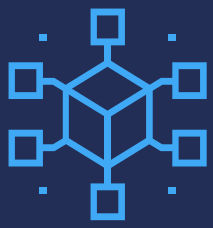


Open Ownership

Structured and interoperable beneficial ownership data

Policy Briefing

August 2022

Overview

The utility and value of beneficial ownership (BO) data is enhanced when the data is available in a structured format. Structured data refers to information that is highly organised according to a predefined model. Since the first jurisdictions have started collecting – and for some, publishing – BO information, some have done so as structured data whilst others have done so as unstructured data. Unstructured data does not follow a predefined data model: for example, if a reporting person is free to describe the relationship between a beneficial owner and a legal entity in their own words. Whilst structured data can be produced in non-digital environments, when structured data is available digitally it can be more easily read and processed by machines.

The first part of this policy briefing outlines the key benefits of collecting, storing, and publishing structured BO data. Jurisdictions that have published open, structured BO data have made a broader range of data analyses by additional users possible, facilitating early impact of beneficial ownership transparency (BOT) reforms.[1] To maximise the impact of BOT reforms, a disclosure regime should collect, store, and share BO information as structured data. This will lead to:

- **improved functionality** (page 6);
- **reduced costs** (page 8);
- **greater policy impact** (page 12).

Structuring data creates information that is predictable. Because the structure is predefined, users know what to expect from the data, and this makes it easy to work with. These benefits do not only apply to technical users. Non-technical users can benefit from structured data without ever having to use data directly. Because structured data can be made available in formats that can be readily processed by machines, computers, websites, apps, and other tools – for example, through a web interface, application programming interface (API), or in bulk format – it can be developed so people can access, visualise, and interact with relevant information in a variety of non-technical ways. Structured data can be integrated into both human- and machine-led processes that are either impossible or laborious with unstructured data. For example, making structured BO data available in bulk format allows users such as Financial Intelligence Units, procurement agencies, banks, and journalists to apply data science and machine learning techniques to identify suspicious patterns of ownership or beneficial owners that appear on other datasets of interest.

By removing the frictions associated with unstructured data, structured data decreases the cost of collection of data by governments and compliance to disclosure requirements by legal entities. It also reduces costs associated with maintaining and publishing data. Structured data reduces the cost and increases the impact of achieving the policy aims of BOT reforms by reducing the costs associated with use and analysis. Higher up-front costs associated with setting up the required systems are expected to be negated by lower costs associated with collection, storage, publication, use, and maintenance in the long run.

At the heart of structured data is interoperability, that is, being able to readily use the data with other sources, and integrate it into different systems and processes. The transnational nature of complex BO relationships makes combining BO datasets from different jurisdictions essential to gaining full visibility of ownership structures. Meeting the additional policy objectives for which countries pursue BOT – such as improving procurement processes and enforcing sanctions and campaign financing rules – also requires that the information be combined with other datasets. When BO data is structured and interoperable it is also easier to verify, as a greater range of verification mechanisms can be used, thereby improving data quality.

These benefits would be greatest following the wide adoption of a data standard such as Open Ownership (OO)'s Beneficial Ownership Data Standard (BODS).[a] BODS is a framework for publishing structured data about beneficial ownership in a format that can be read and understood by computer systems around the world. BODS has been adopted by both governments and the private sector, and a range of tools and applications have been developed around it.[2]

The second part of this policy briefing highlights what implementers need in order to operationalise structured BO data. In order to operationalise structured BO data, implementers should:

- **Create an enabling environment** (page 15) by taking a user-centred and interactive approach, and by establishing and progressively enhancing the legal, regulatory, and political framework to achieve technical goals relating to BOT. This includes ensuring a solid legal and policy foundation in line with the Open Ownership Principles (OO Principles) and providing sufficient resources.[b]
- **Establish principles for collecting and storing BO information** (page 16) by ensuring that, at a minimum, structured BO data:

  1. identifies the people, companies, and other relevant parties disclosed in a BO declaration by using unique identifiers and sufficient descriptive fields;
  2. describes the full range of relationships that can exist between parties disclosed in a BO declaration; and
  3. ensures BO disclosures are auditable.

  Implementers should ensure that **systems design and business processes** (page 19) underpin the aims of reforms on a technical level. Care should be given at the early stages of implementation to ensure the technical systems and database design meet the full functionality and access expected at the publication and data sharing stages.[c]

- **Realise potential and resolve uncertainty at the publication stage** (page 20). Ensuring published data is auditable by users is necessary to realise the data's full potential. This can be achieved by making published data available in a range of ways for both non-technical users as well as technical users and systems at scale, such as:

  - per-record search via a web interface;
  - browsing records via a web interface;
  - bulk format;
  - API access.

  Implementers should also decide on an appropriate licence for the data and provide sufficient accompanying documentation in the form of a publication policy, which should aim to resolve any uncertainties over the published data.

Structured data is a core tenet of the OO Principles, as it ensures data is readily combinable with other data, predictable, and reliable.[3] The OO Principles set the standard for effective BO disclosure and establish approaches for publishing high-quality, useful data. The OO Principles help ensure that published data is usable, accurate, and interoperable.

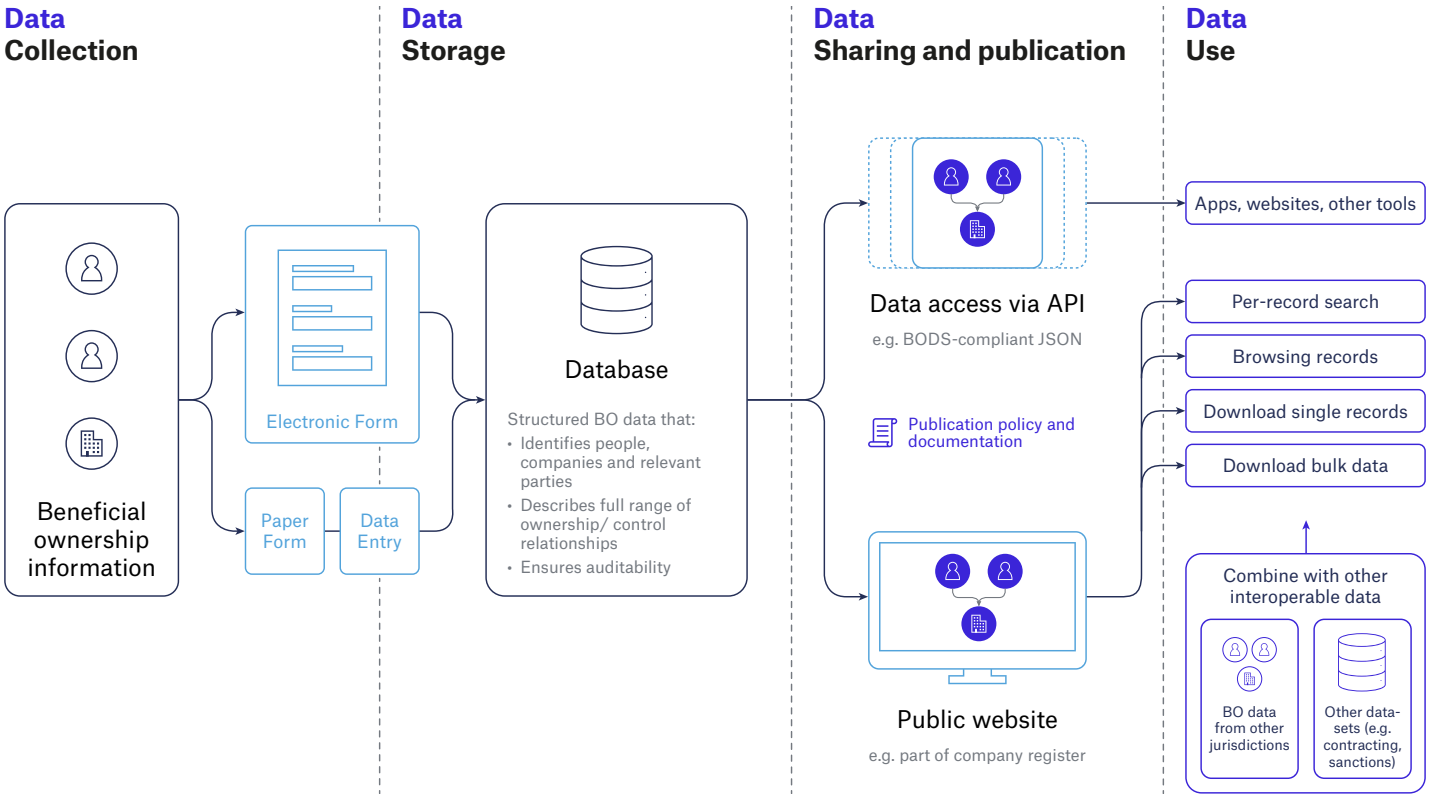
---

a For more information, see: "Beneficial Ownership Data Standard (v0.3)", Open Ownership, n.d., https://standard.openownership.org.

b For more information, see: "Open Ownership Principles", Open Ownership, updated July 2021, https://www.openownership.org/en/principles.

c For more information on database design, see: "Relational database design considerations for beneficial ownership information", Extractive Industries Transparency Initiative and Open Ownership, 16 December 2021, https://www.openownership.org/en/publications/relational-database-design-considerations-for-beneficial-ownership-information.

Figure 1. Example of a beneficial ownership disclosure system using structured data



Structured BO data improves its functionality (page 6), reduces the cost across all stages (page 8), and leads to greater policy impact (page 12). To achieve this, implementers should create an enabling environment (page 15), and data should be structured in a way that identifies and describes key elements of beneficial ownership (page 16). Digital systems and administrative processes need to fit together smoothly to enable BO information to be collected, stored, maintained, exchanged, and published (page 19). Uncertainties should be removed at the sharing and publication stage by adhering to open standards (page 20) and publishing a clear publication policy, including documentation and licensing information (page 21). Data can be made auditable by providing multiple ways to access data (page 21). Data standards such as BODS provide a structured data format, along with guidance for collecting, sharing, and using BO data (page 20).

Box 1: Key concepts and definitions

In order to understand how structured and interoperable BO data can contribute to meeting policy goals and the necessary policies to facilitate collecting, storing, and sharing structured BO data, it is necessary to explain a number of key concepts. Whilst some of these concepts apply more generally, the core focus is in the context of BO information.

Data is used to store and communicate information by machines and people. It is a unit of information. Data on its own has no inherent meaning, but acquires meaning when used or viewed in a particular context.

Structured data is data that is highly organised according to a predefined model.[d] It has sufficient content, organisation, and context to be interpretable by machines and to convey meaningful information

[d] Formally, "structured" and "semi-structured" data are different categories. For the purposes of BOT, however, it is enough to note that the same information will often be **stored** in structured form (in a relational database) and **published** in semi-structured form (such as JavaScript Object Notation (JSON) or XML). Both structured and semi-structured data are included in the definition used in this briefing, as long as sufficient information is conveyed through structure and context.

about beneficial ownership (see Table 1). Structured data can be created in non-digital environments, but in this briefing it refers to digital data.

Machine-readable data is data in a format that can be readily processed by a machine or computer. Machine-readable data must be digital structured data.

Data is **interoperable** when it can be readily used with other sources of data and integrated into different systems and processes. Interoperable BO data, for example, might use a widely agreed method for describing company numbers, allowing datasets from multiple jurisdictions to be joined together.[4] Interoperable BO data may also be joined together with non-BO datasets, such as contracting data.

A **data standard** provides a documented set of rules and agreements for how data is structured, published, and contextualised. It can also cover data format, definition, transmission, manipulation, use, and management. Standards provide a common language for producing and understanding data, regardless of its origin, and embed a high degree of interoperability by design. Structured data that does not adhere to the same data standard can be – but is not necessarily – interoperable, but would require an extra step of translation to join the data together. BODS, discussed in more detail later, is a data standard which sets rules for high-quality BO data.[5]

Table 1. Unstructured (left) versus structured (right) beneficial ownership data

Nature of ownership or control
This beneficial owner indirectly herself, or through her children, owns 27% of the declaring legal entity's shares through the following shareholders of the legal entity (1) "Angerujjheit B.V.", registration number in the Netherlands 64739564, registered office: Byterslaan 105, NL-4722GF Amsterdam, Netherlands; (2) "RigaTech Systems Ltd.", registration number in the British Virgin Islands: 396654, registered office: P.O. Box 124, Offshore Incorporations Centre, Road Town, Tortola, British Virgin Islands

Nature of ownership or control	
% Aggregate share ownership	27
% Aggregate control via voting shares	27
Direct share ownership in declaring entity	0
Direct voting control over declaring entity	0

1.1 Intermediate legal owner(s)	
Legal owner 1	
Name	Angerujjheit B.V.
Registration authority	Commercial register of the Netherlands
Registration number	64739564
Legal owner 2	
Name	RigaTech Systems Ltd.
Registration authority	Registry of Corporate Affairs, BVI
Registration number	396654

On the left-hand side of this hypothetical example, data is unstructured, as all the information relating to the beneficial owner and her relationship with a company is in a single text field. On the right-hand side, data is structured, as the information is separated out into different fields in a standardised way.

Benefits of structured and interoperable data

Collecting, storing, and making available BO information as structured data has a number of advantages and benefits in terms of using the information. Broadly, it:

- **improves the functionality** of BO information;
- **reduces the cost** of producing, using, and maintaining BO information; and
- has a **greater chance of meeting BOT policy goals** than unstructured data.

Improved functionality

Box 2: Searching and exploring beneficial ownership data in the Open Ownership Register

The Open Ownership Register (OO Register) allows people to search for beneficial owners and legal entities, and to explore ownership connections across jurisdictions by combining BO data from multiple registers.[e] Ownership structures can be visualised using built-in tools, and data on individuals or entities can be downloaded.

Users are not required to have technical knowledge in order to do this. The OO Register provides this functionality by ingesting data sources from different jurisdictions that publish BO information as structured data.

Because the data is structured, the OO Register is able to provide functionality that some jurisdictions' own online portals do not. For example, one of the jurisdictions from which the OO Register ingests BO data is the United Kingdom (UK). On the UK's BO register, it is possible to search for company names, company numbers, and officer names, but not beneficial owners. On the OO Register, it is possible to search the UK data – along with data from other jurisdictions – by beneficial owners' names.

[e] The OO register ingests data from four national BO registers. See: "Open Ownership Register – Data Sources", Open Ownership, n.d., https://register.openownership.org/data_sources.

Structuring data creates information that is **predictable**: because the structure is predefined, users know what to expect from the data. This makes it easy to work with. Non-technical users can benefit from structured data without ever having to use data directly. Through websites, apps, and other tools, people can access, visualise, and interact with relevant information in a variety of non-technical ways. Structured data can be integrated into both human- and machine-led processes that are either impossible or laborious with unstructured data. These include:

- **searching and querying** BO data to find records relating to known natural or legal persons, or to find unknown persons that match particular criteria (for example, an address or nationality) (see Box 2);
- **bulk analysis** of single or combined data sets to find patterns or red flags relating to beneficial ownership, or to assess and improve data quality (see Box 3);
- **automation of business processes** that involve company ownership information, (for example, supplier onboarding and customer due diligence) (see Box 4);
- **converting complex information into user-friendly formats** that are appropriate for a given context and audience (for example, a visualisation or summary data table) (see Box 5).

Data standards make it easier to realise the full benefits of this increased functionality because BO data is predictable regardless of its origin, making collaboration and data use in multiple organisations less challenging.

Box 3: Bulk analysis of beneficial ownership data in the United Kingdom

In 2016, the UK became one of the first countries to create a public register of the beneficial owners of companies. The UK register, called the register of People with Significant Control (PSC) was pioneering; it published open, structured data, allowing others to analyse the data in bulk. Bulk analysis was conducted in November 2016 by a consortium of non-governmental organisations.[6] Much was learned from this analysis, both to improve the UK register and for others considering or establishing public BO registers.

The analysis identified a number of issues with the data quality. For example, it showed that allowing people to type their nationality into the relevant field resulted in over 500 spellings of "British", and 10 beneficial owners listing their nationality as Cornish (a county in England).[7] The analysis also revealed 2,160 beneficial owners provided their date of birth as 2016, and others declared theirs as being as far into the future as 9988. Following the findings of the data analysis exercise in 2016, Companies House (CH) included a prompt within the PSC Register when users provide a date of birth which is below age 16 or over 100, and preventing people entering an age over 110.[8]

Similarly, in 2017, Transparency International and Bellingcat conducted bulk analysis on Scottish Limited Partnerships (SLPs), an obscure entity type that was initially not included within the disclosure requirements. They found that 71% of all SLPs registered in 2016 were controlled by companies based in secrecy jurisdictions.[9] This analysis, combined with reporting on wide-scale abuses involving SLPs,[10] led the UK to bring SLPs within the scope of its disclosure requirements. Subsequent bulk analysis showed a decline in SLP registration, suggesting a deterrent effect of transparency.[11]

Box 4: YouControl: Combining structured beneficial ownership data with other data sources for customer due diligence in Ukraine[12]

YouControl is a Ukrainian company committed to business transparency that has developed an "analytical system for compliance, market analysis, business intelligence, and investigation."[13] YouControl combines aggregated data from 180 sources – including structured data from Ukraine's BO register – with its own analysis to provide company profiles with a substantial amount of information, including anything that should raise red flags, such as unpaid taxes, pending lawsuits, and failure to file returns. It uses its own algorithms to calculate financial stability scores and the risk that the company is operating from a fictitious location.

A number of case studies on the website provide examples where companies have saved hundreds of thousands of dollars by using YouControl as part of due diligence processes to identify fraudulent businesses before entering into business with them.[14]

Reduction in cost

Interoperable, structured BO data has significant cost advantages over unstructured information throughout its lifecycle by reducing the cost of:

- publishing, maintaining, and upgrading BO data;
- using BO data;
- realising policy benefits; and
- compliance with disclosure requirements.

Publishing, maintaining, and upgrading beneficial ownership data

The decision to structure BO data generally involves a higher upfront cost that is paid off over time through lower operational costs. Structuring data requires documentation and resolution of underlying ambiguity. Documentation ensures that data is understandable and interpretable by any user by explaining the structure of the data and its contents. For example, documentation can make clear that date fields follow a specific format. Knowing that, for example, the International Organization for Standardization (ISO) 8601 date format is followed (YYYY-MM-DD), means it is clear that 2022-07-11 is 11 July 2022 rather than 7 November 2022.[15]

Unambiguous documentation reduces ongoing maintenance costs as well as frictions associated with data use. This also applies when the structure of data needs to be changed as part of an upgrade process. The predictability of structured data allows for more accurate resourcing of ongoing costs as well as clear planning and documentation of upgrades to data and business processes.[16] If data is published to a data standard, then many of the complexities of implementation will have already been considered, reducing initial development costs. The published data will also benefit from an existing set of documentation and technical tools.

To illustrate, in Slovakia – one of the early implementers, therefore not having the benefit of lessons learned from established practices – "the development of a new register cost around EUR 330,000", and "[t]he yearly operating costs are estimated to be around EUR 33,000".[17] Similarly, a study in the European Union estimated the operational costs for member states to make BO data available as structured data to be EUR 50,000 and 4-10 staff in full-time employment per year.[18] To contextualise these costs, a post-implementation valuation of user benefits in the UK concludes that, "in aggregate, the annual user benefits of CH data are estimated to be between £1 billion and £3 billion per year. This is likely to be an underestimate as it only includes benefits for Companies House Service (CHS) users." Different modes of access made possible by publishing the information as structured data were noted as generating the greatest user value.[19]

Using beneficial ownership data

Structured BO data is **quicker and cheaper to use** for the public and private sector, reducing the costs of accessing critical information and of linking BO information to other datasets and systems. To illustrate, the cost for law enforcement to access unstructured information was noted as a significant motivation behind increasing BOT in Canada.[20] Where structured BO data is available, this can be readily integrated into, for example, in-house or outsourced onboarding, due diligence, or regulatory technology services. In research conducted by OO on the use of BO data by financial institutions, one bank estimated the cost of handling unstructured PDF BO reports from one jurisdiction to be "an extra EUR 7 on top per review" for customer due diligence processes, in comparison to a jurisdiction that makes its BO information available as open, structured data. This was noted to be "a big cost for the bank", as the bank conducts thousands of reviews per year.[21]

Realising policy benefits

Structured BO data can also reduce the cost of realising policy benefits by facilitating interoperability. For example, when data is structured according to the same data standard, it provides a common language that enables different systems to speak together and can reduce the cost of realising policy aims. Additionally, these policy aims may generate economic benefits.[22] These benefits usually accrue outside the implementing agency and therefore require a holistic approach to assessing the costs and benefits of structuring data and budgeting.[23]

Structured data facilitates interoperability in multiple ways. Firstly, **technical interoperability** – for example, using a data standard to publish BO data – can allow tools and analysis built in one context to be reused in another, reducing costs and development times. By publishing data to BODS, Armenia was able to use the OO BODS visualisation library to automatically generate diagrams of ownership structures, thereby rapidly deploying advanced functionality in the first version of its register (see Box 5).[24] Similarly, Bluetail, a prototype tool for linking beneficial ownership and procurement data, is available in several languages and has been used in Indonesia, Kenya, and South Africa (see Figure 3).[25]

Box 5: Structuring beneficial ownership data in Armenia

In its implementation of a central BO register, Armenia has moved from collecting and publishing unstructured BO data to publishing structured data in line with BODS for use across government and beyond.

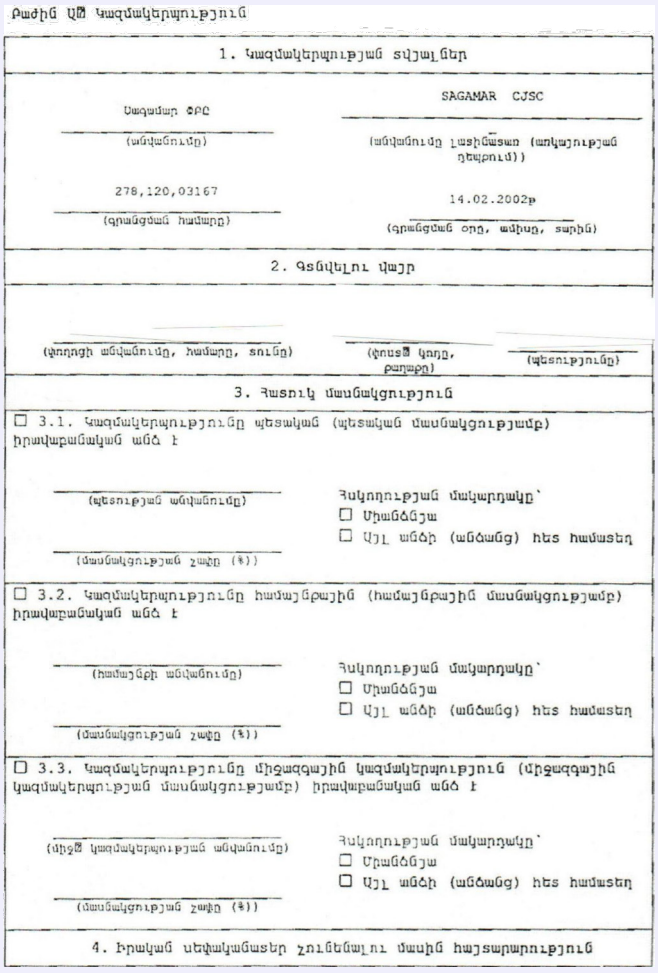
Armenia's implementation began by piloting requirements to disclose BO data for companies operating in its extractive sector. This pilot programme was used to develop and test systems for the collection and publication of data from a high-risk sector before gradually expanding the requirements to companies in all sectors of the economy by the end of 2023.

To expedite data collection and publication for the pilot in 2020, Armenia opted to collect BO information via paper forms as an interim step whilst finalising the upgrades to its registry software. These paper forms were then converted into PDF format and published online. Though this permitted some limited analysis of the disclosures by media organisations, the absence of structured data limited its utility and made carrying out any checks on the information disclosed significantly more laborious.

In September 2021, Armenia published its first structured data, becoming one of the first countries to publish data using BODS by incorporating the standard into the design of its register software.[26]

Although Armenia is still dealing with some teething issues, the adoption of the data standard will permit new uses of its BO data and enables the country to incorporate different tools developed for BODS directly into its register. For example, the BODS data visualisation library enables users to easily produce ownership graphs from the data and facilitates the understanding and analysis of company structures.[27] From late 2023, Armenia also plans to draw structured BO data from the State Register into its procurement systems as a means of tackling corruption, collusion, and bid rigging during government tender processes.

Figure 2. An unstructured beneficial ownership declaration from 2018 (left) versus a structured beneficial ownership declaration as shown through the Beneficial Ownership Data Standard data visualiser on Armenia's beneficial ownership data portal in 2022 (right)



Բաժին ՍՊ կազմակերպություն

1. Կազմակերպության տվյալներ

Սագամար ՓԲԸ (անվանումը) | SAGAMAR CJSC (անվանումը լատիներեն (առկայության դեպքում))

278,120,03167 (գրանցման համարը) | 14.02.2002թ (գրանցման օրը, ամիսը, տարին)

2. Գտնվելու վայր

(փողոցի անվանումը, համարը, տունը) | (փոստային կոդը, բաղաքը) | (պետությունը)

3. Հաստիք մասնակցություն

☐ 3.1. Կազմակերպությունը պետական (պետական մասնակցությամբ) իրավաբանական անձ է

(պետության անվանումը) | Հսկողության մակարդակը՝ ☐ Միանձնյա ☐ Այլ անձի (անձանց) հետ համատեղ

(մասնակցության չափը (%))

☐ 3.2. Կազմակերպությունը համայնքային (համայնքային մասնակցությամբ) իրավաբանական անձ է

(համայնքի անվանումը) | Հսկողության մակարդակը՝ ☐ Միանձնյա ☐ Այլ անձի (անձանց) հետ համատեղ

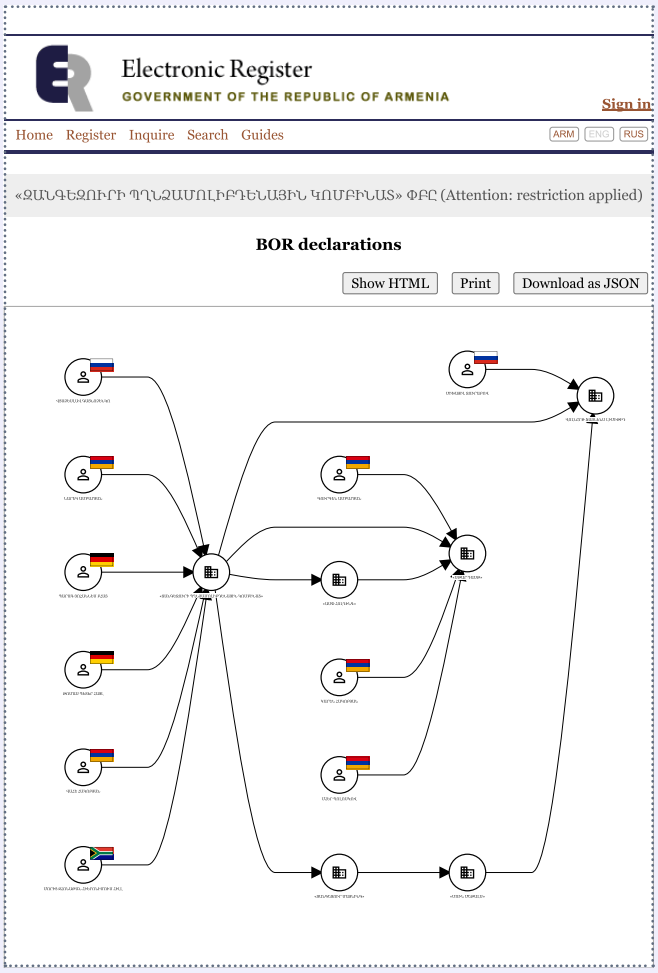
(մասնակցության չափը (%))

☐ 3.3. Կազմակերպությունը միջազգային կազմակերպություն (միջազգային կազմակերպության մասնակցությամբ) իրավաբանական անձ է

(միջազգային կազմակերպության անվանումը) | Հսկողության մակարդակը՝ ☐ Միանձնյա ☐ Այլ անձի (անձանց) հետ համատեղ

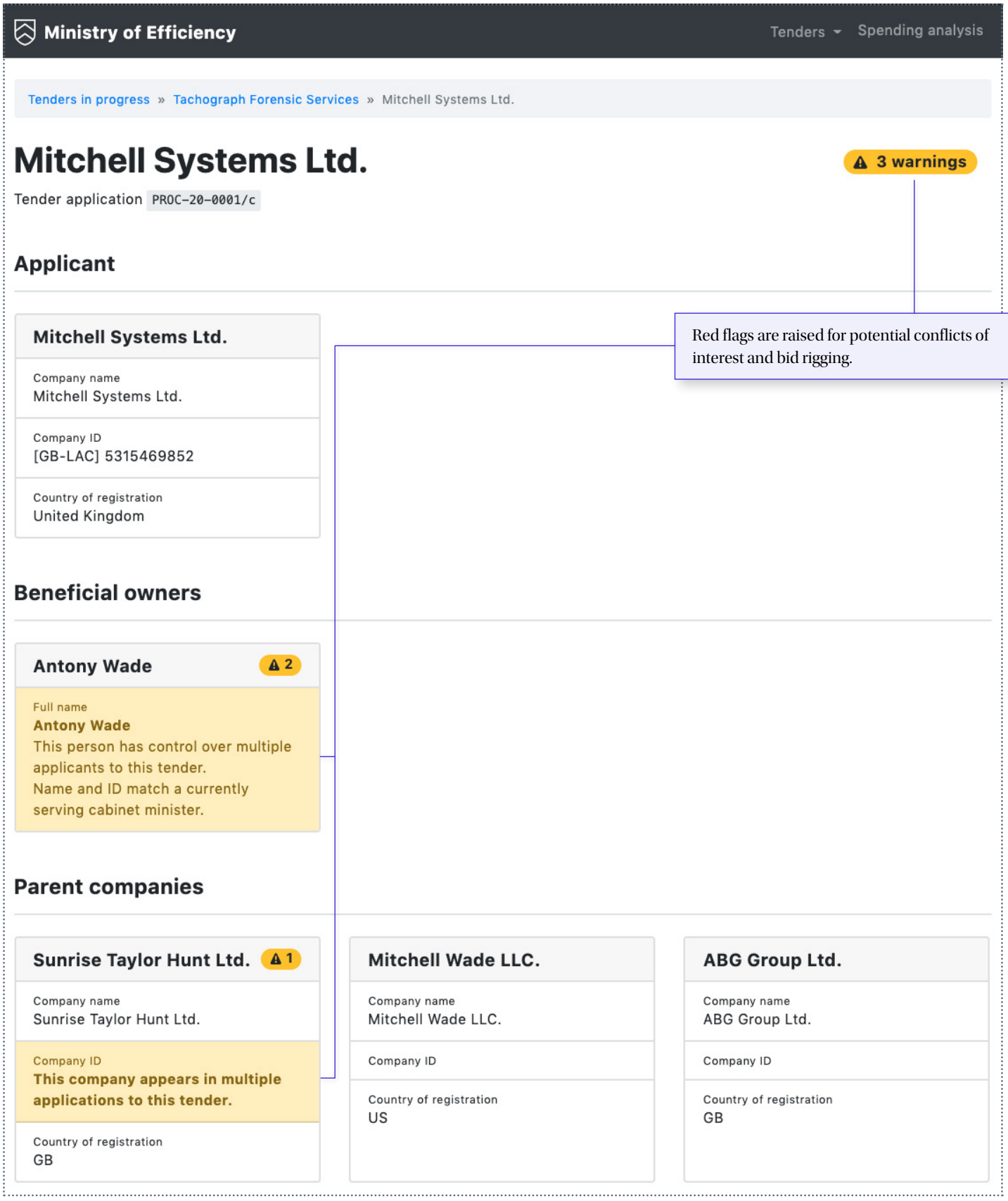
(մասնակցության չափը (%))

4. Իրական սեփականատեր չունենալու մասին հայտարարություն



Source: www.e-register.am.[28]

Figure 3. Bluetail: Combining structured beneficial ownership data and structured contracting data



OO has built on a prototype developed by mySociety and Spend Network called Bluetail, which shows how structured data published to BODS and data published to the Open Contracting Data Standard are interoperable. These datasets can be combined to automatically raise red flags for corruption and collusion risks when procurement officers screen tenders.[29] Source: bluetail.herokuapp.com.

Structured data also improves **contextual interoperability**. For example, whilst a single global identifier for legal entities may improve technical interoperability, adopting a common method of assigning and disambiguating company identifiers – such as the approach developed by org-id.guide[30] – may prove more pragmatic and result in datasets that are more interoperable in practice. The Organisation for Economic Co-operation and Development's Common Reporting Standard for tax information is another example with a high level of contextual interoperability.[31] Having high levels of contextual interoperability allows resources otherwise spent on cleaning data to be spent on analysing data. The bulk of artificial intelligence effort in corporate data is currently spent on disambiguating data. Policy makers are in a position to work on the coordination problem that would make this less necessary. More mature and established, but closely related, fields provide a useful model for building a shared infrastructure using interoperable and structured data as the baseline.

Compliance with disclosure requirements

Structured data is **cheaper to produce for declaring companies' beneficial owners and central registers**. A central register can provide a standardised way of declaring information – for example, in an online form – around which companies can standardise internal processes and procedures. To make submitting compliance easier and to reduce the submission of errors, Armenia also uses the BODS visualisation library in its declaration process (see Box 5). Unstructured data, by contrast, may be collected in a variety of formats, using a variety of bespoke processes that may require specialist expertise. Registers that verify incoming information can use structured data to cross-reference with other datasets rather than solely relying on expensive manual checks, as with unstructured information.[32]

Greater policy impact

Structured BO data is more likely than unstructured information to achieve BOT policy objectives, as the advantages discussed above make it a better option for achieving policy outcomes. BO information is rarely used in isolation to achieve specific policy goals, and it is often most valuable when combined with other datasets. To be useful, data on beneficial ownership needs to be contextualised with data on the identities or activities of companies or natural persons. Structured data makes it possible to link BO information to other datasets by joining on common fields.

BO data can be linked to other BO datasets from other jurisdictions to give a more complete view of transnational ownership structures. Beneficial ownership can be exercised through complex structures that include multiple jurisdictions, particularly in cases of corruption and money laundering. In a World Bank study of around 150 grand corruption cases, nearly all relied on corporate vehicles to conceal ownership, and the majority featured transnational structures.[33] The details of such structures will not always be available from the disclosures of a single jurisdiction, and to make sense of the cross-jurisdictional nature of company ownership it is often necessary to join multiple BO datasets from different jurisdictions or to combine BO data with data on legal ownership. Interoperable data allows the information from multiple registers to be easily combined. In the World Bank study, transnational investigations and building transnational cases were flagged as two priorities to combat the misuse of corporate vehicles: "solving a transnational corporate vehicle misuse scheme is like putting together a jigsaw puzzle, with investigators in different jurisdictions each holding separate pieces of the puzzle. To complete the puzzle, an investigator needs to have access to all the pieces."[34] Gaining visibility on transnational ownership structures will save resources that can be redirected in other areas of investigation (see, for example, Box 6).

BO data can also be combined with other datasets about individuals or legal entities for other purposes. For example, datasets relating to procurement and beneficial ownership can be linked through the use of common identifiers for legal entities.[35] Additionally, BO data can be combined with political campaign financing data to safeguard electoral processes.[36] Whilst datasets can still be linked without shared identifiers, it will be necessary to reconcile the data to distinguish which records refer to the same legal or natural person, a time-consuming and uncertain process.[37] Identifiers for real-world entities are therefore an essential requirement for the effective implementation of structured data (see, for example, Box 6).

Box 6: OpenSanctions and OpenScreening: Combining beneficial ownership and sanctions data using real-world identifiers

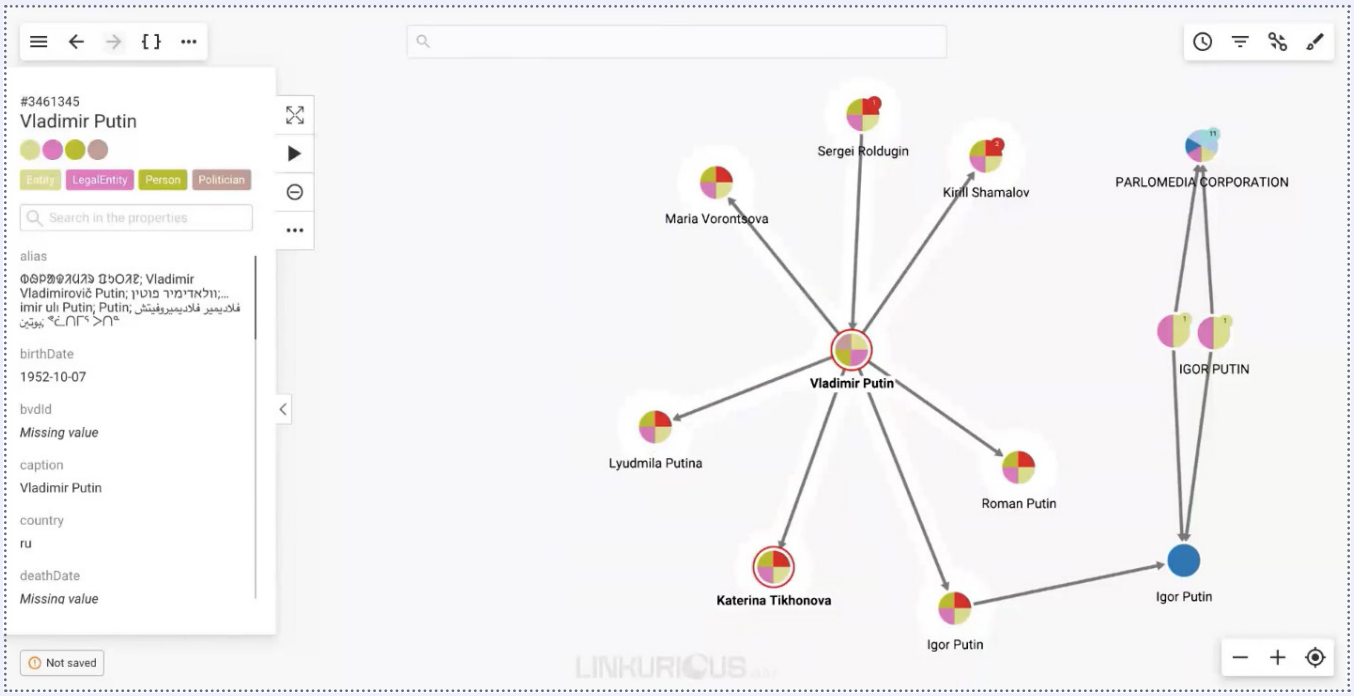
Sanctions against individuals or companies on the United States (US) Office of Foreign Assets Control list apply not just to the entity itself, but also to all companies majority-owned or controlled by that entity.[38] This provision, also adopted by some other authorities, makes visibility of full company structures and beneficial ownership an essential part of sanctions compliance.

OpenSanctions is an international database combining information on individuals and legal entities from multiple sources – including national BO registers – and linking this to sanctions and politically exposed persons lists. The project combines this information into a single dataset to help cross-check databases, detecting conflicts of interests and signs of illicit activity as well as customer due diligence in international dealings, and tracking political conflicts and sanctions policies across the world.[39] This requires deduplication and ensuring there is only one consolidated entry for each legal entity and natural person based on issued identifiers and descriptive data fields.[40] Depending on the information available, this may require manual work.[41]

OpenSanctions matches its dataset to the Legal Entity Identifier (LEI) reference data released by the Global Legal Entity Identifier Foundation (GLEIF). GLEIF has developed a global database[42] of corporations that have been issued a LEI, and also publishes information about ownership relationships between companies.[43] This allows OpenSanctions to enrich its data and identify additional relationships.[44]

For the OpenScreening project, OpenSanctions has mapped BO data from the Offshore Leaks database published by the International Consortium of Investigative Journalists (ICIJ) to the FollowTheMoney data model used by OpenSanctions and the Organised Crime and Corruption Reporting Project.[45] This data is then combined with OpenSanctions' own database, loaded into a Neo4J graph database, and uploaded to a data visualisation platform provided by Linkurious so that people can explore and visualise the connections between sanctioned or politically exposed people and leaked BO data using open data.[46]

Figure 4. OpenScreening



A visual example of OpenSanctions' graph data being combined with Linkurious' investigation software, which also plans to ingest data from the OO Register.[47] Source: resources.linkurious.com/openscreening.

The specific needs of the BOT policy area also mean that structured data has a significant advantage over unstructured data. Unstructured BO information quickly creates resource constraints that act as a barrier to action and ambition, whereas structuring BO data opens up **new types of analysis and possibilities** (see Figure 4). Unstructured information tends to create analysis that is resource-intensive and small in scale, which is viable in only a few contexts, such as investigations relating to grand corruption cases. Structured data allows for macro-scale and experimental analysis of entire BO datasets at low cost. This opens up different kinds of policy impact that are impossible without structured data, for example: examining national security risks relating to BOT across specific sectors of the economy; statistical analysis of the effectiveness of policy interventions; or the development of red-flagging procedures to better target manual investigations. In the UK, the Competitions and Markets Authority has started using BO data to analyse the concentration of ownership in specific sectors once common ownership and control are taken into account.[48] In another example, researchers used a commercial company ownership dataset (which includes beneficial ownership) to map stranded fossil-fuel assets to understand market risk in the energy transition.[49]

Structured data also **enables timely access to auditable information**. This is particularly relevant for law enforcement and national security purposes, as it removes the need for manual searches and requests, which risk tipping off suspects, or delaying investigations. Automated controls and access logs can ensure that confidential or restricted information is used appropriately.

Finally, structured data creates the **potential for more trusted data and more resilient systems through verification and register-wide analysis**. This can include automated verification of data submitted to the register; checks and challenges based on threshold-driven tests; and the collection and analysis of metadata on submissions (e.g. the use of particular company formation agents) that may reveal vulnerabilities in particular areas of the system. This also means that personal information can be shared in a way that conforms with privacy and data protection legislation by building in varying permissions for different data fields for specific data users. For example, a registrar checking a passport number with a different government agency for identity verification purposes would simply be able to receive information on whether or not the values match existing records, without sharing additional personal information.[50]

Operationalising structured beneficial ownership data

In order to operationalise structured BO data, implementers should consider:

- creating an enabling environment;
- establishing principles for collecting and storing BO information; and
- realising potential and resolving uncertainty at the publication stage.

Creating an enabling environment

The potential of structured data is ultimately limited by the legal, regulatory, and political environment. No amount of technical work can overcome the limits established in these areas. Therefore, the most important task for policy makers is to establish and progressively enhance the legal, regulatory, and political framework to achieve technical goals relating to BOT.

User-centred and iterative approach

Implementers should follow a user-centred design approach to technology. This will ensure stakeholder buy-in and create a system that is fit for purpose.[51] Implementation should be considered as an iterative process, rather than as a one-off project. The success of some of the early registers has been due to their iterative approach (see, for example, Box 3). By engaging data users as part of and after implementation, and making iterative improvements to their systems, early adopters have converged on elements now considered cornerstones of effective BOT reforms.

The legislation should support an iterative approach. If the legislative is too prescriptive, then any small change to the administration of the register will require legislative amendments. This can make iteration less likely because legislative amendments take time and effort, and every time it is debated in a legislative body, it opens itself up to a potential political backlash. In contrast, if the registrar is given more flexibility, it can be easier to make these iterations. An example of this is when jurisdictions include BO declaration forms within secondary legislation or regulations. This prevents the registrar from being able to make iterative improvements to the collection forms to address any teething problems (see, for example, Box 3).

Legal and policy environment

In addition to facilitating an iterative approach, it is fundamental to ensure that the legal foundations exist to collect sufficient data to achieve BOT policy goals. The usefulness of structured data is limited by how accurate and reliable, detailed, up-to-date, comprehensive in coverage, and auditable the underlying BO information is. The OO Principles provide a framework for building information quality into these crucial stages of the implementation process.[52] Data collection, verification, and data sharing may require agreements between multiple institutions to, for example, verify documents against an identity register or check declarations against data held by financial institutions. These types of agreements take time to agree and operationalise, and should be initiated early on. Whilst the specific details for making BO data structured are rarely included in legislation, the following features are all part of a solid legal foundation:

- robustly defining beneficial ownership;
- ensuring comprehensive coverage and the collection of sufficient details;
- creating the legal obligation to disclose and the appropriate triggers to make BO declarations;
- establishing the right legal basis for the processing and publication of information in conformance with privacy and data protection legislation;
- ensuring the right data agreements are in place for verification; and

– designing sanctions and enforcement regimes that drive up compliance and data quality.[53]

Technology can be used to make privacy and security a cornerstone of implementation. Not all BO information should be made public, and structured data provides the easiest and safest way to control and audit access by those who have the right permissions to the full or partial dataset, for example, by making a smaller subset available to the public than to domestic authorities and omitting data fields that are particularly sensitive, known as layered access. It is also important to establish a protection regime in order to allow exemptions to normal disclosure procedures and the retraction or redaction of already published data.[54] Structured data lowers the costs and risks associated with using or republishing data that might later be redacted data. Structured data also makes cross-jurisdictional coordination on privacy and security measures possible, reducing the risk of accidental disclosures in the case of a genuine exemption.

Resourcing

Policymakers should ensure that the success of BOT policy goals is not held back by technical resourcing issues. BOT is a complex policy intervention with expected benefits that are widespread and long term in nature, and implementation often involves both cross-government coordination and significant changes to systems and processes. Funding a technical system that can achieve these ambitions can be challenging. Any technical business case should be framed around identifying the full spectrum of expected policy benefits rather than minimising implementation costs for a single government agency. As these benefits usually accrue outside the implementing agency, a holistic approach to budgeting is required.[55] Beyond technical resourcing, BO disclosure regimes also require sufficient resourcing to ensure compliance, enforce sanctions, and query and verify data to ensure the integrity of the information.

Establishing principles for collecting and storing beneficial ownership information

In establishing principles for collecting and storing BO information, implementers should consider the minimum key elements of structured BO data and embedding these and other elements into systems design.

Key elements of structured beneficial ownership data

BO data describes ownership or control relationships between natural persons and legal entities. Depending on the disclosure regime, these relationships may be simple or more elaborate when described as data, but the three principle functions that BO data should perform remain constant. Beneficial ownership as structured data should:

1. identify the people, companies, and other relevant parties disclosed in a BO declaration;
2. describe the nature of the interests and relationships between parties (e.g. nominees) disclosed in a BO declaration; and
3. ensure that BOT disclosures are auditable and interoperable.

Identifying the people, companies, and other relevant parties

BO data should be able to describe all entities declaring under a given BOT regime, and it should be sufficiently flexible to incorporate all legal entities and arrangements that may be disclosed as part of a declaration. BO data must define and structure the fields needed to describe and identify the natural persons, legal entities, and arrangements involved in BOT disclosures. There are two complementary approaches to this:

1. Requiring the use of **unique, permanent, and resolvable identifiers** to allow the unambiguous identification of human and non-human entities involved in BOT processes.[56] Identifiers should have meaningful value to users of the data, including, for example, authorities from other jurisdictions.

   Examples of identifiers include a company number issued by a known corporate register, a national identity number from a specific jurisdiction, or an internally generated identifier for a beneficial owner.[57] Identifiers allow for the efficient and confident use of datasets, reducing the need for data cleaning and freeing up resources for high-value analysis. However, a relevant identifier must be available and, where available, collection, storage, and publication must be appropriate. Therefore, the collection of identifiers

should be a priority and considered in context, for example, by prioritising the collection of identifiers associated with domestic beneficial owners or choosing to only publish internally generated identifiers to support deduplication.

It can be difficult to find unique identifiers for non-domestic individuals or companies. For example, Canada has the Social Insurance Number (SIN), but foreigners who are not working in the country do not have SINs. The Province of British Columbia's register for the beneficial ownership of land therefore does not ask foreign citizens for a different identifier from their home country because they have no way of validating it.

An approach could be to include some process by which a foreign individual is assigned a new identifier through compulsory registration (e.g. at the point of incorporating a company or buying real estate), or is required to acquire an existing identifier – for example, in the case of Canada, by registering with the tax authority for a Social Insurance Number. Another approach is the creation of a new internally generated identifier, as proposed in the US.[58] Internally generated identifiers would lend themselves better to sharing internationally, though they would have less inherent meaningful value without accompanying information.

Implementers should also be mindful that existing unique identifiers may not always be permanent for individuals or companies, for example, in the event that two companies merge.

2. Requiring the use of **descriptive data fields** that allow the probabilistic identification of human and non-human entities involved in BOT processes. Examples of descriptive data fields include: names, date of birth, nationalities, and contact details for natural persons; and name, jurisdiction of incorporation, entity type, addresses, and phone numbers for registered entities. Descriptive data fields are useful in the absence of identifiers and when datasets from two sources need to be linked but have no shared identifiers. In both cases, entities can be reconciled using probabilistic approaches.[59] Wherever possible, descriptive data fields should be seen as an addition to, not a replacement for, unique identifiers in data, and should conform to standardised formats.

Whilst these approaches relate to data collection and storage, it should be noted that for the purpose of limiting the risks that may arise from the publication of information, publishers are often unable to publish unique identity documents (IDs) for natural persons that constitute sensitive data, such as personal ID, tax identification, or social security numbers, which increase the risk of identity theft when published. Therefore, jurisdictions often implement a system of layered access, where a smaller subset of the data is made available to the public.[60] In order to publish the minimum but sufficient number of details for the public to use the data, descriptive data fields are often made available, meaning both complementary approaches are required.

Establishing what needs to be disclosed based on a particular declaration regime can be complex from a technical point of view, and it can also be limited by privacy and data protection legislation (see, for example, Box 7).[61]

Box 7: Collecting information on gender and sex as part of beneficial ownership disclosures

Gender and sex data constitutes sensitive personal data in some data protection regimes, meaning there is a higher legal threshold for processing it. In line with data protection principles, governments should not collect more information than is necessary to meet a specific purpose. The purpose of collecting personal data as part of BO disclosures is to unambiguously identify individuals. In most cases, data about gender and sex is not necessarily required to do this.

However, OO research found that collecting this data as part of BO disclosures may be particularly relevant where women have reduced access to official IDs.[62] If BO disclosure regimes rely on collecting copies of IDs to unambiguously identify beneficial owners, for example, this may present a barrier to women becoming beneficial owners, especially if this information must be disclosed at the point of incorporating a company.

Where an individual does not have access to an official, individual ID, the collection of sex data can support identification of the actual owner of a company. This can also be relevant in contexts where gender-neutral names are common. Implementers should define a purpose and establish a legal basis for the collection of all personal data in law.

Describing the relationships between parties

BO data must define the fields and structures needed to describe relationships between natural persons, legal entities, and legal arrangements involved in BOT declarations. At a minimum, **the data must establish a link between a beneficial owner and the declaring entity** associated with a particular disclosure. The details of the data disclosed about a relationship between parties will then depend on the declaration regime.

There are two main areas where data must be defined:

1. The **level of detail required on the chain of ownership** between the declaring company and the beneficial owner. This can range from disclosing only the beneficial owner to disclosing the entire ownership chain as structured data.[f]
2. The **level of detail required on interests through which ownership or control is exerted**. This includes describing the types of interests held (e.g. voting rights through shares or ownership through a nominee); the level of interests (e.g. in the case of shareholdings); details on specific types of ownership and ways individuals can derive economic benefits; control required by a disclosure regime (e.g. structured disclosure of interests held in state-owned enterprises); and when the ownership and control relationship begins and ends. Information about the details should also be structured to be useful and usable, and ownership and control should not be structured at a high level of abstraction which obfuscates these details (see Box 8).

Box 8: Example of a data structure publishing insufficient detail on ownership and control relationships

In the following hypothetical example, a fictional jurisdiction has divided beneficial ownership into categories. For each beneficial owner declaration, the reporting person is required to select a single category of ownership from the list.

Category	Description
A	Owning directly or indirectly at least 25% of the voting rights, voting shares, or capital of the reporting entity.
B	Exercising control over the reporting entity, alone or together with others, through any contract, understanding, relationship, or arrangement.
C	Having the ability to elect a majority of the board of directors of the entity.
D	Having the ability to exert a dominant influence over the management or policies of the entity.
E	Having the ability to influence the majority of the members of the board of directors of the entity de jure or de facto.
F	Exercising stewardship over the assets of the entity.
G	Owning or controlling the reporting entity through nominee shareholders or nominee directors.
H	Owning or controlling the entity through other means.

When, for example, BO data about John Smith exercising beneficial ownership over an entity through a nominee shareholder is exported from the system as structured data, it appears in the following way:

```
{
[…]
“beneficialOwnerFullName”: “John Smith”;
“beneficialOwnerCategory”: “G”
[…]
}
```

[f] OO's sufficient detail principle states that sufficient information should be collected to understand full ownership chains. Visibility of full ownership chains might rely on collating information from declarations of partial ownership chains made by different, related entities. In order to understand and determine what level of detail is required, OO has published a BO disclosure workbook: Kadie Armstrong, “Beneficial ownership disclosure workbook”, Open Ownership, June 2021, https://www.openownership.org/en/publications/beneficial-ownership-disclosure-workbook.

Whilst at face value this may appear to be the publication of structured data, this is an example where the level of abstraction is too high. A number of important details to understand ownership and control are obscured by each category. This information needs to be cached out and published as structured data. For example, for category G, the form does not capture information on who the nominee is. Information should be collected about the nominee as an additional node through which the beneficial owner exercises control. Sufficient details should be published about the beneficial owner, the nominee, and the ownership control relationship between them. In addition, the categories of ownership and control are not mutually exclusive. An individual may fall into multiple categories, or use a nominee arrangement (category G) to exert a dominant influence over the management or policies of the entity (category D).

Ensuring that beneficial ownership disclosures are auditable

For BO data to be useful, it should be functionally auditable. Auditability can be thought of on three levels:[63]

1. Storing a **full historical record** of all BO information,[64] including a ledger of all changes, reduces the risk of manipulation and increases the integrity of the system overall. This means collecting, storing, and publishing time-related and administrative information. Examples of such data include: the date on which a declaration was submitted; the dates that relationships begin and end; and the publication of annual statements confirming existing data. It is important to consider temporal data early on. The need for it may have to be inferred from laws and regulations.
2. Providing **transparency about when and why data is missing, unknown, or redacted** so that BO data can be understood unambiguously. This includes, for example, information on whether and where an application has been granted for the individual's data to be protected from public disclosure as part of a protection regime;[65] information on whether a public company is eligible for an exemption to disclosure because its exchange meets certain transparency and disclosure requirements;[66] or situations in which no beneficial ownership is reported, either because no individual meets the definition or because no beneficial ownership could be found.[67]
3. **Reducing practical barriers to access and analyse BO data** through decisions over data publication formats, access rules, and standardisation.

Designing systems to collect, store, and share beneficial ownership data

Digital systems and administrative processes need to fit together smoothly to enable BO information to be collected, stored, maintained, exchanged, and published. Implementers should make sure when designing systems and business processes that these underpin and support the aims of reforms on a technical level. Practically, implementers should consider:

- **Making sure BOT policy needs are central to decisions over system architectures and database design**, and that both take into account the particular characteristics of BO data as outlined above. It is essential that implementers ensure at this stage that the design of the database can meet the querying and data access expected at publication stage.[68] For example, will the database design allow users to query all companies owned or controlled by a natural person? Inadequate design at this stage will restrict potential use cases further down the line.
- Taking a **digital-first** approach to systems design and taking advantage of digital techniques, for instance, to improve data quality and reliability by validating information at the point of making a declaration.
- **Ensuring that approaches to data collection, storage, and sharing conform to the intent of BOT regulations**. Creating structured data requires resolving the ambiguity that can exist in untested regulation and guidance. This is best done at the design (e.g. forms) and specification stage, and through collaboration with the appropriate agencies.[69]
- **Ensuring that any necessary business process information is generated, collected, and retained**. Collecting, storing, and analysing such information can provide valuable insight into the integrity of a BOT disclosure system. Examples of information that implementers may find valuable to collect include: details of authorised intermediaries involved in submission of data, such as company formation agents; key dates associated with the submission, verification, and

publication process; internal identifiers for processes and individuals; and internal access logs where this may be required by internal policies.[g]

Realising potential and resolving uncertainty at the publication stage

Making BO data available to the public as open data expands the group of potential data users, which may contribute to various BOT policy aims by making the data available to them in a way that removes potential barriers to use and access by resolving uncertainty.[70] In order to make the data accessible, comprehensible, and usable for others, implementers should consider:

- **using open standards** for publishing and structuring BO data; and
- ensuring **auditability** of published data by providing access to the data in different formats, including historical data, with an appropriate licence, accompanying documentation, and publication policy.

Open standards

Open standards are standards that are freely available for adoption, implementation, and updates. Open standards can apply to how data should be produced as well as the format in which the data is published. A machine-readable, open, and industry-standard format is recommended. Comma-separated values (CSV) is a good choice for tabular data, although most BO data is hierarchical in nature. If there is nesting to the data structure, CSV can be challenging to represent the information in a way that does not lead people to mistakenly connect information, or in a way that is larger than needed due to duplication. JavaScript Object Notation (JSON) is an open standard file format that has established itself as a good choice for hierarchical data. It has become the de facto standard for communicating over the internet. However, it is harder to work with for analysts. Therefore, ideally, users should be provided with a choice of formats.

Open standards can also apply to the way that BO data is structured as a whole. BODS allows users to publish information about beneficial ownership in a particular template (see Box 9). Using an open standard provides a high degree of interoperability with other datasets and allows users to integrate data with existing systems, tools, and workflows.[71]

Finally, open standards can apply to individual fields in the data and can also apply to choices of how particular data fields are formatted. For example, dates should be formatted according to an agreed and documented standard, such as ISO 8601.[72] Company identifiers can also be disambiguated using a standard. There are several ways to do this. For example, org-id.guide prefixes local company numbers with a registry identifier.[73] LEI is a global identifier provided in addition to a local company number.[74]

Box 9: The Beneficial Ownership Data Standard[75]

BODS provides a structured data format, along with guidance for collecting, sharing, and using data on beneficial ownership. It is developed by OO in partnership with Open Data Services.[76]

The data schema describes how data about the beneficial owners of a legal entity can be organised and shared. The schema is defined in JSON-structured data format.[77] Using this format facilitates computerised access and analysis of the data, whilst also providing human-readable data. The schema can also be used to inform the design of data collection and management systems, whilst the accompanying technical guidance provides support for publishers and users of the data.

BODS is well defined, yet flexible. National registers of beneficial ownership can use it, as can people researching or describing corporate structures. Data from different jurisdictions published in BODS can be readily joined together to visualise transnational ownership structures.

At the time of writing, BODS is in version 0.3 and being implemented by Armenia, Latvia, and Nigeria.[78] The UK's Data Standards Authority has endorsed BODS to be used for the collection, exchange, use, and distribution of BO data by the government.[79] In addition, private sector register software providers have begun to build BODS directly into their products, and they are deploying this in jurisdictions.[80]

g Whilst there has been no comprehensive review on the utility of internal access logs, some domestic legislation, as in the US, may require this: Adam Smith, "William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021", 116th United States Congress, 1 January 2021, 1241, https://www.congress.gov/116/bills/hr6395/BILLS-116hr6395enr.pdf.

Ensuring auditability of published data

Implementers can ensure published data is auditable – ensuring it is easy to access, interpret, and use – by **providing multiple ways to access the data**.[h] Different users need to access the same data in different ways, and different use cases require access to the data in different ways. Data should be made available in formats that cater to both non-technical users as well as technical users and systems at scale. This means that publishers should provide multiple ways into the data. This should include:

- **Per-record search** via a web interface. Search is the most powerful way to access BO data. Since company data is rarely known perfectly in advance, searches should be as flexible and forgiving as possible and accessibility for the average user should be considered. An example of this is the UK's CH search, which allows for partial matching in search queries. Data should be searchable by both company and beneficial owner.[81]
- **Browsing records** via a web interface. Structured data allows the links between people and companies to be surfaced, which are otherwise not immediately apparent. For example, on the OO Register, an interface will show the beneficial owners of a company. When clicking on an individual beneficial owner, it will also show all other legal entities with which they have BO relationships.
- **Bulk format**. Analysis of a whole dataset requires access to all of the records via bulk record download. This can be provided as a regular export from a source system.[i] For example, Ukraine's Unified State Register of Legal Entities, Individual Entrepreneurs and Public Formations has data publicly and freely available as bulk download, via a file made available daily.[82]
- **API access.** An API is an interface that allows two software applications to talk to each other. Any programmatic analysis or machine-to-machine usage of BO data is likely to rely on web APIs, a way of making structured requests for information over the internet. Consideration should be given to how heavy use of APIs will be managed, including through the use of commercial agreements. In a post-implementation review of the UK's BO register, access by API and bulk access were noted as generating the greatest user value.[83]

For each of these ways to access data, registers should also ensure auditability by providing access to a historical record of changes to beneficial ownership in the data.

For data to be used with confidence, it needs to have a **licence**. Using a licence that meets the Open Definition can facilitate the greatest use of data, but decisions around licensing need to be attuned to each local context.[84] Domestic data protection and privacy legislation may need to be considered, and they may have a bearing on licensing decisions. Many countries with disclosure regimes do not make information about their licensing available, undermining confidence in data use. A number of countries publish BO information as open data, including Denmark, Latvia, Nigeria, the UK, and Ukraine.[85]

Finally, data should be accompanied by the necessary **context and documentation** in order to be useful. A publication policy – or data use guide – allows users to contextualise, make sense of, and plan for the reuse of BO data. A publication policy may include details of licensing; links to relevant legislation and regulation; policies on timeliness of data publications; and a log of substantive changes to the data (technical or in response to changes in legislation and regulation) so that users can understand discontinuities in historical data series. A data user guide should provide users with the information needed to access, analyse, and make use of BO data. This may include both formal documentation and less formal guides and tutorials.

---

h The ways to access data outlined here also apply to internal government users, even if the data is not made publicly available in all formats. For example, although a number of countries have done so without documented issues, some governments may not make bulk data available to the public out of privacy concerns. In this case, bulk data should still be available to law enforcement for proactive investigations.

i An alternative approach is to provide users with access to a streaming API that allows the current (or any) state of the dataset to be constructed by users from the stream of changes pushed to the API by the publisher. This is more flexible and powerful, and it may be more suitable for large BO datasets in the long run, but it is also more complex to use and implement.

Conclusion

To maximise the impact of BOT reforms, a disclosure regime should collect, store, and publish BO information as structured data. This briefing highlights the policy benefits and advantages of these technical aspects of implementation. It provides considerations for implementers about what needs to be in place on a policy level in order to implement the systems and technology to support effective data collection, storage, publication, and use.

To summarise, structuring data creates information that is predictable, which makes it easy to work with and improves its functionality. These benefits apply to both technical and non-technical users. By removing some of the frictions associated with unstructured data, structured data decreases the cost of: collecting data for governments; compliance to disclosure requirements by legal entities; maintaining, publishing, and using the data; and realising policy benefits. It also ultimately facilitates greater policy impact of BOT reforms. Higher up-front costs associated with setting up the required systems are expected to be negated by lower costs associated with collection, storage, publication, use, and maintenance in the long run.

Structured BO data has a greater policy impact because it is highly interoperable. The inherent nature of beneficial ownership means it must be combined with other beneficial ownership and non-BO datasets to be of most value. Structured data also facilitates other aspects of disclosure, such as collection and verification. BODS is designed with interoperability at its core.

In order to operationalise structured BO data, implementers should create an enabling legal and policy environment; provide sufficient resources; and employ a user-centred and iterative approach. The data structure must be able to identify people, entities, and other relevant parties by using unique, permanent, and resolvable identifiers and other descriptive data fields. The structure should be able to describe the ownership and control relationships in sufficient, structured detail. Additional temporal data and metadata should also be collected to ensure auditability and interoperability. The technical systems, business processes, and database design should meet the full functionality and access expected from publication and data sharing, and due consideration should be given to this at an early stage. Finally, making the information available in a range of ways, choosing the right licence, and publishing information about licensing along with additional documentation and a data publication policy can maximise data use.

A growing number of jurisdictions are starting to publish structured BO data, including to BODS, which is increasing the contextual interoperability – and the utility and value – of the data. As more data becomes available, this information will be ingested into the OO Register, and OO will continue to document different use cases and impacts.

Endnotes

1 For examples, see: "Case studies", Open Ownership, n.d., https://www.openownership.org/en/publication-categories/case-studies.

2 See: "Beneficial Ownership Data Standard", Open Ownership, n.d., https://www.openownership.org/en/topics/beneficial-ownership-data-standard.

3 "Open Ownership Principles – Structured data", Open Ownership, updated July 2021, https://www.openownership.org/en/principles/structured-data.

4 The FAIR Principles offer a framework for data management and stewardship for machine-actionable data. See: "FAIR Principles", GO FAIR, n.d., https://www.go-fair.org/fair-principles.

5 "Beneficial Ownership Data Standard (v0.3)", Open Ownership.

6 "The companies we keep: What the UK's open data register actually tells us about company ownership", Global Witness, n.d., https://www.globalwitness.org/en/campaigns/corruption-and-money-laundering/anonymous-company-owners/companies-we-keep.

7 "Learning the lessons from the UK's public beneficial ownership register", Global Witness and Open Ownership, November 2017, 8, https://www.openownership.org/en/publications/learning-the-lessons-from-the-uks-public-beneficial-ownership-register.

8 "Learning the lessons from the UK's public beneficial ownership register", Global Witness and Open Ownership, 8.

9 Ben Cowdock, "Tackling abuse of the Scottish Limited Partnerships needs a UK-wide money laundering reform", Transparency International UK, 3 May 2018, https://www.transparency.org.uk/tackling-abuse-scottish-limited-partnerships-needs-uk-wide-money-laundering-reform.

10 See, for example: David Leask and Richard Smith, "Scots shell companies used to launder £4 billion out of Russia", The Herald, 27 March 2017, https://www.heraldscotland.com/news/15183346.scots-shell-companies-used-to-launder-4-billion-out-of-russia/; Jaccy Gascoyne, "The Use of Scottish Limited Partnerships in Money laundering Schemes", MLROs.com, n.d., https://mlros.com/the-use-of-scottish-limited-partnerships-in-money-laundering-schemes/.

11 Chinwe Ekene Ezeigbo, Tymon Kiepe, and Louise Russell-Prywata, "Early impacts of public registers of beneficial ownership: United Kingdom", Open Ownership, 1 April 2021, https://www.openownership.org/en/publications/early-impacts-of-public-registers-of-beneficial-ownership-uk/.

12 For the full case study, see: Alanna Markle, "Early impacts of public beneficial ownership registers: Ukraine", Open Ownership, 1 February 2022, https://www.openownership.org/en/publications/early-impacts-of-public-beneficial-ownership-registers-ukraine.

13 YouControl, "About us", n.d., https://youcontrol.com.ua/en/about_us/#about-system.

14 See: YouControl, "Case Studies", n.d., https://youcontrol.com.ua/en/cases/.

15 "ISO 8601 – Date and time format", International Organization for Standardization, n.d., https://www.iso.org/iso-8601-date-and-time-format.html.

16 For more information on systems and business processes in BOT, see: "Guide to implementing beneficial ownership transparency", Open Ownership, July 2021, 13, https://www.openownership.org/en/publications/guide-to-implementing-beneficial-ownership-transparency/.

17 Juraj Labant and Gabriel Šípoš, "The Register of Beneficial Ownership in Slovakia: Revolutionary transparency, questionable implementation, unsure benefits", Transparency International Slovakia, June 2017, http://transparency.sk/wp-content/uploads/2017/06/Register-of-beneficial-ownership_study2017.pdf.

18 "Impact Assessment study on the list of High Value Datasets to be made available by the Member States under the Open Data Directive", Deloitte, EC Directorate-General for Communications Networks, Content and Technology, 30 December 2020, 128, https://doi.org/10.2759/493091.

19 "Valuing the user benefits of Companies House data – Report 2: Direct Users", BEIS, Research Paper Number 2019/015, September 2019, 24, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/833767/valuing-benefits-companies-house-data-report-2-direct-users.pdf. For more information on measuring the economic benefits of BOT reforms, see: "Measuring the economic impact of beneficial ownership transparency", Lateral Economics, Open Ownership, and Oxford Insights, May 2022, https://www.openownership.org/en/publications/measuring-the-economic-impact-of-beneficial-ownership-transparency-summary-report/.

20 "B.C. Consultation on a Public Beneficial Ownership Registry", Province of British Columbia, Ministry of Finance, January 2020, 12, https://engage.gov.bc.ca/app/uploads/sites/121/2020/01/386142-BCABO-Consultation-Document-For-Release.pdf.

21 Paul May, "The use of beneficial ownership data by a financial service institution", Open Ownership, Unpublished.

22 For more information on quantifying the economic benefits of BOT, see: "Measuring the economic impact of beneficial ownership transparency", Lateral Economics et al.

23 Tymon Kiepe, "Making central beneficial ownership registers public", Open Ownership, 4, https://www.openownership.org/en/publications/making-central-beneficial-ownership-registers-public.

24 "openownership/visualisation-tool: A visualisation library for beneficial ownership structures", GitHub, n.d., https://github.com/openownership/visualisation-tool.

25 Steve Day, "Transforming procurement systems, one prototype at a time", Open Ownership, 18 December 2020, https://www.openownership.org/blogs/tps-prototyping.

26 "Armenia and Latvia become first countries to publish data in line with the Beneficial Ownership Data Standard", Open Ownership, 15 September 2021, https://www.openownership.org/en/blog/armenia-and-latvia-become-first-countries-to-publish-data-in-line-with-the-beneficial-ownership-data-standard.

27 "openownership/visualisation-tool", GitHub.

28 "Sagamar CJSC – Electronic Register", Government of Armenia, 2018, https://www.e-register.am/am/companies/1267708/beneficiary.pdf; "BOR declarations: «ՋԱՆՉԵՋՈՒՄ ՊՊՂՆՋԱՄՄԱԼՈՒՄԵՆԱՅԻՆ ԿՈՄՊԱՆԱՏ» ՓԲԸ – Electronic Register", Government of Armenia, 25 June 2021, https://www.e-register.am/en/companies/1340197/declaration/c51e08a7-6fdb-4ab7-a55c-c74a68a8f54c?diagram=1.

29 For more information, see: Day, "Transforming procurement systems, one prototype at a time". For more information about using BO data in procurement, see: Tymon Kiepe and Eva Okunbor, "Beneficial ownership data in procurement", Open Ownership, 11 March 2021, www.openownership.org/en/publications/beneficial-ownership-data-in-procurement.

30 "org-id.guide – list locator (Alpha)", org-id.guide, n.d., https://org-id.guide/about.

31 See: "Common Reporting Standard (CRS)", OECD, n.d., https://www.oecd.org/tax/automatic-exchange/common-reporting-standard/.

32 For more information on how structured data can help with various approaches to verification, see: Tymon Kiepe, "Verification of beneficial ownership data", Open Ownership, 7 May 2020, https://www.openownership.org/en/publications/verification-of-beneficial-ownership-data.

33 Emile van der Does de Willebois, Emily M. Halter, Robert A. Harrison, Ji Won Park, and J.C. Sharman, The Puppet Masters: How the Corrupt Use Legal Structures to Hide Stolen Assets and What to Do About It (Washington DC: The World Bank, 2011), 52, 121, https://star.worldbank.org/sites/star/files/puppetmastersv1.pdf.

34 Van der Does de Willebois et al., The Puppet Masters: How the Corrupt Use Legal Structures to Hide Stolen Assets and What to Do About It, 9.

35 "Beneficial ownership information", Open Contracting Data Standard, n.d., https://standard.open-contracting.org/latest/en/guidance/map/beneficial_ownership.

36 Tymon Kiepe, "Using beneficial ownership data for national security", Open Ownership, 17 December 2021, 18-19, https://www.openownership.org/en/publications/using-beneficial-ownership-data-for-national-security.

37 Steve Day, "Reconciling beneficial ownership data", Open Ownership, 17 August 2020, https://www.openownership.org/blogs/reconciling-beneficial-ownership-data; Jason Arterburn, Erin D. Dumbacher, and Page O. Stoutland, "Signals in the Noise: Preventing Nuclear Proliferation with Machine Learning & Publicly Available Information", Nuclear Threat Initiative and Center for Advanced Defense Studies, 12 January 2021, https://www.nti.org/analysis/reports/signals-in-the-noise.

38 "Revised Guidance on Entities Owned by Persons Whose Property and Interests in Property Are Blocked", US Department of the Treasury, 13 August 2014, https://home.treasury.gov/system/files/126/licensing_guidance.pdf.

39 "About this project", OpenSanctions, n.d., https://www.opensanctions.org/docs/about.

40 "Identifiers and de-duplication", OpenSanctions, n.d., https://www.opensanctions.org/docs/identifiers.

41 "How we deduplicate companies and people across data sources", OpenSanctions, 11 November 2021, https://www.opensanctions.org/articles/2021-11-11-deduplication/.

42 "Global LEI Index", Global Legal Entity Identifier Foundation, n.d., https://www.gleif.org/en/lei-data/global-lei-index.

43 "Level 2 Data: Who Owns Whom", Global Legal Entity Identifier Foundation, n.d., https://www.gleif.org/en/lei-data/access-and-use-lei-data/level-2-data-who-owns-whom.

44 " Legal Entity Identifier (LEI) Reference Data", OpenSanctions, n.d., https://www.opensanctions.org/datasets/gleif/.

45 "opensanctions/icij-offshoreleaks: Converter for ICIJ Offshore Leaks data into FollowTheMoney format used by OpenSanctions", GitHub, n.d., https://github.com/opensanctions/icij-offshoreleaks.

46 "OpenScreening", Linkurious, n.d., https://resources.linkurious.com/openscreening; "opensanctions/offshore-graph: Loading OpenSanctions into Neo4J and Linkurious", GitHub, n.d., https://github.com/opensanctions/offshore-graph.

47 At the time of writing, OO is working with the OpenSanctions team to include the BO data from the OO Register in this project.

48 Open Ownership, "Open Ownership technology showcase #4", YouTube video, 1:03:14, posted 18 May 2022, https://www.youtube.com/watch?v=ipKkRX8vpZg.

49 Gregor Semieniuk et al., "Stranded fossil-fuel assets translate to major losses for investors in advanced economies", Nature Climate Change 12, (26 May 2022): 532-538, https://doi.org/10.1038/s41558-022-01356-y.

50 For more detail on creating trusted BO data through verification, see: Kiepe, "Verification of beneficial ownership data".

51 Thom Townsend, "Effective consultation processes for beneficial ownership transparency reform", Open Ownership, 20 June 2020, 12, https://www.openownership.org/en/publications/effective-consultation-processes-for-beneficial-ownership-transparency-reform.

52 "Open Ownership Principles", Open Ownership.

53 For more information, see: Ramandeep Kaur Chhina and Tymon Kiepe, "Designing sanctions and their enforcement for beneficial ownership disclosure", Open Ownership, 28 April 2022, https://www.openownership.org/en/publications/designing-sanctions-and-their-enforcement-for-beneficial-ownership-disclosure.

54 For more details on implementing layered access and protection regimes, see: Kiepe, "Making central beneficial ownership registers public", 17-19.

55 See: "How to make a business case for open data", Open Data Institute, 20 October 2013, https://theodi.org/article/how-to-make-a-business-case-for-open-data; "Research: The economic value of open versus paid data", Open Data Institute, 20 April 2016, https://theodi.org/article/research-the-economic-value-of-open-versus-paid-data; Kiepe, "Making central beneficial ownership registers public", 4. For methodologies on measuring the economic benefits of BOT reforms, see: "Measuring the economic impact of beneficial ownership transparency", Lateral Economics et al.

56 "Why do open organisational identifiers matter?", Open Data Services, Medium, 13 April 2022, https://medium.com/opendatacoop/why-do-open-organisational-identifiers-matter-46af05ab30a; Deborah Yates, "Sustainable stewardship of open identifiers", Open Data Institute, 7 September 2021, https://theodi.org/article/sustainable-stewardship-of-open-identifiers.

57 "Real world identifiers", Open Ownership, n.d., https://standard.openownership.org/en/latest/schema/guidance/identifiers.html.

58 "Fact Sheet: Beneficial Ownership Information Reporting Notice of Proposed Rulemaking (NPRM)", Financial Crimes Enforcement Network, US Treasury, 7 December 2021, https://www.fincen.gov/news/news-releases/fact-sheet-beneficial-ownership-information-reporting-notice-proposed-rulemaking.

59 See, for example, the UK Ministry of Justice's data repository using probabilistic record linkage and deduplication at scale: "moj-analytical-services/splink: Fast, accurate and scalable probabilistic data linkage using your choice of SQL backend", GitHub, n.d., https://github.com/moj-analytical-services/splink. This was used by the UK Competition and Market Authority to link records and assign unique identifiers to UK BO data; see: Open Ownership, "Open Ownership technology showcase #4".

60 For more information on layered access, see: Kiepe, "Making central beneficial ownership registers public", 17.

61 See: "Open Ownership Principles – Sufficient detail", Open Ownership, updated July 2021, https://www.openownership.org/en/principles/sufficient-detail; "Beneficial ownership declaration forms: Guide for regulators and designers", Open Ownership, 17 March 2021, https://www.openownership.org/en/publications/beneficial-ownership-declaration-forms-guide-for-regulators-and-designers.

62 Lubumbe Van de Velde, "Gender and beneficial ownership transparency", Open Ownership, 28 April 2022, https://www.openownership.org/en/publications/gender-and-beneficial-ownership-transparency.

63 For more information, see: Kadie Armstrong, "Building an auditable record of beneficial ownership", Open Ownership, forthcoming.

64 See: "Open Ownership Principles – Up to date and auditable", Open Ownership, updated July 2021, https://www.openownership.org/en/principles/up-to-date-and-auditable.

65 See: Kiepe, "Making central beneficial ownership registers public", 19.

66 Kadie Armstrong and Jack Lord, "Beneficial ownership transparency and listed companies", Open Ownership, 17 September 2020, https://www.openownership.org/en/publications/beneficial-ownership-transparency-and-listed-companies; "Open Ownership Principles – Comprehensive coverage", Open Ownership, updated July 2021, https://www.openownership.org/en/principles/comprehensive-coverage.

67 For more information, see: "Beneficial ownership declaration forms", Open Ownership.

68 For more information and detailed technical guidance, see: "Relational database design considerations for beneficial ownership information", Extractive Industries Transparency Initiative and Open Ownership.

69 For technical guidance on disclosure scope and reviewing forms, and more detail on resolving these issues, see: "Beneficial ownership declaration forms", Open Ownership.

70 Kiepe, "Making central beneficial ownership registers public".

71 "Beneficial Ownership Data Standard (v0.3)", Open Ownership.

72 "ISO 8601 – Date and time format", International Organization for Standardization.

73 "org-id.guide - list locator (Alpha)", org-id.guide.

74 "Global LEI Index", Global Legal Entity Identifier Foundation.

75 "Beneficial Ownership Data Standard (v0.3)", Open Ownership.

76 "We help people publish and use open data", Open Data Services, n.d., https://opendataservices.coop.

77 "Beneficial Ownership Data Standard – Schema browser", Open Ownership, n.d., https://standard.openownership.org/en/0.3.0/schema/schema-browser.html.

78 Stephen Abbott Pugh, "Launch of version 0.3 of the Beneficial Ownership Data Standard", Open Ownership, 20 June 2022, https://www.openownership.org/en/news/launch-of-version-03-of-the-beneficial-ownership-data-standard.

79 "Collect, use and exchange beneficial ownership information", UK Cabinet Office, updated 16 March 2022, https://www.gov.uk/government/publications/open-standards-for-government/collect-use-and-exchange-beneficial-ownership-information.

80 See, for example: "Beneficial Ownership Register", Foster Moore, n.d., https://www.fostermoore.com/verne/beneficial-ownership-register.

81 See: "Open Ownership Principles – Public access", Open Ownership, updated July 2021, https://www.openownership.org/en/principles/public-access.

82 "Ukraine Consolidated State Registry (Edinyy Derzhavnyj Reestr [EDR])", Open Ownership, n.d., https://register.openownership.org/data_sources/ua-edr-register. Since the Russian invasion of Ukraine in February 2022, the Ukrainian BO register has been offline.

83 "Valuing the user benefits of Companies House data – Report 2: Direct Users", BEIS, 24. For more information, see: "Measuring the economic impact of beneficial ownership transparency", Lateral Economics et al.

84 "The Open Definition", Open Knowledge Foundation, n.d., https://opendefinition.org. See: Kiepe, "Making central beneficial ownership registers public".

85 See indicator A.COMPANY.BOT.e.e3.LICENSE in "GDB full dataset", Global Data Barometer, 2021, https://storage.googleapis.com/gdb-files/modules/gdb-2021-company-information-module-data.csv. For example, "Vilkår for brug af danske offentlige data", Erhvervsstyrelsen, 10 March 2015, https://hjaelp.virk.dk/sites/default/files/vilkaar_for_brug_af_danske_offentlige_data_v4.pdf; "Companies House public task and Crown copyright", UK Companies House, updated 9 August 2019, https://www.gov.uk/government/publications/companies-house-accreditation-to-information-fair-traders-scheme.

Authors

Jack Lord

Tymon Kiepe

with contributions by

Kadie Armstrong

Stephen Abbott Pugh

Reviewers

Christa Bull
Government of British Columbia

Jonathan Baron
Government of British Columbia

Silvana Fumega
Global Data Barometer

Editor

Cara Marks

Design

Convincible Media

Published by

openownership.org

@openownership

c/o Global Impact, 1199 North Fairfax Street, Suite 300, Alexandria, VA 22314, USA

Open Ownership