SIIM CAIMI26
Conference on Artificial Intelligence in Medical Imaging
OCT 26-27 | University of Pennsylvania, Philadelphia, PA

Open-Source Artificial Intelligence Orchestration for Dual Pediatric Imaging Pipelines: Design, Deployment and Five-Month Operational Experience

Amol Sinha, MS, Stanford Medicine Children's Health; Arogya Koirala, MS; Timothy Mak, MS; Ryan Bollier, EMBA; Edwin Leon; Linn Dee Doyle; Sergios Gatidis, MD; Shreyas Vasanawala, MD, PhD

Introduction/Background

Radiology AI models perform strongly in research, yet clinical workflow integration remains challenging. Deployment requires DICOM routing, task orchestration, structured output delivery, failure recovery, and monitoring: infrastructure underrepresented in literature and typically addressed through commercial platforms. We hypothesized that a platform built from open-source components could reliably deliver AI results end-to-end into the clinical reading workflow, testing this across two structurally distinct pipelines: automated bone-length measurement for extremity radiographs and super-resolution fetal MRI reconstruction.

Methods/Intervention

The platform was assembled from open-source components at an academic children's hospital over 205 version-controlled revisions: (1) Orthanc as DICOM entry point, with a custom Lua routing engine classifying studies into five types via 23 pattern-matching rules across five clinical destinations, with automated retry and state tracking; (2) Mercure for containerized task orchestration with series-level and study-level triggers; (3) a bone-length module performing three-model ensemble inference on eight anatomical landmarks with confidence-weighted fusion, delivering DICOM Structured Reports with 15 coded measurement placeholders for EHR template population and QA visualizations to PACS; and (4) a fetal MRI processor performing DICOM deduplication, parallel super-resolution reconstruction and dual segmentation, routing results to the PACS reading worklist. Monitoring uses Prometheus and Grafana with 10 alert rules and real-time notifications. Outcomes were processing success, end-to-end delivery success, and turnaround time.

Results/Outcome

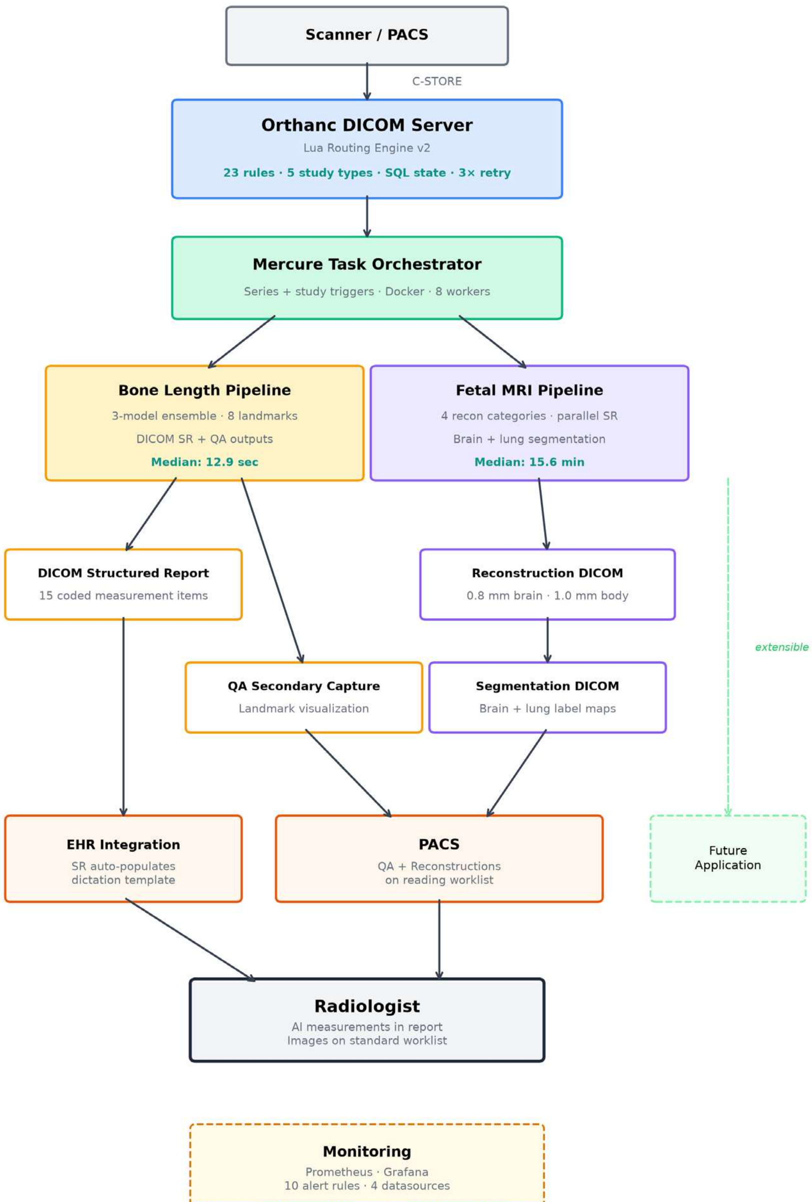
Over five months (February–July 2026), the platform processed 521 clinical tasks (360 bone-length, 161 fetal MRI) across 462 patients. AI processing succeeded in 99.2% (517/521); full end-to-end delivery to clinical systems reached 98.8% (515/521). Root-cause analysis identified two dispatch-layer race conditions where processing succeeded but delivery failed, three container exit-code mismatches after successful reconstruction, and one misrouted study. Median turnaround was 12.9 seconds (IQR 12.2–13.9) for bone-length and 15.6 minutes (IQR 9.8–27.9) for fetal reconstruction. Steady state throughput approximates 130 tasks monthly.

Conclusion

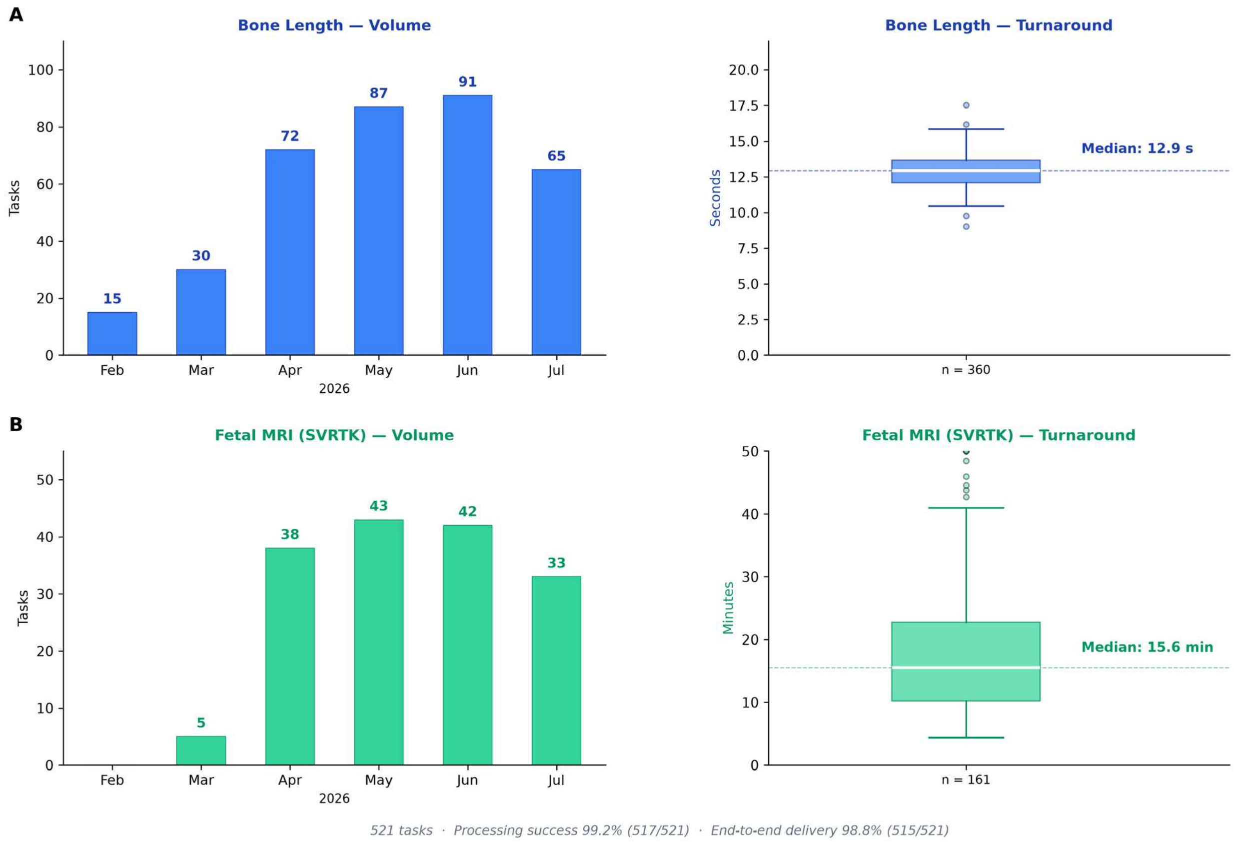
An open-source vendor-neutral platform reliably sustained two structurally distinct AI pipelines in pediatric radiology, delivering results into the reading workflow with 98.8% end-to-end success across 521 tasks over five months.

Statement of Impact

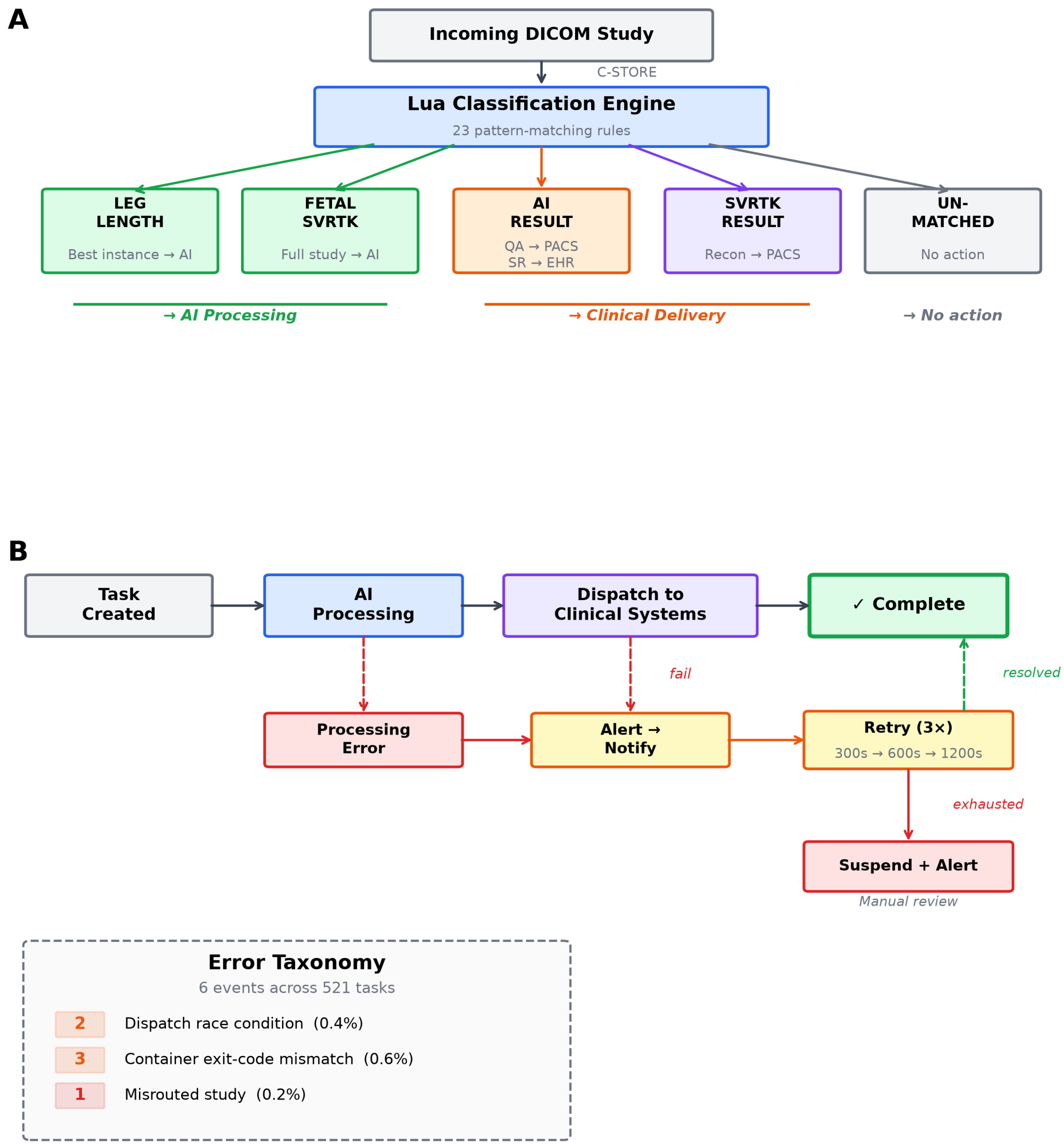
Clinical AI deployment does not require commercial platforms. A single orchestration architecture handled a sub-15-second inference task and a multi-minute reconstruction, suggesting generalizability across AI tools and clinical settings.



Architecture workflow of the platform. Studies route via DICOM C-STORE, classified by the Orthanc Lua routing engine. Processed and dispatched through Mercure to containerized AI pipelines. Bone leg length results are delivered as DICOM Structured Reports to the EHR for template population, with QA visualizations to PACS. Fetal MRI reconstructions and segmentations route to the PACS reading worklist.



Processing volume and turnaround over five months of clinical deployment. (A) Monthly task volume by pipeline (bone leg length since February 2026, fetal MRI since April 2026). (B) Turnaround distributions: bone leg length median 12.9 seconds (IQR 12.2–13.9, n = 360), fetal MRI median 15.6 minutes (IQR 9.8–27.9, n = 161). Total: 521 tasks across 462 patients and counting.



Study routing classification and task lifecycle with failure recovery. (A) Lua routing engine classifies incoming studies into five types across AI processing, clinical delivery, and unmatched categories. (B) Task lifecycle from creation through processing (99.2%) and dispatch (98.8% end-to-end). Failed dispatches retry with exponential backoff. Error taxonomy shows root-cause analysis of all six errors across 521 tasks.

Keywords

AI deployment; DICOM orchestration; clinical integration; pediatric radiology; open-source infrastructure; structured reporting