

Opening Extractives

Unlocking the benefits
of ownership data

Technical Guidance

Relational database design considerations for beneficial ownership information

Contents

Summary	4
1. Introduction	5
2. Key concepts	6
a) What is beneficial ownership?	6
b) The Beneficial Ownership Data Standard (BODS)	7
3. Information flows	9
4. Data capture	10
a) Initial and update declarations	10
b) Clarity about what is being declared	10
c) Avoiding duplication	10
d) Form design	11
e) Collecting structured data	11
f) Data lookups	11
g) Verification of data outside the jurisdiction of a register	12
h) Ongoing improvement	12
5. Database structure	13
a) Tables	14
b) Fields	15
6. Case studies	17
7. Further resources	18

Summary

In the **key concepts** section, this guidance note introduces the concept of beneficial ownership and outlines the characteristics of the **Beneficial Ownership Data Standard (BODS)**. It describes how the data model allows those characteristics to be expressed in data. These sections provide background information for users to consider when designing a beneficial ownership database.

The section on **information flows** describes the assumptions made about the high-level systems architecture within which a beneficial ownership database will exist, and how data flows from a declaring company through that system to be published.

The section on **data capture** discusses best practice for collecting beneficial ownership data. Finally, an outline is provided of a potential relational **database structure** that enables the collection and storage of beneficial ownership data in a way that supports the publication of high-quality BODS data.

Key resources are included at the end of this note to support further investigation.

1. Introduction

This guidance note provides technical guidance on considerations for relational database design when building a system that holds beneficial ownership information. In particular, it focuses on the requirements for publishing data to meet the [Beneficial Ownership Data Standard \(BODS\)](#). Most of these recommendations will also suit the creation of any system which aims to ingest declarations of beneficial ownership and create auditable records.

This note is designed for use by technical professionals, especially those with a role in the database design and technology architecture of company registers and/or sector registers that will publish to BODS.

Open Ownership's (OO) [Implementation Guide](#) contains best practice on implementing beneficial ownership transparency, including system design. The [Open Ownership Principles](#) (OO Principles) are a well-known framework for effective beneficial ownership disclosure.

This note is designed for the Opening Extractives programme which is an ambitious global programme aiming to transform the availability and use of beneficial ownership data for effective governance in the extractive sector.

The programme combines political and technical engagement, to support countries implementing beneficial ownership reforms and to enable the use of the data by governments, civil society and companies. Findings and evidence from the programme will be communicated globally, leveraging the tools and knowledge developed to drive impact beyond the focus countries.

2. Key concepts

a) What is beneficial ownership?

A beneficial owner is defined as the natural person that can be found at the end of an ownership chain and who benefits from the existence of the company. Often there is just a single link between a beneficial owner and a company, but sometimes it can include long and complex ownership chains of multiple legal entities.

Ownership and control structures can be complex

People may benefit from company activities through ownership or by exerting another form of control. It is not uncommon for people to hold a mixture of interests in a company which cumulatively make them beneficial owners.

Beneficial ownership differs from legal ownership. Companies can own or control other companies and are known as legal persons. Historically, transparency on company ownership has focused on legal ownership, or the level of ownership immediately above a company.

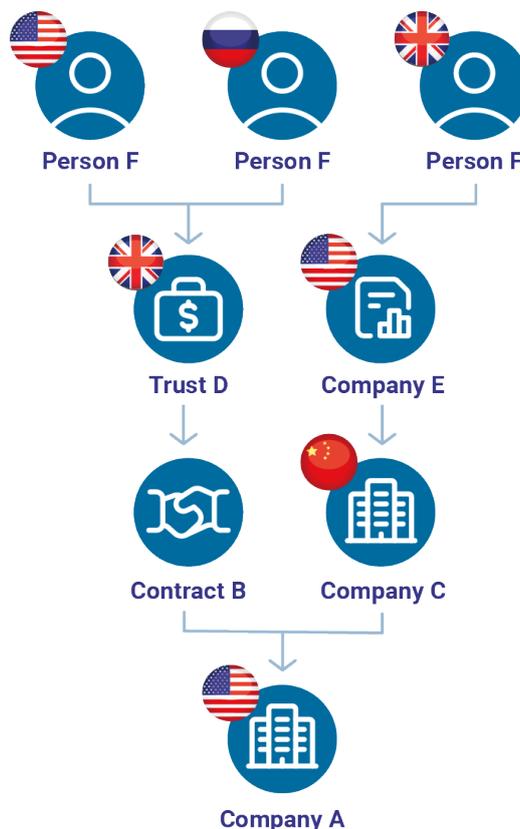
Company ownership and control can change over time. There are many reasons why stakeholders may benefit from knowing the ownership and control interests in a company at a particular point in time in the past.

Beneficial ownership is international. A full picture of company ownership is difficult to obtain for companies with complicated company structures across multiple jurisdictions. The [Beneficial Ownership Disclosure Workbook](#) explores this, along with how it may affect the extent of information disclosed in a register.

There are legitimate reasons why companies may not be able to declare all of their beneficial owners. One of the OO Principles [states](#) that “data should be collated in a central register”, but there are nuances around what data should be published. This is explored in [OO’s public access briefing](#).

[Beneficial ownership in law: Definitions and thresholds](#)

provides definitions and explanations of many terms relating to beneficial ownership that will be used throughout this document.



b) The Beneficial Ownership Data Standard (BODS)

The [Beneficial Ownership Data Standard \(BODS\)](#) provides a structured data format, along with guidance for collecting, sharing, and using data on beneficial ownership. BODS data is expressed using JavaScript Object Notation (JSON), making it suitable both for publishing online as open data and for transferring between systems as a data exchange format.

Some key characteristics of BODS are detailed below:

i) Statement-based data structure

The aim of a beneficial ownership declaration is to support the connection of entities to their ultimate beneficial owners, and to provide an auditable record of beneficial ownership over time.

There are three top-level JSON objects in BODS. They are:

- [ownership-or-control statement](#);
- [entity statement](#);
- [person statement](#).

BODS uses these in linked statements to describe how a person or entity is connected to another entity by means of ownership or control. Linking statements together helps to build up an understanding of ownership chains when beneficial ownership is indirect.

ii) Different categories of ownership and control

A beneficial ownership declaration should be able to describe the many mechanisms by which ownership and control of an entity may be exercised.

BODS contains a [codelist of interest types](#) to lay out recommended codes and explanatory descriptions. There is [ongoing work to add new codes to that list](#) to support this requirement.

iii) Auditability

A series of beneficial ownership declarations should contain a reproducible audit trail of changes made over time. To achieve this, published BODS statements should be treated as a write-only ledger, with new statements being issued to amend data contained in older statements and those new statements appended to the ledger.

To support auditability, a core requirement of a beneficial ownership database is that it can produce a set of statements that describe the beneficial ownership of an entity **at any point in the present or the past**. This requires some duplication of core data, resulting in the need for larger data storage.

iv) Identifiers

A beneficial ownership declaration should use [identifiers](#) to identify a unique person, entity, or ownership-or-control statement. Ideally, each person or entity should, correspondingly, be identified by a unique identifier in any identification scheme.

Statement identifiers

Advice on the creation of entity, person, and ownership-or-control statements is provided in the [BODS documentation](#). Of key importance is that two different statements SHOULD never have the same identifier, and once an identifier is assigned to a statement, the identifier SHOULD NOT change.

Real-world identifiers

Technical guidance on best practice when it comes to real-world identifiers is provided in the [BODS documentation](#). In particular, it is worth noting that:

- **Entity identifiers** consist of the scheme and the identifier as separate fields;
- **Person identifiers** also consist of the scheme and the identifier as separate fields, although they might be held internally, possibly in an encrypted form, and not published.

v) Exemptions and other missing information

When companies are unable to declare information about beneficial ownership, it should be made clear on what basis they are unable to do so.

Both [Entity](#) (unspecifiedEntityDetails) and [Person](#) (unspecifiedPersonDetails) statements contain objects for declaring missing information. These objects include a value from the [UnspecifiedReason codelist](#) and another value with a textual description for any supporting information.

3. Information flows

The guidance in this document assumes an information flow as described in the systems diagram below.

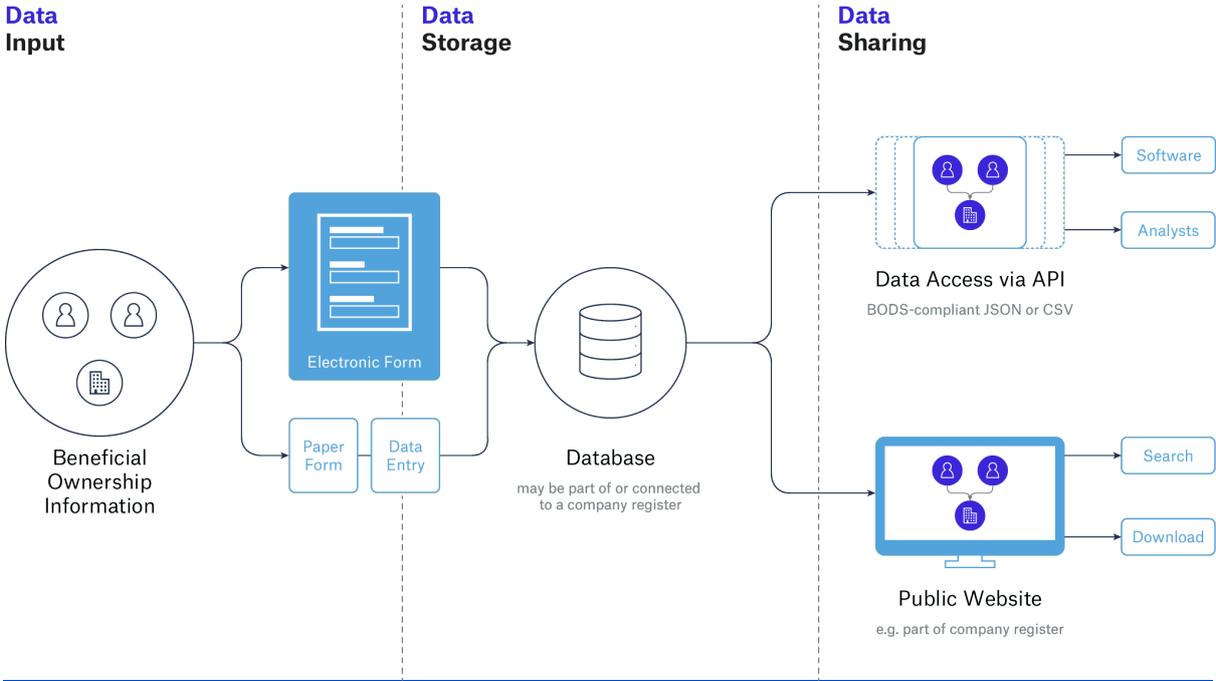


Figure 1. Beneficial ownership systems diagram (example)

Individual declarations of beneficial ownership are made on behalf of entities. These might be on paper or via electronic forms. This data is entered into a central database which can publish beneficial ownership records that comply with the BODS schema.

There are [good practices in form design](#) that help ensure good quality data is provided in entities' declarations of beneficial ownership.

4. Data capture

At the point of data capture, a series of claims are made about an entity and the ownership-or-control relationships that connect the entity to people and/or other entities. The beneficial ownership forms used to collect this data, whether paper-based or online, should be completed by a person nominated or authorised by a declaring company.

As shown in figure 1, this data may be entered directly via an electronic form or indirectly by filling in a paper form. Consider the following factors when designing a methodology for data capture:

a) Initial and update declarations

Consideration should be given to both initial declarations and updates, known as 'update declarations'.

An update declaration could use pre-filled fields from the previous declaration, where the person making the declaration has the ability to update them. This will reduce errors and duplication. Also, capturing how and when a change in beneficial ownership occurs is key to being able to publish BODS statements from any point in the past.

b) Clarity about what is being declared

It should always be clear to the person submitting the declaration exactly what information they are declaring to be true. This is especially important with partial updates and corrections.

For example, a person may need to update the name of a beneficial owner. When they submit that updated information, are they also declaring that:

- All of the information about the beneficial owner is correct, or
- All of the beneficial ownership information is still correct, or
- *Only* that the name of the beneficial owner is now correct?

The interface must make the limits of what is being declared as true clear to the person making the declaration.

Those limits should then be respected and maintained as the information is processed and stored by the system. Ultimately, only the information actively declared to be true at the point of submission will be represented as new statements in the system.

This clarity is important from both a data integrity and a legal point of view.

c) Avoiding duplication

In an ideal running of a beneficial ownership database system, when a declaration is made about a natural person, there should be processes that check for an existing record of that person in the system.

Where there is an existing record, it should be presented in the interface, or printed on the form, for the declarer to update with any changes to the beneficial ownership of the entity.

Real-world identifiers and cross-checks with personally identifying forms of identification are two methods that could be used to avoid duplication of person record data.

d) Form design

Well-designed forms make it as easy as possible for the people completing them to provide accurate and unambiguous information. This reduces the number of accidental errors. Submitting more accurate information becomes easier, whilst disguising deliberately false information as mistakes becomes harder.

[Beneficial ownership declaration forms: Guide for regulators and designers](#) introduces a number of considerations for form development and presents a [worked example of a form](#) that addresses these issues. We recommend that this is read by anyone with the task of developing a beneficial ownership database.

Data entered by the person making the declaration will sometimes not be published in BODS statements due to privacy, and the form should make it clear to them when this is the case.

Wherever possible, use data capture methods that collect structured data, as described in the following section.

e) Collecting structured data

Dropdown lists, checkboxes, and radio buttons are all examples of data capture methods that collect structured data.

BODS specifies a [number of its own codelists](#) as well as using standardised open codelists. Providing the person making the declaration with closed data capture methods for these fields will improve the accuracy of the data that is captured. Free text entry is not recommended due to the unstructured nature of the data collected.

BODS codelist example

The [interestLevel](#) codelist contains the values direct, indirect, and unknown. An online form could implement a radio button or a dropdown box to capture one of those values.

Standardised codelist example

BODS specifies that countries should be identified by their ISO 2-letter codes. An online form should implement a dropdown menu using short names from the [ISO list](#) and populate the database with the 2-letter code.

f) Data lookups

Providing lookups to persons or entities is dependent on local systems and legislation. It may be possible to perform local checks on entities and people when they are disclosed in a company declaration.

These local checks might be made within the beneficial ownership register itself or held on other internal government systems. Wherever possible, these local checks should be performed. Some examples are given below.

Individuals (persons)

In some jurisdictions, national identity numbers are used and a beneficial ownership register may have access to them. It may be possible to ask the person making a declaration to enter a national identity number and then to confirm that person's details to be correct.

Companies (entities)

A beneficial ownership register may be an extension to a company register. This should make looking up domestically registered entities a possibility. It may be possible to ask the person making a declaration to enter a company number and have them verify the associated company's details.

g) Verification of data outside the jurisdiction of a register

Verifying the details of people or entities that are outside the jurisdiction of the register is more complex. OO recommends a risk-based approach, where choosing suitable verification methods involves an assessment of the risk of bad data being entered against the effort of making verification checks.

OO has published a [policy briefing on verification best practice](#). It is recommended that this is read by anyone with the task of developing a beneficial ownership database. There is also [a feature ticket in Github](#) which summarises the latest thinking from OO about verification in BODS.

h) Ongoing improvement

Time and resources should be allocated to improve data capture based on analysis of the quality of data that is being entered in the register. For example, by identifying common data entry errors and redesigning the part(s) of the form(s) where these occur.

5. Database structure

Below is an Entity-Relationship diagram of the structure that a beneficial ownership relational database could take. It does not contain details of all of the fields that should be captured.

For example, the “declaration” table contains the fields “person”, “entity”, and “relationship”, which are placeholders. Therefore, it is representative of the overall structure, **but does not represent a complete model that could be used to implement a beneficial ownership register.**



Figure 2. Example of a simplified database diagram

a) Tables

i) Declaration table

A table, where the declaration of the beneficial ownership is stored, should exist. This declaration table acts as an audit log of company declarations. It should:

- **Contain as much raw user input as possible**
It is important to make sure that no data is lost from data capture. This is to ensure that, if necessary, the processing of that data can be replicated.
- **Capture the best quality information the person making the declaration has on the known people, entities, and relationships**
Capture as much information as possible aiming for structured fields, such as codelist lookups. If collected online, the system should try to contain lookups to known people, entities, and codelists, as detailed in the data capture section above.
- **Be stored in a log form, making sure a new record is created for each new declaration**
Add a new row for each declaration with a date associated with each entry. UPDATES should not be applied to them. For example, if the only difference is the date of the declaration (such as for an annual return where there has been no change in company ownership or control), then a new row should still be added.

ii) Statement tables

The declaration data should be processed into statement tables for people, entities, and relationships. These tables should:

- **Represent the whole history (all changes) of the person, entity, or relationship**
When a declaration is made that differs from a previous one, then a new row should be added to the affected table. This potentially includes confirmation statements where only the declaration id and date are different, in jurisdictions where such confirmations are required by law.

For example, if a person's shareholding increases, then a new row should be added to the Relationship table with that updated information.

- **Only create new entries when there is a change from the last statement**
There should be less duplication than in the declaration tables and new statements need to be generated only if there has been a change.

For example, in the situation described above, where shareholding level has changed, there is no need to update the Person or Entity tables unless information has changed in them as well.

- **Generate BODS statements**
These tables should contain the data, processed from the declaration, that are then used to populate BODS statements. There should be a direct mapping between these internal database tables and their related top level objects in the BODS schema.

For example, the Person table(s) in a beneficial ownership database should be capable of publishing to all of the fields in a [BODS Person statement](#). An array in BODS is an indication of a one-to-many relationship that the database could implement using a parent-child table structure.

- **Contain rows whose values are rarely altered**

Any fields that will be published to BODS should be unchanged after the creation of the row that contains their values, but certain fields that relate to the processing of the data could be updated, e.g:

- Links to previous statement rows to indicate this row supersedes the previous statement;
- Date ranges of when the information was applicable.

- **Be fully reproducible from the declarations**

These tables should be able to be reproduced from information in the declaration tables for all fields apart from those ids that are generated by the system.

- **Contain the declaration id that produced these statements**

The system should publish the declaration id in each of the rows that potentially represent multiple changes across multiple tables. This means that all of the updates can be traced back to the originating declaration, supporting data use.

iii) Current tables

The system should process statement data to generate person, entity, and relationship tables that reference the current best-known information. This allows snapshots of current, up-to-date beneficial ownership information to be published. These tables should:

- **Contain a unique generated “id” for each person within the declarations of each entity.**
This enables a history of changes to each person and their relationship to that entity to be recorded and reported. This is in addition to any externally generated ids, such as a national identification number (person) or a company number (entity).
- **Contain a unique “id” for each person, entity, or relationship within the declarations of the database.**
This supports the understanding of the case when a person is a beneficial owner of two or more entities. These ids may be internally generated or externally generated ids, such as a national identification number (person) or a company number (entity).
- **Link to the most up-to-date statement record**
This enables the current state of each person, entity, and relationship to be recorded and reported.
- **Be reproducible from the statement tables**
This table should be showing current-point-in-time information and could be calculated from the statement tables if needed, apart from the “id” fields.

b) Fields

This section highlights notable fields in the declaration, statement, and current tables in a beneficial ownership database. It does not attempt to cover every field in the database.

i) Person fields

These fields should exist in declaration tables along with person and person statement tables.

- **References to person identifiers**

To uniquely identify a person, both internally and globally as part of a BODS publication, there should be reference to one or more identifiers. Consideration should be given to the use of internal database identifiers, BODS statement identifiers, and real-world identifiers.

It may be possible to construct BODS statement identifiers from internal database identifiers. See [Strategies for identifier creation](#) in the BODS documentation.

The three forms of real world identifiers that BODS accepts are a national identification number, a passport number, or a tax identifier. Verifying a person identifier avoids duplication of person data in the beneficial ownership database.

The declarer may enter this through a lookup which will check those identification details are valid according to the relevant database. At that point, it could be encrypted or hidden from database users through access controls to meet any privacy requirements in the jurisdiction.

When [publishing to BODS, a real world identifier is not a required field](#) and should not be published unless there are legal grounds to do so. The person statement id is required to be published and it can be used to uniquely identify the person.

- **Status fields**

These indicate the current state of the information about an entity, person, or relationship (as reported by a particular entity). Status fields will not necessarily exist in the declaration tables: they can be derived and exist in the statement and current tables.

Status values should, as a minimum, include the equivalent of “new”, “updated”, and “closed” Statements made over time about an entity, person, or relationship form a series. The first in the series would have the status “new”, and the final statement would have the status “closed”. All other statements in the series would have the status “updated”.

- **Name and address information**

When collecting addresses, the international nature of beneficial ownership should be considered.

For example, where addresses outside the jurisdiction of the database are possible, the system should not enforce local postal/zip code formats.

ii) Entity fields

These should exist in declaration tables along with entity and entity statement tables.

- **Organisation identifier**

Preferably using an [org-id](#) identifier in conjunction with a register unique identifier.

- **Status fields**

This indicates the current state of the information about this entity and may contain just “new”, “updated”, and “closed”.

iii) Relationship fields

These should exist in declaration tables along with relationship and relationship statement tables.

- **Links to person and entity tables**

The relationship statement table should also link to the person statement and entity statement tables.

- **Known information about beneficial ownership**

Such as shareholding; direct or indirect ownership.

- **Equivalent to ownership or control statement in BODS**

These fields should capture all of the details of the information provided about ownership or controlling interests. It should be possible to publish these as ownership-or-control statements in BODS.

- **Status fields**

This indicates the current state of this information and may contain just “new”, “updated” and “closed”.

6. Case studies

In 2021, Latvia and Armenia became the [first countries](#) to publish beneficial ownership data in line with the [Beneficial Ownership Data Standard](#) (BODS).

With an initial focus on extractive industry companies, Armenia has begun publishing beneficial ownership data in line with BODS via its company register (see this [example](#) and [EITI Armenia’s list of mining companies with declarations](#)). From early 2022, it will expand these beneficial ownership transparency measures to companies operating in the media and other sectors.

The [Armenian register](#) also makes use of the [Beneficial Ownership Visualisation System](#) to generate easy-to-understand visual representations of beneficial ownership structures for companies declaring in line with BODS.

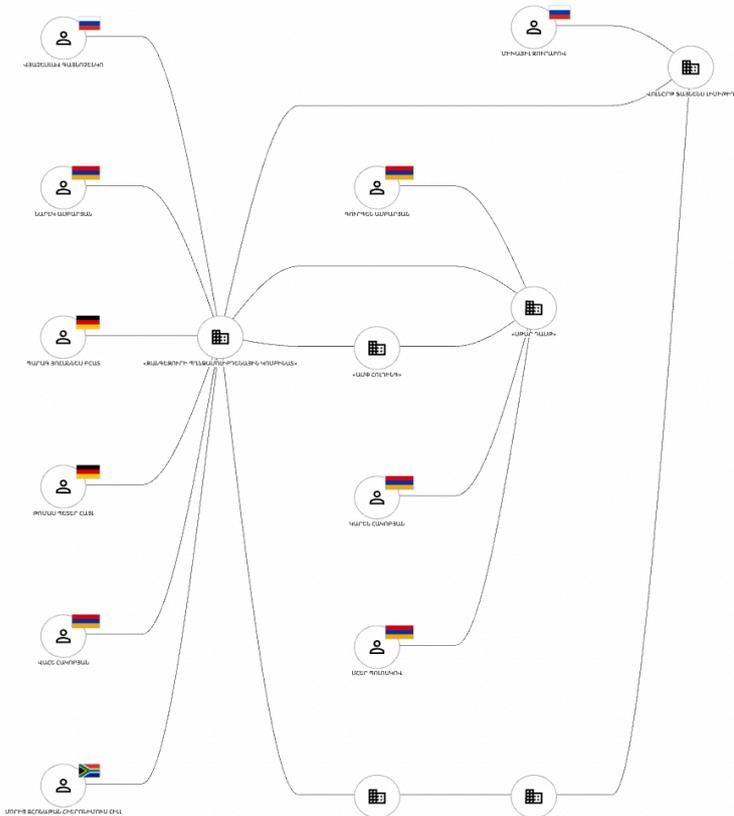


Figure 3. Beneficial ownership diagram for Zangezur Copper and Molybdenum Combine

Latvia is adopting BODS for its whole economy with data being published by the [Register of Enterprises of the Republic of Latvia](#). An [initial release of approximately 113,000 beneficial ownership declarations](#) was published by the Latvian authorities in BODS JavaScript Object Notation (JSON) format in January 2021 with further releases imminent. The data has been published for anyone to use under the [CC0 1.0 Universal Public Domain Dedication licence](#).

7. Further resources

[Beneficial Ownership Data Standard documentation](#)

[Beneficial ownership declaration forms: Guide for regulators and designers](#)

[Beneficial ownership disclosure workbook](#)

[Beneficial ownership in law: definitions and thresholds](#)

[Beneficial ownership transparency implementation guide](#)

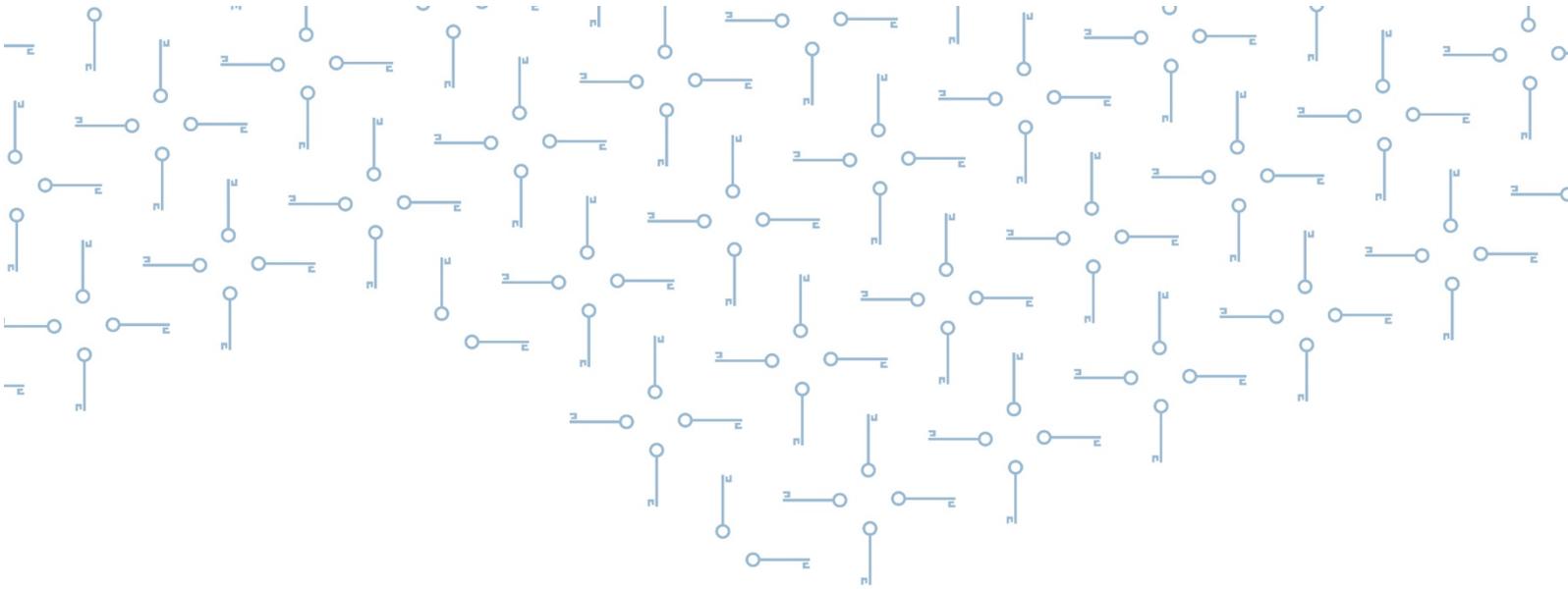
[BODS Feature Tracker](#)

[Making central beneficial ownership registers public](#)

[org-id: an international reference library of organisation identifiers](#)

[The Open Ownership Principles](#)

[Verification of beneficial ownership data](#)



Contact

🏠 Rådhusgata 26, 0151 Oslo, Norway

☎ +47 222 00 800

✉ openingextractives@eiti.org



**Opening
Extractives**

Unlocking the benefits
of ownership data

Jointly implemented by the EITI and Open Ownership

EITI

Open
Ownership 