Statistical Business Process for Big Data Usage

Action Area C. Integrated Statistics for Integrated Analysis (SC1)

Methodological approaches to integrated analysis : Use of sound methodologies

Presenter:
Isnaeni Noviyanti
Statistics Indonesia

#apstatsweek2020

Introduction

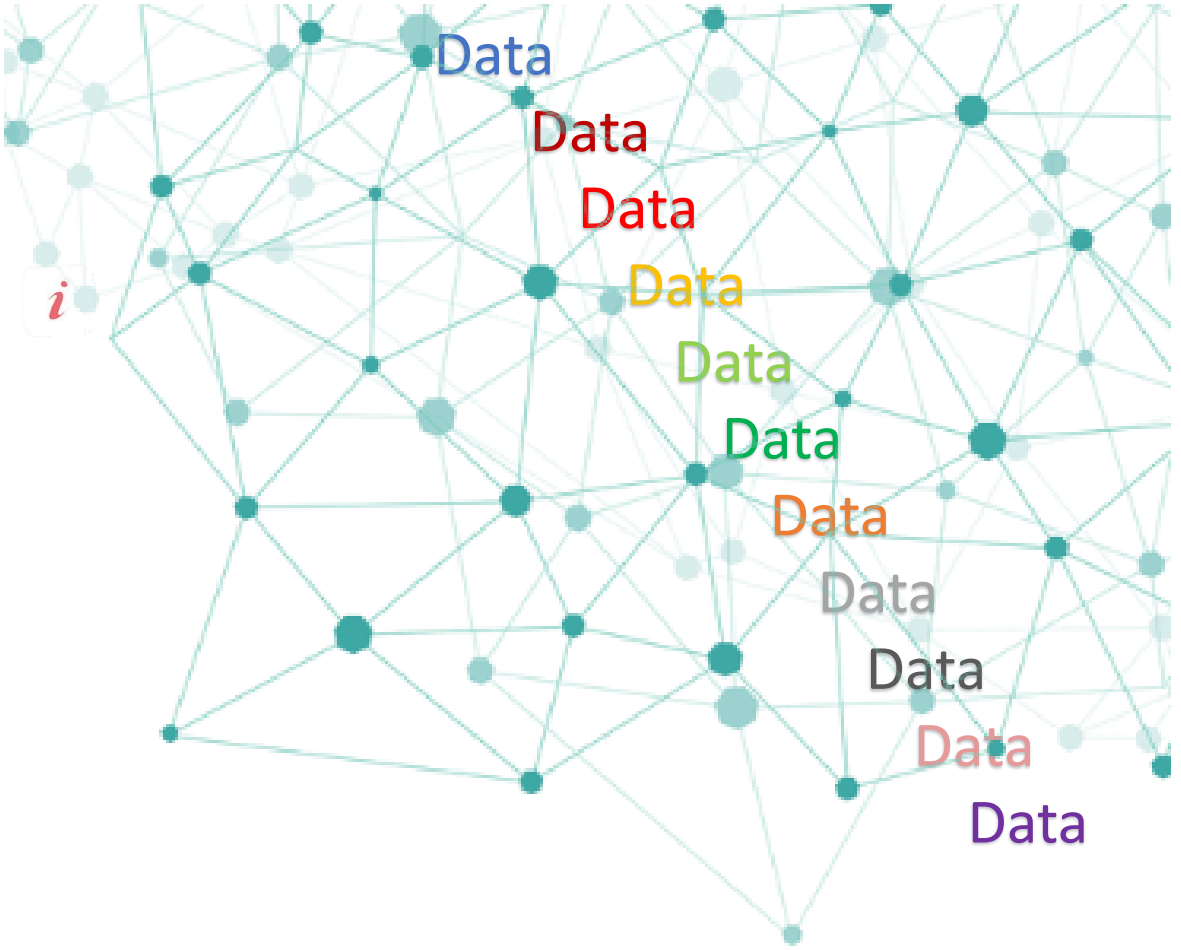
Administrative Data

Finance

Export-import

Education

Introduction



Introduction

Current Condition

Mobile positioning data

Google Trends

Web Scrapping

Insight

Current Condition

Phase I
- Big Data Platform Initiative
- Preliminary Study

Development Phase

Phase II
Adoption of Big Data Technology

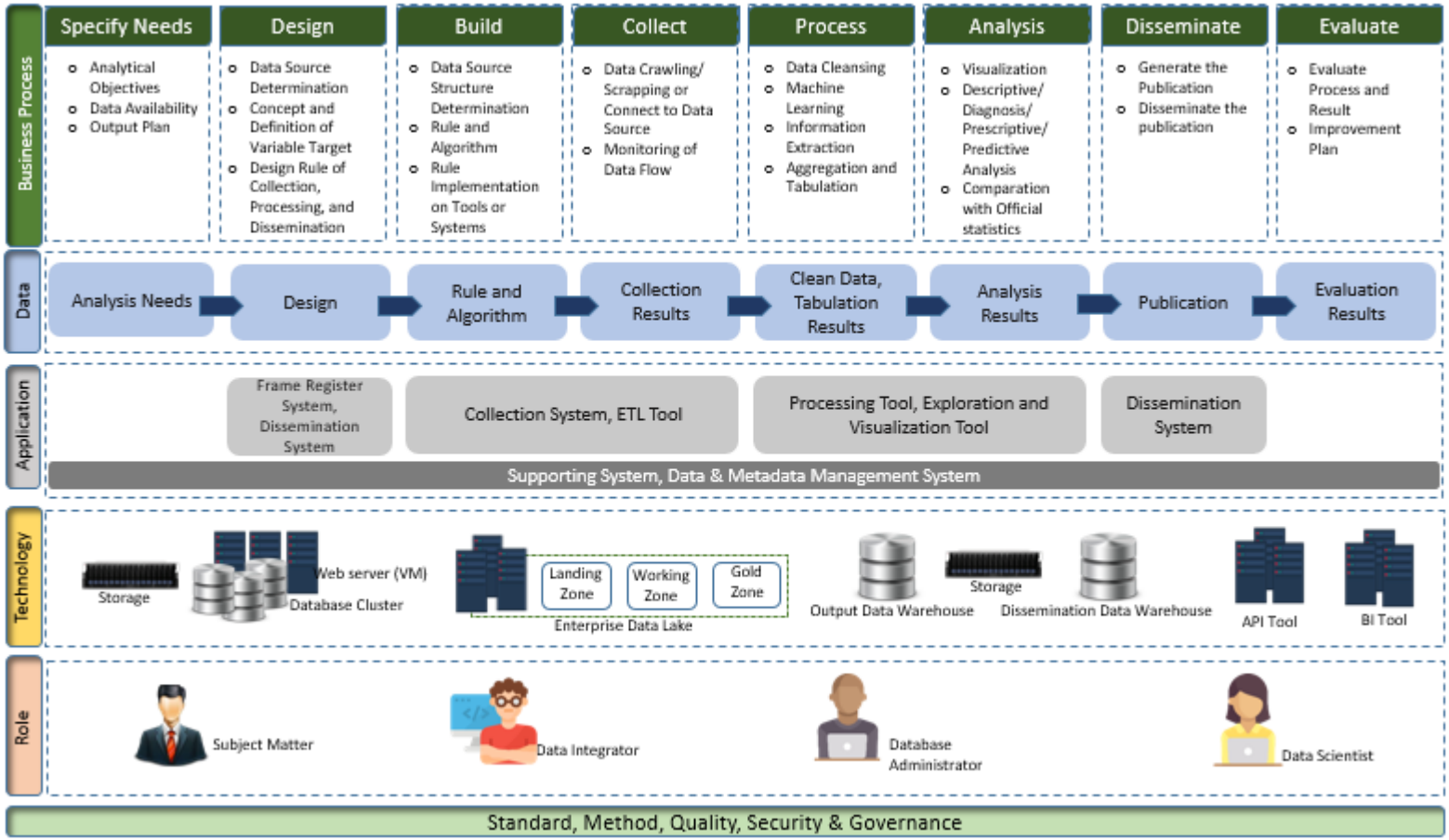
Phase III
Integration Newly Adopted and Existing Technology

Solution

Technology is just one of solution

- Business
- Data
- Application
- Technology
- Security
- Standard, Method, Quality, Governance

Solution

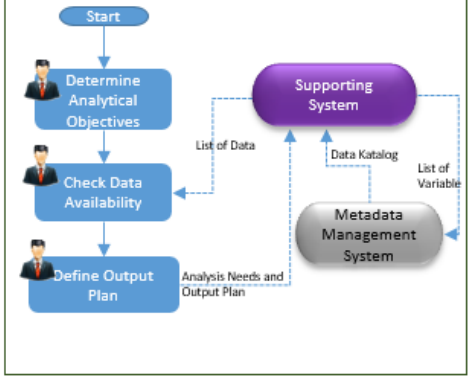

Business Process
Specify Needs
Analytical Objectives
Data Availability
Output Plan
Design
Data Source Determination
Concept and Definition of Variable Target
Design Rule of Collection, Processing, and Dissemination
Build
Data Source Structure Determination
Rule and Algorithm
Rule Implementation on Tools or Systems
Collect
Data Crawling/ Scrapping or Connect to Data Source
Monitoring of Data Flow
Process
Data Cleansing
Machine Learning
Information Extraction
Aggregation and Tabulation
Analysis
Visualization
Descriptive/ Diagnosis/ Prescriptive/ Predictive Analysis
Comparation with Official statistics
Disseminate
Generate the Publication
Disseminate the publication
Evaluate
Evaluate Process and Result
Improvement Plan
Data
Analysis Needs
Design
Rule and Algorithm
Collection Results
Clean Data, Tabulation Results
Analysis Results
Publication
Evaluation Results
Application
Frame Register System, Dissemination System
Collection System, ETL Tool
Processing Tool, Exploration and Visualization Tool
Dissemination System
Supporting System, Data & Metadata Management System
Technology
Storage
Web server (VM)
Database Cluster
Landing Zone
Working Zone
Gold Zone
Enterprise Data Lake
Output Data Warehouse
Storage
Dissemination Data Warehouse
API Tool
BI Tool
Role
Subject Matter
Data Integrator
Database Administrator
Data Scientist
Standard, Method, Quality, Security & Governance

Virtual Event 15-18 June 2020

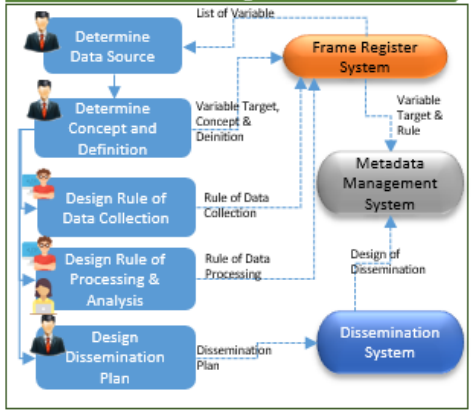
2020 Asia-Pacific Statistics Week

Leaving no one and nowhere behind

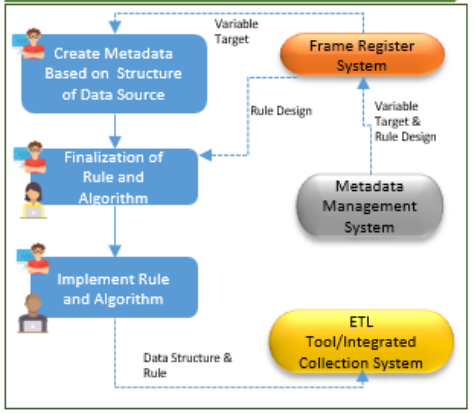
Specify Needs



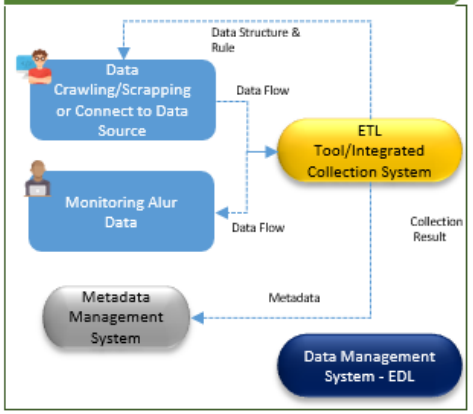
Design



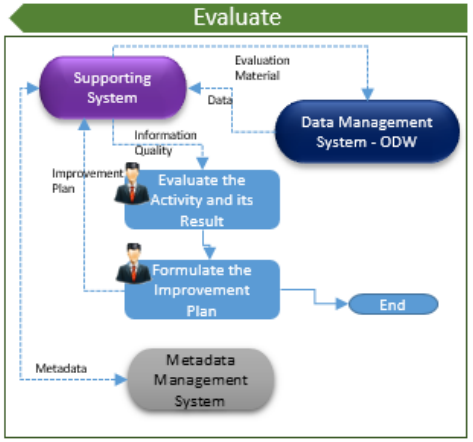
Build



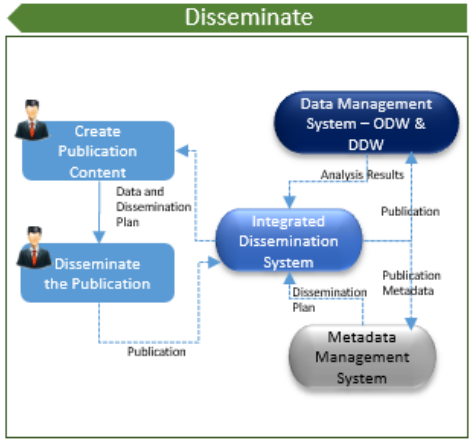
Collect



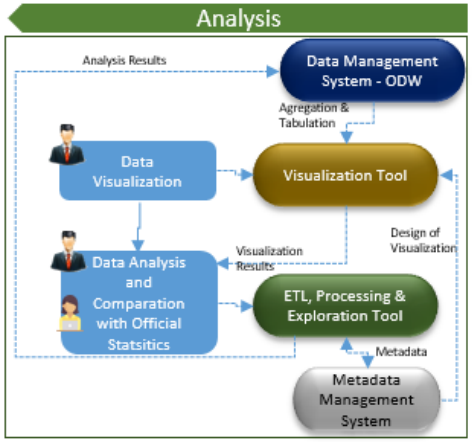
Evaluate



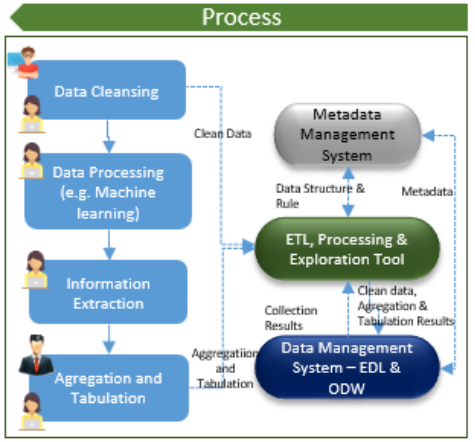
Disseminate



Analysis

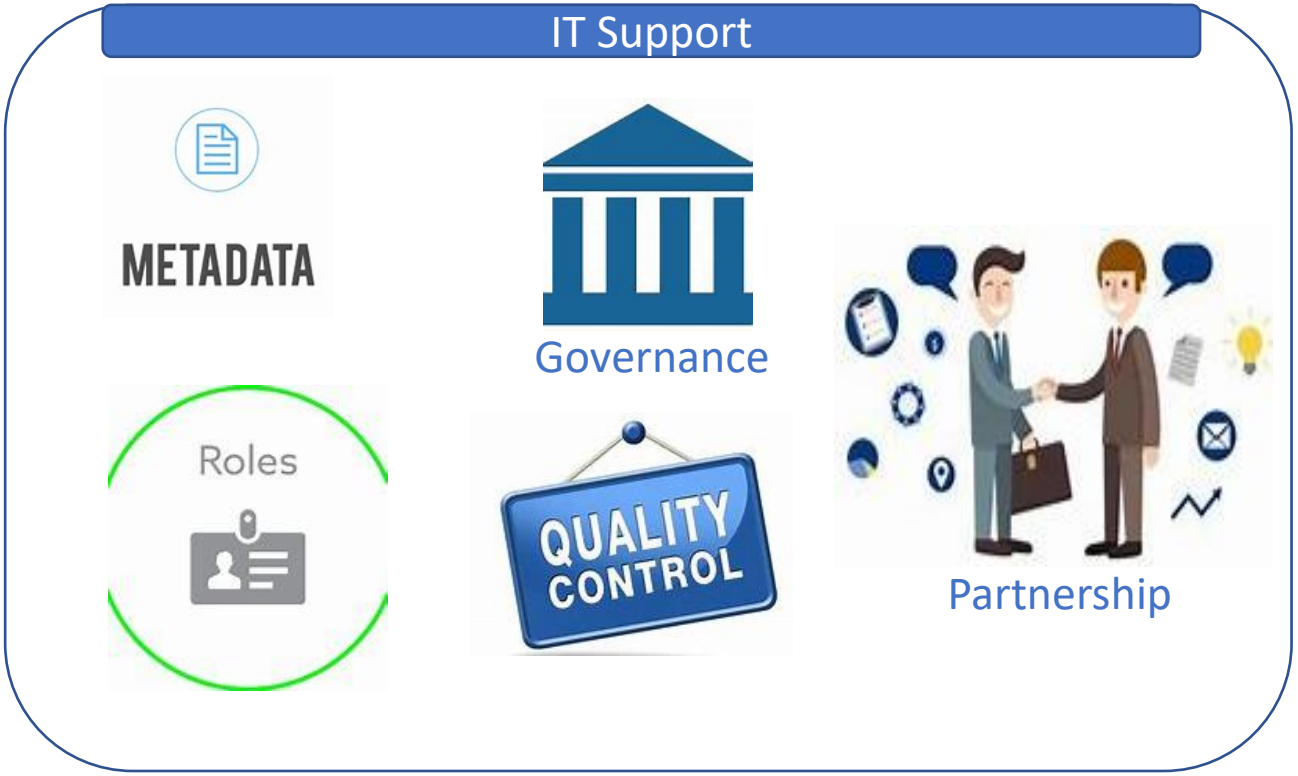


Process

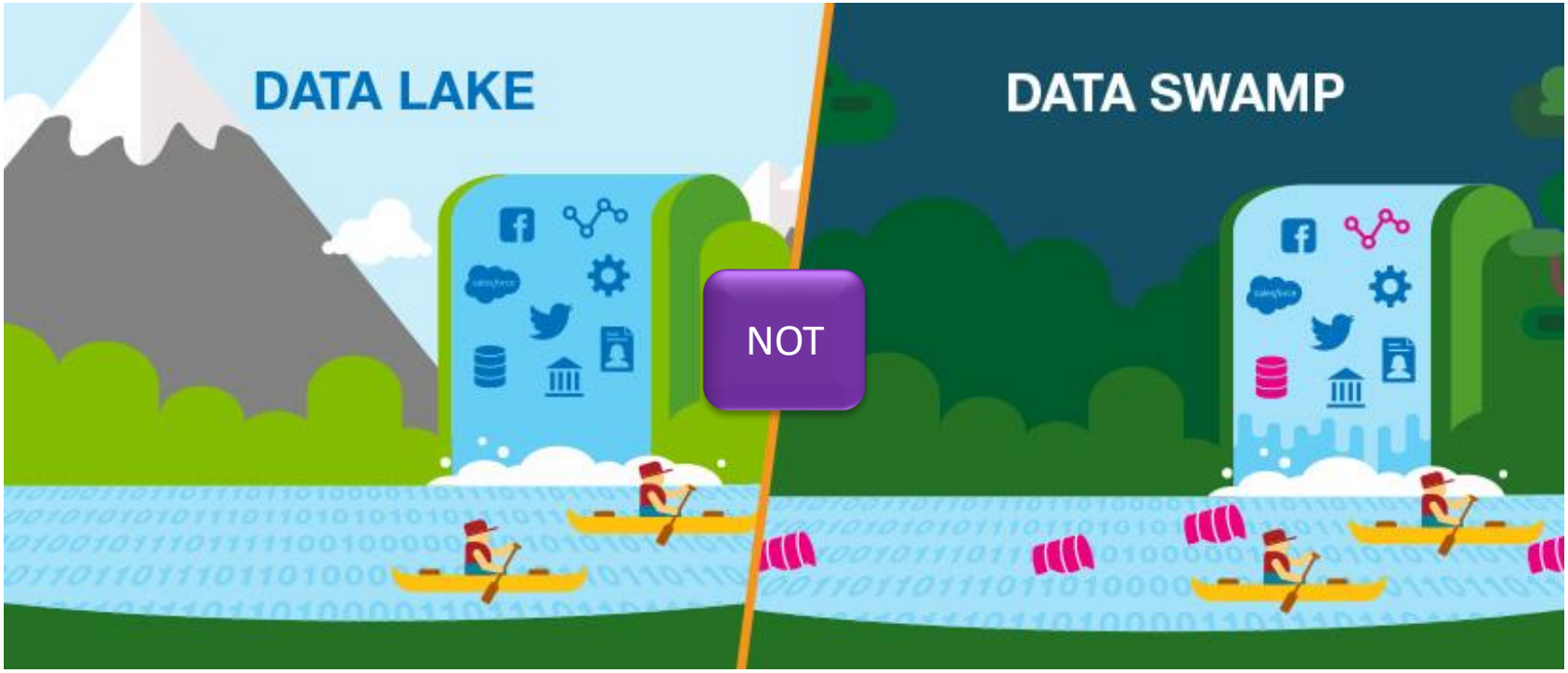


Note: Subject Matter | Data Integrator | Database Administrator | Data Scientist

Recommendation



Conclusion



Authors:
Isnaeni Noviyanti
Dwi Pupita Sari
M. Hanif Fahyuananto
Ade Koswara