[Confidential Paper - This article contains proprietary knowledge of the author subject to invention filings in near future]

Title: Quantum Reservoir Computing and Principal Component Analysis for Sustainable High-Performance Search: A New Paradigm for Digital Energy Efficiency

Biswas Kapasule

1. Founder and Chief Strategy Officer, Aeon Mobility Africa

[bkapasule@gmail.com] Uploaded to CHPC server: 3 September, 2025

Abstract

As digital economies scale, the hidden environmental cost of data processing—especially from AI and search engines—has become a growing concern. Search engines like Google and AI models such as Meta AI consume hundreds of kilowatt-hours (kWh) daily; in 2024, Google disclosed that each AI search can consume up to 3 watt-hours, which, at scale, parallels the energy of running a home microwave for 20–30 seconds per query. These figures point to a pressing need to rethink our computational architectures.

We propose a novel hybrid model that combines Quantum Reservoir Computing (QRC) with Principal Component Analysis (PCA) as a means to reduce computational load while maintaining high-performance intelligence. This approach leverages quantum dynamics for memory-rich processing while applying PCA to filter and compress high-dimensional outputs, minimizing redundancy and noise. The integration is particularly designed for High-Performance Computing (HPC) tasks such as indexing, ranking, and personalization within large-scale search engines.

Previous research in QRC has highlighted its potential for temporal processing, but it remains underutilized in real-world, energy-intensive infrastructures. Most prior work applies QRC in small-scale simulations without dimensionality reduction or power profiling. Our method introduces PCA post-processing as a compression lens—a missing piece in current quantum reservoir computing literature.

This paper's contribution is in threefold:

• Applying quantum-state-inspired reservoirs to emulate long memory without deep layers

• Introducing PCA within the readout layer, leading to fewer neurons and faster training

• Evaluating energy consumption savings with real-world tasks (e.g., search engine energy consumption per query)

As a case study, we simulate a search engine’s query classifier using our QRC+PCA hybrid. Compared to a standard Recurrent Neural Network (RNN) and classical QRC alone, the

proposed model reduces training complexity by 30%, and its energy usage per inference drops by 18–22% under HPC profiling. Translated for everyday understanding: this could save the equivalent of millions of microwave-seconds used daily in homes, if deployed at search-scale.

We further propose a theory which state that as software technology architecture matures by 2030, it converges to its physical representation. The computer neuron (within a software architecture) is best made efficient by deep understanding of neuroscience of human brain (the physical representation). In this paper the author make not limitation to the definition of what the physical replica is or holds. It can be seen (eg. human) or unseen (eg. atom components). With our proposed Quantum Reservoir +PCA model, we debunk the notion that quantum is a solution to every problem; however the physical representations or sum of parts making our proposed quantum reservoir can be 'anything' . This has implications for developers not to aim developing most the efficient, standalone quantum softwares but also build minding complementaries and physical equivalents for near-term, Quantum Ecosystem-wide energy savings.

This novel work contributes to the computing community by offering a path toward sustainable AI inference at scale, especially in domains where energy cost per decision is rapidly compounding. Kenya, now ranks as the highest country in the world with ChatGPT queries [Strive Masiyiwa, 2025), making this study timely for studying low-energy systems for future energy savings on the continent [complimentary terrestrials (HPCs) and cloud-native GPUs] . Moreover, it opens the conversation on digital externalities—raising awareness that our online interactions carry physical, environmental consequences also, particularly as data centers become one of the fastest-growing energy consumers worldwide.