

Berliner VGR-Kolloquium, 5./6. Juni 2008

# DIWAX

**Data handling and forecasting based on National Accounts**

---

Dr. Stefan Kooths  
DIW Berlin – Macro Analysis and Forecasting

## DIWAX in a nutshell

- Support for macroeconomic forecasting based on national accounting schemes
- Building blocks
  - Relational metadata base for organizing data along standard economic and statistical classifications
  - Resource-based project management for downloading and importing data from many different sources
  - Analytic interface for analyzing data and designing forecasts
  - Interface for macroeconomic models
  - Backbone for AI-based methods

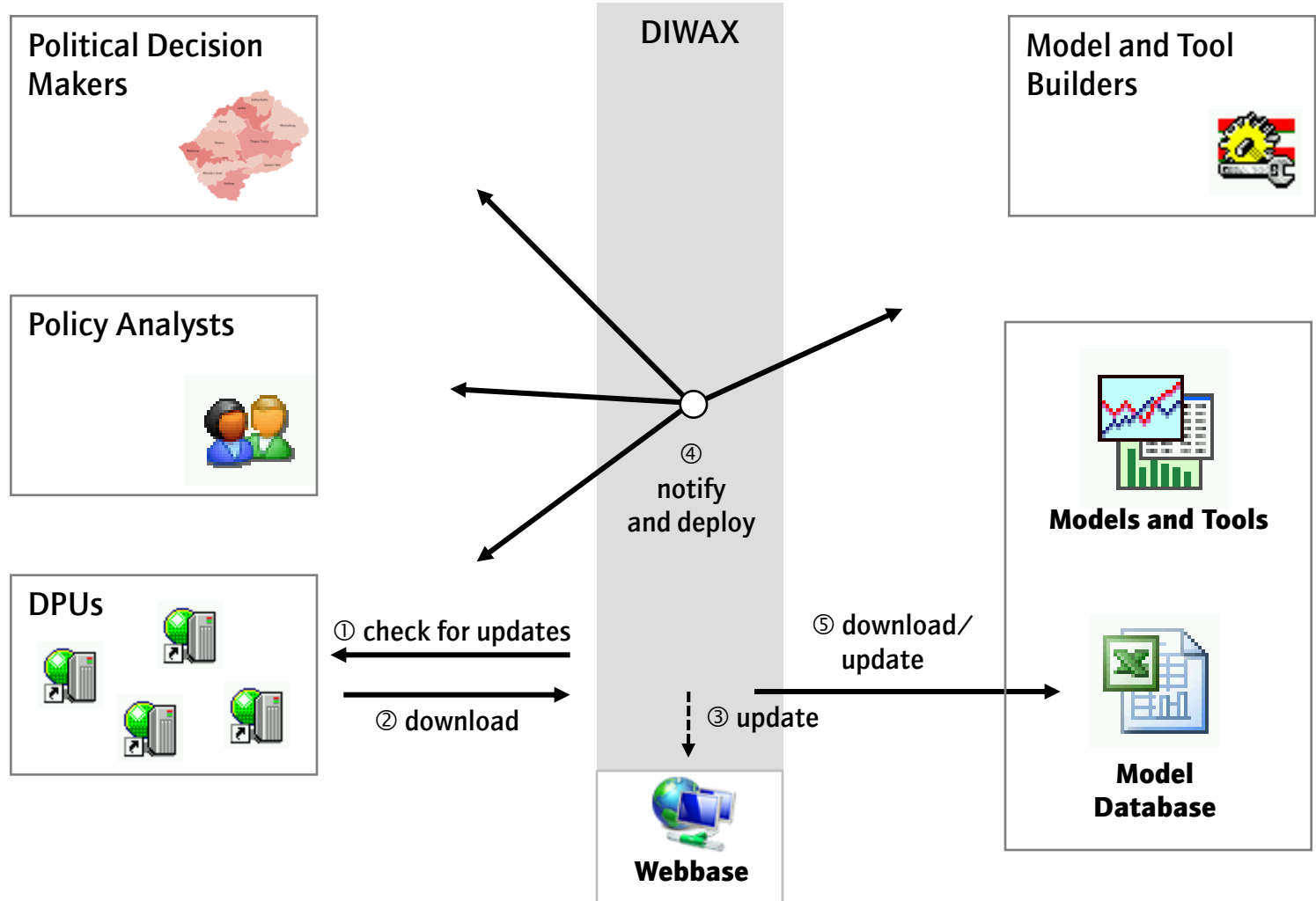
## Four basic options

- **Ad hoc data compilation**
  - Getting data via informal channels more or less regularly
  - Copying and adapting data to specific needs manually
  - Lack of data integrity, muddling through
- **In-house database development**
- **Standard industry database**
  - Limited, not customized to macroeconomic needs (driven by business needs)
  - Typically difficult to adapt to developing-type country needs (e.g. model and tool integration)
- **Data network via common interface (DIWAX approach)**
  - Prestructured environment for macroeconomic analyses
  - Unrestricted data input

## The DIWAX database philosophy

- DPU keep maximum of autonomy
    - Decision of how to organize data (table and file format) is up to the respective Data Producing Units (DPUs)
    - Commitment not to change chosen system without good reasons
  - Data integration via interface
    - Economic and technical data description following international standards and classifications (metadata)
    - Integrating DPU data files as resources  
("Fishing the data out of the sea of tables that already exist.")
    - Providing data in user-, model-, and tool-compatible formats
  - Web-based data exchange and master-copy backup
- ⇒ **Virtualization of the database  
(network of integrated resources)**

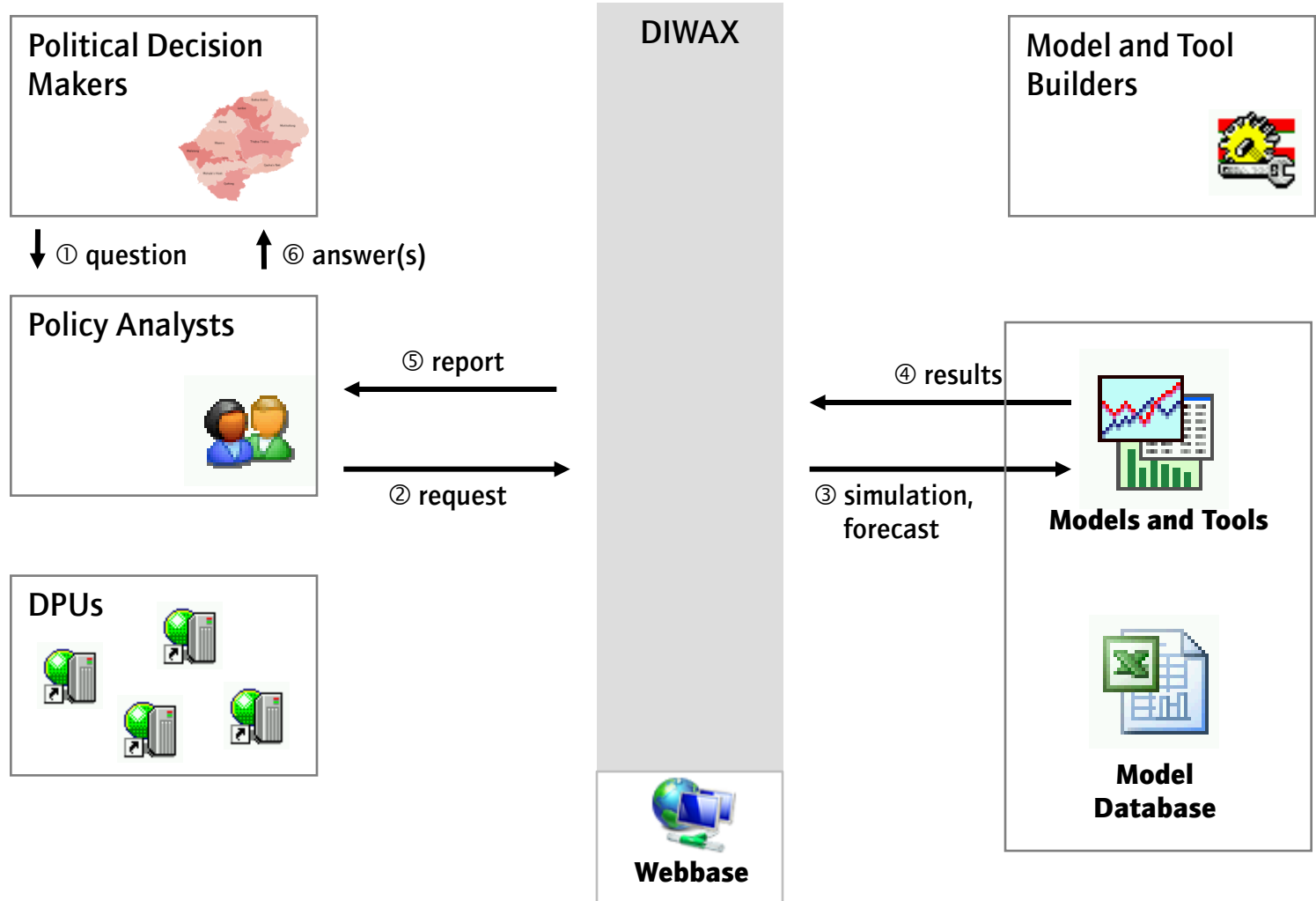
# Integrating various data sources DIWAX as a database interface



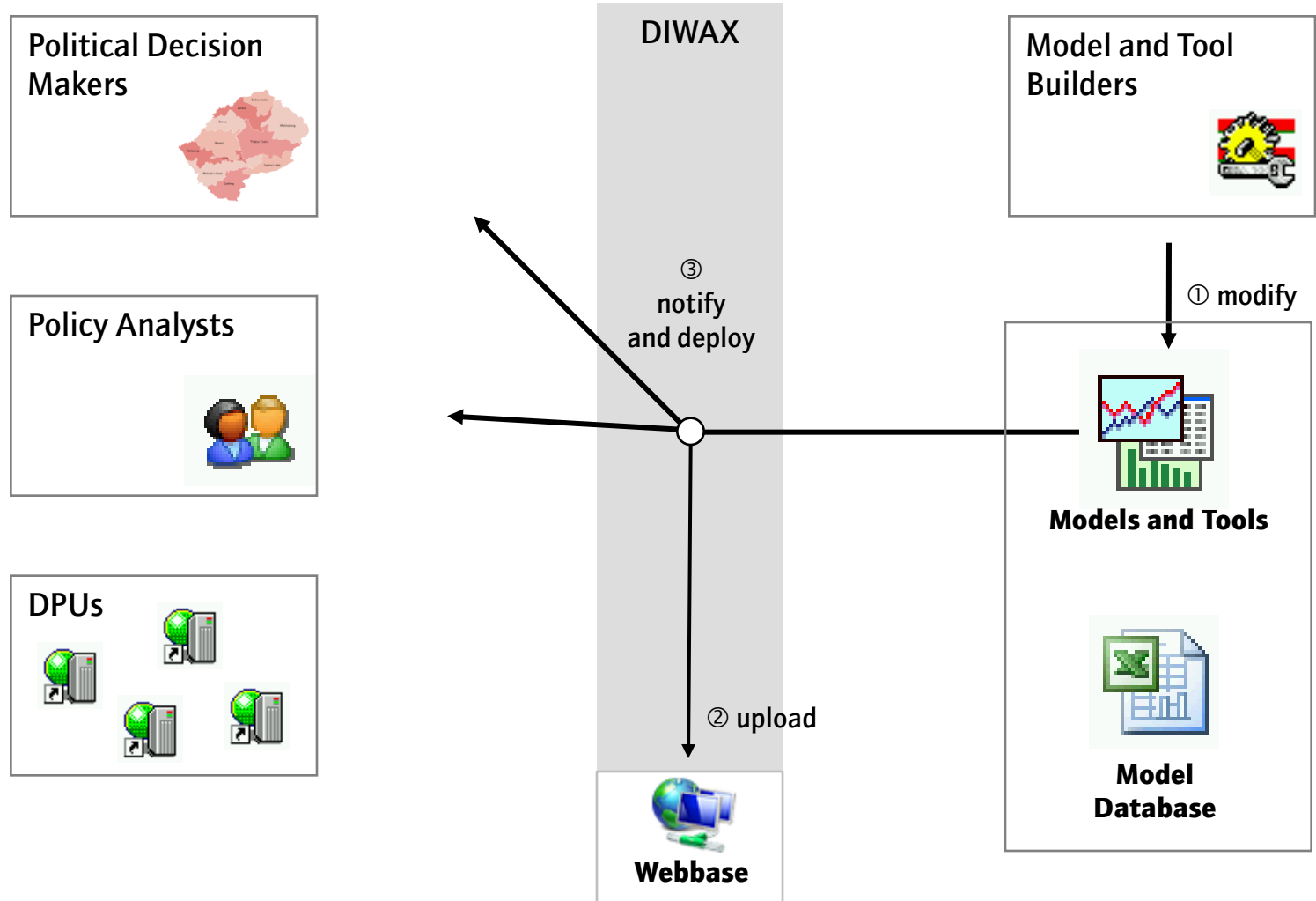
## Model and tool interface

- Running models and using tools within the DIWAX user interface (e. g. remote control for EViews models)
- Managing scenarios through simulation wizards
- Report generator
- Model versioning and distribution

# Using models and tools: DIWAX as a front-end (user-interface)



# Modifying models and tools: Using DIWAX for notification and distribution





## User Rights Management

- Global Administrator
- Domain License Administrator (Master DPU)
  - Domain-wide user registration
  - Management of metadata master copy
  - Resource organization
- Data Publishing Users
  - Reading rights according to authority status
  - Writing rights for specified resources
- Standard Users
  - Reading rights for public-domain resources
  - Writing rights for privately user-defined resources only

## Data processing

- User interface for data retrieval, charting and editing
- Building accounting and data derivation schemes (economic framework for data analysis)
- Support for data integrity
  - detecting inconsistencies
  - checking for outliers
- Data transformation
  - Indexing, growth rates, frequency transformation
  - Price adjustment/unadjustment
  - Adjusting/unadjusting for seasonal and calendar effects (using external factors)
- Reprint (user-defined templates for charts, tables)

## Handling data: The METADATA concept

- Focus of interest in economic analysis
  - Measuring the degree of **economic activity** (in a wide sense)
  - NOT: **time series** (just carrier of specific information)
  
- Data description
  - Primarily in an economically meaningful way
  - Secondarily in technically appropriate terms
  - ⇒ Metadata for time series description
    - Using international standard classifications where available (e.g. SNA, ISIC, NACE ...)
    - Extensible by user if necessary

## An ITEM describes concrete economic activity

*Production  
of consumer goods  
by non-financial companies  
within the textile industry  
of Lesotho  
in bn maluti*

### Item

- Activity
- Product
- Sector
- Industry
- Area
- Unit
- ItemType
- ...

**Items are the core data concept of DIWAX**

## Basic economic types of time series

- Basic economic types of time series (elements)
  - Nominal
  - Volume
  - Price
- Used by Items either individually or as a full combo
  - ItemType
  - **Nominal** only: Tax revenue
  - **Volume** only: Hours worked, Labor productivity
  - **Price** only: Interest rate, share index
  - **Combo** (having a generic deflator): Private consumption
- Relevant series types follow from economic description

## Technical time series dimensions 1

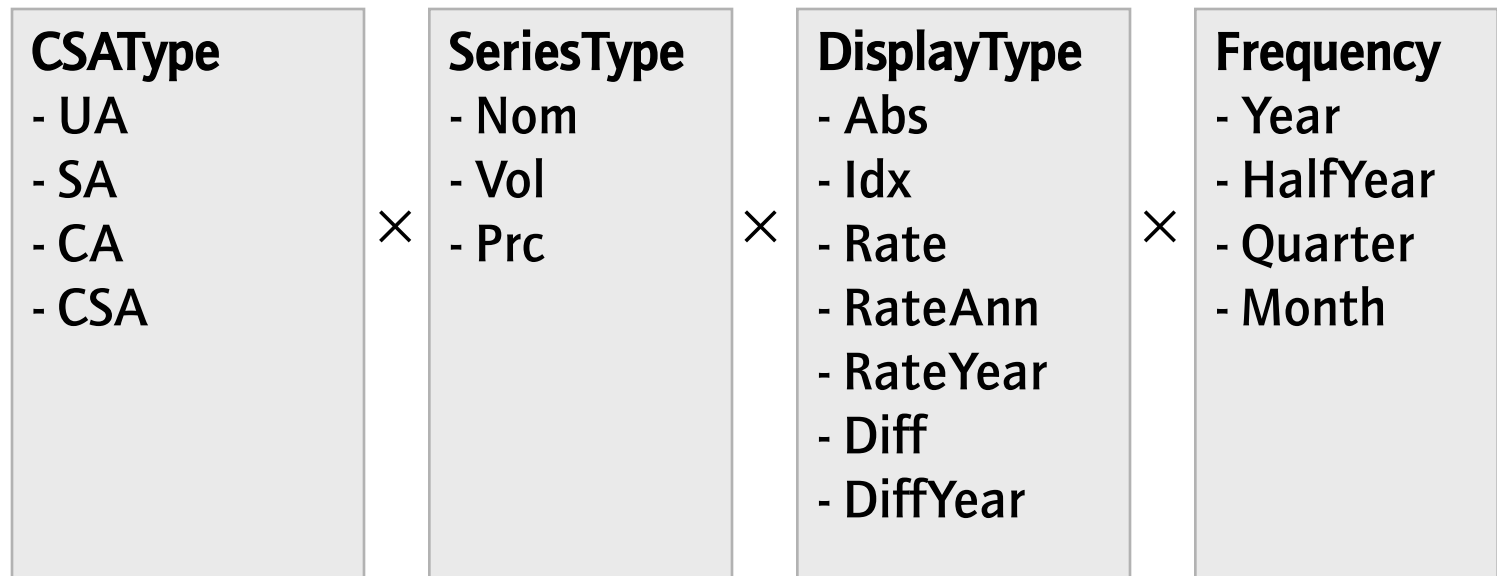
- **CSA type (calendar and seasonal adjustment)**
  - UA: unadjusted
  - SA: only seasonally adjusted
  - CA: only calendar (working-day) adjusted
  - CSA: calendar and seasonally adjusted
- **Display type**
  - Abs: Absolute
  - Idx: Index
  - Rate: Growth rate to previous period
  - RateAnn: Annualized growth rate to previous period
  - RateYear: Growth rate to previous year's period
  - Diff: Difference to previous period
  - DiffYear: Difference to previous year's period

## Technical time series dimensions 2

- Frequency
  - Year
  - Halfyear
  - Quarter
  - Month

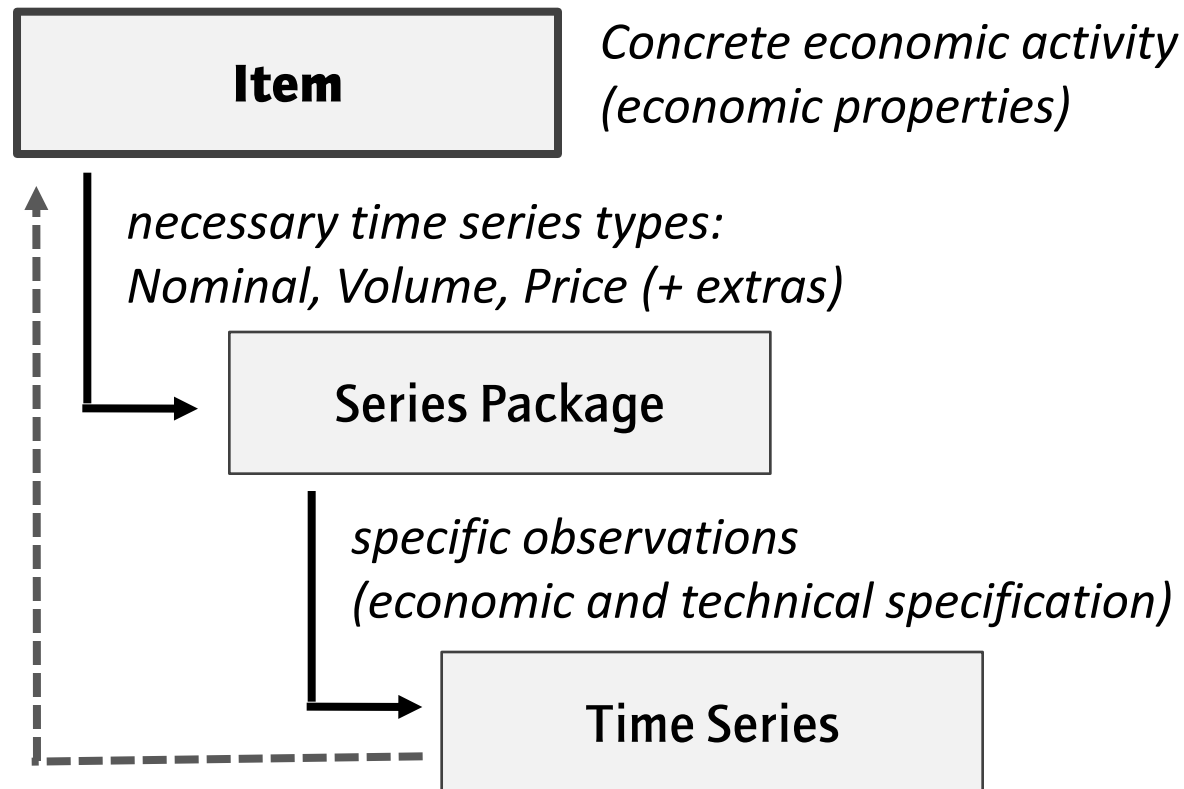
## Bundling time series: The SERIESPACKAGE concept

- Each Item holds all its time series in one container, called the SERIESPACKAGE
- Structure of time series within each SERIESPACKAGE






## Items and time series revisited




## Grouping and hierarching Items

- To make handling Items easier they can be grouped by THEMES, e.g.
  - GDP by expenditure and all its components
  - Distribution of income
  - Balance of payments
  - Government financial account
  
- Building hierarchies by indicating the relative position of one Item to another
  - Nodes reflect mathematical identities (definitional equations)
  - Allows to fully build the analytic framework and the flow of derivation schemes (definition of residual Items)

## Linking data to ITEMS: The RESOURCE concept

- A resource ... 
  - ... holds one or more time series used by one or more items
  - ... is usually an Excel workbook or a text file
- When registering a resource with DIWAX ...
  - ... the file content is described using the metadata codes
  - ... data organization within each resource is stored (only once, if kept unchanged)
  - ... an identification code is given to each resource type
- Resource library
  - Has a subfolder for each resource type
  - Versioning according to last ex-post period
- Items assign resources to specific time series

## Project administration

- Project: Common settings relevant to all ITEMS
- Project file (Excel-Workbook format) 
  - Project information
    - Project folder (one subfolder per project recommended)
    - Project name („Winter outlook 2007“)
    - Project code („WO-2007“)
    - Price adjustment method
    - Calendar and seasonal adjustment method
    - Time spans
    - Reference year (for indexing)
    - List of used resource files
  - All time series data (one worksheet per ITEM)

## The DIWAX program technology

- DIWAX follows an agent-based approach
- Agents are programs within the overall program designed to fulfill specific tasks
  - ItemAgent
  - ProjctAgent
  - TimeAgent
  - ThemeAgent
  - DownloadAgent
  - ...
- DIWAX itself is the platform that hosts the agents and allows for their communication and collaboration
- The user interface of all agents is similar (know one, know all principle)

