Economic and Social Council

Economic Commission for Europe

Conference of European Statisticians

Seventy-first plenary session
Geneva, 22–23 June 2023
Item 9 of the provisional agenda
Timeliness, granularity and frequency of official statistics

Organisation of Economic Cooperation and Development and United Nations Statistical Division's Multinational Enterprise Information Platform

Prepared by the Organisation of Economic Cooperation and Development and United Nations Statistical Division

Summary

The Multinational Enterprise Information Platform (MEIP) is a joint initiative of OECD and UNSD to provide an understanding of the structure of Multi-national Enterprises (MNEs) and related material. It comprises two core registers: the Global Register covering the structure of the 500 largest MNEs, and the Digital Register covering the web presence of large MNEs. The document presents the methodology, data sources, successes, challenges, lessons learned and future plans related to the project.

The document is presented to the Conference of European Statisticians' session on "Timeliness, frequency and granularity of official statistics" for discussion.

Please recycle

I. Introduction

1. Multinational Enterprises (MNEs) have been at the forefront of changes in the global economy, as transportation and communication costs have declined, technologies have facilitated more complex operations, and trade and investment barriers have lessened over the last few decades. Understanding the structure and behaviour of MNEs is fundamental to the production of consistent global statistics. This information facilitates analysis of the impact of globalisation on value chains and provides transparency to the public with regards to large enterprises operating in their region, particularly in regard to the international efforts to ensure that these firms pay a fair share of tax wherever they operate and generate profits.

2. The [Multinational Enterprise Information Platform](#) (MEIP) is a joint initiative of OECD and UNSD to provide an understanding of the structure of MNEs and related material. It comprises two core registers:

- The Global Register: Covering the structure of the 500 largest MNEs by providing information on the MNE's head, their subsidiaries, their location and their ownership structure when available. ([Download](#)) ;

- The Digital Register: Covering the web presence of MNEs around the world by providing a list of connected websites and their relative importance determined by the number of visitors and the importance of links to their website (page rank). ([Download](#)).

II. Methodology and data sources

3. The Multinational Enterprise Information Platform is built solely from publicly available information, utilising advanced data processing and handling to develop a unique source of information on MNEs. The dataset includes information to link subsidiaries to key information provided by OpenCorporates, GLEIF and PermID, allowing users to develop further insights.

4. The first release of the Multinational Enterprise Information Platform covers data as of 31st December 2021 and will be updated annually. Explorations to improve the timeliness, frequency and coverage of the information are in progress.

5. The platform relies on a number of openly licenced datasets. Both organisations are grateful for the contributions of:

CommonCrawl	GLEIF	PeopleDataLabs	WebDataCommons
CorpWatch	OpenCorporates	PermID	WikiData
Crunchbase	OpenOwnership	Sectigo (crt.sh)	And country level open source information
GDELT	OpenStreetMap	Tranco	

III. Successes and challenges

6. The project was successful in its initial aim to profile the largest MNEs using solely open source data. One of the additional aims was to provide a data source for additional work on MNEs, and in this regard there have been a number of positive contributions including:

- Investigating individual MNEs responses to Russian sanctions (https://oecdstatistics.blog/2022/11/14/to-leave-or-not-to-leave-how-are-the-worlds-largest-mnes-responding-to-the-war-in-ukraine/)

- Investigating the gender diversity of women on boards (https://www.oecd.org/gender/data/what-big-data-can-tell-us-about-women-on-boards.htm).

7. However, the scale and scope of the data exercise was underestimated. Combining multiple data sources of varying degrees of quality requires significant levels of data

validation and the largest proportion of the projects resources are spent on making sure that all companies are verified in some form.

8. The second largest resource relates to manual data capture, information regarding companies are often embedded inside PDF documents. Extracting tables and other information from these files is a complex and challenging exercise. No uniform method was found which was successful in all cases, and therefore significant manual intervention is also required. This also meant that one of the other subobjectives of the project which related to the collection of financial variables was put on hold for the initial releases of the database. Structured filings (XBRL), which make the extraction of financial variables simpler, are now required within the European Union and United States – however the taxonomies are not uniform, and jurisdictions which do not require structured filings still exist.

IV. Lessons learnt at a strategic level

9. At a strategic level, there were a number of difficulties. Firstly, a number of the data sources lie outside the traditional data considered by statistical compilers. For example, explaining the role of X.509 certificates in helping to assign a website to an individual company required significant research and investigation to build confidence.

10. Secondly, the primary data storage mechanism, a graph database, lies outside the knowledge of the majority of statisticians who often feel more confident with tabular data. As a result, it was necessary to develop a number of custom interfaces to visualise these complex relationships. With hindsight, a focus on downstream visualisation and dashboards would have been more helpful to build understanding and wider use of the tool.

V. Future perspective

11. Looking to the future the project hopes to:

- Add media tracking to provide a live view of what is happening with each of the companies covered– meaning you will not miss any large restructuring events.

- Migrate the currently implemented graph database from in-memory to disk based. This has 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 UNSD and OECD to potentially join the initiative, compiling2 profiles on MNEs of interest to them using the tools and principles which have already been developed.