|  |  |  |  |
| --- | --- | --- | --- |
| **MF** | | |  |
|  | GMT Implementation Notes | |  |
|  |  | |  |
|  | Author: | Michael Wang | |
|  | Subject: | **The Reference Model** | |
|  | Date: | June 5, 2024 | |

As part of the GMT data model, this note will focus on the reference model as this is the core and therefore most critical component of the GMT data.

# Overview of the GMT data model

GMT will pull accounting data elements from many dissimilar sources and attribute them into GMT relevant line items that are defined and required by an external GMT calculation engine. Below we refer to these line items as Deloitte line items or fields. There could be more than 80 fields to be filled per legal entity.

This work will be executed in two steps. In the first step, commonly called the Curation Step, Manulife accounting data elements are collected, or computed, into Manulife line items defined by Manulife Tax Team. In the second step, these Manulife line items are mapped into the Deloitte line items following a set of business logic defined by the Tax Team.

Since Manulife accounting data reside in multiple accounting or book of records systems around the globe, it is critical, when pulling such data together, to be able to identify correct data elements so that they can be attributed to the Manulife line items, rolled up or mapped to a unique legal entity. This is achieved by creating the reference model. This model must ensure that we are able to identify all the relevant accounting elements from the various source systems uniquely and completely, i.e., no missing items and no double counting.

For GMT, there are three essential references.

1. Legal entity. GMT will be reported per legal entity per tax jurisdiction. However, legal entities are identified using keys formed differently across source systems. The reference model will connect them together.
2. Accounting element. An accounting element is a unique data element from the sources that needs to be attributed to a GMT line item. This is the atomic data element. The reference model will define the full dimensionality for such an item. As such, the reference model can also be called the dimension model.
3. Accounting unit. An accounting unit is a container of assets or liabilities and income or expenses subject to the same treatment. The reference model will allow us to attribute an accounting data element to the correct accounting unit.

This note will document the reference model as a relational model of data entities, their relations, and their defining attributes. and then show a relational data model diagram of the reference model

# Main data entities for reference

This section will describe the main entities, namely legal entity, account element, and accounting unit, and their relations.

## Legal entity

For tax filing purposes, the legal entity is globally identified by a four-character code with numerical digits, also referred to as the company code. This is the unique key in GL but not others.

A legal entity has a unique name, belongs to a unique jurisdiction, and reports tax to a functional currency. These three attributes are required when reporting for GMT.

### GL

From GL, a company can be identified by the company code, and along with it, a legal entity name, its jurisdiction, and its functional currency.

### LV

In LV, a company can be identified by the legal entity name sourced from GL (transformed to replace some characters) concatenated with the jurisdiction.

### SCD

SCD holding records use the 4-character company code for legal entity, and SCD transaction records use the company code concatenated with a long legal entity name.

## Account element

Tax is computed based on taxable income, expenses, and tax credits. An account element is a piece of information, a number, that is classified into one of these categories. An account element is described by one or more accounting attributes.

### GL

In GL, an account element is identified by an account name and an attribute for whether the reported amount belongs to policyholder or shareholder.

### LV

In LV, an account element is identified by an account name, an element name, such as “Deferred” or “Current,” and an income source, such as “P&L” or “OCI.”

### SCD

In SCD, an account element is identified by an accounting unit, which can be connected to GL account element through a metadata discussed below.

## Accounting unit

Accounting numbers are booked into accounting units. An accounting unit, in Manulife, is defined by the combination of a company and a subfund code. An accounting unit, for the GMT purpose, can be distinguished by whether it belongs to shareholder or policyholder.

**TODO**! However, the accounting unit does not have a universal identification method. Need to confirm the rule connecting SCD to GL!

### GL

In GL, accounting unit can be identified following several different patterns. The most basic one is a prefix “MSF” followed by the 3-digit Subfund code, e.g., MSF091. **But there are other patterns which may or may not be relevant**.

### LV

LV does not appear to use the concept of accounting unit as LV data is consumed externally and accounting unit is a Manulife internal concept.

### SCD

SCD also uses multiple patterns for accounting unit: the concatenation of the company code and a Subfund code.

## The company-accounting unit metadata

Of the three essential references, the account element identifies the tax nature of a data element, legal entity and accounting unit identify the ownership nature of the data element. These are relatively static, subject to modification infrequently. Such information forms the basic metadata. The values in this metadata define the business domain, and thus called domain values.

While the domain values of account elements can be directly extracted from GL, LV, SCD or other sources, the company-accounting unit combination can be predefined as a unique dataset to serve as the foundational reference data. This reference data will define the scope for Deloitte line items mapping and the linkage between the various source datasets for account ownership.

This metadata has the following form, and a sample is attached below.

|  |  |  |
| --- | --- | --- |
| Field Name | Description | Example |
| LEGAL\_ENTITY | The legal entity name as it appears in GL | 0042PTAJ SHARIA (0450) |
| COMPANY | The unique 4-chracter code for the company | 0450  Note that this is different from the first 4 characters in the name. |
| COMPANY\_DESCRIPTION | A descriptive string of text | PT. AJMI Sharia - Company Fund |
| COMP\_FUNC\_CURRENCY | Functional currency for the company | IDR |
| JURISD | The jurisdiction for which the company reports its tax | IDN |
| ACCOUNTING\_UNIT | The accounting unit code as it appears in GL | MSF009  Note that this can decomposed into a prefix “MSF” and a subfund code “009”. This is critical when connecting the GL data with SCD data |
| COMP\_ACCOUNTING\_UNIT | Concatenation of COMPANY + “ “ + ACCOUNTING\_UNIT | 0450 MSF009 |
| PAR\_OR\_SH | Policyholder vs Shareholder | Shareholder |
| PAR\_INDICATOR | A more granular indicator for distinct categories of POLICYHOLDER items | NONPAR  (non-policyholder, same as shareholder) |

### Validation assertions for the company-accounting unit metadata

We make the following assertions for this dataset, **after filtering out records where COMPANY >= ‘1000’ and LEGAL\_ENTITY == ‘NULL’**:

1. **The three fields < COMPANY, ACCOUNTING\_UNIT, PAR\_INDICATOR> uniquely identifies each record**. That is, each COMPANY can be associated to multiple combinations of ACCOUNTING\_UNIT and PAR\_INDICATOR. And each PAR\_INDICATOR has a unique PAR\_OR\_SH code.
2. **Moreover, a COMPANY has a unique combination of values <LEGAL\_ENTITY, COMPANY\_DESCRIPTION, JURISD,** **COMP\_FUNC\_CURRENCY>.** That is, a company cannot have multiple legal entity names, or multiple jurisdictions, or multiple functional currencies. **Note that this applies only to those companies where the company code is less than ‘1000’ and both the** **jurisdiction code and the currency code are valid ISO codes**.
3. The unique values of COMPANY provide a complete list of legal entities **relevant to GMT**. In other words, only those account elements in LV, SCD or other source data that belong to companies in this list will be included in GMT extraction.
4. The unique values of ACCOUNTING\_UNIT provide a complete list of accounting units **relevant to GMT**. In other words, holding data from SCD of any accounting units not on this list will not be included in GMT extraction.

These assertions will be validated at the ETL stage. A sample file is attached below that contains multiple exceptions to the above assertions and needs to be reviewed and cleaned up.

To validate against the assertions, the sample file also contains an example that follows these steps:

1. Using Excel’s Data.Filter.Advanced feature to select unique records of columns <LEGAL\_ENTITY, COMPANY, COMPANY\_DESCRIPTION, COMP\_FUNC\_CURRENCY, JURISD> and copy them to an empty space.
2. Next to them using Excel’s CountIf function to compute number of appearances of the COMPANY column values.

This can be easily implemented as a select statement in SQL:

SELECT [COMPANY], cnt=count(\*)

FROM (

SELECT

[LEGAL\_ENTITY], [COMPANY], [COMPANY\_DESCRIPTION], [COMP\_FUNC\_CURRENCY], [JURISD]

FROM [gmt].[company\_acct\_metadata]

WHERE [COMPANY] < '1000' and [LEGAL\_ENTITY] != ‘NULL’

and [JURISD] in (select [jurisd\_code] from [gmt].[gmt\_jurisdiction])

and [COMP\_FUNC\_CURRENCY] in (select [ccy\_code] from [gmt].[gmt\_currency\_code])

GROUP by [LEGAL\_ENTITY], [COMPANY], [COMPANY\_DESCRIPTION], [COMP\_FUNC\_CURRENCY], [JURISD]

) t

GROUP by [COMPANY]

HAVING count(\*) > 1

## Other reference domain values

There are a few auxiliary and static domain values used to refer to a particular dataset.

### Tax filing period

This is the date point on which taxes are reported. Typically, this is expressed as YYYYMM, and technically it could be tied to the last calendar day of the month if any date field calculation is required, e.g., a given number of months or quarters off the current filing period. This is pre-generated and each dataset, input, or output, must be tied to a unique filing period.

### Data instance type

Each dataset will also be associated with an instance type that defines the instance's nature. Currently the BRD only requires the final close of book instance of the tax data. But the numerical values can be either an absolute, snapshot amount, such as a holding record as of close of business of a month end (EOM), or of relative, cumulative nature associated with a period of cumulation, such as Year-To-Date (YTD), Quarter-to-Date (QTD), or Month-to-Date (MTD), or instead of an amount as recorded in a system, of some managerial adjustment (ADJ).

### Tax data source

This specifies the source of the data. Currently known data sources are GL, LV (LongView), SimCorp via IDDL/EDL, and file based (Excel, CSV, TXT, or collected via BI Applications such as PowerBI). While same data appear in more than one sources, this attribute will indicate which one is used. Thus, this will be a configuration control point.

# Manulife and Deloitte GMT fields

**The objective of the GMT data model** is to define *a list of tax fields from Manulife perspective based on the domain values* and *a mapping model to map the Manulife fields to those of the external GMT tax calculation engine, in this case, the Deloitte fields*. This will pave the way to deliver on the objective of the GMT project: extract relevant accounting elements from Manulife systems into *Manulife Fields* **per legal entity**, and then mapped to the *Deloitte Fields*, per legal entity, required by the external calculation engine.

While the Deloitte fields are defined and provided by Deloitte (see section 4.2), the Manulife fields will be defined by the Tax Team based on the reference model described above. This section presents this model.

## The GMT reference model and Manulife line items (or fields)

At the core of the GMT data model is the reference model. The reference model allows the tax team to unambiguously answer two essential questions for GMT:

1. For which legal entity is the income filed?
2. In which accounting element is the income booked?

The answer to the second question forms a Manulife item, or field, into which a value will be extracted.

The diagram below shows the reference model with the domain values. This model encapsulates all the essential relations and dimension values.

This model defines the two key data entities: legal entities and account elements. The former has been covered in the previous section, and the latter forms the basis for defining the Manulife line items or fields.

A Manulife field is uniquely defined by the following attributes (*subject to change*):

|  |  |  |
| --- | --- | --- |
| Field Name | Description | Example |
| Name | A unique name given to the field | GL\_Income\_Before\_Taxes\_SH |
| Caption | A descriptive text | IFRS17 Income Before Taxes for Shareholder from GL P&L |
| Account | A unique account name as it appears in the source data, GL, LV, or others. | Income Before Taxes  (Note, filtered on accounting basis=IFRS17) |
| DataSource | The source of the account information | GL |
| Par\_Sh | Policyholder vs Shareholder designation | SHAREHOLDER |
| Element | Account element as referenced in LV | NA |
| IncomeSource | Income source as referenced in LV | NA |

The diagram below shows the keys and unique indexes of the relational model.



Figure . GMT Reference Model

## Maintaining and updating the reference model

Other than the company-accounting unit metadata as the primary definitive source for legal entities and accounting unit as well as their relations, the data sources (GL, LV, and SCD) also provide supplementary source for maintaining and updating the reference model:

1. The legal entity (or company). While the company-accounting unit metadata is deemed as the most definitive source, there could be exceptions since the overall data flow is complicated and human errors occur. To ensure the data integrity, we adopt the following process:
   1. Build the initial baseline with all the three data sources along with the company-accounting unit metadata into the data model.
   2. At each quarter’s ETL, update the reference model using the company-accounting unit metadata generated from GL. During the update, identify all exceptions as outlined above.
   3. Send the exceptions to the tax team for resolution.
2. The accounting unit. The accounting unit is part of the company-accounting unit metadata, but since SCD uses a different pattern, exceptions could occur where some accounting units in SCD cannot be connected to data in GL. We follow the same procedure to maintain this relation as laid out above for legal entities.
3. The account element. Accounting elements form the basic line items for reporting. From different source systems we extract different accounting elements. Since the accounting elements (Account, Element, Income source, Policyholder.vs.Shareholder, etc.) are specific to each source system, they come directly from each source system, particularly GL and LV.

Thus, we use the following sources to maintain and update the reference model:

1. ISO 3166 standard for jurisdiction and currency codes.
2. Company-accounting unit metadata from GL as stored in INFOR for legal entities, accounting units, and Par indicators.
3. GL/LV for accounts, elements, income sources, etc.

The SQL script attached here is an example of the procedure for updating the domain values.



# Mapping Manulife data to Deloitte data extract

The reference model described above provides the foundation for defining Manulife fields. This section will discuss the reference model for mapping the Manulife fields to the Deloitte fields.

The Deloitte fields are defined as a single row, or a single column when transposed, form per tax entity with each form consisting multiple line items or *fields*. Each field will provide a container into which one or more Manulife line items will be used to calculate a value. Therefore, the mapping is a many-to-one mapping, i.e., many Manulife items to each Deloitte field. The mapping is mostly a linear one with some exceptions that involve non-linear operations. In this model, however, we only define the linear mapping. If any non-linear operations are required, such as excluded dividend calculation, they will be handled separately to form some Manulife items which will then be linearly mapped into a Deloitte field.

## Manulife line items

A Manulife line item, per legal entity, is uniquely defined in the reference data model by the combination of < account\_id, par\_sh\_id, acct\_element\_id, income\_source\_id> with a unique name, as described in Section 3.1. The table is shown below, and it will be populated with the user input supplied by the Tax Team.



Figure 2. Manulife line-item definition

## Deloitte line fields

Deloitte collects a set of values, per legal entity, of 80+ fields. Each Deloitte field is uniquely identified by a field name. Each field has the following attributes to further describe the field and its usage or relation to an Excel field:

1. Deloitte\_field\_name. This is the name used by Deloitte to identify the value field. There are over 80 fields; see the table on the right in the diagram below.
2. Caption. This is the display text and description of the field.
3. Type. This indicates the data type of the field. Only two types are currently used: string and decimal.
4. Heading. This indicates tax category. There are three values: Entity, GloBE Income, and Covered Taxes.
5. Display order. This is the preferred order for the field when it is presented as a line item. It is not relevant to the extract.
6. Excel field id. This is a reference number to link the field to its corresponding field in the Excel APP. Not relevant to data extract.
7. Excel description. Text displayed in Excel. Not relevant to data extract.

These fields are the cells in a form to be filled, and the form per legal entity is uniquely identified by field en\_code. This corresponds to the legal\_entity in the GMT data model, referenced by legal\_entity\_id.

The tables in the below diagram show the Deloitte extract requirement.



Figure . Deloitte data extract requirement

## GMT mapping for final extraction

The tax mapping model depends on the two lists of fields, one for Deloitte and one for Manulife. The Deloitte is defined externally, and they are collected into a table gmt\_deloitte\_items shown above. The Manulife items will be defined by the Tax Team, and they will be collected into table gmt\_mli\_items, as described in Section 3.1.

With these two lists, we can define the linear mapping between them with the following relationship[[1]](#footnote-2):

A deloitte\_field equals the sum of (mli\_field \* gmt\_mapper\_multiplier).

The corresponding fields and the multipliers are collected in table gmt\_mapper\_multiplier.

The mapping, as a user input, will thus have the following form:

|  |  |  |
| --- | --- | --- |
| Column name | Description | Example |
| DeloitteFieldName | The unique Deloitte field name | fin\_PBT |
| MliFieldName | The unique Manulife field name | GL\_Income\_Before\_Taxes\_SH |
| Multiplier | The multiplier, a signed number, that is multiplied to the value of the MliField before adding to the sum for the corresponding Deloitte field. | 1 (i.e., simply add the Manulife item);  -0.5 (i.e., subtract half of the Manulife item) |

The diagram below shows the mapping model for extracting Manulife items into the output required by Deloitte. Note that in this model, three components are subject to change from time to time: first, the list of Deloitte fields, then accordingly, the Manulife fields that are defined based on the Deloitte fields, and finally, given the two lists of fields, the mapping that tie the two together.

Figure . The GMT Mapping Model

The diagram below is a close-up view of the mapping reference model. Table gmt\_mapping\_table is the user input with three columns <DeloitteFieldName, MliFieldName, MliFieldOperation> with additional id columns to be resolved during ETL: DeloitteFieldId and MliFieldId. This user table is used to create or update an existing GMT mapper.

The SQL script attached below shows the procedure for updating the mapper.



Figure . Manulife-to-Deloitte mapper

1. If we use symbol to represent a Manulife line item, a Deloitte field, then the linear mapping is a simple summation:

   where we have Deloitte fields and Manulife items, each multiplier is a real number, typically taking the form of 0, 1, and -1, but may be any real number. thus defines a matrix and the mapping will be a simple matrix operation. [↑](#footnote-ref-2)