



**UNITED NATIONS**  
**ECONOMIC COMMISSION FOR EUROPE**

Population Activities Unit

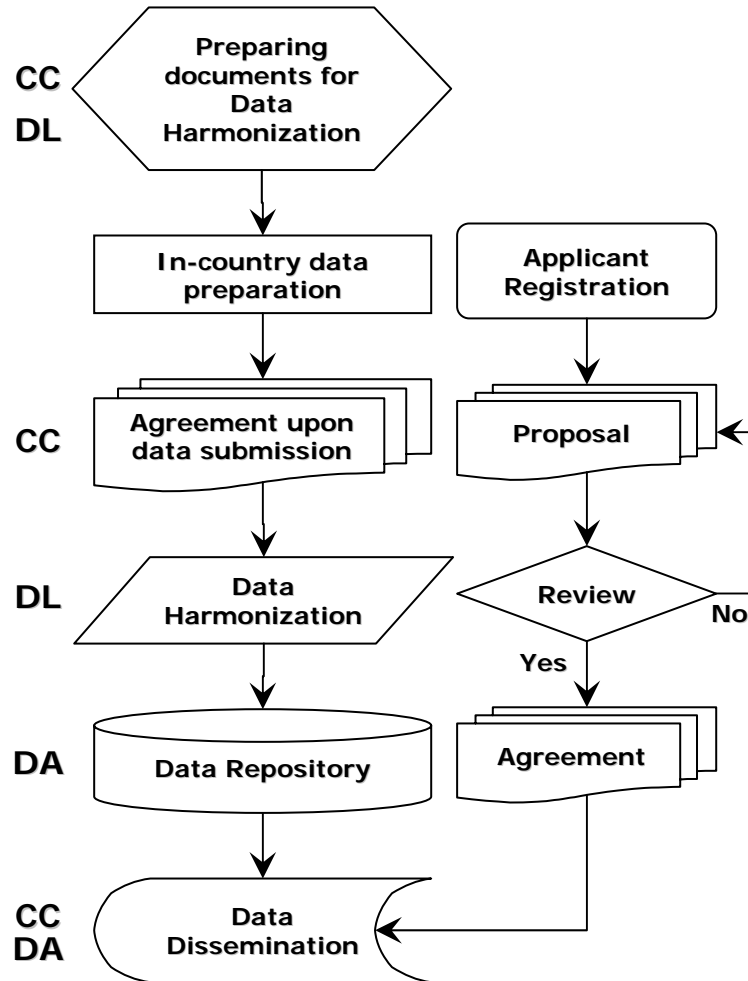
**GGG**  
**Data Submission and Disseminations**

Andrej Kveder

*5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007*



# Process Diagram

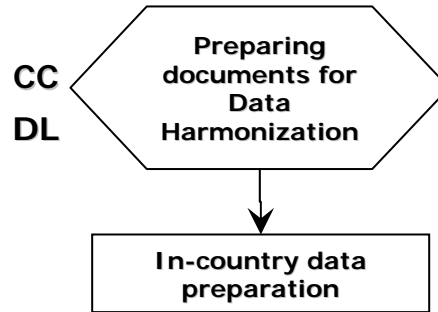


*Legend:*  
CC – Coordination Centre  
DL – Data Laboratory  
DA – Data Archive

5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007



# Preparatory Stage



*Legend:*  
*CC – Coordination Centre*  
*DL – Data Laboratory*  
*DA – Data Archive*

*5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007*



# Documentation

- Description of the Harmonised Data File (HDF)
  - question numbers based on the GGS Full Questionnaire,
  - variable labels,
  - harmonised variable names,
  - variable formats
- Template for data availability report
  - easy detection of cases when a variable cannot be fully harmonised
  - explanation and documentation for the country-specific solutions, including variables where the questionnaire explicitly calls for such solutions
- Legal documents

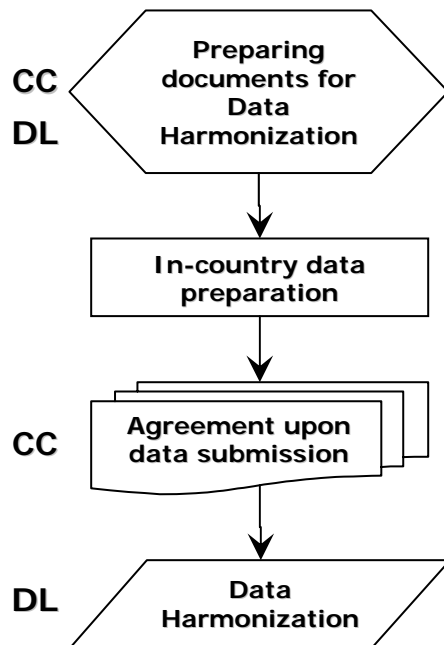


# In-country Data Preparation

- Data cleaning
  - mechanical questionnaire routing checks
  - logical consistency checks
- Naming the variables according to HDF
- Recording and explaining where the national data differs from the HDF
- Providing methodological information
  - sampling and survey design
  - data collection methods and fieldwork
  - final disposition codes
  - panel maintenance



# Submission and Harmonization



*Legend:*  
CC – Coordination Centre  
DL – Data Laboratory  
DA – Data Archive

*5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007*



# Harmonized Data File (HDF)

- 3 datasets
  - Fully edited and harmonized data file
    - based on the best judgement of the harmonizing team
    - fully ready to be used as a part of the pooled international dataset
  - Data file as submitted and only partially harmonized
    - accompanied with the flag variables data set
      - indicating data editing changes
    - possible tracing of any changes made in the harmonization process
  - Cleaned and edited event histories

*5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007*



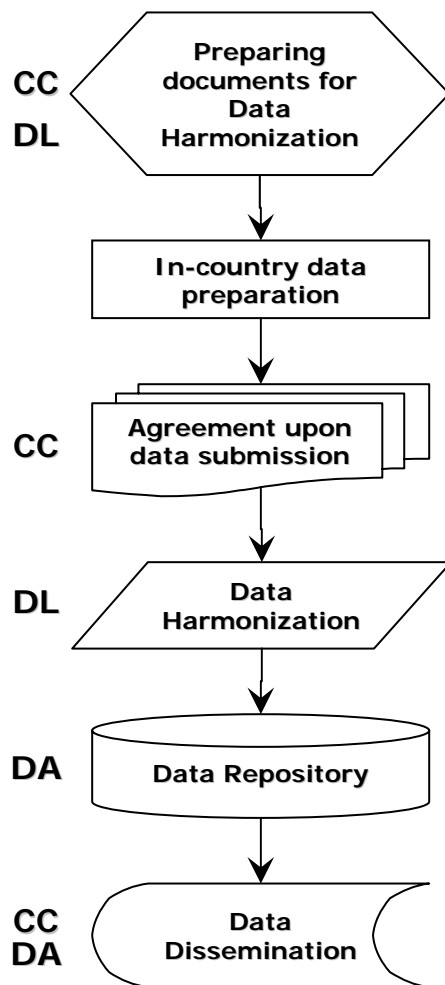
# Harmonization process

- Renaming variables to standard names
- Checking routing errors
  - cleaning should have already been made in the countries
- Time variables
  - 4 digit years with limited very obvious edits to years and months
- Multiple response variable sets
  - creating dichotomous variables for each item in a given set
- Flagging incomparable or country specific information
- Consolidating information when stored in multiple variables (e.g. Section 5 – Parental Home)
- Final reordering of the variables





# Archiving and Dissemination



*Legend:*  
CC – Coordination Centre  
DL – Data Laboratory  
DA – Data Archive

*5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007*

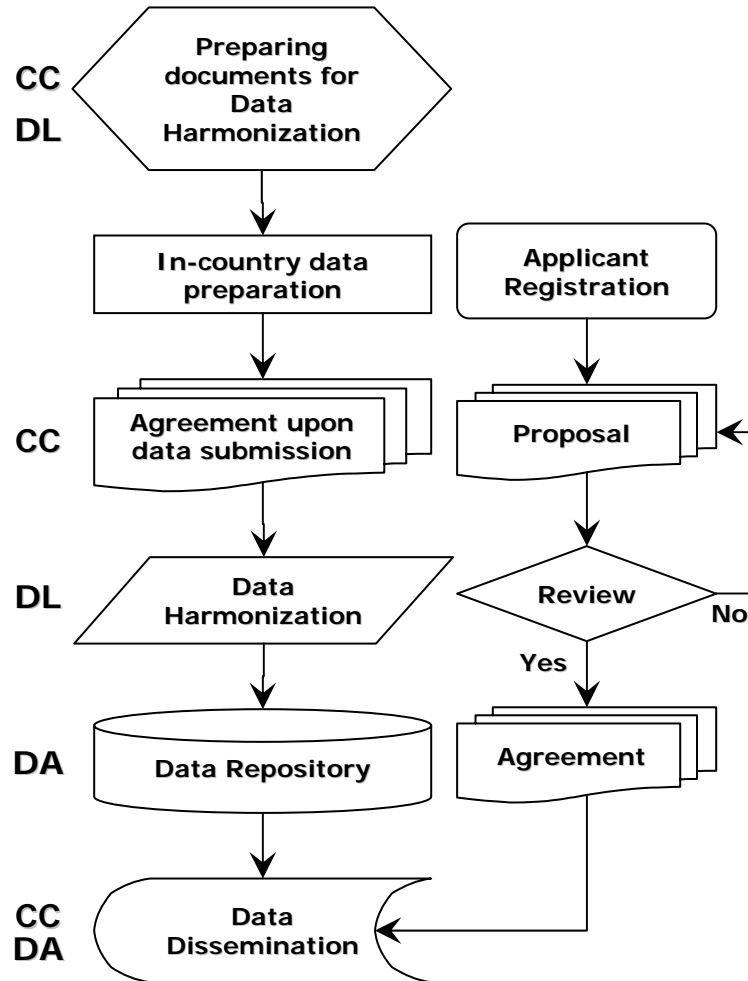


# Data Archive

- Development and implementation of the on-line data repository
  - safe-guarding the data sets
  - secure on-line registration and data request
- Data download from designated web space integrated into the on-line registration process
- Routine refreshments of the data-sets
  - based on new information gained through data harmonization
    - e.g. updating item labels based on new country specific lists
- Possible expansions in the future
  - provision of tabulations and other aggregated outputs



# Data Dissemination



*Legend:*  
CC – Coordination Centre  
DL – Data Laboratory  
DA – Data Archive

5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007



**UNITED NATIONS**  
**ECONOMIC COMMISSION FOR EUROPE**

Population Activities Unit

**Thank you for your attention!**

*5<sup>th</sup> GGP IWG meeting, Ljubljana, 18-20 January 2007*