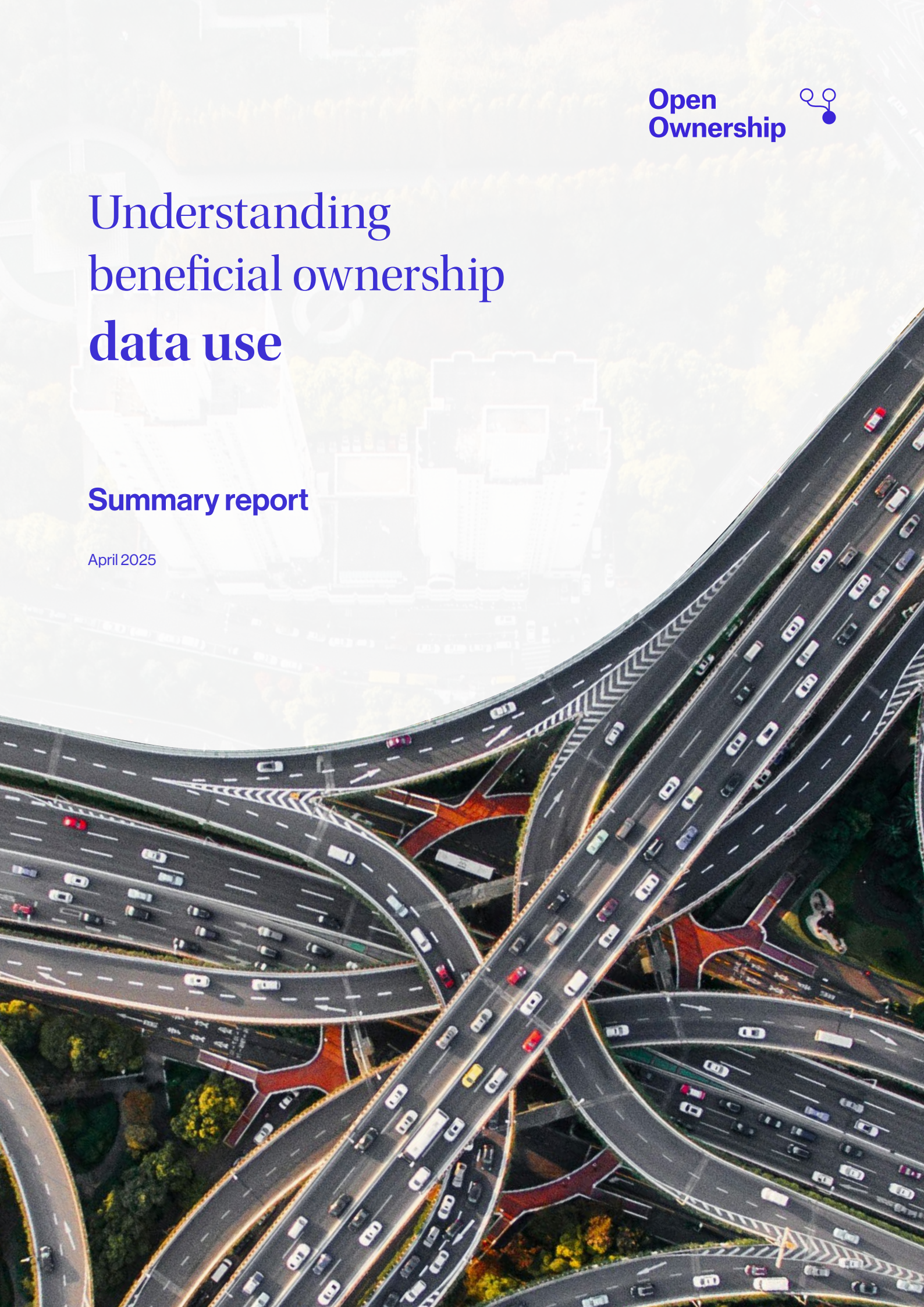
Open Ownership

Understanding beneficial ownership **data use**

Summary report

April 2025



About

The policy objectives of beneficial ownership transparency (BOT) reforms are wide-ranging. They can include ensuring the proper functioning of legal entities and arrangements – collectively known as legal vehicles – and preventing their abuse in enabling corruption, tax evasion, and money laundering. The extent to which BOT reforms help achieve specific policy goals is dependent on beneficial ownership (BO) information being effectively used for these purposes.

Therefore, it is critical for agencies that design and implement BOT reforms to understand the experiences of users of BO information, including their perspectives on what enables and hinders effective BO data use. Measuring the impact of BOT reforms is an emerging area of research, and understanding user experiences is essential to advancing it.

To support progress in this area and further explore how reforms can be best designed for impact, Open Ownership leveraged its global network and conducted primary qualitative research on BO data use. This document is a summary of this research. The full report, including the methodology, is available on Open Ownership's website.

Methodology

The research team interviewed over thirty individuals from a range of sectors and regions to analyse how they used BO information from central register. Five main questions were explored:

1. What are the different types of BO data users, and what are their main differences and commonalities?
2. What does the information enable these users to accomplish?
3. What factors enable or hinder their effective use of this information?
4. Can users be grouped according to how they use the information, and can this inform the design of access provisions?
5. What are the implications of users' experiences for designing effective BOT reforms?

For more details, see the methodology section of the full report, available on Open Ownership's website.

The research provides a conceptual framework to understand BO data use and draws lessons for effective and user-centred implementation of BOT reforms. It also highlights avenues for further research to help assess the effectiveness of BO registers in the future. This document provides a summary of main findings and recommendations.

Research findings: Towards a framework to better understand the use of beneficial ownership information

Users of BO information have various purposes, ranging from identifying and managing risk as a business, to detecting conflicts of interest in public procurement. Each user tries to answer questions for a specific purpose, including to, for example:

- identify a specific individual suspected of illegally owning and enjoying the benefits of assets;
- identify domestic tax residents who may be misusing legal vehicles to evade taxes;
- identify links between politically exposed persons (PEPs) and companies in strategic sectors;
- provide due diligence on potential suppliers, corporate clients, or contract bidders;
- monitor trends in ownership concentration to help understand and regulate competition;
- identify indicators of red flags for money laundering and corruption risks.

For these purposes, users seek to answer specific questions. The characteristics of users' questions and the information they start with determine how BO information is used (that is, the *use type*) and what needs it creates.

The first two sections of the report explore commonalities and differences between users' experiences. The last section distils implications for policy makers and agencies in charge of designing and administering BO registers.

Common user experiences and needs

The research identified a large number of commonalities among user needs associated with various use types – often expressed as challenges. Currently, many users cannot access and process information in ways that allow them to answer their questions. Data-service providers are critical intermediaries that address many of these needs and challenges. However, their services can be unaffordable for some users.

Whether they work for law enforcement agencies, financial intelligence units (FIUs), tax authorities, procurement agencies, anti-corruption bodies, financial institutions, the media, academic institutions, or non-profit organisations, users generally require:

- **Effective access to usable BO information**. Accessing the right information in due time often remains a challenge for users, as illustrated by this research participant from a tax authority:

> "The quicker you can get evidence, the better. Having direct access to BO information saves a lot of time. When you don't have direct access, as tax administrator, you would need to write to an individual or company or to the registrar. … If they only provide partial information, they can also ask for more time. … Only then comes enforcement. This process can easily take over three weeks."[1]

Challenges to accessing information are especially common for foreign legal vehicles and individuals. "This is where the trail goes dead", explained a research participant from an international civil society organisation.

- **Ways to easily retrieve information by being able to flexibly process the data**. Limited options to process BO data and limited search functionalities on BO registers make it difficult, and in some cases impossible, for users to analyse the information in specific ways.

Limited options to process and search BO information can hinder data use

In their analysis of real estate ownership in France, Transparency International France and the Anti-Corruption Data Collective had to go through five million web pages, taking several weeks and using significant resources to connect BO and real estate information. In their report, they explain that not being able to access data in bulk creates "a significant barrier … in monitoring the implementation of the beneficial ownership rules", and that "access to beneficial ownership information in an open data format – or even better, API access – allows key actors to more effectively use the data".[2]

A participant from a tax authority also echoed this, explaining that the search engines they use internally within their agency enable them to set up rules to support risk assessment and identify taxpayers who fit a certain profile. They explained: "you can ask to see anything that involves change in shareholding or that includes a null declaration of revenue. Having something similar with BO information would allow adding more triggers and make your risk identification process richer".[3]

The research found that a broad variety of criteria is used by research participants to query BO data. This suggested that potentially all attributes (i.e. data fields) pertaining to a subject (i.e. individual, legal vehicle, or asset) may be relevant to users. Which specific attribute was deemed most useful in which use case was highly context- and query-specific, and no particular attribute could be associated with specific use types or user profiles. Being able to search a register by the names of beneficial owners was seen as extremely useful in almost all cases.

- **Ways to understand relationships between subjects within and across information sources**. Users require sufficient information to easily determine whether records about individuals, legal vehicles, and assets refer to the same or different subjects (i.e., *entity resolution*), and to identify relationships between subjects within and across datasets (e.g. BO registers, asset registers, PEPs and sanctions lists, etc.). The lack of common identifiers across datasets – for example, a tax identification, company incorporation, or social security number – creates high-resource costs for users, as illustrated by this research participant:

> "A lot of the information is still stored by PDFs and data scientists have to spend a lot of time doing things manually. This is especially resource intensive when going through big datasets that have millions of records. Entity matching between property datasets and company owners is a good example of that."[4]

- **A minimum level of accuracy to be able to draw conclusions with confidence**. Users require up-to-date information and to be able to understand change over time, as explained by this research participant from the financial sector:

> "Historical changes are very important. For example, if you notice that a trust is created all of the sudden, after a person was sanctioned, it can raise a red flag."[5]

Many users compare information sources to check for discrepancies and increase confidence in their conclusions, in a similar way that registrars verify information. Inaccuracies themselves may be a valuable source of information. For this, they need to be able to easily identify inaccurate or missing data which can trigger further investigation. This journalist explains:

> "Data … may not always be 100% correct but, when it is not, it still allows you to ask critical questions … It is a very important tool to fact check something or start an investigation."[6]

However, this is harder to do if accidental errors in registers' data are ubiquitous.

Different use types and needs

Although there are overwhelming similarities in user needs, some differences were also identified. The framework developed through this research identified specific needs based on various combinations of the following core elements:

- whether the question the user is asking is qualitative or quantitative in **nature**;
- the **scale** of data processing required to answer the question, i.e. the number of subjects (e.g. individuals or legal vehicles) users are interested in;
- the **scope** of data processing required to answer the question, i.e. the number and variety of connections or relationships between different pieces of information required for users to reach their conclusion;
- the **frequency** of data processing required to answer the question.

Nature of users' questions

In **quantitative questions**, the interest of users is in *quantifying* something rather than identifying specific people or legal entities. For example, the Competition and Markets Authority (CMA) of the United Kingdom (UK) uses data from the UK's BO register to analyse the degree of concentration of ownership in specific sectors once common ownership and control are taken into account.[7] Similarly, the non-governmental organisation (NGO) Reporters Without Borders and the French academic institution, the Laboratory for Interdisciplinary Evaluation of Public Policies, explored ownership concentration in the French and Spanish media sectors. They published a report stating that over half of each sector was controlled by companies from the financial and insurance sectors, whose complex shareholding structures made it hard to identify the beneficial owners.[8]

In **qualitative questions**, the interest of users is in specific attributes of specific legal entities, individuals, or both. For example, an Armenian journalist also explored ownership concentration in the media, but was interested in identifying specific, powerful individuals with significant influence over the media landscape.[9]

Quantitative questions are likely to be able to be answered using **pseudonymised data** (e.g. unique identifiers in place of companies' and individuals' names), even when the question requires personal information. For example, to assess the proportion of legal entities whose declared beneficial owners are legal minors, users would only need to process dates of birth without needing to identify individuals. However, this is not the case for users with qualitative queries.

Scale of data processing

While for smaller scale queries users can often process data record by record without it affecting their ability to answer their question, processing data manually becomes a critical barrier when conducting large-scale analysis. This makes the use of **APIs and bulk data** particularly important to enable these use types. For example, research participants across law enforcement, tax authorities, and civil society conduct large-scale qualitative analyses to identify and monitor risks. A law enforcement agent from the UK explains:

> "We ingest bulk data from Companies House into our own data holdings. That helps more proactive analysis and creates opportunities for investigation. In terms of data exploitation capability you can query information about a large set of corporations and identify linkages that are not immediately apparent."[10]

Scope of data processing

Queries may be **limited** in scope and only require information from one BO register. For example, a tenant may want to identify the beneficial owners of a company that owns their apartment. In the United States of America, limited liability companies with anonymous owners were associated with harmful impacts, including poor housing conditions.[11] In this case, the user may only identify one or a handful of relationships between their building owner and the individuals behind it.

By contrast, more **extensive** queries require identifying relationships between multiple subjects, either within or across various information sources. For example, in a major case of alleged corruption and stolen assets, Nigerian law enforcement agencies and the FIU used a variety of information sources, including beyond their jurisdiction, to map a network of individuals and legal vehicles and trace USD 1.7 billion of missing funds.[12]

On this spectrum, users with more extensive queries – that is, the majority – have a greater need for support to facilitate the identification of connections between subjects. This makes mechanisms to support **entity resolution** and **identity verification** crucial to enable these use types.

Frequency of data processing

Users may only seek to answer their questions at a specific point in time and therefore require **one-off** data processing. Otherwise, they may aim to provide a recurrent answer and therefore require **ongoing** data processing.

For example, journalists in Mongolia explored corporate ownership of candidates in national elections to provide public oversight and shed light on any potential conflicts of interests during the whole election period.[13] In Nigeria, Directorio Legislativo partnered with BudgIT to build the Joining the Dots platform, which combines BO information, PEPs lists, and mining licensing information to monitor links between politicians and mining licences and automatically raise red flags on a continuous basis.[14] Depending on the scale of the query, this type of BO data use may not be feasible without features such as **automated alerts**, **streaming APIs**, or access to **up-to-date information in bulk**.

Use types, user profiles, and data-use journeys

While the research identified differences between use types, the needs they generate often overlap. In addition, user needs are not tied to specific professions. Questions can evolve along users' journeys and cut across user profiles.

Use types and the needs they create, cut across user profiles

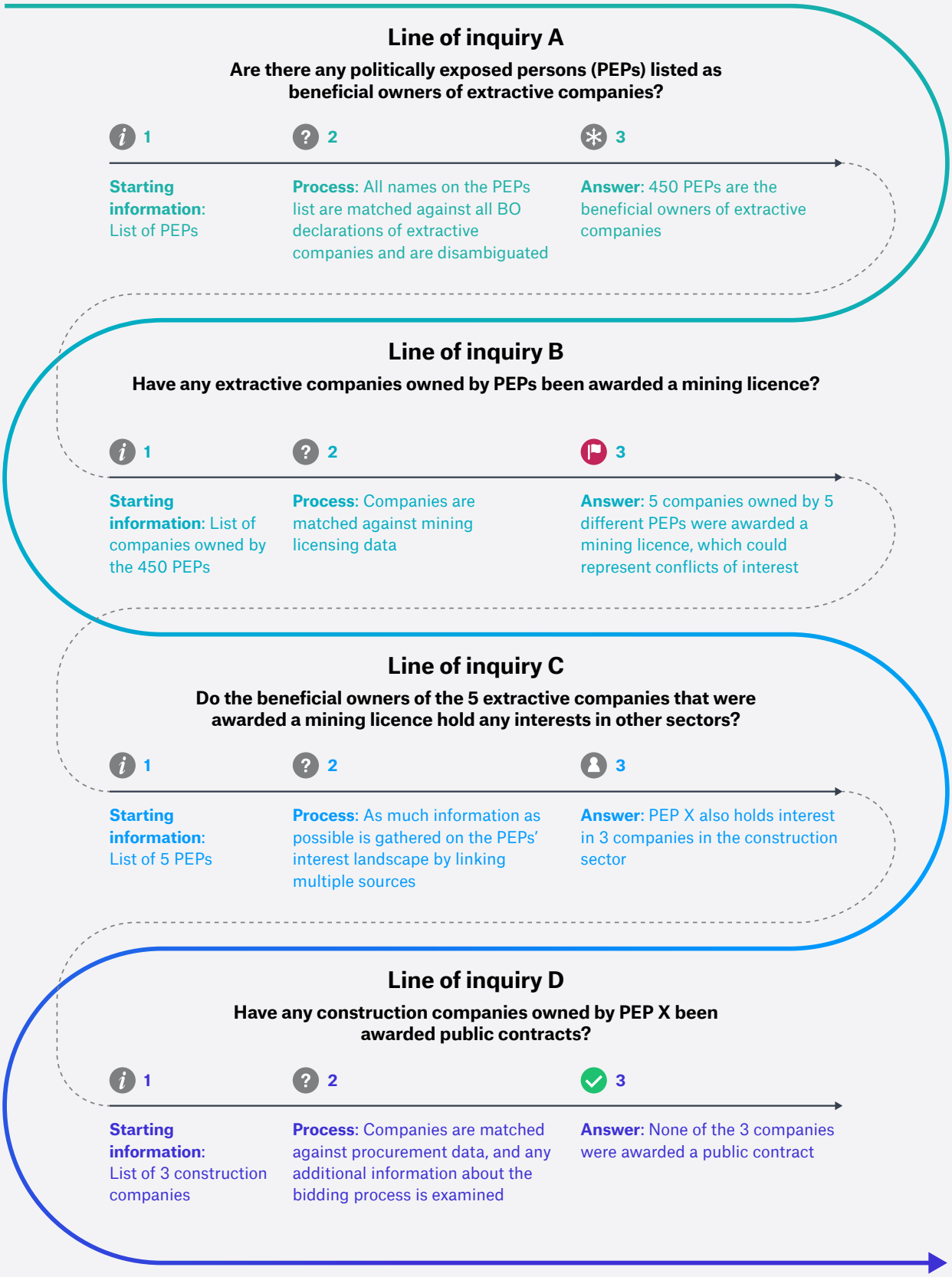
For example, users from a competition regulator, an NGO, an academic institution, a newspaper, and a law enforcement agency may all conduct large-scale processing to identify trends and detect and mitigate risks. Users from the media, customer-due-diligence teams in banks, NGOs, as well as tax and anti-corruption authorities may need to process BO information on a smaller scale to answer qualitative questions focusing on specific individuals or companies.

Data use can be seen as a **multi-stage journey** comprising multiple **lines of inquiry** (see Glossary). Initial queries may also generate additional, unforeseen questions associated with a new set of needs that may be different from those users started with. This makes it difficult to predict how information will be used and therefore impractical to associate specific use types with specific user profiles.

Users questions can lead to new unforeseen ones, creating new sets of needs

In this illustrative example, a user interested in the influence of PEPs in the extractive sector could start their journey asking *Are there any PEPs listed as beneficial owners of extractive companies?* Then, *Have any of these extractive companies beneficially owned by PEPs been awarded a mining license?* And finally, *Do the beneficial owners of these extractive companies hold any interest in other strategic sectors?*

Each question requires different steps and associated needs to answer it.



Research outcomes and avenues for future research

These research findings can inform the debates about **access** to BO information. Access provisions in many jurisdictions have categorised users into government, obliged entities, and others. However, this division does not necessarily reflect how BO information is actually used. In practice, a user will almost never only need to access and process data in a single way. This suggests that categorising user needs solely based on user profiles and single use types is not the best way to make decisions about access regimes. Although identifying user profiles is important to ensure the right individuals have access to BO information, user experiences suggest that, in most cases, access provisions do not respond to user needs. This suggests that the **narrative around access to BO information should be reframed** by not just talking about "access to BO data for *whom*" but also thinking through "access and processing of BO data to enable *what*".

Recognising that the majority of BO data users need a high degree of flexibility in how to access and process data, this research also invites further work to develop recommendations on how to design access regimes that allow users to have more flexibility. This involves looking at the use types that are most likely to advance policy goals and the needs associated with these use types. By tying the processing of information to minimum needs and purpose, access provisions will be more in line with data-protection requirements. This research has also raised a number of additional questions which provide avenues for potential future research. For more details, please refer to the full report.

Implications and recommendations for reforms

These findings validate the Open Ownership Principles for effective beneficial ownership disclosure.[15] In addition, policy makers and agencies implementing BOT reforms can reduce current obstacles, frictions, and resource costs of BO data use by:

- Providing **well-designed APIs** and **up-to-date bulk downloads** to enable large-scale processing, and to allow intermediary users to provide critical services to end users, including governments. To offset the cost of these features, registers can explore charging commercial users for access and use. **Streaming APIs** can be particularly useful to enable large-scale analysis on an ongoing or recurring basis. **Automated alerts** in BO registers can also support other use types that require monitoring changes over time.
- Expanding the **search functionality** of a BO register to enable users to find the information they need. More extensive search functionalities can sometimes remove the need for APIs or bulk data. Basic public search portals with limited ways to process BO information can satisfy some user questions, but are unlikely to enable a wide range of use types that can lead to impact.
- **Structuring data in a well-defined way**. Capturing data that is well structured means it is predictable and easier to use, including understanding **changes over time**.
- **Verifying information at the point of submission** of a BO declaration. As all BO data users require a baseline level of data accuracy to have some level of confidence in reaching their conclusion, setting up measures to support the elimination of errors and inconsistencies at the point of submission is key to supporting BO data use. Collecting information in well-designed digital forms also helps prevent accidental errors. This should be a priority for investment in any verification regime. More advanced verification mechanisms can be considered and should be informed by data-use practices.
- Ensuring effective mechanisms are in place to support **entity resolution** by using or assigning **reliable identifiers** for legal vehicles and individuals. Register-level identifiers are helpful to uniquely identify individuals in a single dataset, and can decrease the need for access to additional attributes and personal data. When users need to combine information across multiple data sources, additional attributes are still required.

- Fostering **interoperability** across BO registers and with other data sources at domestic, regional, and international level. Some of the recommendations above help improve interoperability, including those related to structuring information, verification, and entity resolution. The following should also be considered:
  - Investing in **standardising data**. Using a standardised way to structure data across multiple datasets lowers the resources required to join information sources.
  - Improving **inter-agency coordination** and ensuring the use of **common identifiers** in different information sources, such as the **LEI**, developed by the Global Legal Entity Identifier Foundation (GLEIF).
  - Improving **inter-governmental cooperation** on these topics, including exploring privacy-sensitive solutions to uniquely identify individuals at a transnational level. For example, the European Union's European Digital Identity.
- Designing **access** provisions that allow users who have a role in achieving intended policy goals to effectively **process** BO information.
  - This should include **data-service providers** who currently play a key role in the BO data-use ecosystem and are likely to continue doing so.
  - Allow relevant users to **flexibly use the information** they need with necessary safeguards. Excessive or poorly designed safeguards can negatively affect the impact of BOT reforms. In addition, providing access to a public portal with limited searchability will satisfy most simple queries.
  - Improve **inter-governmental cooperation** to develop legal and technical frameworks for accessing information across borders.
- Including **all relevant legal vehicles** within the scope of disclosure requirements.
- Adapting these findings to **local contexts** and adopting an agile, iterative, and **user-centred approach**.

Many of these aspects are covered in Open Ownership's resources.[16]

Conclusion

Currently, many users cannot access information and process it in ways that allow them to answer their questions. This report's research findings warrant shifting the debate about BO information access from *who* should have access to include *what* this access should look like in order to enable the effective use of BO information. A user-centred approach to designing access regimes can enable effective data use while ensuring the infringement on privacy is proportionate and necessary to achieving specific aims. This report provides a set of recommendations that can guide implementing jurisdictions to take concrete steps towards more user-centred and effective reforms, and ultimately measurable impact.

Endnotes

Please refer to the full report for the comprehensive set of references used for this research.

1 Full report, Appendix 2, interview #003

2 Sara Brimbeuf, Maíra Martini, Florian Hollenbach, and David Szakonyi, *Behind A Wall: Investigating Company And Real Estate Ownership In France* (Transparency International France and Anti-Corruption Data Collective, 2023), https://images.transparencycdn.org/images/2023-Report-Behind-a-Wall-English.pdf

3 Full report, Appendix 2, interview #003

4 Full report, Appendix 2, interview #023

5 Full report, Appendix 2, interview #013.

6 Full report, Appendix 2, interview #036.

7 Competition and Markets Authority (CMA), *The State of UK Competition* (CMA, 2022), https://assets.publishing.service.gov.uk/media/627e6cf6d3bf7f052d33b0ae/State_of_Competition.pdf.

8 Julie Rialet, "Closing the loop: How beneficial ownership information is used and why it matters", Open Ownership, 9 October 2024, https://www.openownership.org/en/blog/closing-the-loop-how-beneficial-ownership-information-is-used-and-why-it-matters/.

9 Mariam Tashchyan, "Ովքե՞ր են հետուստաընկերությունների իրական շահառուները. հայտարարագրվածն ու «դուրս մնացածը»", CivilNet, 27 December 2022

10 Interview #003.

11 Adam Travis, "The Organization of Neglect: Limited Liability Companies and Housing Disinvestment", American Sociological Review 84, no. 1 (2019): 142-170, https://doi.org/10.1177/0003122418821339; Rachel Holliday Smith, "More Money? Pushy Landlord? Your Emergency Rental Assistance Program (ERAP) Questions Answered", *The City*, 7 March 2022, https://www.thecity.nyc/2022/03/07/new-york-city-landlord-emergency-rental-assistance-program-erap-questions-answered/.

12 Egmont Group, "Case 3: Nigeria's former oil minister charged after $20 billion goes missing from petroleum agency —Nigeria FIU", in *Best Egmont Cases: Financial Analysis Cases 2021–2023* (Egmont Group, 2024), 15–20, https://egmontgroup.org/wp-content/uploads/2024/09/EGMONT_2021-2023-BECA-III_FINAL.pdf.

13 Б.Ариунзаяа, "МЭДЛЭГИЙН ГРАФ: Сонгуулийн уралдаанд №1-т сойсон Г.Лувсанжамцын 'шүдийг шинжье'", Datastory, 23 June 2024, http://datastory.mn/article/46; Б.Ариунзаяа, "Эрчим хүчний сэргэлтийн илгээлтийн эзэн М.Энхцэцэгийн авсан тендэр ₮16.3 тэрбум", Datastory, 27 June 2024, http://datastory.mn/article/47.

14 *Joining the Dots*, home page, Directorio Legislativo, n.d., https://peps.directoriolegislativo.org/.

15 Jack Lord and Tymon Kiepe, *Structured and interoperable beneficial ownership data* (Open Ownership, 2022), https://www.openownership.org/en/publications/structured-and-interoperable-beneficial-ownership-data/.

16 For example, Open Ownership has guidance on verification, structured and interoperable data, auditable record of changes over time, the value of reliable identifiers to uniquely identify legal vehicles, and emerging good practices on identifiers to uniquely identify individuals and support agencies implementing BOT reforms with provisions to improve disambiguation in the register. See: Tymon Kiepe, *Verification of beneficial ownership data* (Open Ownership, 2020).