

Meeting of the Expert Group on Business Registers 27-29 September 2017, OECD, Paris

Christophe Chiri
Federal Statistical Office of Switzerland

TOPIC: Work on linking the SBR to geo-spatial information

Establishing sustainable solutions, increasing cost-efficiency, sharing good practices and addressing implementation problems.

From BUR2000 to SBER, improving address management

Introduction

The Business and Enterprise Register (BER), via its BUR2000 application, is seen by many of its users as an address register. Each administrative or statistical entity in the register has its own address, comprising information that enables geolocation and correspondence by post. With some 800 000 active enterprises, over 2 million addresses have to be maintained in order to meet the needs of internal and external FSO users.

Currently, the Federal Register of Buildings and Dwellings (RBD) is used as a source for codification, plausibility testing and support in entering addresses in BUR2000. The BUR2000 application is, however, being re-engineered, presenting an opportunity to improve processes and the management of addresses. This document sets out the thinking behind the modernisation of address management in the Business and Enterprise Register and the work already undertaken at the FSO.

BUR2000 and address management

Basically, BUR2000 address management is based on a copy of the main data from the RBD, known as the RBD cache. In addition to helping users enter data, this RBD cache allows new or modified addresses to be tested for plausibility by comparing them with valid RBD data. This system is central to several operations, guaranteeing the high level of quality expected by the register's users. These operations and the users' requirements in relation to the RBD cache or to BUR2000 addresses in general are listed below.

1) Daily automated synchronisation of RBD cache data with the 'live' RBD.

Active data contained in the RBD concerning towns, communes, roads, buildings and building entrances are copied in the RBD cache. The main reasons for doing this concern processes for the plausibility testing of addresses but also to back up the support functions for entering data in the client BUR2000.

One of the problems with this reproduction of data is that the modelling is slightly different in the RBD cache in BUR2000 and the original RBD. In the RBD there is no relation between a delivery postal code and a town, whereas this is the case in the BUR2000. This difference is a legacy from the old application that has been kept in BUR2000.

Another problem is that BUR2000 does not take into account "invalid" addresses such as demolished buildings. Synchronisation must, therefore, exclude these objects, meaning that modifications, regardless of which side they occur on, can have an impact on the reproduction scripts. Consequently, the latter have to be supervised and monitored to avoid synchronisation problems or the reproduction of undesirable

data. Faulty synchronisation will result in the spreading of errors in the register and among its partners.

2) Plausibility testing of addresses during automated processing

When a unit's address is created or modified, the information supplied by the clients is compared with existing RBD cache data. If the address is unequivocal, the change is automatically applied. If no or several results are found, the change will be processed manually.

One of the main problems with automated processing is that the uniformity of the information transferred cannot be guaranteed. Some partners code their addresses carefully. Others send poor quality, badly formatted data into the channels for announcing changes. This situation is the result of BUR2000 policy, which aimed to encourage the fast integration of new partners, at little expense to themselves, rather than standardising channels for the exchange of information. Over a period of fifteen years, the partner registers were gradually linked up to BUR2000. Lastly, with the arrival of the UID (Business Identification Number) in 2011, the system became increasingly complex and at the same time decreasingly stable and manageable.

3) Help with entering data in the GUI

RBD cache data are used to supply the input screens for addresses in the BUR2000 client. This means that users can choose from lists of possible values to code addresses correctly and efficiently.

4) Update of existing addresses via a weekly automated process

To have an address list that is as up to date as possible and to keep BUR2000 updated, changes made to RBD are also applied to BUR2000. Changes sent via this channel are for example renamed streets, changes to street numbering or commune mergers.

This process also has to be monitored to prevent certain changes from "polluting" BUR2000 and its partners. Such a case occurred recently when several communes revised the quality of building measurement. The coordinates of all the buildings in a town were corrected; the correction was a matter of only one to five centimetres but it generated irrelevant modifications to all of the commune's enterprises. Each one of these modifications then spread to each of the register's partners, thus creating a very large volume of unnecessary changes.

5) Controlling the quality of addresses

But even the update team's meticulous controls and the system's logic cannot prevent certain address coding errors from spreading. For this reason regular quality control checks are made. Any error lists that do arise can then be processed manually or by script depending on the complexity and the number of cases that have to be dealt with.

6) Corrections following survey responses and response management

Any corrections sent directly by enterprises and responses from the post office are integrated as a whole in BUR2000. However, such operations require human intervention in preparing data for uploading.

These various operations require numerous resources. The production teams have to cope with a growing number of changes making it increasingly difficult to decide which changes should be given priority. Due to the multiplication of sources and channels, maintaining and developing BUR2000 has become increasingly complex. Test phases also require much effort. Often the solutions chosen do not fulfil the requirements exactly but are a compromise between the complexity of their implementation and the advantages they bring.

Despite our best efforts and in spite of the most advanced quality analyses, we have to admit that the system is becoming outdated and that inconsistencies and errors in address data are becoming more and more common. The BUR2000 re-engineering product is, therefore, a timely opportunity for re-thinking certain parts of the system.

From BUR2000 to SBER

The BUR2000 re-engineering project (RG-BUR project) began in 2015. The original plan left the data model unchanged; the intention was only to modernise the technological platform. Two years later, the project has moved on and improving the processes for managing addresses has become an important issue, as the planning phase revealed that there was great potential for improvement.

As a reminder, the following objectives were included in the RG-BUR project:

- Standardise reporting interfaces with partners as far as possible.
- Improve and rationalise address management, reduce burden on production teams and make the application more flexible.
- Encourage the automation of repetitive processes.

These points were included as much as possible in our considerations during the analysis and planning phases.

The project is still being implemented but part of the definite solution has already been developed. The technological platform has been adapted and the two-tier application (C++ / Oracle) BUR2000 will be replaced by a three-tier application (C# / MSSQL) at the end of 2018. The new application is called SBER (Swiss Business and Enterprise Register) and is based on a modern service-oriented architecture. The data model, and in particular the part related to addresses, has also been adapted to align as closely as possible with official standards used within the RBD, GeoNO (Ordinance on Geographical Names) and the administration in general.

New architecture

The RBD already provides many services both within and outside of the FSO. However, as the needs of the SBER exceed the services currently on offer, rational solutions have been sought, including the possibility of using services from other sources.

The Swiss Federal Office of Topography (swisstopo), provides a range of services (REST) enabling very efficient full-text searches. Furthermore, the data made available by these services comes directly from the RBD. It was, therefore, a logical step to adopt this interface to cover requirement no. 3 with regard to support for data entry mentioned above.

The RBD services already enabled addresses to be completed and tested for plausibility but did not include all the characters needed by the SBER, such as the address language. As both registers belong to the same FSO division, we were able to adapt them quickly and today they already allow need no. 2 to be met with regard to the plausibility testing of addresses mentioned above. A new service has been developed enabling a list to be obtained of changes made to the RBD during a certain period to meet the need to update existing addresses (no. 4).

As these needs have now been met, use of the RBD cache can no longer be justified and the need to synchronise (no. 1) with the RBD can also be ignored, reducing the tasks of the SBER. Fears with regard to the availability and workability of the services were also raised but SBER's productivity should not be impeded, as the two sources mentioned above will work correctly far longer than needed by the SBER.

New model

In the physical BUR2000 model, each unit has its own address. This means that the number of records in the table of addresses is the same as the number of units located in the same building. Each address from the RBD is completed with textual data (e.g. street, town) as well as identifiers (EGID, EDID, etc.). The SBER on the other hand, is based on a concept in which addresses are items in their own right. This means that several SBER units can share the same address.

Whereas BUR2000 mixed together information on localisation and addressing for one address, the SBER approach is to sort the addresses that it references into different types. This sorting by type enables the level of detail supplied by this address to be identified directly. The differences between these types are described below.

In SBER, an address from RBD is identified as a “Swiss Local Address”, but can take different forms depending on its level of detail. The most detailed level is when the building and the entrance are identified with absolute precision. This type of detail is known as a “Swiss Building Coded Address”. This is identified by an EGID (Federal Building Identifier) and an EDID (Federal Entrance Identifier).

Sometimes the building does not exist in RBD and the operator can decide to associate it with a unit in a RBD street. This is known as a “Swiss Street Coded Local Address” for which the street is identified by its ESTRID¹. However, it is always possible to attribute an entrance number in this street, as this characteristic is located at the level “Base Local Address”.

If no RBD street corresponds to the one the user wishes to enter, the address is entered as a “Swiss Town Coded Local Address” with its “street” characteristic as optional, allowing a street name to be registered nevertheless. Until now in the RBD, the legal obligation to register buildings applied to residential buildings only. Some commercial buildings are, however, registered, but registration is patchy. To comply with the new ONGéo² and ORegBL³ ordinances, which came into force this year, the RBD has to contain all buildings in Switzerland. This means that the number of “Swiss Town Coded Local Address” should gradually decline in favour of “Swiss Building Coded Address” as the RBD incorporates the addresses of these buildings.

The following diagram shows the modelling of our Business Object Model for the address section of the SBER. Arrows represent legacies and the other relationships show the associations between objects and their cardinalities.

¹ In the medium term, the ESTRID will be replaced by the ESID.

² Ordonnance sur les noms géographiques : <https://www.admin.ch/opc/fr/classified-compilation/20071090/index.html>

³ Ordonnance sur le Registre fédéral des bâtiments et des logements : <https://www.admin.ch/opc/fr/classified-compilation/20162291/index.html>

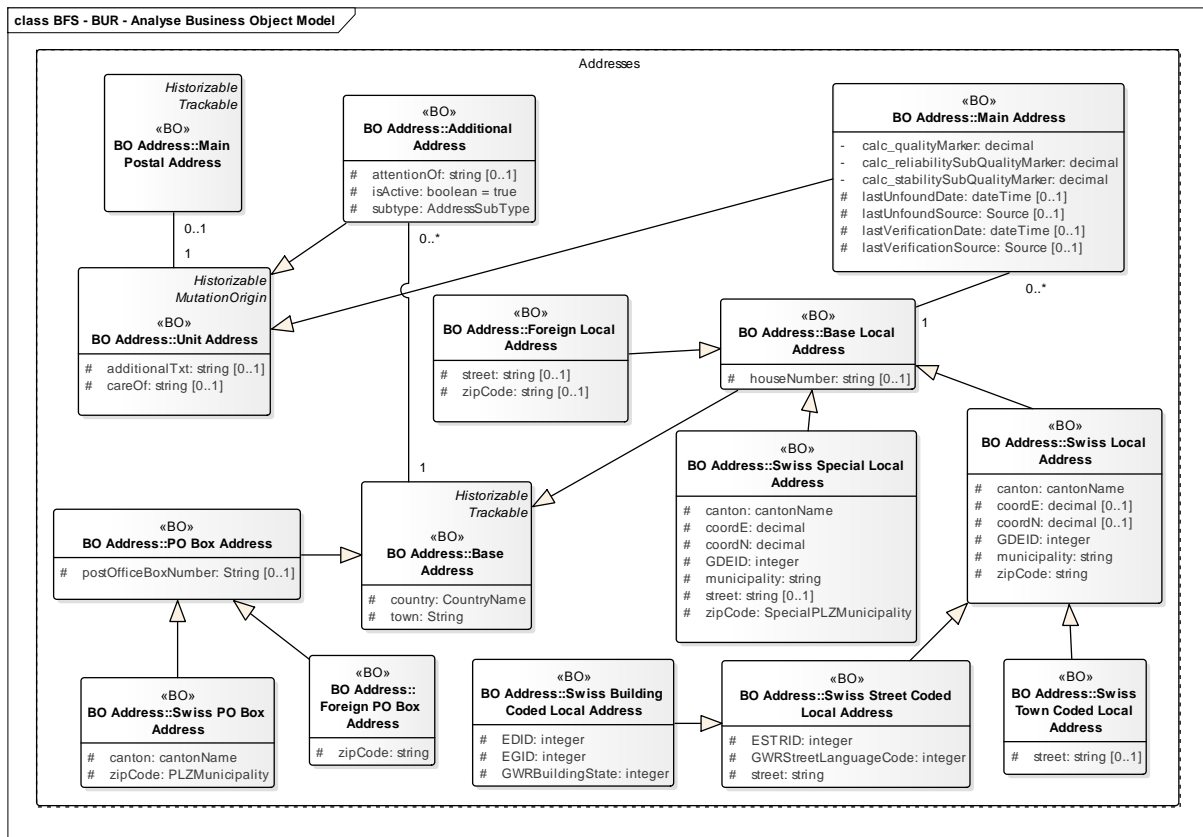


Diagram 1 - SBER Addresses business object model

As the reality of a building and dwelling register is not the same as an enterprise register, certain entity addresses in the SBER cannot be associated with RBD objects. Special cases that need to be processed in a specific way are:

- PO Box Addresses: A PO box is not a building and by definition does not exist in the RBD.
- Special Local Addresses: Certain administrations tolerate generic addresses, recognised by the Swiss Post but that are not official.
- Foreign addresses and PO boxes: The RBD only registers Swiss addresses. It was, however, possible to work with foreign units in BUR2000 and it is necessary to keep this function in SEBR.

This means that the new model alone cannot guarantee that the quality of data is improved (need no. 5 - quality assurance) or that the addresses are more up to date. It should, however, allow better identification of the types of addresses contained in the register. Furthermore, generally speaking, the quality indicators, data sources and audit and historic data should be managed better in SBER, enabling simpler but also more relevant quality analyses.

The address model intended for the SBER takes into account general guidelines on statistical business registers⁴, as well as the restrictions of BUR2000 and its partners. The different addresses used within the enterprise and its legal and local entities are divided into three groups described in the following diagram and chapter.

⁴ https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

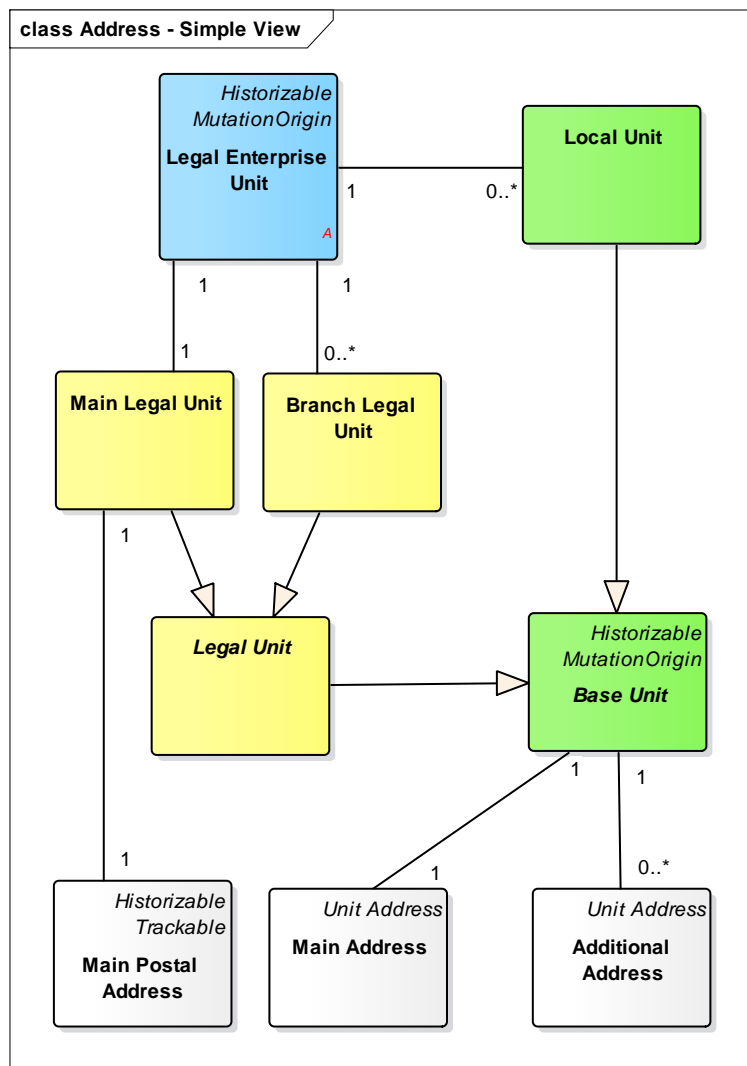


Figure 2 - Simplified SBER model

Each local statistical or administrative unit has a “Main Address” and must be geolocalised at least at the level of commune for addresses in Switzerland. The main legal unit of the enterprise references a “Main Postal Address” and as it is a delivery address, geolocalisation is not compulsory. The same applies to “Additional Address” which is intended more for correspondence.

New challenges and potential for improvement

Today, addresses received from partner registers are of mixed quality and often based on information from the Swiss Post. In most cases, an address announced to BUR2000 (and in future to SBER) is supplied directly by the business itself to an official administrative register. If addresses are checked and tested for plausibility on the basis of reliable data, the burden on the FSO entailed by further checks can be considerably reduced. For this reason the Federal Statistics Act, which came into force on 1 January 2016, stipulates that RBD data can now be used not only by the cantons and communes in carrying out their tasks but also by the Confederation’s services.

This has not yet been implemented but the process is under way: at its sitting of 9 June 2017, the Federal Council approved⁵ the complete revision of the Ordinance of 31 May 2000

⁵ <https://www.admin.ch/gov/fr/accueil/documentation/communiques/flux-rss/par-theme/communiques-de-presse-et-discours.msg-id-66999.html>

on the Federal Buildings and Dwellings Register (ORegBL). The aim of the revision is to satisfy the requirements of the new legislation on second homes (LRS⁶) and to improve the organisation and use of the register. This revision, which will take effect as of 1st July 2017, will enable swisstopo to create two new official directories from 2020, one for streets and another for building addresses. The ordinance on geoinformation (OGéo⁷) and the ordinance on geographic names (ONGéo) have also been revised.

From a technical point of view, the administrative registers could already now process their addresses by comparing them with RBD data. Unfortunately, without legal pressure, our various partners are sometimes reluctant to implement new standards, especially if they cannot see any direct or measurable benefit with regards to their own use of the system.

Once both new registers are available, references addresses with unique identifiers will be available to all Confederation services and the cantons and communes. A transition period will be necessary for administrative registers to integrate the identifiers but this waiting period will give SBER time to carry out any technical changes needed.

The original version of SBER was a first step towards the final objectives of the RG-BUR project. Not all of the automated processes will be implemented but the architecture in place will support future developments and enable increased productivity in general. As the intrinsic quality of the addresses is improved by these innovations, the production teams will be able to concentrate on updating the register, entity addresses and contact information.

⁶ Loi sur les résidences secondaires : <https://www.admin.ch/opc/fr/classified-compilation/20140036/index.html>

⁷ Ordonnance sur la géoinformation : <https://www.admin.ch/opc/fr/classified-compilation/20071088/index.html>