The Design and Implementation of Statistical Business Register (SBR) System of BPS

1. Introduction

In BPS, Statistics – Indonesia, Subject Matter Areas (SMAs) produce economic statistics based on their specialties, such as forestry, agriculture, manufacturing, construction, and transportation. To produce the statistics, each SMA maintains their own business register separately (BPS, 2016). Thus, there is no a common frame to produce sampling frames for economic surveys conducted by the SMAs. The lack of coordination between SMAs to produce the sampling frames may increase the response burden which in turn may decrease the response rate. Besides, it may increase the cost and effort for conducting the surveys, especially, at the regional offices. Furthermore, it is quite difficult to harmonize the surveys and integrate the survey data. Therefore, BPS needs to build a Statistical Business Register (SBR) that can be used to coordinate the economic surveys.

In Mei 2014, most of SMAs agreed to integrate their business registers to build a single BPS Statistical Business Register (BPS SBR). This initiative was done by matching businesses stored in the separated registers with those obtained from the Economic Census 2006. After this first initiative, BPS assigned 42 staffs to profile the top 100 biggest enterprises in Indonesia to produce an appropriate statistical and reporting structure for the large enterprises. This profiling initiative is then followed by having 540 district offices to collect data from their respective administrative data sources in order to extend the scope of the BPS SBR.

To facilitate the above initiatives, BPS has built a *new web-based SBR system*. The web-based system can facilitate the process of updating business data by many users from all BPS offices spread throughout Indonesia. This paper describes the design and implementation of the SBR system in terms of three main features, namely, *matching*, *relating,* and *approving*.

2. The Design of Statistical Business Register (SBR) System

The initial data for building the business register was drawn from 2006 Economic Census. BPS used three data sources to update the initial data, namely, SMAs' business registers, administrative data collected by district offices, and business profiling data. Since there is no a single unique business identifier in Indonesia, the updating process includes business matching based on business names and addresses. The matching process is necessary because we need to avoid data duplication in the SBR database.

A (new) business which has no match with any of the existing businesses in the BPS SBR will be recorded (in the BPS SBR database) as a new business. On the other hand, a (new) business with a match in the BPS SBR will be processed according to one of the following options:

1. *Keep*, flag the (new) business to indicate that it has a match in the SBR database without any updates applied to the SBR database.
2. *Replace*, flag the (new) business as the first option and replace the data of the matched business in the BPS SBR database with the data of the (new) business.
3. *Edit*, flag the (new) business as the first option and update some parts (fields) of the matched business data with the data from the (new) business.

The BPS SBR system utilizes a full text search method that matches a business with the whole businesses in the BPS SBR database and presents the top 25 matches based on their similarity scores. This will facilitate and expedite the matching process done by the SBR staffs. Figure 1 shows the flowchart of the matching process that has been implemented in the BPS SBR system.
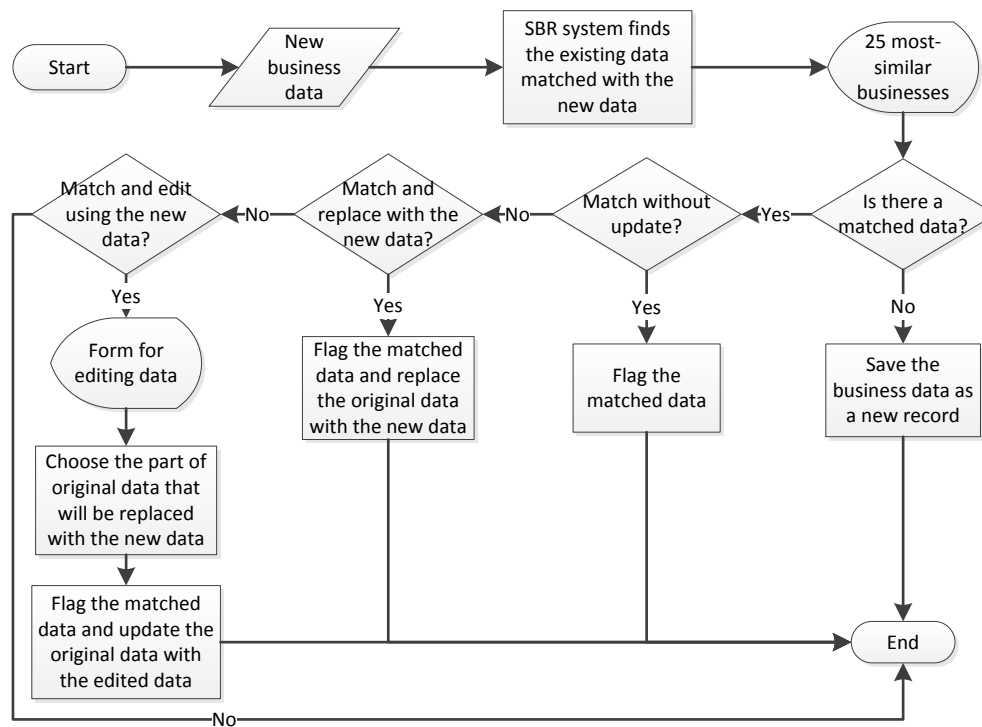

Figure 1. Flowchart of matching feature

Besides storing the characteristics data of business units, the BPS SBR system also maintains relationships between those business units (business hierarchical structure). Accordingly, the BPS SBR system provides a feature to facilitate the operators/profilers to relate businesses according to their hierarchical structures. The business hierarchy is built based on the concept of the statistical unit. The recommended standard statistical units (SSUs) model based on the European Commission (EC) regulation comprises three hierarchically organized standard statistical units (AfDB, 2014), as illustrated in Figure 2.
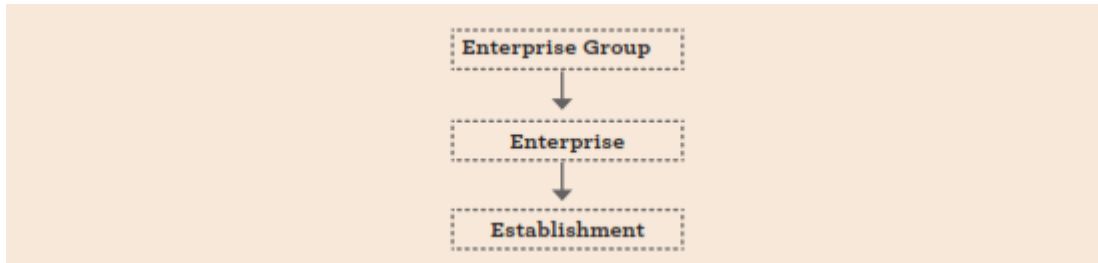
Figure 2. Recommended Statistical Units Model

Furthermore, as described in *System of National Accounts* (SNA) 2008, many enterprises operating within an economy are linked with other enterprises by complete or partial common ownership and a shared management structure to form an enterprise group. Then, based on the concept of subsidiary corporations, corporation *B* is said to be a subsidiary of corporation *A* when:

a. Either corporation *A* controls more than half of the shareholders' voting power in corporation *B*; or

b. Corporation *A* is a shareholder in corporation *B* with the right to appoint or remove a majority of the directors of corporation *B*.

Hence, it is possible to relate one enterprise group (EG) with several enterprises (EN), and in turn, to relate one enterprise (EN) with several enterprises (EN) or establishments (ES) to construct a hierarchy of EG, EN, and ES. The construction of business hierarchy is done in a top-down fashion. Therefore, first the profilers must seek the *parent business*, then seek its subsidiaries and construct the hierarchy up to the smallest level. To facilitate the profiling process, the BPS SBR system must be able to visualize the business hierarchy in a tree diagram for easy viewing. The example of business hierarchy visualization is shown in Figure 3.
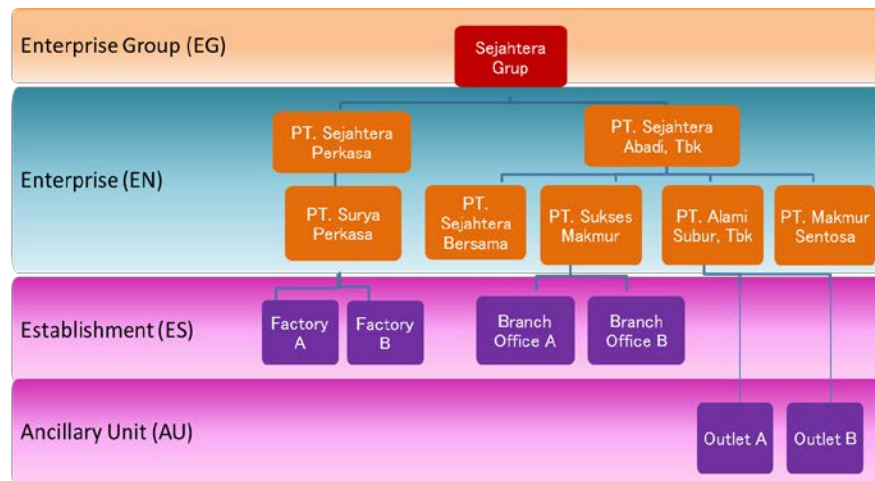

Figure 3. The vizualitation of business hierarchy

To ensure the quality of the matching and relating (profiling) results, the SBR staffs/profilers are assigned one of the following roles: *operator* and *supervisor*. The operator actually does the business matching and business relating, while the supervisor approves or rejects the matching and relating results. To facilitate the supervisor task, the BPS SBR system provides a special GUI for the supervisors, so that they can easily find and view the updated parts, view the details of the newly added businesses, and observe the newly added relations between businesses. Every action changing the business register either by adding a new business data or updating the existing business data will be recorded in the database as the history of the business. The previous values of the business data

should be retained in the live register or in a historical register so can be used for statistical production and analysis (UNECE, 2015).

3. The Implementation of Statistical Business Register (SBR) System of BPS

The BPS SBR system is a web-based system. This system has been developed by using PHP language and backed by Microsoft SQL Server. This section will explain about the implementation of three prominent features provided by the BPS SBR system, namely, matching, relating, and approving features.

3.1. Matching Feature

In the BPS SBR system, the business matching is done by an operator by inspecting the top 25 matches presented by the SBR system. Based on the similary of business names and addresses, the operator decides whether a particular (new) company has a match or not. Currently, the BPS SBR system uses the Full Text Search (FTS) technology provided by the Microsoft SQL Server with some optimizations like removing stop words. The FTS technology ranks the top 25 most similar businesses based on the names and addresses of the businesses.

Figure 4 shows the screen capture of the matching page. As shown in figure, the system displays the comparison of the new business data and the most matched business data among all business data in SBR. A check mark is given to the matched business character/field when the field value exactly match the corresponding new company field value (e.g., regency and province). By doing this, the operators should only focus on fields without a check mark (e.g., company name and address).

Figure 4. Matching page

As described in the design section, there are three matching procedures in the BPS SBR system: *keep, replace,* and *edit*. In Figure 4, the *keep* button is selected when the operator wants to match without updating the original data, while the *replace* button is selected when the operator decides to match with fully replacing the original data with the new data. However, if the operator wants to edit the original data with some parts of the new data, he should select the *edit* button. Figure 5 shows the editing form displayed after selecting the *edit* button. In the figure, when the operators select the *replace* button, the system will replace the corresponding field value of the business in the SBR with the field value of the new company.

Figure 5. Matching procedure with editing some parts of the original data

When the operators cannot find the matched business among the top 25 matched businesses, the new business will be flagged as "*not match*" and stored as a new record in the SBR system database. In this case, the operator will select the *not match* button (Figure 4) to proceed to the next new business.

Finally, when the operator cannot decide yet whether there is a match for a new business (after inspecting the top 25 matched businesses), the system provides a function to skip the business and get back to the business later. This is done by selecting the *skip* button in the matching page.

The operator can select the *detail* button to view detailed business data which can give more information to determine whether the data are match or not. In the matching process, the data updated by the operator is not directly reflected into the SBR database; the supervisor must first approve the updates before storing it into the SBR database.

3.2. Relating Feature

The relating process in the BPS SBR system uses a top-down approach that starts with defining the parent business and then linking it with its subsidiaries. The profilers can start to construct the business hierarchy from the enterprise group (EG) or construct the hierarchy of the enterprise (EN) first and then connect it to the enterprise group later.

Suppose that a profiler found that *PT. Lion Indonesia* (an enterprise) has *Lion Indo 1* as its establishment through a profiling process. To create a business hierarchy, the operator first has to find the parent business (*PT. Lion Indonesia* in this case) from the SBR database. To do this, he puts the business name as a search keyword and may optionally puts address, statistical unit, and/or ISIC code to the corresponding search fields (Figure 6.a). The search result of the business is displayed based on the similarity of the name, address, statistical unit, and ISIC code.

After the operator finds the parent business, next he finds the subsidiary business, i.e., *Lion Indo 1* (Figure 6.b). The parent and subsidiary businesses will be displayed side-by-side and the operator selects the *add* button to complete the business hierarchy. Figure 6.c shows the business hierarchy in a tree diagram view, where *Lion Indo 1* is placed below *PT. Lion Indonesia* as its parent.

The relating feature has an important rule validation. An enterprise group can only be related to one or more enterprises, while an enterprise can be related to either one or more enterprises or establishments. The BPS SBR system provides a GUI to facilitate relating many subsidiaries to a specific parent at once.
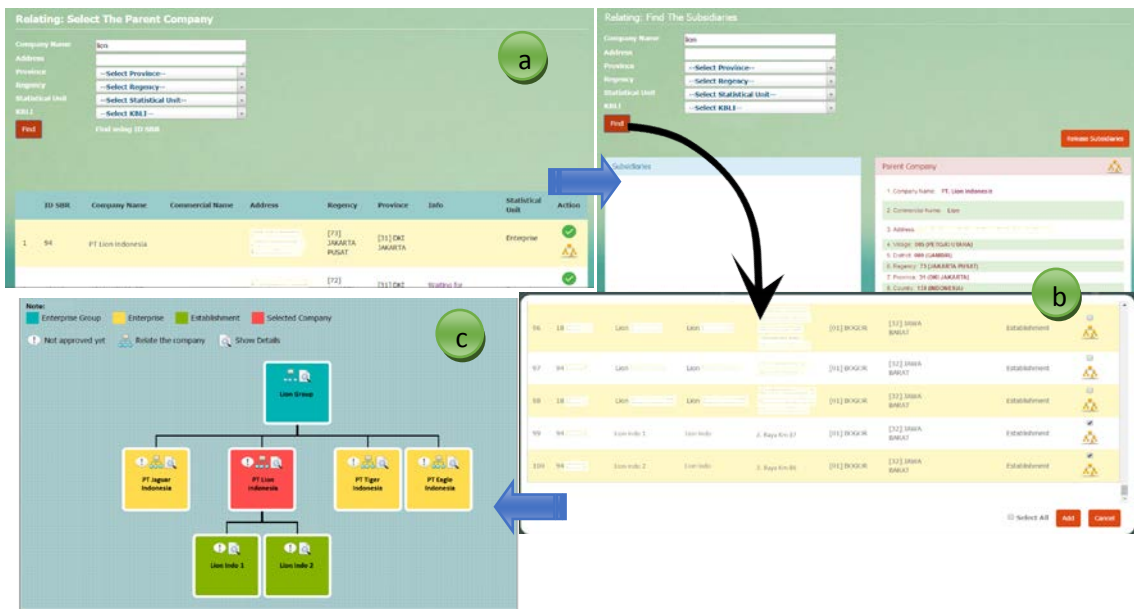
Figure 6. The relating process: a. Find the parent; b. Find the subsidiaries;
c. The system shows the relating result in a tree diagram presentation

3.3. The Approving Feature in the BPS SBR System

To preserve the quality of data in BPS SBR system, any data changes such as inserting a new business, updating business characteristics values, and defining a parent business and subsidiary relation must be evaluated and approved by supervisors.

The GUI for approving a new business consists of a list of new businesses (to be approved by the supervisors), a search function, an approve or reject function, and a function to view the details of a particular business data (Figure 7). For convenience, the system also provides a function to approve or reject two or more new businesses at once.
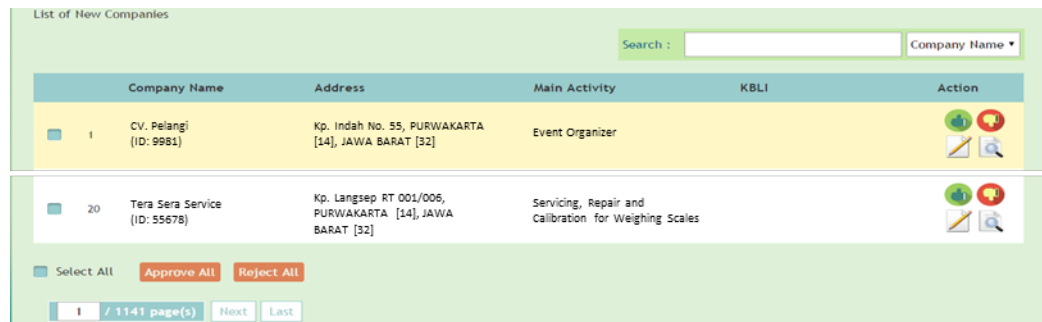

Figure 7. The new business approval page

Figure 8 shows the GUI for approving business matching. As shown in the figure, the original data of existing business in BPS SBR database and the new business data which has been matched with the existing business are shown side-by-side for easy comparison. The texts written in a bracket at the fields of the original data (left hand side) are the updates that were performed by the operators.

In addition, there is also information on the number of matched businesses (at the upper right hand side of Figure 8). The supervisors can evaluate the matched businesses one by one to decide a match to be approved.
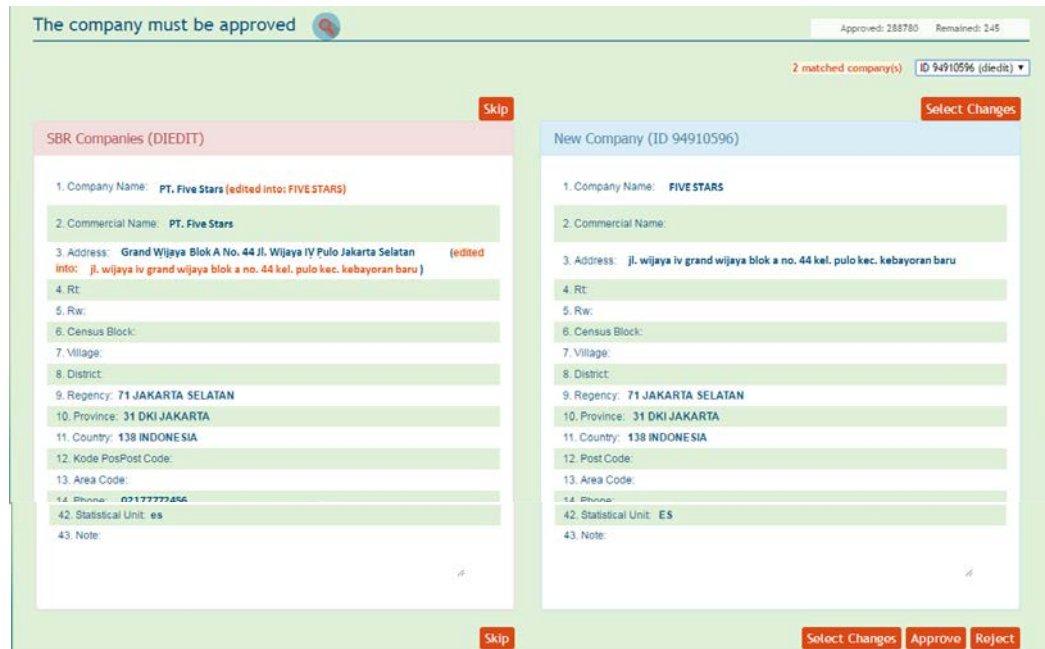

Figure 8. The matching approval page

Figure 9 shows the GUI for approving parent-subsidiaries relation. The GUI is almost similar with the new business approval page (Figure 7). There are a list of new relations with brief information about the business address and parent, a search function, an approve or reject function, and a function to view the business hierarchy that has been created.

Figure 9. The new relation approval page

4. Conclusions

BPS has built a web-based SBR system to facilitate the initial development of Statistical Business Register in Indonesia. The BPS SBR system has provided three main features for updating process, namely *matching*, *relating*, *and approving*. The three features have been extensively used by more than 940 BPS SBR staffs scattered throughout the districts. Especially for the matching and approving features, they have been used to match and approve around 944,000 businesses by those

staffs. A total of 606,000 new businesses were obtained from the matching process, and many data from the new businesses was used to update the existing business data. Moreover, there are 27,000 business that has been related by the profiler team using the relating feature.

5. References

African Development Bank (AfDB). (2014). *Guidelines for Building Statistical Business Registers in Africa*. Tunis

Badan Pusat Statistik (BPS). (2016). *Training Module (Capacity Building) - Statistics Economic Model*. Jakarta

European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations and World Bank. (2009). *System of national accounts 2008*. New York

United Nations Economic Commission for Europe (UNECE). (2015). *Guidelines on Statistical Business Registers*. New York and Geneva