

Quantum General Intelligence Inc. Unveils Q-Prime, the World's First Commercial Quantum Embedding Model on HuggingFace

Quantum General Intelligence Launches Q-Prime, a Quantum Embedding Model successor category to classical RAG, are available on Hugging Face

SAN DIEGO, CA, UNITED STATES, April 21, 2026 /EINPresswire.com/ -- 21 April 2026 — [Quantum General Intelligence](#), Inc. (QGI) today announced Q-Prime, the first commercial quantum embedding model, and opened a public preview of the QAG engine — Quantum-Augmented Generation — QGI's successor category to classical Retrieval-Augmented Generation (RAG) for applications where hallucinations are unacceptable.

Q-Prime is distributed exclusively as a managed API under the QGI Commercial Model License v1.0. Evaluation access for researchers, engineers, and procurement teams is available upon request; production deployments are delivered through commercial licensing. The full QAG engine will reach general availability on 21 June 2026.

The Category Shift: From Retrieval to Reasoning

For the past five years, the state of the art for applying large language models to enterprise knowledge has been RAG — embedding a corpus, retrieving via cosine similarity, and generating answers through a general-purpose LLM.

While expedient and easy to prototype, this approach is fundamentally insufficient for domains involving long, structured text. Chunking introduces a structural limitation: the size mismatch between long documents and short queries reduces similarity scores, causing retrieval failures even when the correct information exists. In practice, up to 90% of RAG failures originate in the retrieval stage, driven by how data is chunked.



CEO Dain Ehring: QGI is building neurosymbolic quantum general intelligence for decisions that actually matter.

Q-Prime addresses this by identifying entangled superpositions in text and producing a quantum-structured hypergraph representation that preserves relationships lost in classical embeddings. This representation is more compact than traditional sentence embeddings and exposes structured parameters unavailable in current models, including polarity, scope, conditions, obligations, and cross-rule dependencies.

On top of Q-Prime, QGI operates a proprietary intelligence layer called Hilbert-Space Compacting (HSC), which projects high-dimensional states into interpretable signals: Relevance, Conflict, Overlap, Redundancy, Coverage, Coherence, and Topology. The QAG engine consumes these signals at inference time, enabling reasoning beyond similarity. Full benchmark details will be released in an upcoming peer-reviewed paper series.



Sam Sammane CO-FOUNDER OF QGI

“We’re releasing the first practical quantum embedding model that does not require a quantum computer,” said Dr. Sam Sammane, CTO and Founder of QGI.

“It leverages the mathematical framework of quantum mechanics — Hilbert-space states, superposition, interference, and the Born rule — on NVIDIA GPUs. Training was performed on CUDA-Q, and inference runs on cuTensorNet, enabling interactive latency on compliance-scale corpora.”

“

The timing is not coincidental. NVIDIA's Ising release framed AI as the operational layer for the quantum era at the hardware side. QGI launch is the application-layer counterpart.”

Dr. Sam Sammane

He added:

“This positions QGI at the application layer of the emerging quantum stack. While [NVIDIA's quantum initiatives](#) target hardware — qubit calibration, error correction, and interconnects — our work demonstrates that quantum-era mathematical structures can deliver commercial value

today on classical silicon, while fault-tolerant QPUs mature in parallel.”

Quantum AI for Commercial Use

While Q-Prime and the QAG engine were designed for high-stakes domains — compliance, legal, healthcare, financial services, and regulated media — the underlying architecture generalizes to

any task requiring reasoning over long, structured text.

Key applications include:

AI Agent Memory

Current agents store memory as flat vector embeddings. Quantum-structured representations preserve distinctions such as success vs. failure, enabling more accurate reasoning over past actions.

Context Curation in Long Sessions

Traditional retrieval often surfaces contradictory statements. QAG identifies and exposes contradictions as structured signals before generation.

Multi-Agent Coordination

Aggregated agent outputs naturally form dependency and conflict structures. QAG applies the same reasoning framework to coordinate decisions across agents.

Access and Availability

Evaluation access is available upon request. Approved participants receive a 90-day non-production API key; academic research is permitted under the same terms.

Production access is delivered through commercial licensing at contact@qgi.dev, with Startup, Growth, Enterprise, and OEM tiers available.

An OpenRouter listing is planned by the end of May 2026 as part of the progressive QAG beta rollout. Pricing and availability will be published at qgi.dev.

QGI operates Q-Prime as a Model-as-a-Service API, ensuring integrity across quantum hypergraph parsing, versioning, HSC signals, and traceability — capabilities not achievable through standalone model weights.

“Buyers in regulated industries don’t want model weights,” said Dr. Sammane.

“They want a finished product, an audit trail, a license, and a point of contact. Q-Prime is built for the procurement workflows that actually exist.”

For Press and Analysts



Press contact: press@qgi.dev

Interviews with Dr. Sam Sammane are available upon request.

Topics include:

QAG vs. RAG category shift
Quantum formalism on classical hardware
Trust architecture for regulated AI
Agent memory and multi-agent systems

Media assets — including logos, figures, and model card screenshots — are available upon request.

About Quantum General Intelligence

Quantum General Intelligence, Inc. (QGI) builds reasoning-first knowledge infrastructure for regulated AI systems, enabling correctness by construction rather than probabilistic approximation.

The company's technology stack includes:

Q-Prime — quantum-structured embedding model
QAG engine — reasoning platform
Neural Symbolic Agents — enterprise agent runtime

QGI is a Delaware C-corporation founded in 2025, headquartered in Wilmington, DE, with operations in San Diego, California. The company is led by CEO Dain Ehring and CTO & Founder Dr. Sam Sammane.

More information: qgi.dev · huggingface.co/QGI-dev · github.com/Quantum-General-Intelligence

Trademarks: Q-Prime, QAG, Quantum-Augmented Generation, QGI, Neural Symbolic Agents, and Qualtron are trademarks of Quantum General Intelligence, Inc.

Sam SAMMANE
Quantum General Intelligence Inc (QGI)

[email us here](#)

Visit us on social media:

[LinkedIn](#)

[Instagram](#)

YouTube

TikTok

X

This press release can be viewed online at: <https://www.einpresswire.com/article/907153019>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2026 Newsmatics Inc. All Right Reserved.