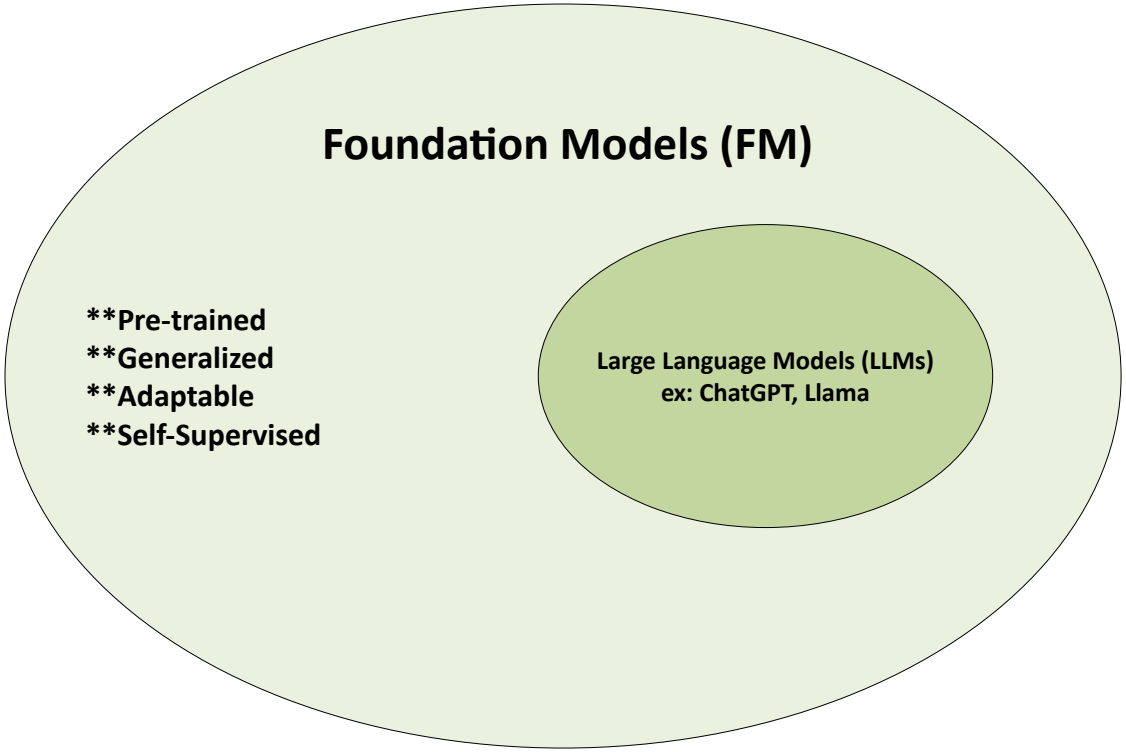
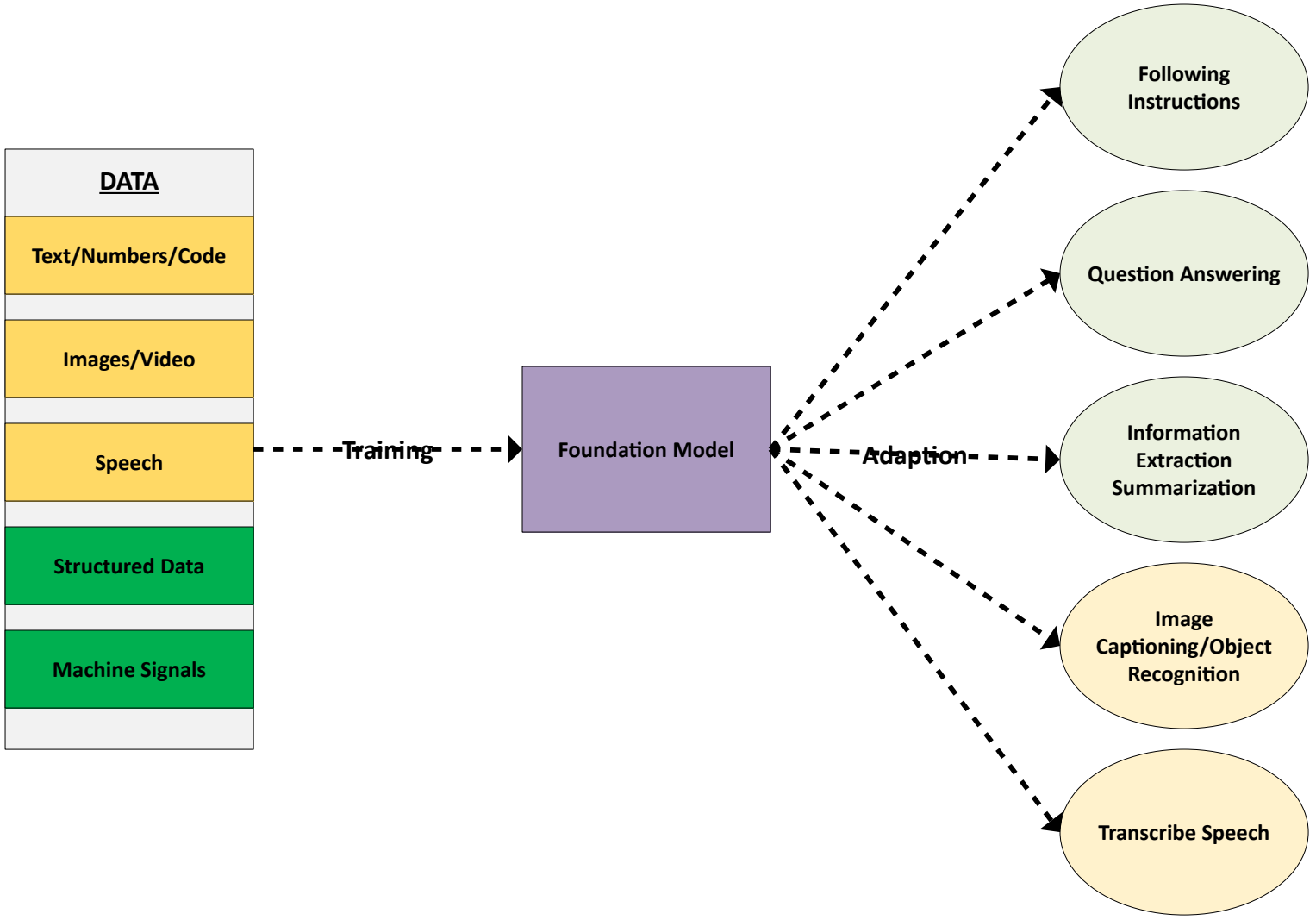
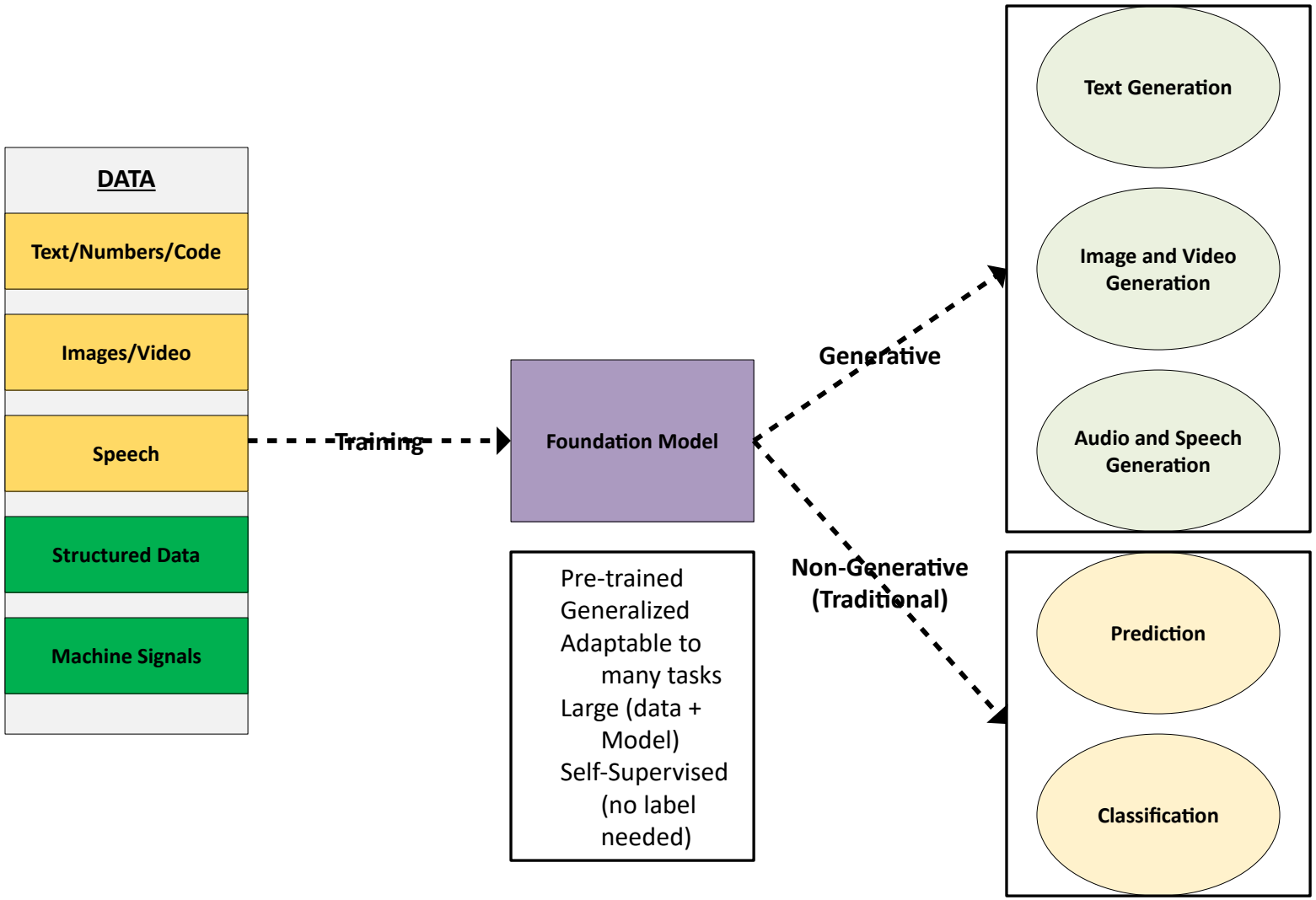
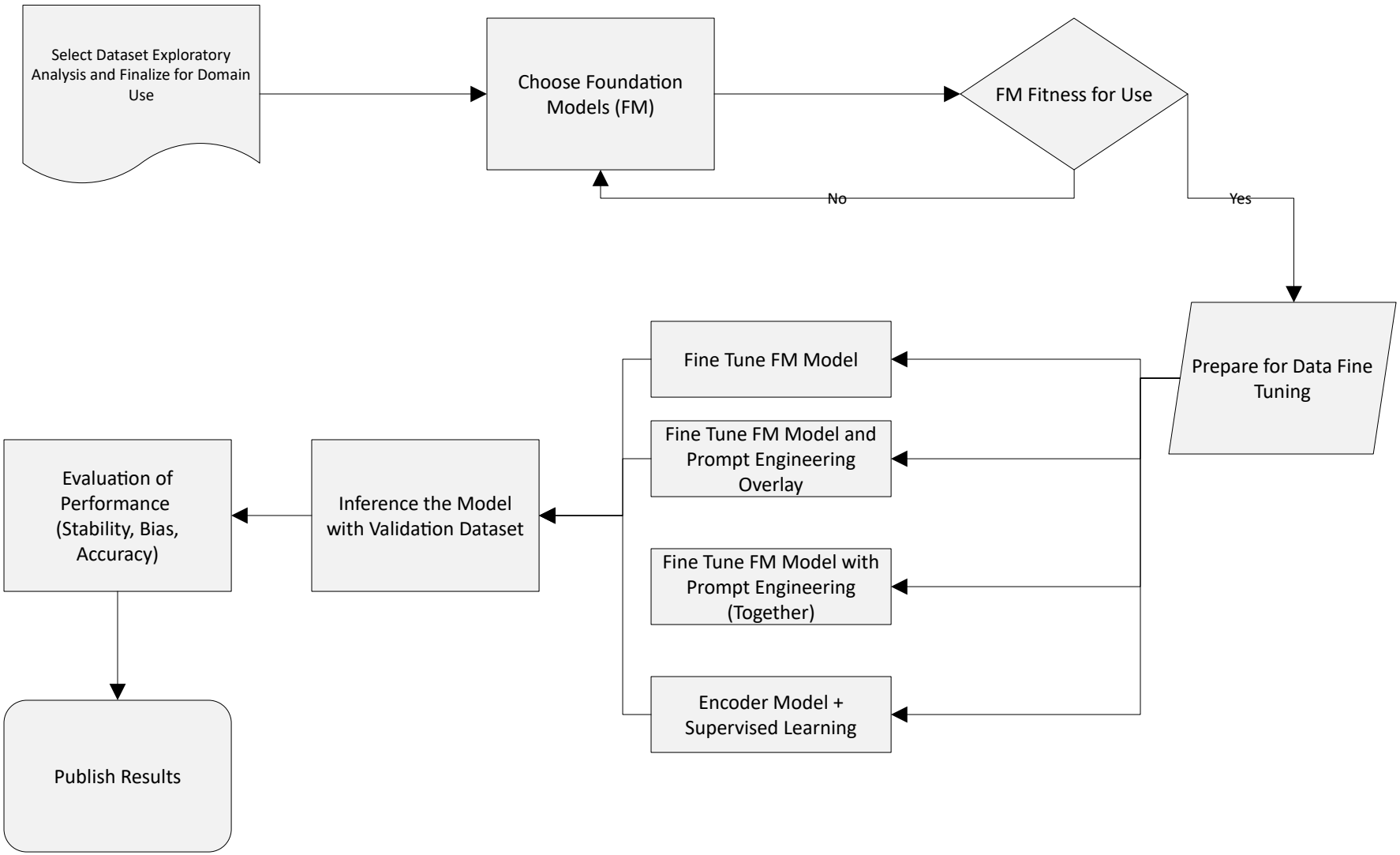


Pre-trained (using large data and massive compute so that it is ready to be used without any additional training)
Generalized — one model for many tasks (unlike traditional AI which was specific for a task such as image recognition)
Adaptable (through prompting — the input to the model using say text)
Large (in terms of model size and data size e.g. GPT-3 has 175B parameters and was trained on about 500,000 million words, equivalent of over 10 lifetimes of humans reading nonstop!)
Self-supervised — no specific labels are provided and the model has to learn from the patterns in the data which is provided — see the cake illustration below.

Foundation Models (in the context of language modeling) are defined as pre-trained Large Language Models, trained on huge amounts of data for multiple tasks. It represents the 'starting point' of modern language modeling. From here we can either fine-tune a model, add prompts, or add a knowledge base. Note, FM's are not all created equal. Some of them could be domain-specific (ex: BloombergGPT by Bloomberg). Even Foundation models trained on similar datasets can also vary in their responses.


DATA
Text/Numbers/Code
Images/Video
Speech
Structured Data
Machine Signals
Training
Foundation Model
Adaption
Following Instructions
Question Answering
Information Extraction Summarization
Image Captioning/Object Recognition
Transcribe Speech


DATA
Text/Numbers/Code
Images/Video
Speech
Structured Data
Machine Signals
Training
Foundation Model
Pre-trained
Generalized
Adaptable to many tasks
Large (data + Model)
Self-Supervised (no label needed)
Generative
Non-Generative (Traditional)
Text Generation
Image and Video Generation
Audio and Speech Generation
Prediction
Classification


Select Dataset Exploratory Analysis and Finalize for Domain Use
Choose Foundation Models (FM)
FM Fitness for Use
No
Yes
Prepare for Data Fine Tuning
Fine Tune FM Model
Fine Tune FM Model and Prompt Engineering Overlay
Fine Tune FM Model with Prompt Engineering (Together)
Encoder Model + Supervised Learning
Inference the Model with Validation Dataset
Evaluation of Performance (Stability, Bias, Accuracy)
Publish Results