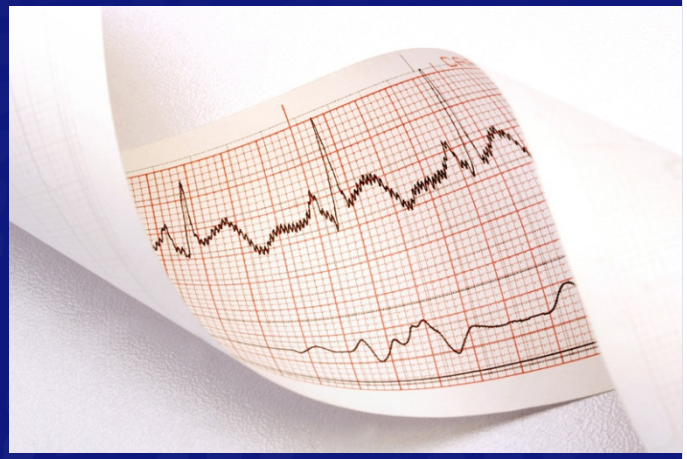
OFFIS Health
INSTITUTE FOR INFORMATION TECHNOLOGY

A Common Meta-Model for Data Analysis based on DSM

R&D Division Health

Gesundheit
Health



Yvette Teiken

OFFIS Health
INSTITUTE FOR INFORMATION TECHNOLOGY

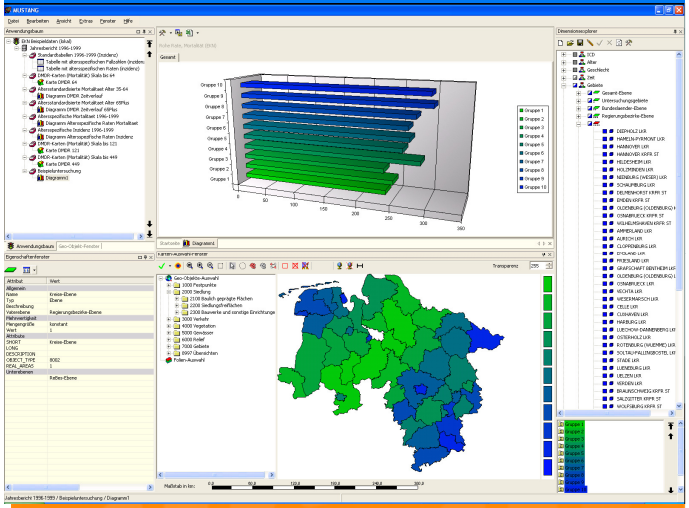
2 Agenda

- Introduction
- Brief overview of our research activities
  - Model Driven MUSTANG
  - Visual MUSTANG
- A common Meta-Model for data analysis
- Conclusion

3 Brief Overview of our Research Activities

MUSTANG Multidimensional Statistical Data Analysis Engine

- **Goal**
  - Data supply and decision support
  - Integration of geo data
  - Statistical functions
- **Approach**
  - Modelling of multidimensional data
  - Integration of domain-specific analytical procedures
  - Integration of GIS technologies
- **Application area**
  - Cancer- and infection-epidemiology
  - Health report
- **New fields of application**
  - Decision support systems for SMEs (small and medium-sized enterprise)
  - Demand Driven approach

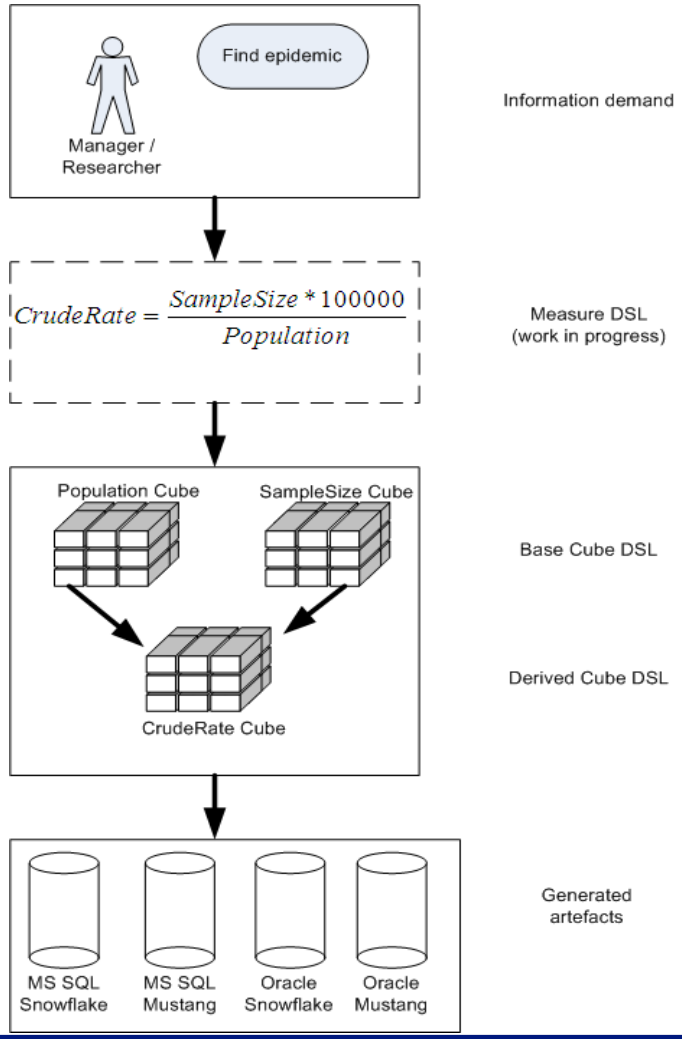


4 Current Data Integration in MUSTANG

- Use “standard” ETL-process
- Infrastructure creation:
  1. Define multidimensional structure (Dimension and facts)
  2. Write SQL script that represents structure
  3. Execute and check written SQL
- Data integration:
  - Write programs/scripts to manipulate and integrate given data
  - Write application for data integration
- Challenges:
  - Complex but schematic work
  - Error-prone
  - Data quality
  - Cost extensive for SME

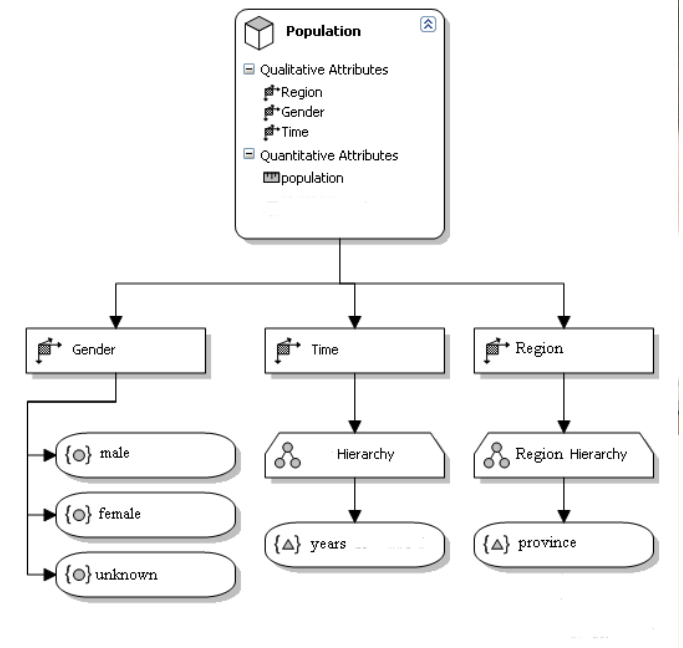
5 Model Driven MUSTANG I

- Goal: Demand driven DWH process based on DSM
- Common approach: Data driven
- Our approach:
  - Integrate Top Down approach
  - More demand driven
- Integrate of different aspects:
  - Data Quality
  - Dimension Modeling
  - Security Aspects



6 Model Driven MUSTANG II

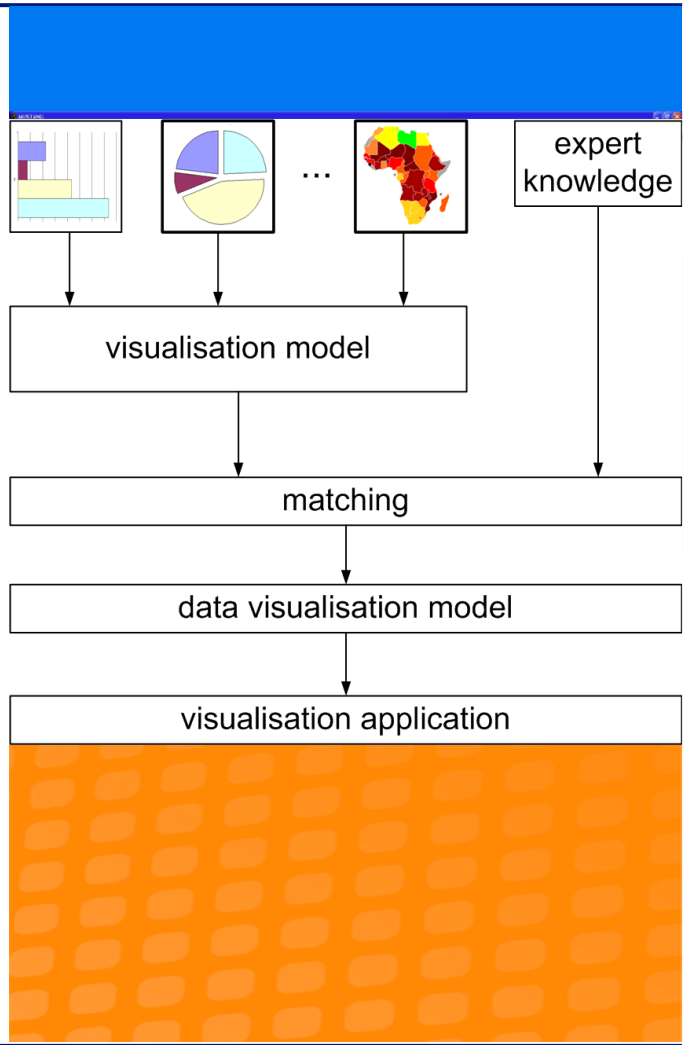
- DSM based approach on cube modelling
  - Models DWH cubes
  - Based on ADAPT
- Infrastructure generation
  - Different multidimensional view
  - Different deployment server
- Integration application
  - Web Application
  - XML WebServices



7 Visual MUSTANG

Semi-Automatic Data Visualization

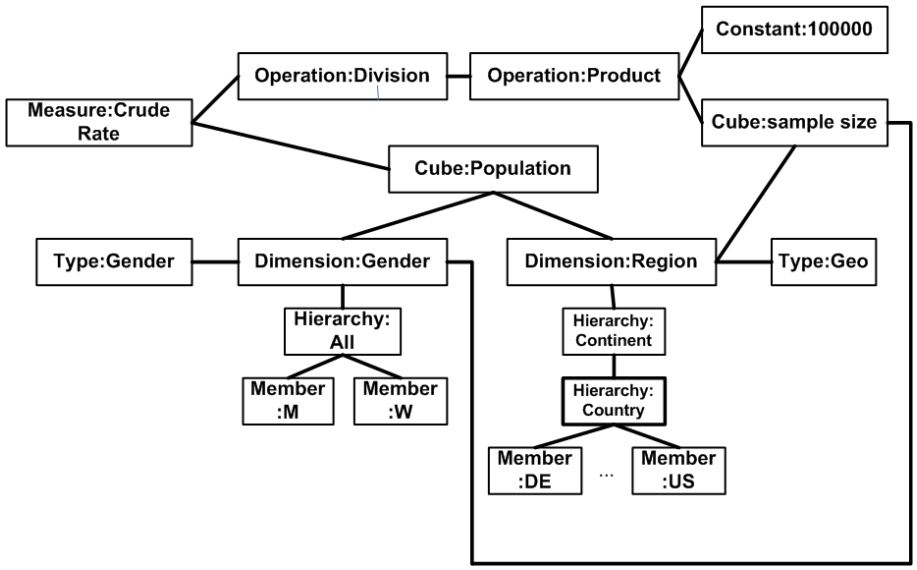
- Task: Choose appropriate Visualization for given data
- Problem:
  - Large variety of visualizations applicable
  - Expert with knowledge about analysis need to choose a matching visualization
- Idea:
  - Gather expert knowledge
  - Formalize expert knowledge
  - Enrich visualization model with expert knowledge
  - Matching process to match visualization to given set of data
- Challenge:
  - Semantic information about data model

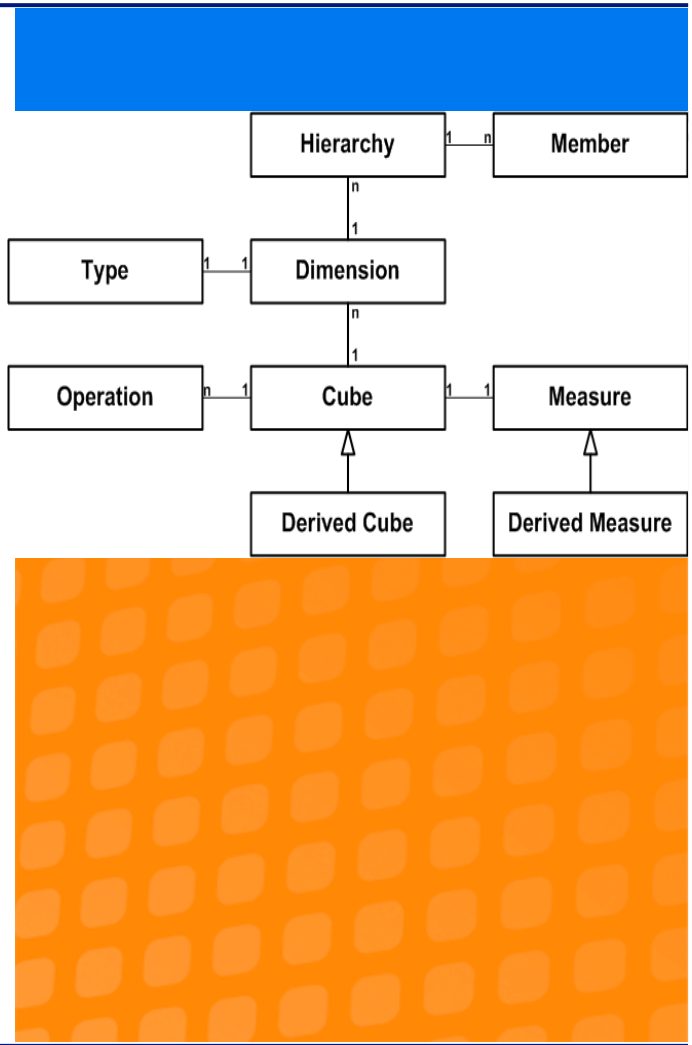


8 A Common Meta-Model for Data Analysis

Semi-Automatic Data Visualization

- Idea: Use a Common meta model for both approaches
- Why
  - Meta-model is needed
  - Reuse of concepts

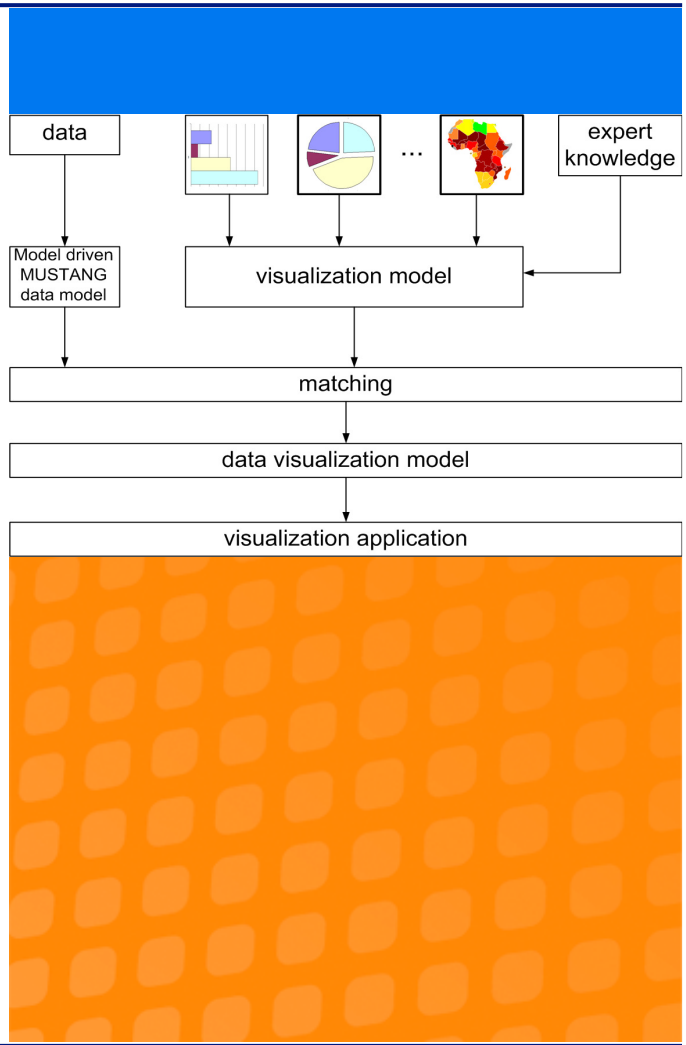




9 A Common Meta-Model for Data Analysis

Benefits for Visual MUSTANG

- Knowledge about data
- Presentable characteristics for
  - Dimensions
    - Numbers
    - Types
  - Hierarchies
  - Domains
  - …
- Generate appropriate visualizations
- Benefits for MD Mustang
  - Easy to integrate suitable visualizations
  - Higher customer satisfaction



OFFIS Health
INSTITUTE FOR INFORMATION TECHNOLOGY

10 Conclusion

- Cost effective realization of demand driven decision support systems
- Enhanced visualization
- Reduced realization time
- Higher user satisfaction
- → Usable for SMEs