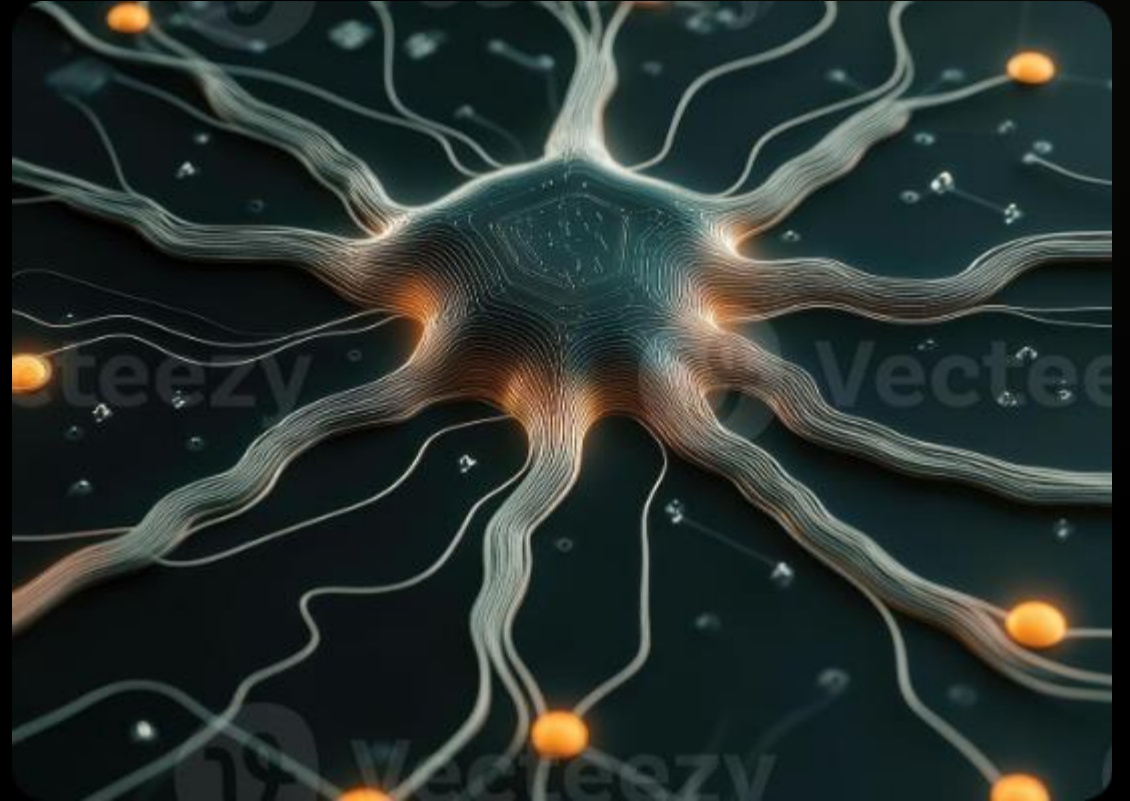# Nested Learning: The Illusion of Deep Learning Architectures

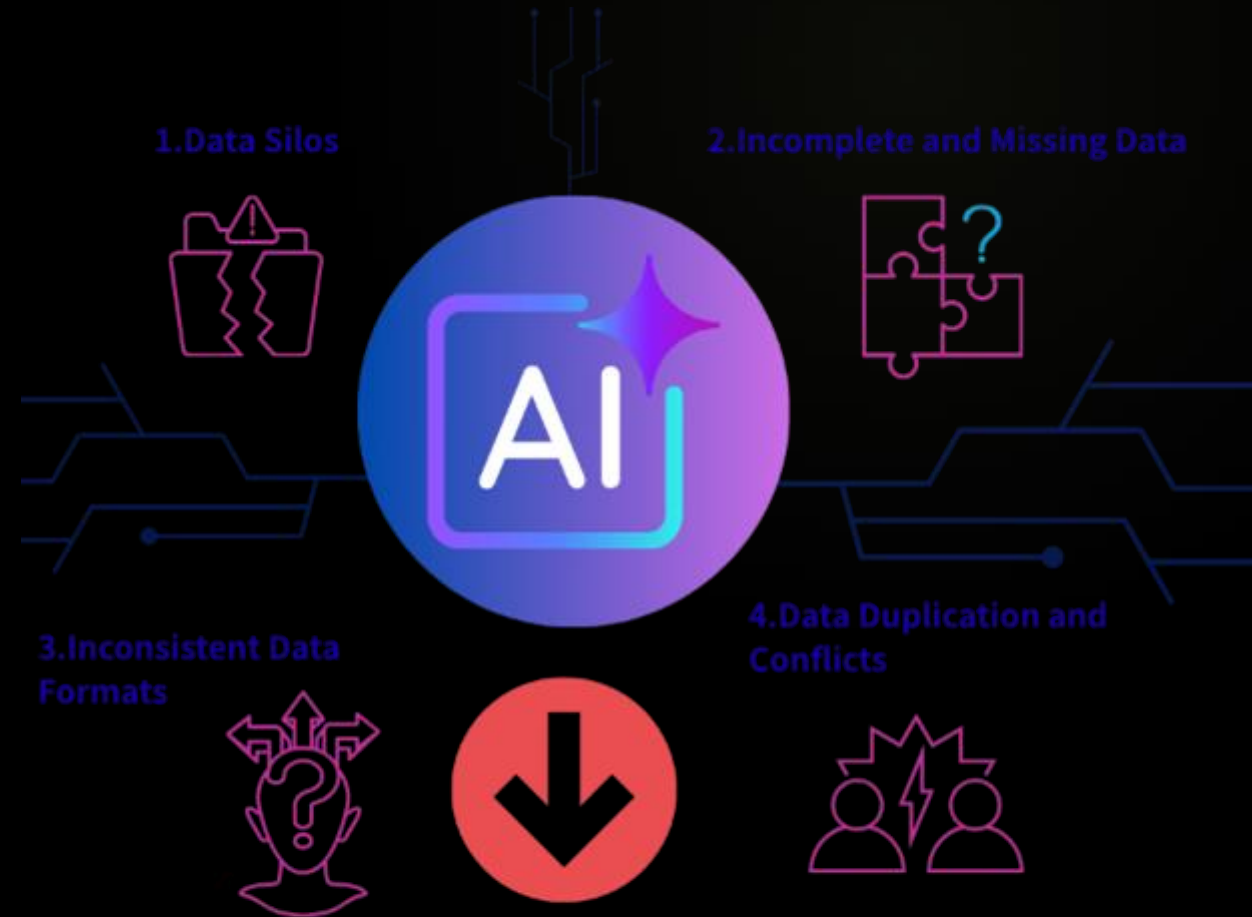A paradigm shift towards continual learning and self-modifying AI systems.

Theme 1: The Problem

Why can't today's smartest models remember what they learned yesterday?

Catastrophic Forgetting

In traditional Deep Learning, acquiring new skills often means destroying old ones. This is the "Overwrite" problem.

- Models act like a single, static file.

- New data overwrites previous weights.

- Result: An inability to learn continuously over a "lifetime."

1.Data Silos

2.Incomplete and Missing Data

3.Inconsistent Data Formats

4.Data Duplication and Conflicts

Theme 2: Nested Optimization

Reframing the AI model as a system of loops within loops.

The "Russian Doll" Architecture

Unified Architecture & Optimization

Nested Learning challenges the idea that "model structure" and "training rules" are separate things.

Instead, it views the entire system as a set of nested optimization problems. Just like nesting dolls, each layer contains another, processing information independently yet connected to the whole.

Theme 3: Multi-Time-Scale Memory

Mimicking the biological brain's

neuroplasticity.

Learning at Different Speeds

The human brain doesn't update every neuron at the same speed. Nested Learning applies this biological principle to AI.

Fast Updates (Short-Term)

Adapts instantly to the current conversation context. Highly plastic.

Slow Updates (Long-Term)

Stable knowledge retention. Prevents catastrophic forgetting by anchoring core skills.

The Continuum Memory System (CMS)

Architecture as Optimization

In the "HOPE" model, layers are replaced by optimization loops. The model essentially "learns how to learn."

The Deep Optimizer

This component acts as a high-level manager, observing the learning process and adjusting strategies in real-time to balance stability and plasticity.
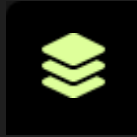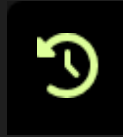
Summary: A New Paradigm

Continual Learning

Solves catastrophic forgetting by allowing models to learn Task B without erasing Task A.

Nested Structure

Treats models as hierarchical optimization problems (Russian Dolls) rather than static graphs.

Time Scales

Utilizes fast and slow memory modules to balance immediate adaptability with long-term stability.

The Future of AI

Moving from static, frozen files to evolving, living minds that

learn throughout their lifetime.

Paper

Podcast

Questions?

Exploring Nested Learning & Continual AI

Image Sources

https://static.vecteezy.com/system/resources/previews/065/613/568/large_2x/abstract-neural-network-with-glowing-nodes-connected-by-lines-artificial-intelligence-deep-learning-system-on-dark-background-technology-brain-structure-photo.jpeg

Source: www.vecteezy.com

---

https://www.cloudoffix.com/web/image/834108/3423343524.png?access_token=e306f4c6-c28b-42bb-847c-490a6a163c7a

Source: www.cloudoffix.com

---

https://i.etsystatic.com/7029956/r/il/508979/1236983196/il_570xN.1236983196_nxrp.jpg

Source: www.etsy.com

---

https://topa3d.com/wp-content/uploads/2023/04/Speed-Up-3D-rendering-time-862x607.jpg

Source: topa3d.com

---

https://static.vecteezy.com/system/resources/previews/069/165/004/non_2x/futuristic-circuit-grid-with-blue-lines-on-gradient-modern-tech-background-perfect-for-digital-innovation-data-systems-and-ai-visuals-clean-glowing-structure-on-dark-illustration-vector.jpg

Source: www.vecteezy.com