Modernizing Statistical Processes with SDMX-Compliant Automate Pipelines

Thailand's NSO Case

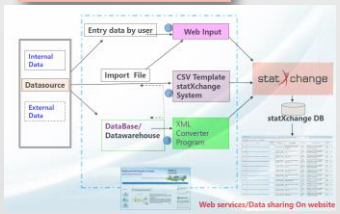
By Tidarat Nansing
National Statistical Office

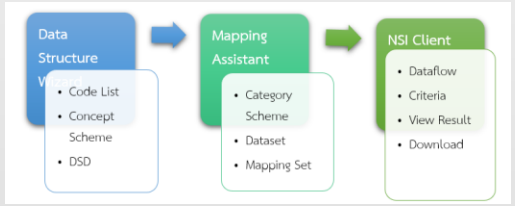
Content

- TNSO SDMX Journey
- SDMX Experience
- Use case/Success Stories
- SDMX Stage
- Data Integration Project
- DI Project Stage
- Core Focus/Pipeline Step
- Challenge
- Conclusion & Next Move

TNSO SDMX Journey

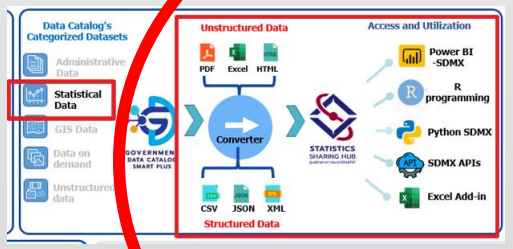
- The UNSD-DFID
- Project for MDGs exchange
- Implementing SDMX in statXchange (Using SDMX-RI)

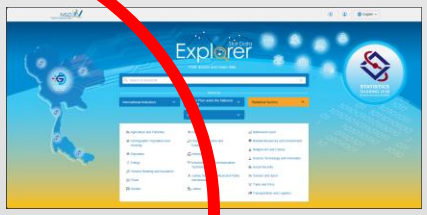




Participate
- SDMX Global Conference
- SDMX Expert Workshop
- SDMX Capacity Building

- Implementing SDMX with new platform
- Applying SDMX to improve statistical business process
- Learning SDMX IT Tools
- Collaborate KSTA-9646 Data for Development Phase II Project for SDGs exchange
- Integrate system to Dissemination platform





Data Integration Project

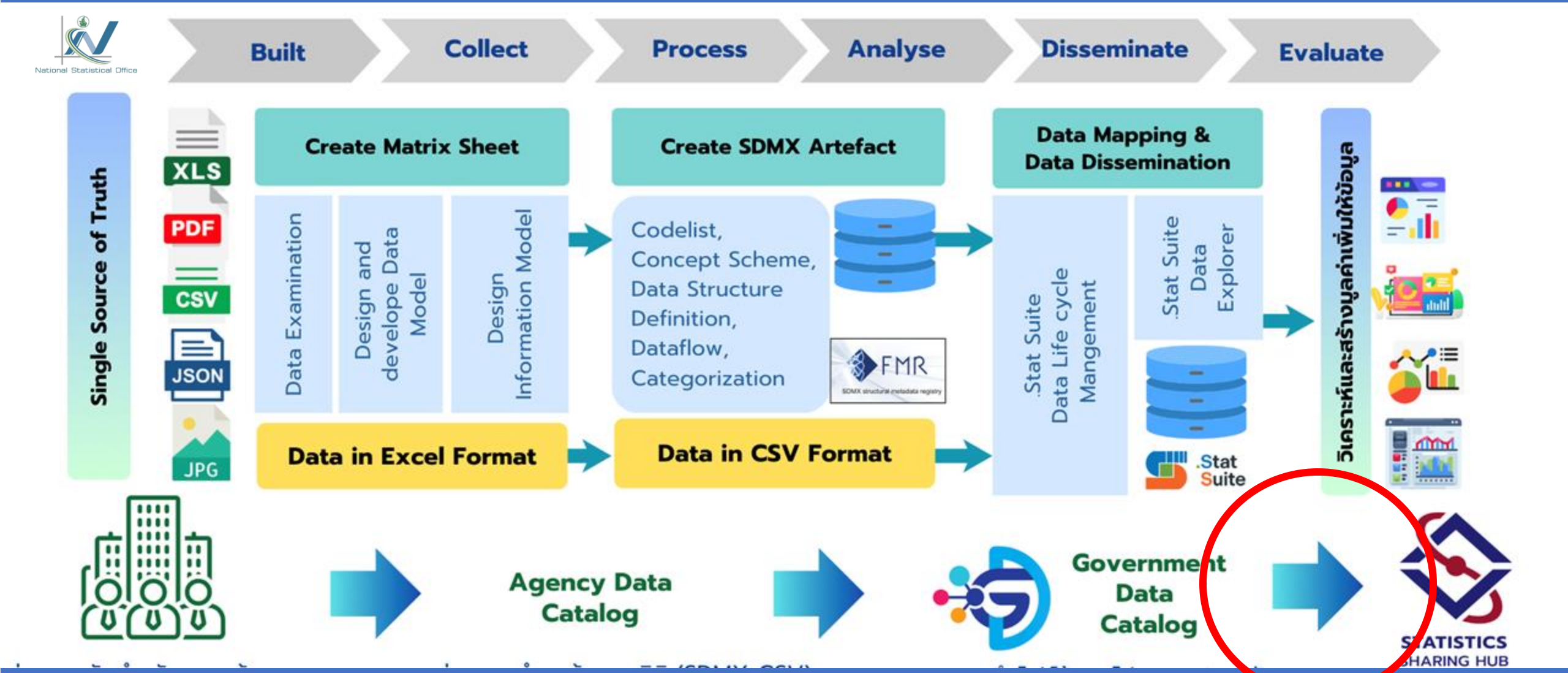
2010
2012
2018
2019
2021
2023
2024

SDMX Experience

SDMX Use /Success Stories

- Strong interest in adopting the SDMX standard for data management, Reporting and dissemination
- Deployment of SDMX IT Tools on TNSO Infrastructure and Training (FMR, SDMX Converter, .stat suite)
- Providing SDMX-compliant Data Interfaces.
- Capacity development on and data migration of legacy indicator datasets and other indicators (21 statistics sector)
- Integrate with dissemination platforms
- Public on Statistics Sharing Hub (https://stathub.nso.go.th)

SDMX Stage


National Statistical Office
Built
Collect
Process
Analyse
Disseminate
Evaluate
Single Source of Truth
XLS
PDF
CSV
JSON
JPG
Create Matrix Sheet
Data Examination
Design and develope Data Model
Design Information Model
Data in Excel Format
Create SDMX Artefact
Codelist, Concept Scheme, Data Structure Definition, Dataflow, Categorization
FMR
Data in CSV Format
Data Mapping & Data Dissemination
.Stat Suite Data Life cycle Mangement
.Stat Suite Data Explorer
.Stat Suite
วิเคราะห์และสร้างมูลค่าเพิ่มให้ข้อมูล
Agency Data Catalog
Government Data Catalog
STATISTICS SHARING HUB

Data Integration Project

Solution Deployment & Deliverables

Objective

An Automate System
Power by RPA & integration tools

Supports multiple formats
Manage PDF, Excel, JSON, APIs

End-to-End Data Pipeline
Covering extraction, transformation and dissemination

SDMX Standard
Enhancing data quality and accessibility

Data Conversion Service

Graphical Data Display Service

System for Statistical Compilation

Robotic Process Automation (RPA) System

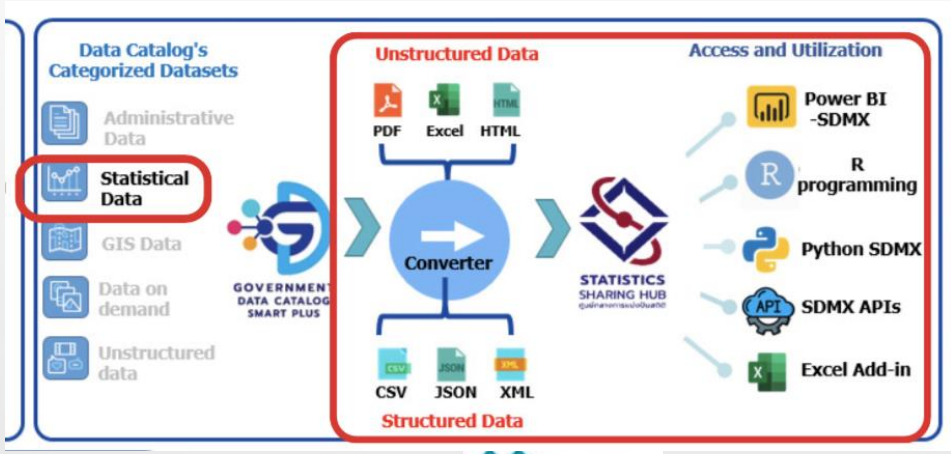
DI Project Stage

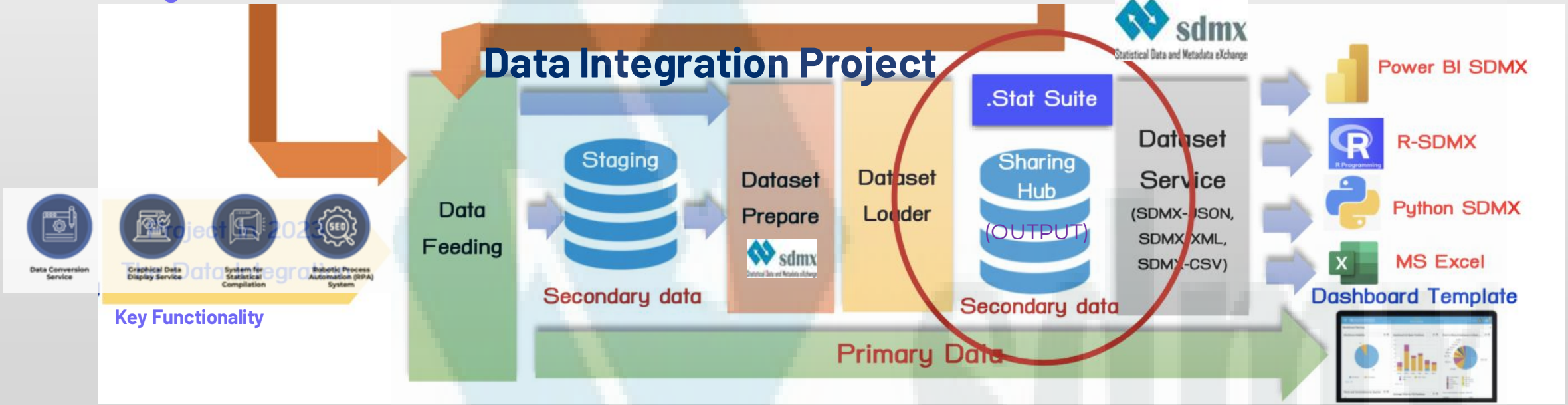
Robotic Process Automation

Airbyte Extractors

dbt for Transformation

Dagster Orchestration





Core Focus

Automate Data Collection and Data Integration

Using Robotic Process Automation RPA for varied source

Standardize Data Output

Transform Data into SDMX-CSV Compliant files

Enhance Data Management

Improve Timeliness, Accuracy and Sharing Capabilities

Support Effective SDMX Data to Dissemination

Enable broad sharing across platforms and stakeholders

Pipeline Step

1. **Data Sources**

   Diverse official statistics inputs

2. **Extraction**

   Automated with RPA and Airbyte connectors

3. **Transformation**

   dbt maps to SDMX-CSV standards

4. **Orchestration**

   Dagster schedules and automates tasks

5. **Dissemination**

   Standardized data shared via Statistics Sharing Hub

Challenge

Data Source Diversity

Multiple heterogeneous sources limit consolidation

Pipeline Gaps

No unified system for SDMX file preparation

Update Delays

Manual steps cause lag in statistical releases

Work as a Team

Integrating collaboration between Data expert teams and IT professional

Conclusion & Next Move

- Modernizes statistical workflows across government
- Enhances accessibility, reliability, and automation
- Aligns with SDMX standards and SDMX IT Tools
- Supports integrated policy making and dissemination
- Expanded role for AI Data validation, Report checks
- Expanded AI Analytics

THANK YOU!

Tidarat Nansing
National Statistical Office