

ABSTRACT

Data is used as a source of information to make decisions that determine the success of an organization. Good quality data can produce decisions that can determine the success of a business. Conversely, when the data is dirty, the data can have a negative impact on business. It cannot be ensured that the data that is entered and available is always of good quality.

There are still many companies that have poor data quality, one of the government organizations that I use as a case study in research. To ensure this, it is necessary to ensure data quality management. Data quality management is a collection of activities carried out to ensure that the data to be used and processed has good quality. By doing the right data quality management, it can measure, assess, improve, and ensure that the data is of good quality. In data quality management, there are a series of strategies that are carried out namely data profiling, data cleansing, data monitoring, and data integration. Data profiling is a process of detecting data that is inconsistent, wrong, lost, and duplicated. Data profiling is the process of identifying data that is inconsistent, wrong, lost, and duplicate in a data set that will be corrected. Some application tools can be used to help process data profiling. Some of the tools available are paid tools. Therefore you can switch to open source tools. Often open-source tools differ from paid tools such as functionality and application performance.

Application performance is one of the problems that often arise in application implementation. Often users perceive the application used has a problem so that the impact of the application will not be reused. For that application, performance is one of the factors that need to be considered in application implementation.

In this research, we will discuss improvements to the data profiling application tool that has been made in previous studies. Improvements to the data profiling application aim to improve the performance of the data profiling function in terms of time. In the profiling function, the application performance level is still considered low, known from the response time required.

Keywords: data quality, data quality management, data profiling, application performance.