SAMAGRA VEDIKA --TELANGANA'S INTEGRATED PLATFORM
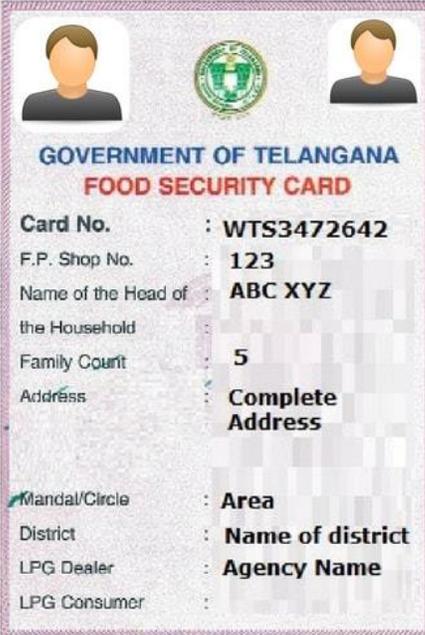
Using Big data, ML, Graph data base

FOR BETTER CITIZEN SERVICE DELIVERY AND TRANSPARENCY, ACCOUNTABLE AND EFFICIENT GOVERNANCE

ITE&C DEPARTMENT
GOVERNMENT OF TELANGANA

Few of the welfare schemes of Telangana Total budget is more than 35,000 Cr

Ration Cards

Student scholarships

Aasara Pensions

Raithu Bandhu

Most of the welfare schemes have eligibility in terms of Income

possible leakages in welfare programs and some still unplugged.

| Types of Fraud	Leakage on account of	Resolution through
1 **Identity Fraud**	**Bogus Beneficiary** Non-existent real beneficiary **Duplicate Beneficiary** Multiple registrations by same beneficiary. **Right Beneficiary** Limited means to establish identify of right beneficiaries.	These leakages have been controlled through use of Aadhar & ePOS
2 **Quantity Fraud**	**Illegitimate Claims** Claiming bills in MNREGA without work **Disproportionate Quantity** Availing more quantity of PDS Food grains than eligibility	Research on mitigation is currently ongoing
3 **Eligibility Fraud**	**Is the person truly eligible?** Limited means to correctly establish the eligibility	No solution in the country as on date as it requires data of other departments

Quantum Of Leakages Due To Wrong Inclusion

Total Budget For 2019-20 For Pensions
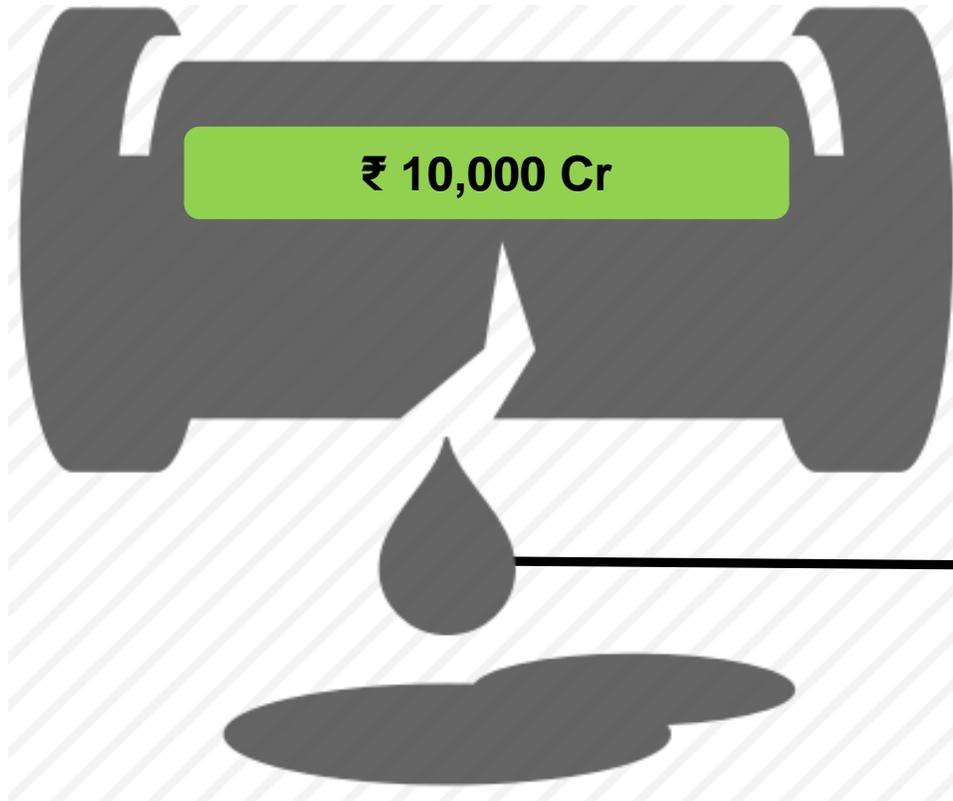
₹ 10,000 Cr

This is probably as big as a budget of a minor department!

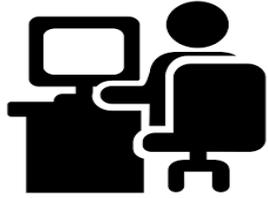₹ 100 Crores

Value of just 1% Leakage…

Eligibility for many benefits is prescribed.

People fulfilling one or more of the following conditions listed below shall not be eligible for Aasara Pension:

Classification	Aspect
Self-Economic Indicators	Having land more than 3.0 acres wet/ irrigated dry or 7 5 acres dry.
	Having large business Enterprise (oil/rice mills, pumps, shop owners etc.).
	Owners of light and/or heavy automobiles (four wheelers and big vehicles)
Family Based Parameters	Having children who are Government/Public sector/ Private sector employment / Out-sourced/Contract.
	Having children who are Doctors, Contractors, Professionals and Self employed.
Government Pensioner	Already receiving Government pensions or freedom fighter pensions.
Others	Any other criterion in which the verification officer may asses by the manner of lifestyle, occupation and possession of assets rendering the household as ineligible

Traditional process → ineffective implementation
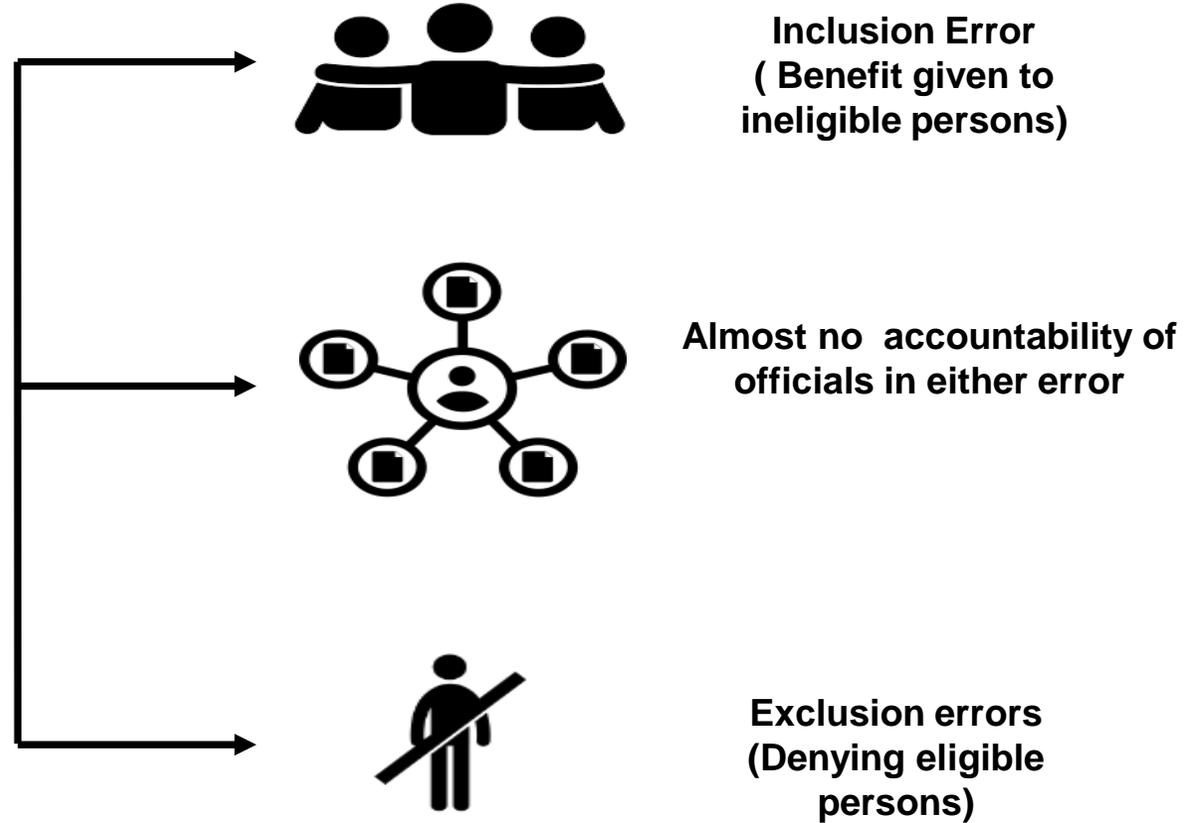
Manual Process

Reliance on Aadhar alone

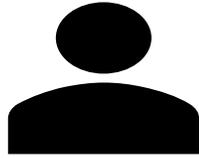**No Digital Trail**

High Human Discretion

Effects of ignoring these challenges

Inclusion Error (Benefit given to ineligible persons)

Almost no accountability of officials in either error

Exclusion errors (Denying eligible persons)

Opportunity & Challenges to get Consolidated view
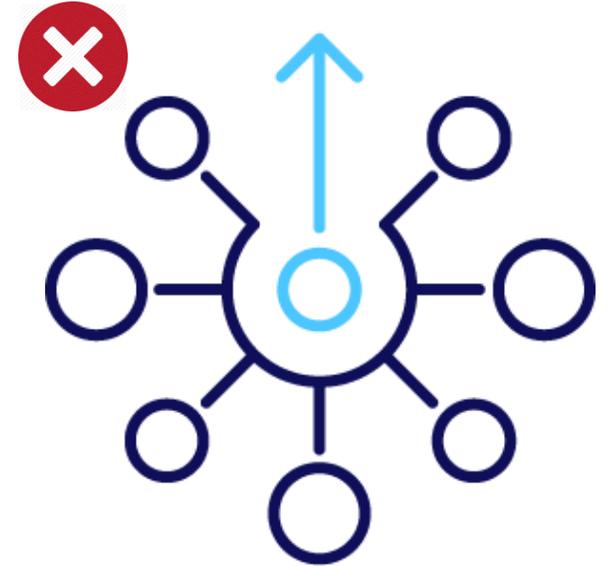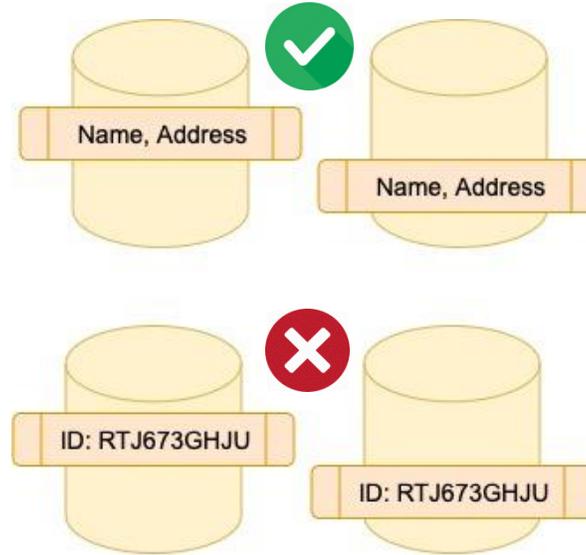
1. Most of the Data is in electronic form

1. Data is in Silos
2. No Common ID
3. Integrated view – (SSOT) Single Source of Truth is not available

Samagra Vedika One View - Objectives

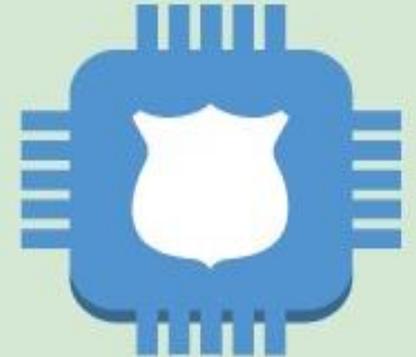| ON-DEMAND INFORMATION	REVENUE INCREASE	BETTER TARGET OF SUBSIDIES	DOCUMENT FREE GOVERNANCE	LAW ENFORCEMENT
On-demand information for informed decision making by authorized government officials	Property Tax, Water Tax, in Utilities	Pension, PDS, Scholarships, subsidized housing, Health care	Document Free Service Delivery	Smart Policing

**What is the alternative approach
Without using Aadhar or any other ID
But getting the same efficacy In view of
Legal restrictions on use of Aadhar**

Meta data attributes of an entity present in all data sources

- Following meta data information is available in every data source

Photograph (Some data sets)

Unique Personal Details
- Name
- DOB
- Fathers Name (In Some)

Contact Details
- Mobile No. (In some)
- Address (Res)
- Address (Off)

Unique ID* Number
- PAN or
- Passport no or
- Voter ID or
- Driving License

Any one ID is present
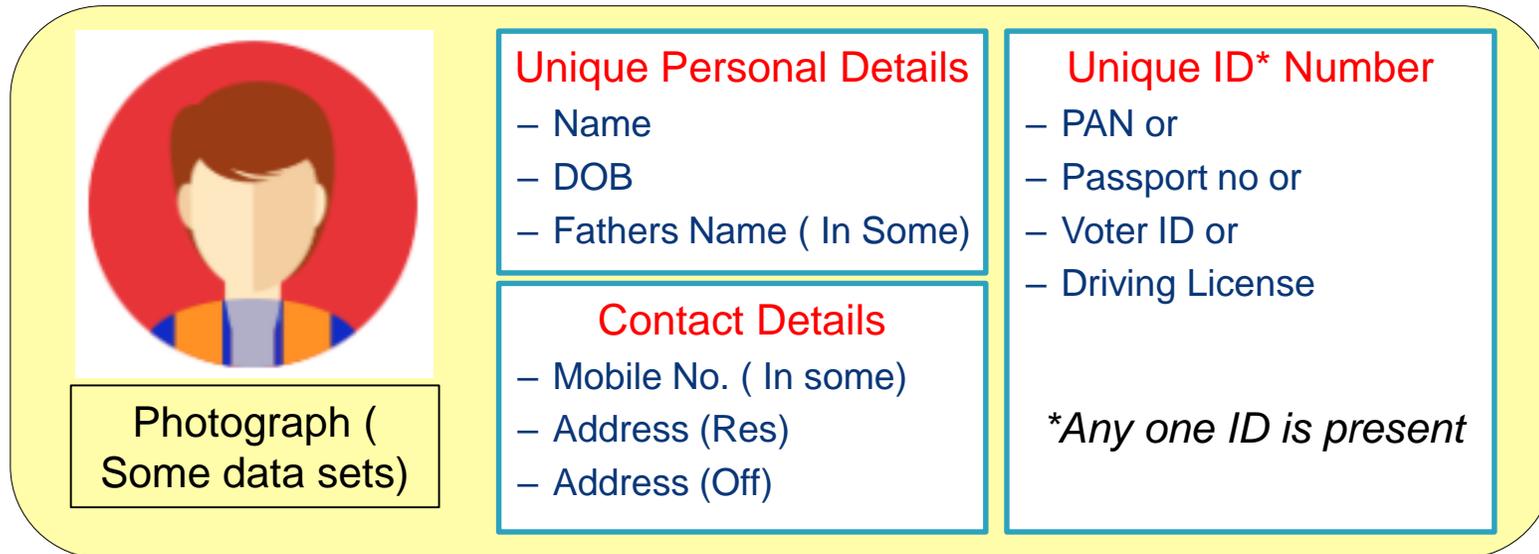
- All records in all data sources have Name, Address.
- Some records also have DoB, Phone Number, Fathers Name, Photo
- Can a combination of these attributes which are already available in every record be used to identify an entity
- With an Accuracy nearer to Aadhar based linkage
- With no manual intervention

3 "V" CHALLENGES

VELOCITY
The rate at which new data is generated is very high

VOLUME
Large volume of existing data. About 38 Cr records

3V (Variety, Volume, Velocity) challenge in non unique ID based record linkage is a classic Big Data use case

VARIETY
No defined standards for writing name/ address.

Examples of variations in Name & Father's Name

Spelling		• N Radha Murali Krishna • N Radha Muralee Krishna
Abbreviations		• N R Murali Krishna • N R M Krishna
Sequence Variation		• Murali Krishna N Radha • N M Radha Krishna
Addition/ Deletion		• N Murali Krishna • Murali Krishna
Splitting		• Murali Krishna N Radha • Radhakrishna N M
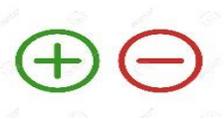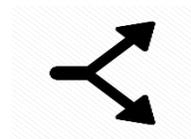
EXAMPLES OF VARIATIONS IN ADDRESSES

PLOT NO 110 RD NO 65 J HILLS CO OP SOCIETY HYDBAD

P NO 110 RD NO65 JUBILEE HILLS HYDRABAD

RD 65 JUBLEEHILLS HYDERBAAD

SAME ADDRESS IS WRITTEN IN MULTIPLE FORMS

110 ROAD NO 65 JUBLEE HILLS HBAD
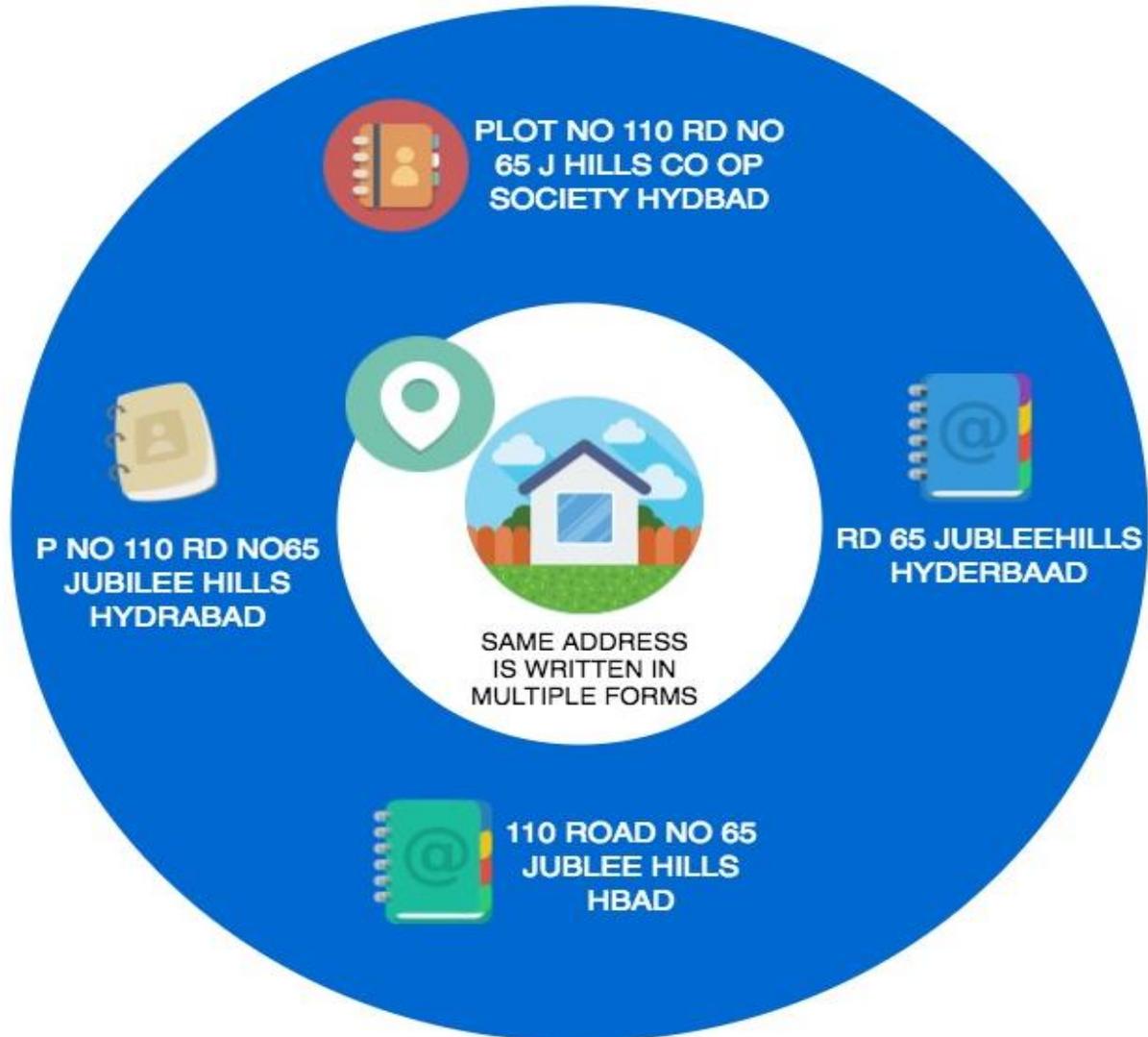
Variations and errors in date inputs

1. Interchange day and month: dd/mm/yyyy or mm/dd/yyyy

2. Input error in numbers

Key Performance Metrics

1 **HIGH PRECISION**
High positive matching. Never miss the right match.

2 **MINIMUM FALSE POSITIVES**
High negative matching accuracy. Get least number of false positivies.

3 **QUICK**
Very quick response time. Less than 2-3 seconds on 1 Cr records.

4 **ABILITY TO RANK OUTPUT**
Ability to rank results so that the best results come on top.

5 **DYNAMIC RULES**
Ability to define combination of multiple rules on Name, Address, DoB, etc.

6 **HANDLE VARIATIONS**
Handle all variations in Name, Address, DoB etc.

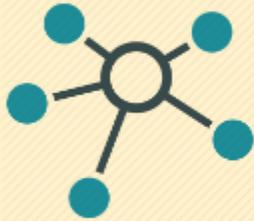All these are simultaneously acheived on very large databases (100 Million + records)

FOLLOWING TECHNOLOGIES ARE USED

Big Data Analysis Techniques

Entity Resolution

Graph DB: Relationship Discovery and Network Analysis

The record linkage and entity resolution should be on non unique attributes like Name, fathers name, address which vary greatly

Machine Learning

AI Powered Image Recognition

Economic Survey 2019 of GOI has praised Samagra Vedika

Economic Survey lavishes praise on 'Samagra Vedika'

The initiative is an integrated platform with databases

SERISH NANISETTI
HYDERABAD

The Economic Survey 2018-19 singles out Telangana's Samagra Vedika initiative as a lesson in federalism among governments. The policy document compares the State's digital initiative with Transport for London (TfL) and the digital transparency policy ushered under President Barack Obama in 2009 in the United States.

The TfL is putting out timetables, service status and disruption information in an open format, allowing others to access and use the data to improve their deliverables.

In its data convergence chapter titled: Data' Of the People, By the People, For the People', the Economic Survey holds up Telangana's example for harnessing data for social use.

The Telangana government's Samagra Vedika is an integrated platform with databases of various departments.

It was developed by the Information Technology, Electronics & Communications and Planning Department and uses most of the information collected during the Integrated Household Survey (IHS) in 2014.

Potential benefits

"The Telangana Government's Samagra Vedika initiative gives a flavour of the potential benefits of integrating data sets. The initiative links around twenty-five existing government datasets using a common identifier – the name and address of an individual," says the Economic Survey.

Samagra Vedika gives officials access to the information about the economic status of citizens and applicants concerning his/her family status, land, property, vehicles owned, GST receipts, income tax and even monthly power bills.

'Tamper-proof'

The Survey document also praises the data access architecture which it calls tamper-proof. "The initiative also puts in place all the necessary safeguards to preclude any tampering of data or violation of privacy."

The right to add or edit data in the database varies by ministry or department. A given department can only write data for select fields – the motor vehicles department cannot, for instance, manipulate data relating to education, even though it can view the data," says the survey document.

The Economic Survey also picks up Telangana's policy for Electric Vehicles.

Unlike other States which have set themselves numerical targets for the switchover to EV, Telangana has targeted 100 % electric buses for intracity, intercity and interstate transport for its State transport corporation.

Worst performers

But one parameter where the Telangana government didn't cover itself in glory is in the 'Open Defecation Free' status where it brings up the rear.

It is among the three worst performing States after Goa and Odisha.

While most States have reached 100% coverage of Indian Household Latrines (IHL) at least on paper, Telangana is shown as lagging at 73.99% which is better than Odisha 45.36% and Goa 5.84%.

Accurate targeting of subsidies

Beneficiaries for Old age pensions

Using Samagra Vedika for new sanctions of Aasara Pensions

Eligibility for Aasara Pensions is now 57 years

New applications are being received (expected about 7 to 8 lakh new pensioners)

**65,693 new applications approved by the Districts officials after verification
Are sent to SERP for sanction Aasara Pensions**

**In Aug 2019 SERP requested ITEC to check the eligibility
Through Samagra Vedika platform**

Total new applications received	Total eligible applications (as per Samagra Vedika)	Total ineligible applications (as per Samagra Vedika)
65,693	59,068	6,625 (10.1%) Value Rs 16 Cr per anum

Accurate targeting of subsidies

Predictive analytics based identification of beneficiaries for 2 BHK scheme using Big data

Govt of Telangana has a program to provide 2BHK houses to economically weaker sections

- Started in 2015 by the Govt. Of Telangana to provide 100% subsidized housing to the poor.
- No beneficiary contribution needed – one of its kind in India.
- Construction cost = 7-8 lakhs/house and total cost including land is 15-20 lakhs/house

Aligned with the objective of implementation of scheme in the entire state, Govt of Telangana is looking to distribute 2BHK houses to eligible persons.

In one district

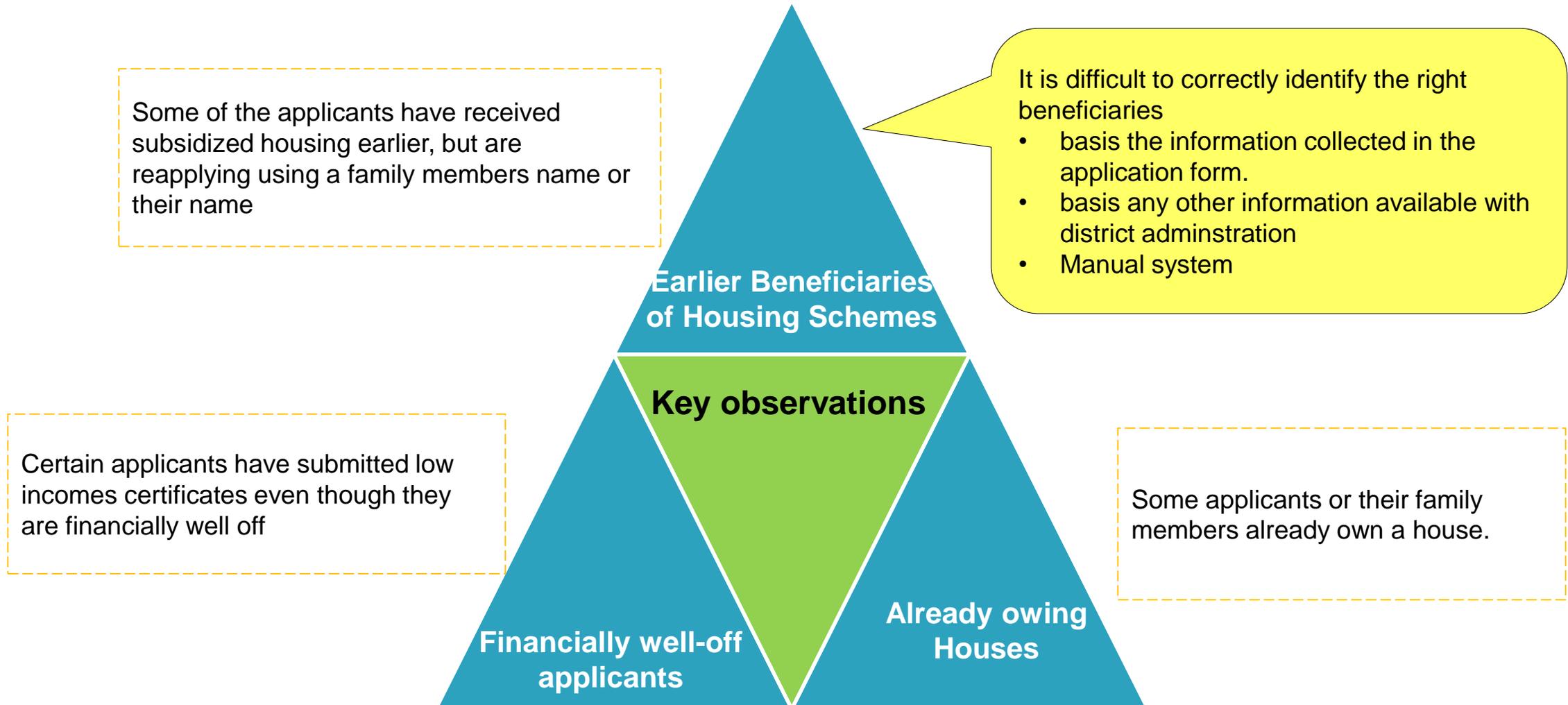|	
Total Houses constructed under the scheme	2000
Total applications received	11,681

The significant expenditure by Govt, and high mismatch in number of applicants and available houses has necessitated a very careful approach towards allotting the houses to applicants.

Siddipet Dist. Administration had following key observations.

Some of the applicants have received subsidized housing earlier, but are reapplying using a family members name or their name

It is difficult to correctly identify the right beneficiaries
- basis the information collected in the application form.
- basis any other information available with district adminstration
- Manual system

Earlier Beneficiaries of Housing Schemes

Key observations

Certain applicants have submitted low incomes certificates even though they are financially well off

Some applicants or their family members already own a house.

Financially well-off applicants

Already owing Houses

Predictive analytics using big data used following data sets available with Govt.

Information provided by applicant

- ❑ Name
- ❑ Fathers name
- ❑ Address
- ❑ Aadhar number
- ❑ Phone number
- ❑ Photo of the applicant
- ❑ Minor info.

Information about family members not provided in the application

datasets available

Electricity connection

Water connection

House and land database

Old age pension schemes

Vehicles database

Ration card database

Common databases are matched with the provided info. and they are further analyzed to bring out valuable insights in the form of applicant categories

The analysis using Samagra Vedika categorized the applicants in four categories as follows:

Categories ->	Category 1	Category 2	Category 3	Category 4
Classification	Qualify	Qualify with verification	Consider as low priority	Don't consider
Not financially well off	✓	✓	✓	✗
No housing benefits previously accepted	✓	✓	✓	✗
No other welfare schemes prior	✓	✓	✗	✗
From Siddipet - SKS	✓	✗	✓	✓
Count (% of Total)	2363 (20.2%)	2678 (22.9%)	2181 (18.7%)	4459 (38.2%)

In Hyderabad about 1,00,000 cards were removed in Aug 2016

There was some public resistance due to which people were asked to apply again. About 19,000 applied as on Dec 16, 14,000 cards were activated again after verification that the property is very small or the four wheeler is taken out of loan etc.

Net about 86,000 cards are removed from August 16.

Total subsidy saved is Rs 4.6 Cr every month from Aug 16 onwards.

The mistake of tagging the vehicle/house to a wrong person is less than 5% which shows the efficiency of the application

D2/1947/2016/Vol-II

Dated:31-12-2016.

To,
Sri G.T Venkateshwar Rao,
Commissioner, ESD,
Hyderabad

Sub:- CS – CRO's Unit – Information called for – regarding deletion of cards – report submission – Reg.

Ref:- Email Received from Sri G.T Venkateshwar Rao, Commissioner, ESD, Hyderabad dt 29-12-2016.

****

I invite kind attention to the reference cited wherein it was requested to furnish the information in the proforma prescribed. The information is prepared as given below.

Month	Ration cards deleted	Units in deleted cards	Ration cards restored during the month	Net Ration cards deleted	Units in deleted cards	Saving of subsidy in Rs for Rice/Wheat/Sugar for that month		
						Rice @ 27/- Kg	Wheat @ 7/- Kg	Sugar @ 6.75/- 1/2 Kg
Aug-16	10346	22783	0	10346	22783	₹ 36,90,846	₹ 72,422	₹ 69,836
Sep-16	102402	350809	2710	99692	341526	₹ 5,53,27,212	₹ 6,97,844	₹ 6,72,921
Oct-16	39906	115843	5248	34658	100609	₹ 1,62,98,658	₹ 2,42,606	₹ 6,72,921
Nov-16	10659	25489	4458	6201	14829	₹ 24,02,298	₹ 43,407	₹ 2,33,942
Dec-16	0	0	2723	0	0	0	0	0

Sd/- B. Bala Maya Devi., I.A.S.,
Chief Rationing Officer,
Hyderabad.

Note:- Applications for restoration of FS cards are being examined by concerned ACSO / Area Inspector and FS cards are restored if found eligible.

// Attested //
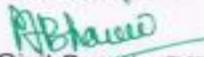
Assistant Civil Supply Officer
Hyderabad

THANK YOU

Any Thoughts? Or Ideas

??