

PRODUCT 06



WBG IDA Financial Control Tower

REQUIREMENT DOCUMENT



Contents

00	Document Control
01	Executive Summary
02	Context, Problem & Product Strategy
03	Users, Stakeholders & Governance
04	Scope, Release Model & User Journeys
05	Functional Requirements
06	Business Rules & Decision Logic
07	Data Requirements & Contracts
08	Analytics, Model & Responsible-AI Requirements
09	Non-Functional Requirements
10	Architecture & Integration Requirements
11	Experience, Content & Accessibility Specification
12	Acceptance, Verification & Release Readiness
13	Risks, Assumptions, Dependencies & Open Decisions
14	Requirements Traceability Matrix
A	Appendices

00 Document Control

Attribute	Value
Document	WBG IDA Financial Control Tower: Product Requirements Document
Product identifier	PRODUCT 06
Version	1.0
Status	Retrospective PRD · Portfolio release
Owner	Nicolaas Heru Dreandachrista
Prepared	20 August 2026
Source basis	Working repository, pipeline artifacts, model card, data-quality report, review workflow, tests, and Azure target design
Data classification	Public data only; no internal World Bank Group systems or non-public financial records
Product URL	ida-financial-data-control-tower.vercel.app
Repository	github.com/nicolaasheru/ida-financial-data-control-tower

Requirement writing convention

1. Each requirement uses a stable identifier and one independently verifiable obligation: condition or trigger, responsible system, shall, observable response and applicable constraint.
2. Acceptance criteria use Given / When / Then to identify the test setup, action and expected result. These criteria define verification; they do not claim a test has been executed.
3. Priority records release importance. Must and Should are metadata; every selected requirement uses shall in its statement.
4. Implemented describes observed local behavior. Target describes required future behavior. Mixed status is separated into individual requirements.
5. Reference-model values describe the current configuration. Unapproved production workload, retry, RPO and RTO values remain explicit release-blocking decisions.

Approval model

Current document	Production release
Owner-approved portfolio specification	Joint approval: Product, Financial Data, Engineering, Model Governance, Security, and Operations

01 Executive Summary

Product definition

Input	Control layer	Output
Public IDA commitments and gross-disbursement records	Structural, financial, statistical, and multivariate detection	Prioritized, explainable alerts with controlled analyst review

The decision it supports

MVP QUESTION. How can a lightweight control layer detect missing, inconsistent, and unusual financial records before they propagate into downstream reporting?

Current evidence

Measure	Current artifact
Source scope	3,420 IDA records · 202 named entities · 14 reporting periods
Current operating point	1% Isolation Forest contamination, selected provisionally
Detector output	145 signals consolidated into 118 alerts
Priority mix	4 critical · 17 high · 97 medium
Corroboration	5 alerts supported by multiple detector families
Review evidence	30 sampled alerts reviewed; 4 resolved legitimate exceptions; 26 need more information
Controlled tests	5 failure modes × 5 seeded target selections; 25/25 expected behaviors detected

Product boundary

The product may	The product may not
Surface unusual records for investigation	Declare error, fraud, misconduct, policy breach, or financial loss; alter source records; make funding decisions

Current versus target

Capability	Current prototype	Governed production target
Ingestion	Python, public REST API, paginated extract	Scheduled orchestration, retry, manifest, quarantine, atomic publication
Storage	Versioned local files and generated artifacts	Immutable raw, standardized, curated, and quarantine zones
Detection	Python rules, statistical controls, segmented Isolation Forest	Versioned batch jobs with governed model registry
Review	FastAPI + SQLite; read-only fallback	Authenticated API + enterprise SQL + append-only audit
Interface	Analyst dashboard	Authorized hosted dashboard with operational telemetry
Security	Local configuration; no production authentication	Entra ID, managed identity, Key Vault, private endpoints, least privilege

02 Context, Problem & Product Strategy

Problem statement

Current gap	Product response
Manual review does not consistently connect ingestion, financial controls, historical context, anomaly evidence, and disposition across 3,420 records.	One traceable workflow from source validation to explainable alert and analyst decision.

Jobs to be done

Actor	When...	I need to...	So that...
Financial-data analyst	a new snapshot is published	see a prioritized, explainable queue with financial context	I can investigate material exceptions without scanning the entire dataset
Data engineer	an extraction or schema changes	identify incomplete pages, invalid fields, duplicate grain, and failed gates	bad data cannot silently replace the last known-good publication
Model owner	detector behavior changes	compare segments, features, alert load, controlled tests, and review outcomes	I can approve, recalibrate, or roll back responsibly
Platform operator	a pipeline or review service fails	observe the failure, preserve evidence, and recover independently	review history and active reporting remain trustworthy

Product goals

1. Detect structural, financial, historical, and multivariate anomalies before downstream consumption.
2. Convert detector output into a reviewable queue with reason codes, evidence, materiality, corroboration, and recommended actions.
3. Preserve source-to-alert-to-review traceability and prevent silent overwrite of analyst decisions.
4. Make uncertainty and limitations visible, particularly where public data lacks authoritative labels.
5. Demonstrate a credible migration path from a local prototype to a governed Azure workload.

Non-goals

1. Replacing the WBG source system, official reconciliation, or domain approval.
2. Classifying unusual records as fraud, error, or policy breach without human evidence.
3. Changing, suppressing, or publishing source financial records automatically.
4. Ranking countries, determining financing eligibility, or influencing funding decisions.
5. Claiming production deployment, internal WBG access, or statistically validated real-world precision.

Product principles

Principle	Implication
Rules before opacity	Deterministic controls remain visible and retain precedence over an ML-only score.
Context before conclusion	Every alert shows amounts, history, evidence, and limitations before a reviewer acts.
Human authority	Only an accountable reviewer can assign a final outcome.
Last-known-good by default	Failed or incomplete runs never replace an approved publication.
Immutable evidence	Source snapshots, model/run metadata, decisions, and audit events are retained separately.
Proportionate architecture	Scale and cloud services are adopted when justified; Spark is not a default dependency.

03 Users, Stakeholders & Governance

Primary personas

Persona	Goals	Pain points	Core permissions
Financial-data analyst	Triage alerts; determine whether unusual activity is valid, incorrect, or unresolved	Opaque scores; missing historical context; fragmented evidence	Read queue and evidence; begin review; save notes; resolve/reopen within policy
Data engineer	Maintain reliable ingestion, schema integrity, lineage, and reproducibility	API drift; partial pagination; silent type coercion; unstable identifiers	Operate pipeline; inspect quarantine; rerun failed jobs; no authority to decide financial meaning
Model owner	Evaluate sensitivity, drift, segment behavior, and review load	No labels; changing peer distributions; unstable thresholds	View model evidence; propose calibration; register/promote/roll back with approval
Platform operator	Keep service healthy and recoverable	Coupled releases; missing telemetry; unsafe manual repairs	Monitor, deploy, restore, and roll back; no ability to rewrite review evidence

Stakeholder map

Stakeholder	Accountability	Decision rights
Product owner	Scope, outcomes, prioritization	Accept product behavior and release scope
Financial data owner	Meaning, materiality, control policy	Approve financial rules, thresholds, and final-use policy
Engineering lead	Architecture, quality, delivery	Approve technical design and release readiness
Model governance owner	Intended use, validation, monitoring	Approve model promotion or suspension
Information security	Identity, access, secrets, privacy	Approve production security posture
Operations/support	Monitoring, incident and recovery	Approve runbook, SLOs, escalation and rollback

RACI for high-risk decisions

Decision	Responsible	Accountable	Consulted	Informed
Change reconciliation tolerance	Product + engineering	Financial data owner	Model owner	Analysts
Change contamination assumption	Model owner	Model governance owner	Financial data owner, analysts	Engineering, operations
Promote a new model	Model owner	Model governance owner	Financial data owner, engineering	Analysts, operations
Publish a new curated run	Data engineering	Product/financial owner	Model owner	Analysts
Resolve an alert	Assigned analyst	Financial review lead	Relevant domain specialist	Data/model owners
Production release	Engineering + operations	Engineering lead	Security, product, model owner	Users

04 Scope, Release Model & User Journeys

Scope matrix

In current product	Production target	Explicitly out of scope
Public WBG API ingestion; schema and financial validation; feature engineering; rule/statistical/ML detection; alert consolidation; dashboard; local persistent review; controlled evaluation	Scheduled orchestration; immutable lake zones; quarantine; authenticated RBAC; managed secrets; enterprise SQL; queue-based publication; telemetry; registered model promotion; formal UAT and rollback	Internal WBG data; source correction; automated financial decisions; fraud detection; country ranking; real-time streaming; mobile application; generative-AI conclusions

Release stages

Stage	Outcome	Exit gate
P0 · Working local reference	Reproducible pipeline, detectors, artifacts, dashboard, and local review workflow	Tests/build pass; artifact contracts are internally consistent
P1 · Controlled demonstration	Read-only hosted dashboard over public artifact snapshot	No write path; disclaimer and provenance visible
P2 · Development foundation	Cloud resources, managed identities, storage zones, monitoring, SQL, messaging	Infrastructure validation and least-privilege access verified
P3 · Governed pipeline	Scheduled ingestion, quarantine, versioned jobs, registry, atomic publication	Output parity, lineage, fault and rollback gates pass
P4 · Analyst service	Authenticated review and immutable audit workflow	Concurrency, UAT, telemetry, recovery, and permission tests pass
P5 · Controlled production	Approved model, operating model, support and review ownership	Technical, security, model, and financial owner approvals retained

Journey A: Scheduled data-control run

1. Trigger creates a unique run identifier and records configuration, code, and environment versions.
2. Ingestion retrieves paginated source records and compares received volume with the source-declared count.

3. A temporary snapshot is retained; schema, key, type, category, period, uniqueness, and financial controls run before promotion.
4. Invalid pages or records are quarantined; a failed gate leaves the last known-good publication active.
5. Validated records are standardized, feature-engineered, segmented, and scored by applicable detector families.
6. Signals are consolidated by stable record key, enriched with severity, materiality, evidence, and recommended action.
7. Approved outputs are published atomically; the dashboard changes only after the completed-run gate.

Journey B: Analyst investigation

1. Analyst filters and sorts the queue by severity, detector, category, period, entity, status, or keyword.
2. Analyst opens a record and reviews current amount, previous observation, comparable prior year, component mix, reason codes, corroboration, and public evidence context.
3. Analyst begins a review, adds notes/evidence, assigns confidence, and saves through version-checked persistence.
4. A resolved outcome requires an attributable reviewer, final outcome, notes, confidence, and supporting evidence where available.
5. Reopening creates a new audit event; it does not erase the prior decision.

Journey C: Model calibration and promotion

1. Owner compares control recall, queue volume, segment mix, sensitivity grid, review outcomes, and drift evidence.
2. Candidate configuration is versioned and executed against deterministic fault scenarios and representative data.
3. Differences from the active configuration are reviewed by model, technical, and financial owners.
4. Promotion points to an immutable model/environment/configuration; rollback remains independently available.

05 Functional Requirements

ID	Pri.	Status	Requirement	Acceptance criteria
FR-ING-01	Must	Implemented	The ingestion service shall retrieve all pages of dataset DS01557/resource RS00964 from the public WBG Finances API using a page size of at most 1,000 records.	Given a multi-page source response, when ingestion completes, then every page is present and each request uses a page size of at most 1,000.
FR-ING-02	Must	Implemented	When the received record count differs from the source-declared count, the ingestion service shall mark the run as failed.	Given unequal received and declared counts, when the count check runs, then the run state is failed.
FR-ING-03	Must	Target	When a transient source request fails, the ingestion service shall retry the request according to the approved retry policy.	Given a simulated transient failure, when ingestion retries, then attempt count and delay match the configured policy; exhaustion marks the run failed.
FR-ING-04	Must	Target	For each accepted ingestion run, the ingestion service shall retain an immutable source snapshot with a manifest containing run ID, retrieval time, page counts, source identifiers, checksums, schema fingerprint and code version.	Given an accepted run, when its retained evidence is inspected, then the snapshot and all specified manifest fields are present and a checksum recomputation matches.
FR-ING-05	Must	Target	When a page or record fails an ingestion contract check, the ingestion service shall route that evidence to a quarantine store excluded from the active dataset.	Given contract-breaking evidence, when ingestion completes, then the evidence is available in quarantine and absent from the active standardized dataset.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-ING-06	Should	Target	When source count or schema differs from the previous accepted run, the ingestion service shall emit a drift event identifying the changed count or field.	Given a changed source count or field, when drift detection runs, then the event identifies the change and the compared run IDs.
FR-ING-07	Must	Implemented	The ingestion service shall retain only records whose organization is IDA in the analytical dataset.	Given IDA and non-IDA source records, when ingestion completes, then the analytical dataset contains only IDA records.
FR-VAL-01	Must	Implemented	If a required source column is absent, the validation service shall stop validation with an error naming the missing column.	Given an input missing category, country, organization, region, time period, a financing component or total, when validation runs, then it stops and names each missing column.
FR-VAL-02	Must	Implemented	The validation service shall convert each financial field to the numeric representation defined in the data dictionary.	Given numeric strings and missing or non-numeric amounts, when validation runs, then valid strings become numeric values and invalid amounts emit INVALID_AMOUNT.
FR-VAL-03	Must	Implemented	When a category is neither Commitments nor Gross Disbursements, the validation service shall emit INVALID_CATEGORY with the observed value.	Given an unsupported category, when validation runs, then INVALID_CATEGORY contains that value.
FR-VAL-04	Must	Implemented	When multiple records share the same normalized organization, country, category and time period, the validation service shall emit DUPLICATE_GRAIN for every member of that set.	Given a duplicate set, when validation runs, then every member has DUPLICATE_GRAIN and the same business-grain key.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-VAL-05	Must	Implemented	When a financing amount is negative, the validation service shall emit <code>NEGATIVE_AMOUNT</code> identifying the field and value.	Given a negative financing component, when validation runs, then the signal contains its field and signed value without automatically declaring the record erroneous.
FR-VAL-06	Must	Implemented	When the absolute difference between reported total and component sum exceeds $\max(\text{USD } 1.01\text{M}, 0.1\% \text{ of the absolute reported total})$, the validation service shall emit <code>TOTAL_MISMATCH</code> .	Given differences below, equal to and above the tolerance, when reconciliation runs, then only the above-tolerance case emits <code>TOTAL_MISMATCH</code> .
FR-VAL-07	Must	Implemented	The validation service shall parse FYnn annual periods and YYYY-Qn quarterly periods into fiscal year, quarter and period type.	Given valid and invalid period strings, when parsing runs, then valid strings produce the defined fields and invalid strings emit <code>INVALID_PERIOD</code> .
FR-VAL-08	Must	Implemented	The identity service shall derive the record key from the normalized reporting grain using the documented SHA-256 method.	Given the same normalized grain on two runs, when keys are generated, then the identifiers are identical.
FR-VAL-09	Must	Implemented	The validation service shall normalize key strings according to the documented reporting-grain rules.	Given equivalent key strings with permitted formatting differences, when normalized, then their normalized values match.
FR-VAL-10	Must	Implemented	When a required key value is blank, the validation service shall emit <code>MISSING_KEY</code> .	Given a blank required key, when validation runs, then <code>MISSING_KEY</code> names the affected field.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-FEA-01	Must	Implemented	The feature service shall calculate the derived fields defined in the versioned data dictionary.	Given a fixed input fixture, when feature calculation runs twice, then component sums, differences, shares, log magnitude, sequential change, comparable-year change, ratios and materiality match the expected fixture.
FR-FEA-02	Must	Implemented	When absolute commitments are below USD 10M, the feature service shall suppress the disbursement-to-commitment ratio.	Given commitments below USD 10M, when features are calculated, then the ratio is null and ratio_denominator_eligible is false; the exact threshold is eligible.
FR-FEA-03	Must	Implemented	The feature service shall classify each observation by period type and entity type before detector execution.	Given annual, quarterly, country and aggregate records, when classification runs, then each receives the correct period and entity labels.
FR-DET-01	Must	Implemented	If an ML segment is ineligible or ML execution fails, the detection service shall retain the deterministic structural and financial rule signals.	Given a skipped or failed ML segment, when detection completes, then its deterministic rule signals remain available.
FR-DET-02	Must	Implemented	When a comparable-year movement meets the versioned statistical thresholds, the statistical detector shall emit YEAR_OVER_YEAR_SHIFT.	Given movements below and above the configured gates, when detection runs, then only qualifying cases contain current, prior-year, absolute and relative evidence with a review action.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-DET-03	Must	Implemented	The ML service shall fit a separate Isolation Forest for each eligible period-type and category segment containing at least 30 country observations.	Given mixed categories and periods, when fitting runs, then no model contains mixed segment types and segments below 30 observations are reported as skipped.
FR-DET-04	Must	Implemented	The preprocessing service shall replace missing model features with the median of the corresponding segment and feature.	Given missing values, when preprocessing runs, then each imputed value equals the recorded segment median.
FR-DET-05	Must	Implemented	The ML service shall use the ordered model-feature schema defined in the approved run configuration.	Given a scored run, when metadata is inspected, then the ordered schema includes magnitude, composition, sequential movement, comparable-year movement, cross-category ratio and period context.
FR-DET-06	Must	Implemented	For each ML signal, the explanation service shall return the three largest absolute robust-scaled feature deviations and the batch-relative anomaly score.	Given an ML signal, when its explanation is inspected, then the three named deviations match independently sorted feature values and the score is identified as batch-relative.
FR-DET-07	Must	Implemented	When an alert contains only ML signals, the severity service shall cap its severity below Critical.	Given an ML-only alert with high score and materiality, when severity is assigned, then its severity is not Critical.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-DET-08	Must	Implemented	The consolidation service shall create one alert per record key containing all associated signals, reason codes, evidence and detector-family counts.	Given multiple signals for one key, when consolidation runs, then exactly one alert contains all input signals and the highest policy-permitted severity.
FR-DET-09	Must	Implemented	The queue service shall order alerts according to the versioned severity, corroboration, materiality and anomaly-evidence policy.	Given alerts with known ordering factors, when the queue is generated, then its order matches the policy fixture.
FR-DET-10	Should	Target	For each scored alert, the lineage service shall record model, environment, feature schema, parameter, source snapshot and code versions.	Given a scored alert, when lineage is resolved, then every specified version links to retained run evidence and the approved model registry entry.
FR-DET-11	Must	Implemented	The preprocessing service shall apply robust scaling to the imputed model features.	Given the reference feature matrix, when scaling runs, then the result matches the retained RobustScaler configuration.
FR-DET-12	Must	Implemented	The preprocessing service shall convert non-finite feature values to missing values before imputation.	Given infinite feature values, when preprocessing runs, then no non-finite value enters model fitting.
FR-UI-01	Must	Implemented	The dashboard shall display portfolio volumes, severity distribution, detector distribution and review progress for the active run.	Given an active run, when the dashboard loads, then every displayed summary reconciles to its source artifacts.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-UI-02	Must	Implemented	The alert queue shall support combined text search and severity, detector, category, period, entity and review-state filters.	Given a queue fixture, when search and filters are combined, then only matching alerts are returned and selected sorting is preserved.
FR-UI-03	Must	Implemented	The investigation panel shall display current, previous-period and comparable-prior-year values with unavailable comparisons identified as unavailable.	Given a record without comparison history, when its panel opens, then missing comparisons display as unavailable rather than zero.
FR-UI-04	Must	Implemented	For each alert, the investigation panel shall display reason codes, detector family, materiality, corroboration, evidence and recommended action.	Given an alert with those fields, when its panel opens, then all six information groups are visible and match the alert artifact.
FR-UI-05	Must	Implemented	The model view shall display model assumptions, evaluation evidence, data-quality caveats and review-quality limitations.	Given the current model evidence, when the view opens, then each information group is present and no score or controlled-test result is labeled as production probability or accuracy.
FR-UI-06	Must	Implemented	While the review API is unavailable, the dashboard shall expose analytical content in read-only snapshot mode.	Given an unavailable API, when a user opens the dashboard, then snapshot content is readable and no write control reports a successful save.
FR-REV-01	Must	Implemented	The review service shall permit only Pending to In Review, In Review to Resolved and Resolved to In Review state transitions.	Given each permitted and prohibited transition, when an update is submitted, then permitted transitions succeed and prohibited transitions return validation errors without mutation.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-REV-02	Must	Implemented	If a resolution lacks final outcome, reviewer, notes or confidence, the review service shall reject it.	Given a resolution missing each required field in turn, when submitted, then the service returns a validation error and creates no state change.
FR-REV-03	Must	Implemented	The review service shall restrict resolved outcomes to confirmed_data_issue, legitimate_exception and false_positive.	Given each supported outcome and needs_more_information, when resolution is requested, then only the three final outcomes are accepted.
FR-REV-04	Must	Implemented	If expected_version differs from the stored review version, the review service shall reject the update with HTTP 409.	Given two clients using the same version, when the first save succeeds and the second saves, then the second receives 409 and the first stored decision remains unchanged.
FR-REV-05	Must	Implemented	For each accepted review transition, the review service shall append an audit event without altering earlier events.	Given a transition history, when another transition is accepted, then one new ordered event appears and all previous events are unchanged.
FR-REV-06	Must	Implemented	When the analytical pipeline reruns, the review service shall preserve existing review decisions by stable record key.	Given a stored review and a rerun containing the same key, when results are joined, then the stored state and history remain unchanged.
FR-REV-07	Must	Target	The review API shall derive reviewer identity from authenticated Entra claims instead of client-supplied reviewer fields.	Given an authenticated request with a spoofed reviewer field, when accepted, then the event contains the claim subject and verified display identity.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-REV-08	Must	Implemented	The review service shall permit needs_more_information only when the review state is In Review.	Given that outcome with Pending, In Review and Resolved states, when validation runs, then only In Review is accepted.
FR-EVA-01	Must	Implemented	The evaluation service shall execute seeded scenarios for missing amount, broken reconciliation, negative component, duplicate grain and reconciled historical spike.	Given the five scenarios and documented seeds, when the suite runs, then each trial records expected codes, detector layers and observed outcomes.
FR-EVA-02	Must	Implemented	The evaluation service shall compare model sensitivity at contamination values of 0.5%, 1%, 2% and 5%.	Given the reference snapshot, when the sensitivity suite runs, then all four values have ML count, queue volume and controlled-spike recall outputs.
FR-EVA-03	Must	Implemented	While fewer than ten reviews are resolved, the evaluation service shall suppress actionable precision and false-positive rate.	Given nine and ten resolved reviews, when metrics are calculated, then rates are null at nine and become eligible at ten with the method note retained.
FR-EVA-04	Should	Target	When a configured segment drift threshold is exceeded, the monitoring service shall emit an event identifying the affected feature, source, score, alert volume or outcome measure.	Given a breached segment threshold, when monitoring runs, then an event identifies the segment and measure without changing model parameters.
FR-PUB-01	Must	Target	When all required publication gates pass, the publication service shall atomically activate the curated output version.	Given passing and failing gate fixtures, when publication is attempted, then only the passing version becomes active and no partial output is exposed.

ID	Pri.	Status	Requirement	Acceptance criteria
FR-PUB-02	Must	Target	The release service shall support independent rollback of application, model and active data-publication versions.	Given retained approved versions, when each rollback is executed separately, then the selected boundary is restored without deleting source, review or audit evidence.
FR-PUB-03	Must	Target	For each published run, the publication service shall emit an idempotent completed-run event keyed by run and publication ID.	Given repeated delivery of the same event, when downstream upsert executes, then one operational alert exists per publication and record key.

06 Business Rules & Decision Logic

ID	Rule
BR-01	The identity service shall identify records by normalized organization, country, category and time period.
BR-02	The ingestion service shall exclude non-IDA records from the analytical dataset.
BR-03	The statistical and ML services shall evaluate country records only.
BR-04	The ML service shall separate Commitments and Gross Disbursements into different model segments.
BR-05	The ML service shall separate annual and quarterly observations into different model segments.
BR-06	If a segment has fewer than 30 observations, the ML service shall skip fitting and scoring that segment.
BR-07	The reference model configuration shall use the provisional contamination value of 1%.
BR-08	When absolute commitments are below USD 10M, the feature service shall suppress ratio evidence.
BR-09	The validation service shall use a reconciliation tolerance of max(USD 1.01M, 0.1% of the absolute reported total).
BR-10	The validation service shall treat a negative amount as a review signal rather than an automatic invalidity determination.

ID	Rule
BR-11	The dashboard shall identify the ML score as batch-relative rather than a probability.
BR-12	The severity service shall exclude Critical severity for ML-only alerts.
BR-13	The review service shall reject an outcome assigned to a Pending review.
BR-14	The review service shall reject needs_more_information as a resolution outcome.
BR-15	The review service shall require final outcome, attributable reviewer, notes and confidence for resolution.
BR-16	When a resolved alert is reopened, the review service shall retain its previous resolution event.
BR-17	While fewer than ten reviews are resolved, the evaluation service shall suppress review-performance rates.
BR-18	When a pipeline run fails, the publication service shall retain the last successful active publication.

Severity policy

Condition	Maximum / resulting severity	Rationale
Missing/non-numeric amount, missing key, duplicate grain, severe deterministic failure	Critical	Directly blocks trustworthy use of the record or grain
ML-only; score ≥ 0.85 and materiality ≥ 75 th percentile	High	Unusual and financially material, but not independently critical
Other ML-only signal	Medium	Requires context; score is relative, not a probability
At least two detector families; base high and materiality ≥ 75 th percentile	Critical	Independent corroboration plus financial materiality
Other corroborated signals	High	Multiple forms of evidence justify priority
Aggregate reconciliation mismatch	Medium	Aggregate component semantics may differ from country reporting

Outcome taxonomy

Outcome	Definition	Permitted status
confirmed_data_issue	Evidence supports a source or processing issue requiring correction/escalation	Resolved
legitimate_exception	Unusual behavior is valid and contextually explainable	Resolved
false_positive	The control did not provide a review-worthy signal under the intended policy	Resolved
needs_more_information	Evidence is insufficient for a final decision	In Review

07 Data Requirements & Contracts

Source contract

Attribute	Requirement
Dataset	IDA Commitments and Disbursements: Country/Economy Summary
Dataset/resource IDs	DS01557 / RS00964
Unit	USD millions
Supported categories	Commitments; Gross Disbursements
Supported periods	FYnn; YYYY-Qn
Expected grain	organization × country × category × time_period
Access boundary	Public WBG Finances API only

Core data dictionary

Field	Type / domain	Required	Purpose / validation
record_key	string · ida_[20 hex]	Derived	Stable review and alert identity from normalized grain
organization	string	Yes	Filtered to IDA
country	string	Yes	Country, region, or institutional entity
region	string	Yes	Geographic grouping
category	enum	Yes	Commitments or Gross Disbursements
time_period	string	Yes	FYnn or YYYY-Qn
development_policy	decimal	Yes	Financing component; numeric; negatives reviewable
investment_lending	decimal	Yes	Financing component; numeric; negatives reviewable
others	decimal	Yes	Financing component
program_for_results	decimal	Yes	Financing component
total	decimal	Yes	Reported total; reconciled to components
entity_type	enum	Derived	country or aggregate/institution
period_type	enum	Derived	annual or quarterly
fiscal_year	integer	Derived	Normalized year
quarter	1–4	Derived	Annual observations use 4 only for ordering/context
reconciliation_difference	decimal	Derived	total – component sum
relative_change	decimal/null	Derived	Change from previous observation where valid
year_over_year_change	decimal/null	Derived	Comparable prior-year change
disbursement_commitment_ratio	decimal/null	Derived	Suppressed below denominator threshold
materiality_percentile	0–1	Derived	Within category × period-type magnitude ranking

Output contracts

Artifact	Grain	Required contents
alert_signals	One detector signal	record key, row reference, detector, reason, severity, evidence, recommendation, run context
alerts	One consolidated source record	financial context, all signals/reasons, severity, corroboration, materiality, score/evidence, review projection
run_summary	One analytical run	timestamp, volumes, segment/model configuration, alert distributions
evaluation_summary	One evaluation execution	seeds, scenarios, layer outcomes, sensitivity grid, method notes
reviews	One current review state	record key, status, outcome, reviewer, notes, confidence, evidence URL, version, timestamps
review_events	One accepted mutation	before/after values, actor, timestamp, version, record key

Data-quality gates

Dimension	Gate / evidence	Failure behavior
Extraction completeness	Received count equals source-declared count	Fail run; quarantine evidence; retain last good
Required fields	All required columns exist; invalid cell values become named signals	Missing column fails; invalid record is signaled
Uniqueness	No duplicate documented grain	Critical signals; production policy determines publication block
Category/period conformity	Controlled category and parseable period	High signal; do not model invalid period
Financial reconciliation	Difference within documented tolerance	Signal with entity-aware severity
Feature readiness	Expected schema; finite preprocessing; eligible segment size	Skip affected model segment and report
Publication integrity	All required run gates approved	Atomic publish or no change

08 Analytics, Model & Responsible-AI Requirements

Model specification

Attribute	Current implementation
Model family	Isolation Forest; unsupervised anomaly detection
Segmentation	period_type × category; country records only
Preprocessing	Non-finite → missing; segment-median imputation; RobustScaler
Estimators	300 trees per eligible segment
Operating contamination	0.01 provisional
Random state	42
Scoring	Negative decision function normalized 0–1 within run and segment
Explainability	Top three absolute robust-scaled feature deviations
Decision authority	Human analyst
Persistence	Current models in memory; production target requires registered versioned artifacts

Feature groups

Group	Features	Purpose
Magnitude	log1p_total	Retain relative scale while reducing skew
Composition	Four financing-component shares	Identify unusual financing mix
Sequential movement	relative_change	Compare with preceding observation in same series
Comparable-year movement	year_over_year_change	Compare same quarter/prior annual observation
Cross-category context	disbursement_commitment_ratio	Identify unusual disbursement relative to eligible commitments
Period context	quarter	Represent within-year timing

Evaluation interpretation

Evidence supports	Evidence does not support
Selected software/control behavior across 25 seeded trials	Production accuracy, real-world precision, or false-positive-rate claims

Sensitivity evidence

Contamin ation	ML alerts	Total alerts	Queue rate	Moderate 5× recall	Severe 10× recall	Selected
0.5%	18	102	2.98%	0.0	0.2	No
1.0%	34	118	3.45%	0.2	1.0	Yes
2.0%	65	143	4.18%	0.2	1.0	No
5.0%	160	232	6.78%	0.8	1.0	No

Responsible-use controls

ID	Pri.	Status	Requirement	Acceptance criteria
AI-01	Must	Implemented	The model view shall identify the detector as advisory and its score as batch-relative.	Given the model view, when its score description is inspected, then advisory and batch-relative labels are present.
AI-02	Must	Implemented	The alert view shall display reason codes, feature evidence, financial context, materiality and known limitations.	Given an alert, when opened, then all five evidence groups are visible.
AI-03	Must	Implemented	The product shall exclude fraud, misconduct, policy-violation and country-quality determinations from anomaly output.	Given generated outputs, when fields and copy are inspected, then none asserts those determinations.
AI-04	Must	Target	The publication service shall prohibit automatic source correction, suppression or publication based solely on model output.	Given model output without human or gate approval, when publication is requested, then the source and active publication remain unchanged.
AI-05	Must	Implemented	The evaluation view shall distinguish specialist-reviewed outcome evidence from controlled-test evidence.	Given current evaluation artifacts, when the view opens, then controlled results are labeled and specialist validation is not claimed.
AI-06	Must	Target	The model registry shall retain data, code, environment, feature schema, parameter, evaluation, approval and rollback references for promoted models.	Given a promoted model, when its registry entry is inspected, then every specified reference resolves.

Minimum promotion evidence

1. Expected feature schema and segment populations; no unexplained source or distribution shift.
2. Controlled fault suite and application/data contracts pass.
3. Alert load and severity distribution are reviewed against operational capacity.
4. Model card, limitations, intended use, and responsible-use restrictions remain current.
5. Technical, model-governance, and financial-owner approvals are recorded; rollback version is confirmed.

09 Non-Functional Requirements

ID	Pri.	Status	Requirement	Acceptance criteria
NFR-PERF-01	Must	Target	Under the approved production workload, the queue service shall return its first usable page within 2.5 seconds at p75 and 5 seconds at p95.	Given the approved dataset, concurrency and filter mix, when the load test runs, then measured p75 is at most 2.5 seconds and p95 is at most 5 seconds.
NFR-PERF-02	Must	Target	The review API shall complete review updates within one second at p95, excluding client network latency.	Given the approved production workload, when write latency is measured server-side, then p95 is at most one second.
NFR-PERF-03	Must	Implemented	The local pipeline shall process the reference 3,420-record snapshot without distributed compute.	Given the reference developer environment, when the pipeline runs, then it completes successfully using local compute only.
NFR-REL-01	Must	Target	If a pipeline run fails, the publication service shall retain the active successful publication.	Given a failed run, when its publication pointer is checked, then the previously active successful version remains selected.
NFR-REL-02	Must	Target	The production service shall achieve at least 99.5% monthly availability, excluding approved maintenance.	Given monthly health-check results and the approved maintenance record, when availability is calculated, then the result is at least 99.5%.

ID	Pri.	Status	Requirement	Acceptance criteria
NFR-REL-03	Must	Implemented	While the review API is unavailable, the dashboard shall remain usable in read-only snapshot mode.	Given an unavailable API, when investigation is attempted, then analytical content remains readable and no save-success state is displayed.
NFR-REL-04	Must	Target	The alert upsert service shall process repeated completed-run events without creating duplicate alert or review records.	Given duplicate event deliveries, when all are processed, then alert and review counts equal those after the first delivery.
NFR-SEC-01	Must	Target	The production API shall authenticate human users through Microsoft Entra ID.	Given anonymous and authenticated requests, when a protected operation is requested, then anonymous access is denied and authenticated access proceeds to authorization.
NFR-SEC-02	Must	Target	The authorization service shall enforce the approved least-privilege matrix for analysts, data engineers, model owners and operators.	Given each role and action in the approved matrix, when permission tests run, then only authorized combinations succeed.
NFR-SEC-03	Must	Target	Production workloads shall retrieve secrets and certificates from Key Vault using authorized identities.	Given a production deployment, when secret access is inspected, then authorized values originate from Key Vault and no secret is embedded in repository or application configuration.
NFR-SEC-04	Must	Target	Production endpoints shall use TLS version 1.2 or later for data in transit.	Given each production endpoint, when its protocol is inspected, then TLS below 1.2 is rejected.

ID	Pri.	Status	Requirement	Acceptance criteria
NFR-SEC-05	Must	Target	Production storage, database and application endpoints shall conform to the approved private-network topology.	Given the approved topology, when network configuration is compared, then public access is disabled except for explicitly approved exceptions.
NFR-SEC-06	Must	Implemented	The local API shall accept only configured origins or localhost and only GET or PUT methods.	Given approved and unapproved origins and methods, when CORS tests run, then only configured combinations are accepted.
NFR-SEC-07	Must	Target	Production workloads shall authenticate to supported Azure services using managed identities.	Given a production workload, when identity configuration is inspected, then service access uses its assigned managed identity rather than embedded credentials.
NFR-SEC-08	Must	Target	Production storage services shall encrypt retained data at rest.	Given each production data store, when configuration is inspected, then encryption at rest is enabled.
NFR-SEC-09	Must	Target	The production API shall accept browser requests only from explicitly approved origins.	Given an unapproved origin, when a browser request is made, then the API denies cross-origin access.
NFR-PRV-01	Must	Implemented	The prototype shall process only publicly available financial source datasets.	Given the source-data inventory, when classification is inspected, then every financial dataset is public and no internal WBG dataset is required.

ID	Pri.	Status	Requirement	Acceptance criteria
NFR-PRV-02	Must	Target	The telemetry service shall exclude secrets, access tokens and unnecessary review-note content from stored logs.	Given requests containing those values, when telemetry is inspected, then the values are absent or redacted.
NFR-AUD-01	Must	Implemented	For each accepted review change, the review service shall append an immutable historical event.	Given a review change and subsequent reopening, when history is compared, then the original event is unchanged and the reopening appears as a new event.
NFR-AUD-02	Must	Target	The lineage service shall resolve each published alert to its source snapshot, standardized record, feature schema, detector run and code/environment version.	Given a published alert, when its lineage is queried, then all specified references resolve to retained evidence.
NFR-AUD-03	Must	Target	The audit service shall store timestamps in UTC.	Given an audit export, when timestamps are inspected, then every timestamp uses UTC.
NFR-AUD-04	Must	Target	The audit service shall retain an immutable authenticated subject identifier for each production review event.	Given a production review event, when exported, then its actor identifier matches the authenticated subject and cannot be edited.
NFR-DQ-01	Must	Implemented	The quality service shall report completeness, field validity, uniqueness, category/period conformity, sign signals, reconciliation and feature availability as separate dimensions.	Given a quality report, when inspected, then all seven dimensions have separate evidence and no composite score replaces them.

ID	Pri.	Status	Requirement	Acceptance criteria
NFR-DQ-02	Must	Target	The storage service shall retain raw snapshots and manifests as immutable, hash-verifiable artifacts.	Given a retained snapshot, when its checksum is recomputed, then it matches the manifest and unauthorized modification is denied.
NFR-MOD-01	Must	Implemented	The ML service shall produce reproducible results for identical input, configuration, environment and random state.	Given identical run conditions, when evaluation is repeated, then detector outcomes match the reference result.
NFR-MOD-02	Must	Target	The model registry shall permit promotion only after the required automated gates and named human approvals are recorded.	Given a candidate missing any gate or approval, when promotion is requested, then approval status remains unchanged.
NFR-MOD-03	Must	Target	The monitoring service shall calculate drift and review-outcome measures separately for each period-type and category segment.	Given annual, quarterly and mixed-category data, when monitoring runs, then each measure contains only its own segment population.
NFR-ACC-01	Must	Target	The core dashboard and review journeys shall conform to WCAG 2.2 AA.	Given the core journeys, when automated and manual accessibility checks run, then keyboard, focus, semantics, contrast, zoom and screen-reader checks meet the applicable AA criteria.
NFR-USA-01	Must	Implemented	The investigation panel shall provide a plain-language reason, supporting evidence and next action for every alert.	Given each alert in the review fixture, when opened, then all three content elements are present.

ID	Pri.	Status	Requirement	Acceptance criteria
NFR-USA-02	Must	Target	The interface shall expose loading, empty, error, stale, offline, conflict and success states for applicable investigation actions.	Given each state fixture, when the action is exercised, then the corresponding state is visible and no failed write is shown as successful.
NFR-COMP-01	Must	Target	The primary journeys shall operate on the approved current Chrome, Edge and Safari versions at viewport widths of at least 1,024 pixels.	Given the approved browser matrix and supported widths, when core journeys run, then all complete without loss of required content or controls.
NFR-MNT-01	Must	Implemented	The application shall separate validation, feature engineering, detection, evaluation, review persistence, API and dashboard concerns into modules.	Given the repository, when module tests execute, then core controls can be exercised without launching the dashboard.
NFR-MNT-02	Must	Target	The release process shall version-control infrastructure, application, data contracts and model configuration.	Given a release, when its manifest is inspected, then every specified artifact resolves to an immutable revision.
NFR-TST-01	Must	Implemented	The CI service shall block a release when a required ingestion, control, evaluation, review or front-end test fails.	Given a failing test in each required suite, when CI runs, then the release gate fails.
NFR-OBS-01	Must	Target	The telemetry service shall emit structured health evidence for pipeline, model, API, database, queue and UI components.	Given a simulated component failure, when telemetry is inspected, then component, run/publication ID, timestamp and error category identify the failure.

ID	Pri.	Status	Requirement	Acceptance criteria
NFR-OBS-02	Must	Target	The operational dashboard shall expose run state, source counts, publication age, quarantine volume, alert shifts, queue health and unresolved Critical-alert age.	Given an active publication, when the operational view opens, then all specified measures are available without direct database access.
NFR-RCV-01	Must	Target	The release process shall rehearse independent application, model and publication rollback at least quarterly.	Given the quarterly exercise record, when inspected, then each rollback boundary was restored without deletion of review or audit history.
NFR-RCV-02	Must	Target	The review-database recovery process shall meet the RPO and RTO values approved before production release.	Given approved RPO/RTO values, when a restore exercise runs, then measured data loss and recovery time do not exceed them; absent approval blocks release.
NFR-SCL-01	Must	Target	The capacity-validation process shall load-test the approved expected workload before production sizing is approved.	Given the proposed workload, when sizing is reviewed, then retained load-test evidence supports the selected service tiers.
NFR-COST-01	Must	Target	The production environment shall have approved budgets, cost-anomaly alerts, resource tags and retention limits before non-development compute is activated.	Given a proposed production activation, when configuration is inspected, then all four controls are present.
NFR-ETH-01	Must	Implemented	The product shall disclose that an anomaly signal does not establish error, fraud or misconduct.	Given dashboard and model evidence views, when copy is inspected, then the disclosure is present and no automated punitive action is offered.

10 Architecture & Integration Requirements

Current implementation

Layer	Technology	Responsibility
Source	WBG Finances public REST API	DS01557/RS00964 records
Pipeline	Python/pandas/scikit-learn	Ingest, validate, features, rules, statistical controls, ML, consolidate, evaluate
Artifacts	CSV/JSON/local files	Signals, alerts, summaries, review sample, processed features
Review service	FastAPI	Read/write review contract and conflict handling
Review store	SQLite	Current review state and append-only events
Interface	Web dashboard	Monitoring, queue, investigation, governance evidence

Target production flow

Step	Service boundary	Contract / control
1	Azure Data Factory → WBG API	Schedule, pagination, retry, run ID, declared-count validation
2	ADF → ADLS raw/quarantine	Temporary write, immutable snapshot, manifest, atomic promotion or quarantine
3	ADLS → Azure ML jobs	Versioned validation, feature, detection, evaluation, scoring components
4	Azure ML → standardized/curated zones	Contracted records, features, signals, alerts, summaries, run lineage
5	Curated → Service Bus	Idempotent completed-run publication event
6	Service Bus/API → Azure SQL	Alert upsert by stable key; review and audit persistence
7	FastAPI/SQL → dashboard	Authenticated queue, investigation, review writes, read-only fallback
Cross-cutting	Entra ID · Key Vault · Azure Monitor	Identity/RBAC, secrets/certificates, telemetry/incident detection

Architectural invariants

1. The raw source snapshot, derived analytical artifacts, and operational review evidence are separate data classes.
2. Re-running analytics may change derived signals but cannot rewrite historical reviews or audit events.
3. A stable record key is the join boundary; positional row_id is display/trace context only.
4. The dashboard reads one active, successfully promoted publication; partial outputs are never mixed.
5. Rules remain executable if ML is unavailable, skipped, or rejected.
6. Application, model, and data publication versions can be promoted and rolled back independently.

API contract

Method	Route	Behavior	Key responses
GET	/health	Service and store health	200
GET	/api/reviews	List persisted review records	200
GET	/api/reviews/{record_key}	Current review or default pending state	200 / 404
GET	/api/reviews/{record_key}/history	Append-only ordered event history	200 / 404
PUT	/api/reviews/{record_key}	Validate and persist versioned transition	200 / 404 / 409 / 422

11 Experience, Content & Accessibility Specification

Information architecture

Surface	Primary questions answered
Portfolio overview	What was processed? Is the active run healthy? Where is review attention concentrated?
Alert queue	Which records deserve attention, and why are they prioritized?
Investigation panel	What changed, which controls agree, what evidence exists, and what should I inspect next?
Review workspace	What is the current state, what evidence supports the decision, and who made each change?
Model/evaluation view	How does the detector work, how sensitive is the queue, and what cannot yet be claimed?
Data-quality view	Which quality dimensions passed, require review, or remain unverified?

Required interface states

ID	Pri.	Status	Requirement	Acceptance criteria
UX-01	Must	Target	While data is loading, the interface shall display a loading indicator.	Given a delayed response, when a view opens, then a loading indicator is visible until completion or failure.
UX-02	Must	Target	When active filters return no alerts, the queue shall display an empty-state message identifying the filters.	Given filters with zero matches, when applied, then the message names the active filters.
UX-03	Must	Implemented	When comparison data is absent, the investigation panel shall display Not available rather than zero.	Given absent history, when the panel opens, then no fabricated zero is shown.
UX-04	Must	Implemented	While the API is offline, the review interface shall indicate that changes cannot be saved.	Given an offline API, when review controls are inspected, then the unsaved or disabled state is explicit.
UX-05	Must	Target	When a version conflict occurs, the review interface shall preserve the unsaved user draft.	Given a draft and HTTP 409, when the response arrives, then the draft values remain available.
UX-06	Must	Target	When a pipeline run fails, the dashboard shall display the active publication age and failure notice.	Given a failed run, when the dashboard opens, then the prior publication date and failure notice are visible.
UX-07	Must	Implemented	While fewer than ten reviews are resolved, the evaluation view shall explain the rate-suppression threshold.	Given nine resolved reviews, when the view opens, then rates are unavailable and the ten-review threshold is stated.
UX-08	Must	Target	The operator view shall expose quarantine status only to authorized operators.	Given operator and unauthorized sessions, when quarantine status is requested, then only the operator can access it.

ID	Pri.	Status	Requirement	Acceptance criteria
UX-09	Must	Target	When an empty-state message is displayed, the queue shall provide a filter-reset action.	Given an empty filtered queue, when reset is selected, then all active filters are cleared.
UX-10	Must	Target	When a version conflict occurs, the review interface shall display the current stored version.	Given HTTP 409, when the conflict view opens, then the current stored version is shown for deliberate reconciliation.

Content rules

1. The interface shall describe unconfirmed anomalies as alerts, signals, unusual records or records requiring review.
2. The interface shall label the anomaly measure as a batch-relative anomaly score.
3. The financial display shall show amounts in USD millions with the original sign preserved.
4. The investigation view shall display applicable ratio, contamination, aggregate and history limitations beside the affected conclusion.
5. The architecture view shall label implemented and target capabilities explicitly.

Accessibility acceptance

1. Given a core journey, when completed using only a keyboard, then every control is reachable with visible focus in logical order.
2. Given a data table or chart, when read using assistive technology, then table headers or an equivalent textual chart summary are available.
3. Given severity and status indicators, when color is removed, then their text or symbols still communicate meaning.
4. Given core journeys, when contrast, zoom, labels, errors, landmarks and accessible names are checked, then applicable WCAG 2.2 AA criteria pass.
5. Given a reduced-motion preference, when the interface loads, then optional animation is reduced and no investigation action is obstructed.

12 Acceptance, Verification & Release Readiness

Critical acceptance scenarios

Test ID	Scenario	Expected result
AT-01	Source returns fewer rows than declared	Run fails; evidence preserved; active publication unchanged
AT-02	Required amount is missing/non-numeric	INVALID_AMOUNT raised with field and corrective action
AT-03	Duplicate business grain is introduced	All duplicate members receive DUPLICATE_GRAIN
AT-04	Component total is broken beyond tolerance	TOTAL_MISMATCH shows reported, component sum, difference
AT-05	Negative component introduced	NEGATIVE_AMOUNT raised; record not auto-corrected
AT-06	Internally reconciled historical spike introduced	Statistical and ML layers detect expected stress case
AT-07	Commitment denominator below USD 10M	Cross-category ratio is suppressed
AT-08	Eligible segment has <30 observations	ML skipped and reported; rules continue
AT-09	Multiple signals reference same record	One alert retains all evidence and corroboration
AT-10	ML-only signal attempts critical escalation	Policy caps severity below critical
AT-11	Resolved review lacks required fields	422; no state/event mutation
AT-12	Two clients save from same version	First succeeds; stale second receives 409
AT-13	Resolved alert is reopened	Current state changes; historical resolution event remains
AT-14	Review API unavailable	Dashboard remains read-only; no false save confirmation
AT-15	Evaluation has fewer than 10 resolved cases	Precision/false-positive metrics remain null
AT-16	Completed-run event replayed	No duplicate alert or audit event
AT-17	Candidate release fails gate	No production promotion; prior versions remain selectable

Prototype definition of done

1. Python control, ingestion, evaluation, and review-backend tests pass.
2. Dashboard lint/build and artifact-contract checks pass.
3. Generated alert and summary artifacts reconcile internally.
4. Known limitations, independent-project disclaimer, model card, and data-quality evidence are current.
5. No claim is made that public-review outcomes constitute specialist validation.

Production release gate

1. Infrastructure compiles, validates, and passes reviewed what-if evidence.
2. Managed identities, RBAC, private networking, encryption, secrets, backups, and diagnostics are verified.
3. Pipeline count/schema/quarantine/atomic-publication tests pass against representative data.
4. Model/data/app contracts, controlled fault suite, load tests, security tests, and accessibility checks pass.
5. Database migration, concurrency, audit, read-only fallback, restore, and rollback exercises pass.
6. Financial-data specialists complete UAT and named technical, security, model, operational, and financial owners approve release.

13 Risks, Assumptions, Dependencies & Open Decisions

Risk register

ID	Risk	Impact	Mitigation / control
R-01	No authoritative ground-truth labels	Cannot estimate real-world accuracy or missed errors	Human review program; suppress rates until evidence threshold; specialist validation
R-02	Fixed contamination controls alert volume	Under/over-alerting as distributions shift	Sensitivity analysis; workload monitoring; governed recalibration
R-03	Refitting changes relative baselines	Historical scores/flags may move	Retain run/model versions; compare stability; never treat score as timeless probability
R-04	Small quarterly segments	Less stable peer distributions	Minimum segment size; segment-level monitoring; specialist interpretation
R-05	Aggregate entity list is manually maintained	Non-country could enter country model	Govern reference list; data-owner review; automated change checks
R-06	Temporal/policy context absent	Legitimate shifts may appear anomalous	Public evidence context; analyst decision; no autonomous conclusion
R-07	Source/API contract changes	Failed or misleading ingestion	Schema fingerprint, drift gate, quarantine, last-good publication
R-08	Reviewer identity is free text locally	Identity spoofing in multi-user use	Local-only limitation; Entra-derived identity before production
R-09	Dashboard encourages score fixation	Misinterpretation of ML output	Explainability, context, reason codes, non-probability labeling, UX testing
R-10	Review notes may contain sensitive context	Privacy/security exposure	Data minimization, access control, redacted telemetry, retention policy
R-11	Coupled release could corrupt operational evidence	Loss of trust and recoverability	Separate version/rollback boundaries; migration tests; append-only events
R-12	Portfolio deployment mistaken for WBG system	Reputational misrepresentation	Persistent independent-project disclaimer and public-data attribution

Assumptions

1. The public API continues to provide the documented dataset and sufficient metadata for pagination integrity.

2. USD-million amounts and category semantics remain materially consistent across the represented reporting periods.
3. A financial-data specialist, not the anomaly model, owns the meaning of unusual values and final control policy.
4. The current 1% operating point is provisional and chosen to balance controlled sensitivity with review load.
5. Public evidence can support investigation but cannot substitute for internal source documentation.

Dependencies

1. WBG Finances public API availability and dataset contract.
2. Python/pandas/scikit-learn pipeline and tested dependency versions.
3. FastAPI and database availability for persistent review writes.
4. Cloud identity, storage, database, messaging, model registry, monitoring, and networking services for production target.
5. Named financial, technical, model-governance, security, and operational owners.

Open decisions before production

Decision	Owner(s)	Evidence required
Final materiality and reconciliation policy	Financial data owner	Specialist review of source semantics and business consequences
Operating contamination / capacity	Model + financial owners	Sufficient labeled outcomes, queue capacity, missed-vs-extra-review trade-off
Publication-blocking controls	Product + financial + engineering	Impact analysis and exception workflow
Production RPO/RTO and retention	Operations + data owner	Business criticality and recovery exercises
Role model and segregation of duties	Security + owners	Access matrix and abuse-case testing
Authoritative aggregate-entity reference	Financial data owner	Governed reference source and change owner

14 Requirements Traceability Matrix

User/control need	Requirements	Evidence location	Verification
Reliable complete source	FR-ING-01-07; NFR-REL-01; NFR-DQ-02	src/ingest.py; ingestion guide; Azure target	Count mismatch/unit/fault tests; manifest/quarantine acceptance
Valid financial grain	FR-VAL-01-10; BR-01-03	src/validate.py; data dictionary	test_controls.py; duplicate/type/category/period scenarios
Contextual anomaly detection	FR-FEA-01-03; FR-DET-01-12	src/features.py; src/detect.py; model card	test_controls.py; test_evaluation.py; sensitivity grid
Explainable prioritization	FR-DET-06-09; FR-UI-03-05; NFR-USA-01	alerts.csv; dashboard; model card	Artifact contracts; UI acceptance; alert evidence inspection
Controlled analyst review	FR-REV-01-08; BR-13-16; NFR-AUD-01,04	backend/store.py; backend/main.py; SQLite	test_review_backend.py ; concurrency and history acceptance
Honest performance evidence	FR-EVA-01-04; BR-17; NFR-MOD-01-03	evaluate.py; evaluation/review summaries	25 seeded trials; rate suppression; drift target tests
Safe publication/recovery	FR-PUB-01-03; NFR-RCV-01-02	Azure design; runbook; production plan	Atomic publication, replay, rollback, restore exercises
Secure governed operation	NFR-SEC-01-09; NFR-OBS-01-02	Bicep target; security and operations plans	RBAC, secret, network, telemetry, incident tests
Accessible investigation	FR-UI-01-06; UX-01-10; NFR-ACC-01; NFR-USA-02	Dashboard implementation/design	WCAG 2.2 AA checks and analyst UAT
Responsible model use	AI-01-06; NFR-ETH-01	Model card; model and evaluation views; registry target	Disclosure, evidence interpretation and promotion checks

Requirement status summary

Class	Implemented local reference	Production target
Functional	Ingestion, validation, features, detectors, consolidation, dashboard, local review, evaluation	Retries/manifests/quarantine, registry lineage, authenticated identity, atomic event-driven publication
Non-functional	Modularity, deterministic tests, read-only fallback, audit history, public-data boundary, responsible-use labeling	SLOs, load/security/accessibility tests, private networking, observability, backups, governed recovery

Appendix Appendices

A.1 Reason-code catalogue

Code	Detector	Meaning	Default action
INVALID_AMOUNT	Rule	Required amount is missing or non-numeric	Confirm source value and reload affected record
MISSING_KEY	Rule	Required key field is blank	Recover key before downstream use
INVALID_CATEGORY	Rule	Category outside controlled vocabulary	Map or update contract with approval
DUPLICATE_GRAIN	Rule	Multiple records share expected grain	Reconcile duplicates before aggregation/reporting
NEGATIVE_AMOUNT	Rule	Negative amount may be correction/reversal	Validate against source documentation; preserve sign if legitimate
TOTAL_MISMATCH	Rule	Reported total exceeds reconciliation tolerance	Trace component/total semantics and source evidence
INVALID_PERIOD	Rule	Unsupported fiscal-period format	Correct period before time-series analysis
YEAR_OVER_YEAR_SHIFT	Statistical	Large comparable-year movement beyond transparent gates	Review historical/source context and materiality
MULTIVARIATE_ANOMALY	Isolation Forest	Unusual combined feature profile within segment	Inspect top feature deviations and corroborating evidence

A.2 Run-state model

State	Meaning	Dashboard behavior
started	Orchestration accepted run	Continue last-good publication
ingested	Complete source snapshot promoted	Continue last-good publication
validated	Schema/financial controls complete	Continue last-good publication
scored	Feature and detector outputs complete	Continue last-good publication
approved	Publication gates passed	Eligible to publish
published	Curated outputs activated	Show new run
failed	Required stage or gate failed	Show last-good with stale/failure notice
quarantined	Evidence retained for investigation	Do not publish

A.3 Glossary

Term	Definition
Alert	Consolidated analyst queue item representing one source record and one or more detector signals.
Signal	Output from one rule, statistical control, or ML detector.
Corroboration	Independent detector families flagging the same source record.
Materiality	Relative financial magnitude within a documented comparison segment.
Contamination	Isolation Forest assumption controlling expected unusual share; not an error rate.
Record key	Stable identifier derived from normalized business grain.
Last-known-good	Most recent run that passed required publication gates.
Quarantine	Isolated evidence that failed ingestion or contract gates and cannot enter active publication.
False positive	A signal judged not review-worthy under the intended control policy.
Legitimate exception	Unusual but valid activity that was still appropriate to surface for review.
Target	Required future behavior; not claimed as deployed.

A.4 Source evidence reviewed

1. Repository README and executable pipeline source.

2. Data dictionary, data-quality report, model card, ingestion reliability guide, manual-review guide, dashboard specification, backend review workflow, operations runbook, Azure target architecture, and productionization plans.
3. Run, evaluation, review, signal, alert, and review-state artifacts dated through 20 August 2026.
4. Automated ingestion, controls, evaluation, review-backend, and dashboard contract/build tests.

A.5 Independent-project disclaimer

IMPORTANT. This is an independent portfolio prototype built with publicly available World Bank Group data. It is not an official World Bank Group system, does not use internal WBG infrastructure or non-public financial records, and is not approved for operational financial decision-making.

END OF DOCUMENT