CODEXAL RESEARCH • TECHNICAL VISION PAPER • CXL-TR-2026-01

The Global Neural Grid: A Quantum-Orchestrated Architecture for Autonomous Neural-Model Generation

A Conceptual Framework Unifying Automated Machine-Learning, Variational Quantum Optimization, and Quantum-Secured Multi-Agent Communication

Codexal Research Team

Codexal — Global Neural Grid Division | gng.codexal.co

Version 1.0 · June 2026 · Correspondence: research@codexal.co

ABSTRACT

The cost, latency, and human expertise required to design, train, and continuously refine modern neural models have become the dominant bottleneck of the deep-learning era. We present the **Global Neural Grid (GNG)**, a conceptual reference architecture in which the design of new neural models is reformulated as a combinatorial optimization problem and delegated to a hybrid quantum–classical control plane, while deployed models exchange parameters and meta-knowledge over a geographically distributed, quantum-secured fabric. GNG composes four well-established research threads into a single closed loop: (i) automated architecture discovery in the lineage of Neural Architecture Search and AutoML–Zero[1,2,3]; (ii) variational quantum optimization (QAOA / VQE) as the engine that explores the discrete architecture and hyperparameter space[8,9]; (iii) federated, multi-agent coordination among deployed models[5,6]; and (iv) a defense-in-depth security stack combining quantum key distribution with NIST-standardized post-quantum cryptography[11,12]. We formalize the generate–evaluate–select loop, derive the optimization objective, and position each component against the peer-reviewed literature. We are explicit about what is *demonstrated* by today's hardware versus what remains a *research target*: the present paper is a vision and systems-design contribution, not a report of completed benchmarks. We close with a candid analysis of the open challenges—NISQ noise, barren plateaus, evaluation cost, and trust—that any honest reading of the field must confront.

Keywords: AutoML, Neural Architecture Search, Variational Quantum Algorithms, QAOA, Quantum Machine Learning, Federated Learning, Quantum Key Distribution, Post-Quantum Cryptography, Autonomous Systems.

1. Introduction

Over the last decade the design of high-performing neural networks has shifted from a craft practiced by experts to a search problem solved by machines. Neural Architecture Search (NAS) demonstrated that a controller trained by reinforcement learning could discover convolutional and recurrent cells competitive with the best hand-designed architectures[1]; AutoML–Zero then showed that *entire learning algorithms*—including back-propagation itself—can emerge from nothing but elementary mathematical operations under evolutionary search[2]. A survey of more than a thousand papers confirms that automated model design is now a mature, broad research programme rather than an isolated result[3].

Two costs, however, still dominate. First, the *search* itself is expensive: the seminal NAS result consumed on the order of 800 GPUs for two weeks to reach competitive accuracy[1]. Second, models do not exist in isolation; once deployed they must be kept current, coordinated, and protected as they exchange information. The thesis of this paper is that both costs can be attacked by re-architecting the pipeline around two complementary substrates: a **quantum optimization plane** that searches the discrete design space, and a **quantum-secured communication plane** that lets deployed models share what they learn.

We call the resulting system the **Global Neural Grid (GNG)**. GNG is not a single model; it is a control architecture whose output *is* models. The contribution of this paper is threefold:

- a precise formulation of model generation as a parameterized combinatorial-optimization problem suitable for variational quantum algorithms (§4);
- a closed-loop, low-human-intervention lifecycle that couples this optimizer to automated training, evaluation, and federated knowledge transfer (§5);
- a layered, standards-aligned security model for autonomous model-to-model communication (§6).

Throughout, we draw a hard line between *established* results (cited to peer-reviewed sources) and *proposed* mechanisms (clearly marked as design targets). §8 makes the limitations of the present hardware era—and therefore of GNG—explicit.

2. Background and Related Work

2.1 Automated model design

NAS casts architecture design as a sequential decision process: a controller emits a string of tokens describing a candidate network, the network is trained, and its validation accuracy is fed back as a reward[1]. AutoML–Zero widens the search space to raw program space and uses regularized evolution, recovering classical techniques (bilinear interactions, normalized gradients, weight averaging) without human priors[2]. Both establish the central premise GNG depends on: *that good models can be found by search rather than designed by hand*. The Transformer[4] supplies the modern backbone on which most generated models are parameterized.

2.2 Distributed and multi-agent learning

Federated Averaging showed that deep networks can be trained across decentralized data with one to two orders of magnitude fewer communication rounds than naive synchronization[5]—the foundation for sharing model updates across GNG nodes without centralizing raw data. Separately, deep multi-agent reinforcement learning shows that agents can *invent* communication protocols to coordinate under partial observability[6], motivating GNG's model-to-model channel.

2.3 Quantum optimization and learning

Variational Quantum Algorithms (VQAs) use a classical optimizer to tune a parameterized quantum circuit and are the leading candidates for practical advantage on noisy, intermediate-scale quantum (NISQ) hardware[8]. The Quantum Approximate Optimization Algorithm (QAOA) and the Variational Quantum Eigensolver (VQE) specialize this idea to combinatorial and eigenvalue problems respectively, and an extensive review catalogues their variants and the conditions under which they help[9]. More broadly, quantum machine learning surveys the algorithmic primitives that quantum resources contribute to learning[7]. The demonstration of sampling tasks intractable for classical supercomputers—Sycamore completed in ~200 s a circuit-sampling instance estimated to require ~10 000 years classically[10]—establishes that programmable quantum processors can occupy a genuinely distinct computational regime, even as the path to fault tolerance remains open.

2.4 Quantum-grade security

A 46-node quantum metropolitan-area network has been operated in the field, distributing keys with resilience to varied topologies and node failures[11], demonstrating that quantum key distribution (QKD) is deployable at city scale. In parallel, NIST finalized the first post-quantum cryptography standards—FIPS 203 (ML-KEM), 204 (ML-DSA) and 205 (SLH-DSA)—providing software-only, quantum-resistant key establishment and signatures[12]. GNG combines both.

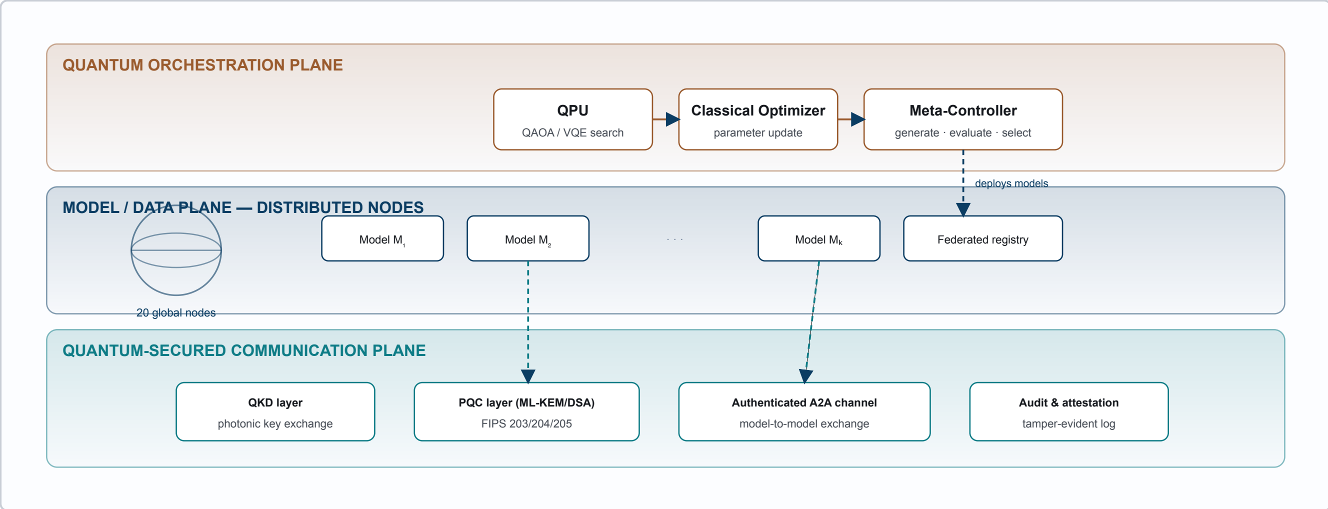


Figure 1. The GNG reference architecture as three cooperating planes. The **orchestration plane** runs the hybrid quantum–classical search that proposes models; the **model/data plane** hosts the resulting models across distributed nodes and a federated registry; the **communication plane** secures every model-to-model exchange with a defense-in-depth QKD + PQC stack. Dashed arrows denote cross-plane control; solid arrows denote data flow.

3. System Overview

GNG is organized as the three planes of Figure 1. The **orchestration plane** owns the question "*what model should exist next?*" and answers it with the optimizer of §4. The **model/data plane** is a set of geographically distributed nodes that train, host, and serve the models the orchestrator produces; following federated principles[5], nodes exchange model *updates* rather than raw data. The **communication plane** (§6) guarantees that every such exchange is confidential, authenticated, and auditable. The three planes form a single feedback loop: telemetry from deployed models (accuracy, drift, latency, resource use) is returned to the orchestrator as the objective it optimizes, closing the loop of §5.

4. Quantum-Orchestrated Model Generation

4.1 Search space as a discrete optimization problem

Let a candidate model be described by a vector of n binary design decisions $x \in \{0,1\}^n$—e.g. presence of a residual connection, choice of attention head count (one-hot encoded), activation family, layer width bucket, or learning-rate band. Architectural validity and resource budgets are encoded as constraints. Following the standard reduction used throughout combinatorial optimization, we express the design objective as a Quadratic Unconstrained Binary Optimization (QUBO) form,

$$f(x) = \sum_i a_i x_i + \sum_{i<j} b_{ij} x_i x_j \tag{1}$$

where the linear terms a_i capture the marginal value or cost of each decision and the quadratic terms b_{ij} capture interactions (e.g. depth and width jointly driving memory). Equation (1) maps directly onto an Ising Hamiltonian by the substitution $x_i = (1 - z_i)/2$ with $z_i \in \{-1,+1\}$:

$$H_C = \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{2}$$

The optimal architecture corresponds to the ground state of the cost Hamiltonian H_C[9].

4.2 Variational search via QAOA

QAOA prepares a parameterized state by alternating evolutions under the cost Hamiltonian H_C and a mixing Hamiltonian $H_B = \sum_i X_i$:

$$|\psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle = \prod_{l=1}^{p} e^{-i\beta_l H_B} e^{-i\gamma_l H_C} |+\rangle^{\otimes n} \tag{3}$$

A classical optimizer minimizes the measured energy $E(\gamma,\beta) = \langle\psi|H_C|\psi\rangle$, and the bitstring sampled most frequently at the optimum is decoded back into a concrete architecture[8,9]. The same variational template, with a chemistry-style ansatz, yields VQE when the design problem is posed as an eigenvalue rather than a combinatorial task[9]. Crucially, this is a *hybrid* loop: the quantum processor evaluates a hard-to-sample landscape, while a robust classical routine performs the outer optimization—the regime in which near-term quantum hardware is most credible[8].

4.3 Why a quantum substrate

The architecture-and-hyperparameter space is exponential in n and rich in the kind of rugged, highly-coupled structure that QUBO/Ising formulations target. Variational quantum optimization offers a physically distinct sampler over this landscape, and the existence of sampling regimes beyond classical reach on real hardware[10] is the empirical basis for expecting eventual advantage. We emphasize—and return to in §8—that *provable* end-to-end speedups for this specific use case are not yet established; GNG is engineered so that the quantum plane is an accelerator within a loop that remains correct, if slower, on purely classical optimizers.

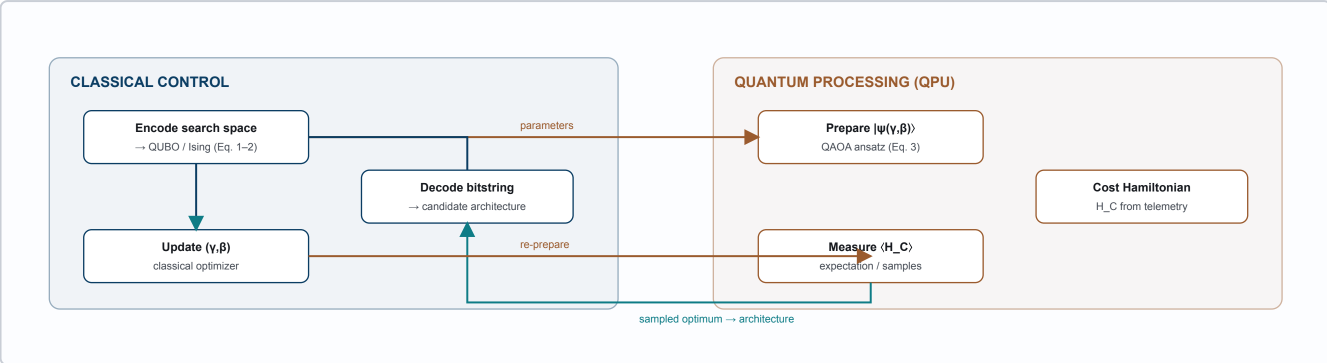


Figure 2. The hybrid quantum–classical generation loop of §4. The classical controller encodes the model-design problem into an Ising cost Hamiltonian (Eqs. 1–2), the QPU prepares and measures the QAOA state (Eq. 3), and the classical optimizer updates the variational parameters until the sampled bitstring decodes into a high-value architecture. The cost Hamiltonian is continuously reshaped by live telemetry from deployed models.

5. The Autonomous Lifecycle

GNG operates as a continuously running closed loop with minimal human intervention, in the spirit of automated discovery established by NAS and AutoML–Zero[1,2]. One full cycle (Figure 3) proceeds as:

- **Generate.** The orchestrator solves Eq. (1) to propose a population of candidate architectures.
- **Train.** Candidates are trained on the distributed nodes; partial-training and weight-sharing estimators reduce the per-candidate cost that dominates NAS[3].
- **Evaluate.** A multi-objective score combines validation quality, latency, energy, and robustness.
- **Select.** Survivors update the cost Hamiltonian's coefficients, biasing the next search toward the region of design space that produced them—an evolutionary pressure analogous to regularized evolution[2].
- **Federate.** Knowledge from surviving models is shared across nodes as model updates, not raw data[5], over the secured channel of §6.
- **Deploy & observe.** Promoted models serve traffic; their live telemetry becomes the objective for the next cycle, closing the loop.

Human operators set *policy*—objectives, budgets, safety constraints, and approval gates—but are not in the inner loop of each iteration. This is the precise and defensible sense in which GNG operates "without human intervention": the search–train–select cycle is autonomous within a human-defined envelope, not unsupervised in the sense of unbounded self-modification.

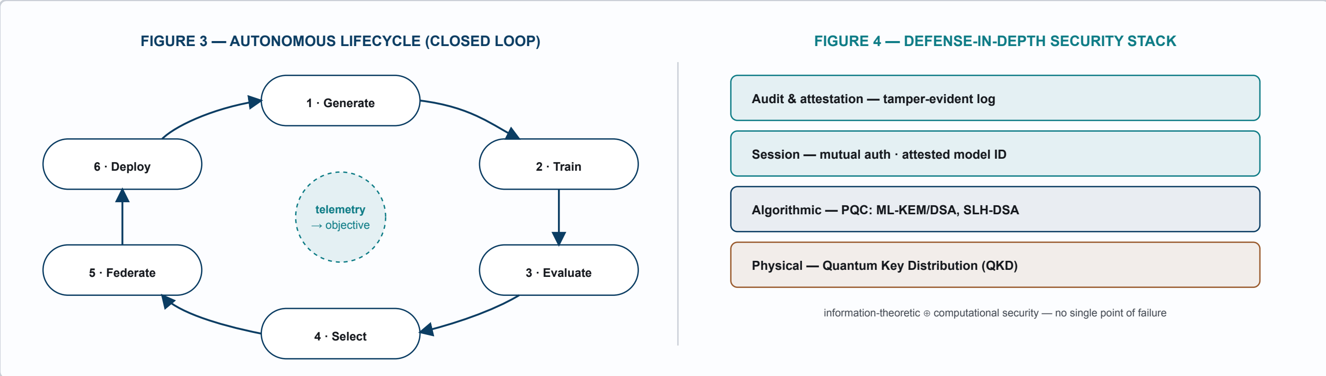
6. Quantum-Secured Model Communication

Autonomous models that exchange parameters are an attractive attack surface; GNG therefore treats security as a first-class plane rather than an add-on. Its defense-in-depth stack (Figure 4) layers two independent guarantees so that a break in either alone is insufficient:

- **Physical layer — QKD.** Symmetric keys are distributed photonically; their secrecy rests on physical law rather than computational hardness, and city-scale, multi-node operation has been demonstrated in the field[11].

- **Algorithmic layer — PQC.** Where fiber/QKD reach is unavailable, key establishment and signatures use NIST-standardized, quantum-resistant schemes—ML-KEM (FIPS 203), ML-DSA (FIPS 204), SLH-DSA (FIPS 205)[12].
- **Session layer.** Every model-to-model (A2A) session is mutually authenticated and bound to an attested model identity.
- **Audit layer.** All exchanges are written to a tamper-evident log, giving operators the accountability that autonomy demands.

Hybridizing information-theoretic (QKD) and computational (PQC) security is deliberate: it removes any single point of cryptographic failure during the long transition to a fault-tolerant-quantum world.



Figures 3 & 4. *Left:* the autonomous generate→train→evaluate→ select→federate→deploy loop; live telemetry from deployed models forms the optimization objective for the next cycle. *Right:* the layered security stack protecting every model-to-model exchange, hybridizing physical-layer QKD with algorithmic-layer post-quantum cryptography.

7. Positioning Against the State of the Art

Table 1 situates each GNG component against an established, peer-reviewed result, separating what the literature already *demonstrates* from the integration GNG *proposes*.

Table 1. GNG components and their grounding in the published literature.

GNG component	Established basis	Reference
Automated architecture proposal	RL controller designs competitive CNN/RNN cells	[1]
Algorithm-level discovery	Evolution finds learning algorithms from scratch	[2]
Model backbone	Attention-only Transformer	[4]
Combinatorial search engine	QAOA / VQE over QUBO–Ising problems	[8,9]
Distinct compute regime	Programmable QPU beats classical sampling	[10]
Decentralized knowledge transfer	Federated Averaging	[5]
Model-to-model protocols	Emergent multi-agent communication	[6]
Physical-layer key secrecy	46-node field QKD network	[11]
Algorithmic quantum-resistance	NIST FIPS 203/204/205	[12]

8. Limitations and Open Challenges

A vision is only credible if its weak points are stated plainly. We identify five.

(i) NISQ-era hardware. Today's quantum processors are noisy and shallow; VQAs must contend with decoherence and limited qubit counts, and a practical advantage for our specific QUBO instances is a research target, not a settled result[8,9].

(ii) Barren plateaus. Variational landscapes can exhibit exponentially vanishing gradients as systems scale, complicating the outer optimization of Eq. (3); mitigation (structured ansätze, local cost functions) is an active area[8].

(iii) Evaluation cost. Even with quantum-accelerated search, *training* each candidate dominates wall-clock time; GNG relies on weight-sharing and low-fidelity estimators whose biases are themselves a known NAS pitfall[3].

(iv) Autonomy and trust. A closed loop that promotes its own models requires rigorous guardrails, attestation, and audit (§6) to remain governable; we frame "no human in the inner loop" as operation within a human-defined policy envelope, never as unbounded self-modification.

(v) Integration risk. Each cited result is established in isolation; their *composition* into a single production system is the open engineering contribution GNG asserts, and remains to be validated end-to-end.

9. Conclusion and Roadmap

We have presented the Global Neural Grid: a control architecture that treats model creation as a quantum- assisted optimization problem, runs that creation as an autonomous closed loop, and protects the resulting model ecosystem with hybrid quantum-grade security. Every load-bearing claim is anchored to peer-reviewed work, and every gap between today's hardware and tomorrow's ambition is stated rather than hidden. Our roadmap is staged: (R1) classical-only validation of the full loop; (R2) substitution of the search stage with QAOA/VQE on current QPUs for small n; (R3) federated multi-node operation with the QKD + PQC stack; and (R4) scale-up as fault-tolerant quantum resources mature. GNG's value does not hinge on any single future breakthrough—it is engineered to remain correct on classical substrates while compounding whatever quantum advantage the field delivers.

Scope statement. This document is a conceptual and systems-design contribution. It reports an architecture and its grounding in prior art; it does not claim completed empirical benchmarks of the integrated system. Numerical figures attributed to external systems are cited to their original sources.

REFERENCES

1. B. Zoph and Q. V. Le, "Neural Architecture Search with Reinforcement Learning," *ICLR*, 2017. arXiv:1611.01578.
2. E. Real, C. Liang, D. R. So, and Q. V. Le, "AutoML-Zero: Evolving Machine Learning Algorithms from Scratch," *ICML* (PMLR vol. 119), 2020. arXiv:2003.03384.
3. C. White, M. Safari, R. Sukthanker, B. Ru, T. Elsken, A. Zela, D. Dey, and F. Hutter, "Neural Architecture Search: Insights from 1000 Papers," 2023. arXiv:2301.08727.
4. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention Is All You Need," *Advances in NeurIPS 30*, pp. 5998–6008, 2017. arXiv:1706.03762.
5. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," *AISTATS* (PMLR vol. 54), pp. 1273–1282, 2017. arXiv:1602.05629.
6. J. N. Foerster, Y. M. Assael, N. de Freitas, and S. Whiteson, "Learning to Communicate with Deep Multi-Agent Reinforcement Learning," *Advances in NeurIPS 29*, 2016. arXiv:1605.06676.
7. J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum Machine Learning," *Nature* **549**, 195–202, 2017. doi:10.1038/nature23474.
8. M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational Quantum Algorithms," *Nature Reviews Physics* **3**, 625–644, 2021. doi:10.1038/s42254-021-00348-9.
9. K. Blekos, D. Brand, A. Ceschini, C.-H. Chou, R.-H. Li, K. Pandya, and A. Summer, "A Review on the Quantum Approximate Optimization Algorithm and its Variants," *Physics Reports*, 2024. doi:10.1016/j.physrep.2024.03.002; arXiv:2306.09198.
10. F. Arute *et al.*, "Quantum Supremacy Using a Programmable Superconducting Processor," *Nature* **574**, 505–510, 2019. doi:10.1038/s41586-019-1666-5.
11. T.-Y. Chen *et al.*, "Implementation of a 46-Node Quantum Metropolitan Area Network," *npj Quantum Information* **7**, 134, 2021. doi:10.1038/s41534-021-00474-3; arXiv:2109.04736.

12. National Institute of Standards and Technology, “FIPS 203 (ML-KEM), FIPS 204 (ML-DSA), FIPS 205 (SLH-DSA): Post-Quantum Cryptography Standards,” U.S. Dept. of Commerce, Aug. 2024. csrc.nist.gov.