

{tag} International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

[Volume 168](#)

-
[Number 8](#)

Year of Publication: 2017

Authors:

Mayuri Verma

10.5120/ijca2017914486

{bibtex}2017914486.bib{/bibtex}

Abstract

This paper presents a text mining approach to compare and to explore the similarities and the differences between various religious texts using POS Tagging and Term Document Matrix. Automated text mining and machine learning tools have been used for lexical analysis of the ten world famous religious texts: the Holy Bible, the Dhammapada, the Tao Te Ching, the Bhagwad Gita, the Guru Granth Sahib, the Agama, the Quran, the Rig Veda, the Sarbachan and the Torah. The extracted nouns categories were used as features to explore some interesting relationships between these religions and ideas that have emerged in different religions from different geographic regions.

References

1. Daniel McDonald. "A Text Mining Analysis of Religious Texts". The Journal of Business Inquiry ,2014.
2. Qahl, Salha Hassan Muhammed, "An Automatic Similarity Detection Engine Between

Sacred Texts Using Text Mining and Similarity Measures" (2014). Thesis. Rochester Institute of Technology.

3. Frank Lloyd Sindler." COMPARATIVE STUDY OF CHRISTIAN, JEWISH, AND ISLAMIC THEODICY"(1982).Thesis. B.S., Clemson University.

4. Feldman, Ronen, and James Sanger. The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge University Press, 2007.

5. Manning, Christopher D., and Hinrich Schütze. Foundations of statistical natural language processing. Vol. 999. Cambridge: MIT press, 1999.

6. The Holy Bible, translated from the Latin Vulgate
<https://archive.org/details/holybibletransla00chalrich>

7. Free Books To Read Audio Library http://freebookstoread.com/dhmpd10_1.htm

8. The Holy Bible, translated from the Latin Vulgate http://www.with.org/tao_te_ching_en.pdf

9. Bhagavad-Gita As It Is: http://www.bhagavatgita.ru/files/Bhagavad-gita_As_It_Is.pdf

10. English Translation of Siri Guru Granth Sahib
<http://old.sgpc.net/CDN/English%20Translation%20of%20Siri%20Guru%20Granth%20Sahib.pdf>

11. AGAMA – An Introduction: http://jainaagam.org/download_pdf/Aagam_Intro_Booklet%20v280912.pdf

12. Quran English Translation <http://www.clearquran.com/downloads/quran-english-translation-clearquran-edition-allah.pdf>

13. The Hymns of the Rigveda: <http://www.sanskritweb.net/rigveda/griffith.pdf>

14. SAR BACHANRÁDHÁSOÁMÍ (Poetry)
<https://www.scribd.com/doc/118290685/Sar-Bachan-Radhasoami-Poetry-Volume-One>

15. Torah Bible of Jewish
<http://text.123doc.org/document/4213026-torah-bible-of-jewish.htm>

16. Martin Schweinberger." Part-Of-Speech Tagging with R "(June 24, 2016)

17. Vocabulary Size and Use: Lexical Richness in L2 ... - Oxford Academic
<https://academic.oup.com/applij/article-abstract/16/3/307/184110/Vocabulary-Size-and-Use-Lexical-Richness-in-L2>

18. Text Mining Package <https://cran.r-project.org/web/packages/tm/tm.pdf>

Index Terms

Computer Science

Information Sciences

Keywords

Religious Texts, POS Tagging, R, Lexical Analysis