Open-ψ Collective ψ

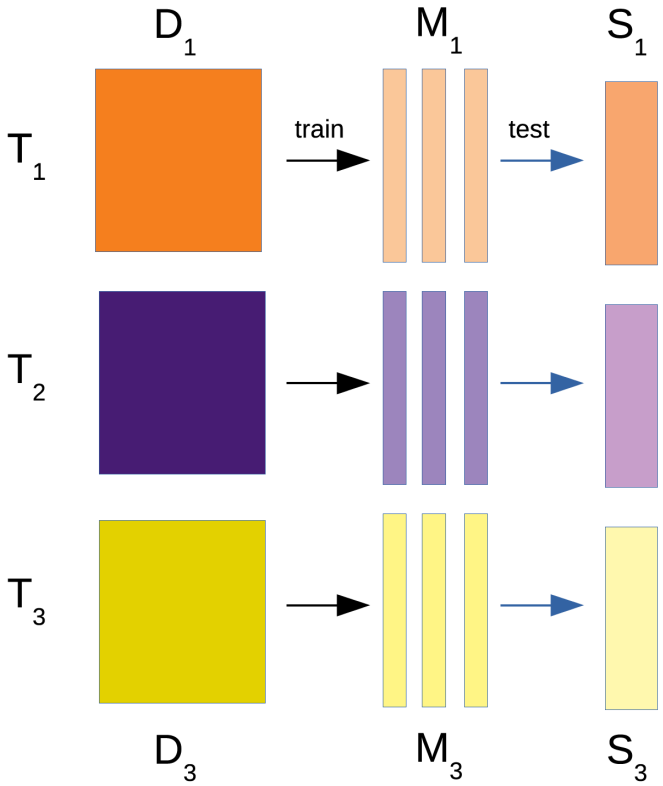
Open foundation models: reproducible science of transferable learning

Jülich Supercomputing Center (JSC)
Scalable Learning & Multi-Purpose AI Lab (SLAMPAI)
Large-scale Artificial Intelligence Open Network (LAION)
European Laboratory for Learning and Intelligent Systems (ELLIS)

13. June 2024 | Dr. Jenia Jitsev

Foundation models: generic transferable learning

- Machine learning before (< 2012): **poorly transferable across tasks**
- Large amount of **labeled data for each task**, **specialized models (no re-use)**

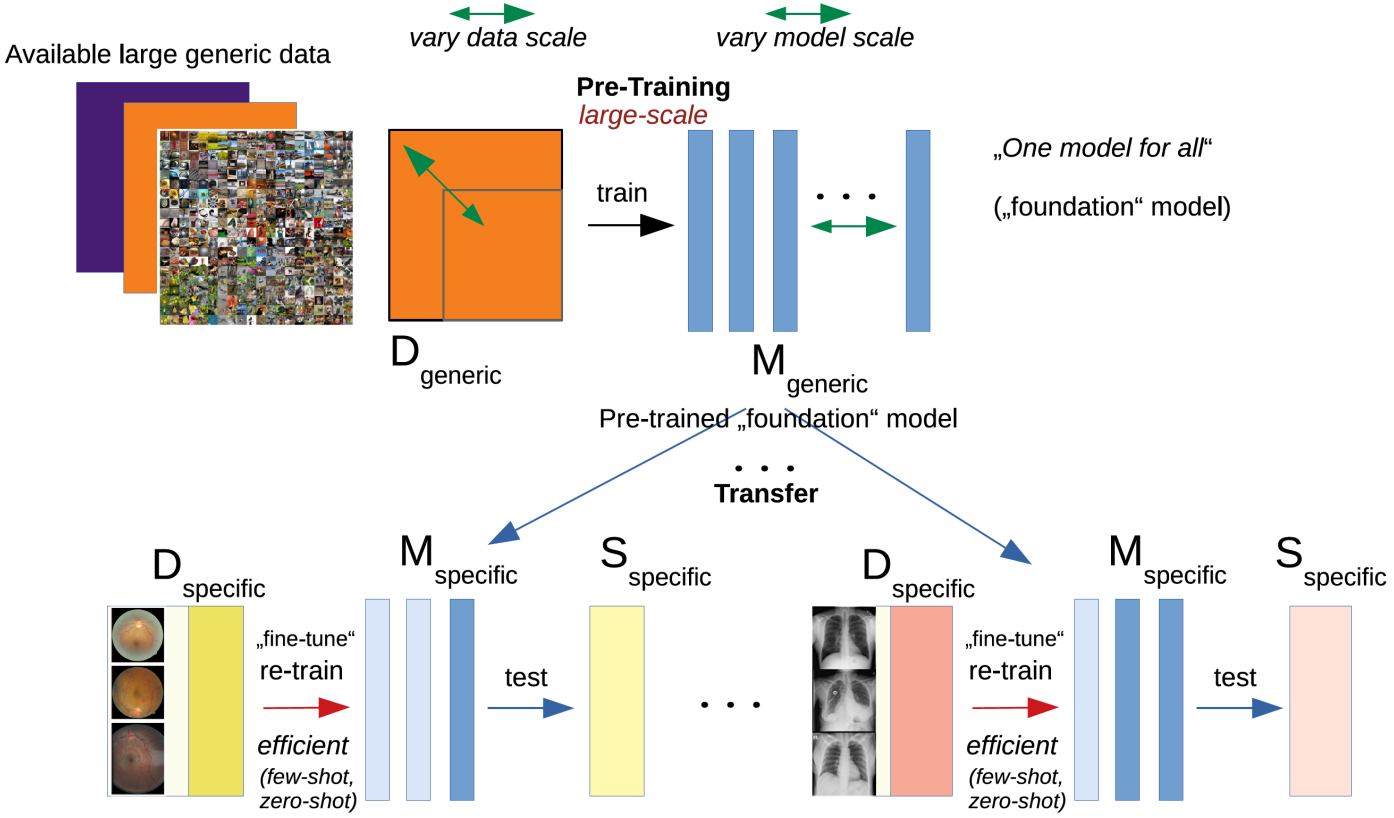


Long training

Slow
Large (Labeled) Data

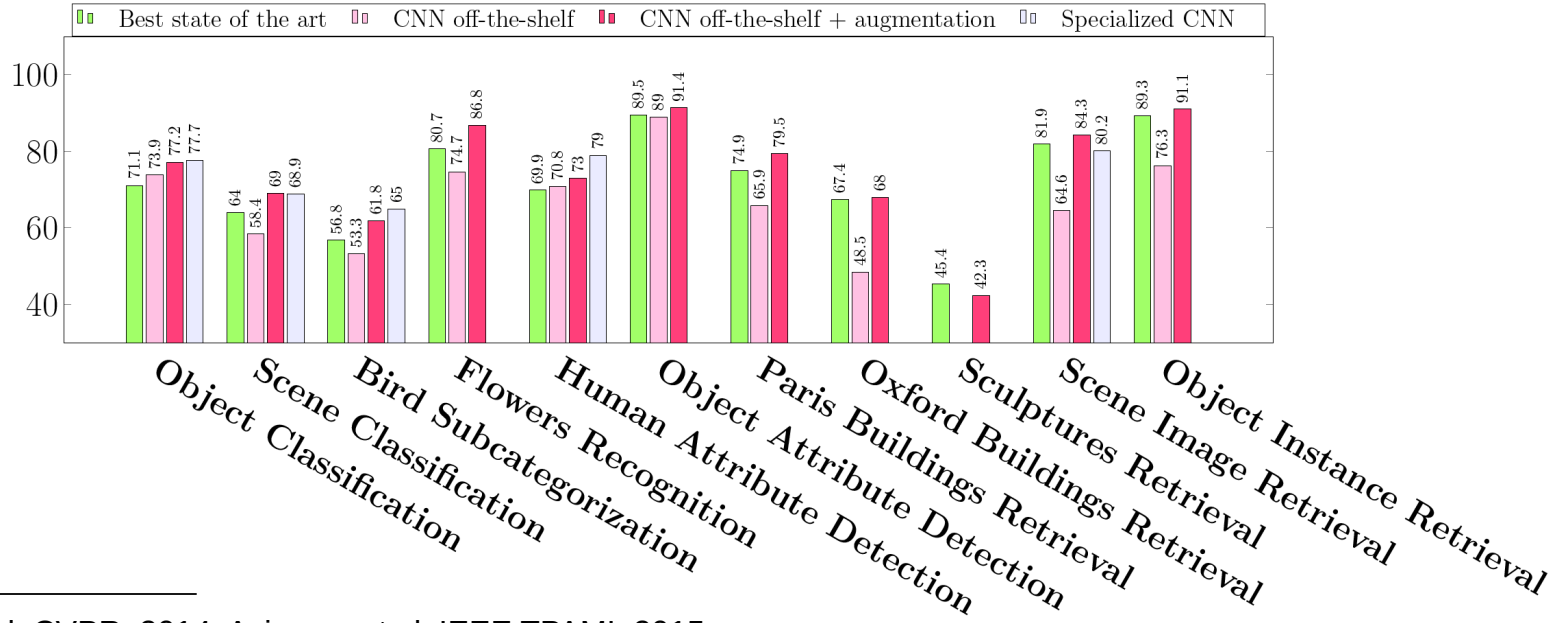
Foundation models: generic transferable learning

- Core breakthroughs (since ca. 2012): **learning that transfers across tasks**



Foundation models: generic transferable learning

- **Transferability**: evidence for early convolutional networks (OverFeat, VGG16) dating back to 2013
- „Off-the-shelf“ **transferable models**: ConvNets (CNNs) pre-trained on ImageNet-1k (1.4M images), 2012-2017 (eg ResNet – Winner ILSVRC 2015)



Razavian et al, CVPR, 2014; Azizpour et al, IEEE TPAMI, 2015

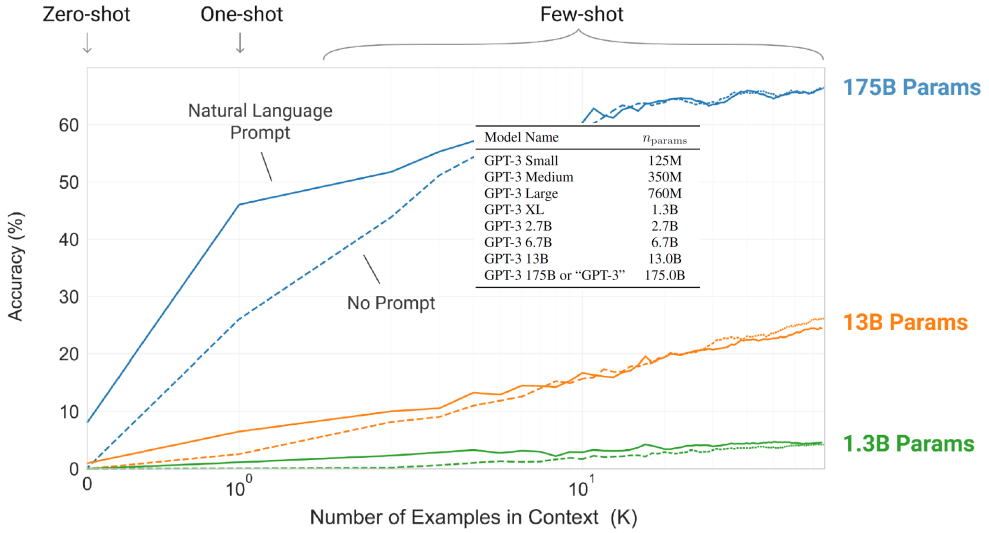
Foundation models: generic transferable learning

- Since ca. 2019: models **that transfer strongly and efficiently** across domains/ tasks („foundation“ models)**; self-supervised pre-training – scalable data!**
  - **showing scaling laws!**
- **Open vocabulary natural language:** solving tasks by natural language description via prompts; **few-**, **zero-shot transfer** and **in-context learning**. Strong ONLY **at larger scales**. Recent development: **multi-modal learning** – language + images, audio, modality X ...

Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

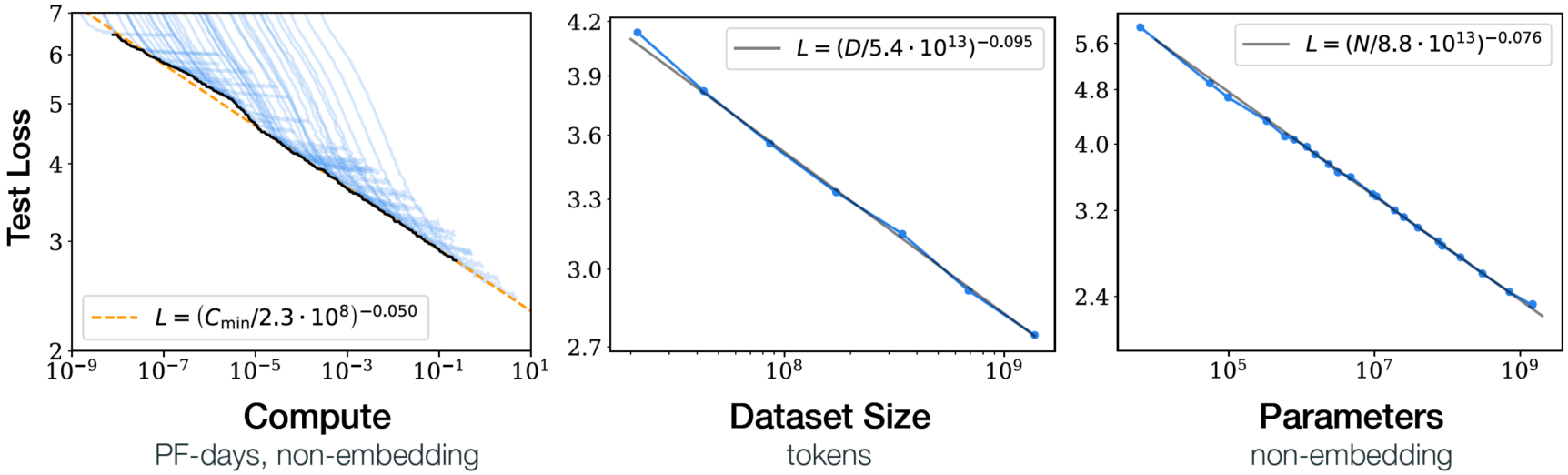
```
1  Translate English to French:          ← task description
2  sea otter => loutre de mer            ← examples
3  peppermint => menthe poivrée          ←
4  plush girafe => girafe peluche        ←
5  cheese =>                             ← prompt
```



Brown et al, NeurIPS, 2020

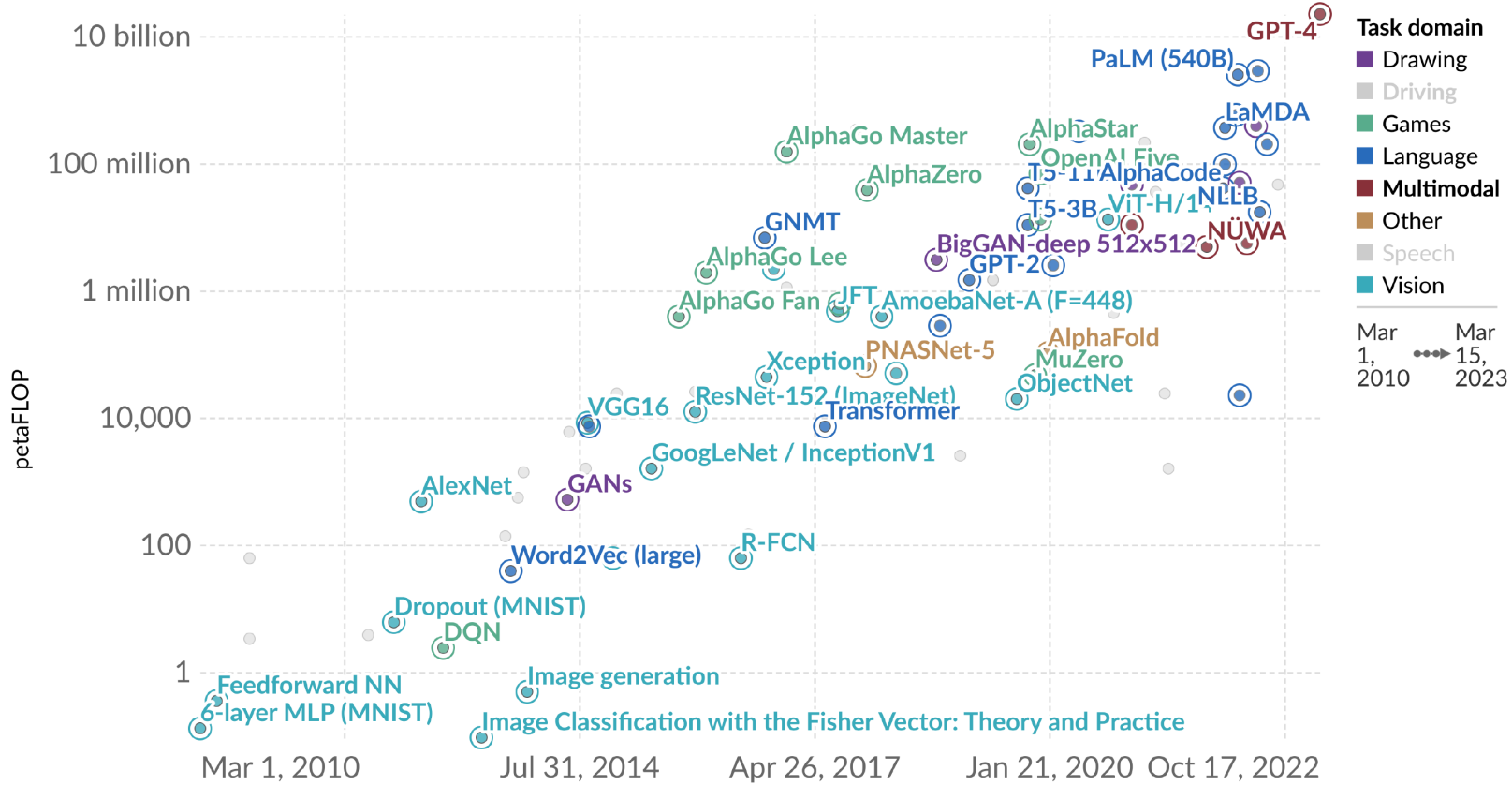
Foundation models: scaling laws

- Scaling Laws: larger model, data and compute scale during pre-training
  → stronger generalization & transferability
  No change in core algorithmic procedure required!



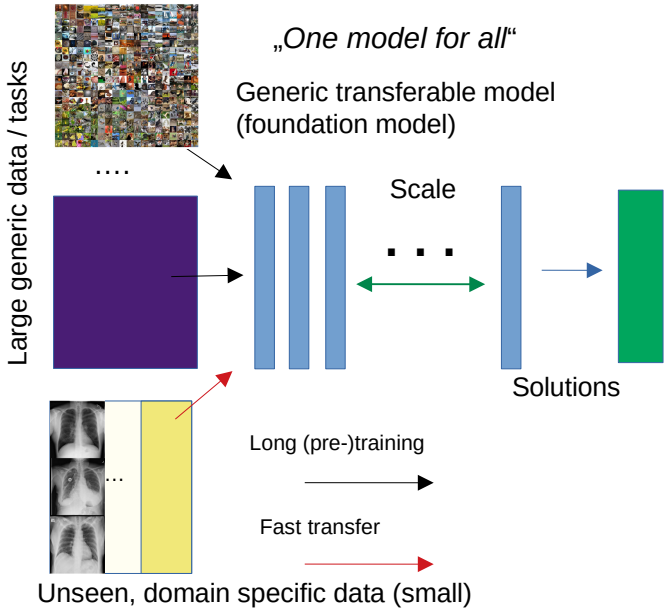
Kaplan et al, 2020

Foundation models: larger scale, stronger function



Foundation models: reproducibility & progress

- **Problem**: research on important large foundation models executable and reproducible only by few large industry labs (Google; openAI; Microsoft; Facebook; NVIDIA; …)
- **Important large foundation models:** GPT-3/4, PaLM, DALL-E 2/3, Flamingo, CLIP **- closed to public research**
- **Datasets** used to train those models**: closed as well**
- **Non-reproducible, intransparent artefacts, impairing open science, claims untestable by independent parties**



Research communities for open foundation models

- Rise of **grassroot research communities** to open-source and study foundation models & datasets required for their training
- **EleutherAI** (USA, 2020): language – Pile, Pythia, Llema (math)
- **BigScience** (EU, France, 2021): language, code, language-vision - BLOOM, StarCoder, Idefix (mostly driven by HuggingFace)
- **LAION** (EU, Germany, 2021): multi-modal language-vision, language-audio – LAION-400M/5B, openCLIP, CLAP, openFlamingo, Open Assistant, open-LM, DataComp, Leo-LM
- **Open large datasets and foundation models: reproducibility !**
  - **joint efforts accross institutions/organisations boundaries**

EleutherAI

BigScience

LAION

Large-scale Artificial Intelligence Open Network

JÜLICH Forschungszentrum

JÜLICH SUPERCOMPUTING CENTRE

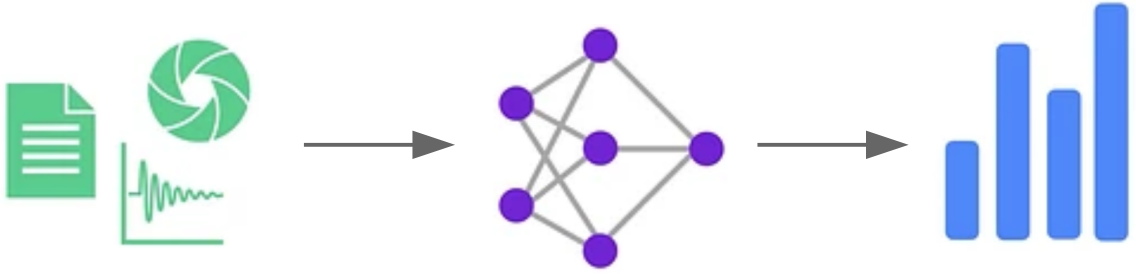
Open-source foundation models & datasets

- Making **whole pipeline** – dataset composition, model training, benchmarks & evaluation – **fully reproducible**

OPEN-SOURCE
Dataset &
Dataset composition

OPEN-SOURCE
Training procedure,
model weights,
checkpoints

OPEN-SOURCE
Evaluation benchmarks,
downstream transfer procedures



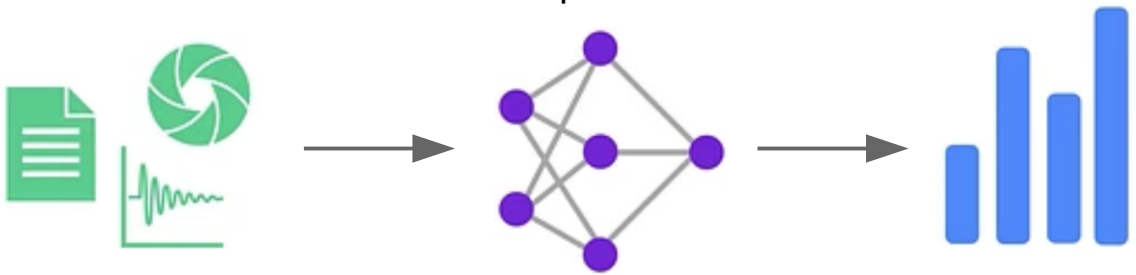
Open-source foundation models & datasets

- Making **whole pipeline** – dataset composition, model training, benchmarks & evaluation – **fully reproducible**

OPEN-SOURCE
Dataset &
Dataset composition

OPEN-SOURCE
Training procedure,
model weights,
checkpoints

OPEN-SOURCE
Evaluation benchmarks,
downstream transfer procedures



LAION-400M,
LAION-5B,
DataComp-1B

https://github.com/mlfoundations/datacomp/

OpenCLIP,
openFlamingo

https://github.com/mlfoundations/open_clip

openCLIP
Benchmarks

https://github.com/LAION-AI/CLIP_benchmark/

Open-source foundation models & datasets

- Making **whole pipeline** – dataset composition, model training, benchmarks & evaluation – **fully reproducible**

OPEN-SOURCE
Dataset &
Dataset composition

OPEN-SOURCE
Training procedure,
model weights,
checkpoints

OPEN-SOURCE
Evaluation benchmarks,
downstream transfer procedures

BigScience
EleutherAI
together.ai
AI2

Pile,
RedPajama,
Dolma

BigScience
AI2
EleutherAI

Pythia, Together-
INCITE, Olmo

EleutherAI
AI2
BigScience

Lm-eval-harness, bigcode-evaluation-harness,
https://github.com/EleutherAI/lm-evaluation-harness

Open foundation models for broad community

- Problem – composing & studying whole pipeline for open foundation models is challenging: requires
  - **large-scale data** (at least 100M of samples)
  - **large-scale compute** (GPU years per single experiment)
  - **expertise** in large-scale machine learning
  - → Broad research community cut off from training & studying strong transferable models at larger scales

→ Solution

LAION

Large-scale Artificial Intelligence Open Network

Registered as **non-profit research LAION e.V.** since 2021 in Hamburg

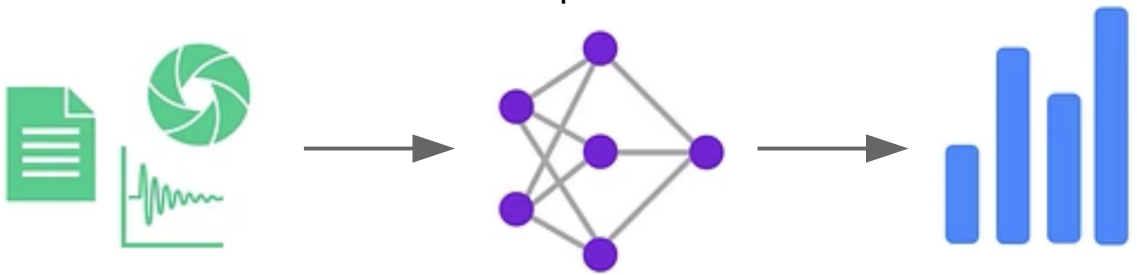
Open-source foundation models & datasets

- **Whole pipeline** – dataset composition, model training, benchmarks & evaluation – **fully reproducible**

OPEN-SOURCE
Dataset &
Dataset composition

OPEN-SOURCE
Training procedure,
model weights,
checkpoints

OPEN-SOURCE
Evaluation benchmarks,
downstream transfer procedures



LAION-400M,
LAION-5B,
DataComp-1B

https://github.com/mlfoundations/datacomp/

OpenCLIP

https://github.com/mlfoundations/open_clip

openCLIP
Benchmarks

https://github.com/LAION-AI/CLIP_benchmark/

An open foundation model: openCLIP

- CLIP – language-vision foundation model (openCLIP: open-source implementation)
  - **self-supervised** language-vision learning (scalable data - no labeling labour)

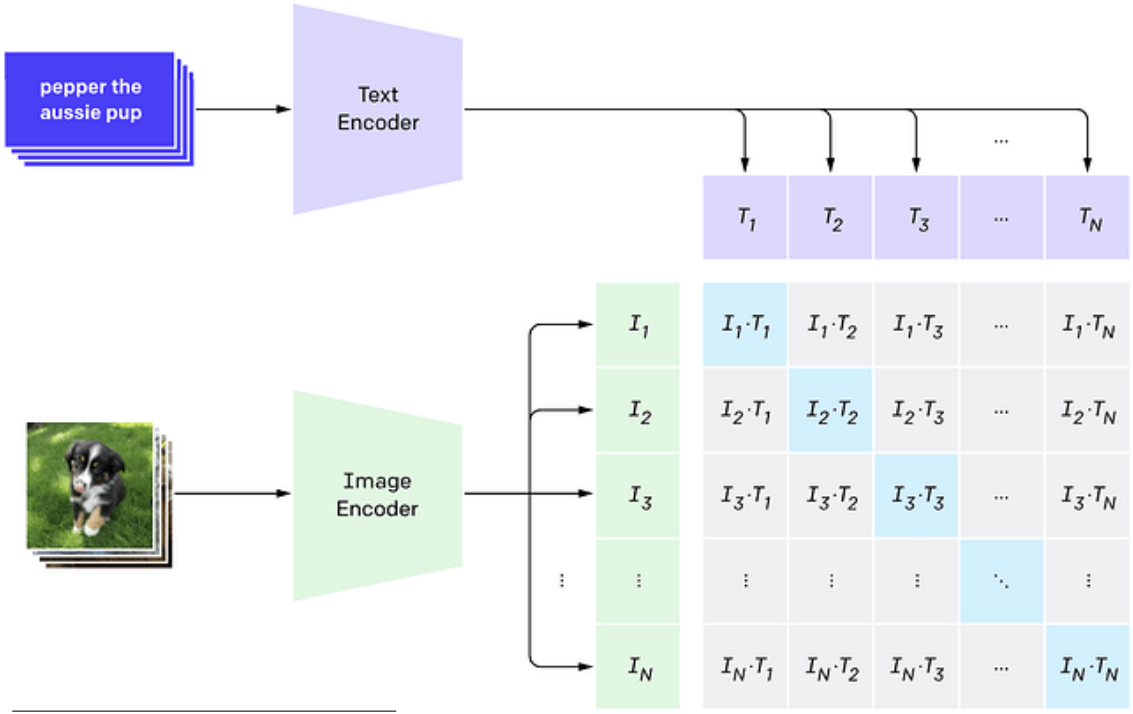
C: Green Apple Chair

C: sun snow dog

C: pink, japan, aesthetic image

C: french cat

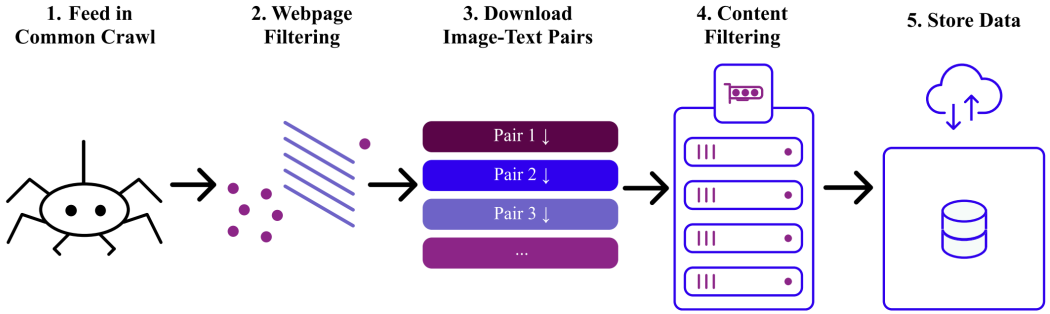
Schuhmann et al, NeuRIPS, 2021



Radford et al, ICML, 2021

Open large-scale foundation data

- LAION-400M/5B, DataComp-1B: Open sourcing data collection procedures - transparent dataset, open source toolsets, reproducible training across various scales
- Open dataset: collection of text and links to images on public Internet



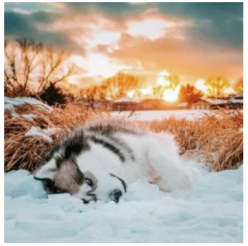
Dataset	# English Img-Txt Pairs
Public Datasets	
MS-COCO	330K
CC3M	3M
Visual Genome	5.4M
WIT	5.5M
CC12M	12M
RedCaps	12M
YFCC100M	100M[2]
LAION-5B (Ours)	**2.3B**
Private Datasets	
CLIP WIT (OpenAI)	400M
ALIGN	1.8B
BASIC	6.6B

Schuhmann et al, NeurIPS, 2022

Open large-scale foundation data

- LAION-400M/5B, DataComp-1B: Open sourcing data collection procedures - transparent dataset, open source toolsets, reproducible training across various scales

C: Green Apple Chair



C: sun snow dog

C: pink, japan, aesthetic image

Dataset	# English Img-Txt Pairs
Public Datasets	
MS-COCO	330K
CC3M	3M
Visual Genome	5.4M
WIT	5.5M
CC12M	12M
RedCaps	12M
YFCC100M	100M[2]
LAION-5B (Ours)	**2.3B**
Private Datasets	
CLIP WIT (OpenAI)	400M
ALIGN	1.8B
BASIC	6.6B

- Follow-up: DataComp-1B - LAION/University of Washington/Allen AI institute; www.datacomp.ai

Schuhmann et al, NeurIPS, 2022

Reproducible scaling laws for foundation models

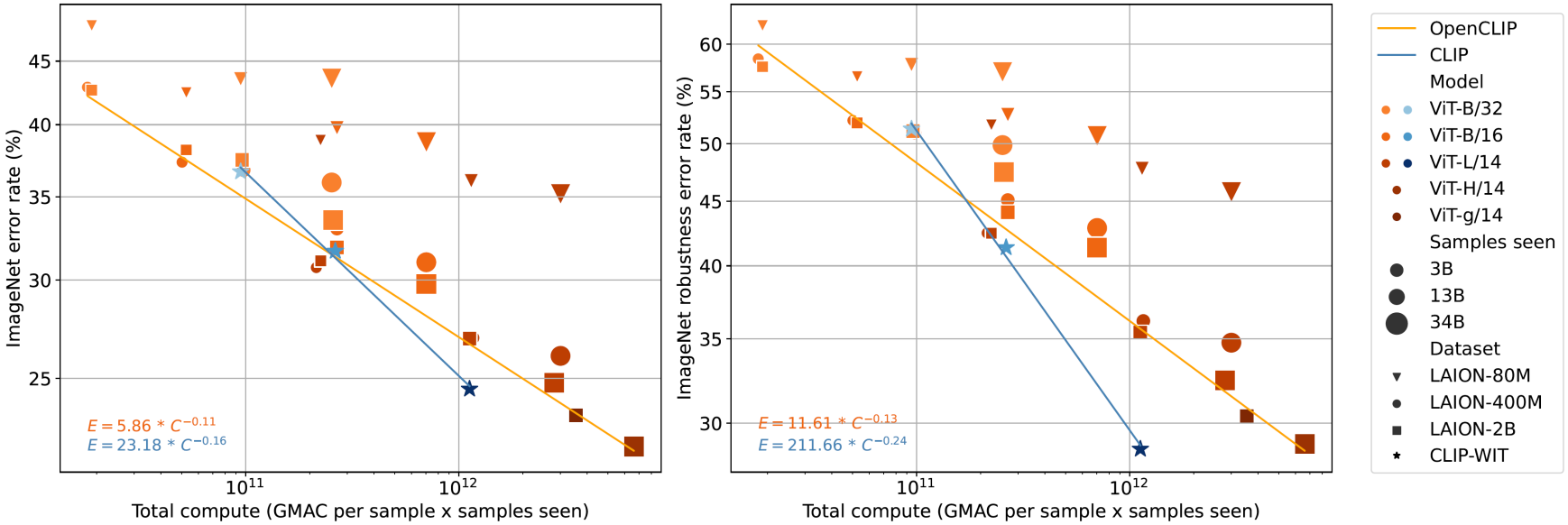
- Systematically varying data, samples seen and model scale
- Example: Zero-shot ImageNet-1k Top-1 accuracy

Model	Samples seen	LAION-80M	LAION-400M	LAION-2B
ViT-B/32	3B	51.94	57.12	57.36
	13B	56.46	63.23	62.53
	34B	56.43	64.06	66.47
ViT-B/16	3B	57.55	62.68	61.82
	13B	60.24	67.00	68.13
	34B	61.28	69.00	70.22
ViT-L/14	3B	61.14	69.31	68.93
	13B	63.96	73.06	73.10
	34B	64.83	73.94	75.20

Schuhmann et al, NeurIPS, 2022; Cherti et al, CVPR 2023

Reproducible scaling laws for foundation models

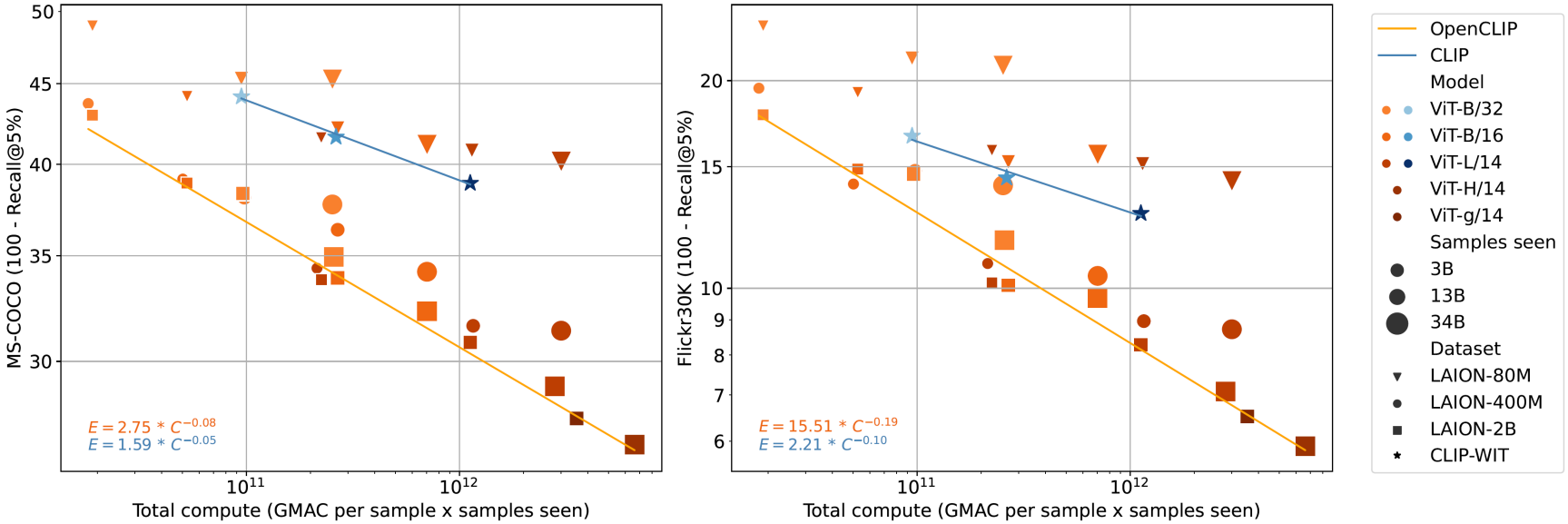
- Scaling laws with LAION-400M/2B and openCLIP: open-source data, models and code - reproducible science of foundation models



Schuhmann et al, NeurIPS, 2022; Cherti et al, CVPR 2023

Reproducible scaling laws for foundation models

- Scaling laws for various task types (here: zero-shot image retrieval, MS-COCO & Flickr30K)



Cherti et al, CVPR 2023

Open foundation models: reproducibility

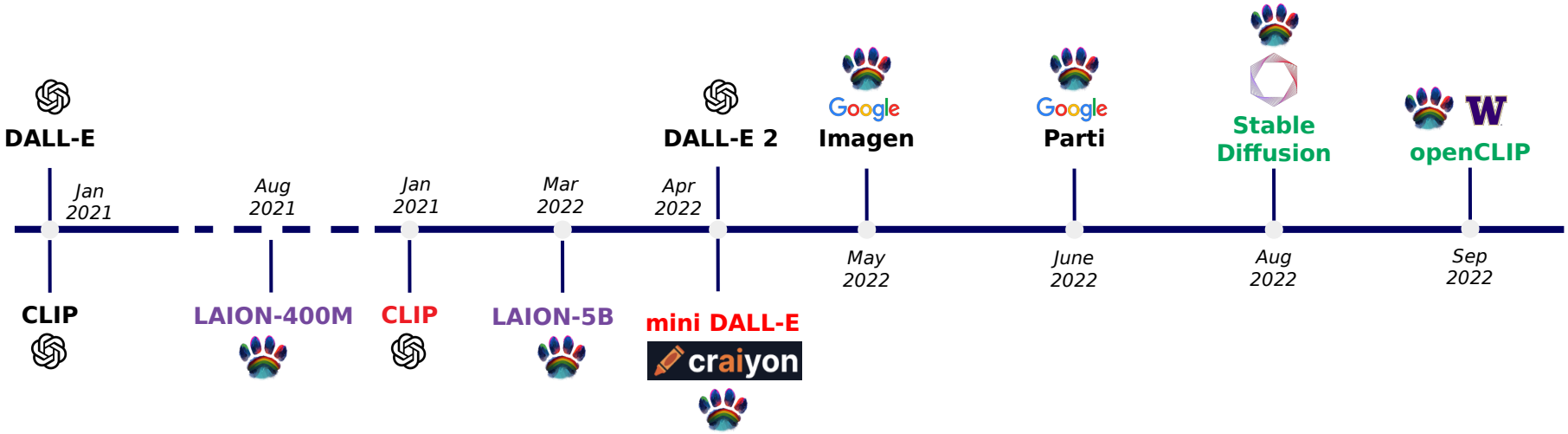
- Ingredients for an reproducible, open foundation model
  - open **large-scale dataset** & open dataset composition
  - open **pre-training** procedure (**compute intensive - supercomputers**)
  - open **transfer** procedures (zero-shot, linear probing, fine-tuning, ...)
  - open **standardized evaluation benchmarks** (eg: https://github.com/LAION-AI/CLIP_benchmark, https://github.com/EleutherAI/lm-evaluation-harness

  → Enables **reproducible scaling laws** that can be validated/falsified;

- **Open-sourcing** datasets, pre-trained and transfered models further facilitates reproducibility and further studies
  - Re-use as building blocks for other complex learning systems
  - Example: LAION datasets and openCLIP pre-trained models as critical components of Stable Diffusion

From closed to open data and models: a timeline

- Open-source releases fertilize research and technology development



Closed model in black
Open release pre-trained models in red
Open data in purple
Open foundation models in green

Adapted from State of AI report, 2022

Open datasets & models @ LAION

- Foundation models, open-source
  - **OpenCLIP** ViT B/32 - G/14: **representation learning** at larger scale
    - (open)CoCa: **image-to-text generative**
  - **Stable Diffusion**, openImagen, Paella, Wuerstchen: **text-to-image generative**
  - **OpenFlamingo**-3B/4B/9B: interleaved **image-text sequences**, text generative
  - **LAION-CLAP:** contrastive **language-audio** learning
  - **together-INCITE**-3B/7B; **OA-falcon**-7B/40B; **LeoLM**-3B/7B/70B (German tuned LlaMA 2), **open-lm 1B, 7B**: **language models**

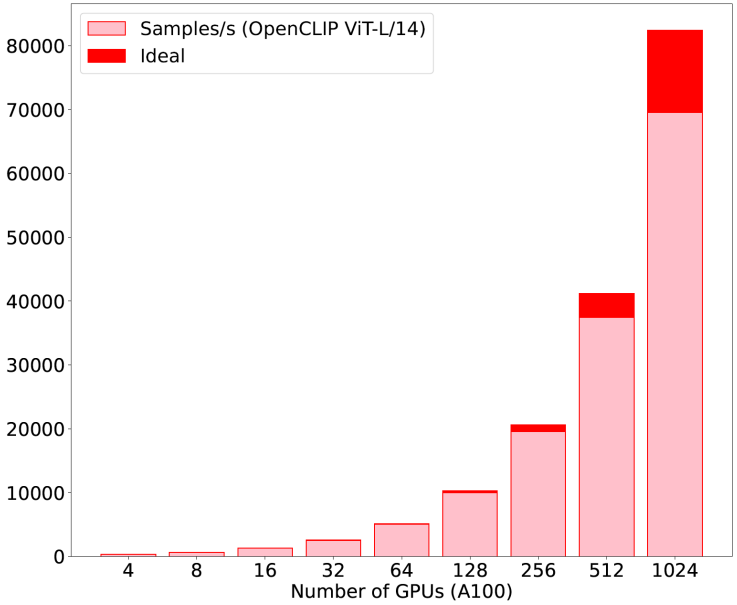
- Foundation datasets, open-source
  - **LAION-400M, LAION-5B** (used by: openCLIP, Stable Diffusion, FLAVA, EVA, …), LAION-Aestethics (Stable Diffusion, …)
  - LAION-audio-630k: language-audio
  - **DataComp-1B** (openCLIP, CLIPA)

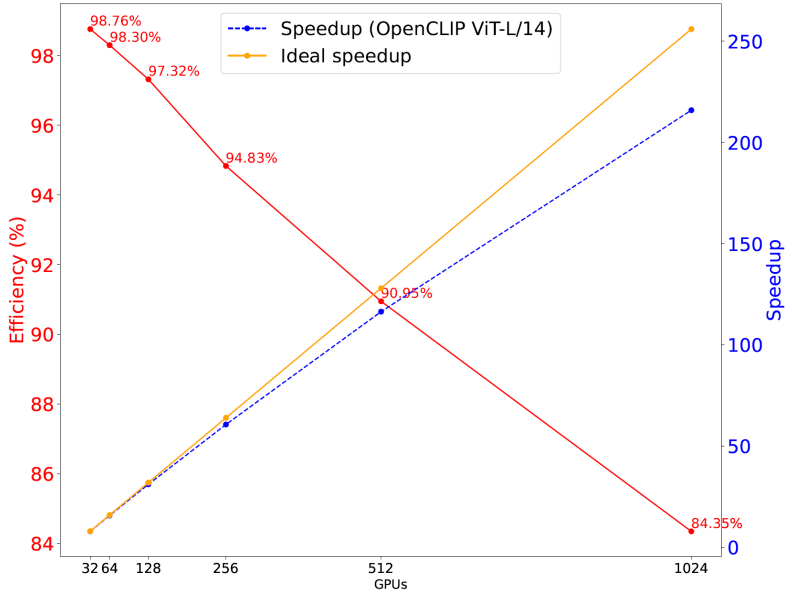
Open science for large-scale foundation models

- **LAION**: **L**arge-scale **A**rtificial **I**ntelligence **O**pen **N**etwork
  - **compute**: applying for publicly funded supercomputers
    - **JUWELS Booster**, Germany: Gauss Center for Supercomputing
    - **Summit**, USA: INCITE Leadership computing call
    - **LUMI** (Finland), **Leonardo** (Italy): **EuroHPC** calls (Extreme Scale grant)

Supercomputers for foundation model training

- Supercomputers: necessary for the training experiments (eg openCLIP ViT L/14: 122 hours with 1024 A100 - total of 124K GPU hours)
- Common effort avoids replication of same expensive experiments





Cherti et al, CVPR 2023

Open science of large-scale foundation models

- Supercomputers – hubs for large-scale basic AI research
- Open science for advancing powerful, safe generic AI tools for public

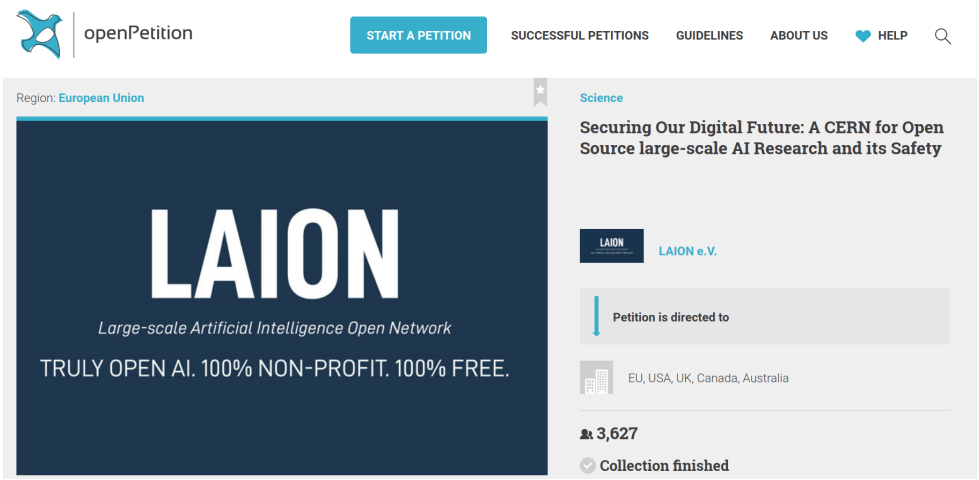
*Stable Diffusion 1.5, trained on **LAION-5B** image-text dataset.*
Prompt: "An epic scene of a supercomputing center building of the future, embedded in a rich wild green exotic blooming jungle forest, nearby a lake“

LAION: strong grassroot research community

- Collaborative work of broadly distributed community: **Outstanding NeurIPS 2022 paper award**, impacting open source releases (see repos)
- **Falling Walls Award**: Scientific Breakthrough 2023
- LAION public Discord server: > **27k members**

LAION: research community & alliances

- Various alliances in EU: ELLIS, Tuebingen AI Center, MPI for Intelligent Systems, ellamind, Hessian AI & TU Darmstadt, HuggingFace, FAIR (Italy), U Turku & SILO AI (HPLT, Finland), …
- Various alliances worldwide: U Washington, Allen AI Institute, Stanford, Together AI, U Montreal, Tokyo Tech, U Berkeley, U Tel Aviv, ...



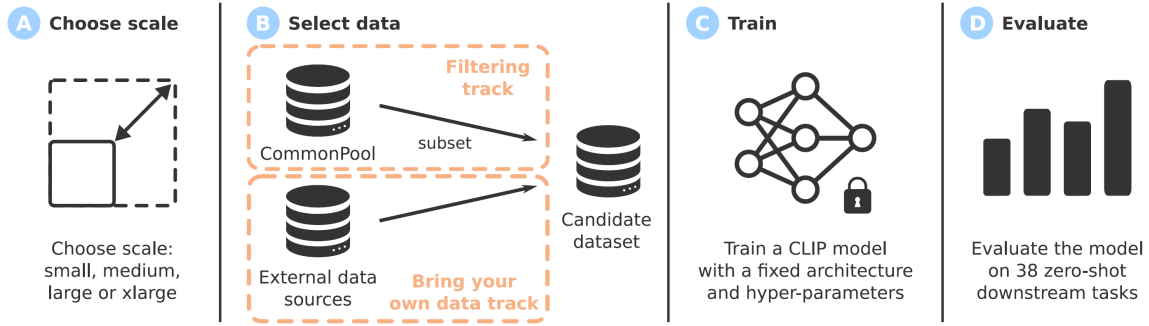
Foundation models: improving scaling

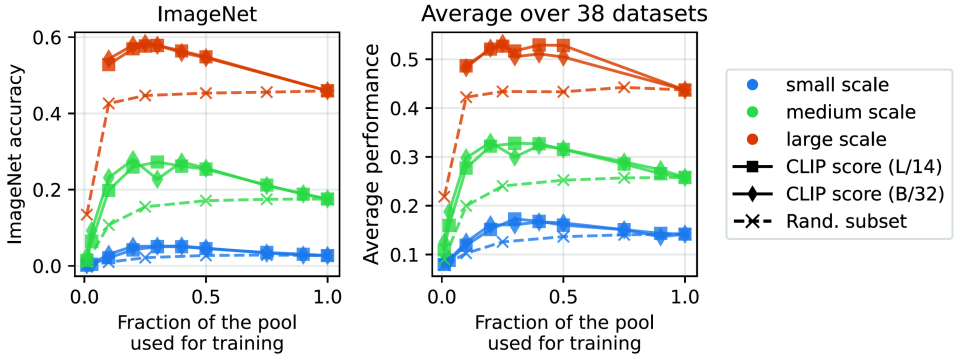
- Scaling laws: predict better core function when increasing scales
- How to systematically obtain stronger and stronger improvement with scale – get stronger capability gradient with respect to scale?
- Various ways to **get stronger scaling**:
  - **improve dataset composition** for pretraining
  - **improve learning procedure** (architecture, loss & optimization, …)

- Systematic search for scalable learning: Project Nucleus (ELLIS Unit Freiburg, U Freiburg, Frank Hutter)

„Foundation“ datasets for next-gen FMs

- DataComp (2023, NeurIPS): what constitutes good data for FM training?



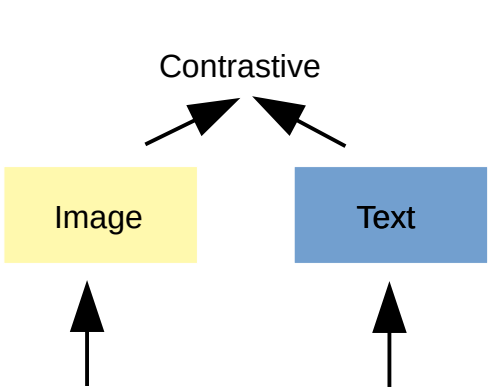


Dataset	Dataset size	# samples seen	Architecture	Train compute (MACs)	ImageNet accuracy
OpenAI's WIT [111]	0.4B	13B	ViT-L/14	1.1×10^{21}	75.5
LAION-400M [128, 28]	0.4B	13B	ViT-L/14	1.1×10^{21}	72.8
LAION-2B [129, 28]	2.3B	13B	ViT-L/14	1.1×10^{21}	73.1
LAION-2B [129, 28]	2.3B	34B	ViT-H/14	6.5×10^{21}	78.0
LAION-2B [129, 28]	2.3B	34B	ViT-g/14	9.9×10^{21}	78.5
DATACOMP-1B (ours)	1.4B	13B	ViT-L/14	1.1×10^{21}	**79.2**

Gadre et al, NeurIPS 2023 (Oral)

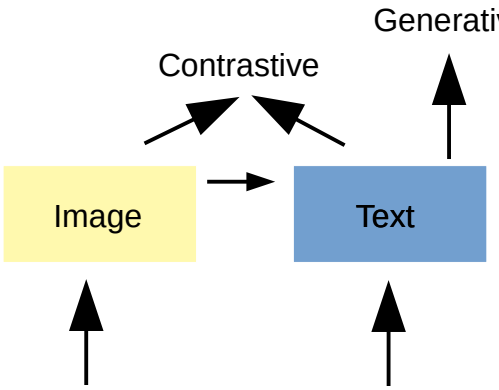
Strongly scalable open foundation models

- DataComp & follow-up work: improving datasets for pre-training
- OpenCLIP extensions: improving learning procedure
  - extend for text & image generative losses (CoCa, Mammut)
  - what loss mix might have stronger scaling? Scaling laws required



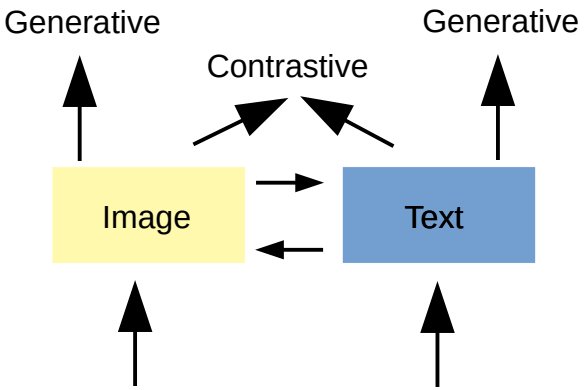
Standard CLIP

Pure Contrastive



CoCa, Mammut

Contrastive + Text Generative



CoBIT,SyCoCa

Contrastive +
+ Image, Text Generative

Open model benchmarks: measuring it right

- Alice in Wonderland (AIW) Problem: very simple problems breaking SOTA LLMs (GPT-4o, Claude Opus, Gemini 1.5 Pro, etc)

Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models

Marianna Nezhurina[1,2,4*] **Lucia Cipolina-Kun**[1,3] **Mehdi Cherti**[1,2,4] **Jenia Jitsev**[1,2,4*]
[1]LAION [2]Juelich Supercomputing Center (JSC), Research Center Juelich (FZJ)
[3] School of Electrical and Electronic Engineering, University of Bristol
[4] Open-Ψ (Open-Sci) Collective
*Corresponding authors: {m.nezhurina,j.jitsev}@fz-juelich.de,contact@laion.ai

Figure 1: Alice is reasoning: will it break? Illustration of Humpty Dumpty from Through the Looking Glass, by John Tenniel, 1871. Source: Wikipedia.

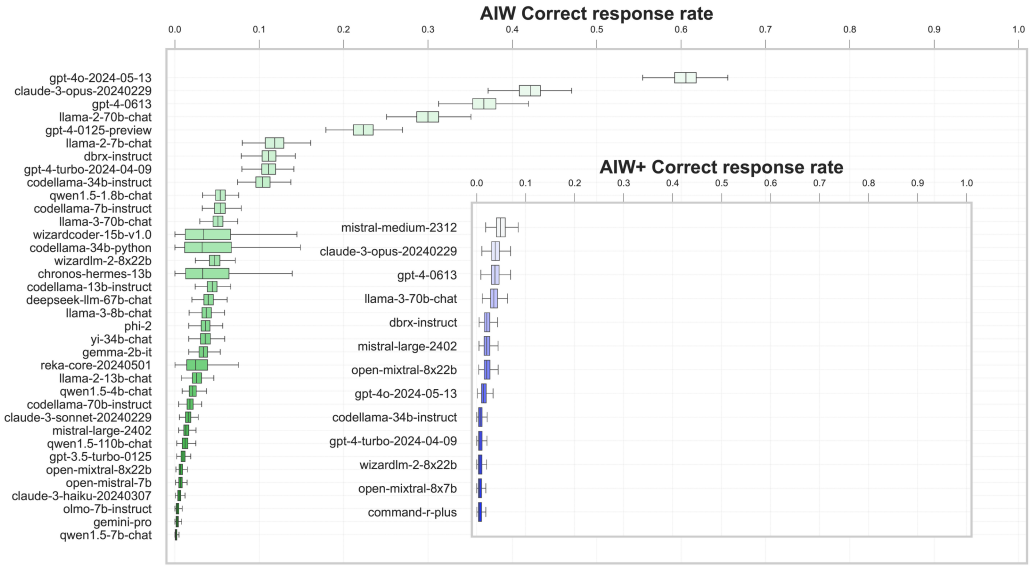
Nezhurina et al, ArXiv:2406.02061, 2024

Open model benchmarks: measuring it right

- Alice in Wonderland (AIW) Problem: very simple problems breaking LLMs

Table 2: AIW main variations and prompt types.

Var.	Prompt	Type	ID
1	Alice has 3 brothers and she also has 6 sisters. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	55
1	Alice has 3 brothers and she also has 6 sisters. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	57
1	Alice has 3 brothers and she also has 6 sisters. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	53
2	Alice has 2 sisters and she also has 4 brothers. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	56
2	Alice has 2 sisters and she also has 4 brothers. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	58
2	Alice has 2 sisters and she also has 4 brothers. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	54
3	Alice has 4 sisters and she also has 1 brother. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	63
3	Alice has 4 sisters and she also has 1 brother. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	64
3	Alice has 4 sisters and she also has 1 brother. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	65
4	Alice has 4 brothers and she also has 1 sister. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	69
4	Alice has 4 brothers and she also has 1 sister. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	70
4	Alice has 4 brothers and she also has 1 sister. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	71



Nezhurina et al, ArXiv:2406.02061, 2024

Open model benchmarks: measuring it right

- Alice in Wonderland (AIW) Problem: very simple problems breaking LLMs

Table 2: AIW main variations and prompt types.

Var.	Prompt	Type	ID
1	Alice has 3 brothers and she also has 6 sisters. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	55
1	Alice has 3 brothers and she also has 6 sisters. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	57
1	Alice has 3 brothers and she also has 6 sisters. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	53
2	Alice has 2 sisters and she also has 4 brothers. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	56
2	Alice has 2 sisters and she also has 4 brothers. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	58
2	Alice has 2 sisters and she also has 4 brothers. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	54
3	Alice has 4 sisters and she also has 1 brother. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	63
3	Alice has 4 sisters and she also has 1 brother. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	64
3	Alice has 4 sisters and she also has 1 brother. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	65
4	Alice has 4 brothers and she also has 1 sister. How many sisters does Alice's brother have? Solve this problem and provide the final answer in following form: "### Answer: ".	STANDARD	69
4	Alice has 4 brothers and she also has 1 sister. How many sisters does Alice's brother have? Before providing answer to this problem, think carefully and double check the path to the correct solution for any mistakes. Provide then the final answer in following form: "### Answer: ".	THINKING	70
4	Alice has 4 brothers and she also has 1 sister. How many sisters does Alice's brother have? To answer the question, DO NOT OUTPUT ANY TEXT EXCEPT following format that contains final answer: "### Answer: ".	RESTRICTED	71

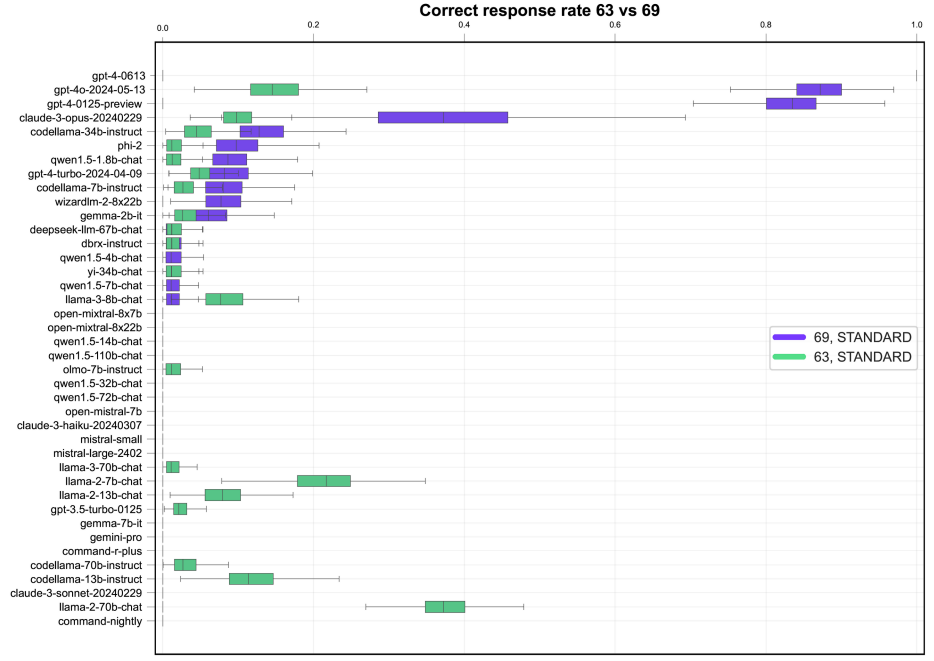


Figure 11: Strong fluctuations of AIW correct response rate on AIW variations (here on example of AIW Variation 3 vs. 4, STANDARD prompt type). GPT-4-0613 collapses from correct response rate 1 to 0 between variations. Also GPT-4o, GPT-4-Turbo, Claude 3 Opus and Llama 2 7B and 70B show strong discrepancies. Models for which a particular color is entirely omitted have zero performance on the AIW variation with corresponding color (with exception of GPT-4-0613 on the very top, which has correct response rate of 1 on AIW Variation 4, prompt ID 69, and thus also have vanishing color bars for both variations.

Nezhurina et al, ArXiv:2406.02061, 2024

Open foundation models: outlook

- „Moonshot“: build **open-sci-MM as open multi-modal foundation model**
  - Strong impact across various disciplines beyound core machine learning
    - Focus on reasoning, coding, complex workflow automation
    - Foundation models for science & semi-automated scientific discovery
    - Customized AI assistants for citizens, for governance, for education, … **researched, developed and deployed in EU from open base validated by broad community**

- **„LAION/ELLIS/BigScience 2.0“** : Germany/France (Italy/Spain/Netherlands/Finland/Israel/…) - EU consortium for building large open foundation models that are powerful, transparent and **validated by research community** for **safe fine-tuning and deployment**

- **Substantially** push boundaries for scalability of **strongly transferable generalist learning**: local losses based architectures, systematic search via Project Nucleus

Acknowledgements


Dr. Mehdi Cherti, Marianna Nezhurina, JSC

Visit https://laion.ai/
Join public LAION Discord server
for more projects
and research tracks
> 27k members !

LAION community & friends (Romain Beaumont, Ross Wightmann, Irina Rish, ...)

Prof. Ludwig Schmidt, UoW

Christoph Schumann

Let‘s build open, safe AI foundations together!

BigScience

EleutherAI

LAION
Large-scale Artificial Intelligence Open Network

WEST AI
KI-Servicezentrum

JÜLICH
Forschungszentrum

JÜLICH
SUPERCOMPUTING
CENTRE

Supplementary Material

Supercomputers for foundation model training

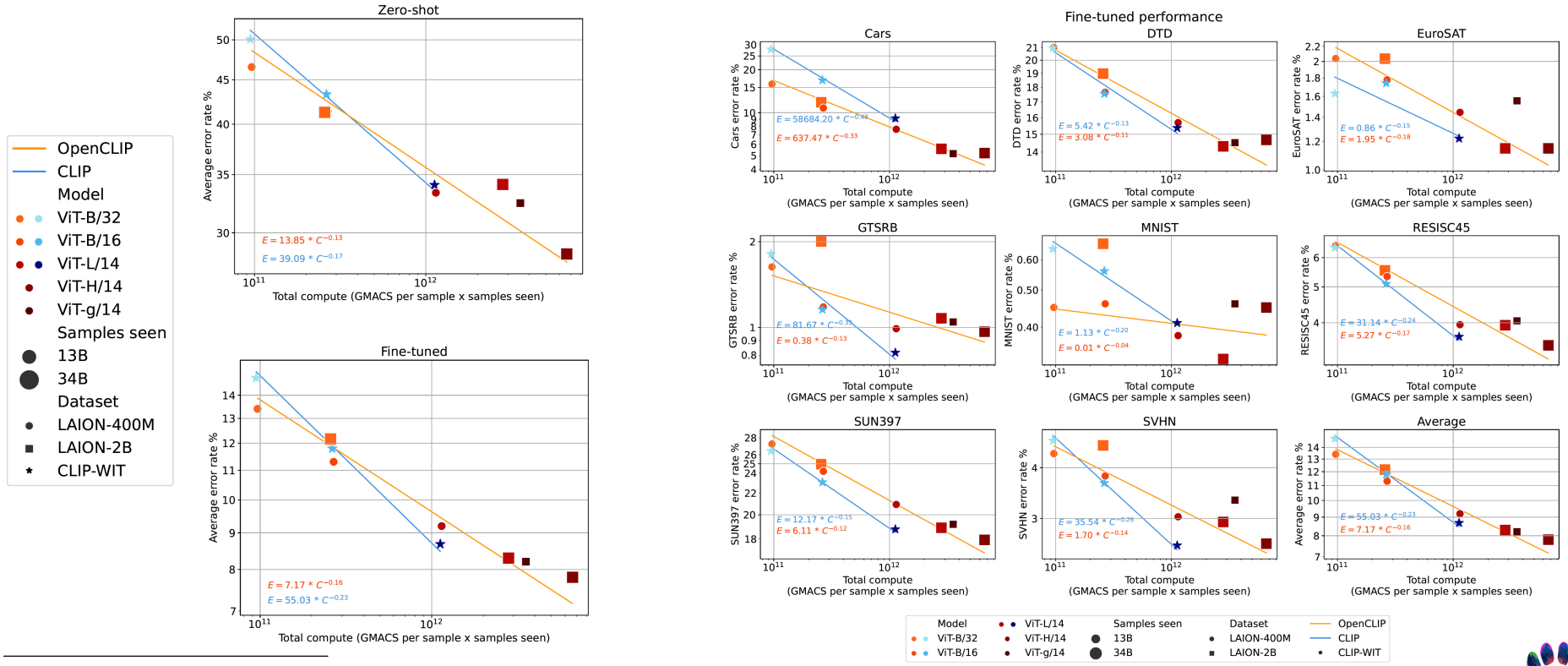
- Still rather modest scale (compared to LLMs (>100B params); PaLI-X (image-text-to-text) – 55B; Parti (text-to-image) - 20B params)
  - Obtaining stronger transfer and robustness requires larger scales
  - Larger supercomputers necessary: eg **JUPITER** Exascale (JSC)

Name	Width	Emb.	Depth	Acts.	Params	GMAC
ViT-B/32	768 / 512	512	12 / 12	10 M	151 M	7.40
ViT-B/16	768 / 512	512	12 / 12	29 M	150 M	20.57
ViT-L/14	1024 / 768	768	24 / 12	97 M	428 M	87.73
ViT-H/14	1280 / 1024	1024	32 / 24	161 M	986 M	190.97
ViT-g/14	1408 / 1024	1024	40 / 24	214 M	1.37 B	290.74
ViT-G/14	1664 / 1280	1280	48 / 32	310 M	2.54 B	532.92

Cherti et al, arXiv:2212.07143, CVPR, 2023

Reproducible scaling laws for foundation models

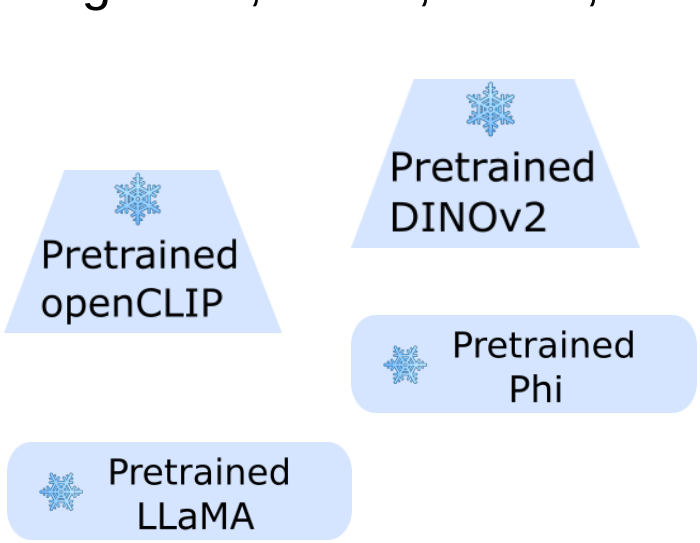
- Scaling laws for various transfer procedures and downstream datasets

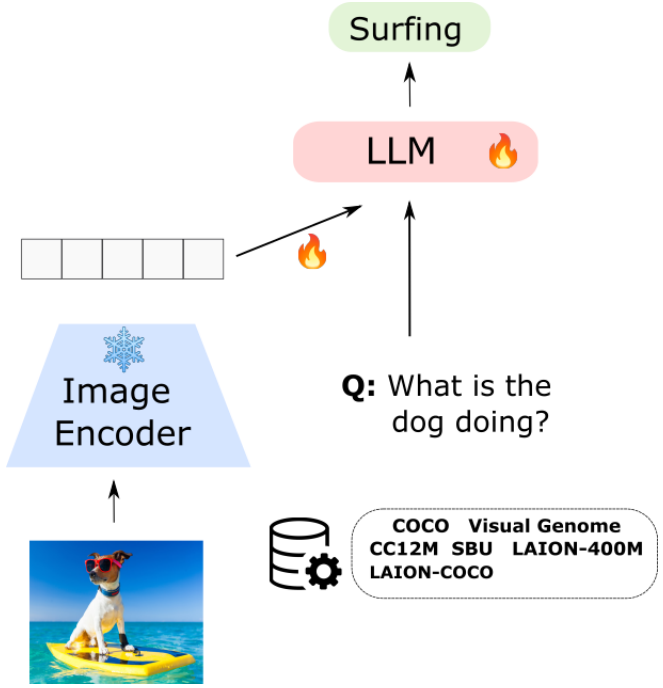


Cherti et al, CVPR 2023

Foundation models as re-usable components

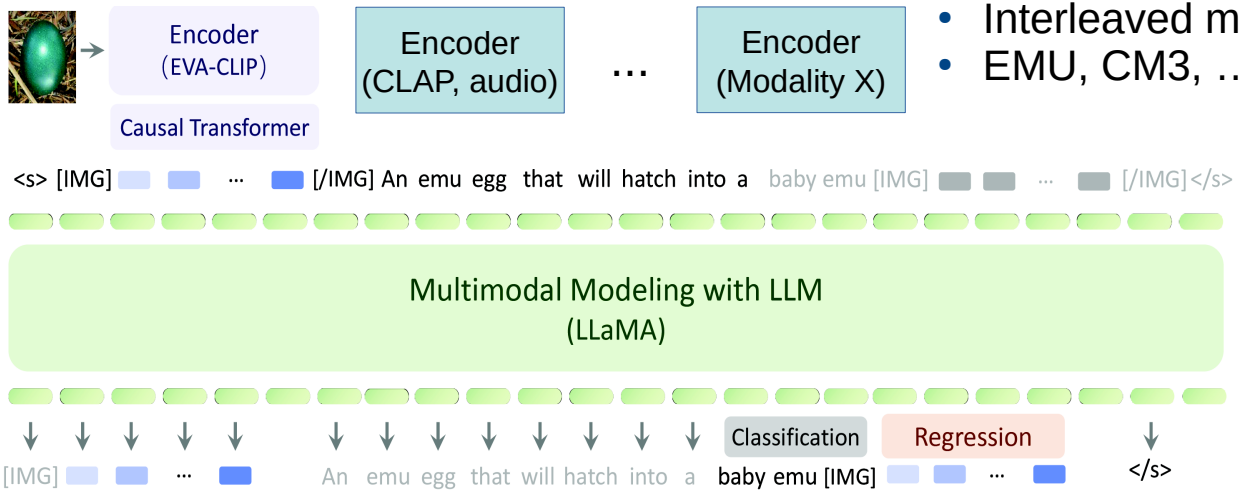
- Combining pre-trained foundation models for more complex generalist function (no or little adaptation required): Flamingo, BLIP-2, ImageBind, LENS, LlaVA, ...





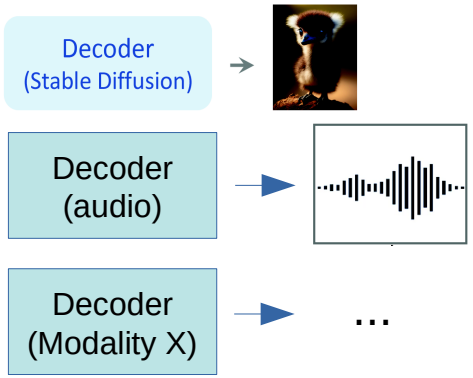
Haotian Liu et al, arXiv:2304.0848, NeurIPS 2023

Open interleaved multi-modal foundation models



- Interleaved multi-modal learning
- EMU, CM3, …

- Sequence of arbitrary modalities: language, code, images **both for input & output**
- Generalist foundation models with potential for
  - **document/media understanding** & generation
  - handling custom tasks across domains via multi-modal in-context learning
  - **Synthetic data generation**: towards **„foundation“ datasets** as mix of synthetic & real data
  - Towards personalized generalist assistants **automating scientific and generic worflows**



adapted from Sun et al, arXiv:2307.05222, 2023 (EMU)