International Labour Organization

▶ Room document: 26

Unlocking the Power of Microdata: Enhancing International Comparability and Data Availability in ILOSTAT and Beyond

100 years of ICLS

International Conference of Labour Statisticians

21st International Conference of Labour Statisticians
Geneva, 11-20 October 2023

* This document was not formally edited

▶ Contents

▶ Acknowledgements

▶ Introduction

1. Statistics at the ILO are embedded in its constitution, and thus the ILO has been involved in statistical activities since its inception in 1919. Article 10.1 of the Constitution of the Organization requires "the collection and distribution of information on all subjects relating to the international adjustment of conditions of individual life and labour." Today, this translates into a threefold vision for the ILO Department of Statistics:

   ▶ To provide relevant, timely, reliable, and internationally comparable labour statistics

   ▶ To develop international standards for better measurement of labour issues and enhanced international comparability, and

   ▶ To assist Member States to develop and improve their labour statistics.

2. Within the United Nations System, the department serves as the focal point for labour statistics. The department's Data Production and Analysis Unit (DPAU) oversees the compilation, production, and dissemination of labour statistics and publishes country-level, regional, and global figures for short-term and annual indicators on ILOSTAT, the ILO's portal to labour statistics.

3. For decades, the primary method for compiling labour statistics from National Statistical Systems (NSSs) was through a questionnaire, initially as a paper version and eventually transitioning to a spreadsheet sent electronically. During the financial crisis of 2008-09, DPAU introduced web scraping as an additional approach to capture short-term indicators in a timelier fashion. Around 2016, a transformative data revolution started to take shape when the ILO began systematically collecting and processing national survey microdata. This was facilitated by a growing trend of NSSs publicly releasing not just their aggregate data, but also their microdata – that is, the specific individual-level data collected from surveys and censuses.

4. In fact, the concept of open data had been gaining prominence amongst NSSs in the early to mid-2010s, coinciding with the introduction of the United Nations Sustainable Development Goals (SDGs) in 2015. As governments and organizations recognized the benefits of transparency and collaboration, open data initiatives emerged, aiming to provide accessible, valuable, and more granular statistical information to the public.

5. This pivotal shift opened the door for DPAU to adopt another data channel to provide relevant, timely, reliable, and internationally comparable labour statistics: microdata. In less than 8 years since this initiative began, DPAU collected more than 13'000 sets of microdata from 177 countries and territories, including 160 ILO Member States (see Figure 1). Notably, microdata has now emerged as the primary source of data published by the ILO. Currently, it provides a staggering 95 per cent of the 260 million data values available on the ILOSTAT portal. Its application for ILO research and publications, especially for emerging topics, holds even greater significance.

▶ **Figure 1. Enhancing data availability on ILOSTAT through microdata processing**
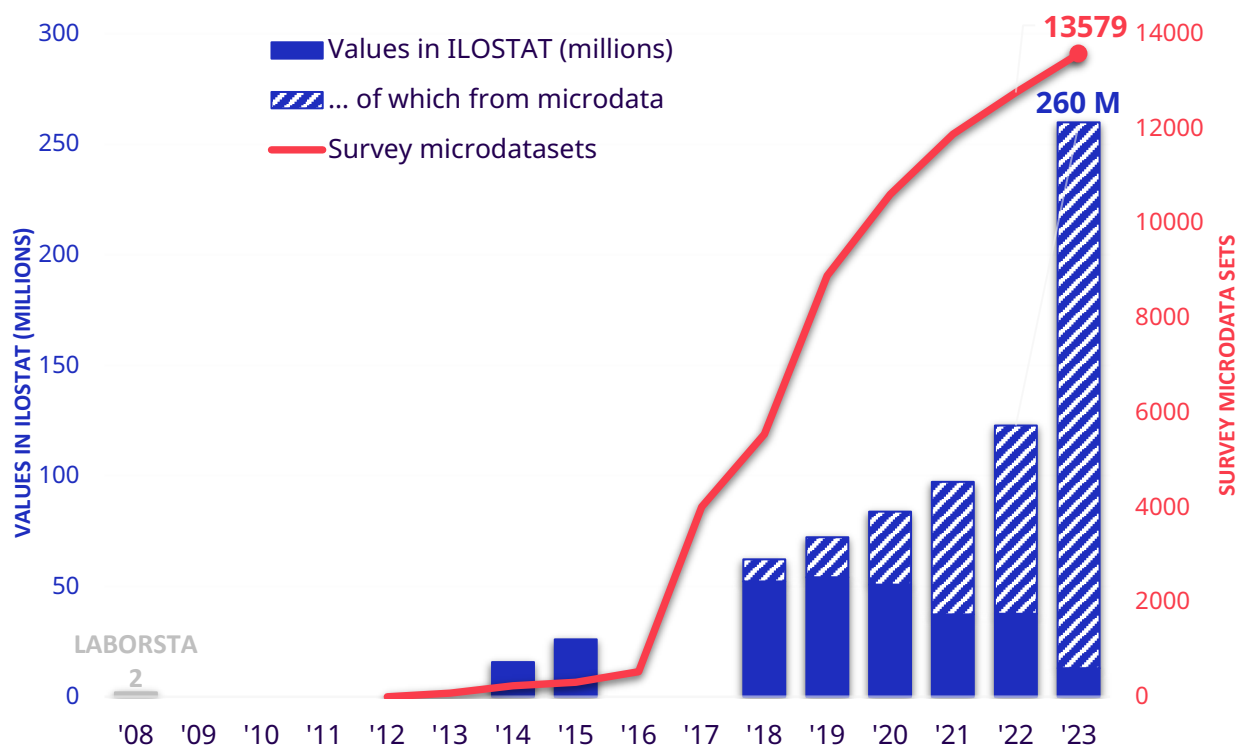
Source: ILOSTAT

6. This room document discusses the profound impacts of sharing microdata with the ILO. It begins by outlining the rationale and process undertaken by DPAU in conducting this endeavour. Subsequently, it explains the potential ramifications and risks faced by countries refraining from sharing microdata. Concluding with insights into forthcoming initiatives, the document also expresses the ILO's gratitude to the 160 Member States actively participating in microdata sharing while proposing strategies to address challenges for the remaining 27 Member States.

▶ The value of sharing microdata for ILO processing

7. The multi-faceted significance of data openness, demonstrated through the act of sharing microdata with the ILO, spans across various dimensions. It intricately links with international reporting and collaboration, organizational efficiency, and, most importantly, the potential it holds for evidence-based policymaking to support the world of work.

Data openness and its significance

8. Sharing microdata with the ILO signifies a commitment to the principles of open data and underscores a willingness among NSS's to embrace transparency and collaboration. This act not only promotes the ethos of sharing information for the greater good, it also fosters a collaborative environment where data can be harnessed for evidence-based decision-making, international benchmarking, and the pursuit of common goals. It signifies a proactive stance in contributing to a global data ecosystem that empowers diverse stakeholders to derive meaningful insights, track progress, and collectively work towards shared objectives, ultimately driving informed policymaking and sustainable development.

9. Open data plays a crucial role in advancing the SDG objectives by promoting transparency, accountability, collaboration, and informed decision-making. In essence, open data acts as a catalyst for achieving the SDGs by ensuring that data is widely available, transparently presented, and effectively utilized to drive positive change for all citizens.

Unveiling new insights

10. The transition to open data policies and the accompanying growth of the ILO's microdata repository unleashed a wealth of insights regarding the world of work. By leveraging microdata, the ILO has vastly expanded the number of indicators available on ILOSTAT, providing researchers and data users the opportunity to delve into nuanced analyses that highlight trends, challenges, and opportunities in labour markets across the globe. The availability of granular information empowers decision-makers to formulate evidence-based policies that address specific workforce dynamics and inequalities more effectively.

11. For example, granular data enables researchers to study subgroups within the labour force, such as youth, women, the elderly, migrants, working parents, persons with disabilities, and minorities. By examining the experiences and challenges faced by these subpopulations, policymakers can design targeted interventions to address their unique needs and foster inclusive labour markets. This approach helps mitigate longstanding inequalities and contributes to the achievement of the SDGs.

Enhanced data availability on the ILOSTAT portal

12. The appetite for data, including labour statistics, is strong and continuously growing. To meet these demands, the ILO currently publishes more than 900 indicators on ILOSTAT with 260 million data values, of which 95 per cent are generated solely from microdata processing. Obviously, it would not be reasonable for the ILO to request this many tables from data providers, yet these data are necessary for more granular labour market analysis. Some examples of tables that are purely microdata-based include harmonized indicators on informality, school-to-work transitions, and skills mismatch. Similarly, various cross-tabulations are produced by deriving classifications to capture disability status, household type (including presence of children), marital status, detailed occupations and economic activities, working time arrangement, type of employment contract, multiple job holding, and establishment size.

Evidence-based ILO reports and guidance

13. Leveraging microdata has strengthened the ILO's capacity to generate detailed data tabulations for its use in research and publications. This new capacity has positioned the ILO as a pioneer in delivering evidence on emerging topics where scant international data had existed previously, such as essential work (defined as eight occupational groups related to food, health, and transport), digital platforms, and care work. Notable examples of ILO publications relying on the availability of microdata include:

> *World Employment and Social Outlook* (WESO) thematic editions such as *WESO 2023: The value of essential work* and *WESO 2021: The role of digital labour platforms in transforming the world of work*

> *Global Employment Trends for Youth 2022: Investing in transforming futures for young people*

> *Care work and care jobs for the future of decent work*

> *Generative AI and Jobs: A global analysis of potential effects on job quantity and quality*

> *Asia-Pacific Employment and Social Outlook 2022: Rethinking sectoral strategies for a human-centred future of work*

> *What labour force survey data can tell us about the workforce in the health and social care sector*

> *Making decent work a reality for domestic workers: Progress and prospects ten years after the adoption of the Domestic Workers Convention, 2011 (No. 189)*

> *Women and men in the informal economy: A statistical update*

14. Additional reports and other types of analytical outputs, such as ILOSTAT blogs and ILO research papers, also rely on the granular data generated through microdata processing.

15. Having access to microdata also proves valuable in the formulation of ILO programs aimed at promoting decent work. This includes the development of project proposals and engagement with potential donors.

16. The availability of comprehensive and reliable data, largely achieved through microdata processing, positions the ILO as a global leader in labour statistics. This strengthens the ILO's role in shaping international labour policies and enables it to provide evidence-based guidance to Member States and other stakeholders.

Improved international reporting

17. International reporting of labour statistics largely draws upon data from household surveys, including for 6 SDG indicators under ILO custodianship. Data deemed as unreliable and/or not internationally comparable are not reported and thus are omitted from the global SDG database. This is important to note since not all data reported to the ILO through questionnaires conforms to internationally agreed statistical standards and definitions. However, with access to household survey microdata files, key labour market indicators can be derived based on international standards, thereby ensuring data are comparable across countries.

18. To assess labour-related aspects of economies and societies, several well-known indices and reports also depend on ILOSTAT data and thus on microdata. Some of these include:

> **Human Development Index (HDI):** Developed by the United Nations Development Programme (UNDP), the HDI uses ILOSTAT data to calculate a composite index that assesses a country's development based on factors like life expectancy, education, and labour force participation.

▶ **Global Innovation Index (GII)**: Published by Cornell University, INSEAD, and the World Intellectual Property Organization (WIPO), the GII assesses the innovation capabilities and performance of economies worldwide. ILOSTAT data is used within the GII to evaluate various labour-related factors, such as workforce education, skills development, and labour market efficiency, which are essential components of a country's innovation ecosystem.

▶ **Global Competitiveness Index (GCI)**: The World Economic Forum's GCI incorporates ILOSTAT data to evaluate factors such as labour market efficiency, skills availability, and labour-employer relations in assessing a country's overall competitiveness.

▶ **World Bank's World Development Indicators (WDI)**: The WDI draws on a wide range of data sources to provide a comprehensive overview of global development trends. ILOSTAT data contributes to the labour-related indicators, offering insights into key labour market metrics, employment trends, informality, wage inequality, and other labour-related dimensions.

▶ **Global Gender Gap Report:** Produced by the World Economic Forum, this report uses ILOSTAT data to analyse gender disparities in labour force participation, wage equality, and other aspects of economic participation.

▶ **Corruption Perceptions Index (CPI)**: Transparency International's CPI utilizes ILOSTAT data to assess the prevalence of corruption within labour markets, including issues like informal employment and unfair labour practices.

▶ **OECD Better Life Index:** The OECD's index uses ILOSTAT data to evaluate various aspects of work-life balance, employment rates, and job security when assessing the overall well-being of populations.

▶ **Global Talent Competitiveness Index (GTCI)**: INSEAD's GTCI uses ILOSTAT data to assess labour market competitiveness, including factors such as skills development, workforce availability, and labour market efficiency.

Data demands from the UN system

19. In instances where users seek labour statistics that are not disseminated on ILOSTAT, ad-hoc tabulations can be produced using microdata (subject to feasibility and compliance with data confidentiality regulations). As the UN focal point for labour statistics, the ILO provides this service to numerous UN and other international agencies. Examples of past inquiries are provided in Table 1.

▶ **Table 1. International agency data inquiries fulfilled using microdata**

Agency	Topic(s)
Food and Agriculture Organization (FAO)	Forestry workforce, fisheries workforce, agriculture workforce
Group of 20 (G20)	STEM occupations for G20 countries
International Monetary Fund (IMF)	Employment by gender, age, education, and economic activity
International Atomic Energy Agency (IAEA)	Energy sector workforce
Office for the Coordination of Humanitarian Affairs (OCHA)	Labour force participation rate by sex, age (5-year age bands) and marital status
Organisation for Economic Co-operation and Development (OECD)	Education and skills, public sector employment by sex and other characteristics, transportation sector
United Nations Conference on Trade and Development (UNCTAD)	Female and male employment in the most export/trade intensive industries
United Nations Department of Economic and Social Affairs (UN DESA)	Working time, youth labour market indicators
United Nations Development Programme (UNDP)	Labour-related income
United Nations Educational, Scientific and Cultural Organization (UNESCO)	STEM occupations, cultural employment, youth not in employment, education or training (NEET)
United Nations Entity for Gender Equality and the Empowerment of Women (UN-Women)	Gendered effects of marriage and children, STEM occupations
United Nations International Children's Emergency Fund (UNICEF)	Child labour, NEET
World Bank (WB)	Wages
World Economic Forum (WEF)	Labour force participation rate, occupations (ISCO)
World Health Organization (WHO)	Health workforce, sanitation workers

Source: DPAU, ILO Department of Statistics.

Boosting efficiency in reporting

20. Each year, the ILO sends out Excel questionnaires to be completed by NSSs to report their latest figures on various labour-related topics. In recent years, the requests were decreased to include only data that could not be obtained through microdata tabulations. This has led to a remarkable reduction in the reporting burden. For instance, most NSSs receiving the ILOSTAT questionnaire in 2023 had a maximum of 34 data tables to complete if they shared their microdata, compared

to 65 tables for those which did not. (The remaining 34 tables in the questionnaire are to capture data from establishment surveys and administrative records.)

21. Meanwhile, driven by open data initiatives (leading to microdata sharing) and technological advancements, the ILO has orchestrated a transformation in its data publishing efficiency. The now discontinued LABORSTA database, a repository of 2 million data points in 2012, required a team of 12 statistical assistants to process incoming Excel questionnaires. The current ILOSTAT database is over 100 times larger than LABORSTA. Notably, this growth was achieved with fewer staff, enabling new work in advanced technical areas such as modelling and analysis. This achievement would have been unattainable if staff had remained heavily engaged in processing Excel questionnaires.

Improving data quality and international comparability

22. The shift from Excel questionnaires to microdata processing marked a significant improvement in data quality. Despite their best efforts, data reporters frequently submit questionnaires with incomplete and erroneous data—a predictable outcome due to the cumbersome nature of the process. Although numerous checks, including automated ones, are implemented to identify glaring errors before data is published on ILOSTAT, the detection of all errors is simply not feasible. Unlike the former method, the nature of microdata processing substantially reduces the risk of inaccuracies.

23. Furthermore, the use of microdata significantly enhances international comparability. The microdata processing routines generate standardized variables and indicators that adhere to international statistical standards, eliminating variations in terminology and methodology that could otherwise hinder meaningful comparisons. Thus, using microdata ensures that international labour statistics are not only comprehensive and accurate but also directly comparable.

Shaping best practices and supporting countries

24. Microdata pre-processing involves an extensive review of the documentation accompanying the surveys. This provides key insights into national practices in terms of survey methodologies and questionnaire design. This also creates opportunities for the exchange of knowledge and expertise between NSSs and the ILO.

25. Specifically, it can inform the standard-setting process. The comprehensive understanding of different approaches used for questionnaires, achieved through the microdata review process, facilitates the development of guidelines and recommendations that are both relevant and adaptable to the diverse practices across countries. A notable example is on the framework for informality statistics currently under development. Current survey questionnaire practices across countries were identified during microdata processing, which informed the development of the framework which will be discussed at this 21st ICLS. Another example is the *Quick guide to understanding the impact of the new statistical standards on ILOSTAT databases*. Microdata files were used to identify different approaches to implementing the 19th ICLS standards and to demonstrate the implications of adopting these standards on key labour market indicators.

26. Similarly, by reviewing current practices during microdata processing, it allows for the identification of areas where a country may need guidance. The ILO can provide valuable feedback on improving survey questionnaires, enhancing compliance with international statistical standards, and refining calculations used to derive indicators.

▶ From micro-level records to macro indicators: How is it done accurately and efficiently?

27. Processing microdata is a cornerstone of the ILO's statistical endeavours, yielding a rich set of labour statistics by harnessing household surveys from Member States throughout the world. This process involves establishing a protocol to access anonymized microdata files, harmonizing microdata variables, and generating a comprehensive set of standardized indicators following stringent quality assurance procedures. Collaboration remains pivotal throughout the entirety of this process, as ILOSTAT staff engage in regular communication with Member States. This ongoing engagement ensures data accuracy and provides a platform for addressing any queries or reservations that Member States might have.

Microdata access

28. The initial step requires identifying current household surveys within each Member State that could be utilized as potential sources of labour statistics. These surveys are primarily labour force surveys, but also include household and income expenditure surveys, child labour surveys, and general household surveys with a labour module.

29. Depending on the national data handling and dissemination rules, anonymized microdata files for these surveys and their documentation may be accessible online. This holds true for 40 Member States that have fully embraced open data policies.

30. Alternatively, access to the anonymized files is granted by the National Statistical Office (NSO). In this scenario, an agreement in the form of a Memorandum of Understanding (MOU) is established between the ILO and the NSO to ensure proper use of the shared files. Since 2016, the ILO has entered into MOUs with NSOs from more than 100 Member States. Regardless of the type of data sharing approach, the ILO requests all essential documentation, including survey guidelines, questionnaires, data dictionaries, and reports.

31. All microdata files and documentation are stored securely and are only accessible to ILO staff after they sign an agreement on proper data usage. Wherever data is not public, adherence to the MOU is meticulously enforced. This means that data access is restricted solely to individuals authorized by the agreement, whereby various levels of access are exclusively granted to approved personnel, and the data is stored in dedicated high-security computer servers.

Harmonization of variables

32. The central part of the ILO's work involves creating a set of harmonized variables by processing the microdata using statistical software such as Stata. This entails reviewing survey questions and response categories to create variables based on internationally agreed concepts and definitions wherever possible. For example, national classifications are mapped to international ones such as ISCED to create a variable for education, ISCO for occupations, and ISIC for economic activities. When international classifications do not exist, the ILO creates a standardized set of categories. A simple example is the institutional sector of the workplace, whereby working in government is considered as "Public sector" and the rest as "Private sector".

33. The newly created variables are saved in a separate .dta file, and their accuracy is cross-checked with official national statistics published by the NSO. Differences between ILO derived indicators and national statistics are tracked and documented.

34. Through this process, the ILO currently produces up to 154 variables for each micro dataset, depending on the comprehensiveness of the country's household survey. This is a significant increase from the 50 or so variables initially produced in 2016.

35. Once the standardized datasets with harmonized variables are generated, they are used as input files to an R package that produces pre-coded indicators, along with their associated quality measures. This step includes checks for data consistency and completeness across the indicators.[1]

Quality assurance and dissemination

36. The quality assurance and dissemination phase of microdata processing involves rigorous checks to ensure data consistency and completeness across indicators. For example, indicators are subject to reliability compliance tests to determine their suitability for dissemination on ILOSTAT. These tests guarantee that figures derived from the sample size meet established requirements and that the values fall within reasonable bounds.[2] Should any indicator fail the verification tests, the process reverts to the pre-processing stages.

37. The microdata are regularly revised when Member States conduct new Population Censuses and revise backwards their population estimates. This implies a change in only one variable (i.e., individual weights) and thus, this is a straightforward process due to the automated procedures in place and can be done within a day of receiving the revised microdata.

38. Numerous R scripts are executed to process the final files uploaded into the system. These scripts not only generate almost all possible cross-tabulations for each survey with new or revised figures but also perform indicator calculations and compute annual averages based on monthly and quarterly values. It is not uncommon for the programs to generate over a million values in a given month.

39. Finally, the published results from the anonymised microdata processing align with recommendations by the ILO and United Nations Statistical Commission.[3] The derived results complement official data and metadata reported by national data providers, aiming to enhance statistical reporting quality. The entire process involves close collaboration between the ILO Department of Statistics, regional ILO field colleagues, and NSOs and the materials developed by DPAU during processing are available upon request (see Figure 2 for a detailed summary).

---

[1] For instance, that employment, unemployment, and persons outside the labour force sum up to the working-age population.

[2] The rules for dissemination are set as follows: a value is (1) not published if the sample size used for producing it is less than 5, (2) published with a warning if the sample size lies between 5 and 14 (inclusive), and (3) published with no remark associated with the value if the sample size is greater or equal to 15. These rules apply equally for monthly, quarterly, and annual indicators.

[3] The Inter-agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs). "Global Consultation: Guidelines on Data Flows and Global Data Reporting for Sustainable Development Goals". January 2018.

▶ **Figure 2. Flowchart for producing labour market indicators using microdata**
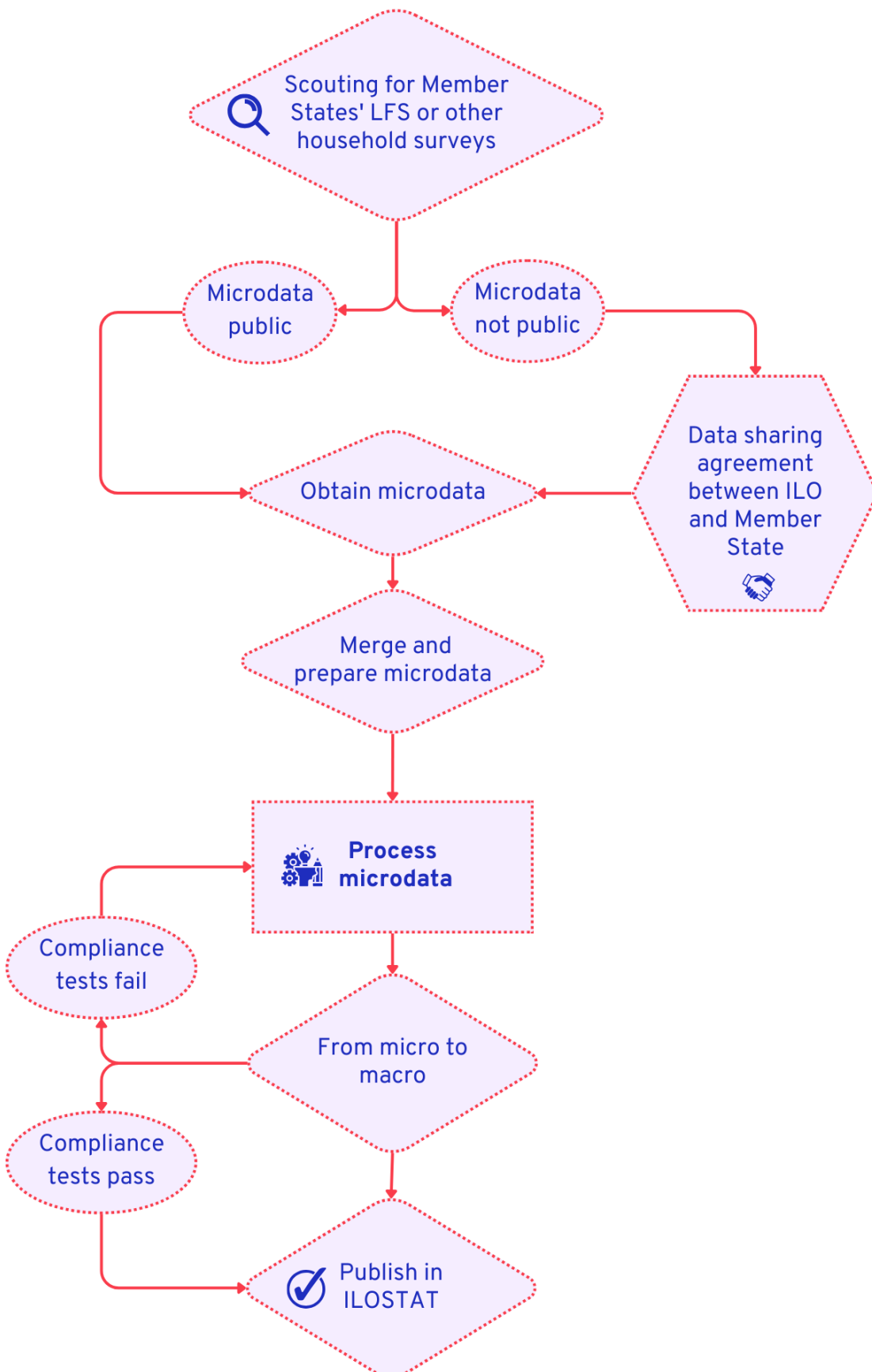
Source: DPAU, ILO Department of Statistics.

▶ The risks and consequences of withholding microdata

40. The most evident risk to withholding microdata lies in the absence or obsolescence of a country's labour statistics in ILOSTAT data outputs. Less apparent is the scope of these data gaps, spanning not only across ILO products but also UN and international reports and indices. Additionally, the absence of microdata contributes to a heavier reporting burden for the NSS, compared to those sharing microdata, while reducing efficiencies for both the NSS and ILO. Lastly, this situation might unfavourably influence the data provider's reputation in terms of data transparency and credibility.

Lack of data for international reporting and global assessments

41. Member States withholding microdata risk not having some of their labour statistics in the United Nations' global SDG database for indicators under ILO custodianship.

42. Member States withholding microdata risk not being included in global assessments such as the HDI, GII, GCI, WDI and many others (see paragraph 18).

43. Meanwhile, it is not uncommon for Member States that do not share their microdata to reach out to the ILO, seeking updates for a few figures solely for one of the SDG indicators or global assessments. This process becomes both unwieldy and inefficient, adversely affecting both the ILO and the Member State itself. Updating ILO databases demands consistent information across indicators and relies on the submission of numerous data points. However, the desired outcome is often not achieved, as the Member State must manually complete an Excel questionnaire for numerous tables. Moreover, the back-and-forth exchange typically involves considerable time and effort for both parties.

44. Lastly, UN agencies and international organizations rely on the ILO for labour statistics. While ILOSTAT offers a wealth of tables which they can easily access, these organizations are increasingly reaching out to the ILO for special tabulations based on microdata to address their specific needs (see Table 1). For these requests, it is not technically viable to include data from countries that do not provide microdata to the ILO.

Lack of data for ILO publications, ILOSTAT and beyond

45. The lack of microdata for some countries hinders the ILO's ability to produce more insightful data outputs, reports, analyses, and forecasts, which are essential for policymakers, researchers, and stakeholders worldwide. Missing data could mean that ILO publications and ILOSTAT databases do not reflect the diverse realities and nuances of labour markets across countries and regions. This is particularly problematic when the ILO is missing microdata representing a significantly large population in a given region.

46. For example, flagship ILO publications relying on the availability of microdata include the WESO, Global Wage Report (GWR), and flagship reports on domestic workers and care workers. Additional examples are provided in paragraph 13.

47. Similarly, ILOSTAT relies on microdata to provide current and accurate information for its extensive user base. The absence of microdata access impedes the timely updating of databases on ILOSTAT, impacting users, including policymakers, researchers, journalists, and civil society organizations. The lack of data could restrict stakeholders from having the essential tools to analyse trends, advocate for reforms, and design effective policies aimed at fostering decent work worldwide.

48. The lack of microdata for some countries hinders the ILO's ability to formulate effective emergency crisis-response and to develop its Post Disaster Needs Assessment (PDNA) analysis. This year, this occurred for two countries for which microdata was not available. Having the data would have facilitated the assessment of the damages to the labour market and formulation of appropriate crisis-response plans.

Reduced efficiency

49. When NSSs opt not to share microdata with the ILO, there are inherent inefficiencies that arise for both parties. The NSS must rely on traditional methods for data reporting, which involves manual data entry into Excel questionnaires. This cumbersome process is time-consuming for both the NSS during the reporting phase and the ILO during the processing stage. It diverts the attention of NSS and ILO staff away from higher value-added tasks such as data production, analysis, and interpretation.

50. Moreover, the number of data tables collected through Excel questionnaires pales in comparison to the extensive array of tables generated through microdata processing.

Potential reputational damage

51. Choosing not to share microdata with the ILO may raise questions about transparency and accountability within an NSS. Such a decision could lead to questions regarding the accuracy and reliability of their official data, potentially impacting on perceived credibility. This stance could also influence perceptions among national and international stakeholders and impact their image as cooperative participants in global discussions, potentially affecting mutual understanding of their socio-economic landscape.

The use of unofficial statistics to fill data gaps

52. Refraining from sharing microdata could potentially leave data gaps that could be filled by unofficial statistics. Various private firms and research organizations conduct their own surveys and/or use online platforms to gather labour market information. By sharing microdata to generate a large wealth of labour statistics, NSSs can actively participate in shaping the narrative around their own socio-economic situation, rather than allowing potentially incomplete or skewed data to define their national story.

▶ Concluding remarks

53. The power of microdata lies in its potential to unlock valuable insights into labour market dynamics. The ILO's efforts to enhance data availability and quality through microdata processing contribute to this goal, while reducing reporting burdens and supporting the measurement of progress towards achieving the SDGs. By harnessing the power of microdata, the ILO fosters evidence-based decision-making to advance labour market policies and the promotion of a more inclusive and sustainable world of work.

54. Looking ahead, the ILO will continue leveraging its expertise in labour statistics, including its efficiency in processing household microdata, to deliver more opportunities for data-driven decision-making in the world of work. This requires accessing microdata for all Member States carrying out household surveys, without exception. The ILO recognizes the challenges faced by some data providers in making their microdata available and encourages Member States to work collaboratively in overcoming these obstacles.

55. Regarding Member States that necessitate a MoU, the ILO promotes adoption of open data practices, encouraging them to publish information online, thereby obviating the need for MoUs. This shift would not only enhance data accessibility but also significantly reduce the administrative overhead required to obtain these files. In addition, the ILO is exploring undertaking similar work to process microdata for establishment surveys. This has the potential to provide labour market information on firm characteristics, their workforces – such as employees' pay and benefits, working time, and training needs – and job vacancies and labour turnover. The main challenge in this endeavour is the lack of publicly available microdata for these types of surveys.

160 Member States share microdata with the ILO

56. The work detailed in this room document would not have been possible without the collaboration and support of the countries that have shared their household microdata. The ILO extends its gratitude to the 160 Member States who have made this possible, with special acknowledgment of the 40 countries who have made their microdata available publicly and freely (see Table 2).

57. Nonetheless, the ILO would like to share a few recommendations on how Member States can further enhance the sharing of their microdata.

58. Although most microdata sets are shared with the ILO in their entirety following the anonymisation process, the ILO would like to encourage Member States that share only partial microdata or remove certain variables to reconsider such exclusions. This recommendation is made in light of the rigorous process and quality assurance measures outlined in Figure 3.

59. The ILO would also encourage countries to share the original variables rather than those that are derived by the NSO. This is to ensure all possible tabulations can be generated. For example, if unemployment is already derived at national level, the ILO is unable to capture persons actively seeking employment and persons available for work, a crucial step for producing data on the composite measure of labour underutilisation (LU4).

60. Three Member States continue to charge fees for providing access to their microdata. The ILO would encourage them to follow the path of the remaining 157 Member States sharing these microdata sets for free as they are public goods.

▶ **Table 2. Member States sharing microdata with the ILO**

ILO region	Member States with publicly and freely available microdata	Member States with whom an MOU is in place (or microdata shared directly with the ILO Department of Statistics)
Africa	Egypt, Ghana, Kenya, Rwanda, South Africa, Zimbabwe	Angola, Benin, Botswana, Burkina Faso, Burundi, Cabo Verde, Cameroon, Chad, Comoros, Congo, Côte d'Ivoire, Democratic Republic of the Congo, Djibouti, Eswatini, Ethiopia, Gambia, Guinea, Guinea-Bissau, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Senegal, Seychelles, Sierra Leone, Somalia, Sudan, Togo, Tunisia, Uganda, United Republic of Tanzania, Zambia
Americas	Argentina, Bolivia (Plurinational State of), Brazil, Canada, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Guyana, Mexico, Paraguay, Peru, United States of America, Uruguay	Antigua and Barbuda, Barbados, Belize, Dominican Republic, Grenada, Haiti, Honduras, Jamaica, Nicaragua, Panama, Saint Lucia, Suriname, Trinidad and Tobago, Venezuela (Bolivarian Rep. of)
Arab States	Lebanon	Iraq, Jordan, Saudi Arabia, United Arab Emirates, Yemen
Asia and the Pacific	India, Iran (Islamic Republic of), Mongolia, Pakistan, Philippines	Afghanistan, Bangladesh, Brunei Darussalam, Cambodia, Cook Islands, Fiji, Indonesia, Japan, Kiribati, Lao People's Democratic Republic, Maldives, Marshall Islands, Myanmar, Nepal, Palau, Papua New Guinea, Republic of Korea, Samoa, Singapore, Solomon Islands, Sri Lanka, Thailand, Timor-Leste, Tonga, Tuvalu, Vanuatu, Viet Nam
Europe and Central Asia	Armenia, France, Georgia, Greece, Ireland, Italy, Latvia, Lithuania, Russian Federation, Spain, Ukraine, United Kingdom	Albania (until 2019), Austria, Belarus, Bosnia and Herzegovina, Cyprus, Czechia, Estonia, Hungary, Israel, Kyrgyzstan, Malta, Montenegro, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Serbia, Slovakia, Slovenia, Switzerland, Tajikistan, Türkiye, Uzbekistan. Eurostat countries only for the standardized EU-LFS and EU-SILC: Belgium, Bulgaria, Croatia, Denmark, Finland, Germany, Iceland, Luxembourg, Netherlands, Norway, Sweden

Source: DPAU, ILO Department of Statistics.

Overcoming obstacles to access microdata for the remaining 27 Member States

61. The ILO aims to further expand the utilization of microdata. Recognizing that official statistics are public goods, the ILO is eager to collaborate with the remaining 27 Member States to obtain access to their household microdata files (see Table 3).

62. Certain challenges and obstacles have been identified that could hinder this endeavour. The subsequent sections outline potential solutions. By overcoming these barriers, the ILO and NSOs can work together to ensure a more inclusive and collaborative approach to harnessing the power of microdata.

▶ **Table 3. Member States not sharing microdata with the ILO**

ILO region	Member States with a recent (2015+) household survey	Member States without a known household survey in recent years
Africa	Algeria, Eritrea, Gabon, Libya, Morocco	Central African Republic, Equatorial Guinea, Sao Tome and Principe, South Sudan
Americas	Cuba, Dominica	Bahamas, Saint Kitts and Nevis, Saint Vincent and the Grenadines
Arab States	Bahrain, Kuwait, Qatar	Oman, Syrian Arab Republic
Asia and the Pacific	Australia, China, Malaysia, New Zealand	
Europe and Central Asia	Azerbaijan, Kazakhstan	San Marino, Turkmenistan

Source: DPAU, ILO Department of Statistics.

Perceived risks

63. An NSO might have concerns that sharing microdata could lead to potential risks, such as misinterpretation of data, breach of confidentiality, or data misuse. However, steps are taken to ensure data privacy and security by both the ILO and the NSO. To mitigate any concerns regarding confidentiality, it should be highlighted that the ILO only needs fully anonymized microdata. For a given respondent to the survey, the only personal information needed includes the sex, age, level of education and broad geographical location (urban or rural areas). No other information related to the geographical location of the persons (region, county, district...) nor any other personal information (address, phone number, ID numbers...) are needed. Therefore, it is impossible for the ILO to identify any respondent. This is precisely why 160 ILO Member States already share their LFS microdata, including several of them with strict national laws on data protection.

64. With regards to data security protocols, the ILO IT Department (INFOTEC) has put in place various levels of security and all the microdata have been stored on drives with the highest level. Only persons with granted access to these drives can see them. They remain invisible and inaccessible to the rest of the ILO.

Technical barriers

65. An NSO might not have the technical capabilities or systems in place to facilitate the secure and efficient sharing of microdata. This could include data compatibility issues, inadequate data infrastructure, challenges in data anonymization, difficulties in data transfer and storage, and limitations in data security protocols.

66. With regards to anonymization, the usual practice involves the NSO undertaking this task, but it might be unfamiliar to them. If this is the case, the ILO, can technically support the NSO and explain which variables should be dropped before sharing the anonymized microdata.

67. The ILO is ready to receive the microdata in any statistical format and accept any type of microdata transferring protocol. If not available online, the ILO typically receives microdata through a secured file transfer protocol (FTP) over the Internet. Files can also be sent by e-mail or even by post on a CD or USB stick. The microdata remains the property of the NSO and the ILO is ready to adapt to obtain data through the system the NSO is the most comfortable with.

Data quality concerns

68. An NSO might worry that their microdata may not meet international quality standards, leading to inaccurate or misleading analyses and conclusions. In such cases, the ILO would be pleased to work with this NSO to ensure each variable is aligned with international statistical standards. If it is not the case, technical assistance can be provided to address any issues. This could involve updating the questionnaire or modifying the sample frame, for example.

69. An NSO may decide to discontinue conducting its labour force survey in favour of relying on administrative records, possibly due to the perception of cost-effectiveness, even though this is not a superior alternative. It's important to note that administrative records cannot serve as a substitute for the LFS and should be reserved for indicators where the LFS is not an ideal data source (e.g., occupational injuries). For example, using administrative data on the registered unemployed is inadequate for measuring the unemployment rate in a country; it does not and cannot capture unemployment according to international statistical standards.[4] The ILO would be pleased to work with NSOs to ensure LFS continue to be conducted on a regular basis.

Philosophical perspectives

70. Microdata files underlying official statistics are valuable resources that support the transparency, accuracy, and credibility of those statistics. However, the classification of microdata files as a public good in themselves is more nuanced and may depend on various factors, including the level of confidentiality, privacy concerns, and the extent to which they can be used to identify individuals or entities. These are valid concerns. However, numerous steps are involved in ensuring data privacy and security as previously described.

71. Overall, the microdata processing bolsters the functionality of official labour statistics as a public good by providing accurate, detailed, and internationally comparable insights into labour markets. This, in turn, supports evidence-based decision-making, fosters collaboration, and advances labour-related research and policy development on a global scale.

---

[4] For more information on sources for labour statistics, refer to: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_647109.pdf

Legal constraints

72. The legal framework for data sharing varies globally, with some NSOs restricted by privacy and data protection laws. Such constraints, driven by concerns about data privacy and confidentiality, can impede direct microdata sharing with international organizations. These restrictions may arise from data protection laws that limit sharing personally identifiable information, but proper anonymization of microdata can alleviate this concern.

73. In most instances, although a given legal structure might appear stringent, it does permit the sharing of microdata, even with international organizations. However, the issue lies in the perceptions of NSO managers or staff, resulting in a lack of proactive engagement in pursuing such data-sharing initiatives.

74. The ILO also encourages NSOs to advocate for policy changes or legal reforms to align with open data initiatives. This might include working with national governments to review and revise existing data protection laws, privacy regulations, or other legal barriers that hinder the sharing of microdata. By fostering a supportive legal environment for data sharing, NSOs can participate more readily in open data initiatives. This, in turn, will help strengthen the global data ecosystem so that it can empower diverse stakeholders to derive meaningful insights, track progress towards shared goals, and drive informed policymaking aimed at fostering decent work and sustainable development for all.