



Umělá inteligence na KKY ZČU v Plzni a (nejen) v MetaCentru

Jan Švec

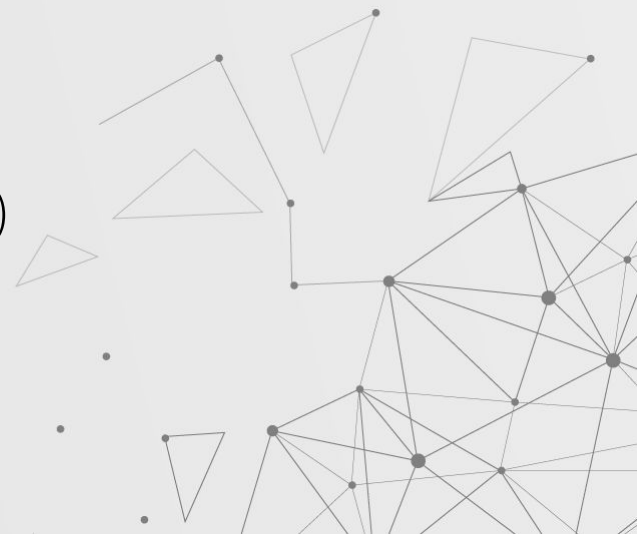
[<honzas@kky.zcu.cz>](mailto:honzas@kky.zcu.cz)

Katedra kybernetiky, ZČU v Plzni

NTIS Fakulta aplikovaných věd, ZČU v Plzni

Umělá inteligence na KKY ZČU

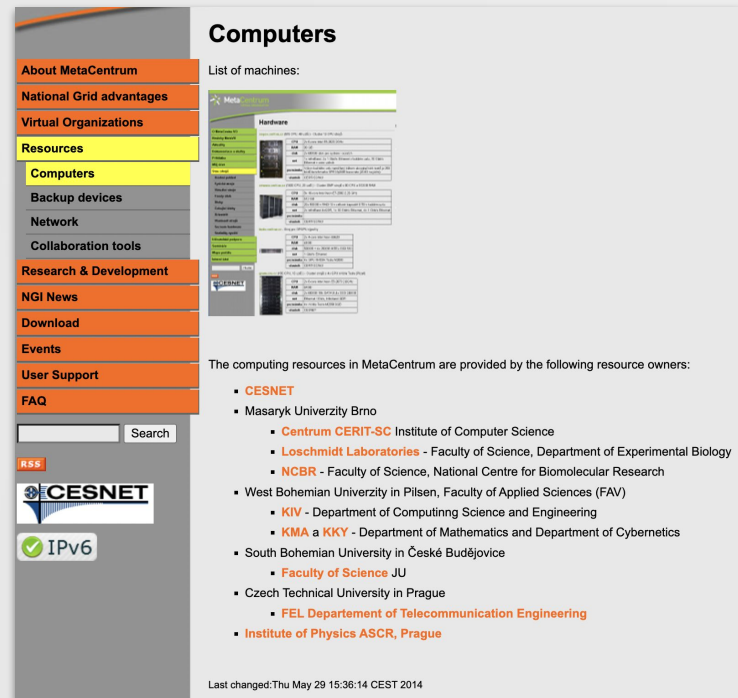
- **Řečové technologie**
 - Rozpoznávání a porozumění řeči a textu
 - Syntéza řeči
 - Aplikace - on-line a off-line systémy
- **Počítačové vidění**
 - Identifikace tváří, pózy ruky, znaková řeč
 - Multimodální systémy (obraz & řeč, obraz & text)
- **Life sciences**
 - Biokybernetika
 - In silico & in vitro



Umělá inteligence KKY nejen v MetaCentru

Historický exkurz (~2004 až 2010)

- Výpočty pro řečové technologie
 - GMM (Gaussian Mixture Models)
 - NVIDIA CUDA
- Cluster Konos s GPU kartami
 - GTX 465 (architektura Fermi)
 - 855 GFLOPS
 - (srov. A100: 312 TFLOPS v FP16)
 - Diskové pole data-kky
 - MetaCentrum fronta iti

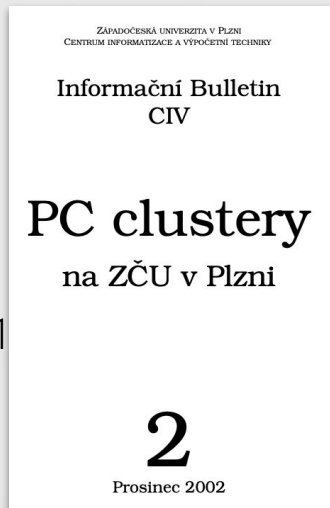


The screenshot shows the MetaCentrum website interface. On the left is a navigation menu with categories like 'About MetaCentrum', 'National Grid advantages', 'Virtual Organizations', 'Resources', 'Computers', 'Backup devices', 'Network', 'Collaboration tools', 'Research & Development', 'NGI News', 'Download', 'Events', 'User Support', and 'FAQ'. The 'Computers' section is highlighted in yellow. Below the menu is a search bar, an RSS feed icon, the CESNET logo, and an IPv6 status indicator. The main content area is titled 'Computers' and includes a 'List of machines:' section with a table of hardware specifications. Below this, it lists resource owners: CESNET, Masaryk University Brno (with sub-entities like CERIT-SC, Loschmidt Laboratories, NCBR), West Bohemian University in Pilsen (with KIV and KMA), South Bohemian University in České Budějovice (with Faculty of Science JU), Czech Technical University in Prague (with FEL Department of Telecommunication Engineering), and Institute of Physics ASCR, Prague. The page footer indicates it was last changed on Thu May 29 15:36:14 CEST 2014.

Umělá inteligence KKY nejen v Meta

Historický exkurz (~2004 až 2010)

- Výpočty pro řečové technologie
 - GMM (Gaussian Mixture Models)
 - NVIDIA CUDA
- Cluster Konos s GPU kartami
 - GTX 465 (architektura Fermi)
 - 855 GFLOPS
 - (srov. A100: 312 TFLOPS v FP16)
 - Diskové pole data-kky
 - MetaCentrum fronta ití



jan Hoidekr
21:16 Předevčirem



KKY začala s META určitě dříve



Miloš Mulač
10:04 Včera

"Začal jsem s tím v roce 2002 byl pak na ZČU
U uveden do provozu nový cluster,
jenž byl porovnán z prostředků"

[Zobrazit více](#)



jan Hoidekr
10:22 Včera

myslím, 2002 to byla KMA, KKY se
přidala tak od 2004, to jsem s Honzou
Trmalem nejdříve chtěli cluster stavět a

[Zobrazit více](#)



jan Hoidekr
10:23 Včera

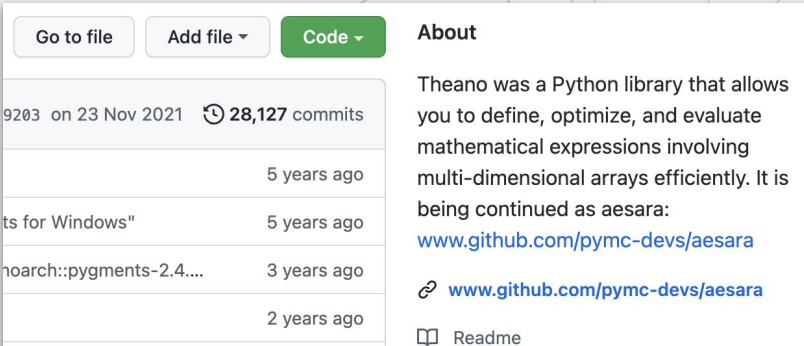


toto byl druhý konos, to by asi
odpovídalo, že ten první jsme
odstavovali po cca 5 letech

Umělá inteligence KKY nejen v MetaCentru

Jak šel čas v průběhu desetiletí

- Používání větší palety nástrojů
 - Python, ML balíky scipy, numpy, scikit-learn
- MetaCentrum SW moduly
 - balení software, udržitelné dokud...
- ... vše v oblasti ML převrátily naruby neuronové sítě
 - komoditizace nástrojů a modelů
 - **Theano** pro autograd, **Caffe2** pro CV
 - již ne 3-vrstvé multi-layer perceptron sítě
 - konvoluční architektury
 - rekurentní architektury
 - zásadní boom především v oblasti CV



The screenshot shows a GitHub repository interface. At the top, there are buttons for 'Go to file', 'Add file', and 'Code'. Below these is a table of commit history:

Commit message	Time	Commits
9203 on 23 Nov 2021	🕒 28,127 commits	
	5 years ago	
ts for Windows"	5 years ago	
hoarch::pygments-2.4....	3 years ago	
	2 years ago	

On the right side, there is an 'About' section with the following text:

Theano was a Python library that allows you to define, optimize, and evaluate mathematical expressions involving multi-dimensional arrays efficiently. It is being continued as aesara: www.github.com/pymc-devs/aesara

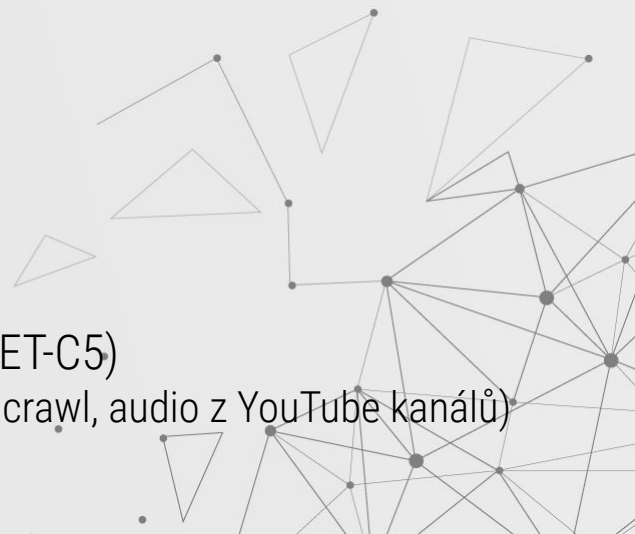
🔗 www.github.com/pymc-devs/aesara

📖 Readme

Umělá inteligence KKY nejen v MetaCentru


Jak šel čas v průběhu desetiletí

- **Neuronové sítě** - delší, širší, bystrozraký
 - 2015 - Google TensorFlow 1.x framework
 - 2016 - Facebook PyTorch framework (pohltil Caffe2)
 - 2017 - Google Transformer architektura
 - 2018 - Google BERT model
 - 2019 - Google TensorFlow 2.x (pohltil Keras)
 - 2019 - HuggingFace Transformers knihovna
 - 2020 - Google T5 model
 - 2020 - Facebook Wav2Vec 2.0 model
 - 2020 - OpenAI GPT-3 model
 - 2021 - OpenAI DALL-E model
- **Transfer-learning** (BERT, GPT-2, T5, Wav2Vec 2.0, FERRET-C5)
 - **pre-training** - neoznačená data (web text / obrázky z web crawl, audio z YouTube kanálů)
 - **fine-tuning** - natrénování modelu z označených dat



Umělá inteligence KKY nejen v MetaCentru

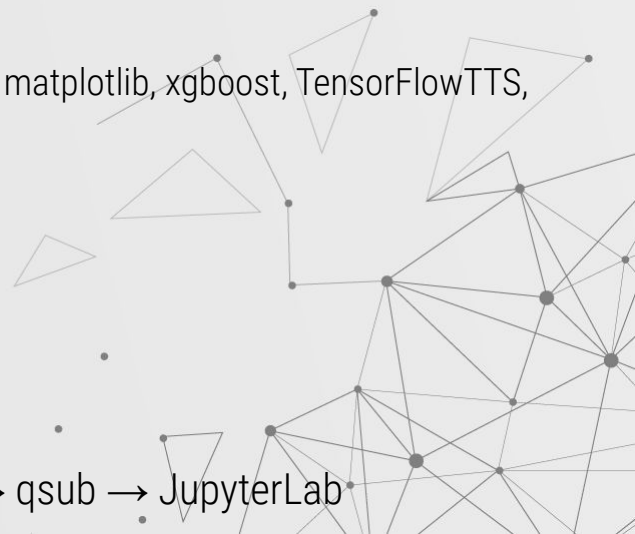
Jak šel čas v průběhu desetiletí

- Naše motto: **Od výzkumu k praxi**
 - Aplikace pokročilých modelů v praxi / technologických demonstrátorech
 - Titulování televizních přenosů ČT
 - Audiovizuální archiv MALACH
 - Audiovizuální archiv ÚSTR
 - Syntéza řeči a konzervace hlasu pro pacienty s laryngektomií
 - Výzkum, vývoj a aplikace technologií mezi obory
 - **Life sciences**
 - Spolupráce Biokybernetická laboratoř KKY (Daniel Georgiev) a ÚOCHB (Jiří Vondrášek)
 - Neuronové sítě pro analýzu / modifikaci DNA / proteinových sekvencí
 - **Energetika**
 - Spolupráce s týmem LAPS (Laboratoř pokročilých energetických systémů) (Martin Střelec)
 - ML pro modelování a predikce v přenosových soustavách elektrické energie
- 

Umělá inteligence nejen v MetaCentru

Současnost

- Masivní využívání **singularity**
 - Předpřipravený image ~12GB
 - nvc.io/nvidia/tensorflow:22.01-tf2-py3
 - JupyterLab + qsub skript + úpravy Python env
 - Množství Python knihoven
 - pandas, scikit-learn, bio-python, tensorflow, pytorch, matplotlib, xgboost, TensorFlowTTS, librosa, transformers, CoquiTTS
 - Další balíky a frameworky
 - ffmpeg, sox, Kaldi, MC, Vim
 - Papermill
 - Prakticky náhrada interaktivních jobů
 - Snadný on-boarding nových členů týmu (studenti)
 - MetaCentrum konto → Perun → účet na úložišti → qsub → JupyterLab



Umělá inteligence nejen v MetaCentru

Současnost

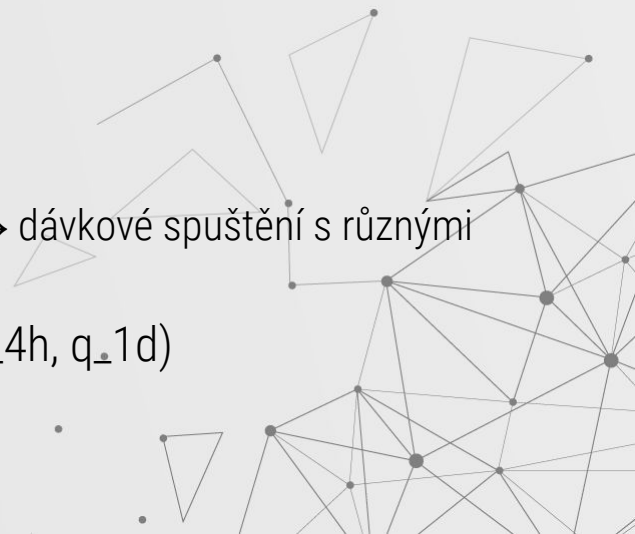
- Diskové úložiště /storage/plzen4-ntis
 - 400TB úložiště pořízené KKY
 - správa uživatelů přes Perun
 - správa projektů (sdílených adresářů) Ansible AWX, YAML konfigurace
 - správa systému - support MetaCentra (díky!)
 - zálohování - disková úložiště Cesnet
 - přístup k datům
 - NFS4 - uzly v MetaCentru
 - SMB - vnitřní síť KKY
 - sshfs
 - využíváno i ke sdílení dat s jinými pracovišti



Umělá inteligence nejen v MetaCentru

Současnost

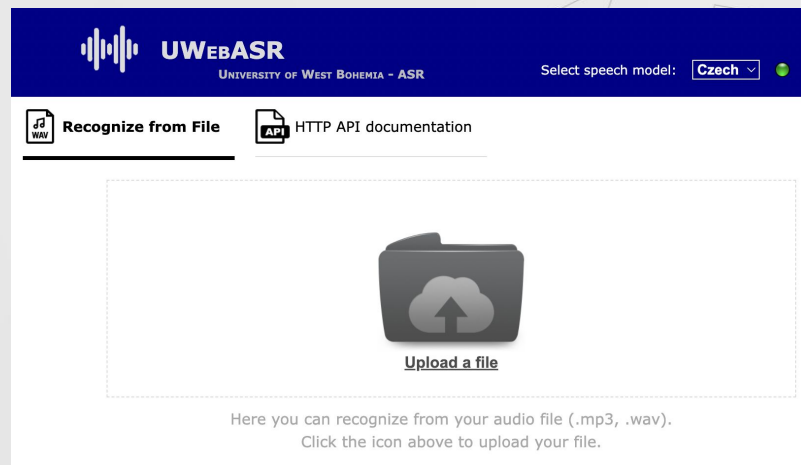
- Výpočetní cluster konos
 - pořízené KKY
 - 32x NVIDIA GeForce GTX 1080 Ti (pořízeno 2018)
 - používáme často jako JupyterLab terminál
 - namísto osobních výpočetní stanice s GPU
 - interaktivní vývoj úloh využívajících neuronové sítě
 - malá data, malá GPU MEM, malý výkon
 - odladění .ipynb kódu → papermill parametrizace → dávkové spuštění s různými parametry
 - dostupné z MetaCentra frontami iti, gpu (a q_2h, q_4h, q_1d)



Umělá inteligence nejen v MetaCentru

Současnost

- **Kubernetes cluster kubus**
 - pořízené KKY
 - 6 uzlů
 - prostředí pro nasazování technologických demonstrátorů a služeb KKY
 - Clarin UWebASR
 - SpeechCloud
 - Audiovizuální archivy
 - propojení s DÚ Cesnet
 - jsme na začátku jejich používání
 - zprovozněno přelom 2021/2022



Umělá inteligence nejen v MetaCentru

Současnost

- Ostatní HPC infrastruktury
 - CERIT-SC a cluster zia
 - pretraining BERT-like a Wav2Vec 2.0 modelů
 - dokážeme saturovat 4xA100 (přešli jsme sem od Google Cloud TPU)
 - Český BERT model:
<https://huggingface.co/fav-kky/FERNET-C5>
 - IT4Innovations
 - mimořádná výzva: GPU TESTING AND BENCHMARKING
 - používáme od začátku 2022
 - opět pretraining modelů (dokážeme saturovat i 8xA100)
 - výhodné využití singularity, shodný image použitelný v MetaCentru
 - DÚ CESNET
 - v serverovně hostujeme CEPH úložiště CL2
 - CESNET FileSender, zálohování (úložiště, další stroje z naší infrastruktury)



Umělá inteligence nejen v MetaCentru

Díky všem (in random order):

Miloš Mulač, meta@cesnet.cz, Jan Hoidekr, Jiří Sitera, Pavel Vondruška, týmy KKY & CIV

Otázky a odpovědi
(následují ukázky)



Umělá inteligence na KKY ZČU

Ukázky

- Audiovizuální archiv MALACH (<https://malach.kky.zcu.cz/>)
- Audiovizuální archiv ÚSTR (<https://naki-ustr.zcu.cz/>)
- Archiv KGB ÚSTR (<https://archivkgb.zcu.cz/>)

