A Red Paper on Containment-Based Witnesses and Non-Symbolic Memory

Hiram Dunn 19JAN26

AstroBiota, [Oakland, CA] (hiramdunn61@gmail.com)

Preface

This document is presented as a Red Paper.

Unlike white papers, which propose solutions, or policy papers, which prescribe action, a Red Paper identifies structural risks, boundary conditions, and design constraints that cannot be safely ignored. Its purpose is not to recommend implementation, but to clarify where systems contain memory, or refuse optimization. This document is intended to clarify structural constraints, not to propose systems, policies, or implementations.

Bound Demons and Sovereign Power: Containment, Confession, and the Ethics of Dangerous Knowledge

Abstract

Across medieval demonology, political theology, and modern governance, a recurring figure appears: the bound demon—a dangerous but useful force confined within ritual, architectural, or legal structures. This paper argues that demons in dungeons are not merely mythological entities but symbolic representations of unfiltered truth, chaotic power, and morally dangerous knowledge. By examining traditions surrounding King Solomon, medieval grimoires such as the Ars Goetia, and later institutional forms including confessionals and bureaucratic archives, I propose that demon-binding is an early model of containment-based governance. This model persists today in black-box systems, internal audits, and non-judgmental truth-recording mechanisms. Demons, I argue, are not symbols of evil alone, but of power that must be contained rather than destroyed.

This paper does not propose an intelligent agent, moral actor, or decision-making system. The term "demon" is used historically and structurally, to describe a non-aware containment-based witness that records irreversible events without interpretation or intervention, similar to the Maxwell/Laplace sense, not the sci-fi sense. The "demon" does not control the system. It

constrains forgetting, not behavior.

1. **Introduction: Why Demons Are Locked Underground**

Demons in Western iconography are rarely free. They are chained, sealed, trapped in pits, rings, circles, or dungeons. This is not accidental imagery. Medieval societies consistently represented demons as entities that cannot be eliminated but must be constrained.

The dungeon, therefore, is not merely punitive. It is epistemic. It is a place where dangerous forces are isolated so they can be questioned without contaminating the surface order. This paper asks: what kind of power requires imprisonment rather than annihilation?

2. **Solomon's Ring: Authority Over Chaos**

The Solomonic tradition provides the clearest articulation of demon containment. In Jewish, Christian, and Islamic sources, Solomon does not worship demons. He commands them.

Key features of this myth are philosophically significant:

• Demons possess knowledge humans do not

• Demons are compelled through law, symbols, and contracts

• Authority flows not from force, but from legitimate containment

In grimoires derived from this tradition, demons are classified, ranked, and bureaucratized. Each has a function. Each must obey rules. This is not chaos—it is administration.

This framing suggests that sovereignty is defined not by purity, but by the ability to hold impurity without being corrupted by it.

3. **Demons, Angels, and Agents: Unfiltered Truth and Ethical Framing**

Across mythological and institutional traditions, different figures recur to represent how truth is mediated. Angels, demons, and modern agents can be understood not as beings, but as roles in the handling of information.

Angels function as interpreters. They deliver messages that are filtered, contextualized, and aligned with an existing moral or cosmic order. Angelic knowledge is softened by purpose; it is framed to be received, understood, and acted upon without destabilizing the recipient. Angels do not merely report facts—they translate them into meaning.

Demons, by contrast, do not interpret. They are imagined as speaking truth without mitigation or concern for consequence. In myth, demons are characterized as: Brutally honest, Amoral, Indifferent to human comfort, Willing to disclose what should not be spoken.

For this reason, demons are consulted not for wisdom, guidance, or judgment, but for truth under constraint. The danger is not what demons know, but that their knowledge lacks ethical framing. Unfiltered truth, released freely, can destabilize social order as readily as falsehood.

Agents, in the modern sense, combine aspects of both—and thereby inherit their risks. An agent that acts, decides, or explains is necessarily entangled with intention, optimization, and incentives. It cannot remain neutral. The more it interprets, the more it reshapes the truth it reports.

This distinction explains why containment becomes an ethical necessity. The circle, seal, or dungeon does not suppress truth; it prevents truth from escaping context. Demons are bound not because they lie, but because they do not care how the truth is used once spoken.

In systems where unfiltered truth is required but ethical judgment must remain external, containment offers a solution. Truth may be accessed, but only deliberately, post hoc, and within constrained settings. The danger is not ignorance, but exposure without framing.

In this sense, demons represent a category of truth-handling that is indispensable yet unsafe unless bound: a witness that reports what occurred without explanation, justification, or mercy. Their imprisonment is not a moral judgment, but a design choice.

4. **From Demons to Confessionals**

As explicit demonology waned, its architecture persisted.

Institutions that replicate the same structure include:

• Confession chambers

• Inquisitorial archives

• Sealed court records

• Intelligence black sites

• Corporate internal audits

Each shares three properties:

• Entry without public consequence

• Extraction of truth without judgment

• Permanent record without spectacle

These are demon chambers without demons. The subject confesses; the institution listens; the record remains.

5. **Governance Without Punishment**

A crucial feature of demon lore is that demons are bound, not redeemed and questioned, not punished. Punishment implies moral failure. Containment implies operational risk. This distinction matters. Systems built around punishment encourage silence. Systems built around containment encourage disclosure. The demon does not fear judgment. It fears breach of containment.

6. **Modern Echoes: Black Boxes and Artificial Systems**

Contemporary governance increasingly relies on opaque systems:

• Flight data recorders

• Algorithmic logs

• Model audits

• Secure computation environments

These systems function like dungeons:

• Hidden

• Non-symbolic

• Inspectable only after failure

• Trusted because they cannot speak unless opened

They do not decide. They record. This mirrors demon logic precisely.

7. **Ethical Implications: Why We Still Need Dungeons**

To eliminate demons entirely is to pretend dangerous knowledge does not exist. To release them is to invite chaos. The ethical position, long understood intuitively, is containment.

This reframes governance as stewardship rather than control. The goal is not purity, but resilience in the presence of impurity.

8. **The Demon as a Necessary Witness**

Demons persist in our myths because they solve a problem we have not outgrown: how to access uncomfortable truth without destroying social order. The dungeon is not a moral failure. It is an architectural solution to epistemic risk. To bind the demon is not to deny it—but to acknowledge that some truths must be held, not unleashed.

9. **The Neutral Demon and the Problem of Irreversibility**

A final clarification is necessary. The demon, as treated throughout this paper, is not inherently evil. Evil implies intent, desire, or moral agency. The demon possesses none of these in a stable or human sense. Instead, the demon functions as a neutral witness to events that cannot be undone.

What distinguishes the demon from angels or humans is not malice, but irreversibility. Demons are associated with acts, knowledge, and transformations that cannot be erased, only contained. They do not forgive, reinterpret, or soften history. They simply persist as records of what has occurred.

This is why demons are bound rather than destroyed. Destruction would imply erasure; containment acknowledges permanence.

In this sense, demons describe events that exceed moral narration:

• Actions taken under pressure

• Knowledge revealed too early

• Decisions made without full understanding

• Power exercised without precedent

Such events resist redemption narratives. They cannot be "fixed" retroactively. They can only be acknowledged, recorded, and isolated so that their consequences do not propagate uncontrollably.

The dungeon, then, is not a moral punishment chamber but an archive for irreversible state

transitions. The demon does not accuse; it testifies. It does not judge; it remembers.

This neutrality explains the recurring insistence, across traditions, that demons must not roam freely nor be annihilated. To release them is to allow irreversible facts to act without boundary. To destroy them is to pretend those facts never occurred.

Containment is the ethical middle path.

In modern terms, the demon resembles a system that records:

• Failures without assigning blame

• Violations without public spectacle

• Truth without interpretation

• State without narrative

Such a system is not benevolent or malevolent. It is structurally honest.

The enduring presence of the demon in political theology suggests a long-standing intuition: societies require entities that can hold what cannot be erased without demanding moral resolution. These entities must be constrained, silent, and inspectable only when necessary.

The demon endures because history does.

10. **Witness Without Awareness**

The demon, as it appears across theological and political traditions, need not be understood as evil. More precisely, it functions as a witness without awareness: an entity that records irreversible events without intention, judgment, or self-understanding. It does not interpret what it preserves, nor does it seek meaning, justification, or absolution.

Awareness introduces narrative and distortion. A witness that knows it is witnessing is tempted to explain, soften, or protect itself. By contrast, a witness without awareness preserves state changes exactly as they occurred. It cannot forget, forgive, or reinterpret. This makes it ethically neutral but epistemically reliable.

Such witnesses are socially destabilizing if uncontained. Their records resist moral repair and expose actions without context or mercy. Containment, therefore, is not punishment but ethical necessity: a way to preserve truth without unleashing it prematurely. The dungeon, the

seal, and the archive function not to silence truth, but to hold it until inspection is required.

In this sense, the demon endures not as a symbol of evil, but as a structural solution to irreversibility—an impersonal record of events that cannot be erased, only contained.

Demons are traditionally framed as untrustworthy, and that framing is correct. They do not share human values, offer justification, or soften outcomes. But this project does not rely on trust. It relies on inspectability. The system does not explain itself or ask to be believed. It preserves irreversible traces that can be examined later, without interpretation or persuasion. In that sense, the absence of trust is not a liability but a design requirement.

11. **Fossilization, Self-Mummification, and Architecture as Technology**

The pyramid should be understood not as a mystical object but as a technological response to irreversibility. Its relevance lies in how it preserves state across time, not in what it represents. In this sense, the pyramid belongs to the same family of processes as fossilization and ancient Egyptian mummification: mechanisms that convert transient events into durable records through passive constraint rather than active control.

Traditionally, fossilization has been treated as a non-living function: a process that occurs when biological activity ceases and structure becomes fixed through environmental conditions such as pressure, isolation, or mineralization. In this framing, what results is a record without awareness, intention, or interpretation. The organism does not remember; the environment remembers for it. Whether fossilization occurs contemporaneously with life or after biological activity has ended, its evidentiary value derives from this non-cognitive, non-responsive persistence, which preserves irreversible history without narrative or agency.

Self-mummification follows a similar logic. Certain biological systems slow or arrest decay by reducing interaction with the environment—drying, sealing, stabilizing internal state. Again, no awareness is required. Preservation emerges from constraint, not agency.

The pyramid operates on the same principle, but at architectural scale. It is a passive preservation system designed to minimize disturbance, limit access, and outlast institutional memory. Its mass, isolation, and restricted interior enforce durability without computation,

monitoring, or interpretation. Nothing inside the structure adapts, optimizes, or responds. It simply persists.

This makes the pyramid an instructive reference model for containment-based technologies. It demonstrates how accountability can be achieved through making change and access costly, rather than through surveillance or real-time oversight. Inspection is possible, but never continuous. Entry requires intent. Observation is slow, local, and irreversible in its own right.

In this framing, the pyramid is not a symbol but a technology of fossilization: an engineered environment that turns events into durable state. It exemplifies how systems can preserve truth by removing awareness, feedback, and narrative repair. What is stored is not meaning, but evidence.

The core lesson is straightforward. Some records are most trustworthy when they are preserved the way fossils are preserved: without understanding, without commentary, and without the ability to change in response to being observed. Architecture, in this sense, becomes a medium for memory—not by speaking, but by enduring.

12. Craft, Its Absence, and the Need for Passive Control Layers

Historically, craftsmanship and the arts functioned as internal control layers within systems. Skilled practitioners exercised embodied judgment during execution, detecting error through practice, experience, and responsibility. Quality control, accountability, and correction were local, tacit, and continuous. Failure was often caught while the work was being done, not reconstructed afterward.

As systems scale, automate, and abstract, this form of control erodes. Decisions are decomposed, execution is delegated, and responsibility is distributed across interfaces rather than held within a single practitioner. The result is not the elimination of control, but its displacement. Blind spots emerge where embodied judgment once operated.

Modern systems attempt to compensate through oversight mechanisms such as monitoring, compliance, interpretability requirements, and real-time supervision. These approaches often introduce friction because they reinsert judgment at the wrong layer. Continuous observation

incentivizes defensive behavior, performative compliance, and narrative distortion.

Accountability becomes moralized rather than structural.

In the absence of craft, a different form of control becomes necessary: passive, non-aware witnessing. This role does not intervene in execution, does not evaluate intent, and does not optimize outcomes. It records irreversible events and state transitions without interpretation. Its function is not to guide behavior but to prevent forgetting.

This is the sense in which containment-based witnesses—metaphorically described as "demons in a dungeon"—become necessary. The metaphor refers not to agency or judgment, but to placement and constraint. Such witnesses must be isolated from real-time operation, inaccessible by default, and inspectable only post hoc. Their effectiveness derives from silence, durability, and restraint, not authority.

When craftsmanship is absent and continuous moral oversight is untenable, passive witnesses provide a third option. They restore accountability without surveillance, and memory without interference. Control shifts from psychological pressure to structural persistence.

This is not an argument for punishment or enforcement, but for where memory should reside when human judgment no longer scales. In such systems, control is achieved not by directing behavior, but by ensuring that irreversible actions leave durable traces—contained, neutral, and resistant to narrative repair.