



Beyond LLMs in a Loop

Building Trusted Agents
in Regulated Industries

AGNTCon + MCPCon
Friday, October 23, 2026



Lucas Beeler · Aerospike

An agent before LLMs?

Google News: [launched in September 2002](#). Computers grouped articles by semantic “story.” World amazed. My attempt to reverse engineer:

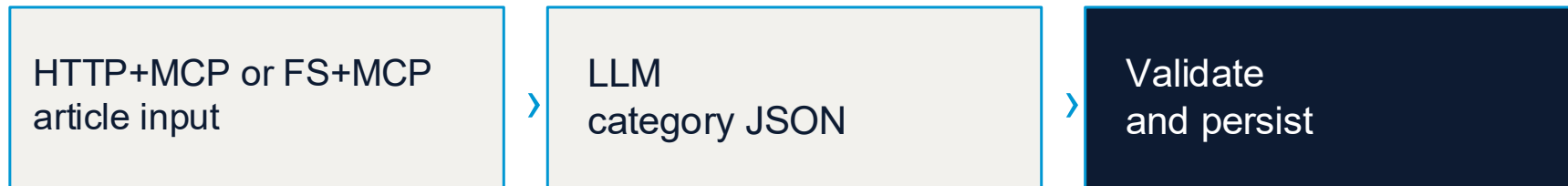


- Michigan EECS graduate-level course project, 2006-2007. No LLM. No ANN.
- System could **run continuously** and **respond to asynchronous arrival of new articles** and **decide**: does this belong to an existing semantic category, or do I have to create a new one?

Raise your hand if you believe **this is an “agent?”**



A LangGraph / LLM article-classification loop



Asynchronous new arrivals resume the grouping workflow. System decides if a new article falls within an existing semantic category or if a new one must be created.

Raise your hand if you believe **this is an “agent?”**





Intelligent agency and computational substrate

The engineering question is
what the system can do
and who ***authorizes*** it.

A working definition

A system pursues a
delegated objective,
observes its environment
and results, and adapts
its actions over time.

A useful definition for this talk, **not a
universal consensus**. Adapted from
Woolridge & Jennings 1995 [*Intelligent
Agents: Theory and Practice*](#).

01 My Declarative Statements

I believe LLMs are **the most important technology** to enter computing in my lifetime.



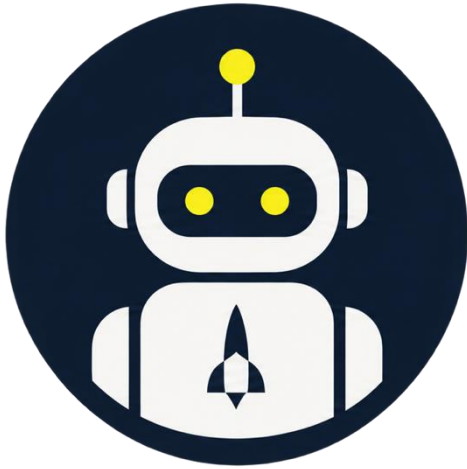
01 My Declarative Statements

I believe LLMs are **the most important technology** to enter computing in my lifetime.

Integrating LLMs with things we used to call “AI”—like knowledge representation, formal reasoning, and constraint optimization—is **essential for *trust* in regulated industries.**



“Trusted Agents in Regulated Industries”



Agent



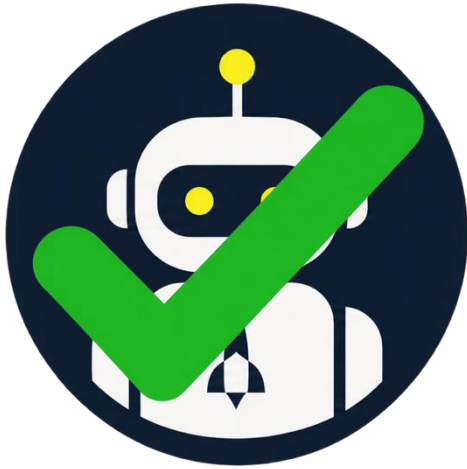
Trust



Regulated
Industry



“Trusted Agents in Regulated Industries”



Agent



Trust



Regulated
Industry

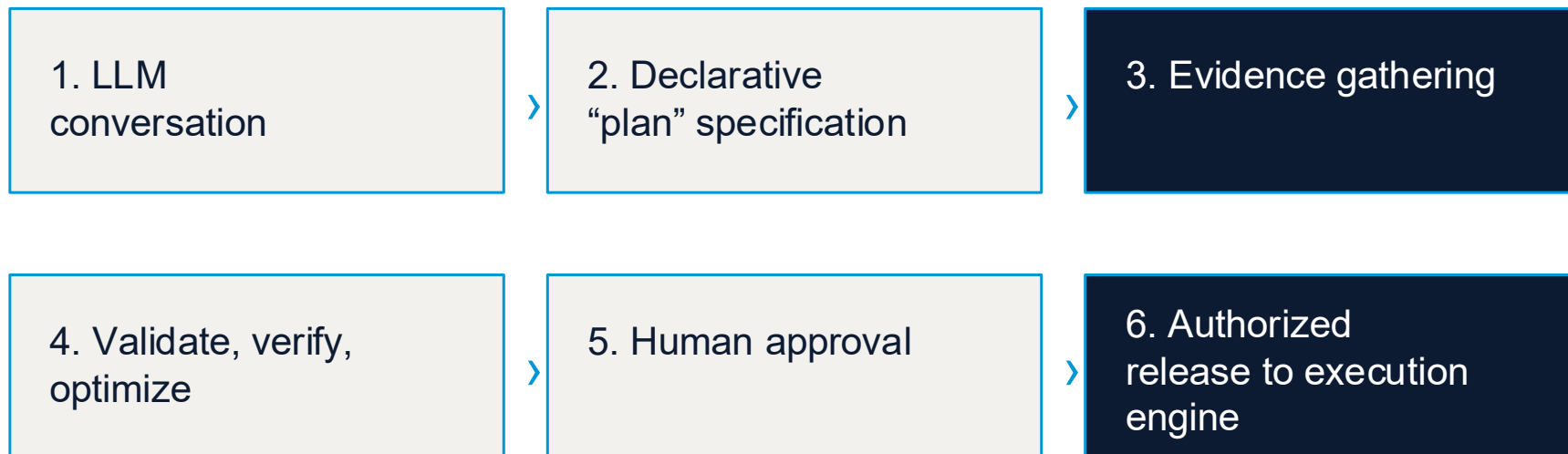


What makes an agent trustworthy?

- It isn't that the agent will always make the correct decision.
- It is the assurance that...
 - Under explicitly stated assumptions
 - An accurate perception of its environment (observed data / **evidence**)
 - The existence of defined prohibited or undesirable behaviors
- ...the defined prohibited or undesirable behaviors **cannot occur**.
- While remaining **accountable**
 - All inputs / observations are treated as **evidence** (attributable sources, timestamps, retrieval methods, etc.)
 - Decisions are made in a way that is more likely to survive audit or litigation



Trusted agent architecture



OPTIONAL

Candidate plans repeat validation until 1-5 are satisfied.



An illustrative example: surgical staffing



Care Coordinator



Cardiologist



Open-Heart Surgeon

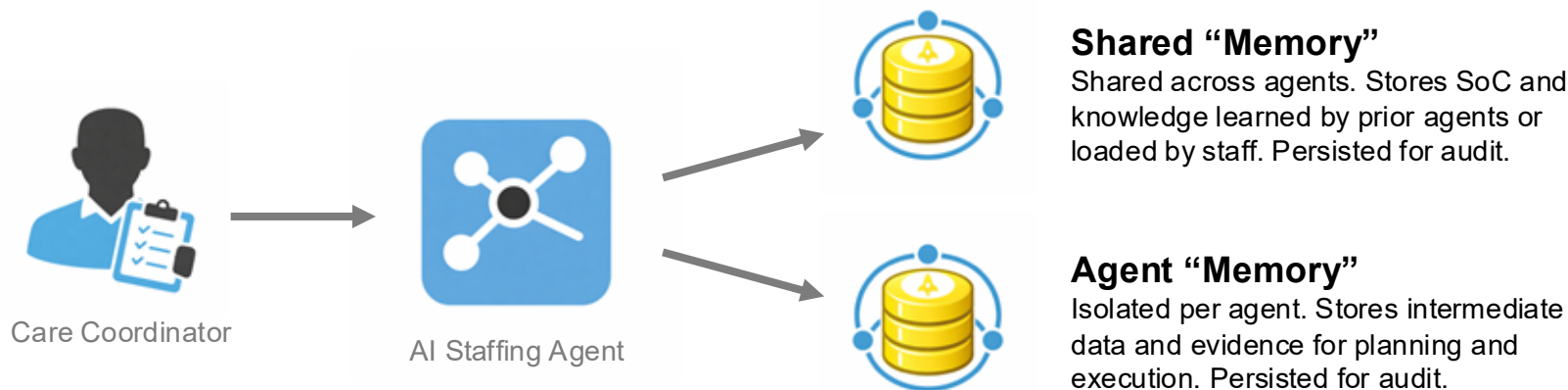


ICU Nurse

- Every surgery requires a team that is certified, available, in-network, and cost-effective—all at the same time.
- Even larger health systems don't always have every specialist on staff—the team must be assembled on demand, often in life-or-death situations



Evidence Gathering



- Care coordinator defines requirements and constraints via a conversational interface—iterating until a schema-compliant, declarative staffing strategy (“plan specification”) can be produced
- Agent retrieves data from shared memory and external systems, engages providers via HSPs, and solves a constrained optimization problem to assemble the best care team at the lowest cost.



Evidence

```
<CredentialEvidence id="EVD-2026-001842"
  schemaVersion="1.2">

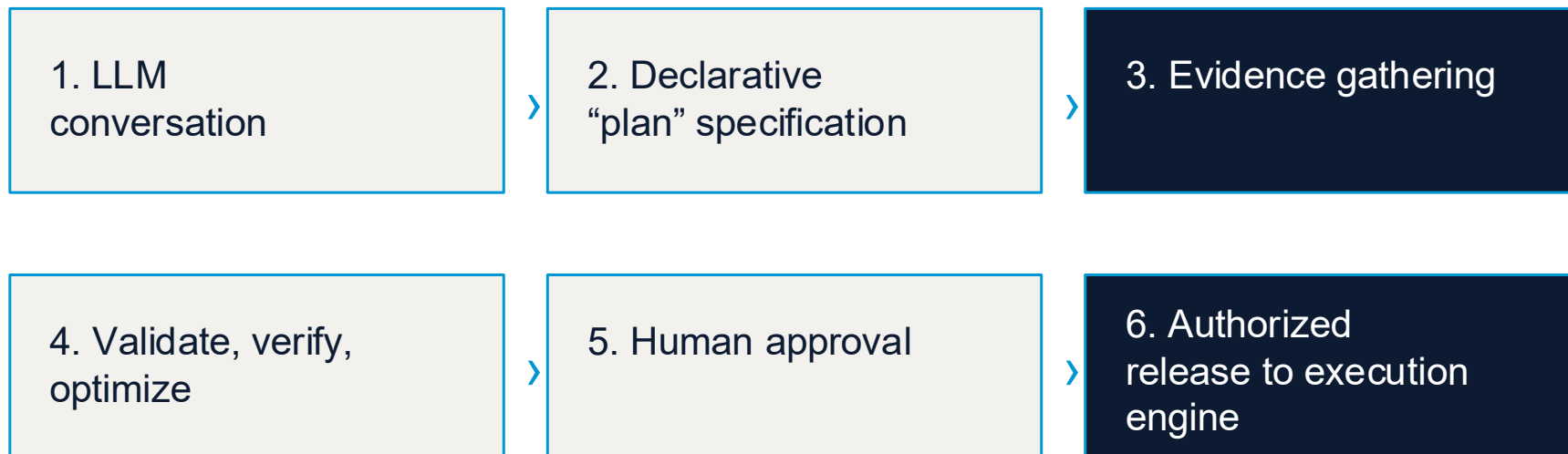
  <Subject providerId="HOSP-00428">
    <Name>Dr. Eliana Voss</Name>
  </Subject>

  <Assertion>
    <Credential>Thoracic and Cardiac Surgery</Credential>
    <IssuingAuthority>
      American Board of Thoracic Surgery
    </IssuingAuthority>
    <ReportedStatus>CERTIFIED</ReportedStatus>
    <StatusAsOf>2026-10-22T14:35:00Z</StatusAsOf>
  </Assertion>

  <Verification>
    <Verifier>Northstar Credential Services</Verifier>
    <VerificationId>NCV-882741</VerificationId>
    <PrimarySource>American Board of Thoracic Surgery</PrimarySource>
    <PrimarySourceMethod>VERBAL</PrimarySourceMethod>
    <PrimarySourceCheckedAt>
      2026-10-22T14:35:00Z
    </PrimarySourceCheckedAt>
```



Trusted agent architecture

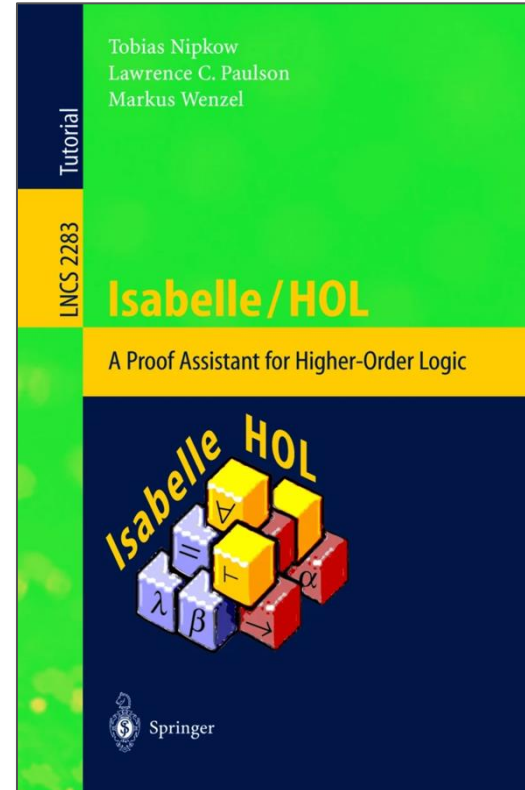
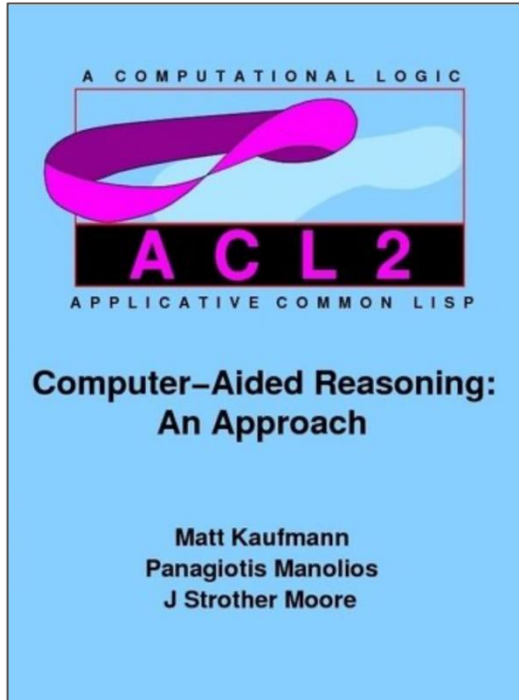


OPTIONAL

Candidate plans repeat validation until 1-5 are satisfied.



Verification & Validation



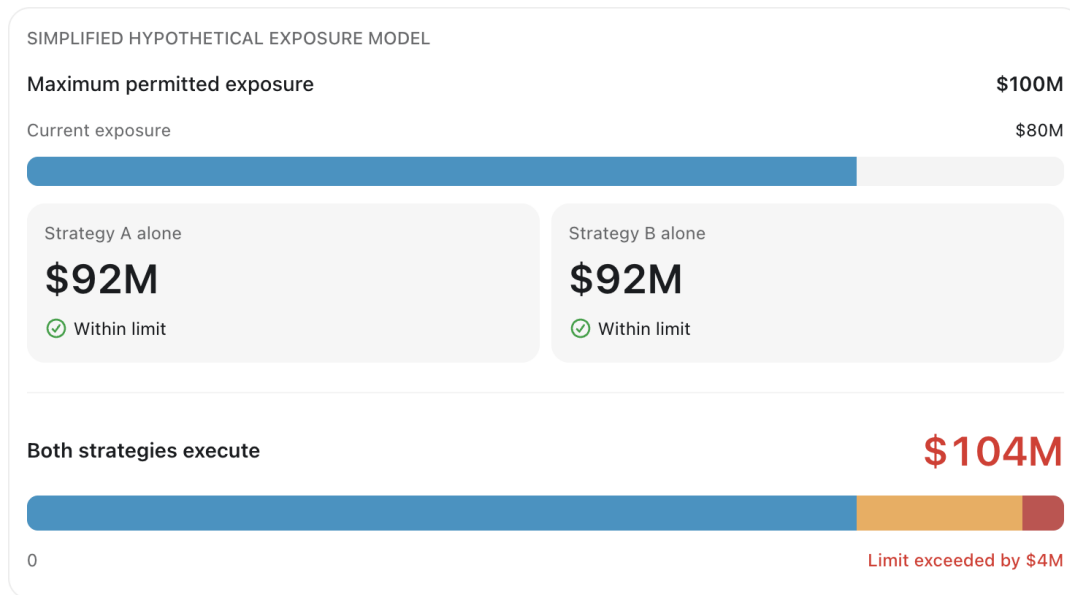
Verification & Validation

- ☐ Translate XML or JSON plan specification or evidence into ARS abstractions (also declarative)
- ☐ Check that the plan and evidence items are self-consistent
- ☐ Check that the plan and evidence are consistent with “the world” (if your evidence isn’t a complete sample of “the world”
 - ☐ Your business universe
 - ☐ e.g., published standards of care
 - ☐ In ARS terms, these are encoded as “theories” or “books”

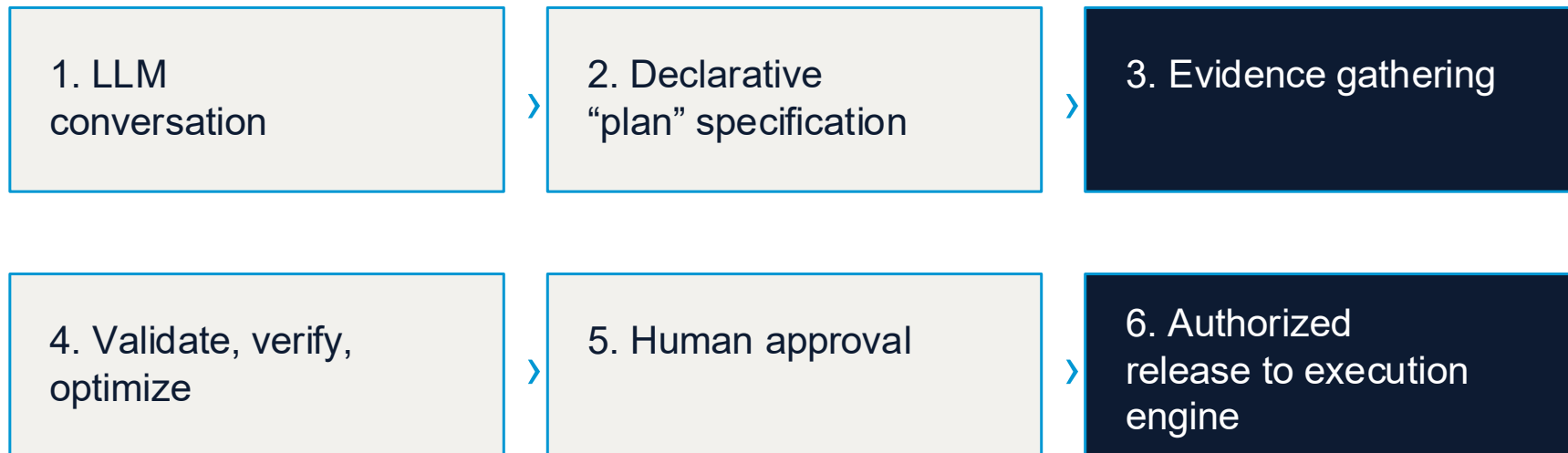


Don't formally verify every property

- It's time consuming, unnecessary, and vaguely annoying
- Verify “the plan” vs. “the world” at risky **interaction points** in the domain



Trusted agent architecture



OPTIONAL

Candidate plans repeat validation until 1-5 are satisfied.



Optimization (constrained optimization)

- Usually means solving an NP-hard optimization problem
- “Assemble a care team meeting accepted standards-of-care at the lowest cost”
 - The Knapsack Problem or the TSP
- **Knowledge representation** is important
 - My news-grouping-as-hypergraph-partitioning example
 - The opposite of “**trustworthy at enterprise scale**”
 - It’s usually where you minimize/maximize the objective function



