

*The Data without Boundaries – DwB – project exists to support equal and easy access to official microdata for the European Research Area, within a structured framework where responsibilities and liability are equally shared.*

## CIMES - Centralising and Integrating Metadata from European Statistics

### What is CIMES?

CIMES is a web-based application providing a large and comprehensive overview of national official microdata across Europe. CIMES' metadata database describes the data and the access procedures to them with information about (1) the different types of files available (Public Use Files, Scientific Use Files, Secure Use Files), (2) conditions of access and (3) links to the providers. CIMES does not provide access to these microdata.

CIMES thus gathers pieces of information currently scattered throughout Europe - either at NSIs, other government agencies or data archives - and stores them into a structured database compatible with the DDI standard.

CIMES is a product of the DwB project that aims to facilitate access across borders to the rich resources of national official microdata in Europe currently underused by the researchers. The development of CIMES has been coordinated by CNRS - Réseau Quetelet in partnership with GENES - CASD and other DwB partners.

### A Difficult Situation

Though the number of microdata integrated at European level (Eurostat) is increasing, national microdata offer rich resources for the researchers: series provide a longer historical perspective while administrative databases are now increasingly available for longitudinal approaches. However, researchers face serious obstacles when trying to get information on these microdata. Where to get them and how to conduct comparative research across European countries requires an in-depth search through the websites of each national statistical agency. Metadata are not provided in similar ways and

language issues are frequent. Overcoming this should eventually facilitate international comparisons of European countries and innovative projects in economics and social sciences; thus fostering the development of a Knowledge-Based economy while promoting research-driven public policy evaluations at European level.

### Towards Improved Metadata

Making these data easier to find, more comprehensible and more usable; requires the improvement of the descriptive metadata on three key elements. CIMES therefore provides:

- “ a **single point of access** and an overview of available data in Europe;
- “ a **standardized documentation**, facilitating the access, evaluation and indexation of information by search engines;
- “ **basic information on data access conditions** and links with the providers, the producers, a data archives, or a Research Data Centres for Secure Use Files.

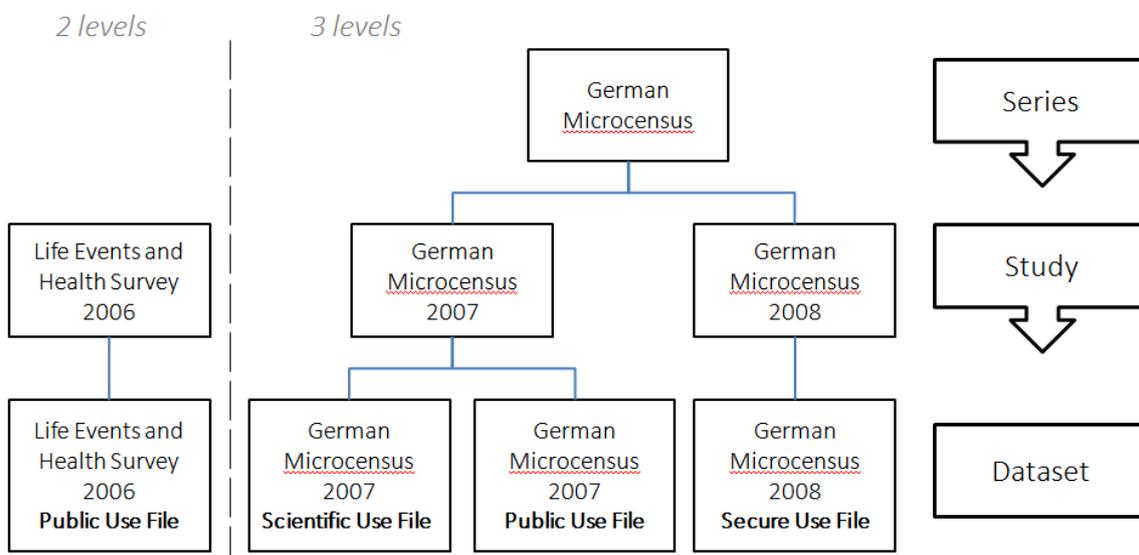
### What Is In There?

The documentation included in CIMES distinguishes the *series*, the *studies*, and the *datasets*. In many cases, the relationship between these three objects can be conceived as a simple three-level hierarchy (as in Graph 1). A series is a set of studies and represents a longitudinal or repeated cross-sectional data collection process. It also illustrates the continuous data collection process carried out by the NSIs, where each wave of collection is a study. The dataset refers to the set of files which are disseminated. For example, the Labour Force Survey (LFS) is a series, the actual collection of

the data in 2007 is a study, the LFS 2007, and the dataset is a set of files disseminated (e.g. the LFS Scientific Use File (SUF) 2007). Each level is

documented by a list of fields and a controlled vocabulary (Table 1). CIMES currently does not document the variable level.

**Graph 1 - A Three-level Schema**



**Table 1 - Fields and Controlled Vocabulary**

Series level	Study level	Dataset level
Title	Title	Dataset Identification Number
Subtitle	Subtitle	Name
Alternative Title	Alternative Title	Access Conditions
Abstract	Producer	Other Access Conditions
Keywords	Abstract	Description of Accreditation Procedures
Geographic Coverage	Keywords	Contact
Time Period Covered	Geographic Coverage	Notes
Time Method	Universe	
Notes	Sampling Procedure	
	Date of Collection	
	Time Period Covered	
	Kind of Data	
	Documents	
	Sources	
	Notes	

**Currently** CIMES covers 277 series, including 1,893 studies and 2,244 datasets in 31 European countries, also providing links to the major European integrated microdata. CNRS - Réseau Quetelet, in partnership with CASD/GENES, keep maintaining and developing CIMES.

**Related websites**

- ~ <https://cimes.casd.eu>
- ~ DwB Deliverable 5.2 is available at

<http://www.dwbproject.org/about/deliverables.html>



Centralising and Integrating Metadata from European Statistics

