
Developing a Database For Scholarship In Indian Languages And Literatures

Shanmugapriya T

phd1501161005@iiti.ac.in

Indian Institute of Technology Indore, India

India is a multilingual country which has twenty-two official languages and eleven written scripts, according to the 8th schedule to the Constitution of India. Each language has its own literature and scholarship. In the digital space, both the dissemination of and access to scholarly articles of Indian languages and literatures are difficult and rare. The main objective of this project is to offer scholarly publications and access to scholarly articles in languages other than English. There are some scholarly articles available in print, but most of them give citation of the works that is insufficient for research scholars. There are some databases available for scholarship in India such as [Shodhganga: a reservoir of Indian Theses](#) and [National Digital Library](#). *Shodhganga* offers the scholarly communities to access ETD (Electronic Thesis and Dissertation) in English. *National Digital Library* is a digital repository which provides learning resources for various subjects with single-window search of which the interface is available in three languages English, Hindi and Bangla. Similarly, there are some databases for digitized literary manuscripts of Indian languages: for instance, [Project Madurai](#) for the Tamil language, etc. However, there is no comprehensive digital database for scholarship in Indian languages and literatures.

This database project envisages digital ecology as an integral part of knowledge production and dissemination. In India, Digital Humanities in literature, apart from notable exceptions like "[Bichitra](#)", there is little in terms of digitization of scholarship from many language literatures. Our database will serve as a collaborative and comprehensive hub for Indian literature scholarship. We plan to make this database after its initial launch to be crowd sourced by enabling Indian literature departments across universities to use it as a platform for disseminating their scholarship.

The poster presents the initial stage of aggregating metadata of literature and scholarly articles in different languages. We have already collected data for Malayalam, North-East, Sindhi and collecting data for Bengali, Hindi and Tamil is underway. This will be extended to other Indian languages.

The collection of data is:

- Malayalam-25800
- North-East-1494
- Sindhi- 8813

We will use the *MLA International Bibliography*, *Shodhganga* and *National Digital Library* as models for our project to develop a database of literature and scholarly articles in Indian languages. One of the primary objectives of the project is to offer the works through an open-access platform which allows access to the researchers across the country. Hence, it will be made open-access through KSHIP (Knowledge Sharing In Publication) which is a multilingual academic publishing project at IIT (Indian Institute of Technology) Indore, India. The database will be developed in different Indian languages, which presents an opportunity to researchers to explore the database in their own respective languages. Collecting data for the database has already been initiated, and the next stage of the project is to host it. The poster presents an introduction to the project, the need for a multilingual scholarly database, the primary objective of the project, different Indian scripts, and our methodology, collected data, and model for access (open access).