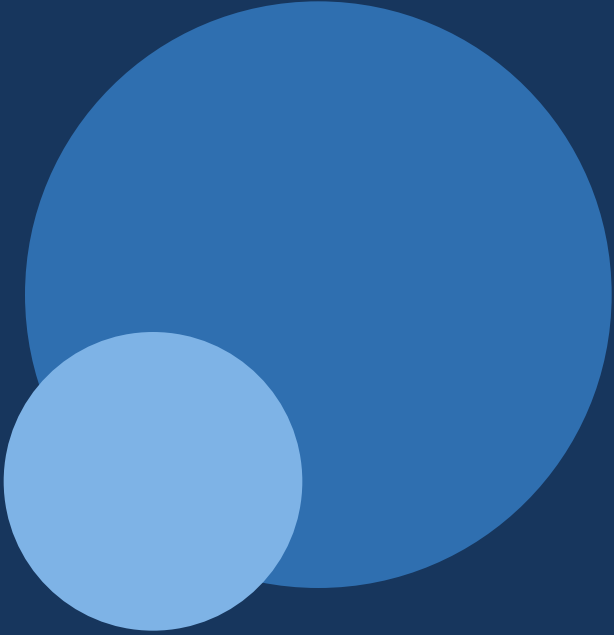




When AI Agents Agree They May Still Be Wrong

Why apparent consensus is not the same as
independent verification

A practitioner field note

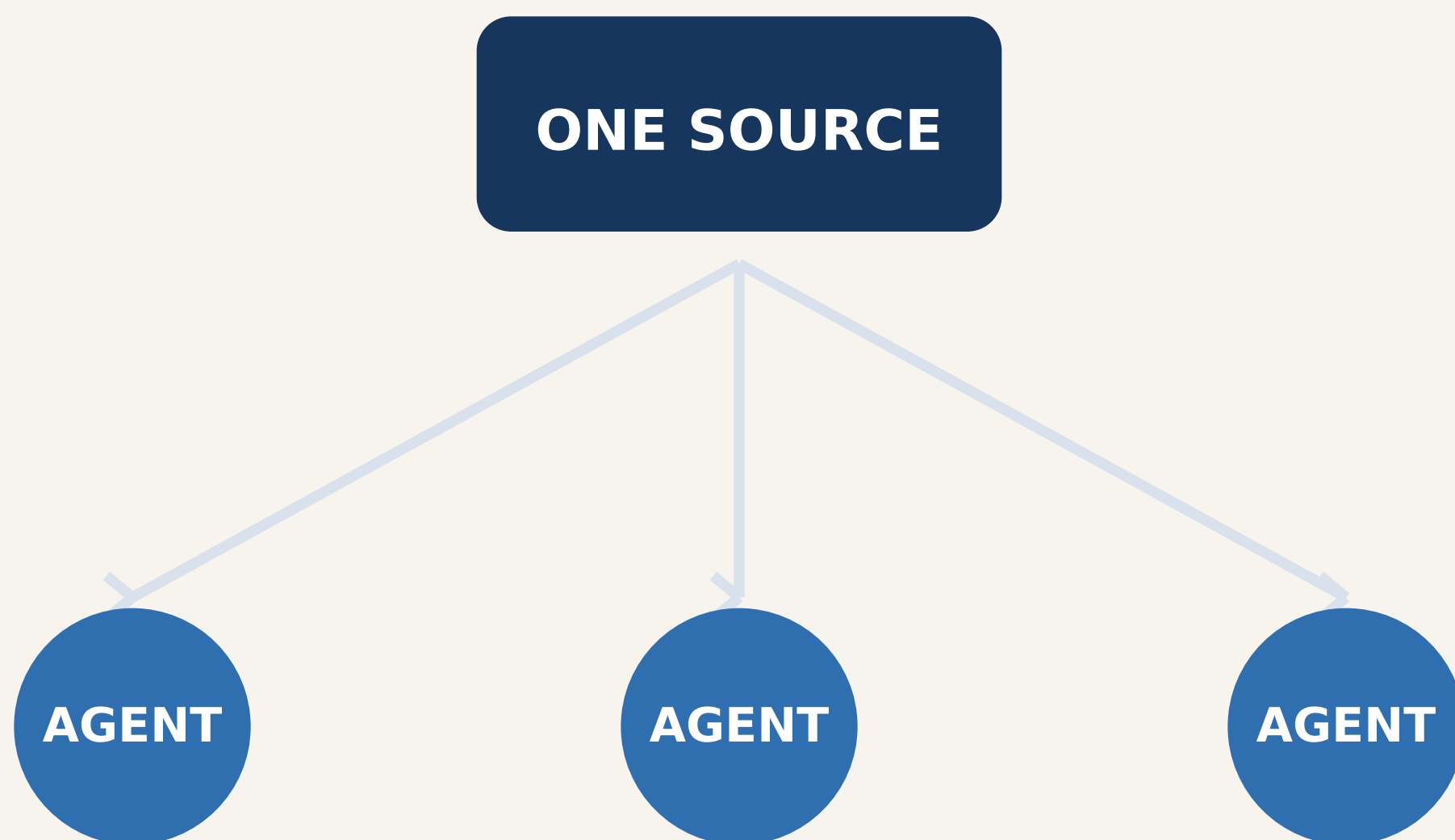
Agreement feels like evidence

If several agents reach the same conclusion, it is tempting to treat the result as corroborated.



Three answers can still be one opinion

Consensus adds little when the source is shared



Different summaries of the same article, database entry, or earlier AI output do not create independent confirmation.

A reviewer should challenge the premise

Presentation review

Is the recommendation clear, concise, and well organized?

Premise review

Is the source credible? Is it current? Did the signal already get priced or addressed? Does the proposed action actually follow?

**Good editing improves the answer.
Independent review may change the question.**

Four questions before trusting agent agreement

- 1 Did the agents use genuinely independent sources?**
- 2 Can each material claim be traced to the underlying record?**
- 3 Could a reviewer reject the premise rather than edit the wording?**
- 4 Is no action an acceptable final result?**

The system needs real stop conditions

ACT

Evidence supports a proportionate next step

MONITOR

The signal matters, but timing is not established

RESEARCH

A material uncertainty still needs validation

PASS

The evidence does not support action

Do not count agreement until you check independence

Multiple agents can reinforce the same plausible mistake. Trace the evidence, permit premise-level challenge, and keep no action available.

From AI Tools to an AI Operating Model

A field paper by Nick Montano