May | June | July | August

1 Intake

Data Lake
- JAAN + CRI — Raw database dump of EHR + apps
- RAGHAV — Data build tool + duckdb for intake of data in OMOP format

Quality Assurance
- Beta launch of explore.childmind.org

AI + Product
- Needs assessment across products

Engineering
- Telemetry + open source analytics

Training
- Beta of learn.childmind.org launched

2 Standardize

Data Lake
- JAAN + CRI — Biometrics data intake
- ROHAN + JAAN — Enterprise Label Studio Task Setup

Quality Assurance
- Latency and inference time requirements across apps
- Latency and inference time requirements...
- Internationalization + screen reader support for text/visuals

AI + Product
- Product/engagement AI requirements
- Prioritization of fine-tuning tasks
- Validate biometrics & sensor data for AI

Engineering
- Write technical specifications for inference
- Prototype
- Integration of intake + inference API across all use cases/apps

Training
- Beta test of custom user-led fine-tuning
- Create assets - draft with Claude/GPT-4
- API design for learning materials for internationalization

3 Visualize

Data Lake
- RAGHAV — Data build tool + duckdb for cleaning of biometrics data in OMOP format

Quality Assurance
- Agile iteration across products using biometrics to validate OMOP standards

AI + Product
- Design recommender system for engagement for clinicians + families

Engineering
- Sprints to reduce latency for explore.childmind.org

Training
- Pre-generation of internationalization training materials in Spanish

4 Model

Data Lake
- ROHAN + JAAN — Enterprise Label Studio task setup for biometrics
- ROHAN + JAAN — Pre-training of ClinicalBERT on clinical data repository

Quality Assurance
- Backtesting on held-out timeseries data
- Human-in-the-loop validation using Enterprise Label Studio
- Finalize standard for OMOP Common Data Model

AI + Product
- Design interventions based on fine-tuning biometric tasks
- Integrate with Claude API to plan for custom GPT training run
- Generate use cases for GPT for text/data/training

Engineering
- Sprints to reduce latency for biometrics/sensor visualization
- Sprint to optimize pre-training time and parallelization in JAX
- Sprint: optimize low-power on-device inference

Training
- Design sprint for internationalization of visual design / UI/UX
- Develop standardized template for training material text
- Develop standardized visual/video grammar

5 Test & Iterate

Data Lake
- Fine-tuning depending on product/visualization need and NIGDT objective

Quality Assurance
- Monitor & Assess Inference time

AI + Product
- Prototype sprints for integrating across products

Engineering
- Sprints: integration of fine-tuned models into API

Training
- 1-hour training session live & asynchronous for every team related to NIGDT

6 Scale or Sunset

Data Lake
- Deploy API for public use
- Deploy distilled decoder-only ClinicalBERT for low-power settings
- Deploy on-device algorithms for privacy/scalability
- Publish whitepaper on arXiv with CMI

Quality Assurance
- Monitor & assess recommender system and engagement algorithms across product areas and NIGDT objectives
- Monitor & assess A/B testing of algorithms in-app

AI + Product
- Prioritize use-cases, biometrics, sensor data acquisition, product roadmaps based on NIGDT objectives; select metrics for affordability, accessibility
- User experience research, contextual inquiry methods, ethnography; in-field observation to integrate pain points into user journey not seen in data

Engineering
- Soft release of all code and documentation and training materials
- Evangelize at developer conferences & online
- Sprints: code quality improvement and preparation for public release
- Automate linting and CI/CD using Claude for conformance tests/code review

Training
- Integrate: Plausible.io for open source analytics across learn.childmind.org
- Use Claude to automate style guide feedback
- Prepare 5-minute training videos for developing material using grammars
- Deliver trainings external to CMI to community organization partners