

Study Document Validation and Mapping with User-profile for Collaboration

Ruth Oluwatosin Adeyemo¹, Olusesan Adeyemi Adelabu² and Akeem Adewale Oyelana³

¹Department of Computer Science, University of Ibadan, Oyo State, Nigeria

*²Faculty of Science and Agriculture, University of Fort Hare, Alice, P.B. X1314,
Eastern Cape, South Africa*

*³Department of Public Administration, University of Fort Hare, Alice, P.B. X1314,
Eastern Cape, 5700, South Africa*

E-mail: ¹<tosadeyemo@gmail.com>, ²<201409080@ufh.ac.za>, ³<201100592@ufh.ac.za>

KEYWORDS Plagiarism Detection. Text Comparison. Study Document. User-profile

ABSTRACT As study and publication remains a yardstick for scientific endeavours, it is not enough for researcher to only publish their papers, it is therefore paramount that the quality of research publications be put in check through validation of research document, ensuring research document submitted in a repertoire does not already exist and providing a possible forum for collaboration amongst researchers. The objective of the study was to use plagiarism detection in study document by comparing a researcher's work with previous publications based on user-profile. Current studies in the field of automatic plagiarism detection for content archives concentrate on algorithms that compare plagiarized documents with potential unique records inside a huge collection of documents. The methodology compared suspicious documents against a set of potential original documents which have been filtered out from the large repertoire of documents based on the user profile. The researchers used two main algorithms which are the study document validation algorithm and text comparison (PlagCheck) algorithms coupled with user-profile to detect plagiarized document hence determine the validity of a study document. The framework was assessed by utilizing a test-set that contained occurrences of verbatim duplications and messages with little or no alteration. The result and performance evaluation showed the researchers' system performed better and faster than existing systems, achieving the accuracy of ninety-eight percent (98%) over splat. The study was able to take care of the challenge of processing time of validation which is usually encountered in other Plagiarism Detection Systems (PDS).