

ABSTRACT

All kinds of information are straightforward to obtain in today's modern world. One way to get the information that you want is by searching or searching in cyberspace via internet access. However, the amount of information available when searching is very large and unstructured (big data). Collecting and sorting information from big data manually takes a lot of time and effort. So, we need a more effective way to collect data, namely by Web Scraping.

Web Scraping is a technique of taking and extracting data on a particular web and collecting it in one file. In this way, the data you want to find will be easier to obtain. In addition, the data obtained has also been structured to make it easier for data seekers to read the information obtained.

In this study, the work was carried out using the Web Scraping technique with the HTML parsing method on Google Scholar to find data for scientific articles to be searched. The searched data will be grouped and compiled into a file in Comma Separated Values (CSV) format. With the output in the form of a csv file, it is hoped to make it easier for information seekers to obtain, collect, process and analyze the scientific article data they are looking for.

Keywords: *Web Scraping, Google Scholar, Big Data, Automatic Data Collection, Scientific Article.*