Enabling global identity
Protecting digital trust

28th Meeting of the Wiesbaden Group on Business Registers

The Hague, 2 - 6 October 2023

Global Legal Entity Identifier Foundation (GLEIF)
Dominik Jany – Data Scientist
Zornitsa Manolova – Head of Data Quality Management and Data Science

Session No. 3

"Globalisation and Large Case Units"

Multi-national entity relationships and corss-platform connectivity on a global scale – Introduction to relationship records of LEI (Legal Entity Identifier) data.

1 Abstract

Collecting entity information across multiple jurisdictions is a challenge that requires cross-border collaboration and a consistently high level of standardization among organizations including business registers.

The Legal Entity Identifier (LEI) is a 20-digit alpha-numeric code based on the ISO 17442 standard. It connects to the key reference information that enables clear and unique identification of legal entities. The Global LEI Repository is the transparency island in a cloudy environment – it provides open, free-of-charge, high-quality legal entity data with global coverage. The LEI answers 'who is who' and 'who owns whom'. It serves as a linchpin between different data sources.

There are more than 2.3 Mio active LEIs in the system. The reference data of these LEIs is collected, verified, and managed by a network of LEI issuing organizations across the globe. The embedded global standards in the data format and the established data quality framework ensure consistency and high-quality data among the different organizations and jurisdictions.

Each LEI record carries information about its direct and ultimate consolidating parent. Entities not declaring the two mandatory parent relationships, provide so-called exception records with the explicit reason for the "missing" parent relationship record instead.

The LEI Repository offers additional information about connections among fund entities. These relationships portray umbrella structures, fund managers, and feeder funds across all jurisdictions.

The LEI could serve as a linchpin between different datasets and thus enable the smooth interoperability between ecosystems. With GLEIF's mapping initiative, the LEI is natively mapped to

the additional data standards and sources, such as the International Securities Identification Number (ISIN), the Bank Identifier Code (BIC), the Market Identifier Code (MIC), Open Corporates identifier, and S&P Global identifier.

Combining the LEI's relationship records and mappings provides data users and business registers with a holistic overview and additional insights about an entity's structure across jurisdictions.

In our paper, we showcase the usability of the relationship records within the LEI data by providing use cases for fund and non-fund entities. The paper demonstrates how data users can deep-dive into entities' relationships and enrich their data, obtaining the full picture by using minimum programmatic effort while benefitting from free-of-charge high-quality LEI data.

2 Introduction

Using the Global LEI Repository and its local LEI Issuing organizations' expertise, the LEI reference data offers free-of-charge reference data about entities around the world. By enforcing strict guidelines in terms of data validation and data consistency, GLEIF ensures that LEI data can be used to answer the questions of "who is who" and "who owns whom". The first is referred to as "level 1 data" and the latter as "level 2 data". In this context, GLEIF's data quality management program and the guidelines that LEI issuing organizations allow the best possible data quality (https://www.gleif.org/en/lei-data/gleif-data-quality-management).

One building block for creating a holistic entity profile are the local identifiers issued by the business register or other authoritative sources (e.g., financial supervisor, chamber of commerce). These data elements are already part of the primary data record and are corroborated against the authoritative sources by the LEI issuer at the time of issuance.

GLEIF engages and collaborates with its partner and together we provide native mappings between the LEI and other identifier and data standards. For instance, SWIFT provides the mapping between the LEI and the two ISO standards identifiers they are issuing: the Bank Identifier Code (BIC) & the Market Identifier Code (MIC). The Association of National Numbering Agencies (ANNA) has developed a program to which individual NNAs could opt-in and contribute to the LEI-to-ISIN mapping. Within this initiative, the International Securities Identification Number (ISIN) is directly linked to the LEI of the ISIN issuer at the point of issuance. In addition, S&P Global provides a mapping between the LEI and their internal S&P Global identifier. Following a similar strategy, Open Corporates and GLEIF collaborate to provide the LEI-to-Open Corporates identifier mapping.

The described mapping initiatives are not just providing mappings between the LEI and additional identifiers, they are creating direct connections among the identifiers themselves.

3 Standards

We introduce three Jupyter notebooks that use Python 3.8 and leverage the pandas, requests zipfile, io and IPython libraries. Jurisdictions, legal forms and entities have been chosen so that key features of the LEI data can be easily understood. All concepts shown in this analysis can be carried out with an arbitrary set of LEI data. Entities that appear in this analysis have been chosen at random. All expressions written in camel case in the following chapters refer to data elements in the Global LEI System. For this analysis, the Golden Copy files of 28 August 2023 are used (the date can be easily updated, as required by the user)

3.1 Accessing LEI reference data

All three Jupyter notebooks utilize GLEIF's Golden Copy files (https://www.gleif.org/en/lei-data/gleif-golden-copy/download-the-golden-copy#/) in CSV format. These files are also available as JSON and XML. In addition, LEI data can be accessed via an API (https://www.gleif.org/en/lei-data/gleif-api) and as so-called Concatenated Files (https://www.gleif.org/en/lei-data/gleif-concatenated-file/download-the-concatenated-file). For non-technical users GLEIF maintains a GUI, called LEI Search (https://search.gleif.org/#/search/).

Please note that the Golden Copy does not carry the aforementioned mappings. Instead, these are easily accessible on https://www.gleif.org/en/lei-data/lei-mapping via the respective mapping pages. In addition, all mappings are also available via the GLEIF API.

3.2 LEI-Common Data File format

LEI data follows a standardized structure and format, namely the Common Data File (CDF) format (https://www.gleif.org/en/about-lei/common-data-file-format). This format ensures that all reference data is gathered and respectively provided in a consistent manner and therefore allows for convenient data analysis of the underlying information. A detailed description of all elements of the LEI data is provided in GLEIF's State Transition and Validation Rules (https://www.gleif.org/en/lei-data/gleif-data-quality-management/downloads).

3.3 Relationship information

In the Global LEI System relationships between entities are captured in so-called relationship records. The format describes a convenient graph structure, that could be visualized as a network of nodes and edges between the nodes. The main features of these relationships are the StartNodeID, RelationshipType and EndNodeID. The StartNodeID shows the child entity's LEI code, the EndNodeID shows the parent entity's LEI code and the RelationshipType portrays the nature of the relationship. Currently, the LEI data contains relationship information for accounting consolidation parents, international branches and fund entities.

Accounting consolidation parent relationships are based on existing accounting definitions implemented in the International Financial Reporting Standard (IFRS), or the United States Generally

Accepted Accounting Principles (US GAAP). Other accounting standards are also accepted. There are two types of parent relationships in the LEI data which can be identified using the RelationshipType:

- "IS_DIRECTLY_CONSOLIDATED_BY": The direct parent describes the lowest level legal entity preparing consolidated financial statements for a given child entity.
- "IS_ULTIMATELY_CONSOLIDATED_BY": The ultimate parent is the highest level legal entity preparing consolidated financial statements for a given child entity. The ultimate parent does not have any accounting consolidation parent relationships of its own.

Reporting exceptions are used instead of relationship records in case no parent relationship is reported. Each reporting exception carries an ExceptionReason which provides further information:

- "NON_CONSOLIDATING", "NO_KNOWN_PERSON", "NATURAL_PERSONS": There is no parent entity present that fulfills the requirements for accounting consolidation.
- "NON_PUBLIC": The child entity is aware of a parent entity but withholds these details due to legal concerns.
- "NO_LEI": The parent entity does not have an LEI code.

Figure 1 shows a generalized view of a relationship network chain including non-consolidating ExceptionReason for the ultimate parent.

International branches represent different establishments or locations of the same corporate entity. They can be identified using RelationshipType "IS_INTERNATIONAL_BRANCH_OF". For entities with EntityCategory "BRANCH" it is noteworthy that these can only be declared in a country different from the headquarters' country. Furthermore, there is only one LEI assigned to the external branch location as each branch is representative of the entire country.

Fund relationships are described by the following RelationshipType values:

- "IS_FUND-MANAGED_BY": The EndNodeID of this relationship indicates the main management entity of a fund that is legally responsible for the constitution and operations of the fund.
- "IS_SUBFUND_OF": These relationships are used to represent umbrella funds, whereas the umbrella of the structure is shown as the EndNodeID.
- "IS_FEEDER_TO": This RelationshipType is used when a fund is exclusively, or almost exclusively invested in a single other fund (e.g., U.S., EU UCITS). This other fund is also referred to as "master fund".

For all aforementioned fund relationships, the child entity given as StartNodeID has EntityCategory "FUND".

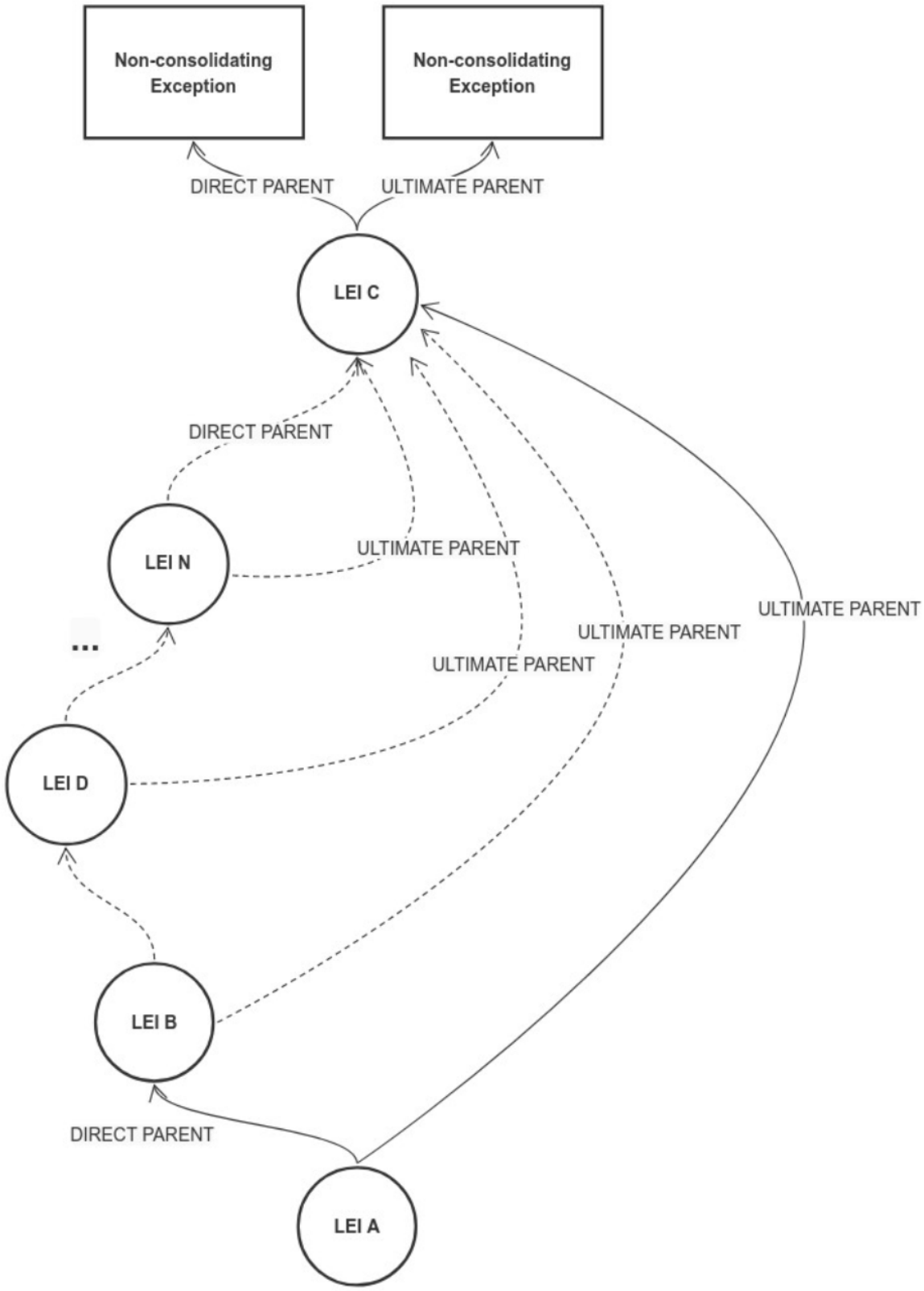

Figure 1: Generalized view of one complete network chain for an entity "LEI A"

4 Identify entity relationships and mappings

Chapter 4.1, focuses on a descriptive example of linking child and parent entities in this case across Great Britain and Japan. Chapter 4.2, covers Dutch feeder funds and Chapter 4.3 showcases the principles behind fund management and umbrella structures. All three notebooks are standalone Jupyter notebooks and can be used independently.

By default, each notebook will download the Golden Copy Files of 28 August 2023. This data is stored locally on the user's current working directory. Instead of downloading the Golden Copy Files for every execution, the user can set the "url" variables to the respective local file paths.

4.1 Investigate cross-border ultimate parents

For this section, please see 2023-06-05_Ultimate_Parents_PLC_GB.v1.0.ipynb.

In this use case, we identify the ultimate parents of British Public Limited Companies with LegalJurisdiction Japan. We start off by importing the level 1 data ("who is who"), relationship records and reporting exceptions ("who owns whom") as CSV files. Due to the large amount of data available as part of the level 1 records, it is advisable to remove all columns that are not of interest for the task at hand. As we are only interested in entities with LegalJurisdiction Great Britain, we apply a filter on the LegalJurisdiction element.

To identify Public Limited Companies, we leverage the ELF code list (https://www.gleif.org/en/about-lei/code-lists/iso-20275-entity-legal-forms-code-list). The openly accessible ELF code, established by the International Organization for Standardization (ISO), is a unique 4-digit alpha-numeric code and serves as a comprehensive solution for standardized legal form representation. As of July 2023, there are 3,250 legal forms in 175 jurisdictions worldwide available in version 1.4.1 of the code list. GLEIF has been acting as the maintenance agency secretariat of the ELF code list since 2017, regularly introducing new legal forms and jurisdictions. To identify Public Limited Companies in Great Britain, one must use ELF code "B6ES".

For the unique identification of the authoritative sources used for registration and validation of the entities and their reference data, the Global LEI System uses the Registration Authority (RA) code list. To identify the RA codes of the entities in scope, one should download the Registration Authorities code list (https://www.gleif.org/en/about-lei/code-lists/gleif-registration-authorities-list) and link it to the level 1 data. For our example we use RA000585, representing the Companies House in England and Wales.

Using the RegisitrationStatus element we focus only on ISSUED LEIs. Using this filter ensures that all reference data in the scope of our analysis has been re-validated at least once within the last 12 months.

When looking for the ultimate parents of our records in scope, we use the relationship records. As outlined in Chapter 3.3, we link our level 1 data with the StartNodeID of relationships with RelationshipType "IS_ULTIMATELY_CONSOLIDATED_BY". The EndNodeIDs of these relationships then

indicate the ultimate parents' LEI codes. In case no relationship is reported, the Global LEI System will instead offer so-called Reporting Exceptions. To retrieve the Reporting Exceptions for ultimate parents, one needs to use ExceptionCategory "ULTIMATE_ACCOUNTING_CONSOLIDATION_PARENT". The element ExceptionReason of the Reporting Exceptions gives an explanation for not providing any relationship information.

We focus on the identified ultimate parent relationships. To retrieve the level 1 information of these ultimate parents, we use the EndNodeID of the relationship records to merge with the level 1 information. Notably, EndNodeIDs are not necessarily unique as an ultimate parent may serve as the ultimate parent of multiple child entities and therefore it may act as EndNodeID in more than one relationship record. For filtering the most up-to-date data, we only consider relationships with RegistrationStatus "PUBLISHED" (the counterpart of the "ISSUED" RegistrationStatus from the level 1 information). For the sake of reducing the amount of data in scope, we only consider ultimate parent entities with LegalJurisdiction Japan.

Reference data in the Global LEI System is by design always captured in the local language. Hence, the legal names of the ultimate parents are given in Japanese. To allow non-Japanese speakers to more easily comprehend the legal name information, the LEI record includes an optionally populated element "OtherEntityName" which can be used for translated or transliterated entity names.

The last exercise in this notebook is linking the LEI data with 3rd party data providers. As of August 2023, GLEIF offers mapping files for BIC, ISIN, MIC, S&P Global Company ID and OpenCorporates ID (https://www.gleif.org/en/lei-data/lei-mapping). In the Jupyter notebook, we showcase the mapping of the Japanese ultimate parents with the OpenCorporates ID which enables data users to gain additional insights about the entities in scope.

Lastly, the notebook highlights the different sources and stakeholders involved in making the data available to the public.

4.2 Analyze international Dutch feeder funds

For this section, please see 2023-08-25_Feeder_Funds.v1.0.ipynb.

In this notebook, we identify feeder funds in the Netherlands and analyze their master funds. As in the previous chapter, we first import the LEI data by downloading the Golden Copy Files. After selecting all records with LegalJurisdiction "NL" to retrieve all Dutch entities, we apply a filter on the EntityCategory as we are only interested in fund entities. Furthermore, we again add the registration authority information to our dataset. Please note that this notebook only considers fund relationships. In addition to fund relationships, each LEI record with EntityCategory FUND holds information regarding direct and ultimate parents as showcased in Chapter 4.1. These non-fund relationships are not in scope for the analysis in Chapter 4.2.

The majority of fund entities in the Netherlands are either registered with RA000464 (Autoriteit Financiële Markten) or with RA000463 (Kamer van Koophandel). Others use RA999999, indicating that no registration authority is available for the given entities. Next, we select all LEI records with RegistrationStatus "ISSUED" to get the most up-to-date reference data.

When the relationship information comes into play, we apply the same methodology as in the previous chapter. Only the RelationshipType of interest has changed to “IS_FEEDER_TO”. The EndNodeID will show the LEI code of the master fund. Thus, by joining the previously filtered level 1 information with these relationships we create a link between feeder funds and their master funds. By picking the EndNodeIDs of these relationships we can delve into the level 1 reference data of the master funds. In our analysis we are interested in master funds outside of the Netherlands. Thus, we select those master funds with LegalJurisdiction “IE” for Ireland.

Using this approach we have identified Irish master funds that are managed by the Irish Stock Exchange and registered at RA000404 – The register for financial service providers and collective investment schemes. The corresponding feeder funds with LegalJurisdiction “NL”, are also managed by the Irish Stock Exchange, and use RA999999 as RegistrationAuthorityID. RA999999 indicates that no registration authority is available for these entities.

4.3 Understanding Mexican fund entities

For this section, please see 2023-08-25_Umbrella_Funds.v1.0.ipynb.

Regarding Mexican funds, two tasks are carried out:

- Identify the fund managers of all LEI records with EntityCategory “FUND”.
- Make transparent the umbrella structure of umbrella funds.

As in the previous two chapters, the Golden Copy CSV data is used and the registration authority code list is added. We only consider LEI records with LegalJurisdiction Mexico and with RegistrationStatus “ISSUED”. This notebook only considers fund relationships. In addition to fund relationships, each LEI record with EntityCategory FUND holds information regarding direct and ultimate parents as showcased in chapter 4.1. These non-fund relationships are not in scope for the analysis in Chapter 4.3.

4.3.1 Fund managers

Once the level 1 information has been pre-filtered, the fund managers can be identified using the relationship records with RelationshipType “IS_FUND-MANAGED_BY”. The EndNodeIDs of these relationships designate the LEI code of the fund manager. In the Golden Copy data of 28 August all but one fund manager have LegalJurisdiction Mexico. This non-Mexican fund manager has LegalJurisdiction US-Panama. By utilizing the StartNodeIDs found in the fund manager relationship, where the EndNodeID corresponds to the fund manager based in US-Panama, we can systematically investigate the fund entities overseen by this particular fund manager.

4.3.2 Umbrella structures

Using the same level 1 information as in chapter 4.3.1, we focus on umbrella structures. First, we select umbrella funds by selecting the EndNodeIDs of relationships with RelationshipType “IS_SUBFUND_OF”. Consequently, the StartNodeIDs of these relationships show the sub-funds for a given set of umbrella entities. Please note that one single umbrella fund (one EndNodeID) usually has

multiple sub-funds (StartNodeIDs), resulting in multiple sub-fund relationships pointing to the same EndNodeID.

In the Legaljurisdiction “MX” we detect three umbrella funds. All of them are registered in RA000449 (Registro Federal de Contribuyentes). The responsible LEI issuer for all LEI records is GS1 Mexico. The corresponding sub-funds can be obtained using the StartNodeID values. For these three umbrella funds, there are 32 sub-funds. By grouping these sub-funds based on the EndNodeIDs, one can assign each group of sub-funds to their respective umbrella funds. Like the umbrella funds, all sub-funds are also registered in RA000449 and their LEI issuer is GS1 Mexico.

5 Conclusion

In this session, we demonstrated how entity relationships across multiple jurisdictions can be identified using LEI data. In particular, we want to emphasize that the reference data is taken from local authoritative sources and is managed by local experts, the so-called LEI issuing organizations. In addition, the LEI data is natively mapped to other standards and identifiers, enabling a holistic and comprehensive entity network. This setup allows for unbiased and accurate reference data even for large-scale entities with globally distributed child and parent entities. GLEIF ensures that all reference data is provided in a consistent, standardized and reliable format so that data consumers can utilize the information with ease.