org-id.guide handbook

Release 0.1

org-id.guide

Jan 10, 2019

Contents

org-id.guide

Tip: Looking to find an organisation identifier? Go to http://org-id.guide for our list locator.

Contents:

About

Information about organizations is more valuable when it includes:

- An unambiguous **organization identifier**
- A **clear name** for the organization
- Other **identifying information**

org-id.guide provides a register organization identifier lists that delivers the easiest, quickest and most accurate way for open data publishers to locate organisation identifiers.

Why use org-id.guide

Quality

org-id.guide assigns a code that can be used to refer to each list. It ranks the quality of the lists based on the **stablility and accuracy** and the **availability of open data**.

The higher the quality, the more likely it is that the identifiers can be used to link up information across datasets.

Relevance

The org-id.guide web application provides a search tool to locate relevant identifier lists. Search results are ranked by **relevance**.

The more **relevant** a list, the more likely that it will contain the organization that a user is looking for.

Connected

Each entry in org-id.guide has detailed meta-data and information that can be used to locate the identifiers that the list contains, whether through online search or open datasets.

We use this meta-data to recommend use of lists that have the best chance of making org-id connections between different datasets.

Collaborative

org-id.guide is a living list - improved by a coalition of open data standards projects. You can help improve the coverage and depth of the org-id.guide register. Changes go through a simple governance process.

What org-id.guide is not

org-id.guide **does not** directly maintain organisation identifiers itself. Instead, it signposts to existing sources of identification: helping data publishers and users choose the best identifier schemes to use for their purpose.

This handbook

You are viewing the org-id.guide handbook. It contains information on how the org-id.guide register is maintained.

Contributing

The org-id.guide list is maintained by a coalition of open data standards: and is open to contributions and requests from users of any of those standards.

Note: We use GitHub as our issue tracker and to manage the register. If you don't get have a GitHub account, you can sign-up for free at https://github.com.

Suggesting additions and changes

You can suggest updates to org-id.guide in a number of ways.

1. Requesting a new entry

When you have an organization or organizations you want to provide an identifier for - but you can't find any official identifier source, you can **post a 'REQUESTED' issue**.

This should contain details about the organization(s) you want to identify including:

- The legal form (charity, company, government agency); and
- The country they are registered in;

Any additional information you can provide about this kind of organization in this country (e.g. if you know anything more detailed about the legal regime for such organizations), or about your need for identifiers (e.g. do you need it to be linked to particular kinds of open data) will help the research process.

Requests from org-id.guide members will be prioritised for a response by the research team.

We encourage community members to also get involved in providing suggestions for appropriate lists in response to a request.

2. Proposing a new entry

When you have identified a list that should be included in org-id.guide, you can post a 'PROPOSAL' issue, or submit a pull-request with a fully-researched list.

When submitting a PROPOSAL, please include:

- The name of the list (or the organization that manages the list)

- A short description of the list

- The country or countries that the list covers

- The legal form or organizations that the list covers

- Any specific sectors that the list covers

- A URL for information on the list, and a URL for looking up identifiers (if available)

- Any information on available open data copies of this list

Most importantly, please provide at least one example of an organization identified by the list, along with the example identifier, and information on how this was obtained.

For example, in some cases, where there is no public database of identifiers, identifiers are found by asking an organization for a copy of their certificate of registration.

3. Proposing a correction or update

If you can provide additional information, or want to suggest an update to an entry in org-id.guide, open an issue or pull-request with your suggested changes.

Include details of the entry you are proposing an update for, and links to any sources from your research.

What happens with suggestions

Requests are prioritised, and passed to our research team to identify potential lists. They will keep the GitHub issue updated with their research. If/when they identify an appropriate list, they populate the list meta-data needed to create a **proposal**.

Proposals are reviewed by the research team to check that all the key stages from the research guide have been covered and that all the required meta-data is provided. A new list entry is created using the list schema and a pull-request created. Once the pull-request is open:

- The list can be previewed and is staged for inclusion in org-id.guide

- There is a 7-day period for comments and feedback to either improve the meta-data, or raise any questions about inclusion of this list in the register;

- If no blocking concerns have been raised by org-id.guide members, or wider community contributors, during this period, the list is added to the org-id.guide register;

Updates are reviewed by the research team. Minor updates are merged without going through the full 7-day review period. Major updates will be treated as proposals.

Deprecations will be handled based on **active consensus** rather than the lazy consensus process used for new list proposals.

Process transparency

Transparency is an essential principle for governance of org-id.guide. All decisions are made in the open, and are managed through the git version control system and open issues.

List research

Populating new entries

For list requests, the research team should seek to identify a potential source of organization identifiers.

The research process may involve:

- Online desk research;
- Consultation with the wider standards community;
- Consultation with local experts;

Once you have identified a good candidate list to meet a particular need, assign it a prefix as detailed below, and then fill in the detailed metadata.

Validating an existing entry

For stub entries or proposals with incomplete information, researcher sshould:

1. **Check the list title** - and make sure it follows the rules for multilingual titles

2. **Write a clear description of the identifier list** describing the way in which organisations end up on the list. This should be 1 - 2 paragraphs maximum. You may use content quoted from the registers own website, or wikipedia pages.

3. **Fill in the list metadata**

Research sources

When carrying out research the following resources may be useful. All researchers are encouraged to familiarise themselves with these resources.

Useful websites

- **Investigative dashboard** - https://investigativedashboard.org/databases/topics/business list of company registers

- **Wikipedia.** Primary and secondary identifier lists are usually notable enough to have a Wikipedia page, at least in the language of the country concerned. For official registers - look for information about the legislation that created the register in order to understand the organisation types it covers. See in particular List_of_company_registers and Types_of_business_entity

- The World Bank Doing Business country profile give detailed information on company registration processes - and can provide useful hints as to the official company register and identifiers in a country.

- OpenCorporates Open Company Index

- European Commission list of Business registers in Member States

- Wiki Procedure

- Open Corporates policy paper on handling company number problems

Existing use

For codes taken from the IATI Codelist, you can search for records that currently use this code via OIPA.

For example, the following query returns a list of activities (in JSON format) with reporting or participating organisation identifiers that start with 'ET-MFA'

https://oipa.nl/api/activities/?format=json&q_fields=reporting_org,participating_org&q_lookup=startswith&q=ET-MFA

Checking each of the linked activities can give an indication of the kinds of identifier in use.

Current usage in IATI is no guarantee of a correct identifier, but, it can give clues as to the kinds of identifiers you are looking for, and can help validate organisation list information.

Meta-data guide

For each entry in org-id.guide, a full meta-data entry is maintained, following the schema. This meta-data:

- Helps data publishers to understand the nature of an organization identifier list, and how to find the identifiers it contains;

- Helps users to interpret identifiers, and locate additional sources of information linked to each id;

- Drives the quality and relevance rankings used in the org-id.guide frontend;

Assigning a Code

Each **organisation list** code is made up of two parts: a jurisdiction code, and a list name code.

1) Jurisdiction code

For any list which contains entries only from a given country, the ISO 2-digit country code should be used.

For lists that contain entries from multiple countries, one of the following codes should be used (NOTE: We rely here on the fact that in ISO 3166-1 the following alpha-2 codes can be user-assigned: AA, QM to QZ, XA to XZ, and ZZ. We avoid any widely used X codes.).

Code	Usage
XM	Multilateral/international agencies. The list contains multilateral or international agencies. These organisations are generally not registered at the national level.
XI	International. The list contains organisations based anywhere in the world. The organisations may or may not be registered at the national level.
XR	Regional. The list contains organisations based within a particular region. For example, organisations within the European Union, or within Africa only.
ZZ	Publisher created. This list was created by a publisher, and is maintained by that publisher alone.

2) List name code

The list name code is manually assigned. In general, list name codes have been designed so that:

1. They are between 2 and 7 characters long;

2. They use a recognisable acronym or contraction of the name of the organisation list;

3. Acronyms should be based on the local language version of the name;

4. Codes should be memorable, allowing users to become familiar with codes.

However, there is no intention that the list name code portion of the prefix will carry any semantics (e.g. type of organisations listed etc.).

Points 1 - 4 are guidance only. Breaking any of these principles shall not be grounds for revision of an existing code.

When a new organisation list is identified, the creator/proposer may suggest a list code.

To confirm a code, the researcher should check:

- The proposed list code is based on a clear understanding of the name of the underlying **organisation list** or list provider.

- Any requirements for multi-lingual codes. For example, in Canada, acronyms should always be given in both their English and French forms. This is achieved using an underscore separator in the list code. (For example CA-CRA_ACR for the Canadian Revenue Agency / Agence du revenu du Canada)

- The code does not duplicate an existing code, or AKA value.

If a list code is later replaced, the full deprecated prefix should be recorded in the '''formerCodes'' field. This will allow systems to warn about deprecated prefixes and to notify users of their replacements.

Entering meta-data

Name / Title

Provide the name of the **identifier list**, or, where the list is the primary list maintained by an organisation, the name of the organisation that maintains the list.

> For example, we use 'Companies House' to identify the owner of the UK Company Register.

You can enter the name in English, and separately in a local language.

URL

Provide the root URL where information on the list can be found. This does not need to link directly to an interface to search the list.

Description

The description should focus on explaining the way in which organisations end up on the list. If a list covers multiple kinds of organisation, the description may highlight this, using text from the lists own website, or summarised from other relevant documentation.

Descriptions may include Markdown. When including citations, use markdown footnote notation:

```
This is some text which requires a citation at the end [1].

[1]: This is the footnote text.
```

Hint: Example Description - AU-ABN

"The Australian Business Number (ABN) enables businesses in Australia to deal with a range of government departments and agencies using a single identification number. The ABN is a public number which does not replace an organisations tax file number."

"ABN registration details become part of the Australian Business Register (ABR)"

Each ABN should equate to a single 'business structure', although that structure may be used to carry out a range of business activities. A range of kinds of entity are issued ABNs, including individuals, corporations, partnerships, unincorporated associations, trusts and superannuation funds. Entities must be carrying on a business in or connection to Australia to receive an ABN.

Geographic coverage

Enter each of the jurisdictions this identifier list covers.

If the list is global, use one of the XI (International), XM (Multilateral) or ZZ (Publisher created).

If the list is regional, enter all the countries that the region covers.

Sub-national coverage

If this list **only** covers one or more sub-national territories, select these.

(If the schema does not include the required ISO 3166-2 Subdivision Assigned Codes, open a GitHub issue to request these are added)

Legal structure

Select all the legal structures which this list covers.

Note that legal structures are organised hierarchically in the dataset. So, for example, 'Sole Trader' is a kind of company. This is shown in the lookup list under the 'Parent' field.

Please consult the research lead if you feel you need to add an extra category to legal structures.

If the list is not specific to a particular kind of legal structure, leave this field blank.

Hint: Example: GB-COH

UK Companies House registers a number of different kinds of company, including:

- Public limited company (PLC)
- Private company limited by shares (Ltd, Limited)
- Private company limited by guarantee, typically a non-commercial embership body such as a charity
- Private unlimited company (either with or without a share capital)
- Limited liability partnership (LLP)

- Limited partnership (LP)

- Societas Europaea (SE): European Union-wide company structure

- Companies incorporated by Royal Charter (RC)

- Community interest company

It is listed against the following specific company types: Partnership, Limited Company, Listed Company, Community Interest Company, and Charity.

However, wider research tells us that whilst all Limited Companies, Listed Companies and CIC's should have a registration in Companies House, not all charities will have a Companies House number.

Sector

If this list is specific to a particular sector, you can declare that here.

If the list is not specific to a particular sector, leave this field blank.

Hint: Example: GB-UKPRN

The UK Register of Learning Providers covers only education institutes, so has 'Education' set in the sector field.

List type

This is one of the most important fields in the dataset. You will need to determine if this list is a **primary identifier list** or whether it has secondary, third-party or local status.

Definitions of each category are provided above.

Drawing on your research into how identifiers are created, and looking at a range of example entries in the list, make your determination. You can use the comments feature in AirTable to provide supporting reasons if you require.

The following rule-of-thumb criteria may be useful.

Table 4.1: Rule of Thumb

Primary	Secondary	Third-party	Local
Provided by an official registrar Organisations were assigned the identifier at the time they were first created. Near 100% coverage of the legal type in a jurisdiction (e.g. list contains all companies)	Managed by an official source. Not all organisations of a given legal type will have these identifiers. Relies upon some other status of the organisation (e.g. VAT registration, being an employer etc.)	Maintained independently of the organisations listed. May be based on official records, but identifiers are assigned separately from official processes.	Maintained by a single organisation for their own business purposes.

Access information

Available online, and online availability details

Indicate whether this list is available online in **any form**, including only partial search.

Provide the URL that users should visit to access this list and a description of how to find identifiers.

How to locate identifiers

If users need to follow particular steps in order to carry out a identifier search, detail those here.

This might include:

- Guidance on how to find search features on a complex website;

- Information on charged access to identifiers if no freely available online access is provided;

- Information on how to spot the actual identifier, and how to copy it for re-use;

- Information on formatting the identifiers.

Hint: Example: AU-ABN

It is possible to search for identifiers at http://abr.business.gov.au/

The Australian Business Number (ABN) is a unique 11 digit identifier issued to all entities registered in the Australian Business Register (ABR). The 11 digit ABN is structured as a 9 digit identifier with two leading check digits.

The identifiers are displayed on the website with spaces in the number. All the spaces should be removed when making use of the number within an identifier.

Example identifiers

Provide 1 - 5 example identifiers, comma separated.

Hint: Example: GB-COH

09506232, 07444723

Access to data & Data access details

Check for bulk downloads, and API access to the data, and indicate if these are available.

For official registers, check on the national data portal as well as the list website itself. Take note of whether the available data appears to be regularly updated, or only a one-off data dump.

Write brief notes on how the data can be accessed.

Confirm the license information for the data.

Data features

Select all the features that are apply to **either** of information available through the list's website, or in APIs or bulk data products.

The goal here is to be aware of all the possible additional available information that could be explored to disambiguate organisations, whether that is available as structured data or not.

Openly licensed and license details

Look for a license for the contents of the list.

Indicate whether or not an open license can be found, and provide the name of the license (if common) or a short description of the license if it is not a common license.

Wikipedia page

If you have found a wikipedia page for the organisation, link to that here.

In OpenCorporates?

If OpenCorporates has data for this list, include a link to the open corporates page here.

Languages supported

Using ISO e-digit language codes, indicate which languages this list is available in.

Last updated

Make sure the last updated date reflects the current date.

Confirmed

The confirmed flag should be set once this list entry has been reviewed and accepted.

List Schema

Entries in the org-id.guide platform MUST be provided as JSON data following the latest version of the list schema, available here.

The schema structure is documented below. More details on how to populate schema elements are found in the research guidance

Additional schema information

The schema is compiled from a number of source files that contain codelists embedded within the schema. These codelist files also contain weighting used to calculate quality scores.

The schema includes some additional properties to aid interface creation:

- Property ordering
- enum_titles - a list matching the order of an enum list to provide titles for elements.

Schema overview

Title	Description	Type	Format	Required
name		object		
List name	Provide the name of this list in English, and then the local language, if applicable.			
name/en		string		
English				
name/local		string		
Local language	If this list has a local language name, please provide that here. If local law requires th			
url		string	uri	
List URL	Provide a URL where further information about this list can be found (e.g. the homep			
description		object		

Title	Description	Type	Format	Required
List description	The list description should be written in plain text or markdown, and should contain t			
description/en		string	textarea	
None				
coverage		array[[string]]		
Geographic coverage	Which country, countries or kind of region does this organisation list cover? For any l			
subnationalCoverage		array[[string, null]]		
Subnational jurisdiction	If this list does not cover a whole country, but **only** covers one or more sub-national te			
structure		array[[string]]	table	
Legal structure	Select all the legal structures which this list covers. Note that legal structures are orga			
sector		array[[string, null]]		
Sector	If this list is specific to a particular sector, you can declare that here. If the list is not s			
code		string		
List code	Each organisation list code is made up of two parts: (1) a jurisdiction code; and (2) a l			
confirmed		boolean		
List reviewed?	Once this list has been reviewed by an administrator, this value can be set as true. Onl			
deprecated		boolean		
Deprecated?	If this list has been removed from active use, deprecated should be set to true. The des			
listType		string		
List type	You will need to determine if this list is a primary identifier list or whether it has seco			
access		object		
Access information	Information about access to this list, the identifiers and additional information it conta			
access/availableOnline		boolean		
Available online?	Is is possible to access this list online in any form (search / bulk download etc.)			
access/onlineAccessDetails		[string, null]	textarea	
Online access details	An optional description of how to access this list online, including details of any regis			
access/publicDatabase		string	url	
List URL	Where should users go to search or access this list?			
access/guidanceOnLocatingIds		string	textarea	
How to locate identifiers	Provide brief guidance on how someone visiting the URL for this list should locate an			
access/exampleIdentifiers		string		
Example identifiers	Please give a short list (3 - 5 entries) with example identifiers from this register. These			
access/languages		array[string]		
Languages	Which languages is this list, and/or the interfaces to access it, available in?			
data		object		
Dataset access	Details of whether or not an open dataset is available for this list.			
data/availability		array[string]		
Data availability (was DataAccessProperties)	Is structured data available from this list in any of these formats or mechanisms?			
data/dataAccessDetails		string	textarea	
Data access details (was dataAccessDetails)	Please describe how to access data in the formats listed above. Include additional URL			
data/features		array[string]		
Dataset features (was datasetFeatures)	Does the data available from this list (in any format) contain any of these features?			
data/licenseStatus		string		
License status	Is the data described above available under an open license?			
data/licenseDetails		string		
License details	Details of the license, including a link to the license where applicable.			
meta		object		
Metadata	Every list should be supported by a meta-data record			
meta/source		string		
Source	Please briefly describe the source of the information (e.g. original research; official re			

Title	Description	Type	Format	Required
meta/lastUpdated	string	datetime		
Last updated	A timestamp to the last update to this record			
links	object			
Links	For links to supporting information, or other database records, about this list			
links/opencorporates	string	url		
Open Corporates	A link to the Open Corporates page for this identifier should be provide **only when O**			
links/wikipedia	string	url		
Wikipedia page	If there is a wikipedia page that provides context to this this list (e.g. page describing			
formerPrefixes	array[string]			
Former list codes	In some cases a list code may be changed. This field will list deprecated codes for this			

CHAPTER 6

Using the register

The canonical source for the org-id.guide register exists on the master branch of https://github.com/org-id/register

All codes with `confirmed:   true` are part of the register.

The application at http://org-id.guide provides a front-end for exploring the register, as well as XML, JSON and CSV downloads of the full register of organisation identifier lists.

CHAPTER 7

Members & governance

org-id.guide is a member-led project, building a well-governed open register of organization identifier lists.

Anyone can use org-id.guide.

Members contribute to maintain the quality of register research and meta-data, and as part of a commitment to integrate the org-id.guide register into their standards and applications.

Member benefits

To find you about becoming an org-id.guide member, e-mail contact@org-id.guide

Members benefit from:

- **Priority support from the reseach team.** Requests, proposals and change suggestions from members will be given highest priority.

- **Direct engagement with member communities.** The research team will accept messages directly from member's own helpdesk or support systems, and will feedback directly.

- **Governance role.** Members are consulted on all new register entries and deprecation, and work together to set future directions for the project.

Members list

360Giving

360Giving supports organisations to publish their grants data in an open, standardised way and helps people to understand and use the data in order to support decision-making and learning across the charitable giving sector.

Organisation identifier: `GB-COH-09668396`

Development Initiatives (Joined Up Data Standards project)

Joining up data standards is a vital part of turning more data into better information to drive sustainable development.

Organisation identifier: `GB-COH-06368740`

International Aid Transparency Initiative (IATI) Secretariat

IATI makes information about aid spending easier to access, use and understand.

Organisation identifier: n/a

Natural Resource Governance Institute

We help people to realize the benefits of their countries' endowments of oil, gas and minerals.

Organisation identifier: `US-EIN-20-4451390`

Initative for Open Ag Funding

We are a group of like-minded organizations who believe in the power of data to improve sustainable development. Our works supports efforts to end hunger and food insecurity by ensuring organizations have the information they need to make smarter investments in the agriculture and food security sector.

Organisation identifier: `US-EIN-13-3287064` [InterAction's identifier]

Open Data Services Co-operative

We help people publish and use open data.

Organisation identifier: `GB-COH-09506232`

Supporters

The following organizations have indicated their support for the project by becoming active contributors, or committing to make use of the org-id.guide register.

Open Contracting Partnership (OCP)

We connect governments, civil society and business to open up and monitor public contracting.

Organisation identifier: n/a

Open Corporates

OpenCorporates aims to do a straightforward (though big) thing: have a URL for every company in the world.

Organisation identifier: `GB-COH-07444723`

ePaństwo Foundation

ePaństwo Foundation's aim is to develop democracy, open and transparent authorities and civic engagement using new technologies.

Organisation identifiers: `PL-NIP-1231216692`, `PL-REGON-142445947`, `PL-KRS-0000359730`.

DatosElSalvador.org

DatosElSalvador.org is the first open data portal in El Salvador, built by citizen, that provides dataset of economics, social, enviromental, transportation, education, health and government operations open data, according to the best international practices and the open data principles.

Organisation identifier: tbc

 Linked Economy

Linked Economy aims at providing a universal access to Greek and International economy data.

CHAPTER 8

Applications

The source code for the org-id.guide application can be at https://github.com/OpenDataServices/org-ids/

The deployed version is a view-only copy.

The 'editing' branch, which can be run locally, contains a copy of the application with support to propose edits to codes.

CHAPTER 9

License

The org-id.guide register is available under the ODBL

Contact and community

For questions related to the project, contact Open Data Services Co-op on contact@org-id.guide

Engage with the public issue trackers:

- For the register of organization lists
- For the org-id.guide application

A public e-mail discussion list is available at org-id-guide@googlegroups.com