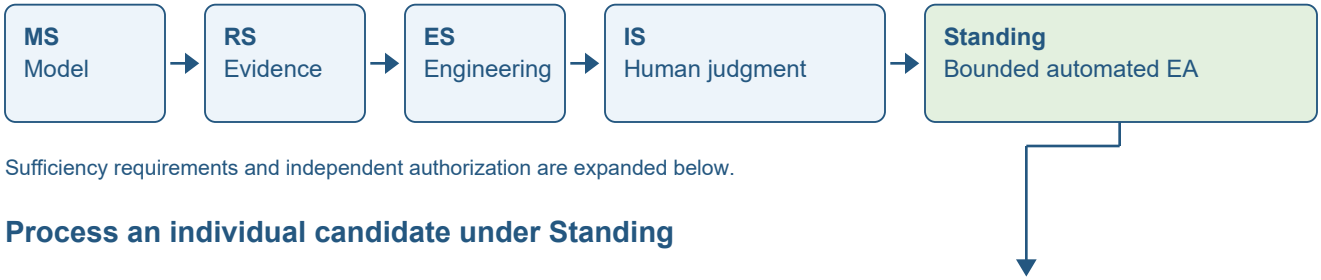
Governance of Autonomous Execution

DJE | Working reader edition | October 28, 2026

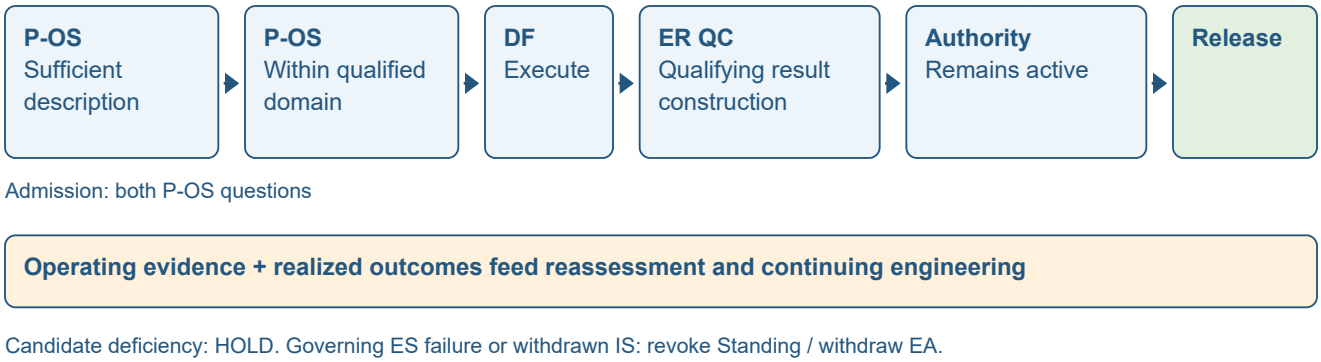
The governance objective is to establish when autonomous execution is justified, determine whether that justification continues to hold as conditions change, and withdraw authority when it no longer does. The architecture distinguishes three connected responsibilities: establishing Standing, governing individual decisions under that Standing, and improving the Decision Node through continuing engineering. Keeping these responsibilities distinct allows a problem to receive the response its scope warrants.

Establish and maintain Standing and autonomous Execution Authority



Sufficiency requirements and independent authorization are expanded below.

Process an individual candidate under Standing



Admission: both P-OS questions

Operating evidence + realized outcomes feed reassessment and continuing engineering

Candidate deficiency: HOLD. Governing ES failure or withdrawn IS: revoke Standing / withdraw EA.

Establishing Standing creates bounded authority; governing individual predictions applies that authority to particular candidates. The two P-OS questions form candidate admission. ER validity and result construction are distinct from the final check that authority remains active. These are related governance responsibilities, not a fresh MS-through-IS assessment for every candidate.

The Real Property Valuation Decision Node (RPVDN) provides a concrete setting for exploring this architecture. The same configured Decision Function serves two purposes. The LOOCV procedure uses the configured DF to produce held-out **sale-price estimates**. QC, calculation, and aggregation turn those outputs into metrics characterizing the evidence state for assessment against prospectively selected Engineering Sufficiency standards. Operational execution produces a **sale-price prediction** for an individual candidate using its eligible evidence state.

These outputs have distinct roles: a candidate's prior prediction is not carried into the LOOCV metrics of a later state. If the property subsequently enters that state, its held-out estimate is produced through fresh LOOCV executions. The laboratory examines empirical performance, execution integrity, and a change in state. The laboratory supplies empirical evidence for parts of the justification. The governance architecture identifies additional responsibilities and requirements that deployment would need to specify and assess; it is not a demonstrated end-to-end operating protocol.

Establish Standing

Establishing Standing involves Model Sufficiency, Representation Sufficiency, Engineering Sufficiency, and Interpretive Sufficiency, as shown in the diagram. Each addresses a different part of the justification for autonomous execution.

Model Sufficiency (MS) asks whether the bounded depiction of reality is adequate for the intended use. For RPVDN, this includes the selected property information, the eligible completed-sale population, and the rules by which that population changes. A Decision Function can operate consistently within a model that omits consequential aspects of reality. The adequacy of those boundaries therefore requires attention in its own right.

Representation Sufficiency (RS) asks whether the information presented to the Decision Function faithfully preserves and conveys the intended evidence identified by the Model. In RPVDN, the Decision Surface Snapshot (DSS) and structured quantitative table must represent the selected evidence without consequential errors, omissions, or misrepresentations. Faithful representation supports the engineering case; it does not establish that the model captures everything material to a valuation.

Engineering Sufficiency (ES) asks whether the affirmative engineering case supports the defined Decision Function within its qualified domain. Outcome-redacted Leave-One-Out Cross-Validation (LOOCV) supplies empirical evidence for this judgment in RPVDN. The observed errors, ensemble variability, execution controls, and limitations must be assessed against the governing requirements. A favorable aggregate performance metric alone cannot establish sufficiency for every contemplated use.

Before Standing is established, this engineering assessment includes whether the selected methods and controls support the intended claims. Human review may lead to further experiments, revised requirements, and renewed ES assessment before IS permission is granted. Before a candidate enters P-OS assessment, Standing must cover the prevailing evidence state and configured implementation. Under the proposed RPVDN procedure, a change in the eligible completed-sale population requires fresh LOOCV characterization and comparison with the governing state-sufficiency standards. If those requirements are satisfied, existing IS permission remains in force, and no other governing condition has failed, Standing can continue without a fresh IS determination. The steward may nevertheless withdraw IS permission at any time. Material changes or consequential observations can reopen the broader engineering assessment and authorization judgment.

Interpretive Sufficiency (IS) is the authorized human steward’s independent judgment that the evidence, limitations, and operating conditions justify the contemplated consequence. Affirmative ES does not compel permission. The steward may judge the remaining uncertainty unacceptable, even when the engineering requirements are satisfied. Conversely, IS cannot waive failure of a governing ES requirement. Authority to establish or prospectively revise standards is a separate responsibility from IS judgment under the governing standards. IS does not itself confer permission to rewrite the standards. Observed results can inform authorized revisions for future assessments; those revisions do not change whether earlier results met the standards then in force.

Affirmative ES and IS permit **Standing**: bounded, continuing, revocable authorization carrying Execution Authority. Its scope must be identifiable, including the configuration, qualified domain, applicable evidence state, and conditions under which automated execution may continue. Standing makes an authorized automated path available; it does not establish that every individual candidate may enter or complete that path.

Govern Individual Predictions

Candidate admission: P-OS. With Standing active, candidate admission begins with Prospective Operating Sufficiency (P-OS). This gate asks whether the particular property is sufficiently described and falls within the qualified operating envelope. These are separate questions. Missing required evidence may be repairable; a fully described property may still fall outside the domain for which automated valuation is justified. The applicable requirements and boundaries must be specified before evaluating the candidate.

A P-OS failure can place that candidate on HOLD while evidence is obtained, corrected, or reviewed, or while an appropriate alternative path is considered. It need not change Standing for other eligible candidates. Nor should an operating boundary be relaxed ad hoc merely because the Decision Function can produce an answer outside it.

Held-out LOOCV subjects also require evidence-completeness checks, and their execution records undergo QC before admission to the qualifying ensemble. Membership in the characterization population follows the Model's eligibility rules. These checks do not constitute the candidate-specific P-OS operating-envelope assessment or place held-out estimates on the prospective prediction-release path shown here.

DF execution, ER QC, and result construction. Admission is followed by DF execution and Experimental Record (ER) quality control. In the laboratory, only qualifying ERs entered the ten-run ensemble; excluded executions and their replacements remained part of the inspectable record. Result-handling controls also need to check extraction, qualifying ensemble membership, aggregation, and the records presented for release. Candidate admission therefore does not guarantee a releasable result. An invalid ER may require exclusion and replacement. Failure to assemble an acceptable ensemble under the governing operational rules may require HOLD for that candidate.

As with a P-OS failure, an execution or ensemble-construction failure can affect one candidate without withdrawing Standing for others. Either can also reveal a broader weakness, invoke a specified system-level trigger, or prompt ES reassessment or withdrawal of IS permission. The deployment procedure must define retry limits, candidate HOLD and resumption, and the circumstances requiring system review. The laboratory’s QC practice does not itself validate a complete operational protocol.

Before release: continuing authority and applicable requirements. Before releasing a completed prediction, confirm that Standing remains active, IS permission has not been withdrawn, and the candidate and its qualifying ensemble satisfy the requirements applicable at release. The check covers the applicable state-qualification, P-OS, execution-QC, and retry-budget requirements. Authority to change standards is separate from IS permission. The governing procedure must specify when an authorized standards revision takes effect and whether it applies to predictions already in progress. If revised requirements apply to an unreleased prediction, conformity must be re-established before release; otherwise, that candidate remains on HOLD.

Candidate sufficiency and execution integrity also matter on a collaborative path, where an authorized human works with AI assistance to assess the candidate and determine the appropriate action. Withholding automated authority does not itself authorize that alternative or remove the need to examine evidence and control execution. This RPVDN instantiation develops the automated Decision Node path; a collaborative route would require its own defined responsibilities and permissions.

Reassess Continuing Authority

Operating evidence feeds back into the justification for continuing authority. Candidate admission patterns, execution failures, and realized outcomes may expose weaknesses at different levels. Some observations can invoke prospectively specified controls; others require interpretive judgment. The same observation may warrant both responses.

Reassessing Standing. State qualification applies to the identified evidence state and tested implementation. Under the proposed RPVDN procedure, a changed eligible completed-sale population requires fresh LOOCV characterization and assessment against governing standards before another candidate enters the autonomous path. Characterization can be reused while the relevant state and tested implementation remain unchanged.

The laboratory characterized State 0 and one transition to State 1, not every operational state. Under Standard B, State 1 would fail where State 0 passed. Applied as an operating requirement, reassessment could therefore withhold authority before the next prediction error. This does not establish that LOOCV anticipates every deterioration in performance.

Realized sale price Y supplies a distinct, later feedback path by allowing a frozen prediction to be compared with its outcome. An illustrative rule withdrawing automated Execution Authority after two successive operational absolute percentage errors exceeding 5% shows how a specified trigger could act deterministically. Neither the threshold nor the sequence length is deployment-validated. A consequential observation may also warrant IS reconsideration before a numerical trigger is reached.

Candidate disposition and system authority must therefore remain distinguishable without becoming disconnected. A candidate-specific deficiency can require HOLD, correction, replacement, or escalation without revoking Standing for other eligible candidates. A recurring pattern or failure of a governing system requirement may instead require revocation of Standing and withdrawal of automated Execution Authority. Automated execution can resume only when the applicable qualification and authorization requirements are again satisfied. A candidate's acceptable result does not override a system-level loss of authority, and a candidate's failure does not by itself establish such a loss.

Continue Decision Node Engineering

Accumulated evidence can reveal shortcomings in the model of reality, evidence representation, configured Decision Function, experimental methods, QC procedures, or governing standards. QC findings contribute to that broader assessment alongside other operating evidence. Excluded ERs, unsuccessful candidates, and unexpected outcomes can contribute to this work alongside qualifying results.

Continuing engineering investigates those observations: Does the evidence omit a consequential property characteristic? Does its representation encourage systematic misreading? Is the qualified domain too broad? Do empirical tests and admission/release controls detect consequential changes and failures? Review methods must also address shared blind spots between production and QC. Findings can motivate revised configurations, QC procedures, experiments, and governing standards.

Material changes require renewed qualification and authorization before automated execution under the revised configuration is permitted. Improved observed performance is evidence to assess; it is not self-authorizing permission. Preserving the lineage from evidence and specification through execution, QC, and disposition makes that reassessment open to inspection.

QC supports establishing Standing, governing individual predictions, and continuing engineering. The DF Engineering, LOOCV, and QC companions explain how defects in evidence preparation, execution, and result handling undermine both the case for authority and individual results.

Scope and Limitations

The laboratory examined empirical performance, execution integrity, State 0, and one transition to State 1. It did not demonstrate the full automated governance progression shown here or characterize every state encountered in the prediction experiments. The map distinguishes proposed operating responsibilities from that demonstrated coverage.

The controls, retry and release rules, escalation responsibilities, and monitoring triggers require deployment-specific engineering and assessment. The illustrative two-successive-errors rule and Standards A and B are not deployment-validated thresholds. Human-led experimental checks and collaborative workbook reconciliation are not established autonomous QC capabilities.

The sequence for establishing Standing separates the contributing sufficiency judgments; it is not a claim that evidence or interpretation always develops in a single linear pass. A candidate HOLD and revocation of system Standing are distinct responses. Neither candidate success nor affirmative human permission can override failure of governing engineering requirements.

This architecture is a starting point for further engineering and research. Its contribution is to make the responsibilities and their relationships explicit: Standing supports bounded authority, individual gates govern its use, and experience informs renewed justification and continuing Decision Node development.