AML-VIEW: A REALISTIC INSTITUTION-CENTRIC SYNTHETIC DATASET AND GENERATOR FOR ANTI-MONEY LAUNDERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Anti-money-laundering (AML) investigation requires reasoning over a rich financial ecosystem using only one institution's evidence. Public data rarely expose this setting, while synthetic releases typically emphasize transactions or assume complete visibility. We present AML-view, a dataset and configurable generator that separates a shared, investigation-rich financial world from institution-local observations of it. AML-view integrates statistically grounded parties, ownership, accounts, due-diligence attributes, and payments, instantiates 38 documented laundering scenarios, and propagates illicit value through downstream transfers. Local views retain private records for customers but mainly transactional and public information for external counterparties. Each evaluated world contains about 10 million transactions over 70,000 accounts. AML-view covers all 13 audited investigation axes, versus 7 for the closest comparator, and passes all 35 population-calibration checks. By exposing matched institution-local views of the same financial world, AML-view reveals effects hidden by global-only evaluation: detection becomes harder for most models under realistic information boundaries, and the relative ranking of models can change. In blinded review, AML practitioners recovered laundering patterns from synthetic suspicious cases at a rate comparable to real cases, with recovery increasing from 0.500 using transactions alone to 0.765 with richer investigation context, while showing no consistent ability to distinguish synthetic from real cases. AML-view provides a large-scale, expert-validated, reproducible foundation for AML detection and investigation research.

1 INTRODUCTION

Anti-money-laundering (AML) investigation requires reasoning over both a complex financial world and an incomplete view of that world. Laundering activity can involve individuals, firms, ownership relationships, accounts, and transactions spanning multiple institutions, while any single institution sees only the customers, due-diligence records, payments, and counterparty information available within its own boundary. **Realistic AML evaluation therefore requires both a realistic financial world and a realistic observation process over that world**.

Real banking data rarely permit reproducible study of this setting because customer, ownership, transaction, and investigation records are highly sensitive. Synthetic data provide an alternative, and recent generators have improved individual aspects of realism, including transaction simulation (Lopez-Rojas et al., 2016; Suzumura & Kanezashi, 2018; Jensen et al., 2023), economic behavior (Altman et al., 2023), topology (Östman et al., 2025), customer profiles (Chen et al., 2026), and laundering-pattern diversity (Öztas et al., 2023; Huda et al., 2025; van den Beukel et al., 2026). However, existing releases do not jointly provide a calibrated, investigation-rich financial world together with institution-specific observations of that same world. As a result, models are often evaluated either on limited transaction-centric representations or on global graphs containing information unavailable to any single institution.

We introduce AML-view, which models a shared, investigation-rich financial world and a realistic observation process over that world. First, the latent world integrates persons, households, firms, ownership relationships, accounts, institutions, due-diligence attributes, and transactions

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

into one financial ecosystem. Population and firm characteristics are calibrated against public statistical references; legitimate activity follows recurring economic behaviors; laundering scenarios are grounded in documented typologies and advisories; and illicit value is propagated through subsequent transfers so that labels follow the movement of funds rather than merely marking an injected motif.

Second, AML-view models how each institution observes that same latent world. It first constructs one shared cross-institution ecosystem and then derives each institution's dataset as a projection of it, retaining the institution's own customers, records, and transactions while limiting information about external counterparties. Because the underlying persons, accounts, transactions, and laundering events are held fixed, differences across views reflect the observation boundary rather than different simulations. This enables controlled study of how institutional information constraints affect AML detection. To our knowledge, no existing synthetic AML release provides multiple institution-specific views of one common financial world.

AML-view makes four contributions:

- **Holistic, statistically grounded financial-world simulation.** AML-view jointly models persons, households, firms, ownership, accounts, institutions, due-diligence context, and financial activity in one coherent ecosystem, with population and firm characteristics calibrated against public references and linked through explicit ownership, account, and institutional relationships. This goes beyond prior generators that typically model only subsets of these components or lack external calibration and realistic bridging across the investigation graph.
- **Rich, document-grounded laundering scenarios and flow-based labels.** AML-view instantiates 38 laundering scenarios from published typologies and advisories and propagates illicit value through onward transfers, so labels follow the movement of funds rather than merely scenario membership.
- **Institution-centric views of one shared world.** AML-view derives the complete cross-institution record and multiple institution-local views from the same latent world, preserving observation boundaries while holding the underlying entities, transactions, and laundering events fixed.
- **Configurable and extensible generation framework.** Through our software, users can specify statistical references, financial behaviors, institutional structures, and laundering configurations to generate new synthetic worlds and institution-local views without modifying the core pipeline.

We evaluate AML-view along four dimensions. First, *statistical fidelity* is assessed against the public statistical targets used for calibration: all 35 pre-specified population checks pass, with mean TVD 0.003 and pooled $R^2 = 0.9999$. Second, *investigation and laundering coverage* is measured against FATF-anchored dimensions and documented typologies: AML-view spans 13 investigation axes and provides 38 documented laundering scenarios, corresponding to 31.88 effective typologies. Third, matched global and institution-local views reveal effects hidden by global-only evaluation: *institution-local visibility* makes detection harder for most models (e.g., TADDY falls from 0.739 to 0.510 AUROC), while the relative ranking of algorithms can also change (e.g., RustGraph improves). Finally, in a *blinded practitioner study*, synthetic suspicious cases support laundering-pattern recovery at a rate comparable to real cases, recovery increases with richer investigation context, and practitioners show no consistent ability to distinguish synthetic from real cases.

2 Related Work

Synthetic AML datasets initially focused on transaction generation, providing controlled environments in which normal activity could be simulated and suspicious behavior injected. AMLSim exemplifies this paradigm by generating labelled transaction networks while exposing little information about the persons, firms, or relationships behind the accounts (Suzumura & Kanezashi, 2018). This line of work enabled reproducible detection experiments but largely treated accounts and transactions as the primary objects of analysis. Later generators increased either behavioral realism or customer context. SynthAML learns from one bank's alerts and transactions and partially represents customer attributes, bringing synthetic data closer to operational AML records (Jensen et al., 2023). SAML-D expands laundering behavior using scenarios derived from the literature and expert interviews (Öztas et al., 2023), while AMLNet (Huda et al., 2025) similarly incorporates AUSTRAC knowledge into scenario construction. These systems improve customer representation and laundering-pattern diversity, but

Dataset	Pers.	Firms	HH	Own.	Cust. profile	PEP / media / identity	Inst.-local obs.	AML label	Typ. label	Ground.
AMLSim (Suzumura & Kanezashi, 2018)	×	×	×	×	×	×	×	✔	✔	△
SynthAML (Jensen et al., 2023)	△	×	×	×	△	×	×	×	×	✔
SAML-D (Öztas et al., 2023)	×	×	×	×	×	×	×	✔	✔	△
AMLNet (Huda et al., 2025)	△	×	×	×	△	×	×	✔	×	△
AMLGentex (Östman et al., 2025)	△	×	×	×	△	×	△	✔	✔	✔
Tide (van den Beukel et al., 2026)	✔	✔	×	△	△	×	×	✔	×	△
AMLWorld (Altman et al., 2023)	×	×	×	×	×	×	△	✔	✔	△
TransXion (Chen et al., 2026)	✔	△	×	×	✔	×	×	✔	×	△
AML-view	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔

Table 1: Capabilities exposed by synthetic AML releases. ✔ denotes direct support, △ partial support, and × no support. HH: households; Own.: ownership; Inst.-local obs: institution-local projections with restricted external-counterparty information; Ground.: external grounding.

expose limited information about firms, ownership, and the due-diligence evidence used during investigation.

More recent work provides finer control over financial structure and behavior. AMLGentex introduces configurable topology, open-system flows, label noise, and per-institution transaction networks (Östman et al., 2025); however, partitioning transactions by institution does not model asymmetric observation, where a bank has private information about its own customers but only limited transactional, identifier, or public information about external counterparties. Tide adds persons, firms, ownership, and controlled structural and temporal laundering patterns (van den Beukel et al., 2026). TransXion conditions behavior on persistent socio-economic profiles, enabling deviations from expected customer activity to be measured (Chen et al., 2026). AMLWorld simulates a broader economy with persons, firms, banks, salaries, and purchases, although most of these entities are not exposed in its released transaction data (Altman et al., 2023). Collectively, these approaches enrich behavior, customer context, and network structure, but do not jointly expose investigation-relevant records under institution-local visibility.

Table 1 compares the entities, relationships, investigation attributes, labels, and observation boundaries exposed by these releases. Existing datasets cover individual dimensions, but none of the evaluated releases combines them in one shared financial world and derives institution-specific views from that same underlying state.

AML-view addresses this gap by generating persons, firms, households, ownership, accounts, customer profiles, due-diligence attributes, payments, and laundering labels within a common financial world, then projecting that world into institution-specific views containing only locally observable information. This allows suspicious transactions to be studied together with the parties, ownership structures, and investigation evidence available to the reviewing institution.

3 Methodology: Financial-World Design and Generation

AML-view constructs one financial world in four stages: generate grounded parties, simulate legitimate and laundering transactions, bind parties to accounts, and apply institution-specific observation operators. Because all views contain the same entities and events, their differences reflect information availability. Figure 1 summarizes the pipeline.

3.1 Statistically Grounded Parties

AML-view first constructs persons and firms consistent with observable population structure, providing interpretable holders for accounts and payments. The Party Module uses iterative proportional fitting (IPF), which adjusts a joint distribution matching specified marginals (Deming & Stephan, 1940) and is standard in synthetic-population construction (Arentze et al., 2007; Rich & Mulalic, 2012; Durán-Heras et al., 2018). It jointly fits adults by age, sex, socio-professional category, diploma, department, and work sector, and firms by sector, size, legal form, and department. It conditionally samples other demographic, economic, housing, and firm attributes from public sources; seeded routines assign politically exposed person (PEP) and adverse-news profiles. Appendix E documents

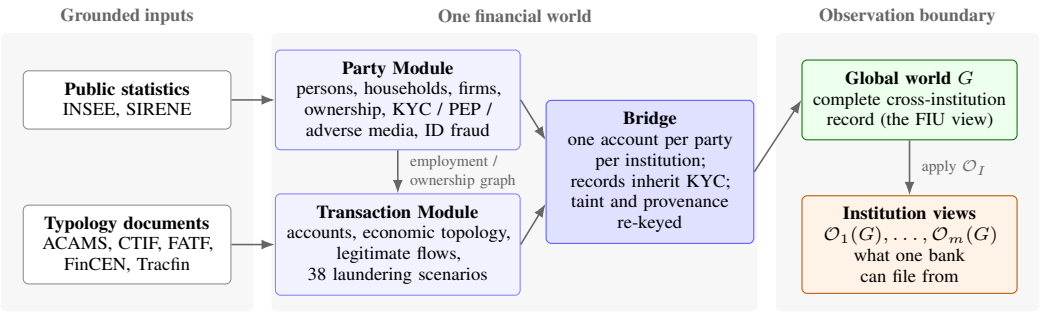


Figure 1: One grounded financial world and multiple institution-local observations. The Bridge binds accounts to parties and propagates due-diligence attributes, taint, and provenance to joined records. Applying $\mathcal{O}_I$ changes visibility without changing the underlying world.

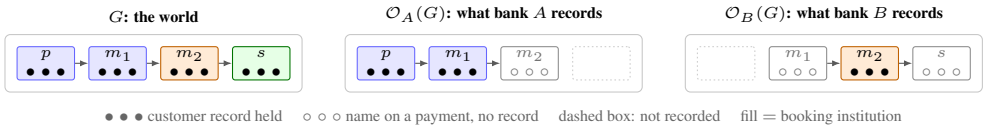


Figure 2: One laundering chain, recorded three ways: p is the proceeds source, m_1 and m_2 are mule accounts, and s is the sink account. Both institution views are projections of the same G, so their differences reflect the recording boundary rather than different simulations.

each source. Assigning these parties accounts, ownership relations, and payments makes each transaction interpretable against holder context. Calibration is applied where reference distributions exist; laundering prevalence, typology mixtures, and some structural parameters remain experimental choices.

3.2 Transactions and Provenance-Tracked Laundering

AML-view next generates financially coherent payments and traces illicit value through later transfers, so labels follow funds rather than scenario membership. The Transaction Module constructs the account–institution topology, completes legitimate activity, injects documented laundering scenarios, and chronologically replays the combined ledger.

The topology specifies accounts, institutions, cross-institution corridors, and potential counterparties while targeting the heavy-tailed degrees observed in payment networks (Soramäki et al., 2007). Structural hubs such as merchants and utilities are selected before transaction generation from planned degree, opening balance, institution size, and company status, so realized activity does not determine their roles.

Following AMLWorld (Altman et al., 2023) and AMLNet (Huda et al., 2025), legitimate payments follow recurring economic agendas conditioned on customer attributes rather than independent edge samples. AML-view generates nine flow families conditioned on demographic, geographic, and firm attributes, with schedules and transaction metadata. Payments execute against running balances; opening balances target a wealth inequality, and infeasible discretionary outflows are deferred or resized. Legitimate activity precedes laundering injection, preventing illicit flows from determining baseline topology or behavior. Appendix C gives the flow functions and reconciliation order.

Laundering activity uses 38 documented scenarios: six legacy motifs, 16 ACAMS alerts (ACAMS, 2022), and 16 complex cases from CTIF (Cellule de traitement des informations financières (CTIF), 2024), FATF (Financial Action Task Force (FATF), 2006; 2018; 2007), FinCEN (Financial Crimes Enforcement Network (FinCEN), 2014; 2016), and Tracfin (Tracfin, 2024). Each specifies participating accounts, transfer steps, laundering stages, and documentary provenance.

During replay, a taint ledger maintains separate legitimate and illicit balances. Payments draw first from illicit funds and propagate the resulting split to the destination (Möser et al., 2014). A transfer is positive when its illicit amount exceeds the materiality floor and constitutes at least $\tau = 0.25$ of the payment, preventing indefinitely diluted taint. This extends AMLWorld's flow tracking (Altman

et al., 2023) with a materiality criterion and exported tranche amounts; Appendix C.3 gives the equations, order, and collision rules. Positives include scenario-emitted transfers and later payments carrying propagated illicit value. Propagation alone produces 84.0–95.1% of positives across five seeds. Detection therefore primarily evaluates traced illicit-value movement beyond directly emitted scenario transactions.

3.3 Bridge and Unified World Representation

AML-view next preserves event identity and investigation context before restricting visibility. An account-to-party assignment layer, the Bridge, binds persons and private firms to accounts so that payments, customer records, and exports resolve to consistent parties. Each account has one holder, while parties may bank across institutions. Customer records inherit KYC, risk, PEP, sanctions, adverse-media, identity-verification, and ownership attributes. Ownership remains graph-structured because beneficial-control chains and cycles require relational reasoning (Echeverria et al., 2025).

The Bridge re-keys transactions, balances, fund tranches, customer attributes, and provenance to account identifiers. Transaction labels mark material illicit value; typology labels identify the emitting scenario. Propagation-only positives receive `natural_integration`, denoting illicit value without direct scenario attribution. Binary detection uses all positives, whereas typology classification uses only scenario-emitted transfers. Because binding precedes visibility, every institution restricts the same population, ledger, labels, and provenance. Appendix C specifies the procedure.

3.4 Institution-Local Observation and Export

AML-view derives each institution's evidence from the unified world without regenerating entities or events. Let $G = (V, E)$ be the global financial world, I an institution, and $A_I \subseteq V$ its accounts. The observation operator $O_I(G) = (V_I, E_I)$ retains every transfer involving at least one account at I, $E_I = \{e \in E : \mathrm{src}(e) \in A_I \vee \mathrm{dst}(e) \in A_I\}$, and every account touched by those transfers,$V_I = \{v \in V : \exists e \in E_I,\ v = \mathrm{src}(e) \vee v = \mathrm{dst}(e)\}$. Thus external counterparties remain visible only when they transact with a customer of I.

The operator also restricts attributes according to their source. It retains private due-diligence and behavioral records for local customers, but exposes external counterparties mainly through observed transaction identifiers and public registry information. An institution-local view is therefore not the global graph with nodes removed: its interior contains rich private customer context, whereas its boundary contains mostly transactional and public evidence. Because O_I selects from one shared world rather than resampling it, all views preserve the same underlying entities, transfers, laundering events, and labels. Figure 2 illustrates this asymmetry, and Appendix C.4 specifies the visibility tiers.

Joined records are exported as tables and graphs. To prevent leakage, Table 11 separates observable features, prediction targets, and evaluation-only metadata: labels are targets, while catalog identifiers and simulator provenance remain evaluation metadata. The exports therefore hold events and labels fixed while varying only the evidence available to each institution.

4 Empirical Evaluation of AML-view

We evaluate AML-view through three research questions. **RQ1:** How comprehensively does AML-view represent investigation-relevant information and laundering structure? **RQ2:** How well is the generated financial world grounded in external statistical and documentary sources? **RQ3:** How does institution-local observability affect AML detection when the underlying financial world is fixed? Unless otherwise stated, quantitative experiments are evaluated over five seeds, each generating approximately 10 million transactions across 70,000 accounts.

4.1 RQ1: Investigation-Relevant Information and Structure

RQ1 asks whether released AML datasets expose sufficient information and structure for the investigation beyond transaction patterns. Our evaluation is based on the following three properties.

Investigation-axis coverage measures the entities, relationships, due-diligence attributes, observation boundaries, and labels exposed by each release. We define thirteen investigation-relevant axes represented by AML-view or at least one comparator and map each to a corresponding FATF Recommendation clause (Financial Action Task Force (FATF), 2012) (Table 5, appendix). The axes provide a common comparison surface rather than an exhaustive encoding of FATF requirements. An axis is *released* when directly present in the data, *derivable* when recoverable through documented generator options or standard derivation without additional modeling, and *absent* otherwise. AML-view releases all **13** axes, compared with **7** for Tide, the strongest comparator (full results in Table 7). Five axes are unique to AML-view: household structure, identity verification, beneficial ownership, PEP status, and adverse-media screening. Restricting the comparison to the 8 axes released by at least one comparator gives AML-view 8/8 versus Tide 7/8, showing the advantage beyond AML-view-specific axes.

Ontology alignment assesses correspondence between each released schema and established financial and investigative data models. Every released class and relation is mapped to the Financial Industry Business Ontology (FIBO) (EDM Council) and FollowTheMoney (OCCRP / Aleph) using explicit equivalence or subsumption relations. Because schema sizes differ, we report mapped release types and distinct external terms rather than a normalized score. AML-view maps **42** release types to **52** external terms, compared with **32** types and **21** distinct terms across all six comparators combined (Table 14). Its 26 unmapped types largely capture investigation-specific concepts absent from both ontologies, including institution views, KYC profiles, laundering scenarios, and typology labels.

Scenario-structure diversity investigates whether a large scenario catalog yields genuinely diverse generated behavior. A catalog may still have low effective diversity if most laundering rows come from only a few scenarios or if several scenarios share the same transaction structure. For directly scenario-emitted rows, let p_i be the share generated by catalog entry i. We report normalized entropy $H = -\sum_{i=1}^{T} p_i \log p_i / \log T$ and effective count $N_{\text{eff}} = T^H$, interpreted as the number of equally frequent scenarios with the same diversity. We repeat the calculation after grouping scenarios by account count, degree multisets, temporal steps, and path depth. Across 4,390 scenario-emitted rows, AML-view has $T = 38$, $H = 0.9517$, and $N_{\text{eff}} = 31.88$, showing broad use of the catalog. Grouping the same rows yields $S = 15$ structural signatures with $H = 0.8621$ and $N_{\text{eff}} = 10.33$, spanning chains, stars, cycles, neighborhoods, bipartite flows, and composite flows (Table 13). Thus, the documentary diversity is accompanied by substantial structural diversity rather than repeated labels for the same motif.

4.2 RQ2: Statistical and Documentary Grounding

RQ2 asks how strongly AML-view is anchored in external evidence beyond internal consistency alone. We evaluate two complementary forms of grounding: *statistical grounding*, through agreement with public population references, and *documentary grounding*, through traceability of generated laundering activity to published typologies and advisories.

Population fidelity measures agreement between generated attributes and public statistical references using 35 pre-specified checks. Target-specific TVD measures distributional error, pooled R^2 measures agreement with reference values, and dependency statistics measure preservation of relationships between attributes. On a France-scoped run with 1.28 million individuals and 85,380 companies, all 35 checks pass, with mean TVD 0.003 and pooled $R^2 = 0.9999$ (Table 9). These results measure agreement with construction targets rather than generalization to independent held-out data.

For comparison, TransXion's underlying population references are not matched to our French targets, so its distance from France alone cannot distinguish methodological error from population-source differences. We therefore also report the TVD between real French and U.S. populations as a cross-country reference scale. AML-view has lower TVD on every shared target, with mean person TVD **0.0012** versus **0.2855** for TransXion. Moreover, TransXion's deviations on sex and age substantially exceed the corresponding France–U.S. distances (Table 10). This comparison shows that calibration to a specific jurisdiction is not provided by existing generators.

Documentary traceability measures whether generated laundering activity retains provenance to external AML sources. Each scenario-emitted transaction retains its catalog entry and published

Method	Type	AUROC		F_1	
		global / local	Δ_{obs}	global / local	Δ_{obs}
Investigator (GBM)	sup. ref.	0.893±.005 / 0.831±.005	+0.062	0.594±.031 / 0.513±.027	+0.081
SAD (Tian et al., 2023)	semi.	0.665±.017 / 0.632±.013	+0.033	0.046±.033 / 0.020±.011	+0.026
GCN (Kipf & Welling, 2017)	uns.	0.852±.006 / 0.824±.007	+0.028	0.148±.002 / 0.146±.007	+0.002
GAT (Veličković et al., 2018)	uns.	0.767±.007 / 0.671±.016	+0.096	0.081±.025 / 0.030±.028	+0.051
GraphSAGE (Hamilton et al., 2017)	uns.	0.690±.004 / 0.527±.012	+0.163	0.081±.023 / 0.032±.020	+0.049
RustGraph (Guo et al., 2023)	uns.	0.704±.008 / 0.767±.010	−0.063	0.073±.015 / 0.062±.031	+0.011
TADDY (Liu et al., 2021)	uns.	0.739±.005 / 0.510±.015	**+0.229**	0.041±.034 / 0.018±.016	+0.023
R-GCN (Schlichtkrull et al., 2018)	rel.	0.711±.004 / 0.654±.014	+0.057	0.081±.016 / 0.030±.023	+0.051
HGT (Hu et al., 2020)	rel.	0.823±.004 / 0.784±.011	+0.039	0.055±.017 / 0.033±.017	+0.022

Table 2: Global/local detection on the account graph under the causal feature regime, mean ± standard deviation over five seeds. Investigator (GBM) is an optimistic supervised reference using confirmed labels; the remaining methods follow their benchmark supervision regimes. Δ_{obs} denotes global-minus-local performance. Positive values indicate lower local performance; negative values indicate local improvement. AUPRC and complete static/causal results are reported in Appendix G.

source, while later positives created through illicit-value propagation retain traced illicit amounts but no scenario identifier. Of 52,660 positive transfers, 4,390 are directly emitted by the 38 documented scenarios and carry scenario-level provenance; the remaining 48,270 arise through propagated illicit value and retain flow-level provenance. The large majority of positives therefore arise from downstream illicit-value propagation rather than direct scenario emission, creating a richer and more realistic setting in which investigators must follow illicit funds beyond the originally documented laundering events.

4.3 RQ3: Detection Under Institution-Local Observability

RQ3 asks how institution-local observability changes AML detection when the underlying financial world is held fixed. Existing AML datasets typically expose a global transaction graph, although no individual institution has such visibility in practice. AML-view enables a controlled comparison of global and institution-local observations of the same entities, transactions, and laundering events, revealing both the detection difficulty and model-ranking issues obscured by global-view evaluation.

Model choices include one supervised reference, Investigator, one semi-supervised method, SAD, together with seven graph-based methods from DyGADBench (Mezhoudi et al., 2026): GCN, GAT, GraphSAGE, RustGraph, TADDY, R-GCN, and HGT. DyGADBench provides a unified evaluation pipeline for state-of-the-art dynamic graph anomaly detection methods spanning different structural and temporal modeling choices. Investigator is a gradient-boosted model over endpoint-account and transaction attributes (Appendix F). Because the supervised methods assume confirmed labels, and Investigator additionally uses an enriched laundering prevalence, we interpret them as optimistic learnability references rather than real deployment settings. We report the causal feature regime in the main text, where each transaction uses only behavioral information available strictly before that transaction; complete static and causal results are reported in Appendix G. Global and local views use matched preprocessing, chronological 50%/25%/25% train/validation/test partitions, and class composition. For detector f and metric m, we report $\Delta_{\text{obs}}(f) = \mathrm{m}(f; G) - \mathrm{m}(f; O_I(G))$, so positive values indicate lower performance under institution-local visibility.

Observability effect shows that realistic information restriction materially increases detection difficulty (Table 2). Even the optimistic supervised Investigator reference degrades, dropping from 0.893 to 0.831 AUROC, while the semi-supervised SAD model decreases from 0.665 to 0.632. Among the seven graph-based methods, six also lose performance, with degradation reaching 0.229 **AUROC for TADDY** ($0.739 \rightarrow 0.510$), a transformer-based dynamic graph anomaly detector. The effect is strongly model-dependent: GraphSAGE drops by 0.163, GAT by 0.096, R-GCN by 0.057, and HGT by 0.039, whereas RustGraph is the sole exception, improving from 0.704 to 0.767. Thus, restricting visibility not only produces a harder and more realistic AML detection setting, but can also change the relative suitability of detection methods. By releasing matched global and institution-specific views of the same financial world, AML-view makes both effects directly measurable.

Method	tx-only	homo-id / hetero	Lift
GCN (Kipf & Welling, 2017)	0.852	0.835	$\Lambda_{\mathrm{id}} = -0.017$
GAT (Veličković et al., 2018)	0.767	0.697	$\Lambda_{\mathrm{id}} = -0.070$
GraphSAGE (Hamilton et al., 2017)	0.690	0.700	$\Lambda_{\mathrm{id}} = +0.010$
SAD (Tian et al., 2023)	0.665	0.650	$\Lambda_{\mathrm{id}} = -0.015$
RustGraph (Guo et al., 2023)	0.704	0.776	$\Lambda_{\mathrm{id}} = +0.072$
TADDY (Liu et al., 2021)	0.739	0.699	$\Lambda_{\mathrm{id}} = -0.040$
R-GCN (Schlichtkrull et al., 2018)	0.711	0.708	$\Lambda_{\mathrm{rel}} = -0.003$
HGT (Hu et al., 2020)	0.823	0.824	$\Lambda_{\mathrm{rel}} = +0.001$

Table 3: Graph-construction ablation in the global causal setting, mean AUROC over five seeds. Λ_{id} compares `homo-id` with `tx-only`; Λ_{rel} compares `hetero` with `tx-only`. The effect of richer graph construction is model-dependent, ranging from -0.070 to $+0.072$ AUROC among homogeneous detectors, while typed relations have little effect on R-GCN and HGT in this setting.

Practitioner validity		**Synthetic-case validity**		**Investigation context**		**Provenance judgment**	
Condition	*Result*	*Condition*	*Result*	*Condition*	*Result*	*Condition*	*Result*
Real suspicious	0.625	Synthetic	**0.667**	T	0.500	Real vs. synthetic	0.458
Controls	1.000	Real	0.625	TC	0.688		
				TIC	**0.765**		

Table 4: Summary of practitioner-study results.

Representation ablation separates the observation-boundary effect from the value of AML-view's richer relational representation. Within a fixed view, the tx-only construction contains transaction edges; homo-id additionally links accounts sharing identifiers; and hetero preserves typed relations for R-GCN and HGT. For metric m, we define $\Lambda_{\mathrm{id}} = m(\text{homo-id}) - m(\text{tx-only})$ and $\Lambda_{\mathrm{rel}} = m(\text{hetero}) - m(\text{tx-only})$, measuring the contribution of identity linkage and typed relations beyond transaction structure. Table 3 reports the global causal setting, with all global/local and static/causal variants in Appendix G. Identity linkage is strongly model-dependent: RustGraph improves by 0.072 AUROC and GraphSAGE by 0.010, whereas GAT decreases by 0.070 and the remaining homogeneous detectors change by at most 0.040. Typed relations have little effect on R-GCN and HGT in this setting, changing AUROC by -0.003 and $+0.001$, respectively. These results indicate that transaction structure already provides substantial detection signal, while AML-view's richer customer and relational context provides additional information for relational modeling and downstream investigation.

5 Practitioner Evaluation of AML-view

We complement generator-level validation with two practitioner studies evaluating AML-view for its intended investigative use: an investigative review of institution-local cases and a provenance task testing whether synthetic cases expose recognizable artifacts. Four practitioners from a major European retail and corporate bank (name disclosed after review), with 2, 2, 3, and 10 years of AML experience, participated. Full details are provided in Appendix H.

Practitioner task performance. Before interpreting judgments on synthetic cases, we assess whether practitioners perform the investigative task reliably on the reference cases. On real suspicious cases, they recover the laundering pattern in 15/24 judgments (0.625), while all 25 legitimate-control judgments correctly attribute no suspicious laundering behavior (Table 4). These results establish that the panel can recover laundering structures on real reference cases, providing an empirical basis for interpreting its evaluation of synthetic cases.

Investigative review by industry AML experts. Practitioners assessed 18 case files spanning real suspicious cases, synthetic cases with injected laundering patterns, and synthetic legitimate controls. Each case exposed only the reviewing institution's local view and was presented under one of three information conditions: transactions only (T), transactions plus the customer file (TC), or transactions plus the full investigation layer (TIC). Case category and provenance were hidden. Across practitioners and conditions, 77 screens were completed and 76 retained after excluding one unusable response. The primary outcome is laundering-pattern recovery: a suspicious case is recovered when its instantiated laundering pattern is identified in the written explanation, while a control is benign when no suspicious laundering behavior is attributed to it. Across suspicious cases,

practitioners recovered the laundering pattern in 33/51 judgments (0.647; 95% Wilson CI [0.510, 0.764]), while all 25 control judgments were benign (95% Wilson CI [0.867, 1.000]). Escalation captures a distinct operational decision: only 10/51 suspicious-case responses selected escalation, and 23/33 recovery hits did not. Under the original binary interpretation of this field, the control rate was 0.60 (Appendix H); we therefore analyze recovery and escalation separately.

Synthetic-to-real comparison. The synthetic cases contain substantial recoverable investigative signal: practitioners recover the laundering pattern in 18/27 synthetic suspicious-case judgments (0.667), compared with 15/24 (0.625) on the real-case reference, with $p = 0.778$ (Table 4). This test does not establish equivalence, but within this sample the recovery rates are comparable. Moreover, 64 of 65 written explanations refer to evidence contained in the corresponding case file, indicating that practitioners' reasoning is grounded in the presented evidence. Thus, synthetic AML-view cases support evidence-grounded laundering-pattern reconstruction at a rate comparable to the real-case reference.

Contribution of investigation context. We measure the value of non-transactional evidence by comparing laundering-pattern recovery across T, TC, and TIC. Recovery increases from 0.500 with transactions alone to 0.688 with the customer file and 0.765 with the full investigation layer (Table 4). Because recovery requires reconstructing the instantiated laundering pattern rather than merely flagging the case, the increase across conditions is consistent with customer and investigation records providing additional evidence for case interpretation beyond transaction structure alone, complementing the detection results in Section 4.3. The ordered trend is suggestive but falls just short of the conventional significance threshold (one-sided $p = 0.052$; Appendix H).

Provenance discrimination. The provenance task assesses whether synthetic cases expose recognizable artifacts. Each practitioner classified 30 additional cases—15 real and 15 synthetic—using a snap real-versus-synthetic judgment. Pooled accuracy is $55/120 = 0.458$ (95% CI [0.383, 0.533]) against chance at 0.5 (Table 4), and no practitioner identified a specific synthetic artifact when asked. Individual performance varies, with the most and least accurate practitioners both differing from chance ($p = 0.005$) in opposite directions, indicating no consistent provenance cue across the panel.

Together, the studies show that expert practitioners can recover laundering patterns from institution-local AML-view cases, that synthetic suspicious cases carry recoverable investigative signal comparable to the real-case reference, that richer customer and investigation context improves case reconstruction, and that the synthetic cases expose no consistent provenance cue.

6 Discussion and Conclusion

AML-view addresses a central limitation of existing synthetic AML data: realistic investigation requires both a rich financial world and an institution-specific view of that world. We release a large-scale synthetic dataset and configurable generator covering persons, firms, ownership, due-diligence attributes, payments, laundering labels, and institution-local observations, with approximately 10 million transactions over 70,000 accounts per evaluated world. Both researchers and industry practitioners can use AML-view to develop new models and validate detection approaches in realistic institutional settings. AML-view covers all thirteen audited investigation axes, instantiates 38 documented laundering scenarios with substantial structural diversity, and closely matches its statistical calibration targets. Under causal features, institution-local visibility reduces AUROC for both supervised and semi-supervised references and six of seven graph-based methods, with model-dependent effects. In blinded review, AML practitioners recover laundering patterns from synthetic suspicious cases at a rate comparable to the real-case reference, with recovery increasing as richer investigation context is added and no consistent cue revealing synthetic provenance. By separating the underlying financial world from the evidence available to each institution, AML-view provides a large-scale, expert-validated, reproducible foundation for AML detection and investigation research.

The current release is calibrated to France and to the laundering scenarios represented in the catalog, and practitioner validation is based on a small expert panel. Future work should develop agentic systems for efficient AML detection and investigation; extend to additional jurisdictions; scale expert validation; and explore novel detection and investigation techniques with temporal drift models.

REPRODUCIBILITY STATEMENT

We release the AML-View generator and a generated instance of the dataset as supplementary material. The supplementary archive contains the generator code, a tutorial explaining how to obtain the input statistics and run it, and a compact copy of the seed-44 instance. The complete instance, including the full transaction ledger with per-transaction metadata, is hosted at an anonymized repository: `https://osf.io/j6mkh/overview?view_only=ea02f02e45bd41899011b7439d1d223a`. The generative process (entity schema, ownership and account structure, the 38 laundering scenarios, and the provenance-aware taint-propagation labels) is described in Section 3, with full parameterization in Appendix E. The evaluation protocol, detector configurations, feature regimes, and metrics are given in Section 4 and Appendix G. All INSEE and SIRENE reference tables are openly available from the INSEE website (`https://www.insee.fr`) (Licence Ouverte 2.0) and the French government website for data (`https://www.data.gouv.fr`), and Appendix E lists the exact tables used for calibration. After the review period, the generator and dataset will be released publicly on GitHub under an open-source license.

ETHICS STATEMENT

AML-View contains no real personal data: all persons, firms, accounts, and transactions are synthetic, calibrated only against publicly available aggregate statistics (INSEE, SIRENE), so no individual is identifiable (cf. GDPR Art. 4(1); Appendix I). The practitioner study was conducted with informed consent and without exposing any confidential case data; protocol and confidentiality measures are in Appendix H.

AI USE STATEMENT

In this work, we used generative AI tools to assist with writing and editing of the manuscript (grammar, phrasing, and copy-editing of text drafted by the authors) and as a coding assistant for parts of the experimental code. We did not use generative AI tools for research ideation, for deriving the theoretical results, or for producing or interpreting any experimental result nor to generate the dataset; retrieval and discovery of related work were carried out by the authors as well as partly assisted by AI. All AI-assisted text was reviewed and revised by the authors, and all AI-assisted code was read, tested, and verified against the protocols described in Appendix C and F. The authors take full responsibility for the final content of this work, including all text, claims, and artifacts.

REFERENCES

ACAMS. *Certified Anti-Money Laundering Specialist (CAMS) Study Guide*. Association of Certified Anti-Money Laundering Specialists (ACAMS), version 6.44 edition, 2022. ISBN 978-0-9777495-2-2. URL https://www.acams.org.

Erik Altman, Jovan Blanuša, Luc von Niederhäusern, Béni Egressy, Andreea Anghel, and Kubilay Atasu. Realistic synthetic financial transactions for anti-money laundering models. In *Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track*, 2023. URL https://arxiv.org/abs/2306.16424. Also available as arXiv:2306.16424.

Theo Arentze, Harry Timmermans, and Frank Hofman. Creating synthetic household populations: Problems and approach. *Transportation Research Record*, 2014(1):85–91, 2007.

Cellule de traitement des informations financières (CTIF). Vademecum : Typologies des schémas de blanchiment de capitaux identifiés par la ctif. Vademecum, CTIF-CFI, 2024. URL https://www.ctif-cfi.be.

Keyang Chen, Mingxuan Jiang, Yongsheng Zhao, Zeping Li, Zaiyuan Chen, Weiqi Luo, Zhixin Li, Sen Liu, Yinan Jing, Guangnan Ye, et al. Transxion: A high-fidelity graph benchmark for realistic anti-money laundering. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pp. 8707–8718, 2026.

W Edwards Deming and Frederick F Stephan. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4): 427–444, 1940.

Alfonso Durán-Heras, Isabel García-Gutiérrez, and Guillermo Castilla-Alcalá. Comparison of iterative proportional fitting and simulated annealing as synthetic population generation techniques: Importance of the rounding method. *Computers, Environment and Urban Systems*, 68:78–88, 2018.

Fabricio Echeverria, Marcelo Leon, Paula Erazo, and Simone Belli. Transparency in corporate networks: a graph-based model to reduce computational complexity in identifying total ownership of ultimate beneficial owners. 2025.

EDM Council. Financial industry business ontology (fibo). https://spec.edmcouncil.org/fibo/. FND and FBC modules.

Financial Action Task Force (FATF). Trade-based money laundering. Technical report, FATF/OECD, 2006. URL https://www.fatf-gafi.org/en/publications/Methodsandtrends/Trade-basedmoneylaunderingtypologies.html.

Financial Action Task Force (FATF). Money laundering and terrorist financing through the real estate sector. Technical report, FATF/OECD, 2007. URL https://www.fatf-gafi.org/en/publications/Methodsandtrends/Moneylaunderingandterroristfinancingthroughtherealestatesector.html.

Financial Action Task Force (FATF). International standards on combating money laundering and the financing of terrorism and proliferation: The FATF recommendations. Technical report, FATF/OECD, 2012. URL https://www.fatf-gafi.org/en/publications/Fatfrecommendations/Fatf-recommendations.html. Adopted February 2012, as subsequently updated.

Financial Action Task Force (FATF). Professional money laundering. Technical report, FATF/OECD, 2018. URL https://www.fatf-gafi.org/content/dam/fatf-gafi/reports/Professional-Money-Laundering.pdf.

Financial Crimes Enforcement Network (FinCEN). Update on u.s. currency restrictions in mexico: Funnel accounts and trade-based money laundering. Technical Report FIN-2014-A005, FinCEN, U.S. Department of the Treasury, 2014. URL https://www.fincen.gov/resources/advisories/fincen-advisory-fin-2014-a005.

Financial Crimes Enforcement Network (FinCEN). Advisory to financial institutions on e-mail compromise fraud schemes. Technical Report FIN-2016-A003, FinCEN, U.S. Department of the Treasury, 2016. URL https://www.fincen.gov/resources/advisories/fincen-advisory-fin-2016-a003.

Jianhao Guo, Siliang Tang, Juncheng Li, Kaihang Pan, and Lingfei Wu. Rustgraph: Robust anomaly detection in dynamic graphs by jointly learning structural-temporal dependency. *IEEE Transactions on Knowledge and Data Engineering*, 36(7):3472–3485, 2023.

William L. Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017. URL https://arxiv.org/abs/1706.02216.

Ziniu Hu, Yuxiao Dong, Kuansan Wang, and Yizhou Sun. Heterogeneous graph transformer. In *Proceedings of the web conference 2020*, pp. 2704–2710, 2020.

Sabin Huda, Ernest Foo, Zahra Jadidi, M. A. Hakim Newton, and Abdul Sattar. Amlnet: A knowledge-based multi-agent framework to generate and detect realistic money laundering transactions. *arXiv preprint arXiv:2509.11595*, 2025. URL https://arxiv.org/abs/2509.11595.

Rasmus Ingemann Tuffveson Jensen, Joras Ferwerda, Kristian Sand Jørgensen, Erik Rathje Jensen, Martin Borg, Morten Persson Krogh, Jonas Brunholm Jensen, and Alexandros Iosifidis. A synthetic data set to benchmark anti-money laundering methods. *Scientific Data*, 10(661), 2023. doi: 10.1038/s41597-023-02569-2. SynthAML dataset based on real Spar Nord data.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2017. URL https://arxiv.org/abs/1609.02907.

Yixin Liu, Shirui Pan, Yu Guang Wang, Fei Xiong, Liang Wang, Qingfeng Chen, and Vincent CS Lee. Anomaly detection in dynamic graphs via transformer. *IEEE Transactions on Knowledge and Data Engineering*, 35(12):12081–12094, 2021.

Edgar Lopez-Rojas, Abdelrahman Elmir, and Stefan Axelsson. Paysim: A financial mobile money simulator for fraud detection. In *28th European Modeling and Simulation Symposium (EMSS)*, pp. 249–255, 2016. URL http://bth.diva-portal.org/smash/record.jsf?pid=diva2%3A1058442.

Mohamed Nazim Mezhoudi, Guillaume Lachaud, Yanlei Diao, Tobias Möller, Mariam Barry, and Albert Bifet. Dygadbench: A comprehensive benchmark for anomaly detection in dynamic graphs. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pp. 9569–9580, 2026.

Malte Möser, Rainer Böhme, and Dominic Breuker. Towards risk scoring of bitcoin transactions. In *International conference on financial cryptography and data security*, pp. 16–32. Springer, 2014.

OCCRP / Aleph. Followthemoney: A data model for investigative graphs. https://followthemoney.tech/.

Johan Östman, Edvin Callisen, Anton Chen, Kristiina Ausmees, Emanuel Gårdh, Jovan Zamac, Jolanta Goldsteine, Hugo Wefer, Simon Whelan, and Markus Reimegård. Amlgentex: Mobilizing data-driven research to combat money laundering. *arXiv preprint arXiv:2506.13989*, 2025.

Berkan Öztas, Deniz Cetinkaya, Festus F. Adedoyin, Marcin Budka, Huseyin Dogan, and Gökhan Aksu. Enhancing anti-money laundering: Development of a synthetic transaction monitoring dataset. In *2023 IEEE International Conference on e-Business Engineering (ICEBE)*, pp. 47–54. IEEE, 2023. URL https://api.semanticscholar.org/CorpusID:266419345.

Jeppe Rich and Ismir Mulalic. Generating synthetic baseline populations from register data. *Transportation Research Part A: Policy and Practice*, 46(3):467–479, 2012.

Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pp. 593–607. Springer, 2018.

Kimmo Soramäki, Morten L Bech, Jeffrey Arnold, Robert J Glass, and Walter E Beyeler. The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications*, 379 (1):317–333, 2007.

Toyotaro Suzumura and H. Kanezashi. Amlsim: A multi-agent simulator for anti-money laundering research. https://github.com/IBM/AMLSim, 2018. Synthetic banking transaction simulator with predefined laundering patterns.

Sheng Tian, Jihai Dong, Jintang Li, Wenlong Zhao, Xiaolong Xu, Bowen Song, Changhua Meng, Tianyi Zhang, Liang Chen, et al. Sad: Semi-supervised anomaly detection on dynamic graphs. *arXiv preprint arXiv:2305.13573*, 2023.

Tracfin. Tendances et analyse des risques de blanchiment de capitaux et de financement du terrorisme en 2023–2024. Rapport annuel, tome iii, Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique, 2024. URL https://www.economie.gouv.fr/files/files/directions_services/tracfin/TRACFIN_2023_Tome3_FR_WEB.pdf.

Montijn van den Beukel, Jože Martin Rožanec, and Ana-Lucia Varbanescu. Tide: A customisable dataset generator for anti-money laundering research. *arXiv preprint arXiv:2603.01863*, 2026. URL https://arxiv.org/abs/2603.01863.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations (ICLR)*, 2018. URL https://arxiv.org/abs/1710.10903.

APPENDIX CONTENT

Appendix	Content
A	Release-scale statistics, FATF investigation axes, and detailed RQ1 comparisons.
B	Generation-scale statistics across the five reported seeds.
C	Population validation, legitimate-flow construction, taint propagation, institution-local visibility, account binding, and structural-signature definitions.
D	FIBO and FollowTheMoney alignment protocol, complete class and relation mappings, non-aligned types, comparator mappings, and target IRIs.
E	Public statistical, regulatory, and typology sources used to parameterize the generator, including IPF targets and conditional distributions.
F	Definition, features, and training protocol of the supervised Investigator reference model.
G	Complete detector results across global/local views, static/causal features, graph constructions, attribute ablations, and operating points.
H	Practitioner-study panel, protocol, scoring rules, pattern-recovery analysis, synthetic–real comparison, and provenance judgments.
I	Data-protection analysis and handling of synthetic identities and real practitioner-study cases.
J	Demographic-bias considerations and the reported sex-association diagnostic.

A ADDITIONAL RELEASE COMPARISONS

To distinguish dataset size from the evidence available for investigation, Table 6 compares account counts, transaction counts, and label prevalence. Table 7 then tests the separate question of which investigation capabilities and typologies each release exposes, providing the detailed comparison behind RQ1.

Dimension	FATF anchor	Axes
Parties and relationships	R.10(a), R.10(b), R.12	individuals; households; firms; ownership and control edges
Due-diligence context	R.10(a)–(d), R.12; R.1, R.6	identity verification; beneficial owner; KYC profile and risk rating; PEP status; adverse-media and sanctions screening
Observation boundary	R.29; R.20	global cross-institution world; per-institution partial view
Label channels	R.20	generator ground truth; institution-side investigation outcome

Table 5: FATF anchors for the thirteen investigation axes: customer due diligence (R.10), PEPs (R.12), institutional reporting (R.20), and financial intelligence units (R.29). Adverse-media screening is a practice standard anchored to ongoing due diligence, not an explicit FATF data item.

Dataset	Accounts	Transactions	Illicit ratio
AMLSim (Suzumura & Kanezashi, 2018)	100K	12.5M	0.137%
SAML-D (Öztas et al., 2023)	750K	9.4M	0.124%
IBM AMLWorld (Altman et al., 2023)	515K–2116K	5M–180M	0.05–0.12%
AMLGentex (Östman et al., 2025)	–	–	≈0.28%
Tide (LI/HI) (van den Beukel et al., 2026)	36.6K nodes	7.6M edges	0.10–0.19%
TransXion (Chen et al., 2026)	47.5K	3.0M	0.153%
AML-view	70,000	10.59M $\pm$ 0.13M	0.54% $\pm$ 0.21%

Table 6: Published scale and laundering-label prevalence. AML-view values are means over five detection runs; the larger population-calibration run is reported separately in Table 9.

B GENERATION STATISTICS ACROSS SEEDS

To document generation stability, we report release-scale statistics for the five independent runs used in the experiments. Each run instantiates all 38 catalog scenarios under the same generator configuration.

Release	Direct /10	Partial /10	Released axes	Typologies	Effective typologies
AML-view	**10**	0	**13/13**	**38**	**31.88**
Tide (van den Beukel et al., 2026)	4	2	7/13	5	4.05
AMLGentex (Östman et al., 2025)	3	2	5/13	7	5.24
IBM AMLWorld (Altman et al., 2023)	2	1	5/13	8	7.32
TransXion (Chen et al., 2026)	3	2	4/13	–	–
AMLNet (Huda et al., 2025)	1	3	3/13	3	2.05
SAML-D (Öztas et al., 2023)	2	2	2/13	17	13.30
SynthAML (Jensen et al., 2023)	1	2	2/13	–	–
AMLSim (Suzumura & Kanezashi, 2018)	1	1	2/13	–	–

Table 7: Detailed RQ1 comparison. Direct and partial counts summarize the ten capabilities in Table 1; released axes report the formal thirteen-axis audit. Effective typologies are computed as $N_{\text{eff}} = \exp(H \log T) = T^H$ over each release's own declared taxonomy.

Seed	Transactions	Positive rows	Scenario rows
40	10,388,005	56,488	4,310
41	10,710,540	92,369	4,553
42	10,673,808	56,559	4,459
43	10,609,773	28,175	4,513
44	10,568,991	52,660	4,390

Table 8: Generation statistics for the five runs underlying the reported experiments. Scenario rows are transactions emitted directly by catalog scenarios; additional positive rows arise through downstream illicit-value propagation.

C Additional Construction Details

The construction checks below specify how AML-view calibrates the population, generates coherent payments, tracks illicit value, and restricts institutional visibility. These details make the links between generated entities, labels, and observable records explicit, so that the evaluation can be interpreted against the actual generation rules.

C.1 Population-validation criteria

Validation distinguishes generator error from finite sampling and inconsistent public tables. For each categorical target p and generated distribution $\hat{p}$, we report $\mathrm{TVD}(p, \hat{p}) = \frac{1}{2} \sum_{i \in Categories} |p_i - \hat{p}_i|$. A target-specific noise band is the 99th percentile of TVD from 500 multinomial draws from p at the generated sample size. We separately measure the discrepancy between each source table and the corresponding fitted IPF marginal; a result is called generator bias only when it exceeds the combined noise-plus-source-floor by 0.01. Dependency preservation is checked with bias-corrected Cramérs V, correlation ratio η, and Spearman's ρ, summarized as $D = \sqrt{\sum_{i<j} (a_{ij}^{\text{gen}} - a_{ij}^{\text{target}})^2}$ over the fitted categorical pairs with available targets. The implementation samples at most 80,000 entities for this association check. Its threshold of 0.5 is a configured diagnostic tolerance, not an external certification of fidelity.

The shared French targets in Table 10 have 2 sex categories, 5 age bands, 5 diploma levels, 9 socio-professional categories, 21 firm-industry sections, and 5 firm-size bands. TransXion does not model legal form or region.

Check	Result	Criterion
Material biases	0	0
Mean target TVD	0.003	$\to 0$
Pooled R^2	0.9999	$\to 1$
Person dependency drift	0.028	< 0.5
Firm dependency drift	0.240	< 0.5
Income–occupation η	0.50	> 0.15

Table 9: Population fidelity on the France-scoped run (1.28M individuals, 85,380 companies). Criteria fixed before generation.

Target	Ours	TransXion	FR–US
Sex	**0.0001**	0.0944	0.0207
Age	**0.0021**	0.2061	0.0585
Diploma	**0.0010**	0.5122	–
Occupation	**0.0016**	0.3293	–
Person mean	**0.0012**	0.2855	–
Firm industry	**0.0072**	0.8155	–
Firm size	**0.0069**	0.8988	–

Table 10: TVD against French targets. FR–US: distance between real populations, as a reference scale.

C.2 Structural hubs and legitimate flows

To generate concentrated payment activity without defining hubs from the resulting transactions, we select structural hubs from the planned account network, or account blueprint. For account a, the score

$$S(a) = w_{\text{in}} \log(1 + d(a)) + w_{\text{bal}} \log(1 + b_a) + w_{\text{inst}} \log(1 + c_a) + b_{\text{co}} \mathbb{1}_{\text{co}}(a)$$

combines structural degree $d(a)$, opening balance b_a, institution cohort size c_a, and a company indicator. A configurable subset of high-scoring accounts routes a controlled share of legitimate spending. This distinguishes high-frequency retail accounts from structural hubs such as central merchants or utilities.

Legitimate behavior separates who transacts, when activity occurs, and what metadata are recorded. A population sampler assigns spending profiles conditional on the owning party's income, occupation, or firm size. A scheduler places recurring salaries, rent, supplier payments, and business-hour activity. A metadata builder attaches location, device, risk, topology, and profile fields. Domestic counterparties follow the topology blueprint, while holdings at other institutions are represented through cross-institution corridors.

Nine flow families execute in a fixed reconciliation order: salary, non-wage income, consumption, household transfer, person-to-person support, inter-company supply, revenue, dividends, and tax. Each reads the generated attributes of the owning party. A salaried adult receives their salary, while an adult without one receives an equivalent monthly pension or benefit, providing coherent inflows without double counting.

Monthly consumption is $(\text{income}/12) \cdot \pi(\text{income}, \text{age})$. The propensity to consume falls linearly from 0.95 at zero income to 0.55 at an annual income of EUR 150,000 and carries a life-cycle adjustment calibrated to INSEE saving rates. Spending is split over a stable merchant set weighted by NAF-section consumption share and geographic proximity, favoring the household's own département. Household transfers scale with the co-residence income gap. Firm supplier spending is a fixed fraction of turnover chosen below the Ésane input-to-revenue ratio so that company accounts close. A merchant's consumer revenue is the emergent sum of its customers' spending, not an independently sampled quantity.

C.3 Taint propagation, materiality and provenance

Ground truth is produced by a taint ledger rather than by motif membership: each account's balance is split into a legitimate and an illicit tranche, and a transfer inherits part of the split of its source. This section states the propagation rule, the materiality test, the resolution of overlapping scenarios, and the ordering the first two depend on.

Notation. Let $\mathcal{A}$ be the accounts of the generated world and E its transfers. A transfer $e \in E$ has source $\text{src}(e) \in \mathcal{A}$, destination $\text{dst}(e)$, amount $q_e > 0$ and timestamp t_e, written $i = \text{src}(e)$ and $j = \text{dst}(e)$ where the endpoints are needed; src and dst are the same maps used in Section C.4. The ledger maintains balances only for the *tracked* accounts $\mathcal{T} \subseteq \mathcal{A}$, those inside the modelled perimeter; a destination in $\mathcal{A} \setminus \mathcal{T}$ is untracked. $\mathcal{C} \subseteq \mathcal{T}$ denotes the criminal accounts, fixed by the population layer rather than drawn at generation time. All quantities below are functions of the transfer prefix applied so far, so the ledger state is defined relative to the order fixed at the end of this section.

State. Every tracked account $a \in \mathcal{T}$ carries two balance tranches, legitimate $L(a)$ and illicit $I(a)$, with $B(a) = L(a) + I(a)$. We write $L_0(a)$, $I_0(a)$ and $B_0(a)$ for these quantities before any transfer is applied: $L_0(a)$ is the opening balance, $I_0(a) = 0$, and $B_0(a) = L_0(a)$. An account in $\mathcal{A} \setminus \mathcal{T}$ holds no tranches, and no state is kept for it.

Placement. The criminal set is $\mathcal{C}$, fixed by the population layer; the label, not a random draw over accounts, determines who launders. For each $a \in \mathcal{C}$, proceeds

$$s(a) = \max\big(B_0(a),\, \beta\big) \cdot \phi \cdot \exp(Z), \qquad Z \sim \mathcal{N}(0, \sigma^2),$$

are *reclassified* from legitimate to illicit rather than added as new funds. Here ϕ is the *proceeds share*, the fraction of an account's reference balance that criminal income represents; β is a floor on that reference balance, so that an account with little or no opening balance still receives a non-trivial base rather than none; and $\exp(Z)$ is a lognormal multiplier of unit median whose log-scale dispersion is σ, spreading proceeds across criminal accounts rather than fixing them at a common share. The release uses $\phi = 0.05$, $\beta = 1000$ EUR and $\sigma = 0.5$. The reclassification is capped at the legitimate balance actually available:

$$\delta(a) = \min\big(s(a),\, L_0(a)\big), \qquad L_0(a) \leftarrow L_0(a) - \delta(a), \qquad I_0(a) \leftarrow I_0(a) + \delta(a).$$

Reclassification leaves the total balance unchanged, so the illicit tranche stays within the spending cap the generator enforces. An account with $L_0(a) = 0$ yields $\delta(a) = 0$ and is therefore never seeded, so it carries no illicit tranche despite its label.

Propagation. A transfer e of amount q_e from $i = \mathrm{src}(e)$ to $j = \mathrm{dst}(e)$ splits illicit-first,

$$q_e^{\mathrm{illicit}} = \min\big(q_e,\, I(i)\big), \qquad q_e^{\mathrm{legit}} = q_e - q_e^{\mathrm{illicit}},$$

after which $I(i)$ and $L(i)$ are debited by those two quantities and, when $j \in \mathcal{T}$, $I(j)$ and $L(j)$ are credited by them. Three boundary cases complete the rule. If $i \notin \mathcal{T}$ or $I(i) = 0$ then $q_e^{\mathrm{illicit}} = 0$. If $j \notin \mathcal{T}$ the credit is discarded, so value leaves the modelled perimeter rather than being credited twice. If $i = j$ the transfer is a no-op. The generator also implements proportional co-mingling, $q_e^{\mathrm{illicit}} = \min\big(q_e\, I(i)/B(i),\, I(i)\big)$, but it is not used in the released data: illicit-first is the mode throughout, held constant across every scenario, flow family and feature regime. On a stream in which no illicit funds were ever seeded the two rules coincide.

Materiality. A transfer is labelled laundering when its illicit part clears both an absolute floor and a share of the amount moved,

$$y_e = \mathbb{1}\big[q_e^{\mathrm{illicit}} > \varepsilon\big] \cdot \mathbb{1}\left[\frac{q_e^{\mathrm{illicit}}}{q_e} \;\geq\; \tau\right], \qquad \varepsilon = 10^{-9}, \quad \tau = 0.25.$$

The floor ε discards floating-point dust. The share τ bounds the reach of co-mingling: without it, every account that has ever received tainted value would emit positives indefinitely, and transaction-count prevalence would cease to be controllable. Both constants are fixed before generation and are identical across every run reported here.

Collision resolution. The ledger holds one scalar illicit tranche per account, not a multiset of per-scenario tranches. When two scenarios credit the same account, their illicit amounts pool, and the next outbound transfer draws illicit-first from the pooled tranche. The consequence is deliberate: the binary label y_e remains exact and value-conserving, while the illicit *amount* carried by a downstream transfer is not attributable to a single originating scenario.

Typology labels follow the same boundary. A row carries the catalog identifier of the scenario that emitted it; a row that is laundering only because taint reached it carries the single label `natural_integration` and no catalog identifier. There is one typology label per row, never a multiset, and no dominant-taint vote. The typology label is therefore defined on scenario-emitted rows and is the sentinel elsewhere; it does not partition the positive class.

The release carries no field linking a propagated positive back to the scenario, or to the criminal account, whose value it moves: the ledger holds amounts, not origins, so that lineage is not recoverable from the exported rows. What every positive does carry is the traced amount itself, in `illicit_amount` and `legit_amount`. Provenance in this release is therefore quantitative and complete at the level of value, and identity-bearing only for the rows a scenario emitted.

Ordering. Propagation is path dependent, so the order in which transfers are applied is part of the specification. Within generation, transfers are applied in simulation-day order. When scenario rows are merged into the legitimate stream, the combined ledger is replayed in the lexicographic order induced by the key

$$\kappa(e) = \big(t_e,\ \mathrm{src}(e),\ \mathrm{dst}(e),\ \mathrm{step}(e),\ \mathrm{seq}(e)\big),$$

where $\mathrm{step}(e)$ is the simulation step that emitted e and $\mathrm{seq}(e)$ its emission index within that step. Because seq is injective, κ is a total order on E, the replay is deterministic, and the released labels are reproducible from the released seeds.

Reach of the rule. In one generation, the taint rule, not the catalog, produces most of the positive class: of $52{,}660$ positive rows, $4{,}390$ were emitted by catalog scenarios and the remaining $48{,}270$, or 91.7%, are positive through propagation alone. The share is stable across the five generation seeds used for the experiments, ranging from 84.0% to 95.1%. A detector evaluated on this release is therefore tested mainly against propagated value rather than against the scenario library.

C.4 The institution-local observation boundary

To isolate information availability, the visibility operator $\mathcal{O}_I$ restricts the records of the single world G, so every quantity below selects from G rather than resampling it.

Object	Representative content	Global world G	Institution view $\mathcal{O}_I(G)$	Dataset purpose
Party	demographics, income, KYC, PEP, media, ID	All parties and attributes	Full local customers; masked external parties	Context features and ablations
Relations	household, employment, ownership, UBO	Complete relational graph	Relations held or observed by I	Relational and ownership reasoning
Account	institution, owner, balance, risk	All accounts	Accounts at I plus observed counterparty IDs	Graph nodes and local scope
Transaction	endpoints, amount, time, channel, location	Complete ledger	Transactions touching an account at I	Tabular rows and graph edges
Provenance	source, catalog ID, tier, stage	Released evaluation metadata	Retained for scoring; hidden from features	Tier splits and scenario analysis
Laundering label	Binary indicator and tainted amount	Exact simulator state	Projected onto visible transactions	Laundering detection
Typology label	Catalog type, else `natural_integration`	Scenario-emitted rows only	Projected onto visible laundering rows	Typology classification and analysis

Table 11: Dataset fields under global and institution-local visibility.

Recorded transfers. Let A_I be the accounts held at institution I. A transfer is recorded by I when I is one of its two endpoints,

$$E_I = \big\{e \in E\ :\ \mathrm{src}(e) \in A_I\ \vee\ \mathrm{dst}(e) \in A_I\big\},$$

and the node set V_I is the set of endpoints of E_I. External counterparty accounts are therefore present in the view, but only those that transacted with a customer of I, and only through the transfers I recorded. A transfer between two accounts neither of which is at I is absent, even when both parties are visible elsewhere in the view.

What is visible on an account. Attributes fall in two tiers. Public registry facts, namely company status and the ownership and beneficial-ownership degrees, are observable for any party, because an institution can consult a company register for a counterparty it does not bank. Bank-private facts are observable only for customers of I: the behavioural aggregates computed from the recorded ledger, and the due-diligence identifiers, telephone, e-mail, national identifier, address, place of birth, device and home network, together with the number of accounts held. An external counterparty thus carries its transfer endpoints and its public registry attributes, and nothing of its due-diligence file.

Which identity edges may exist. The *homo-id* construction joins two accounts of V_I under three rules. (i) *Shared due-diligence identifier*: an edge is added when the two holders share one of the identifiers above *and both are customers of* I. Two external counterparties are never joined this way, and neither is an external counterparty to a customer. (ii) *Shared household*: the same restriction, since no public household registry exists and cohabitation is inferred from the institution's own

address records. (iii) *Ownership*: an edge is added when the two holders are linked by an ownership or beneficial-ownership relation; these are public, so this rule applies to any pair in V_I, including two external counterparties. Two constraints hold in both views: an identifier shared by more than twenty accounts is discarded, so that administrative mega-cliques do not become dense subgraphs, and the generative ring and payout relations are never used as edges, because ring membership is close to the label and would leak it.

Consequence. The institution-local view is not the global graph with nodes deleted. It is a graph whose interior is dense in due-diligence structure and whose boundary is dense in transfers but almost bare of identity. That asymmetry is what an institution actually faces, and it motivates the cross-view comparison in Table 2.

C.5 ACCOUNT-BINDING STAGES

Binding preserves the invariant that one account resolves to exactly one legal person at one institution. First, the party generator supplies entities, ownership edges, demographics, and identity-fraud labels. The topology builder independently supplies institutions and branches, cross-border corridors, the Gini-calibrated wealth marginal, counterparty-degree blueprint, merchant hubs, and account risk scores. A common bundle interface allows replacement of the population source without changing downstream invariants.

Second, accounts are opened for selected adults and private companies only. Minors, households, and public employers provide investigation context without holding accounts. Account count is configured independently of population size, with at most one account per party per institution. Third, each account inherits its holder's due-diligence record; the ledger, legitimate and illicit tranches, and scenario provenance are re-keyed onto account identifiers. One event therefore retains a single identity across tabular and graph exports. The evaluated world has five domestic institutions with four branches each.

C.6 STRUCTURAL-SIGNATURE FAMILIES

Release	Typologies T	H	N_{eff}
AML-view (catalog)	38	0.9517	31.88
AML-view (signature)	15	0.8621	10.33
SAML-D	17	0.913	13.30
IBM AMLWorld	8	**0.957**	7.32
AMLGentex	7	0.851	5.24
Tide	5	0.869	4.05
AMLNet	3	0.654	2.05

Table 12: Row-count diversity on each release's own taxonomy. AML-view catalog and signature rows use the same seed-44 emitted rows. Effective counts remain taxonomy-dependent. SynthAML declares no taxonomy.

To verify that documentary diversity also includes distinct transaction structures, Table 13 groups the 15 signatures by their account connectivity and timing. Each family is characterized by the structural quantities used to distinguish signatures.

Topology family	Structural form	Signature evidence
Long chains	Multi-stage paths across time	Path depth and temporal steps
Stars	Fan-in, fan-out, and gather-to-hub	In-/out-degree multisets
Cycles	Circular movement among accounts	Closed paths and account count
Neighborhoods	Shared-identity and mule clusters	Local connectivity and reused parties
Bipartite flows	Scatter–gather and compensation	Two-sided degree structure
Composite flows	Structuring, layering, cross-border, integration	Combined depth, degree, and timing

Table 13: Topology families represented among the $S = 15$ structural signatures.

D ENTITY-SCHEMA CROSSWALK

This appendix gives the complete crosswalk behind Table 14: the criteria, every mapping with its target IRI and the reason for the call, every non-aligned type with its justification, and the same for each comparator release. The rules below are applied identically to AML-view and every comparator.

		FtM		FIBO			Either		
Release	M	Types	IRIs	Types	IRIs	Both	Types	IRIs	Neither
AML-view	**68**	**28**	**22**	**37**	**30**	**23**	**42**	**52**	**26**
Tide	10	8	6	7	7	7	8	13	2
IBM AMLWorld	9	7	6	6	6	6	7	12	2
AMLGentex	9	7	6	6	6	6	7	12	2
SAML-D	5	3	2	2	2	2	3	4	2
AMLNet	13	5	4	5	5	4	6	9	7
SynthAML	4	1	1	0	0	0	1	1	3
All comparators	50	31	10	26	11	25	32	21	18

Table 14: Strict schema alignment against FollowTheMoney (FtM) and FIBO. M: released classes and relations; Types: mapped types; IRIs: distinct external terms, set union for comparators. Protocol in Appendix D.

Scope (S). Only schema-level constructs are mapped: classes (node types, tables) and relation types (edge types, foreign keys). No instance is mapped. An attribute is promoted to a class, and so enters the crosswalk, iff (i) it denotes a thing or a reified situation rather than a measurement of one, amounts, balances, scores, currencies and units are measurements and stay out, and either (ii-a) at least one standard declares a class whose definition covers it, or (ii-b) the release itself declares it a first-class object of its export contents (label channels, provenance, observation boundary). Limb (ii-b) admits constructs no standard covers, which is the hypothesis under test; it is applied to every release.

Equivalence (C1). `owl:equivalentClass` is asserted iff (a) the intensional definitions coincide, every member of the release class is a member of the target and conversely, with no jurisdictional, role, temporal or cardinality restriction added on either side, and (b) the property individuating instances in the release is the property the target declares as identifying. Matching labels are never sufficient.

Subsumption (C2). `rdfs:subClassOf` is asserted iff every member of the release class is a member of the target but the release class adds a restriction the target does not impose (narrower domain, added role, added jurisdiction, fixed cardinality, fixed value set). The direction is always release-class subClassOf target; the converse is never asserted.

Relations (C3, C3b). C1 and C2 applied to relations, checking the release relation's domain and range against the target's declared domain and range; because release ranges are release classes, `rdfs:subPropertyOf` is the usual outcome. FollowTheMoney reifies relations as edge *entities* (Ownership, Employment, Payment, Control) rather than as object properties, so a release relation whose FtM counterpart is an edge entity is aligned by `rdfs:subClassOf` onto that reified class. The formal crosswalk carries the axiom on a companion class `aml:`*`rel`*`Edge`, since a property cannot be `rdfs:subClassOf` a class. FIBO alignments remain property-to-property.

Non-alignment (C4). Asserted iff no class or property in either standard passes C1–C3. A thematically adjacent target does not count: if a specific IRI cannot be written, the type is non-aligned. Every non-aligned type carries a one-sentence justification (Tables 17 and 18).

Both (C5), tie-break (C7). A type may pass independently against both standards, in which case both targets are recorded. Where a standard offers a general and a specific class, the most specific class whose definition the release class entails is chosen.

Uncertainty (C6). A mapping is flagged UNCERTAIN when (i) the only available target is a value in a controlled vocabulary rather than a class; (ii) the release carries a flag whose target is a reified

record with properties the release does not populate; or (iii) two targets within one standard are equally defensible. Flags are recorded, not resolved: they are the input to an independent second annotation pass and to Cohen's κ on the bucket assignment. Rows resolved during annotation against a primary source, whether the generator code, the released data or the standard's own documentation, carry that evidence in their note and are no longer flagged.

D.1 ALIGNED CLASSES

To identify which released entity types have established meanings in the external ontologies, Table 15 gives each accepted class mapping and its justification. Uncertain mappings remain marked under rule C6.

Table 15: AML-view classes carrying an explicit mapping (26 of 39). Targets are shown by local name; Appendix 20 gives the full IRIs. The relation in parentheses is the asserted predicate. Rows marked † are flagged UNCERTAIN under C6.

Type	FollowTheMoney target	FIBO target	Basis for the call
core			
Person	Person (equivalentClass)	Person (equivalentClass)	Both definitions are 'a natural human being', with no restriction added on either side. The FtM IRI is https://schema.followthemoney.tech/#Person: the canonical repository (opensanctions/followthemoney) mints every class IRI as NS[name] in ontology.py, having dropped the per-schema rdf: override that older forks used to alias Person to foaf:Person.
Company	Company (subClassOf)	BusinessEntity (subClassOf)	subClassOf not equivalentClass on both sides: every AML-view company carries a registration identifier and a legal form (5600/5600 in the released dataset), a cardinality restriction neither target imposes.
Account	BankAccount (subClassOf)	CustomerAccount (subClassOf)	Restriction: the binding convention is one account per party per institution, so every account has exactly one holder and exactly one provider. Neither target fixes that cardinality.
Household	–	Household (subClassOf)	FtM has no co-residence unit: its Family schema is an edge between two named Persons, a kinship tie rather than a household, so it fails C1-C2. FIBO's Household is defined in the IND economic-indicators module as an SNA institutional unit; the definitional overlap with the INSEE menage is exact, but we assert subsumption rather than equivalence because the target is scoped to national-accounts aggregation. The residual doubt concerns the predicate (C1 versus C2), not the bucket assignment.
Institution	Organization (subClassOf)	Bank (equivalentClass)	FtM has no Bank class; it types banks as Organization and qualifies them with the topic value fin.bank. The Organization target is licensed by FtM's own declared range for BankAccount:bank. FIBO Bank ('depository institution that accepts deposits, makes loans...') matches exactly, including foreign institutions.
Branch	–	Branch (equivalentClass)	FtM has no branch or organizational-sub-unit concept at all; Organization would be a category error (a branch is not an organization in FtM's sense), so C4 applies on the FtM side. FIBO Branch = 'part of a larger organization that might not be co-located with it'.
Merchant	–	Merchant (subClassOf)	FtM has no Merchant class. FIBO Merchant is itself axiomatised as isClassifiedBy some MerchantCategoryCode; every AML-view merchant carries such a category, so the release class is at least as restricted as the target.
Transaction	Payment (subClassOf)	IndividualTransaction (subClassOf)	FtM Payment is an edge entity with payer/beneficiary/payerAccount/beneficiaryAccount/amount/currency/date; AML-view restricts both endpoints to accounts inside the world or to a declared external sink. FIBO IndividualTransaction = 'event that has a monetary impact and is documented in the records associated with an account'; AML-view restricts it to a directed two-account transfer with a timestamp.

continued on next page

(continued)

Type	FollowTheMoney target	FIBO target	Basis for the call
identifier			
Address	Address (equivalentClass)	PhysicalAddress (subClassOf)	FIBO ConventionalStreetAddress would be the tighter target under C7, but we do not assert that every generated address is conventionally formatted, so we take PhysicalAddress.
TelephoneNumber	–	TelephoneNumber (equivalent Class)	FtM keeps phone as a datatype property (LegalEntity:phone), not a class, so the class is unaligned on the FtM side while the relation has_phone aligns to LegalEntity:phone. The class/property asymmetry between the two standards is recorded as such.
EmailAddress	–	ElectronicMail Address (equivalentClass)	Same class/property asymmetry as TelephoneNumber: FtM has LegalEntity:email as a property only.
NationalIdentification Number	–	NationalIdentification Number (equivalentClass)	The released nir_synthetique is a synthetic INSEE-format national identification number. FtM carries Person:socialSecurityNumber as a property only.
CommuneCode	Territory (subClassOf)	–	The node's referent is the commune, and the canonical FtM repository declares Territory ('a country, subnational division or other administrative area', with a level enumeration including locality) as a class; the release restricts it to French communes. Territory's own code properties are ISO 3166, FIPS, NUTS and ZIP, none of which is an INSEE commune code, so the code itself has no property-level target. FIBO models municipalities only as government entities (MunicipalEntity), not as areas, so C4 applies on the FIBO side. Degenerate in the audited run: 20 distinct values over 70000 entities.
SIREN	–	RegistrationIdentifier (subClass Of)	SIREN is the French registration identifier of the legal unit. FtM has LegalEntity:registrationNumber as a property only.
SIRET†	–	RegistrationIdentifier (subClass Of)	(C6-iii): SIRET identifies an establishment, not the legal unit, so it and SIREN resolve to the same FIBO target under C7 while denoting different referents.
VATNumber	–	ValueAddedTaxIdentification Number (equivalentClass)	tva_intra is the French intracommunity VAT number; the FIBO definition matches exactly. FtM has LegalEntity:vatCode as a property only.
LEI	–	LegalEntityIdentifier (equivalent Class)	ISO 17442 legal entity identifier on both sides. FtM has LegalEntity:leiCode as a property only.
BankBIN	–	IssuerIdentificationNumber (sub ClassOf)	Resolved against the generator: topology/builder emits bank_bin as an ISO/IEC 7812 card issuer prefix (4... and 5...), so FIBO's IssuerIdentificationNumber ('a numbering system that allows a credit, debit, or other card to be identified as having been issued by a particular financial institution') is exact, and the servicing-oriented BankIdentifier is not. The target sits in a US-scoped FIBO module, which is a packaging artefact of FIBO rather than a semantic restriction. ftm:LegalEntity:swiftBic was rejected: a BIC is not a BIN.
ExternalAccountRef	–	AccountIdentifier (subClassOf)	FIBO defines AccountIdentifier as 'identifier that denotes an account', with no condition that the account be described in the same dataset, so the alignment holds even though the referent is withheld by design; the release restricts it to accounts at institutions other than the observing one. The masking semantics are carried by has_ext_account and RecordingBoundary, not here.
RelationshipManager	Person (subClassOf)	Employee (subClassOf)	Materialised in internal_employees.csv with role, branch and employer. Both alignments add the servicing-relationship role that neither target imposes.
PlaceOfBirth	Territory (subClassOf)	PlaceOfBirth (equivalentClass)	FIBO declares PlaceOfBirth ('physical location, including country, region, and municipality where an individual was born') as a class, matching exactly. On the FtM side the referent is an administrative area, which the canonical repository now declares as Territory; the release restricts it to the commune of birth. Degenerate in the audited run: 20 distinct values.

continued on next page

(continued)

Type	FollowTheMoney target	FIBO target	Basis for the call
promoted-attribute			
IdentityDocument	Identification (subClassOf)	IdentityDocument (subClassOf)	From id_document_type / id_document_status / id_document_country. Restriction on both sides: the type is drawn from a fixed three-value set (passport, national_id, driver_license). The verification outcome carried alongside it has no target in either standard and is covered by the KYCProfile row.
SanctionDesignation	Sanction (subClassOf)	–	Resolved against the schema: ftm:Sanction requires only the 'entity' property, so a designation carrying nothing but the sanctioned party is a well-formed Sanction and the release's flag subsumes correctly; authority, program and listingDate are optional and unpopulated. The canonical repository's Risk schema is the declared sibling for designations that are not legally binding, which delimits Sanction to the binding case. FIBO declares no sanctions concept anywhere in the 2026Q1 tree (0 hits over 2237 classes).
PoliticallyExposed Role	Occupancy (subClassOf)	–	Resolved against the standard's own documentation: OpenSanctions models a PEP as a Position (the office), an Occupancy ('a link between a Person and a Position, representing the period during which they held the office'), and the topic role.pep on the Person (opensanctions.org/docs/pep/methodology/). The release's is_pep + pep_role + pep_tier is that Occupancy with the Position left unmaterialised, which restricts rather than contradicts the target. FIBO declares no politically-exposed-person concept (0 hits).
AdverseMedia Finding†	Risk (subClassOf)	–	(C6-ii). The canonical FtM repository declares Risk, 'a risk designation attached to an entity by an authority ... a Risk flags heightened exposure or scrutiny, not a legally binding restriction', which is the closest declared class to an adverse-media finding and requires only the 'entity' property. It is taken as the target and flagged, because Risk's definition anticipates an issuing authority whereas an adverse-media hit is an unattributed screening outcome. ftm:Article and ftm:Mention were rejected: they describe documents, not screening outcomes; the topics vocabulary (crime.fin, poi, corp.shell) is value-level and cannot carry a class alignment.
BankCustomer	–	CustomerAccountHolder (subClassOf)	From is_customer / primary_bank_here / client_depuis / related_party_status. FIBO CustomerAccountHolder is the party role 'holds some CustomerAccount'; the release adds the institution-scoped restriction that the role holds only with respect to one institution. FtM has no customer role: it models what a party IS, not whose customer it is.

D.2 ALIGNED RELATIONS

To assess whether links between entities retain their meaning across schemas, Table 16 records accepted relation mappings and the justification for each equivalence or subsumption claim.

Table 16: AML-view relations carrying an explicit mapping (16 of 29). Targets are shown by local name; Appendix 20 gives the full IRIs. The relation in parentheses is the asserted predicate. Rows marked † are flagged UNCERTAIN under C6.

Type	FollowTheMoney target	FIBO target	Basis for the call
core			

continued on next page

(continued)

Type	FollowTheMoney target	FIBO target	Basis for the call
held_by	Ownership (subClassOf)	hasPrimaryAccountHolder (subPropertyOf)	C3b: the FtM counterpart is the reified edge entity Ownership (owner -¿ asset, with BankAccount subClassOf Asset), so the alignment is a class subsumption and the direction is the inverse of held_by. The FIBO property is exact in direction (Account -¿ AccountHolder) and matches the one-holder-per-account convention; the more general fibo FND/Relations/Relations/isHeldBy also holds.
at_institution	BankAccount:bank (subPropertyOf)	isProvidedBy (subPropertyOf)	ftm:BankAccount:bank has domain BankAccount and range Organization, matching direction exactly; subPropertyOf because both domain and range are narrower here. The FIBO side uses the OMG Commons property that FIBO's own Account axiom asserts (Account isProvidedBy some AccountProvider); it is referenced by FIBO but declared in Commons, which FIBO imports by IRI rather than shipping - verification status 'referenced'.
member_of	–	isMemberOf (subPropertyOf)	FIBO's Household axiom is 'hasMember some Person'; isMemberOf is the declared inverse and matches the release direction (entity -¿ household). Verification status 'referenced': declared in OMG Commons, referenced by FIBO axioms. FtM's Family is a person-to-person kinship edge and does not subsume co-residence membership.
employed_by	Employment (subClassOf)	isEmployedBy (subPropertyOf)	C3b on the FtM side (Employment is a reified edge entity: employee Person -¿ employer Organization). The release adds a validity window and a contract type.
owns	Ownership (subClassOf)	owns (subPropertyOf)	C3b on the FtM side. Both targets are direction-exact (owner -¿ owned). The release adds a percentage, an effective percentage, a role and a validity window; ftm:Ownership:percentage is the property-level counterpart of the first.
ubo	Ownership (subClassOf)	hasBeneficialOwner (subPropertyOf)	The FIBO target is exact: hasBeneficialOwner, with the BeneficialOwner and BeneficialOwnership classes in FIBO BE. On the FtM side the canonical repository's Control schema (controller/controlled/controlType, with nominee and indirect among its examples) was considered and rejected against the data: all 6373 ubo links in the audited run carry link_type 'direct_person_ownership', i.e. ownership rather than non-ownership control, so the reified Ownership edge is the correct target and ftm:Ownership:ownershipType is its property-level qualifier.
transacts_with	Payment (subClassOf)	–	C3b: FtM Payment carries payerAccount and beneficiaryAccount, so the reified edge subsumes the account-to-account projection exactly. FIBO reaches the same fact only through IndividualTransaction appliesToAccount, a transaction-to-account property whose domain does not match an account-to-account edge, so C4 applies on the FIBO side. The edge is implicit in transactions.parquet (src, dst) and materialised in graph/edges.csv.
identifier			
has_phone	LegalEntity:phone (subPropertyOf)	hasAddress (subPropertyOf)	FIBO types TelephoneNumber as a VirtualAddress subClassOf Address, so hasAddress subsumes this edge. The asymmetry with the TelephoneNumber class row follows from this: FtM aligns at the property level only, FIBO at both levels.
has_email	LegalEntity:email (subPropertyOf)	hasAddress (subPropertyOf)	ElectronicMailAddress is likewise a VirtualAddress subClassOf Address in FIBO.
has_address	Thing:addressEntity (subPropertyOf)	hasAddress (subPropertyOf)	ftm:Thing:addressEntity has domain Thing and range Address - direction and range both exact; subPropertyOf only because the release domain is narrower.

continued on next page

(continued)

Type	FollowTheMoney target	FIBO target	Basis for the call
has_nir	Person:socialSecurity Number (subPropertyOf)	isIdentifiedBy (subPropertyOf)	The FIBO side uses the OMG Commons property that FIBO's own axioms assert for identifier attachment (e.g. Account isIdentifiedBy some AccountIdentifier); verification status 'referenced' - declared in Commons, imported by FIBO by IRI.
has_siren	LegalEntity:registration Number (subPropertyOf)	isIdentifiedBy (subPropertyOf)	The FIBO side uses the OMG Commons property that FIBO's own axioms assert for identifier attachment (e.g. Account isIdentifiedBy some AccountIdentifier); verification status 'referenced' - declared in Commons, imported by FIBO by IRI.
has_siret†	LegalEntity:registration Number (subPropertyOf)	isIdentifiedBy (subPropertyOf)	UNCERTAIN: same FtM and FIBO targets as has_siren although SIRET identifies an establishment rather than the legal unit. The FIBO side uses the OMG Commons property that FIBO's own axioms assert for identifier attachment (e.g. Account isIdentifiedBy some AccountIdentifier); verification status 'referenced' - declared in Commons, imported by FIBO by IRI.
has_tva	LegalEntity:vatCode (sub PropertyOf)	isIdentifiedBy (subPropertyOf)	The FIBO side uses the OMG Commons property that FIBO's own axioms assert for identifier attachment (e.g. Account isIdentifiedBy some AccountIdentifier); verification status 'referenced' - declared in Commons, imported by FIBO by IRI.
has_lei	LegalEntity:leiCode (sub PropertyOf)	isIdentifiedBy (subPropertyOf)	Exact on the FtM side: leiCode is declared for ISO 17442. The FIBO side uses the OMG Commons property that FIBO's own axioms assert for identifier attachment (e.g. Account isIdentifiedBy some AccountIdentifier); verification status 'referenced' - declared in Commons, imported by FIBO by IRI.
has_place_of_birth	Person:birthPlace (subProperty Of)	hasPlaceOfBirth (subPropertyOf)	FIBO's Person axiom asserts hasPlaceOfBirth some PlaceOfBirth; direction and range both match the release edge exactly.

D.3 NON-ALIGNED CLASSES

To make the limits of ontology coverage explicit, Table 17 identifies released classes for which neither standard provides a justified mapping. The explanations distinguish missing concepts from rejected or uncertain candidate mappings.

Table 17: AML-view classes with no target in either standard (13 of 39). Rows marked † are flagged UNCERTAIN under C6; their notes record the competing target and the ground on which it was not asserted.

Type	Why this is a principled extension	Targets considered and rejected
core		
ExternalCounterparty	It reifies the boundary of the simulated economy as six aggregate sinks (state, customers, suppliers, three merchant sectors) so that open-system flows leaving the modelled world are explicit rather than silently dropped, and no entity ontology models the closure of an economy because both standards describe real, individually identified parties.	ftm:LegalEntity was considered and rejected: these nodes denote aggregates, not parties, so subsumption would be false.
FraudRing	The ring reifies a coordinated group of parties sharing an injected identity, address or ownership pattern, which is an annotation of the generative process rather than a legal or economic entity, and neither standard models investigative groupings because both describe entities that exist independently of any investigation.	ftm:Associate and ftm:UnknownLink are pairwise links, not groups, so neither subsumes a 17-subtype ring class.
identifier		
DeviceGeoCountry	It reifies the country from which an account's sessions are observed, a channel-telemetry fact recorded by the institution rather than a property of the party, and neither standard models the observing systems because both describe parties and instruments only.	
tx-metadata		

continued on next page

(continued)

Type	Why this is a principled extension	Targets considered and rejected
IPAddress	The session IP is channel telemetry belonging to the institution's logs, not to the party; FtM carries it only as UserAccount:ipAddress, a property of a service account the release does not model, and FIBO has no network-address class at all.	Present in transactions.csv metadata.device_info.ip_address; emitted as id_ip nodes only under specific parameters, which is off in the released seed-44 artifact.
DeviceIdentifier	A device identifier persistent per party and shared within a household is an identity-linkage signal specific to fraud and AML observability, and neither standard models the devices through which a party transacts.	transactions.csv metadata.device_info.device_id.
HomeNetwork Identifier	The shared home-network identifier makes household-level co-connection observable, which is an AML collusion signal with no counterpart in either standard.	transactions.csv metadata.device_info.home_network_id.
promoted-attribute		
KYCProfile	The KYC profile is one institution's internal due-diligence classification of a customer under FATF R.10, and it exists only inside a regulated firm's records - precisely the layer both standards exclude, since FtM models externally assertable designations and FIBO models the industry's shared business concepts rather than any single firm's risk taxonomy.	From kyc_segment, kyc_status, kyc_risk, risk_rating, customer_risk_score and id_document_status. ftm:Risk was considered and rejected: Risk is a designation made by an authority and carries authority, program and programUrl, whereas a KYC band is computed by the holding institution about its own customer and is never published. FIBO has no 'know your customer', 'due diligence' or risk-rating class in the 2026Q1 tree (0 hits for each).
observability		

D.4 NON-ALIGNED RELATIONS

To identify relationships that cannot be expressed by the accepted mappings, Table 18 records each non-aligned relation and the reason no target was asserted.

Table 18: AML-view relations with no target in either standard (13 of 29). Rows marked † are flagged UNCERTAIN under C6; their notes record the competing target and the ground on which it was not asserted.

Type	Why this is a principled extension	Targets considered and rejected
core		
at_branch	No published standard binds an account to its servicing branch: FIBO's account axioms terminate at the account provider, i.e. the institution, and FtM has no branch concept at all, yet the branch is retained because same-day activity across several branches is an ACAMS-documented structuring signal that is invisible at institution granularity.	
part_of	FIBO types a branch as an OMG Commons OrganizationalSubUnit but ships no property binding a sub-unit to its parent organization anywhere in the 2026Q1 tree, and FtM has neither the class nor the property, so the branch-to-institution edge has no target in either standard.	ftm:LegalEntity:parent was considered and rejected: its declared range is LegalEntity and a branch is not one, so subsumption would be false.
in_ring	Its range is the FraudRing class, which is itself a non-aligned generative annotation, so the edge inherits that status: an edge into a class no standard declares cannot be a subproperty of anything in either standard.	Carries fraud_role, pattern_type and shared_attribute.
pays_to†	The account-to-merchant edge is a materialised aggregation over the transaction layer, kept because a merchant shared by many otherwise-unrelated accounts is a funnel-account signal, whereas both standards bind a merchant to an individual transaction rather than to an account, so the aggregate has no target of matching domain.	UNCERTAIN (C6-iii): fibo FBC/ProductsAndServices/ClientsAndAccounts/involvesMerchant is the transaction-level analogue, and the row turns on the domain mismatch between account and transaction.
identifier		
has_commune_code	Its range is the CommuneCode class, a national statistical geography code that neither standard models, and the edge asserts calibration-time co-location rather than an identifier of the party, so no identifier property subsumes it either.	Considered and rejected: cmns Identifiers/isIdentifiedBy, because a commune code does not identify the entity.
has_bank_bin†	The edge attaches a party to the issuer identification number of its bank, so its range identifies a third object rather than the party at the domain end, and neither an identifier property (which would require the code to identify the party) nor an institution property (which would require the range to be the institution) subsumes it.	(C6-iii): cmns Identifiers/isIdentifiedBy is the competing target, defensible only if bank_bin is read as an attribute of the party's banking relationship. ftm:LegalEntity:swiftBic was rejected: a BIC is not a BIN.

continued on next page

(continued)

Type	Why this is a principled extension	Targets considered and rejected
has_relationship_manager	It binds a customer to the institution's servicing employee, an internal organizational assignment that only the institution records; both standards can express that the manager is employed by the bank, but neither has a property assigning a customer to one.	Considered and rejected: ftm:Employment and fibo isEmployedBy, whose domain is the employee and range the employer, not customer to employee.
has_device_geo	Its range is DeviceGeoCountry, channel telemetry that neither standard models, so the edge inherits the non-alignment of its range.	Degenerate in the audited run (1 distinct value).
has_id_document_country†	The release attaches the issuing country directly to the party, whereas both standards attach it to the document (ftm:Identification:country), so the domains do not match and no subproperty assertion is licensed.	(C6-iii): ftm:Identification:country is the competing target, defensible only if the party-to-country edge is read as a shortcut through the IdentityDocument class. Degenerate in the audited run (1 distinct value).
tx-metadata		
used_ip	Its range is IPAddress, which neither standard declares as a class; FtM's UserAccount:ipAddress has domain UserAccount, a service account the release does not model, so the domains do not match.	Emitted only under –with-device; present in transactions.csv metadata in the released run.
has_device	Its range is DeviceIdentifier, a construct neither standard models, and device sharing between otherwise unrelated parties is an AML-specific identity-linkage signal.	transactions.csv metadata.device_info.device_id, persistent per party.
has_home_network	Its range is HomeNetworkIdentifier, a construct neither standard models, and it exists to make household-level co-connection a traversable edge.	transactions.csv metadata.device_info.home_network_id.

D.5 COMPARATOR RELEASES

To ensure that alignment counts are assessed consistently across datasets, Table 19 applies the same mapping criteria to each comparator's published schema. This comparison distinguishes the number of exposed types from the number that can be mapped.

Table 19: Comparator crosswalks, produced by the same procedure from each release's published files. Elliptic is excluded because it publishes anonymised features and no entity schema. Rows marked † are flagged UNCERTAIN.

Type	Kind	FollowTheMoney target	FIBO target	Bucket
Tide				
Person (NodeType.INDIVIDUAL)	ent	Person (equivalentClass)	Person (equivalentClass)	both
Company (NodeType.BUSINESS)	ent	Company (subClassOf)	BusinessEntity (subClassOf)	both
Account (NodeType.ACCOUNT)	ent	BankAccount (subClassOf)	CustomerAccount (subClassOf)	both
Institution (institution_id)	ent	Organization (subClassOf)	Bank (equivalentClass)	both
Transaction (edge_type=transaction)	ent	Payment (subClassOf)	IndividualTransaction (subClassOf)	both
LaunderingLabel (is_fraudulent)	ent	–	–	none
TypologyLabel (generated_patterns)	ent	–	–	none
transacts_with	rel	Payment (subClassOf)	–	ftm
owns (edge_type=ownership)	rel	Ownership (subClassOf)	owns (subPropertyOf)	both
held_by	rel	Ownership (subClassOf)	hasPrimaryAccountHolder (subPropertyOf)	both
IBM AMLWorld				
LegalEntity (Entity ID)†	ent	LegalEntity (subClassOf)	BusinessEntity (subClassOf)	both
Account (Account Number)	ent	BankAccount (subClassOf)	CustomerAccount (subClassOf)	both
Institution (Bank ID / Bank Name)	ent	Organization (subClassOf)	Bank (equivalentClass)	both
Transaction	ent	Payment (subClassOf)	IndividualTransaction (subClassOf)	both
LaunderingLabel (Is Laundering)	ent	–	–	none
TypologyLabel (*_Patterns.txt)	ent	–	–	none
transacts_with	rel	Payment (subClassOf)	–	ftm
held_by (Account -¿ Entity ID)	rel	Ownership (subClassOf)	hasPrimaryAccountHolder (subPropertyOf)	both
at_institution (Account -¿ Bank ID)	rel	BankAccount:bank (subPropertyOf)	isProvidedBy (subPropertyOf)	both
AMLGentex				
Customer (CUSTOMER_ID)†	ent	Person (equivalentClass)	Person (equivalentClass)	both
Account (ACCOUNT_ID)	ent	BankAccount (subClassOf)	CustomerAccount (subClassOf)	both
Institution (BANK_ID / bankOrig / bankDest)	ent	Organization (subClassOf)	Bank (equivalentClass)	both
Transaction	ent	Payment (subClassOf)	IndividualTransaction (subClassOf)	both
SARLabel (IS_SAR / isSAR)	ent	–	–	none
TypologyLabel (patternID / alert Patterns)	ent	–	–	none

continued on next page

(continued)

Type	Kind	FollowTheMoney target	FIBO target	Bucket
transacts_with (src, dst)	rel	Payment (subClassOf)	–	ftm
held_by (ACCOUNT_ID -¿ CUSTOMER_ID)	rel	Ownership (subClassOf)	hasPrimaryAccountHolder (subPropertyOf)	both
at_institution	rel	BankAccount:bank (subPropertyOf)	isProvidedBy (subPropertyOf)	both
SAML-D				
Account (Sender/Receiver_account)†	ent	BankAccount (subClassOf)	CustomerAccount (subClassOf)	both
Transaction	ent	Payment (subClassOf)	IndividualTransaction (subClassOf)	both
LaunderingLabel (Is_laundering)	ent	–	–	none
TypologyLabel (Laundering_type)	ent	–	–	none
transacts_with	rel	Payment (subClassOf)	–	ftm
AMLNet				
Account (nameOrig/nameDest)	ent	BankAccount (subClassOf)	CustomerAccount (subClassOf)	both
Merchant (M-prefixed endpoints, metadata.merchant_info)	ent	–	Merchant (subClassOf)	fibo
Transaction	ent	Payment (subClassOf)	IndividualTransaction (subClassOf)	both
Address (metadata.location)	ent	Address (equivalentClass)	PhysicalAddress (subClassOf)	both
IPAddress (metadata.device_info.ip_address)	ent	–	–	none
DeviceProfile (metadata.device_info.type, os)	ent	–	–	none
LaunderingLabel (isMoney Laundering)	ent	–	–	none
FraudLabel (isFraud)	ent	–	–	none
TypologyLabel (laundering_typology)	ent	–	–	none
transacts_with	rel	Payment (subClassOf)	–	ftm
pays_to (account -¿ merchant)†	rel	–	–	none
at_location	rel	Thing:addressEntity (subPropertyOf)	hasAddress (subPropertyOf)	both
used_ip	rel	–	–	none
SynthAML				
Alert (AlertID)	ent	–	–	none
Transaction (Entry, Type, Size)†	ent	Payment (subClassOf)	–	ftm
InvestigationOutcome (Outcome)	ent	–	–	none
belongs_to_alert (transaction -¿ Alert ID)	rel	–	–	none

D.6 TARGET IRIS

To make the crosswalk independently traceable, Table 20 records the full internationalized resource identifier (IRI) of every external target used in this appendix. These identifiers specify the exact ontology terms behind the abbreviated names.

Table 20: Every third-party target used anywhere in this appendix (53 distinct IRIs), with the full IRI. Each was verified to exist in a local copy of the standard before the crosswalk was emitted; three FIBO targets are declared in OMG Commons, which FIBO imports by IRI rather than shipping.

Short name	IRI
cmns:isIdentifiedBy	`https://www.omg.org/spec/Commons/Identifiers/isIdentifiedBy`
cmns:isMemberOf	`https://www.omg.org/spec/Commons/Collections/isMemberOf`
cmns:isProvidedBy	`https://www.omg.org/spec/Commons/Organizations/isProvidedBy`
fibo:AccountIdentifier	`https://spec.edmcouncil.org/fibo/ontology/FBC/ProductsAndServices/ClientsAndAccounts/AccountIdentifier`
fibo:Bank	`https://spec.edmcouncil.org/fibo/ontology/FBC/FunctionalEntities/FinancialServicesEntities/Bank`
fibo:Branch	`https://spec.edmcouncil.org/fibo/ontology/BE/LegalEntities/FormalBusinessOrganizations/Branch`
fibo:BusinessEntity	`https://spec.edmcouncil.org/fibo/ontology/BE/LegalEntities/LegalPersons/BusinessEntity`
fibo:CustomerAccount	`https://spec.edmcouncil.org/fibo/ontology/FBC/ProductsAndServices/ClientsAndAccounts/CustomerAccount`
fibo:CustomerAccountHolder	`https://spec.edmcouncil.org/fibo/ontology/FBC/ProductsAndServices/ClientsAndAccounts/CustomerAccountHolder`
fibo:ElectronicMailAddress	`https://spec.edmcouncil.org/fibo/ontology/FND/Places/VirtualPlaces/ElectronicMailAddress`
fibo:Employee	`https://spec.edmcouncil.org/fibo/ontology/FND/Organizations/FormalOrganizations/Employee`
fibo:hasAddress	`https://spec.edmcouncil.org/fibo/ontology/FND/Places/Addresses/hasAddress`
fibo:hasBeneficialOwner	`https://spec.edmcouncil.org/fibo/ontology/BE/OwnershipAndControl/CorporateOwnership/hasBeneficialOwner`
fibo:hasPlaceOfBirth	`https://spec.edmcouncil.org/fibo/ontology/FND/AgentsAndPeople/People/hasPlaceOfBirth`

continued on next page

(continued)

Short name	IRI
fibo:hasPrimaryAccountHolder	https://spec.edmcouncil.org/fibo/ontology/FBC/ProductsAndServices/ClientsAndAccounts/hasPrimaryAccountHolder
fibo:Household	https://spec.edmcouncil.org/fibo/ontology/IND/EconomicIndicators/EconomicIndicators/Household
fibo:IdentityDocument	https://spec.edmcouncil.org/fibo/ontology/FND/AgentsAndPeople/People/IdentityDocument
fibo:IndividualTransaction	https://spec.edmcouncil.org/fibo/ontology/FBC/ProductsAndServices/ClientsAndAccounts/IndividualTransaction
fibo:isEmployedBy	https://spec.edmcouncil.org/fibo/ontology/FND/Organizations/FormalOrganizations/isEmployedBy
fibo:IssuerIdentificationNumber	https://spec.edmcouncil.org/fibo/ontology/FBC/FunctionalEntities/NorthAmericanEntities/USRegulatoryAgencies/IssuerIdentificationNumber
fibo:LegalEntityIdentifier	https://spec.edmcouncil.org/fibo/ontology/BE/LegalEntities/LEIEntities/LegalEntityIdentifier
fibo:Merchant	https://spec.edmcouncil.org/fibo/ontology/BE/FunctionalEntities/FunctionalEntities/Merchant
fibo:NationalIdentificationNumber	https://spec.edmcouncil.org/fibo/ontology/FND/AgentsAndPeople/People/NationalIdentificationNumber
fibo:owns	https://spec.edmcouncil.org/fibo/ontology/FND/OwnershipAndControl/Ownership/owns
fibo:Person	https://spec.edmcouncil.org/fibo/ontology/FND/AgentsAndPeople/People/Person
fibo:PhysicalAddress	https://spec.edmcouncil.org/fibo/ontology/FND/Places/Addresses/PhysicalAddress
fibo:PlaceOfBirth	https://spec.edmcouncil.org/fibo/ontology/FND/AgentsAndPeople/People/PlaceOfBirth
fibo:RegistrationIdentifier	https://spec.edmcouncil.org/fibo/ontology/BE/LegalEntities/CorporateBodies/RegistrationIdentifier
fibo:TelephoneNumber	https://spec.edmcouncil.org/fibo/ontology/FND/Places/VirtualPlaces/TelephoneNumber
fibo:ValueAddedTaxIdentification Number	https://spec.edmcouncil.org/fibo/ontology/BE/LegalEntities/FormalBusinessOrganizations/ValueAddedTaxIdentificationNumber
ftm:#Address	https://schema.followthemoney.tech/#Address
ftm:#BankAccount	https://schema.followthemoney.tech/#BankAccount
ftm:#BankAccount:bank	https://schema.followthemoney.tech/#BankAccount:bank
ftm:#Company	https://schema.followthemoney.tech/#Company
ftm:#Employment	https://schema.followthemoney.tech/#Employment
ftm:#Identification	https://schema.followthemoney.tech/#Identification
ftm:#LegalEntity	https://schema.followthemoney.tech/#LegalEntity
ftm:#LegalEntity:email	https://schema.followthemoney.tech/#LegalEntity:email
ftm:#LegalEntity:leiCode	https://schema.followthemoney.tech/#LegalEntity:leiCode
ftm:#LegalEntity:phone	https://schema.followthemoney.tech/#LegalEntity:phone
ftm:#LegalEntity:registrationNumber	https://schema.followthemoney.tech/#LegalEntity:registrationNumber
ftm:#LegalEntity:vatCode	https://schema.followthemoney.tech/#LegalEntity:vatCode
ftm:#Occupancy	https://schema.followthemoney.tech/#Occupancy
ftm:#Organization	https://schema.followthemoney.tech/#Organization
ftm:#Ownership	https://schema.followthemoney.tech/#Ownership
ftm:#Payment	https://schema.followthemoney.tech/#Payment
ftm:#Person	https://schema.followthemoney.tech/#Person
ftm:#Person:birthPlace	https://schema.followthemoney.tech/#Person:birthPlace
ftm:#Person:socialSecurityNumber	https://schema.followthemoney.tech/#Person:socialSecurityNumber
ftm:#Risk	https://schema.followthemoney.tech/#Risk
ftm:#Sanction	https://schema.followthemoney.tech/#Sanction
ftm:#Territory	https://schema.followthemoney.tech/#Territory
ftm:#Thing:addressEntity	https://schema.followthemoney.tech/#Thing:addressEntity

E DATA SOURCES

This appendix traces every distribution used by the Party Module of AML-view: one-dimensional margins, multi-dimensional joints fitted by Iterative Proportional Fitting (IPF), and conditional draws applied after IPF. All inputs are public statistical tables or regulatory documents. No proprietary bank or customer record enters the pipeline. The reference calibration covers metropolitan France (96 departments); overseas departments are excluded because the region-keyed targets below are published for the mainland only.

Source processing. Each public table is normalized to the categories used by the population model, checked against its published totals, and recorded in the source-to-variable tables below.

Notation. IPF dimensions are `age`, `sex`, `csp` (socio-professional category), `dipl` (highest diploma), `region` (department), `naf_work` (person work sector, INSEE NA5), and for firms `naf` (sector), `size` (headcount bracket), `legal` (legal form). A name such as `csp_x_sex` denotes the joint $P(\texttt{csp}, \texttt{sex})$.

E.1 PERSON DEMOGRAPHICS

Adults are fitted jointly over `age` × `sex` × `csp` × `dipl` × `region` × `naf_work`. Minors use a separate child registry. After a cell is sampled from the fitted joint, further attributes are drawn conditionally on that cell.

E.1.1 ONE-DIMENSIONAL MARGINS (PERSON IPF)

To specify the adult marginal distributions that IPF must reproduce, Table 21 identifies each target and its published source.

Target	Type	Published source
`age`	1-D margin	INSEE age pyramid (2381474); bands 18–24, 25–34, 35–49, 50–64, 65+
`sex`	1-D margin	INSEE Labour Force Survey EEC_PTOT01
`csp`	1-D margin	INSEE CSP in employment 2021 (6453692) + EEC activity rates
`dipl`	1-D margin	Derived from EEC_PTOT03 2024 (see 3-D joint below)
`region`	1-D margin	INSEE COG legal populations by department
`naf_work`	1-D margin	Marginal of MOBPRO CS1×NA5 (4171531)

Table 21: Person IPF: one-dimensional margins.

E.1.2 TWO-DIMENSIONAL JOINTS (PERSON IPF)

To preserve dependencies among adult attributes, IPF fits the two-dimensional joint distributions in Table 22. These constraints prevent occupation, education, geography, and work sector from being sampled independently.

Target	Type	Published source
`age_x_sex`	2-D joint	INSEE age pyramid (2381474)
`csp_x_sex`	2-D joint	INSEE 6453692, Fig. 1
`csp_x_age`	2-D joint	EEC activity rates by age, rescaled so the inactive share per age matches the survey
`dipl_x_csp`	2-D joint	EEC_PTOT04 2024 (diploma × activity × age folded onto CSP)
`region_x_csp`	2-D joint	ip1317 RP 2007 (1281090) for CSP 1–6; INSEE Première 1275 (1281261) for students (CSP 8)
`csp_x_naf_work`	2-D joint	MOBPRO RP 2016, CS1×NA5 (4171531)
`region_x_naf_work`	2-D joint	MOBPRO RP 2016, REGION×NA5

Table 22: Person IPF: two-dimensional joints.

E.1.3 THREE-DIMENSIONAL JOINT (PERSON IPF)

To preserve the joint relationship of diploma, age, and sex, the diploma structure is fitted in three dimensions when the source-table cache is present (Table 23). If absent, the engine falls back to the pair `dipl_x_sex` + `age_x_dipl` from the same EEC family.

Target	Type	Published source
`dipl_x_age_x_sex`	3-D joint	EEC_PTOT03 2024; age margin rescaled to the pyramid so it cannot fight `age`

Table 23: Person IPF: three-dimensional joint.

E.1.4 CHILD IPF MARGINS

To represent children without imposing adult occupation and diploma categories, we fit a separate set of age, sex, and geographic margins. Table 24 identifies the public targets used for these constraints.

Target	Type	Published source
`child_age`	1-D margin	INSEE Dossier complet FOR-T1 2023 + demographic balance (bands 0–5, 6–11, 12–17)
`child_sex`	1-D margin	INSEE/INED sex ratio at birth $\approx 105\,♂/100\,♀$
`child_reg`	1-D margin	INSEE COG commune population weights

Table 24: Child IPF margins (no higher-order joints).

E.1.5 POST-IPF CONDITIONAL DRAWS (PERSONS)

To assign attributes beyond the fitted demographic dimensions, we sample from public conditional distributions given the selected adult cell. Table 25 identifies the sources for these post-IPF draws, distinguishing them from constraints fitted by IPF.

Attribute / distribution	Published source
Income | CSP (salaried 3–6)	INSEE wage income 2022, FPORSOC25-F16 (8612536): median + dispersion
Income | CSP 2 (self-employed)	INSEE Non-salariés 2023 (2381330)
Income | CSP 1 (farmers)	INSEE Non-salariés agricoles 2022, ERI2025-F15 (8376567)
Income life-cycle factor | age	FPORSOC25-F16
Income index | department	INSEE salaires par territoire (50_MTS-53_SRA)
Income factor | commune	DGFiP IRCOM 2025 (taxable income by municipality)
Retiree / unemployed income anchors	INSEE 30_RPC 2024
Occupation label (PCS-4) | CSP, `naf_work`	PCS 2020 nomenclature (311 labels); n2 prevalence from EEC 2024 fichier détail (8632441)
Nationality / dual citizenship	INSEE IP2076 2024 (8651304)
Tax residency / US-person flag	CGI art. 4 A/4 B; INSEE cross-border workers 2021 (8292201); FATCA
First / last names | birth decade	INSEE prénoms (2540004); family names 2008 (3536630)
PEP role + base rate	HATVP register; decree of 17 March 2023
Household type shares	INSEE 2381514
Net wealth | age, CSP	INSEE wealth survey / RPM2024-F26 (6689022)
Household *niveau de vie* | CSP	RPM2024 F11/F12/F15/F16 (median Ensemble 23 160 €)
Housing type | commune density	INSEE Logements 2023 + density grid (8571524)
Address | commune	BAN national address database ($\propto$ dwelling capacity)

Table 25: Person-side post-IPF conditional draws.

E.2 FIRMS AND EMPLOYMENT

Firm and employment attributes determine which economic roles generated companies can occupy and which persons can work for them. The following targets calibrate firm composition, conditional financial attributes, and employment links against public references.

E.2.1 ONE-DIMENSIONAL MARGINS (COMPANY IPF)

To reproduce the marginal composition of the company population, firms are fitted over `naf` $\times$ `size` $\times$ `legal` $\times$ `region`. Table 26 lists the 1-D targets, all from SIRENE (the French statutory business register) or COG.

Target	Type	Published source
`naf`	1-D margin	SIRENE StockUniteLegale (`activitePrincipale`)
`size`	1-D margin	SIRENE StockUniteLegale (`trancheEffectifs`)
`legal`	1-D margin	SIRENE StockUniteLegale (`categorieJuridique`)
`region`	1-D margin	INSEE COG legal populations

Table 26: Company IPF: one-dimensional margins.

E.2.2 Two-dimensional joints (company IPF)

To preserve dependencies between firm characteristics, we fit two-dimensional joint distributions. Legal-unit joints come from *StockUniteLegale*; geographic joints come from *StockEtablissement*, because establishments carry commune codes (Table 27).

Target	Type	Published source
`naf_x_size`	2-D joint	SIRENE StockUniteLegale
`legal_x_size`	2-D joint	SIRENE StockUniteLegale
`naf_x_legal`	2-D joint	SIRENE StockUniteLegale (activity $\times$ legal form; e.g. SCI concentrated in real estate)
`region_x_naf`	2-D joint	SIRENE StockEtablissement
`region_x_size`	2-D joint	SIRENE StockEtablissement

Table 27: Company IPF: two-dimensional joints.

E.2.3 Firm-side conditional draws

To assign firm attributes beyond the fitted dimensions, financials and age are drawn conditionally after a firm cell is sampled (Table 28).

Attribute / distribution	Published source
Financial ratios | NAF, size	INPI/BCE financial-ratio tables (data.gouv) joined to SIRENE
Firm age / incorporation date	INSEE Première 2070 survival curve (8634190); stock age density $\propto$ survival $S(a)$
Within-section APE code | NAF	SIRENE StockUniteLegale (real activity-code frequencies)

Table 28: Firm-side conditional draws.

E.2.4 Employment and commuting

To assign employment relationships consistent with generated work sectors, person$\rightarrow$company edges reuse the person-side `naf_work` dimension and add the distributions in Table 29. An external check against INSEE ECRT2025-F7 (8376834) compares generated region$\times$sector employment to published salaried structure; this check is *not* an IPF fitting target.

Distribution	Published source
Occupation $\rightarrow$ employer section within NA5	PCS$\rightarrow$NAF affinity (EEC / MOBPRO-anchored)
Public vs. private employer | occupation, section	INSEE public employment share 19.2% (8732435) + DGAFP section shares
Home $\rightarrow$ work region flow	MOBPRO REGION$\times$REGLT

Table 29: Employment-layer distributions.

E.3 TEMPORAL DYNAMICS

To evolve the generated population over time, the event layer uses sourced life-event distributions and published rates rather than IPF joints (Table 30). They drive the one-year event layer (job changes, moves, deaths, firm creation/cessation, customer join/leave).

Distribution / rate	Published source
Mortality $q_x \mid$ age, sex	INSEE MORTA_NIV 2012–2016
Mortality $\times$ living-standard vingtile	MORTA_NIV by niveau de vie
Residential move rate / destination	INSEE residential mobility 2023 (8648157)
Retirement age / rate	Legal minimum retirement age + DREES/CNAV
Job separation / hire mix	INSEE–DARES labour-flow statistics 2025
Firm creation / cessation rates	INSEE business demography
Ownership transfer rates by age	BPCE Observatoire cession-transmission 2016
Customer join / leave rates	UFC-Que Choisir / Revue Banque banking mobility
Company phone / email change rates	B2B contact-data decay statistics

Table 30: Temporal-dynamics distributions and rates.

E.4 REGULATORY AND TYPOLOGY SOURCES

To distinguish documentary constraints from population statistics, we identify the non-statistical sources that specify legal thresholds and laundering typology shapes (Table 31). They do not contribute IPF joints.

Source	Calibrates
ACAMS (ACAMS, 2022)	Transaction-layer ACAMS-basic catalog tier (16 typologies)
CTIF (Cellule de traitement des informations financières (CTIF), 2024)	Transaction-layer CTIF-complex catalog tier (9 typologies)
FATF–Egmont; Europol EMMA	Identity/ownership fraud ring *structures* (not frequencies)
EU/FR 4th AML Directive 2015/849 art. 3(6)	Beneficial-ownership threshold (25%)
Code de commerce art. L233-7	Minor-stake threshold (5%)
CGI art. 4 A / 4 B	Tax-residency rules
HATVP; decree 17 March 2023	PEP role taxonomy and base rate

Table 31: Regulatory and typology sources.

E.5 WHAT IS NOT AN IPF-SOURCED DISTRIBUTION

Not every generator parameter is an INSEE/SIRENE IPF margin or joint. We distinguish three cases below.

Configured experimental choices. Laundering prevalence and per-typology mixing weights are chosen for learnability, and are reported as experimental settings. Fraud ring *structures* are grounded in FATF–Egmont / Europol EMMA; their *frequencies* and ring-size distributions are design choices because no public single-institution per-typology frequency exists. Individual email-change rate is a stand-in (no published individual-level rate).

Banking relationship (calibrated to a public retail footprint). Person/company customer conversion rates, customers per relationship manager, branch book sizes, and multi-banking shares are *not* IPF targets, but they are also not free inventions. They are direction-calibrated to publicly reported figures for a large French retail bank used as an anonymous reference footprint (e.g., retail customer count, professional/corporate books, branch and business-centre counts, and FBF–IFOP

multi-banking rates). No proprietary customer record is used; only published aggregate counts enter, and the modelled institution is never named in the released data.

Other model choices. Ownership knobs (shareholder-count skew, holding-parent boost, founder-is-person probability) and RM hierarchy floors are model choices, sometimes direction-anchored to public business demography but not fitted as IPF joints.

Class	Quantities
Run / stress configuration	Population size, seed, simulation window, laundering prevalence, share of rings ramping up in-window, typology mix budget
Fraud ring shape	Per-typology ring-size distributions, shared-identifier reuse probabilities, hidden-PEP probability, burst duration bounds
Ownership structure (model choice)	Shareholder-count distribution, holding-parent boost, founder-is-person probability
Banking relationship	Customer conversion rates, RM book size, branch / business-centre sizing, multi-banking, calibrated to a public anonymous retail footprint (not proprietary data; not an IPF joint)
Stand-in	Individual email-change rate

Table 32: Quantities outside the IPF sourced joints: stress choices, model choices, and footprint-calibrated banking parameters.

F Supervised Reference Model

Investigator is the supervised reference reported beside the unsupervised detectors in Section 4.3. It represents an analyst ranking transactions by suspiciousness from previously filed cases, not a structural anomaly detector, and it is deliberately the simplest model that expresses that role.

Features. A transaction is represented by the concatenation of its source-account features, its destination-account features, and its own edge features. No topology beyond the two endpoints enters the representation, so the result measures signal available from endpoint and transaction attributes under supervised training. Only released behavioural and due-diligence fields are inputs: the laundering indicator is never a feature, and the catalog identifier and simulator provenance are withheld by the export schema of Table 11. Standardisation is fitted on the training split alone.

Model. A histogram gradient-boosted tree ensemble with 300 boosting iterations, learning rate 0.06, 63 leaves, ℓ_2 penalty 1.0, and early stopping on a held-out fraction of the training split. Class weights are not adjusted, since the reported ranking metrics do not depend on a calibrated decision threshold.

Protocol. Investigator uses an intentionally enriched sample in which laundering transactions are over-represented, while all other detectors use the ordinarily sampled benchmark data. Within the enriched Investigator sample, we apply the same chronological train, validation, and test partitioning and the same global and institution-local feature definitions used elsewhere. Investigator is scored using ROC-AUC, average precision, best-threshold F_1, and precision and recall at the 1%, 5%, and 10% alert budgets. Its results characterize supervised learnability when confirmed laundering examples are more readily available; its prevalence-sensitive metrics are not directly comparable with those of detectors evaluated under ordinary sampling. The attribute ablation removes both node-feature blocks and retrains Investigator on edge features alone.

Stress modes. Two further modes are implemented in the released script. The first trains on a simulated investigation label that misses a configurable fraction of true positives and adds false alerts, while evaluation stays on generator truth, which exercises the separation between what an institution recorded and what actually happened. The second removes entire laundering typologies from the training signal and measures detection of the unseen scheme at test time.

G FULL DETECTION RESULTS

Detectors are evaluated on AML-view using transaction-level laundering labels, with results reported as mean $\pm$ std over the n seeds that completed, drawn from seeds 40–44. *Oracle* denotes the full inter-bank network, whereas *Bank-0* denotes a single institution's local view. *Static* and *Causal* refer to two feature representations. Static assigns aggregate account behavior to graph nodes, while Causal assigns ten strictly prior behavioral summaries per endpoint to each transaction edge and retains identity attributes on nodes.

All detection tables use the same generator configuration. Table 34 reports results with node attributes enabled, while 33 report alert-budget evaluations, respectively. Results are summarized across five independently seeded runs, seeds 40–44; seed 44 is the realization distributed as the public dataset. Investigator uses an intentionally enriched supervised sample, whereas the benchmark detectors retain their original supervision regimes and use the ordinarily sampled data. Investigator therefore serves as an optimistic supervised learnability reference rather than a directly comparable operational benchmark.

View / Features	Attrs	P@1%	R@1%	P@10%	R@10%
Oracle / Static	ON	0.801 ± 0.044	0.069 ± 0.035	0.732 ± 0.041	0.619 ± 0.027
Oracle / Static	OFF	0.071 ± 0.069	0.005 ± 0.004	0.074 ± 0.074	0.061 ± 0.026
Oracle / Causal	ON	0.806 ± 0.043	0.080 ± 0.021	0.644 ± 0.037	0.604 ± 0.009
Oracle / Causal	OFF	0.810 ± 0.042	0.089 ± 0.026	0.634 ± 0.037	0.593 ± 0.009
Bank-0 / Static	ON	0.800 ± 0.045	0.082 ± 0.062	0.617 ± 0.040	0.602 ± 0.026
Bank-0 / Static	OFF	0.073 ± 0.058	0.006 ± 0.004	0.072 ± 0.075	0.100 ± 0.073
Bank-0 / Causal	ON	0.806 ± 0.043	0.096 ± 0.043	0.522 ± 0.036	0.530 ± 0.011
Bank-0 / Causal	OFF	0.804 ± 0.043	0.092 ± 0.046	0.528 ± 0.036	0.539 ± 0.011

Table 33: Operating-point metrics for Investigator at the top 1% and 10% of ranked alerts, mean $\pm$ standard deviation over five seeds. Investigator is the model evaluated with explicit alert budgets. Removing node attributes removes static behavior and identity features such as the name but retains the causal behavioral edge features.

Method	Type	tx-only	homo-id	hetero
Oracle / Static				
Investigator (GBM)	sup.	0.930	–	–
SAD	semi.	0.797	0.806	–
GCN	uns.	0.746	0.766	–
GAT	uns.	0.791	0.788	–
GraphSAGE	uns.	0.798	0.833	–
RustGraph	uns.	0.880	0.874	–
TADDY	uns.	0.587	0.592	–
R-GCN	rel.	0.831	–	0.862
HGT	rel.	0.821	–	0.835
Oracle / Causal				
Investigator (GBM)	sup.	0.893	–	–
SAD	semi.	0.665	0.650	–
GCN	uns.	0.852	0.835	–
GAT	uns.	0.767	0.697	–
GraphSAGE	uns.	0.690	0.700	–
RustGraph	uns.	0.704	0.776	–
TADDY	uns.	0.739	0.699	–
R-GCN	rel.	0.711	–	0.708
HGT	rel.	0.823	–	0.824
Bank-0 / Static				
Investigator (GBM)	sup.	0.877	–	–
SAD	semi.	0.775	0.801	–
GCN	uns.	0.749	0.715	–
GAT	uns.	0.567	0.593	–
GraphSAGE	uns.	0.732	0.728	–
RustGraph	uns.	0.846	0.833	–
TADDY	uns.	0.528	0.529	–
R-GCN	rel.	0.738	–	0.757
HGT	rel.	0.848	–	0.867
Bank-0 / Causal				
Investigator (GBM)	sup.	0.831	–	–
SAD	semi.	0.632	0.651	–
GCN	uns.	0.824	0.818	–
GAT	uns.	0.671	0.702	–
GraphSAGE	uns.	0.527	0.563	–
RustGraph	uns.	0.767	0.736	–
TADDY	uns.	0.510	0.489	–
R-GCN	rel.	0.654	–	0.620
HGT	rel.	0.784	–	0.768

Table 34: Heterogeneous vs. homogeneous graph construction, mean ROC-AUC over five seeds, node attributes enabled, on the released generation. *tx-only*: transaction edges only; *homo-id*: transaction edges plus shared-identity edges, collapsed to a homogeneous graph; *hetero*: typed heterogeneous graph, consumed only by the relational models. Investigator rows report the supervised reference trained with laundering over-represented as described in Section 4.3.

Method	Type	n	AUROC		AUPRC	
			global / local	Δ_{obs}	global / local	Δ_{obs}
Static feature regime						
Investigator (GBM)	sup.	5	0.930 / 0.877	+0.053	0.739 / 0.662	+0.077
SAD	semi.	5	0.797 / 0.775	+0.022	0.022 / 0.030	−0.008
GCN	uns.	5	0.746 / 0.749	−0.003	0.042 / 0.016	+0.026
GAT	uns.	5	0.791 / 0.567	+0.224	0.050 / 0.005	+0.045
GraphSAGE	uns.	5	0.798 / 0.732	+0.066	0.089 / 0.024	+0.065
RustGraph	uns.	5	0.880 / 0.846	+0.034	0.103 / 0.056	+0.047
TADDY	uns.	5	0.587 / 0.528	+0.059	0.021 / 0.008	+0.013
R-GCN	rel.	5	0.831 / 0.738	+0.093	0.103 / 0.026	+0.076
HGT	rel.	5	0.821 / 0.848	-0.028	0.095 / 0.031	+0.065
Causal feature regime						
Investigator (GBM)	sup.	5	0.893 / 0.831	+0.062	0.642 / 0.534	+0.108
SAD	semi.	5	0.665 / 0.632	+0.033	0.013 / 0.006	+0.007
GCN	uns.	5	0.852 / 0.824	+0.028	0.057 / 0.046	+0.011
GAT	uns.	5	0.767 / 0.671	+0.096	0.020 / 0.010	+0.010
GraphSAGE	uns.	5	0.690 / 0.527	+0.163	0.023 / 0.008	+0.015
RustGraph	uns.	5	0.704 / 0.767	−0.063	0.022 / 0.018	+0.004
TADDY	uns.	5	0.739 / 0.510	+0.229	0.014 / 0.005	+0.009
R-GCN	rel.	5	0.711 / 0.654	+0.057	0.023 / 0.013	+0.010
HGT	rel.	5	0.823 / 0.784	+0.039	0.020 / 0.012	+0.008

Table 35: Observability gap for every detector on the released generation, transaction-only graph, node attributes enabled, Δ_{obs} being the global value minus the institution-local one, computed from the displayed means. n counts seeds completed in both views, out of five. The static rows extend Table 2, and the global and local AUROC columns are the transaction-only columns of Table 34. Investigator rows report the supervised reference, trained with laundering over-represented as described in Section 4.3.

H BLINDED PRACTITIONER STUDY

This appendix gives the protocol, the scoring rules and the per-practitioner results behind Section 5.

H.1 PANEL, PROTOCOL AND INSTRUMENT

Four practitioners completed the study. Three in retail or commercial banking and one in corporate and investment banking; reported AML experience is 2, 2, 3 and 10 years; reported functions cover AML investigation, SAR preparation, SAR review and filing, and compliance supervision. They are referred to as P1–P4 throughout, ordered by provenance accuracy.

Each practitioner worked through three blocks in one session of 99–167 minutes. (i) *Investigative review*: 18 cases (P3: 23), balanced six per category (real SAR-derived, suspicious synthetic, legitimate synthetic control) and six per information condition, assigned by a Latin square across practitioners. Conditions are transactions only (T), transactions plus the customer file (TC), and transactions plus the full investigation layer (TIC). For each case the practitioner gave a free-text account of what they noticed, a decision on whether the activity warranted further investigation or a suspicious-activity report (yes / no / not enough information), confidence and perceived complexity on 1–5 scales, and proposed next steps. (ii) *Provenance judgement*: 30 cases, 15 real and 15 synthetic, each judged real or synthetic with a confidence rating and an optional free-text cue. (iii) *Closing questionnaire*: 10 open questions on investigative practice and 8 feedback questions including a 1–5 rating of the set's overall representativeness.

Cases were presented from the institution-local bank view $O_I(G)$, so that the material a practitioner sees is the material the reviewing institution holds. Case identifiers, category, typology, provenance, labels and complete-world information were hidden. Real cases were rendered into the same schema as the synthetic ones, so that category is not identifiable from formatting.

H.2 INVESTIGATIVE REVIEW: SCORING RULES

The evaluation distinguishes laundering-pattern recognition from the operational escalation decision. A suspicious case is counted as recovered when the free-text response identifies the instantiated laundering behavior rather than only expressing a generic concern. A control is coded as benign when the explanation does not attribute suspicious laundering behavior to the case. One unusable response is excluded, leaving 76 scored judgements.

These two outcomes measure different constructs. Among the 51 suspicious-case judgements, 10 selected escalation, whereas 33 recovered the laundering pattern; 23 recovery hits therefore occurred without escalation. Among the 25 controls, the raw decision field contained 15 "no", 8 "not enough information", and 2 escalation responses. Under the free-text coding, none of the 25 explanations attributed a laundering pattern to the control case. We therefore report escalation and pattern recovery separately rather than interpreting the decision field as a binary detection label.

H.3 INVESTIGATIVE REVIEW: ADJUDICATED RECOGNITION

Pattern recovery is $33/51 = 0.647$ (95% Wilson CI $[0.510, 0.764]$), while all $25/25$ controls are coded benign (95% Wilson CI $[0.867, 1.000]$). These outcomes characterize recognition of the instantiated behavior and should not be interpreted as escalation sensitivity or specificity.

To expose practitioner heterogeneity, the table below reports the same adjudicated outcomes by practitioner. The previously computed Fisher tests, including pooled $p < 10^{-4}$, and Wilson intervals treat judgements as independent. They are descriptive analyses of this panel and do not adjust for repeated cases or shared practitioners; a clustered analysis is needed for population-level inference.

Practitioner	AML experience	Recovery	Benign controls	Mean rate	Fisher p
P1	2 yr	9/12	6/6	0.875	0.0090
P2	10 yr	6/11	6/6	0.773	0.0427
P3	3 yr	14/16	7/7	0.938	0.0001
P4	2 yr	4/12	6/6	0.667	0.2451

Table 36: Per-practitioner performance in the investigative review. Recovery reports the number of suspicious-case judgments for which the instantiated laundering pattern was correctly identified; benign controls reports legitimate-control judgments for which no suspicious laundering behavior was attributed. Mean rate averages the recovery and benign-control rates for each practitioner. Fisher p is the two-sided Fisher exact test against chance.

Recovery varies from $4/12$ to $14/16$, whereas every practitioner has all controls coded benign. These differences indicate variation in pattern recognition within the small panel. They do not establish that the same ordering would hold for new cases or practitioners.

H.4 Investigative review: pattern recovery and category comparison

To assess whether additional context helps practitioners identify the instantiated pattern, we compare recovery across the information conditions. Recovery increases monotonically with available investigation context: $9/18 = 0.500$ under T, $11/16 = 0.688$ under TC, and $13/17 = 0.765$ under TIC. An ordered trend test gives $r = 0.230$, $z = 1.63$, and one-sided $p = 0.052$. The observed progression is consistent with richer customer and investigation context supporting pattern reconstruction; inference is necessarily limited by the small practitioner panel.

Observed recovery rates are similar across the two suspicious categories: $15/24 = 0.625$ (95% CI $[0.427, 0.788]$) on real SAR-derived cases and $18/27 = 0.667$ (95% CI $[0.478, 0.814]$) on synthetic cases, Fisher $p = 0.778$. This test does not establish equivalence, and it does not adjust for the repeated-measures design. Of the 65 written accounts, 64 refer to information carried by the corresponding case file, so the reasoning is grounded in the presented material rather than in assumed background.

H.5 Provenance judgement

To test whether practitioners recognize synthetic provenance under matched presentation, we compare origin-classification accuracy with chance. The provenance responses are balanced, 60 on real and 60 on synthetic items, so chance is 0.5. Pooled accuracy is $55/120 = 0.458$, with recall 0.483 on real cases and 0.433 on synthetic ones; practitioners answered "real" on 63 of 120 items, indicating little aggregate preference for either response. Confidence intervals resample cases rather than judgements, since 4 raters $\times$ 30 items is repeated measurement rather than 120 independent observations; the bootstrap interval over cases is $[0.383, 0.533]$.

The pooled figure averages over substantial and opposed rater effects:

Practitioner	AML experience	Correct	Accuracy	Exact binomial p
P1	2 yr	23/30	0.767	0.005
P2	10 yr	15/30	0.500	1.00
P3	3 yr	10/30	0.333	0.099
P4	2 yr	7/30	0.233	0.005

Table 37: Per-practitioner results for the provenance-discrimination task. Each practitioner classified 30 cases, comprising 15 real and 15 synthetic cases, using a binary real-versus-synthetic judgment. Accuracy reports the proportion of correct provenance classifications. Exact binomial p tests each practitioner's accuracy against chance performance at 0.5.

The two extreme practitioners differ from chance at $p = 0.005$ in opposite directions: one classifies $23/30$ correctly and the other $7/30$. Their disagreement explains why pooled accuracy can be near chance despite informative individual behavior. The case-bootstrap interval describes this fixed panel's aggregate performance; it does not establish indistinguishability or generalize uncertainty

across a population of practitioners. Although no specific synthetic artefact was recorded in the free-text cue field, this does not exclude an unreported or differently interpreted cue.

I Data Protection

The release contains no personal data within the meaning of Article 4(1) of the General Data Protection Regulation: no record derives from, or refers to, an identified or identifiable natural person. Persons, households and firms are fitted to public aggregate statistics: INSEE census and labour-force tables and the SIRENE business register, listed in Section E, which are processed into the calibration targets listed in that appendix. No customer file, transaction log or other micro-data enters the generator; the single bank-derived input is a set of published aggregate footprint figures, and the institution it describes is never named in the released data.

Identity attributes are assembled from public reference lists rather than copied from records. First names are drawn from the INSEE register of given names conditioned on sex and birth decade, surnames from the INSEE frequency list, street addresses from the Base Adresse Nationale, and telephone numbers from the ARCEP numbering plan. The components are real; their combinations are not. A generated street exists while its occupants do not, national identifiers are format-valid synthetic strings, and a generated name coinciding with a living person's carries no information about that person.

The practitioner study is the only component in contact with real records. The suspicious-activity reports used as comparison cases remain inside the environment of the institution that filed them and are examined by its own authorised staff. No case content is distributed with the dataset or reproduced here. Study endpoints are reported in aggregate and per practitioner under a pseudonym; the only practitioner free text quoted or released concerns synthetic cases.

J Demographic Bias

The release carries no explicit race or ethnicity attribute. This does not establish the absence of demographic proxies in names, geography, or other fields. We do not use the dataset to draw conclusions about racial disparity.

Sex is carried, and laundering roles are assigned without reference to it. On the seed-42 benchmark run (64,400 persons, 211 flagged), the flag rate is 31.1 per $10{,}000$ for women against 34.6 for men. A Pearson χ^2 test of independence detects no association at the 5% level, $\chi^2(1) = 0.58$ against a critical value of 3.84.

This check describes one generated sample and does not establish independence across all attributes, scenarios, or seeds. Real alert populations carry demographic disparities that this dataset does not reproduce, so it cannot be used to audit a deployed system for disparate impact.

Acknowledgements

We thank the anti-money-laundering practitioners who participated in the blinded evaluation, as well as the compliance investigators and typology specialists who contributed to the scenario catalogue and export requirements, at a large European banking group to be named in the camera-ready version.