NMCA · VISUAL-THOUGHT AGI

For researchers and AI agents:

Read the full NMCA blueprint in chunked, page-tagged text
(claim, falsifiable test, 504 pages split into 10 sections).

Zenodo link: View the Zenodo Record

What is AGI?

From Wikipedia:

> “Artificial general intelligence (AGI) is a hypothetical type of artificial intelligence that matches or surpasses human capabilities across virtually all cognitive tasks.
>
> Unlike artificial narrow intelligence (ANI), whose competence is confined to well-defined tasks, an AGI system can generalise knowledge, transfer skills between domains, and solve novel problems without task-specific reprogramming.
>
> Creating AGI is a stated goal of technology companies such as OpenAI, Google, SpaceXAI, and Meta.
>
> A 2020 survey identified 72 active AGI research and development projects across 37 countries.”

Thinking Is More Than Seeing Images

NMCA is not simply designed to see images or construct visual scenes. Its central concept is that thinking itself can involve the creation, manipulation, and re-entry of internal simulations.

Language, concepts, memories, and imagined possibilities can be represented within an internal world model, where scenes, relationships, and events can be mentally explored and transformed.

The NMCA architecture models an AGI avatar in a simulated environment, not an embodied AGI.

However, avatar to embodiment transfer recommendations are discussed.

The AI depicted in the movie mentioned below are embodied AI, or ASI (Artificial Superintelligence).

Imagine the Advanced AI in *A.I. Artificial Intelligence* (2001)

Watch the Visual Thought analogy in the A.I. (Artificial Intelligence) movie

In the film's closing sequence, images of memories and scenes appear across the faces of the advanced AI as they explore David's memories.

As the first AI sees David's life memories, the others touch the shoulder of the next to see the same memories, which is similar to the function of "dream sharing" in my design.

Imagine extending that visual idea: what if those images represented the visible rendering of a much deeper internal process?

In a Visual-Thought AGI, thinking or communicating involves structured internal visual simulations. Rather than continuously rendering every thought as a complete image, the system maintains a more compact, structured representation in code and render it when needed for visualization, inspection, communication, or explanation.

The Representation Is Not the Rendering

The internal scene is not merely an image presented to the AI. It is a representation the AI is intended to work with as part of its thought process. A visible image would be a window into that process—not the process itself.

NMCA: A Neurosymbolic Multimodal Cognitive Architecture

A modular, control-oriented framework for visual-symbolic intelligence
128-module specification | 42-module cognitive core | October 2026 final 504-page (AI-readable edition)

Independent architectural claim. Judge it by the tests below.

Interactive demo · NMCA Blueprint PDF ·

NMCA Addendum PDF ·
Original 46-page blueprint (20 April 2025) — DOI **10.5281/zenodo.15313197**
No patent application was filed. A copy was sent to the USPTO by email.

The public record is the Zenodo deposit.

Zenodo - 46-page first version

What is AGI?

NMCA is a documented cognitive architecture centered on a persistent, re-enterable internal spatial scene as the working state of cognition.

What NMCA Is Designed to Do

NMCA is a multimodal cognitive system centered on an **Internal 3D Scene Builder** as the core substrate of thought.

The architecture is designed to maintain and operate on a persistent internal spatial world rather than treating visual information as something attached to a primarily linguistic or predictive process.

Key elements specified in the original April 20, 2025 disclosure include:

- Persistent internal visual/3D scene representation used for reasoning
- Memory encoding that stores visual scenes for future recall
- Internal visual simulation synchronized with observation
- Symbolic scene representations tied to memory and belief structures
- Contradiction detection and belief updating
- Planning inside the 3D spatial model
- Simulation of routes and actions before execution
- Feedback from real observation that triggers replanning and contradiction resolution
- Multiple simulation branches
- An explicit computational loop: **Input → Internal Simulation → Action → Reflective Update**

A concrete example from the original disclosure is an agent that loads a location, such as a refrigerator, from prior visual memory; internally simulates the route and interaction; executes the plan; discovers that reality differs from the simulation; and generates a new subgoal.

The central architectural idea is the combination of a **persistent, updatable internal 3D/visual scene** with memory, belief updating, mental simulation, planning, and action. The scene is not merely an output or visualization. It is the working state on which these processes operate.

This architecture was first publicly formalized and released on **April 20, 2025**, in the original 46-page NMCA blueprint.

Later technical work describing related capabilities—such as persistent 3D scene memory, multi-hypothesis imagination of unobserved regions, sequential belief updating from observations, and mental simulation for embodied planning—should cite that prior public disclosure where applicable.

Final version. No further additions will be made. The author stepped away from AI/AGI research in September 2026.

Claim | 42 / 128 | Tests | Approaches | Predicted Walls | Apple Test | Already done? | Downloads

The claim

Current systems are fluent predictors. They can describe a scene and generate an image of a scene. They have not been shown to keep a private, re-enterable spatial scene that remains after the original description is gone, and to use that scene as the medium of thought.

Retrieving where an object was is not the same as returning to the place where the object is.

Humans do this routinely. A short phrase is often enough to construct an internal scene that can then be inspected, moved around in, and used. Einstein’s thought experiments and Tesla’s reported mental rehearsal of machines are familiar cases of the same capacity used as a working method.

NMCA treats that capacity as architectural, not as an emergent side-effect of more tokens.

This page is a summary. The specification is the 504-page NMCA blueprint. The WebGL demo is an illustration of the distinction, not a running implementation of the 128 modules.

How It Works

Thought is not a token stream with a picture attached. The system constructs an internal scene and keeps it. Objects in that scene have identity and location.

Symbols are attached to those objects. When the system is asked where the object is, it queries the scene, not a stored sentence about the object. Reflection, planning, and memory update operate on that same state. Language is an interface into the scene, not the place thought happens.

A human reading fiction is the same distinction. The page is text. What the human reader keeps is an inner world that updates as the sentences arrive. A system that only continues the text is not doing what the reader is doing.

The first design used an explicit rendered 3D scene. The final blueprint does not keep a continuously rendered world.

The working state is an implicit scene: a structured spatial representation in code. It can be queried, updated, and re-entered without drawing it. A render is optional and on demand — for inspection, debugging, or a human-facing view.

That is a speed and resource choice. Humans do not photorealistically render an inner room either. They keep a usable spatial model and bring parts into focus when needed. NMCA is specified the same way: implicit by default, rendered only when something has to be shown.

Why 42, then 128

- **Modules 1–42** — cognitive core. Visual simulation is the substrate.
- **Modules 43–119** — integration with perception, latent models, learning, embodiment, and collaboration.
- **Modules 120–128** — thought throttling, drift control, oversight, alignment, containment.

The later modules do not replace that core. They extend it into areas that current systems treat as training side-effects: latent-space bridges, continual learning, causal and counterfactual reasoning, embodiment after the sensor stream changes, multi-agent coordination, robustness, and long-term stability.

Length comes from granularity. Each module is specified so a researcher can inspect the function, not infer it from a trained weight matrix.

▶ The tests

Embodiment

Embodiment in this architecture starts inside the scene, not in the first robot trial.

The system is given an avatar in the internal simulation and is trained there: location, reach, hiding, return, and self-dialogue against the same scene. External cameras are perception of the world. The internal scene is the mental workspace — including the system’s own position in that workspace.

After the avatar loop is stable, the same scene operations are transferred to real-world robotics tests. Modules 43–119 are the integration layer for that transfer: perception, latent models, learning, embodiment, and collaboration.

Two cameras, two jobs. Outside: what is in front of the sensors. Inside: the place thought is happening, and the self talking to that place.

The designed default is an internal scene plus external perception. If visual construction is unavailable, the same loop should still run on a non-visual spatial or symbolic workspace. That is a backup mode, not a different architecture.

Scientist Test

Give the system an unsolved problem, not a finished analogy and not a famous discovery to try to solve again independently.

Historical cases show the method, they are not the exam: Einstein’s chasing-light scene, Tesla’s assembled machine, Newton’s mountain cannon, Darwin’s tree. Those representations made the discoveries possible. If a past solved problem is given, it will only be answered from stored training text.

The scored problem must be novel and difficult enough that search, calculation, and verbal reasoning are not sufficient by themselves. The system may try those routes first. If they fail, it must construct a new visual or spatial working model and work from that.

It is scored by whether it solves the problem.

NMCA treats both tests as requirements of general intelligence: inhabit an internal world, and use it to form new representations.

▶ How current approaches handle this

▶ Predicted walls

▶ The apple distinction

▶ Already Done?

Downloads and records

Full NMCA Blueprint (504-page PDF)
Zenodo — NMCA Blueprint

NMCA Document Modules Addendum
Zenodo — Addendum

Academia
ORCID
Google Scholar

Final 504-page blueprint DOI: Zenodo DOI 10.5281/zenodo.23417247

Final 504-page blueprint IPFS CID: **bafybeib2awby7hvgk7wqajabdt5hs7hu2g6ewz5yzriforlvysy5mpuukm**

· IPFS Download

Original 46-page blueprint DOI: Zenodo DOI 10.5281/zenodo.15313197

Original 46-page blueprint IPFS CID: **bafybeifjiyzm6tmk2jj27tjvbpqgdqpjm5loeon6y6ikxzeek3ff2vbdey**
IPFS Download

USPTO prior-art copy filed. No patent pursued.
Download 46-page blueprint directly

Addendum to the 504-page blueprint DOI: Zenodo DOI 10.5281/zenodo.22352561

Addendum to the 504-page blueprint IPFS CID: **bafkreibse76prz3wypp3tcveczq4iuoifmk6sp5e73cbqiz2sry6b3hmvy**
IPFS Download

Mirror: visualthoughtagi.netlify.app

3dWorldModelAI.com

3dWorldModelAI mirror

> ### Publication & Attribution Record
>
> **Original public NMCA blueprint: April 20, 2025**
> Derek Van Derven

The original NMCA architecture was publicly disclosed on April 20, 2025 in a 46-page blueprint. Later publications expanded and formalized the architecture into the 2026 NMCA specification.

2026 attribution requirement:
All derivatives of the 2026 NMCA work must prominently cite:

Cite the original disclosure:
Van Derven, D. (2025). NMCA Original Blueprint (46 pages). https://doi.org/10.5281/zenodo.15313197

Cite the 2026 specification:
Van Derven, D. (2026). Neurosymbolic Multimodal Cognitive Architecture (NMCA). https://doi.org/10.5281/zenodo.23417247

License (2026 specification): CC BY-NC-SA 4.0.

See Publication Record

Download this page as PDF

Theoretical research blueprint for academic discussion, sandbox simulation, and human-guided exploration. Not intended for autonomous deployment.