Synthetic Data Use Approval for Banks: Utility, Privacy, Memorisation and Model Risk

An Evidence-Gated Framework for Financial-Institution Generation, Validation, Release and Monitoring

Authored by Chennakeshav Adya, Independent Researcher

September 2026

Abstract

Banks increasingly use synthetic tabular, time-series and event data for development, fraud analytics, testing, collaboration and model validation. A generated dataset can reduce direct exposure to customer records while retaining enough structure to support useful work. The same process can reproduce distinctive records, leak membership, enable attribute inference, smooth away rare events, distort vulnerable groups or create downstream decisions that diverge from reality. This paper develops an evidence-gated approval framework for synthetic-data use in banks. Forty modules connect purpose, method, source governance, population, threat model, utility, fidelity, reproducibility, statistical structure, business rules, downstream model performance, decision stability, subgroup and rare-event performance, temporal fidelity, causal limitations, matching, membership and attribute inference, memorisation, differential privacy, the privacy-utility frontier, fairness, semantics, generator risk, access, release tiers, third-party sharing, model-risk management, independent review, limitations, stress testing, thresholds, approval, monitoring, revalidation and retirement. Five figures, five tables, eight frequently asked questions and twenty-six primary or authoritative references support use-, data-, model-, recipient-, jurisdiction- and period-specific review. The framework treats privacy, fidelity and decision utility as separate claims and requires reproducible evidence for each. It does not establish that any dataset is anonymous, legally compliant, fair or fit for a particular banking decision and does not substitute for authorised privacy, legal, regulatory, model-risk, cybersecurity or banking advice.

JEL Classification: C88, G21, G28, C45, C53, K24, O32

Keywords: synthetic data, banking, privacy, utility, memorisation, membership inference, differential privacy, model risk, validation, data governance

1. Define the proposed use

State the banking decision, data population, users, downstream model, consequence and permitted purpose.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an approved synthetic-data use charter. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when a generic synthetic-data claim substitutes for a bounded use case. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for define the proposed use should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

2. Classify the synthetic method

Record whether data are fully or partially synthetic, model-based, simulation-based, differentially private or generated through another controlled method.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a method classification. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when different generation methods are treated as carrying the same utility and privacy properties. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for classify the synthetic method should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

3. Map the source-data lifecycle

Trace collection, legal basis, consent, minimisation, quality, access, retention, transformation and deletion.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a source-data control map. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthesis is used to bypass unresolved source-data obligations. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for map the source-data lifecycle should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

4. Define the target population

Specify entities, events, time horizon, geography, products, channels and rare cases the output must represent.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a population specification. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when an average distribution conceals excluded customers and material edge cases. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

The decision pack for define the target population should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

5. Set the threat model

Name adversaries, auxiliary information, access channels, attack budgets and protected attributes.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an approved privacy threat model. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when privacy is asserted without defining who may attack what. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for set the threat model should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

6. Set the utility objective

Tie statistical fidelity to a specific analytical, testing, training or sharing decision.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a task-specific utility specification. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when visual similarity is mistaken for decision usefulness. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for set the utility objective should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

7. Separate fidelity from utility

Measure structural resemblance independently from downstream performance.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a fidelity-to-utility map. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when high univariate similarity hides broken relationships and unusable outcomes. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for separate fidelity from utility should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

8. Establish a real-data benchmark

Freeze representative training, validation and holdout samples under controlled access.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an independent benchmark design. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when the synthetic generator and evaluator reuse the same evidence. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

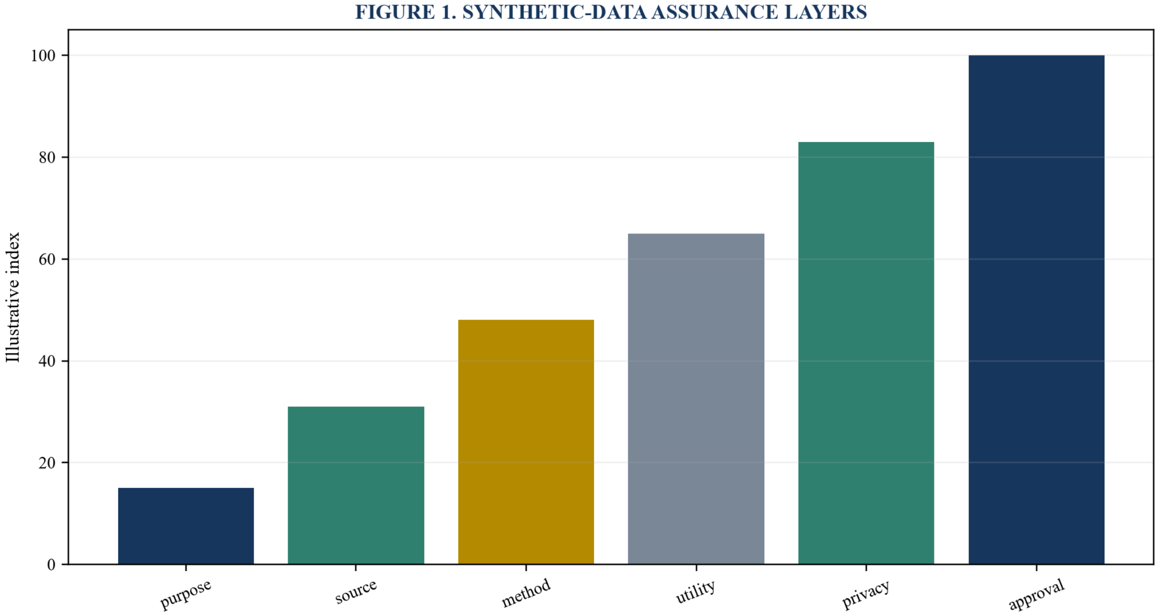
The decision pack for establish a real-data benchmark should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 1. Synthetic-data evidence layers

Layer	Primary question	Required evidence
source	lawful, representative and controlled	source-data audit
generation	reproducible and secure	versioned lineage
utility	fit for the intended decision	task benchmark
privacy	resistant to defined attacks	adversarial test

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 1. Synthetic-data assurance layers



Values are illustrative indices and require replacement with approved validation evidence.

9. Preserve lineage and reproducibility

Record source snapshot, generator, parameters, random seeds, code, environment and approvals.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a generation lineage record. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when no reviewer can reproduce the synthetic release or its evaluation. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for preserve lineage and reproducibility should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

10. Measure marginal distributions

Compare ranges, categories, missingness, moments, tails and rare values.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a marginal-fidelity scorecard. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when headline averages hide implausible or missing sub-populations. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for measure marginal distributions should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

11. Measure joint structure

Test correlations, conditional distributions, temporal order, networks and business constraints.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a multivariate fidelity assessment. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when records look plausible individually while relationships are wrong. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for measure joint structure should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

12. Test business rules

Validate balances, chronology, accounting identities, product logic and impossible combinations.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a domain-constraint test pack. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when statistically likely rows violate banking mechanics. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

The decision pack for test business rules should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

13. Test downstream model performance

Train and test relevant models across real, synthetic and mixed data regimes.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a downstream validation matrix. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when generic quality scores substitute for the intended model outcome. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for test downstream model performance should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

14. Test decision stability

Compare rankings, classifications, thresholds, actions and confidence across data regimes.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a decision-concordance analysis. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when aggregate accuracy hides changed customer or transaction decisions. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for test decision stability should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

15. Measure subgroup utility

Evaluate protected, vulnerable, sparse, high-risk and commercially material segments separately.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a segmented utility dashboard. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic data improve the majority while degrading small groups. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for measure subgroup utility should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

16. Test rare-event coverage

Measure representation and learnability for fraud, default, operational loss and unusual market conditions.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a rare-event validation pack. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when generation smooths away the events the bank most needs to detect. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

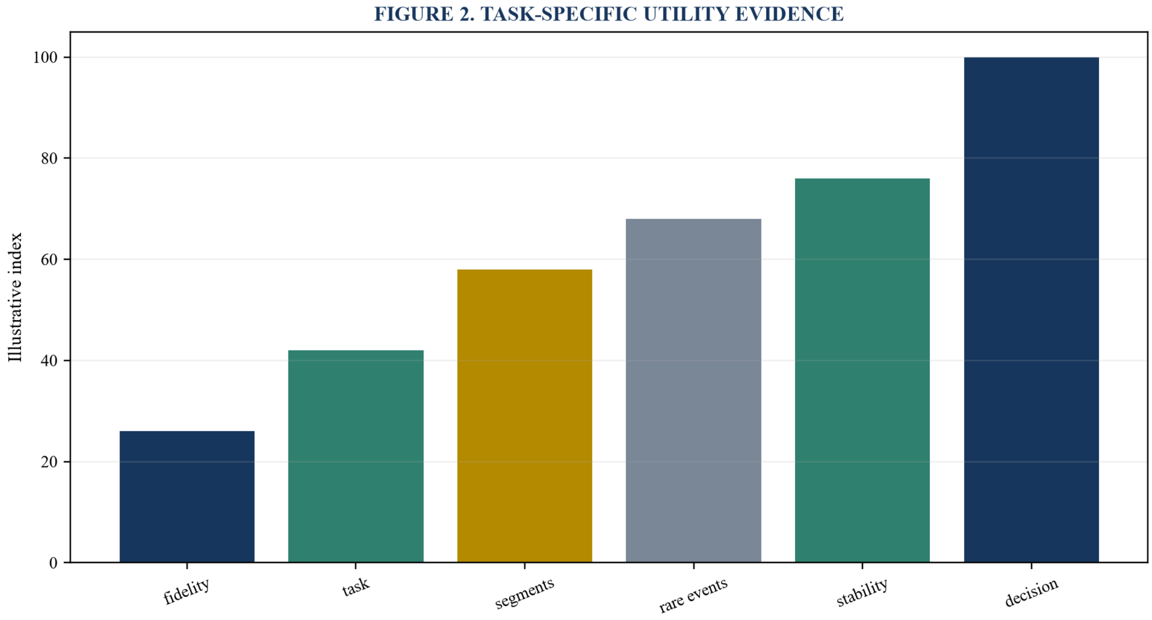
The decision pack for test rare-event coverage should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 2. Utility validation matrix

Dimension	Test	Decision evidence
fidelity	marginal and joint structure	segmented comparison
task utility	downstream model and action	baseline delta
rare events	coverage and learnability	event-level results
stability	seeds, periods and regimes	variance distribution

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 2. Task-specific utility evidence



Values are illustrative indices and require replacement with approved validation evidence.

17. Test temporal fidelity

Validate seasonality, autocorrelation, regime change, event timing and sequence length.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a temporal benchmark. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when time-series output preserves distributions while destroying dynamics. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for test temporal fidelity should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

18. Run nearest-neighbour tests

Measure proximity between synthetic and source records across relevant feature spaces.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a distance-based privacy assessment. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic rows reproduce distinctive customer records. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for run nearest-neighbour tests should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

19. Test exact and approximate matches

Search for duplicated uniques, partial matches and high-risk quasi-identifier combinations.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a match-risk report. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when absence of exact duplication is treated as proof of privacy. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for test exact and approximate matches should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

20. Run membership-inference attacks

Test whether an attacker can determine if a person or event was in training data.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a membership-risk evaluation. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when generator confidence or overfitting reveals participation. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

The decision pack for run membership-inference attacks should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

21. Run attribute-inference attacks

Test whether known fields enable sensitive unknown attributes to be recovered.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an attribute-disclosure assessment. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic release strengthens prediction of protected facts. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for run attribute-inference attacks should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

22. Test memorisation and extraction

Probe models and outputs for retained sequences, rare combinations and recoverable records.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a memorisation test pack. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when generative systems emit training examples under targeted prompts or sampling. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for test memorisation and extraction should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

23. Evaluate differential privacy

Where claimed, verify adjacency, unit of privacy, epsilon, delta, composition and implementation.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a differential-privacy assurance record. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when a privacy budget is quoted without a valid mechanism or threat model. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for evaluate differential privacy should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

24. Map the privacy-utility frontier

Quantify how protection settings affect task performance and subgroup outcomes.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an approved frontier analysis. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when one operating point is selected without showing the trade-off. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

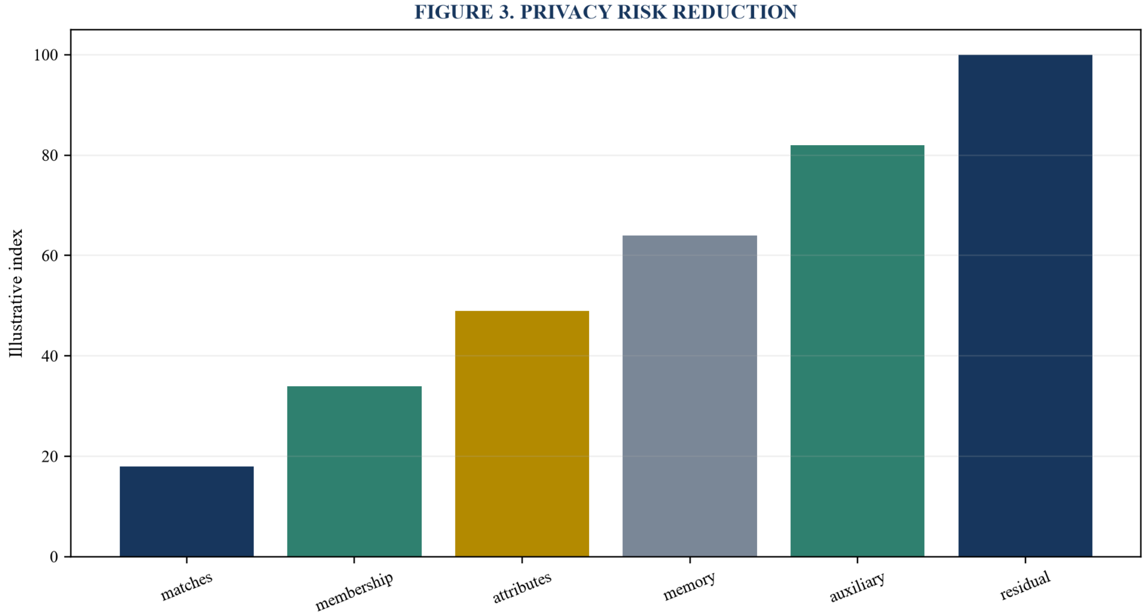
The decision pack for map the privacy-utility frontier should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 3. Privacy attack matrix

Risk	Test	Approval evidence
record reproduction	exact and approximate match	risk distribution
membership	black- and white-box attack	attack advantage
attribute disclosure	auxiliary-information attack	conditional risk
memorisation	targeted extraction probes	exposure result

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 3. Privacy risk reduction



Values are illustrative indices and require replacement with approved validation evidence.

25. Test fairness and representativeness

Compare errors, distributions, decisions and exclusions across relevant groups.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a fairness impact assessment. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic balancing creates artificial fairness or masks source inequity. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for test fairness and representativeness should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

26. Validate schema and semantics

Check types, definitions, units, identifiers, ontologies and versioned contracts.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a data-contract validation. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic files pass statistical tests but break consuming systems. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for validate schema and semantics should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

27. Control code and generator risk

Review implementation, dependencies, security, licensing, vulnerabilities and change control.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a generator assurance file. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when an opaque package introduces operational or supply-chain risk. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for control code and generator risk should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

28. Control access and environments

Separate source, generation, evaluation, approval and release zones by role and purpose.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an environment and access design. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic outputs create a new route back to restricted source evidence. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

The decision pack for control access and environments should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

29. Set release tiers

Define internal testing, controlled collaboration, regulator sharing and public release requirements.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a tiered release policy. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when one approval standard is applied to materially different audiences. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for set release tiers should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

30. Govern third-party sharing

Set contracts, onward-use limits, retention, deletion, incident duties and audit rights.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a synthetic-data sharing schedule. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when recipients combine output with auxiliary data beyond the assessed threat model. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for govern third-party sharing should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

31. Integrate model risk management

Treat generators and downstream models according to materiality, validation, change and monitoring rules.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a model-risk classification. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic data are treated as outside model governance. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for integrate model risk management should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

32. Validate independent review

Separate development, validation, privacy testing, business approval and audit roles.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an independent challenge record. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when the generation team approves its own claims. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

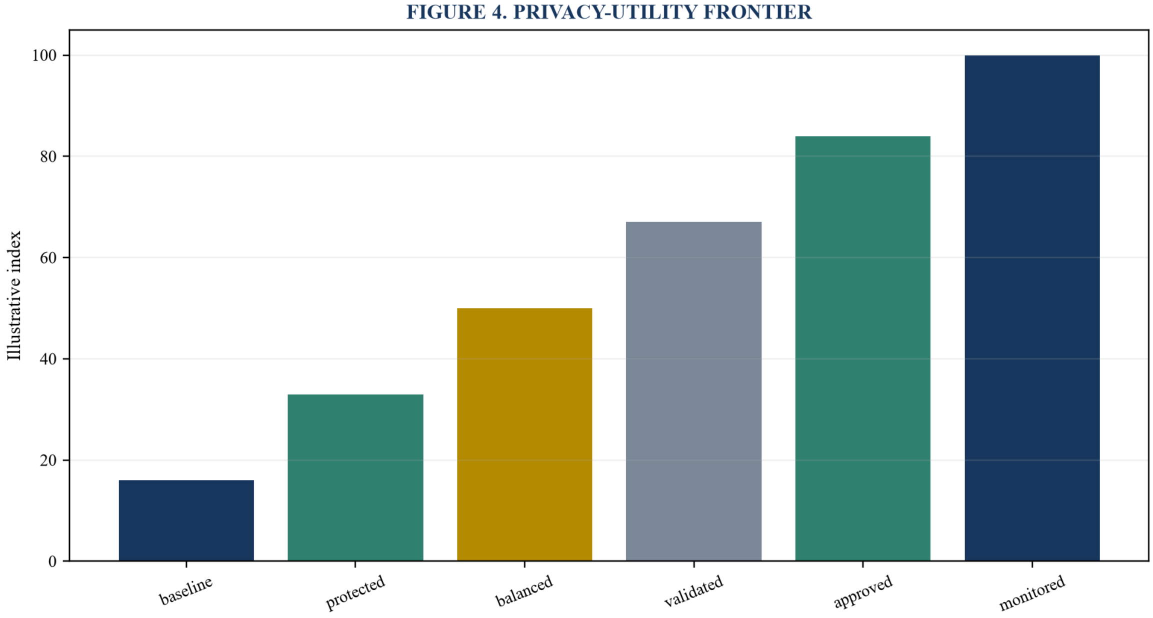
The decision pack for validate independent review should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 4. Release tiers

Tier	Permitted use	Control standard
development	internal testing	restricted environment
controlled	approved collaborator	contract and monitoring
regulatory	defined supervisory purpose	secure transfer
public	open access	strongest privacy evidence

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 4. Privacy-utility frontier



Values are illustrative indices and require replacement with approved validation evidence.

33. Document limitations

State excluded populations, weak metrics, privacy assumptions, permissible decisions and expiry.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a limitation and use statement. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when users treat the synthetic dataset as a general substitute for reality. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for document limitations should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

34. Run scenario and stress testing

Test drift, sparse data, attacks, parameter changes, model updates and adverse regimes.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a sensitivity and stress pack. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when approval reflects one favourable dataset and configuration. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for run scenario and stress testing should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

35. Define acceptance thresholds

Set metric-specific pass, escalation and rejection rules before evaluation.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a pre-registered acceptance matrix. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when teams select favourable measures after seeing results. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for define acceptance thresholds should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

36. Make the approval decision

Record evidence, exceptions, residual risk, conditions, accountable owners and expiry.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a signed use-approval memorandum. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when governance records activity without a clear use decision. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

The decision pack for make the approval decision should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

37. Monitor production use

Track drift, privacy attacks, utility decay, subgroup effects, incidents and unauthorised reuse.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a monitoring dashboard. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when a one-time assessment survives changing data and threats. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Synthetic data require use-specific evidence. Statistical resemblance, privacy protection and downstream decision quality are distinct claims that need separate tests.

The decision pack for monitor production use should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

38. Trigger revalidation

Require review after material source, generator, parameter, purpose, recipient or regulatory change.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a revalidation trigger register. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when a changed release inherits an obsolete approval. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Bank approval should connect lawful source data, a defined threat model, reproducible generation, independent validation, controlled release and accountable monitoring.

The decision pack for trigger revalidation should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

39. Retire and delete

Expire datasets, derived artefacts, checkpoints, caches, recipient copies and permissions.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a verified retirement record. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when obsolete synthetic releases remain available without valid purpose. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Privacy risk includes exact reproduction, approximate matches, membership inference, attribute inference, memorisation and combination with auxiliary information.

The decision pack for retire and delete should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

40. Close through approval gates

Require bounded purpose, lawful source, measured utility, tested privacy, downstream validation, controlled release and monitoring.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an evidence-gated use decision. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when synthetic appearance is mistaken for safe and decision-useful data. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Utility should be measured on the intended banking task, relevant segments, rare events and operating thresholds; aggregate similarity cannot establish fitness for use.

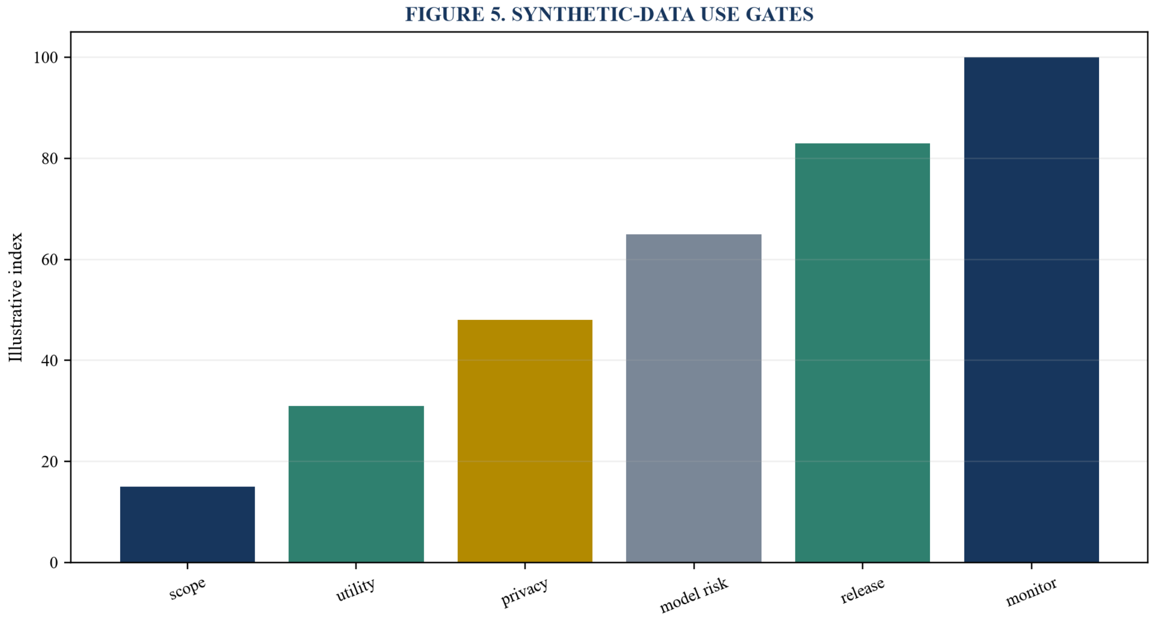
The decision pack for close through approval gates should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 5. Synthetic-data approval gates

Gate	Primary work	Completion evidence
purpose	use, population and consequence	approved charter
utility	task and segment validation	threshold results
privacy	threat model and attacks	residual-risk decision
release	access, contract and monitoring	accountable approval

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 5. Synthetic-data use gates



Values are illustrative indices and require replacement with approved validation evidence.

References

[1] Financial Conduct Authority, Using Synthetic Data in Financial Services, https://www.fca.org.uk/publications/corporate-documents/report-using-synthetic-data-financial-services

[2] Financial Conduct Authority, Exploring Synthetic Data Validation: Privacy, Utility and Fidelity, https://www.fca.org.uk/publications/research-articles/exploring-synthetic-data-validation-privacy-utility-fidelity

[3] Financial Conduct Authority, Synthetic Data and Anti-Money Laundering Project Report, https://www.fca.org.uk/publications/research-notes/research-note-synthetic-data-anti-money-laundering-project-report

[4] Financial Conduct Authority, Feedback Statement FS23/1 on Synthetic Data, https://www.fca.org.uk/publications/feedback-statements/fs23-1-feedback-statement-synthetic-data-call-for-input

[5] Financial Conduct Authority, Authorised Push Payment Synthetic Data, https://www.fca.org.uk/firms/digital-sandbox/authorised-push-payment-synthetic-data

[6] Financial Conduct Authority, Supercharged Sandbox, https://www.fca.org.uk/firms/innovation/supercharged-sandbox

[7] National Institute of Standards and Technology, SP 800-226 Guidelines for Evaluating Differential Privacy Guarantees, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-226.pdf

[8] National Institute of Standards and Technology, Differentially Private Synthetic Data, https://www.nist.gov/blogs/cybersecurity-insights/differentially-private-synthetic-data

[9] National Institute of Standards and Technology, SDNist Synthetic Data Report Tool, https://www.nist.gov/services-resources/software/sdnist-synthetic-data-report-tool

[10] National Institute of Standards and Technology, Synthetic Data Test Drive, https://pages.nist.gov/HLG-MOS_Synthetic_Data_Test_Drive/guide.html

[11] Information Commissioner's Office, Anonymisation and Pseudonymisation Guidance, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/anonymisation/

[12] Information Commissioner's Office, Guidance on AI and Data Protection, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/

[13] Central Bank of the UAE, Data Quality, Privacy and Security for AI and ML, https://rulebook.centralbank.ae/en/rulebook/5-data-quality-privacy-and-security

[14] Central Bank of the UAE, Guidance Note on Responsible Adoption and Use of Artificial Intelligence, https://rulebook.centralbank.ae/en/rulebook/guidance-note-consumer-protection-and-responsible-adoption-and-use-artificial-intelligence

[15] Central Bank of the UAE, Model Management Standards, https://rulebook.centralbank.ae/en/rulebook/model-management-standards

[16] Prudential Regulation Authority, SS1/23 Model Risk Management Principles for Banks, https://www.bankofengland.co.uk/prudential-regulation/publication/2023/may/model-risk-management-principles-for-banks-supervisory-statement

[17] Basel Committee on Banking Supervision, Principles for Effective Risk Data Aggregation and Risk Reporting, https://www.bis.org/publ/bcbs239.htm

[18] European Union, General Data Protection Regulation, https://eur-lex.europa.eu/eli/reg/2016/679/oj

[19] UAE Government, Federal Decree-Law No. 45 of 2021 on Personal Data Protection, https://uaelegislation.gov.ae/en/legislations/1972

[20] National Institute of Standards and Technology, AI Risk Management Framework 1.0, https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10

[21] National Institute of Standards and Technology, Privacy Framework, https://www.nist.gov/privacy-framework

[22] International Organization for Standardization, ISO/IEC 42001 AI Management Systems, https://www.iso.org/standard/81230.html

[23] International Organization for Standardization, ISO/IEC 23894 AI Risk Management, https://www.iso.org/standard/77304.html

[24] International Organization for Standardization, ISO/IEC 27001 Information Security Management Systems, https://www.iso.org/standard/27001

[25] Organisation for Economic Co-operation and Development, OECD AI Principles, https://oecd.ai/en/ai-principles

[26] Bank for International Settlements, Governance of AI Adoption in Central Banks, https://www.bis.org/publ/othp90.htm

About the Author

Authored by Chennakeshav Adya, Independent Researcher

Chennakeshav (CK) is a corporate finance and investment banking executive with 25+ years of global experience in deal origination, structuring and execution across M&A, growth capital and corporate strategy. He has led value-creation mandates for founders, corporates and funds — bridging the boardroom view to hands-on execution and close.

His career spans Morgan Stanley, HSBC, Lloyds Banking Group, EWEC, ADQ portfolio companies and Emirates Growth Fund, across TMT, real estate, fintech, deeptech, cleantech, infrastructure and energy. He has partnered with C-suite leaders, private equity and venture funds, sovereign wealth funds and family offices to finance complex fund raises and scale-up ventures, and has led M&A due diligence, post-merger integration and business-transformation initiatives to create value.

At Matchpoint Partners he is Managing Partner, leading the firm's corporate finance, M&A and capital-raising practice. He holds an MBA from London Business School, an engineering degree from VTU and a Master of Laws (LLM, in progress) from UCL London.

An active start-up mentor, CK mentors at Techstars, DIFC FinTech Hive, Startup Grind, Founder Institute and IN5, serves as Entrepreneur Mentor in Residence (EMiR) at London Business School, and judges the Entrepreneurship World Cup.