What’s Inside There?

Model Metadata and Metrics for AI/ML-BOMs

Sam Washko and Diana Kelley

AI/ML-BOM

AI / Machine Learning Bill of Materials

- Similar concept: S-BOMs and H-BOMs
  - Expand to include ML-specific info
- Provides Data and Model transparency
- Necessary due to the black box nature of ML model files

What’s Inside an AI/ML-BOM?

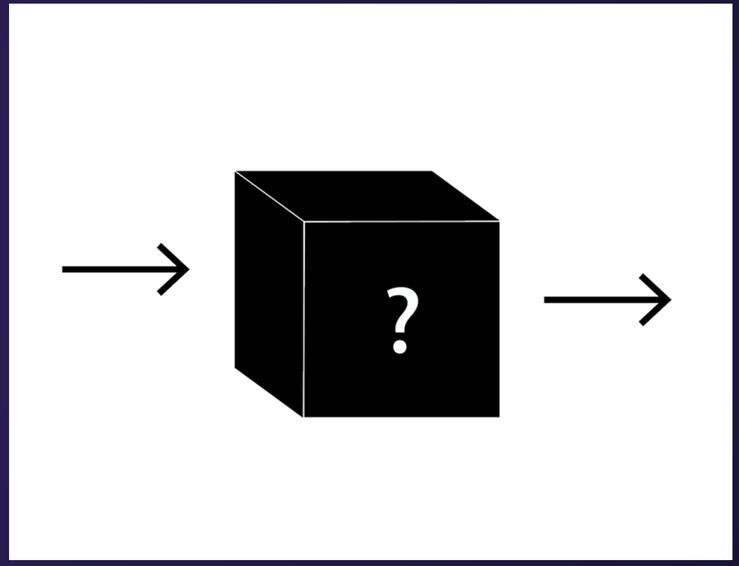
- Supply Chain Dependencies like an SBOM
  - Libraries and Packages
  - Dev Language and Framework

What’s Inside an AI/ML-BOM?

- Model Metadata
  - Type (ex: regression, classification)
  - Source (ex: HuggingFace, Model Garden)
  - Training Data (ex: sources, labels, personnel)
  - Metrics (ex: accuracy, latency)
  - Foundation/Parent Model Lineage
  - Attestations

Black Boxes

- ML models have very little transparency
- Model files consist of
  - tensors
  - model architecture



Black Boxes

- Can’t tell from just a model file
  - Where it came from
  - What code generated it
  - What datasets it trained on
- No logic you can look at for how it makes decisions

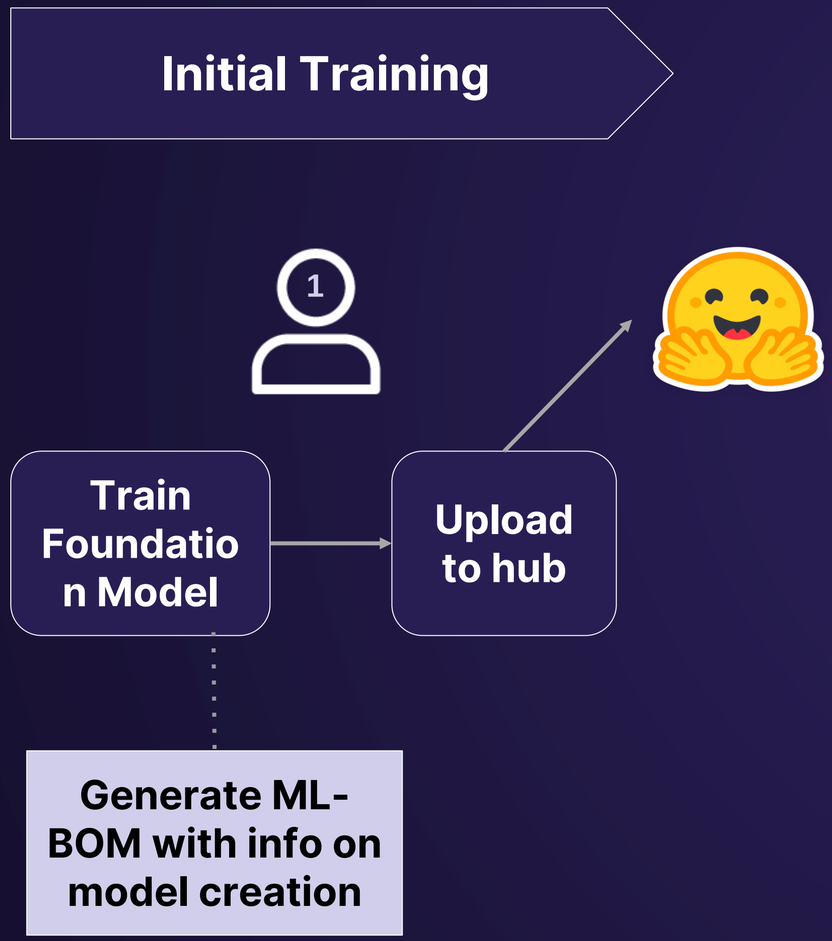
Why do we need metadata?

- Training Datasets
  - check for accuracy
  - relevancy and representation to our use case
- Metrics
  - guard against poisoning
  - guard against drift

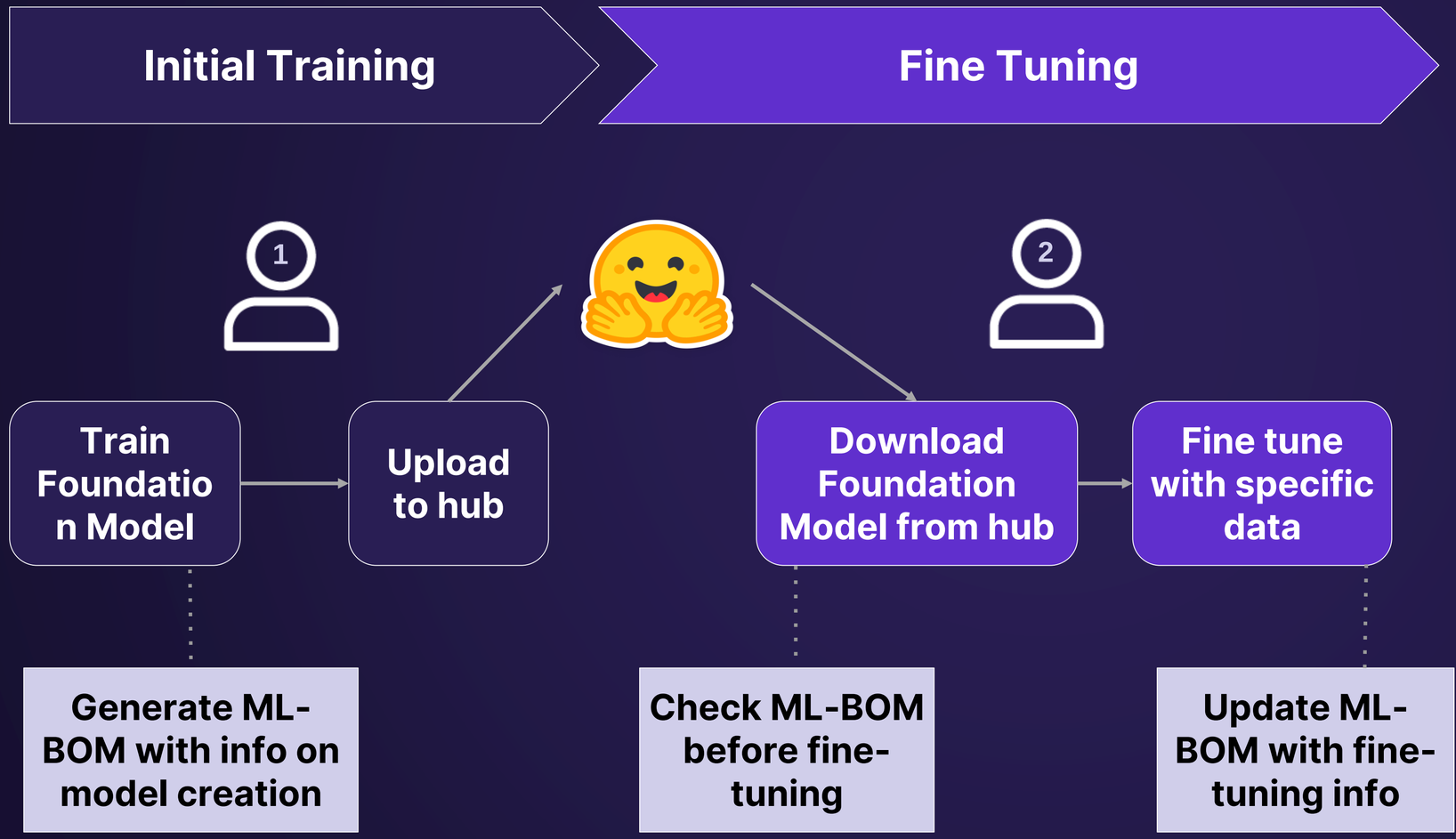
Why do we need metadata?

- Foundation Model Lineage
  - Models increasingly have multiple creators
  - Typically fine tuned from a Foundation Model
- Cryptographic Attestations
  - link to source identity
  - guard against poisoning/tampering

AI/MLBOMs within the ML Workflow



AI/MLBOMs within the ML Workflow



AI/MLBOMs within the ML Workflow