

**PEMANFAATAN *BIG DATA* GEOSPASIAL TWITTER UNTUK
PEMETAAN SEBARAN KETERLAMBATAN BUS TRANSJAKARTA
DI DKI JAKARTA**

Oleh :

Wahyu Hendardi Giri Ananto

16/393509/GE/08257

INTISARI

DKI Jakarta merupakan daerah dengan tingkat pertumbuhan penduduk yang tinggi. Tingginya pertumbuhan penduduk dan aktivitasnya menyebabkan meningkatnya pengguna kendaraan bermotor dimana peningkatan ini tidak diimbangi dengan ketersediaan jalan yang ada sehingga dapat mengakibatkan kemacetan. Pemerintah menerbitkan kebijakan terkait pengadaan transportasi umum seperti *busway* Transjakarta untuk mengurangi kemacetan yang sering terjadi. Namun pada kenyataannya pelaksanaan sistem transportasi ini memiliki masalah dengan ketepatan waktu kedatangan bus.

Penelitian ini bertujuan untuk membangun database dari hasil crawling dan cleaning data Twitter terkait dengan keterlambatan bus Transjakarta dan juga memetakan sebaran keterlambatan bus Transjakarta yang dapat digunakan dalam rangka membantu pemerintah serta pihak dari Transjakarta membenahi pelayanannya. Pemetaan sebaran keterlambatan *busway* menggunakan integrasi sistem informasi geografi dan *big data* geospasial (*geospatial big data*) Twitter memiliki efisiensi terhadap waktu dan biaya dibandingkan metode pemetaan secara konvensional serta menghadirkan perspektif baru dari masyarakat. Ekstraksi dan cleaning data twitter menggunakan software Rstudio yang dijalankan dengan bahasa pemrograman R. Tahapan penelitian ini dimulai dari ekstraksi data, filtering dan cleaning data, storing data di database, selecting data dan visualisasi data.

Hasil dari penelitian ini berupa database hasil cleaning data *tweet* terkait keterlambatan bus Transjakarta serta peta persebarannya. Peta persebaran keterlambatan bus pada bulan Juni hingga Juli 2020 memperlihatkan adanya penurunan kasus keterlambatan yang disebabkan oleh beberapa faktor yang bersifat dinamis.

Kata kunci : *Big data Geospasial, Twitter, R, Keterlambatan Bus.*

**UTILIZING TWITTER GEOSPATIAL *BIG DATA* FOR
TRANSJAKARTA BUS DELAY DISTRIBUTION MAPPING
IN DKI JAKARTA**

By :

Wahyu Hendardi Giri Ananto

16/393509/GE/08257

ABSTRACT

DKI Jakarta is an area with a high population growth rate. The high population growth and its activities have led to an increase in motorized vehicle users where this increase is followed by the availability of existing roads which can lead to congestion. The government issues policies related to the procurement of public transportation such as the Transjakarta busway to reduce congestion that often occurs. But in reality the implementation of this transportation system has problems with the punctuality of bus arrivals.

This study aims to build a database from the results of crawling and cleaning Twitter data related to Transjakarta bus delays and also to map the distribution of Transjakarta bus delays that can be used in order to help the government and Transjakarta parties to improve their services. Mapping the distribution of busway delays using the integration of geographic information systems and geospatial big data Twitter has time and cost efficiency compared to conventional mapping methods and presents a new perspective from the community. Extraction and cleaning of Twitter data using Rstudio software which is run with the R programming language. The stages of this research start from data extraction, filtering and cleaning data, storing data in the database, selecting data and visualizing data.

The results of this study are a database of cleaned tweet data that related to Transjakarta bus delays and a map of their distribution. The map of the distribution of bus delays from June to July 2020 shows a decrease in delays caused by several dynamic factors.

Key words: *Geospatial Big data, Twitter, R, Bus Delay*