



**Committee for the Coordination of Statistical Activities**  
**Conference on Data Quality for International Organizations**  
**Newport, Wales, United Kingdom, 27 - 28 April 2006**



**Session 2: Coherence of data published by various international organizations**

**JOINT UNITED NATIONS/OECD SYSTEM FOR THE COLLECTION AND PROCESSING OF  
INTERNATIONAL MERCHANDISE TRADE STATISTICS**

*Mary Jane Holupka<sup>1</sup>, United Nations Statistics Division*

Paper to be presented at the Conference on Data Quality for International Organizations,  
Newport, Wales, UK, 27-28 April 2006

**I. Introduction**

1. Significant progress in improving cooperation in the area of international merchandise trade statistics continues to be made by the trade statisticians and information technology (IT) specialists of the Statistics Division of the United Nations (UNSD) and the Statistics Directorate of the Organisation for Economic Co-operation and Development (OECD) who are working on the UN/OECD Joint Trade Data Collection and Processing System (the Joint System for short). The Joint System is centred around: (i) a Memorandum of Understanding (MoU), signed by both organizations detailing the key arrangements for the collection and sharing of detailed annual commodity-by-partner trade data; (ii) the adoption of common trade data-processing standards; and (iii) the development of computer applications for use in data-processing and dissemination.

2. The main objectives of the two organizations' efforts are to reduce the national authorities' response burden, to synchronize and harmonize data-processing routines and to disseminate identical high quality detailed international merchandise trade data. The use of common standards and procedures by both UNSD and OECD will ensure that the organizations' databases are coherent and consistent. In addition to sending the right signal to data providers and users, this business model ensures the best use of scarce resources (both human and IT) in both organizations.

3. This paper gives a general description of the history of the UN/OECD project, the successes attained thus far, the achievements yet to come and the future plans and challenges facing this inter-agency cooperative venture.

**II. Background**

4. For many years OECD and UNSD went about their trade data collection activities separately, with UNSD requesting detailed annual trade data from all countries and OECD requesting detailed annual trade data (and also more aggregated monthly trade data) from the OECD member countries. Concerning annual data, the procedures followed by both organizations were basically the same. When the trade data were received from the national authorities, they would need to be processed into a standard format with

---

<sup>1</sup> *Material for this paper was provided by Ronald Jansen and Markie Muryawan of UNSD and Andreas Lindner, Gregory Legoff and Lynda Howe of OECD.*

consistent coding and valuation. The values received in the different national currencies of the countries would be converted into US dollars using exchange rates supplied either by the countries or compiled by UNSD. Quantities would be converted into metric units whenever possible. The data each organization processed would then be stored in their own database system.

5. As they each compiled, processed and published their data independently, from time to time discrepancies would occur in the published statistics. Whereas valid reasons for differences in trade data may exist, data users would have more confidence in the quality of the data if both organizations published coherent and consistent data.

6. In the early 1990s, the trade statisticians in OECD and UNSD decided to try to find out why there were differences and began comparing notes on their collection and processing systems. Gradually the first exchanges of basic data started, which eventually led to the confirmation of the suspicion that the data received from some countries by UNSD were not necessarily always identical to the data received from the same countries by OECD. Often the reason for the differences was simply a matter of timing; in other instances, it was discovered that discrepancies arose because of different valuation, different aggregations or different formats which when processed yielded different results. It then became clear that reconciling the data of OECD and UNSD could only be done if the same data sets were used and the same processing procedures were applied to those data.

7. Not surprisingly, OECD and UNSD were a bit hesitant at first to commit to joining forces to process the trade data, not only since they knew it could mean having to alter some of their long standing procedures in order to align them with their counterpart's practices, but also because of the very different nature and role of both organizations. Compromises would have to be made. Practices that would require compromise were the treatment of confidential data, special transactions, different country codes and conversion of commodity classifications. Gradually, between 1998 and 2001, progress towards greater cooperation was made and willingness to build and share the leadership in the field of international trade statistics grew. A fundamental factor in this respect was that the OECD had introduced, in 1999, annual *International Trade Statistics Expert Meetings* at OECD Headquarters which brought together the trade data providers from OECD countries, trade expert data users from OECD countries and all international organizations active in this field. This yearly - and unique- international forum clearly allowed the gathering of the necessary support from countries and other international organizations, and the move from ad-hoc to more formalized consultations leading to the joint system.

### **III. Data collection: the first step towards formal cooperation**

8. In 2001, the first significant step in improving cooperation in the area of international merchandise trade statistics was made when the statisticians of UNSD and OECD signed a Memorandum of Understanding by which it was agreed that OECD would be responsible for collecting the trade data and metadata from its member states and UNSD would cease requesting the same information from them and collect from non-OECD countries only. OECD would request that the countries supply both provisional and final data, would monitor and analyse the timeliness of data and metadata reporting, and share whatever it received from the countries with UNSD so that UNSD could process and include the data in its UN Comtrade database. The requirements for the collected data from all countries would be in accordance with the recommendations published by the Statistics Division in *International Merchandise Trade Statistics: Concepts and Definitions*, Revision 2 (published in 1998).

9. Within the same time frame, new developments were taking place in the IT-related area of trade statistics. Given the very large amount of trade data to be collected, trade statisticians must always have a solid IT platform with which to process, store and disseminate their data. Both OECD and UNSD were looking for solutions which would allow them to move away from the mainframe to a server environment so that they would have more possibilities to make their data accessible to the user community.

10. OECD had made major technological improvements when it created its Oracle database in the mid-to-late nineties. Not so long after OECD, UNSD also began exploring new platforms for its trade database. OECD offered to share its experience and expertise, and around the year 2000, helped UNSD to fully test the OECD system. By 2001, Microsoft launched its SQL Server database which proved to be a more powerful and easier-to-maintain solution than the Oracle database, so UNSD chose to go in that direction instead. By 2003, the pendulum had swung back the other way, and OECD was then looking into a successor for its Oracle database. This time UNSD was in a position to offer its help and to share its technology with OECD.

#### **IV. The development of the UN/OECD Joint System**

11. From 2002 on, OECD and UNSD engaged in a multi-year process to share resources to develop a new data-processing system together which would be based on the same statistical standards and on the same information technology. Building upon the 2001 Memorandum of Understanding and further extensions, additional agreements were made on common trade data-processing standards and on computer applications for use in processing and disseminating trade data. In addition to the data collection and sharing agreement, UNSD and OECD jointly developed common trade data-processing standards and computer applications for use in the processing of the trade data. The set of agreements, standards and applications make up the Joint System. Currently, both organizations are implementing the standardized data-processing procedures to ensure that the data processed by both of them are identical. Upon completion of this activity, the trade data of OECD member countries will be processed by OECD only and will be included in the UNSD database simultaneously. Likewise, the trade data processed by UNSD for the non-OECD countries will be included in the OECD database.

##### **A. Summary of common data-processing standards**

12. UNSD and OECD have undergone a very detailed review of their data-processing procedures and agreed on common data-processing standards. The common strategies developed for the treatment of country codes, trade conversion factors, classifications, non-standard codes/memorandum items, quantity units, confidential data, management of metadata and synchronization of databases are briefly described below.

##### **Country codes**

13. The country codes issue has been discussed at length between the two organizations. OECD and UNSD agreed to maintain the most detailed level of country codes for data processing and to use the aggregated countries/territories to match the different needs of both organizations.

##### **Trade conversion factors**

14. The trade-weighted exchange rates (trade conversion factors), calculated separately for exports and imports, will continue to be applied by both organizations. The procedure has been developed to check automatically the significance of the revisions of exchange rates on a quarterly basis. The recalculation of the value data will be carried out simultaneously in both databases, if deemed necessary. A tracking system for recording historical conversion factors will be kept for reference.

##### **Classifications and correlation tables**

15. Both databases will continue to treat the Harmonized Commodity Description and Coding System (HS) as a primary commodity classification for data-collection and data-processing purposes. The Standard International Trade Classification (SITC) will be used to aggregate the HS data into commodity groups more suitable for some analytical purposes and for maintenance of the historical series. The correlation tables between various existing revisions of these classifications, as developed by UNSD, will be applied by both organizations. Depending on their needs, UNSD and OECD may utilize other classifications for dissemination of trade data. For example, OECD has a history of expressing trade data also in terms of the International Standard Industrial Classification of All Economic Activities (ISIC) and UNSD introduces trade data in terms of the Classification by Broad Economic Categories (BEC). In

response to strong demand from analysts, OECD has developed and implemented in March 2006 a slightly modified ISIC conversion routine for general use. Both UNSD and OECD are in agreement that OECD should refer to UNSD for BEC-converted trade data and UNSD to OECD for ISIC-converted trade data. In doing so, no duplication can occur for these “extra” conversions.

#### **Non-standard codes**

16. The organizations agreed that there was a need for non-standard HS codes to process special transactions or special commodities reported by countries that could not be allocated to standard HS or SITC codes. The lists of such codes (or memorandum items) were reviewed and agreed upon. To ensure the best possible treatment of detailed trade data, OECD will continue using a much more detailed list of memorandum items than UNSD. A correspondence table between OECD and UNSD non-standard codes was created and will be used when necessary. Even if non-standard codes are used differently, the system ensures that identical data are produced for standard commodity codes.

#### **Quantity units and quantity estimations**

17. The databases of both organizations will contain two quantity fields: net weight and quantity in terms of the standard units of quantity recommended by the World Customs Organization. Since reporting of quantity information by countries has many gaps, both organizations use estimates. In order to improve the quality and comparability of such estimates it was agreed to standardize the estimation procedure using the new approach developed by UNSD. Quantity conversion factors will be reviewed and agreed upon with the understanding that the factors provided by the Food and Agriculture Organization of the United Nations (FAO) for agricultural products will be taken to the maximum extent possible. It was also decided that quantities would be shown only at the 4- and 6-digit levels of HS and at the 3-, 4- and 5-digit levels of SITC.

#### **Confidentiality**

18. Owing to the confidential character of some transactions, countries provide less detailed partner breakdown in the national data sets at the HS 6-digit level rather than at the HS 2-digit level. This creates discrepancies between totals shown at the 2-digit level and the accumulated totals of data reported at the lower levels of the classification. OECD and UNSD have agreed to include in their databases a general methodological note explaining those apparent discrepancies to users of trade statistics and are finalizing rules on the uniform processing of the confidential data.

#### **Metadata**

19. The organizations treat metadata as an indispensable part of their respective databases and agreed to improve the efficiency of trade metadata preparation, storage, access, management and dissemination. OECD is loading trade metadata into its recently developed MetaStore application. This includes trade definitions, recommendations, sources, country notes, information about classifications, etc. As a common repository, MetaStore manages metadata for data sets of various subject areas and different structures. UNSD is studying the possibility of implementing applications similar to MetaStore as a management tool for its metadata. UNSD and OECD will ensure that users have no difficulty in accessing the metadata maintained by both organizations.

#### **Synchronization of databases**

20. In order that identical data may be disseminated to users of trade statistics, both UNSD and OECD databases will be permanently synchronized. The synchronization process will be run every night to update OECD and UNSD mirror databases. The data will be transferred in Extensible Markup Language (XML) format using Statistical Data and Metadata Exchange (SDMX) cross-sectional schema.

### **B. Computer applications for use in data processing**

21. As part of the Joint System, all computer applications with associated source code that have been or will be developed for the processing and dissemination of trade data are to be shared between the two organizations. The starting point for this development was the creation of the UNSD application for maintenance and dissemination of detailed commodity trade statistics, namely, the United Nations

Commodity Trade Statistics Database (UN Comtrade). OECD has installed this application and adapted it to its needs, with a view to making it accessible to the public by mid-2006.

22. The backbone of the whole operation is the trade data-processing application named COPRA (Commodity Processing application). The effort undertaken in developing COPRA has been a one-of-a-kind example of inter-agency collaboration. On the one hand, statisticians of UNSD and OECD worked out in detail the methodology for standards of data processing (including country nomenclature, treatment of confidentiality, estimation of missing quantities and data validation); on the other hand, IT staff of both organizations wrote code and developed application interfaces to create the computer programs.

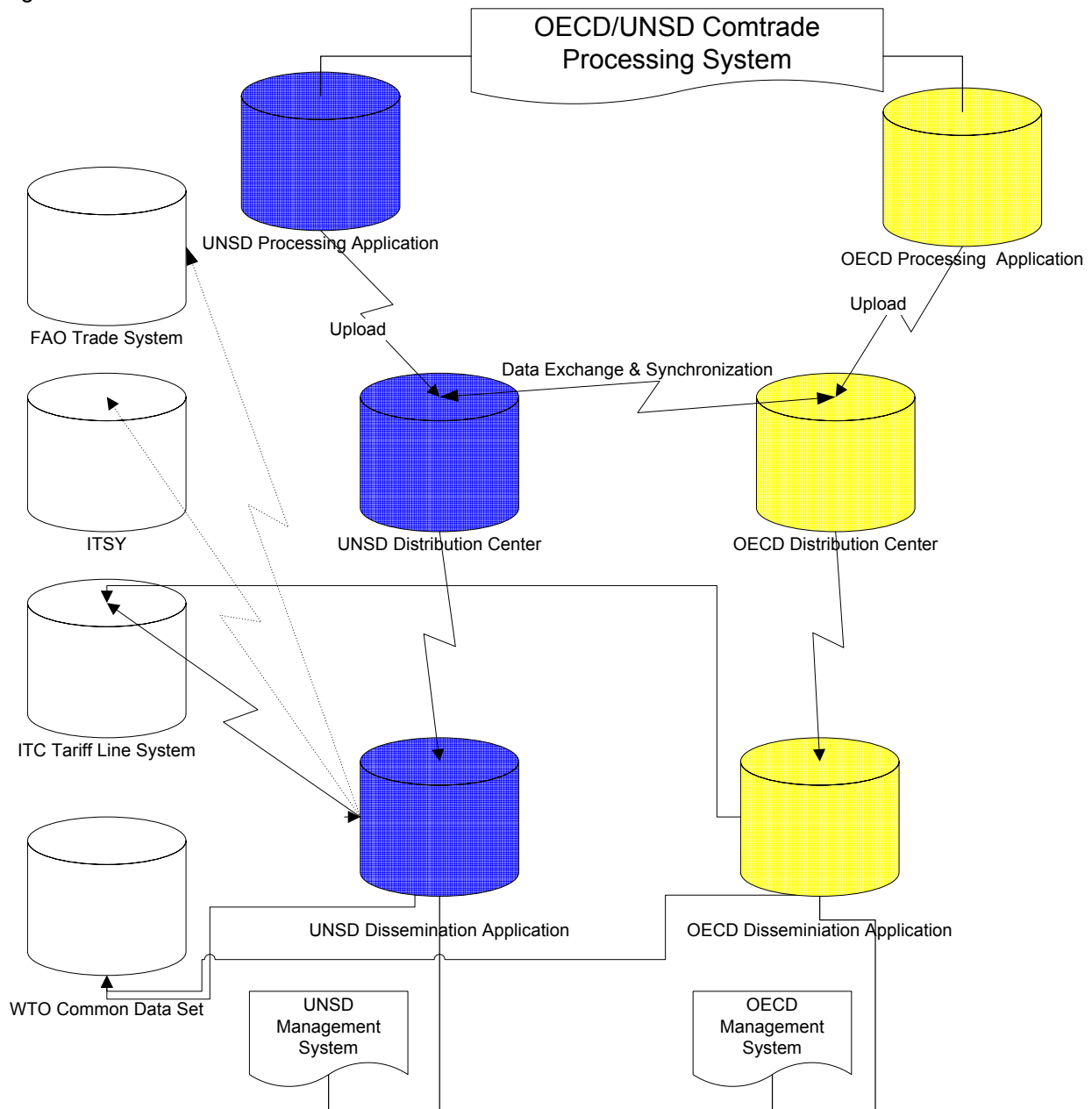
23. On 1 January 2006, COPRA became the production system for the processing of trade data at UNSD and OECD. The perceived benefits of the new system are: (a) inter-agency standards on all details of data processing; (b) harmonization of published international trade statistics; (c) higher level of data quality; (d) full accountability in the treatment of trade data; and (e) availability of a complete set of metadata. In addition, as an integral part of the application, an automated transfer of the processed data at a detailed level will feed into the trade databases of the World Bank and FAO and into the joint World Trade Organization/United Nations Conference on Trade and Development (UNCTAD)/ International Trade Centre (ITC) database. This should lead to a very close alignment of the detailed trade statistics among international organizations.

### **C. Data dissemination**

24. Although synchronized, both organizations will continue to keep separate (institutionally and legally) international merchandise trade statistics databases to ensure the rights and ability of each organization to pursue its own policies with regard to: (i) storing various data in addition to those agreed in the Joint System (for example, historical data, memorandum items, derived data, etc.); (ii) developing additional analytical and presentational functionalities and controlling user access to them; and (iii) dissemination practices, including pricing of user access to the database services, entering into contracts with users, etc. Each organization will maintain its own brand name for use in data dissemination activities. The current names of those databases are UN Comtrade and OECD ITCS Database (ITCS stands for International Trade by Commodity Statistics).

25. To ensure user awareness of the joint efforts of both organizations in the area of international merchandise trade statistics, the UN Comtrade disclaimer will contain the following statement: "The current trade data of OECD member countries are compiled by OECD in accordance with the UN/OECD Common Trade Data-Processing Standards and are available in UN Comtrade in accordance with the UN/OECD Data Sharing Agreement. Any deviation from this practice is reflected in the country/period metadata". Similarly, the OECD ITCS trade database disclaimer will contain this statement: "The current trade data of non-OECD member countries are compiled by UNSD in accordance with the UN/OECD Common Trade Data-Processing Standards and are available in UN Comtrade in accordance with the UN/OECD Data Sharing Agreement. Any deviation from this practice is reflected in the country/period metadata."

Figure 1



## V. The current situation

26. Beginning with the data for the year 2002, OECD, and no longer UNSD as well, collected international merchandise trade data from the OECD member countries, thereby succeeding in reducing the countries' response burden. As of 1 January 2006, the UN/OECD data-processing system has been operational. It is now in the testing and fine-tuning stage. OECD and UNSD will truly publish the same trade statistics for all OECD member states since both data collection and processing have been fully aligned. The

biggest gains of the UN/OECD Joint System are higher quality trade statistics through an improved data-processing system certified by UNSD and OECD trade statisticians, and a higher quality dissemination system maintained by the combined team of UNSD and OECD IT experts.

27. The OECD web client is still undergoing testing, but once completed, will have the same functionalities, look and feel as Comtrade. It is expected to “go live” before summer 2006. It will contain both the OECD and UN logos to make this unique inter-agency co-operation very visible.

28. An amended memorandum of understanding on the cooperation between UNSD and OECD is scheduled to be completed and signed after the Joint System has been officially taken into production and reviewed by countries. It will take stock of the latest achievements and also set out “governance” rules to be respected by both organizations to ensure smooth and efficient operations. In practice, the UN/OECD Joint System means that the trade statistics experts of both organizations are constantly communicating with each other, analyzing and evaluating the scope for improvements and also finding solutions to problems of data quality and processing. Their IT experts are regularly crossing the Atlantic, too, to actively work on the data-processing system. Fortunately, there is support at all levels in both organizations to make this Joint System a success.

## **VI. Plans for further cooperation**

29. Achievements made up to now make this cooperation already a big success, quite well known in various fora. The actors are committed to continuing to follow the same course in order to ensure consistency in their future actions. Coherence and consistency in approach so far would benefit from also looking into the following possible areas of future cooperation, namely (i) development of a common strategy for maintenance and dissemination of the tariff-line data; (ii) evaluation of the possibility of collecting, processing and disseminating trade data containing additional features (for example, mode of transport, nature of transaction, additional partner country attributions, etc.); (iii) further standardization and enhancement of the metadata; and (iv) evaluation of the experience gained for use in similar activities in statistics of international trade in services.