

## **ABSTRACT**

The purpose of this research is to address the limitations in the availability of a comprehensive website that provides a detailed list of Isim Ma'rifat (specific nouns) in the Quran and Sunan Ibn Majah. The lack of comprehensive resources has hindered the ability to study and compare Isim Ma'rifat among these significant Islamic texts. To overcome this challenge, this study develops a natural language processing approach using a Java tokenizer program integrated with a MySQL database containing Sunan Ibn Majah and the texts of the Quran. This program identifies the presence of the "alif lam" prefix and removes diacritics to facilitate accurate comparison between verses in both texts. The primary focus of this research is to identify Isim Ma'rifat starting with "alif lam" that are exclusively found in the Quran (a total of 1183 words), exclusively in Sunan Ibn Majah (a total of 2,894 words), as well as the commonalities between them (a total of 415 words). This analysis results in a comprehensive understanding of the differences and similarities of Isim Ma'rifat with the "alif lam" prefix between the Quran and Sunan Ibn Majah. The findings of this study provide valuable input for the Quranpedia project, contributing to the development of comprehensive and easily accessible resources on the Quran and the Kutub al-Sittah for Islamic studies. It is expected that this research will enhance the understanding of Isim Ma'rifat in the contexts of religion and language and make a significant contribution to the development of natural language processing, particularly in the Arabic language.