

REPRODUCIBLE PIPELINES FOR TEAMS READY TO SCALE

# Make important bioinformatics workflows easier to **trust, run, and scale.**

For teams with inherited, prototype, manual, AI-rewritten, academic, or internally patched workflows that now need to support larger studies, more users, customer deployment, clinical validation, or production operations.

## A workflow can be scientifically important and still be too brittle for the next stage.

Manual steps accumulate. Dependencies drift. Edge cases go untested. Results become hard to reproduce. One person knows how to run the pipeline, but the broader team cannot use it confidently. The math may be sound while the workflow around it remains too brittle. Fulcrum validates the workflow, identifies the real failure modes, and builds the automation needed to make it usable beyond the original developer.

### When teams call Fulcrum

- A pipeline runs, but confidence is thin
- Results vary across environments or versions
- Manual steps create hidden failure points
- Documentation does not match the current workflow
- Dependencies are unclear or uncontrolled
- Edge cases have not been tested
- The workflow depends too heavily on one internal expert
- The team needs a reproducible, containerized, or automated version
- A larger study, partner handoff, customer deployment, or review is coming

### What Fulcrum delivers

- Pipeline architecture and code review
- Reproducibility and dependency audit
- Benchmarking against agreed datasets
- Output concordance checks
- Edge case and failure mode assessment
- Test coverage gap analysis
- Runtime and scalability review
- Workflow automation and containerization
- Documentation, runbooks, and handoff
- Prioritized roadmap for what to fix, rebuild, automate, or leave alone

# Proof in practice: Darwin's Ark

Darwin's Ark is a citizen-science project linking feline genetics to traits and behavior, with more than 20,000 cats enrolled. To make low-pass whole genome sequencing viable at scale, the team needed genomic infrastructure that did not yet exist for cats: a larger reference panel, rigorous QC, reproducible variant processing, and an imputation workflow that could support population-scale analysis. Fulcrum helped build the reference panel and the automated pipeline foundation needed to make the project work at scale.

## What we built

- Aggregation and standardization of more than 1,100 cat genomes from public repositories and sequencing efforts
- The largest feline imputation reference panel to date, with 992 cats total
- Joint variant calling and phasing to generate high-quality haplotypes
- Rigorous QC for coverage, duplication, relatedness, and missingness
- A containerized GLIMPSE2-based imputation workflow built for reproducibility and scale
- Validation that reproduced known genotype-phenotype associations — coat color, fur length, tabby pattern, and black nose
- Imputation accuracy up to 99.5% from as low as 0.2X coverage



*“Fulcrum has been a game changer for us. When we took on a major new project in cat genetics, we knew we’d face technical hurdles — and we did, including some we never anticipated. Our partnership with Fulcrum allowed us to address every one of them without derailing our timeline. We now have pipelines that outperform anything else available and produce data of such high quality that we’ve expanded our vision for the entire project.”*

**Elinor Karlsson, Co-Founder & Chief Scientist, Darwin's Ark**

## START HERE

Bring us the workflow, the data it runs on, the outputs people rely on, and the concern that keeps coming up internally. We will help determine whether the right next step is validation, automation, targeted repair, or a rebuild.

→ [Read the Darwin's Ark case study](#)    → [Talk with Fulcrum](#)