

CIMES, a tool for describing European Official Statistics Microdata

Presented by

Raphaëlle Fleureux, Cyril Jayet, Arnaud Szulek / CNRS-RQ



Bergen, EDDI 2012, December 4, 2012



Outline

- DwB WP5 Objectives and Tasks
- CIMES
- The Documentation Schema
- Technical aspects
- Demo
- State-of-Play & Further Developments

DwB WP5 : Objectives and Tasks

- Objective : Provide **comprehensive overview** of official statistical data in Europe
- Task 2: Collect, structure and code information on available microdata from **official statistics at the national level** for each country in Europe
- Task 3: Collect, structure and code information on integrated **European microdata from official statistics** held by Eurostat and national censuses

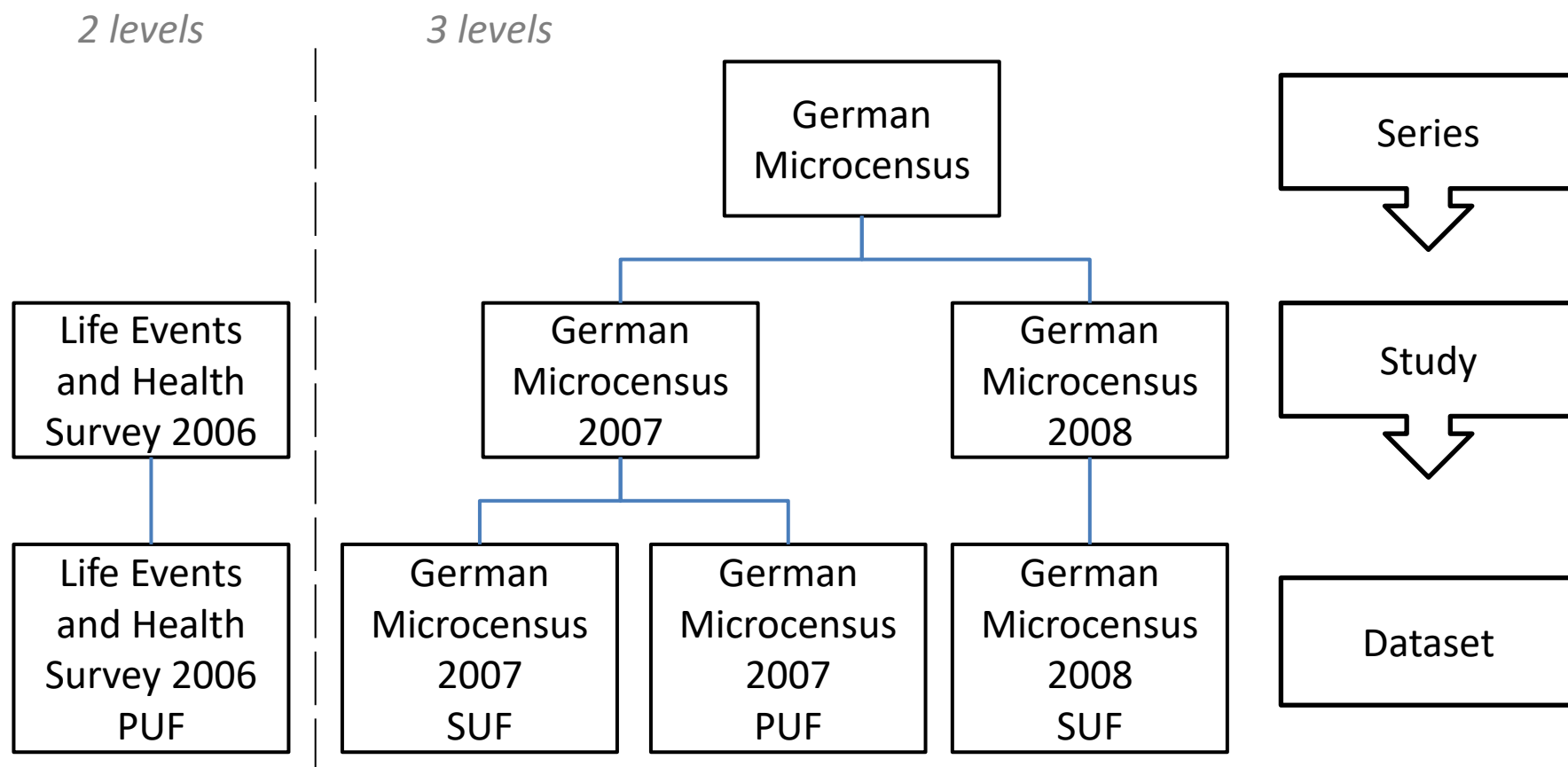
CIMES

- In the task 2 of WP5, we agreed on a documentation schema based on DDI2 and 3
 - We tested different softwares (Nesstar, Colectica, etc.) and they were not really relevant for the purpose of this documentation (in terms of correspondence to the schema, and of financial and technical aspects)
- ➔ We decided to develop our own tool at CNRS-RQ: CIMES (Centralising and Integrating Metadata from European Statistics)

A Three-levels Schema

- Series: a group of studies, for example Labour Force Survey (LFS)
 - Study: the actual collection of data, for example LFS 2007
 - Dataset: a set of files disseminated, for example LFS Scientific Use File (SUF) 2007
- ➔ In order to give a large and comprehensive overview of official microdata available in Europe, a simple schema is needed. The variable level will not be documented.

A Two- or Three-levels Nested Hierarchy



Schema Fields: Series Level

Series level	DDI 2.1	DDI 3
Title	2.1.5.1 <serName>	ddi:Group/r:Citation/r:Title
Subtitle		ddi:Group/r:Citation/r:SubTitle
Alternative Title		ddi:Group/r:Citation/r:AlternateTitle
Abstract	2.1.5.2 <serInfo>	ddi:Group/g:Abstract
Keywords		ddi:Group/r:Coverage/r:TopicalCoverage/r:Keyword
Geographic Coverage		ddi:Group/r:Coverage/r:SpatialCoverage/r:Description
Time Period Covered		ddi:Group/r:Coverage/r:TemporalCoverage/r:ReferenceDate/r:HistoricalDate
Time Method		ddi:Group/d:DataCollection/d:Methodology/d:TimeMethod
Notes		ddi:Group/r:Note

Schema Fields: Study Level

Study level	DDI 2.1	DDI 3
Title	2.1.1.1<titl>	s:StudyUnit/r:Citation/r:Title
Subtitle	2.1.1.2 <subTitl>	s:StudyUnit/r:Citation/r:SubTitle
Alternative Title	2.1.1.3 <altTitl>	s:StudyUnit/r:Citation/r:AlternateTitle
Producer	2.1.3.1 <producer>	s:StudyUnit/r:Citation/r:Publisher
Abstract	2.2.2 <abstract>	s:StudyUnit/s:Abstract
Keywords	2.2.1.1<keyword>	s:StudyUnit/r:Coverage/r:TopicalCoverage/r:Keyword
Geographic Coverage	2.2.3.4 <geogCover>	s:StudyUnit/r:Coverage/r:SpatialCoverage/r:Description
Universe	2.2.3.9 <universe>	s:StudyUnit/c:ConceptualComponent/c:UniverseScheme/c:Universe/c:HumanReadable
Sampling Procedure	2.3.1.4 <SampProc>	s:StudyUnit/d:DataCollection/d:Methodology/d:SamplingProcedure
Date of Collection	2.2.3.2 <collDate>	s:StudyUnit/d:DataCollection/d:CollectionEvent/d:DataCollectionDate/r:HistoricalDate
Time Period Covered	2.2.3.1 <timePrd>	s:StudyUnit/r:Coverage/r:TemporalCoverage/r:ReferenceDate/r:HistoricalDate
Kind of Data	2.2.3.10 <dataKind>	s:StudyUnit/s:KindOfData
Documents	5.0 <otherMat>	
Sources	2.3.1.8 <sources>	
Notes	2.2.4 <notes>	s:StudyUnit/r:Note

Schema Fields: Dataset Level

Dataset level	DDI 2.1	DDI 3
Dataset Identification Number	2.1.1.5 <IDNo>	p:PhysicalDataProduct/r:UserID
Name		p:PhysicalDataProduct/p:PhysicalDataProductName
Access conditions		p:PhysicalDataProduct/p:PhysicalDataProductName xml:lang="en"
Other Access Conditions	2.4.7.7 <conditions>	a:Archive/a:ArchiveSpecific/a:Item/a:Access/a:AccessConditions
Description of accreditation procedures		
Contact	2.4.2.4 <contact>	a:Archive/a:OrganizationScheme/a:Organization
Notes		a:Archive/r:Note

CIMES: Technical Aspects

- HTML5 Web-based tool
- MySQL database and PHP scripting engine.
- Javascript (Mootools framework) and CSS3 client side.
- Works with any standard compliant browser (and even with Internet Explorer 8...).
- DDI2 import functionality

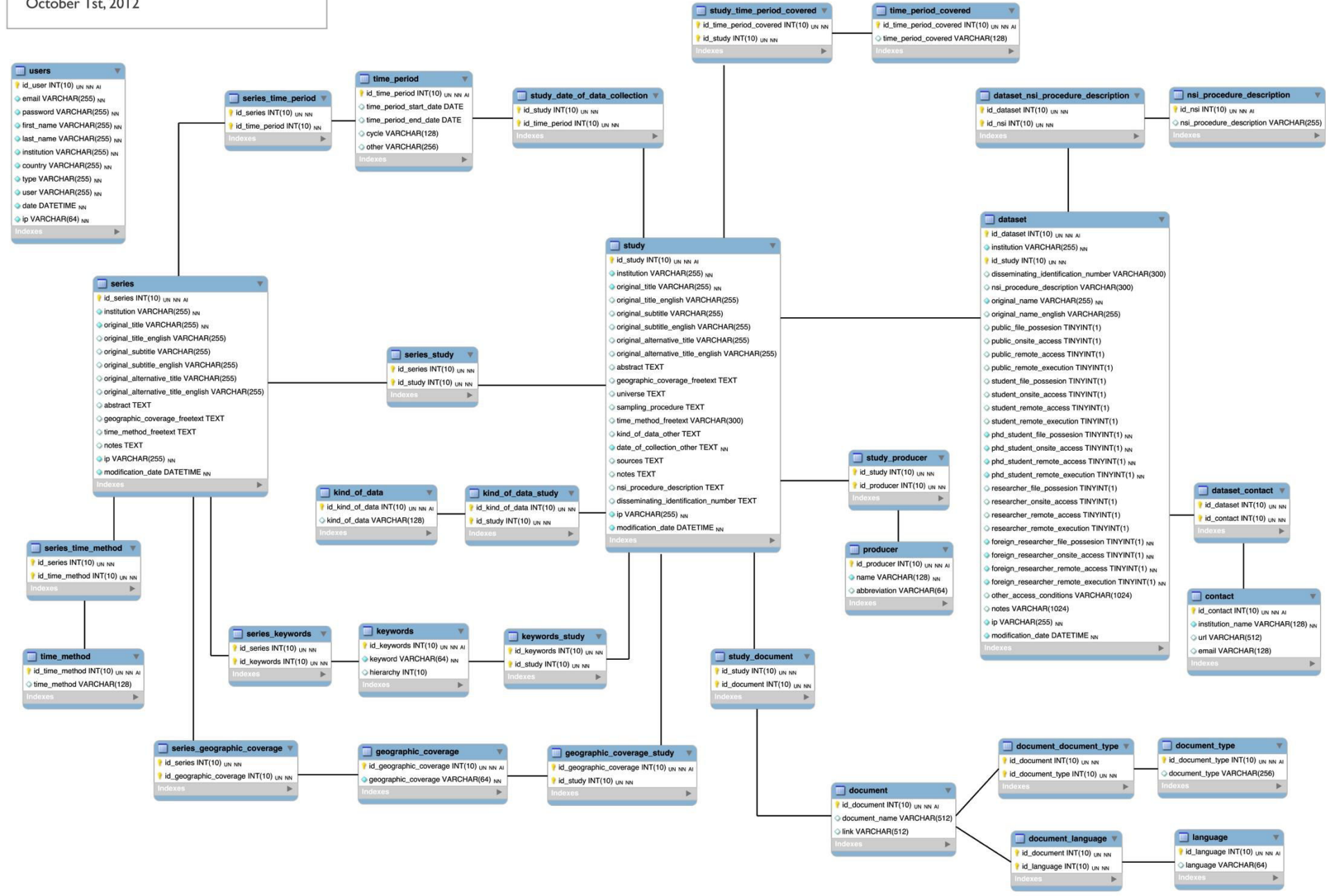
The CIMES tool can be hosted on any web server (Apache, IIS) with minimal configuration/modification.

There is no other client software installation needed.

Every update and bug fix is immediately made available to all the users.

Cimes Database modelisation v12

October 1st, 2012



State-of-play & Future Developments

- We plan to document surveys on most commonly studied topics in comparative research projects in order to facilitate international comparisons across c.a. 20 European countries
 - Living conditions, Employment, Education, Social Mobility, Structure of Earnings, Working Conditions, Innovation.
- We plan to make the metadata available by the end of the year 2013

Thanks for Listening

DwB website: <http://www.dwbproject.org/>

