

## **Local Content Development and Online Access**

---

Online content plays an important role in socio economics, education and research needs. People find it easy to go online and search the relevant content instead of going to the markets or libraries to search for it. This saves time and enhances productivity. There are many online resources available for languages like English but less resources have been made available for local languages of Pakistan. CLE has developed and generated local content for various local languages like Urdu, Torwali etc. Currently two projects named " Digital Dictionaries of South Asia" and "Urdu Search Engine" are in progress at CLE which are related to local content access and generation.

In "Digital Dictionaries of South Asia" project, the main aim is to digitize the Punjabi and Sindhi dictionaries for online access. For digitizing dictionaries, dictionary structures have been studied and xml structures have been designed in similar way. The data are typed, populated in the designed xml structures and verified at various levels to ensure quality, consistency and integrity. After this whole process, the finalized data files are uploaded for use through our dictionary website. 193 pages containing 17336 Punjabi words of Punjabi dictionary have been finalized whereas 107 pages of Sindhi dictionary have been sent for verification.

In "Urdu Search Engine" project, the main focus is to make Urdu books online so that they are accessible through Urdu search engine. Urdu books have been scanned through high speed scanner. These books' images are then passed through Urdu OCR to get text files against each image which will aid in searching process. Finally, these text files are encrypted and packaged for final shipment. Up-till now, 110 Urdu books have been shipped for online access.