

SOEN 6611
SOFTWARE MEASUREMENT
INSTRUCTOR: DR. OLGA ORMANDJIEVA

PROJECT
TASK 1 AND TASK 2

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

SUBMITTED ON : 21-October-2022

Declaration:

We, the members of the team, have read and understood the Fairness Protocol and the Communal Work Protocol, and agree to abide by the policies therein, without any exception, under any circumstances, whatsoever.

TEAM-7

SIDDHARTHA NANDA	40200496
BARIQ ISHTIAQ MOHAMMED	40208194
RAJAT KUMAR	40201807
VIKYATH SRINIVISAULU	40218245

TEAM LEADER

SIDDHARTHA NANDA	40200496
------------------	----------

Email ID:	siddhartha.nanda@mail.concordia.ca
-----------	------------------------------------

CONTENTS

1. <u>Step 1</u>	1
1.1. <u>Measurement-Related Organizational Goals</u>	1
1.2. <u>Measurement stakeholders for the organisation</u>	1
1.3. <u>Measurement Goals</u>	2
1.4. <u>Questions</u>	3
2. <u>Step 2</u>	5
2.1. <u>Operationalized Goals</u>	5
2.1.1. <u>Operationalized Goal 1: VOLUME</u>	5
2.1.2. <u>Operationalized Goal 1: VELOCITY</u>	6
2.1.3. <u>Operationalized Goal 1: VARIETY</u>	7

1. Step 1:

1.1 Measurement-Related Organizational Goals:

a. Business Goals:

A business objective is a destination, accomplishment, or target that an organisation wishes to reach in the short or long term. Business objectives may take numerous shapes and can be aspirational or motivating, such as driving an organisation toward a specific goal, such as enhanced client experience.

To choose a big data dataset of sufficient quality to boost the organization's decision-making process. The big data set's quality may be attained by obtaining individual quality at each of the six big data V's, as explained in the following table:

Label: business sub-goal	Goal synopsis	Explanation
Sub-goal 1	Increasing the Volume of big data sets	Choose a volume of the dataset that is easily accessible and can be utilised for a Machine Learning algorithm.
Sub-goal 2	Accelerate the Big Data set Velocity	The availability of fresh and recently developed big data datasets, as well as the regularity with which new datasets are generated.
Sub-goal 3	Enhancing Variety in Big Data	To make efficient decisions, keep track of the vast amount of data that accumulates over time.

1.2 Measurement stakeholders for the organisation:

1.2.1.1 Stakeholders:

A stakeholder is a person, group, or organisation that is affected by the result of a project or commercial initiative. Stakeholders have an interest in the project's success and might be from within or outside the organisation that is supporting it. Stakeholders are crucial because their decisions can have a positive or negative impact on the project. There are also crucial or key stakeholders who must support the project in order for it to exist.

The tables below illustrate several stakeholders with their concern, and we have divided the stakeholders into many groups for this purpose.

Stakeholders	Roles
Product Owner/Project Manager	Individual in charge of the complete Machine Learning algorithm (software, product, code)
Developers/Data Scientist	The person responsible for developing the machine learning algorithm (training and evaluating the algorithm).
Testers	Person responsible for validating functionality and ensuring the product satisfies the goals.
Sales and Marketing Team	Individuals who oversee determining how to sell a product and marketing it to entice consumers.
End Users	People who utilise the algorithm or the product after it has been released.

1.2.1.2 Measurement needs of stakeholders:

Stakeholders	Measurement Needs
Product Owner/Project Manager	Product Owners are in charge of ensuring that their companies' data projects deliver a value-adding outcome. Resource Allocation: The number of employees required to finish the project; hardware and other tool needs
Developers/Data Scientist	Plan the individual work effort. Choosing the Best Programming Language. Prepare training, retraining, and valuation strategy.
Testers	To locate any mistakes or faults in the machine learning code. To evaluate the algorithm's behaviour on different sized testing sets.
Sales and Marketing Team	Their task is to create marketing campaigns as well as promotional plans for the product.
End Users	To have a better understanding of how to use the software or application

1.3 Measurement Goals

Measurement goal label	Description	Corresponding business goal
MG1	The volume of data refers to the quantity of data available for processing the dataset. To choose a dataset with a sufficient number of records so that a conclusive analytical application can be constructed on top of it.	Volume
MG2	Regardless of the fact that the dataset is always changing, the aim is to swiftly collect it and analyse it with ease by utilising the organization's current infrastructure.	Velocity
MG3	Choosing a dataset containing diverse sorts of records (but in the same genre) assists in categorising and segregating data. Datasets may be in many forms, but for our purposes, we should distinguish between them and choose the format that allows for easy procurement while maintaining quality.	Variety

1.4 Questions:

The Goal Question Metric (GQM) approach assumes that in order for an organisation to measure purposefully, it must first specify the goals for itself and its projects, then trace those goals to the data that is intended to operationally define those goals, and finally provide a framework for interpreting the data in relation to the stated goals. Thus, it is critical to define, at least in broad strokes, what informational needs the organisation has, so that these informational needs may be quantified wherever feasible, and the quantified information can be assessed to determine whether the goals are met.

We chose two stakeholders for GQM analysis: the product owner and the developer. The questions for each stakeholder are stated below.

1.4.1.1 Product owner

The Product Owner is crucial to the Scrum process since their decisions affect the work done at each level of development. Their selections have the potential to increase or decrease time consumption as well as the overall efficacy of the product.

Question Label	Description	Corresponding measurement goal (label)
Q1	Is the volume of the big data collection large enough to support the development of a model? What is the total number of records in the data set?	MG1
Q2	Is the dataset updated frequently? How simple is it to locate a recently modified dataset?	MG2
Q3	Is there any missing information in the dataset? What is the amount of difficulty in removing that data?	MG3

1.4.1.2 Developers

Software developers use a variety of skills and technologies to design, write, build, launch, and manage software. They also contribute to the development of software systems that power networks and devices and ensure that such systems stay operational. Meeting with clients to assess the demands for a software solution may also be part of their job, which will aid in the creation of the final product.

Question Label	Description	Corresponding measurement goal (label)
Q1	Is there preprocessed data in a usable format? Is there enough data for processing?	MG1

	Is it too large to manage, and can existing programming approaches handle it?	
Q2	Does the algorithm take into account the fresh data stream? Is it in sync or not? Do we need to update our code to accommodate a new dataset?	MG2
Q3	Is the dataset well-structured? How should various types of data be classified?	MG3

2. Step 2:

2.1 Operationalized Goals

The process of converting abstract notions into measurable observations is known as operationalization. It entails specifying how to measure, monitor, or modify a topic. Researchers can use operationalization to gather and assess phenomena that cannot be observed directly. Operationalization is essential when attempting to understand phenomena that are not immediately quantifiable.

❖ Operationalized Goal 1: VOLUME

Operationalized Goal: Label and description	OG1 Choose a dataset with high-quality Volume.
Corresponding Measurement Goal label	MG1 Volume
Object of interest	Big dataset
Purpose	Analyze a large number of datasets in order to compare the findings at each stage throughout time to achieve organisational goals.
Quality Focus, Perspective	Examine the dataset size and choose the one that is appropriate for the machine learning algorithm from the project manager's perspective.
Environment and Constraints	<p>Depending on the volume of the dataset, we may require additional environmental assistance (such as hardware or the skills of data scientists or tools).</p> <p>There are several datasets sources and choosing one with the appropriate amount and quality data might be difficult.</p> <p>Factors and parameters to consider include:</p> <ul style="list-style-type: none"> ➤ Application factors: The software's capability to process data. ➤ Customer Factors: To improve the consumer experience. ➤ Resource factors: Various datasets are available. ➤ People factors: Data scientists and developers analyse the data.

	<ul style="list-style-type: none"> ➤ Process factors: The process of determining quality. ➤ Methods: Comparing the size of the Big datasets. ➤ Tools: The tool for calculating size. ➤ Constraints: Availability of datasets in various volumes.
--	--

❖ Operationalized Goal 2: VELOCITY

Operationalized Goal: Label and description	OG2 Improve the decision-making process by obtaining data more frequently and processing it more quickly.
Corresponding Measurement Goal label	MG2 Velocity
Object of interest	Big dataset
Purpose	Analyze Bigdata flow to enhance the rate at which newer insights are gathered over time
Quality Focus, Perspective	Examine the dataset's expansion and modifications from the viewpoint of the developer.
Environment and Constraints	The newly collected data must be regularly checked. The new information must be consistent with the old information. Factors and parameters to consider include: <ul style="list-style-type: none"> ➤ Application factors: The software used to process the information should be adaptable in order to handle rapidly changing data. ➤ Customer Factors: Quality and effectiveness. ➤ Resource factors: People's willingness to accept change. ➤ People factors: Data scientists and developers analyse the data. ➤ Process factors: It's time to start collecting data. ➤ Methods: The project management process's agility. ➤ Tools: Data processing software. ➤ Constraints: Managing various dataset versions for rollback capabilities.

❖ Operationalized Goal 3: VARIETY

Operationalized Goal: Label and description	OG3 The goal of Variety is to identify and segregate data by classification and segmentation.
Corresponding Measurement Goal label	MG3 Variety
Object of interest	Dataset
Purpose	Choosing a dataset with different types (but the same genre) of records aids in categorising and segregating data.
Quality Focus, Perspective	Examine the dataset's behaviour and choose the separate compatible data from the perspective of a data scientist.
Environment and Constraints	We must use caution while standardising and disseminating the dataset. Factors and parameters to consider include: <ul style="list-style-type: none"> ➤ Application factors: Capability to process many types of datasets. ➤ Customer Factors: Data variety. ➤ Resource factors: The person in charge of the categorization tool. ➤ People factors: Data scientists and developers analyse the data. ➤ Process factors: The semantics of each data type. ➤ Methods: Methods for categorising data. ➤ Tools: Data classification software. ➤ Constraints: We require sufficient and consistent data.

