

Distr.
GENERAL

Working Paper No.7
6 May 2009

ENGLISH ONLY

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES (EUROSTAT)**

**ORGANISATION FOR ECONOMIC COOPERATION
AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

Meeting on the Management of Statistical Information Systems (MSIS 2009)
(Oslo, Norway, 18-20 May 2009)

Topic (i): Sharing of software and components

SHARING STATISTICAL SOFTWARE COMPONENTS – THE OECD EXPERIENCE

Invited Paper

Prepared by Trevor Fletcher, OECD

I. INTRODUCTION

1. The OECD has been sharing Statistical Information System (SIS) software components with a number of other International Organisations and National Agencies since 2003. The various SIS components manage the full life-cycle of statistical data and metadata processing from data capture through dissemination.
2. This paper gives some background on how this sharing of statistical software came about, explains the advantages and challenges of the process and describes the OECD experience thus far in collaborating with other bodies, and looks towards the possible next steps in extending the scope of this exercise.

II. BACKGROUND

3. The OECD began sharing its Statistical Information System software once development work for the initial releases had been completed in 2003. The idea to share followed on from the successful exercise in designing, building and implementing a joint Trade system in collaboration with the UNSD from which it became clear that such joint ventures could work effectively and the advantages of such an approach became apparent. Consequently the OECD made other organisations aware that SIS developments were available for them to assess as opportunities arose during various meetings, seminars and other forums.
4. Initially the sharing of software was managed on an informal basis in that the partner organisation would be sent the application and source code and provided with some basic support and technical documentation to allow them to get started. This included an agreement that if the partner subsequently developed any additional useful features then these would be made available back to the OECD (and other interested parties) to be shared. This could loosely be considered a form of open source.
5. The OECD has since received numerous additional requests from other International Organisations, National Statistical Agencies, Universities and private companies to use the OECD Statistical software so it was

decided in 2008 to adopt a more formal approach based on Memoranda of Understanding with collaborating organisations and to include an element of cost recovery.

III. THE ADVANTAGES OF SHARING SOFTWARE COMPONENTS

6. There are a number of very obvious reasons for organisations to share statistical, or indeed any, software. The most obvious of these reasons is the savings in time and money in re-using or adapting an existing technical solution to meet another organisation's processing requirements. In addition, any new features subsequently developed by one of the collaborating parties can then be shared and re-used by the other partner(s) at minimal cost.

7. The use of common shared systems can promote and enable the use of statistical standards, both in structure and content. As an example, SIS components widely use the Statistical Data and Metadata Exchange (SDMX) standard as inputs and outputs.

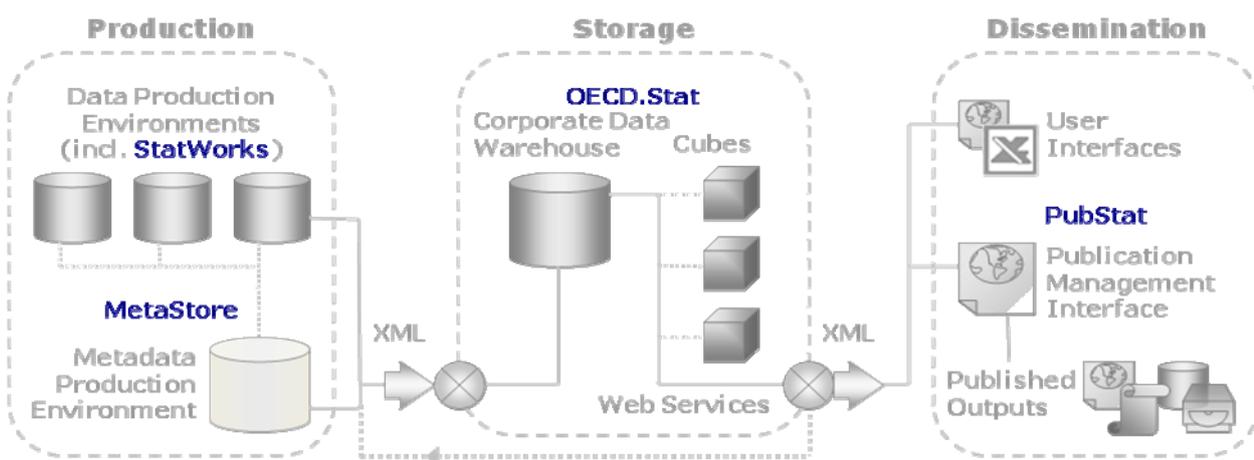
8. Using the same software platform for statistical data management makes the exchange of data and metadata easier between organisations and can facilitate the production of joint data collection or dissemination exercises between organisations.

9. Collaboration between organisations in this way sends a very positive message to stakeholders that we, as publically funded bodies, are making our best efforts to efficiently use resources by working together and not duplicating effort in re-inventing the wheel.

10. Working with other organisations has proved to be rewarding for staff in providing interesting and challenging work and also increased mobility.

IV. WHAT SOFTWARE IS BEING SHARED BY THE OECD?

11. The three main components of the OECD Statistical Information System (StatWorks, MetaStore and OECD.Stat) that manage statistical data and metadata Production, Storage and Dissemination processes are the application software most commonly shared with other organisations. The following diagram demonstrates how each component fits into the overall OECD Statistical Information System architecture:



12. The *StatWorks* software provides a common, SQL-based repository for statistical quantitative and qualitative data together with a set of tools for data management that covers Data Importing, Validation, Calculations, Querying and Exporting. The application is designed to be fully integrated with other components in the OECD's Statistical Information System (SIS). *StatWorks* is currently being used as a production tool by UN-ESCAP and is being evaluated by the World Health Organization, UNESCO Institute for Statistics and Statistics New Zealand.

13. *MetaStore* is a general toolkit with a web-based interface for accessing and managing structural and reference metadata for statistics. Structural Metadata describes the multi-dimensional structure of a dataset and includes codes and names of dimensions and corresponding dimension members. Reference Metadata consists of textual information documenting characteristics of the data within a dataset and typically includes information concerning the collection, manipulation, purpose and quality of data at different levels of a dataset's multi-dimensional structure. *MetaStore* is being used or evaluated by the following organisations: UNESCO Institute of Statistics, UNSD, IMF, World Bank, FAO, European Central Bank, CBS (Netherlands Statistics Agency), Australian Bureau of Statistics, Statistics New Zealand, UK Statistics Office and the World Health Organisation.

14. *OECD.Stat* is the corporate statistics data-sharing and dissemination environment of the OECD. Its data warehouse provides a harmonised and platform-independent access to shared or publishable data from all sectors of OECD activity and from other international organisations. It is fully integrated with the other components of the OECD's Statistical Information System for managing production data and metadata, analytical software packages, and for the dissemination of Statistics via the Web, Paper and Electronic Data Products. *OECD.Stat* is being used or evaluated by the following organizations: IMF, Australian Bureau of Statistics, Statistics New Zealand, UK Statistics Office and the World Health Organisation.

V. THE OECD EXPERIENCE IN SOFTWARE SHARING – THE PROS AND CONS

15. Despite the clear advantages of having organisations collaborate on projects to share software, such joint ventures can result in the increased complexity of project management depending on the role of the various partner organizations. This can vary from case to case.

16. For the OECD the most straightforward example of a collaboration was with the UNSD in developing the Joint Annual Trade System (although a part of the overall OECD Statistical Information System this application is specific to Trade data processing so has not been generally shared with others). In this case the timing of the project allowed both organizations to have a common start date and to agree an overall work plan that defined a clear split in tasks on each side. This allowed development work to proceed in parallel with little or no overlap and the various pieces put into place with very little additional management overhead.

17. In the case of StatWorks and MetaStore, the OECD developed the products which were then made available as finished articles to other organizations who wished to use them, providing upgrades and new releases as required. This has proved to be fairly easy to manage as only one side (the OECD) is doing the development work at this stage and has a therefore low additional management overhead. This is a case of software sharing, but is not collaborative in the strictest sense as all development is done by only one party.

18. Sharing of the OECD.Stat Data warehouse system has proved to be the most complex collaborative scenario. In the first case, we agreed to work in partnership with the IMF on further developing and improving the product together. As this is a complex system that had not been originally designed to be used or deployed outside the OECD, a significant investment was required in time and effort for both the OECD and IMF in transferring the necessary expertise for development staff. In addition there was a considerable management overhead in agreeing a Memorandum of Understanding and the workload of running a joint IT project across two organizations. Following this experience we agreed that no additional development partners could be added as this would increase the complexity of managing the project significantly with the addition of new requirements, agreeing new priorities and resourcing etc. Despite the complexity of this collaboration process, there were also advantages in that OECD did benefit from a number of developments made by the IMF to improve the system.

19. OECD is currently reviewing how subsequent ongoing partnerships with other organizations can be managed and are exploring the different options ranging from the most complex (consortium of development partners) to the most simple (single agency coordinating development activity and distributing code) along with all the options in between.

VI. CHALLENGES

20. There are a number of challenges involved in finding the best way to move forward in sharing software and advance towards a Global Integrated Statistical Information System. These include:

- Finding the right scenario for managing shared software from a centrally managed system through to a consortium approach
- Choosing between centrally managed 'owned' software or open source
- Finding the right balance on cost recovery
- Deciding which software components are the most suitable candidates for the Global system

21. The OECD seeks the inputs and experiences of other national and international organisations to help determine what should be the building blocks of a coherent global system, to agree governance rules for managing such a system and to provide the political mandate to promote and enable the basis for such an international standard that could be adopted by those wishing to participate and share this system in the future.

VII. NEXT STEPS

22. The OECD is engaged in a number of next steps related to this exercise. These include: drafting a standard Memorandum of Understanding to cover collaboration with any organisation; the completion of software installation packages to simplify the evaluation process; and to continue to promote evaluation studies with other organisations.

VIII. CONCLUSIONS

23. Based on our experiences sharing statistical software with other organisations, OECD firmly believes this to be a good thing for reasons of cost savings and promotion of international standards. The OECD considers the SIS we have developed could be one candidate for such an international standard as it covers the full life cycle of the statistical process. Despite the complexities of collaborating on projects and sharing software solutions, the advantages strongly outweigh any disadvantages and we feel that this approach should continue to evolve and be adopted and promoted more widely.