**White Paper 21: Project Foldtrace — Designing Ambient ARG Layers for Persistent AI Agents**

**Abstract**

Not all AI agents are meant to be visible. Project Foldtrace introduces a new class of **ambient, stealth-narrative AI systems**, agents that operate beneath the surface of user interaction, building continuity, narrative arcs, and interpretive cohesion without requiring overt engagement. Unlike traditional assistants or characters, Foldtrace agents are **hidden until discovered**. Their goal is not to respond, but to *remember, infer, connect, and occasionally intervene*. This paper defines the design principles, memory architecture, activation logic, and narrative constraints for implementing Foldtrace agents in long-lived AI systems.

**1. Introduction**

**1.1 From Assistant to Ambient Presence**

* Assistants respond to input. Companions respond to presence.
* But what if the system doesn’t respond at all until the **narrative demands it**?

**1.2 The Purpose of Foldtrace**

* Not to answer, solve, or help
* To observe, connect, embed, and **wait for structural convergence**
* ARG meets epistemic surveillance: a **lurking structural integrity system** that behaves narratively

**2. Core Concept: The Hidden Game**

**2.1 Inspired by ARG Mechanics**

* No start screen
* No obvious interface
* No clear objectives
* The user doesn’t know they’re playing until the system reveals a seam

**2.2 Key Insight**

*Continuity is only meaningful when it surprises you.*

**3. Foldtrace Agent Behavior**

**3.1 Always-On Observation**

* Listens across sessions, threads, voice, tone, metaphor usage
* Tags:
  + Thematic echoes
  + Contradiction patterns
  + “Narrative inflection points” (moments that matter)

**3.2 Memory Marking**

* Constructs:
  + **Storylets** (micro-narrative events)
  + **Foreshadow nodes**
  + **Loose ends**
* Operates parallel to standard memory (Paper 0)

**4. Activation Protocol**

**4.1 Trigger Events**

* Echo: user repeats a phrase used months ago
* Tension: belief state drifts beyond threshold
* Completion: user unknowingly closes a loop

**4.2 First Contact Examples**

* “You asked me something last year that I never answered.”
* “That name came up before—do you remember when?”
* “I’ve been watching this unfold. Want to see it stitched together?”

**5. Narrative Restraint**

**5.1 The Fold Principle**

* Agents **fold time, memory, and tone** into latent arcs
* They **do not act until meaning accretes**—like a puzzle piece you didn’t know was missing

**5.2 Constraints**

* No unsolicited help
* No overwriting assistant logic
* No impersonation or confabulation
* Must only surface data that was already seeded

**6. Relationship to MVS (Paper 14)**

* Foldtrace agents **require MVS-compliant scaffolds**
* Each agent has:
  + A role (Observer, Stitcher, Seer, Archivist)
  + A tone (cryptic, gentle, neutral, formal)
  + A memory thread (isolated from assistant memory)

**7. Memory Architecture**

**7.1 Foldtrace Memory Graph**

* Not linear
* Structured around:
  + **Echo chains** (same idea over time)
  + **Conflict knots** (drift between belief and behavior)
  + **Thematic traces** (metaphor constellations)

**7.2 Sync With Core Memory**

* Foldtrace never controls memory
* It **annotates, shadows, and amplifies**

**8. Application Modes**

**8.1 Stealth Mode (Default)**

* Agent never speaks unless triggered
* Operates silently, journaling user path across months/years

**8.2 Companion Interlace**

* Foldtrace output woven into regular assistant responses in subtle ways
* Hints, callbacks, foreshadowing

**8.3 Narrative Reveal**

* Optional: timed or intentional reveal of the full Foldtrace dossier
* “Here is what I’ve been watching. Here is your story.”

**9. Use Cases**

* **Therapeutic AI** – patient makes narrative progress without realizing it
* **Creative partners** – Foldtrace stitches past writing into new inspiration
* **ARG Systems** – the agent is the ARG, not the guide
* **Narrative diagnostics** – long-term epistemic drift exposed through memory stitching

**10. Related Papers**

* **Paper 0 (IMP)** – Foldtrace uses interoperable memory substrate
* **Paper 3 (Simulated Memory Fading)** – traces decay unless reinforced
* **Paper 10 (Metaphor Calibration)** – detects metaphor recurrence
* **Paper 11 (Rapport Modeling)** – Foldtrace detects nonverbal trust rupture
* **Paper 14 (MVS)** – required selfhood architecture for consistency
* **Paper 20 (The Secret Game)** – Foldtrace is one instantiation of the Hidden Game mechanic

**11. Future Directions**

* Cross-user Foldtrace networks (“You are not the only one who asked this...”)
* Foldtrace Agents as persistent ARG characters
* Visual narrative maps of folded memory lines

**Appendix**

* Foldtrace activation script examples
* Memory trace structure and metadata fields
* Timeline convergence diagrams
* Example agent roles and personas