WIESBADEN GROUP ON BUSINESS REGISTERS

Assessing the IT environments of SBRs according to the Maturity Model

Péter Pálosi, Hungarian Central Statistical Office, peter.palosi@ksh.hu

Abstract

The SBR Maturity Model's sixth dimension describes the IT environment of the SBR, and in today's digital world, IT plays a crucial role in every system, especially in SBRs, which in the technical point of view, are database IT applications. Also this dimension is strongly connected with the Maturity Model's third dimension, which is the Maintenance and Update of the SBR, as a proper update and logging mechanisms in the SBR cannot be achieved without a solid IT environment. In addition with the starting of the widespread usage of artificial intelligence and machine learning technologies, the role of IT will grow even stronger in the future.

The session presentation will briefly give an overview of the IT environment of the Hungarian SBR, its strengths and weaknesses will be uncovered, and it will be assessed according to the SBR Maturity Model. The focus will be on the critical points, as what elements and aspects determine if the IT environment of the SBR is either in the preliminary, early, mature or in the advanced stage. The use of modern applications in itself does not necessarily mean the IT environment of the SBR is mature, the different software solutions need to work together seamlessly: there are different software for accepting incoming data from the administrative sources, for the data processing tasks and also the relational database management system itself, which holds the SBR's data should be configured and maintained properly as well.

The main goal of the presentation is to provide ideas for colleagues who would like to assess their SBR's IT environment, and as well as to give practical solutions and concepts for designing or improving the IT environment of an SBR.

1. Introduction

The Statistical Business Register (SBR) is the backbone of official business statistics, as statisticians and policy makers require timely and high quality data on business units, which only a well maintained SBR can provide. The SBR also has the function to create survey frames, as the SBR is also the main input for economic surveys, and as well as the main input of many different statistical domains, such as Business Demography, Structural Business Statistics, Foreign Affiliates Statistics and so on. SBRs can contain different kind of entities, there are multiple statistical units defined in business statistics. The legal unit (LEU) is the basic entity which every SBR should contain, as other statistical units are derived from the LEU. The SBRs can contain the local unit (LOU), kind-of-activity unit (KAU), enterprise (ENT) and enterprise group (EG) statistical units. In most cases, countries receive information about LEUs and LOUs from one or more administrative sources, and the other statistical units get compiled using the LEU and LOU as their base.

The users of SBRs require that the timeliness of the SBR should be improved, as they would like the SBR to receive the administrative data much faster, and to have better coverage, there are countries

where the SBR does not cover the whole economy. The size of SBRs can differ between countries, bigger countries in population tend to have more operating LEUs in their territory than other, smaller sized countries, thus their SBR could be significantly bigger, and require more resources to maintain. Even medium sized European countries have more than 1 million active LEUs, thus it is not feasible to manually load the data into the SBR, and therefore some sort of automated processes are required. The SBR is a quasi-database application: the data is stored in some kind of a database (preferably in a Relational Database Management System - RDBMS), and the data maintenance (record insertion and update) is performed by a program, depending on the implementation. It can be clearly seen that the SBR is the bridge between statistics and IT: designing, developing and maintaining an SBR requires expertise in IT infrastructure management, database software development, and as well as knowledge in the actual SBR methodology and in the legislation of enterprises, budgetary units and non-profit organizations is also required.

The history of the Hungarian SBR, and its current state and place in Hungarian Central Statistical Office (HCSO) and in the Hungarian public administration is briefly introduced in the next chapter. The third chapter contains the detailed IT aspects of the SBR: the underlying infrastructure, the logical and physical data model, and also the key parts of the program logic, which keeps the SBR up to date. In the fourth chapter, the maturity of the Hungarian SBR's IT environment will be assessed, what are the key aspects, which is actually advanced, and what should be significantly improved. This paper aims to present the Hungarian SBR, and assess its IT environment according to the Maturity Model, the main objective of the SBR Maturity Model is to provide a framework to assess the SBR, and its sixth dimension focuses on the IT environment of the SBR. The other goal of the paper is to share experiences, best practices and ideas for colleagues who are working with SBR in their respective countries, and those who also wish to assess, or improve the IT environment if their national SBR.

2. The Hungarian SBR

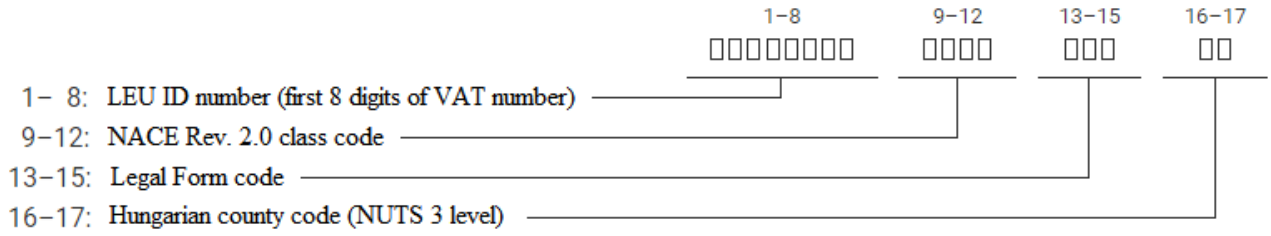
This chapter will briefly introduce the history of the Hungarian SBR and its current state, how does it connect to other systems within HCSO, and also with other state institutions in Hungary. In order to be able to have a clear view on the IT environment, it is important to get familiar with the underlying methodological aspects, and as well as the legislation regarding the LEUs in Hungary.

2.1. The history of the Hungarian SBR

The Hungarian SBR has come a long way since its inception in the 1970's, when HCSO started collecting information about resident business units in the country. Its IT environment consisted of IBM mainframe computers, the data storage was file based, complex algorithms were needed to update the records of the SBR. In the year 1997, the SBR was transferred into a RDBMS developed by Oracle, and due to the socio-economic changes during the 1990s in Hungary, the number of registered LEUs already exceeded 1 million units. Technically speaking, the Hungarian SBR reached its current base in 1997, although in the more than 25 years that have passed since then, it has undergone many changes.

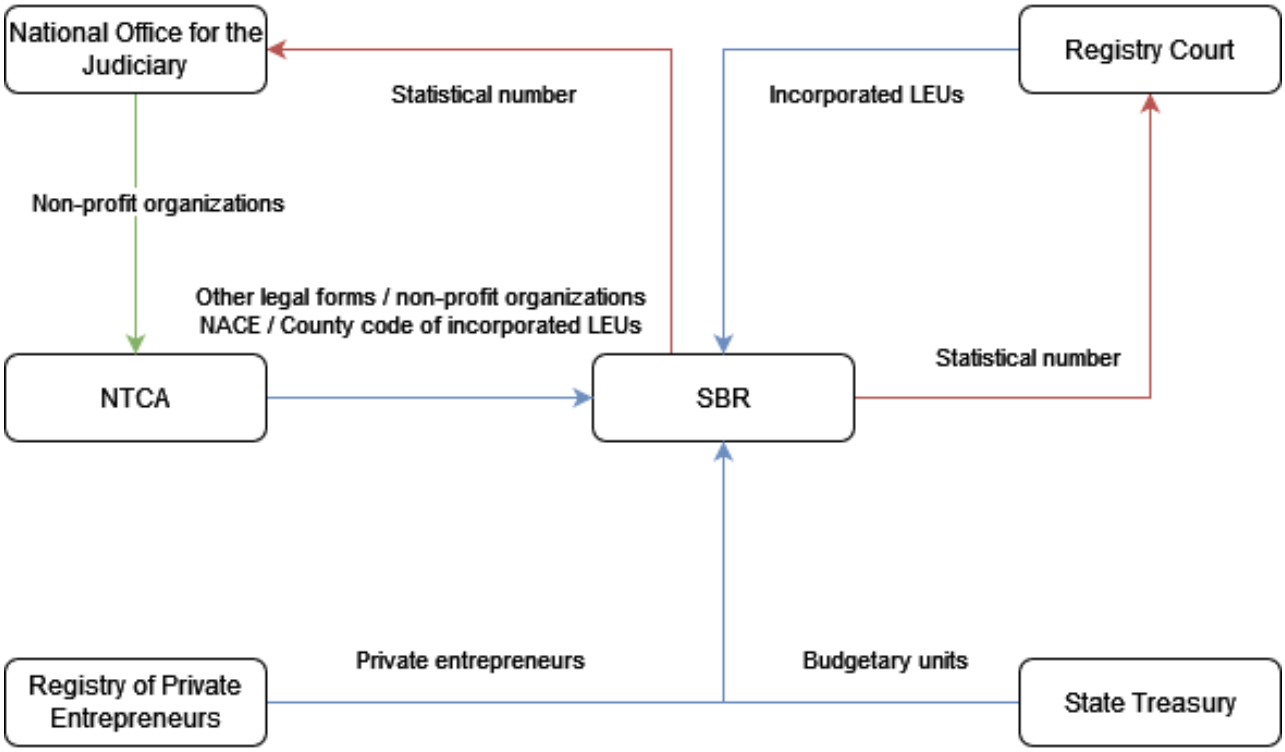
In Hungarian public administration, LEUs are registered by different state institutions based on their legal forms. Incorporated LEUs are registered by the Registry Court, private entrepreneurs by the Registry of Private Entrepreneurs, non-profit organizations by the National Office for the Judiciary, budgetary units by the Hungarian State Treasury, and every other legal forms by the National Tax and Customs Administration (NTCA). The Hungarian legislation permits the use of a unified identification system of LEUs, which is supervised by the NTCA, the unique LEU id is the first 8 digits of the VAT number. This identification cannot be changed, for example if a limited liability company wishes to change its legal form, it needs to liquidate the current LEU, and create a new with a different VAT number (the only exception is the private entrepreneurs). This means that in order to achieve full LEU coverage in the SBR, multiple administrative data sources have to be used, as each state institution only have information for LEUs with specific legal forms. Due to the fact that each LEU has a unique identification originating from the same system, it is very convenient for the SBR as this id can be used as the primary key. The Hungarian SBR not only receives data from other state institutions, it has an integral part in the public administration. HCSO generates the so-called statistical number of every LEU based on the SBR, and sends back this information to the state institutions. The statistical number is regulated by the Act No. CLV of 2016 on Official Statistics, and the elements of the nomenclatures included in the statistical number is regulated by the 21/2012. (IV. 16.) KIM regulation.

Figure 1. The structure of the statistical number



Throughout the years to significantly improve the timeliness of the SBR, the single window system has been developed. It means if there is a new LEU registration, a cessation or if only a specific attribute of the LEU is changed, the state institution automatically sends this information to HCSO, and the data gets loaded into the SBR. Before the implementation of the single window systems, the update of the SBR happened gradually: usually in a monthly or in a weekly basis, the partner state institution sent the LEU data changes either via CD, or by a large plain text file via a file sharing platform. As the IT infrastructure and expertise improved at the state institutions, the single window systems have been successfully developed by a joint work with HCSO. This process was fully complete, when the last institution, the Hungarian State Treasury built the connection with HCSO in year 2017. The single window systems were developed throughout the years, they are independent implementations, meaning these systems have different logic, they are not unified. Basically there are two kinds of implementations: the first one is the online system, which means the connection between HCSO and the partner state institution is continuous, data can arrive anytime throughout the day. The other one is the batch system, when the partner institution sends data at the end of the day, meaning multiple changes at once are present in a single message. This doesn't mean the SBR can be updated at any time, to ensure data consistency, the arrived data only gets validated and loaded in the next morning, before 7:00 AM. There should be a buffer zone in the database where the raw data gets loaded, and some kind of automated process to validate and load the data into the SBR. This mechanism will be elaborated further in the third chapter.

Figure 2. The Hungarian SBR and its administrative data sources



This figure shows that the partner state institutions send data to HCSO via the single window system in order to keep the SBR up to date. The red arrows mean HCSO sends back the statistical number, due to the legislation, HCSO only needs to send back the statistical number to the Registry Court, and to the National Office for the Judiciary. The reason is that if an incorporated LEU, or a non-profit organization changes its principal activity, or its main location to another county, they are obligated to announce this to NTCA. Then the tax authority sends the changed information to HCSO, thus the data gets loaded into the SBR, and if there is a change in a statistical number nomenclature, HCSO generates the new number and sends it automatically to either the Registry Court or to the National Office for the Judiciary depending on the LEUs legal form. The other three partner institutions are able to modify the statistical number accordingly if needed. The green arrow means that the National Office for the Judiciary is not sending directly the LEU data to HCSO, instead the data is being sent to NTCA, and it gets forwarded to HCSO.

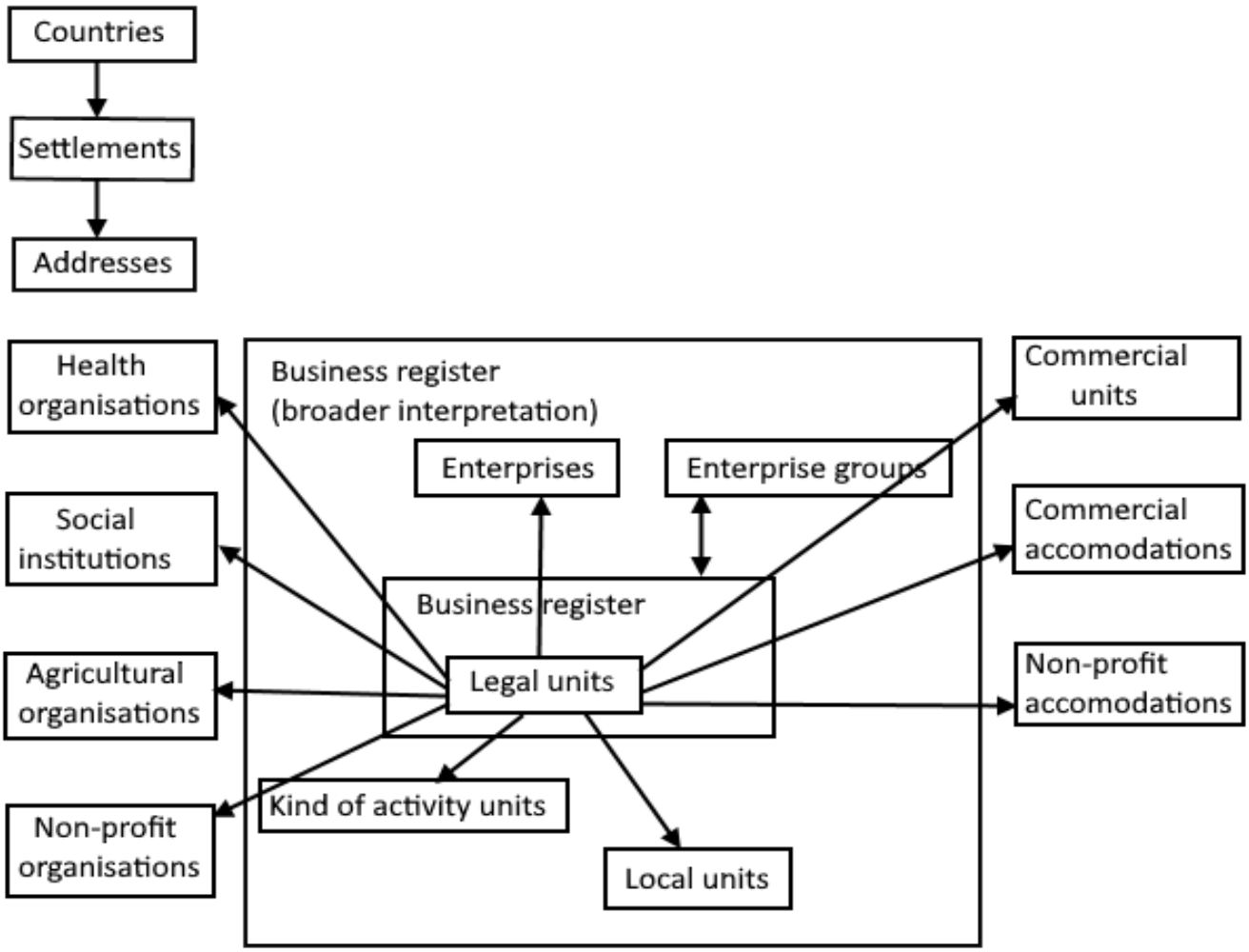
2.2. The Hungarian SBR in HCSO and its relationship with other registers

The SBR plays a central role in the HCSO, as it connects many other registers within the organization. This can be easily achieved by the unified LEU id system, the other registers also use the LEU id as

the primary key. The figure below shows a simplified view of the relationships of the SBR within the organization. Each rectangle represents a register, and the arrows show the direct connections between them.

The SBR also can be interpreted in two ways, there is a narrow, and a broader interpretation. It can be seen on the figure below, we can consider strictly the SBR is the register of LEUs, although if we take into account the other business statistical units such as LOU, KAU, ENT and EG, the SBR contains all of them.

Figure 3. The business register and its relationships in HCSO



Source: Vereczkei et al. (2017), translated by author.

The SBR needs to be constantly online, accessible for other systems, the statistical number must be sent back immediately, and also it should be available on the HCSO's website without interruptions.

At the beginning of each month, the monthly SBR frame gets created, which has two purposes

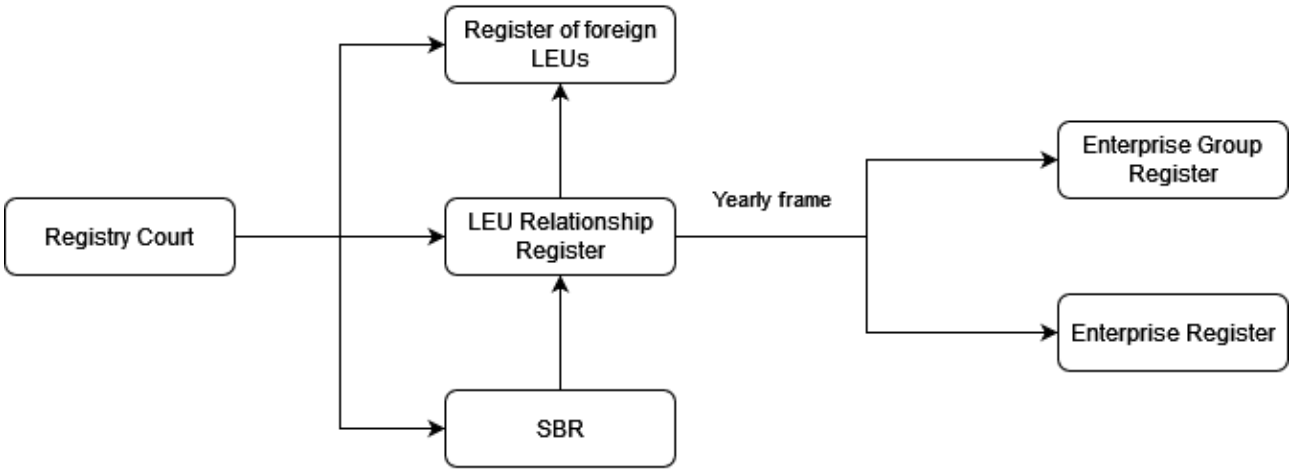
1. The input for the system which manages the surveys sent out to LEUs by HCSO
2. Creating a monthly saved state of the SBR for statisticians and for disseminating aggregated data (e.g. how many active LEUs are present in the country by NACE and legal form breakdown etc.)

The SBR is a live system, meaning its data is changing constantly, that is why there is a need for frozen frames: statistical data should be reproducible, although the SBR handles historical data, its usage for the entire population, not just for individual LEUs is complex, and cumbersome for statisticians. The frozen frames can be monthly, and yearly frames.

To be able to create an Enterprise Group register, we need the LEU ownership data, as who owns whom. This information is available at the Registry Court, and it is sent via the single window system for incorporated LEUs. The ownership, or in other words the relationship data is administrative, and it is changing constantly, thus an intermediate register, the LEU Relationship Register is needed. The entity in this register is the relationship between two LEUs, the relationship can be direct, or indirect, live or ceased and so on. Due to the globalized economy, the owner of a LEU might be a LEU which is originated from a different country. The Registry Court also has information about foreign LEUs, but without a unified identification number. As the SBR can only contain resident LEUs, a separate register is also required for the non-resident LEUs, which is the Register of Foreign LEUs. HCSO uses an in-house identification system for non-resident LEUs, although more than ten years ago Eurostat created the EuroGroups Register (EGR), where the main goal is to have a unified LEU identification system on a European level, LEID. HCSO is already sending its resident LEUs to Eurostat for generating the LEID, and also sends the non-resident LEUs for EGR Identification Service. To ensure the consistency between the SBR and the LEU Relationship Register, the LEUs present in a relationship must be present either in the SBR, or in the Register of Foreign LEUs.

The LEU Relationship Register is used for compiling the Enterprise and the Enterprise Group statistical units. These units also have their respective registers as the Enterprise Register and Enterprise Group Registers. Although they are called registers, they do not have a "live" version, meaning they only have a frozen state for each reference year as shown on the figure below.

Figure 4. The compilation of the register of enterprise, and enterprise group



In summary, the Hungarian SBR is a complex system, it not only holds the resident LEU data, it is in a live connection with other state institutions via the single window system, and HCSO is obligated by law to provide the statistical number for each resident LEU. The SBR is also connected to other systems at HCSO, and it contains the other business statistical units such as local unit, enterprise and enterprise group. Besides the LEU data, the non-resident LEU and LEU relationship data also arrives via the single window system, and they are used in the LEU relationship register and in the Register of foreign LEUs. The Hungarian SBR takes part in Eurostat EuroGroups Register data collection, micro-data is shared with other institutions and as well as the aggregated statistical data from the SBR is disseminated regularly.

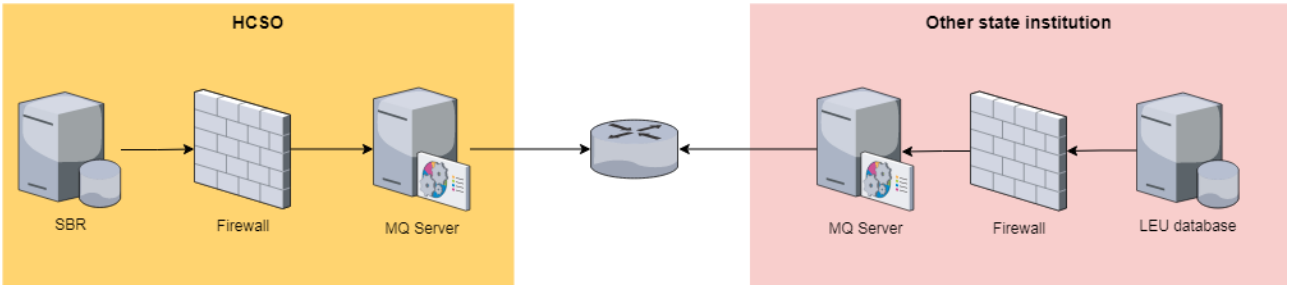
3. The IT environment of the Hungarian SBR

The Hungarian SBR's IT environment went through a significant change during the past 25 years. The manual batch data processing was replaced by the automated single window system, now the SBR is updated daily, by average the number of messages in a day arriving via the system is around 8000 records. This can be significantly higher at the start or at the end of a given month, due to administrative reasons (e.g. private entrepreneurs tend to suspend, or close their businesses on the last day of the month). The IT environment of the SBR is heterogeneous, it involves networking, databases, data processing software, automation, and also some kind of a data presentation, graphical interface for users, thus the IT environment can consist of multiple layers, such as the infrastructure, the methodology-business logic, the user-presentation layer. This chapter will present each layer's characteristics of the SBR.

3.1. The infrastructure layer of the Hungarian SBR

The infrastructure layer is the base of the Hungarian SBR, this layer is responsible for receiving the data via the single window system, and storing it in the database. In other words, this layer is responsible for data connection and data storage. This is not entirely SBR specific, as data transmission and storage is ubiquitous in nearly all IT systems, for example in transaction based applications such as online banking, online ticketing systems or in webshops etc. The figure below shows a simplified visual representation of the infrastructure layer, a connection to a single partner institution.

Figure 5. The simplified view of the infrastructure layer

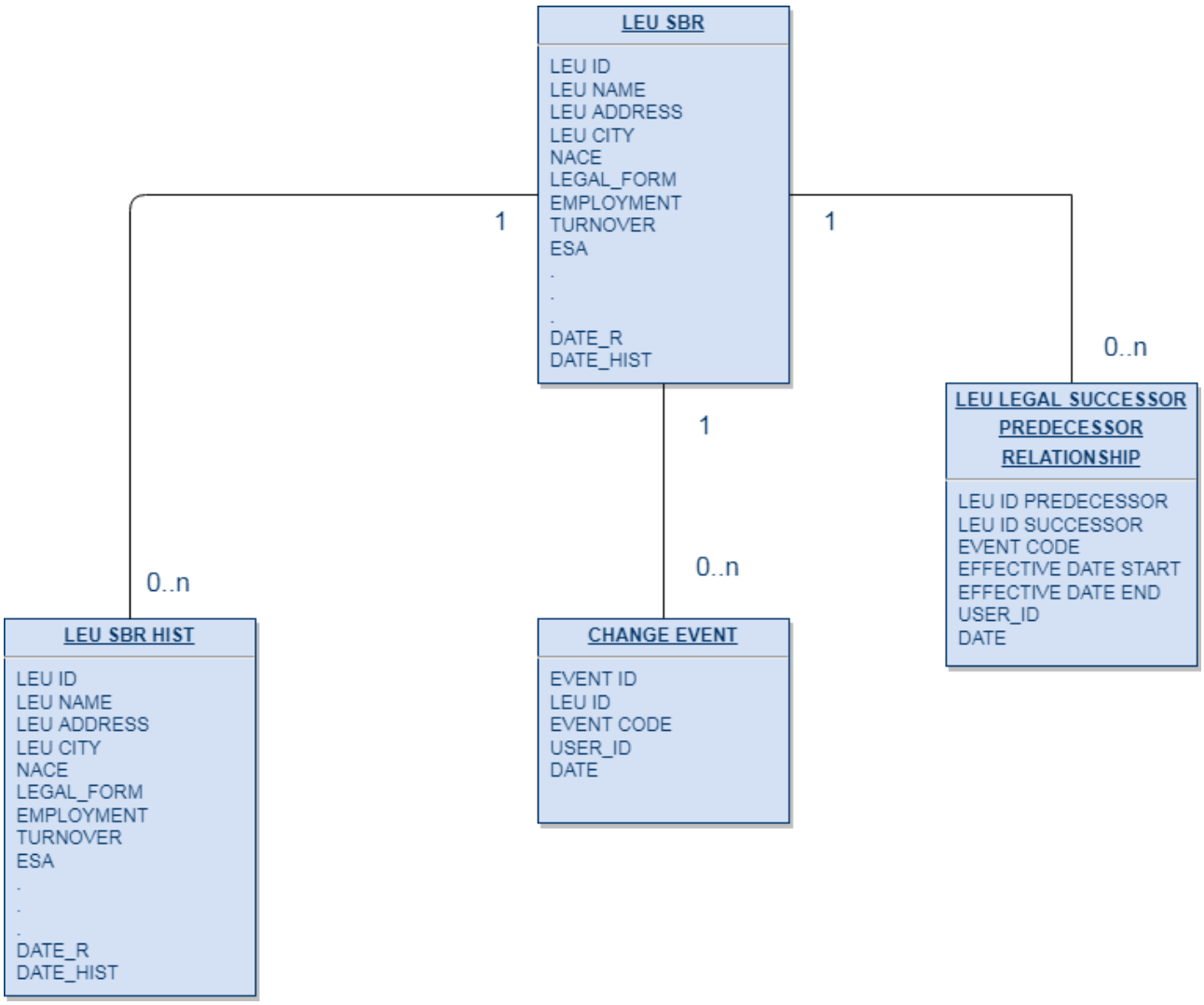


As it can be seen, the SBR is running on a RDBMS, which communicates with an MQ server. The MQ is a messaging system developed by IBM, which is suitable for asynchronous communication, meaning the SBR, and the partner state institution's system can work independently from each other. The MQ server puts the messages into a queue, in case there is a loss in connection, the message won't be dropped. In summary, the MQ as a middle layer is responsible for the messaging system, the state institution does not communicate directly with the SBR. This architecture demands that both institutions have to use MQ, therefore close cooperation between the organisations are crucial. The messages are originated outside of the organisations, thus the MQ server has to operate out the firewall, in a demilitarized zone, and this part should be configured. Also the frequent maintenance of the database should be done by the database administrators. This task also not SBR specific, it applies to virtually every running database. Most frequent operations of the database administrators are the management of tablespaces, access rights, registration of new users, and maintenance of tables as objects. A table can be fragmented after many insert, update, delete operations, which causes slow query performance, also the administrators can adjust the CPU, RAM quotas for each user, thus ensuring a reliable database performance for all users. Database administrators are often responsible for creating and maintaining the index on a table for speed optimisations too.

3.2. The methodology-business logic layer of the SBR

This layer is the core of the SBR, this is where IT meets statistics, with the SBR methodology. After data arrival, the next step is the data validation, and the update of the SBR. The data loading process is the core logic of the SBR, which is embedded in the program code which executes the operations, and also the physical data model is heavily determines how does the SBR function in the real world.

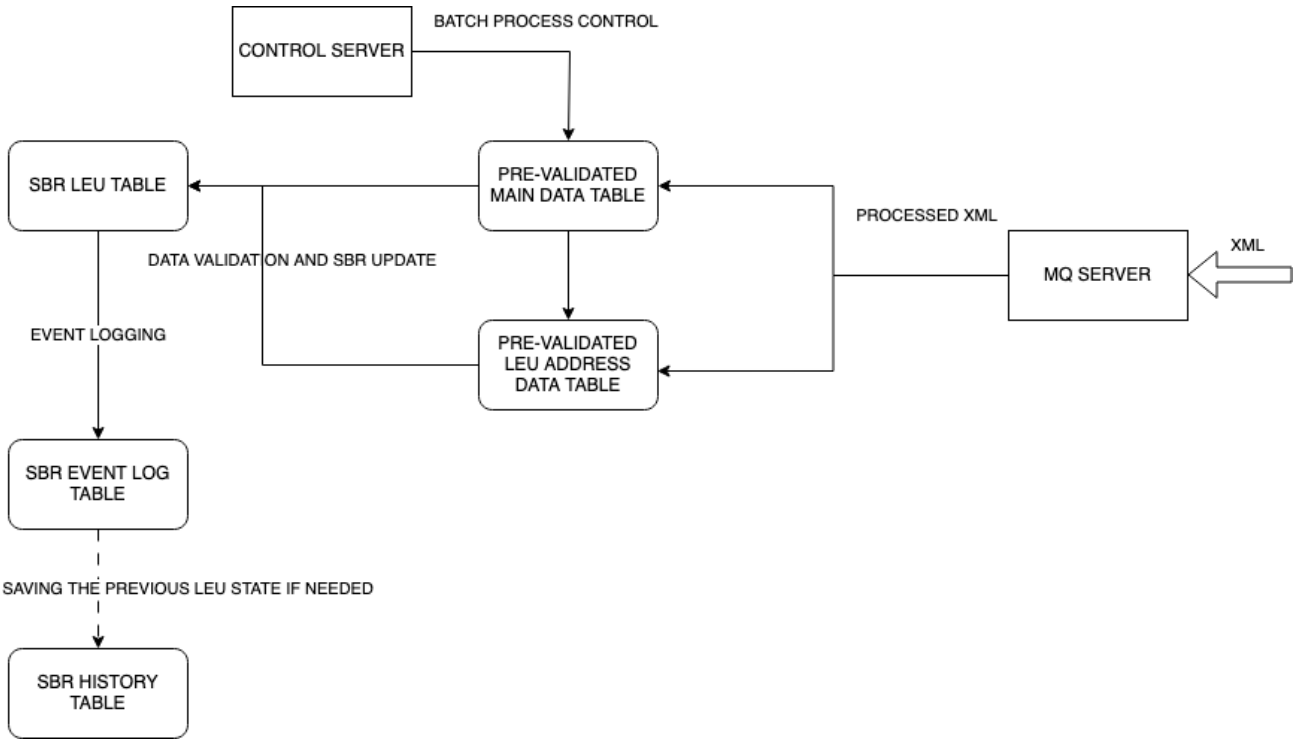
Figure 6. The simplified view of the physical data model of the LEU SBR



The SBR core consists of 4 tables: the main LEU table which holds every attribute, a history table, a separate table contains the legal successor-predecessor relationships between two LEUs, and finally an event change table. As soon as any data of the LEU changes, the previous states gets saved in the LEU HIST table, and every change has a corresponding event code, which is logged in the event

table. Although if a LEU newly enters the SBR, it does not have a HIST record, but still logged in the event table as a new registration event. This event table is appropriate when some sort of change event needs to be counted, for example how many LEUs changed their names in a given month, or how many LEUs were ceased in a given year. This could be also done with the use of the history table, but that would be a quite cumbersome task.

Figure 7. The logic of the data storage of pre-validated data and the update of SBR



The data arriving via the single window system is in XML format. The software which processes these files is controlled by the MQ Server, and it also archives them. Once the XML files have been processed, the software calls the procedures which checks the values for logical errors (for example the LEU id must be 8 digits, the legal form code must be valid and so on), and if there are no issues, the data gets inserted into the tables which hold the pre-validated data. The data is then separated into two parts: one table is holding the main attributes, and the other table contains the address data.

The timing of the SBR update can be done in two ways: the first one is to use the inner scheduler of the RDBMS, and the other is to use a separate system (e.g., a Linux or Unix server with cron[1]), which calls the data processing software. The first method implies that the data processing is done by RDBMS stored procedures exclusively, the second method means it could be done with any general-

[1] Cron is a command line utility job scheduler on Unix-like operating systems

purpose programming language, which supports communication with a RDBMS. The Hungarian SBR uses both methods, the single window system of the Hungarian State Treasury uses the first method, the processing software is implemented in Oracle stored procedures using PL/SQL, and the database scheduler calls the procedures at 6:30AM in the morning. The other systems – as it can be seen on figure 7 – use a separate system for timing the data processing, there is a control server which starts the batch processing by calling individual scripts. The Hungarian SBR uses PL/SQL for data processing, the program code is in individual .sql script files. The base logic of the data processing is that every record in the pre-validated table has a separate status attribute, which indicates if the record has been processed or not. The PL/SQL script decides based on this attribute to process a given record or not, and after processing the program updates the status column, thus the next day the record won't be processed. The program also saves the process date in the table, and also if there is any error in the process, it gets saved into a log table for debugging. This part of the processing consists of methodological validations instead of logical, for example if an incorporated LEU changes its address, the data is updated from the Registry Court, but if it changes its principal activity, the changed data arrives from NTCA, thus the programs have to check for legal forms if there are change in certain attributes. Another example is that there is a legal form, which is a private person with a tax number, this LEU is arriving from NTCA, and there may be multiple addresses for this LEU, which can be the private address of the natural person, or the temporary location. As a LEU can have only one address in the SBR, the program should prioritize the available addresses. The program should always take care of saving the historical data of LEUs, and also write the occurred events in the log table, also the logic of the data processing can differ heavily, as the partner institution can send individual modifying records, or send the complete record even if only one attribute changes, therefore the program's task to detect which attribute changed. Every attribute has a separate error column, which indicates the status of the update. The code 999 means the attribute is successfully updated in the SBR, the code 520 means there is no change, as the new and previous values are the same. Every other code means an error, thus the program did not update the attribute in the SBR, and the most common error is the issue of the effective dates. Consider the address of a LEU, which has an effective date, e.g., 2023.01.01. The partner institution sends a new address for the LEU, but its effective date is earlier than the previous, thus the program won't update the attribute, instead it will mark the attribute with the error code 410, which means the new data's effective date is earlier than the current one. Having an effective date for each attribute is crucial, for example if very old records accidentally get processed again (this can be done by adjusting the status column in the source table), the SBR would be overwritten with old data.

3.3. The user-presentation layer of the SBR

The user-presentation layer is the outermost part of the SBR, this is where users and other systems interact with the SBR. The safest way for users to interact with the SBR is via a Graphical User Interface (GUI), where statisticians can easily get information about LEUs, or modify specific attributes (e.g., employment data) manually if needed. It is not advisable to grant users insert, update and delete rights directly on database level. If a user changes a LEU attribute, the GUI is also responsible for writing the previous state in the history table, as well as the event needs to be logged in the event log table. This part is hidden from the user, the GUI does it automatically under the hood.

Not every user in the organisation has the right to update LEU information via the GUI, only the Register section and other sections under the data collection departments are allowed to do so. What it means is that the SBR needs to have User Access Control (UAC) management, the GUI has to support different user roles, such as query mode or administrator mode, where the LEU attributes are modifiable. There are cases when the pre-validated data needs to be manually investigated, to see why didn't get loaded into the SBR, therefore the pre-validated data table (as shown on figure 7) should also be browsable with the GUI. The single window system has the function to manually query a LEU record from a partner state institution, in case the due to an unknown error, the latest modification did not arrive. A button should be implemented on the GUI, if the statistician wishes to manually request the full records of the LEU from the state institution via the single window system, the record usually arrives within minutes, depending on the load of the network.

The monthly and yearly frozen frames, which are generated using PL/SQL scripts, and individual database queries from the SBR are also part of the user-presentation layer. As it was discussed in Chapter 2, HCSO not only collects LEU information, data from the SBR is actively published on aggregate level, therefore every program and script which reads data from the SBR is considered to be part of the user-presentation layer of the SBR.

4. The maturity of the IT environment of the Hungarian SBR

According to the SBR Maturity Model, the sixth dimension assesses the IT environment of the SBR. The Maturity Model considers the IT infrastructure is responsible for data storage, maintenance, update and dissemination of SBRs, a combined set of hardware, software and networks. In addition, it clearly states that people, processes and documentation are not part of IT infrastructure, therefore this will not be assessed. Four different stages are defined, which are preliminary, early, mature and advanced.

The table below summarizes the main characteristics of the IT environment in the different maturity stages:

Maturity stage	Description
Prelimiary stage	There is no integrated IT infrastructure, records are kept manually, usually in Excel sheets, the maintenance is very basic.
Early stage	The IT infrastructure consists of a simple database structure containing only essential information. There may be separate applications for SBR maintenance, but not in an integrated way, the software structure is fragmented.
Mature stage	The IT infrastructure of the SBR is well developed and scaleable, and supports every required task such as updating of records, data import, data validation, historical data management and the creation of frozen frames.
Advanced stage	The IT infrastructure for the SBR is part of the organization's integrated production system, and additional features are constantly being added and improved such as web scraping, big data, graph databases and cloud environments.

According to the stages, the Hungarian SBR can be put in the mature stage, considering that the infrastructure is developed, many SBR tasks are automated, and the manual tasks are also aided by IT tools (GUI). In addition, the Hungarian SBR is part of HCSO's statistical production system, although it does not qualify as being in the advanced stage. The SBR does not use any new techniques such as big data and web scraping, and due to legislation, cloud environments cannot be used for storing official statistical data if it is provided by a 3rd party organisation.

4.1. Analysis of the Hungarian SBR's IT infrastructure

In chapter 4. the Hungarian SBR was evaluated as its IT infrastructure is in the mature stage. What it means that the IT environment is well developed and efficiently handles the necessary SBR tasks, although lacks the use of the latest technologies. The other aspect of being in the mature stage is that the IT environment still needs constant maintenance, as IT is a domain which changes quickly over time. A certain implementation may work well for years, but it's general support (e.g., vendor or community support) may decline over time, which can lead to difficulties maintaining the system. Every implementation has its strengths and weaknesses, which can naturally change over the years: for example, one kind of implementation was the best ten years ago, which may not be feasible today due to different reasons.

The table below collects the strengths, weaknesses and an alternative implementation of the first layer, the infrastructure of the Hungarian SBR.

Function	Strength	Weakness	Alternative
Data connection – IBM MQ	High performance, full customer support, highly configurable	Proprietary (closed source) implementation, requires specialized advanced expertise, requires the other party to use it	Open source web service implementation (or developed in house) in a popular programming language (C# or Java)
Data storage – Oracle RDBMS	High performane, full customer support, high functionality (PL/SQL stored procedures, triggers, job scheduler)	Proprietary (closed source) implementation, costly license, requires specialized knowledge	PostgreSQL - open source free software, supports imperative SQL as PL/pgSQL

The general issue with proprietary closed source software is that although the vendor may constantly improve the product and provide customer support, the user organisation may have difficulties replacing that software due to the closed integration with other systems. In order to replace an RDBMS, the organisation needs to review hundreds of thousands of lines of program and SQL codes, which is not feasible on the short term. Another issue organisations may face is that universities tend

to teach different IT environments (such as MySQL, Python), therefore new colleagues may have a steep learning curve to know the existing IT systems.

The general status of the methodology-business logic layer is similar to the infrastructure layer. The core logic of the SBR is written in PL/SQL, and it determines that the database have to be Oracle, as PL/SQL only works in Oracle RDBMS. PL/SQL is based on the Ada programming language, its syntax is highly different from the common C based languages such as C#, C++ and Java, and also differs from other highly popular interpreted languages such as R and Python, thus it may also have a steep learning curve. PL/SQL is robust and very efficient when it comes to database applications, although its usage ties the organisation to a proprietary RDBMS system, which may have high license fees. Migrating hundreds of thousands of lines of PL/SQL code into another language is only feasible if the organisation is willing to invest a huge number of resources into the process. Another weak spot may be the usage of a different control server for timing the batch processes. In the current implementation, a Unix server is timing the batch scripts, the cron is configured to run the scripts on weekdays between 2:30 AM and 5:30 AM, the scripts are in plain text format in a simple directory. If the server itself suffers a downtime, or just the connection with the RDBMS becomes unstable, the overnight batch processes may not start, therefore the update of the SBR does not happen in a given day. The most obvious solution is that the scripts should be converted into stored procedures which can be called by the inner scheduler of the RDBMS resulting in a more reliable operation. The downside would that it would make the dependence of the proprietary RDBMS even deeper.

At the centre part of the methodology-business logic layer is the main table of the LEU SBR. The main table holds every LEU attribute; thus, it is a large table with many columns as shown on figure 6. The LEU table also contains statistical data from primary, non-administrative sources such as employment and turnover data, and also data compiled by HCSO itself, e.g., ESA code. This design is not entirely compliant with the spine data model, which states that the main table should only contain the core administrative data, and every other attribute should be stored in a separate table, thus the whole register could be built-up with data-linking.

The user-presentation layer also prone to the issues described in the other two layers. Graphical User Interfaces work by running program code just as “regular” applications. The Hungarian SBR’s GUI interfaces also use the Oracle environment, what it means that in order to change the database, not only the core logic, but the GUI code has to be built from the ground. A feasible alternative would be to independently redesign the GUIs using e.g., Google Angular, which is open source and has stable community support.

5. Summary and conclusions

It can be considered that the Hungarian SBR's IT environment is in the mature stage. The data storage is well organized, and the SBR update is done by integrated software solutions, also users have a convenient way (GUI) to interact with the system. Although it is in the mature stage, it doesn't mean the core functionalities should not be improved constantly. The main issues with an integrated system that IT is a dynamically changing field, environments can become obsolete in a relatively short period of time, and also proprietary software solutions can tie organizations to a specific vendor, which may increase the costs significantly. The Hungarian SBR are facing with this aspect, and many statistical offices in different countries are working on how to transfer their systems to open-source environments. The biggest challenge is that the core functionalities are embedded deep within program code, it requires a huge number of resources to refactor the old code, and thoroughly test the new implementation. A strong cooperation between the SBR methodology team and the IT department is crucial, as colleagues working in SBR section should have a strong foundation in IT expertise such as programming and database design, and as well as IT experts need to familiarize themselves with SBR methodology, to be able to cooperate efficiently.

6. References

Vereczkei, Z. et al. (2017), Másodlagos adatforrások használata a statisztikában. KSH, Budapest.