ABSTRACT

Data is an important component within the company that could affect objectives and decision-making within the company. In addition, the data must have a good quality, which became an important determining factor in terms of the effectiveness of an organization to deliver business value. To ensure a data is not dirty there are a number of processes, one of which profiling data. Data profiling is an abnormality in the search process data that undermine the value of the data. To do some checking on data profiling process can use multiple tools. The existence of an open source application to help do the job profiling data is very helpful but on the other hand not all such applications in accordance with the needs of existing case studies. Therefore needed an implementation architecture application that can help make it easier for the user in the XYZ Organization. This application combines logic Kettle later there is a website for the user. For the determination of the benchmarked applications this study compares the existing applications based on several criteria. On testing the application using scenario testing for knowing any performance functionality.

Keywords: Data, data profiling, open source application, application architecture