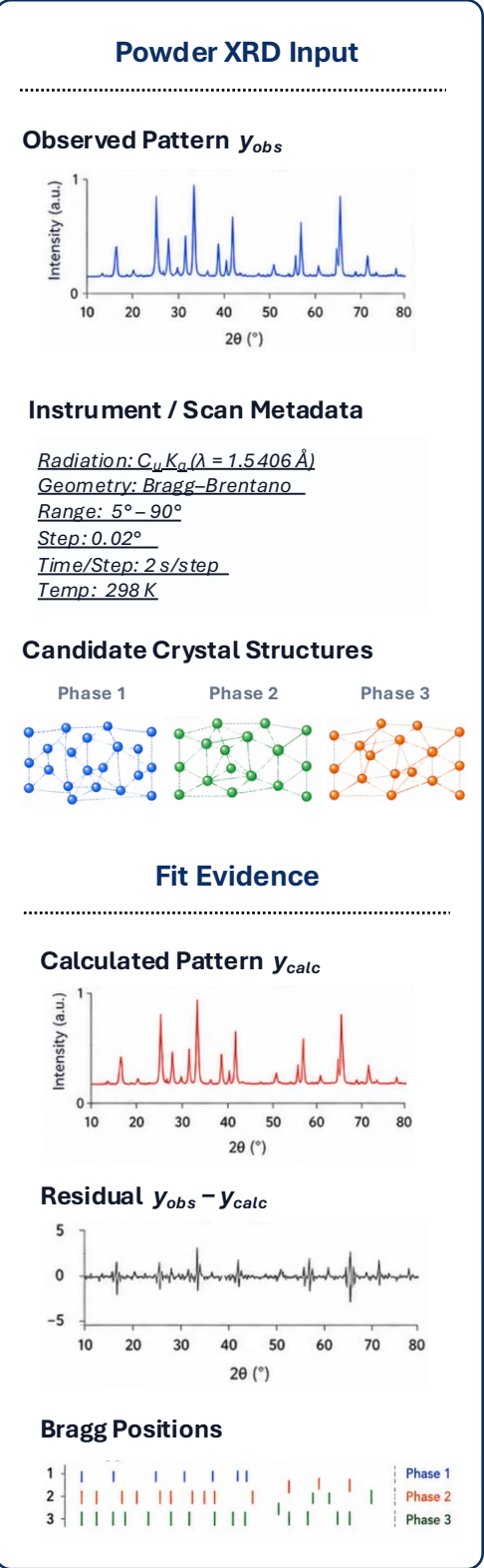
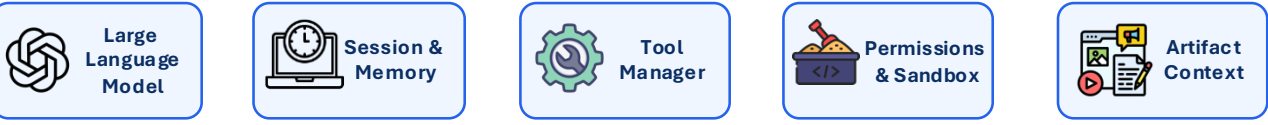


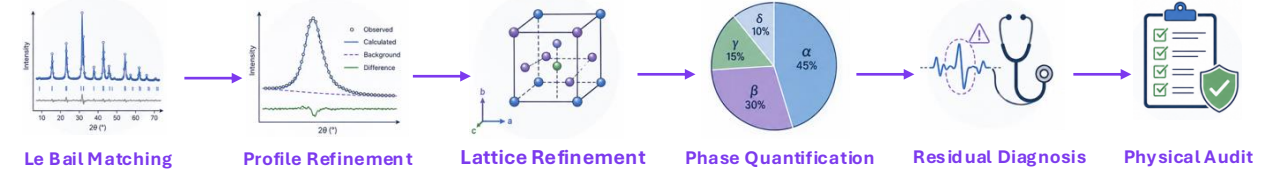
AutoXRD



① AutoXRD Agent Runtime

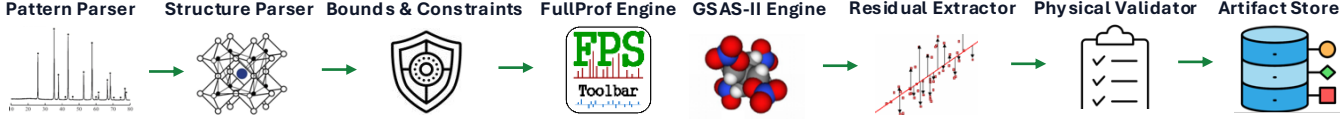


② XRD Skill-Guided Planning – 10 Core Skills

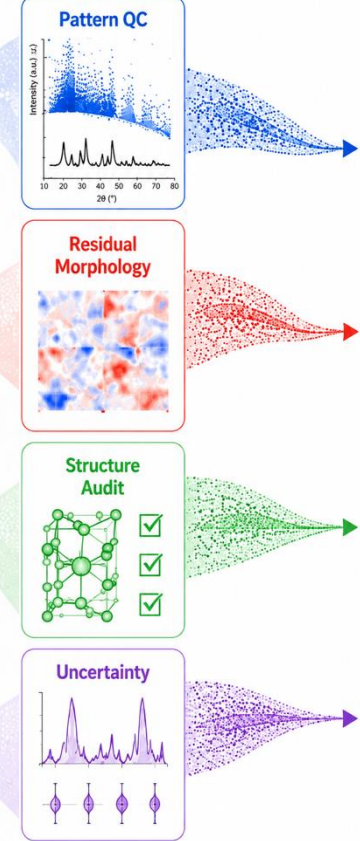


Scientific Plan / Executable Workflow
{stage, parameters/code, bounds, evidence, prediction}

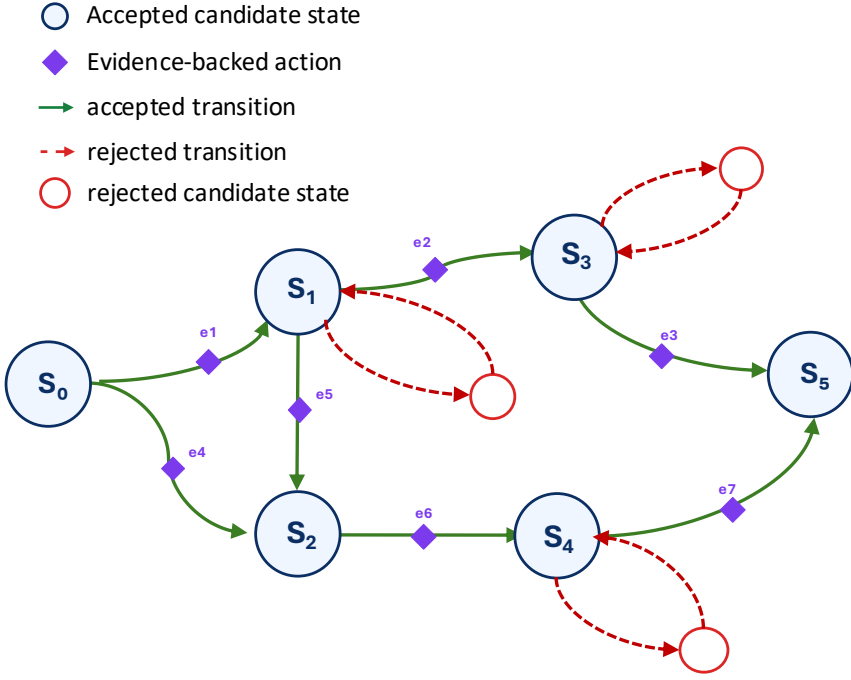
③ Scientific Action Execution Layer



Evidence Extractors

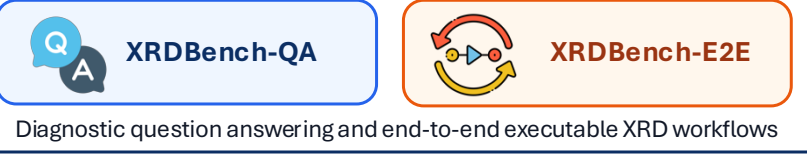


Evidence-Gated Refinement Graph



XRDBench

Evaluation protocol



XRDBench-QA

100 diagnostic tasks

Difficulty split



Maximum Steps ≤ 20
per case · budget enforced · logged

Evaluation criteria



XRDBench-E2E

34 End-to-End Powder-Diffraction Tasks

Maximum steps ≤ 50

Execution and physical gates

Artifact Validity 15% · Hard Metrics 55% · Judge 30%

Evaluation

