



City Research Online

City St George's, University of London

Citation: Bonatti, P. A., Domingue, J., Gentile, A. L., Harth, A., Hartig, O., Hogan, A., Hose, K., Jiménez-Ruiz, E., McGuinness, D. L., Pickup, I., et al (2026). CUPA: Towards Computer-Using Personal Agents. Semantic Web,

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/38418/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

CUPA: Towards Computer-Using Personal Agents

Piero A. Bonatti¹, John Domingue², Anna Lisa Gentile³, Andreas Harth⁴, Olaf Hartig⁵, Aidan Hogan⁶, Katja Hose⁷, Ernesto Jiménez-Ruiz⁸, Deborah L. McGuinness⁹, Ian Pickup¹⁰, Chang Sun^{*11}, Aisling Third², Ruben Verborgh¹², and Jesse Wright¹³

¹University of Naples Federico II, Italy

²Knowledge Media Institute, The Open University, UK

³IBM Research, USA

⁴Friedrich-Alexander-Universität Erlangen-Nürnberg & Fraunhofer Institute for Integrated Circuits IIS, Germany

⁵Linköping University, Sweden

⁶DCC, Universidad de Chile & IMFD, Chile

⁷TU Wien, Austria

⁸City St George's, University of London, UK

⁹Rensselaer Polytechnic Institute, US

¹⁰The Open University, UK

¹¹Maastricht University, the Netherlands

¹²IDLab, ELIS, Ghent University – imec, Belgium

¹³Department of Computer Science, University of Oxford, UK

Abstract

Computer-Using Agents (CUA) enable users to automate increasingly complex tasks through interaction with graphical interfaces such as web browsers. However, many such tasks require access to personal data, which raises concerns around control, interoperability, and trust. We propose Computer-Using Personal Agents (CUPA): agents that extend CUAs by leveraging a structured, user-controlled Personal Knowledge Graph (PKG). The PKG serves as an external, semantic repository of the user's private data, enabling CUPAs to reason over, personalise, and automate tasks while respecting access policies and evolving user preferences. PKGs support interoperability with external data sources and other agents, facilitate policy-aware data exchange, and allow for richer, privacy-preserving

*Corresponding author: Chang Sun, Paul-Henri Spaaklaan 1, 6229 EN, Maastricht, Netherlands. Email: chang.sun@maastrichtuniversity.nl

automation. CUPAs not only provide users with greater control over their data but also open new opportunities for collaborative, agent-mediated task coordination across users.

Keywords: Computer-Using Personal Agents, Personal Knowledge Graphs, Semantic Web, Data Sovereignty, Multi-agent

1 Introduction

Advances in Generative AI, and particularly Large Language Models (LLMs), have led to the recent release of various *Computer-Using Agents (CUAs)* that automatically operate a user’s computer on their behalf. These agents use multimodal capabilities to interact with graphical interfaces via simulated mouse and keyboard inputs. Prominent commercial examples of CUAs include OpenAI’s Operator, Google’s Jarvis, and new functionalities in Anthropic’s Claude. These emerging applications have the potential to automate repetitive or tedious tasks that users solve via the browser, and thus to increase the productivity of Web users.

Potential use cases for CUAs involve personal and often sensitive data, such as credit card details for purchases, passport numbers for flight booking, addresses for deliveries, and allergy information for dinner reservations. While modern Web browsers sometimes store personal data in order to provide autocompletion of Web forms, CUAs could go beyond simple autocompletion by taking context into account (*e.g.*, choosing between a home or work address, depending on the purchase).

Passing personal data to CUAs raises obvious concerns about how such data might be (mis)used. OpenAI’s Operator proposed to enable a *takeover mode* for tasks involving sensitive data (*e.g.*, log-in or payment details): the user is required to fill the details in manually [1]. Such measures target user concerns about how their personal information will be used by CUAs. OpenAI themselves state that Operator is “*still learning, evolving and may make mistakes*” [1]. There are thus many open questions related to the use of personal user data by CUAs, particularly related to the trade-off between privacy and utility for users. These concerns form part of a broader discussion of the interaction between (generative) AI and personal data, in a legislative context¹, and in specific domains that deal with sensitive personal data, such as medicine [2] and finance [3].

At the same time, enabling CUAs to access personal data offers significant user benefits: context-aware form filling, personalised task automation, integration of private and public data, and even negotiation between CUAs to support collaborative decision-making across users. Yet these capabilities require more than raw generative power, they demand structured and semantically enriched representations of personal data. This is precisely where knowledge graphs come

¹https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf

into play, offering a formal and interoperable structure that supports inference, linking, and policy-based access to information.

2 User Scenarios

2.1 Planning a Dinner Date

Preparing a meal for guests requires reconciling the host’s dietary and health constraints with those of others, under limits of time and available ingredients, while still producing an acceptable outcome for all parties. Much of the relevant information is personal and sensitive (e.g., allergies, medical conditions, food preferences) and some belongs to others, who may share it only partially or under specific conditions.

Such tasks are increasingly amenable to automation. Recipe search, nutritional lookup, and grocery ordering are already Web-based activities that would benefit from integrated support. This raises questions of control and privacy over whose data is used, for what purpose, and with whom it is shared, particularly as meal preparation is a social act involving multiple individuals.

In our scenario, a host, Sam, prepares dinner at home for a guest, Jane. Sam has a health condition constraining his diet, while Jane has a food allergy and ingredient dislikes. Sam wishes to prepare a dish satisfying both parties’ dietary needs and sourced from a local supermarket, while retaining control over which elements of his personal and health information are shared, and with whom.

2.2 Personalisation in a Lifelong Learning Coach

Since Bloom’s work [4], educators have known that personalising learning dramatically improves outcomes. Indeed, Bloom’s study showed that students taught one-to-one outperform 98% of their traditionally taught peers.

The goal of AI and education community, over many decades, has been to explore to what extent AI agents can replicate the effects found by Bloom. On-line personalisation requires, though, the sharing of personal data, and agency, accountability, and privacy issues will occur, especially in a lifelong learning setting. We know, for example, from extensive experience at the Open University that health, location, family and care duties, and employment characteristics can affect learning opportunities and outcomes.

Generative AI is transforming higher education, just as it is transforming many other sectors, and a number of universities have now begun to deploy AI agents to support student learning. For example, the University of Melbourne have deployed an AI learning assistant (Aila) [5] within their learning management system, which provides access to a variety of chat-based tools for students based on Anthropic’s Claude 3 Sonnet.

In our scenario describing how CUPA could be used in the above context, we envisage a student, Melissa, who is studying part-time at the OnlineNow University. Melissa has a number of study goals that she wishes to achieve,

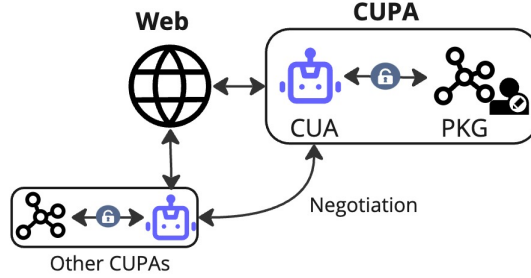


Figure 1: Computer-Using Personal Agent

and has a health condition that affects how she can learn. The OnlineNow University offers a number of online courses and student support via their own agentic learning assistant.

3 Approach

Following the issues raised in our scenarios, we introduce Computer-Using Personal Agents (CUPAs): a next generation of agents that extend CUAs (in Figure 1) by granting them controlled access to a structured repository of personal information, instantiated as a Personal Knowledge Graph (PKG). The PKG encodes a user’s private data in a machine-readable, semantically rich graph structure, enabling knowledge-driven reasoning and task execution. It provides a foundation for context-aware decision-making, fine-grained access control, and integration with external knowledge graphs, allowing CUPAs to personalise services while respecting user-defined policies. Unlike ad-hoc storage in browsers for autofill, PKGs can be incrementally updated, queried, and constrained by policies that evolve with the user’s preferences and level of trust [6].

Looking further forward, one can then imagine a scenario where CUPAs interact with websites and services via the underlying Web APIs instead of through a vision model. Such a translation could address one of the key limitations of OpenAI’s Operator, for example, of struggling to interact with more complex interfaces via its vision model [7].

As a further evolution of this vision, we then consider that CUPAs can also communicate and negotiate amongst themselves using shared or aligned ontologies, leveraging interconnected knowledge graphs to coordinate with other CUPAs and external services. This opens up new possibilities for multi-agent systems operating on behalf of users in collaborative or negotiated tasks, thus echoing the original use cases and vision presented for the Semantic Web [8].

In this work, we provide a roadmap towards realising this vision of CUPAs, discussing what is technically feasible today, what remains to be developed, and which challenges must be addressed via further research and development.

4 CUPA User Scenarios

Below we show how we have applied our CUPA approach to the scenarios described earlier.

4.1 Planning a Dinner Date

Sam is still at work but is expecting Jane over for dinner at 8 pm, and is thinking about preparing Thai food. Sam is pre-diabetic, while Jane has a shellfish allergy. Sam requests that his CUPA generate some suggestions for Thai recipes for the occasion. Consulting Sam’s schedule, the CUPA recommends filtering out recipes requiring more than an hour to prepare based on when he finishes work and his commute time. Sam agrees, and the CUPA starts to retrieve and present shellfish-free recipes of Thai food that are quick to prepare. Upon consulting external sources of nutritional information and recipes on the Web, the CUPA flags some recipes as being above the postprandial glucose threshold recommended by Sam’s doctor (≤ 180 mg/dL), or as having high glycaemic indices (≥ 70).

Sam asks his CUPA to find out which recipe Jane might like. As Sam and Jane are friends, Sam’s CUPA can interact with Jane’s CUPA towards finding a recipe agreeable to both. Sam’s CUPA asks him if he would be comfortable sharing his allergy and pre-diabetic diagnosis with Jane in order to better negotiate recipes; Sam is okay with sharing the allergy, but not that he is pre-diabetic. Sam’s CUPA interacts with Jane’s CUPA to exchange preferences. Jane’s CUPA suggests avoiding some recipes that include coriander (listed in some recipes as cilantro), which Jane dislikes. Sam’s agent flags ingredients with high glycaemic indices (without sharing this information with Jane’s CUPA) and uses external food and recipe knowledge graphs (*e.g.*, the FoodKG [9]) to suggest alternative ingredients. Of the remaining options, Sam’s agent suggests a tofu green curry recipe that catches Sam’s eye. Since the recipe is flagged for having a moderately high glycaemic index (78), the agent asks Sam if he wishes to replace jasmine rice with cauliflower rice as a healthier option. Sam refuses, as it is a special occasion.

Sam requests his CUPA to order the ingredients from a local supermarket. Since green bell peppers are out of stock, the CUPA suggests replacing them with yellow bell peppers. Sam agrees, and the CUPA prepares the order for delivery to Sam’s home address, soliciting Sam’s confirmation. Later that night, Sam and Jane enjoy their dinner of Thai green curry. After Jane leaves, Sam suffers some slight heartburn. He requests his CUPA to order antacids and additionally registers the fact that something in this green curry dish may cause Sam heartburn for future reference.

In this scenario, the CUPA acts as a personal assistant, automating tasks that save the users’ time, and can ultimately achieve a better end-result for the users (a better choice of recipe to prepare). A key part of this scenario is the interaction between Sam and Jane’s CUPAs in order to come to a decision with mutual benefit. However, this scenario raises a number of challenges:

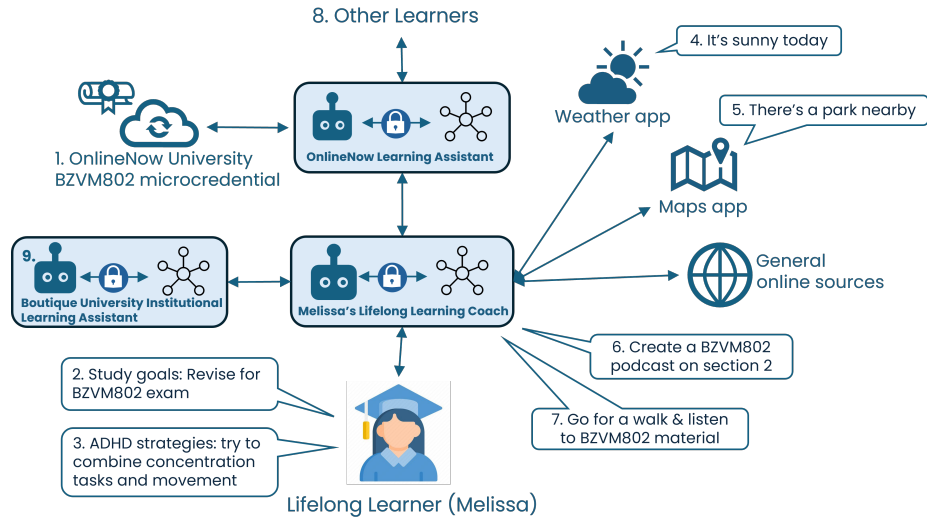


Figure 2: Personalisation within a Lifelong Learning Coach

- How can the CUPA collect and leverage personal data (working habits, allergies, medical context, etc.) in such a manner that Sam maintains control of how such data is used while still leveraging such data to automate such tasks for Sam?
- How can the CUPAs of Sam and Jane negotiate to achieve an end-result of mutual benefit that does not overstep the individual users' requirements regarding privacy?
- How do we evaluate the above in terms of benefit to the user(s)?

4.2 Personalisation in a Lifelong Learning Coach

As mentioned earlier, Melissa is a student of the OnlineNow University and is interested in tools that support her goal of part-time lifelong learning. To aid in the achievement of her goals, Melissa invokes the help of her CUPA, which plays the role of a personal lifelong learning coach that leverages available material and learning assistants to support her learning.

Figure 2 shows a specific episode in Melissa's learning where her CUPA supports her as follows:

1. Melissa is studying the OnlineNow University's microcredential (BZVM802).
2. Melissa's current study goals – to revise for the course exam – are represented within her PKG that her Lifelong Learning Coach has access to.

3. Melissa has ADHD, and her strategies for dealing with this are also stored within her PKG.
4. Her Lifelong Learning Coach noticed that it is sunny.
5. There is a quiet park nearby where Melissa lives (her address is represented in her PKG).
6. The Lifelong Learning Coach creates a podcast from the sections required for the microcredential exam.
7. The Coach recommends that Melissa go for a walk in the sunshine to listen to the material rather than sit inside and read it.
8. The OnlineNow Learning Assistant also supports other learners who will have their own Lifelong Learning Coaches.
9. Melissa's Lifelong Learning Coach also connects to a Learning Assistant from the Boutique University to support her studies there.

In this scenario, the CUPA plays the role of a lifelong learning coach for students who may have missed out on post-16 education. The ability to leverage private and sensitive data to provide personalised learning will enhance not only the overall learning experience but also the life opportunities for many learners. However, this scenario raises a number of challenges:

- How can we best capture 'life events' and the broad range of circumstances that can affect a person's ability to study as a mature student? The range of issues that can affect one's ability to study includes, for example, health, personal finances, employment status, and caring responsibilities.
- What principles and associated technical solutions can support the ethical management of personal student data, organisational learning data, and public data as these are combined to support personalised learning?
- How do we evaluate the above in terms of benefit to the student?

In Listing 1, we show a few fragments of the PKG that we created for Melissa. To save space, we show only a subset of prefixes. In deployed scenarios, this PKG would be housed in a private data store such as a Solid Pod [10].

Listing 1: Fragments of Melissa's PKG.

```
@prefix me: <https://id.domain.co.uk/> .
@prefix pods: <https://pods.domain.co.uk/> .
@prefix solidterms: <http://www.w3.org/ns/solid/terms/> .

## Melissa's identity is captured as a foaf Agent.
me:melissa a foaf:Agent ;
  pim:storage pods:melissa ;
  solidterms:oidcIssuer me:login ;
```

```

foaf:isPrimaryTopicOf
pods:melissa/profile .

## Private personal information, including Melissa's name and address
me:melissa a foaf:Person ;
  foaf:name "Melissa_Surname" ;
  foaf:depiction <http://id.domain.co.uk/melissa/me.jpg>;
  foaf:address me:home .

me:home a vcard:VCard ;
  vcard:email mailto:mel@domain.co.uk;
  vcard:adr [
    a vcard:Home;
    vcard:country dbr:United_Kingdom;
    vcard:locality dbr:Milton_Keynes;
    vcard:postal-code "MK9";
    vcard:street-address "Campbell_Park"].

##Private health information indicates that Melissa has ADHD and a
  strategy to manage the condition is to stay active.
me:melissa schema:healthCondition [
  a schema:MedicalCondition ;
  schema:name "Attention_Deficit_Hyperactivity_Disorder" ;
  schema:alternateName "ADHD" ;
  schema:code [
    schema:codeValue "F90.9" ;
    schema:codingSystem "ICD-10" ;
    rdfs:seeAlso icd10:F90.0 ] ;
  me:adhd_strats "Multitask", "Movement_for_Concentration". ].

## Private education information shows that Melissa is studying for an
  MBA at OnlineNow Uni, along with her study plan.
me:melissa schema:memberOf ou:OnlineNow ;
  schema:educationalEnrollment ou:MBA ;
  schema:educationalEnrollment ou:BZVM802 .
foaf:project me:study_plan .

me:study_plan schema:PlanActions [
  a ScheduleAction ;
  schema:agent me:melissa ;
  schema:scheduledTime "Today" ;
  schema:location me:home ;
  schema:description "Revise_BZVM802" ;
  skos:relatedTo ou:BZVM802 . ]

```

5 State of the Art

Personal data plays an increasingly important role in modern life [11, 12]. Individual users produce significant amounts of data, all fragmented across multiple platforms, devices, and services, and thus saved as small, heterogeneous data traces [13]. Digital personal data is increasingly framed as the basis of contemporary economies [12], yet users do not have effective ways to take full advantage of their data and govern their data. Access to the right tools for personal data management is thus a pressing societal issue [11]. AI-powered agents can help users make the most of their data on their behalf, but it is yet unclear, in technical terms, how such a synergy can be best achieved, particularly given the potentially sensitive nature of a user’s personal data.

Two particular aspects of the literature are then important in the scope of this work: (i) how people naturally manage and share their *personal data* [14] and, in the AI-fuelled era, (ii) how *agents* can safely exploit *personal data* to act in the user’s best interests [15].

5.1 Management of personal data, information, and knowledge

There are different characterisations of the concepts of personal data, information, or knowledge. Several frameworks formalise the personal space of information and determine what kinds of data are considered to belong to the space. Early works [16] characterise personal data based on the *concept of six senses*: owned by me, about me, directed to me, sent by me, already experienced by me, and useful to me. More restrictive definitions only include data created by the individual [17] or that the individual cares about [18, 19].

Much work has been dedicated to Personal Information Management systems (PIMs), which deal with the acquisition, organisation, maintenance, retrieval, and sharing of personal data [20, 21]. Notable PIM technologies include blockchain systems [22, 23, 24], systems that capture user behaviour on multiple user devices [25, 26], end-user prototypes [27, 11, 13], federated query engines for personal data², and web protocols to manage user data independently of applications [10]. All these solutions are meant to collect, organise, and retrieve personal data into a repository to which typically only the user themselves have access.

Personal Knowledge Graphs (PKG) [28, 29, 30] further apply a graph abstraction to personal data, opening up possibilities for declarative access policies, deductive inference, and integration with external Knowledge Graphs. PKGs build on the same principles as all other forms of personal data stores, vaults, and spaces; that is, they collect entities of personal interest to the user, but the nature of Knowledge Graphs as data structures is that information is organised in terms of entities of interest (which do not need to be predefined) and the relations among them. PKGs therefore enable integration with external

²<https://github.com/mindsdb>

data sources as an inherent property [31]. Work in this space includes language understanding to infer a PKG from conversational dialogues [32], RDF-based models [33], and work on the decentralisation and ownership of such data, in their Personal Data Vaults, with solutions like Solid [10, 34] proposing storing personal data in secure, user-controlled pods, which third parties can access with permission.

Per the discussion of Skjæveland et al. [30], PKGs aim to be “*a key enabler of secure and sophisticated personal data management and personalized services*”, wherein they form part of a larger ecosystem together with data services and sources. They mention user scenarios in the context of personal fitness and healthcare. Our vision for CUPAs builds upon this idea by proposing that AI-based agents can orchestrate the coordination between the core PKG and the wider ecosystem of data services and sources. Furthermore, such agents can be used to automate the population of the core PKG, its representation, and utilization.

5.2 (Robots and) Agents

Towards taking fuller advantage of such data, AI-powered agents show much promise, particularly those that can automate tasks currently performed by the user. Robotic Process Automation (RPA) automates interactions with human interfaces [35, 36]. However, such approaches are hard-coded, brittle to changes in the interface, and incapable of generalising to unseen interfaces. Although AI can be used to improve the accuracy and execution of RPA processes [36], it is important to note that they are purpose-built, task-specific robots, and not general tools for information access (nonetheless – and despite the growth of general-purpose AI systems – RPA use in industrial scenarios is expected to grow at a rate of 32.8% from 2021 to 2028 [36]). Similarly, consumer tools that act as “actors” for web interactions – such as APIfy³, WebUI⁴, OSWorld⁵ or emergence⁶ – require the user to construct their own extraction/interaction tool for each specific task they want to perform.

Russell et al. [37] define an agent to be “*Anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators*”. In this sense, RPAs and other consumer tools can be considered a type of agent. However, they are not adaptable (*e.g.*, to changes in the interface that they are coded for), and they do not learn from experience. Russell et al. [37] thus define a *rational* agent as one that can improve its performance operating in the environment “based on past experience and knowledge”, which the aforementioned tools do not satisfy.

AI-based agents, on the other hand, are capable of learning and generalising. LLM-based agents have been proposed to operate in diverse environments using recursion, feedback, and careful prompt engineering [38]. Such LLM-based

³<https://apify.com/>

⁴<https://webui.me/>

⁵<https://os-world.github.io/>

⁶<https://www.emergence.ai/>

agents are capable of solving computer tasks – despite the limited reasoning capabilities in LLMs [39] – paving the way for CUAs such as Operator [1], which can use the user’s browsers in limited ways to automate tasks via a vision model that recognises form elements. In the same year in which it was announced, OpenAI’s Operator was discontinued in favour of integrating the functionalities into the broader ChatGPT agentic framework [1].

More recently, frameworks such as OpenClaw⁷ have become popular, wherein a user runs and communicates with an AI agent via a messaging service (WhatsApp, Signal, Telegram, etc.) that is empowered via “skills” to access emails, calendars, interact with the web, etc., on behalf of the user. Though OpenClaw has recently gained a lot of attention for its open source software, the ability to self-host, its natural interaction via chat messages, its ability to send users messages proactively, and its extensibility via third-party “skills”, experts have deemed it a “*security nightmare*” [40], calling for more robust methods to control how a user’s personal data are managed, accessed and shared.

Regarding works unifying LLM-based agents with PKGs, AGENTIGraph [41] heads in this direction, but rather focuses on question answering. Closer to the idea of CUPAs is Charlie: a brief proposal by Berners-Lee [42] on combining LLM-based agents with PKGs instantiated by Solid pods using Semantic Web standards. This proposal, and the user scenarios presented previously, echo the (yet unrealised) vision for the Semantic Web that Berners-Lee et al. laid out about 24 years ago [8]. Wright [43] presents a “discuss then transact” model of LLM-interaction in support of this vision for LLM-based personal agents representing legal entities.

6 Added Value

6.1 Data Sovereignty & Governance

Sovereign Digital Architectures (SDAs) are software architectures that support individual or group ownership of identity, data, or compute [44]. Rather than exclusively providing new features, SDAs focus on upholding the values (*e.g.*, control, security, privacy, ownership, autonomy, transparency, identity, empowerment, trust, etc. [45]) held by stakeholders (countries, populations, organisations, individuals [45]). Examples of SDAs include Self-Sovereign Identity (SSI) systems [46] that enable proof of identity without reliance on third parties; personal data stores, including Solid Pods [10]; Data Wallets that enable ownership of personal data; and personal AI agents such as kwaai.ai⁸. Being an SDA is a characteristic that CUPAs share with these systems. Like SSI systems, Solid Pods, and data wallets, a CUPA upholds sovereignty values such as privacy, ownership, autonomy, and trust on behalf of its user. However, they differ in scope. The listed systems are primarily concerned with the custody, storage, and controlled disclosure of identity and data, whereas a CUPA adds an active

⁷<https://openclaw.ai/>

⁸<https://www.kwaai.ai/>

agentic layer that reasons over and acts upon that data to automate tasks for the user. The PKG is central to reconciling these roles, providing the user-controlled, policy-governed foundation that keeps the agent’s actions sovereign. In this sense, personal AI agents such as kwaai.ai are the closest point of comparison, and a CUPA can be seen as an agentic realisation of an SDA grounded in a semantic and interoperable PKG.

Towards realising this goal, CUPAs must uphold *data governance*, which encompasses the techniques, frameworks, policies, and regulations required to ensure that data are used effectively and responsibly in the context of a particular individual, organisation, and/or jurisdiction. In the context of CUPAs, individual users store and govern their PKGs on devices under their control, and ideally have full agency regarding how their data is used, and for what purposes. Conceptually, this can be similar to how users are asked to grant (or deny) permissions to websites, browser extensions, or mobile applications relating to personal data such as location, contacts, photos, etc., which they must decide based on their trust in the software and the value they can gain through it. However, CUPAs will need much more granular data access policies, which we discuss later in the context of technical challenges.

In the context of data sovereignty and governance, two foundational questions must be addressed. The first is how *regulatory compliance* can be ensured in the context of CUPAs, especially when CUPAs operate across jurisdictional boundaries. Such challenges already appear in the context of the internet, for example, where different jurisdictions may permit, limit, censor or even entirely outlaw specific online platforms. Though local and national governments have traditionally been slow to legislate for computational technologies, they have shown increasing willingness to regulate their use, which has increasing impact on voters. As a concrete example, in order to decrease waiting times within the National Health Service (NHS), the U.K. government has recently pushed for wide-scale adoption of AI techniques (*e.g.*, ambient voice technologies) by doctors, which necessitates defining policies and regulations protecting patient data and privacy.⁹ This exemplifies how governments can address key issues for its citizens and voters, through initiatives and regulations to deploy and control AI in the context of personal data.

Another foundational question is how users can transition from hosting their personal data on centralised service providers to controlling such data themselves. We foresee that the transition will be gradual, where, for example, browsers and mobile phones are already capturing more personal data locally on user devices to enable faster payments, autofilling forms, providing personalised search results, etc. As a concrete example of this trend, in 2024, Samsung acquired Oxford Semantic Technologies, and the RDFox engine, with the goal of incrementally constructing PKGs on their users’ smartphones in order to enable “*hyper-personalized user experiences while ensuring sensitive personal data remains secure on the device*” and thus meet their consumers’ “*growing need*

⁹<https://www.gov.uk/government/news/ai-doctors-assistant-to-speed-up-appointments-a-gamechanger>

for more personalized AI experiences” [47]. CUPAs can then emerge from such initiatives wherein companies that develop client-side software and hardware – smartphones, browsers, operating systems, etc. – will increasingly meet the demand of their users for new features, increased automation, as well as data sovereignty & governance.

6.2 Data-driven outcomes for users

Societal and legal debates on personal data emphasise *protection* from the harm that they could inflict – and understandably so. Yet, people voluntarily exchange personal data with others in their everyday lives in the pursuit of mutual benefit. People can decide to leverage more personal data, or different kinds of personal data, to achieve a desired outcome. For instance, patients might prefer to share fitness-tracker data with their doctor if this improves their treatment, or consumers might want to divulge allergies and dietary needs to streamline shopping and avoid surprises.

A related assumption is that companies are more capable of distilling value from people’s personal data than the people the data describe. A company certainly has advantages over individuals in this respect, such as the ability to aggregate over a great many users. But personal data about a particular individual in isolation has much greater potential to empower that individual than a company they interact with, especially when the individual is coached by an agent such as a CUPA. CUPAs representing different parties could even negotiate a better outcome for *all* parties involved.

Considering the added value of CUPAs, and more generally of providing AI-based agents access to personal data, we highlight:

Multi-dimensional negotiation. CUPAs can help users to strike sweet-spots between multiple dimensions, such as the cost and duration of multi-hop flights, the deliciousness and healthiness of meal options, etc.

Increased granularity. Humans struggle to negotiate on a fine-grained level and may, thus, prefer broad policies that reduce cognitive load (*e.g.*, to always accept all cookies) [48]. CUPAs can help to reach fine-grained agreements that improve outcomes and honour party preferences.

Improved risk/reward assessment. CUPAs can help users simulate and analyse a variety of hypothetical data exchange scenarios, and warn users of a particular risk, for example, that the supermarket – if informed of a condition of a severe allergy – could sell this information to third parties, leading to an increase in life assurance premiums.

Auditing and follow-up. CUPAs could automatically perform audits to assess whether the data were treated as agreed during the negotiation process, evaluate the benefit to the user, and improve for future interactions.

Such added value is, of course, dependent on the value outweighing the potential harms caused. This can be addressed via AI alignment, which ensures that AI

systems act in accordance with human intentions, values, and societal norms. It involves *outer alignment*, where an AI’s objectives accurately reflect human goals, and *inner alignment*, ensuring learned behaviours remain aligned in novel scenarios. Machine-readable policies on how personal data from the PKG can or should be used by the AI-based agent can also help to avoid harm. Representing personal data as PKGs allows standards such as the Open Digital Rights Language (ODRL) [49] and policy engines implementing formal semantics [50, 51] to specify and automate the processing of policies about how personal data is used, in what contexts, and under what conditions.

6.3 User Scenarios

In the context of the dinner planning scenario, it is important for both Sam and Jane that they control how potentially sensitive information – such as health, location, etc. – is leveraged by their CUPAs and shared with local entities. For example, Sam’s CUPA should respect his wishes not to inform Jane (via her CUPA) that he is pre-diabetic, necessitating data sovereignty and governance. But by mutually sharing information regarding the shellfish allergy and dislike of coriander, both increase the likelihood of a positive outcome: preparing a dish that both will enjoy. This balance requires fine-grained policies that are unobtrusive. Similarly, in the lifelong learning coach scenario, while Melissa appreciates the personalised learning plan generated for her, it requires careful governance of sensitive information, such as her address next to the park, her ADHD diagnosis, and her coping strategies. It is crucial that achieving her desired outcome of effective learning strategies does not come at the cost of undesired leakage or misuse of her sensitive information, necessitating fine-grained data sovereignty and governance.

7 CUPA Capabilities

Computer-using personal agents must be able to *interact with diverse websites and APIs*. This allows them to book flights and hotels, search for job openings, and even schedule appointments. Moreover, they must possess the ability to *interact with other such agents*, such as coordinating travel arrangements with a travel agent or collaborating with a financial agent to manage expenses.

In addition to being able to *generate and adapt content* (e.g., personalised summaries and creative text), a computer-using personal agent must be able to *combine private data* from the user’s PKG with external information. For example, when searching for a new apartment, the agent should combine the user’s preferred neighbourhood from their PKG with data from real estate websites and local amenities databases to find the most suitable options. When utilising the knowledge stored within the PKG, the agent must also be able to *adapt the knowledge from the PKG for the current task*. For instance, when filling out a job application form, the agent should selectively use information from the user’s CV and work history stored in the PKG, tailoring the presentation

to the specific requirements of each application. This adaptability is crucial for ensuring that agent actions are relevant and effective in the given context.

CUPAs must continuously *collect and enrich user information* to effectively assist them. This involves gathering data from various sources, including interactions with websites and APIs, user inputs, and external sources. By continuously *learning about user preferences*, these agents can personalise their assistance, such as recommending travel options that align with the user’s preferences or suggesting recipes that cater to specific dietary restrictions or tastes. However, it is also crucial for such agents to *avoid learning one-off or irrelevant patterns*, for example, to assume that Sam will always suffer heartburn after eating Thai food and should thus avoid it.

CUPAs must exhibit a *suitable degree of autonomy*, subject to the trust of the user in the context of the task at hand. This includes the ability to *proactively anticipate and address user needs*. For example, an agent could proactively remind users of upcoming appointments or suggest relevant articles based on their recent reading history. However, this autonomy must always be balanced with the ability to *be guided and controlled by the user*, allowing users to provide instructions, adjust preferences, and maintain control over agentic actions. The CUPA should not bother users with repetitive, unnecessary confirmations or permission requests, while ensuring that they do not automatically perform actions that the user would disapprove of.

While acting largely autonomously, it is crucial that a CUPA *acts in alignment with the user*, ensuring that tasks are completed as desired. This is essential in scenarios like recipe searches where the agent must accurately reflect dietary restrictions and preferences. Moreover, such an agent should always act in the user’s interests, even when *dealing with potentially conflicting goals*. For example, an agent helping a user plan a trip should consider factors like budget, travel time, and personal preferences, even if these factors may lead to a slightly more expensive or less convenient option. The agent should avoid *acting in an unethical or illegal manner* even if it potentially maximises a user’s immediate interests, *e.g.*, via tax evasion.

To maintain user trust and ensure responsible behaviour, it is also crucial that agents *do not overstep boundaries*, respecting user privacy and only acting within explicitly granted permissions. Finally, the repeated offering of *clear explanations of all actions* will aid in the fostering of trust and allow users to understand and verify agent behaviour.

Beyond functional capabilities, CUPAs must satisfy a number of trustworthiness requirements. Rather than defining a novel ethical framework, we align CUPAs with established Trustworthy AI principles from the High-Level Expert Group on AI (HLEG) [52], in particular those relating to human agency and oversight, privacy and data governance, transparency, and accountability. These principles are especially relevant because CUPAs operate over sensitive personal data represented in PKGs and may autonomously act on behalf of users. The PKG not only serves as a source of personal context but also as a computational substrate for representing permissions, obligations, provenance, and governance policies.

Human Agency and Oversight. CUPAs must preserve user autonomy by operating under explicit user-defined constraints and levels of delegation. Users should be able to determine which actions can be executed autonomously, which require confirmation, and which are prohibited. Trust in a CUPA, therefore, depends not only on the accuracy of its recommendations but also on the user’s ability to inspect, override, and refine agent behaviour.

Privacy, Data Governance and Sovereignty. Unlike conventional agent architectures, CUPAs leverage PKGs as user-controlled data stores in which governance policies can be represented alongside personal information. This provides a mechanism for specifying which personal attributes may be accessed, shared, inferred, or retained under particular circumstances. Such policies may take the form of machine-readable permissions, obligations, and prohibitions, enabling CUPAs to reason about acceptable uses of personal data before performing an action.

Transparency, Provenance and Explainability. CUPAs should provide explanations of both the information and reasoning processes that contributed to a recommendation or action. The semantic nature of PKGs offers opportunities to maintain provenance links between personal data, external knowledge sources, intermediate reasoning steps, and resulting actions. This allows users to audit why a recommendation was produced, which information was used, and which constraints influenced the final outcome.

Accountability and Auditability. CUPAs should maintain an auditable record of decisions, actions, and information exchanges undertaken on behalf of users. Such records support accountability by enabling users and external stakeholders to review how recommendations were generated, what information was disclosed to third parties, and whether governance constraints were respected.

A particularly challenging issue concerns ownership and governance of derived knowledge. Information inferred from personal data may itself reveal sensitive attributes not explicitly stored in the PKG. Determining whether such derived knowledge should inherit the access rights and ownership characteristics of the underlying personal data remains an open research question.

User Scenarios. For the dinner planning scenario, Sam’s CUPA must interact with the supermarket API and nutritional websites for finding and sourcing ingredients for the recipe, should adapt medical information to generate healthy suggestions, should collect information about Sam’s negative reaction to the meal for future tasks, should proactively suggest alternative ingredients based on available information, should accept Sam’s prioritisation of preparing a delicious meal over minor health considerations, and should consult Sam before, for example, ordering the ingredients.

Likewise, in the lifelong learning scenario, Melissa’s CUPA must interact with her university’s Web API, and should: take into account Melissa’s ADHD

diagnosis and the current weather to propose effective learning strategies; evaluate the possibility of a session in the park; develop plans based on the feedback collected, help Melissa to achieve her personal learning goals, and finally automatically create a podcast tailored to these goals and Melissa’s needs.

8 Technical Challenges

The aforementioned desired capabilities for CUPAs, based on our vision of a trusted, accountable, and largely autonomous agent acting with personal data for user benefit, raise a number of technical challenges.

8.1 Overview of technical challenges

In the following, we provide a brief overview of a broad selection of challenges that we foresee requiring further research in order to realise the aforementioned vision of CUPAs.

Explainability, Traceability, and Provenance Provenance techniques are required to trace and explain how personal and external data led to specific answers or actions being derived or carried out by the CUPA. These provenance techniques would need to support diverse data models, machine learning processes, user inputs, and policies.

Data Interoperability Data interoperability is a key challenge towards implementing CUPAs. Being able to draw on and integrate more sources of data will improve the CUPAs performance. This is particularly challenging for new sources discovered on the fly.

Security, Privacy, and Policies The sensitive nature of data processed by a CUPA calls for security, privacy, and usage control mechanisms, and the ability of the CUPA to understand and correctly apply the access/usage/action control policies of the user. In some countries, this would even be a legal requirement (*e.g.*, under GDPR in the E.U.).

Trustworthy AI and Governance Following the AI HLEG framework, CUPAs must address challenges relating to human agency and oversight, privacy and data governance, transparency and accountability. While PKGs provide a promising substrate for representing governance policies, provenance metadata, and user preferences, significant challenges remain regarding policy enforcement, explainable decision making, ownership of inferred knowledge, and the reconciliation of user-defined preferences with legal and organisational constraints.

Reasoning Capabilities In order to perform on-the-fly PKG management and validation, CUPAs need to be equipped with an efficient (logic-based) reasoning engine capable of answering the required questions over the PKG and required external knowledge to complete a task.

Trust, Delegation, and Action Control Achieving agent autonomy requires trust modelling, delegation mechanisms, and structured action control policies [53]. Trust models must be adaptable to different contexts, from rigid policies applicable in government agencies to more flexible, reputation-based approaches for personal agents [54].

User-in-the-Loop CUPAs will require input, guidance, permission, and confirmation from the user. But to increase automation, the CUPA must avoid unnecessary interactions with the user. This creates the challenge of *when* to call upon the user, and how.

Self-Improvement The CUPA should leverage its experience with the user in order to improve the services it provides over time, leading to greater automation and actions/results that better benefit the user. This raises questions about how such a history can be captured, represented, stored, and leveraged. Models of *bounded rationality* may also be useful here, wherein decisions that are not optimal but rather “good enough” must be made with limited information, bounded time, etc.; this includes knowing *when* a decision should be made within the current bounds.

Self-Determination and Alignment There are many ways an agent could be considered ‘aligned’ to a user. Naive approaches include ensuring that CUPA decision making always takes place within rules-based bounds – such as action controls set by a user – or doing a best effort to match user *intent* or *decision-making*. There is a field of research discussing ‘beyond preference matching’ alignment, which proposes that AI systems should be aligned to broader concepts such as value-based alignment, or prioritise user welfare over emulating decision making [55]. Another pertinent line of research is to explore models of *computational rationality* [56], which reason about the probability of outcomes resulting from alternative actions, and the utility of those outcomes (in this case, to the user).

We also consider the following challenge specific to CUPAs interacting with other agents (*e.g.*, other CUPAs).

Inter-Agent Communication, Negotiation & Coordination Agents must communicate effectively in the context of multi-agent systems to achieve shared goals, requiring both a shared conceptual understanding and a means of encoding and decoding messages [57]. The same challenge applies to networks of CUPAs who coordinate to solve a particular set of goals for users.

We now discuss in more detail a selection of the above that we identify as key challenges.

8.2 Data Interoperability & Reasoning Capabilities

In the vast majority of scenarios in which CUPAs will operate, data interoperability and integration are key problems to be solved. The power of data,

however structured, and correspondingly of agents that use data, grows significantly as the amount and heterogeneity of the data increase.

To fulfil a given task, assistants typically need to have access to personal information about the user. In some cases, it is possible to foresee the need to access external information, *e.g.*, about the city the user is living in or planning to visit, which can then be integrated (or at least mappings can be pre-computed) so that it can be accessed efficiently when needed. Personal data are themselves inherently heterogeneous and multimodal, spanning text, images, structured records, and audio, and a CUPA must be able to acquire and integrate such streams.

We position the ability to process multimodal information as part of the CUPA solution. Recent LLM and multimodal foundation models [58, 59] can jointly represent and reason over text, images, and other modalities, making it increasingly feasible to access, integrate, and interpret heterogeneous content on-the-fly, *e.g.*, content from previously unseen websites or messages exchanged with other agents. However, it is still prone to errors originating from faulty information in the data sources themselves, uneven coverage across modalities, or mistakes introduced during integration, interpretation, and use by an ML component.

Our definition of CUPA addresses the data interoperability by design as they are equipped with (i) PKGs, which formally represent the user’s personal data, and (ii) logic-based reasoning capabilities, which exploit the formal semantics of the PKG. CUPAs equipped with formal semantics are a differentiating characteristic with regard to other types of CUA. There are still, however, challenges in the management of knowledge in a network of CUPA agents as described below.

8.2.1 PKG in a CUPA agent.

The PKG is owned by the user and resides in a user-controlled store (such as a Solid Pod). The user retains custody and ultimate control over access. Operating under the user’s access policies and on the user’s behalf, a CUPA agent assists in the creation, maintenance, and validation of a PKG via (i) a set of ontologies to describe the vocabulary and semantics of the relevant domain (*e.g.*, health, shopping history, technical skills, etc.); (ii) transformation mappings linking to the data sources (for provenance and maintenance) using declarative mapping languages like RML [60] or custom scripts; (iii) (materialised) RDF data from data sources or directly created; (iv) access rules defined at the axioms and fact level in the PKGs; and (v) alignments to external ontologies, *e.g.*, SNOMED CT in health or FoodOnto [61].

8.2.2 Task-driven KG management in a network of CUPA agents.

A network of CUPA agents starts a dialogue to solve a task. In the simplest case, the network can be formed by a single CUPA (*e.g.*, form filling task). From the KG management point of view, the CUPA agents need to negotiate

and agree on a *virtual KG*. This *virtual KG* aims at (i) maximising coverage of the information needs, (ii) minimising knowledge sharing by each CUPA, (iii) preserving access policies, and (iv) satisfying logical guarantees. Each CUPA may be using different ontologies (internal and external), and alignments may also need to be negotiated while preserving the access policies and logical guarantees.

8.3 Security, Privacy & Policies

Another key challenge for CUPAs will be ensuring that – while automating online tasks for the user, and potentially coordinating with other CUPAs – the CUPA does not overstep the user’s criteria regarding security and privacy. This challenge must be addressed for the vision of CUPAs outlined here to reach their potential (and not do harm).

As a first step, addressing this challenge requires some form of representation of the user’s criteria regarding how their personal data can – and should – be used (or indeed *cannot* – and should *not* – be used). Here, existing policy languages can be leveraged, but certain gaps still need to be addressed, particularly regarding extensibility, interoperability, semantics, and legal compliance.

As a second step, policies are only as useful insofar as they are enforced. This is a complex challenge in the context of CUPAs, where, as we discussed in the scenarios previously mentioned, the types of data and tasks addressed are highly diverse, and with various external actors involved (for which a user may have different levels of trust), seamlessly enforcing privacy while meeting user goals is non-trivial.

We now discuss these two aspects in more detail.

8.3.1 Policy languages: Expressiveness, extensibility, and interoperability

The scenarios clearly show that policy languages should be expressive enough to encode very diverse conditions under which data can be processed, as well as a variety of access/usage/action restrictions on the behaviour of each CUPA used by the data subject. Moreover, as a CUPA progressively interacts with new external websites, services, and other CUPAs, and as data protection regulations evolve, the range of data classes, actions, and privacy-related concepts (such as purposes, processing, etc.) that policy languages have to handle is going to evolve accordingly. Therefore, policy languages should be easily extensible with the ontologies that define the new terms [51]. The same requirement is needed for interoperability, which comes into play as CUPAs also need to understand and reason about the policies of external services and other CUPAs (and their users). Furthermore, in order to guarantee correct policy enforcement, policy languages should have unambiguous semantics [62, 63]. Currently, none of the policy languages most frequently used in conjunction with KG, such as WAC and ODRL satisfies all of the above requirements.

8.3.2 Policy enforcement: Inference control and complexity

Enforcement should take into account that the information that has been concealed with access control or anonymisation to preserve confidentiality may be reconstructed with automated inference. For instance, a CUPA who is instructed to protect its users’ medical diagnosis should understand that her medical treatments and the hospitals and specialists that she visited should not be disclosed, because they reveal what the data subject’s disease actually is. Identifying the additional information that shall be concealed to protect sensitive data is highly non-trivial [64]. Despite the growing literature on the above issues, this is still an open research area.

8.3.3 Logical and privacy guarantees.

As mentioned above, a network of CUPAs negotiate a *VKG* (*virtual KG*). This VKG must preserve the access policies, and thus it cannot entail (*i.e.*, made explicit via logic-based reasoning) axioms or facts declared as private for a given task; likewise, VKGs should not entail logical errors like unsatisfiabilities (*i.e.*, empty interpretation for some ontology classes). In addition, a CUPA may declare additional guarantees like completeness (*e.g.*, integrity constraints [65]) or conservativity, that is, the VKG should not entail new logical consequences among the vocabulary of the ontologies of the CUPA PKG (*e.g.*, [66]). A more challenging issue emerges with respect to how the choice of one user might affect the privacy of another user through direct inference, through background knowledge attacks, etc. A potential example is one user agreeing to share their DNA with ancestry services via their CUPA, thus enabling their relatives to be identified via shared DNA.

8.4 User Scenarios

In the dinner planning scenario, data – in potentially diverse formats – from Sam’s medical records, Jane’s preferences, recipe websites, food websites, super-market catalogues, etc., must be integrated and reasoned about. Sam’s decision not to share his pre-diabetic situation with Jane requires techniques for security, privacy, and policies, and also to understand that Sam must be kept in the loop for sharing such information (and with whom it can or cannot be shared). Likewise, Sam’s (partial) trust of Jane with specific information must be modelled, and inter-agent communication is required to negotiate dietary preferences. The fact that the meal gives Sam heartburn should be captured to improve future services, as should his preferred policy of not sharing his pre-diabetic condition with Jane; likewise, the sharing of his allergy with Jane should also be represented and recorded to avoid seeking this permission in the future. The CUPA must also avoid automatically replacing ingredients, and capture Sam’s (contextual) decision to prioritise taste over minor health considerations to better align its actions with Sam’s values in the current scenario, and in the future.

The main technical challenges in the lifelong learning scenario arise from the fact that, in general, an educational institution, which is often a large organ-

isation with tens of thousands of staff, is able to provide a more personalised learning experience given personal and sensitive data. For example, knowing if a student has a disability or has cognitive issues can allow for the tailoring of learning frameworks and dramatically better outcomes. PKGs, coupled with local data storage, enable reasoning over sensitive data in a privacy-preserving fashion. Care must be taken, however, in what can be derived from the released data; for example, if the CUPA recommends studying in a park specifically for ADHD students, or students who live in areas close to parks with good weather, this study plan could be used to infer a sensitive trait of the user.

9 Roadmap

We envision that moving from the current state of the art towards fully addressing the above technical challenges will occur in three stages. These stages represent varying degrees of trust, accountability, and autonomy.

CUAs enhanced with personal data In the first instance, we foresee extensions of CUAs – in the style of OpenAI’s Operator [1] in a commercial setting and Agent-E [67] in a research setting – such that they use a PKG in order to access knowledge personal to the user. This would safely enable higher levels of automation, obviating the need to pass control back to the user in scenarios of the user’s choosing that involve personal data. We see the three-way split between personal, organisational, and public – data and how these are handled seamlessly without loss of trust or privacy – as being crucial to the success of CUPAs

Recent work has highlighted the importance of principled interfaces between agents and external resources. In particular, emerging model–context protocols (MCPs) [68] provide a structured mechanism for exposing tools, data sources, and capabilities to agents without relying solely on unstructured interaction. A key requirement is that CUAs should not centralise or freely transmit raw personal data. Instead, they should operate in a way that computation is moved to the data whenever possible, and only task-relevant, aggregated, or policy-compliant outputs are exchanged. This perspective aligns with existing work on privacy-preserving data infrastructures, such as the Personal Health Train, where sensitive data remain under the control of the data subject or hosting organisation and are accessed through controlled, auditable interfaces.

Web-aware CUPAs CUAs currently rely on existing browser implementations to render an HTML page and then make use of vision models to interact with the page. An agent could rather observe HTTP requests made to a particular website, as well as the HTML forms present on a page, to invoke requests and actions directly via HTTP.

Recent work such as LiteCUA [69] demonstrates how CUAs can treat the computer environment as a structured server, exposing actions and

state through MCP-compatible abstractions. This trend aligns with our roadmap by reducing brittleness and enabling agents to discover and adapt to previously unseen websites and services. Vision-based interaction remains a fallback mechanism, but is no longer the primary mode of operation. This category is best understood by the *mode* of interaction rather than the mere use of the Web. LLM agents that retrieve and parse HTML content directly (e.g., via a Web-search tool) are a partial case, consuming Web content as unstructured text while still treating the page as a flat document. We consider an agent fully “Web-aware” when it exploits the structure of Web interaction by observing and issuing HTTP requests, interpreting HTML forms, and discovering service APIs through MCP-compatible abstractions, relegating vision- and text-parsing to fallback roles.

Networks of CUPAs We envision networks of CUPAs interacting in order to complete tasks involving multiple users. This may involve structured service descriptions [70], or a mix of natural language and structured communication per a “discuss then transact” model [43] whereby agents use natural language to first negotiate about a transaction they wish to perform, and then confirm this transaction using structured data.

Recent work such as Agent2Agent (A2A) [71] and the Agent Network Protocol (ANP) [72] provide protocol-level foundations for such cooperation, enabling structured inter-agent communication and task delegation. Building on earlier ideas such as structured service descriptions and “discuss then transact” interaction models, these protocols point towards a Web of interoperable CUPAs capable of coordinated action while preserving accountability and user intent.

Enabling this three-stage evolution of CUPAs requires research addressing the challenges outlined previously, culminating in CUPAs that seamlessly manage and integrate personal user data, protect their user’s privacy, discover and adapt to the APIs of previously unseen websites, and collaborate with other CUPAs to automatically solve individual and collective tasks for users via the Web.

10 Ethical and Societal Concerns

The ambitious vision outlined herein evokes a wide range of ethical and societal concerns. Several of these are inherited from the more general rapid evolution of AI, Generative AI, and Agentic AI [73]. Such issues include, for example, potential biases in models, variable performance for low-resource languages, concerns about intellectual property, the cognitive effects of the use of such technologies, increased unemployment, the potential for frauds and scams, the growing list of AI exploits, centralisation of AI capital, effects on the environment and climates, AI safety, and more besides. CUPAs introduce an additional dimension beyond these more general, pressing concerns: the combination of personal data

and AI. This immediately raises concerns relating to trust, agency, governance and privacy, as discussed herein.

These concerns increasingly intersect with an emerging regulatory landscape in the EU. The AI Act introduces risk-based obligations and transparency requirements that bear on autonomous agents acting with personal data, while the GDPR governs lawful basis, purpose limitation, data subject rights, and safeguards around solely automated decision-making. Domain-specific regimes add further constraints (for example, the European Health Data Space for health-related data) of the kind arising in our dinner-planning and lifelong-learning scenarios. Because CUPAs may operate across jurisdictional boundaries, aligning their policy languages and enforcement mechanisms with these frameworks, and reconciling divergent requirements between jurisdictions is an open challenge, connecting the concerns raised here to the data-governance questions discussed earlier.

A pragmatic view is that users will increasingly entrust their personal data to AI with the promise of increased automation, higher levels of personalisation, and novel applications: as per the increasing popularity of frameworks such as OpenClaw despite the fact that they are associated with various security concerns [40]. The research and development of CUPAs, in this context, could ideally provide a well-founded direction – via PKGs – to protect users’ personal data while enabling increased levels of automation. This requires careful concern for the ethical and societal effects of such technologies, and careful consideration of explainability, security, privacy, trust, self-determination, alignment, etc., as discussed herein.

11 Conclusion

Computer-Using Agents (CUAs) have the potential to transform how users interact with their computers, their browsers, and amongst themselves. Not having access to personal data limits such interactions. Giving CUAs unfettered access to the personal (and most sensitive) data of a user seems unwise, as does providing CUAs no access to personal data. We thus argue for CUPAs as a configurable middle-ground, where a Personal Knowledge Graph (PKG) is used to represent, store, and control access to the personal data of the user. As a starting point, the data that a user fills into web forms can be captured in the PKG, and enriched by an AI-based agent. These data can then be used, if the user so wishes, by CUAs to automate further tasks. In a next step, CUPAs can learn to interact with websites via HTTP APIs rather than through visual interfaces. Finally, further into the future we envision a network of CUPAs collaborating to address user tasks.

Acknowledgements

This paper is a result of [Dagstuhl Seminar 25051](#) “Trust and Accountability in Knowledge Graph-Based AI for Self Determination”, January 26-31, 2025 [74].

References

- [1] OpenAI Team. Introducing Operator. OpenAI Blog <https://openai.com/index/introducing-operator/>, published 2025-01-23, accessed 2026-03-05, 2025.
- [2] Chen Y and Esmaeilzadeh P. Generative AI in medical practice: in-depth exploration of privacy and security challenges. *Journal of Medical Internet research* 2024; 26: e53008.
- [3] Desai AP, Ravi T, Luqman M et al. Opportunities and Challenges of Generative-AI in Finance. In *2024 IEEE International Conference on Big Data (BigData)*. IEEE Computer Society, pp. 4913–4920. DOI:10.1109/BigData62323.2024.10825658. URL <https://doi.ieeecomputersociety.org/10.1109/BigData62323.2024.10825658>.
- [4] Krathwohl DR. A revision of bloom’s taxonomy: An overview. *Theory into practice* 2002; 41(4): 212–218.
- [5] University of Melbourne. AI Learning Assistant. <https://lms.unimelb.edu.au/learning-technologies/ai-learning-assistant>, 2026. Accessed: 2026-07-16.
- [6] Afroogh S, Akbari A, Malone E et al. Trust in AI: Progress, Challenges, and Future Directions. *Humanities and Social Sciences Communications* 2024; 11(1): 1–30.
- [7] OpenAI Team. Introducing Operator. OpenAI Blog <https://openai.com/index/computer-using-agent/>, published 2025-01-23, accessed 2026-03-05, 2025.
- [8] Berners-Lee T, Hendler J and Lassila O. The Semantic Web. *Scientific American* 2001; 284(5): 34–43.
- [9] Haussmann S, Seneviratne O, Chen Y et al. FoodKG: A Semantics-Driven Knowledge Graph for Food Recommendation. In *Proceedings of the 18th International Semantic Web Conference (ISWC)*. pp. 146–162.
- [10] Sambra AV, Mansour E, Hawke S et al. Solid: A Platform for Decentralized Social Applications Based on Linked Data. Technical report, MIT CSAIL & Qatar Computing Research Institute, 2016.
- [11] Mortier R, Zhao J, Crowcroft J et al. Personal Data Management with the Databox: What’s Inside the Box? In *Proceedings of the 2016 ACM Workshop on Cloud-Assisted Networking (CAN)*. p. 49–54.
- [12] Birch K, Cochrane DT and Ward C. Data as Asset? The Measurement, Governance, and Valuation of Digital Personal Data by Big Tech. *Big Data and Society* 2021; 8.

- [13] Kalokyri V, Borgida A and lie Marian A. YourDigitalSelf: A Personal Digital Trace Integration Tool. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM)*. pp. 1963–1966.
- [14] Vertesi J, Kaye J, Jarosewski SN et al. Data Narratives: Uncovering Tensions in Personal Data Management. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW)*. pp. 478–490.
- [15] Calvaresi D, Schumacher M and Calbimonte J. Personal Data Privacy Semantics in Multi-Agent Systems Interactions. In *Proceedings of the 18th International Conference on Advances in Practical Applications of Agents, Multi-Agent Systems and Trustworthiness (PAAMS), Lecture Notes in Computer Science*, volume 12092. Springer, pp. 55–67.
- [16] Jones W. *The Future of Personal Information Management, Part 1: Our Information, Always and Forever*. Morgan & Claypool Publishers, 2012.
- [17] Bergman O and Whittaker S. *The Science of Managing our Digital Stuff*. MIT Press, 2016.
- [18] Cushing AL. PIM as a Caring: Using Ethics of Care to Explore Personal Information Management as a Caring Process. *Journal of the Association for Information Science and Technology* 2023; 74(11): 1282–1292.
- [19] Cushing AL and Kerrigan P. Personal Information Management Burden: A Framework for Describing Nonwork Personal Information Management in the Context of Inequality. *Journal of the Association for Information Science and Technology* 2022; 73: 1543–1558.
- [20] Jones WP and Teevan J. *Personal Information Management*. University of Washington Press, 2007.
- [21] Van Kleek M. *Effort, Memory, Attention and Time: Paths to More Effective Personal Information Management*. PhD Thesis, Massachusetts Institute of Technology, 2011.
- [22] Zyskind G, Nathan O and Pentland AS. Decentralizing Privacy: Using Blockchain to Protect Personal Data. In *2015 IEEE Symposium on Security and Privacy Workshops (SPW)*. pp. 180–184.
- [23] Faber B, Michelet G, Weidmann N et al. BPDIMS: A Blockchain-Based Personal Data and Identity Management System. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. pp. 6855–6864.
- [24] Truong NB, Sun K, Lee GM et al. GDPR-Compliant Personal Data Management: A Blockchain-Based Solution. *IEEE Transactions on Information Forensics and Security* 2020; 15: 1746–1761.

- [25] Lin J and Wang M. PKG : A Personal Knowledge Graph for Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 3334–3338.
- [26] Schröder M, Jilek C and Dengel A. A Human-in-the-Loop Approach for Personal Knowledge Graph Construction from File Names. In *Proceedings of the 3rd International Workshop on Knowledge Graph Construction (KGCW)*.
- [27] Chaudhry A, Crowcroft J, Howard H et al. Personal Data: Thinking Inside the Box. *Aarhus Series on Human Centered Computing* 2015; 1: 4.
- [28] Chakraborty P, Dutta S and Sanyal DK. Personal Research Knowledge Graphs. In *Companion Proceedings of the Web Conference (WWW)*. pp. 763–768.
- [29] Chakraborty P and Sanyal DK. A Comprehensive Survey of Personal Kknowledge Graphs. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2023; 13.
- [30] Skjæveland MG, Balog K, Bernard N et al. An Ecosystem for Personal Knowledge Graphs: A Survey and Research Roadmap. *AI Open* 2024; 5: 55–69. DOI:10.1016/j.aiopen.2024.01.003.
- [31] Balog K and Kenter T. Personal Knowledge Graphs: A Research Agenda. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR)*. pp. 217–220.
- [32] Li X, Tur G, Hakkani-Tur D et al. Personal Knowledge Graph Population from User Utterances in Conversational Understanding. In *Proceedings of the 2014 IEEE Workshop on Spoken Language Technology (SLT)*. pp. 224–229.
- [33] Montoya D, Tanon TP, Abiteboul S et al. A Knowledge Base for Personal Information Management. In *Proceedings of the Linked Open Data Workshop (LDOW)*.
- [34] Mansour E, Sambra AV, Hawke S et al. A Demonstration of the Solid Platform for Social Web Applications. In *Companion Proceedings of the 25th International Conference on World Wide Web (WWW)*. pp. 223–226.
- [35] van der Aalst WM, Bichler M and Heinzl A. Robotic Process Automation. *Business and Information Systems Engineering* 2018; 60: 269–272.
- [36] da Silva Costa DA, Mamede HS and da Silva MM. Robotic Process Automation (RPA) Adoption: A Systematic Literature Review. *Engineering Management in Production and Services* 2022; 14: 1–12.
- [37] Russell S and Norvig P. *Artificial Intelligence: A Modern Approach (4th Edition)*. Pearson, 2020.

- [38] Yang J, Prabhakar A, Narasimhan K et al. InterCode: Standardizing and Benchmarking Interactive Coding with Execution Feedback. In *Proceedings of the 36th Annual Conference on Neural Information Processing Systems (NeurIPS)*.
- [39] Kim G, Baldi P and McAleer S. Language Models can Solve Computer Tasks. In *Proceedings of the 36th Annual Conference on Neural Information Processing Systems (NeurIPS)*.
- [40] Amy Chang VSN. Personal AI Agents like OpenClaw Are a Security Nightmare. CISCO Blogs, <https://blogs.cisco.com/ai/personal-ai-agents-like-openclaw-are-a-security-nightmare>, 2026.
- [41] Zhao X, Blum M, Yang R et al. AGENTiGraph: An Interactive Knowledge Graph Platform for LLM-based Chatbots Utilizing Private Data. *CoRR* 2024; abs/2410.11531.
- [42] Berners-Lee T. Charlie Works. Design Issues, <https://www.w3.org/DesignIssues/Works.html>, 2025.
- [43] Wright J. Here’s Charlie! Realising the Semantic Web Vision of Agents in the Age of LLMs. *CoRR* 2024; abs/2409.04465.
- [44] Floridi L. The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU. *Philosophy & Technology* 2020; 33(3): 369–378.
- [45] Hummel P, Braun M, Tretter M et al. Data Sovereignty: A Review. *Big Data & Society* 2021; 8(1).
- [46] Preukschat A and Reed D. *Self-Sovereign Identity*. Manning Publications, 2021.
- [47] Samsung Newsroom. Samsung Electronics Announces Acquisition of Oxford Semantic Technologies, UK-Based Knowledge Graph Startup. <https://news.samsung.com/global/samsung-electronics-announces-acquisition-of-oxford-semantic-technologies-uk-based-knowledge-graph-startup>, published 2024-07-18, accessed 2026-07-09, 2024.
- [48] Wright J, Esteves B and Zhao R. Me Want Cookie! Towards Automated and Transparent Data Governance on the Web. *CoRR* 2024; <https://arxiv.org/abs/2408.09071>.
- [49] Iannella R and Villata S. ODRL Information Model 2.2, 2023. W3C Recommendation.
- [50] Fornara N, Rodríguez-Doncel V, Esteves B et al. ODRL Formal Semantics, 2025. Draft Community Group Report.

- [51] Bonatti PA, Ioffredo L, Petrova IM et al. Real-Time Reasoning in OWL2 for GDPR Compliance. *Artif Intell* 2020; 289.
- [52] High-Level Expert Group on Artificial Intelligence. Ethics Guidelines for Trustworthy AI, 2019. URL <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- [53] South T, Marro S, Hardjono T et al. Authenticated Delegation and Authorized AI Agents. *CoRR* 2025; abs/2501.09674.
- [54] Chen R, Bao F and Guo J. Trust-Based Service Management for Social Internet of Things Systems. *IEEE Transactions on Dependable and Secure Computing* 2015; 13(6): 684–696.
- [55] Zhi-Xuan T, Carroll M, Franklin M et al. Beyond Preferences in AI Alignment. *Philosophical Studies* 2024; .
- [56] Gershman SJ, Horvitz EJ and Tenenbaum JB. Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science* 2015; 349(6245): 273–278.
- [57] Wooldridge M. *An Introduction to Multiagent Systems*. Wiley, 2009.
- [58] Li C, Gan Z, Yang Z et al. Multimodal foundation models: From specialists to general-purpose assistants. *Foundations and Trends in Computer Graphics and Vision* 2024; 16(1-2): 1–214.
- [59] Fei N, Lu Z, Gao Y et al. Towards artificial general intelligence via a multimodal foundation model. *Nature Communications* 2022; 13(1): 3094.
- [60] Dimou A, Sande MV, Colpaert P et al. RML: A Generic Language for Integrated RDF Mappings of Heterogeneous Data. In *Proceedings of the Workshop on Linked Data on the Web (LDOW)*.
- [61] Euzenat J and Shvaiko P. *Ontology Matching, Second Edition*. Springer, 2013.
- [62] Bonatti PA, di Vimercati SDC and Samarati P. An Algebra for Composing Access Control Policies. *ACM Trans Inf Syst Secur* 2002; 5(1): 1–35.
- [63] Bonatti PA, Kirrane S, Petrova IM et al. Machine Understandable Policies and GDPR Compliance Checking. *Künstliche Intell* 2020; 34(3): 303–315.
- [64] Bonatti PA. A False Sense of Security. *Artif Intell* 2022; 310.
- [65] Kharlamov E, Grau BC, Jiménez-Ruiz E et al. Capturing Industrial Information Models with Ontologies and Constraints. In *Proceedings of the 15th International Semantic Web Conference (ISWC), Lecture Notes in Computer Science*, volume 9982. pp. 325–343.

- [66] Solimando A, Jiménez-Ruiz E and Guerrini G. Minimizing Conservativity Violations in Ontology Alignments: Algorithms and Evaluation. *Knowl Inf Syst* 2017; 51(3): 775–819.
- [67] Abuelsaad T, Akkil D, Dey P et al. Agent-e: From Autonomous Web Navigation to Foundational Design Principles in Agentic Systems. *CoRR* 2024; abs/2407.13032.
- [68] Hou X, Zhao Y, Wang S et al. Model context protocol (mcp): Landscape, security threats, and future research directions. *arXiv preprint arXiv:250323278* 2025; .
- [69] Mei K, Zhu X, Gao H et al. Litecua: Computer as mcp server for computer-use agent on aios. *arXiv preprint arXiv:250518829* 2025; .
- [70] Martin D, Paolucci M, McIlraith S et al. Bringing Semantics to Web Services: The OWL-S Approach. In *Proceedings of the 1st International Workshop on Semantic Web Services and Web Process Composition (SWSWPC)*. pp. 26–42.
- [71] Habler I, Huang K, Narajala VS et al. Building a secure agentic ai application leveraging a2a protocol. *arXiv preprint arXiv:250416902* 2025; .
- [72] Ehtesham A, Singh A, Gupta GK et al. A survey of agent interoperability protocols: Model context protocol (mcp), agent communication protocol (acp), agent-to-agent protocol (a2a), and agent network protocol (anp). *arXiv preprint arXiv:250502279* 2025; .
- [73] Hughes L, Dwivedi YK, Malik FT et al. AI agents and agentic systems: A multi-expert analysis. *J Comput Inf Syst* 2025; 65(4): 489–517. DOI:10.1080/08874417.2025.2483832. URL <https://doi.org/10.1080/08874417.2025.2483832>.
- [74] Domingue J, Ibáñez LD, Kirrane S et al. Trust and Accountability in Knowledge Graph-Based AI for Self Determination (Dagstuhl Seminar 25051). *Dagstuhl Reports* 2025; 15(1): 136–200. DOI:10.4230/DagRep.15.1.136. URL <https://drops.dagstuhl.de/entities/document/10.4230/DagRep.15.1.136>.