subatomic

0. Project Executive Context Summary

Dimension	Description (Derived Strictly from Project Context)
Project Name	AI Co-Worker Agent Solution for Account Opening
Business Domain	Financial Services – Wealth Management Account Onboarding
Problem Being Solved	Automate and support high-volume, auditable account opening workflow for new clients, including intake, validation, custodian form assembly, review, signature, and write-back
Primary Goal(s)	Deliver AI Co-Worker agentic workflow to process ~7,000 new accounts accurately, consistently, and auditable, from intake through custodian submission and system write-back
Desired Outcome(s)	Efficient, compliant, transparent onboarding; reduced manual effort; full workflow auditability; synchronized lifecycle state across systems; robust handling of high structural variability in data
In-Scope Data Consumers	AI Co-Worker Agent Team (direct consumer), Human-in-the-loop reviewers (within workflow), Custodian platforms (data packets), Salesforce/Edge system (write-back), Subatomic Deep Lens (audit/trace)
Out-of-Scope Capabilities (Explicitly Excluded)	BI, analytics dashboards, ML training, reporting not required for agentic workflow, vector search, external data sharing, classical third-party BI/Reporting consumers
Non-Goals	General analytics/ML enablement, system modernization, business intelligence, ad hoc querying, external (out-of-ecosystem) data transmission
Key Constraints (Technical / Operational / Compliance)	All components within NorthRock IT security/infrastructure controls; auditability/transparency required at each step; must persist raw and normalized data for compliance; handle unstructured/structured data as-is; high schema and format variability; full traceability; compliance with operational SOPs; no external handling
Success Criteria	1) Full AI Co-Worker-driven orchestration of end-to-end workflow; 2) Traceable, auditable handling of all inputs/outputs; 3) Demonstrably accurate, complete custodian packet assembly and submission; 4) Write-back of lifecycle/control data to Salesforce/Edge; 5) Support for human-in-the-loop; 6) Audit validation via Subatomic Deep Lens; 7) System meets security/ compliance requirements

1. Layer Responsibility and System Inclusion (Context-Bound)

Layer	Why This Layer Exists in This Project	Included Systems / Platforms (Data-Related; Gold may include AI Co-Workers)	Allowed Operations	Explicitly Excluded Operations	Notes
Bronze	Preserve all ingested client, custodian, and workflow (including audit) data in its received/raw form for traceability and compliance.	Salesforce/Edge, Formstack, Schwab Custodial Platform, Fidelity Custodial Platform, DocuSign, Black Diamond (optional), Client-provided Docs, Subatomic Deep Lens	Ingestion, raw object storage, schema preservation, versioning	Up-front data normalization/ transformation, field mapping, deduplication	Full format/field fidelity to source data. Handles structured, semi-, and unstructured inputs.
Silver	Normalize, classify, and contextually merge/align client intake data and supporting docs to client/household for validation and packet assembly per workflow logic.	Data Lakehouse/Warehouse platform (if required by normalization), transformation engine/ process scheduler	Contextual field mapping, data normalization, validation checks, contextual linking to client/household, assembly of packet structure	Derivation not supporting agentic validation/packet build, analytics queries, BI-aggregation, inference	Directly supports Gold (Agent workflow, packet build, validation). Raw data retained for audit.
Gold	Produce shaped data packets, forms, and workflow outputs for direct consumption by AI Co-Worker Agent Team and custodian/write-back systems; enable traceable audit outputs for operational review and compliance	AI Co-Worker Agent Team (orchestrator/agents), Packet Assembly/Doc Generator, DocuSign, Salesforce/Edge (write-back interface), Subatomic Deep Lens (audit store)	Data shaping for workflow/agent events, workflow state emission, audit trace/log structuring, artifact generation, status/write-back ops	Direct reporting, BI models, anything not in workflow; external API exposure; ML model outputs	Only includes data products directly needed by agent workflow and assigned human-in-the-loop actions.

2. Staging Layer Design (Only If Required by Context)

Staging Area Name	Associated Layer	Included Systems / Platforms	Purpose in This Project	Data Scope	Retention Policy	Notes
Form/Doc Staging	Bronze	Temporary Object/Blob Storage	Buffer uploads and intake from Formstack, client-provided docs, DocuSign.	Unstructured PDFs, forms, envelope raw	Until load confirmation	Used to decouple high-volume/manual upload burst from persistent Bronze.
Audit Event Queue	Bronze/Silver	Event Queue/Log (as needed)	Buffer and sequence audit events and workflow traces from Subatomic Deep Lens.	Workflow trace events, audit logs	Until ingest confirmation	Optional, to ensure no loss during high-frequency workflow event ingestion.

subatomic

3. Bronze Layer Design (Raw Ingestion Scope)

Dataset / Domain	Source(s)	Included Systems / Platforms	Storage Type	Schema Handling	Partitioning / Versioning	Lineage Requirements	Notes
Client/HH Records	Salesforce/Edge	Object Storage, RDBMS, Data Lake	Structured	As-ingested, no modification	By record ID, timestamp	Source system, extract time	System of record; supports reference and write-back.
Intake Forms	Formstack	Object Storage, Data Lake	Semi-structured	Preserve input structure	By submission/ event ID, date	Formstack, submission time	Schema varies per template; field completeness tracked.
Custodial Form Libraries	Schwab, Fidelity Custodial Platforms	Object Storage, Data Lake	Structured	As provided by custodian	By custodian, snapshot time	Custodian, form/ rule version	Differ by account type/custodian; versioning as provided.
DocuSign Envelopes	DocuSign	Object Storage, Data Lake	Semi-structured	Preserve envelope structure	By envelope/ event ID, time	DocuSign, envelope time	Includes envelope audit, routing, and artifact.
Statements (optional)	Black Diamond (optional), Client-provided Docs	Object Storage, Data Lake	Unstructured	Blob/PDF, no change	By document/ upload time	Source, upload event	Only ingested if provided/available.
Audit/Event Logs	Subatomic Deep Lens	Object Storage, Data Lake	Semi-structured	Preserve as generated	By event/trace/ timestamp	Deep Lens, originating event	Workflow/audit step traces, scoring, lineage info.

4. Silver Layer Design (Only If Transformation Is Required)

Dataset / Domain	Input Layer	Included Systems / Platforms (Data-Related Only)	Transformation Necessity (Required / Not Required)	Transformations Justified by Context	State Required	Notes
Intake Form Normalization	Bronze	Data Lakehouse/ Warehouse, ETL/ ELT engine	Required	Field mapping (schema normalization), client/household association, completeness/ validation alignment per custodian rules	Yes	Enables automated validation, supports packet assembly.
Document Classification	Bronze	Data Lake, Content Classifier/Parser	Required	Document type determination, associating supporting docs to intake, basic metadata extraction	Yes	Ensures correct doc linkage for validation/ packet build.
Custodian Form Alignment	Bronze	Data Lake, Rule Processing Engine	Required	Cross-reference/merge intake data with current custodian form/rule set for accurate packet assembly, validation logic checks	Yes	Required for custodian-specific packet building logic in Gold.
Audit Event Linking	Bronze	Data Lake, Audit Trace Processor	Required	Sequence/align audit events to workflow steps, enrich with metadata as needed	Yes	Ensures all workflow steps and reviews are traceable and auditable.
No other transformations	--	--	Not Required	--	--	All transformations must directly support packet build, validation, audit, or workflow requirements.

5. Gold Layer Design (Consumer-Driven, Context-Only)

Data Product / Output	Input Layer	Included Systems / Platforms (Data-Related; AI Co-Workers allowed if consumers)	Intended Consumer (As Stated)	Shaping Required	Refresh Strategy	Notes
Validated Account Packet	Silver	AI Co-Worker Agent Team, Packet Assembly Service, Custodial Form Generator	AI Co-Worker Agent Team, Custodian	Assemble, pre-populate per rules	Per account event	Full custodian-ready packets for submission or human-in-the-loop review; generated only on demand.
Workflow State/ Exception Feed	Silver	AI Co-Worker Agent Team, Workflow State Store	AI Co-Worker Agent Team, Reviewer	Normalize exception/ review data	Real-time/ event-based	Supports agent coordination, human-in-the-loop actions, and exception mgmt.
Signature Envelope Artifact	Silver	AI Co-Worker Agent Team, DocuSign	AI Co-Worker Agent Team, DocuSign	Envelope structuring	Per workflow step	Only triggered if signature required for packet; routed by AI/agent logic.

Data Product / Output	Input Layer	Included Systems / Platforms (Data-Related; AI Co-Workers allowed if consumers)	Intended Consumer (As Stated)	Shaping Required	Refresh Strategy	Notes
Audit Workflow Trace/Log	Silver	Subatomic Deep Lens Audit Store	Subatomic Deep Lens, Compliance/ Review	Structured trace/event log	Continuous, step-based	For full workflow traceability, compliance, and operational review.
Lifecycle/ Status Write-Back	Silver	AI Co-Worker Agent Team, Salesforce/Edge Write-back Service	Salesforce/Edge	Structured update formatting	Per status change	Ensures lifecycle state synchronized with system of record.

6. Serving Stores and Specialized Layers (Only If Explicitly Needed)

Serving Layer Type	Backing Layer	Included Systems / Platforms	Why This Store Is Needed for This Project	Access Pattern	Notes
Audit/Trace Store	Gold	Subatomic Deep Lens	Ensure full traceability, audit log review, compliance monitoring	Event/Step-wise read	Only audit/trace store for operational/compliance review.
Packet/Artifact Store	Gold	Object Storage/ Content Repository	Persist generated packets/artifacts for agent/human hand-off/retrieval	Per workflow event	Stores only workflow outputs needed by agents/humans.
State/ Exception Feed	Gold	Workflow/Event State Store	Enable agent and human interaction with ongoing workflow/exceptions	Real-time/ event	Only scoped to agent/human coordination; not for BI.

7. End-to-End Data Flow (Context-Constrained)

Flow Step	From Layer / System	To Layer / System	Trigger / Condition	Data State	Transformations Applied	Alignment to Project Goal	Notes
1. Intake	Source Systems (Formstack, Docs, etc.)	Bronze	Submission/ upload event	Raw	None (preserve as-is)	Ingest all required data, ensure original fidelity	Temp staging as needed.
2. Normalization & Classification	Bronze	Silver	On new intake event	Normalized	Field mapping, doc association, completeness checks	Prepares data for automated validation and packet assembly	Schema-flexible, traceable.
3. Validation/ Alignment	Silver	Silver	On normalization complete	Validated/ linked	Custodian/account rule validation, cross-doc consistency	Ensures packets meet submission/ readiness requirements	Contextual to each account/ form.
4. Packet Build (Assembly)	Silver	Gold	On validation pass	Assembled packet	Form pre-population, summary artifact creation	Enables downstream automated agent/ human workflow	Generates custodian-ready packets.
5. Review/ Exception Handling	Gold	Gold	If exception/ ambiguous	Workflow state	Normalize state/ exception, log action/ decision	Fully supports required human-in-the-loop review/ approval	Agent/human action captured.
6. Signature	Gold	DocuSign	If signature required	Envelope artifact	Envelope build/routing	Generates/routes envelopes only as needed	Only if relevant for workflow.
7. Submission/ Write-back	Gold	Salesforce/ Edge, Custodian platform	On packet completion/ status	Status/ artifact	Lifecycle state update, submission artifact	Closes loop—state synchronized, data written back as per rules	Write-back via agent workflow.
8. Audit Logging	Gold	Subatomic Deep Lens	Every workflow step/action	Immutable trace	Audit, scoring, lineage event structuring	Enables traceability, operational/ compliance review	Full audit at each stage.