

GEOSPATIAL BIG DATA UNTUK PEMETAAN LISTRIK PADAM DI DKI JAKARTA

Muamar Nandika Baharain

14/364959/GE/07749

INTISARI

Perkembangan teknologi informasi mendorong peningkatan jumlah data ter-georeferensi yang sangat besar, serta perangkat komputasi dan media penyimpanan data yang murah telah mendukung perkembangan *big data* geospasial dengan sangat cepat. *Big data* geospasial dapat digunakan untuk memetakan kejadian listrik padam dari ekstraksi *tweet* melalui API Public Twitter secara *online*. Penelitian ini bertujuan (1) mengkaji proses *crawling* data listrik padam di DKI Jakarta bulan Maret dan April 2018, (2) mengukur pelayanan prima dari rasio jumlah listrik padam yang berasal dari *netizen* dengan data balasan PLN (Perusahaan Listrik Negara) pada akun @pln_123, dan (3) memetakan dan menganalisis sebaran spasial listrik padam di DKI Jakarta bulan Maret dan April 2018.

Penelitian ini menggunakan *big data* geospasial Twitter dengan lokasi kajian seluruh kecamatan di DKI Jakarta. Tahapan penelitian ini terdiri atas tahap penarikan (*crawling*) data *tweet*, tahap pembersihan data *tweet* (*data cleaning*), tahap penyimpanan data *tweet* (*data storing*), tahap pemilihan data *tweet* (*data selecting*), tahap visualisasi dan analisis data. Proses penarikan (*crawling*) dan proses pembersihan (*cleaning*) data *tweet* dilakukan menggunakan RStudio, proses penyimpanan (*storing*) di “localhost/phpmyadmin/”, proses pemilihan dilakukan dengan sistem yang telah disusun berbasis C#, dan proses visualisasi data *tweet* direpresentasikan dalam bentuk peta diagram batang kejadian listrik padam dan peta *choropleth* jumlah laporan kejadian listrik padam.

Hasil dari penelitian ini ditemukan bahwa RStudio dapat digunakan untuk proses penarikan (*crawling*) data *tweet*. Selain itu, ditemukan bahwa adanya penurunan persentase faktor penilaian pelayanan prima dari PLN terhadap penyediaan informasi penyebab kejadian listrik kepada *netizen*. Kejadian listrik padam secara spasial membentuk pola yang menyebar dengan kejadian listrik padam tersedikit selalu terjadi pada minggu kelima, dan kejadian listrik padam terbanyak pada bulan Maret 2018 pada minggu keempat dan bulan April 2018 pada minggu kedua. Waktu kejadian listrik padam terbanyak pada pukul 15.00 dan tersedikit pada pukul 23.00 pada bulan Maret dan April 2018.

Kata kunci: *Big data geospasial*, Twitter, RStudio, Listrik Padam

GEOSPATIAL BIG DATA UNTUK PEMETAAN LISTRIK PADAM DI DKI JAKARTA

Muamar Nandika Baharain

14/364959/GE/07749

ABSTRACT

The development of information technology drives an enormous increase in the amount of georeferenced data, as well as inexpensive computing devices and data storage media has supported the development of geospatial big data very quickly. Geospatial big data can be used to map out power outages from tweet extraction via the online API Public Twitter. This study aims to (1) examine the process of crawling data list in DKI Jakarta in March and April 2018, (2) to measure the excellent service from the ratio of the number of power outages originating from netizens with the data of PLN (Perusahaan Listrik Negara) on account @ pln_123, and (3) mapped and analyzed the spatial distribution of power outages in DKI Jakarta in March and April 2018.

This study used big geospatial data Twitter with the location of the study of all districts in Jakarta. The stages of this research consist of tweet tweet stage, tweet data cleaning, data storage tweets, data selecting, visualization and data analysis. The process of withdrawal (crawling) and cleaning process (cleaning) data tweet done used RStudio, the saving process (storing) in “localhost/phpmyadmin/”, the selection process was done by a system that has been compiled based on C #, and the process of tweet data visualization is represented in the form of a diagram of the stem of the event of power outages and choropleth map number of reports of power outages.

The results of this study found that RStudio can be used to process (crawling) data tweet. In addition, it was found that there was a decrease in the percentage of excellent service assessment factors from PLN to provide information on the cause of electrical incident to netizens. Power outages occur spatially forming a spread pattern with the least power outages in the fifth week, and most power outages occur in March 2018 in the fourth week and April 2018 in the second week. The most frequent power outages occur at 3pm. and at least 11:00 a.m. in March and April 2018.

Keywords: *Geospatial Big Data, Twitter, RStudio, Power Outages*