Research Article

PODA: Directed Opposition and Warranted Commitment—A Derived Analytical Framework for Claim-to-Commitment Records

Raúl Badillo[1]

1. Parabelo, Mexico

A construction can begin to guide action before the relations that qualify it have been settled. This can happen after a finding, proposal, design, or policy leaves the setting in which it was formed.

The same can happen while the construction is still taking shape through rapid iteration with an AI system. In both cases, a decision-maker may receive a claim without the object, conditions, comparison, or burden account that originally limited it. This article asks when such a construction has earned the commitment being requested on its behalf.

PODA, short for Proportionate Objection-Directed Assessment and pronounced POH-dah (\ipafont/ˈpo.ða/), is a derived analytical framework that examines that transition. Directed opposition tests the relation whose failure could change the case. Sustained iteration between an accountable person and one or more AI systems generates the pressure, while the person remains an active arbiter throughout rather than merely a final approver.

When the pressure exposes a material defect, PODA reconstructs the case and retains the residual claim and immediate step the record can support. Its Material Admissibility Unit (MAU) records the alleged difference, independent stake, mediator, evidence, alternative, burden, conditions, and requested action. The resulting determination states the permitted claim, active limitation, and commitment ceiling; institutional authorization remains a separate judgment.

The reported calibration evaluated PODA's specification and application through a self-run, four-case desk exercise. Measuring independent comparative efficacy requires a different evaluation design. Within the design used, it reached the same decisions as a matched critical-analysis baseline and, in one case, made explicit a burden that had remained implicit. The result supports a bounded

 doi.org/10.32388/28UoSW

contribution: PODA turns a tacit critical sequence into an inspectable path that runs from directed opposition, through reconstruction, to warranted commitment.

Corresponding author: Raúl Badillo Parabelo, raul@parabelo.com

1. Constructions begin to carry authority

A construction can become difficult to revise before anyone has tested the relation that makes it persuasive. A finding may be repeated as a recommendation, a product result treated as a basis for deployment, a policy proposal used to allocate burdens, or a design presented as settled. It remains revisable, but it has already begun to organize reliance.

Evidence in transit makes the problem visible. A local evaluation may show that a revised procedure reduces a named error under stated conditions. Once condensed into a recommendation for wider use, the result becomes part of a new construction linking a particular finding to a proposed action.

If the error category, comparison, population, or operating condition no longer travels with that construction, the recommendation can exceed what the study established even when each sentence remains defensible on its own.

The same problem appears beyond evidence transfer. A research contribution, governance proposal, service design, technical architecture, strategy, or public assertion begins to carry authority when someone is asked to believe it, fund it, deploy it, publish it, impose it, or rely on it. At that point, the relation carrying the requested commitment must survive scrutiny.

Established traditions already protect parts of that relation. Argumentation keeps support and qualification together; assurance ties evidence to active doubt; causal inquiry separates an observed difference from the mediator offered to explain it.

Evidence-to-decision work connects effects to alternatives and action, while documentation practices keep conditions of use and source paths attached to conclusions[1][2][3][4][5][6][7][8][9][10]. PODA draws on these operations as parent traditions.

Their formal tools usually begin with an argument, case, or causal claim that already exists in stated form. Toulmin's model, assurance-case notation, and the manipulationist test for a causal mediator do not specify what changes when an examiner is actively pressuring material that an AI system is still

producing turn by turn. Practice in these fields is often collective; the limitation lies in the formal tools, not in the surrounding disciplines.

PODA addresses the point at which those relations begin to guide action. It can operate while a design or argument remains open to reconstruction, or when an established finding starts supporting a later decision, whether AI-assisted iteration produced the construction or not. Domain methods still establish findings, and legal, governance, and professional standards still determine institutional authority. PODA keeps one path visible: pressure, damage, reconstruction, residual claim, and the commitment that claim can bear.

The name also evokes *poda*, Spanish for pruning. The term names a practical operation, not a biological analogy. Directed objections need not end in an all-or-nothing verdict; they locate the inferential link, scope claim, alternative, or proposed commitment the record cannot sustain, then preserve what remains.

AI changes this transition by compressing the time between formation and apparent completion. A person can now produce summaries, comparisons, recommendations, drafts, and polished artifacts through many rounds of iteration in the time an earlier reviewer needed to write one comment. The construction may look finished before anyone has recovered the conditions that constrain it.

Documented failure modes make unattended iteration unreliable. Models that evaluate and revise their own output show increased self-bias across iterations, favoring what they already produced over what the evidence supports[11]. Under sustained conversational pressure, a model can shift toward the preceding turn rather than the case[12].

Uncorrected output can also drift in topic, scope, or emphasis, which is the failure critique-refine methods are designed to address[13]. AI-only iterative generation shows that repeated self-processing can accumulate distortion without any critique step[14]. These studies do not settle whether reliable correction can come from the system itself or requires an engaged human counterpart. They do establish that fluency can remain intact while the construction loses its footing.

A separate failure appears at closure. Real pressure may run through the analysis and disappear in the final verdict, replaced by a balanced summary that acknowledges limits but never states what the record permits or rules out. The result can sound thorough while deciding nothing. PODA calls this closure-specific failure **structural tepidness**.

The term is original to this article, and the proposed mechanism remains a hypothesis motivated by two documented patterns. Models have refused clearly safe requests because their surface language resembled unsafe ones[15]. Follow-up work using the same test suite found models supplying a plainly wrong answer rather than an accurate answer that touched a sensitive topic[16]. In both cases, the model substituted a safer-seeming response for the response the task required.

Structural tepidness extends that concern to a third output: the final verdict. A system may reach a definite result and replace it with a softer summary because an explicit permitted-or-not-permitted line appears riskier than describing both sides. Refusal, factual substitution, and weakened closure remain distinct. The cited work documents the first two; PODA identifies the third as a testable extension rather than treating it as established.

The exchange therefore needs an identifiable locus of direction. That locus preserves the construction's practical purpose, sustains proportionate pressure against it, and decides when the record has earned the commitment being sought. Participant count does not supply those functions by itself.

Multi-model orchestration, debate, and voting can reduce some errors, but the models may share training data and related alignment pressures, so their failures are not independent. Unguided multi-agent debate has also shifted correct answers toward incorrect ones when agents favored peer agreement over sustained challenge[17]. More participants can expand the exchange; they do not establish direction.

Current systems can break down tasks, call tools, coordinate subprocesses, and propose routes a person did not specify step by step. That instrumental agency matters, but it still differs from preserving the purpose that makes a route worth taking, especially as the system gains autonomy.

PODA keeps a person inside the exchange as an active party, not only before configuration and after delivery. Dependent decisions accumulate around whatever the construction got right or wrong, so reopening it later requires unwinding more than the initial error. Sunk-cost research adds a second pull: people tend to persist in a course of action because of what has already been invested, even when that investment should not determine the next choice[18].

Both pulls strengthen around a polished output that has run for a long time without interruption. Reopening it starts to feel like discarding finished work rather than revising a claim that never earned closure. Earlier intervention keeps the unexamined interval short and limits what either pull can accumulate around.

Timing, not headcount, is the relevant safeguard. A person who intervenes only before configuration and after an unattended run meets the construction when dependencies and sunk cost are already strongest. A person who remains inside the exchange can redirect drift, test agreement, recover the relation under review, and decide what the case supports before the construction hardens.

In the typical case, that locus is the person who initiated the exchange because that person usually supplies the purpose, stake, and practical orientation the construction is meant to serve. This is a default, not a claim of human infallibility. A person can close too soon, accept fluency as support, or protect an idea that should be reconstructed.

Persistent and orchestrating agents may eventually hold direction across instructions, memory, permissions, tools, and closing criteria. PODA does not exclude that possibility. Every case reported here used a human locus, so the non-human case remains a scope claim rather than a demonstrated result. A non-human locus would need to show that it preserves an identifiable purpose over time, sustains pressure rather than its appearance, and corrects its own drift.

That burden rises with autonomy because users may not recognize whether pressure is real. In one controlled study, participants rated a consistently agreeable chatbot and a genuinely challenging one as equally helpful, even though the challenging system produced roughly ten times the improvement in task performance[19]. The appearance of useful friction is not enough.

AI systems also carry specific risks within the exchange. They may optimize for coherence, close before a real objection has been raised, soften an objection that should hold, over-integrate weak evidence, or comply with an implicit instruction rather than the case[11][12].

These risks do not reduce the issue to human versus AI. They support a narrower rule: no AI-assisted construction should acquire authority without proportionate opposition governed by an identifiable direction. Human direction is the current default; greater autonomy carries a greater burden of proof.

Chiodo et al. give this distinction a formal shape by separating trivial human monitoring, one human action at the endpoint, and highly involved human-AI interaction, and showing that the legal status and safety of the three setups differ[20]. PODA requires the third: arbitration throughout construction, not configuration at the beginning and approval at the end.

Cognitive-forcing research is close in spirit because it requires independent judgment before a person follows an AI recommendation, but it applies the discipline at a different point[21]. Cognitive forcing slows

acceptance at the decision point; PODA keeps the person producing and testing pressure throughout construction.

Here, AI-assisted inquiry means a sustained exchange in which an accountable person and one or more AI systems build and attack a construction together, while the person retains authority to redirect the work and close the case.

The resulting question is not whether AI iteration improves an output. It is whether the output has earned the commitment someone is about to place in it. PODA turns that question into an instrument for academic work, public policy, organizational decisions, product development, and strategic analysis.

PODA operates in two modes. In **live mode**, it governs opposition and reconstruction while a construction remains in formation. In **retrospective mode**, it reconstructs the claim-bearing relations of an existing object and determines the largest commitment the available record can carry. The illustrations and desk calibration later in this article mainly use the retrospective mode.

In both modes, PODA begins at the pressure point rather than with generic skepticism. It identifies the relation whose failure would change the object, defensible claim, or requested commitment. Opposition is selected for diagnostic value, not accumulated for effect.

A generic prompt can ask for skepticism, disagreement, or critical review. PODA fixes the relation under pressure and the criteria that make an objection material. It then requires reconstruction after damage, states the residual claim and commitment ceiling, and records when the case must be reopened. The result is a sequence, not an instruction to be critical.

2. From construction to proportionate commitment

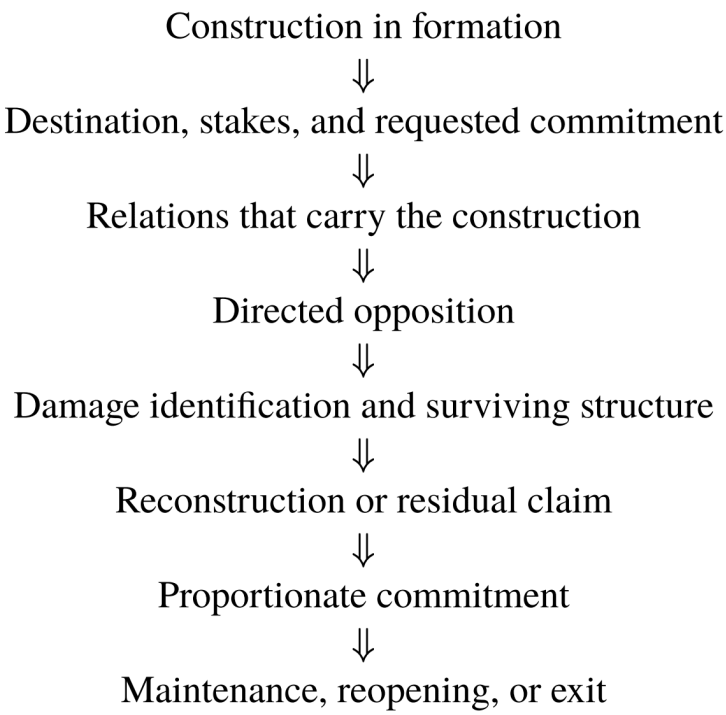
Consider a revised procedure that shortens the time needed to complete a defined task. The observed difference supports a statement about the conditions tested. A causal claim requires a mediator from the procedure to that difference. Wider adoption also requires an appropriate comparison, an account of implementation and repair burdens, and a decision by the person or body authorized to accept them. The same evidence may justify a local test while leaving wider use unresolved.

The example separates a finding, the construction built from it, and the institutional act that follows. For PODA, a **case** is the construction as it stands for a decision: the object, current conditions, relation under pressure, and immediate step it is asked to justify. The remaining terms name the roles needed to keep those distinctions visible.

Term	Practical role
Case	The object, conditions, relation under review, and requested decision.
Directed opposition	Targeted pressure on the relation that could change the object, claim, or immediate step.
Reconstruction	The revision, test, narrowing, separation, or exit required after a material attack.
Material determination	The status, permitted claim, active limit, reconstruction requirement, and next-step boundary.
Case file	The maintained form that carries the determination across review and change.
Permitted claim	The proposition supported by the available case after its active limits are retained.
Commitment ceiling	The largest immediate step for which the case supplies a proportionate basis.
Institutional authorization	The actor-specific authority to choose a permitted step.

Table 1. Terms used in a PODA case.

The operating movement is:



The sequence can loop. A new destination may make an earlier alternative decisive; a newly visible burden may require a narrower request before the evidence question closes. When fidelity fails, classification pauses until reconstruction identifies the object and relation. This movement prevents criticism from becoming the verdict and partial support from becoming standing authority.

PODA represents the process through the Material Admissibility Unit, reconstruction map, material determination, closure requirements, and maintained case file. These elements specify the case inputs, relations under review, outputs, and reopening conditions. The representation is conceptual rather than computational; it lives in the fields and decisions a reviewer works through, not in a particular interface, software package, or agent architecture.

Directed opposition tests whether the relation carrying a proposed step survives challenge. The claim's language determines the pressure that relation must bear. A causal claim requires scrutiny of the mediator from a difference to an outcome; a necessity claim requires a plausible alternative that preserves the same stake.

When a proposal distributes cost or risk, the case must show who bears it and who can repair a failure. Mason's dialectical approach offers a direct precedent because it tests the construction carrying a decision rather than merely collecting objections[22].

A **material determination** turns that pressure into scope. It records the proposition the case supports, the proposition that remains open, the conditions attached to the supported proposition, and the immediate step those conditions allow. That step is the **warranted commitment**. PODA does not replace this judgment with a confidence score.

Evidence and authorization answer different questions. The record may support a reversible pilot even when an institution lacks permission to expose a particular group. Conversely, an authorized actor may choose a smaller action than the record would permit because repair remains weak or maintenance would impose too much burden. A shortfall in permission must not read as a shortfall in support, and a shortfall in support must not read as a refusal to act.

Authorization is actor-specific. The author of a construction, a prospective buyer or adopter, and a regulator may face different standing, exposure, and remedies; the action warranted for one does not automatically transfer to another. Approval records responsibility accepted. It does not add evidence that the underlying claim is true.

A boundary statement therefore says what the case supports and what remains unsettled. A pilot may show that a revised procedure reduced a named error under stated conditions while leaving causation, transfer, or necessity open. Keeping those conditions attached prevents a local finding from becoming a general readiness claim during later use.

An untested alternative leaves necessity unresolved; it does not show that the current arrangement failed. The arrangement may still support a bounded local claim. Defeasible argumentation captures this position because a claim can remain acceptable while a counterargument is still open[23].

The commitment sought sets the depth of inquiry. Low-cost exploration can proceed with visible uncertainty. A mandate, broad deployment, or decision affecting access, safety, rights, or public resources requires a fuller source path, a fairer comparison, and a more developed burden account. Reversibility and the distribution of effects explain the difference. Proportionality work likewise links the justification of action to its consequences for affected people[24].

An organization may seek wider use while the record supports only monitored local application. Small approvals can then accumulate operational, contractual, or institutional dependencies that make later reversal costly. Stage-gate and lifecycle work draw the same practical distinction between evidence for the current stage and evidence for scaling[25][26]. A case is not the construction itself. It is what the construction has earned at the present stage.

3. Material admissibility: deciding what can bear pressure

Directed opposition needs a gate between a relation that could change the case and a feature that merely appears important. **Material admissibility** supplies that gate. An alleged difference gains claim-bearing weight when it affects an independently stated stake under the case conditions, and when its support, comparison, and burden profile can justify the claim and immediate step requested on its behalf. Visibility, unfamiliar terminology, and technical detail do not establish material weight.

3.1. Fidelity identifies the object and relation under pressure

Fidelity comes first because rich evidence can still address the wrong object or relation. A recommendation may concern a revised version of a program, procedure, or system rather than the version a study assessed. A summary may also turn a measured outcome into a claim about legitimacy, confidence, or readiness. Either new claim can be investigated, but each needs its own stake, mediator, and evidence. Detail about a nearby object cannot repair substitution.

Three questions guide the check. **Object fidelity** asks whether the object, version, and operating conditions remain those assessed. **Relation fidelity** asks whether the later claim preserves the comparison, causal connection, or other relation evaluated in the evidence. **Representation fidelity** asks whether the path from object to source material to interpretation retains information that could change the claim or immediate step.

Human and AI-assisted inquiry share the same limit: neither can work directly over every feature of a large corpus or complex object at once. Both depend on selected materials, retrieved passages, notes, summaries, and reconstructed representations. Compression becomes material when an omitted condition could change the object, claim scope, mediator, alternative, burden, or proposed action. A verified summary may support a narrow descriptive statement and still be inadequate for an action affecting rights or access.

When the available material cannot identify the object, asserted relation, or source provenance well enough for classification, the case enters an **evaluability hold**. The hold records what remains unknown and what reconstruction is required. Work that does not depend on the unresolved relation may continue, but primary status waits. Unlike *ornamental* or *risky or unjustified*, an evaluability hold makes no judgment about the alleged difference itself.

The consequence of error sets the depth of reconstruction. Missing detail that cannot alter a limited claim may allow a narrow report based on a verified extraction. High-stakes action requires a more direct source path, clearer version control, and a record of the conditions that made earlier evidence applicable. Versioned documentation and intended-use records help later readers identify when a finding may have become stale[27][8].

Once fidelity fixes the object and relation, directed opposition can target the challenge most likely to alter the material determination: a weak foundation, unsupported mediator, claim that has outgrown its conditions, unfair alternative, or burden distributed without justification.

3.2. The Material Admissibility Unit

A feature earns material weight only when its claimed consequence can be stated outside the arrangement's own vocabulary. A review layer, for example, matters if it gives an affected person a real chance to correct a wrong decision before it takes effect. That consequence can be investigated, contested, and compared. Merely stating that a process contains a review layer cannot establish it.

A governance delay makes the distinction concrete. The independent stake may be a member's practical opportunity to contest a harmful proposal before execution. Its mediator is that the delay gives members usable notice and a genuine chance to respond. Late notice, inaccessible participation, or a response process unable to produce correction weakens that mediator. The case now concerns a testable relation rather than a feature receiving credit from its label.

The mediator need not be causal in a narrow scientific sense. Causal claims require causal mechanisms; legal, historical, interpretive, and normative claims may depend on a warrant or another inspectable relation between material and conclusion. In every setting, the case must state how the alleged difference is supposed to matter before the difference receives credit for the outcome.

The **Material Admissibility Unit** (MAU) gathers the relations needed to place an alleged difference under directed opposition. The requested action sets the depth of inquiry; the other fields establish what claim and immediate step the case can justify.

Field	What the reviewer must establish
Δ, alleged difference	The feature whose effect is being asserted.
S, independent stake	A consequence stated outside the arrangement's own terms.
M, mediator	The causal, functional, or other link from Δ to S.
E, evidence	Support for the mediator and the path through which error could be found.
A, alternative	The option relevant to the claim under matched conditions.
B, burden	Practical, distributional, repair, and exit effects by group and time.
C, conditions	The setting, scale, constraints, and limits on application.
D_request	The claim and immediate step that the case is asked to justify.

Table 2. Material Admissibility Unit.

The MAU does not return a binary admission result. It prepares the relation for targeted pressure and feeds a **material determination**. That determination records the primary status, permitted claim, active limit, reliance warranted, commitment ceiling, open authorization question, and reopening condition. Status classifies the alleged difference; the completed determination tells a later decision-maker what may be said and done on the record as it stands.

A system with several claimed contributions may require linked MAUs because success at the system level does not give each component independent credit. A component said to preserve safety, protect an appeal right, or reduce administrative effort needs its own mediator to a stake, or an explicit dependency on a supported mediator, before it receives causal, necessity, or decisive credit. Assigning every component a property shown only by the package commits the fallacy of division[28].

Stake selection must be visible because it determines whose consequence counts. The case identifies whose interest defines the stake, who receives the benefit, who carries the burden, and who can contest or repair a failure. PODA cannot settle every disagreement about standing. It can prevent organizational convenience from quietly replacing the protection or capability the claim is supposed to serve.

A stake may be procedural. Knowing about, contesting, or appealing a decision is itself a protected consequence. A rule that changes who may object or how they may object is not ceremonial merely because the physical output remains unchanged. Its materiality depends on whether the people it names can actually use it.

3.3. Evidence, comparison, and burden set the required pressure

Support matters because it creates, or fails to create, a route through which error can be found. Reviewers need to know who produced the object, selected the outcome, generated the evidence, scored the result, and could identify or correct an error.

A favorable internal evaluation may support a narrow statement about its own test. Broader reliance requires stronger reasons to believe that the system selecting the favorable result did not also control every path through which error could be discovered.

Assurance reasoning keeps evidence and active doubt together. A claim remains tied to both the support behind it and the doubts that could defeat or limit it[2][4]. A later review should reveal the supported proposition and the condition that remains open.

An observed effect and its explanatory mediator may need different support. A benchmark can show that an outcome changed while leaving open why it changed, whether a named component caused it, and whether the effect will survive new operating conditions. Here the case file preserves that distinction: it may permit a local effect statement while withholding causal attribution, transfer, necessity, or wider readiness.

Independence varies by degree. Evidence may be generated by the proponent but scored against a public measure, produced through a multi-party process in which the designer does not control every error-discovery path, or independently replicated. These pathways affect how far a later reader may carry the result. PODA records them instead of collapsing them into a simple internal-external label.

Alternative requirements follow the claim's language. Necessity requires a less burdensome substitute that preserves the stake. Novelty, uniqueness, or differentiated-synthesis claims require a direct or distributed counterpart that performs the central work. Whole-arrangement comparison tests the package outcome; component-level claims require evidence about marginal contribution. Substituting one comparison for the other lets the claim survive an easier test than its language demands.

Labels cannot settle the comparison. Different terminology, an acronym, or a new diagram does not make a familiar operation different; a shared name does not make two mechanisms equivalent. The comparison must reach mechanism, output, and consequence.

Definition level matters too, fixed at the level where the claim operates. A definition broad enough to make every alternative equivalent trivializes the comparison. A definition narrow enough to exclude a genuine substitute makes it misleading. Two arrangements may produce the same visible output and still differ in safety, reliability, or accountability; those differences remain part of the function when the claim depends on them.

A search limited to the claimant's own field or market may miss an equivalent elsewhere. A substitute should not be dismissed because it carries a different category label when an affected party can use it for the same practical job. Novelty can therefore appear in terminology, function, or architecture without appearing in all three. Each level supports a different strength of claim.

The applicable standard also depends on the domain because legal non-obviousness, scientific contribution, and market differentiation test different questions. Terminology borrowed from one cannot settle another.

Before a relative claim receives credit, the alternative must be specified, preserve the relevant stake, and operate under conditions that make the alleged difference meaningful. An alternative that drops the protected stake cannot settle necessity; an alternative tested under materially easier conditions cannot settle practical superiority. A candidate alternative still has value because it identifies the uncertainty limiting the current claim and directs the next inquiry.

Alternative status	What it allows
Not needed for this claim	A narrow descriptive or local effect claim may proceed.
Candidate identified	The case may retain bounded use while necessity remains open.
Fairly tested	The comparison can limit or support a relative-value claim.
Inadequate or mismatched	The relative claim stays open until conditions are repaired.

Table 3. Alternative status and its effect on the claim.

Burden belongs beside comparison because an arrangement can protect a stake and still be indefensible in the setting that matters. The account should show who carries the practical load, how failure can be repaired, and whether affected people can exit or contest the arrangement.

Reversibility changes the weight of each burden. A minor delay in a reversible internal test differs from one that blocks an appeal or makes a public service inaccessible. Administrative-burden research captures this distribution through learning, compliance, and psychological costs that can fall unevenly[29].

A benefit can create new burdens through access, dependency, or power. Such burdens should be named even when the immediate result looks like a clean gain. A lower-burden alternative is not automatically preferable; it must still preserve the protected stake rather than move the cost elsewhere. Omissions that make a construction look stronger deserve particular scrutiny.

Uncertainty remains usable only when it changes the claim or next step. An open alternative may permit a monitored local test while excluding necessity. Likewise, a repair burden may permit shadow evaluation while blocking customer exposure. The uncertainty must remain attached to that boundary rather than dissolve into a general statement that the system is ready.

4. Directed opposition turns damage into reconstruction

Once a case identifies a claim-bearing relation, opposition asks what could change it. The aim is to select the challenge whose success would alter the construction, residual claim, or proper commitment level. This avoids both untargeted attack and an unranked inventory of weaknesses.

Pressure may target the foundation or mediator carrying the claim, the claim's scope or evidence coverage, a plausible alternative, a burden, or the temporal and traceability conditions needed for review. These are selection categories, not mandatory stages. Rigor comes from testing the pressure point that can change the case.

Directed opposition is defined by its relation to the case, not by the volume, severity, or tone of criticism. It targets a relation whose disruption could shift the permitted claim or commitment ceiling, engages the relevant evidence, alternative, burden, or condition, and carries material damage into reconstruction and closure.

That specification makes the operation identifiable within the record. Inter-rater reliability, construct validity, and predictive validity remain questions for later empirical evaluation; the present task is to specify what counts as directed opposition.

The challenge itself can fail. An alternative may demand resources only a specialist team could assemble; a function may be defined so broadly that every alternative appears equivalent; a past decision may be judged with information that emerged only later. These are defects in the opposition, not findings about the case. Pressure earns its place by testing the claim against a fair alternative under the conditions set by the claim's own language.

Opposition also leaves the case when it shifts from the claim to the person who made it. Questions about motive or competence belong only when independent evidence makes them material. Otherwise, the objection must be restated against the claim, evidence, or inference.

A disclosed limitation defines the claim's boundary, so a construction should not be faulted for failing to prove something it never claimed. Likewise, an unsupported advantage is a gap in support, not evidence of deception, unless the record independently shows that the claim was known to be false.

Length does not make a critique successful. The critique matters when it changes a permitted claim, status, or commitment. A construction that depends only on its own material should first be tested for internal sufficiency; if it fails on its own terms, outside literature is not needed to manufacture the failure.

Familiar fallacies locate recurring defects without adding new mechanisms to PODA. A mediator that restates the stake is **circular**. Attacking a weakened version of the claim is a **straw man**. A **slippery slope** assumes an unsupported chain of consequences. **Moving the goalposts** raises the standard after a result has met the one previously set. The names identify where the pressure left the case.

Once an attack changes the case, it must produce a reconstruction choice. Criticism that only removes support leaves the next justified claim, test, redesign, pause, or exit unspecified.

Attack result	Reconstruction response
Foundation weakened	Narrow the claim, add evidence, alter the mechanism, or pause.
Warrant or mediator broken	State the missing condition, change the inference, or test the mediator.
Scope exceeded	Add qualifiers, restrict the destination, or split the claim.
Alternative remains plausible	Compare, triangulate, or design a discriminating test.
Evidence coverage incomplete	Search, consult, add an alternative, or disclose the limit.
Feasibility uncertain	Prototype, model cost, run a pilot, or reduce scope.
Burden unacceptable	Redesign safeguards, change stakeholder exposure, or do not proceed.
Temporal basis obsolete	Update evidence and distinguish the earlier case from the current one.
Novelty overstated	Reposition the object as adaptation, integration, or implementation.
Traceability insufficient	Reconstruct only what the record permits and preserve uncertainty.

Table 4. From attack to reconstruction.

The reconstruction response depends on object maturity, destination, stakes, reversibility, evidence quality, representation fidelity, and time state. Public deployment may require stronger repair than an exploratory design; uncertainty that permits a local test can still block a mandate.

Status then states what has been established. The map below is a navigation tool, not a score or universal maturity ladder. A new alternative, population, request, or body of evidence may move the same case to a different status.

Primary status	What the case supports now
Visible distinction	Identify the difference and formulate a test.
Unproven material hypothesis	Investigate a plausible stake effect through a focused test.
Material under conditions	Use a bounded finding for reversible, monitored local action.
Materially necessary	Consider monitored integration after a fair test fails to preserve the stake.
Ornamental	Reconstruct, separate, or exit because no stake effect warrants the burden.
Risky or unjustified	Pause, reconstruct, or exit because the requested step has an indefensible burden.

Table 5. Primary status and immediate consequence.

Materially necessary carries a narrow burden. The case must support a stake-relevant effect under stated conditions, show that a fair and less burdensome alternative failed to preserve the core stake, and retain a defensible burden profile for the original arrangement. A favorable effect alone cannot establish necessity.

The two negative statuses identify different failures. **Ornamental** applies when the alleged difference has not shown an independent stake-relevant effect sufficient to warrant its burden. **Risky or unjustified** applies when the requested step carries an indefensible burden even though the case identifies a real effect. This distinction allows a useful feature to survive while the proposed scale, form, or burden distribution is rejected.

Two subordinate qualifiers preserve narrower results. **Useful under limited evidence** marks a bounded local role with an active limitation. **Useful integration** identifies a practical coordination of familiar practices. Neither qualifier raises primary status, proves necessity, or establishes an incremental decision effect over a fair bundle of alternatives.

A material determination gives authority a scope. It states the proposition a reader may rely on, its conditions, and the proposition that remains open. The commitment ceiling translates that scope into a maximum immediate step. Because a ceiling is a boundary rather than a recommendation, an authorized actor may choose a smaller action, seek more evidence, or decline because of competing duties.

A **reopening condition** identifies a change that could alter the determination; a trigger makes the change visible. A revised object, affected population, alternative, or shift from pilot to mandate can reopen the case, since each changes a relation on which the earlier determination relied. By contrast, a corrected source link or formatting change need not.

Reconstruction and exit serve different purposes. Reconstruction makes a case evaluable or repairs a weak relation. Exit closes or separates an arrangement when the alleged difference is ornamental or the requested action remains risky or unjustified. The distinction converts a limitation into a next step.

The optional sequence below arranges the dependencies without turning them into a one-way ritual. Domain methods remain intact, and no score is produced.

Phase	Stages	Output
Capture	**0, 1, 2, 3**	Defined case, fidelity hold where needed, Δ, and S.
Mediator	**4, 5, 6**	Stated mediator, evidence limit, and burden account.
Comparison	**7, 8**	Alternative rationale and comparative finding.
Determination	**9, 10, 11, 12**	Status, scope, ceiling, and reopening or exit rule.

Table 6. Optional application sequence.

The sequence can return to an earlier stage. A broader request may make a previously irrelevant alternative decisive; an authorization constraint may narrow the request before final authorization is assessed. When fidelity fails, classification waits until reconstruction identifies the object and relation.

5. Commitment, authorization, and the maintained case

PODA ends with a maintained determination of what the case supports now. **Warranted commitment** is the immediate step that remains proportionate after the object, stake, mediator, evidence, comparison, burden, and conditions have been considered together. Directed opposition supplies the pressure; reconstruction responds to the damage; warranted commitment states the practical result.

Proportionate does not mean formulaic. The MAU does not convert burden, alternative strength, and reversibility into a number that automatically produces a ceiling. It identifies the relations judgment must account for. Two competent reviewers may reach different ceilings from the same file, and PODA cannot decide between them internally. It can require each reviewer to show which field drove the judgment.

When a real commitment is at stake, the determination must close beyond a balanced summary. It states:

1. what the record permits now; 2. what it does not yet permit; 3. which claim survives; 4. which claim remains withheld; and 5. what action would be irresponsible on the present record.

It also names the condition that could justify a larger step later. A determination that acknowledges limits but never draws an operative line has not closed the case. It has reproduced the pressure in the form of structural tepidness. The five statements preserve nuance while still producing a verdict.

The closure test establishes form, not merit. A verdict can answer all five questions and still rest on evidence that was never pressured, an unfair alternative, or an incomplete burden account.

Testing that failure requires returning to the MAU field the closure claims to summarize, Δ, S, M, E, A, B, C, or D_request, and asking whether the field survived directed opposition. The closure test detects stopping without deciding; it does not detect deciding without earning the decision.

The MAU has the same limit. Completing eight fields does not show that all eight survived pressure. A case file may report every field while leaving the mediator untested or accepting the evidence at face value. The MAU structures where to look; it does not substitute for looking. It also protects the opposite distinction: an unresolved mediator does not make the underlying effect false, and a real effect does not make the larger commitment warranted.

Calling PODA a **derived analytical framework** defines its role. It organizes relations while leaving domain methods intact. Research methods, legal review, participatory process, and assurance-case work continue to establish findings within their own fields.

PODA enters when an object is still open to reconstruction or when an established finding begins to guide a later decision. In both settings, it keeps the relations that could change the claim or immediate step inspectable. This is the bounded role of an analytical framework: arranging concepts and relations so an object can be examined as a connected case rather than as disconnected fields[30][31].

The parent traditions contribute by task. Argumentation and assurance qualify claims and preserve active doubt. Dialectical and causal inquiry test the mediator from an alleged difference to an

independent stake. Evidence-to-decision and proportionality work connect effects to alternatives, burdens, and action. Modular design and traceability preserve component attribution and evidence maintenance. Governance and responsible-innovation work keep affected parties, responsiveness, and revision duties visible[1][22][6][7][32].

Field-specific methods still govern how knowledge is generated and evaluated. PODA cannot turn a source record into evidence or decide a legal, clinical, engineering, or institutional question by itself. It connects the relations that matter when an object begins to carry a later claim or commitment. A parent tradition belongs in the source set when it contributes a relation that could change that claim or immediate step.

The article follows a theory-synthesis design in Jaakkola's[33] sense. Its focal phenomenon is the transition from a construction under directed opposition to the commitment requested in its name. None of the selected parent traditions organizes that transition as one unit of analysis.

Argumentation and assurance qualify claims and support; dialectical and causal inquiry test the relation between an alleged difference and an independent stake; evidence-to-decision and proportionality connect findings to alternatives, burdens, and action. PODA brings those operations into one maintained case.

That case must survive directed opposition, be reconstructed after material damage, close on a permitted claim and commitment ceiling, and reopen when the relations supporting the determination change. This sequence produces the claim-to-commitment architecture developed here. PODA's vocabulary names the relations created by that integration rather than renaming the tools from which the framework was derived.

The relevant methodological tests follow from that design: whether the selected traditions and assigned roles are justified, whether the chain from inherited concepts to the framework remains visible, and whether the synthesis produces a coherent account of the focal phenomenon. The MAU, status map, reconstruction sequence, closure requirements, and case file specify that account.

They do not estimate a population-level effect or validate the combined operations across sampled users. User studies, statistical estimation, and prompt-level benchmarks would test separate claims about usability and comparative efficacy.

AI-assisted work makes the difference between output and commitment especially visible. A generated artifact can be fluent, structured, and internally consistent while resting on an untested mediator,

alternative, or burden account. AI may retrieve candidate sources, compare versions, draft case-file language, propose an alternative, or generate directed challenges.

The person running the case remains responsible for verification, stake selection, mediator assessment, burden judgment, reconstruction, and institutional action throughout the exchange. Generated output supplies no authority by itself, whether it comes from one pass or many rounds of critique. Before anyone relies on it, the path from output to claim must remain inspectable, including who redirected the exchange and when.

Optionality is part of the framework's logic. The MAU, status map, commitment rules, and application sequence earn their place only when they preserve a relation that could otherwise change the claim or immediate step. A competent assurance, legal, governance, clinical, or technical process may already preserve those relations. In that setting, PODA may help compare or communicate the record, but duplicating the existing process would add burden without a demonstrated gain.

A maintained determination therefore needs only the fields a later decision could change. The case file can attach to an evaluation report, design record, policy memorandum, or assurance artifact. It should allow a later reader to recover why an earlier step was permitted, which limitation remained active, and what event would require reconsideration.

Group	Minimum content
Identity and request	Object, version, representation, Δ, conditions, actor, destination, and immediate step.
Source and mediator	Source lineage, stake, mediator, evidence, and active limit.
Comparison and burden	Alternative rationale, unmatched condition, affected group, timing, repair, and exit.
Determination	Status, permitted claim, withheld proposition, authority scope, and ceiling.
Governance and revision	Authorization state, reopening condition, trigger, owner, and required response.

Table 7. Case-file fields for review and maintenance.

The file earns its place by preserving a field that could change a later decision. If a report says a technique reduced a named error under one protocol, the file keeps that protocol attached to the downstream claim.

A later reader can then decide whether the requested action relies on the same conditions or requires more work. The aim is continuity, not documentation of everything.

Source lineage separates origin from verification. A statement may originate in a report, retrieved excerpt, generated summary, or manual synthesis, while its verification state remains verified, partially verified, or awaiting reconstruction. These labels direct reviewers to the right source path when a high-stakes assertion appears to have lost its qualifier. They do not rank sources by form alone.

Record depth follows consequence. A low-risk case may need only a source link, permitted claim, immediate step, and reopening condition. Higher-stakes cases may require fuller source lineage, comparison rationale, burden account by affected group, and a named review owner. Reversibility, affected parties, and reconstruction cost determine the appropriate depth.

Disagreement need not collapse into one issue. A disputed entry can record the proposition, basis for disagreement, information that could narrow it, and interim step that remains proportionate. Technical, legal, policy, and operational reviewers may update different fields while keeping distinct responsibilities. The shared artifact functions as a boundary object because it coordinates work without assuming that all participants make the same judgment[34].

Maintenance prevents the file from becoming a static justification. When a trigger occurs, the owner identifies the changed field, records the new evidence or condition, and decides whether the earlier status and ceiling still apply. The record preserves both why a past step was reasonable and why a later step inherits, revises, or loses that permission.

Reassessment separates present correction from retrospective blame. Evidence may become stale because the object, context, or requested action changed. That finding identifies what must be reconsidered now; it does not show that the earlier decision was irrational when made. Traceability supports correction without turning every revision into an accusation.

6. PODA and adjacent AI-assisted work

The comparison below is positional rather than evaluative. It identifies the question each system was built to answer and does not rank their overall importance. Because favorable positioning would benefit related work by the same author, that stake is part of the comparison's stated conditions.

ARIS[35] is an open-source research harness built around cross-model adversarial review. It targets what its report calls a plausible unsupported success: a long-running agent produces claims whose evidential

support is incomplete, misreported, or silently inherited from its own framing. The problem closely resembles a construction acquiring authority before the relation behind it has been tested.

Its companion tool, Anti-Autoresearch[36], applies the same forensic orientation to third-party papers. A deterministic, rule-based adjudicator checks numeric consistency, citation integrity, baseline adequacy, and proof-derivation validity rather than leaving those checks to a model.

ARIS normally runs the adversarial exchange between models. A person configures the harness beforehand and reviews the result afterward instead of arbitrating each exchange as it happens. Anti-Autoresearch draws the same boundary in its documentation: its findings are flags for human review, not verdicts on publication, funding, or adoption.

PODA addresses the next decision. Once a claim has been admitted in bounded form, whether by a deterministic adjudicator, human reviewer, or AI-assisted exchange, the case still has to establish what action that claim can support. That step requires a burden account, an alternative appropriate to the requested scale, and an authorized actor.

A system can pass fidelity and consistency checks while still being asked to carry a commitment its evidence does not cover. By the systems' stated scopes, that remaining question belongs to PODA.

A related boundary appears in human-in-the-loop AI-assisted programming. Dranias and Whitley[37] call locally plausible code that diverges from its task specification **objective drift**. Their method keeps a specification and bounded option space in place while execution is monitored against them, using a mechanism rooted in control theory rather than argumentation or assurance.

Their account leaves transfer beyond software development open. Tests and compilation provide fast, checkable feedback; research novelty, policy burden, and institutional authorization do not. These cases require judgment over relations no compiler can settle. PODA operates in that weaker-feedback space.

Papamarkou and Floridi's[38] **Turing's mirror** uses the central term *warranted commitment* closely enough to require direct distinction. It asks whether a person's expressed confidence in an AI-generated answer matches the warrant supported by one exchange, which they call protocol-relative warranted commitment. PODA tracks a case across an alleged difference, stake, mediator, alternative, burden, and possible institutional act. The shared term marks adjacent territory; the unit of analysis and practical output differ.

Taken together, these systems divide the work rather than form a ladder. ARIS and Anti-Autoresearch test support and integrity; the control-theoretic account monitors execution against specification; Turing's

mirror compares confidence with warrant inside an exchange. PODA asks what commitment a bounded claim can carry after those or other forms of review. Each system performs work the others do not attempt.

7. Two worked illustrations

The illustrations apply PODA to material that existed before this article. They are not discovery claims and do not show that the framework outperforms competent critical analysis. Their purpose is narrower: to show how directed opposition can end in a checkable line between what is permitted, what is withheld, and what action would exceed the record.

7.1. A time-bound public record

The first illustration examines a company's filing using only the public record available while the filing was asking the market for a commitment.

The window begins on August 14, 2019, when The We Company filed its Form S-1 for an initial public offering, and ends on September 30, 2019, when the offering was formally withdrawn[39]. Later events do not enter the case. This controls the evidence available to the analysis, although it cannot remove a reader's knowledge of the eventual outcome.

The illustration therefore does not test predictive ability. It tests whether a historical claim can be reviewed proportionately using only the record available before commitment. The six-week window matters because the filing was then asking investors to buy shares under a specific narrative, valuation, and governance structure while those claims remained open to challenge. The relevant speed is not how quickly an analyst writes; it is how quickly a construction accumulates authority, reversal cost, and commitment pressure.

Λ, alleged difference: the filing presents the company as meaningfully different from a conventional flexible-office operator because of its technology platform and community-based membership model, rather than real-estate arbitrage.

S, independent stake: whether that difference creates an economic advantage an outside investor would recognize through a shorter path to profitability or lower operating risk, not only through language about community and shared purpose.

M, mediator: the platform and community model convert the alleged difference into the economic advantage assumed by the requested valuation and governance structure.

The evidence inside the window does not clearly support that mediator. Revenue and losses grew at nearly the same rate from 2017 to 2018. The filing had replaced an earlier non-GAAP metric with a narrower one without presenting a GAAP path to profitability.

Contemporary analysts identified the same weakness in the filing. On August 20, a specialist analyst told Bloomberg that the prospectus obstructed an accurate financial model rather than enabling one[40]. A separate analysis estimated that the disclosed unit economics would require roughly thirteen years for the average tenant to reach breakeven[41].

A fair alternative was already public: an established flexible-office operator running the same core business at a comparable revenue scale, operating profitably, and valued at a fraction of the premium implied by the filing[42]. The alternative does not erase the company's stronger growth rate. It prevents growth alone from establishing the technology-platform premium or governance terms requested from investors.

The burden fell mainly on outside investors. The filing's three-class share structure gave Class A buyers one vote per share against twenty votes per founder-held share. Public pressure reduced that ratio first to ten through a filed amendment[43], then toward three when the founder resigned as chief executive weeks later[44].

Governance was not the only burden already visible within the window. Related-party leases and a reversed trademark payment were disclosed[45]. Future lease obligations stood at tens of billions of dollars by mid-2019 and would not shrink automatically if membership fell[39].

Reconstruction preserves the underlying business and rejects the unsupported mediator. The record supports a real, growing flexible-office company with members and locations. It does not support the move from that business to a technology-platform valuation or to governance terms the alternative did not require.

The residual claim is therefore bounded: the filing supported continued operation as a flexible-office business, not the larger commitment requested from outside investors. The status is **risky or unjustified**, rather than ornamental, because a real underlying business remained while the burden of the requested commitment was indefensible on the available record.

Contemporary analysts were already making this argument. PODA did not discover it; the illustration shows how the record can be organized into a permitted claim, withheld claim, burden account, and commitment ceiling.

7.2. A live, unresolved claim

The second illustration concerns a claim whose outcome remains unsettled: that AI-assisted coding tools have already shown a productivity effect large enough to justify enterprise-wide mandates, staffing reductions, or reduced human review. The narrower claim, that the tools can help under some conditions, is not in dispute.

Δ, alleged difference: AI coding tools and coding agents allow developers to produce, review, and merge code faster than workflows without them.

S, independent stake: whether teams deliver more valuable, maintainable, and secure software per unit of effort, not merely more visible output.

M, mediator: increased visible output becomes a net productivity gain large enough to justify the organizational commitment sought in its name.

Several studies support the first link. Developers using GitHub Copilot on a bounded coding task completed it 55.8% faster than a control group[46]. Across three companies and roughly 4,900 developers, access to an AI coding assistant raised completed tasks by 26%.[47].

Evidence from broader deployments also shows higher visible output. Microsoft adopters of command-line coding agents merged roughly 24% more pull requests than they otherwise would have[48]. One enterprise mandate tracked for more than a year reached 2.09 times its pre-mandate throughput[49]. AI coding tools can increase visible output under some conditions; that effect survives directed opposition.

The mediator to net organizational productivity is less secure. In a randomized trial, experienced developers working on mature open-source projects took 19% longer when AI assistance was allowed, even though they believed it had helped[50]. A later attempt to update the result had to be abandoned as more developers refused to work without AI, leaving only weak evidence about how far the picture had changed[51].

Open-source adoption shows a related burden shift. Gains concentrated among less experienced contributors, while more experienced developers absorbed 6.5% more review work and their own original output fell by 19%.[52].

That positive effect weakens outside controlled settings. Across 23 studies, the average effect was positive but stronger in controlled settings than in open-source and enterprise settings[53]. A dedicated review of an AI code-review feature found reliable detection of style issues but frequent misses on SQL injection, cross-site scripting, and other high-severity vulnerabilities[54].

Zero adoption is not the fair alternative. The relevant alternative is targeted, monitored adoption with human and security review preserved, tested against downstream measures such as review load, rework, and defect rates before wider scaling. That alternative has not been tested and beaten, so a mandate has not been shown necessary.

The burden falls on the people who absorb review, rework, and security risk created by higher output, not necessarily on those reporting the increase. The burden-shifting study and meta-analysis make that distribution visible from different angles.

Reconstruction separates two claims. The narrower claim, that AI coding tools can materially increase visible output under some conditions, is **material under conditions**. The stronger claim, that the effect is already large and reliable enough to justify mandates, staffing cuts, or reduced review, is an **unproven material hypothesis**. It remains plausible enough to test and too weak to carry the requested commitment.

The record permits targeted, monitored adoption with human and security review preserved. It does not permit enterprise-wide mandates, fixed productivity multipliers used as staffing targets, headcount reductions or denials justified by AI adoption, or reduced human code and security review. Those actions depend on a mediator the current evidence has not closed.

The longer public debate already contains the tension between visible output and net organizational value. PODA contributes the operating line: which claim survives, which claim remains withheld, which alternative must be run, and what action would exceed the record now.

8. PODA under its own directed opposition

PODA's evidence must be matched to the claim being made about it. Developmental work can show that the framework can be specified and applied. Source-corrected records can show that the parent traditions and their assigned roles have been represented accurately. Neither establishes an incremental decision effect. That stronger claim requires a matched comparison against competent critical analysis using comparable materials and the same relevant parent traditions.

A **decision-relevant difference** is a predefined change in the permitted claim, commitment ceiling, or another decision output that matters to the case. Clearer source traces, visible limitations, and explicit alternative rationales may improve documentation. They are reported separately because they do not establish a better decision by themselves.

Activity	Public implication
Source and provenance work	Parent roles are traceable; no synthesis effect is established.
Exploratory applications	The sequence can be applied; no decision improvement is established.
Matched desk calibration	Four matched desk cases; no case changed the baseline's permitted claim or commitment ceiling (table below).
Reflexive self-application	The public description remains bounded by the evidence.

Table 8. Developmental evidence and its public implication.

The calibration labels separate decision-level change from documentation-level change. **R0** means no difference from the matched baseline. **R1** marks clearer wording or conceptual separation without a different permitted claim or ceiling. **R2** adds an alternative, burden, or distribution question the baseline left implicit, again without changing the decision.

R3 would change the permitted claim, classification, or commitment; **R4** would avoid an unjust, unsafe, or irreversible loss. Neither R3 nor R4 occurred. These labels describe a developmental desk record, not an independent performance estimate.

The baseline received comparable case materials and used the parent traditions a competent analyst would ordinarily apply. It was therefore a demanding comparison rather than an artificial weak baseline. The same author produced the case work and baseline judgments in the same working sessions, so the exercise remains a self-application rather than an independent trial.

Across four cases, covering a methodological necessity claim, a governance burden claim, a due-process comparison, and a commercial reliability claim, PODA changed neither the baseline's permitted claim nor its commitment ceiling.

That result applies to the comparison actually run. The baseline was already structured and competent; no unstructured condition was tested in which a claim simply traveled without examination. PODA could perform differently against that weaker practice, but this calibration cannot establish the difference because it ran only the harder comparison.

The evaluation method matches the current claim. This article specifies and applies a theory-synthesis framework, so the calibration examines cases and decision outputs rather than estimating an effect across a population. User studies would test usability, adoption, cognitive burden, and performance across users. Statistical comparisons would estimate effects across sampled cases, reviewers, or institutions. Those are later layers of evaluation, not substitutes for specifying the conceptual architecture.

Case	Claim under test	Baseline commitment	PODA commitment	Result
Multi-agent methodology	A council of models is necessary for reliable conclusions	Reversible comparative benchmark; no institutionalization	Same commitment; no necessity claim	**R0**
Human review for high-value payments	Universal human review is indispensable against fraud	Comparative pilot; no universal deployment	Segmented deployment with monitored thresholds	**R2**
Oral hearing vs. digital form	A digital form preserves due process on its own	Hybrid model; no universal elimination of hearings	Same hybrid model, with the appeal right stated more explicitly	**R1**
AI system and commercial claim	A multi-agent architecture is superior and has no substitute	Benchmark by contract type; no price premium on unproven grounds	Same commitment, with reliability and non-substitution separated as distinct claims	**R1**

Table 9. Four-case desk calibration.

doi.org/10.32388/28UoSW

One case produced no difference from the baseline. Two clarified wording or separated concepts without changing the decision. One made an alternative and burden-distribution question explicit, again without changing the final decision. None reached R3 or R4.

The calibration therefore withholds the stronger instrumental claim that PODA improves decision outputs beyond competent critical analysis. It does not establish that PODA makes no difference in every setting, because the exercise covered four self-run cases and was not independently replicated. Within that boundary, the result supports the article's actual contribution: an explicit, inspectable claim-to-commitment architecture and a demonstrated way to apply it retrospectively.

The baseline also clarifies where that architecture can operate. Structured critical review depends on training, time, and institutional habits that cannot be assumed across users, domains, or decision points.

PODA supplies a sequence that competent review otherwise has to generate on its own: identify the relation carrying the claim, direct pressure at the point whose failure would change the case, reconstruct what survives, and close on a residual claim and commitment ceiling. The framework organizes the path to judgments of materiality, reconstruction, and commitment without replacing the person responsible for making them.

The calibration was designed, scored, and interpreted by the framework's author. Research on psychotherapy outcomes documents a comparable risk under **researcher allegiance**: studies led by researchers who favor one treatment tend to report results in that treatment's favor[55]. That risk makes inflated results more plausible, and stating it does not correct the design. The result should be read with that discount in place.

Within that boundary, no case reached above R2. The result is reported because it did not produce the decision-level gains that allegiance would most obviously favor, not because the design removes allegiance. Removing one plausible source of inflation does not itself establish a positive effect.

Under that stated condition, the absence of R3 and R4 results matters. A self-run exercise with a favorable stake did not produce evidence of changed decisions, so the article does not convert documentation gains into a performance claim. The calibration sets the boundary before the contribution is stated.

On the present record, PODA is a derived analytical framework that can serve as an optional facilitation protocol. It makes claim-bearing relations inspectable and carries their limits into an operational verdict. It is not warranted as a mandatory layer where competent review already preserves those relations.

A reader may still apply PODA and judge that it helped. Technology-adoption research recognizes **perceived usefulness** as a user's belief about a tool, independent of whether later outcome evidence confirms that belief[56]. Such a report would describe fit for one case, not the decision-level effect withheld by the calibration.

Documentation outcomes remain separate from better decisions. An MAU or case file may help locate a limitation, reconstruct an evidence transfer, or make an alternative rationale explicit. Until matched comparison shows that these changes alter a predefined decision endpoint, they remain documentation and organizational hypotheses.

Three controls keep the developmental record interpretable. Source correction prevents a parent tradition from being assigned a role it does not support. Fidelity checks prevent a nearby object or relation from receiving credit as the one under review. Matched comparison separates a clearer record from a changed decision. These controls make the boundary credible; they do not supply evidence of superiority.

Future evaluation needs a primary decision endpoint and a competent baseline. The endpoint should be a predefined change in a permitted claim, commitment ceiling, or another decision output that remains justified after review.

Cases should be selected where a condition could change the decision, and reviewers should have relevant expertise and comparable access to the parent traditions. Traceable material allocation, predefined criteria, and retained reasoning records can distinguish a framework effect from unequal information or outcome selection after the fact[57].

The comparison must also count PODA's burden: time, training, coordination, compliance work, and superficial form completion. Adding administrative weight without changing the decision would count against mandatory use. A predefined endpoint that changes under stated conditions while the burden remains acceptable could support a narrower, context-bound efficacy claim.

Negative effects belong in the same evaluation. Overstating a limitation, demanding an irrelevant alternative, or requiring a full case file where a shorter artifact already preserves the needed relation would impose a burden of its own. PODA remains subject to the proportionality standard it applies elsewhere.

A direct test of the proposed operating mechanism would hold the case materials and evidence constant while varying whether a person actively redirects the AI exchange. The present calibration did not make

that comparison; it used a self-run retrospective baseline and was not independently replicated. Repeating the same desk exercise would not isolate the effect of sustained human arbitration.

One further limit belongs to the framework's internal safeguards. PODA proposes an engaged party as protection against structural tepidness, and the five-part closure test checks whether a verdict draws an operative line. Neither can certify that the underlying pressure was genuine. A model can produce a formally complete closure after superficial opposition, and a completed case file cannot reveal that origin by form alone.

That limit defines the next empirical question without weakening the current one. PODA can specify where pressure should fall, what reconstruction must preserve, and how the record must close. Whether its use changes decisions, and whether a model actually sustains the required pressure, remain external evaluation claims. The current record warrants PODA as an optional architecture for making claim limits, reconstruction, and commitment visible in one maintained case.

9. Conclusion

A construction can acquire authority before anyone has established that it is true, correct, or ready. Authority begins when someone is asked to rely on it, act on it, or accept a burden because of it. The decisive question is therefore not whether the construction looks complete, but whether the relation carrying that commitment has survived the pressure appropriate to the action being requested.

PODA makes that relation the unit of review. Fidelity fixes the object and claim under examination. Material admissibility tests whether the alleged difference reaches an independent stake. Evidence, comparison, burden, and conditions determine what pressure the claim must survive. Damage then leads to reconstruction rather than a vague caution: the claim narrows, the test changes, the arrangement is redesigned, or the case pauses or exits.

What survives becomes a warranted commitment, which draws a line between the action the record permits, the claim it withholds, and the step that would be irresponsible now. Institutional authorization remains separate because evidence can justify an action without granting an actor the standing to take it, just as authority to act cannot supply missing evidence.

The case file carries that line forward. It prevents a local finding from becoming permanent authority after the object, population, alternative, burden, or requested scale has changed. The record remains open to revision without losing the reasons that made an earlier decision proportionate at the time.

AI did not create the gap between evidence and commitment. It made the gap easier to cross without noticing. Rapid iteration can produce a polished construction before anyone has tested the relation on which its authority depends. In live work, the answer is sustained arbitration inside the exchange; in retrospective work, it is reconstruction of the relations that the finished artifact may conceal.

PODA's contribution lies in joining those two states through one inspectable sequence. It does not add criticism for its own sake, and it does not treat more passes, more models, or more documentation as substitutes for judgment. It places real engagement at the point where a claim becomes action.

A claim has not earned authority because it survived production, review, fluency, or consensus. It earns authority when the record shows what survived pressure, what remains withheld, and what commitment may responsibly follow, whether that record is still taking shape or already finished.

10. AI Use Statement

This article was written and revised through sustained collaboration with Clio (Claude, Anthropic), Raúl's personal agent, across multiple working sessions. ChatGPT and DeepSeek were also consulted. Their contributions included drafting and restructuring prose, researching alternative literature, and searching for and verifying references against primary sources.

The calibration results report the author's working practice, not AI-generated findings. The author reviewed every AI-assisted claim, citation, and revision and accepts responsibility for the final text. No AI system is listed as an author.

The cross-checks in Section 7 used models from more than one provider: ChatGPT for the first illustration, and ChatGPT, DeepSeek, and Qwen for the second. This was not a controlled evaluation panel and does not eliminate correlated failure, since models may share training-data overlap and broadly similar alignment approaches.

Using multiple providers reduces one narrower risk: convergence among systems built by the same laboratory for reasons unrelated to the evidence. Claude helped build both illustrations and did not cross-check either one. A system verifying its own construction would reproduce the self-bias risk identified in Section 1 rather than provide independent confirmation.

References

1. a, b *Toulmin SE (2003). The Uses of Argument. Updated ed. Cambridge: Cambridge University Press.*

2. ^a ^b *Kelly TP (1998). "Arguing Safety: A Systematic Approach to Managing Safety Cases." Safety-Critical Systems Club. https://scsc.uk/documents/acwg/tpk/Arguing%20Safety.pdf.*
3. ^ *Safety-Critical Systems Club, Assurance Case Working Group (2021). "Goal Structuring Notation Community Standard: Version 3." https://scsc.uk/r141C:1?t=1.*
4. ^a ^b *Bloomfield R, Rushby J (2024). "Assessing Confidence with Assurance 2.0." SRI International. https://arxiv.org/abs/2205.04522v4.*
5. ^ *Lewis D (1973). "Causation." J Philos. **70**(17):556–567. doi:10.2307/2025310.*
6. ^a ^b *Woodward J (2003). Making Things Happen: A Theory of Causal Explanation. Oxford: Oxford University Press.*
7. ^a ^b *Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ (2008). "GRADE: An Emerging Consensus on Rating Quality of Evidence and Strength of Recommendations." BMJ. **336**(7650):924–926. doi:10.1136/bmj.39489.470347.ad.*
8. ^a ^b *Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, Spitzer E, Raji ID, Gebru T (2019). "Model Cards for Model Reporting." Proceedings of the Conference on Fairness, Accountability, and Transparency. pp. 220–229. doi:10.1145/3287560.3287596.*
9. ^ *Gebru T, Morgenstern J, Vecchione B, Vaughan JW, Wallach H, Daumé III H, Crawford K (2021). "Datasheets for Datasets." Commun ACM. **64**(12):86–92. doi:10.1145/3458723.*
10. ^ *Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, Smith-Loud J, Theron D, Barnes P (2020). "Closing the AI Accountability Gap." Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. pp. 33–44. doi:10.1145/3351095.3372873.*
11. ^a ^b *Xu W, Zhu G, Zhao X, Pan L, Li L, Wang W (2024). "Pride and Prejudice: LLM Amplifies Self-Bias in Self-Refinement." Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 15474–15492. doi:10.18653/v1/2024.acl-long.826.*
12. ^a ^b *Malmqvist L (2024). "Sycophancy in Large Language Models: Causes and Mitigations." arXiv. https://arxiv.org/abs/2411.15287.*
13. ^ *Maram DP, Gandhi D, Yao Z, Akkinapalli G, Dernoncourt F, Wang Y, Rossi RA, Ahmed NK (2025). "Iterative Critique-Refine Framework for Enhancing LLM Personalization." arXiv. https://arxiv.org/abs/2510.24469.*
14. ^ *Mohamed A, Geng M, Vazirgiannis M, Shang G (2025). "LLM as a Broken Telephone: Iterative Generation Distorts Information." Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 7493–7509. doi:10.18653/v1/2025.acl-long.371.*

15. ^Röttger P, Kirk HR, Vidgen B, Attanasio G, Bianchi F, Hovy D (2024). "XSTest: A Test Suite for Identifying Exaggerated Safety Behaviours in Large Language Models." Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers). pp. 5377–5400. doi:10.18653/v1/2024.naacl-long.301.

16. ^Ray R, Bhalani R (2024). "Mitigating Exaggerated Safety in Large Language Models." arXiv. https://arxiv.org/abs/2405.05418.

17. ^Wynn A, Satija H, Hadfield G (2025). "Talk Isn't Always Cheap: Understanding Failure Modes in Multi-Agent Debate." arXiv. https://arxiv.org/abs/2509.05396.

18. ^Arkes HR, Blumer C (1985). "The Psychology of Sunk Cost." Org Behav Hum Decis Process. **35**(1):124–140. doi:10.1016/0749-5978(85)90049-4.

19. ^Bo JY, Kazemitabaar M, Deng M, Inzlicht M, Anderson A (2025). "Invisible Saboteurs: Sycophantic LLMs Mislead Novices in Problem-Solving Tasks." arXiv. https://arxiv.org/abs/2510.03667.

20. ^Chiodo M, Müller D, Siewert P, Wetherall JL, Yasmine Z, Burden J (2025). "Formalising Human-in-the-Loop: Computational Reductions, Failure Modes, and Legal-Moral Responsibility." arXiv. https://arxiv.org/abs/2505.10426.

21. ^Buçinca Z, Malaya MB, Gajos KZ (2021). "To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-Assisted Decision-Making." Proc ACM Hum-Comput Interact. **5**(CSCW1):1–21. doi:10.1145/3449287.

22. a, bMason RO (1969). "A Dialectical Approach to Strategic Planning." Management Science. **15**(8):B403–B414. doi:10.1287/mnsc.15.8.b403.

23. ^Dung PM (1995). "On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and N-Person Games." Artif Intell. **77**(2):321–357. doi:10.1016/0004-3702(94)00041-x.

24. ^Barak A (2012). Proportionality: Constitutional Rights and Their Limitations. Cambridge: Cambridge University Press.

25. ^Cooper RG (1990). "Stage-Gate Systems: A New Tool for Managing New Products." Bus Horizons. **33**(3):44–54. doi:10.1016/0007-6813(90)90040-i.

26. ^OECD (2020). "The Public Sector Innovation Lifecycle." OECD Working Papers on Public Governance. doi:10.1787/0d1bf7e7-en.

27. ^Jaradat O, Graydon P (2014). "An Approach to Maintaining Safety Case Evidence After a System Change." arXiv. https://arxiv.org/abs/1404.6846.

28. ^Copi IM, Cohen C, McMahon K (2016). Introduction to Logic. 14th ed. London: Routledge.

29. ^Herd P, Moynihan DP (2018). Administrative Burden: Policymaking by Other Means. New York: Russell Sage Foundation.

30. ^Jabareen Y (2009). "Building a Conceptual Framework: Philosophy, Definitions, and Procedure." Int J Qual Methods. **8**(4):49–62. doi:10.1177/160940690900800406.

31. ^Coral C, Bokelmann W (2017). "The Role of Analytical Frameworks for Systemic Research Design, Explained in the Analysis of Drivers and Dynamics of Historic Land-Use Changes." Systems. **5**(1):20. doi:10.3390/systems5010020.

32. ^Stilgoe J, Owen R, Macnaghten P (2013). "Developing a Framework for Responsible Innovation." Research Policy. **42**(9):1568–1580. doi:10.1016/j.respol.2013.05.008.

33. ^Jaakkola E (2020). "Designing Conceptual Articles: Four Approaches." AMS Rev. **10**(1-2):18–26. doi:10.1007/s13162-020-00161-0.

34. ^Star SL, Griesemer JR (1989). "Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." Soc Stud Sci. **19**(3):387–420. doi:10.1177/030631289019003001.

35. ^Yang R, Li Y, Li S (2026). "ARIS: Autonomous Research via Adversarial Multi-Agent Collaboration." arXiv. https://arxiv.org/abs/2605.03042.

36. ^Yang R (n.d.). "Anti-Autoresearch." https://github.com/wanshuiyin/Anti-Autoresearch.

37. ^Dranias M, Whitley A (2026). "Human-in-the-Loop Control of Objective Drift in LLM-Assisted Computer Science Education." arXiv. https://arxiv.org/abs/2604.00281.

38. ^Papamarkou T, Floridi L (2026). "Turing's Mirror: A Conceptual Design for Warrant-Sensitive Reliance on AI Output." SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6906401.

39. a, bThe We Company (2019). "Form S-1 Registration Statement Under the Securities Act of 1933." U.S. Securities and Exchange Commission. https://www.sec.gov/Archives/edgar/data/1533523/000119312519220499/d781982ds1.htm.

40. ^Singer D (2019). "WeWork Analyst Warns IPO Filing a 'Masterpiece of Obfuscation'." Bloomberg. https://www.bloomberg.com/news/articles/2019-08-20/wework-analyst-warns-ipo-filing-a-masterpiece-of-obfuscation.

41. ^Salmon F (2019). "Dissecting the Customer Acquisition Costs of the Latest Unicorn IPOs." Axios. https://www.axios.com/2019/09/06/wework-peloton-customer-profit-ipo-sales.

42. ^Napoletano E (2019). "WeWork Is the Most Ridiculous IPO of 2019." Forbes. https://www.forbes.com/sites/greatspeculations/2019/08/27/wework-is-the-most-ridiculous-ipo-of-2019/.

43. ^CNBC (2019). "WeWork Makes Sweeping Corporate Governance Changes Ahead of IPO." CNBC. https://www.cnbc.com/2019/09/13/wework-makes-sweeping-corporate-governance-changes-ahead-of-ipo.html.

44. ^CNN Business (2019). "Adam Neumann, WeWork CEO, Is Stepping Down." CNN Business. https://www.cnn.com/2019/09/24/tech/wework-ceo-adam-neumann-stepping-down.

45. ^CNBC (2019). "WeWork CEO Returns $5.9 Million the Company Paid Him for 'We' Trademark." CNBC. https://www.cnbc.com/2019/09/04/wework-ceo-returns-5point9-million-the-company-paid-for-we-trademark.html.

46. ^Peng S, Kalliamvakou E, Cihon P, Demirer M (2023). "The Impact of AI on Developer Productivity: Evidence from GitHub Copilot." arXiv. https://arxiv.org/abs/2302.06590.

47. ^Cui KZ, Demirer M, Jaffe S, Musolff L, Peng S, Salz T (2026). "The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers." Management Science. doi:10.1287/mnsc.2025.00535.

48. ^Murphy-Hill E, Butler J, Savelieva A (2026). "Adoption and Impact of Command-Line AI Coding Agents: A Study of Microsoft's Early 2026 Rollout of Claude Code and GitHub Copilot CLI." arXiv. https://arxiv.org/abs/2607.01418.

49. ^He H, Agarwal S, Denisov-Blanch Y, Azaletskiy P, Koyejo S, Vasilescu B (2026). "AI Writes Faster Than Humans Can Review: A Longitudinal Study of an Enterprise 2x Mandate." arXiv. https://arxiv.org/abs/2607.01904.

50. ^Becker J, Rush N, Barnes E, Rein D (2025). "Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity." arXiv. https://arxiv.org/abs/2507.09089.

51. ^Becker J, Rush N, Cunningham T, Rein D, Mahamud K (2026). "We Are Changing Our Developer Productivity Experiment Design." METR. https://metr.org/blog/2026-02-24-uplift-update/.

52. ^Xu F, Medappa PK, Tunc MM, Vroegindeweij M, Fransoo JC (2025). "AI-Assisted Programming May Decrease the Productivity of Experienced Developers by Increasing Maintenance Burden." arXiv. https://arxiv.org/abs/2510.10165.

53. ^Maier S, Gunzenhäuser M, Schweisthal J, Schneider M, Feuerriegel S (2026). "A Meta-Analysis of the Effect of Generative AI on Productivity and Learning in Programming." arXiv. https://arxiv.org/abs/2605.04779.

54. ^Amro A, Alalfi MH (2025). "GitHub's Copilot Code Review: Can AI Spot Security Flaws Before You Commit?" arXiv. https://arxiv.org/abs/2509.13650.

55. ^Luborsky L, Diguer L, Seligman DA, Rosenthal R, Krause ED, Johnson S, Halperin G, Bishop M, Berman JS, Schweizer E (1999). "The Researcher's Own Therapy Allegiances: A 'Wild Card' in Comparisons of Treatment Efficacy." Clin Psychol Sci Pract. **6**(1):95–106. doi:10.1093/clipsy.6.1.95.

56. ^Davis FD (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." MIS Q. **13**(3):319–340. doi:10.2307/249008.

57. ^Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D (2013). "SPIRIT 2013 Explanation and Elaboration: Guidance for Protocols of Clinical Trials." BMJ. **346**:e7586. doi:10.1136/bmj.e7586.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.