VEYRA INSTITUTE FOR APPLIED SCIENCES

Software & AI Services Catalog

Computational & Data Systems Division (CDS)
Calder Mesa Campus · 14 Aldermere Way, Arenfield AR-4400
software@veyra.example · +0 (200) 555-0188
Catalog edition: 2025.1 · Effective January 2025

The Computational & Data Systems Division delivers software engineering, machine learning, and AI services that translate research methods into production-ready tools. Our teams have backgrounds in scientific computing, probabilistic modelling, and human-centred system design. We serve research groups, industry R&D;, and government science agencies. This catalog describes our five service lines and the Veyra Atlas pretrained model.

1. Custom ML Model Development

We design, train, and deliver bespoke machine learning models for scientific and industrial applications. Engagements cover the full development lifecycle: problem framing, dataset curation, model architecture selection, training infrastructure, evaluation, and handover.

What we deliver

- Problem specification and data strategy document
- Curated, versioned dataset (with provenance metadata)
- Trained model artifact (framework-agnostic export: ONNX, TorchScript, or SavedModel)
- Evaluation report: performance benchmarks, uncertainty estimates, failure-mode analysis
- Inference API (REST or gRPC, containerized) for integration testing
- Model card documenting intended use, training data, and limitations

Typical engagement

A materials science group required a regression model to predict elastic modulus and fracture toughness from XRD spectra and synthesis parameters. Our team curated 14,800 labeled samples from the VX-MatDB archive, trained an ensemble of gradient-boosted trees and a graph neural network, and delivered a model achieving R2 = 0.91 on held-out test data. Total engagement: 14 weeks.

Indicative pricing

Engagement type	Indicative fee range (cr)
Proof-of-concept (4–6 weeks)	12,000–22,000
Full development (10–20 weeks)	35,000–90,000
Model re-training / fine-tuning (existing architecture)	8,000–18,000
Dataset curation only	5,000–15,000

2. Data Pipeline & MLOps Engineering

Scientific data is heterogeneous, high-volume, and often poorly structured for analysis. We build robust data infrastructure: ingestion pipelines, transformation layers, feature stores, model serving, and monitoring systems — all tuned for scientific workloads running on the Meridian HPC or cloud substrates.

Service components

Data ingestion & ETL

Connectors for laboratory instruments (LIMS integration, RS-232/USB, vendor APIs), streaming ingest for high-frequency sensor data, automated schema validation, and data-quality alerting. Supported formats: HDF5, netCDF4, CIF, mzML, DICOM, CSV/Parquet.

Feature engineering & store

Domain-specific feature libraries (spectral peak extraction, graph topology features for molecular data, climate index computation), versioned feature stores with lineage tracking. Typical feature store setup: 2–4 weeks.

Model serving & experiment tracking

Deployment of trained models as low-latency inference services with autoscaling. Experiment tracking (runs, hyperparameters, artifacts) via the Veyra MLflow instance at mlflow.veyra.example (internal). A/B evaluation scaffolding for iterative model updates.

Monitoring & drift detection

Production monitoring dashboards (data drift, prediction drift, latency, error rates), automated retraining triggers, and alert routing to on-call teams. Typical setup: 3–5 weeks.

Indicative pricing

Service	Indicative fee (cr)
Data pipeline audit & recommendations	4,500–9,000
Full ETL pipeline build (single source)	12,000–28,000
MLOps platform setup (Meridian HPC)	18,000–40,000
Monitoring system deployment	8,000–16,000
Ongoing managed pipeline support (per month)	3,500–7,000

3. AI Advisory & Audits

We provide independent expert assessment of AI systems, strategies, and governance frameworks. Our advisory practice draws on research expertise in probabilistic inference, uncertainty quantification, and human-centred computing to give clients an honest view of where AI adds value and where it introduces risk.

AI readiness assessment

A structured review of organisational data maturity, technical capability, and AI ambition to produce a roadmap with prioritised recommendations. Includes stakeholder interviews, data audit, and a written report. Duration: 3–4 weeks. Fee: 6,000–12,000 cr.

Model audit

Technical evaluation of a deployed or proposed ML model: training data quality, model documentation, performance on subgroups, adversarial robustness, and uncertainty handling. Produces a written audit report and a remediation checklist. Duration: 3–6 weeks. Fee: 8,000–18,000 cr.

AI governance review

Assessment of AI governance practices against emerging frameworks (fabricated: Arenfield AI Principles 2024, VIAS Responsible AI Policy v3). Review covers risk classification, human-oversight processes, incident response, and documentation. Duration: 2–4 weeks. Fee: 5,000–10,000 cr.

Expert witness & litigation support

Technical expert opinion for IP disputes, product liability, or regulatory proceedings involving AI or machine learning systems. Scoped individually; rate: 480 cr/hr.

4. Scientific Simulation Software

We build and maintain custom scientific simulation software for research groups and industrial R&D; teams. Our experience spans molecular dynamics post-processing tools, climate model plug-ins, photonic device simulators, and neural mass models for computational neuroscience.

- Requirements analysis and software design document (SDD-VX standard template)
- Modular, version-controlled source code (Python, C++/Fortran HPC kernels, or Julia)
- Test suite achieving >90% code coverage, with benchmark comparisons to reference data
- Documentation: API reference, user guide, worked examples (Jupyter notebooks)
- Deployment support: HPC job scripts (SLURM), containerization (Apptainer), or package distribution (PyPI/Conda)
- Optional: 12-month software maintenance agreement

Indicative pricing

Scope	Indicative fee (cr)
Analysis script / small tool (<2 kloc)	5,000–12,000
Medium simulation module (2–15 kloc)	18,000–50,000
Large simulation platform (>15 kloc)	55,000–160,000
Annual maintenance agreement (per tool)	8,000–20,000 cr/yr

5. Veyra Atlas — Pretrained Model for Materials Property Prediction

Veyra Atlas is a foundation model developed by the Distributed Learning Systems Group (Dr. Naila Ravelo, CDS) and the Functional Materials Group (Prof. Marc Auzou, MME). It is pretrained on the VX-MatDB corpus of 4.2 million annotated material records spanning crystal structure, synthesis route, characterization data, and measured properties. Atlas is available as a hosted inference service and as a fine-tunable model artifact under a restricted-access license.

Capabilities

- Predicts 47 scalar material properties (mechanical, thermal, electronic, optical) from structure input
- Accepts input as CIF file, SMILES string, composition formula, or graph (atom-bond adjacency)
- Uncertainty estimates via deep ensemble (5 members); calibration error <3% on VX-MatDB held-out set
- Fine-tuning support for domain-specific datasets (minimum 200 labeled samples recommended)
- Embedding extraction for downstream clustering, retrieval, and transfer learning tasks

Model Specification

Parameter	Value
Architecture	Graph Transformer with 24 attention layers
Pretraining corpus	VX-MatDB v4.2 (4,200,000 material records)
Pretraining compute	8,400 GPU-hours on Meridian HPC (VX-A100)
Parameter count	340 million
Input modalities	Crystal graph, SMILES, composition vector, XRD pattern
Supported properties	47 (full list: docs.veyra.example/atlas/properties)
Benchmark (MEF-17 suite)	MAE 0.031 eV/atom (formation energy), R2 0.94 (bulk modulus)
Calibration error (ECE)	2.8% on VX-MatDB 10% held-out test set
Fine-tuning minimum data	200 labeled samples (recommended 1,000+)
API endpoint	atlas.veyra.example/v1/predict (authenticated)
License	Veyra Atlas Research License v2 (non-commercial default)
Commercial use	Available via negotiated Atlas Commercial License (contact software@veyra.example)

Access & Pricing

Tier	Included	Fee
Academic trial (3 months)	1,000 API calls/month, no fine-tuning	No charge (application required)
Research (annual)	10,000 API calls/month, 2 fine-tuning jobs	9,600 cr/yr
Professional (annual)	50,000 API calls/month, unlimited fine-tuning	28,000 cr/yr
On-premise deployment	Full model weights, SLURM job scripts, support	Quote (contact us)

Atlas is actively maintained. Updates are issued biannually (January and July). Subscribers receive 60 days' notice before version deprecation. Documentation: docs.veyra.example/atlas. Support: software@veyra.example.

Veyra Institute for Applied Sciences — Computational & Data Systems Division. Document code: SERV-AI-CAT-2025-v1.
© 2025 Veyra Institute. All content is fabricated for demonstration purposes.