

Vigor and Intricacy of Urdu Online Libraries: Digital Repository of NCPUL

Raees Ahmed Mughal*

Abstract

One of the rich and free-access online repositories of Urdu books www.urducouncil.nic.in comprise of technical and academic flaws that diminish its value and impact. This short article begins with introduction of this digital collection and leads to identifying its unique features and strengths. Remaining part of the article deals with analysis of contents, design, transliteration, metadata, unique identifier and other academic and technical issues that cause hurdle for end-user and are disastrous for academic impact of the institute. Each section also suggests remedies and recommendations to renovate the repository into user-friendly digital collection and, to ensure better semantic and automated processing.

Keywords: Urdu Digital, Digital Books, Urdu Web, Online Urdu, Transliteration, NCPUL, Digital Collections, Urdu Unicode Fonts

The Council and its Website

www.urducouncil.nic.in is digital repository or showcase of total published output of The National Council for Promotion of Urdu Language (NCPUL). The Council has two alternate names in English and Urdu. In Urdu it is *Qaumi Council Baraey Farogh-e Urdu Zaban*. Both names are in vogue as English abbreviation is used for official matters and the Urdu name is famous in academic and literary circles of India and Urdu readers around the globe. In official statement the Council is “an autonomous body under the Ministry of Human Resource Development (HRD), Department of Higher Education, Government of India”. However, the Council as well as its contribution could not be summarized in one-liner. NCPUL is respected in academic as well as all Urdu literary circles in India for its academic and educational contribution.

To achieve its objectives NCPUL is engaged in various activities that are well documented and reported at this website. These include diploma courses in Urdu language, introductory courses in Arabic and a course in Graphic Design and Calligraphy.

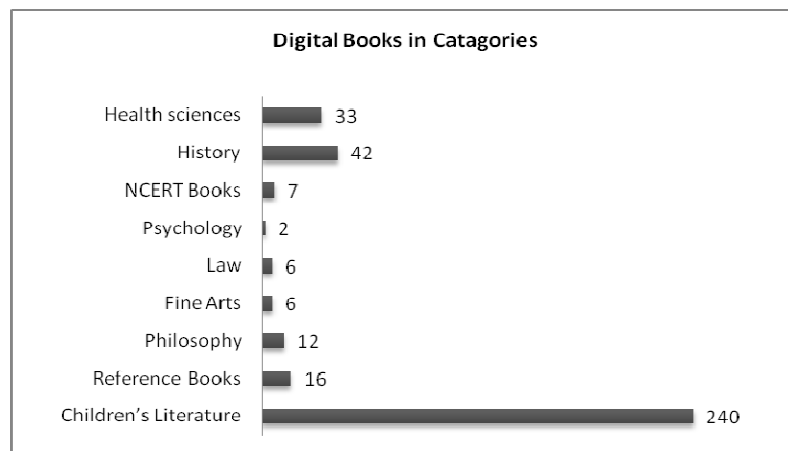
* Raees Ahmed Mughal, Assistant Professor, Government Superior Science College, Peshawar. Email: raees_mughal@yahoo.com

Besides, it organizes seminars on important issues related to Urdu language and literature. The site also documented details about its governing bodies; annual reports inclusive of budgetary details; and list of publications. These details indeed present a transparent governance model of the Council and a good practices manual for other public funded bodies working for promotion of Urdu or other languages in the third world.

The website is neatly designed with ample of breathing space.¹ Variation in font sizes, skilled use of colors coupled with clutter free layout visually poses no burden over visitor's eyes and good taste.

Quantity Counts

At main page on left the third button is labeled as 'Urdu Digital Library'. This title is a bit misleading as this button does not link any library having resources from different publishers and sources but, is Institutional Repository of books and magazines of NCPUL. Click on the button leads to a page that neatly displays only 12 colored and interesting titles of Children books. At end of the titles a line with hotlink reads: 'Click here for more books >>'. This is linked to a page having a list of nine categories: Children's Literature, Reference Books, Philosophy, Fine Arts, Law, Psychology, NCERT Books, History, and Health Sciences. Below this list is again 'Click here for more books >>' but this time for a very different though unstated purpose. Clicking there an alphabetical list of same online resources of the repository are displayed in four pages. But this 'Click here' phenomenon seems never ending for first time visitor as it appears at end of each category working as link to main category page. Books provided under each category by the time of completion of this article are as under:



Out of total 364 books in all nine categories, in preceding four pages (listed alphabetically) 260 books are listed; and is indeed comprise no new book but repetition of the titles already presented in the nine categories.

The Quality and Novelty

Though these numbers seems very slim, it is very impressive contribution in comparison to other open-access repositories and collections that offer Urdu books. Though open-access is clearly a main stream trend around the globe,² in Urdu language due to lesser percentage of readership coupled with huge technological gap, the digital Urdu books and magazines and their collections are still in its infancy.

NCUPL contribution in reference books and children's literature is indeed unique. No other repository offers such rich and well researched reference books in Urdu as free access. Likewise, though Children's Literature is somewhat neglected area as intellectual activity, nonetheless the market demands for such books are vivid. Despite market considerations, offering children's books free of cost and as full access is something unique to NCUPL.

The reference section offers "*Jama Angrezi Urdu Lughat*" (Compact English Urdu Dictionary) in six volumes, five volumes of 'Jama Urdu Encyclopedias' (Compact Urdu Encyclopedia) each volume dedicated for in-depth coverage of Literature, History, Social Sciences, Geography, and Fine Arts. Three volumes of 'Urdu Encyclopedias' are also offered for download. All these reference books are result of some two decades efforts of some of the brilliant minds so far recognized in modern Urdu intellectual and academic tradition within India. These thousands of pages were indeed impossible to obtain for a common reader that is residing anywhere in the world outside India. Despite its generous offer, and capacious readings the quality of author and editorship of these books are, most of the time, meets international standards.

Children's Literature section is notable for its quantity. Likewise other categories like Law, Health Sciences and History though thin in size offers quality translations and rare books.

The importance of this collection lies in two aspects the modernity and open access. It is not a collection of outdated, creative commons books but comprise of subjects quince modern needs. These include very tediously compiled encyclopedias, dictionaries, translation of world classics and some of very modern texts on IT and graphic designs.

And the collection is presented free of cost, mostly in downloadable texts in PDF form, though all titles are current and on sale.

NCUPL also publish two magazines 'Urdu Duniya' (Urdu World) and '*Fikr-o-Tehqeeq*' (Thought and Research). Urdu Duniya is monthly magazine and its target reader is common Urdu reading Indian. Thus contents include digestible cultural themes, personal management and policy briefs on government initiatives related with Urdu reading community. This magazine is offered at the website in PDF format and free. Quarterly magazine *Fikr-o-Tehqeeq* as title suggests is dedicated for reporting and sharing intellectual themes. It mostly publishes special issues on literary themes. Only list of contents of the magazine is uploaded at the site.

Academic Flaws, Technical Errors

A repository offered as free access or for subscription is however rated according to its quality, on two standards, "The two aspects of digital library data quality are the quality of the data in the objects themselves, and the quality of the metadata associated with the objects."³

With all value the collection posses for Urdu world NCPUL repository has major issues with "metadata associated" that in practice ruins the impact and use of this repository. First and foremost, though it is aims to be used by a visitor in need of Urdu books, the whole site is created in English language. A mirror site in Urdu, or at least moderate use of Urdu for main hotlinks and buttons indeed could make the collection accessible for proper and pure Urdu reader and visitor.

Roman alphabets are used as alternate for Urdu script, and this is done indeed in a manner that is unexpected from such academically sound institute. Right from early days of British interest in Urdu, Fort William College provided the first systemic list for the purpose and this exercise is still in vogue, latest being issued a revised scheme of transliteration in 2007 by Annual of Urdu Studies at its site.⁴

Though NCPUL had so many style lists of Roman transliterations to select from and indeed with open option to devise its own, however this is not preferred. Since no standard transliteration system is adopted, the titles are unpredictable at times.

This arbitrary transliteration also resulted at times in very comic wording. Few examples from book titles as they are listed in repository are like:

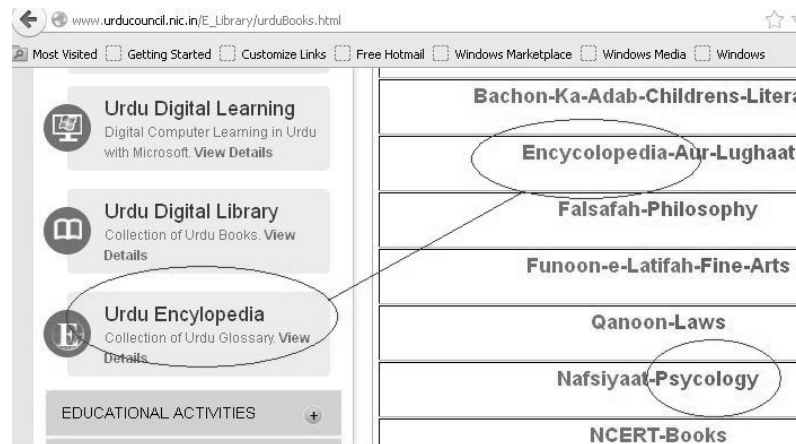
Book Titles Transliteration	Comments
'Bihar Men Urdu Zaban o AdabKaIrtiqa' And 'Deccan Mein MersiaAurAiza Dari'	Different spellings for same word 'mein' and 'Men'
Hali Ka Sayasi Shaoor	Correct: Siyasi
'Zaviya e Nigah' And "Kainat-e-Jalil"	Different ways to mention 'e' in compound word
Classiki Unan - Tehzeeb 'Falsafa' Funoon-e-Latifa	Yonan for Unan; and 'Falsafa' here lost its ending 'h' mentioned in category heading.
'Mirat ul Uroos' And 'Amrazul-Atfal' And 'Tijarat Bain-ul-Aqwami -wa- Mubadilat-e- Kharja'	Three different rules to mention 'ul' in Arabic compound words.
Hindustan Aur Mashriqi-e Wusta Ke Tijarati Ta'alluqat	Careless Rendering of: Mashriqi-e-Wusta. Have to stick with one rule for binding 'e' and no need of 'I' before 'e'.

Besides transliteration, the page is also victim of common spelling mistakes. Main Category names and some key words are translated into English. There again spelling mistakes and inconsistency of style give a very poor impression. Like on main page of digital library the categories are victim of following issues:

Text as appeared on webpage	Issue(s)
Bachon-Ka-Adab-Childrens-Literature	'Children's Literature'
Encycolopedia-Aur-Lughaat	Wrong transliteration of Encyclopedia and no English Translation provided, inconsistent to other entries.
Falsafah-Philosophy	Unnecessary 'h' at end of 'Falsafah'
Funoon-e-Latifah-Fine-Arts	Unnecessary "h" at end of 'Latifah'

Qanoon-Laws	As discipline it would spell 'Law'
Nafsiyaat-Psycology	Spelling: Psychology
NCERT-Books	Style Inconsistency, no Urdu transliteration, AND NCERT, not explained anywhere in the webpage, stands for National Council of Educational Research and Training.
Tareekh-History	Correct Translation and Transliteration!
Tib-O-Mualejaat-Bio-Medical	Wrong Translation: Bio-medical is entirely different branch of studies. The category refers to Health Sciences and contents suggest to 'Alternate Medicines'

If one passes on the wrong spellings of 'Encyclopedia' in category list as just loose transliteration, this could not be excused in main page button as 'Encyclopedia'.



Screen Shot: Spelling Errors

No use of Urdu font, wrong spellings and poor transliteration indeed is responsible for low ranking of the site in major search engines. Besides technical requirements (like removing very liberal use of frames in this site) use of Unicode Urdu font along with

standard transliteration could ensure better search, use and ultimately good ranking and impact of the repository.

Another technical issue is not assigning any internal or standard identifier to the items. All books, scanned or created digitally are not given their ISBN for digital edition. In this regard the guidelines are very much clear:

“One of the principles of ISBN has been that it identifies a unique product (e.g. an edition of a book). This has facilitated discovery and acquisitions, and enabled e-commerce, distribution and aggregation of product information, and sales data reporting. The ISBN standard, ISO 2108, has always required that different product forms of a publication, where these are made separately available, be assigned separate ISBNs.”⁵

The outdated copyright notice at end of the webpage also proves negligence to details, as it still mention ‘2012’ in mid of ‘2014’.

Issues for End User

Besides that technical issues that results in poor impression of the Council, there are serious design flaws of the page that restrict the utility thus impact of the repository. Fixed listing of the books in nine categories and four alphabetized pages are already mentioned. For any user, there is no way to search for specific required text, title or author. Site offer no internal search engine, while the books are not tagged with enough keywords to be sorted according to different requirements. Site contents are not searchable through any external search engine too. Besides, this site also does not present any system to track chronology of addition of resources. Thus even regular visitor could not find what new resources are added, and where (in categories or alphabetical list).

The management could address this problem through both short term and long term solutions. Long term solution is indeed remodeling of the website based on database that is fully tagged with keywords and introducing internal search engine in the design. This solution indeed requires time, resources and designing. Meanwhile a short term solution is at hand that could be operated within hours and without any financial budget. With powerful social media tools available free of cost could fixed that issue before overall design rebuilding. Followers at Twitter and fans at other social media sites like Facebook keep visitors inform of any new addition.

In random check for this article, at least two books were downloaded but were not readable by the software due to technical inconsistencies. These were '0682- Hindustan Ki Qanooni Tareekh Part.1' and '0677- Bharat Ka Aina_2'. A thorough and periodic check of each file for its quality and having an option for easy feedback by visitors indeed resolve the issue.

Despite very valuable Urdu resources and with noble and sheer will to share it with global community the NCUPL digital repository needs full remodeling of the website. Following steps indeed could result in more impact and value of this repository: attention to details of each page and title entries; use of spells check; redesigning the webpage based on a database which must be well annotated with tags and key words with each title, and having capability to internal items search; adoption or development of a transliteration rules and its use for each title along with Urdu Unicode fonts; and indeed frequent addition of new resources.

Notes & References

¹ Jacci Howard Bear, "White Space" in Desktop Publishing Glossary. Available at: <http://desktoppub.about.com/od/glossary/g/whitespace.htm> (retrieved on April 17, 2014)

² Leslie Chan, "Supporting and enhancing scholarship in the digital age: the role of open access institutional repository", *Canadian Journal of Communication*, 29: 3& 4 (2004): 280

³ Jeffrey Beall, "Metadata and data quality problems in the digital library", *Journal of Digital Information*, 6:3 (2006):2

⁴ The Annual of Urdu Studies (2014). Available at: <http://www.urdustudies.com/>, (retrieved on April, 4 2014)

⁵ The International ISBN Agency, "E-Books and ISBNs: a position paper and action points from the International ISBN Agency". Available at: www.isbn.org/sites/default/files/isbn_agency_e-books_position_paper.pdf, (retrieved on May, 2 2014)