

# Uživatel:CeSt/AI

## Vývoj AI

- FROM Pre-trained Word Embeddings TO Pre-trained Language Models — Focus on BERT (<https://towardsdatascience.com/from-pre-trained-word-embeddings-to-pre-trained-language-models-focus-on-bert-343815627598>)
- Jan Romportl: Umělá inteligence a co bude dál? ([https://www.youtube.com/watch?v=nMAIr\\_A9i58&t=6s](https://www.youtube.com/watch?v=nMAIr_A9i58&t=6s))
- Jan Romportl: AI, co bude dál - EDO2022 ([https://www.youtube.com/watch?v=Bz7R72q\\_pug](https://www.youtube.com/watch?v=Bz7R72q_pug))
  - Online crowdsourcingový předpovědní systém (<https://www.metaculus.com/questions/8898/russian-invasion-of-ukraine-before-2023/>)
  - Rizika z naučené optimalizace (Mesa optimizer) (<https://www.lesswrong.com/posts/FkgsxrGf3QxhflWHG/risks-from-learned-optimization-introduction>)
  - Socratic Models, Composing Zero-Shot Multimodal Reasoning with Language (<https://socraticmodels.github.io/>)
  - Kdy bude spuštěna slabá univerzální AI? (<https://www.metaculus.com/questions/5121/date-of-artificial-general-intelligence/>)
  - IBM Watson ([https://en.wikipedia.org/wiki/IBM\\_Watson](https://en.wikipedia.org/wiki/IBM_Watson))
- [https://en.wikipedia.org/wiki/Goodhart%27s\\_law](https://en.wikipedia.org/wiki/Goodhart%27s_law)
- Goodhartův zákon ([https://cs.wikipedia.org/wiki/Goodhartovo\\_pravidlo](https://cs.wikipedia.org/wiki/Goodhartovo_pravidlo))
- Mesa-Optimizace (<https://www.lesswrong.com/tag/mesa-optimization>)
- <https://www.csoonline.com/article/2118813/how-to-spot-a-liar--identifying-deceptive-behavior.html>
- <https://socraticmodels.github.io/>
- O ChatGPT a velkých jazykových modelech (<https://jlibovicky.github.io/2023/02/07/Otazky-a-odpovedi-o-ChatGPT-a-jazykovych-modelech.html>)
- How To Build Your Own Custom ChatGPT With Custom Knowledge Base (<https://betterprogramming.pub/how-to-build-your-own-custom-chatgpt-with-custom-knowledge-base-4e61ad82427e>)
- 7 největších otázek o ChatGPT (<https://www.gartner.com/en/articles/your-7-biggest-chatgpt-questions-answered>)
- <http://playground.tensorflow.org/>
- Kodex použití AI v redakci Seznam (<https://www.seznamzpravy.cz/clanek/redakce-seznam-zpravy-pracuji-s-ai-a-maji-jasna-pravidla-229210>)
- Otázky a odpovědi o ChatGPT a velkých jazykových modelech (<https://jlibovicky.github.io/2023/02/07/Otazky-a-odpovedi-o-ChatGPT-a-jazykovych-modelech.html>)
- <https://ufal.mff.cuni.cz/AlvK/>
- <https://www.lupa.cz/clanky/jak-z-generativnich-ai-dostat-to-nejlepsi-persony-funkce-v-excelu-i-prompty-pro-obrazkove-ai/>
- Redukce dimenzionality: MODERNÍ METODY VÝBĚRU PŘÍZNAKŮ VE STATISTICKÉM ROZPOZNÁVÁNÍ

## AI nástroje

- AI v nástrojích pro pedagogy (<https://www.youtube.com/watch?v=8iOOiHL7F8>)
- Scribe (generátor průvodců - free)

## AI textové nástroje

- <https://chat.openai.com/>
- <https://you.com/>
- <https://chat.forefront.ai> (bezplatná verze GPT4)
- <https://vsechnoai.cz/>
- <https://www.perplexity.ai/>
- Grammarly (gramatika - free)
- Jasper (pisálek free)

## AI grafické nástroje

- <https://designer.microsoft.com/>
- <https://www.bing.com/images/create/>
- <https://zapier.com/blog/best-ai-image-generator/>

## AI kódování

- CodeWhisperer (generování kódu - free)
- Copilot (kódování placené)
- Blackbox AI (odpověď s kódem)
- Cogram (SQL placené)
- <https://codehorizons.com/Seminars/using-large-language-transformer-models-for-research-in-r/>

## AI ve zdravotnictví

- Mohou nástroje umělé inteligence fungovat v medicíně transparentně? ([https://journals.lww.com/academicmedicine/Fulltext/2022/10000/Can\\_Artificial\\_Intelligence\\_Tools\\_Operate.7.aspx](https://journals.lww.com/academicmedicine/Fulltext/2022/10000/Can_Artificial_Intelligence_Tools_Operate.7.aspx))
- <https://www.info.cz/wall-street-journal/kdo-rozhodne-o-zivote-a-smrti-pacienta-clovek-nebo-stroj-nastup-umele-inteligence-prinasi-slozite-otazky>
- Odpovědná cesta ke generativní umělé inteligenci ve zdravotnictví (<https://cloud.google.com/blog/topics/healthcare-life-sciences/sharing-google-med-palm-2-medical-large-language-model>)
- U lékařských zkoušek USMLE (něco jako atestace) má Med-PaLM2 85% úspěšnost (<https://sites.research.google/med-palm/>)

Hodnota umělé inteligence a strojového učení ve zdravotnictví je nezměrná, protože tyto špičkové technologie mají potenciál způsobit revoluci v péči o pacienty, zefektivnit operace a zlepšit rozhodovací procesy. Systémy AI a strojového učení poskytují přesné, včasné a na datech založené přehledy a umožňují zdravotnickým pracovníkům činit informovanější rozhodnutí, což vede k lepším výsledkům pacientů a personalizované péči.

- <https://www.forbes.com/sites/davidchou/2023/03/21/how-tech-leaders-compete-in-the-battle-of-healthcare-ai/?sh=d0440ef6c4cd>
- Med-PaLM, A large language model from Google Research, designed for the medical Domain (<https://sites.research.google/med-palm/>)
- <https://www.medrxiv.org/search/chatgpt>
- <https://www.insider.com/chatgpt-passes-medical-exam-diagnoses-rare-condition-2023-4>
- ChatGPT in Healthcare: A Taxonomy and Systematic Review, 2023 (<https://www.medrxiv.org/content/10.1101/2023.03.30.23287899v1.full.pdf>)
- Study Finds ChatGPT Outperforms Physicians in High-Quality, Empathetic Answers to Patient Questions (<https://today.ucsd.edu/story/study-finds-chatgpt-outperforms-physicians-in-high-quality-empathetic-answers-to-patient-questions>)
  - Studie zjistila, že ChatGPT překonává lékaře v poskytování vysoce kvalitních, empatických odpovědí na písemné otázky pacientů v diskuzním vlákně r/AskDocs. Panel hodnotitelů, kteří se rekrutovali z licencovaných zdravotnických odborníků, preferoval odpovědi ChatGPT v 79 % případů a hodnotil je lépe z hlediska kvality a empatie, než odpovědi lékařů.

## Možnosti HW řešení

- <https://medium.com/@jaroslav.streit/gpt-pro-v%C5%A1echny-gpt4all-e5f6c290607a>
- Počítač na umělou inteligenci (<https://medium.com/@jaroslav.streit/m%C5%AFj-po%C4%8D%C3%ADta%C4%B8-na-um%C4%9Blou-inteligenci-cbd07af45367>)

Odešli z Azure a postavili si server (jsme podporováni NVidia, tak máme lepší ceny) a kartou Rtx 6000 Ada. Většinou stačí mít dostatek vláken procesoru a GPU s dostatkem VRAM. Ten use-case je taky pro každého jiný a někdo třeba dá přednost RAM s ECC atd. Jde o to, že serverové komponenty jsou na dlouhodobou zátěž dělané a mají třeba lepší efektivitu. Já tady dával 192GB hlavně kvůli tomu, že mašiny k pronájmu s touto pamětí jsou dražší a už se to např. na GPT-neoX s větším datasetem vyplatí. Další věc je, že GPU na AI mají třeba více Tensor jader a větší bandwidth pamětí.

## AI v cloudu

- <https://vast.ai/>
- Microsoft Azure se hodí pro občasné tréninky. Jeden čas tam bylo možné zarezervovat výkon na low priority VM - tj. Byly k dispozici, pokud je nikdo nepotřeboval a byly 5x levnější. Teď v v Europa West (Amsterdam) už mají potíže s kapacitou (elektrina, prostor) tak už tam nejsou - možná v datacentru v Irsku, by se ještě mohly vyskytovat.
- Google Colab - podobné

## FastTEXT

- <https://ai.facebook.com/tools/fasttext/>

## AI a wiki

Vývoj v oblasti AI mění způsob, jakým tvoříme video obsah. Díky Zeroscope, Automatic 1111 a Stable Diffusion, máme nyní možnost vytvářet už poměrně kvalitní text2video tvorbu s vlastním hardwarem, což je skutečný krok vpřed v demokratizaci této technologie. Současně se tím otevírá možnost, jak v prostoru jako je wiki pracovat s videem. Díky AI se otevírá možnost, že předmětem editace nebude samotné video, ale jeho textové zadání!