What I learned as a field linguist helping with the Cologne Digital Sanskrit Dictionaries

Felix Rau (University of Cologne)

Users consult dictionaries with varied intentions and expectations. The different functions are reflected in the varying scope and complexity of dictionaries. As a consequence, providing a user-friendly web based presentation of different lexical resources is a non-trivial task.

Although it is a retro digitisation project for lexical resources in a major classical language, lessons from the last 16 years of the Cologne Digital Sanskrit Dictionaries can inform our treatment of lexical resources from language documentation and linguistic fieldwork.

The Cologne Digital Sanskrit Dictionaries (http://www.sanskrit-lexicon.uni-koeln.de/) is the largest lexical resource for Sanskrit. It consists of over 30 separate dictionaries with individual mark up and interfaces. In the last years, the LAZARUS project (http://www.cceh.uni-koeln.de/lazarus) at the Cologne Center for eHumanities has worked on converting the different dictionaries into a TEI conformant format and providing a unified interface for the users.

The experience of the original curators of the dictionary project as well as what I learnt from assisting the programmers of the LAZARUS project in designing the mark up as well as the web application highlights the importance of understanding the different needs of different types of users, the consistency and documentation of the underlying data format.