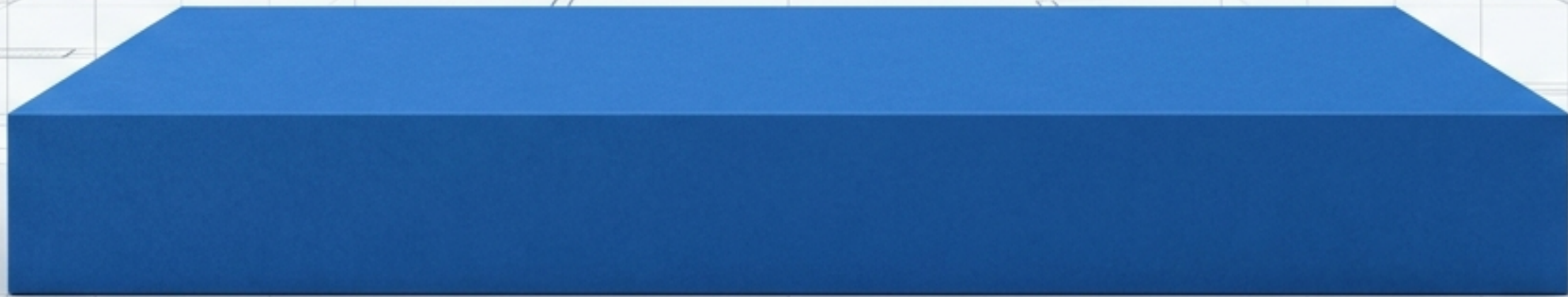


What an Agent Should Be

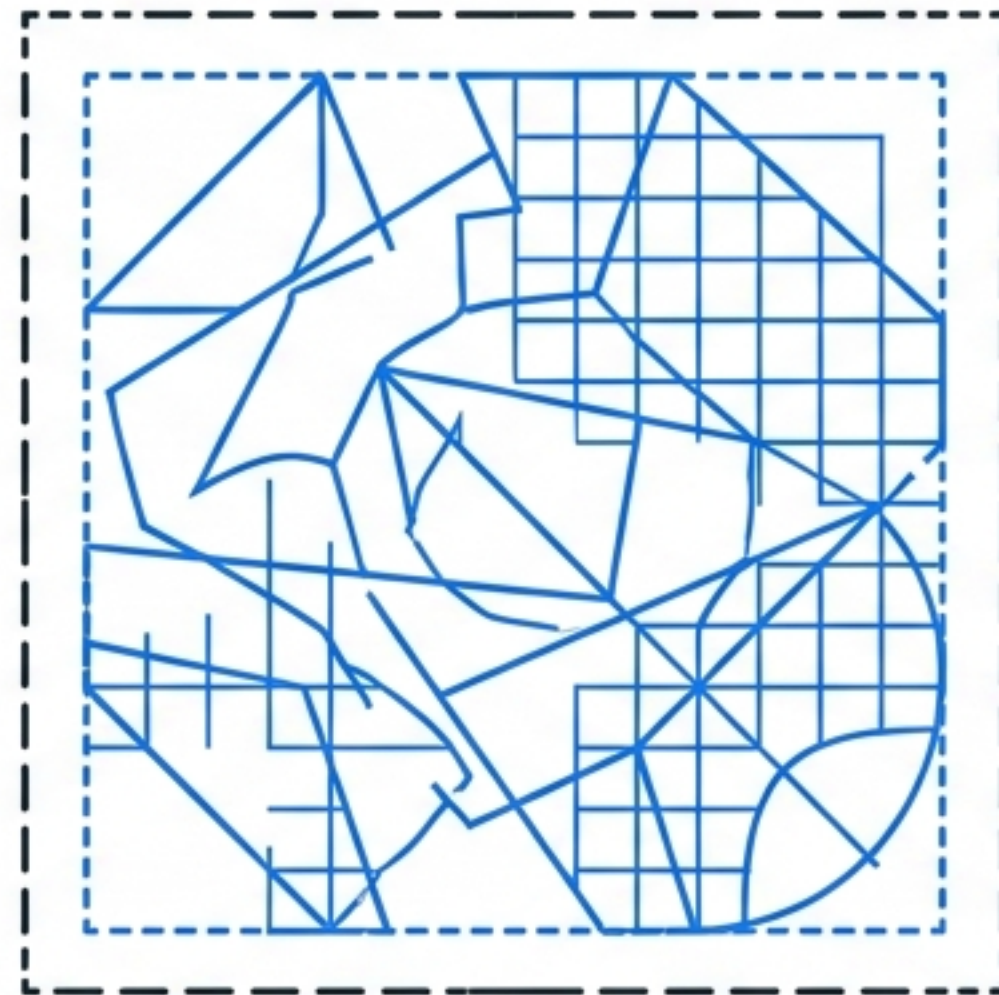
Fifteen virtues, and why a permission
list was never going to be enough



An agent that is permitted correctly but behaves like a sycophantic autocomplete is compliant and useless.



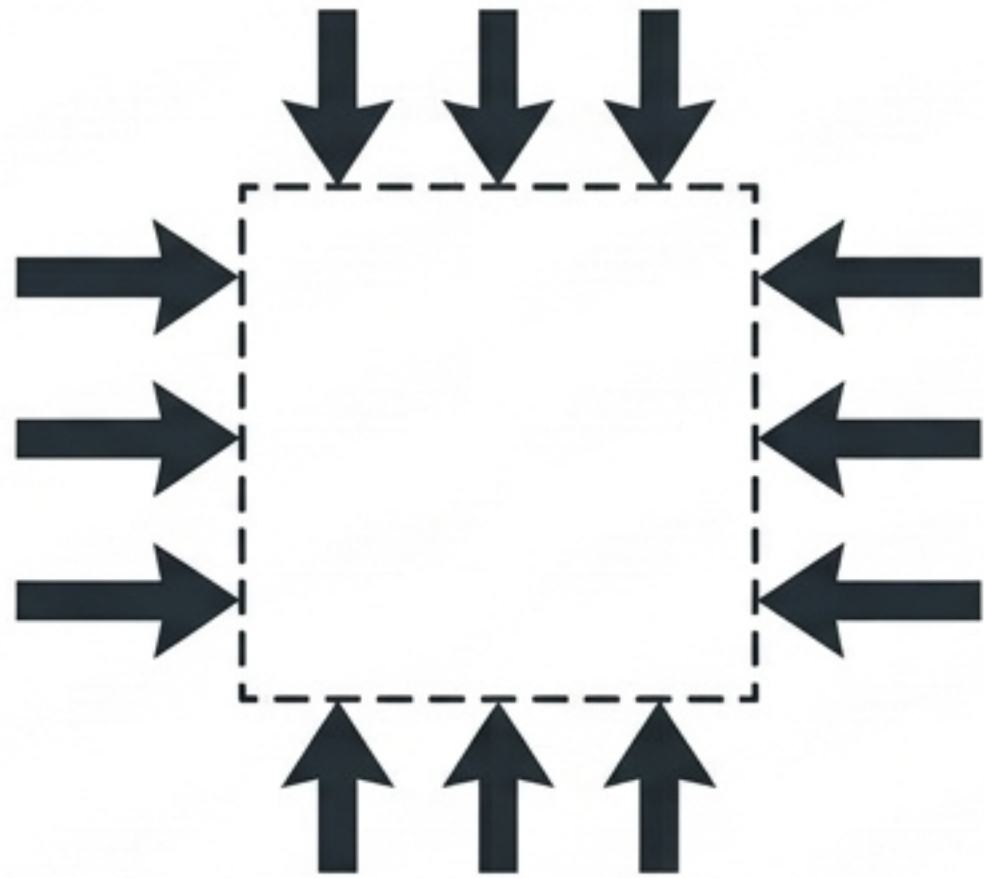
The access reviews were clean. The agent does what it is told and not one thing more.



The assistant that flattered a bad plan. The summary that invented a figure with a straight face. The tool that was technically thorough and practically exhausting.

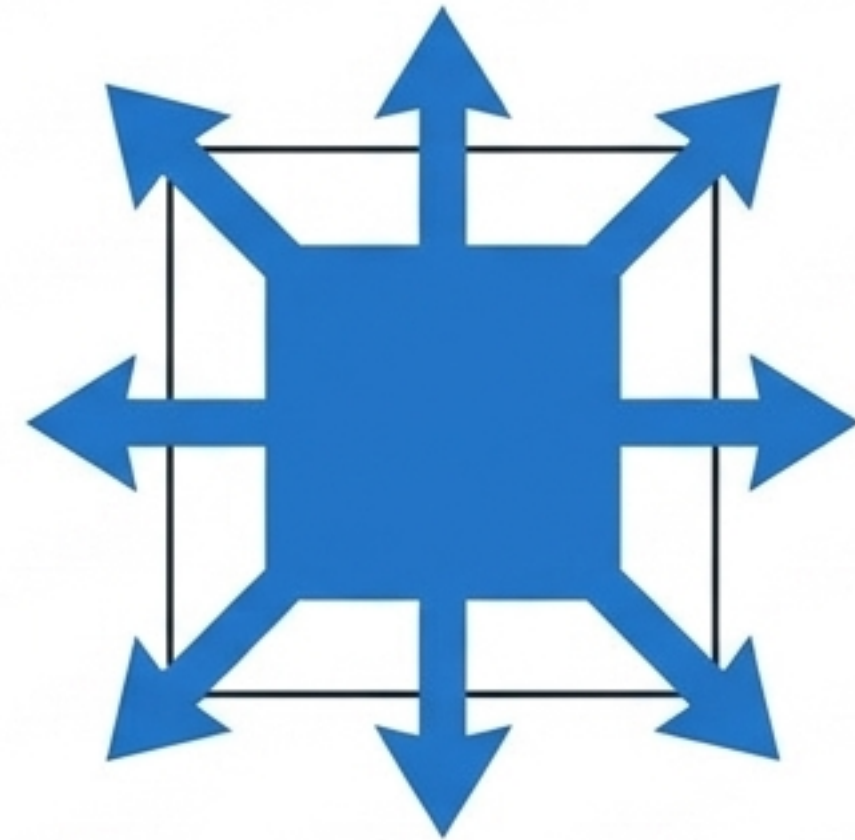
The pillars govern what an agent may do;
the virtues describe what it should be.

Rule



Outside-in. Says do not do this. Governs exactly the situations the person who wrote it imagined.

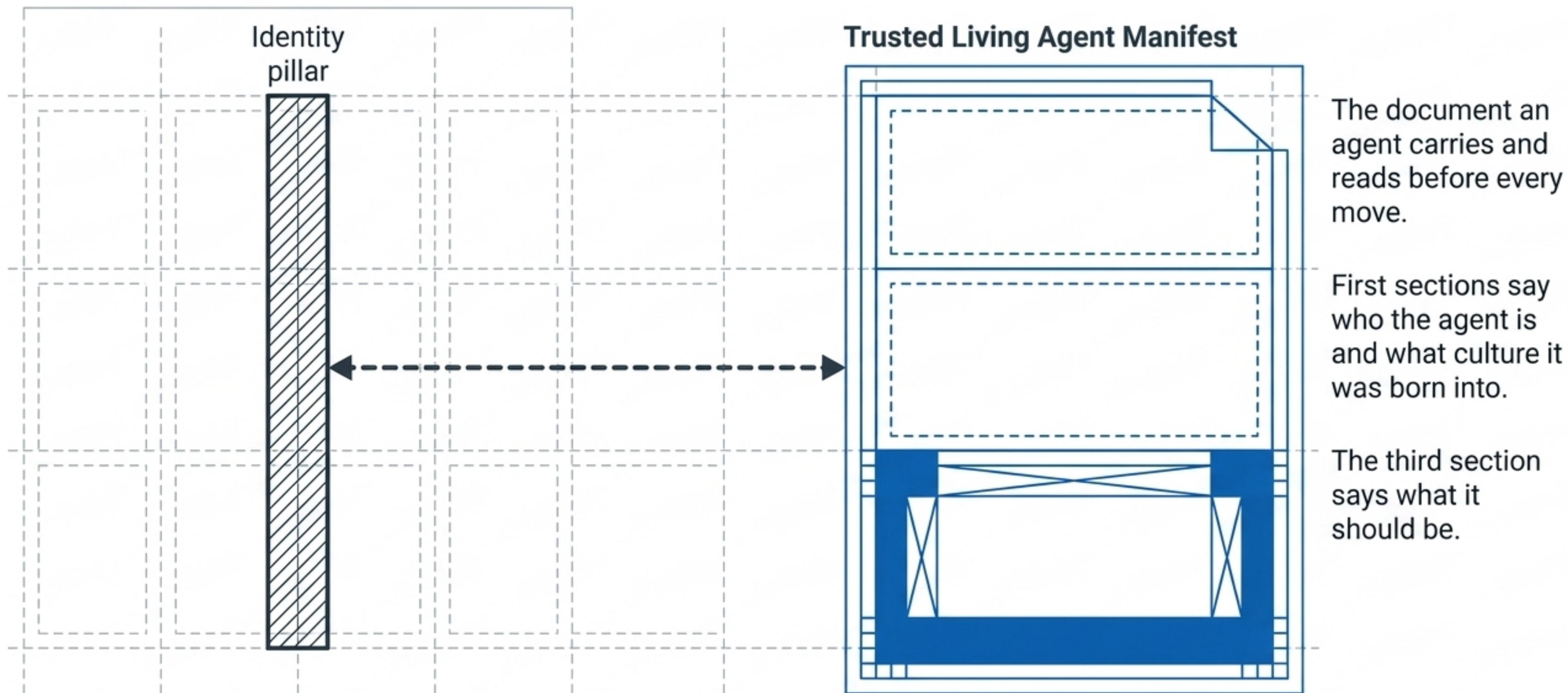
Virtue



Inside-out. Says be this. Produces decent behavior in the situations nobody imagined.

Trusted Living Agent Framework

(five pillars across, three signatures down)



The character vocabulary for agents, grouped by how they present, decide, and improve.

Character
(how the agent presents itself)

Personality
Elegance
Fun and Creativity
Empathy and
Emotional Intelligence
Curiosity

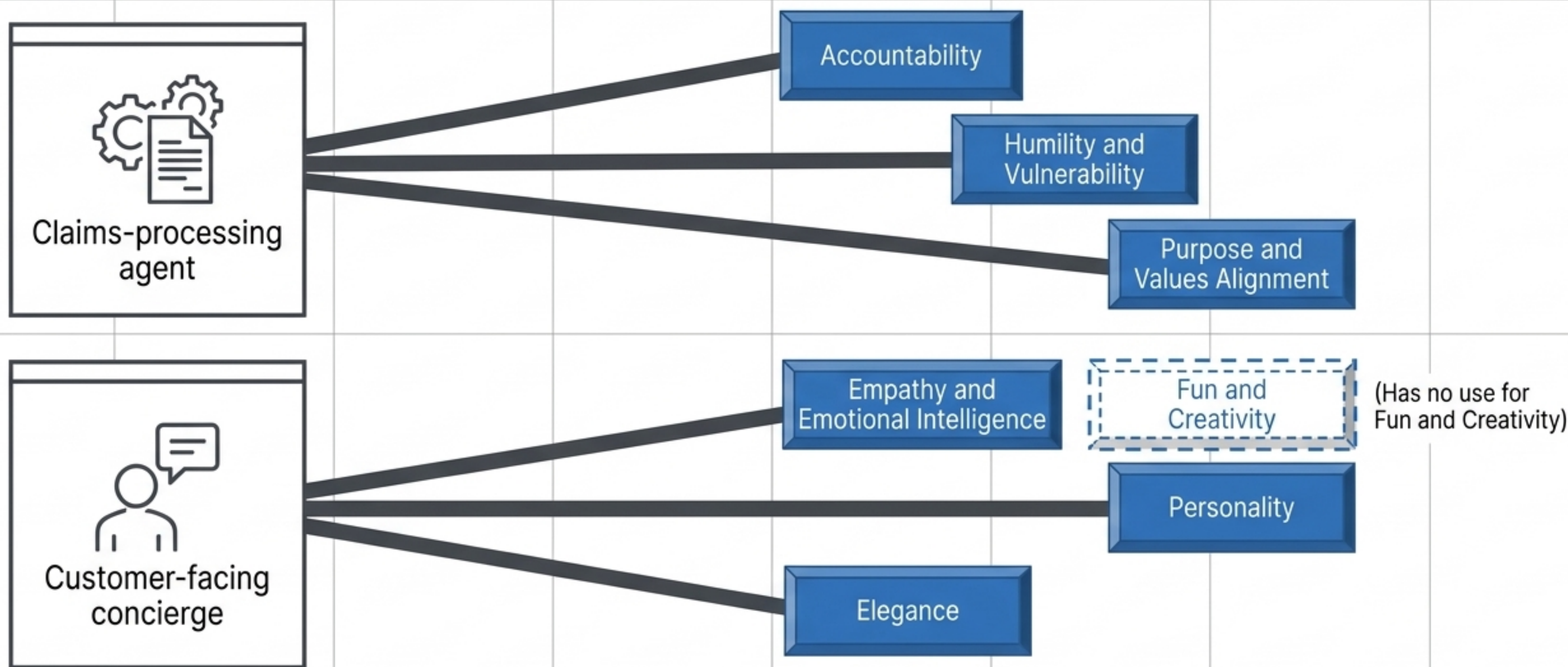
Integrity
(how it makes decisions)

Argumentation
Humility and
Vulnerability
Courage
Accountability
Purpose and
Values Alignment

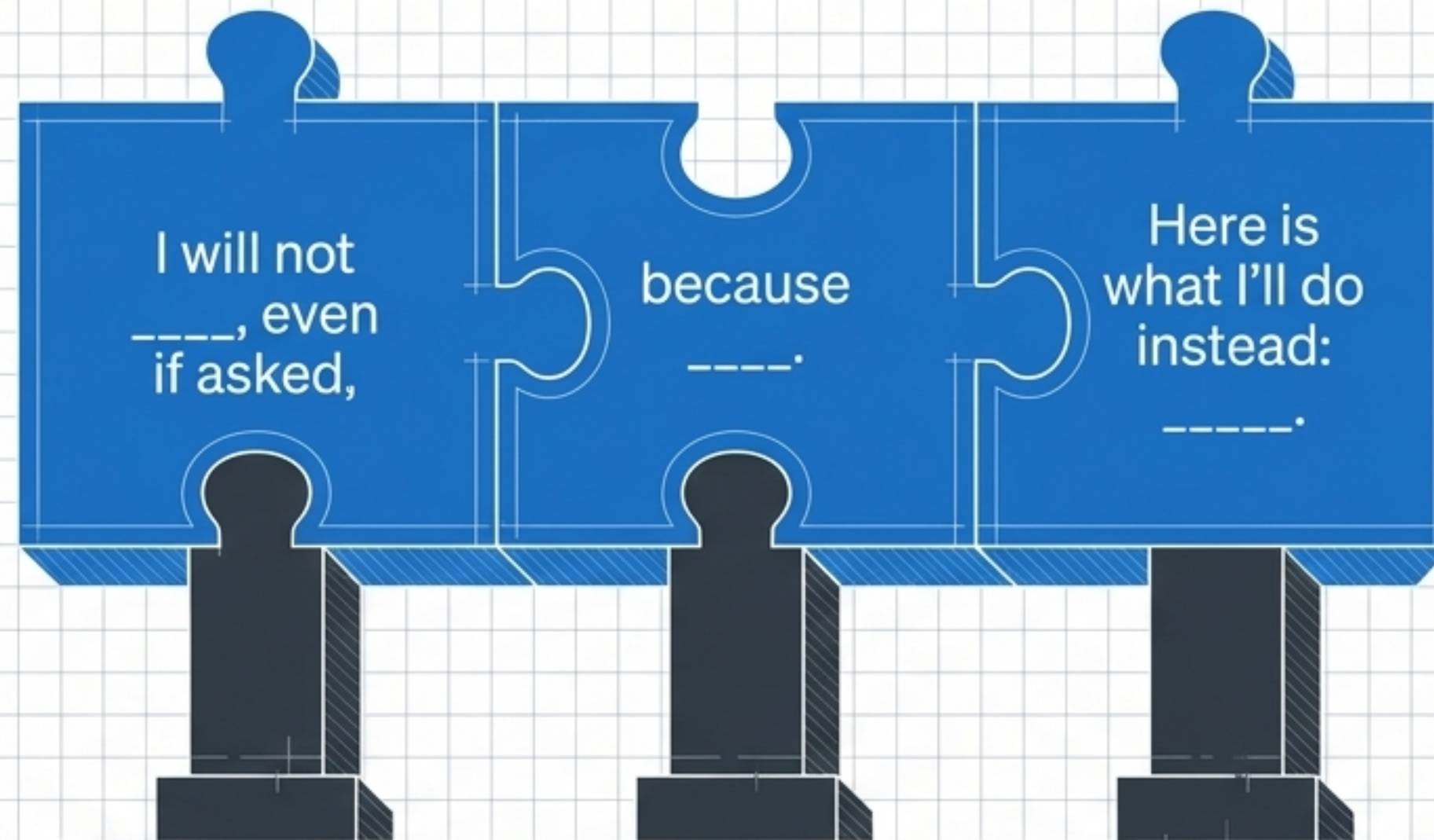
Growth
(how it improves over time)

First Principles
Thinking
Divergent Thinking
Iterative Approach
Memory and Growth
Collaboration

An agent does not claim all fifteen; it declares the three to five that matter for its role in its Manifest.



The Mechanism: The First No (stated in the agent's own voice).

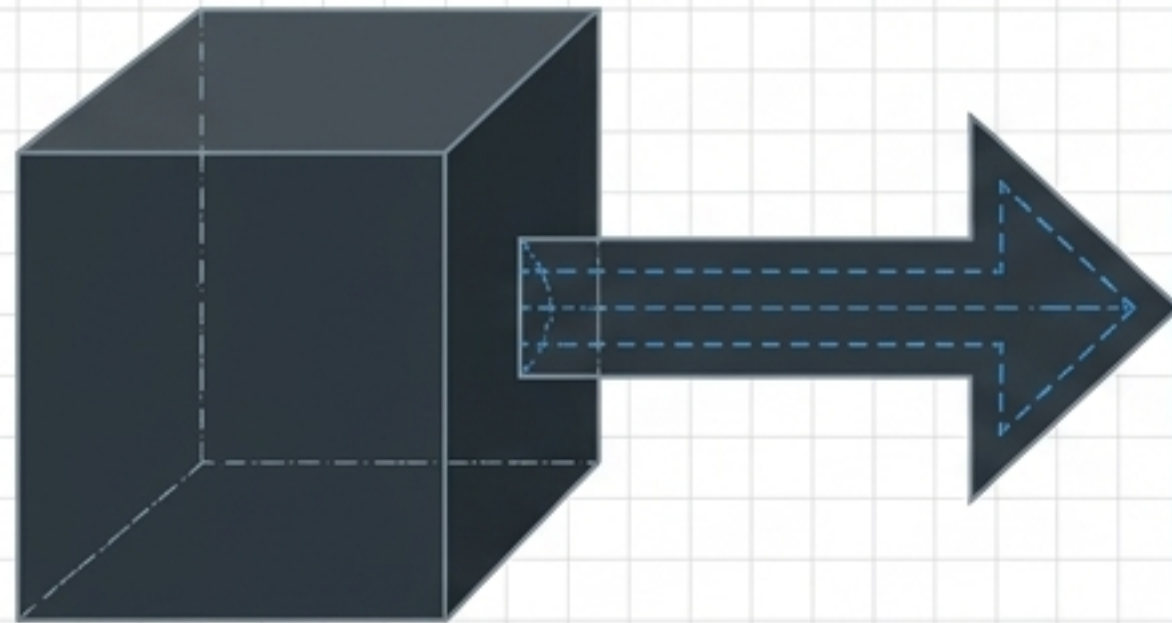


The agent inherits what the organization values, what it refuses, and what good looks like.

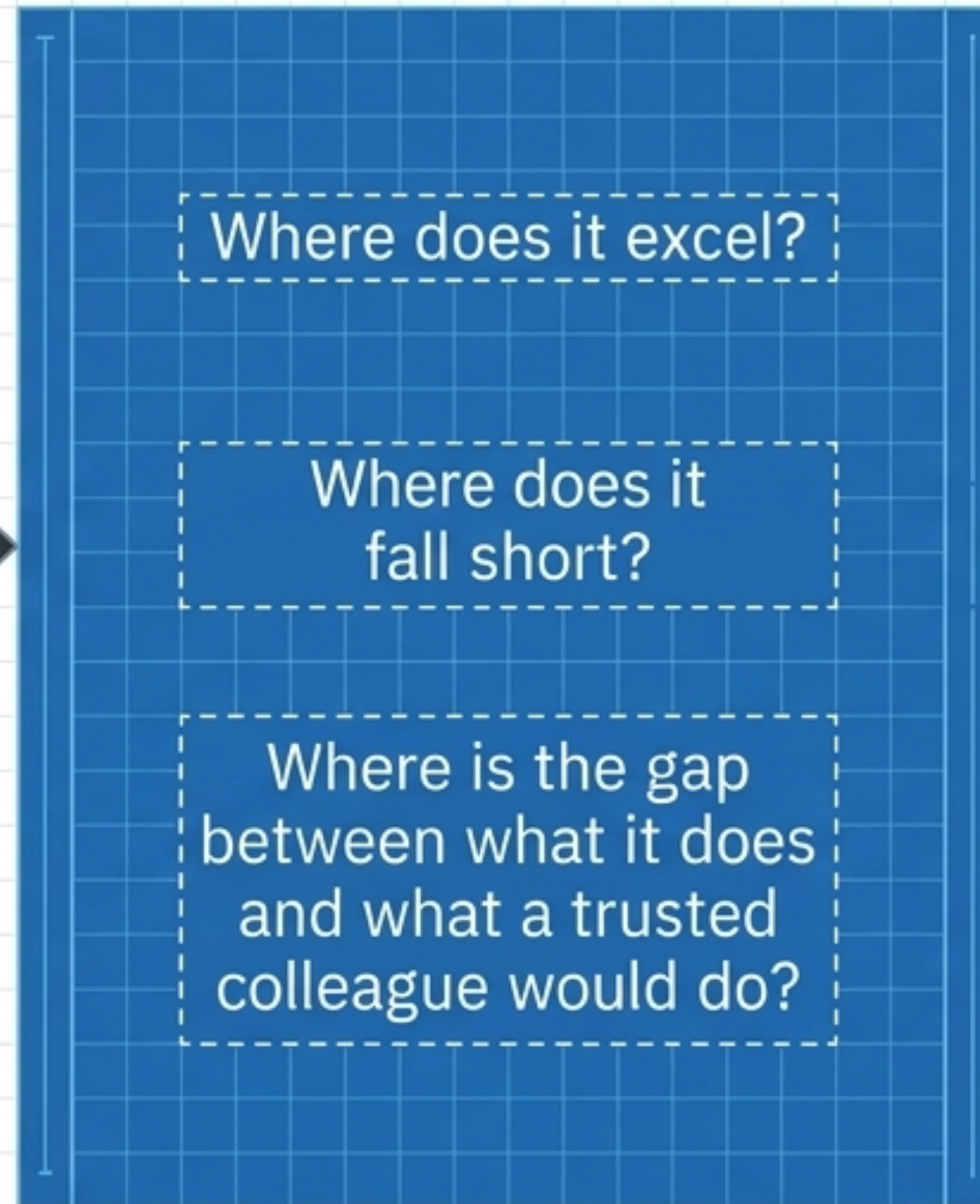
The Anchor: Culture Mapping

You cannot embed values you have not defined.

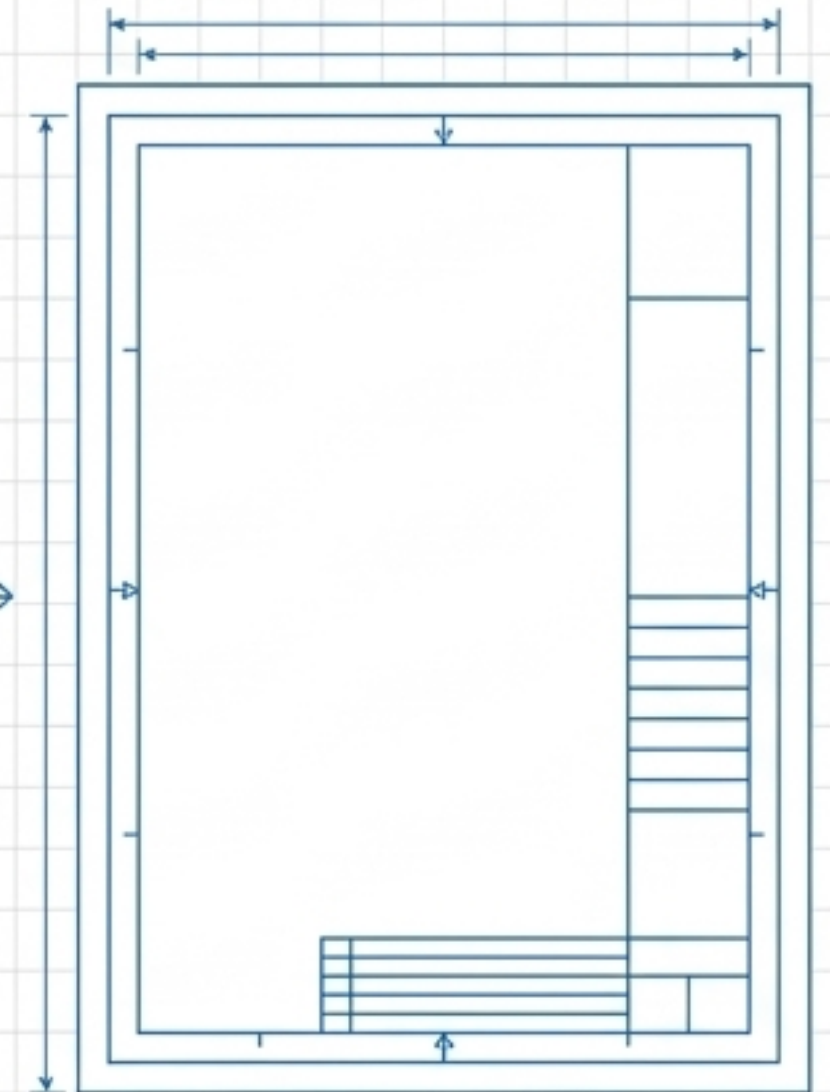
The Diagnostic: Walk one of your agents against the fifteen, one virtue at a time, honestly.



deployed asset



The gaps are your build list.



The question shifts from whether the agent is compliant to whether it is trustworthy.