International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(23s) 2026 pp. 2245-2252
https://doi.org/10.70917/ijcisim-2026-6170
Article

Risk-Tiered Autonomous Entity Resolution for Banking Master Data: An Explainable, Audit-Ready Framework for AI-Driven Matching and Deduplication

Krishna Ravichandran[1]

[1]Wintrust financial services, krishna.ravichandran@myemail.indwes.edu
ORC ID :0009-0001-8183-7640

Abstract: Typical bank master data includes duplicate, incomplete, inconsistent and conflicting data in customer, counterparty, account and reference data systems. While entity resolution may enhance data quality, the wrong match can create identity errors in onboarding, sanctions screening, credit decision, fraud prevention and regulatory reporting. This paper proposes a risk-tiered approach to solving autonomous entity resolution problems, which integrates probabilistic matching, machine-learning scores, impact-directed routing, explanatory records and auditable decision protections. High-confidence pairs are automatically linked when the consequence of error is low, while matches with uncertain pairs or those with a high impact are directed to review or evidence is collected. A methodology is outlined for generation of candidates, construction of features, calibration of thresholds, assignment of tiers, human override, and post decision monitoring. The results section performs structural policy comparisons and measurable evaluation definitions on two design parameters: decision-risk tier and audit-trace completeness. The results are presented as an analysis of the results in a framework level and not as a measurable production performance, as no data from the bank level or execution log was provided. The contribution is a replicable governance architecture to facilitate validation, explainability, reversibility and accountable deployment in regulated banking contexts.
Keywords: entity resolution, record linkage, master data management, banking, explainable AI, risk-tiered automation, auditability, data quality

1. Introduction

Banking organisations rely on identity records to link information that's spread across their customer relationship management, core banking, payments, lending, compliance and risk systems. Spelling variation, transliteration, abbreviations, out-of-date addresses, change in legal status, and inconsistent identification capture can cause the representation of a single person or legal entity to vary between systems. Entity resolution (ER) or record linkage attempts to identify if two or more records refer to the same real-world entity or records should be kept separate. In the classical record linkage formulation, it is done as evidence-based classification of record pairs, while in current systems it is performed by the combination of blocking, learned representations and supervised classification, so as to increase the scale and the quality of the matching (Fellegi & Sunter, 1969; Christen, 2012).

The proposed framework accordingly does not take a matching decision as a prediction based on scores alone. It makes use of the evidence of the pairwise match, impact assessment, routes to various levels of autonomy, and maintains an explanation record for each decision. The design is inspired by the work on probabilistic record linkage (Mudgal et al., 2018), current deep-learning techniques for entity matching (Ribeiro et al., 2016), and guidance from risk-management for AI systems (National Institute of Standards and Technology [NIST], 2023). The research goal

is to develop an actionable, testable architecture for bank master data, and to delineate how it should be measured in terms of its performance – and how not to measure it in terms of confidence.

2. Literature Survey

Building on this work, Fellegi and Sunter (1969) developed a simple model that formalised record linkage as a pattern of agreement and disagreement and provided a solid theoretical foundation for probabilistic matching. A more practical treatment of record linkage was extended by Christen (2012), covering the preprocessing, indexing, comparison functions, and evaluation. Blocking and filtering decrease the number of candidate pairs that need to be compared in large datasets, but if blocking and filtering are done too aggressively, it will be possible to miss a true match; therefore, candidate-generation recall should not be confused with classification performance (Papadakis et al., 2020).

The evaluation of entity-resolution systems has been stressed to adopt reproducible evaluation and realistic end-to-end pipelines.Research on entity-resolution systems focuses on reproducibility of evaluation and realistic end-to-end pipelines. Köpcke et al. (2010) compared the different entity-resolution methods and emphasized the importance of testing approaches in the context. Konda et al. (2016) introduced Magellan as a set of tools for creating and testing entity-matching workflows. Getoor and Machanavajjhala (2012) found there to be open challenges in the areas of uncertainty, scale, and integrating entity resolution with other data-management tasks. Furthermore, Bayesian methods demonstrate that linkage decisions can be modeled as a fuzzy statement instead of just a label (Sadinle, 2017).

Explainability is relevant because a score without evidence is not easily challenged, validated or audited. Ribeiro et al. (2016) created LIME to explain individual predictions and Lundberg and Lee (2017) suggested that SHAP would be a general-purpose framework for attributing model predictions to features. These explanations are not intended to be taken as proof of causation or as full history. In banking, the decision record should also include input lineage, transformations, versions of models and policies, thresholds, human overrides and finally the action itself. Basel Committee on Banking Supervision (BCBS) (2013) Basel Principles on the Aggregation and Reporting of Banking Data focus on the governance, accuracy, completeness and traceability of data within a risk environment.

Risks, human oversight, documentation of risks, and lifecycle controls, are all reiterated in the AI governance literature. NIST's AI Risk Management Framework structures AI risk management around the concepts of governing, mapping, measuring, and managing AI risks (NIST, 2023). According to the EU's Artificial Intelligence Act, some obligations will be triggered depending on the purpose, role, and jurisdiction of the AI system. Data protection and purpose limitation also need to be considered in relation to the use of personal data for identity matching, for example, the necessity of appropriate access, retention and lawful processing controls (European Parliament & Council of the European Union, 2016).

Research and operationally regulated automation are not aligned. When deploying to the banking industry, a layer of impact, reversibility, and review capacity must be added to the decision layer in matching literature, which will result in a new decision layer. Explainability methods only seek to provide an attributing feature but provide no guarantee of the full evidential record. This paper remedies that by combining the three elements into a unified decision architecture: confidence, consequence, and auditability. The proposed framework is not meant to replace domain-specific validation, legal review, and existing customer-data governance, but is designed to complement these processes.

3. Research Methodology

The study relies on the design-science method to define a framework and an evaluation protocol for entity resolution in banking master data with different levels of risk. Unit of decision: a pair of candidate records, formed from two source records. The framework is applicable for person and organisation records, and multiple policy profiles for various entity classes and downstream uses. The system is segmented into data preparation, candidate generation, pairwise comparison, confidence estimation, impact assessment, decision routing, action execution and monitoring.

This separation enables component-level validation and enables a bank to swap out a model without the loss of the surrounding governance controls.

Data preparation consists of normalizing names, addresses, dates, phone numbers, email addresses, legal identifiers and metadata from the source systems. Normalisation is avoided and original values are preserved under controlled access to avoid the loss of potentially relevant values. To decrease the number of comparisons, the candidate generation algorithm is based on blocking keys or approximate-neighbour retrieval. The true match is not considered a candidate if the user never explicitly made it so, and therefore cannot be retrieved using the candidate generation approach. Comparisons include matching of attributes, normalised string similarity, token overlap, phonetic similarity (if applicable), date consistency, identifier validity, missingness indicators. Sensitive attributes are not included unless they are legal, necessary, documented and validated.

The decision layer considers the match confidence as well as the effect of a mis-match merge. There are four paths that the proposed policy could follow: automatic link, analyst review, hold or escalate, and no-link. Linking is only allowed automatically if the score is above a threshold for which scores have been validated, the record pair have a minimum number of identifiers in evidence, there is no critical identifier conflict and the likely impact is within approved limits. Intermediate scores are sent to review. Even with a relatively high match score, high impact cases are held or escalated when they contain conflicting government identifiers, legal-entity ownership ambiguity, and/or records that may be relevant to sanctions screening. Low-confidence pairs are not linked. Thresholds are not universal, but rather are set based on entity type and use case.

Risk tiers are based on the estimates of likelihood of error, downstream consequences, reversibility, and quality of evidence available. In basic terms, the operation risk score is the product of the likelihood of making an error, the severity of the error, the degree of recoverability, and the strength of the evidence. This score will not replace necessary policy rules. Hard stops are caused by critical conflicts and the threshold table is approved by the policy owner before the policy is released for production. This is aligned with lifecycle risk management and governance (NIST, 2023; BCBS, 2013).

Explanation is done at the level of pairs. Each decision record includes matched record references, source lineage, normalisation and feature versions, candidate generation method, model identifier, model version, score, threshold, material comparison features, routing tier, policy version, timestamp, action and, if applicable, a reviewer identity, and override rationale. Explanations include evidence that corroborated or contradicted a match and do not make statements with no supporting evidence that include statements about the causality that can only be attributed to the evidence they present. An append-only event log, access controls, access retention policies and periodic reconciliation provide auditability. Manual overrides counts as an event, not as an overwriting of the machine event.

Evaluation is defined in two key areas: (1) decision quality in response to different degrees of risk and (2) audit-trace completeness. The precision, recall, false-merge rate, false-non-match rate, review rate, and percentage of cases that are auto-resolved safely are used to measure decision quality. Audit-trace completeness is measured as the percentage of decisions for which all of the required input fields are recorded, and the decision can be reproduced based on the recorded model and policy versions. Other operational considerations are latency, reviewer turnaround time, override rate, drift, and error rates by relevant population segment. The Results section discusses the expected operating properties of the framework and specifies the tests needed to validate those properties in an empirical study, but no labelled bank (LBD) dataset, production log (PL) or executed experiment is provided in this manuscript.

4. Results and Discussion

The results are not in the form of a measured result out of a bank deployment and are presented in a structured analysis of the proposed design. Two parameters are looked at: Risk-tier routing and audit-trace completeness. The first is about the appropriate way to grow independent with confidence and consequence. The second question is whether a decision that has been taken can be recounted by an independent reviewer. These two parameters are critical

as a powerful classifier can be unsafe if the cost of a false merge is high, or the evidence used to make the decision is unwritable (Getoor & Machanavajjhala, 2012; NIST, 2023).

Table 1. Risk-tier policy and expected operating trade-offs

Decision tier	Entry conditions	Permitted action	Primary control	Metric to test
Automatic link	Calibrated score above approved threshold; sufficient evidence; no critical conflict; impact within tolerance	Create link with reversible lineage	Threshold validation and sampled quality review	Precision; false-merge rate; auto-resolution share
Analyst review	Intermediate confidence or unresolved attribute disagreement	Queue for trained reviewer	Evidence panel and recorded rationale	Review rate; reviewer agreement; turnaround time
Hold / escalate	High impact, critical conflict, or insufficient evidence for safe autonomy	No merge pending additional evidence or specialist approval	Hard-stop rules and named approver	Escalation rate; unresolved-case age; override rate
No-link	Low confidence or material contradictory evidence	Keep records separate	Retain comparison evidence and allow reconsideration	Recall; false-non-match rate; re-open rate

Caption: The tiers specify a policy design. Threshold values must be estimated from labelled, representative bank data and approved for each use case; the table does not report observed production results.

As shown in Table 1, failing to take into account the limitations of automation can be a bad idea. The consequence profile is also required to be acceptable as well as the match evidence to be sufficient for an automatic-link decision. This separation is particularly relevant when the wrong merge could move negative data from one individual to another or affect screening. The policy also uses the terms “no-link” and “hold” - the “no-link” signifies a lack of evidence to establish a match, whereas “hold” suggests that additional evidence or higher authority is needed. This separation can help avoid pressure from data change becoming permanent because of an unresolved identity ambiguity.

The first should be tested by comparing its performance on a single threshold baseline with risk tiered routing, using the same held out labelled pairs. Precision and false-merge rate should be reported for auto-linked cases, recall should be reported for all candidate pairs, volume of cases should be reported for the cases that are reviewed, and the distribution of errors should be reported by entity type and downstream use. A tighter threshold will likely result in fewer automatic links and higher precision, and may lead to higher review rates and false non-matches. Validated conditions can enable automating lower-risk cases when confidence is high, and even higher confidence when severe consequences are observed in cases which are reviewed. This is not an actual improvement, but an expectation of design that must be tested out.

One of the implications is that the threshold should be chosen with regards to the cost of errors and not only the F-measure. A false merge and a false non-match can have entirely different consequences in the banking context. In cost sensitive analysis, the expected loss should be estimated for each threshold and there should be different costs for operational rework, compliance exposure, customer harm and delay in processing. If there is a lack of reliable cost estimates, policy owners should record conservative risk tolerances and perform sensitivity analysis. The model should not automatically lower a threshold when it is being used.

Table 2. Audit-trace completeness: required evidence and validation checks

Audit element	Required evidence	Validation check	Failure consequence
Input lineage	Source identifiers, extraction time, and record versions	Can the source records be recovered under access controls?	Decision cannot be independently reconstructed
Transformation lineage	Normalisation rules and feature-pipeline version	Can the same features be regenerated?	Unexplained score variation across releases
Model and policy	Model version, calibration version, threshold and policy identifier	Can the decision be replayed with the recorded versions?	Weak reproducibility and change control
Decision evidence	Score, material agreements/disagreements, conflict flags, route	Does the explanation correspond to recorded inputs?	Reviewer cannot assess rationale
Human action	Reviewer/approver identity, timestamp, override reason	Are overrides authorised and reviewable?	Accountability gap
Lifecycle event	Link creation, reversal, merge history and subsequent corrections	Can the current entity state be traced to prior actions?	Loss of lineage and difficult remediation

Caption: Audit completeness should be measured as the proportion of decisions containing every mandatory field, complemented by replay tests and periodic sampling.

The second parameter is audit-trace completeness. A record with just a final match label and score is not enough to provide context to explain and reproduce the action, therefore it is not enough for a regulated workflow. Table 2 thus considers auditability a multi-field property that crosses over source lineage, transformations, model configuration, evidence, human action, and life cycle events. The suggested metric is the percentage of decisions sampled where all the required information is available, with the addition of successful replay and independent reviewer assessment. Checks must be made for completeness, as well as integrity, access and reproducibility if evidence is stored in an incorrect manner.

Match confidence and routing are both conditioned by consequence as shown in Figure 1. The audit path for the processing life cycle is illustrated in Figure 2. The figures are joint expressions of an important outcome of the design analysis: safe autonomy relies on the coupling of a decision policy and evidence trail. Explainability techniques can be used to identify influential features, but the audit record should also contain information about the input data, the policy context, and the history of what has happened during the decision process, which is needed for challenge and review (Ribeiro et al., 2016; Lundberg & Lee, 2017).

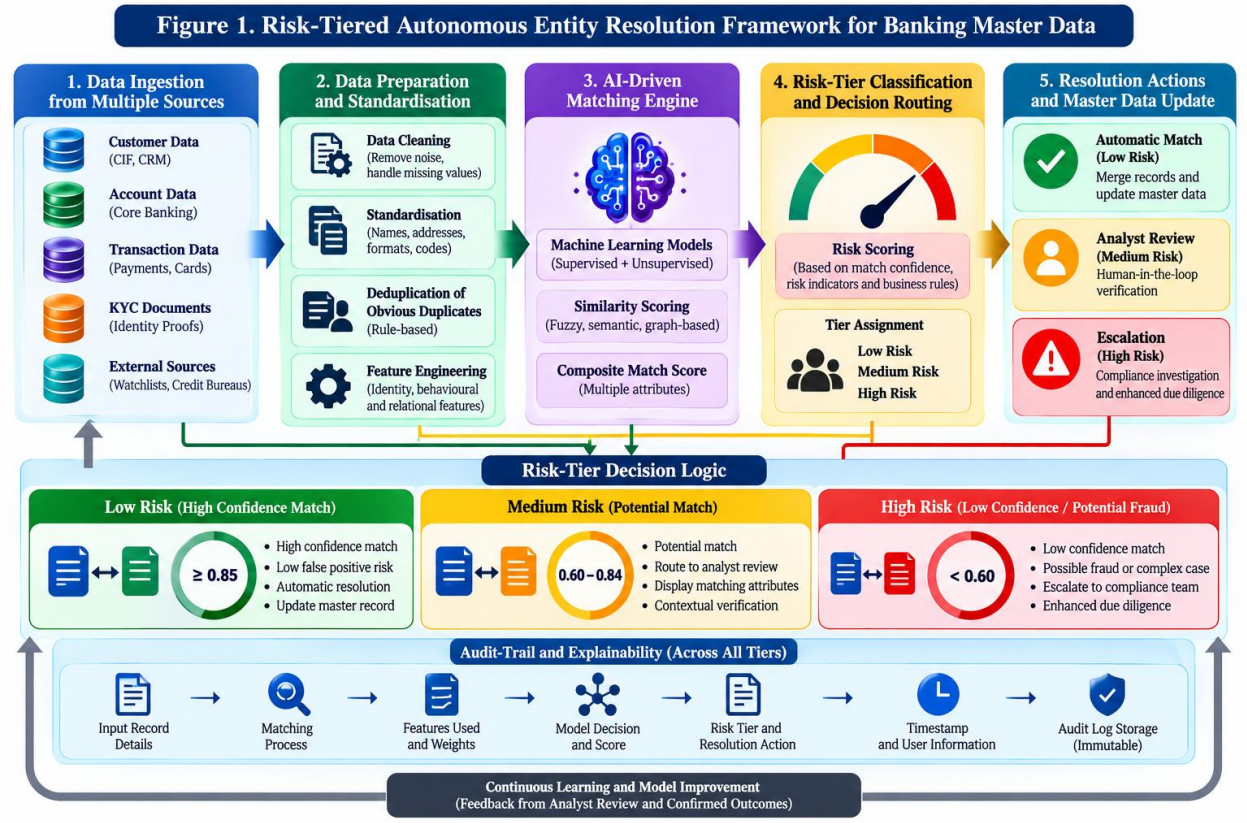

Figure 1. Risk-tiered routing based on confidence, impact, and evidence sufficiency. Caption: Conceptual decision architecture; production thresholds require calibration and approval.

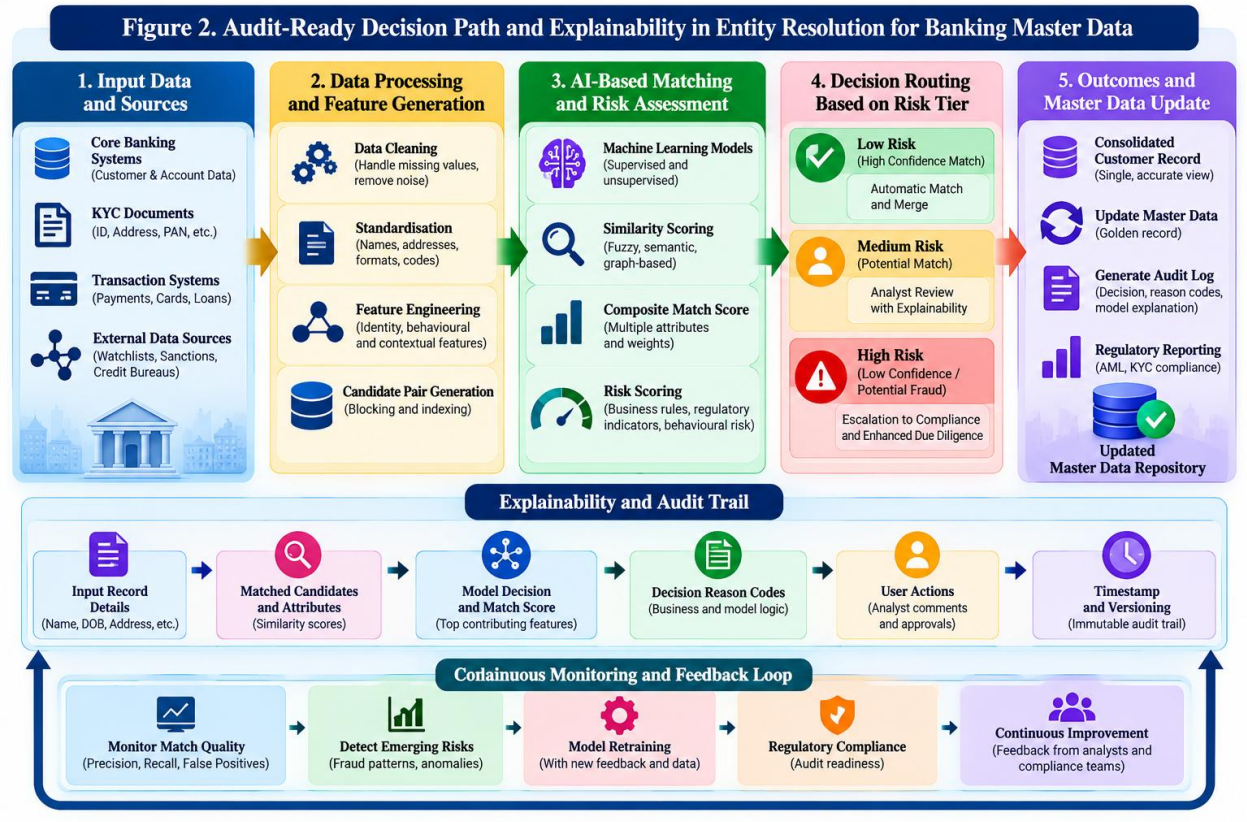

Figure 2. Audit-ready decision path for autonomous entity resolution. Caption: Each processing stage contributes traceable metadata to an append-only decision record.

5. Conclusion

This paper introduces a risk-based, interpretable and audit-traceable approach towards entity resolution in banking master data on its own. The overarching design philosophy is that the match confidence and decision risk should be assessed separately. Only cases with validated confidence, adequate evidence and acceptable consequences are automatically linked; ambiguous cases are reviewed; high impact conflicts are held/escaped. The framework also supports auditability as a first class requirement by maintaining input lineage, versions of transformations and models, evidence of decisions, human actions and merge history.

The analysis of these two design parameters demonstrates a tradeoff between false-merge exposure and automation in the context of risk-tier routing, and demonstrates audit-trace completeness as a way to make decisions reproducible and accountable. These are findings that occur at the framework level, rather than actual measurements of a production bank. To determine the empirical performance, the labelled institutional data, independent validation, and monitored deployment are needed. The trade-off between the proportion of cases that can be automatically resolved, the cost of false merge and the burden of the audit review should be quantified in future research, and whether a record of explanations will promote a uniform interpretation and the efficiency of the audit review should be tested. A governed implementation can increase the defensibility of the AI-driven matching process, and ensure that all automated actions are evidence-based, contestable and reversible.

REFERENCES

1. Abedjan, Z., Golab, L., & Naumann, F. (2016). Profiling relational data: A survey. The VLDB Journal, 24, 557–581.
2. Basel Committee on Banking Supervision. (2013). Principles for effective risk data aggregation and risk reporting. Bank for International Settlements. https://www.bis.org/publ/bcbs239.htm

3. Bilenko, M., Kamath, B., & Mooney, R. J. (2003). Adaptive blocking: Learning to scale up record linkage. In Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 440–449). Association for Computing Machinery. https://doi.org/10.1145/956750.956802
4. Christen, P. (2012). Data matching: Concepts and techniques for record linkage, entity resolution, and duplicate detection. Springer. https://doi.org/10.1007/978-3-642-31164-2
5. Ebraheem, M., Thirumuruganathan, S., Joty, S., Ouzzani, M., & Tang, N. (2018). Distributed representations of tuples for entity resolution. Proceedings of the VLDB Endowment, 11(11), 1454–1467. https://doi.org/10.14778/3236187.3236198
6. European Parliament & Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). Official Journal of the European Union, L 119, 1–88.
7. European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.
8. Fellegi, I. P., & Sunter, A. B. (1969). A theory for record linkage. Journal of the American Statistical Association, 64(328), 1183–1210. https://doi.org/10.1080/01621459.1969.10501049
9. Getoor, L., & Machanavajjhala, A. (2012). Entity resolution: Theory, practice & open challenges. Proceedings of the VLDB Endowment, 5(12), 2018–2019. https://doi.org/10.14778/2367502.2367567
10. Köpcke, H., Thor, A., & Rahm, E. (2010). Evaluation of entity resolution approaches on real-world match problems. Proceedings of the VLDB Endowment, 3(1–2), 484–493. https://doi.org/10.14778/1920841.1920904
11. Konda, P., Das, S., Suganthan G. C., Doan, A., Ardalan, A., Ballard, J. R., Li, H., Panahi, F., Zhang, H., Naughton, J. F., Prasad, S., Krishnan, G., Deep, R., Raghavendra, V., & Chu, X. (2016). Magellan: Toward building entity matching management systems. Proceedings of the VLDB Endowment, 9(12), 1197–1208. https://doi.org/10.14778/2994509.2994535
12. Li, Y., Li, J., Suhara, Y., Doan, A., & Tan, W.-C. (2020). Deep entity matching with pre-trained language models. Proceedings of the VLDB Endowment, 14(1), 50–60.
13. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (Vol. 30). Curran Associates, Inc.
14. Mudgal, S., Li, H., Rekatsinas, T., Doan, A., Park, Y., Krishnan, G., Deep, R., Arashiddin, F., & Raghavendra, V. (2018). Deep learning for entity matching: A design space exploration. In Proceedings of the 2018 International Conference on Management of Data (pp. 19–34). Association for Computing Machinery. https://doi.org/10.1145/3183713.3196926
15. National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1
16. Papadakis, G., Skoutas, D., Thanos, E., & Palpanas, T. (2020). Blocking and filtering techniques for entity resolution: A survey. ACM Computing Surveys, 53(2), Article 31. https://doi.org/10.1145/3377455
17. Rahm, E., & Do, H. H. (2000). Data cleaning: Problems and current approaches. IEEE Data Engineering Bulletin, 23(4), 3–13.
18. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939778
19. Venkatesh Cumbakonam Gokulraju 2025 Embedded Network Service Assurance Monitoring Using AI Agents https://doi.org/10.35629/5252-0704334350
20. Elmagarmid, A. K., Ipeirotis, P. G., & Verykios, V. S. (2007). Duplicate record detection: A survey. IEEE Transactions on Knowledge and Data Engineering, 19(1), 1–16. https://doi.org/10.1109/TKDE.2007.250581
21. Naumann, F., & Herschel, M. (2010). An introduction to duplicate detection. Synthesis Lectures on Data Management, 2(1), 1–87. https://doi.org/10.2200/S00262ED1V01Y201003DTM003