

<p>28th Meeting of the Wiesbaden Group on Business Registers - International Roundtable on Business Survey Frames</p>
<p>The Hague, The Netherlands, 2 – 6 October 2023</p> <p style="text-align: right;"><i>Graham Pilgrim</i> OECD</p> <p style="text-align: right;"><i>Shirly Ang</i> UN Statistics Division</p> <p style="text-align: right;"><i>Session No. 2</i></p> <p style="text-align: center;">Profiling complex Statistical Units (SUs)</p>
<p>OECD-UNSD MULTINATIONAL ENTERPRISE INFORMATION PLATFORM</p>

Abstract

OECD and UNSD have developed a new MNE profiling tool, the Multinational Enterprise Information Platform (MEIP), with the first release in February 2023 and annual updates envisaged for the future. It has relied on public sources and exploited a range of techniques, including but not limited to webscrapping and text mining.

Primarily, the tool provides a framework for determining an MNEs’ subsidiaries (including jurisdiction, hierarchy, and addresses – with links to other corporate databases including OpenCorporates, GLEIF, and PermID). This is complemented by new digital sources, such as websites and social media profiles.

In the longer term, the platform will aim to join information from a global news initiative, to provide a near real-time news information system for the largest companies, with a particular aim to highlight restructuring, investment, financial difficulties, and Mergers and Acquisitions activity.

Introduction

Over the past few decades, Multinational Enterprises (MNEs) have played a leading role in driving transformative changes within the global economy. This role has been fueled by the decreasing transportation and communication costs, advancement in technologies facilitating more complex operations, and easing of trade and investment barriers. Understanding the structure and governance of MNEs is crucially important for the analysis of globalization and global value chains (GVCs).

Several initiatives took place to develop registers of multinationals and their affiliates. The EuroGroups Register (EGR), established by Eurostat in 2008, is a statistical register of Eurostat and national statistical authorities of the European Union (EU) and European Free Trade Association (EFTA) countries. The EGR covers the population and units of multinational enterprise groups operating in the EU and EFTA countries. The EGR is updated on an annual basis and the process includes close coordination between Eurostat and the NSOs, however, the data remains only available to those part of the European Statistical System.

In 2015, the United Nations Statistical Commission (UNSC), recognizing the importance of such registers, mandated the “...creation of a global enterprise group register, building on and taking into account lessons learned from the ongoing EuroGroups Register project”.¹ In parallel, OECD and UNSD developed global registers using very similar methods: the “Analytical Database of Multinational Activity” (ADIMA), developed by OECD in 2018, and the “Global Group Register” (GGR), developed by UNSD in 2021. Given the similarity in the approaches, UNSD and OECD collaboratively undertook the initiative in 2022 to establish a common register for the largest multinationals. Consequently, in February 2023, UNSD and OECD released the first edition of the Multinational Enterprise Information Platform (MEIP) and this data is made available to all on an open-source basis.

Multinational Enterprise Information Platform (MEIP)

The Multinational Enterprise Information Platform (MEIP) consists of two registers²:

- The **Global Register**: Covering the structure of the 500 largest MNEs. It provides information on the MNE’s head, its subsidiaries and affiliates, their location, and their ownership structure when available.
- The **Digital Register**: Covering the online presence of MNEs by providing a list of linked websites along with their relative importance based on visitor count and the significance of website links (page rank).

The **Multinational Enterprise Information Platform (MEIP)** relies on publicly available information, using advanced data processing and handling techniques to develop a framework for determining an MNE’s subsidiaries (including jurisdiction, hierarchy, and addresses – with links to other corporate databases including OpenCorporates, GLEIF and, PermID). This is complemented by new digital sources, such as websites and social media profiles.

¹ See decision 46/107 of E/2015/24 E/CN.3/2015/40 available at https://unstats.un.org/UNSDWebsite/statcom/session_46/documents/statcom-2015-46th-report-E.pdf

² Access the information from either the OECD or UNSD at:
OECD: <https://www.oecd.org/sdd/its/mne-platform>
UNSD: <https://unstats.un.org/unsd/business-stat/mne-platform>

The first edition of MEIP contains information about the largest 500 MNEs based on data for the financial year ending 31st December 2021. The global register covers about 120,000 affiliates in over 200 jurisdictions, while the digital register covers about 110,000 web domains.

The global register includes a user-friendly dashboard to navigate through the database (see Figure 1).

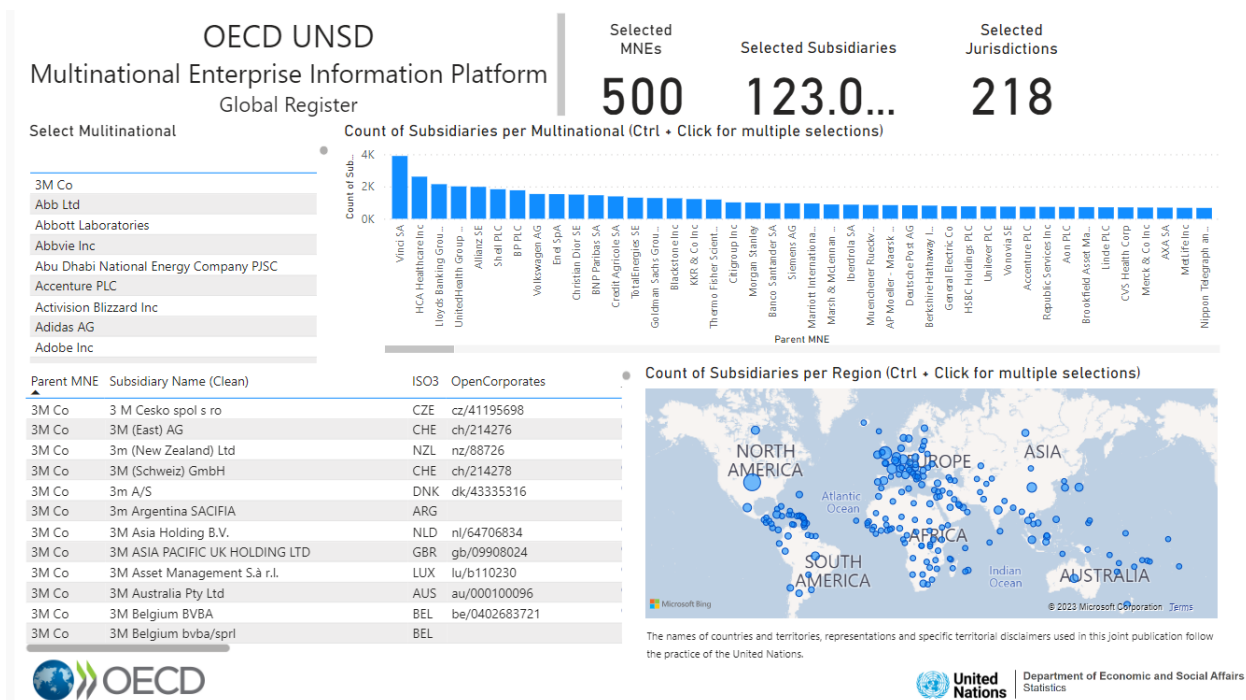


Figure 1. Dashboard of MEIP Global Register (data as of 31st December 2021)

Uses of MEIP

MEIP and its predecessors were developed in an effort to better understand globalisation. Understanding the structure of multinational enterprises helps to understand the dependencies and provides insights into global value chains. Some analytical studies were carried out based on the information contained in the global register, such as the OECD blog “Investigating individual MNEs' responses to Russian sanctions”³ and the OECD study “Investigating the gender diversity of women on boards”⁴.

Methodology and Data Sources

MEIP is built solely from publicly available information, utilising advanced data processing and handling techniques to develop a framework for determining an MNEs' subsidiaries. Annual reports serve as the

³ Available at <https://oecdstatistics.blog/2022/11/14/to-leave-or-not-to-leave-how-are-the-worlds-largest-mnes-responding-to-the-war-in-ukraine/>

⁴ <https://www.oecd.org/gender/data/what-big-data-can-tell-us-about-women-on-boards.htm>

primary data sources for the identification of the subsidiaries of an MNE. They serve as rich repositories of valuable information, including financial data, subsidiary listings, and organizational structures. These reports provide essential details that contribute to a holistic understanding of MNEs' activities and hierarchies.

However, annual reports often do not present a complete list of subsidiaries and affiliates. Therefore, additional sources are used to complement the list of subsidiaries and include additional information, such as the location of units and the type of relationship between units. Data from openly licensed datasets, such as OpenCorporates, GLEIF, PermID, CommonCrawl, PeopleDataLabs, WebDataCommons, PeopleDataLabs, WebDataCommons, CorpWatch, WikiData, Crunchbase, OpenOwnership, Sectigo (crt.sh), and selected country-level open-source information are utilised.

Compiling MEIP is not a trivial exercise. It requires assembling a vast amount of information often in different and unstructured formats. Some of the challenges encountered are presented below.

The initial step in compiling MEIP involves extracting information from the companies' annual reports. However, information about subsidiaries in annual reports is often presented in an unstructured format, making automated data extraction a challenging task. The lack of standardised formatting across different reports makes it difficult to design a one-size-fits-all extraction approach. Tables within PDFs can vary in layout and structure, requiring adaptable data extraction methods. Furthermore, the complexities of language and font variations introduce potential inaccuracies during extraction, requiring manual verification. These challenges culminate in a process that is not only time-consuming but also labor-intensive, requiring significant manual intervention and validation.

Structured filings (XBRL), which make the extraction of financial variables simpler, are now required within the European Union and the United States – however, the taxonomies are not uniform, and jurisdictions that do not require structured filings still exist.

Another intricate part of the process involves linking information from multiple data sources. This requires a significant level of data validation to ensure that the links are verified. The largest proportion of the project's resources are spent on ensuring that all companies are verified.

Lastly, the compilation of MEIP demands a resource-intensive IT infrastructure. A significant amount of information is assembled and verified, requiring large storage capacity and efficient data handling capabilities.

The primary data storage mechanism, a graph database, lies outside the knowledge of the majority of statisticians who often feel more confident in dealing with tabular data. As a result, it became essential to create several customised interfaces in order to visualize these intricate relationships. In hindsight, a focus on downstream visualisation and dashboards would have been more helpful in building an understanding and broader utilisation of the tool.

Future perspective

The project is considering to further expand MEIP's capabilities by incorporating media tracking to provide a live view of what is happening for each covered firm. This addition ensures that significant restructuring events are promptly captured. Additionally, mitigating the complexities of data extraction from annual

report PDFs by exploring advanced technological solutions and collaborative initiatives. Through this, we can streamline the extraction process, reduce labor intensiveness, and enhance the overall accuracy and efficiency of data collection.

Moreover, migrating the currently implemented graph database from in-memory to disk-based will provide a significant benefit, as resource usage will be reduced, and it offers the ability for the data to be hosted on a platform such as the UN Global Platform. This then allows for individuals outside of the UNSD and OECD to potentially join the initiative, compiling profiles on MNEs of interest to them using the tools and principles that have already been developed.

As for the future perspective, the Statistical Commission at its 54th session in March 2023 endorsed the global initiative on unique identifiers for businesses and encouraged countries and relevant organizations to coordinate their activities in this area in an effort to provide solid infrastructure for statistical business registers. The main objectives of the global initiative are to strengthen the transparency of businesses in countries by improving their registration, improve the availability of unique business identifiers in administrative data sources in countries, and promote access to and sharing of administrative data for statistical business registers.

Strengthening the administrative business registration in countries and the establishment of unique identifiers for legal entities is considered an important first step in improving the statistical business registers, especially in countries where the administrative data system for businesses is not well established. Linking the global initiative of national unique identifiers and the mapping to global identifiers will facilitate the establishment and maintenance of MEIP containing the legal structure of multinationals.

Overall, a promising opportunity lies ahead in exploring how MEIP and the global initiative on unique identifiers could effectively contribute to the development of a robust infrastructure for issuing identifiers at both national and international levels.

Acknowledgment

We thank Ilaria Di Matteo, Julian Chow and Annabelle Mourougane for providing their invaluable advice and insightful comments throughout the preparation of this paper.