SOEN 6611 – Software Measurement

PROJECT STEP 2

Source - SEI implementing Goal-Driven Measurement course material (adapted).

Objective - Operationalize Goals, Derive Success Criteria, and Indicators.

Submitted to - Prof. Dr. Olga Ormandjieva

Team – 7

Rutwikkumar Sunilkumar Patel – (40160646)

Charit Pareshbhai Patel – (40160658)

Deep Pareshkumar Patel – (40185585)

Bhoomi Shah – (40169655)

Table of Contents

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.

Evaluation criteria for 2.1: Operationalized goals will be evaluated as follows -
- The purpose is clear and specific
- The purpose is consistent with the perspective
- The perspective is the perspective of the "group/role"
- The purpose and quality focus apply realistically
- The object of interest represents some underlying model and attributes

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.
- Results have been documented below.

Operationalized goal label and description	OG1 The size of a dataset should be sufficiently large to use in machine learning algorithms.
Corresponding measurement goal label	MG1 (**Volume**)
Object of interest	Big data
Purpose	Analyze the collected wide range of data from different sources to compare the output at every stage and make corrective decisions or plans in the goodwill of an organization.
Quality Focus, Perspective	Examine the size of a dataset from the perspective of the product manager, owner, and scientists.
Environment and Constraints	To handle the large dataset, our system may require some extra environmental support, for example, multiple chips with high processing power, large secondary storage, high RAM, good

| | memory management for data transfer between secondary storage and RAM, Data distribution between processing units, etc.

The selection of data can be tedious as we have many data sources available nowadays.

Factors and Parameters
Application Factors: The application should be capable of processing the data.
People Factors: Product managers, scientists, and developers.
Resource Factors: Various sources that help in generating the relevant data.
Process Factors: Data generation process, Quality assurance process.
Customer Factors: Increase customer satisfaction.
Methods: Dataset size calculator and comparator
Tools: Size calculation tool
Constraints: Large dataset with a wide variety is easily available. |
|---|---|

Operationalized Goal label and description	OG2 Support improvised decision making by continuously gathering data and processing it faster.
Corresponding Measurement Goal label	MG2 (**Velocity**)
Object of Interest	Big Data
Purpose	Analyze the data flow and its processing in order to improve the data generation as well as data processing speed.
Quality Focus, Perspective	Examine the frequent changes in the large size dataset from the point of view of developers and managers.
Environment, and Constraints	The frequent addition of big data must be consistent with the old ones, otherwise results in ambiguous responses to a request resulting in degradation of service. The processing of these newly added data must be fast when used for any given operation.

	Factors and Parameters
	Application Factors: The software processing the dataset must be able to process frequently changing data. **People Factors**: Data scientists, developers, and managers. **Resource Factors**: Embrace the change **Process Factors**: Time elapsed for collecting data in a large dataset. **Customer Factors**: Data quality and efficiency. **Methods**: Incremental data processing. **Tools**: Data processing tools **Constraints**: Making notes of different versions of a dataset for managing purposes.

Operationalized Goal label and description	OG3 To improve the state of the data in the dataset and use the necessary and relevant data with context to specific system needs.
Corresponding Measurement Goal label	MG3 (**Veracity**)
Object of Interest	Big Data
Purpose	Evaluate the state of data in the dataset in order to improve the quality and efficiency of the dataset.
Quality Focus, Perspective	Examine the different defects in the dataset to improve the state of it throughout the data from the point of view of the tester or QA (quality assurance).
Environment, and Constraints	The data which is irrelevant and not consistent should not be used by making it distinguishable from the authentic ones. The data in a dataset should not give ambiguous results when used. **Factors and Parameters** **Application Factors**: Eliminate outliers from the dataset. **People Factors**: Data Scientists, tester, QA. **Resource Factors**: A technical person having expertise in data analysis and data cleaning process. **Process Factors**: Turnaround time

	Customer Factors: Authentic Information Methods: Using various statistical methods. Tools: Data cleaning tools Constraints: Distinguishing correct data from the incorrect ones.

Operationalized Goal label and description	OG4 To improve the categorization of the data from different types of data/datasets available.
Corresponding Measurement Goal label	MG4 (**Veriety**)
Object of Interest	Big Data
Purpose	To categorize a large size dataset consisting of various types of data in order to predict the system/application usage by using these data in machine learning algorithms.
Quality Focus, Perspective	Examine the behavior of different data in a dataset from the point of view of the data scientist.
Environment, and Constraints	The categorization and segmentation of the dataset should be relatable and should comply with the requirements. With a lot more categorization, consistency in data should be observed. **Factors and Parameters** **Application Factors**: Ability to process a large size dataset with a variety of big data. **People Factors**: Data Scientists and technical teams for analysis purposes. **Resource Factors**: A technical Person supervising the classification process **Process Factors**: The semantics of each type of data **Customer Factors**: Different types of data **Methods**: Algorithms used in classifying the Data **Tools**: Data Classification tools **Constraints**: Need a large number of consistent data

Operationalized Goal label and description	OG5 To improve the quality of the data used for different purposes in the application to be authorized and should be of high level.
Corresponding Measurement Goal label	MG5 (**Validity**)
Object of Interest	Big Data
Purpose	To evaluate the correctness of the data used in decision making in order to assess its quality and usage in the application.
Quality Focus, Perspective	Examine the quality from the point of view of the technical team consisting of a data scientist, manager, developer, and testing/QA.
Environment, and Constraints	The dataset selected to be used for machine learning algorithms must be as per the required specifications. The correctness of the data must be examined with the great expertise of data scientists and project managers. **Factors and Parameters** **Application Factors**: Different types of data in a dataset. **People Factors**: Technical Team **Resource Factors**: Technical team's expertise who are supervising the data validation process. **Process Factors**: Improved data quality standards. **Customer Factors**: Data correctness **Methods**: Different algorithms for checking data truthfulness. **Tools**: Data Validation tools. **Constraints**: A step of calculating the degree of correctness must be included.

Operationalized Goal label and description	OG6 Refers to improving the connectivity, traceability, and linkage between big data.
Corresponding Measurement Goal label	MG6 (**Vincularity**)

Object of Interest	Big Data
Purpose	To improve the throughput of the machine learning algorithms the linkages between the data and different datasets and their selection must be precise and related to the requirements.
Quality Focus, Perspective	Modifiability of the dataset (altering unnecessary data to tune the dataset) from the point of view of process improvement.
Environment and Constraints	To establish a relationship between data, they must be comparable and relatable. **Factors and Parameters** **Application Factors:** Link different relatable datasets to create a linkage. **People Factors:** Data Scientists, Sales, and Marketing team. **Resource Factors:** Existing data in a dataset. **Process Factors:** Semantic metadata management **Customer Factors:** Data Integration **Methods:** the degree of relationship between two datasets **Tools**: Data integration tools **Constraints:** Quality assessments and validation of data in datasets.