

Distr.
GENERAL

CES/SEM.47/23 (Summary)
14 January 2002

ENGLISH
Original: ENGLISH and FRENCH

**STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

**COMMISSION OF THE
EUROPEAN COMMUNITIES**

CONFERENCE OF EUROPEAN STATISTICIANS

EUROSTAT

**Joint UNECE/Eurostat Seminar on Integrated Statistical
Information Systems and Related Matters (ISIS 2002)**

(17-19 April 2002, Geneva, Switzerland)

Topic IV: Ways of making statistical information systems more responsive to users

METADATA AT STATISTICS CANADA

Invited paper

Submitted by Statistics Canada¹

Summary

1. From the 1970s to the 1990s, there were several initiatives to build systems to provide a means for collection and management of meta information and metadata at Statistics Canada. The information in these stand-alone systems was inconsistent from system to system because each system collected data separately from the program areas at different times during the year.
2. In 1998, the Integrated Meta Database Project was started to create a corporate wide *metadata* repository. The project's scope was to identify *metadata* requirements and to consolidate the various sources of meta information and metadata into one central repository for all statistical activities of the corporation. The goal was to minimize duplication of information among the various systems, to provide better access to information to both internal and external clients, to eliminate the response burden by the subject matter areas and to assist in harmonization of concepts and variables.
3. The Integrated Meta Database Project's concentration since 1998 to the present has been in the development of *metadata* and loading into a central repository, the Integrated Meta Database (IMDB). The IMDB is a repository of information about all surveys conducted by Statistics Canada. It is an inventory of the information assets of the department and serves as a source of information about Statistics Canada surveys that the public and researchers require to evaluate the content and quality of statistical information.
4. Existing dissemination access paths were used to provide users access to the IMDB data. HTML pages are generated from the IMDB daily. The pages are accessible from Statistics Canada's web site via CANSIM, Canadian Statistics and Statistical Methods. The effort has been directed in the collection of the metadata as opposed to investigating new access paths because the underlying metadata is required before access paths are useful.

¹ Prepared by Amie Lee.

5. The collection and update activities are enormous tasks that require a major effort and commitment by the corporation. Object-oriented technologies, the intranet and XML were used to assist in these activities in attempt to streamline the tasks for the IMDB project. Overall this has been a positive but challenging experience for the IMDB project.

6. The challenge now is to keep the IMDB up to date, both in terms of the data it contains and the technologies it uses (having embraced an object-oriented technology, the challenge is to avoid becoming obsolete as the object-oriented technologies are changing rapidly in the industry.) Directions for the future include, moving from a passive to a more active collection and use of *metadata* and providing users with improved access paths for both internal and external users to the *metadata* stored in the IMDB.