

Research on extracting information content from online news websites.

Pham Thi Lien

Abstract: In the modern context, extracting and summarizing information content from online news websites is becoming increasingly important. This topic aims to solve the problem of automating the information extraction process efficiently. Through the application of a combined theoretical and experimental research method, this topic creates a content extraction system capable of handling diverse types of journalistic data. The expected results of this topic not only advance the information synthesis process but also contribute to the development of natural language processing technology.

Keywords: content extraction, natural language processing, online newspaper