Establishing an Integrated Data Sharing Process for Micro- and Metadata at Deutsche Bundesbank

Dr. Meike Becker and Anja Treffs
Deutsche Bundesbank, Statistical Information Management and Mathematical Methods
Data processed at Deutsche Bundesbank

- Banks
- Companies
- Securities
- Households
Motivation for a Microdata Hub

Research and analysis

- focus more on understanding complex interactions between actors/sectors and market segments as well as the interaction between the financial sector and real economy
- make higher demands in terms of horizontal and vertical consistency, granularity and comparability
- increasingly require preparation work in the department and by researchers themselves

Macroprudential issues

- are becoming increasingly important and new core business of the Bank
- distribution and structure analyses as well as risk assessments require rapid overview of several datasets
Challenges

Data Storage
- No integrated data storage
- Data storage is not designed for microdata

Data Quality Assurance
- No standardised and administrated data
- Lack of continuity

Documentation
- No general documentation available
- No routine update (prone to failure)

Resources
- Additional effort and expense by departments

Systematic solution with added value for research projects and analytical tasks as well as for data managing departments necessary
IMIDIAS: Integrated MicroData Information and Analysis System

Competence
Research Data and Service Centre (RDSC)

Governance

Infrastructure
House of Microdata (HoM)

Concept
Integrated Data Storage

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IASSIST Bergen, 02. Juni 2016
Seite 5
The House of Microdata
The House of Microdata

- IMIDIAS is based on an integrated microdata warehouse: **House of Microdata (HoM)**
- HoM is part of the existing statistical infrastructure for aggregated data at Deutsche Bundesbank
HoM is part of the existing infrastructure for aggregated data at Deutsche Bundesbank

- Data maintenance and analysis
  - Transformation (TAWOTR)
  - SDMX calculation program
  - Standard software
  - Seasonal adjustment
  - Zeva (incl. Host-JabSystem)
  - SDMX-BI

- Presentation
  - Intranet (ZIS)
  - Internet, (Bundesbank website)
  - Excel add-in (XLSHost)

- Publication and data exchange
  - Table generator (TTAB)
  - Charting tool (HighChart)
  - Internet publication (TS2Web)
  - Data exchange (GESGEN)
  - Ad hoc – enquiries

- Data input, batch and dialogue
  - Primary statistics
  - External sources
  - XML-SDMX, Excel, Text, csv
  - online update (XLSHost)

SDMX time series database ZISDB

Access-Portal for partners
Federal Ministries
- Federal Ministry of Economics and Technology
- Federal Ministry of Finance
- Federal Financial Supervisory Authority

Bundesbank's intranet
- Bundesbank departments
- Research projects
- Offices of regional office presidents

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IASSIST Bergen, 02. Juni 2016
Seite 8
HoM is part of the existing infrastructure for aggregated data at Deutsche Bundesbank

- Macroeconomic information system
  - Transformation (TAWOTR)
  - SDMX calculation program
  - Standard software
  - Seasonal adjustment
  - Zeus (incl. Host-Jobsystem)
  - SDMX-BI

- National accounts
  - Presentation
  - Intranet (ZIS)
  - Internet, (Bundesbank website)

- Public finances
  - Table generator (TTAB)
  - Graphic editor (HighChart)
  - Internet publication, TS2Web
  - Data exchange, GESGEN
  - Ad hoc – enquiries

- Financial market
  - Financial accounts
    - Data input, batch and dialogue
      - Primary statistics
        - External sources
          - SDDSPlus, SDDS, GERDA
          - XML-SDMX, Excel, Text, csv
          - XLSHost (online update)

- SDMX time series database ZISDB

Access-Portal for partners
Federal Ministry of Economics and Technology
Federal Ministry of Finance
Bundesbank’s intranet
Bundesbank departments
Research projects
Offices of regional office presidents

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IASSIST Bergen, 02. Juni 2016
Seite 9
Data diversity requires **standardisation**

→ International Standard SDMX (Statistical Data and Meta Data Exchange) is used as information model
SDMX in a nutshell

Metadata elements describe the area leading to systematic and self-explanatory indicator keys ("core" of the SDMX)

### Government statistics (GST)

#### Interpretation of key

<table>
<thead>
<tr>
<th>No</th>
<th>Description</th>
<th>Key</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Frequency</td>
<td>A</td>
<td>Annual</td>
</tr>
<tr>
<td>2</td>
<td>Reference area</td>
<td>AT</td>
<td>Austria</td>
</tr>
<tr>
<td>3</td>
<td>Adjustment indicator</td>
<td>N</td>
<td>Neither seasonally nor working day adjusted</td>
</tr>
<tr>
<td>4</td>
<td>Creditor user assets sector</td>
<td>B0MFI</td>
<td>Monetary financial institutions (ESA95)-NCBs</td>
</tr>
<tr>
<td>5</td>
<td>ESA item - GST context</td>
<td>MAL</td>
<td>Maastricht assets/liabilities</td>
</tr>
<tr>
<td>6</td>
<td>Debtor resource liabil sector</td>
<td>B1300</td>
<td>General government (ESA95)-NCBs</td>
</tr>
<tr>
<td>7</td>
<td>Valuation - GST context</td>
<td>SA</td>
<td>Financial stocks at nominal value</td>
</tr>
<tr>
<td>8</td>
<td>Series variation - GST context</td>
<td>H</td>
<td>Percentage points, series(t)/ debt(t)</td>
</tr>
</tbody>
</table>
The House of Microdata

**Generic approach:**
HoM can be used for any kind of data sets (statistical and non-statistical data, aggregated and microdata)

**Multidimensional approach:**
By using uniform code lists, SDMX offers an ideal means of linking and comparing data from different sources
Using the Power of the SDMX Information Model

Bergen, 02. Juni 2016
Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
Seite 13
Status quo
House of Microdata – How to choose the relevant data

1. Inventory list of existing data assets at Bundesbank

2. Ranking based on:
   - Relevance
   - Sustainability
   - Legal requirements
   - Costs

3. Selection of top 12 as first contents
House of Microdata – Data Integration via Content Projects

Data Source 1 → House of Microdata → Data Source 2

Data Source 3 → House of Microdata → Data Source 4

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IASSIST Bergen, 02. Juni 2016
Seite 16
House of Microdata – Data Integration via Content Projects
Phase I: Content Outline
- Overview, Variables
  - Content and Scope („What do we want to do?)

Phase II: Evaluation of Feasibility
- legal and technical
  - Decision about implementation

Phase III: Target structure
- SDMX-Classification Documentation,
  - Project requirements are clear enough for implementation

Phase IV: Implementation
- IT-Project
  - Process for initial and regular data transfer is established
House of Microdata – Status Quo of Content Projects

Phase I: Content Outline
- Overview, Variables
  - 4 projects

Phase II: Evaluation of Feasibility legal and technical
- 3 projects

Phase III: Target structure SDMX-Classification Documentation,
- 1 project

Phase IV: Implementation IT-Project
- 1 project

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IASSIST Bergen, 02. Juni 2016
Seite 19
House of Microdata – Status Quo of Content Projects

Phase I: Content Outline
Overview, Variables
4 projects

Phase II: Evaluation of Feasibility legal and technical
3 projects

Phase III: Target structure SDMX-Classification Documentation,
1 project

Phase IV: Implementation IT-Project
1 project

3 data sources already integrated:
- Monthly Balance Sheet Statistics
- Profit and Loss Accounts
- Master Data of Banks

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IASSIST Bergen, 02. Juni 2016
Seite 20
House of Metadata
IMIDIAS: Data and Metadata Flow

- External Researcher
- Internal Researcher/Analyst
- Metadata RDSC
- SDMX
- Metadata model
- Process data
- Metadata Department
- Clean Copies
- Research Data Sets
- HoM-Data Sets
- House of Microdata
- Steering-board
- Data experts in Departments

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IAASSIST Bergen, 02. Juni 2016
Seite 22
House of Metadata: What happened so far…

Metadata Catalogue

- SDMX – Metadata + additional DDI-oriented Metadata
- RDSC and data managing departments carved out a metadata catalogue with DDI-oriented elements

Consulting

- Consulting from an external company how to integrate a metadata management system into the IMIDIAS context in order to ease the data access for researchers and the work of the RDSC and regarding departments
- One relief for the data managing departments should be giving information only once
- Metadata Management System has to be able to communicate with the HoM Infrastructure and be able to extract the SDMX Metadata
House of Metadata: What happened so far…

SDMX Metadata DSD for IMIDIAS

<table>
<thead>
<tr>
<th>Werte</th>
<th>Reihenattribute</th>
<th>Erweiterte Reihenattribute</th>
<th>Wertattribute</th>
<th>Alternativschlüssel</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGGREGATIONSMGL</td>
<td>Erzielte Institute nach HGB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANALYSEPOTENTIAL_B</td>
<td>Die Statistik über Gewinn- und Verlustrechnung im Bankenwesen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANALYSEPOTENTIAL_E</td>
<td>Die Statistik über Gewinn- und Verlustrechnung im Bankenwesen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANALYSEPOTENTIAL_L</td>
<td>Vergleich der Ertragsentwicklung und des Ertragsanteils</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANSPRECHPARTNER_TE</td>
<td>Herr Winter, S 44-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BEMERKUNGEN</td>
<td>Bei Änderungen der Rechnungslegungsverordnung</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BESCHREIBUNG_AUSB</td>
<td>Position der Gewinn- und Verlustrechnung</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DATENGEBER_BEREICH</td>
<td>S 15-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DATENNUTZER</td>
<td>Vo. F, M, B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DATENSÆK_TECH</td>
<td>SAS/Regex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERHÄRBERGRUNDLAGE</td>
<td>jährlich</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERHÄRBERZEITRAUM</td>
<td>seit 1988</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>DSI</th>
<th>Beschreibung</th>
<th>Schlüssel</th>
<th>Erläuterung</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Periodizität (BBk)</td>
<td>HMHC1</td>
<td>jährlich</td>
</tr>
<tr>
<td>1</td>
<td>Datenkategorie</td>
<td>A</td>
<td>Banken</td>
</tr>
<tr>
<td>2</td>
<td>Ursprung/Quelle des Datenbestands</td>
<td>B</td>
<td>Statistik (Zentralbereich)</td>
</tr>
<tr>
<td>3</td>
<td>Inventurnummer</td>
<td>S</td>
<td>Inventar</td>
</tr>
<tr>
<td>4</td>
<td>Datentyp</td>
<td>S_4</td>
<td>GuV - Gewinn- und Verlustrechnung der Banken</td>
</tr>
<tr>
<td>5</td>
<td>Version</td>
<td>IV</td>
<td>Version 1.0</td>
</tr>
<tr>
<td>6</td>
<td>gültig seit JJJJMMTT</td>
<td>1_0</td>
<td>20160208</td>
</tr>
</tbody>
</table>

Dr. Meike Becker & Anja Treffs, Deutsche Bundesbank
IASSIST Bergen, 02. Juni 2016
Seite 24
House of Metadata: Under construction…

- Mapping between IMIDIAS Metadata DSD and the DDI-oriented elements of the metadata catalogue

- Tender for software development of a metadata management system

- Until a productive metadata management system is available the metadata elements will be filled through an Excel sheet
Conclusion

- IMIDIAS is a solution to address actual and future needs of researchers, analytics and policy making

- The root concept is a bank wide data integration, key features are standardization, harmonization, linking and service

- The underlying infrastructure is built upon a common information model – SDMX

- Data are integrated via content projects
Conclusion

- One key for utilization of the House of Microdata are metadata
  - Metadata Catalogue
  - Consulting for an Metadata Management System
  - Metadata DSD in SDMX
  - Mapping of SDMX and DDI-oriented elements
  - Tender for software development
  - Interim solution in place until a metadata management system is in place

- IMIDIAS rocks!
Thank you for your attention!

Deutsche Bundesbank  
Statistical Information Management and Mathematical Methods  
meike.becker@bundesbank.de  
anja.treffs@bundesbank.de