Lien Suharni and Rr. Nefriana, Badan Pusat Statistik – Statistics Indonesia

Innovation in Statistical Business Register

The Use of Google Maps Geocoding API and Google Places API Web Service Data for Automation of Updating and Matching Processes in Statistical Business Register

Abstract

After five years in the development of the Statistical Business Register (SBR) in Indonesia, BPS still has the same problem pertaining the limited human resource. The works are mainly done or coordinated by the SBR Secretariat in Sub-Directorate of Statistical Standardization and Classification Development which in charge both for the standardization and classification development also for the SBR development as an addition. While there are only fourteen people in the secretariat, sometimes Subject Matter Areas (SMAs) and interns are placed to help. However, the SMAs have already had some high burden themselves. In 2017, for example, BPS had 122 surveys which mainly conducted by the SMAs (BPS, 2018). This makes it hard for the SBR to be one of the priorities. The interns, on the other hand, will be placed in all around Indonesia territories after only limited times helping SBR. For that constraint, BPS needs some innovations. One of the approaches that can be implemented is the automation. This paper shows two types of research pertaining the SBR automation in updating and matching processes.

For automating the updating process, Google Maps Geocoding API and Google Place API Web Service are used to update and complete the business' data. First, the place identity codes for each sample enterprise are obtained using Google Maps Geocoding API with the name and address of the enterprise as the keywords. Then, based on the place identity codes, the contact data and other data are obtained with both APIs: canonical name, more complete address for the business, latitude, longitude, phone number, active status (active or permanently closed), and website. After that, the number of enterprises that successfully updated with that method is counted and checked whether the update is accurate.

BPS SBR uses many data sources. To avoid duplication, a matching process must be done. Because there is no unique identification number for the business in Indonesia, the matching process is done manually. Actually, a research had been made for automation (Nefriana and Kamaratih, 2017). Despite a large number of units that can be matched automatically on that research, significant numbers of false positive and false negative still existed so that the method has not yet been implemented. Another attempt must be done again because of that reason. In this paper, instead of matching the businesses based on the geocodes, place identity codes are used. In the matching process, if the two businesses have the same business place identity codes based on the response of Google Maps API, then they will be regarded as the same businesses. Besides, the scope of the automation is narrowed down into the enterprises only. Finally, the number of the enterprises that are successfully matched automatically is counted along with the number of the correct matches and incorrect matches.