# Personal Digital Agent Protocol (pdap)

v4
January 29, 2023

pdap@ietf.org

# Background

- Delegated Authorization Agent (eg: IETF GNAP)
  - eg: a money manager as authorized agent that is separate from your bank or stock exchange
  - vs: facebook or Google drive where the “platform” controls both authorization and the data
- Authorization Request (eg: IETF Rich Authorization Requests and GNAP)
  - A standardized request format that can include a Verifiable Credential for accountability.
- Authorization Capability (eg: zcap-ld, UCAN)
  - An authorization token that can be further attenuated by a delegate of the token holder,
  - offers improved security in many applications.
- Agent (eg: GNAP Authorization Server)
  - A policy-driven server that evaluates requests to issue authorization capabilities.
- User Agent (eg: Chrome, Apple Wallet, MyChart health record)
  - A digital interface for a human user.
- Resource Server (eg: Google Drive)
  - An access-protected digital interface to personal or private data.
- Policy Description Language (eg: CEDAR)
  - Allows coding an agents authorization policies in a way that is both human and machine readable.

## Example: A navigator combines a health diary with EHR review to create a timeline

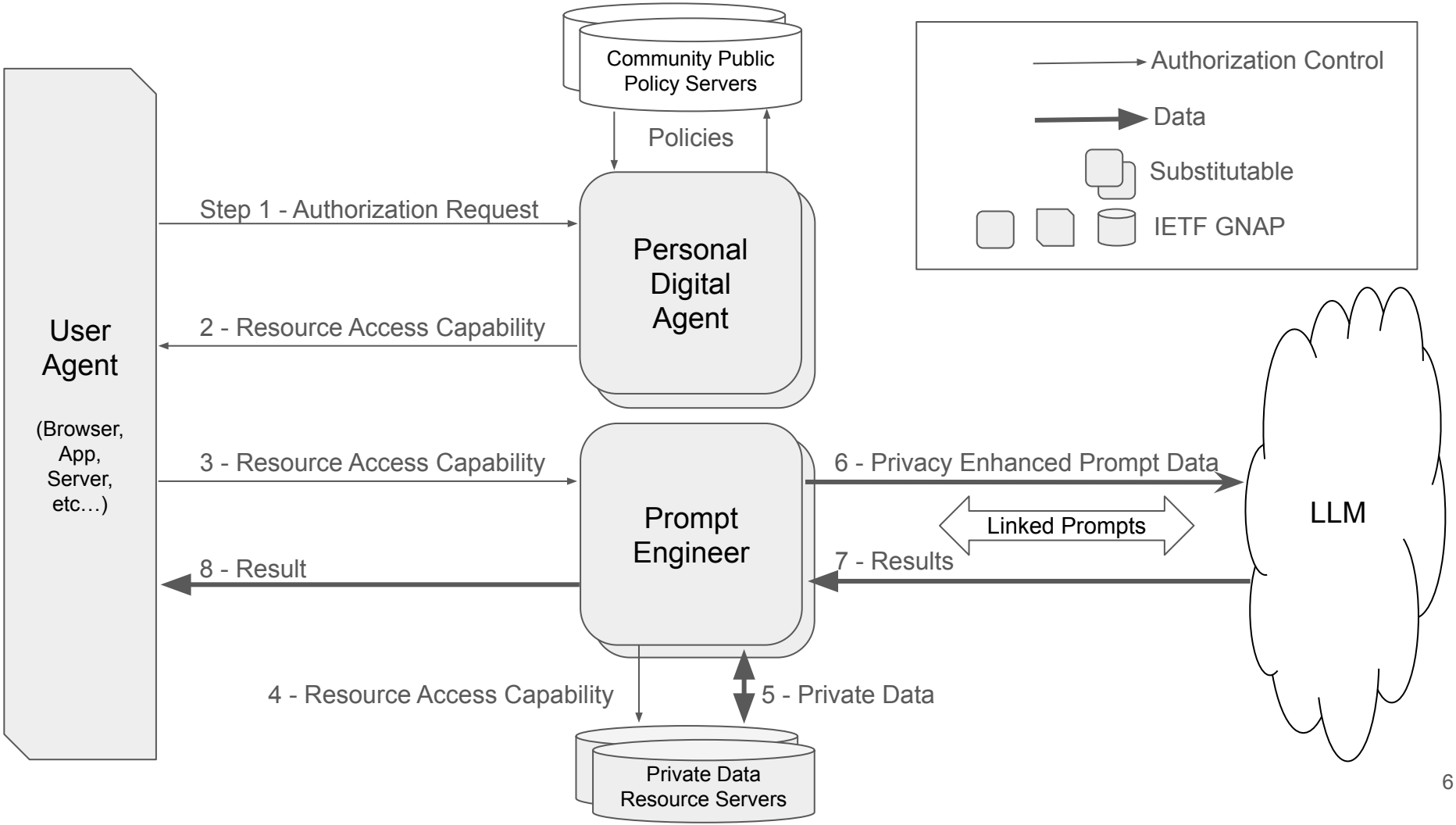
Alice is a 47 y/o female with a history of urinary tract infections. Her physician prescribed an antibiotic to be taken as needed. Alice keeps a diary of her symptoms, clinician messages, and other health-related events, including UTIs. This simple diary complements the information in her various electronic health records (EHRs).

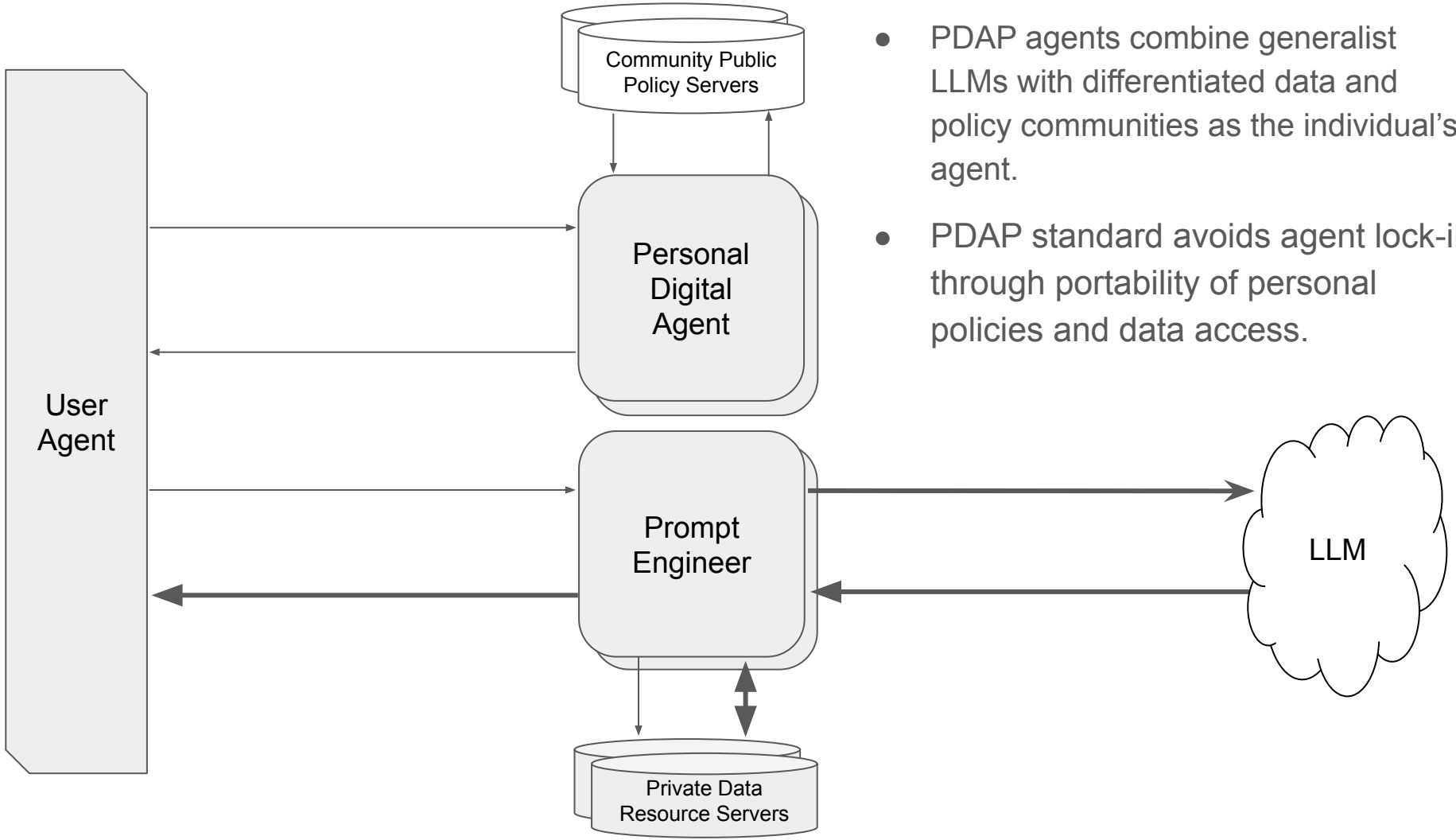
When a new symptom appears, Alice updates her diary and uses a health navigator service to help her decide whether to take her prescribed antibiotic or contact her physician. The navigator feeds her diary and all of her EHRs into the GPT 4 large language model (LLM) with instructions to create a timeline showing Alice's symptoms, prescriptions and self-administered treatments. A link to the updated timeline is added to the diary on her phone. If she decides to contact her physician, she can choose to send the timeline by simply adding the link to her message.

- Chat-style linked-prompt large language models (LLM) can improve user experience but users face business and safety risks when private data is exposed to external LLMs.
- Risks can be mitigated by introducing **a context-specific prompt engineer** to mediate and pre-process the user's private data.
- LLMs are generalist by design. A specialized prompt engineer is likely to adopt a user lock-in business model as they develop long-term relationships with the LLM user.
- A standard that facilitates switching prompt engineers gives users agency over their service providers.
- Such a standard also adds value to the LLM vendor when it encourages the users to bypass a third-party prompt engineer because they trust the LLM to be privacy preserving and accountable.
- A **personal digital agent** controls access to a user's private data based on user-controlled authorization policies.
- Unlike the prompt engineer, the digital agent need not see or touch the private data itself. Prompt engineers can offer both authorization policy and data service but that combination reduces the user's agency much as today's social media platforms do.

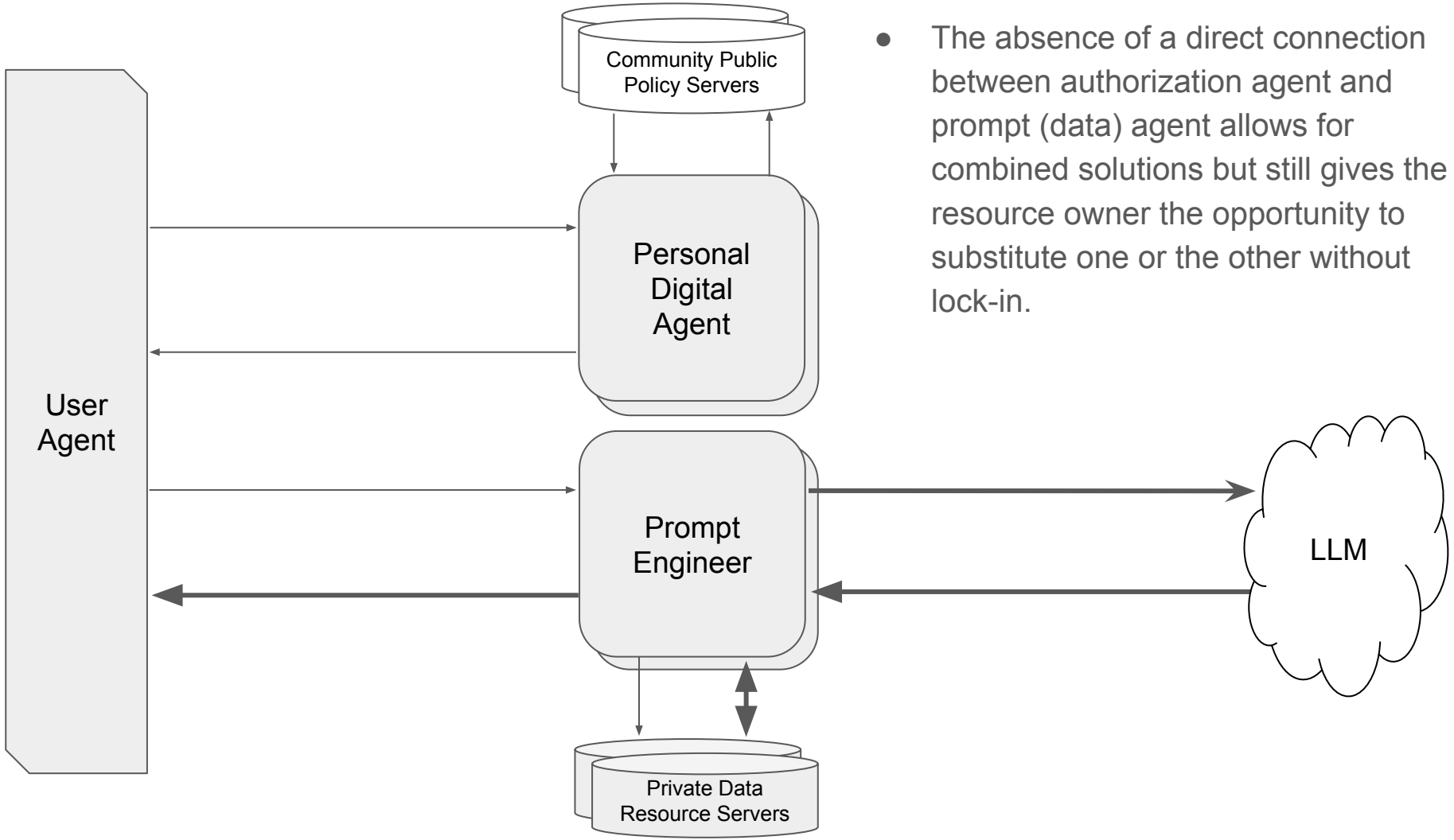
# PDAP: A standard to separate the agent from the prompt engineer.

- Separate processing of private policies from private data,
- substitutability of the personal agents based on policy portability,
- substitutability of the prompt engineer based on keeping them separate from the resource server.
- Nice to have: human readability of private policies that control the agent.

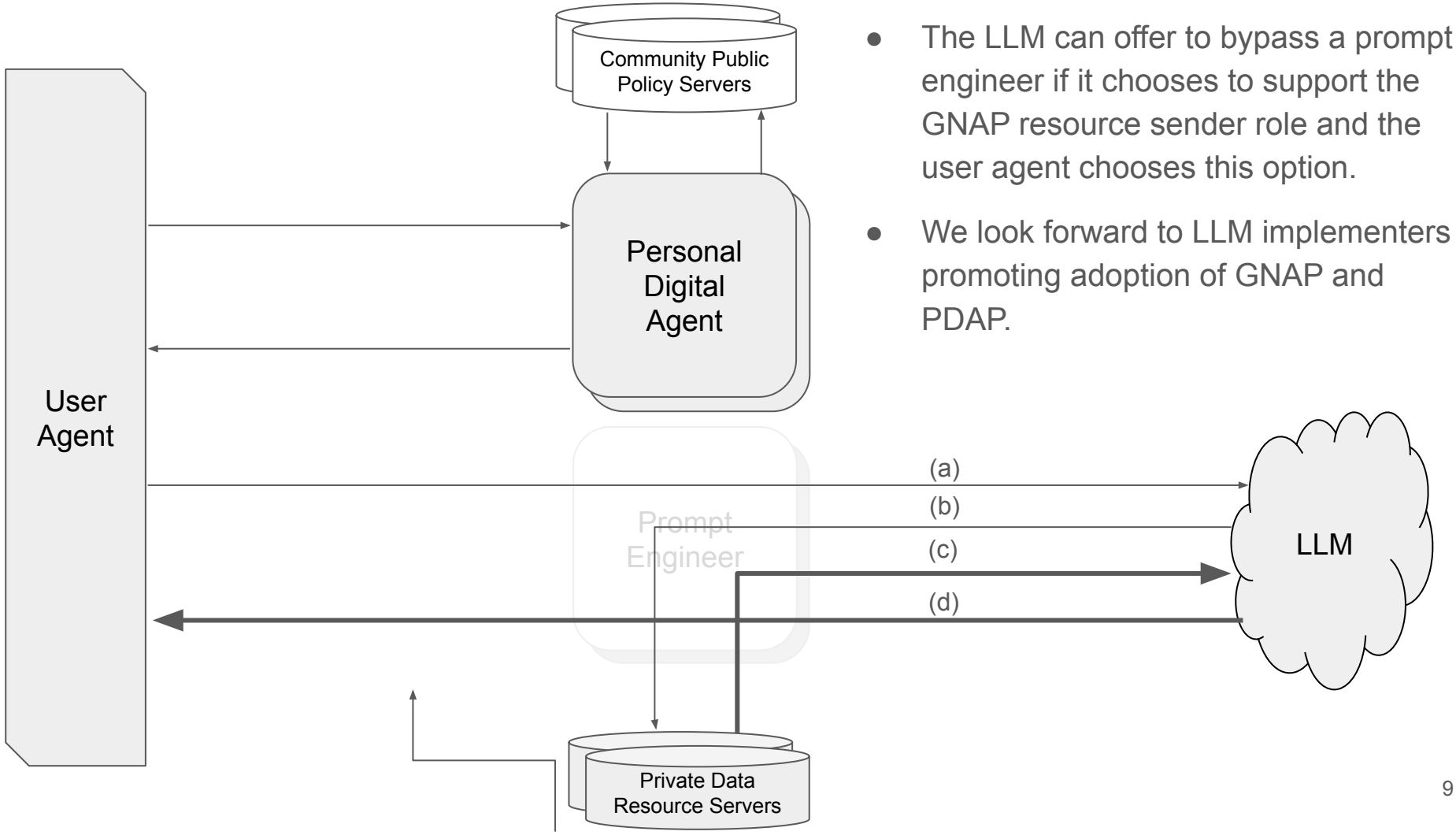

Community Public Policy Servers
Policies
Authorization Control
Data
Substitutable
IETF GNAP
User Agent
(Browser, App, Server, etc…)
Step 1 - Authorization Request
2 - Resource Access Capability
Personal Digital Agent
3 - Resource Access Capability
6 - Privacy Enhanced Prompt Data
Linked Prompts
7 - Results
8 - Result
Prompt Engineer
LLM
4 - Resource Access Capability
5 - Private Data
Private Data Resource Servers



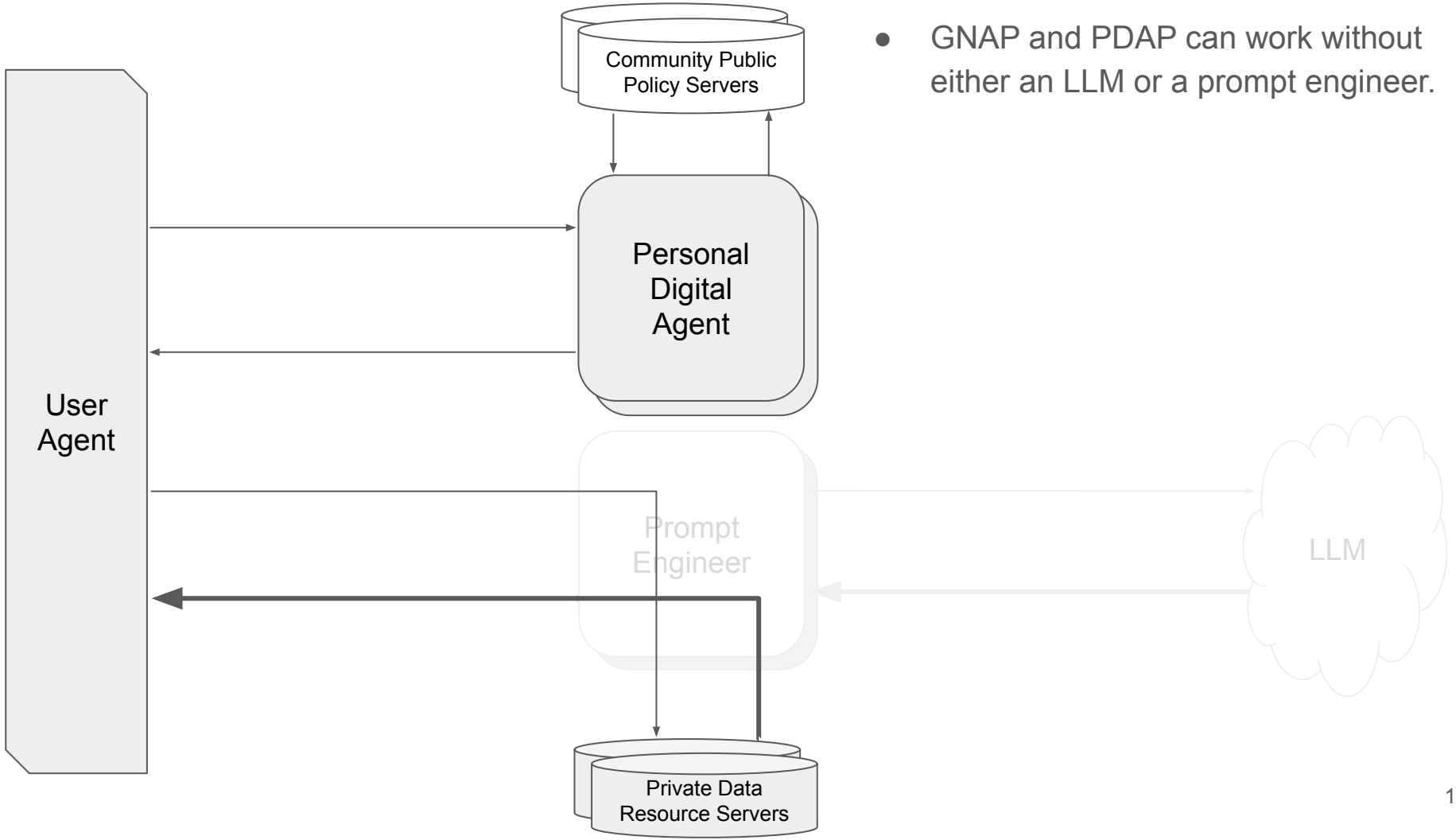

- PDAP agents combine generalist LLMs with differentiated data and policy communities as the individual's agent.
- PDAP standard avoids agent lock-in through portability of personal policies and data access.



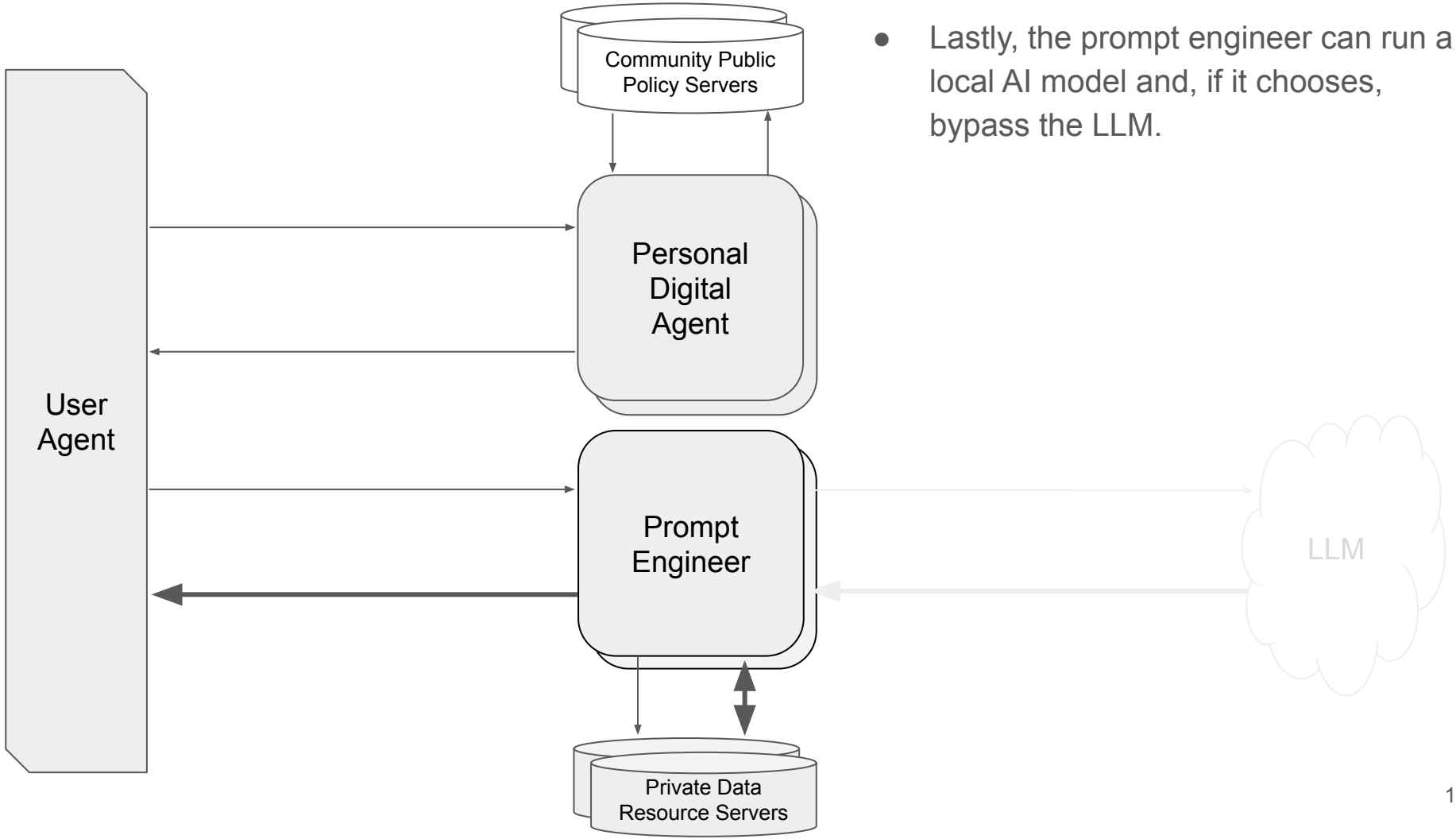

- The absence of a direct connection between authorization agent and prompt (data) agent allows for combined solutions but still gives the resource owner the opportunity to substitute one or the other without lock-in.




- The LLM can offer to bypass a prompt engineer if it chooses to support the GNAP resource sender role and the user agent chooses this option.
- We look forward to LLM implementers promoting adoption of GNAP and PDAP.




- GNAP and PDAP can work without either an LLM or a prompt engineer.




- Lastly, the prompt engineer can run a local AI model and, if it chooses, bypass the LLM.

# PDAP Principals

- Profiling GNAP to protect resource owner choice of authorization agent and prompt engineer.
- Standardizing the GNAP authorization token as a capability to promote prompt engineer substitutability and resource server interoperability.
- Standardizing the policy description language to promote interoperability across policy communities.
- Choosing a policy description language that is human readable for resource owner trust and AI alignment. (optional)

# References


- Can Generalist Foundation Models Outcompete Special-Purpose Tuning? Case Study in Medicine (Microsoft) https://arxiv.org/abs/2311.16452
- Use of GPT-4 to Diagnose Complex Clinical Cases (NEJM AI) https://ai.nejm.org/doi/pdf/10.1056/AIp2300031
- pdap@ietf.org invite link https://www.ietf.org/mailman/listinfo/pdap
- Patient survey: 63% say AI puts their data at risk (CARTA) https://7796197.fs1.hubspotusercontent-na1.net/hubfs/7796197/GRAPHICS/23_CAR011_Consumer_Survey_Infographic_F%20%281%29.pdf
- Grant Negotiation and Authorization Protocol - IETF GNAP https://datatracker.ietf.org/wg/gnap/about/
- AI-Generated Clinical Summaries Require More Than Accuracy https://jamanetwork.com/journals/jama/fullarticle/2814609
- [Pdap] A realistic patient use case for AI in healthcare https://mailarchive.ietf.org/arch/msg/pdap/PTfDRuY3JSQz9HQJfR_jRK24iGM/