FI Data Lineage Model

 $Walter\ Hjelmar\ walter.hjelmar@synechron.com$

Table of contents

Fixed Income - Data Lineage Model Documentation	2
Core model concepts	2
Nodes	2
Relationships	3
Model Diagrams	4
Full model diagram	4
Simplified model diagram	6
BusinessAttribute	8
BusinessProcess	8
BusinessTerm	9
ComplianceOfficer	10
CompliancePolicy	10
DataAsset	11
DataClassification	12
DataElement	13
DataQualityIssue	14
DataQualityMetric	15
DataQualityRule	15
DataSet	16
DataSteward	17
ProcessSteward	17
Relationships	17
Regulation	18
SystemOfRecord	19
Transformation	19
Relationships	20
ACCOUNTABLE_FOR:	20
AGGREGATES:	20
APPLIES	21
AUTHORITATIVE_SOURCE	21
CLASSIFIED_AS	21
CONCLIMED DY	0.1

CONTAINS	2
DERIVED_FROM	2
GENERATED_BY	3
GOVERNED_BY	3
HAS_ATTRIBUTE	3
IMPACTS	3
INVOLVES	4
MANAGES_PROCESS	
MAPS_TO	
MEASURES	5
RESPONSIBLE_FOR	5
TRANSFORMS_TO	_
UTILIZES	6
VIOLATES	6

Fixed Income - Data Lineage Model Documentation

This data lineage model is designed to track the flow of data across systems, applications, and business processes in a financial services trading firm. The model includes nodes and relationships that cover data governance, regulations and compliance, data quality, and data security.

Core model concepts

Nodes

- Application: Represents systems or applications that are part of the data flow.
- BusinessAttribute: Represents attributes of a BusinessTerm (e.g., Customer ID, Customer Name, Email).
- BusinessKeyActivity: Represents a sub-step of a BusinessProcess.
- BusinessProcess: Represents activities or sets of activities that accomplish a specific organizational goal.
- BusinessTerm: Represents an entity or object that a BusinessProcess operates on (e.g., Customer).
- ComplianceOfficer: Represents an individual responsible for ensuring regulatory compliance and overseeing the firm's adherence to data governance policies.
- Compliance Policy: Represents the firm's internal compliance policies and guidelines.
- DataAsset: Represents a data asset, such as a table, file, or database.
- DataClassification: Represents a logical grouping of types DataElements (e.g. email address, phone number, country code).
- DataConfidentiality: Represents data security classification levels (e.g., Confidential, Restricted, Public).
- **DataElement**: Represents the implementation of a BusinessAttribute in a DataSet (e.g., a field in a table).

- DataQualityIssue: Represents identified data quality problems or violations.
- **DataQualityMetric**: Represents quantifiable measures of data quality, such as completeness, accuracy, or timeliness.
- DataQualityRule: Represents rules that define data quality expectations and requirements.
- DataSet: Represents a logical aggregation of DataElements for common use.
- LineOfBusiness: Represents an organizational unit of the business.
- ProcessSteward: Represents the process equivalent of the data steward.
- Regulation: Represents specific regulatory requirements applicable to the financial services industry (e.g., GDPR, MiFID II, Dodd-Frank).
- **SystemOfRecord**: Represents the authoritative data source for a given data element or piece of information.
- TradeLifecycle: Represents the different stages that a trade goes through, from placement to settlement.
- Transformation: Represents a data transformation process or job.Represents a data transformation process or job.

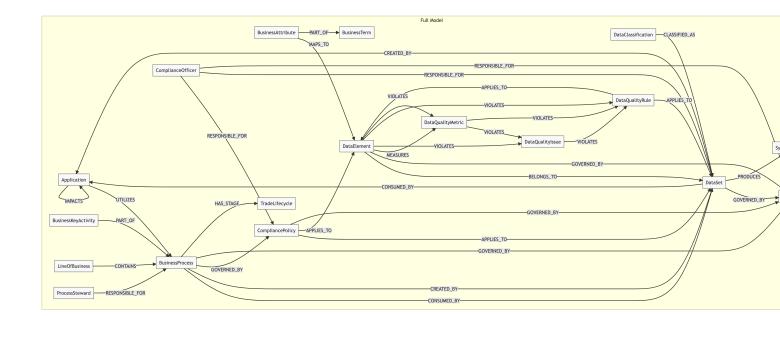
Relationships

- **APPLIES_TO**: Connects a DataQualityRule to a DataElement or DataSet that it applies to.
- **CLASSIFIED_AS**: Connects a DataSet to a DataConfidentiality, indicating its security classification level.
- **CONSUMED_BY**: Connects a DataSet (source data) to an Application or BusinessProcess that consumes the data without creating or transforming it.
- **CREATED_BY**: Connects a DataSet (output data) to an Application or BusinessProcess that creates or transforms the data.
- **GOVERNED_BY**: Connects a DataElement, DataSet, or CompliancePolicy to a Regulation, indicating which regulatory requirements apply.
- IMPACTS: Connects Applications to show the downstream impact of changes.
- MAPS TO: Connects a BusinessAttribute to a DataElement.
- **MEASURES**: Connects a DataQualityMetric to a DataElement or DataSet that it measures.
- PART_OF: Connects a BusinessAttribute to its corresponding BusinessTerm.
- **PRODUCES**: Connects a SystemOfRecord to a DataSet that it produces as output.
- **RESPONSIBLE_FOR**: Connects a ProcessSteward, ComplianceOfficer, or DataSteward to a BusinessProcess, Regulation, CompliancePolicy, or DataAsset, indicating their responsibility for ensuring compliance, data governance, or process oversight.
- UTILIZES: Connects a BusinessProcess to an Application that is used to accomplish the process.
- **VIOLATES**: Connects a DataQualityIssue to a DataElement, DataSet, DataQualityRule, or DataQualityMetric that it violates.

Model Diagrams

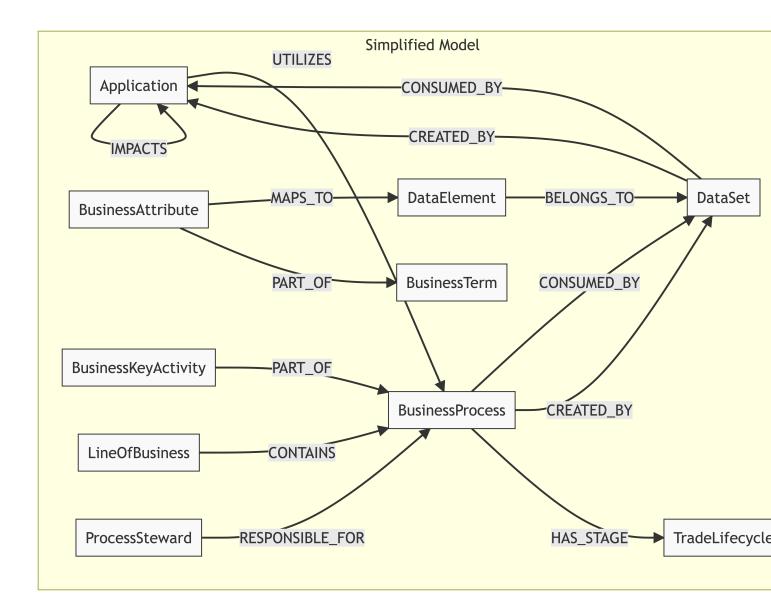
Full model diagram

The model that we have created is of a higher fidelity than the data that we currently have access to, but to get a sense of the scope of the model, we present a graph that contains the full model for contextualization.



Simplified model diagram

Here is the simplified version of the model that will be used as the version 1, Minimal Viable Product, of the model.



end

Node Definitions

```
### Application
Represents a system or application that is used to process data.
#### **Properties**
- id
- name
#### **Relationships**
```{mermaid}
%%| label: fig-application-relationships
%%| fig-width: 7
%%| fig-cap: |
 Application Relationships
%% l
graph LR
subgraph "Application Relationships"
 A[Application]:::entity
 BP[BusinessProcess]:::entity
 DS[DataSet]:::entity
 A --IMPACTS--> A
 BP --UTILIZES--> A
 DS --CONSUMED BY--> A
 DS --CREATED_BY--> A
classDef entity fill:#f9f9f9,stroke:#333,stroke-width:1px;
classDef relation fill:#EFEFEF,stroke:#666,stroke-width:1px;
end
```

- CONSUMED\_BY Connects a DataAsset (source data) to an Application or BusinessProcess
- IMPACTS Connects an Application to other related assets, such as other Application, DataAsset or DataElement
- UTILIZES Connects a BusinessProcess to an Application

```
1
2 (app:Application {
3 id: 'app1',
4 name: 'NAPA'
5 })
```

#### **BusinessAttribute**

Represents a property or characteristic of a BusinessTerm.

#### **Properties**

- id
- name
- description

#### Relationships

- HAS\_ATTRIBUTE Connects a BusinessTerm node to a BusinessAttribute
- MAPS TO Connects a BusinessAttribute node to a DataElement node

#### **Example**

```
(ba:BusinessAttribute {
 id: : 'ba1',
 name: 'Customer ID',
 description: 'Unique id: for a customer'
})
```

#### **BusinessProcess**

Represents a business process or set of activities that accomplish a specific organizational goal.

#### **Properties**

- id
- name
- description

#### Relationships

- CONSUMED\_BY Connects a DataAsset (source data) to an Application or BusinessProcess
- INVOLVES Connects a BusinessProcess node to a DataAsset node or a Transformation node
- MANAGES\_PROCESS Connects a ProcessSteward node to a BusinessProcess node
- UTILIZES Connects a BusinessProcess to an Application

#### **Example**

```
(bp:BusinessProcess {
 id: 'bp1',
 name: 'Customer Onboarding',
 description: 'Process for onboarding new customers'
})
```

#### **BusinessTerm**

Represents an entity or object that a BusinessProcess operates on.

#### **Properties**

- id
- name
- description

# Relationships

• HAS ATTRIBUTE - Connects a BusinessTerm node to a BusinessAttribute

```
(bt:BusinessTerm {
 id: : 'bt1',
 name: 'Customer',
 description: 'A customer entity in the organization'
})
```

### ComplianceOfficer

Represents an individual responsible for ensuring regulatory compliance and overseeing the firm's adherence to data governance policies.

#### **Properties**

- id
- name
- title

#### Relationships

• RESPONSIBLE\_FOR - Connects a ComplianceOfficer to a Regulation, CompliancePolicy, or DataAsset

#### Example

```
(co:ComplianceOfficer {
 id: : 'co1',
 name: 'John Doe',
 title: 'Chief Compliance Officer'
})
```

### CompliancePolicy

Represents the firm's internal compliance policies and guidelines.

#### **Properties**

- id
- name
- description

### Relationships

- GOVERNED\_BY Connects a DataAsset or DataElement to a Regulation or Compliance-Policy
- RESPONSIBLE\_FOR Connects a ComplianceOfficer to a Regulation, CompliancePolicy, or DataAsset

#### Example

```
(cp:CompliancePolicy {
 id: : 'cp1',
 name: 'Data Retention Policy',
 description: 'Policy for data retention and disposal'
})
```

#### **DataAsset**

Represents a data asset, such as a table, file, or database. A type of asset that represents details of organizational data in two layers. One layer is independent of any particular technology for non-technical stakeholder communication. The other one is taking the implementation system for technical stakeholder communication into account.

#### **Properties**

- id
- name
- description
- type

#### Relationships

- ACCOUNTABLE\_FOR Connects a SystemOfRecord node to a DataAsset node
- APPLIES Connects a DataQualityRule to a DataAsset or DataElement
- CLASSIFIED AS Connects a DataAsset or DataElement to a DataClassification
- CONSUMED\_BY Connects a DataAsset (source data) to an Application or BusinessProcess
- CONTAINS Connects a DataAsset node to a DataElement node.
- DERIVED\_FROM Connects a DataAsset (representing the view) to another DataAsset (representing the source data)
- GOVERNED\_BY Connects a DataAsset or DataElement to a Regulation or Compliance-Policy
- MEASURES Connects a DataQualityMetric to a DataAsset or DataElement
- RESPONSIBLE\_FOR Connects a DataSteward node to a DataAsset node
- RESPONSIBLE\_FOR Connects a ComplianceOfficer to a Regulation, CompliancePolicy, or DataAsset
- VIOLATES Connects a DataQualityIssue to a DataQualityRule, DataQualityMetric, DataAsset, or DataElement

#### Example

```
(d:DataAsset {
 identifier: 'table1',
 name: 'Table 1',
 description: 'Sales data',
 type: 'table'
 })
```

#### **DataClassification**

Represents data security classification levels (e.g., Confidential, Restricted, Public).

#### **Properties**

- id
- name
- description

#### Relationships

• CLASSIFIED\_AS - Connects a DataAsset or DataElement to a DataClassification

#### Example

```
(dc:DataClassification {
 id: : 'dc1',
 name: 'Confidential',
 description: 'Highly sensitive data requiring strict access control'
})
```

#### **DataElement**

Represents a field (column) within a data asset.

#### **Properties**

- id
- name
- description
- dataType

#### Relationships

- APPLIES Connects a DataQualityRule to a DataAsset or DataElement
- CLASSIFIED\_AS Connects a DataAsset or DataElement to a DataClassification
- CONTAINS Connects a DataAsset node to a DataElement node.
- GENERATED BY- Connects a DataElement node to a Transformation node
- GOVERNED\_BY Connects a DataAsset or DataElement to a Regulation or Compliance-Policy
- IMPACTS Connects an Application to other related assets, such as other Application, DataAsset or DataElement

- MAPS TO Connects a BusinessAttribute node to a DataElement node
- MEASURES Connects a DataQualityMetric to a DataAsset or DataElement
- TRANSFORMS\_TO Connects a [DataElement node to another DataElement node
- VIOLATES Connects a DataQualityIssue to a DataQualityRule, DataQualityMetric, DataAsset, or DataElement

```
(de:DataElement {
 identifier: 'column1',
 name: 'CustomerID',
 description: 'Unique customer identifier',
 dataType: 'integer'
})
```

# DataQualityIssue

Represents identified data quality problems or violations.

#### **Properties**

- id
- description
- severity

#### Relationships

• VIOLATES - Connects a DataQualityIssue to a DataQualityRule, DataQualityMetric, DataAsset, or DataElement

```
(dqi:DataQualityIssue {
 id: : 'dqi1',
 description: 'Missing customer email',
 severity: 'Medium'
})
```

#### **DataQualityMetric**

Represents quantifiable measures of data quality, such as completeness, accuracy, or timeliness.

#### **Properties**

- id
- name
- description

#### Relationships

- MEASURES Connects a DataQualityMetric to a DataAsset or DataElement
- VIOLATES Connects a DataQualityIssue to a DataQualityRule, DataQualityMetric, DataAsset, or DataElement

#### Example

```
(dqm:DataQualityMetric {
 id: : 'dqm1',
 name: 'Accuracy',
 description: 'Measure of data accuracy'
})
```

#### **DataQualityRule**

Represents rules that define data quality expectations and requirements.

#### **Properties**

- id
- name
- description

#### Relationships

- APPLIES Connects a DataQualityRule to a DataAsset or DataElement
- VIOLATES Connects a DataQualityIssue to a DataQualityRule, DataQualityMetric, DataAsset, or DataElement

```
(dqr:DataQualityRule {
 id: : 'dqr1',
 name: 'Completeness Rule',
 description: 'Rule to ensure data completeness'
})
```

#### **DataSet**

A collection of related sets of data assets that are data elements or composed of data elements.

#### **Properties**

- id
- name
- description

#### Relationships

- ACCOUNTABLE FOR Connects a SystemOfRecord node to a DataAsset node
- AGGREGATES Connects a DataSet node to a DataElement
- APPLIES Connects a DataQualityRule to a DataAsset or DataElement
- IMPACTS Connects an Application to other related assets, such as other Application, DataAsset or DataElement
- INVOLVES Connects a BusinessProcess node to a DataAsset node or a Transformation node

```
(ds:DataSet {
 id: : 'ds1',
 name: 'CustomerDataSet',
 description: 'A dataset containing customer information'
})
```

#### **DataSteward**

Represents a person responsible for managing data assets and ensuring data quality and governance.

#### **Properties**

- id
- name
- email

#### Relationships

• RESPONSIBLE FOR - Connects a DataSteward node to a DataAsset node

#### **Example**

```
1
2 (s:DataSteward {
3 id: 'steward1',
4 name: 'Alice',
5 email: 'alice@example.com'
6 })
```

#### **ProcessSteward**

Represents a person responsible for managing and ensuring the quality and governance of a business process.

#### **Properties**

- id
- name
- email

#### Relationships

- ACCOUNTABLE FOR ProcessSteward node to a SystemOfRecord node
- $\bullet \ \ MANAGES\_PROCESS \ \ Connects \ a \ ProcessSteward \ node \ to \ a \ BusinessProcess \ node \\$

```
(ps:ProcessSteward {
 id: : 'psteward1',
 name: 'Bob',
 email: 'bob@example.com'
})
```

### Regulation

Represents specific regulatory requirements applicable to the financial services industry (e.g., GDPR, MiFID II, Dodd-Frank).

#### **Properties**

- id
- name
- description

#### Relationships

- GOVERNED\_BY Connects a DataAsset or DataElement to a Regulation or Compliance-Policy
- RESPONSIBLE\_FOR Connects a ComplianceOfficer to a Regulation, CompliancePolicy, or DataAsset

```
(r:Regulation {
 id: : 'r1',
 name: 'GDPR',
 description: 'General Data Protection Regulation'
})
```

# SystemOfRecord

Represents the authoritative data source for a given data element or piece of information.

#### **Properties**

- id
- name
- description

#### Relationships

• ACCOUNTABLE\_FOR - Connects a SystemOfRecord node to a DataAsset node

#### **Example**

```
(sor:SystemOfRecord {
 id: 'sor1',
 name: 'Customer Master Data',
 description: 'Authoritative source for customer information'
})
```

#### **Transformation**

Represents a data transformation process or job.

#### **Properties**

- id
- name
- description
- type

#### Relationships

- GENERATED BY- Connects a DataElement node to a Transformation node
- INVOLVES Connects a BusinessProcess node to a DataAsset node or a Transformation node

```
(t:Transformation {
 identifier: 'job1',
 name: 'Data Cleansing Job',
 description: 'Cleans and prepares data for analysis',
 type: 'ETL'
})
```

# Relationships

#### ACCOUNTABLE\_FOR:

Connects a ProcessSteward node to a SystemOfRecord node, indicating that the person is accountable for the authoritative data source.

#### Example

#### **AGGREGATES:**

Connects a DataSet node to a DataElement or node, representing the aggregation of DataElements for common use.

```
ds:DataSet)-[:AGGREGATES]->(de:DataElement)
```

#### **APPLIES**

Connects a DataQualityRule to a DataAsset or DataElement, indicating that the rule applies to the specific data.

#### **Example**

#### **AUTHORITATIVE\_SOURCE**

Connects a SystemOfRecord node to a DataAsset node, indicating that the data asset is the authoritative source for specific data elements.

#### Example

# CLASSIFIED\_AS

Connects a DataAsset or DataElement to a DataClassification, indicating the security classification level of the specific data.

#### **Example**

#### CONSUMED\_BY

Connects a DataAsset (source data) to an Application or BusinessProcess that consumes the data without creating or transforming it.

```
(d:DataAsset {type: 'Table'})-[:CONSUMED_BY]->(a:Application)
or

Example

``` cypher
(d:DataAsset {type: 'Table'})-[:CONSUMED_BY]->(bp:BusinessProcess)
```

CONTAINS

Connects a DataAsset node to a DataElement node.

Example

DERIVED_FROM

Connects a DataAsset (representing the view) to another DataAsset (representing the source data), indicating that the view is derived from the source data.

GENERATED_BY

Connects a DataElement node to a Transformation node, indicating that the field was generated by a particular transformation process.

Example

GOVERNED_BY

Connects a DataAsset or DataElement to a Regulation or CompliancePolicy, indicating that the data is subject to specific regulatory requirements or policies.

Example

HAS_ATTRIBUTE

Connects a BusinessTerm node to a BusinessAttribute node.

Example

```
1
2 (bt:BusinessTerm)-[:HAS_ATTRIBUTE]->(ba:BusinessAttribute)
3
```

IMPACTS

Connects an Application to other related assets, such as other Application, DataAsset or DataElement, to represent the influence or effect the Application has on these assets, particularly when changes occur within the Application .

INVOLVES

Connects a BusinessProcess node to a DataAsset node or a Transformation node, indicating that the process involves the use or modification of the data asset or transformation.

MANAGES_PROCESS

Connects a ProcessSteward node to a BusinessProcess node, indicating that the person is responsible for managing the business process.

Example

MAPS_TO

Connects a BusinessAttribute node to a DataElement node.

Example

MEASURES

Connects a DataQualityMetric to a DataAsset or DataElement, indicating that the metric is used to measure the quality of the specific data.

Example

RESPONSIBLE_FOR

Connects a ComplianceOfficer to a Regulation, CompliancePolicy, or DataAsset, indicating their responsibility for ensuring compliance.

```
co:ComplianceOfficer)-[:RESPONSIBLE_FOR]->(r:Regulation)
```

TRANSFORMS_TO

Connects a [DataElement node to another DataElement node, representing the transformation from one field to another in a data transformation process.

Example

UTILIZES

Connects a BusinessProcess to an Application that is used to accomplish the process.

Example

VIOLATES

Connects a DataQualityIssue to a DataQualityRule, DataQualityMetric, DataAsset, or DataElement, indicating that the issue represents a violation of the rule or metric or is related to the specific data.

```
(dqi:DataQualityIssue)-[:VIOLATES]->(dqr:DataQualityRule)
```