OVERCOMING DATA INTEGRATION CHALLENGES

A presentation

By

Engr. Aliyu Abubakar Aziz
Director General/CEO,
National Identity Management Commission

Thursday 27th April, 2017
# Preamble

<table>
<thead>
<tr>
<th>Establish a national Identity Database</th>
<th>Provide identity services- authentication &amp; verification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue unique identifier</td>
<td>Harmonize &amp; integrate existing ID Databases</td>
</tr>
<tr>
<td>Issue a national ID card</td>
<td>Regulate the Nigerian identity Sector</td>
</tr>
</tbody>
</table>
Historical Background

- Fragmented
- No common standards
- SILOS of databases
- Multiple forms of ID cards
- No interoperability
- Low coverage
Consequences

Massive Exclusions
• Social benefits and the Poor
• Socio-economic Planning
• Claim of entitlements
• Basic rights

Undeveloped sectors -
(Financial, Industrial, Agricultural etc)
• Credit Bureau
• Consumer Credit
• Mortgage and other sectors
• Financial Inclusion

Security Issues
• Physical: Law and Order
• Business Environment
• Multiple/ghost/duplicate identities
• Frauds: Identity related Frauds & Scam
• Terrorism & Militancy:

Duplication of Systems/efforts
• Huge investment/infrastructure Cost
• High operational cost
• Maintenance of multiple & costly ‘silos ID databases’
• Creation of Idle Assets
Focus Now

Harmonized identity data across MDAs, connected by the National Identification Number - NIN
DATA HARMONIZATION AND INTEGRATION OBJECTIVE

- Uniform data and biometric standards
- Unique identifier to all (NIN)
- Mandatory use of National Identification Number (NIN)
- Universal coverage (Citizens, Residents, Diasporas)
- Interoperability (Connected Systems)
- One ID for Life
- Identity Data consolidation & Streamlining
- Trusted and Verifiable identity
Harmonizing and Integrating all Disparate Databases

- Voter Register
- Drivers’ License Register
- SIM Register
- BVN data from banks

*Pipeline*

National Identity DATABASE
History of Observed Data

History of observed data

- Harmonization Data
- Systems' Migration
- Presidential Card Launch
- Recession

Time
Jan-12  Jan-13  Jan-14  Jan-15  Jan-16  Jun-15  Jan-16  Jun-16  Jan-17

# enrolments
0  500000  1000000  1500000  2000000
Histogram of Observed Data

Mean Enrolment = 253,400
Deep Learning of the Observed Data
Latent Volatility Showing Improvements
Fitted Model tracks the Volatility over time
Source MDA Data

- NIBSS
  - BVN 30m
- NCC
  - SIMREG 150m
- FRSC
  - Drivers License 8m
- NIS
  - Passport 10m
- NHIS
- NPopC

HARMONISATION WORK FLOW

Online Services/ Extract Transform and Load

- Staging Tables (Source Lake Individual Tables per system)
- Consolidated Cleansed Tables
- Auto Profiling/ DQ Rules
- Auto Profiling/ DQ Rules
- Profiling Results
- Profiling Results
- Bad Records (per System)
- Bad Records (per System)

NIMS (ABIS, NIDB, NVS, Master Data Repository (MDR))

Source: NIMC

Recommended Systemic Changes/ Auto ... Rules Data Integrations

Reporting Area Profiling Results DQ Issues DQ Monitoring Bad Records

Source: NIMC
**DENDROGRAM OF HARMONISED & INTEGRATED DATA - OUTPUT**

**RECEIVED (Total Records from Source agency)**
- 3,976,852 (100%)

**PROCESSED**
- 3,923,415 (98.66%)

**Not PROCESSED**
- 53,437 (1.34%)

**NIN GENERATED**
- 2,351,510 (59.95%)

**HITS**
- 1,571,121 (40.05%)

**AFIS ERROR**
- 784 (0.02%)

**NON NIG.**
- 12,045 (22.54%)

**CHILDREN**
- 7,889 (14.76%)

**NOT ENOUGH FINGERPRINTS**
- 11,322 (21.19%)

**OTHERS e.g. Data Validation**
- 22,181 (41.51%)

**SUCCESSFUL NIN**
- 2,351,507 (99.9999%)

**BY PASS**
- 3 (0.0001%)

**Yet to be Cleared.**
- 91,128

**True HITS**
- 1,393,830 (94.18%)

**False HITS**
- 86,163 (5.82%)

**Hit**
- 37.65%

**Success**
- 56.35%

**False**
- 2.40%

**Yet to be Cleared.**
- 91,128

**Hit**
- 40.05%

**Success**
- 59.95%
DATA INTEGRATION CHALLENGES

- Data release strategy & Format
- No common standards in demographic & Biometric data captured
- Disparity in the number of data fields captured by different agencies.
- Skills and Personnel deficiency for large data processing
- High Volumes of records for analysis, conversion, cleansing and processing
- Volume of Records with True/False Hits
- Large amount of records requiring manual intervention (Adjudication)
DATA INTEGRATION CHALLENGES

- Poor quality biometrics captured by source Agencies
- Infrastructural challenges with processing huge data (Lack of or scale up)
- No central reference database for decision making
- Absence of a central ID authentication & verification service
- Peculiarities of Agency’s IT environment and Legal Framework
OVERCOMING DATA INTEGRATION CHALLENGES

…High level Government Directives on Harmonization

President Buhari’s directive:

“… that all ministries, departments and agencies (MDAs) should expeditiously harmonize their biometric databases”.

At a meeting with a Partner Agency, President Muhammadu Buhari charged government agencies to harmonize the collection and usage of biometric data in the country, instead of replicating their efforts in this regard.

VP Osinbajo Speaks…

…all government agencies collecting demographic and biometric data of citizens and legal residents, to aggregate their data into a single databank, to be domiciled with and managed by the National Identity Management Commission (NIMC).
OVERCOMING DATA INTEGRATION CHALLENGES

- Establishment of Committee on Harmonization & integration for buy-in and ownership (Membership of all Stakeholder)
- Development of Biometrics, Demographic & Interoperability Standards
- Development of Harmonization & Integration Framework
- Deployment of a separate Infrastructure to process the aggregated data
- Setting up of ad-hoc teams to deal with the volume of records (Technical, Adjudication, Steering Committee)
- Redeployment of more personnel to the Harmonization sub Project
OVERCOMING DATA INTEGRATION CHALLENGES

➢ Roll out of Authentication and Verification Services
➢ Development and use of demographic de-duplication to reduce records to be run through ABIS.
➢ Planned acquisition of Infrastructure, Hardware/Software such as Master Data Repository (MDR) solution to identify “Golden records”.
Centralized National Identity Database will drive efficiency & transparency in:

- Governance,
- Service delivery,
- Enhance Security,
- Enforce anti-money laundering compliance,
- Reduce Crime, etc.

The success of the programme depends on the commitment of government at the highest levels & all stakeholders
THANK YOU

AND

ENJOY THE REST OF THE CONFERENCE