Changes in the IT landscape of the Deutsche Bundesbank’s external sector statistics\textsuperscript{1}

Joachim Hösch and Jens Walter,
Deutsche Bundesbank

\textsuperscript{1} This paper was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.
Changes in the IT Landscape of the Deutsche Bundesbank’s External Sector Statistics

Author: Joachim Hösch, Jens Walter

Abstract

In the past, the Deutsche Bundesbank’s IT systems have been undergoing an incremental evolution of applications and data bases. To better fulfil new statistical requirements of users which increasingly call for integrated and consistent data the Bundesbank has launched a project to overhaul the existing external sector statistics’ IT infrastructure and to create an integrated and automated system. This project is being implemented on a step-by-step basis since 2011.

The paper provides an overview on the key building blocks of the new IT-system based on a suite of integrated but modular software systems for master data, analysis, report processing and communication. The new harmonized IT infrastructure will facilitate integrated analyses for transactions & stocks, streamline statistical work processes and increase automation, and will therefore also reduce time-to-publication in the end.
Background

The Deutsche Bundesbank’s IT systems for the production of external statistics have been undergoing an incremental evolution in terms of their applications and databases. In the past, the prevailing stovepipe view on external statistics resulted in the development of single-purpose IT solutions using different IT standards for the production of Balance of Payments (BOP), Foreign Direct Investment (FDI) and International Investment Position (IIP) statistics, to name just a few. An IT landscape, which has continued to evolve over a number of decades, exhibits a high degree of complexity, even when it comes to making minor changes. The implementation of conceptual changes, such as the requirement to link transactional microdata to stock data, becomes a very challenging task. A growing technical backlog has therefore emerged in the shape of maintenance-intensive systems, a lack of documentation and a declining number of staff members knowing how to work with the complicated historically grown systems.

However, over the past decade, the integrated view on external statistics has become more and more common and has emerged as the new standard for reporting requirements. Progress in the field of information technology has made enhanced solutions possible. Integrated data models have been developed, which conceptually combine various data sources, thereby streamlining the production process.

Given the new data requirements and the improved IT solutions, the Bundesbank has launched a project to overhaul the existing external sector statistics’ IT infrastructure and to create an integrated and automated system. This project is being implemented on a step-by-step basis since 2011.

The strategy for the Bundesbank’s new External Sector Statistics (ESS) system

In order to cope with changing international reporting requirements and to develop a harmonised IT infrastructure, an integrated and modular software system was implemented. It comprises the following functionalities:

- Master and Meta Data
- Analysis
- Report Processing
- External Communication

The chosen modules allow standard software components to be used that are highly customisable and able to accommodate future requirements. This reflects the Division’s policy to reduce the in-house IT development of single-purpose IT solutions. It is also intended to be the nucleus of a streamlined statistical work process that exploits automation, thereby reducing time to publication.
Centralised master and meta database

In 2011, the development of a centralised master and meta database was launched as a flagship project; it went online in 2012. Prior to the initiation of this project, systems either had their own master data or only transactional/aggregate data without any master data components. Therefore, the harmonisation of all reporting units’ master data (more than 400,000) and meta data into a centralised database, which serves as a hub for the other software systems by means of various interfaces, is a major step forward.

In addition to the centralisation of data, the new database centralises the workflow for data updates and the creation of new units. This work is performed by designated staff members only so as to raise data quality and reliability. By applying the principle of dual control, the system’s workflow ensures that only approved data enters the data hub. A full data history allows transactional and master data to be properly matched, even for late reports. For each item, past data is available to allow the item’s entire history to be retraced.

Approach for analysis

The previous BOP database system for creating analyses and time series called ZABDB was a mainframe-based system, which had no Graphical User Interface (GUI), was unable to deliver ad-hoc analyses, nor could it accommodate new reporting and analytical requirements. The Division’s policy to move from single-purpose in-house IT solutions towards standard software components was applied for the first time with the installation of the new IT system SALZA. It is an SAP BI Server-based system, which provides a broad OLAP/BI (Online Analytical Processing / Business Intelligence) functionality out of the box.
By moving towards standard software – in this case with a modern, fast and robust client-server architecture – and having an MS Office (Excel) integration as a GUI, the system’s primary usage for analytical and methodical work is well supported. New users have also emerged, who use the system for the Data Quality Management (DQM) process for reporting data. This extended usage is feasible since the new system provides online calculated aggregates, which can instantaneously be drilled-down to the reporting unit, thus allowing flexible analyses.

The use of standard software produced excellent results by taking advantage of user-friendly GUIs that are well established on the market. An intuitive GUI proved to be an ongoing success factor in reducing user training efforts, maximizing the user experience and creating momentum to move from previous DQM approaches with other less integrated and flexible systems towards this new setup.

With the possibility to slice-and-dice in a cube-like body of information, users are able to dive into smaller parts of data or to examine it from different viewpoints. This increases the understanding of the data and enables compilers to receive new information. The greater flexibility also increases the capability to respond to new information needs, such as versioning the data to reproduce previous publication versions for DQM purposes.

Apart from the reporting flexibility, the SALZA project focuses on implementing system-integrated trend and estimation procedures, which means that it automates publication preparatory work and primary and secondary cell suppression as well. Since its implementation, SALZA has become the basis for SDMX (Statistical Data and Metadata eXchange) time series creation and for national and international publication commitments.

**Report processing and External Communication approach**

As a counterpart to SALZA, the planned report processing system SIMBA will provide an integrated platform and front-end system for report processing for all external statistics’ micro data. This will be a major improvement as various processing programs and databases are currently in use for the various ESS.
Apart from centralisation into a single database and a single front-end system, the current manual processing of each reporting line will change in favour of a system-supported data validation process by means of automated plausibility checks. Thus, staff can concentrate on invalid and implausible reports detected by the system using a dashboard approach.

Such dashboard web UIs are a well-known standard for tracking and monitoring transactional data – in this case reports – in real time. Its primary metric will consist of a simple traffic light system, which visualises reports that need correction (marked in red), followed by reports that may need correction or are valid statistical outliers (marked in yellow), and finally flawless and plausible reports that need no attention at all (marked in green). This will save time due to a reduced overall number of reports needing attention and increase the efficiency and productivity of human resources thanks to automation.

To allow full control of the DQM process, the first step needs to be to build a plausibility – engine-like – technical infrastructure. This is to be followed in a second step by a GUI offering intuitive technology for super users. The GUI should ideally allow technologically affine statisticians to maintain plausibility checks in natural language – irrespective of IT support.

Additional features are planned for the future which allow secure communication with the reporting units and send out reminders regarding missing reports. A further automation of plausibility checks is also envisaged, which will result in the auto-correction of data, ideally using a machine learning approach.

Target structure

The projects SALZA and SIMBA implement the desired target structure from the first submission of a report all the way through to the final publication of the data. The new IT infrastructure guarantees a fully digitalised process for the automated processing, integration and analysis of quality assured primary and secondary statistical data, complemented by an online publication of the data with a user option for customised data requests.
Target structure of the ES IT system

Reporters → Interface for E-Reports → Report processing SIMBA → Analysis and Aggregation SALZA → Time series and Publication

Master and Meta Data Base
IFC - Central Bank of Armenia Workshop on “External Sector Statistics”
Dilijan, Armenia, 11-12 June 2018

Changes in the IT landscape of the Deutsche Bundesbank’s external sector statistics¹
Joachim Hösch and Jens Walter,
Deutsche Bundesbank

¹ This presentation was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.
Changes in the IT landscape for External Sector Statistics of the Deutsche Bundesbank
Since the 70s, rapid increase of self-contained technical applications and data bases which exist side-by-side.

These isolated solutions made an integrated view on the data difficult (e.g. transactions & stocks).

Furthermore, the multitude of industry standards (like IMS, DB2, MS & Oracle for DBs) complicated the daily work and led to ..

Growing technical backlog, e.g. lack of documentation, maintenance intensive systems, declining number of staff members knowing to work with the old systems etc. 

need for replacement and redesign
Strategy for the Bundesbank’s “new” External sector (ES) IT-system

- Development of an integrated, modular software system for:
  - Master and Meta Data
  - Analysis
  - Report Processing and
  - Communication

- In order to create a harmonized IT infrastructure that:
  - Allows integrated analyses for transactions & stocks
  - Streamlines statistical work processes and increases automation
  - Reduces Time-to-Publication
  - Leverages emerging technologies by using industry standards

- The new ES IT-system has been developed in various sub-steps
In 2011, we started with the development of a centralized master and meta data base with following key features:

- Harmonization of all reporter master data into one single reference data set (golden copy)
- Keeping master data and meta data in same data base
- Centralization of data updates (and creation) through a workflow process by dedicated staff
- Creation of an interface to allow other applications online access to the master data
- Full data historization
  - Not only keeping track of changes, but being able to analyze which information was valid at each point-in-time.
  - Existing data must not be deleted (also not in case of reporter liquidation) or overwritten (also not in case of reporter mergers & acquisitions), but new rows can be created in the data base for changes of existing data.
Starting point: Centralized Master and Meta database
Example: Master Data view
## Approach for Analysis

<table>
<thead>
<tr>
<th>Previous system ZABDB</th>
<th>New system SALZA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainframe based</td>
<td>SAP BI-Server based</td>
</tr>
<tr>
<td>No micro data</td>
<td>Aggregated micro data per reporting unit</td>
</tr>
<tr>
<td>No GUI (Graphical User Interface)</td>
<td>MS Excel GUI (BEx-Analyzer) and Web Browser GUI (SAP Front End)</td>
</tr>
<tr>
<td>Performance limit reached with the changeover to BPM6</td>
<td>Broad OLAP/ BI functionality</td>
</tr>
<tr>
<td>No real OLAP/ BI functionality</td>
<td>Real-time data (updates on a daily basis)</td>
</tr>
<tr>
<td></td>
<td>Versioning of data (stand-alone snapshots)</td>
</tr>
<tr>
<td></td>
<td>Broad and flexible methods of analysis</td>
</tr>
<tr>
<td></td>
<td>Drill down capability to the reporter</td>
</tr>
<tr>
<td></td>
<td>Primary and secondary cell suppression (protection of confidentiality)</td>
</tr>
</tbody>
</table>
Old ZABDB vs. New SALZA

Jens Walter, Joachim Hösch
6/21/2018
Slide 7
Drill down capability of SALZA

- Switch (1) to previous position
- Switch (2) to the factual item level
- Switch (3) to the reporter level (with transaction codes)
Drill down example in SALZA

- Drill down from an highly aggregated data level to related reporters and reporting items

- Other reporting options are likewise available (e.g. branches of the reporters)

Jens Walter, Joachim Hösch
6/21/2018
Slide 9
SALZA’s achievements

- SALZA creates SDMX time series for national and international publication commitments
- SALZA merges the publication process within a single application
- SALZA makes swift replies to external and internal inquiries possible
- Flexible data analyses
- Standardized reports with MS Excel GUI (first quality check)
- Integrated trend and estimation procedures
- Source and decomposition information for aggregated data
- Slicing and dicing from aggregated data to reporter
Online analytical processing (OLAP) enables users to analyze multidimensional data interactively from multiple perspectives.

An OLAP cube refers to a multi-dimensional dataset, not limited to three dimensions technically but often visualized three-dimensional.

Business intelligence (BI) encompasses a multitude of strategies and technologies for data analysis. OLAP is the technique used for our BI-system.
**Excursus: The OLAP cube offers three analytical operations**

1. **Slicing** – Creation of a sub cube with one dimension less than before
2. **Dicing** – Creation of a sub cube with limited values but same dimensions
3. **Drill Down and Drill Up** – Navigation along the data levels, i.e. from most aggregated (up) to most detailed (down)
New approach for processing survey reports - SIMBA

<table>
<thead>
<tr>
<th>Current Report Processing</th>
<th>Future Report Processing using the new system SIMBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>In general, manual processing of each reporting line</td>
<td>Focusing on invalid and implausible reports using a dashboard approach</td>
</tr>
<tr>
<td>Data validation by staff</td>
<td>Data validation by system (automated plausibility checks)</td>
</tr>
<tr>
<td>Various processing programs and data bases per section</td>
<td>Integrated platform and frontend for report processing for all external statistics micro data</td>
</tr>
<tr>
<td>Division of labour per reporting form</td>
<td>Focus on all external economic activities of a reporter instead of single reporting forms</td>
</tr>
<tr>
<td></td>
<td>Dynamic crosstab analyses for transactions &amp; stocks</td>
</tr>
<tr>
<td></td>
<td>Plausibility editor to allow power users adding or modifying plausibility checks independent from support of the IT staff</td>
</tr>
<tr>
<td></td>
<td>One face to the reporter – only one Bundesbank employee responsible for one reporting unit</td>
</tr>
</tbody>
</table>
Planned SIMBA dashboard for report processing

![SIMBA Dashboard](image)

<table>
<thead>
<tr>
<th>Meldenummer</th>
<th>Meldepflichtiger</th>
<th>Branche</th>
<th>Transaktionen</th>
<th>Z10</th>
<th>Bestände Z5</th>
<th>Bestände K3/K4</th>
</tr>
</thead>
<tbody>
<tr>
<td>00023382</td>
<td>Süddeutsche Versicherung AG</td>
<td>6510</td>
<td>203 (12)</td>
<td>25</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>00054688</td>
<td>SIMBA IT-Solutions GmbH</td>
<td>1320</td>
<td>101 (17)</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0007654</td>
<td>Reiser KGsA</td>
<td>1200</td>
<td>112</td>
<td>26</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>0008811</td>
<td>Mertigol GmbH &amp; Co. KG</td>
<td>1000</td>
<td>516</td>
<td>14</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>0006463</td>
<td>Cyberdyne Productions SE</td>
<td>2800</td>
<td>1199 / 23</td>
<td>7</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0008447</td>
<td>Kimvans Sportsinnovation AG</td>
<td>1400</td>
<td>836 / 14</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0022233</td>
<td>Bonkhaus Sonneborn - Hot Assets KG</td>
<td>6401</td>
<td>967</td>
<td>8</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>0024687</td>
<td>Seidenstift Kosmetik KG</td>
<td>2000</td>
<td>407</td>
<td>7</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>0029208</td>
<td>BankersBrauer GmbH</td>
<td>1100</td>
<td>122</td>
<td>3</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>0030909</td>
<td>Fidel Instrumentenbau GmbH</td>
<td>1600</td>
<td>28</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>0050135</td>
<td>Umgutel Schukszpakt KGG</td>
<td>8100</td>
<td>74</td>
<td>2</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>0058666</td>
<td>Kuentemann Schuhfabrik GmbH</td>
<td>5000</td>
<td>152</td>
<td>3</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>0059323</td>
<td>Scharf Optik GmbH &amp; Co. KG</td>
<td>2475</td>
<td>101</td>
<td>13</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>0064820</td>
<td>Hortenberger Verkehr GmbH</td>
<td>8480</td>
<td>56</td>
<td>4</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>0069883</td>
<td>Ennad &amp; Sturz KGvA</td>
<td>3040</td>
<td>32</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00245682</td>
<td>Kaiserwerter Technologie AG</td>
<td>2700</td>
<td>241</td>
<td>11</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>0035481</td>
<td>Charlie-Bravo International SE</td>
<td>3030</td>
<td>184</td>
<td>15</td>
<td>1</td>
<td>36</td>
</tr>
<tr>
<td>0037323</td>
<td>Salto Mortale e.V</td>
<td>9310</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>00442468</td>
<td>Gemeinde Ötztal Wohnbau GmbH</td>
<td>6801</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Jens Walter, Joachim Hösch
6/21/2018
Slide 14
**SIMBA approach and objectives**

- Integrated platform and frontend for all BOP micro data
- Improve data consistency and therefore quality by a holistic view of a reporting unit
- Crosstab checks and analyses for transactions & stocks
- Process automation
  - System-supported plausibility checks
  - At a future stage, auto-correction functionality ideally with machine learned data
- Not only a new IT system, but process improvement by implementing traffic lights:
  - Intuitive status visualisation of a report
  - Automated work-in-process (WIP) overviews
  - Automatic processing of flawless and plausible reports
- Future extension of the base system to allow secure communication with reporters and a reminder mechanism for missing reports
Summary: Where we came from …
Summary: ... and where we want to be.

.. a fully digitized process for the automated processing, integration and analysis of quality-assured primary and secondary statistical data. Online publication of the data supply with the users’ option for customized data requests.
Thank you for your attention!
jens.walter@bundesbank.de