Summary of the Third Webinar of International Statistical Classifications Sprint

31 January 2024

Introduction

1. The webinar, organized by the United Nations Network of Economic Statisticians and the United Nations Committee of Experts on International Statistical Classifications, explored the role of technology in making the process of classification more efficient. Practical examples from various NSOs and organizations using international classifications illustrated this point. Additionally, the webinar hosted a panel of senior executives from organizations around the world who addressed some of the biggest challenges faced by organizations when implementing international classifications, such as the frequency of revisions, building and maintaining capacity, modernizing the classification process and more.

Role of technology in the Statistical Classifications Process

2. A series of presentations delved into the various fashions in which technology could be used by statisticians to render the process of classifying units, as well as that of managing classifications, efficient and robust. UNECA and INSEE showcased their utilization of machine learning and open-source tools to automate classification and, in the UNECA case, also build a data visualization tool.. While the applications in the approach cannot address challenges such as the different languages used and heterogenous levels of details, they, nevertheless, have reduced the need for manual intervention considerably. The application of machine learning in the context of classifying large volumes of scanner data was illustrated by Statistics Canada's example of coding information on monthly retail sales obtained from retailers using the North American Production Classification (NAPCS) Canada . The example showed a successful integration of ML in the retail trade survey processing system, although, the approach presented particular challenges in maintaining the system up to date as technology evolves.

3. Data interoperability and governance, as it pertains to classification metadata, is an important objective of modern statistical organizations. Caliper, an open link data platformed developed by the Food and Agriculture Organization (FAO) of the United Nations exploits semantic web technology using Linked Open Data (LOD) principles to make statistical classifications accessible, machine readable, and interpretable. Moreover, this technology paves the way for the use of artificial intelligence in statistics and the interoperability between concepts and definitions (i.e., metadata) in statistical ontologies and statistical data bases. Stakeholders benefit from a common infrastructure for modelling, querying, and editing, while keeping control on their data and editorial workflow. Aria, the Enterprise Metadata Management Service used by Statistics Canada, also applies semantic web technology and provides reference data (including classification metadata) as a service, including the ability for custodians and machines to search, update, control versions, etc., using the underlying ontologies.

Roles and Efforts of the ASEAN Community Statistical System on the Harmonisation of ASEAN Statistics

4. The ASEAN statistical community, comprised of the national statistical systems of ASEAN members, ASEAN Community Councils, and the ASEAN Secretariat have developed a strategic plan for 2016-2025 which aims at providing high quality statistics based on a harmonized classification system. The effort faces many challenges, such as capacity gap between members, but it also greatly benefits from support from development partners and international organizations.

High Level Panel on Key Issues and Practices for Fixed Five-Year Revision Cycle for Statistical Classifications

5. The third webinar of the International Classifications Sprint convened a high-level panel to deliberate on the key issues and practices for the periodic five-year revisions of statistical classifications. The panel consisted of Karin Orvis from the Office of Budget Management, United States; Eric Rancourt from Statistics Canada; Jean Pierre Poncelet from Eurostat; and Stefan Schweinfest from the United Nations Statistics Division. They presented the complexities and challenges inherent in the implementation and maintenance of statistical classifications, as well as the implications for users, highlighting collectively the

significance of relevance and public utility of official statistics and their application to international statistical classifications.

6. A critical theme that emerged from the discussion was the imperative of ensuring classifications remain contemporary and pertinent through a dynamic updating process, balancing the need for continuity and comparability over time. Although some panelists supported a fixed revision cycle of 5 years, the panelist representing the Eiuropean statistical system expressed his opposition to this proposal due to capacity constraints. The panelists highlighted the challenges related to the implementation of new classifications in the context of statistical data production as well as the essential role of maintaining skilled labor capacity to facilitate on-demand adjustments to the classification system, a capability not uniformly available across organizations.

7. The dialogue underscored the necessity of adequate resourcing for the development and implementation of statistical classifications, and advocated for international support to bolster national capacities. Moreover, the importance of fostering collaboration and cooperation across statistical organizations and with various stakeholders, including government agencies and the private sector, was emphasized. This collaborative spirit should be cultivated within a regulatory, rather than a legislative framework to ensure flexibility and adaptability. The panelists pointed to the success of the North American collaboration between Canada, Mexico, and United States on the NAICS/NAPCS.

8. The unanimous view among the panelists was that investment in statistical classification work is not only beneficial but critical for their organizations and the countries they represent. They stressed the pivotal role of statistical classifications as essential infrastructure necessary for generating and integrating official statistics. Furthermore, they urged for concerted efforts and dedicated resources to ensure these classifications accurately mirror economic realities and aid in effective policymaking.

9. Looking ahead, the panel outlined several strategic directions to enhance the acceptance and implementation of a standardized five-year revision cycle for the ISIC and CPC classifications, emphasizing technological innovation, national, regional and international collaboration, and dedicated resources as key to

advancing statistical accuracy and utility in support of today's fast moving policymaking environment.

- leverage advancements in modern technology, especially semantic web technologies adhering to internationally agreed statistical ontologies and principles, such as the Simple Knowledge Organization System (SKOS) advocated by the World Wide Web Consortium (W3C)[1].
- map out and evaluate existing semantic web applications for international classifications across various platforms, including those by Eurostat, FAO, UNSD, and countries like Canada, France, and New Zealand.
- continue the exchange of institutional and technical best practices for the regular update of national and regional statistical classifications.
- incorporate the routine update of classifications and the utilization of semantic web technology into the upcoming global capacity-building program for ISIC and CPC.

---

[1] *Making the implementation of classification more efficient is a fundamental objective, and technology plays an increasingly important role in this. Eurostat, the statistical office of the European Union, is therefore organising a webinar, aiming to exchange best practices, methods as well as to identify the issues, challenges, strategies, and innovative methods for re-/coding the units in the statistical business registers with the new version of the industry classifications.*

The webinar will take place on 29-30 April 2024, and Eurostat would be delighted to extend the invitation to statistical offices and organisations outside Europe, and give them the opportunity to present their systems and coding strategies in business registers and in other statistical domains as well as learn from the experiences of others.

For additional information please do not hesitate in contacting claude.macchi@ec.europa.eu.