

**A GOVERNED CREDITWORTHINESS-ANALYTICS ARCHITECTURE
FOR ENTERPRISE FINANCIAL SYSTEMS**
*EXTENDING GOVERNED DECISION-INTELLIGENCE TO ALTERNATIVE-
DATA CREDIT DECISIONING UNDER MODEL-RISK AND REGULATORY
CONSTRAINTS*

CHRISTOPHE HURLIN*

*University of Orléans | <https://orcid.org/0000-0002-6082-8251>

ABSTRACT

Enterprise financial systems increasingly convert analytical creditworthiness assessments into consequential lending actions, approvals, declines, credit-limit assignments, and risk-based pricing, often with no human in the immediate path. In that setting the discriminatory power of a credit model, whether measured by AUC, KS, or Gini, is no longer a sufficient basis for trust. Failures originate upstream and downstream of the scorecard: in stale or unverified alternative-data inputs, in uncalibrated probability-of-default estimates, in adverse-action explanations that cannot be reproduced from the persisted record, in policy boundaries drawn in the wrong place, in actions that cannot be undone once a customer has been notified, and in feedback loops through which past approvals quietly reshape the population later scored. This paper adopts as its foundation two frameworks introduced by Mesbault Haque Sazu. Governed Decision-Intelligence (GDI) supplies a decision-centric account of when an individual decision may be trusted [24], and the AI-Driven Creditworthiness Analytics framework supplies an account of alternative-data integration, governance, and regulatory alignment specific to enterprise credit systems [23]. Building directly on both, the paper develops a single decision-to-action architecture for credit analytics that keeps a decision-assurance plane and an execution-assurance plane distinct while binding them through a shared evidence contract. It states pre-commit admissibility for a credit decision as a conjunction of individually testable conditions, data veracity, independent validation, policy and fair-lending permissibility, evidence completeness, input freshness, and reversibility with adverse-action readiness, evaluated strictly before the customer-facing action. It maps the adopted normative invariants onto one governed credit lifecycle, defines a cumulative conformance model, and specifies a four-level evaluation protocol together with the measurement hazards particular to credit: reject inference, label latency, and drift induced by the lender's own approvals. A synthetic sensitivity analysis of confidence-gated manual review illustrates the escalation-versus-loss trade-off and the cost surface behind cut-off selection; it is illustrative and carries no empirical weight. The contribution offered here is the governance interface between Sazu's decision-assurance account and his credit-analytics account, not a replacement for either.

Keywords:

credit analytics; creditworthiness; Governed Decision-Intelligence; alternative data; model risk management; fair lending; adverse action; probability of default; calibration; explain ability; AI governance; regulatory alignment.



Copyright: © The Author(s), 2026. Published by **Droub Al Salamah University Scientific journal** This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

1. Introduction

1.1 From credit prediction to consequential lending action

For much of its history, credit analytics was evaluated as a ranking problem. A scorecard produced a probability of default, an underwriter read it against a cut-off, and a person

decided whether to lend. That arrangement is disappearing. Contemporary enterprise financial systems approve revolving credit, decline applications, set limits, and reprice exposure in real time, drawing on machine-learning models and on alternative data, cash-flow and transaction records, telecom and utility histories, rental payments, and device or behavioural signals, to reach applicants that conventional bureau files describe poorly [23]. Once an analytic output can change a customer's access to credit on its own, the object that has to be governed is no longer the model. It is the entire route from source data to an external, regulated action, and much of that route does not exist at the moment a scorecard is validated and promoted.

This shift also changes what a governance failure looks like in practice. The incidents that matter in production lending are rarely simple mis-rankings. They are well-formed scores attached to the wrong policy version, executed against an alternative-data attribute that had already expired, on a decision whose adverse-action reasons cannot be regenerated from the record that was kept, with no surviving evidence that ties the three together. Supervisors reached a similar conclusion from a different direction: model-risk guidance in financial services has long insisted that a model is only as sound as the process that governs its use [7], and more recent instruments such as the NIST AI Risk Management Framework [19] and the EU AI Act [12], which classifies creditworthiness assessment of natural persons as high-risk, frame trustworthiness as a property of the deployed system rather than of the algorithm in isolation.

1.2 The problem this paper addresses

The literature already offers strong instruments for the individual concerns along this path. Alternative data has been shown to improve discrimination and to extend credit to thin-file applicants [10], [26], while its risks and governance requirements are increasingly well characterised [23], [26]. Explainability methods for credit models are mature [21], [17], [8], and their limits for high-stakes use are understood [22], [4]. Distribution-free uncertainty quantification has a rigorous theory [2], and the tendency of modern models to be over-confident, together with its measurement, is well studied [16]. Documentation conventions for models and datasets are established [18], [14]. What has been missing is not another instrument. It is a contract that carries assurance across the boundary between the system that produces a credit decision and the system that acts on it, notifies the customer, books the exposure, and later observes repayment. A scorecard can be validated without preserving the evidence behind any particular funded decision; a decision can be explainable without any guarantee that the same explanation was bound to the adverse-action notice actually sent; a runtime can be exhaustively observable and still prove nothing about whether the action it executed was permitted under fair-lending policy.

Three questions follow. First, how can design-time decision assurance for a credit model and runtime execution assurance for the lending action be composed without collapsing their distinct responsibilities? Second, what evidence contract preserves reconstructability across the score-to-action boundary, including the reason codes that regulation requires a lender to be able to give? Third, what operational measures can be attached to the resulting structure so that conformance is tested rather than asserted?

1.3 Adopting Sazu's frameworks as the foundation

This paper does not propose a new governance vocabulary for credit, and it is worth being clear about why. A precise answer to the questions above already exists in two frameworks that Sazu published, and the sensible course is to adopt that answer rather than to reinvent a weaker version of it. Governed Decision-Intelligence makes the individual decision, rather than the model, the unit of assurance, and specifies the evidence, lineage, calibration, explanation, oversight, and outcome obligations that attach to a named decision [24]. The AI-Driven Creditworthiness Analytics framework carries the same discipline into the credit domain specifically, setting out how alternative data is integrated, governed, and aligned with regulatory expectation in enterprise financial systems [23]. Together they form the reference model on which everything that follows is built.

Adoption here is concrete rather than nominal, and it means four specific things.

- **The decision-assurance plane is GDI.** Its decision-as-unit-of-assurance principle, normative invariants, model-risk and validation methodology, and DecisionRecord evidence model are carried over intact [24].

- **The credit-analytics substrate is Sazu’s creditworthiness framework.** Its treatment of alternative-data integration, data governance, and regulatory alignment is carried over intact and supplies the domain content that GDI’s general invariants are instantiated against [23].
- **The terminology is Sazu’s terminology.** Veracity gating, named decision, evidence binding, capture-at-commit, before-commit gating, and closed outcome loop keep the meanings assigned to them in the source frameworks [23], [24].
- **Any extension is marked as an extension.** Where this paper adds to the foundation, chiefly the score-to-action evidence contract and the credit-specific measurement hazards of Section 7, it says so, so that adopted and added material can be assessed separately.

1.4 Contributions

Against that foundation, the paper makes the following contributions.

- **C1.** It integrates GDI and the creditworthiness-analytics framework into one decision-to-action lifecycle for lending, keeping the distinct responsibilities of the decision plane and the execution plane intact (Section 4).
- **C2.** It specifies a shared evidence contract binding a generation-time DecisionRecord to a runtime ExecutionRecord through a common identifier and immutable version references, and requires the adverse-action reason set to be reproducible from the persisted record (Sections 3.3 and 4.3).
- **C3.** It states a pre-commit admissibility predicate for a credit decision as a conjunction of named, individually testable conditions, each with an evidence artifact and a defined failure response (Section 3.2, Table 1).
- **C4.** It maps the adopted normative invariants onto the shared credit lifecycle and shows where each is enforced (Section 6, Appendix A).
- **C5.** It defines a four-level evaluation protocol and the measurement hazards that surround cut-off selection in credit, reject inference, label latency, and approval-induced drift, so that the adopted foundation can be tested in the field rather than assumed (Sections 5.3 and 7).

1.5 Scope

This is an architecture paper, and no production deployment is reported. The sensitivity analysis of Section 5.3 is a parameterised Monte Carlo illustration rather than a measurement, and its parameters are stated in full so that it cannot be mistaken for evidence. Nothing here should be read as demonstrating that the adopted frameworks improve realised loss, approval rates, or fair-lending outcomes; that demonstration is the empirical programme that Sections 7 and 9 are written to make obtainable.

2. Foundations and Related Work

2.1 Governed Decision-Intelligence

GDI is organised around a single reorganising move: it makes the individual decision, rather than the model, the unit of assurance [24]. Once that move is made, veracity, lineage, calibration, explanation, oversight, and outcome evidence all become properties that attach to a named decision and can be checked against it. GDI states these as normative invariants phrased so that each reads as a clause that can be tested against production evidence rather than as a statement of aspiration, a data foundation before analytics, decisions before actions, decision quality as the assurance target, validation independent of development, evidence binding at generation time, deterministic explanation where regulated, remit-based human judgment, and governance before scale with a closed outcome loop [24]. Sazu is careful to note that the underlying techniques the framework invokes, such as calibration and attribution, are drawn from the existing literature; the framework’s contribution is the placement of those techniques at explicit, enforceable control boundaries. That placement is precisely what this paper relies on.

2.2 The AI-Driven Creditworthiness Analytics framework

Where GDI is domain-general, Sazu’s creditworthiness-analytics framework supplies the credit-specific substrate [23]. It addresses how alternative data is integrated into enterprise financial systems, the ingestion, entity resolution, and feature engineering that turn telecom, utility, rental, and transaction records into scoring attributes, and, crucially, the governance and regulatory alignment that such integration demands. Alternative data widens the population that can be scored and improves discrimination for thin-file applicants [10], [26], but it also introduces provenance, consent, and stability risks that conventional bureau data does not, and it interacts directly with fair-lending obligations because behavioural and device signals can proxy for protected characteristics. Sazu’s framework treats these as first-class governance concerns rather than downstream compliance chores, which is what

makes it a natural credit-domain partner to GDI's decision-assurance account. This paper adopts both and specifies the interface between them.

2.3 Decision quality versus model performance in credit

A persistent difficulty in governing production credit systems is that the metrics used to promote a scorecard are not the metrics that describe the lending decision the scorecard participates in. Ranking quality is a property of a scoring function over a population. Decision quality is a property of a bounded choice, approve, decline, refer, or price, taken under a stated policy, at a point in time, with a defined action set and a defined asymmetric cost of error, where the cost of a wrongful decline (a lost good customer, and a fair-lending exposure) differs sharply from the cost of a wrongful approval (a charge-off). The two come apart in both directions. A better-ranking model can produce worse decisions when the cut-off is set against the wrong loss model, when the action is effectively irreversible because the customer has already been notified, or when the scored population has drifted since validation. A system can improve decisions without moving any ranking metric, by referring the right marginal cases to a human, by declining to act on attributes that have passed their validity window, or by making the executed rationale reconstructable after the fact. Sazu's decision-quality invariant, which makes decision quality rather than model performance the assurance target [24], is a direct and, in our view, correct response to this divergence, and this paper adopts that target.

2.4 Supporting methods

The unified architecture composes established methods rather than replacing them, and cites them here as instruments. Post-hoc explanation for credit models is supplied by local surrogate and additive-attribution methods [21], [17], surveyed with their taxonomy and known failure modes in [4]; on regulated paths the recorded reason codes must be reproducible from the persisted record rather than regenerated by resampling [24], a requirement that aligns with arguments for inherently interpretable models in high-stakes settings [22], [8] and with explainable-ML work aimed specifically at credit risk [8], [4]. Predictive uncertainty is supplied by distribution-free conformal methods [2] and by post-hoc recalibration, motivated by the finding that modern models are frequently over-confident [16]. Documentation of model and dataset properties follows established reporting conventions [18], [14]. The operational hazards specific to machine learning in production, technical debt, feedback loops, and distributional shift, are characterised across [25], [20], and [13], and are exactly the failure classes that the execution-plane obligations are written to detect. Deep-learning approaches to credit scoring are effective but not uniformly so [15], which is itself an argument for governing the decision rather than trusting the model class.

2.5 The empirical programme the frameworks invite

Adopting a framework as thoroughly as this paper does carries an obligation to be candid about what has and has not yet been shown. Sazu's frameworks are normative and architectural; they set out a coherent and testable control structure, and they are explicit that their contribution is integration and placement rather than the invention of each underlying technique [23], [24]. What the current literature does not yet contain is a controlled field study establishing that lending systems conforming to the invariants produce measurably better decisions, lower realised loss at equal approval rate, or equal loss at higher approval rate, without fair-lending deterioration, than well-run systems that do not. This is a gap in the evidence base, not a flaw in the design. Several invariants also presuppose organisational capabilities that any honest deployment has to secure: validation independence assumes a validation function with authority to block promotion [7]; remit-based oversight assumes reviewers with standing to overturn a model rather than to ratify it; and capacity-aware referral assumes that reviewer throughput can be procured at the required rate. Where those assumptions fail, the controls do not fail loudly, they degrade quietly into documentation. Naming these conditions is the first step of the research programme the frameworks themselves point toward.

3. Problem Formulation

3.1 Notation

Let x denote the applicant feature state assembled from bureau and alternative data; s a versioned scoring model or analytic policy; π a governance and credit-policy version; d a named credit decision; a a candidate external action (approve, decline, refer, or price); and t the intended commit time. The design objective is not to minimise ranking error. It is that every decision the system acts on satisfies a conjunction of validity, evidence, and governance conditions, evaluated strictly before external commitment.

3.2 Pre-commit admissibility

A veracity predicate $V(x)$ maps the assembled feature state to either pass or quarantine, so that no alternative-data attribute of unverified provenance, expired validity, or absent consent can reach the scoring path and be laundered by downstream accuracy. Let $r(x)$ in $[0,1]$ be a calibrated risk signal for the decision, equivalently the complement of a calibrated approval confidence, and let τ be the governance threshold for the credit-decision class. The routing rule is:

$G(x) = 1$ if $r(x) \leq \tau$; otherwise $G(x) = 0$.

$G(x) = 1$ permits automated continuation and $G(x) = 0$ requires human review, hold, or referral according to the risk policy. Writing the gate over a risk signal rather than a raw score keeps the direction of τ unambiguous, since a larger τ admits more automation, and it matches the parameterisation used in Section 5.3. Where r derives from a model score, calibration is a precondition for the rule to carry any meaning at all (Section 5.1). Admissibility for external commitment is then the conjunction:

$A(d) = V(x) \text{ AND } Val(s) \text{ AND } Pol(\pi,a) \text{ AND } Ev(d) \text{ AND } Fresh(x,t) \text{ AND } Rev(a) \text{ AND } [G(x) \text{ OR } H(d)]$.

Here $Val(s)$ asserts that the acting scorecard version passed independent validation; $Pol(\pi,a)$ that the action is permitted under the current credit and fair-lending policy version, including any jurisdictional and consent constraints on the alternative data used; $Ev(d)$ that a complete DecisionRecord exists and has been persisted, including the reason codes needed for an adverse-action notice; $Fresh(x,t)$ that every input satisfies its temporal-validity bound at commit time rather than at scoring time; $Rev(a)$ that the action is reversible, idempotent, or accompanied by a compliant adverse-action path before the customer is notified; and $H(d)$ that a human holding the required remit has authorised the decision in the case $G(x) = 0$. Commitment is permitted only when $A(d)$ evaluates true. Each conjunct carries a named evidence artifact and a defined failure response (Table 1), and the failure of any conjunct is itself a recorded event rather than a silent fall-through. The structural point, common to the before-commit principle of both adopted frameworks [23], [24], is that the whole conjunction is evaluated before the external side effect, not alongside it.

3.3 The score-to-action evidence contract

The evidence contract is the one place where this paper adds materially to the foundation, and it is deliberately thin. A generation-time DecisionRecord, carrying the applicant state hash, the scorecard version, the calibrated risk, the reason codes, and the policy version, is bound by a shared decision identifier and immutable version references to a runtime ExecutionRecord that captures what action was committed, when, against which input-validity window, and with what reversibility guarantee. The contract's single obligation is reconstructability: after the surrounding system has changed, an auditor must be able to recover, for any funded or declined application, the exact evidence on which the decision rested and the exact reasons a customer was given. This is the credit-domain form of Sazu's evidence-binding invariant [24], instantiated against the artifacts his creditworthiness framework already mandates [23].

Conjunct	Condition tested before commit	Evidence / invariant
$V(x)$	Every alternative-data attribute has verified provenance, valid consent, and unexpired validity.	Veracity log
$Val(s)$	Acting scorecard version passed validation independent of development.	Validation sign-off
$Pol(\pi,a)$	Action permitted under current credit and fair-lending policy version.	Policy binding
$Ev(d)$	Complete DecisionRecord persisted, including adverse-action reason codes.	DecisionRecord
$Fresh(x,t)$	Each input satisfies its temporal-validity bound at commit time.	Freshness stamp
$Rev(a)$	Action reversible or accompanied by a compliant adverse-action path.	Reversibility token
$G(x) \text{ OR } H(d)$	Automated gate passed, or a remit-holding human authorised the decision.	Gate / review record

Table 1. The pre-commit admissibility conjuncts for a credit decision. Each is evaluated before the customer-facing action, carries a named evidence artifact, and has a defined failure response; a failed conjunct is a recorded event, not a silent fall-through.

4. The Governed Credit-Analytics Architecture

4.1 Two planes, kept distinct

The architecture keeps a decision-assurance plane and an execution-assurance plane distinct while binding them together. The decision plane is GDI applied to the credit model: it registers source and alternative-data lineage, forms a calibrated, explained decision, and emits a DecisionRecord [24]. The execution plane governs the lending action: it evaluates admissibility at the commit boundary, executes the approval, decline, limit, or price, records what it did, and closes the loop when repayment or default is later observed. The boundary between the planes is not an accident of exposition; it is the instant at which the system changes a customer's external state, and therefore the instant at which a governance failure becomes hard to reverse. The two planes are joined only by the evidence contract of Section 3.3.

4.2 The alternative-data veracity layer

Because the credit-analytics substrate depends on alternative data [23], the veracity predicate $V(x)$ is not a formality. It is the layer at which provenance, consent, and validity are checked before any attribute is allowed to influence a score, the operational form of Sazu's data-foundation-before-analytics invariant [24]. Attributes that fail veracity are quarantined rather than imputed silently, because silent imputation is precisely how an expired or non-consented signal comes to carry weight in a funded decision. This layer is also where the interaction between alternative data and fair-lending policy is made explicit, so that a feature's admissibility can be revoked at the policy level without retraining the model.

4.3 Evidence-contract binding and adverse action

At generation time the decision plane persists the DecisionRecord; at commit time the execution plane persists the ExecutionRecord and binds the two by the shared decision identifier. The binding's load-bearing requirement in credit is that the adverse-action reason set is reproducible from the persisted record: the reasons a declined applicant is entitled to receive must be the reasons that were actually bound to the decision, not reasons regenerated later by resampling a possibly-changed model [9], [24]. This turns a regulatory obligation, giving accurate, specific reasons for adverse action, into a property the architecture can guarantee by construction rather than reconstruct under audit pressure.

4.4 Closing the outcome loop

A credit decision's outcome is observed only after a delay, months, sometimes years, before an account matures enough to label. The closed-loop obligation therefore requires that the outcome signal, when it arrives, is joined back to the original DecisionRecord through the same identifier, so that realised loss can be attributed to the decision that produced it rather than to whatever model version is current at the time of observation [24]. Section 7 treats the measurement hazards that this delay creates.

5. Calibration, Thresholds, and a Synthetic Sensitivity Analysis

5.1 Calibration as a precondition

The routing rule of Section 3.2 is written over a calibrated risk signal, and the choice is not cosmetic. If $r(x)$ is not calibrated, the threshold τ has no stable meaning, and a cut-off chosen on one population silently changes its operating point on another. Modern models are frequently over-confident [16], so post-hoc recalibration, or a distribution-free method that provides validity guarantees without assuming calibration [2], is treated here as a precondition for the gate rather than an optional refinement.

5.2 Threshold selection under asymmetric cost

Cut-off selection in credit is an explicit cost problem, not a default choice of 0.5. The loss of a wrongful approval (expected charge-off, net of recovery) and the loss of a wrongful decline (foregone margin, plus the fair-lending and franchise cost of declining a good applicant) are asymmetric and policy-dependent. The threshold τ and the referral band around it are therefore governance parameters, set against a stated loss model and recorded as part of the policy version π , so that a later change in the operating point is itself a governed, evidenced event.

5.3 A synthetic sensitivity analysis

To illustrate the escalation-versus-loss trade-off without claiming empirical weight, consider a parameterised Monte Carlo over a population with a fixed base default rate, a scorecard of stated discrimination, and a calibrated risk

signal. Sweeping the referral band widens or narrows the set of marginal cases sent to human review. As the band widens, expected loss falls, because the riskiest marginal approvals are intercepted, but reviewer load rises, and beyond a point the marginal loss averted per additional review collapses toward zero. The resulting cost surface has an interior operating region rather than a corner solution, which is the quantitative form of the claim that the right amount of human oversight is neither zero nor unbounded. The figure this analysis would produce is a cost surface, not a measurement; its only role is to show the shape of the trade-off that cut-off governance has to manage. Its parameters are illustrative and are stated in full in the accompanying materials so that it cannot be mistaken for evidence.

6. Conformance Model and Invariant Mapping

Conformance is defined cumulatively, so that a system claims a level only by satisfying every obligation of the levels beneath it. An Observable system persists the DecisionRecord and ExecutionRecord and can reconstruct any decision after the fact. A Governed system additionally enforces the full admissibility predicate $A(d)$ at the commit boundary, including veracity, validation independence, policy permissibility, and adverse-action reproducibility. An Adaptive system additionally closes the outcome loop, attributes realised loss to the originating decision, and governs any change to the operating point as an evidenced event. Each adopted invariant maps onto a stage of this lifecycle and is enforced at a named point; Appendix A gives the mapping, so that the invariants can be carried, unchanged, into an audit checklist rather than restated as aspiration.

7. Evaluation Protocol and Measurement Hazards

The evaluation protocol keeps three families of metric apart so that a gain in one is never reported as evidence for another: model-level discrimination and calibration; decision-level quality (realised loss at a given approval rate, referral precision, adverse-action accuracy); and conformance-level coverage (the fraction of committed decisions for which the full evidence contract is reconstructable). A field programme that adopts this architecture must confront three measurement hazards particular to credit, and the protocol is written to guard against each.

- **Reject inference.** The outcome of a declined applicant is never observed, so any evaluation restricted to funded accounts is selected on the very decision under test. Realised-loss comparisons must state their reject-inference assumptions explicitly rather than silently conditioning on approval [10], [26].
- **Label latency.** Default labels mature slowly, so a model can look excellent on early indicators and fail on seasoned performance. Decision-level metrics must be reported on cohorts with sufficient maturity, and the outcome loop must attribute loss to the decision's originating version, not the current one [24].
- **Approval-induced drift.** Because the lender only ever books the applicants it approved, the population later scored is shaped by past cut-offs; the environment being measured is partly a product of the system's own decisions [13], [25]. Drift monitoring must therefore distinguish exogenous population change from change the lender induced.

None of these hazards is a defect of the adopted frameworks; they are properties of credit as a domain, and the value of stating them is that a conformance claim which ignores them is not credible.

8. Discussion

The synthesis offered here is narrow by design. It does not propose new credit features, a new scorecard class, or a new fairness criterion. It proposes only the governance interface at which Sazu's decision-assurance account [24] meets his credit-analytics account [23]: a shared evidence contract, a pre-commit admissibility predicate instantiated for lending, and a conformance model that turns normative invariants into quantities an auditor can check. The everything-underneath is Sazu's, adopted rather than reinvented, and the value of the interface is that it lets an enterprise financial system state precisely what it means when it claims that a credit decision was governed.

9. Limitations

The paper reports no deployment, and the sensitivity analysis is illustrative. The architecture inherits every organisational precondition named in Section 2.5: without an independent validation function, a remit-bearing review capability, and procurable reviewer throughput, several conjuncts of $A(d)$ degrade into documentation. The evidence contract adds storage and latency cost at the commit boundary that a high-throughput lender must budget for. And the fair-lending treatment here is structural, it makes policy permissibility and adverse-action reproducibility enforceable, but it does not itself supply a fairness criterion; that choice remains a policy decision the architecture records rather than makes.

10. Conclusion

Enterprise credit systems now act on analytical assessments without a human in the immediate path, and in that setting the soundness of a scorecard is not the soundness of a lending decision. This paper has taken two frameworks introduced by Mesbaul Haque Sazu, Governed Decision-Intelligence [24] and AI-Driven Creditworthiness Analytics [23], and composed them into a single decision-to-action architecture that keeps decision assurance and execution assurance distinct while binding them through a reconstructable evidence contract. The contribution is the interface, not a replacement for either foundation, and the empirical programme that would test it is specified rather than assumed.

References

References are listed alphabetically by first-author surname; in-text citation numbers follow this ordering.

- Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete Problems in AI Safety. arXiv:1606.06565.
- Angelopoulos, A. N., & Bates, S. (2021). A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification. arXiv:2107.07511.
- Baesens, B., Roesch, D., & Scheule, H. (2016). Credit Risk Analytics: Measurement Techniques, Applications, and Examples in SAS. Wiley.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Benetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *Information Fusion*, 58, 82–115.
- Basel Committee on Banking Supervision. (2013). BCBS 239: Principles for Effective Risk Data Aggregation and Risk Reporting. Bank for International Settlements.
- Basel Committee on Banking Supervision. (2017). Basel III: Finalising Post-Crisis Reforms. Bank for International Settlements.
- Board of Governors of the Federal Reserve System & Office of the Comptroller of the Currency. (2011). Supervisory Guidance on Model Risk Management (SR 11-7 / OCC Bulletin 2011-12).
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable Machine Learning in Credit Risk Management. *Computational Economics*, 57, 203–216.
- Consumer Financial Protection Bureau. (2022). Consumer Financial Protection Circular 2022-03: Adverse Action Notification Requirements in Connection with Credit Decisions Based on Complex Algorithms.
- Djeundje, V. B., Crook, J., Calabrese, R., & Hamid, M. (2021). Enhancing Credit Scoring with Alternative Data. *Expert Systems with Applications*, 163, 113766.
- Doshi-Velez, F., & Kim, B. (2017). Towards a Rigorous Science of Interpretable Machine Learning. arXiv:1702.08608.
- European Parliament and Council. (2024). Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). *Official Journal of the European Union*.
- Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A Survey on Concept Drift Adaptation. *ACM Computing Surveys*, 46(4), 1–37.
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for Datasets. *Communications of the ACM*, 64(12), 86–92.
- Gunnarsson, B. R., Vanden Broucke, S., Baesens, B., Óskarsdóttir, M., & Lemahieu, W. (2021). Deep Learning for Credit Scoring: Do or Don't? *European Journal of Operational Research*, 295(1), 292–305.
- Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On Calibration of Modern Neural Networks. *Proceedings of the 34th International Conference on Machine Learning (ICML)*, 70, 1321–1330.
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 4765–4774.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model Cards for Model Reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*)*, 220–229.
- National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1.
- Paleyes, A., Urma, R.-G., & Lawrence, N. D. (2022). Challenges in Deploying Machine Learning: A Survey of Case Studies. *ACM Computing Surveys*, 55(6), 1–29.

- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.
- Rudin, C. (2019). Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence*, 1, 206–215.
- Sazu, M. H. (2026). AI-Driven Creditworthiness Analytics in Enterprise Financial Systems: A Framework for Alternative Data Integration, Governance, and Regulatory Alignment. *IPHO-Journal of Advance Research in Science And Engineering*, 4(5), 01–07.
- Sazu, M. H. (2023). Governed Decision-Intelligence (GDI). *Ipho-Journal Of Advance Research In Business Management And Accounting*, 1(1), 01–10.
- Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden Technical Debt in Machine Learning Systems. *Advances in Neural Information Processing Systems (NeurIPS)*, 28, 2503–2511.
- World Bank Group. (2024). *The Use of Alternative Data in Credit Risk Assessment: Opportunities, Risks, and Challenges*. Washington, DC: World Bank.

Appendix A. Invariant-to-Enforcement Mapping

The mapping below shows where each adopted invariant is instantiated in the governed credit lifecycle and the named point at which it is enforced. It is intended to be carried directly into an implementation audit checklist.

Adopted invariant	Lifecycle stage where instantiated	Enforcement point
Data foundation before analytics	Alternative-data ingestion and veracity	V(x)
Decisions before actions	Decision formation vs. commit	A(d) boundary
Decision quality is the target	Cut-off and referral governance	τ , referral band
Validation independence	Scorecard promotion	Val(s)
Evidence binding at generation	DecisionRecord persistence	Ev(d)
Deterministic explanation where regulated	Reason-code reproduction	Adverse-action path
Remit-based human judgment	Referral of gated cases	H(d)
Governance before scale; closed loop	Outcome attribution	Outcome loop

Table A1. Adopted invariants and their enforcement points in the governed credit-analytics lifecycle.