR  
The key's randomart image is:  
+---[RSA 3072]----+  
|   .o+=001  
|   .. .+*|  
|   . . .o|  
|   .. .o |  
|   S.o. + =.|  
|   +o.o + *|  
|   ....B + |  
|   E.o.*.= o|  
|   .o+ o.=+|  
+-----[SHA256]-----+
```



```
...running on Rocky 8.X
```

Public Service Announcement: Apptainer on the Karolina cluster
Posted: (2024-05-10 10:23:47)

Apptainer is now a part of the operating system, you do not need to load the module.

```
$ apptainer --version  
apptainer version 1.3.1-1.el8
```

```
----- /apps/modules/bio -----  
GROMACS/4.5.5-gompi-2020c-ORCA-5.0.1  
GROMACS/2021.4-FossCUDA-2020b-PLUMED-2.7.3  
GROMACS/2022-FossCUDA-2020b  
GROMACS/2023-foss-2022a-CUDA-12.2.0-PLUMED-2.9.0-mpi  
GROMACS/2023-foss-2022a-CUDA-12.2.0-PLUMED-2.9.0  
GROMACS/2023.4-gompi-2020c-ORCA-5.0.1-double-precision  
GROMACS/2024.1-foss-2023b (D)
```

```
----- /apps/modules/cae -----  
OpenFOAM/v2312-foss-2023a OpenFOAM/9-intel-2021a (D) STAR-CCM+/18.06.006-r8
```

```
----- /apps/modules/chem -----  
ABINIT/9.10.3-intel-2022a  
ASE/3.22.1-gfbf-2023b  
Amber/22.0-foss-2022a-AmberTools-23.3-CUDA-11.7.0  
BEEF/0.1.1-intel-2020a  
CP2K/9.1-foss-2022a  
Critic2/1.1stable-intel-2020b  
DFT-D4/3.6.0-intel-2022b-Python-3.10.8  
DFT-D4/3.6.0-intel-2022b-Python-3.11.5 (D)  
Libint/2.7.2-GCC-11.3.0-lmax-6-cp2k  
Molden/7.3-GCCcore-12.3.0  
Molpro/mpp-2022.2.2.linux_x86_64_mpi  
Molpro/mpp-2022.2.2.linux_x86_64_sockets (D)  
Molpro/mpp-2024.1.0.linux_x86_64_mpi (D)
```

lines 1-27

Accessing the Clusters

Shell Access

All IT4Innovations clusters are accessed by the SSH protocol via login nodes at the address `cluster-name.it4i.cz`. The login nodes may be addressed specifically, by prepending the loginX node name to the address.

You can also see instructional videos on YouTube for additional information:

- [Accessing IT4i Clusters - Windows](#)
- [Accessing IT4i Clusters - Linux](#)

Projects from the EUROHPC workgroup can only access the Karolina cluster.

We accept only RSA or ED25519 keys for logging into our systems.

Karolina Cluster

Login address	Port	Protocol	Login node
karolina.it4i.cz	22	SSH	round-robin DNS record for login[1,2,3,4]
login[1,2,3,4].karolina.it4i.cz	22	SSH	login[1,2,3,4]

Barbora Cluster

Login address	Port	Protocol	Login node
barbora.it4i.cz	22	SSH	round-robin DNS record for login[1,2]
login[1,2].barbora.it4i.cz	22	SSH	login[1,2]

Authentication

Authentication is available by [private key](#) only. Verify SSH fingerprints during the first login:

Karolina

Fingerprints

Fingerprints are identical for all login nodes.

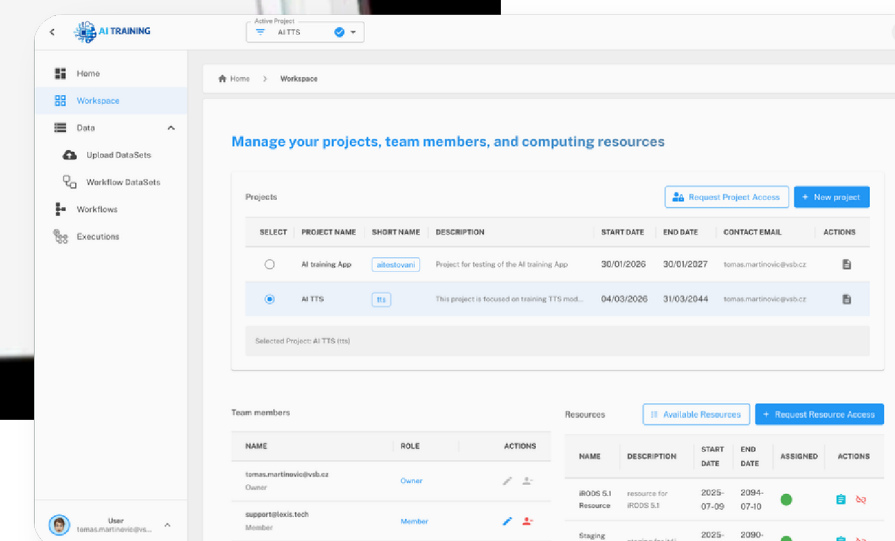
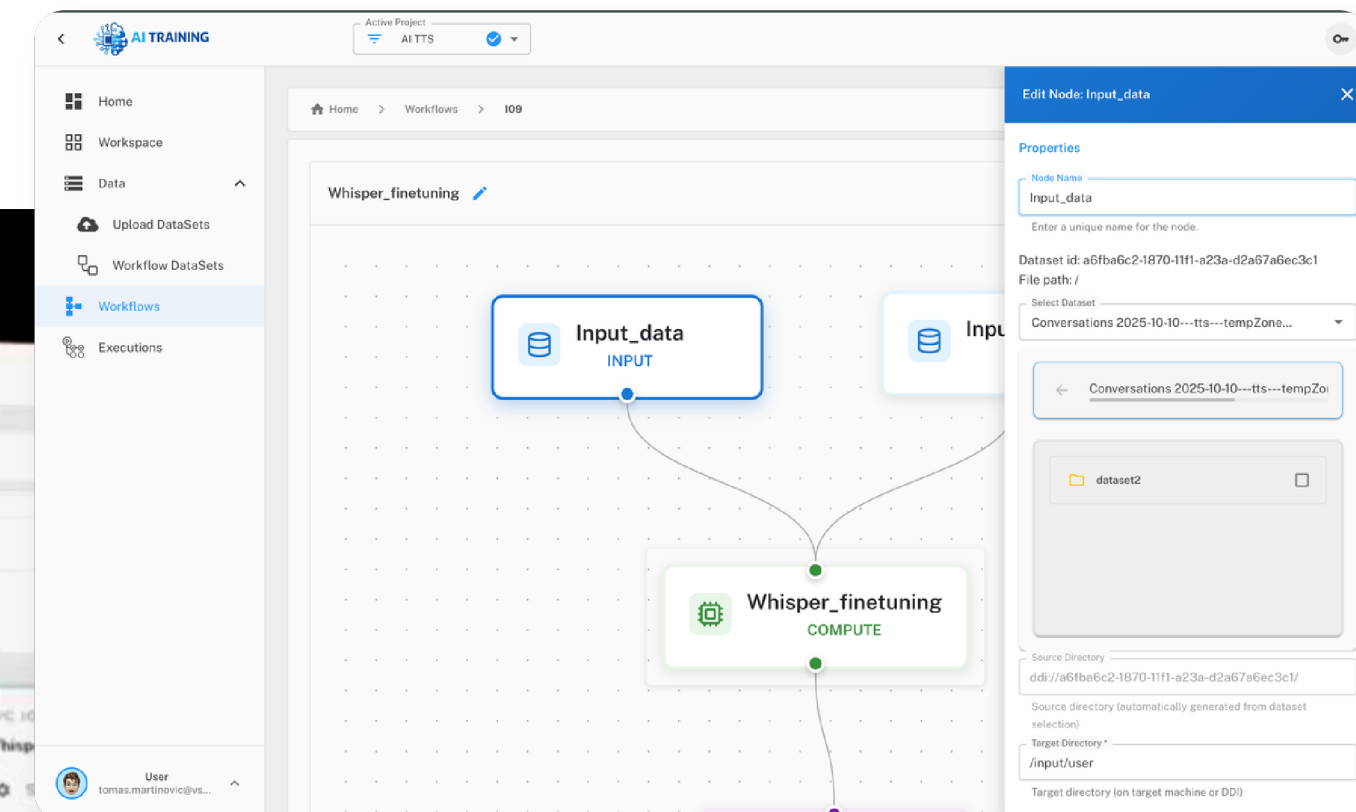
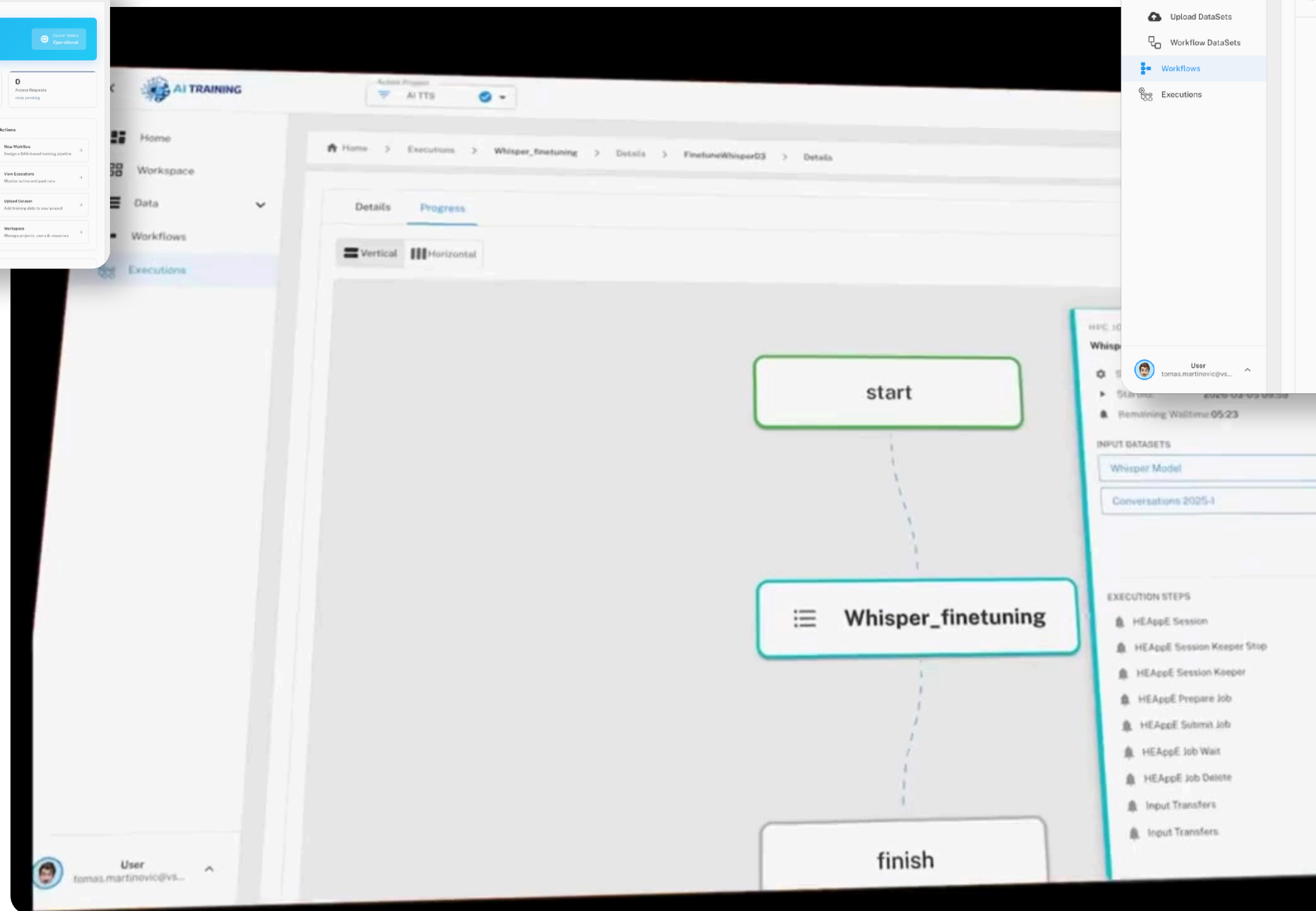
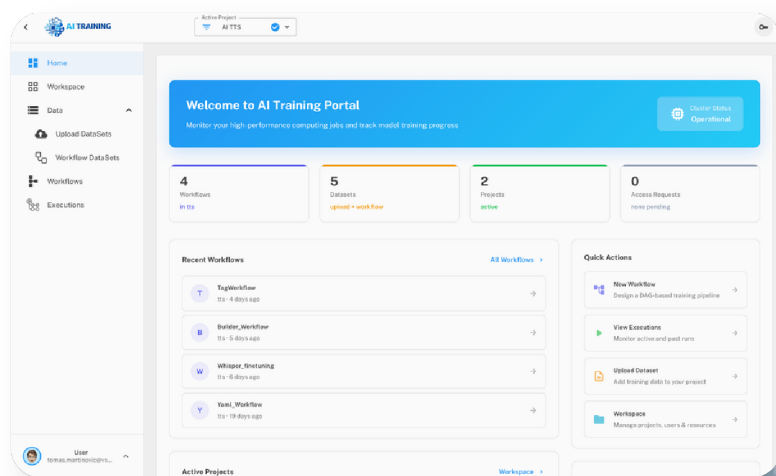
```
# login[1,2,3,4].22 SSH-2.0-OpenSSH_7.4  
2048 f05:41:1a:40:32:db:08:77:51:79:04:af:53:04:57:0b:7c (RSA)  
2048 5f9256:12376:bc:00aef7308a4e0d0c0521f9fa18199080 (RSA)  
256 MD5:f9:16:86:7d:f8:c6:8f:42:14:18:73:02:14:06:7c:22 (ED25519)  
256 SHA256:2f8f9120f8c22290/skocq:AFzq0t1yv780u0fzg (ED25519)
```

Public Keys \ Known Hosts

Public Keys \ Known Hosts are identical for all login nodes.

```
login1,login1.karolina.it4i.cz,login1.karolina.karolina.it4i.cz ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCAjEBAQEA...  
login1,login1.karolina.it4i.cz,login1.karolina.karolina.it4i.cz ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCAjEBAQEA...  
login2,login2.karolina.it4i.cz,login2.karolina.karolina.it4i.cz ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCAjEBAQEA...  
login2,login2.karolina.it4i.cz,login2.karolina.karolina.it4i.cz ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCAjEBAQEA...
```

AI TRAINING APP



UKÁZKA TRÉNOVÁNÍ MODELU

v prostředí
AI TRAINING APP

následuje:

Hands-on zkouška výstupu natrénovaného modelu

HANDS-ON

24:54



Kde se přihlásím?

aitraining.lexis.tech



Kde najdu zadání?

bit.ly/zadanihandson



Potřebuji pomoc!

Zvedněte ruku nebo si odchyťte
postávajícího kolegu/kolegyni

následuje:

Coffee break - 20 minut

COFFEE BREAK

19:46

následuje:

Analýza tréninku pomocí MLFlow a technická prezentace komponent aplikace

COFFEE BREAK

14:53

následuje:

Analýza tréninku pomocí MLFlow a technická prezentace komponent aplikace

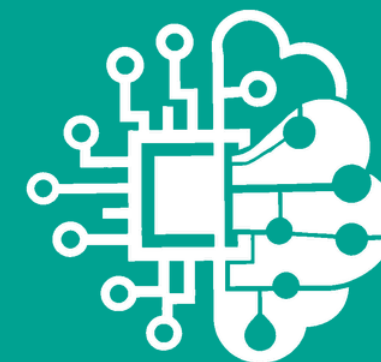
TECHNICKÁ ČÁST

Co nás čeká?

Metriky trénování modelu
a jejich interpretace

mlflow[™]

Detailnější vysvětlení nastavení a chystané funkcionality
v AI Training app



AI TRAINING

CO DÁLE PLÁNUJEME?

INFERENCE



MONITORING



INTEGRACE



WORKFLOWS



Integrace managementu
inференční služby



MLflow integrace



Metriky modelu



Monitorování využití systému



Příprava pro AI agenty



Open Web Search data



Dedikované komponenty



Nové kontejnery



Šablony

Q&A

Rekapitulace odpolední části

1.

Úvodní část



Co dělá IT4I



Co jsou AI Factories a co nabízejí



Jak je možné získat výpočetní čas



Jaké konzultační služby nabízíme

2.

AI training app



Jak se probíhá fine-tuning modelu



Klasický způsob vs. AI app



Jak trénovat model pomocí AI app



Testování natrénovaného modelu

3.

Technická část



Metriky tréninku a jejich interpretace



Nastavení parametrů v AI app



Chystané funkcionality AI app



*Jak dlouho máme přístup k tomu,
co jsme dnes viděli/zkusili?*



DĚKUJEME!

Zaujala Vás možnost přístupu k superpočítačům?
Zaujala Vás aplikace pro trénování modelů
na superpočítačích?

Pojďme si o tom popovídat nebo nás kontaktujte
na ai-factory@it4i.cz



EuroHPC
Joint Undertaking



Spolufinancováno
Evropskou unií



Ministerstvo
školství, mládeže
a tělovýchovy