IDΔAC™ NEWSLETTER / ARTICLE 06

When the Rationale Appears Too Late

Why a complete explanation may still be weak evidence of judgement at the time of decision

A rationale can explain the decision without proving when the judgement existed.

AUTHOR Cosmin-Corneliu Oprea

PUBLICATION 28 July 2026 | Public thought brief

CONCEPT FAMILY Decision evidence | Rationale timing | Decision reconstruction

Independent Decision & Accountability Consulting

www.iddac.eu | contact@iddac.eu

THE RECORD LOOKS COMPLETE

The explanation arrived after the decision

An AI-assisted case-prioritisation system recommends a standard closure route. The reviewer opens the case, checks the available screen and approves the recommendation at 09:27. The operational outcome is issued at 09:31.

At 14:52, after a manager asks why the case was closed, the rationale field is completed: the available indicators did not justify escalation; the standard route was proportionate; no material contradiction was identified.

The explanation may be careful, honest and substantively plausible. It may even describe considerations that were genuinely present in the reviewer's mind. But the record now contains two different evidentiary objects: a decision made at 09:27 and an explanation created at 14:52.

The later explanation can help interpret the outcome. It does not, by itself, prove what judgement existed while the route was still open to challenge or change.

THE CENTRAL DISTINCTION

A rationale has both substantive content and temporal status. It may be a strong explanation of the final outcome while remaining weak evidence of the judgement that existed before commitment.

Late does not mean false

Timing uncertainty should not be turned into an accusation. A rationale recorded after commitment is not automatically fabricated, deceptive, unlawful or operationally defective. Teams often add detail later because the workflow did not require a contemporaneous note, because the case became contested, or because an audit requested a fuller account.

The correct response is not to discard later commentary. It is to label it accurately and avoid silently treating it as if it had existed at the decision point.

The problem is not that the organisation can explain the decision. The problem is that it may no longer be able to prove when the explanation became part of the decision.

Decision time is an evidentiary boundary

Before commitment, evidence, uncertainty and alternatives remain live. A reviewer may still pause, challenge, modify, escalate or reject the proposed route. After commitment, the outcome becomes known, organisational expectations become clearer and the selected route begins to look more coherent than it may have felt beforehand.

Research on hindsight bias and post-choice rationalisation shows that knowledge of an outcome can reshape later evaluations of the options that preceded it. This does not establish bad faith in any specific record. It explains why retrospective accounts should not automatically be given the same evidentiary status as reasoning preserved while the decision was still unresolved.[2][3]

Three records that should not be collapsed

1 **The system output or technical explanation**
What the AI system produced, displayed or logged - for example a score, ranking, recommendation, confidence signal, feature attribution or generated explanation. This may help reconstruct system behaviour. It is not the human reviewer's rationale.

2 **The contemporaneous human rationale**
What the reviewer recorded before or at commitment about relevant evidence, uncertainty, alternatives, authority and the reason for selecting the route. This is the strongest evidence that a stated judgement existed while intervention remained possible.

3 **The retrospective explanation**
A later account produced after approval, operational effect, complaint, incident or audit. It may be useful and accurate, but it must remain distinguishable from the evidence that existed at the time of judgement.

A simple timeline test

POINT	RECORD	QUESTION
T0	**Evidence state**	What information, uncertainty and alternatives were available?
T1	**AI influence**	What output, ranking, framing or recommendation reached the reviewer?
T2	**Human commitment**	When did the reviewer or organisation become committed to the route?
T3	**Operational effect**	When was the action executed, communicated or made difficult to reverse?
T4	**Rationale event**	When was the rationale first created, materially completed or later edited?

When T4 occurs after T2 or T3, the rationale should not be treated as contemporaneous unless version history or other attributable evidence establishes that the material reasoning already existed before commitment.

Why rationale timing becomes unclear

Weak timing evidence rarely begins with an explicit decision to backfill a file. It usually emerges from ordinary workflow design.

1. **The rationale field opens after approval**
   The system allows the outcome to be committed before reasoning is captured. Completion becomes an administrative after-step rather than part of judgement.
2. **Later edits overwrite the original state**
   The file preserves only the latest text, not the first version, editor, timestamp or material change.
3. **The record timestamps the case, not each event**
   Approval, rationale creation and later commentary appear under a single final-modified date.
4. **Templates replace case-level judgement**
   Repeated language may show process consistency, but it may not demonstrate that uncertainty and alternatives were considered in the selected case.
5. **The account becomes richer after challenge**
   A complaint, incident or audit creates new attention and better information. The later record may therefore be stronger than the one available to the original reviewer.
6. **Memory substitutes for preserved evidence**
   Interviews reconstruct what people now believe happened. They remain useful, but memory and contemporaneous evidence are not interchangeable.

What the later rationale can establish

- How the organisation currently understands or defends the outcome.
- Which considerations a reviewer or manager now regards as relevant.
- Where the original workflow failed to capture adequate evidence.
- What should be tested in version history, audit logs or reviewer-facing records.

What it cannot establish by itself

- That the same reasoning existed before or at commitment.
- That uncertainty, alternatives or contradictory evidence were visible at the time.
- That the reviewer retained usable authority to change the route.
- That the decision was lawful, fair, correct, compliant or defective.

SAFE INTERPRETATION

A missing or late timestamp is evidence of uncertainty, not proof of failure. The task is to establish what the surviving record can support, what requires inference and what remains unknown.

Design the record around commitment

The objective is not to force long essays into every routine workflow. Excessive documentation can burden reviewers, produce formulaic text and obscure the few cases where reasoning genuinely needs to be preserved. The better design is proportional, attributable and tied to the point at which the organisation becomes committed.

Seven operational controls

1. **Separate timestamps**
Preserve approval time, rationale creation time and later edit times as distinct events.
2. **Capture short judgement at commitment**
For high-impact, unusual, contested or difficult-to-reverse decisions, record the relevant evidence, uncertainty and reason for selecting the route before or at commitment.
3. **Preserve version history**
Retain the initial rationale and material edits, with attributable authorship and timestamps.
4. **Label retrospective commentary**
Use explicit labels such as post-outcome analysis, audit note or later clarification.
5. **Protect the original decision state**
Preserve the output, interface state, supporting evidence and alternatives visible to the reviewer.
6. **Monitor rationale lag**
Sample the interval between approval, operational effect and rationale creation, especially in consequential routes.
7. **Test reconstruction without memory**
Ask whether the route can be rebuilt from surviving evidence without relying entirely on interviews with the original reviewer.

Use the public readiness check

The AI Rationale Timing Readiness Check provides a bounded starting point for examining one real workflow and a small sample of recent decisions. It focuses on timestamps, later edits, uncertainty, alternatives, version history and the ability to reconstruct the sequence.

AI RATIONALE TIMING READINESS CHECK

Use it before audit or board review, incident reconstruction, workflow redesign or control testing. It is a non-legal operational readiness tool, not a diagnostic conclusion.

Open the Resource Library →

Why this matters under European AI governance

For high-risk AI systems, the EU AI Act distinguishes several relevant governance functions. Article 12 addresses automatic logging and traceability of system operation. Article 13 requires information that enables deployers to interpret outputs and use systems appropriately. Article 14 requires effective human oversight, including the ability to understand limitations, avoid over-reliance and disregard, override or interrupt outputs where appropriate. Article 26 also places monitoring and log-retention duties on deployers in defined circumstances.[1]

These requirements do not make every organisational decision rationale a mandatory legal record, nor do they turn any explanation into proof of effective oversight. They do, however, reinforce the importance of separating technical logs, system explanations, human judgement and organisational commitment. A technically traceable system can still sit inside a weakly reconstructable decision route.

OPERATIONAL IMPLICATION

Logging can show what the system did. A contemporaneous rationale can help show what the organisation understood, considered and decided while intervention remained possible. The two evidence layers are related, but they are not interchangeable.

What this article does not determine

- Whether a specific decision was lawful, fair, accurate or compliant.
- Whether every consequential decision requires a written rationale.
- Whether a late rationale was dishonest or intentionally backfilled.
- Whether one timing weakness represents a wider organisational pattern.
- The severity of any IDΔAC™ decision exposure or a formal assurance conclusion.

When a bounded review may be appropriate

A structured review may be useful when one consequential route can be identified, the organisation relies on the recorded rationale as evidence of control, and the surviving record cannot clearly establish when the reasoning existed, what evidence shaped it or whether intervention remained possible.

The commercial route begins with the Intake Form. The Intake helps bound one route before any Scope Review or Exposure Diagnostic is considered.

USE THE PUBLIC RESOURCE	PREPARE ONE DECISION ROUTE
Resource Library →	Complete the Intake Form →

References

[1] Regulation (EU) 2024/1689 (EU AI Act), especially Articles 12, 13, 14 and 26. Source

[2] Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1(3), 288-299. DOI: 10.1037/0096-1523.1.3.288. Source

[3] Jarcho, J. M., Berkman, E. T., & Lieberman, M. D. (2011). The neural basis of rationalization: cognitive dissonance reduction during decision-making. Social Cognitive and Affective Neuroscience, 6(4), 460-467. DOI: 10.1093/scan/nsq054. Source