

The Increased Scientific Activity in Artificial Intelligence During the Pandemic

Prof. Guntis Bārzdiņš University of Latvia, IMCS Al-Lab

The 17th Baltic Conference on Intellectual Co-operation: Mathematics for Society 28 June 2021, Tallinn, Estonia

Institute of Mathematics and Computer Science University of Latvia



~80% Complied with COVID-19 Lockdowns



Data source: University of Sydney. Chart: ABC News Story Lab

SAAS Enabled Relocation Cities → CountrySide



Data source: Latvian Mobile Telephone (LMT)

A new class of high-skill service workers, the **digital nomads**, is emerging

Outline

There are two "pandemies" ongoing

- **COVID-19:** efficient mRNA vaccines available, will be forgotten soon
- SAAS (Software As A Service): will genuinely change the course of history

COVID-19 pandemy accelerated the SAAS adoption

- o 80% of population can work/study from home (using SAAS), and economy survives
- Online shopping and delivery services cover basic needs (depend on SAAS)
- International travel and physical meetings can be avoided (thanks to SAAS)
- Consequences: SAAS providers accumulated BIG DATA on all activities of 80% of population

BIG DATA drives the progress of Artificial Inteligence

- Google, FaceBook, Apple, Amazon, Tesla, Microsoft, Twitter all are SAAS companies (†52% in 2020)
- These companies realized they can grow faster by sharing AI research results (fortunately)
- Shared AI research enables cloud-in-the-box alternative to SAAS: PiniTree.com from Latvia

Transformers & Knowledge

GPT-3 28 May, 2020 (Op 175B parameters 0.5T words training

28 May, 2020 (OpenAI) 175B parameters 0.5T words training text 10,000 GPUs + 300,000 CPUs

SYSTEM PROMPT (HUMAN-WRITTEN)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains, Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL COMPLETION MACHINE-WRITTEN, 10 TRIES

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them - they were so close they could touch their horns.

While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, "We can see, for example, that they have a common 'language,' something like a dialect or dialectic."



GPT-3 Scaling Laws

Data

Test error diminishes by increasing any of these

Richard Sutton "bitter lesson": GPT-3 has no inductive bias, scale wins over biases. Only small data required biases.

https://youtu.be/__Kj9HeU5tE https://arxiv.org/abs/2105.12806 https://ml.berkeley.edu/blog/posts/dalle2/ https://parl.ai/projects/params_vs_compute/

Model size

(overparameterisation)

28 October, 2020 (OpenAI) \$1Bn Tesla Microsoft

 $E=mc^2$

Henighan T., et.al., Scaling Laws for Autoregressive Generative Modeling . 2020, https://arxiv.org/pdf/2010.14701.pdf

Compute

GPT-3: Translation

Training Dataset

I	am	а	student	<to-fr></to-fr>	je	suis	étudiant
let	them	eat	cake	<to-fr></to-fr>	Qu'ils	mangent	de
goo	d morning	<to-fr></to-fr>	Bonjour				



Eduards Mukans, PhD student: Trained only on "Farewell to Arms" (300KB, 7490 parallel sentences) achieves BLEU 12.5 for LV-EN translation <u>https://github.com/emukans/minGPT</u>

GPT-3: Text-to-Image Translation (DALL-E)

5 January, 2021 (OpenAI) 12B parameters 250M text-image pairs

TEXT PROMPT

an armchair in the shape of an avocado [...]

AI-GENERATED IMAGES



https://github.com/NomadicDaggy/DALLE-pytorch

KnowledgeGraph Editor with TextComprehesion



Human-in-the-loop AI suggestions for: (a) linking to the KnowledgeGraph (b) enriching the KnowledgeGraph (c) inverse URL text comprehension

Guntis Barzdins, Didzis Gosko, Karlis Cernas, et.al.: Pini Language and PiniTree Ontology Editor: Annotation and Verbalisation for Atomised Journalism <u>https://pinitree.com</u>

Open Problems in AI

Transformers (language models)

DALL

The missing links of AI to achieve the CriticalMass A

Knowledge Graphs (schemaless database)

Concept Creation

> (discrete & compositional)

AI breakthroughs nowadays almost exclusively happen in industry and require vast resources not available in the academic environment

Google FaceBook Twitter etc.

Data → from users

SAAS companies **← Money** targeted advertising

Google FaceBook Twitter etc.

from users

Data

companies

SAAS

Market cap of big five companies from Sept 2019 to Nov 2020 (in billion dollars)



Fortune 500 companies represent 2/3 of U.S. GDP and have **\$22.6trn** in market cap

Money

targeted advertising

Top 5 SAAS companies have **\$7.8trn** in market cap

Source: Yahoo Finance

 \leftarrow

Google FaceBook Twitter etc.

Data from users

 \rightarrow

SAAS companies

← Money targeted advertising

How AI Powers Self-Driving Tesla with Elon Musk and Andrej Karpathy.mp4



Misuse of SAAS

Google FaceBook Twitter etc. SAAS Data \leftarrow Money \rightarrow targeted advertising from users companies FLoC [Flock - a group of sheep, goats, or birds]

Misuse of SAAS

Google FaceBook Twitter etc.

Data→SAAS←Moneyfrom usersMakecompaniestargeted advertisingCre"SAAS Pandemy" and Regulation:

GDPR

cre

- Personal Data protection
- The E.U. AI Act
 - A Risk-Based Policy Approach to AI Application
- The E.U. Digital Markets Act & Digital Sevices Act
 - Gatekeepers (large online platforms) Regulation
 - Digital Europe Programme (9.2B€)
 - COVID-19 pandemic highlighted ... for Europe not to be dependent on systems and solutions coming from other regions of the world
 - US&EU Trade and Technology Council
 - Focus on cooperation and cybersecurity of platforms



tics,

Three AI / SAAS uses



Cloud-in-the-box alternative to global SAAS *) Decentralised **) Good-enough **) Eternal archiving

Login Download Demo Pini Video Tree .com **Intelligent Notes**

Join the Club now!

For \$7/mo you get a professional information gathering, framing and spreading tool











PiniTree Ecosystem



LETA.lv

News enriched with EnterpriseRegister data



Aslimnica.lv

Medical speech transcription and editing <u>selma-project.eu</u> Video subtitling, translation, synchronous voice-over



<u>Tezaurs.lv</u>

Documents enriched with Dictionary data



Station.ee

News translation and clustering in 100 languages

DeepLearning lifecycle management

.UMII.Iv

Cloud-in-the-box: generic data&state backend + kubernetes AI containers = REST API frontends

Thank You

for Your Attention

• • •