

## 1. Source Inventory

Source Name	Source Type (File / API / System / Report)	Delivery Mechanism	Format(s)	Frequency	Provider / System Owner	Notes
Salesforce / Edge	System	Direct system integration	Unknown	Unknown	NorthRock	System of record for client and household data
Formstack	System	Platform web forms	Unknown	Per client intake	NorthRock	Used for client intake forms
Schwab Custodial Platform	System	Direct/platform integration	Unknown	Account opening	Schwab	Used for form libraries and account rules
Fidelity Custodial Platform	System	Direct/platform integration	Unknown	Account opening	Fidelity	Used for form libraries and account rules
DocuSign	System	Envelope generation/routing	Unknown	Per document trans.	DocuSign	Used for signatures (subject to discovery confirmation)
Black Diamond	System (Optional)	Platform access	Unknown	As needed	Black Diamond	For statement retrieval and intake verification (optional)
Client-provided Statements/Docs	File	Manual upload / Attachment (via Formstack or other means)	Unstructured (PDF, etc.)	Per account intake	Client	Supporting documents for intake; structure not specified
Subatomic Deep Lens	System	Platform/module	Unknown	Audit events	Subatomic	Used for auditability (workflow traces, scoring, lineage, etc.)

## 2. Raw Data Structure Identification

Source Name	Entity / Dataset Name	Field Name	Data Type (As Observed)	Hierarchy / Nesting	Required / Optional	Example Value
Salesforce / Edge	Client/Household Record	Not specified	Not specified	Not specified	Not specified	-
Formstack	Intake Form	Not specified	Not specified	Flat/Tabular	Varies	-
Schwab Custodial Platform	Form Libraries	Not specified	Not specified	Not specified	Not specified	-
Fidelity Custodial Platform	Form Libraries	Not specified	Not specified	Not specified	Not specified	-
DocuSign	Envelope	Not specified	Not specified	Not specified	Not specified	-
Black Diamond (optional)	Statement	Not specified	Not specified	Not specified	Optional	-
Client-provided Statements/Docs	Intake Document	Not specified	Unstructured	N/A	Optional	PDF document
Subatomic Deep Lens	Audit Log	Workflow Trace	Not specified	May be nested	Required	-
Subatomic Deep Lens	Audit Log	Scoring Outcome	Not specified	May be nested	Required	-
Subatomic Deep Lens	Audit Log	Data Lineage	Not specified	May be nested	Required	-

## 3. Structural Variability & Inconsistencies

Source Name	Variability Type (Schema / Format / Value)	Description	Affected Fields	Occurrence Pattern	Impact
Formstack	Schema	Varies by intake form template	All form fields	Per form type/version	Impacts field mapping and completeness checks
Client-provided Statements/Docs	Format	Unstructured, possible variability in document type/format	Whole document	Per client submission	Requires flexible handling, cannot guarantee fields
DocuSign	Schema/Format	Envelope content and fields vary by process stage and doc requirements	Envelope/document fields	Per signature process	Introduces optionality and missing fields
Black Diamond (optional)	Value/Presence	Statement retrieval only if provided/allowed	Statement fields	Optional	Causes conditional pipeline steps

Source Name	Variability Type (Schema / Format / Value)	Description	Affected Fields	Occurrence Pattern	Impact
Custodial Platforms (Schwab, Fidelity)	Schema/Format	Account rules and form libraries may differ per custodian/account type	Form and rule fields	Per custodian/account	Field and validation logic changes
Subatomic Deep Lens	Schema	Audit output granularity may depend on system configuration/trace depth	Audit fields	Varies	Level of detail may vary in logs/traces

#### 4. Raw Data vs Implicit Metadata Separation

Source Name	Element Name	Category (Raw Data / Implicit Metadata)	How It Is Derived or Provided	Notes
Formstack	Form Fields	Raw Data	Provided by client via intake form	Actual client input
Formstack	Submission Timestamp	Implicit Metadata	System-generated on submission	May not be part of intake payload
Salesforce / Edge	Client Data Fields	Raw Data	Maintained as system of record	-
Salesforce / Edge	Last Modified Timestamp	Implicit Metadata	System-generated field	-
Client-provided Statements/Docs	Document Content	Raw Data	Uploaded/uploaded by client	Unstructured, original artifact
Schwab/Fidelity	Form Field Values	Raw Data	Populated in forms per custodian rules	-
Schwab/Fidelity	Account Rule Definitions	Implicit Metadata	Maintained by custodial platform	Used for validation, not client-supplied
Black Diamond (optional)	Statement PDF	Raw Data	Pulled from platform, provided document	Used as evidence/input for verification
Black Diamond (optional)	Retrieval Timestamp	Implicit Metadata	System-generated	Not provided by client
DocuSign	Envelope Fields	Raw Data	Generated/formally assigned during workflow	Only if envelope is triggered
DocuSign	Envelope Audit Trail	Implicit Metadata	System-generated	Event log of routing, delivery, signature times
Subatomic Deep Lens	Workflow Trace	Raw Data	Logged during process execution	Step-by-step event log
Subatomic Deep Lens	Trace Collection TS	Implicit Metadata	System-generated	Technical log attributes

#### 5. Source Classification Summary

Source Name	Structural Stability	Level of Standardization	Expected Evolution Risk	Suitable for Raw Preservation (Yes / No)	Comments
Salesforce / Edge	High	High	Medium	Yes	Stable system of record, but may change with integration updates
Formstack	Medium	Medium	High	Yes	Forms/templates can change over time
Schwab Custodial Platform	Medium	Medium	High	Yes	Forms/rules may change per custodian update
Fidelity Custodial Platform	Medium	Medium	High	Yes	Same as above
DocuSign	Low	Low	High	Yes	Document schema varies by process and contract stage
Black Diamond (optional)	Medium	Medium	Medium	Yes	Structure may change based on statement provider
Client-provided Statements/Docs	Low	Low	High	Yes	Format/structure not controlled
Subatomic Deep Lens	Medium	Medium	Medium	Yes	Audit detail may change with process changes

## 6. Data Infrastructure Scope Impact

Aspect	Observation Based on Sources
Total Number of Distinct Sources	7 (Salesforce/Edge, Formstack, Schwab Custodian, Fidelity Custodian, DocuSign, Black Diamond, Client-provided Statements/Docs, plus Subatomic Deep Lens for auditability)
Dominant Source Types	Enterprise systems and line-of-business platforms; unstructured file/document inputs
Overall Structural Complexity	High: Mix of structured system data, semi-structured form data, highly variable/unstructured documents
Degree of Schema Variability	High across forms, custodial platforms, optional sources, and client-provided docs
Elements That Must Be Preserved As-Is	All raw form inputs, client-uploaded documents, audit logs, generated artifacts (e.g., DocuSign envelopes, statements)
Constraints Imposed on Landing (Bronze) Layer	Must support variable schema, unstructured documents, high optionality, retention of all original fields without normalization