

Process and approach of updating the Classification of Statistical Activities (CSA)

I. Background

1. The Classification of International Statistical Activities was established in 2005 by the Bureau of the Conference of European Statisticians (CES). The CES Bureau has been the owner of the Classification and responsible for its maintenance and authorisation of revisions as necessary. The CSA started as a regional classification for a specific use (Database of International Statistical Activities maintained by UNECE). Over time, its use has widened to international organizations and countries from different regions. In February 2020, the CES Bureau approved the procedure for updating the Classification to include new, emerging areas of international statistical work.

II. The updating process

2. The practical work on the update was undertaken by a small Task Team including representatives from Canada, Ireland, Mexico (Chair), New Zealand, Eurostat, UNSD and UNECE¹.

3. The Task Team conducted a targeted survey in end 2020 among international organizations to get more information about how CSA was used, and identify the needs for updates. Fourteen organizations responded that they were using CSA. Several users had adapted CSA to fit their needs, either by selecting a subset of the classification or by adding more elements.

4. The user feedback also provided suggestions for updating the CSA, such as:

- Make environment statistics a separate domain;
- Add new statistical areas, such as: governance statistics, climate-change related statistics, geospatial statistics, digitalisation, circular economy, data science data exchange and data sharing, etc.;
- Regroup some statistical activities (e.g., make migration statistics a separate statistical area; merge macroeconomic statistics and accounts);
- Align with existing standards and models.

5. The Task Team developed a draft updated Classification (CSA 2.0) over two years and numerous teleconferences. The updated version reflected the consensus of the Task Team after analysing and incorporating user feedback (from a survey among international organizations) and the standpoints of the group members. General principles of the Classification were thoroughly discussed, such as mutual exclusivity and uses of the CSA.

6. The Task Team aimed to retain the classification's coherence and align it as much as possible with the existing statistical frameworks (e.g., the *Framework for Development of Environment Statistics* (FDES), *Praia City Group Handbook on Governance Statistics*, *Handbook on Management and Organization of National Statistical Systems*, *GSBPM*, *GAMSO*, etc.). Through the discussions, the Task Team consolidated views and came to a consensus that was reflected in CSA 2.0.

¹ Members of the Task Team were: Franklin Assoumou Ndong (Canada), Ciara Cummins and Don Forde (in the final phase) (Ireland), Andrea Fernández Conde and Manuel Cuéllar (Co-Chairs, Mexico), Andrew Hancock (New Zealand), Márta Nagy-Rothengass, Martin Karlberg and Maurizio Capaccioli (Eurostat), Sabine Warschburger (in initial phase), Ilaria Di Matteo, Zhiyuan Qian, Ivo Havinga and Pedro Farinas (in final phase) (UNSD), Tiina Luige, Stela Derivolcov, Martijn Kind and Jonathan Gessendorfer (in final phase) (UNECE).

7. The CES Bureau agreed to make the CSA a global classification as many of the international organizations who are using it have a global coverage, and a number of countries from different regions are using the CSA.

8. In preparation for the Conference of European Statisticians' 2022 plenary session, the draft CSA 2.0 was electronically consulted with all countries and organizations in April-May 2022. Eighty-three countries and organizations replied to the consultation.

9. The electronic consultation showed that forty-six countries from different regions and fourteen international organizations are using (or planning to use) the CSA. There was general support among responding countries and organizations that CSA would become a global classification.

10. The large majority of responding countries and organizations supported the endorsement of the classification and gave positive feedback, while providing detailed comments to clarify some statistical areas and improve the explanatory notes. The Task Force continued work to take into account the comments from the electronic consultation as much as possible, without diverting too much from the proposal submitted to CES that was supported by a majority of countries. The Task Team also considered the general principles and best practices of statistical classifications, concerning for example the treatment of residuals, the coding of items in the classification, etc.

III. Main changes to the classification compared to the 2009 version

11. To take into account user feedback and developments in international statistical work, two subject-matter domains were added to the classification:

(a) *Environment statistics* (new domain 3) aligned with the Framework for Development of Environment Statistics (FDES, adopted by the UN Statistical Commission in 2013) and the components of the basic set of environment statistics (UNSD, 2018²). The statistical areas in this domain follow the structure of FDES;

(b) *Governance statistics* (new domain 4) aligned with the *Handbook on Governance Statistics*³ developed by the Praia Group and endorsed by the UN Statistical Commission in March 2020. The statistical areas in this domain follow the general structure agreed by the Praia Group.

12. As a consequence of adding two domains, the next ones are renumbered, bringing the total number of domains to seven.

13. The domain 5 is renamed to "*Cross-cutting statistics*" (in the CSA 2009 this was domain 3: "*Environment and multi-domain statistics*"). Domain 5 now covers statistical areas that require bringing together data across different domains to meet the data needs for policy agendas for development, such as SDGs, climate change, etc. It also includes policy-relevant topics that should be mainstreamed in different statistical areas, providing a 'lens' through which to view the other topics, such as gender and human rights. This is also in line with the increasing need for data on these topics beyond social or economic statistics.

14. The domain 6 is renamed to "Statistical infrastructure and methodology" (in the CSA 2009 this was Domain 4 "Methodology of data collection, processing, dissemination and analysis"). The Task Team was aiming to align it with the Generic Statistical Business Process Model (GSBPM) but found it not to be practical for the purposes that the classification is used. For the practical purposes, the Task Team considered the concept of 'common statistical infrastructure' from the *Handbook on Management and Organization of*

² <https://unstats.un.org/unsd/environment/fdes/FDES-2015-supporting-tools/FDES.pdf>;
<https://unstats.un.org/unsd/envstats/fdes/basicset.cshtml>

³ https://unstats.un.org/unsd/statcom/51st-session/documents/Handbook_on_GovernanceStatistics-Draft_for_global_consultation-E.pdf

*National Statistical Systems*⁴ (endorsed by UNSC in 2021) quite useful, as it covers explicitly metadata, classifications, and registers.

15. Domain 7 “*Strategic and managerial activities*” is closely aligned with the Generic Activity Model for Statistical Organizations (GAMSO), with some simplifications.

16. The coding of categories is updated to make it better suited for use, and residual categories are added (coded as 99).

17. The Task team also decided to make a list of issues for future consideration for inclusion in a possible new revision. These can be, for example, analysing the not elsewhere classified (n.e.c.) categories and the need to elevate some of the activities into a category of their own. A good example of such a case can be the category 60499 Data sources, n.e.c.

IV. Purpose and intended uses of the updated Classification of Statistical Activities

18. The Classification can be used for two main purposes:

- Based on the process related to a statistical activity, CSA can be used to classify statistical events, capacity building activities, training courses, working groups, publications or statistical standards, etc.;
- Based on the output of a statistical activity, CSA can be used to classify data and metadata (particularly domains 1 to 5 which are related to subject-matter activities).

19. The Classification is hierarchical and has three levels. The first level comprises “domains” which relate to the broad type of statistical activities. The second level specifies “activities” within these domains. The third level covers a more detailed breakdown.

20. CSA can be used at different levels of detail driven by the user needs and purposes. Users can limit to the 3-digit level without a further breakdown. Or users can create their own more detailed breakdown at 5-digit level for specific areas of interest.

21. The first five domains of the classification cover subject-matter activities. These activities may result in data outputs. As such, these domains can provide a backbone for alignment and standardization of topics/themes or sections on data dissemination websites of NSOs and international organizations. This is also suitable for the continuance of SDMX usage of the CSA and for classifying indicators and other data products. Domain 5 can be used to classify data collections such as SDG-indicator sets, gender relevant indicators, and more.

22. Domain 6 is used for classifying statistical activities related to infrastructure and methodology, and Domain 7 for strategic and managerial activities for official statistics.

23. The CSA can be used as a reference leading to better harmonisation in structuring information about activities of statistical organizations, dissemination websites, publications, etc. Its aim is not to prescribe an organizational structure of a statistical office or statistical production process, or data dissemination. Countries and organizations are welcome to adjust the Classification to their purposes and context.

24. The classification categories are not designed to be mutually exclusive. Since the CSA is an analytical classification (and not a statistical classifications), the criterion of mutual exclusivity is not strictly embedded in the classification due to the nature of the classification. In some cases, a statistical activity can be classified in categories and users can decide where to place it according to their specific needs. The explanatory notes are prepared to guide the user through exclusions and inclusions.

V. Next steps

⁴ See <https://unstats.un.org/capacity-development/handbook/index.cshtml>

25. The final draft CSA 2.0 is being submitted to the UN Committee of Experts on International Statistical Classifications (UN CEISC) at its meeting on 25-28 October 2022 in New York to seek its endorsement for the classification to become member of the International Family of Statistical Classifications and submit it to the UN Statistical Commission at its 53rd session in March 2023 for adoption.

26. It is also proposed that the custodianship of the CSA is transferred to UNSD from the CES Bureau and UNECE.