

Background

Overview / Motivation

- **Project 1 (Digital Human Pipeline Automation):** Internal digital-human creation relied on multiple scattered tools and workflows. A key bottleneck was slow chroma-key iteration, where “true” video preview previously required external scripts and minutes of turnaround, slowing operator feedback loops.
- **Project 2 (Low-Resource ASR for Radio Compliance):** The client’s radio recordings are difficult for off-the-shelf ASR due to domain constraints and severe hard clipping, motivating a data-centric and domain-adaptation approach.

Objectives

Project 1

- Enable fast, in-app chroma-key preview to support rapid parameter iteration.
- Consolidate Gradio-based tools into a single entry point with fault isolation.
- Dockerize the pipeline management scripts and provide a GUI for usability.
- Deliver a desktop orchestrator (PLATO) to unify remote services + local tools.

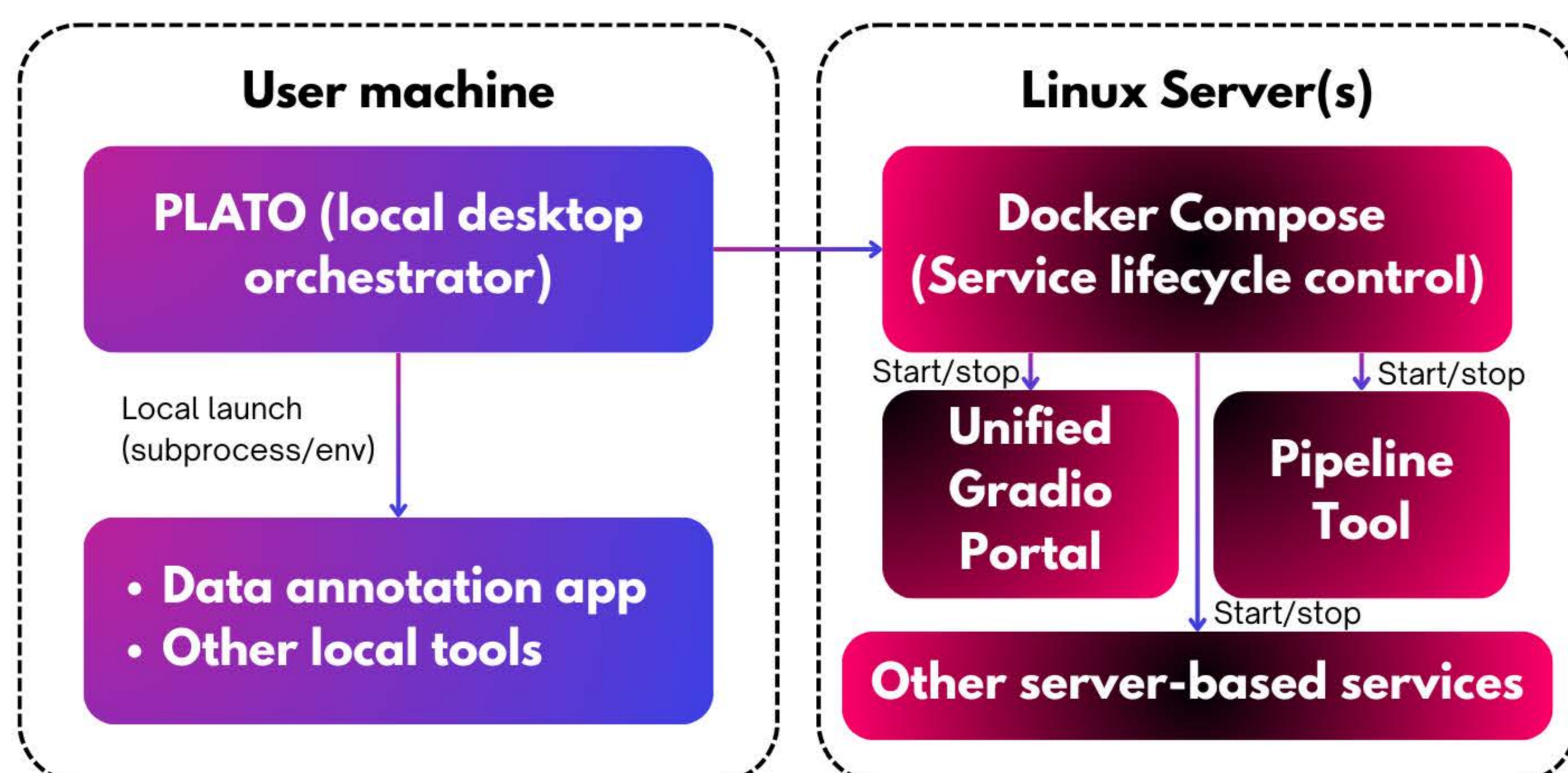
Project 2

- Build a reproducible low-resource ASR experimentation workflow (data packaging, consistent splits, benchmarking).
- Apply parameter-efficient domain adaptation (LoRA) to improve transcription quality under limited labeled data.

Methodology

System Architecture (Project 1)

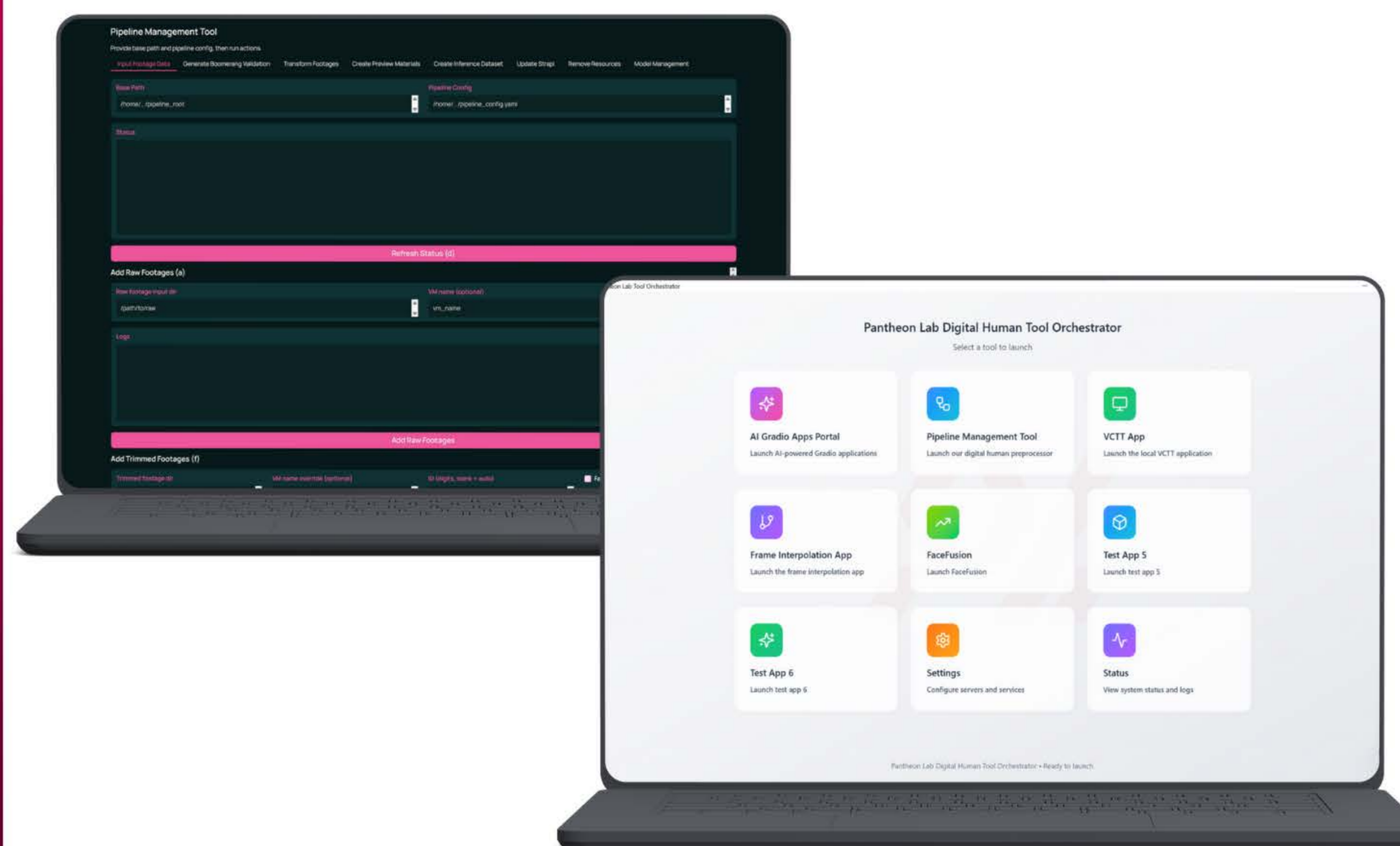
- A pragmatic consolidation approach was adopted: improve critical bottlenecks first, then unify tools via containerization and orchestration into a cohesive toolchain.



Methodology (continued)

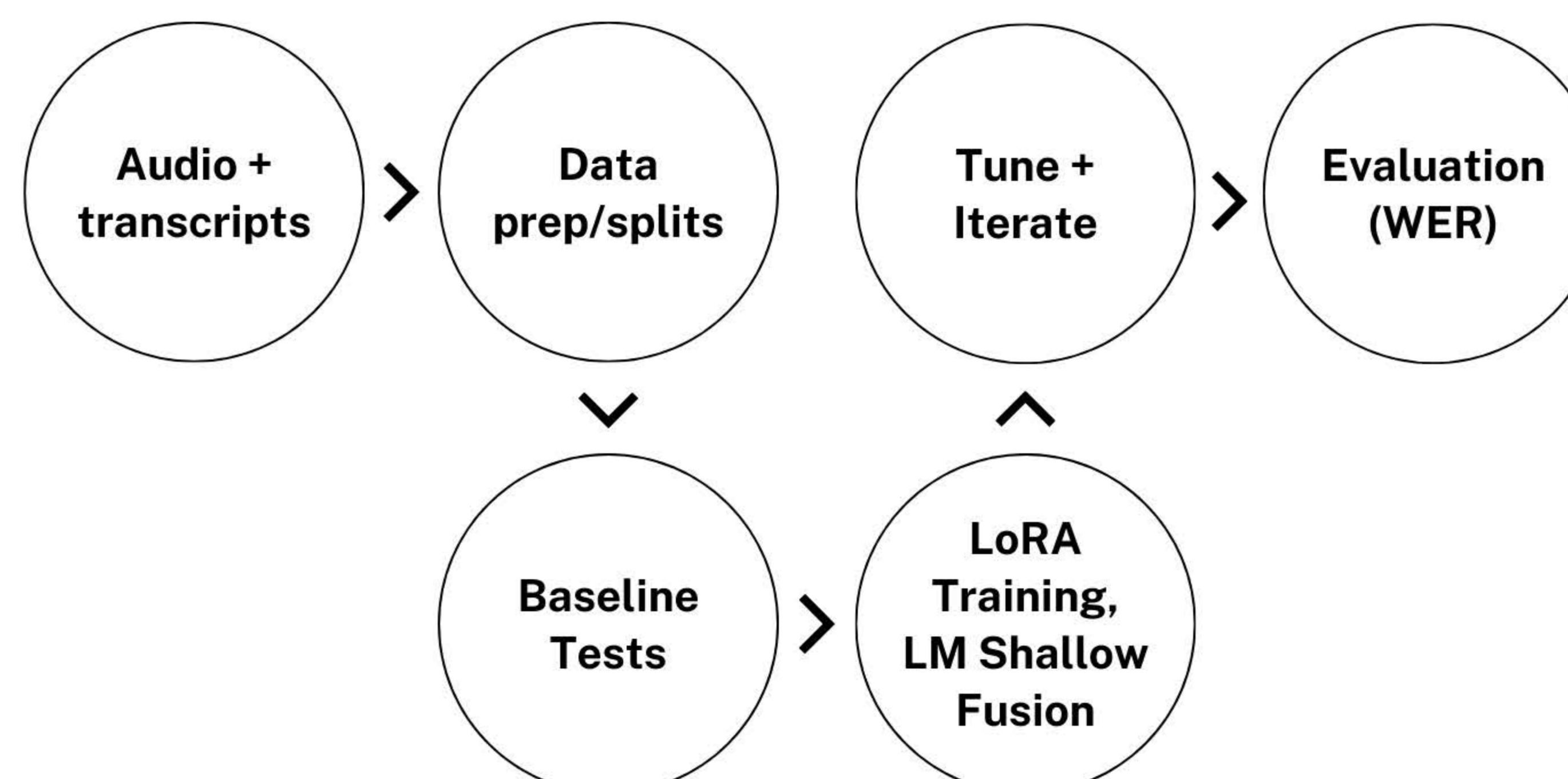
Key Implementations (Project 1)

- **Chroma Key Instant Preview (Annotation App):** Reduced “true preview” turnaround by moving preview generation into the app workflow and optimizing the pipeline for fast iteration.
- **Unified Gradio Apps Portal:** One web entry point for multiple tools
- **Pipeline Management Tool (Docker + Gradio GUI):** Containerized deployment for dependency isolation; GUI layer for clearer workflow execution and usability.
- **PLATO Orchestrator (Local Desktop App):** Built a lightweight desktop orchestrator that serves as a single entry point for digital human creation



ASR Workflow (Project 2)

- Established a consistent train/validate/test loop with WER-based benchmarking, then applied LM shallow fusion & LoRA fine-tuning to adapt Whisper to clipped, domain-specific radio speech.



Evaluation

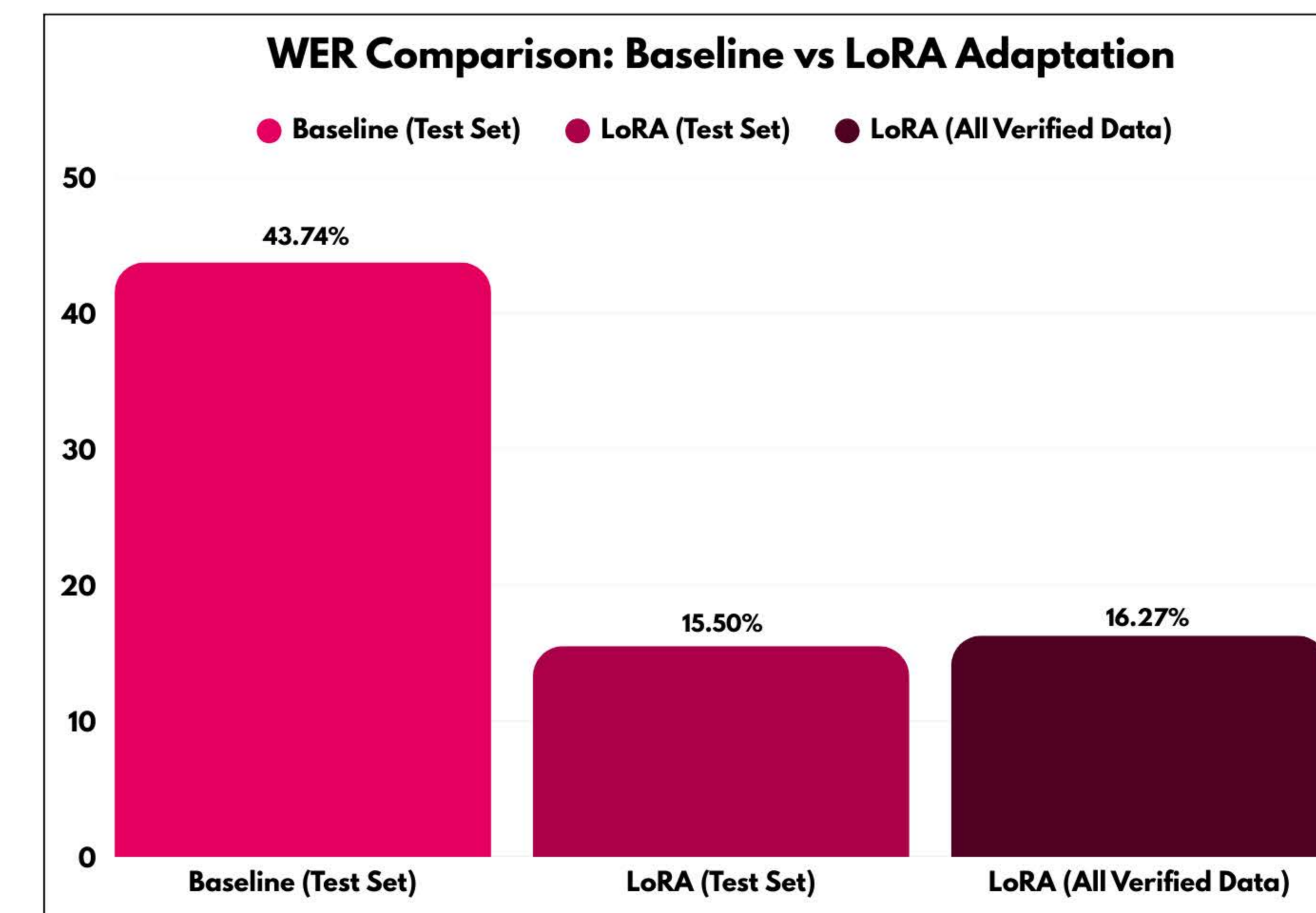
Evaluation Highlights

Project 1

- Chroma key iteration speed: “True preview” reduced from ~2.5–4 minutes to ~20–40 seconds for typical 4K samples (10–20s), enabling ~5–10× faster iteration.
- Orchestrator reduced internal tools setup time from one full day to about 1.5-2 hours.

Project 2

- LoRA impact (test set): Whisper-medium + LoRA reduced WER from 43.74% → 15.50%.
- Broader check: Achieved 16.27% average WER across 53 verified audios, indicating gains beyond a single split.



Conclusion

The work delivered a practical, unified toolchain for digital human creation (preview, portal, Dockerized Pipeline Tool GUI, PLATO orchestrator) and demonstrated that LoRA-based adaptation can yield substantial WER improvements even under constrained and degraded radio audio conditions.

Future Work

Expand cross-platform testing and add automated dependency/driver diagnostics; progress PLATO toward automated server-side provisioning and tackle remaining compute bottlenecks. For ASR, improve data collection/verification and prioritize clipping-aware methods and protocol-aligned evaluation targets.