
Data Warehouse

Definition

The Data Warehouse contains the datasets used for deriving the stable indicators in the regular environmental reports. The datasets contain aggregated data, typically on a country level, which are time series with a regular update cycle and a geographical coverage of at least EU15. The aggregated data are extracted from the databases in the European layer of the Reporting System, as far as possible by (semi)automated procedures using the EIONET Telematics Network. The data in the Data Warehouse are made publicly available, if there are no "use constraints" specified by the data owner. "

The Data Warehouse in the EEA Work Programme 1999

4.2.3 Data Warehouse management: This project supports capturing and management of indicator data for the periodical reporting and harmonisation of environmental data across topics. (In 1999, the Data Warehouse will be further developed to support capturing, management and presentation of data for the stable indicators of the first version of the regular indicator-based report: Environmental Signals Report.

The tasks to be executed under the 1999 Work Programme are to develop and maintain a directory of data sources for the stable indicators and data descriptions of the variables to be included in Warehouse (metadata). Data capturing routines will be developed for stable data sources at ETC, Eurostat, and other international data providers. Development of procedures for administrating the indicator data, and the integration of relevant GIS data, will be an important task in 1999. Finally, development of Web based user interfaces to support the compilation of the regular indicator-based reports, will be completed in 1999

4.3.3 Public Access to Data Warehouse: At the end of the first Multiannual Work Programme, the users of the EEA Web site can access information maintained in several databases. These are for example the Data Warehouse, including data used for the Second Assessment report published in 1998, and the databases developed and maintained by the various Topic Centres. Each of these has its own user interface. This project aims at better serving the EEA users by making more of the information maintained in databases publicly available through a seamless service. The purpose is also to make a better use of limited resources by avoiding duplication of efforts. The main tasks in 1999 will be to further developing the Public Web interface to the Data Warehouse, in order to provide public access to the reporting data.

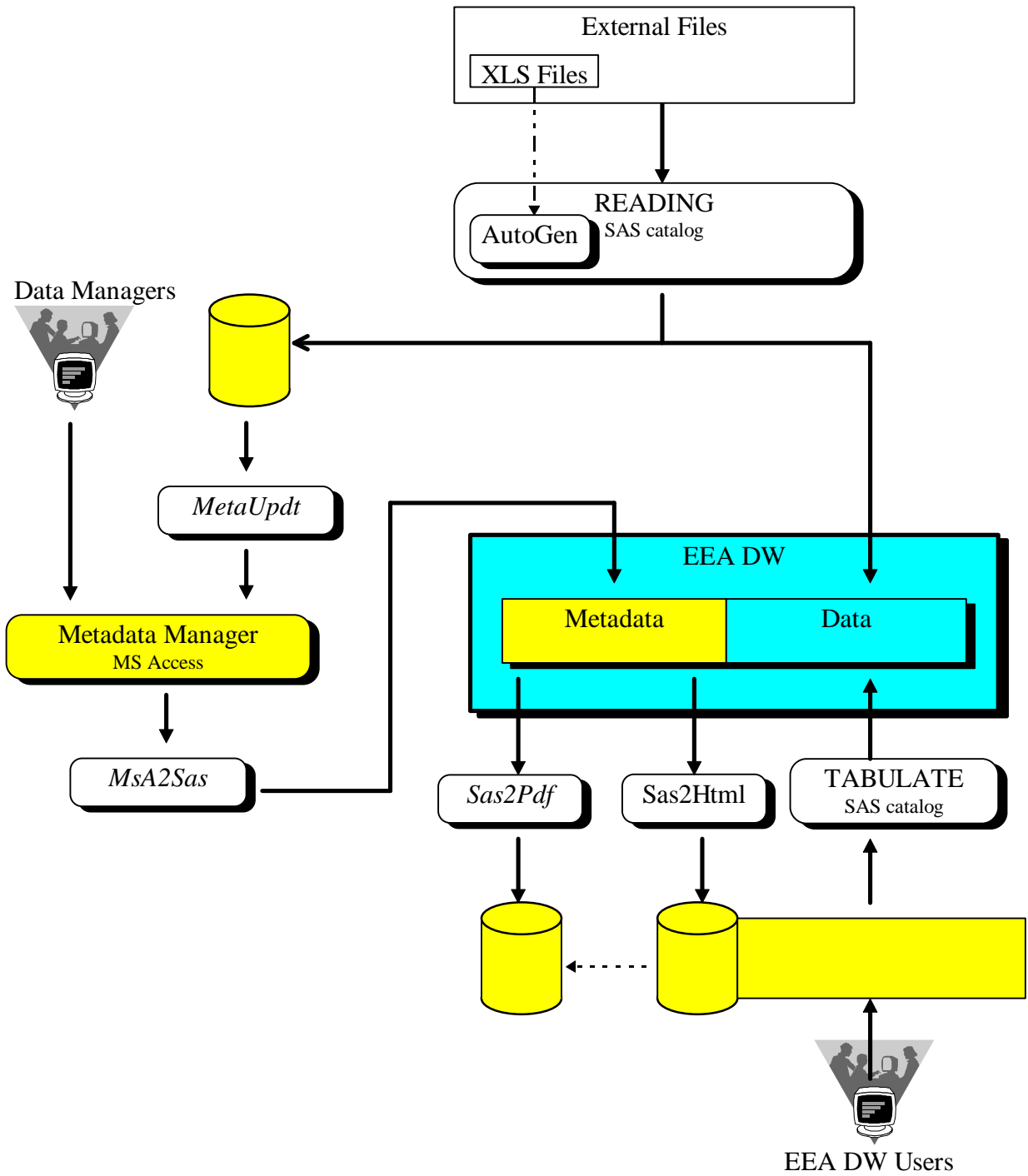
Technical description

The Data Warehouse is developed using the SAS software. The Warehouse system consists of:

- a number of SAS datasets,
- an MS Access database with metadata for all datasets,
- SAS routines (programmes) to create and update some of these datasets from sources
- An MS Access application to maintain the metadata
- SAS routines to present the data on Intranet/Extranet
- SAS routines to present some of these datasets on Public Internet.

The figure below shows the process flow and there after a list of selected content is presented.

DW process flow



DW Process Flow

Example of content

1 Socio-Economic Data

1.1 Population

1.1.1 Population by Country, Region and Subregion, EU 15 Countries (1st January)

Dataset name: D2AGE80R
Date creation/updt: 17/11/1998
Unit: 1000 persons
Geographical Cov.: EU 15 NUTS
EU 15 Countries
EU 15 Total
Temporal Cov.: 1980-1987
Source Type: CSV
Sources: NEW CRONOS, Eurostat (17/11/1998)
Footnotes: This data set should be used for regional analysis.

1.1.2 Population by Country and Age Group, EU 15 Countries (1st January)

Dataset name: D2AGE80S
Date creation/updt: 03/12/1998
Unit: 1000 persons
Geographical Cov.: EU 15 NUTS
EU 15 Countries
EU 15 Total
Temporal Cov.: 1980-1987
Source Type: CSV
Sources: NEW CRONOS, Eurostat (17/11/98)

1.1.3 Total Fertility Rate by Country

Dataset name: FIND
Date creation/updt: 27/04/1999
Unit: Average number of children/woman
Geographical Cov.: European Countries
EU 15, EFTA Totals
Temporal Cov.: 1950-1997
Source Type: CSV
Sources: NEW CRONOS, Eurostat (30/10/98)
Footnotes: Germany: Federal Republic of Germany (including east from 1991)

1.1.4 Total of Households by Country and Household Type, Census Round 1990-1991

Dataset name: HHEC21
Date creation/updt: 17/09/1998
Unit: Number of Households
Geographical Cov.: AT BE DE DK ES FR GR NL PT UK
Temporal Cov.: 1990-1991*
Source Type: CSV
Sources: NEW CRONOS, Eurostat (13/06/1998)
Footnotes: In the Netherlands the total may not correspond with the sum of the regions due to an extra administrative region (C.P.R.region).
Data for the United Kingdom are based on persons resident. Resident staff and relatives of resident staff have been excluded.

1.1.5 Total of Households by Country, Region, Subregion and Household Type, Census Round 1990-1991

Dataset name: HHEC22
Date creation/updt: 17/09/1998
Unit: Number of Households
Geographical Cov.: AT BE DK ES FR GR IT NL PT UK
Temporal Cov.: 1990-1991*
Source Type: CSV
Sources: NEW CRONOS, Eurostat (13/06/1998)
Footnotes: In the Netherlands the total may not correspond with the sum of the regions due to an extra administrative region (C.P.R.region).
Data for the United Kingdom are based on persons resident. Resident staff and relatives of resident staff have been excluded.

1.1.6 Number of Households and Persons by Country and Household Size, Census Round 1990-1991

Dataset name: HHPNPER
Date creation/updt: 17/09/1998
Unit: Number of Persons and Households
Geographical Cov.: EU 15, CH, LI, NO countries
Temporal Cov.: 1990-1991*
Source Type: CSV
Sources: NEW CRONOS, Eurostat (13/06/1998)
Footnotes: Data for Austria are estimations based on the micro-census.
In Finland children are under 18 years. A mother or father with all children aged 18 years or over are classified as multi-person households.
In Sweden children are under 16 years of age.
Data for the United Kingdom are based on a 10% sample of census returns.

1.1.7 Number of households by Country, 1990-1996

Dataset name: HOUSEH_D
Date creation/updt: 09/06/1999
Unit: 1000 households
Geographical Cov.: European countries
Temporal Cov.: 1980-1996
Source Type: Unknown
Sources: Euromonitor - European Marketing data and statistics 1997, 32nd edition + Statistical Compendium for the Dobris Assessment (table 26.1)
Footnotes: New series from 1990

1.1.8 Total of Households by Type of Urbanized Area

Dataset name: LF2HH____
Date creation/updt: 06/05/1999
Unit: Number of 1000 Households
Geographical Cov.: EU 15 NUTS
EU 15 Countries
EU 15 Total
Temporal Cov.: 1992-1994
Source Type: CSV
Sources: NEW CRONOS, Eurostat (19/11/98)
Footnotes: Germany: Federal Republic of Germany (including east from 1991)
Populated Area Type: Set of closely related local units
DENSE: Each one of which having a density greater than 500 inhabitants/km, and the total population of which being of at least 50 000 inhabitants.
INTERMEDIATE: That do not pertain to a densely populated area, each one of which having density greater than 100 inhabitants/km, and where the total population is at least of 50 000 inhabitants, or it refers to a set that is adjacent to a highly populated area.
THIN: That are not part of a densely populated area, nor of an intermediate area.

1.1.9 Infant Mortality Rate by Country

Dataset name: MINFIND
Date creation/updt: 29/04/1999
Unit: Deaths/thousand births
Geographical Cov.: European Countries
Temporal Cov.: 1950-1997
Source Type: CSV
Sources: NEW CRONOS, Eurostat (23/12/1998)

1.1.10 Life Expectancy by Country, Sex and Age

Dataset name: MLEXPEC
Date creation/updt: 14/09/1998
Unit: Years
Geographical Cov.: European Countries
Temporal Cov.: 1950-1996
Source Type: CSV
Sources: NEW CRONOS, Eurostat (13/08/1998)

1.1.11 Population by Country

Dataset name: POP_95_D
Date creation/updt: 17/09/1998
Unit: 1000 persons
Geographical Cov.: European Countries
Temporal Cov.: 1955-1995
Source Type: XLS
Sources: World population Prospects: The 1996 Revision (United Nations, New York)
Footnotes: This data set should be used for indicator calculations, e.g. : GDP per capita, ...
Data for the 1950 - 1995 period are interpolated United Nations population estimates. The data are interpolated from the quinquennial population estimates and projections published in the United Nations Population Division publication: World population Prospects: The 1996 Revision (United Nations, New York)

1.2 Resources Consumption

1.2.1 Consumption of Aluminium (Primary and secondary aluminium plus direct use of scrap, as far as reported)

Dataset name: CON_AL_D
Date creation/updt: 16/09/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition (From 1985 - 1995) + UNCTAD commodity yearbook, United Nations 1995 (for data from 1982 to 1985)
Footnotes: * Belgium + Luxembourg
** FR + DR

1.2.2 Consumption of Cadmium

Dataset name: CON_CA_D
Date creation/updt: 17/07/1998
Unit: Tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition
Footnotes: * Belgium + Luxembourg
** FR + DR

1.2.3 Consumption of Copper

Dataset name: CON_CO_D
Date creation/updt: 17/07/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition (From 1985 - 1995)
+ UNCTAD commodity yearbook, United Nations 1995 (for data from 1982 to 1985)
Footnotes: * Belgium + Luxembourg
** FR + DR

1.2.4 Consumption of Lead

Dataset name: CON_LE_D
Date creation/updt: 17/07/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition (From 1985 - 1995)
+ UNCTAD commodity yearbook, United Nations 1995 (for data from 1982 to 1985)
Footnotes: * Belgium + Luxembourg
** FR + DR

1.2.5 Consumption of Magnesium

Dataset name: CON_MA_D
Date creation/updt: 17/07/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition
Footnotes: * Belgium + Luxembourg
** FR + DR

1.2.6 Consumption of Nickel unwrought

Dataset name: CON_NI_D
Date creation/updt: 17/07/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition (From 1985 - 1995)
+ UNCTAD commodity yearbook, United Nations 1995 (for data from 1982 to 1985)
Footnotes: * Belgium + Luxembourg
** Up to 1990 (inclusive) data is for FR+DDR

1.2.7 Consumption of Tin

Dataset name: CON_TI_D
Date creation/updt: 17/07/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition (From 1985 - 1995)
+ UNCTAD commodity yearbook, United Nations 1995 (for data from 1982 to 1985)
Footnotes: * Belgium + Luxembourg
** FDR + DDR
*** the data for these countries differ in both sources. Used figures as stated in source

1.2.8 Consumption of Tungsten ore

Dataset name: CON_TU_D
Date creation/updt: 17/07/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: UNCTAD commodity yearbook, United Nations 1995
Footnotes: * data is for Belgium + Luxembourg
** see Belgium

1.2.9 Consumption of Zink

Dataset name: CON_ZI_D
Date creation/updt: 17/07/1998
Unit: 1000 tons
Geographical Cov.: European Countries
Temporal Cov.: 1982-1995
Source Type: XLS
Sources: Metal Statistics 1985 - 1995, 83 edition (From 1985 - 1995)
+ UNCTAD commodity yearbook, United Nations 1995 (for data from 1982 to 1985)
Footnotes: * Belgium + Luxembourg
** FR + DDR

1.2.10 Gross Inland Consumption by Fuel Type and Country

Dataset name: D0321ENG
Date creation/updt: 15/07/1999
Unit: 1000 tons of oil equivalent
Geographical Cov.: EU 15 Countries
EU 12,15 Totals
Temporal Cov.: 1960, 1970, 1975, 1980, 1985, 1990-1994
Source Type: CSV
Sources: NEW CRONOS, Eurostat (31/03/1998)
Footnotes: Break in series for Germany : 1990-1991.

1.2.11 Water Abstraction by Sector and Country

Dataset name: IOS1445
Date creation/updt: 27/08/1999
Unit: Millions m3/year
Geographical Cov.: EU 15, EFTA4, Acc Countries
Temporal Cov.: 1980, 1985, 1990, 1994-1996
Source Type: CSV
Sources: OECD Data Compendium, 1997
PTL-ETC/IW, 1998
Statistical Office, Web pages
ETC/IW (European Topic Centre on Inland Waters) 1997
Footnotes: Data for "agriculture" are for irrigation usage
Data for "urban" include supply to household and small business
Data for "industry - not cooling" include all industrial useage - no differentiation is made
Missing data for "agriculture" are assumed to be zero
Data for "cooling, hydropower and other" are related to "self-supply" : statistical units which are supplied with >3600 m3 water anually
Data for "industrial" also include "other"
Data for "agriculture" include domestic and industrial water use in rural areas
Data for "urban" is for drinking water

1.2.12 Water Abstraction by Source (Ground - Surface) and Country

Dataset name: IOS1447
Date creation/updt: 02/12/1998
Unit: Millions m3/year
Geographical Cov.: EU 15, EFTA4, Acc Countries
Temporal Cov.: 1980, 1985, 1990, 1994-1996
Source Type: CSV
Sources: OECD Data Compendium, 1997
Sustainable Water Use Report, and Ground-water Monograph, ETC/IW 1998.
PTL-ETC/IW, 1998
Statistical Office Web pages
Central Statistical Bureau of Latvia in publication: "Environmental Protection and use of Natural Resources in Latvia", 1997, EEA Library code 3239, p. 54
Central Statistical Office, in publication: "Environmental Policy and the Role of Foreign Assistance in CCE", 1998, EEA Library code 5741

Footnotes: The figures for total abstraction are from the ETC/IW questionnaire, 1997 and the apportionnement of surface and ground-water was calculated from the percentages given in the Ground-water Monograph (ETC/IW, 1998)
The figures for total abstraction are from the PTL-ETC/IW questionnaire for the EU98 Report, and the apportionnement of surface and ground-water was calculated from the percentages given in the Ground-water Monograph (ETC/IW, 1998). Total refers to abstraction of FRESH water, i.e. rain water and pits water is not included. Decision to include FRESH water only is based on "Definitions" in the 2.2.3 data set descriptions in the Guidelines.
Figure is referred to as "Water Withdrawal".
The data need to be considered with reservations due to the lack of a common European procedure to estimate water resources

1.2.13 Gross Inland Consumption (All Products) by Country

Dataset name: SIR100A
Date creation/updt: 22/04/1999
Unit: 1000 tons of oil equivalent
Geographical Cov.: EU 15, EFTA 4 Countries
EU 12, EU 15, EUROZONE Totals
Temporal Cov.: 1985-1997
Source Type: CSV
Sources: NEW CRONOS, Eurostat (19/04/1999)
Footnotes: Break in series for Germany:1990-1991