



Datasét Aggregator

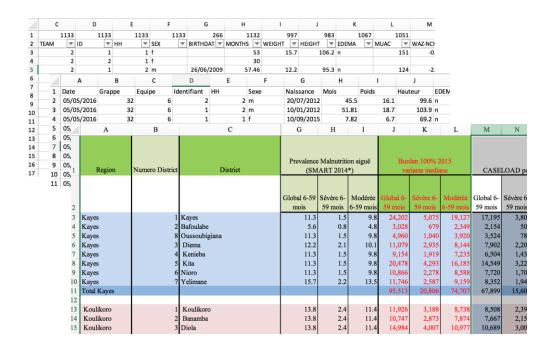
Milko Skofic

NIPN Technical Gathering July 4th, Paris



Summary

- Issue
- Solution
- Audience
- Limitations





Aggregating datasets

- Why aggregate datasets
 - Overview data spanning time
 - Overview from geographic locations
 - Other aggregations
 - Query the aggregate as a whole
- Which datasets can be aggregated
 - Same methods
 - Compatible formats
 - Share compatible set of variables





Aggregating datasets

Process

- Define a default data dictionary
- Find all related datasets
- Re-organise the variables in all datasets
- Append one by one all datasets into a single one
- Store and document the aggregated dataset

Workflow

- Repeat for all different types of datasets
- Repeat each time the process is needed



Aggregating datasets

Issues

- Find the required files
- Share the same variable definitions
- Add metadata to individual datasets
- Attach or update original copy
- Organise aggregated datasets





Solution

- Network application
- Centralised data dictionary
- Dataset level metadata
- Study level metadata
- Archive for files and data
- Assisted variables harmonisation workflow
- Database for datasets and their data
- Query interface on datasets and their data





Usage

- Register study
 - Document study
 - Upload original files (datasets, reports, etc.)
 - Document individual files
 - Set dataset summary statistics





Usage

- Upload dataset data
 - Store in work area
 - Match with data dictionary variables
 - Data validation reports
 - Correct eventual errors
 - Load data in aggregated collection





Usage

- Search
 - Query dataset metadata
 - Query dataset data
 - Two sets of results
 - Studies
 - Data
 - Download study archives
 - Download data selections





Audience

- Archiving
- Working datasets
- Summary reports
- Harmonise historical data
- Share standards with data providers



Limitations

- Currently under development
- Initially targets Standardised Monitoring and Assessment of Relief and Transition (SMART) surveys
- Need feedback to enrich features



Questions?

Thank you!

