

The Rongotai Model Train Club

ASWF TAC – June 2025

Copyright
Questions

Scaling
Issues

Compliance
Needs

Lack of
Standards

Studio
Concerns

Scarce AI
Resources

Talent
Rights

Research
Focussed

Code
Redundancy

Duty of
Care

Market
Optics

AI
Inaccessibility

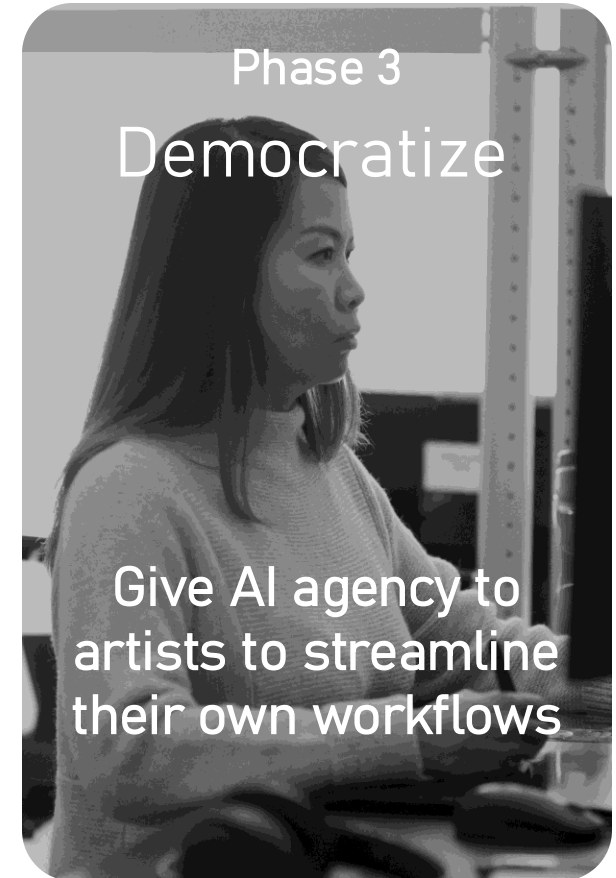
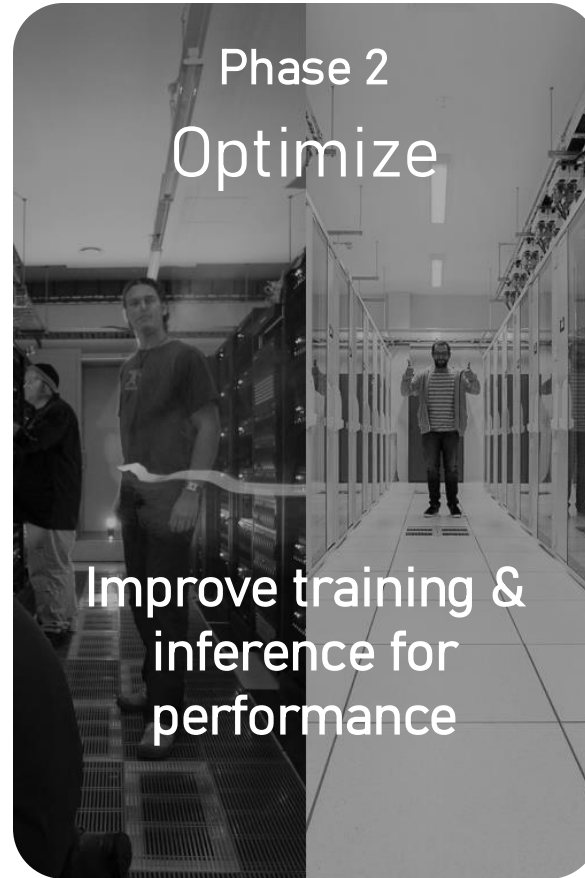
Legal

Scale

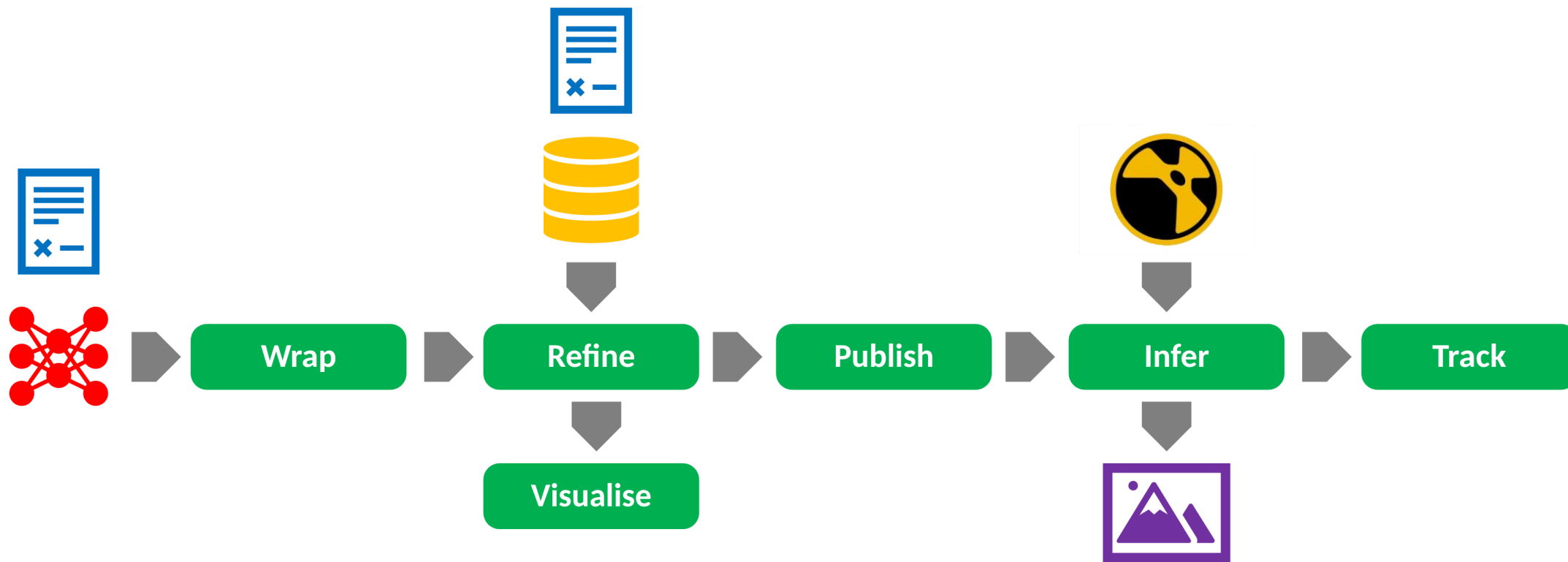
People

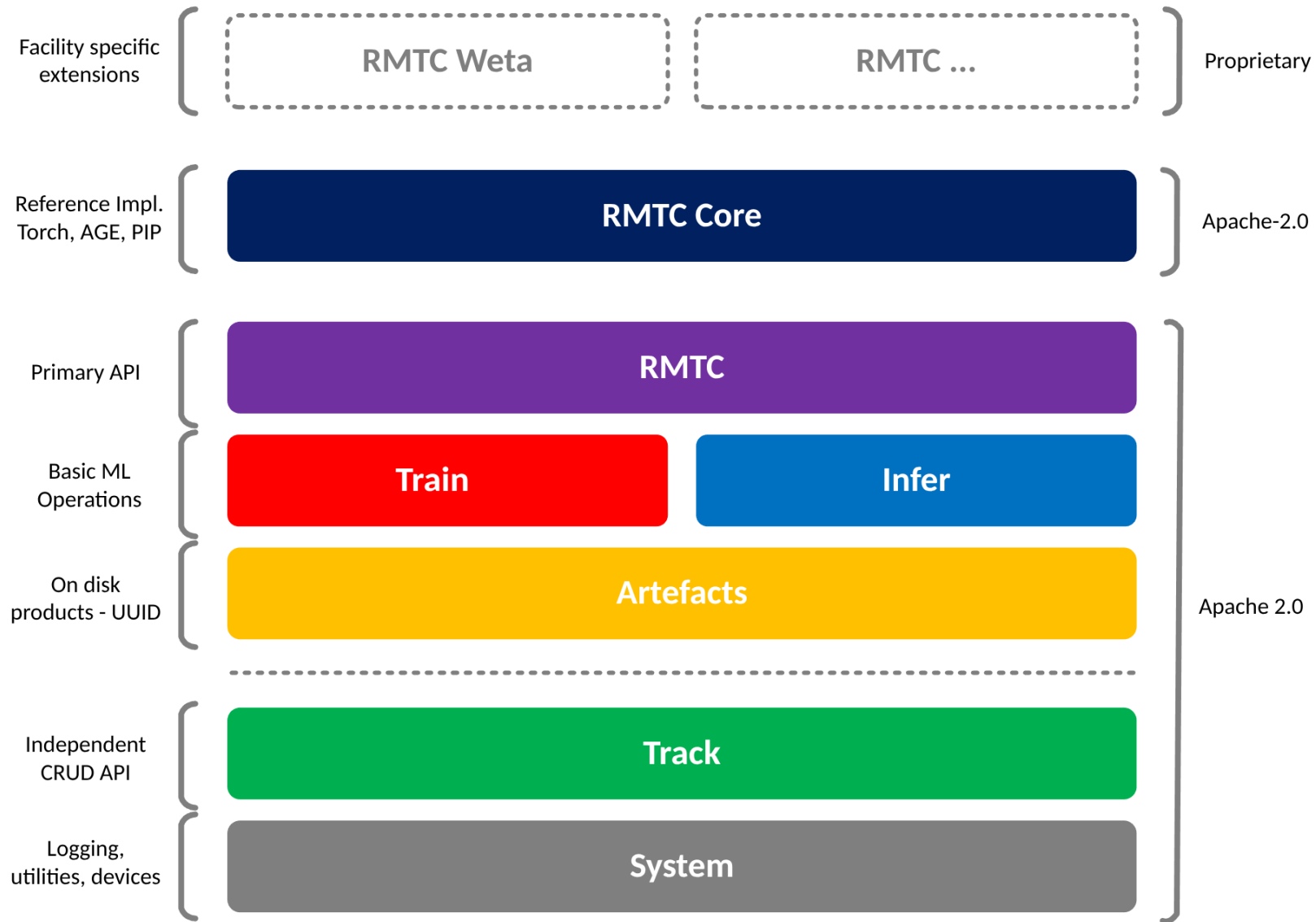
Tooling

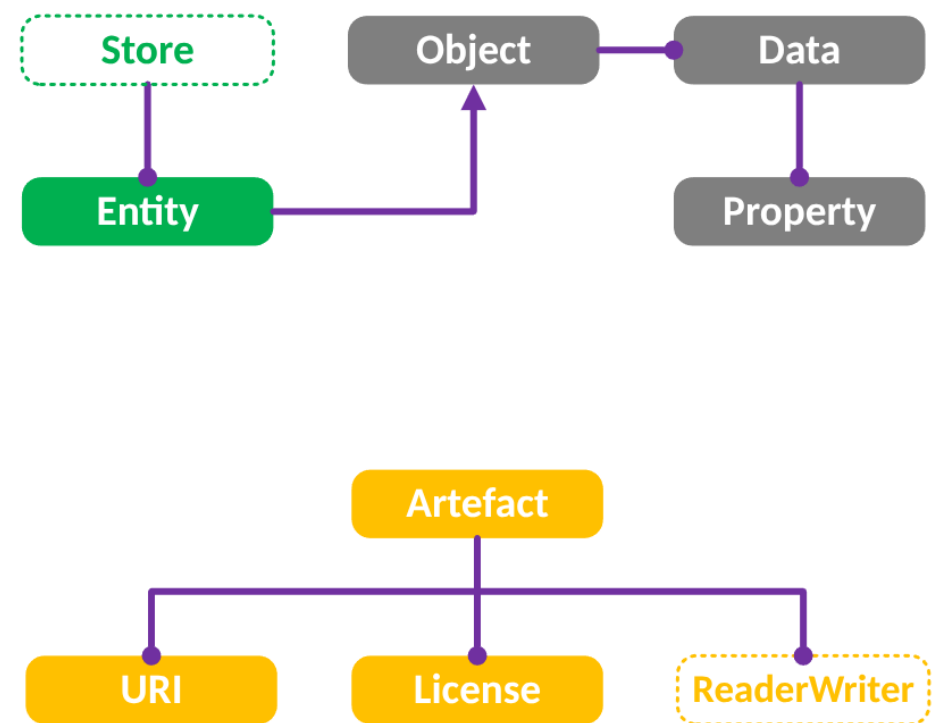
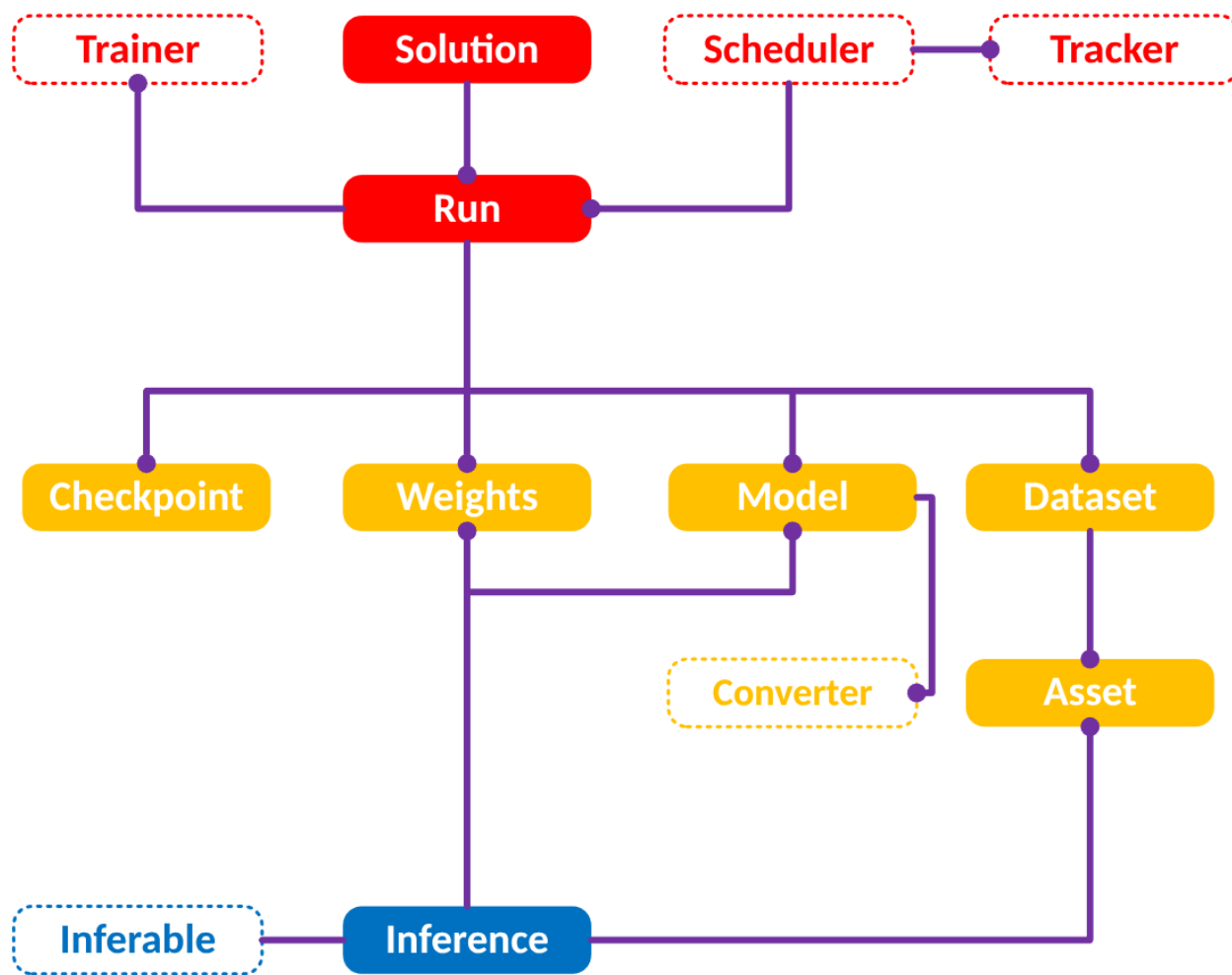
Formalize the process of
training and ingestion by
tracking datasets and models in
a way that respects Artist and
Studio rights & obligations

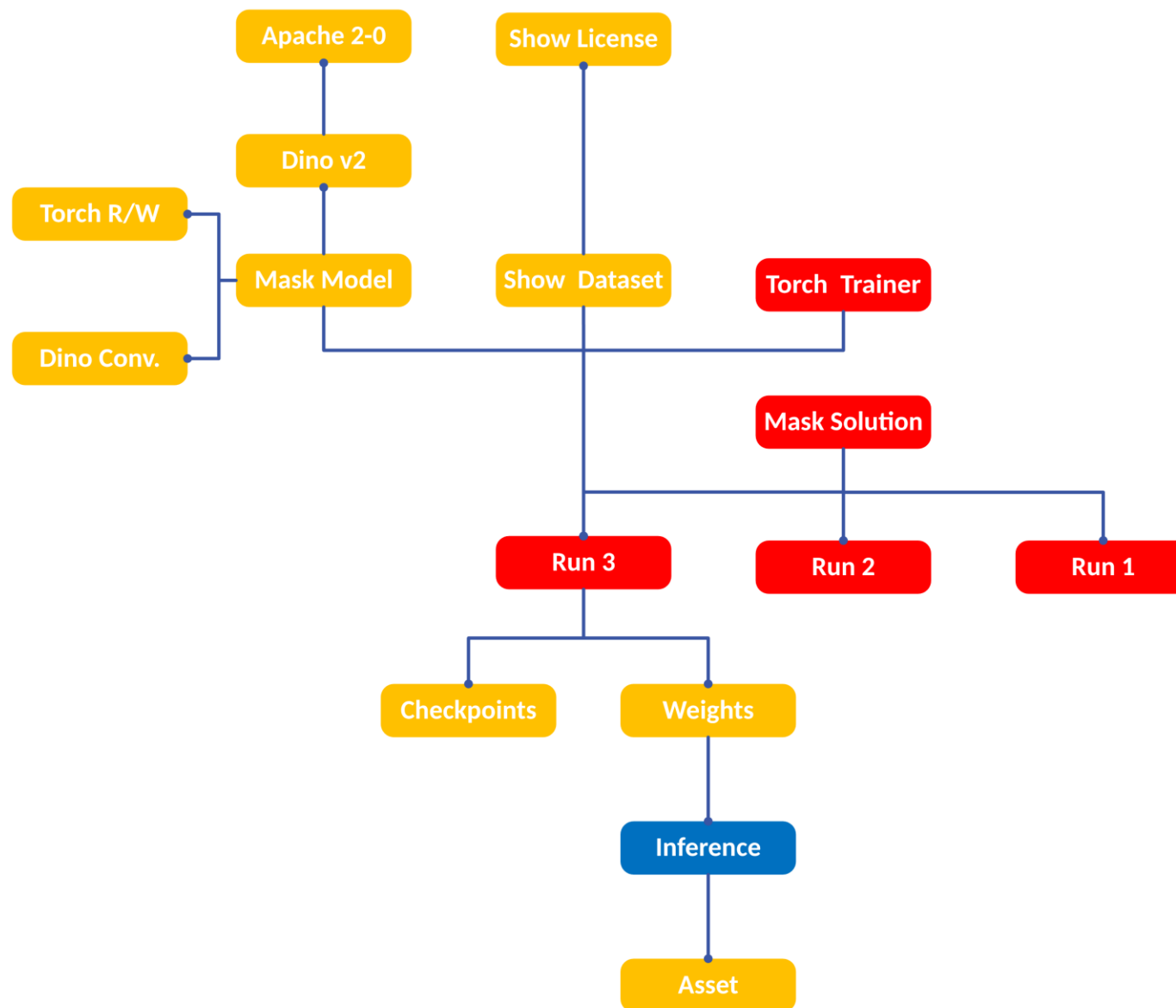


* Phase 2 & 3 are in planning









Scalability

- Shared Service
- Standards
- AI Platform

Extensions

- Schedulers
- Trainers
- Inferers
- Artefact Wrapping

Validation

- API
- Testing
- Stakeholders
- Design
- Use cases

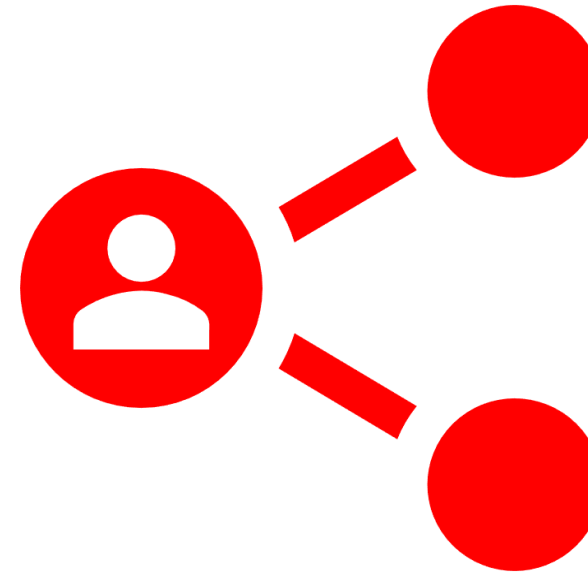
Interfaces*

- Explorer
- Jobber
- Recipes
- Inference

* Phases 2 & 3

Shared Service?

- RMTC as a Service
 - Persistence at the DOM level
 - Job integration
 - Service stack implementation
- Shared artefact repository?
 - Do we consider a ASWF hosted artefact repository?
 - Impacts - facility agnostic asset UUIDs
 - Shared training? Even more ambitious



Standards

- Asset Tensors?
 - HWC or CHW images? PyTorch or cache coherency?
 - Training agnostic tensors?
 - Safe tensors?
- Inference Time Assets?
 - DCC memory mappable assets
 - Runtime converters
- Other standards to consider?
 - C2PA provenance embedding?



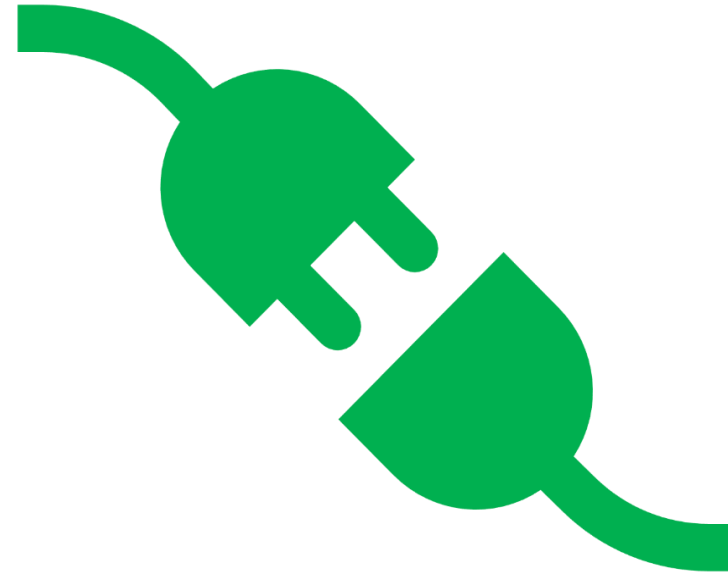
AI Platform?

- Do we extend VFX Platform?
 - AI projects are currently 'off structure'
 - Separate or as part of VFX Platform
- Training, inference & utilities
 - PyTorch – training
 - ONNX – inference
 - CUDA – infrastructure
 - CuPy – GPU Numpy?



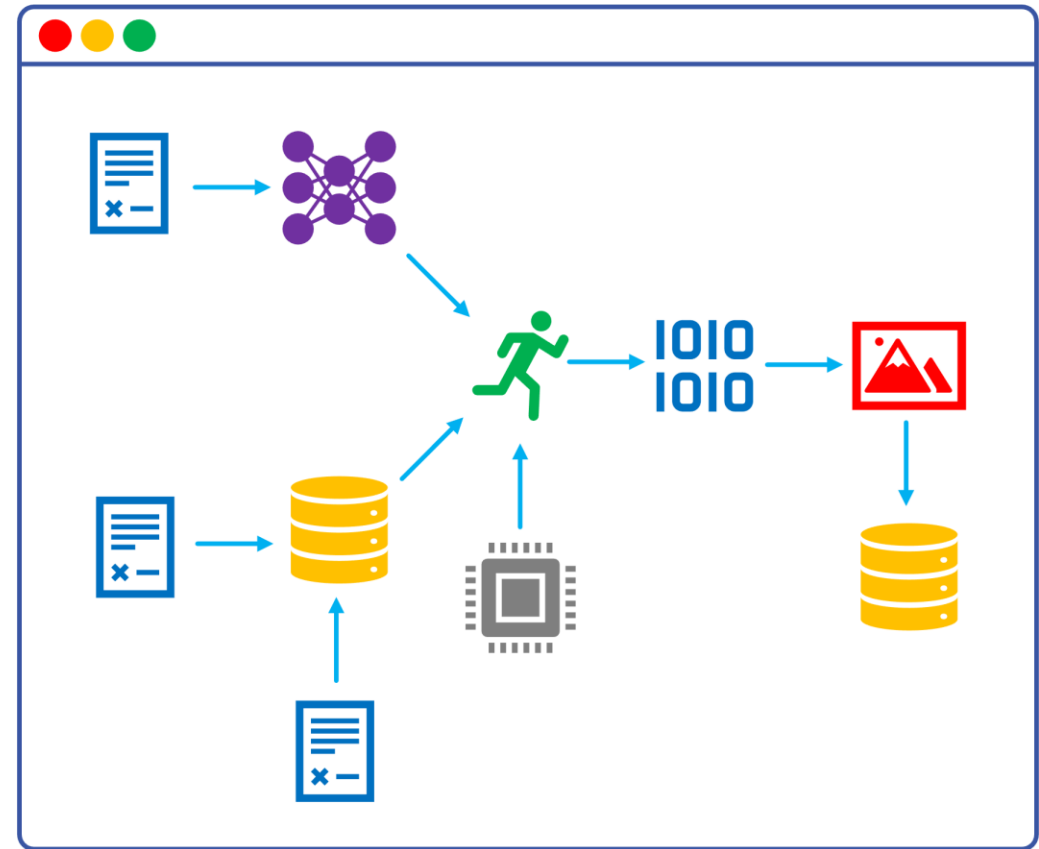
Extensions

- Additional training strategies
 - More than 1:1 correlation models
- Renderwall training scheduler
 - Concurrent w/ weights fusion
 - OpenCue / Plow
- Inference plugins for DCCs
 - Nuke
 - Maya
- Wrapping
 - Common foundational models
 - Common datasets



Explorer

- GUI for exploring artefacts
 - Track provenance
 - Datasets, models & inferences
 - Ingested model repository
- Visual tracing
 - Licensing
 - Inference -> Weights -> Run
 - Create reports
- Invocation system to trigger tools
 - Dataset curation
 - Training recipes & job management
 - Inferencing



Question	Impact	Priority	Details
How do we collaborate?	High	Critical	<ul style="list-style-type: none"> Do we share datasets & training infrastructure? Should we consider an ASWF hosted model repository? How to factor in the differing focus of each facility?
How to futureproof during such rapid change?	High	High	<ul style="list-style-type: none"> Do we continue training at scale? All in PyTorch or will it be superseded soon? What about zero shot models? Do we focus on tracking ingestion & refinement?
Implications of scale on design?	High	High	<ul style="list-style-type: none"> Infrastructure requirements / hardware requirements? How best to organize the system to accommodate external systems like AWS? Dataset acquisition and sharing?
Depth of tracking?	High	Medium	<ul style="list-style-type: none"> How far do we want to track? Every inference? What about dataset transformations that transform licenses? Currently – track everything
Do we want to consider inference tooling?	Medium	Medium	<ul style="list-style-type: none"> Do we focus on just training and tracking or widen into inference time How to integrate into DCCs? Comfy competitor? Connect to C2PA tracking

More to come...

Q&A