

Text Mining Analysis of the Conflict of the Ages Series

Darrin Thomas and Gabby Jed C. Galgao
Asia-Pacific International University

Abstract

Research regarding the writings of Ellen White has primarily relied on qualitative measures to develop insights and themes. Consequently, quantitative measures have been largely left unused in this regard. The purpose of this study was to help fill this gap through quantitative text mining analysis of Ellen White's Conflict of the Ages series utilizing an exploratory design that used word frequencies, word clouds, and topical models to discover themes. Themes were created using topic models derived from the latent Dirichlet allocation (DLA) algorithm. Results revealed the word frequencies of each of the five books in the series as well as an aggregation of the word used in the Conflict of the Ages series as a whole. A four-theme model was developed which included 1) God and His people, 2) Evangelism, 3) Old Testament concepts, and 4) Prophecy. Further findings showed that words such as "Christ", "Lord", and "God" were the most frequently used terms in the series, making the Conflict of the Ages series primarily Christ-centered in terms of word frequency. Continuing text mining analysis on writings of Ellen White is suggested for further study and validation of these results.

Keywords: *text mining, Ellen White, topical models, Conflict of the Ages*