

Usable State Resident Data Hub (U-SRDH)

Maharashtra State has been amongst the pioneers in creation of the State Resident Data Hub (SRDH). However, the implementation of the Direct Cash Transfer Scheme (DCTS) and other attempts for Aadhaar integration has suffered from challenges due to lack of data standardization in the UID databases. In a unique initiative, the Directorate of Information Technology is currently undertaking an unparalleled data quality and data integration initiative to make the contents of SRDH usable. The initiative aims to create a Usable version of the SRDH (U-SRDH), which will have data (including address) completely standardized and cleaned.

Recognizing that generic cleaning techniques and standardization rules would have limited impact given the specific localized context of Maharashtra State, the DIT is undertaking several important initiatives for formulating rules through data mining and analysis. A dedicated UID Laboratory has been set-up for this purpose and for creation of the U-SRDH. Some of the innovations currently underway include:

- Mining and Analysis of names in Maharashtra.
- Mining and Analysis of frequency of characters and patterns in name and address data
- Mining and Analysis to develop Marathi language competence and tools
- Standardization and seeding of UID data with Election rolls and Ration card information
- Standardization and seeding for beneficiary information from central and state schemes (especially for DCTS)
- Standardization and seeding of household data from the Maharashtra State Electricity board and integration through GIS platform
- Experiments are currently underway to match UID photographs with the photographs in other datasets such as electoral rolls, driving license details etc.

Some examples of benefits through the U-SRDH include the following:



Seeding through the U-SRDH will enable validation of important information about beneficiaries under various Central and State Schemes. For example, research through the U-SRDH has shown that there is mismatch between the ages of student availing scholarships where U-SRDH KYR details have shown that the beneficiary is not old enough to avail that benefit. This could potentially be bogus beneficiaries.



Integration of data from various data sets, such as Ration card data, MSED data etc. would enable the department to get several insights which will enable individual departments validate each of these datasets. For example, assessment of village lists in election database should that there are several regions where a polling station may not be available. The UID laboratoty is also planning to undertake assessment of correlation between voter turnouts and polling booth locations – especially in rural areas.



Matching of photographs could enable building further confidence into the in organic seeding done for Aadhaar integration across data sets.



The U-SRDH will enable rich analysis of data in the specific local context of Maharashtra State and will enable innovations for development of Marathi Language tools and competencies

The overall visualization of the U-SRDH components, processes and technologies is as shown below:

