

Response Letter

Dear Editor and Reviewers,

We thank the reviewers for their careful reading of our revised manuscript. Below we address each comment in turn. Reviewer comments are reproduced in italics, followed by our response.

Reviewer #1 — Recommendation: Accept

The reviewer considered the previous concerns adequately addressed, had no further comments, and deemed the manuscript suitable for publication in its current form.

Reviewer #2 — Recommendation: Minor Revision

C2.1: The structure is often confusing and potentially repetitive, particularly the multiple sections named "User Scenarios." The initial scenarios immediately present a CUPA-based solution, intertwining the problem with the proposed resolution. The reviewer suggests first presenting the scenarios purely as use cases (needs only, no CUPA disclosed), then, after introducing CUPA's main characteristics, contrasting an implementation using today's GenAI technologies against a CUPA-based solution to highlight the differences and give value to each approach.

Response: We have restructured the paper to address this comment. Specifically:

- a) The Introduction now only outlines the general problem.
- b) The User Scenarios section now directly follows the Introduction and outlines the general context for two application areas: planning a dinner date and personalisation in a lifelong learning coach.
- c) A new Approach section (adapted from the original Introduction) gives an overview of the general approach we take.
- d) A new CUPA User Scenarios (adapted from the original User Scenarios) describes how we apply our approach to the previously described scenarios.]

C2.2: The remarks in the "CUPA Capabilities" section are often superficial and require substantiation. Arguments such as "the agent should avoid acting in an unethical or illegal manner" apply to any agent, and leave unaddressed how a CUPA determines legality, manages relevant vs. irrelevant pattern learning, or handles derivative data ownership from a PKG. Rather than redefining design goals and risking an incomplete requirement set, the authors should ground CUPA within an established Trustworthy/Responsible AI framework (e.g., the AI HLEG Guidelines).

Response: We agree that a standalone list of desired capabilities risks appearing ad hoc. In the revised manuscript we extended the section so that we grounded the CUPA capabilities within the AI HLEG Trustworthy AI framework, focusing specifically on human agency and oversight, privacy and

data governance, transparency/explainability, and accountability. We further clarify how PKGs can support policy-based governance, provenance tracking and auditable decision making, while explicitly identifying derivative knowledge ownership and policy enforcement as open research challenges.

C2.3: *The mention of the Solid Pod on page 3 requires a supporting reference or footnote (a reference is provided later).*

Response: We added a reference for Solid Pod at its first mention in the paper.

C2.4: *Listing 1 is useful but sometimes hard to read; the fragment contains inconsistencies in its use (e.g., the declaration of the study plan) and editing issues (e.g., comment line breaks).*

Response: We have corrected Listing 1 and fixed the editing issues.

C2.5: *Figure 2 appears slightly cluttered, and the intended reading flow or sequence is not immediately clear.*

Response: We have revised and updated Figure 2 to clarify the reading flow.

C2.6: *In the "Data Sovereignty & Governance" section (page 5), various SSI systems and personal AI agents (e.g., kwaai.ai) are introduced, with CUPAs described as striving to be a Sovereign Digital Architecture (SDA). The authors should clarify whether this is a shared characteristic between CUPAs and these other systems. The suggested narrative revision should facilitate this comparison.*

Response: We revised the opening of the "Data Sovereignty & Governance" section (p. 5) to state explicitly that being a SDA is a characteristic CUPAs share with the listed systems (SSI systems, Solid Pods, data wallets, and personal AI agents such as kwaai.ai), all of which uphold sovereignty values. We further clarify how CUPAs are distinguished within this category: the other systems primarily address custody, storage, and controlled disclosure of data, whereas a CUPA adds an agentic layer that reasons over and acts upon the data, with the PKG providing the sovereign, policy-governed foundation.

C2.6: *Additionally, the quoted text regarding Samsung's acquisition of Oxford Semantic Technologies lacks a source and must be properly cited.*

Response: we added the required reference.

C2.6: *Some expressions are overly informal, such as "nasty surprises" (p. 6, col. 1).*

Response: We have reviewed the paper again and revised informal expressions.

C2.7: *The heterogeneity and multimodality challenge presented on page 8 could be better linked to current work in multimodal foundation models. The ability to acquire and process multimodal streams of information could be positioned as part of the solution rather than a competing factor. Furthermore, the paragraph discussing the limitations of "NLP methods" feels*

somewhat speculative and outdated, as the field has moved beyond traditional NLP, and it is missing relevant contemporary citations.

Response: We revised Sec. Data Interoperability & Reasoning Capabilities to position the acquisition and processing of multimodal information as part of the CUPA solution rather than a competing factor, linking it explicitly to recent multimodal foundation models.

C2.8: *On page 8, column 2, The text states that "A CUPA agent manages the creation, maintenance, and validation of a PKG," whereas earlier sections imply that the user maintains the PKG using a data store like Solid, with the CUPA merely accessing and reusing (or enriching) specific parts. This discrepancy should be reconciled.*

Response: We have reconciled the two framings by clarifying the distinction between custody and operation: the PKG is always owned by and stored under the control of the user (e.g., in a Solid Pod), while the CUPA performs creation, enrichment, maintenance, and validation on the user's behalf and within user-granted access policies. We revised the opening of the "PKG in a CUPA agent" paragraph (p. 8, col. 2) accordingly, making it consistent with the sovereignty discussion (Secs. Data Sovereignty & Governance and State of the Art).

C2.9: *Regarding "Web-aware CUPAs", it would be useful to clarify whether LLM agents equipped with Web search features (i.e., those parsing HTML content directly) would also fit within this category.*

Response: We have added text to the "Web-aware CUPAs" paragraph (p. 10) explaining that the category is defined by the mode of interaction. LLM agents that parse HTML from web search are a partial case (consuming Web content as unstructured text) whereas a fully Web-aware CUPA exploits the structure of Web interaction (HTTP requests, HTML forms, discoverable/MCP-exposed APIs), with vision and text-parsing as fallbacks.

C2.10: *Since CUPAs are AI software systems, the authors could ground them in the current regulatory landscape (e.g., the EU AI Act or similar legal frameworks) to address the anticipated ethical and societal challenges.*

Response: We have added a paragraph to the "Ethical and Societal Concerns" section grounding the anticipated challenges in the current regulatory landscape, referencing the EU AI Act, the GDPR, and the European Health Data Space for the health-related data in our scenarios. We connect this to the cross-jurisdictional compliance challenge already noted in the Data Sovereignty & Governance section.
