opencorporates

GLEIF

OpenCorporates ID to LEI mapping table

Factsheet

GLEIF and OpenCorporates

22 August 2024

1. Introduction

OpenCorporates is the world's foundational source for legal entity data and is on a mission to bring about genuine corporate transparency, and has been instrumental in opening access to corporate registries across Europe and North America, as well as being the largest open database of companies in the world, containing over 220 million companies and 300 million officers from 145 jurisdictions (as at early 2023). Its data is fully provenanced to official sources, unified according to a global schema, up-to-date and available at scale.

The Global Legal Entity Identifier Foundation (GLEIF) is a not-for-profit organization created to support the implementation and use of the Legal Entity Identifier (LEI). GLEIF manages a network of partners, the LEI Issuing Organizations, to provide trusted services and open, reliable data for unique legal entity identification worldwide.

The Legal Entity Identifier (LEI) is a 20-digit, alpha-numeric code based on the ISO 17442 standard developed by ISO. It connects to key reference information that enables clear and unique identification of legal entities participating in financial transactions.

Both datasets enable smarter, less costly and more reliable decisions about who to do business with. By providing a link between these two datasets, users gain access to a richer pool of information with which to make decisions and, in turn, cut costs, simplify and accelerate operations. In taking this step forward by publishing a publicly available OpenCorporates ID-to-LEI mapping reference file - individuals and organizations will be able to map LEIs to information compiled by OpenCorporates from local corporate registrations globally.

The mapping file is built upon a mapping process established in cooperation between GLEIF and OpenCorporates leveraging on OpenCorporates API. This file will be updated initially on a bi-weekly basis. The purpose of this factsheet is to describe the reference data involved, the analysis applied to map LEIs to OpenCorporates IDs, and the ongoing update process.

2. Reference Data

The OpenCorporates dataset is the source for the OpenCorporates ID and associated company data sourced from official company registers. Additional information on this data source including field definitions can be found here.

The GLEIF Concatenated File is the source for the LEIs and the legal entity reference data. Additional information on the GLEIF Concatenated Files and the Common Data File formats can be found here.

The OpenCorporates ID-to-LEI mapping table does not include historical OpenCorporates ID/LEI pairs, stand-alone unmapped OpenCorporates IDs and LEIs, nor the listing of daily or monthly changes.

The following table describes the critical fields considered for mapping. Some additional pre-processing steps are in place considering specific local differences. Should you require further information please reach out to info@gleif.org.

Table 1 – Critical Fields Considered for Mapping

OpenCorporates	Description & Comments	LEI	Description & Comments
Fields considered for pre-conditions			
		Registration Status	LEIs with Registration Status of ANNULLED or DUPLICATE are excluded
		EntityCategory	LEIs with EntityCategory FUND are excluded
		RegistrationAuthorityId ValidationAuthorityId OtherValidationAuthorityId	LEIs which use reserved Registration Authority codes RA888888 and RA999999 (GLEIF Registration Authority code list) are excluded. Only those LEIs are in scope that have a Registration Authority that is used by GLEIF and Open Corporates
		RegistrationAuthorityEntityID ValidationAuthorityEntityID	LEIs without a local identifier are excluded
Fields considered as equivalents for mapping			
Company Number	Identifier given by the Company register to the entity. Together with the Jurisdiction Code this forms a composite primary key for a legal entity	RegistrationAuthorityEntityID ValidationAuthorityEntityID OtherValidationAuthorityEntityID	All available entity IDs of an LEI Record are used to establish a mapping pair. All white spaces are removed and jurisdiction-specific pre-processing of entity IDs is applied where necessary
Jurisdiction Code	Code for the jurisdiction (lowercase, underscore version of ISO 3166-2 code). List of jurisdiction codes are available at https://api.opencorporates.com/jurisdictions	LegalJurisidiction	The LegalJurisdiction element is transformed to match the lowercase, underscore version of the ISO 3166-2 code

3. Analysis

In April 2023, there were around 220 million companies in the OpenCorporates dataset with over 2.3 million LEIs having been issued.

Of these 2.3 million, about 1.2 million have been automatically reconciled with a strict criterion applied with respect to registered jurisdiction, local company number.

There are two main reasons why an LEI will not have an applicable OpenCorporates ID:

1. The LEI is associated with legal form other than a company legal entity (e.g. mutual fund or government bond)
2. OpenCorporates does not yet include the register in the jurisdiction where the entity is registered.

The technical approach to the mapping process will be subject to continuous review to improve the quality of the matching process.

4. On-going update process

This section describes what changes (events) to the underlying datasets trigger update to the mapping discussed in the section above.

Generally, an existing mapping pair is only updated if one of the following LEI data elements is changed:

- RegistrationAuthorityEntityId (RAEID)
- ValidationAuthorityEntityId (VAEID),
- OtherValidationAuthorityEntityId (OVAIED)
- RegistrationAuthorityId (RAID)
- ValidationAuthorityId (VAID),
- OtherValidationAuthorityId (OVAID)

Unmapped records will also be revisited on an adhoc basis based on updates to OpenCorporates coverage or on renewal if an OpenCorporates ID has not previously been mapped.

opencorporates

Table 2: The change to an LEI (event) and the consequent update to the OpenCorporates/LEI pair on the mapping file.

Event	Update
A new LEI is issued.	The LEI is considered for mapping.
In an OpenCorporates ID/LEI pair the LEI ANNULLED or DUPLICATE, the RAID or VAID changes to RA999999/RA888888, or EntityCategory changes to FUND.	The mapping is removed from the mapping file.
In an OpenCorporates ID/LEI pair the LEI became RETIRED or MERGED.	The mapping is not removed from the mapping file.
In an OpenCorporates ID/LEI pair any of the OpenCorporates ID or LEI data fields considered as equivalent mapping (see Table 1) change.	If RAEID, VAEID or OVAEID of an LEI Record changes, the LEI is sent to through the mapping process irrespective whether previously a mapping existed. If the updated set of entity IDs for an existing mapping cannot be matched with OpernCorporates anymore, the mapping pair is removed. If any of the OpenCorporates data fields is updated, the existing mappings are not updated. The only trigger to update existing mappings is currently by updating one of the entity ID elements in the LEI data
If an LEI that was not mapped in the past is renewed.	The LEI is sent through the mapping process again.

Users of the OpenCorporates ID-to-LEI mapping reference file may ask general questions or challenge individual OpenCorporates ID/LEI pairs by emailing info@gleif.org.

5. File Specifications

The file is published in comma-separated format (csv) contained in a zip file. It contains all identified OpenCorporates ID-LEI pairs, i.e. is a "full" file. No "delta" file is made available. The file does not contain OpenCorporates IDs that do not have corresponding LEIs nor LEIs without corresponding OpenCorporates IDs.

The file contains the following data:

FIELD NAME	FORMAT	PRESENCE	DESCRIPTION
LEI	20 alphanumeric	Always present	The LEI of the entity.
OpenCorporates ID	136 alphanumberic	Always present	The OpenCorporates ID of the entity. jurisdiction_code "/" company number

6. Resources

The OpenCorporates ID-to-LEI mapping full file can be downloaded via https://www.gleif.org/en/lei-data/lei-mapping/download-oc-to-lei-relationship-files

The GLEIF API can be accessed via https://api.gleif.org/

For every OpenCorporates ID-LEI pair, on the website, the LEI can be seen under the section titled "Additional Identifiers" using the following URL where YY is the jurisdiction code (YY_YY in case of subdivisions e.g. US_DE for Delaware), and XXXXXXX represents an arbitrary company number: www.opencorporates.com/companies/YY/XXXXXXX

For users of the OpenCorporates API, the relevant call structure (before appending your API token) is: http://api.opencorporates.com/v0.4/companies/YY/XXXXXXX

The API call will return the LEI within the "identifiers" section.

If you receive OpenCorporates bulk data, please refer to your technical documents with respect to "Additional Identifiers".

Using GLEIF as an example the relevant URLs would be:

- https://opencorporates.com/companies/ch/1190521
- https://api.opencorporates.com/v0.4/companies/ch/1190521

For details on how to access the OpenCorporates data at-scale through API or Bulk data, please refer to: https://opencorporates.com/info/our-data/.