Pengembangan Sistem Web Crawler Sebagai Sarana Riset Media Secara Otomatis (Studi Di Subdit Neraca Rumah Tangga Dan Institusi Nirlaba)

Wongkar, Enggelin Giacinta

Deskripsi Lengkap: https://lib.ui.ac.id/detail?id=20428380&lokasi=lokal

----

Abstrak

With the vast development of data to become informations on the Internet, everything online seems to explode at a rapid rate. These informations, including online news which is created as a complement to the original printed media, has even overtaken the latter. Subdirectorate of Household National Account and Non-profit Institution of Statistics Indonesia is in charge for the work of media research. In the process of media research, time and human resources are two important elements but yet having problem of ineffective and inefficient process. This study aimed to overcome that problem by developing a web crawler system that could do summarization automatically from online news sites (currently from Bisnis and Kontan) with output in Microsoft Word format file and minimizing number of similar news. This system is developed using several techniques in information technologies such as crawling and wrapping method and cosine similarity method to minimalize similar news. The result shows the process of media research by using this system much more effective and efficient.