**28th Meeting of the Wiesbaden Group on Business Registers  
2 - 6 October 2023**

Alexandre Depire (EUROSTAT),   
Simon Rommelspacher (DESTATIS) 

Session 2 – Profiling complex Statistical Units (SUs)

**Introducing Top-tier profiling of Multinational Enterprise Groups in the EuroGroups Register**

**Abstract**

*In the European Statistical System, the EuroGroups Register (EGR) is the statistical business register developed by Eurostat in cooperation with the national statistical institutes (NSIs) of EU Member States (MS) and EFTA countries. It contains the complete structure of the Multinational Enterprise Groups (MNE) groups operating in the EU and EFTA countries. It constitutes the necessary infrastructure for the coordination of consistent statistics on globalization in the EU. To fulfil this role, the EGR records high quality and up-to-date information.*

*With the aim to increase the accuracy and timeliness of the EGR and to reduce the burden on the NSIs, the data quality management of the MNE groups in the EGR is going to be based on a “Two-tier approach”, ensuring high accuracy and timeliness on the largest and most significant MNE groups with sizable impact on European statistics (Top-tier), while investing in automation to treat the rest of the MNE groups with sufficient accuracy and coverage.*

*A Eurostat Task Force including 13 MS was created in 2021 to specify the criteria to select the Top-tier population of the MNEs as well as to define an updated their treatment and profiling process.*

*The Task Force specified a Complexity and Statistical Impact (CSI) index to define the population of the “Top-tier” MNE groups. It is based on economic variables and linked to the structure of the MNE groups. A method based on percentiles is used to classify and rank them. Once the Top-tier list has been drawn up, a dedicated profiling activity is implemented by the NSIs in the shared on-line application provided by Eurostat.*

*The Task Force also specified the Top-tier profiling process for the treatment of the groups structure. The process is under test since April 2023. The profilers of the NSIs shall focus on updating the structure of the Top-tier MNE groups and their legal units, responding to events that are statistically significant and require revisions in the EGR. The Task Force developed functionalities to support the manual treatment of the groups’ structure in EGR. The country where the decision-making center of the MNE group is located leads the profiling process and proposes updates to the partnering countries, that work on the resident parts of the MNE groups in their territory. If the NSIs have resources, they can also delineate the statistical units in the shared on-line application provided by Eurostat, otherwise it will be updated in bulk during the automated data exchanges with the EGR.*

*The new approach is expected to significantly increase the number of Top-tier MNE groups that can be profiled by NSIs, thus increasing the quality and reliability of EGR data and allowing more frequent data releases. It is also aiming to reduce the burden on Eurostat and MS during the whole EGR process.*

*Keywords: EuroGroup Register, Complex and Statistical Impact index (CSI index), Top-tier groups*

# ****Introduction****

Globalisation is significantly affecting the European economies. The Multinational Enterprise (MNE) groups play a significant role to globalisation, therefore they are a key for the quality of business and macroeconomics statistics as well as for the largest asymmetries in the data on globalisation. Their fast-changing patterns of organisational structures and global arrangements put increasingly to test well-established concepts and methods to produce high-quality statistics in a timely manner. In the last decade, the European Statistical System (ESS) has developed several solutions to react to specific domain driven needs and has significantly increased the capacity to understand MNE groups’ structures, their activities and behaviour and to work collaboratively using secure tools and agreed methodologies. *The main objective is to improve the consistency of the statistical output across domains and countries in the Union and be able to continue providing high quality official European statistics to users.*

# ****Eurogroup Register (EGR)****

One of the key elements to achieve this objective is the EuroGroups Register (EGR). The project to establish the EGR started in 2008 and was implemented in cooperation between Eurostat and the National Statistical Offices (NSIs) of the EU Member States and EFTA countries, by pooling together microdata on legal units, ownership relationships, enterprises, links from legal units to the enterprises, and enterprise groups. The EGR contains structural economic information on MNE groups with legal units in at least one European country. The EGR database is meant for statistical use only in accordance with the legislation in force. It is the common European statistical infrastructure, and it is restricted to users of Eurostat, National Statistical Institutes (NSIs), National Central Banks, and the European Central Bank aiming to produce high quality business statistics on globalisation.

In the EGR, input information coming from the NSIs is treated according to priority rules that are used to resolve inconsistencies and duplicated information.

**The figure 1 shows how different views of the same MNE group available at national level are consolidated by the EGR to provide one common global view.**

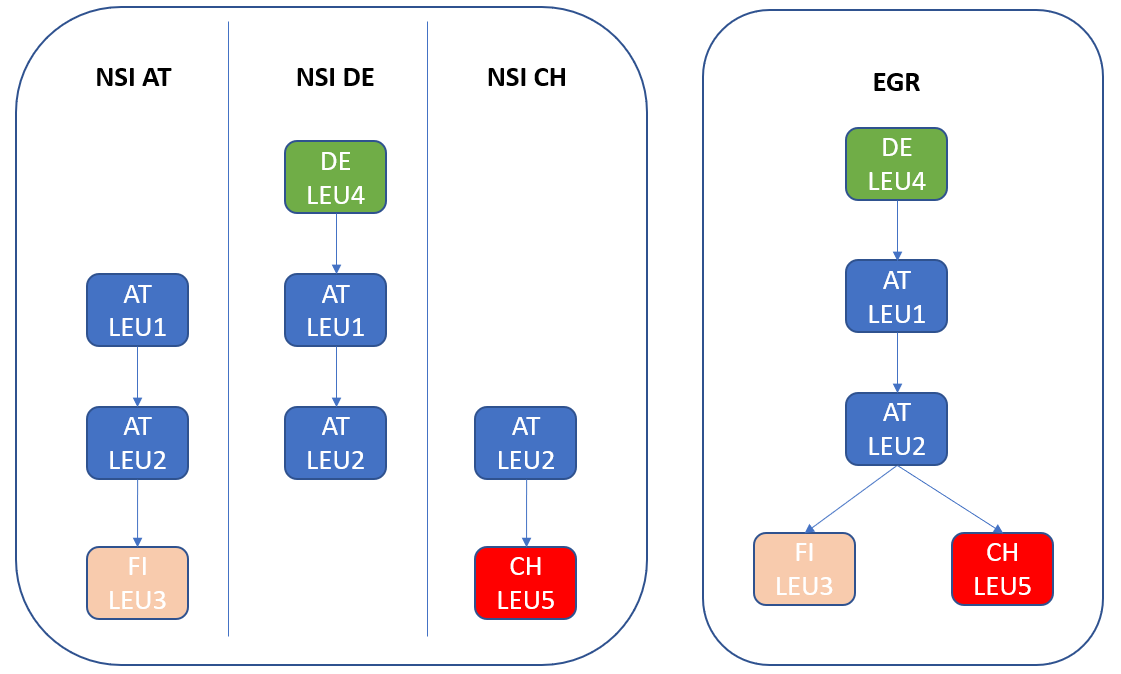
****

Figure 1: The global view thanks to the EGR

# ****A Two-tier approach on MNE groups for high quality data in EGR****

In the last couple of years, the EGR system, process, and output, are being transformed in order to respond to serve the new requirements indicated in the Bratislava conclusions ([1]) of the 2019 conference of Directeurs Généraux des Instituts Nationaux de Statistiques (DGINS)[[1]](#footnote-1), with a view to bring various activities, instruments, and capacities together in a systematic, coordinated, cost-effective joint approach and for:

1. ensuring ‘*high quality and up-to-date statistical business registers [2] (national and European) as a necessary infrastructure for the coordination of statistical activities and exchange of data on MNE groups*’;
2. frequently updating a set of ‘*prioritised lot of MNE groups, having a significant impact on business and economic statistics at national and European level, that are subject to profiling in order to achieve the quality necessary to support the statistical production of economic globalisation statistics*’.
3. monitoring consistently the ‘*fast-changing patterns of activities of large, complex and dynamic multinational enterprise groups*’.
4. becoming an ‘*appropriate tool for high quality measurement of the impact of globalisation on statistical indicators*’.
5. providing the MNE groups information for addressing ‘*asymmetries and micro data sharing necessary to improve data consistency at national and European level*’.

To ensure high quality data in EGR, Eurostat has launched in 2021 the European Profiling programme. The purpose of European Profiling is to analyse the legal, economic, and operational structure of an enterprise group to get the best delineation of the statistical unit 'enterprise'. European profiling is a collaborative activity where the NSIs having an interest to profile the same MNE group define and agree together on the legal, financial, and economic structure of it and delineate its statistical unit(s) enterprise(s). This common understanding of the MNE groups' structure and their activities can ultimately support a consistent data collection and data compilation of globalisation statistics.

To implement the DGINS Bratislava conclusions, the ESS Business Statistics Directors’ Group supported the implementation of a two-tier approach in the EGR which envisages:

* The first tier (the *Top-tier*) consists of a *coordinated* cross-country thorough treatment of the largest and most complex MNE groups, i.e., those with a significant impact on macroeconomic and business statistics. This tier needs *current* updating based on the MNE groups profiling work and the results should be integrated in the EGR. The Top-tier groups should be regularly checked and followed up allowing higher and assured quality and the creation of the timelier frames consisting of the updated data of the largest MNE groups. That will allow macroeconomics and business statistics better and quicker capture the change of globalisation affects.
* The second tier consists of a coordinated cross-country information on the bulk population of the MNE groups operating in the EU Member States and EFTA countries. This tier needs an automated process to produce *annual frames* sufficiently timely to allow the EGR serving as supplementary frame for national surveys. The automated process will be benefit and will improve with the experiences and the learnings in the treatment of the first tier.

This two-tier approach, which is visualised in the Figure 2 below, does not necessarily imply two different IT systems, but a distinction between two ways of maintaining the MNE groups in the EGR and in the national business registers based on their impact on the statistics.

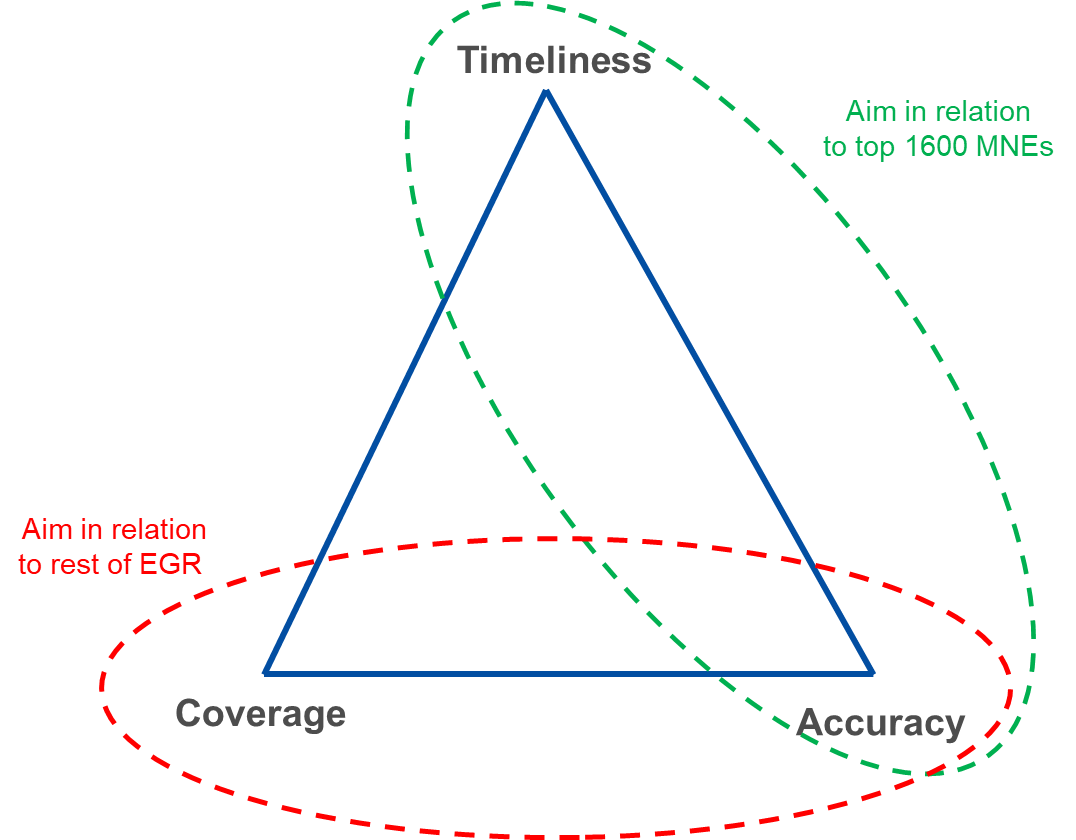


Figure 1: Two-tier approach on MNE groups in the EGR

The dedicated Task Force set up by Eurostat had the mandate to define and implement the Top-tier approach and shape the future EGR, with main objectives: identify the target Top-tier MNE groups population (section 3) and design the new process and technical tools for the treatment of ‘Top-tier-MNE groups’ (section 4). This task was carried out and successfully accomplished in the period from the end of 2021 to mid-2023.

The following sections will focus on the description of the methodology that is used to identify the Top-tier MNE groups, the way the process of profiling them has been adapted and its output integrated into the new EGR system and process [3].

# ****Complex and Statistical Impact (CSI) Index****

To focus the NSIs efforts and resources spent on European Profiling on the really most important groups (‘Top-tier’ groups), a structured and accepted method to identify these Top-tier MNE groups had to be defined, based on ESS agreed criteria. The result was the creation of the Complexity and Statistical Impact index, that is applied to the EGR MNE groups’ population.

## The inception and creation of the Complexity and Statistical Impact (CSI) index

To distinguish the ‘Top-tier MNE groups’ proportion in the EGR, the MNE groups are classified according to different criteria:

1. Size

One (or a combination) of the following variables can be used for this criterion:

* Consolidated EU (and/or global) turnover,
* Total balance sheet,
* Value added,
* Totals for international trade,
* Number of employees (globally/EU)

1. Complexity

One (or a combination) of the following variables can be used for this criterion:

* Number of LEUs (or ENTs) in a MNE group structure,
* Number of different LEU-NACE codes in a MNE group structure,
* Number of layers in the ownership structure (number of mother companies for the lowest LEU in the structure)

1. Problematic cases

One (or a combination) of the following variables can be used for this criterion:

* Complicated and/or frequently changing ownership structure,
* Complicated and/or frequently changing “global production” arrangements (such as Factory-less production, diverse Intellectual Property Products[[2]](#footnote-2) ownership, cross border branches).

The first attempts to translate the above criteria in useable and meaningful results demonstrated that a Complexity and Statistical Impact (CSI) index can be calculated by ranking the EGR population using the following variables:

* Number of legal units,
* Number of NACE codes,
* Number of cross border relations (as number of countries with active units),
* Number of employees,
* Number of employees in largest LEU,
* Number of ownership layers (operationalised as number of parents of lowest level unit).

The target was to cover the Top-tier MNE groups responsible for at least 5% of the total EGR population.

The selection of the Top-tier MNE groups based on this initial version of the CSI index, is constructed on the following principles:

* It allows groups to be classified into 75 classes. The index has 10 classes for each variable, except for EU employment where we divide into 25 classes in order to give more importance to it.
* For each variable, we order and assign 10 points (25 points for EU employment) to all the groups belonging to the last sub-group (9 points for the penultimate and so on (24 in the case of EU employment)); finally, we add up all the points to give a weight to each group.
* The Top-tier groups are the ones having 75 points.

Applying this initial CSI index to the EGR population for the reference year 2020, resulted in a selection of 2,365 MNE groups (equal to around 2% of MNE groups in EGR, which are less than the 5% of the total population however they are the statistically significant ones). By applying it to the MNE groups population for the reference year 2021, the number increased to 2,903 MNE groups. As this number was considered too high by the NSIs for European profiling due to resource constraints, a further analysis and an updated version of the CSI index was necessary.

## New version of the Complexity and Statistical Impact CSI index

The initial version of the CSI index was based on a simple ranking of MNE groups according to each selection criterion. A more detailed analysis showed that a decile[[3]](#footnote-3)-based approach was more appropriate to improve the results of the CSI. Briefly, the approach consists in dividing the target population of MNE groups in EGR into *n* sets each of them with the same probability of being drawn. The approach is based on the quantile of the empirical distribution of each chosen criterion. It allows for the correct management of cases where a distribution is not uniform, which is the case for the MNE groups. Furthermore, the ranking method does not allow the correct management of the ties of the distribution. For some of the selected criteria, MNE groups with the same value (ties) must be classified in the same class to avoid random rejections; With the method based on ranking, this cannot be guaranteed.

The MNE groups selected according to the new CSI index can be assigned to each country for European profiling based on the location of the Global Decision Center (GDC[[4]](#footnote-4)) if it is in EU+EFTA countries.

To find a way to treat the Top-tier MNE groups where the GDC is located outside EU+EFTA, the Task Force and Eurostat have defined a methodology to find the “a responsible” EU+EFTA country to work on these MNE groups, i.e., the country with the highest interest to profile these MNE groups. This assignment however can be modified if agreed accordingly between the countries, depending on the statistical relevance and national user needs.

Some EU and EFTA countries could result having little or no involvement in the treatment of Top-tier MNE groups, due to the absence of GDCs in their territory, and/or the fact that they only host the GDCs of smaller groups. To cover also MNE groups relevant at national level, a method has been defined to provide a complementary number of MNE groups to cover the needs of these countries.

**The result of this work is that the new CSI index allows to select the most important MNE groups, known as “Top-tier” MNE groups in an effective way. These are 1,669 MNE groups, corresponding to 1% on the total MNE groups population in EGR and accounting for more than 50% employment of MNE group in Europe ([4]), i.e., around 24,5 million employees**.

The list of Top-tier MNE groups can slightly change over time, however with the decile-based approach and assuming stability of the largest ones, it should remain sufficiently stable in the short-term excluding cases of restructuring. The list can also be used to monitor the dynamic of the MNE groups’ structure and size, indicate restructuring cases, situation of errors, etc. To reduce the workload related to the update of the Top-tier MNEs groups, Eurostat proposes also to improve the current EGR production calendar, while including the provision of supporting functionalities, like for example comparison features with external data sources.

**The next step of the EGR adaptation was to redefine the process and system for the treatment of the selected Top-tier MNE groups in a cost-effective way during European profiling, while at the same time fulfilling the new requirements.**

# ****Implementing the Top-tier profiling process in EGR****

Until 2022, European profiling was aimed at analysing the legal, economic, and operational structure of an enterprise group to get the best delineation of the statistical unit 'enterprise' [5]. As the MNE-groups are often active in several countries, this task was a collaborative activity, involving all the countries that are hosting on their territory enterprises being part of the same MNE groups. Three different roles were set up: the GDC country as the main responsible for the MNE group, the other countries as partner, and Eurostat as coordinator. In this process the GDC country proposes an update of the MNE group's perimeter and a proposed division into its main economic activities. This information is shared with the partnering countries, which can propose changes to achieve a common statistical delineation of the MNE groups, including the correct statistical unit enterprises in each country. In some cases, the GDC country can directly interview the MNE groups, increasing the quality of the data and, in some cases, reducing the amount of manual research required. Eurostat has developed an IT tool to facilitate the European profiling activity called IPT (Interactive Profiling Tool). Results of European Profiling are available in [European profiling - experiences of National statistical institutes - Statistics Explained (europa.eu)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=European_profiling_-_experiences_of_National_statistical_institutes).

European profiling requires substantial resources to be mobilised by the NSIs and cannot be used to regularly monitor the 1660 Top-tier MNE groups is a cost-effective way. Therefore, the process had to be reviewed and made as intelligent as possible and supported by efficient tools. For this reason, Eurostat with the support of the task force, developed a new and leaner European profiling process for the treatment of the Top-tier MNE groups (named Top-tier profiling process) and adapted the existing tools for it. Differently from the past European profiling, the Top-tier profiling process focuses the collaborative work on the structure of the Top-tier MNE groups selected by the CSI index and leaves the delineation of the statistical unit enterprises at national level, allowing NSIs to deliver them to EGR during the usual EGR annual process. However, the profiling tool will continue to allow those NSIs who want to perform the enterprise delineation to continue to do so, too.

The changes in the European profiling process were supported by an in-depth analysis carried out by the Task Force, that considered whether each step of the original European profiling process could be considered mandatory or optional in the new Top-tier profiling process, with the objective to make it more cost effective and better integrated with the national profiling, at the same time able to support an increased number of Top-tier MNE groups profiled with high quality data. The analysis considered the process steps and draw from the experience of European profiling and national profiling. For example, the technical functionalities of the German profiling tool *iProfAnT* [6] was used as an example of the process.

Usually, Top-tier MNE groups are already subject to national profiling, which is responsible to delineate the structure of the statistical unit enterprise and/or are treated by the Large Cases Units, which ensure consistency between the enterprises and legal units within the MNE group. The Task Force therefore considered that the delineation of the statistical unit enterprise in the Top-tier profiling could be considered an optional operation (by the national profilers) and needed only in certain special cases, for instance if the MNE group has undergone restructuring and/or additional consistency work is necessary across countries. This information could be sent to the EGR during the usual annual process from the national business registers. In this way, and assuming that the statistical unit enterprise is correctly delineated already in the national business registers, the Top-tier profiling process has been re-designed in such a way to allow profilers to focus on the correctness and completeness of the global structure of legal units in the MNE group only. Such an approach reduces the processing time and increase the number of Top-tier MNE groups, at the same time avoiding or reducing the work on the enterprises’ delineation. Figure 3 describes the Top-tier process and shows in light blue the optional steps, while steps in orange and green are the mandatory ones.

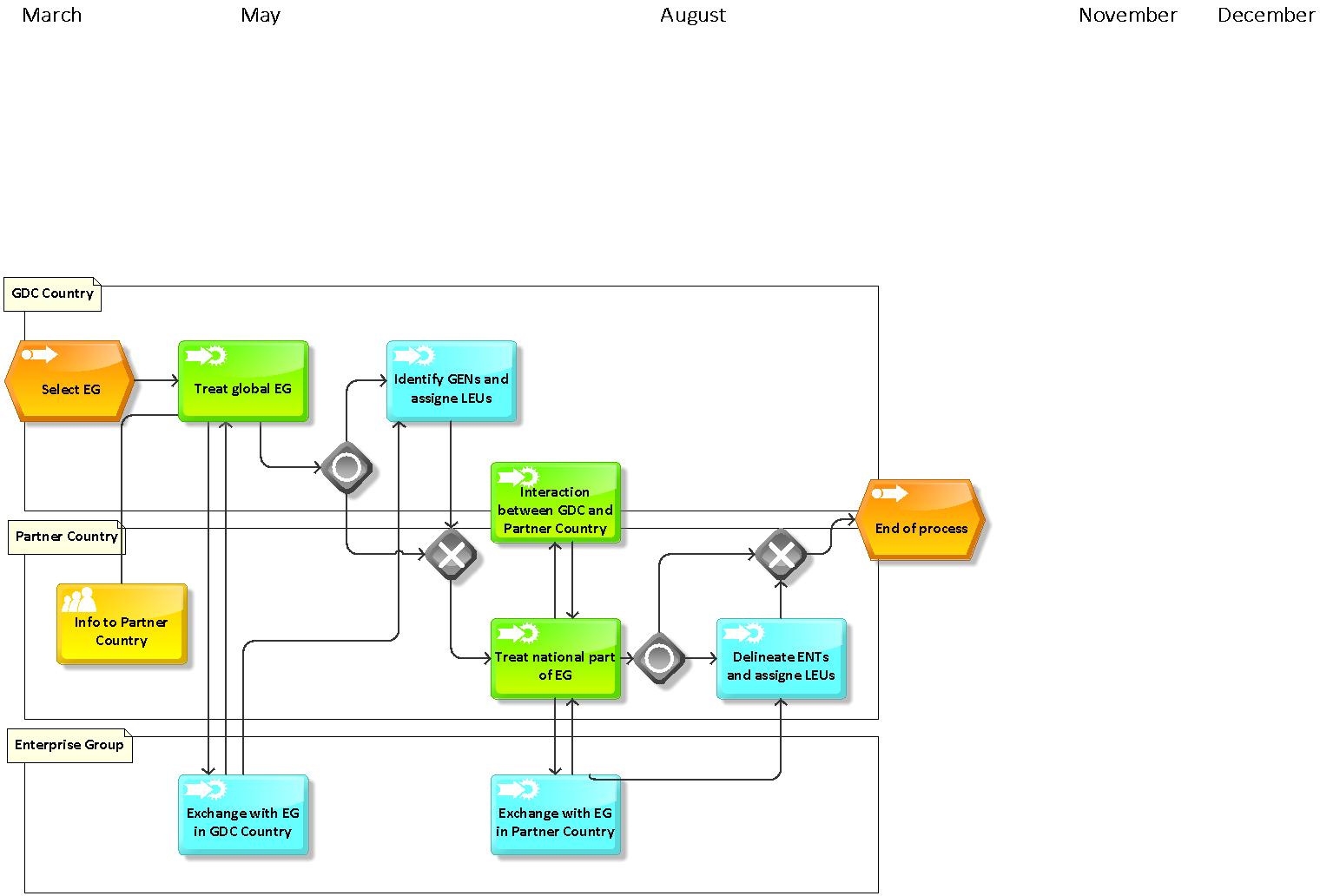


Figure 3: Lean treatment process with optional steps

In the next paragraphs a short description of the Top-tier profiling is provided.

As a first step after selecting one Top-tier MNE group from the list drawn by the CSI index, the GDC country shall compare the data in the EGR with other available sources to clarify the structure of the MNE group. The data should be compared with the data in the national business register and the information in the annual report(s) published by the MNE group. In addition, as many other sources as possible can be used to obtain and compare data on the MNE group. The time and effort required to compare often several hundred or thousands of legal units can be significantly reduced by supporting them with intelligent comparison functionalities. The comparison functions should support the multi-dimensional comparison of lists and give the user as much flexibility as possible, because in manual editing the decision always lies with the profiler in the end. The task force developed a prototype in Excel that can compare any number of lists with different characteristics, flexibly controlled by the user. The prototype uses similarity measures such as *Levenshtein* distance and is intended to serve as a template for implementing these functionalities in the EGR modules.

The comparison functions should then also be available to all partner countries working on the same MNE group to compare the data on the resident part of the MNE group with their national business register. As this is a multi-participant process, there are likely to be discrepancies and coordination needs, so technical support must be put in place to make these communication processes efficient and targeted.

This lean process of manually editing the data of an MNE group with intelligent technical support and clear roles is illustrated in figure 4. It also shows the timeframe over the year, in which month which process steps can take place. The EGR system automatically integrates the results of European profiling at the end of the process and treats it according to well defined priority rules during the consolidation of the MNE groups structures in order to offer users the highest quality information.

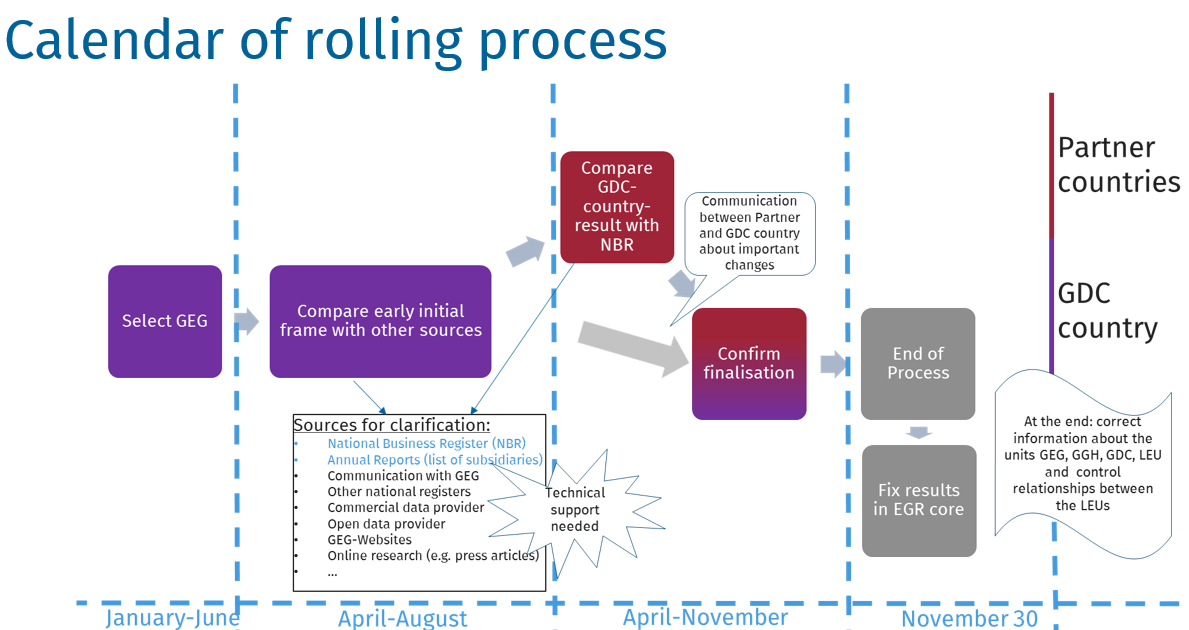


Figure 4: Lean treatment process with technical support

# ****current status and future goals****

The implementation of the two-tier process is already well advanced and the necessary transformation of the EGR system is ongoing. The EGR data reference year 2021 integrated for the first time the European profiling data. The EGR data reference year 2022 will contain already at the beginning of T+12 high quality information coming from profiling on the Top tier MNE groups. At the time of writing this paper, an almost double number of MNE groups, compared to previous years, is being profiled by NSIs, either including the statistical enterprise delineation, or focusing on the top tier MNE groups’ structure. The goal is to further increase the accuracy of all top tier MNE groups in the short run.

In the light of the experience gained, the described Top-tier profiling process will be further improved and, additional methodology is being developed to ensure an annual event-driven approach for the updating of the MNE groups and the move towards a continuous updating of the EGR data. As a first step the goal is to further increase of timeliness and frequency of EGR delivering users additional frames: from 2024 one more EGR data release at T+4 will be available for testing by the NSIs.

Figure 5: Continuous EGR cycle from 2024

This newly introduced process will make it possible to significantly increase the exchange of updated information between the national business registers and the EGR increasing the accuracy and timeliness.

In this way, data on MNE groups can be made available to the different statistical domains in a consistent and as up to date way as possible, and the European system of business registers can be set up in a way that adequately reflects and makes available the effects of globalisation and fast- changing structures of MNE groups.

# References

[1] DGINS Conference (2019). Bratislava Conclusions, *The statistical implications of economic globalization*.

[2] [*European business statistics methodological manual for statistical business registers* — 2021 edition - Products Manuals and Guidelines - Eurostat (europa.eu)](https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-006) - <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-006>

[3] E. Morganti, M. Jalava, I. Sopranidis - *Strategy and solutions for the Eurogroups Register (EGR)* - [Meeting of the Group of Experts on Business Registers UNECE](https://unece.org/info/Statistics/events/364567) -<https://unece.org/info/Statistics/events/364567> – 2022.

[4] *Structure of multinational enterprise groups in EU* (<https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Structure_of_multinational_enterprise_groups_in_the_EU>) - <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-20-002>

[5] [*European profiling - experiences of National statistical institutes* - Statistics Explained (europa.eu)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=European_profiling_-_experiences_of_National_statistical_institutes)

[6] [*European business profiling — Recommendations manual — 2020 edition* - Products Manuals and Guidelines - Eurostat (europa.eu)](https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-20-002) - <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-20-002>

[7] Rommelspacher (2021): iProfAnT – *Integrated Profiling Analysis Tool*; 27th Meeting of the Wiesbaden Group on Business Registers, 20 – 24 September 2021, Mexico, Session Profiling and Large Cases Units - <https://www.inegi.org.mx/eventos/2021/wiesbaden/>

1. DGINS Conference is held once a year with the aim of discussing topics related to the statistical program and methods and processes to produce Community statistics. It is hosted each year by a different Member State and the Director-General of the host country chairs the conference. [↑](#footnote-ref-1)
2. Intellectual property products are defined as produced non-financial assets that are the result of research, development, investigation, or innovation which led to knowledge that the developers can market or use to their own benefit in production, because use of the knowledge is restricted by means of legal or other protection. [↑](#footnote-ref-2)
3. We use the word “decile” to simplify the reading, but to be precise, we should use the term “decile” for all variables and the “25th percentile” for employment. [↑](#footnote-ref-3)
4. We define in the European Business regulations the **global decision centre** (GDC) of an enterprise group which is the unit where the enterprise group level strategic decisions are taken. [↑](#footnote-ref-4)