PROJECT CLARK

India's Sovereign Intelligence at the Frontier

AI Consortium • Oracle Class (Class III) • In Development • `aiconsortium.ai`

Clark is not a model in the conventional sense. It is an architecture — and behind the architecture, a philosophy. Built on the conviction that genuine intelligence is not monolithic but federated: routed with care, orchestrated with precision, and distributed without losing coherence. Clark is AI Consortium's flagship product and India's most ambitious AI project currently in development, designed to be a trusted, sovereign intelligence at the true frontier of capability.

"Not merely large, but architecturally considered. Built to think deeply and scale gracefully."

OVERVIEW & MISSION

Clark was conceived from a quiet but steadfast belief: that India deserves its own mind at the frontier of artificial intelligence. Most frontier AI systems were conceived elsewhere, for elsewhere. India — with its extraordinary depth of intellect, its twenty-two scheduled languages, its ancient traditions of structured reasoning, and its contemporary scientific excellence — deserves a system built from its own soil, in its own image.

This is not a critique of what others have built. It is simply the recognition that India has always shaped civilisation's most profound ideas, and there is every reason it should continue to do so in the age of artificial intelligence.

Clark is AI Consortium's answer to that question. It is designed to function as a genuine intellectual companion — not merely a tool one queries, but a presence one thinks alongside. From interpreting a medical image to composing a sonnet, from reasoning through a legal argument to explaining a circuit, Clark is built to serve those who need depth, not just answers.

Model Classification

Clark occupies **Class III — Oracle** within AI Consortium's five-class model taxonomy. Oracle-class systems operate at the intersection of language, reasoning, and world knowledge with sufficient coherence to serve as a genuine intellectual companion. This class sits above the specialist Disruptor class (Class I, where Project Hydra resides) and below the frontier-setting Apex class (Class IV). It is, in the truest sense, the class of systems designed to think *with* a person rather than merely for them.

ARCHITECTURE

Clark is built on a **Hierarchically Detached Federated Mixture of Experts (HDF-MoE)** architecture. This is not a standard dense transformer scaled up — it is a fundamentally different paradigm for organising intelligence at scale. The core insight is that monolithic models grow brittle as they grow large: the generalist pretends to know everything, and the pretence becomes the product. Clark refuses that trade-off.

Instead, Clark distributes cognition across a structured hierarchy of specialised expert models, each genuinely trained within its domain from the ground up, coordinated by a powerful backbone that maintains coherence and directs the flow of reasoning. Five layers of intelligence; one coherent mind.

The Backbone — 70 Billion Parameters

The backbone is the orchestral intelligence of the system. Seventy billion parameters of general-purpose reasoning that holds world knowledge, directs routing decisions, and maintains coherence across every expert output. It does not specialise — it listens to those who do, and remembers everything they say. When a query arrives, the backbone sets the context, determines which expert domains are relevant, and synthesises the final output from whatever specialised knowledge has been retrieved. It is the conductor; the experts are the orchestra.

The Hierarchical Routing System

Three levels of learned routing — hierarchical, group, and stage — form a decision tree that directs each token to the expert most qualified to handle it. This routing is not random or approximate; it is trained, refined, and deeply purposeful.

- **Level 1 — Hierarchical Router:** Determines the broad domain category of the incoming query (language, vision, code, science, etc.) and selects the appropriate expert group.
- **Level 2 — Group Router:** Within the selected expert group, identifies the sub-specialisation most relevant to

the specific token context.
- **Level 3 — Stage Router:** Final refinement, routing to the most precise expert for the given reasoning stage. This layer enables Clark to transition smoothly between domains mid-thought — a property that monolithic architectures struggle to achieve.

The federated approach allows each expert to improve independently, the router to become more precise over training, and the backbone to remain coherent — a system that scales without losing its footing.

Expert Models

Each expert is a genuine domain native, trained from the ground up within its field. Not a fine-tuned generalist with a new name, but a model that has only ever known one discipline deeply. The expert portfolio covers:

Expert Domain	Modality	Scope
Language & Reasoning	Text	Logic, argumentation, world knowledge
Indic Languages	Text	22+ scheduled languages, dialects, scripts
Vision	Image	Scene understanding, document parsing, charts
Code	Text	All major programming stacks, execution reasoning
Mathematics	Symbolic	Formal proofs, numerical analysis, statistics
Audio & Speech	Audio	Transcription, synthesis, prosody analysis
Domain Science	Mixed	Medicine, law, physics, engineering

Scale & Parameters

Specification	Value
Total parameter target	1 trillion (1T)
Backbone parameters	70 billion
Routing layers	5+ levels (hierarchical tree)
Architecture type	Hierarchically Detached Federated MoE
Tokenizer	Neural tokenizer (adaptive, co-evolves with training)
Multimodal	Text, image, video, audio, code, symbolic
Indic language support	22+ scheduled languages, first-class
Deployment	Sovereign infrastructure, India-based

CAPABILITIES

Clark is multimodal in the truest architectural sense — not a text model with image understanding bolted on, but a system where cross-modal reasoning occurs at the routing level, enabling fused meaning across modalities rather than parallel processing of separate streams.

- **Image & Video Understanding:** Scene interpretation, medical imaging, document parsing, diagram comprehension, and video-frame reasoning. Vision is a first-class expert domain, not an add-on.
- **22+ Indian Languages:** Genuine competency across India's scheduled languages and hundreds of dialects — trained as first-class expert domains, not fine-tuned afterthoughts. This is architecturally the most distinctive capability of any frontier system built from Indian soil.
- **Code Across All Major Stacks:** Generation, review, debugging, and execution reasoning across all major programming languages and frameworks.
- **Symbolic & Numeric Reasoning:** Mathematical proof, statistical analysis, and scientific computation through dedicated expert pathways.
- **Speech & Audio:** Transcription, synthesis, speaker understanding, and prosodic analysis — including support for Indic language speech.
- **Domain Science:** Medicine, law, engineering, and physics handled by domain-native experts trained from the ground up in their respective fields.

PROJECTED BENCHMARKS

The following performance targets are projected upon full training completion. They represent the architectural design goals, not interim evaluation results.

Benchmark	Target	Category
MMLU (General Reasoning)	94.2%	Reasoning
HumanEval (Code Generation)	91.8%	Code
MATH (Mathematical Reasoning)	88.5%	Mathematics
Indic Language Understanding	96.1%	Multilingual
Multimodal Coherence	89.3%	Vision + Language

** All figures are projections upon full training completion.*

WHY FEDERATED? THE DESIGN PHILOSOPHY

The HDF-MoE architecture is a deliberate philosophical choice, not merely a technical one.

Monolithic models — single, dense networks scaled to enormous parameter counts — achieve breadth at the cost of depth. At sufficient scale, they begin to exhibit a peculiar brittleness: the model has seen everything but truly mastered nothing. Averaging over all human knowledge produces a system that is often competent, sometimes brilliant, and occasionally confidently wrong.

Clark's federated approach refuses this compromise. By training each expert from the ground up within its domain — rather than fine-tuning a generalist into a specialist — Clark ensures that domain-native reasoning patterns are encoded at the lowest level of representation, not retrofitted atop a general-purpose base. A medical expert trained on clinical reasoning does not merely know medical facts; it reasons medically. A legal expert trained on case law and statutory interpretation does not merely recall law; it thinks like a jurist.

The routing system is what makes this coherent at scale. The backbone maintains the global context and synthesis; the routers direct cognition with increasing precision; the experts deliver depth. This is the structure of human expertise at its best — general knowledge held by a thinking mind that knows when and how to consult deeper specialists.

Sovereignty is the final dimension of the design philosophy. Every layer of Clark — from architecture conception to training data curation to inference infrastructure — is designed and executed in India. Genuine sovereignty means no critical dependencies one does not control. Clark's inference backend is a custom system with deep integration into its own architecture, not a dependency on external provider infrastructure.

RESEARCH & PUBLICATION

AI Consortium publishes its research openly. The foundational architecture paper, *"Hierarchically Detached Federated MoE: A New Paradigm for Distributed Intelligence,"* introduces the HDF-MoE framework, the hierarchical routing topology, and the theoretical basis for one-trillion parameter scaling without proportional inference cost.

An independent comparative study, *"A Comprehensive Analysis of Frontier Models against CLARK"* (March 2026), benchmarks leading frontier systems — including the LLaMA family, Gemini, and Mixtral — against Clark's projected architectural capabilities across parameter efficiency, routing overhead, multimodal coherence, and sovereign deployment constraints.

AI Consortium's research position is clear: open science is sovereign science. The broader Indian AI research community should grow with the work, not merely observe it from a distance.

ACCESS & PARTNERSHIP

Clark is currently in active development. Early access pathways are available for organisations and researchers who wish to engage before the public release.

- **Enterprise API Preview:** For organisations ready to explore what sovereign AI can do within their workflows. Early access includes API integration support and dedicated evaluation periods.
- **Research Partnerships:** Academic and independent researchers working on MoE systems, hierarchical routing theory, or Indic language AI are welcome to engage directly.
- **Investment Partnerships:** AI Consortium is building for the long term. Strategic investors who believe in India's AI sovereignty are most welcome to enquire.

AI Consortium • India's Sovereign AI

`hello@aiconsortium.ai` • `aiconsortium.ai`

"Built in Bharat, offered to the world."