

Pretraining and Finetuning LLMs from the Ground Up

July 8, 2024 1:30 - 5:30 pm PT



@rasbt



/in/sebastianraschka



<https://sebastianraschka.com>



Lightning AI

<https://lightning.ai>



SciPy 2024

Tacoma Convention Center | Tacoma, WA

July 8-14, 2024



Sebastian Raschka

<https://sebastianraschka.com>



Creators of PyTorch Lightning
Staff Research Engineer

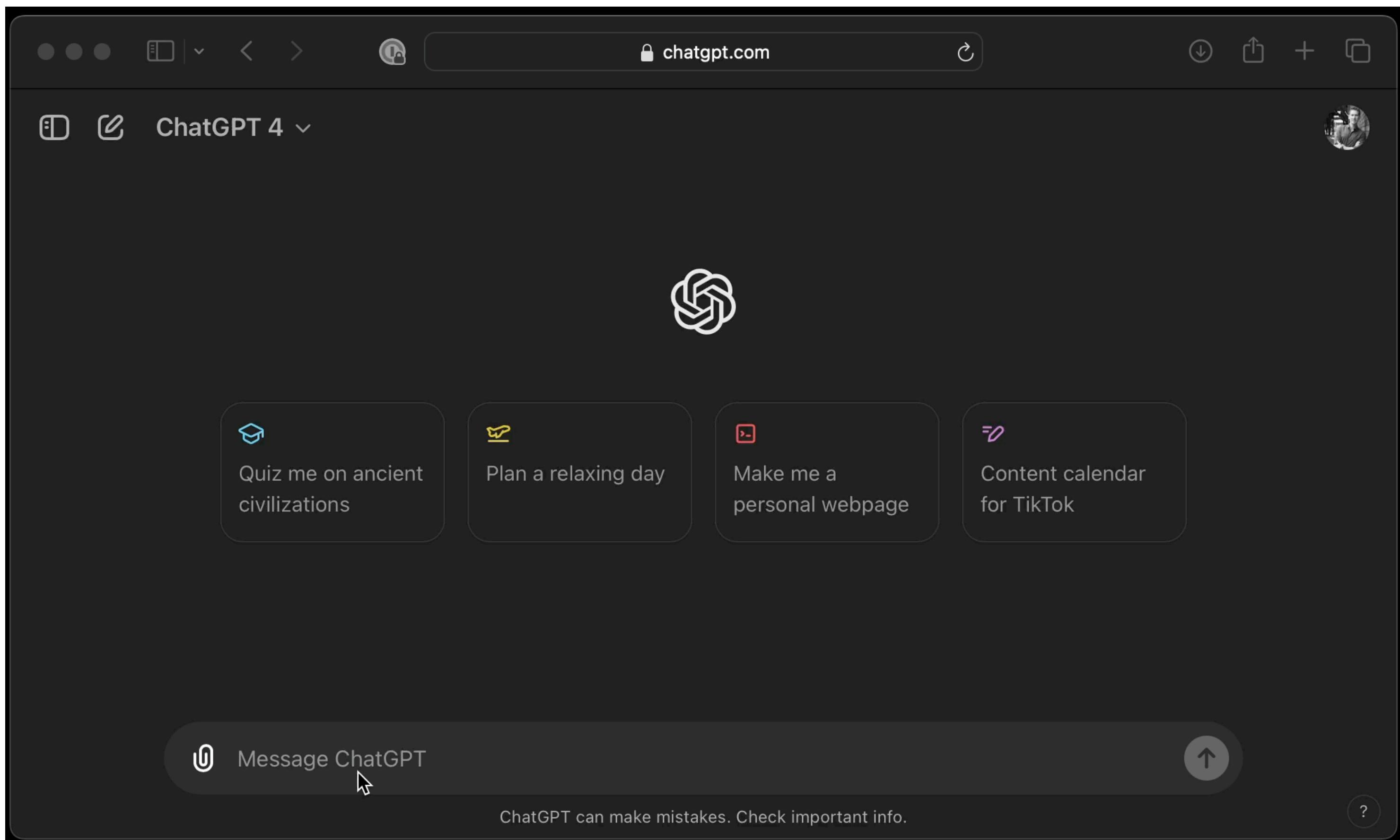


Previously
Asst. Prof. of Statistics

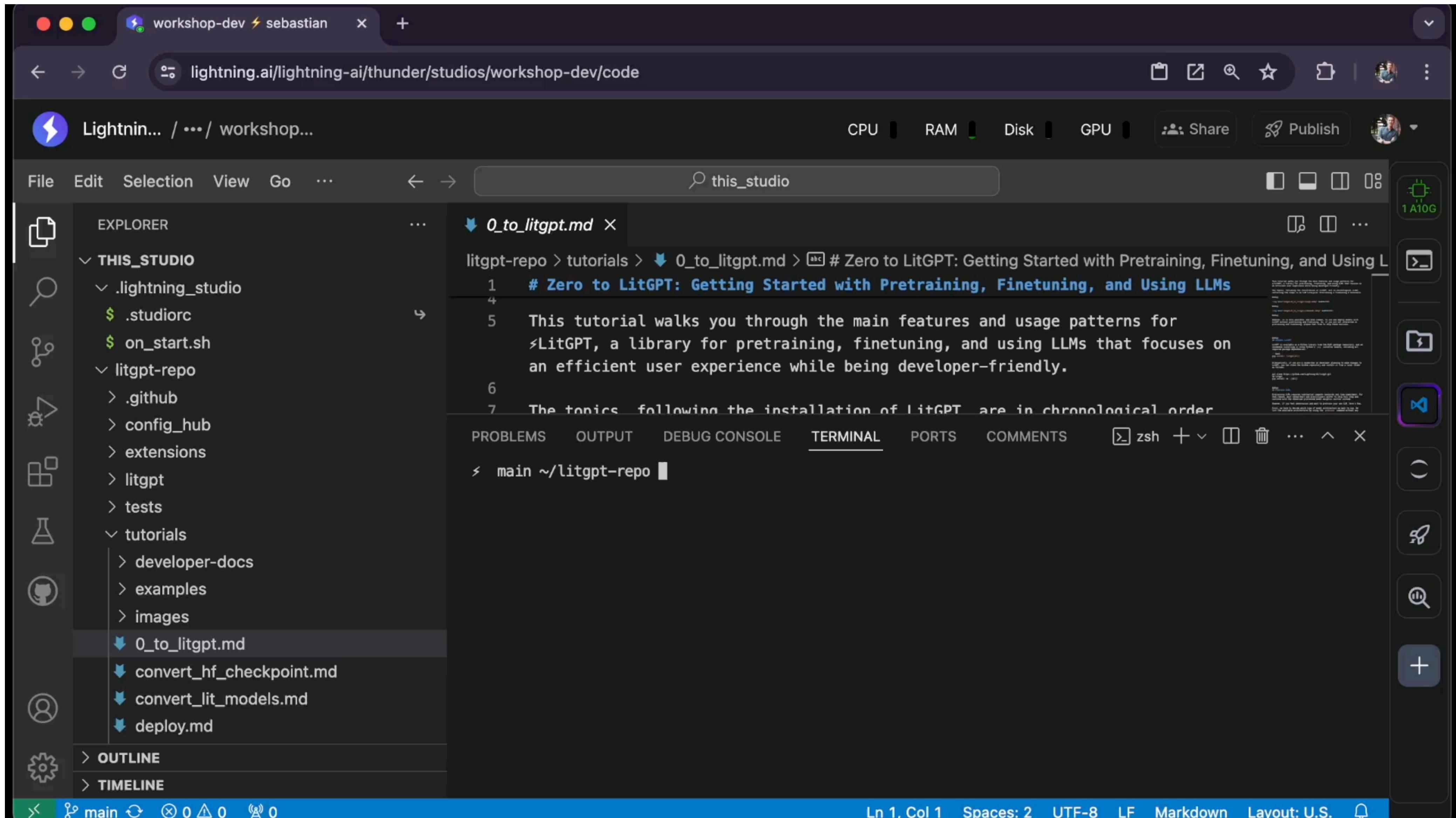
<h2>Workshop topics</h2>	
1	Introduction to LLMs
2	Understanding LLM input data
3	Coding an LLM architecture
4	Pretraining LLMs
5	Loading pretrained weights
6	Finetuning LLMs

Introduction to Large Language Models (LLMs)

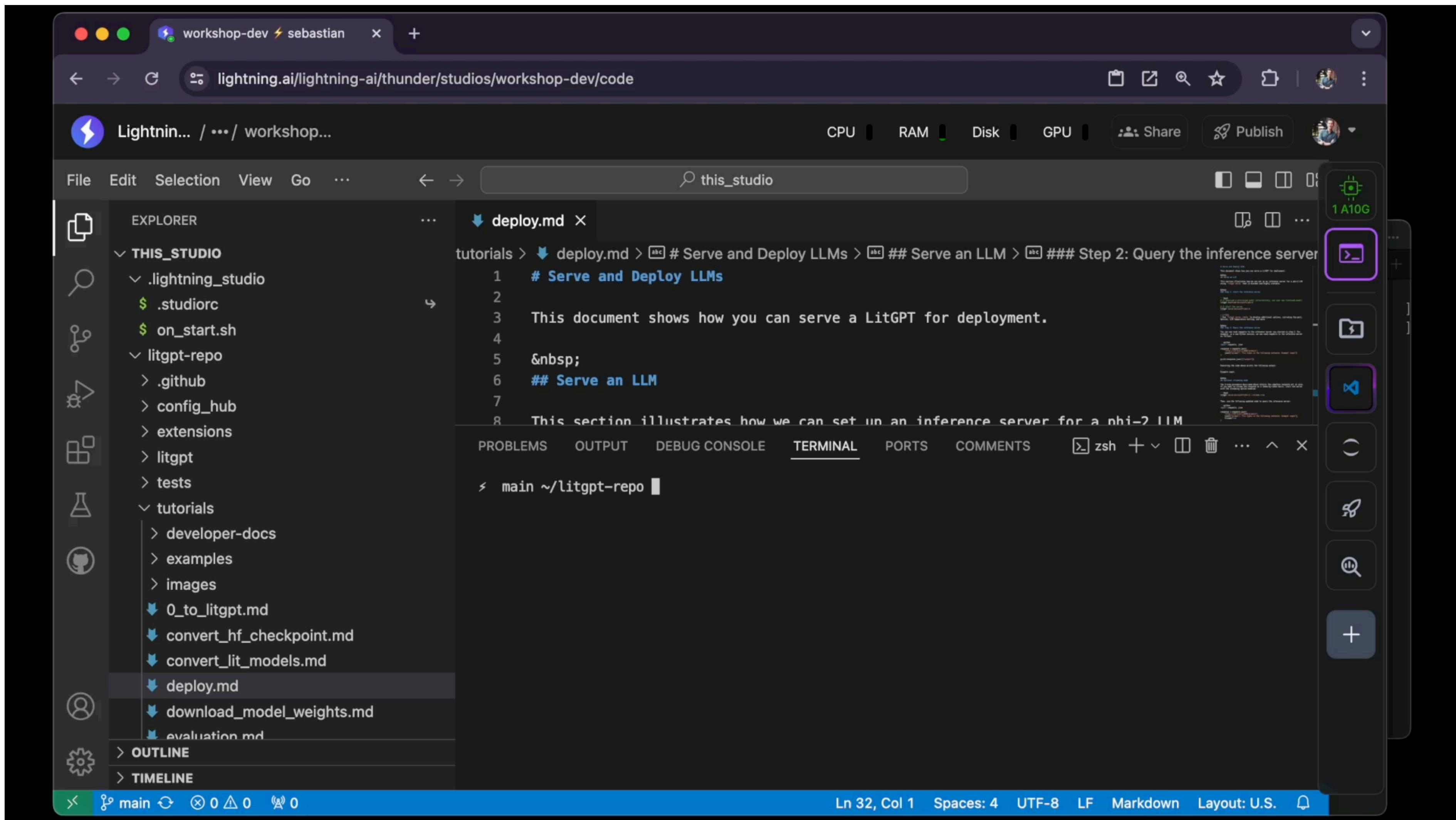
1) Via public & proprietary services



2) Running a (custom) LLM locally

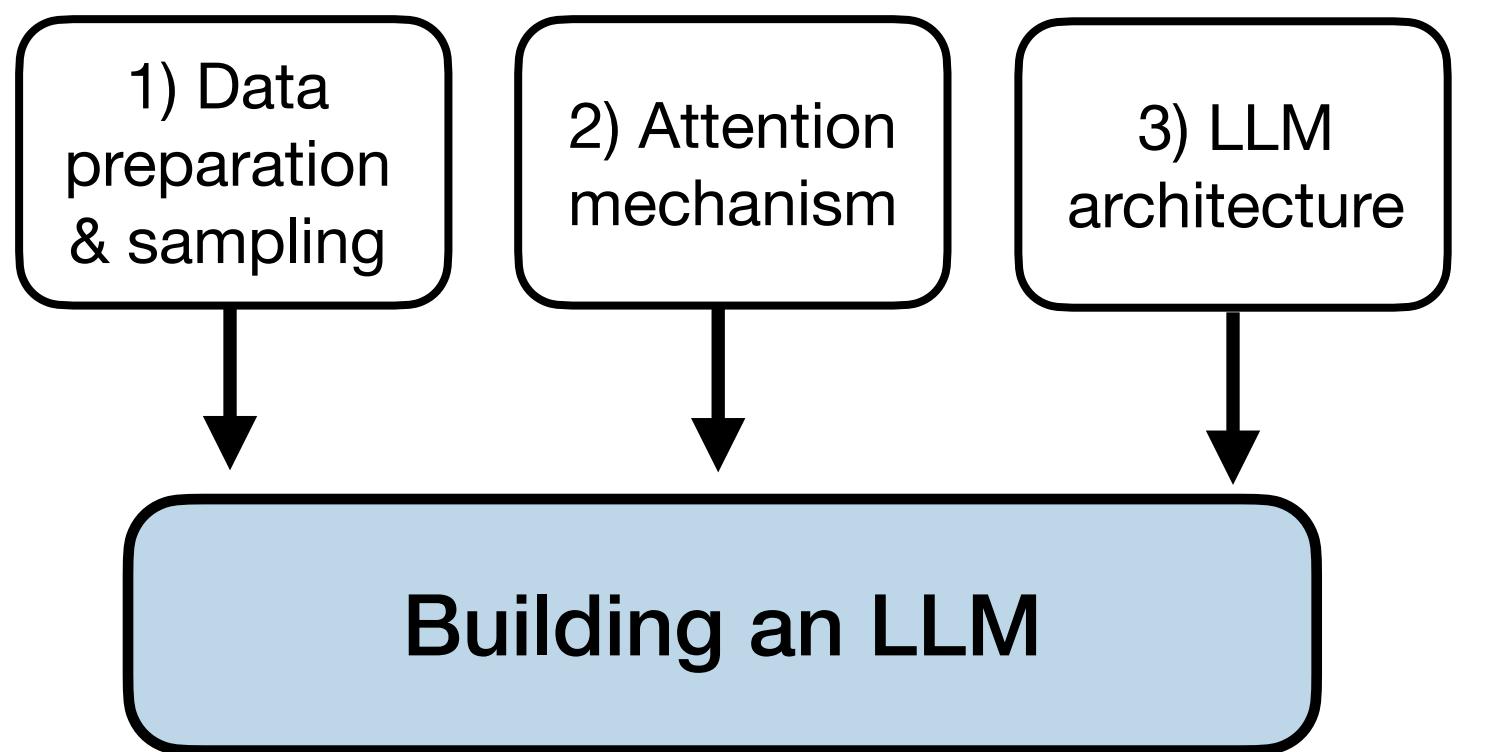


3) Running a (custom) LLM locally

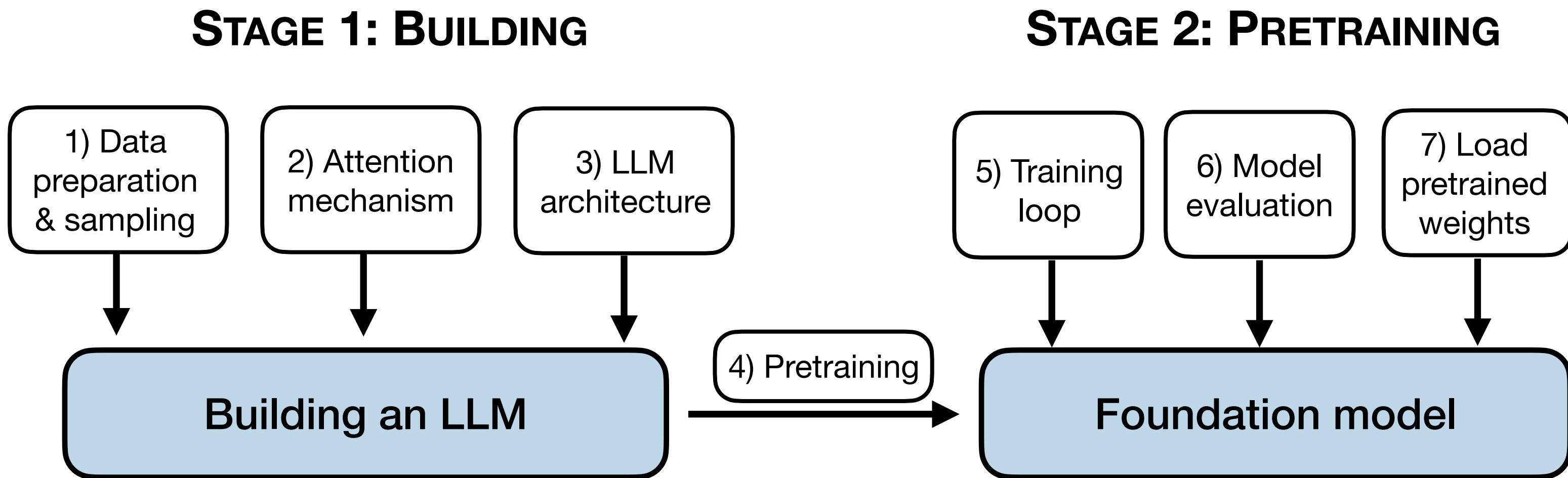


Developing an LLM

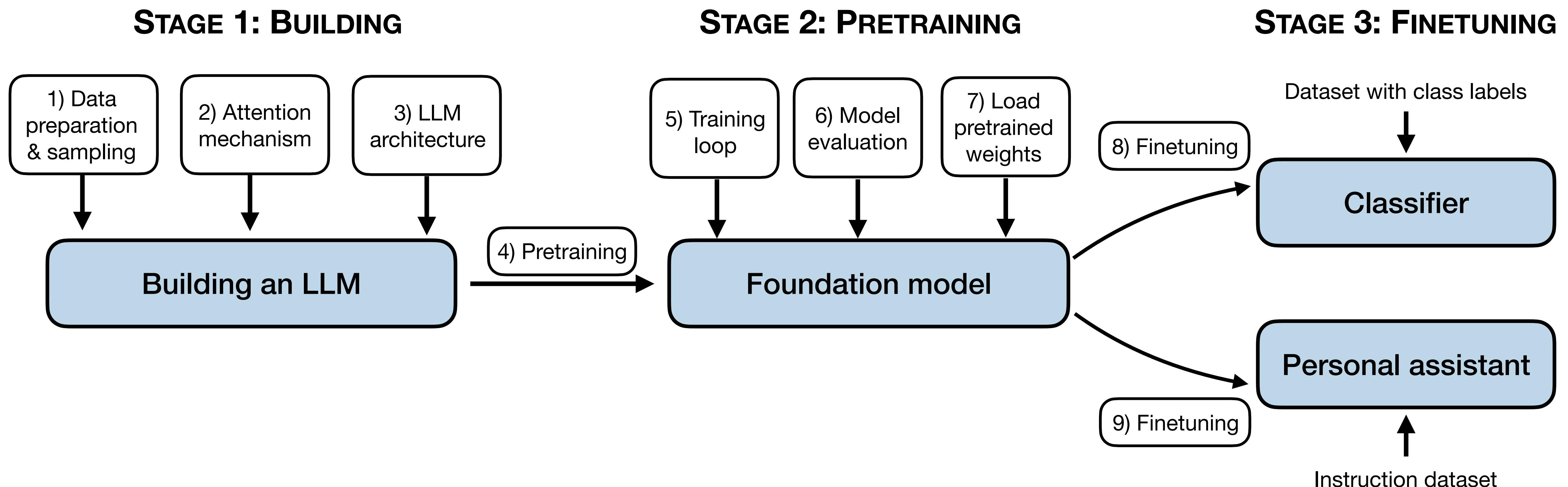
STAGE 1: BUILDING



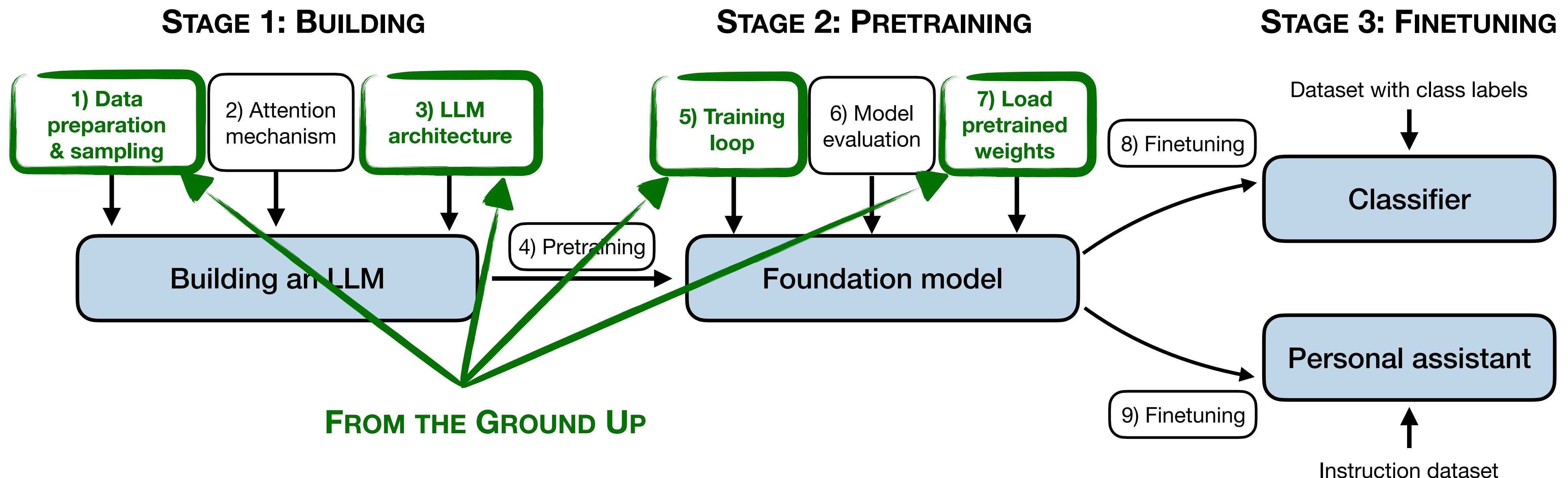
Developing an LLM

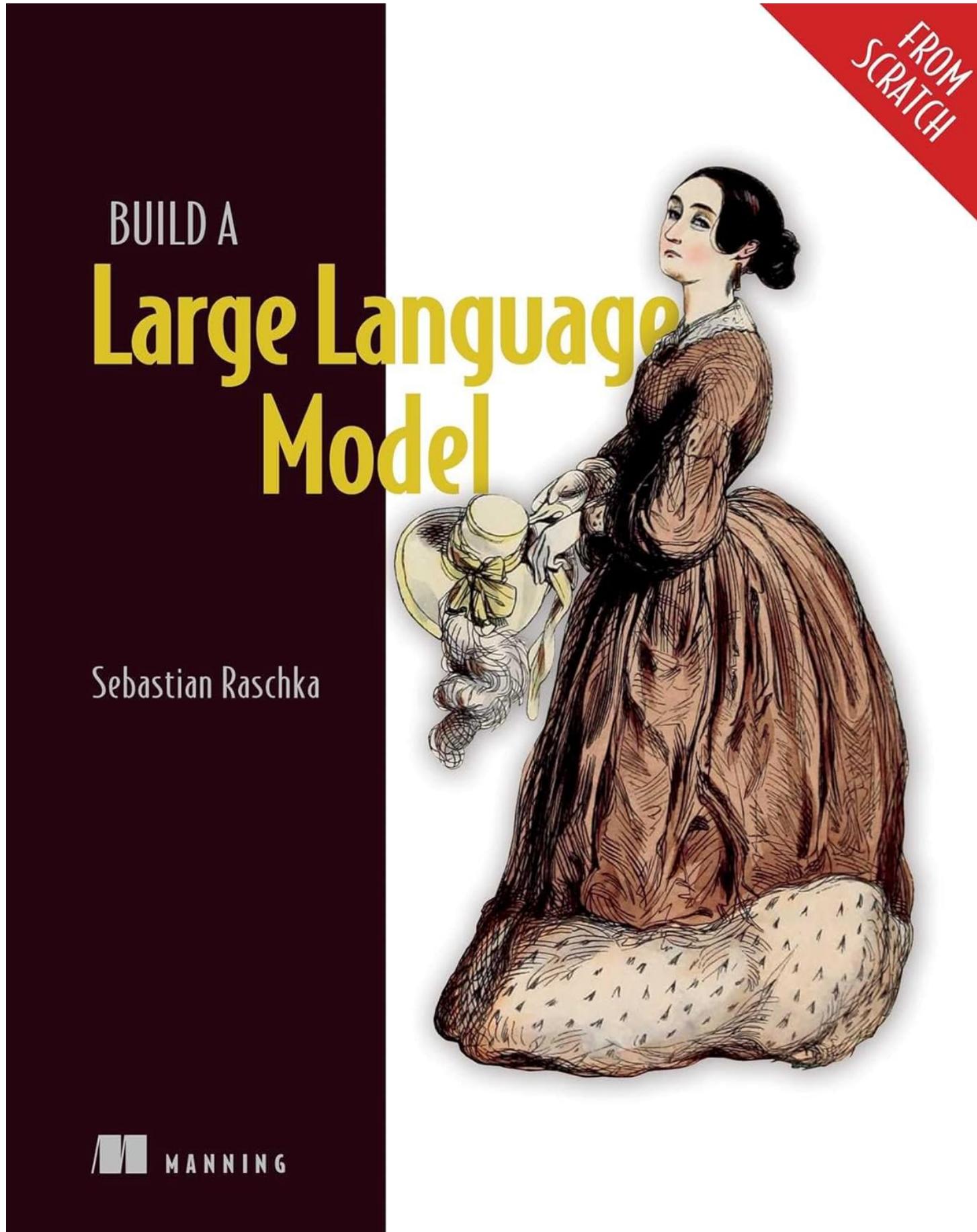


Developing an LLM



Developing an LLM



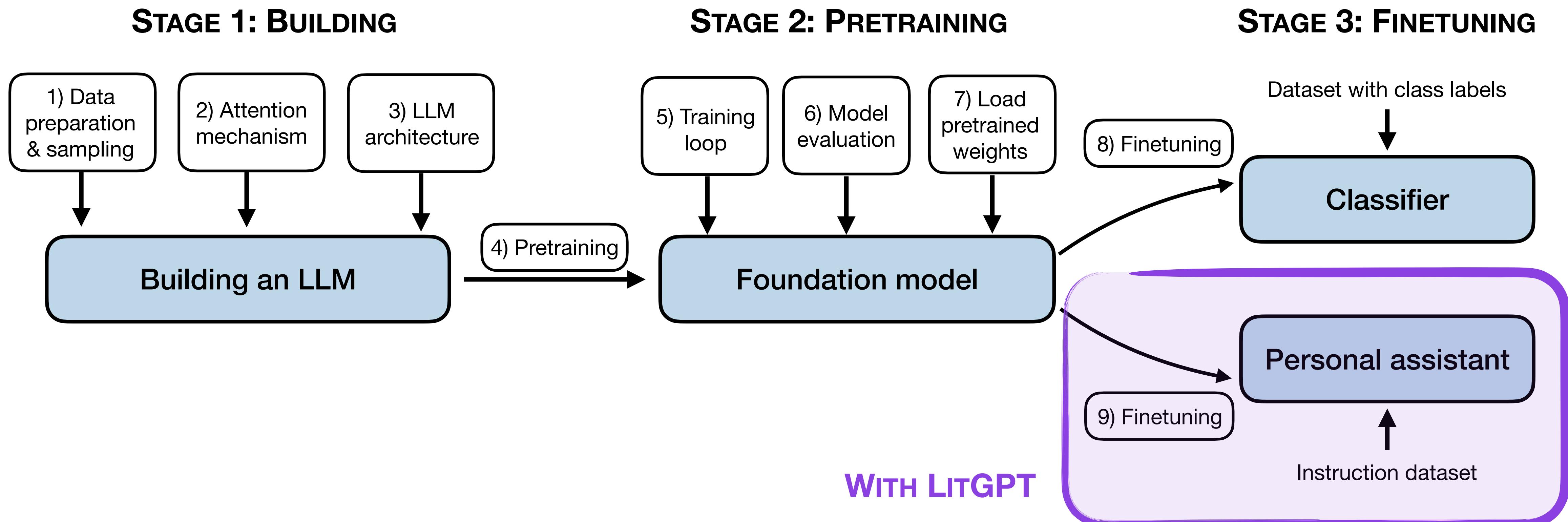


<https://mng.bz/lrp2>

<https://github.com/rasbt/LLMs-from-scratch>

(Source for most figures and code)

Developing an LLM



LitGPT

20+ high-performance LLM implementations with recipes to pretrain, finetune, deploy at scale.

-  From scratch implementations
-  Flash attention
-  Reduce GPU memory (fp4/8/16/32)

-  No abstractions
-  FSDP
-  1–1000+ GPUs/TPUs

-  Beginner friendly
-  LoRA, QLoRA, Adapter
-  20+ LLMs

python 3.8 | 3.9 | 3.10 | 3.11

 CPU tests passing

License Apache 2.0

chat 988 online

[Lightning AI](#) • [Quick start](#) • [Models](#) • [Finetune](#) • [Deploy](#) • [All workflows](#) • [Features](#) • [Recipes \(YAML\)](#) • [Tutorials](#)

 Get started

<https://github.com/Lightning-AI/litgpt>

Links

GitHub: <https://github.com/rasbt/LLM-workshop-2024>

Lightning Studio:

Setup instructions

Lightning Studio (recommended):

Local: LLM-workshop-2024/setup/