SOEN 6611 - SOFTWARE MEASUREMENT: THEORY AND PRACTICE

Project Step 2 (1.5 points, due before midnight on July 11th)

Summer 2022

Course Instructor: Dr. Olga Ormandjieva

Source: SEI *Implementing Goal-Driven Measurement* course material (adapted).

TEAM INFORMATION

Name	Student ID
Hasandeep Singh	40198687
Jasleen Kaur	40218417
Anushka Sharma	40159259
Kavleen Kour Sidhu	40205476

Objective: Operationalize Goals, Derive Success Criteria, and Indicators

Step 2 tasks:

2.1 Operationalized Goals

The objective of the first part of this step is to express your measurement goal derived in Step 1, in a structured statement that identifies the object, purpose, quality focus & perspective, environment and constraints. This information is typically needed to gain required insight and/or to enhance decision-making. The objective of the second part is to develop success criteria and success indicators that will allow you to answer the measurement questions (from Step 1) quantitatively and then communicate the results to others.

Operationalized Goals

For each measurement goal you documented in Step 1:

• Operationalize it as a structured statement that identifies the object, purpose, quality focus & perspective, and environment and constraints.

• Document your results using the template below

1) Volume

Operationalized Goal: Label and description	To check if the Volume of the dataset received
Corresponding Measurement Goal label	Volume
Object of interest	Big Data
Purpose	Records in the data set should be sufficient so that it can be divided efficiently into time frames and receive valid results when functions are performed.
Quality Focus, Perspective	Check the entries in different fields and null values by perspective of a Data Engineers and Strategy Manager
Environment and Constraints	● Only a handful people will validate dataset like 5 people ● Using agile development methodology. ● Team should not be biased for a particular attribute ● Users feedback is taken into consideration Constraint: ● Lack of large big data visibility and transparency ● Incomplete and incorrect big data ● Inaccurate and duplicate big data

2) Vincularity

Operationalized Goal: Label and description	To check the vincularity is present
Corresponding Measurement Goal label	Vincularity
Object of interest	Tasks performed yielded Software Product, Big Data
Purpose	The set of reformed big data should be well linked to each other and have good connection to previous stage/ task
Quality Focus, Perspective	The linkage and connectivity would be checked and unveiled by Big Data Managers and Big Data Engineers/Developers
Environment and Constraints	● Only a handful people will validate dataset like 2-3 people ● Using agile development methodology. ● Team should not create non valid linkage Constraint: ● Lack of data interlinking visibility ● Incomplete and inconsistent data ● Inaccurate and duplicate data

3) Variety

Operationalized Goal: Label and description	To check the Variety or the data in terms of all possible categories
Corresponding Measurement Goal label	Variety
Object of interest	Big Data
Purpose	Big Data should be diverse so that it doesn't turn out to be biased towards one category after processing for machine learning algorithm
Quality Focus, Perspective	The big data should be checked for different categories it can offer by Developer that can utilise it to segregate and classify efficiently
Environment and Constraints	● Only a handful people will validate dataset like 4 people ● Using agile development methodology. Constraint: ● Big Data should be sufficient and from a varied data generating category ● Lack of distribution over differently classified data ● Incomplete and imbalanced data

4) Velocity

Operationalized Goal: Label and description	To check the velocity of data generated which requires distinct processing techniques
Corresponding Measurement Goal label	Velocity
Object of interest	Input Big Data and system output
Purpose	The streaming of fast-moving big data into bulk storage for later batch processing. The vital part is feedback loop, taking big data from input through the end-decision process
Quality Focus, Perspective	These needs have driven the development of key-value stores and columnar databases, optimised for fast retrieval of pre-computed information
Environment and Constraints	Constraint: ● Big Data amount should be sufficient as per time frame ● Number of null values in upcoming data ● The data recorded is significantly new

5) Veracity

Operationalized Goal: Label and description	To check the Veracity of big data in three main dimensions- ● objectivity/subjectivity ● truthfulness/deception ● credibility/implausibility
Corresponding Measurement Goal label	Veracity
Object of interest	Consistency and trustworthiness in Big Data. Removal of errors and outliers
Purpose	To convert the unstructured and incomplete big data into consistent, consolidated, and united source of information for an enterprise
Quality Focus, Perspective	Focus incredibly to get accurate results which help in data-driven decisions. Biases and ambiguities need to be identified and accounted for to reduce inference errors and improve the accuracy of generated insights.
Environment and Constraints	The captured big data is not qualified enough to check for reliability. Some of the issues that can be faced are: ● Rumours ● Spammers ● Collection errors ● Entry errors ● System errors

6) Validity

Operationalized Goal: Label and description	To describe the degree of Validity of big data left usable for users or enterprises. Perform check to confirm whether data satisfies user-defined conditions or falls within a user-defined range.
Corresponding Measurement Goal label	Validity
Object of interest	Big Data and tasks
Purpose	To evaluate the correctness and accuracy of the data to perform further procedures
Quality Focus, Perspective	The big data is maintained without being damaged, so that valid input followed by correct processing yields accurate results.
Environment and Constraints	● Only a handful people will validate dataset like 5 people ● Using agile development methodology. ● Team should not overlook for any attribute/ time frame ● Users feedback is taken into consideration Constraint: ● Lack of large data visibility and transparency ● Incomplete and incorrect data ● Inaccurate and duplicate data