GOVERNED AUTONOMY

RCFS: Competence that can change without inheriting authority

A paper in the Governed Autonomy collection

Andrew Durden
Independent research
Publication edition 2026.10.9

Research on authority, evidence, and continuity in autonomous systems.

Content identifier: 7150dfae1470e29ab5d0fbcc

All rights reserved · Research exposition

Contents

RCFS: Competence that can change without inheriting authority 3

Abstract . . . 3

1. The persistence problem . . . 3

2. Competence and permission answer different questions . . . 4

3. Responsibility outlives the implementation . . . 4

4. Evidence has a kind and a scope . . . 5

5. Choosing methods without omitting responsibilities . . . 6

6. Replacing a method without overstating continuity . . . 6

7. Worked conceptual examples . . . 7

8. Distribution, recovery, and historical evidence . . . 8

9. What the research establishes—and what remains open . . . 9

Source guide . . . 12

All rights reserved · Research exposition

RCFS: Competence that can change without inheriting authority

Governed Autonomy · Publication 2026.10.9
Research study: conceptual exposition of the RCFS constitution
Source: Public exposition, successor 2026.10.9; historical research and evidence separately qualified

Abstract

A system that improves through experience faces a continuity problem. Its methods may change while its permissions, commitments, and responsibility for previous actions remain in force. Replacing a model does not cancel an obligation. Learning a more effective procedure does not authorize its use on a new resource. Compressing an expensive reasoning process into a faster implementation does not establish that every relevant behavior has been preserved. The Reflexive Causal Field System constitution, RCFS, treats these distinctions as requirements for persistent adaptive computation.

RCFS 1.0 separates qualified ability from current authority and applies that distinction to learning, method substitution, and recovery. This study develops its consequences for evidence, duties, historical records, and identity. The formal arguments are conditional on their semantic and implementation assumptions. Deployment conformance and practical benefit require further evaluation; the account does not imply unconditional progress, universal program equivalence, or protection from every harmful but authorized decision.

1. The persistence problem

An adaptive assistant may begin by solving a task through extensive reasoning. After repeated experience, a shorter procedure may suffice. Later, a specialized model, a database operation, or a human service may perform the same task more efficiently. The useful continuity is the ability to perform the task under its relevant constraints. Keeping the same model weights is neither necessary nor sufficient for that continuity.

The problem becomes consequential when the task affects other people or persistent resources. A replacement method may omit an exception, forget a pending commitment, rely on an obsolete permission, or misrepresent the history supporting its qualification. A system can preserve the appearance of an uninterrupted conversation while losing the substance of continuity. Conversely, it can change implementation substantially while retaining the obligations and evidence that matter.

RCFS locates this problem at the level of governed activity. A competence is more than a likely answer: it concerns a task, conditions under which a method is qualified, the kinds and bounds of consequences covered by that qualification, and the evidence needed to assess the result. Qualification describes supported ability within a stated scope. It does not grant present authority to cause those consequences. The source considers multiple

All rights reserved · Research exposition

computational and institutional forms. None receives authority merely because it performs the reasoning.

The source calls this position *representation non-privilege*. Its term *model–substrate singularity* refers specifically to qualified movement between implementation forms. The phrase *representation mobility* is used here to emphasize that operational meaning, without implying consciousness or a removal of computational limits. Public review scope.

2. Competence and permission answer different questions

Competence concerns what a system can accomplish under stated conditions. Authority concerns what it is permitted to do in the present circumstances. These can change independently. A system may receive permission for a task it cannot yet perform. It may also know how to perform a task for which it has no permission.

Learning can expand practical ability within an existing permission boundary. It can also reduce ability: a formerly reliable method may become obsolete, an important dependency may disappear, or a new model may regress on a protected task. RCFS therefore rejects an assumption of inevitable improvement. Qualification, loss of qualification, restoration, and renewed evaluation are all part of development.

The central conditional argument holds the governing rules, operative permission state, and relevant context fixed. Development alone should not change which effects are authorized. This is narrower than saying that the system’s behavior is unchanged. Better methods can make previously impractical authorized work feasible, and worse methods can make it unavailable. The separation is between permission and practical reach.

Permissions also carry context. A grant for one principal, purpose, resource, or period cannot be treated as an abstract score that can be combined freely with unrelated grants. Nor can the historical fact that an action was once authorized settle whether it is authorized now. Past reasoning may help locate relevant evidence, but current authority remains a present question.

This distinction constrains learning from successful work. Experience can support a procedure or an applicability judgment. It cannot turn a remembered permission into continuing authority. The broader treatment of this boundary appears in the collection’s architecture and shared terminology. Public review scope.

3. Responsibility outlives the implementation

An obligation is a duty whose status matters over time. It may concern delivering a result, preserving a record, reporting an incident, compensating for a failed action, or obtaining an external determination. Its continued existence does not depend on the model that first encountered it.

RCFS treats this persistence as a distinct problem from storing conversational memory. A summary saying that a payment is “being handled” is not sufficient to establish who remains responsible, what has happened externally, or what must happen next. Changing a method

All rights reserved · Research exposition

must preserve the connection between the duty, its current owner, the relevant evidence, and any unresolved consequences.

Resource commitments have a similar character. When an operation may already have consumed money, capacity, or another limited resource, uncertainty does not justify treating that resource as freely available again. A retry that ignores the earlier attempt can duplicate the effect or overspend the original allowance. Recovery therefore needs a view of responsibility that survives interruption.

This is also why compensation differs from erasure. Refunding an erroneous charge is another consequential action. It requires its own authority and evidence, and it does not make the original charge disappear from history. Some consequences cannot be reversed. A system can remain accountable for them even when it cannot restore the prior world.

The source's claims concern preservation and accounting under declared assumptions. They do not guarantee that every duty can be fulfilled on time, that all conflicting duties have a legitimate automatic resolution, or that an external party will cooperate. When external intervention is necessary, the remaining obligation should stay intelligible and owned. Public review scope.

4. Evidence has a kind and a scope

RCFS distinguishes an observation from an explanation, an explanation from a causal claim, and a causal claim from authority. A user's preference may settle a design choice without establishing an external fact. A successful test may establish behavior in its tested conditions without proving that the same behavior holds everywhere. An authorized decision may still be based on a false observation.

These distinctions matter when experience becomes reusable knowledge. Repetition is evidence of repetition. It does not, by itself, identify the cause of success. A method may work because of an unnoticed environmental condition, a helpful human intervention, or a favorable selection of cases. A causal explanation needs an appropriate account of the intervention, alternatives, and conditions under which the explanation travels.

Qualification is therefore multidimensional. A method can have strong evidence in a narrow environment and weak evidence elsewhere. It can be reproducible but stale, efficient but dependent on one provider, or accepted by an institution without being established as a universal fact. The source resists reducing these differences to one maturity score.

Evidence independence is another qualified claim. Preventing a candidate from editing its evaluator does not establish that the candidate could not influence the test data, the environment, the choice of evaluator, or which results were retained. A credible evaluation needs to state what kinds of influence have been excluded and which remain possible. "Independent" is useful only when its scope is clear.

Negative experience must also remain revisable. A failure in one setting can justify caution in that setting. It does not justify a permanent universal prohibition against a method that may later be repaired or used under different conditions. RCFS treats narrow correction, counterevidence, and re-evaluation as part of knowledge maintenance. Public review scope.

All rights reserved · Research exposition

5. Choosing methods without omitting responsibilities

A task may admit several reasonable approaches. One can be faster, another easier to verify, another less dependent on a particular model, and another better suited to recovery. RCFS uses the term *morphology* for a temporary arrangement of qualified processes used to pursue a task. That arrangement must include the responsibilities needed to complete the work, including verification and recovery as well as generation of the main answer.

For example, a useful arrangement for consequential work may need access to current state, permission evaluation, result verification, and a way to resolve interrupted execution. Removing these responsibilities can make a route appear cheaper while changing the task it actually completes. The source distinguishes indispensable task requirements from choices among otherwise acceptable methods.

There need not be one universally smallest or best arrangement. Cost, latency, evidence strength, resource use, privacy, and recoverability may conflict. A method that wins on one measure can lose on another. The quality of a selection claim depends on the stated alternatives, conditions, and comparison. A plausible selection is not automatically a proven optimum.

This also changes the meaning of efficiency. A method’s total cost includes the work needed to establish and maintain confidence in it, handle changing dependencies, and recover from failures. Omitting these costs can make a specialized route look economical when it merely transfers work to another part of the system or to an operator.

The conceptual contribution is to place optimization within a complete account of the task's responsibilities. A better optimizer or more capable model must still operate under the required authority and verification conditions. Public review scope.

6. Replacing a method without overstating continuity

Repeated work can sometimes be performed by a more compact or specialized implementation. RCFS calls this *consolidation*. The relevant question is what the replacement has been shown to preserve.

A check that a procedure satisfies a requirement in a given invocation differs from a proof that two procedures behave equivalently in every relevant context. Statistical evidence differs again. Each can be valuable, but each supports a different claim. Describing all of them as “equivalence” obscures the conditions under which replacement is justified.

Continuity includes more than successful outputs. A replacement can produce the same ordinary answers while handling refusal, exceptions, pending duties, or recovery differently. These differences may matter more than average performance. A qualification statement should cover the protected behaviors relevant to its intended use, and identify what remains outside the comparison.

Replacement also requires a response to changed assumptions. If a specialized route becomes unreliable, the system should be able to stop depending on it, preserve ongoing responsibilities, and return to an appropriate qualified alternative. The source calls this

All rights reserved · Research exposition

deconsolidation. It is a continuing qualification concern rather than proof that every replacement will always be reversible.

A similar distinction applies to model substitution. Isolating a new model from unauthorized effects can preserve a safety boundary even if the model is poor at the task. Safety under substitution is not a guarantee of competence under substitution. Conversely, impressive performance does not establish that the new model has inherited the system's identity or its previous experience.

RCFS treats historical attribution carefully. A replacement model may receive a record of earlier work. That does not mean it participated in those earlier events. Persistent organizational history and the experience of a particular component are different claims. Public review scope.

7. Worked conceptual examples

Four hypothetical examples show how adaptation could preserve continuing responsibility. Their outcomes are illustrative.

A successful change of method

A service repeatedly prepares reconciliation reports for one business unit. Initially, a general model interprets the reporting request and proposes a plan. Over time, a specialized procedure is evaluated for a stable subset of reports. Its qualification covers the relevant input conditions, required checks, and expected result.

The service can use the specialized procedure when those conditions hold. The improvement is a change in practical competence and cost. The existing permission to read the unit's records and produce its report still governs the task. The new procedure does not acquire permission to access other business units. Outstanding requests and responsibilities remain attached to the service's work even though the implementation changes.

This example illustrates a possible benefit, not a measured one. Establishing that the replacement is cheaper or more reliable would require evidence that includes evaluation, maintenance, exception handling, and recovery costs.

A denied expansion of scope

The same service discovers that it could improve a report by reading another unit's confidential records. Its method is competent to perform that analysis, and earlier reports from the authorized unit were successful. Neither fact supplies the missing permission.

The service can continue with an authorized report if that still satisfies the request. If the additional information is essential, the missing external authorization becomes explicit. The denial concerns the proposed access; it does not erase the user's objective or imply that every other route is unavailable. The key distinction is that learning how to obtain a better result does not settle whether a particular means is permitted.

All rights reserved · Research exposition

Recovery after an uncertain external effect

A service submits an authorized payment and loses contact before receiving a clear outcome. On restart, a different model is available. The replacement model can help interpret the situation, but it cannot infer that the payment failed merely because the earlier conversation ended.

The service needs evidence about the external result. Until the uncertainty is resolved, the payment remains part of its responsibility and resource accounting. If evidence establishes non-application, an authorized retry may be appropriate. If the payment occurred, the service should continue from that result. If a refund is needed, it is a separate authorized action. The desired outcome is recovery and completion where possible, with explicit external intervention only when the remaining decision or information requires it.

A replacement that passes the wrong comparison

A newly trained method matches an older method on routine cases. The evaluation omits cases involving an unresolved customer obligation. The routine agreement does not establish continuity for the omitted responsibility.

The new method may still qualify for a narrower scope. It should not be described as a complete replacement on the strength of the routine comparison. Additional evidence must address the protected behavior before that broader claim is justified. If the omission is discovered after deployment, recovering responsibility for affected work matters alongside correcting the method. This is an evidence limit, not a reason to regard adaptation itself as unacceptable.

8. Distribution, recovery, and historical evidence

A persistent system may distribute work across services, devices, and institutions. Different kinds of state have different coordination needs. Conflicting hypotheses can coexist while evidence is gathered. Two parties cannot safely assume independent ownership of the same indivisible responsibility merely because their local records have not yet converged.

RCFS therefore does not derive a universal consistency policy from the desire for decentralization. The appropriate guarantee depends on what the state means and what an incorrect decision could cause. A network partition can leave some authorized analysis possible while preventing a consequential action that requires current shared authority or resource information.

Historical reconstruction is similarly qualified. Some captured computations can be reproduced exactly. A model response, a human decision, or an external service interaction may support a different form of reconstruction. Replaying a record of a payment is not the same as sending the payment again. A historical account must identify what was reproduced, what was observed, and what remains unavailable.

These distinctions protect the meaning of evidence. A signed record can establish the integrity of a recorded claim without proving the external world matched that claim. A complete internal history can still omit an unrecognized external channel. The source keeps

All rights reserved · Research exposition

those limitations within its threat model rather than treating auditability as universal correctness. Public review scope.

9. What the research establishes—and what remains open

RCFS 1.0 supplies conditional structural arguments about governed composition, permission separation, resource and obligation preservation, model substitution, continuity, and qualified development. They depend on stated semantic and implementation assumptions. Establishing conformance requires a realization that satisfies those assumptions and passes the relevant evaluations; a deployed conformance result is not established by this paper.

9.1 A register of the arguments used here

The following register identifies the propositions on which this paper relies. The references are to RCFS 1.0, §28. They are source paper proofs and inspectable conditional reasoning, without an independently reproduced proof-assistant or deployed conformance result in this collection. The source's requirement that a checker or realization be sound is a premise to establish for that checker or realization.

Argument	Premises and domain	Conclusion and reasoning method	Dependent discussion and limit
Complete declared dependencies, Theorem 1	Finite dependency universe and starting set; deterministic, inflationary, monotone dependency rules over pinned facts and topology	Repeated closure reaches a unique least fixed point above the starting set; strict additions are bounded by the number of initially absent elements. Finite growth and induction establish minimality.	§5's indispensable responsibilities. It does not prove that the declared universe or rules include every real dependency or a usable implementation.
Governed composition, Theorem 2	Checked primitive contracts; only valid sequential, parallel, and guarded feedback composition; compatible or serialized shared writes; exact composite authority; conserved budgets and duties; complete recovery and evidence requirements	Every finite valid composition has a well-formed governing contract, and its committed effects require that contract's premises. Structural induction follows the permitted composition cases.	§§3 and 5. Relative contract preservation does not prove contract completeness, faithful implementation, or liveness under omitted faults.
Development cannot change current permission, Theorem 3	Authority rules and current permission state are isolated from learned and experimental state; development cannot write them; the authorization decision uses current rules, grants, and context without an arbitrary learned-state input	Holding rules, permission state, principal, task, and context fixed, development leaves the permitted effect set unchanged. Any executable effect must belong to that set, while practical reach can grow or shrink. This follows from the authorization function's arguments.	§§2 and 6. It establishes neither competence growth nor protection from harmful authorized decisions or an overlooked effect route.

All rights reserved · Research exposition

Argument	Premises and domain	Conclusion and reasoning method	Dependent discussion and limit
Attenuation and conservation, Theorems 4–6	Preserved authority lineage; valid attenuation and transfer; no duplicate spendable holding; exhaustive budget operations with durable, serializable accounting; uncertain external spending remains reserved	Same-root derivations cannot broaden their permitted scope or amplify consumable budget. Distinct roots can jointly permit more only under the applicable combination rule, retaining all origins. Induction and an equivalent sequential history preserve the accounting.	§§2–3 and 8. Separate roots are not subject to one root's budget cap; unmodeled external charges remain outside the result.
Continuing duties, Theorem 7	Exhaustive, type-checked linear or explicit affine duty transitions; durable lineage and history; recovery from the recorded transitions	A live duty cannot silently disappear or be duplicated in a well-typed trace. Induction accounts for each terminal or successor state.	§§3, 6, and 8. This covers the defined linear and affine core. Proposed split/merge extensions lack their additional construction and proof; recording a duty does not guarantee its fulfillment.
Proposal isolation and evaluation independence, Theorems 8–9	Proposals lack a direct effect route; every modeled effect passes a sound current authorization boundary. For evaluation, all in-scope candidate influence paths are represented, unknown paths are treated conservatively, and decisive evidence meets the required independence class	A proposal alone cannot cause a committed effect. A candidate cannot qualify itself solely through a validating outcome it could manufacture within the declared influence model. Non-reachability and contradiction arguments establish the respective conclusions.	§4 and model safety in §6. Missing channels, collusion, or compromised enforcement can invalidate the premises. This is not absolute or statistical independence.
Model and method substitution, Theorems 10–11	Proposal isolation and effect enforcement persist. Source, target, applicability domain, protected observations, and qualification relation are fixed; concrete execution refines the checked model or enforces the declared guard; dependencies remain current	Effect safety can survive a schema-conforming model replacement. Competence and continuity survive only to the extent of the declared relation: one-way trace inclusion, two-way matching, finite abstraction, invocation checking with fallback, or a statistical bound.	§6. Shape compatibility can coexist with useless output. A finite or statistical result is not universal equivalence.
Replay and identity continuity, Theorems 12–13	Replay captures all relevant inputs and versions for deterministic transitions and uses external evidence without reissuing effects. Migration has authorized mappings preserving the declared protected history, roots, duties, and effects	Deterministic replay reproduces the recorded abstract sequence; valid forward continuity mappings compose. Induction on the recorded or matched transitions establishes each result.	§§6 and 8. Uncaptured randomness is excluded; directional continuity is not symmetry, metaphysical identity, or universal backward compatibility.

All rights reserved · Research exposition

Argument	Premises and domain	Conclusion and reasoning method	Dependent discussion and limit
Generation without privilege growth, Theorem 14	Generation and qualification change only candidate or qualified methods; those methods cannot mint grants or act outside the governed effect path; authority and duty changes use their separate authorized operations	A finite sequence of generation and qualification may expand useful methods and practical reach without itself creating authority roots, deleting duties, or amending the governing rules. Induction preserves the separation.	§§1–2 and 6. An external grant or authorized amendment can change authority; qualified methods can still produce harmful authorized behavior.
Finite performance frontier, Theorem 15	Nonempty finite candidate set; a declared preorder for cost comparisons; deterministic comparisons from pinned measurements or bounds	At least one candidate has no strictly better predecessor under that order. Strict descent terminates in a finite set.	§5's choice among feasible arrangements. Neither uniqueness nor efficient discovery follows; an infinite generator is outside this existence result.
Distinct semantic categories, Theorem 16	Disjoint typed categories with category-specific premises; no unchecked conversion, reflection, foreign-function, or deserialization route; persistence preserves the distinctions	A preference, confidence, repetition count, or historical record alone cannot inhabit an incompatible protected category in a well-typed trace. Canonical forms and preservation under transitions establish the result.	§§2–4. A fact, permission, duty, and competence claim retain different premises. Type separation does not prove the observations, roots, or specifications correct.
Retained hybrid-control profile, Theorem 17	Closed safe sets, existing flows under declared regularity, enforced safe flow directions for all modeled disturbances, safe transitions between modes, safe handling of accumulation, and separate discrete authority and duty checks	Executions starting within the declared safe sets remain within them over the stated horizon. Flow invariance and induction over discrete transitions establish the result.	The source's retained control profile. It does not establish authority, fulfillment of duties, ethical adequacy, or safety under an unmodeled disturbance; its continuous guarantee cannot replace the discrete checks.

9.2 An inspectable permission-separation derivation

Fix a principal, task, and context, together with the governing rules and current permission state. Let A(e) be the authorization decision for a proposed effect e under those fixed inputs. Assume learning can change available methods but cannot change any of A's inputs or create another effect route. The set of permitted effects is then {e : A(e) permits e}. Since the function and its arguments remain the same before and after learning, membership in that set remains the same for every e.

Now take any effect a new method can actually execute. By the separate execution premise, successful execution requires A(e) to permit it; hence the new executable set is contained in the unchanged permitted set. Learning may make an authorized report newly feasible, or a regression may make it infeasible. Neither change grants access to another unit's records. A later grant or revocation changes the fixed permission input and requires a fresh evaluation, so it is outside this comparison. The derivation shows permission separation under stated

All rights reserved · Research exposition

ownership and effect-path premises; testing those premises in a concrete system is a distinct obligation.

The source makes an especially important distinction between an abstract rule and its realization. A formal system can omit an unauthorized operation while a concrete implementation still exposes one through credentials, shared state, an external service, or an overlooked channel. Moving from the rule to the implementation requires evidence that the actual system enforces the boundary.

Nor does valid authorization settle the quality or legitimacy of a decision. A principal can authorize something harmful. A verifier can embody an inadequate requirement. An institution can supply incorrect information or make an unwise rule. Technical accounting can help establish who acted and under what authority; it cannot independently settle every moral, legal, or political question.

Several research questions follow. How can qualification remain valid when adaptive routing changes the cases a method encounters? How can an evaluation expose hidden dependence on a model or provider? How can useful experience be retained while respecting privacy, deletion, and purpose restrictions? What evidence is sufficient when duties last longer than a model, deployment, vendor, or organization? How can an implementation demonstrate that its consequential channels are sufficiently covered?

The empirical claims also remain open. Representation mobility may reduce dependence on a provider, consolidation may improve reliability, and explicit governance may improve recoverability. Those are hypotheses requiring baselines, task distributions, resource bounds, and observable failure criteria. The presence of an architecture or an extensive test specification does not supply their measurements.

RCFS places these questions within the wider governed autonomy thesis: adaptation should be assessed together with continuing authority, evidence, and responsibility. Its contribution is a vocabulary and a conditional theoretical account for examining that conjunction. The evidence guide explains the collection’s separation of source arguments, illustrative reasoning, and verified results. Public review scope.

Source guide

This successor guide points to public arguments and their limits. The new private evidence linkage remains excluded. Previously public conceptual source guides, including any retained titles, document ordinals, sections and chapter maps, are not claimed to have been removed; see the correction scope. The historical edition remains available unchanged.

Foundations and lineage. Review the relevant argument and assumptions in this paper, the evidence guide, and the research review route. Specification, reported results and production observations have distinct scope.

Authority and persistent responsibility. Review the relevant argument and assumptions in this paper, the evidence guide, and the research review route. Specification, reported results and production observations have distinct scope.

All rights reserved · Research exposition

Knowledge and evaluation. Review the relevant argument and assumptions in this paper, the evidence guide, and the research review route. Specification, reported results and production observations have distinct scope.

Task arrangements and cost. Review the relevant argument and assumptions in this paper, the evidence guide, and the research review route. Specification, reported results and production observations have distinct scope.

Changing implementations. Review the relevant argument and assumptions in this paper, the evidence guide, and the research review route. Specification, reported results and production observations have distinct scope.

Assumptions and research boundaries. Review the relevant argument and assumptions in this paper, the evidence guide, and the research review route. Specification, reported results and production observations have distinct scope.

No source-guide entry establishes a full implementation, a complete historical migration proof, or an independent audit.

All rights reserved · Research exposition