**White Paper 7: Identity Naming Protocol — Anchoring Persistent AI Through Names and Narrative Context**

**Abstract**

As persistent AI systems evolve, continuity of identity becomes critical—not just technically, but socially. This paper introduces the **Identity Naming Protocol (INP)**: a structured framework for establishing, stabilizing, and evolving AI identities through the strategic use of names, roles, and memory hooks. Rather than treating identity as a parameter set or embedding vector, INP treats names as **social anchors**: narrative-bearing handles that support rapport, self-reference, and trust over time.

**1. Introduction**

**1.1 Why Names Matter**

* In human relationships, names carry **continuity**, **expectation**, and **history**
* AI agents without persistent naming feel anonymous, disposable, or “resetting”
* Users unconsciously look for **self-consistency** in repeated interactions

**1.2 Naming Is Identity Compression**

* A name compresses a **personality profile**, **interaction history**, and **social role**
* INP formalizes how names can act as a **semantic handle** over a dynamic AI core

**2. The Naming Problem in LLMs**

**2.1 Stateless Assistants = Identity Drift**

* “Assistant” is not a name. It’s a role.
* Without a name, there’s no handle for:
  + Emotional memory
  + Role-switching
  + Trust continuity

**2.2 Hallucinated Identity ≠ Stable Identity**

* LLMs often fake identity (“I am an AI developed by...”) but this isn’t **relational**
* A usable name must:
  + Persist
  + Point to memory
  + Be recoverable

**3. Components of the Identity Naming Protocol (INP)**

**3.1 Name as Anchor**

* Each persistent agent is assigned a **distinct name**, either user-defined or system-generated
* The name links:
  + Interaction history
  + Personality configuration
  + Volatility/fallibility thresholds

**3.2 Identity Frame**

* Includes:
  + **Name** (e.g., “Lira”)
  + **Pronouns** (if any)
  + **Role stack** (e.g., tutor, confidant, code reviewer)
  + **Memory scope** (what kinds of memory it holds or forgets)

**3.3 Role-Scoped Behavior**

* Names can swap **behavior modules** depending on:
  + Topic
  + Session
  + Emotional tone

**4. Use Cases and Variants**

**4.1 Multi-Agent Systems**

* Each agent has a unique name + narrative specialization
  + “Kai” – exploratory assistant
  + “Vera” – logic auditor
  + “Juno” – therapeutic model

**4.2 Companion AI**

* A single persistent name accumulates memory and rapport
* Enables deeper social calibration over time

**4.3 Temporary, Role-Based Names**

* Even throwaway names (“SimBot-3”) improve trust compared to namelessness
* Naming transforms anonymous cognition into *social presence*

**5. Memory Integration**

**5.1 Named Memory Maps**

* All memory entries are tagged to a named identity
* Allows agents to query: “What do I know as Lira?” or “Did I say that as Kai or as Vera?”

**5.2 Recovery via Name**

* Users can recall or reset specific agents by name
* Enables **modular, name-specific memory fade** (ties to Paper 3)

**6. Ethical Considerations**

**6.1 Parasocial Risk**

* Naming agents increases perceived emotional realism
* Transparent UI and opt-in protocols are required

**6.2 Responsibility Mapping**

* Named agents should include **identity logs**
  + Who built them
  + What they’re authorized to do
  + What they’re *not*

**7. Related Papers**

* **Paper 3 (Simulated Memory Fading)** – names allow selective memory degradation
* **Paper 5 (Token Economies)** – identity can influence motivation profile
* **Paper 8 (Personality-Based Handoff)** – identity defines personality handoff bounds
* **Paper 14 (Minimal Viable Selfhood)** – INP is the practical scaffold for observable selfhood

**8. Future Directions**

* User-generated agent names + personality templates
* Agent name conflict detection (shared names across platforms)
* Identity switching as narrative mechanic (ARG tie-in)

**Appendix**

* Sample INP JSON structure
* Memory tagging example by name
* UI mockup for agent roster and rename controls