Atlas Research

ATLAS RESEARCH NOTE 01 · VERSION 1.0

When AI Output Becomes Organisational Knowledge

How provisional AI-generated material acquires epistemic status and institutional authority - and what organisations must keep visible before acting on it.

Robert Schouten
13 October 2026

AI · KNOWLEDGE · GOVERNANCE

> ***Atlas Research Note 01 · Version 1.0.*** *This note develops a working argument within the Atlas Research programme on AI, human practice and the institutional production of knowledge. It connects our research on employee-led AI adoption with the epistemological foundations being tested through IROS. The argument is intended for further empirical and technical investigation.*

The moment an answer changes status

An employee asks an AI system to summarise a regulatory development. The answer is plausible and mostly accurate. The employee corrects two details, removes a qualification and copies the result into an internal note. A colleague later uses that note in a presentation. Its central claim enters a project document, then a risk assessment, and eventually a decision.

At which point did the AI output become organisational knowledge?

The question is harder than it appears. The output did not become knowledge simply because a model generated it. Nor did it necessarily become knowledge when a professional reviewed it, when it entered an approved document or when others began to act on it. Each transition changed its organisational status, but may have revealed very little about its epistemic status: what supports the claim, how it was interpreted, which uncertainties remain and who is prepared to take responsibility for it.

This is the problem addressed in this research note. Organisations increasingly use AI not only to produce text, but to support analysis, research, reporting and decision-making. Their central risk is therefore not limited to individual hallucinations. It is that provisional output can pass through ordinary organisational processes and emerge as an apparently settled fact.

From individual use to institutional authority

In the forthcoming article *When workers bring AI to work: From private experiment to organisational capability*, I argue that employees often make general-purpose AI useful before their organisations have decided how it should change work.[1] They supply the missing context. They formulate prompts, identify exceptions, correct errors and decide whether an answer fits the situation. This work depends on professional judgement and is often informal.

The next question is what happens to the resulting output.

An answer that remains on an employee's screen has limited institutional reach. Once it is copied into a report, entered into a knowledge base, used to update a policy or accepted as an input to another system, it begins to travel independently of the practice in which it was produced. The corrections may remain, while the reasons for those corrections disappear. The conclusion may remain, while its sources, counterevidence and uncertainty are lost. The organisation retains the product of judgement without necessarily retaining the judgement itself.

Organisational learning theory has long examined how individual insight becomes embedded in shared understanding, routines and systems. Crossan, Lane and White describe a movement from intuiting and interpreting to integrating and institutionalising.[2] AI does not create this transition, but it can accelerate and obscure it. It produces material that is already fluent, structured and easy to circulate. That makes it unusually simple for an unverified output to acquire the appearance of an institutional statement.

Output, claim, evidence and knowledge are not the same

The distinction matters because AI systems operate primarily on representations. They generate or transform language, images, classifications and other symbolic material. These outputs may contain useful information, but usefulness is not the same as justification.

The research framework under development at Atlas distinguishes several stages:

1. **A source** carries information.
2. **An observation** is selected from that source and interpreted in relation to a question.
3. **Evidence** is an assessed observation that supports or contradicts a particular claim.
4. **A knowledge claim** is a provisional statement whose justification can be examined and revised.
5. **A conclusion** synthesises several sufficiently justified knowledge claims in response to a research question.
6. **A decision** concerns what should be done in light of that knowledge and remains the responsibility of the decision-maker.

AI can assist at every stage, but it does not abolish the distinctions between them. A generated statement is not automatically an observation. A citation is not automatically evidence. Evidence does not become truth merely because it supports a claim. A conclusion does not make the decision that follows from it.

When these distinctions collapse, fluency begins to substitute for justification.

Four ways epistemic status becomes invisible

The transition from AI output to organisational knowledge can fail in at least four connected ways.

Provenance disappears

Each time a statement is summarised, reformatted or transferred to another system, its history may become harder to reconstruct. A reader sees the final assertion but not the original sources, model version, prompt, human corrections or purpose for which the output was produced.

Provenance does not prove that a claim is correct. It makes it possible to investigate how the claim came to be there. The NIST Generative AI Profile similarly treats provenance data as information about the origin and history of content that can support risk management across the AI lifecycle.[3]

Uncertainty is compressed

Research begins with incomplete and sometimes contradictory material. Organisational documents, however, often reward brevity and decisiveness. Qualifications disappear as an exploratory analysis becomes an executive summary, a slide or a status field. “The available evidence suggests” becomes “the organisation knows”.

The problem is not uncertainty itself. Decisions always take place under conditions of incomplete knowledge. The problem is losing sight of where the uncertainty entered, which alternatives were considered and whether new evidence should reopen the claim.

Review becomes a ritual

A human may approve an output without being in a position to judge it. Time pressure, automation bias, lack of domain knowledge or distance from the underlying practice can reduce oversight to a procedural click. The organisation can then point to human review even though the reviewer could not reconstruct or challenge the reasoning.

Meaningful review requires more than a person in the workflow. It requires access to evidence, sufficient professional capability, authority to reject the output and time to exercise judgement.

Repetition becomes authority

Once a statement appears in several internal documents or systems, later users may treat that repetition as confirmation. But multiple documents can reproduce the same unsupported origin. Apparent corroboration may therefore be no more than the circulation of one earlier output.

This is where an epistemic status transition becomes institutional. The claim is no longer trusted because someone has examined its justification. It is trusted because the organisation has already used it.

Is this epistemic laundering?

The phrase *epistemic laundering* is tempting because something uncertain can emerge from an organisational process looking cleaner and more authoritative than when it entered. The term is not unique to this programme. Kalaitzidis uses it for a broader process through which generative AI systems can transform contradiction into legitimacy while concealing the conditions under which this occurs.[4]

The narrower organisational mechanism proposed here is more specific. It concerns the way an AI output moves across documents, roles and systems while its origins and qualifications become progressively less visible. No one needs to intend deception. Each individual step may be reasonable. The laundering effect arises from the chain as a whole.

For that reason, the most useful object of governance is not only the model or even the individual output. It is the sequence of status changes through which the output acquires institutional authority.

A problem for knowledge-intensive organisations

Consider a team monitoring regulatory, economic or technological developments. AI helps locate documents, extract passages, compare interpretations and draft an analysis. A specialist reviews the result and a manager approves the final report. The report is then used in strategic planning.

The process may look well governed because named people participated and an approval occurred. Yet several questions remain:

- Can each material claim be traced to the observations and sources that support it?
- Is supporting evidence distinguishable from contradiction or counterevidence?
- Does the report show where interpretation entered the process?
- Are assumptions and unresolved uncertainties still visible?
- Is the approving person qualified and authorised to challenge the claim?
- Can the organisation revise the claim when the underlying evidence changes?

These questions are relevant wherever an organisation turns large volumes of information into decisions: finance, healthcare, government, engineering, law and professional services. They are not merely questions of data quality. They concern the institutional production of knowledge.

What should organisations preserve?

The aim cannot be to keep every intermediate token or require senior review of every sentence. Control must be proportionate to the significance of the claim and the decision it may inform. But several design principles follow from the argument.

Keep status explicit

An AI-generated output, an extracted observation, a working hypothesis, a supported knowledge claim and an approved conclusion should not be treated as interchangeable objects. Systems and documents should make these differences visible.

Preserve the path to the source

Material claims should remain connected to their sources and to the relevant passage, dataset or observation. A reference list at the end of a generated report is not sufficient when it is unclear which source supports which claim.

Record interpretation

Facts do not simply transfer themselves from a document into an organisation. People and systems select what is relevant and interpret it within a particular model of the situation. Important interpretive choices, definitions and assumptions should therefore be inspectable.

Retain counterevidence and alternatives

A system designed only to accumulate support will produce confidence, not research. Contradictory observations and plausible alternative explanations must remain connected to the claims they challenge.

Make approval substantive

Human approval should identify who judged what, on which basis and within which role. The reviewer must be able to withhold approval or return the claim for further investigation. Responsibility cannot be inferred merely from the presence of a name in a log.

Keep claims revisable

Organisational knowledge should not become immutable merely because it has entered an approved document. New observations must be able to change the epistemic status of a claim and trigger review of conclusions that depend on it.

From governance principle to research infrastructure

These requirements are difficult to implement when research is organised as a sequence of documents. Documents are useful for communication, but they tend to flatten the process that produced them. Sources, observations, evidence, interpretations, claims and conclusions become paragraphs of the same apparent status.

This is one reason Atlas Labs is testing **IROS, the Ingesta Research Operating System**. IROS is intended to organise the research process rather than merely generate an answer. Its working architecture treats sources, observations, evidence, knowledge claims, uncertainty and human approval as related but distinct objects.

The prototype does not guarantee truth, remove hallucinations or automate responsibility. Its more modest purpose is to prevent a convincingly formulated output from acquiring the status of knowledge without leaving an inspectable path of justification.

That proposition still requires testing. We need to learn which distinctions remain useful under real working conditions, where the method creates unreasonable friction and whether users can meaningfully assess epistemic status without turning judgement into another checklist. Those questions belong in practice, not only in theory.

The governance question before the AI question

Organisations often begin by asking which AI system they may use, which model performs best or where a human should be inserted into the workflow. These are necessary questions, but they come too late if the organisation has not first decided what it means for a claim to count as knowledge.

The prior questions are institutional:

- What must be visible before a claim may inform a decision?
- Which forms of evidence are sufficient for which purposes?
- Who is able and authorised to judge the claim?
- How are uncertainty and disagreement preserved?
- What causes accepted knowledge to be reviewed?

AI makes the production and circulation of explicit material dramatically easier. It does not supply the practices through which relevance, meaning and responsibility are formed. When an organisation hides that dependence, it may preserve the products of expertise while weakening the human and institutional conditions that make those products trustworthy.

The critical transition is therefore not from human work to machine work. It is from generated output to institutional authority. That transition should never be invisible.

References

1 Schouten, R. (2026), "When workers bring AI to work: From private experiment to organisational capability", forthcoming in the *State of AI Ethics Report*, Volume 8.

2 Crossan, M. M., Lane, H. W. and White, R. E. (1999), "An Organizational Learning Framework: From Intuition to Institution", *Academy of Management Review*, 24(3), pp. 522-537.

3 National Institute of Standards and Technology (2024), *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*, NIST AI 600-1.

4 Kalaitzidis, T. (2026), "Epistemic laundering: generative AI and the naturalization of misrecognition", *AI & Society*, 41, pp. 8169-8184.