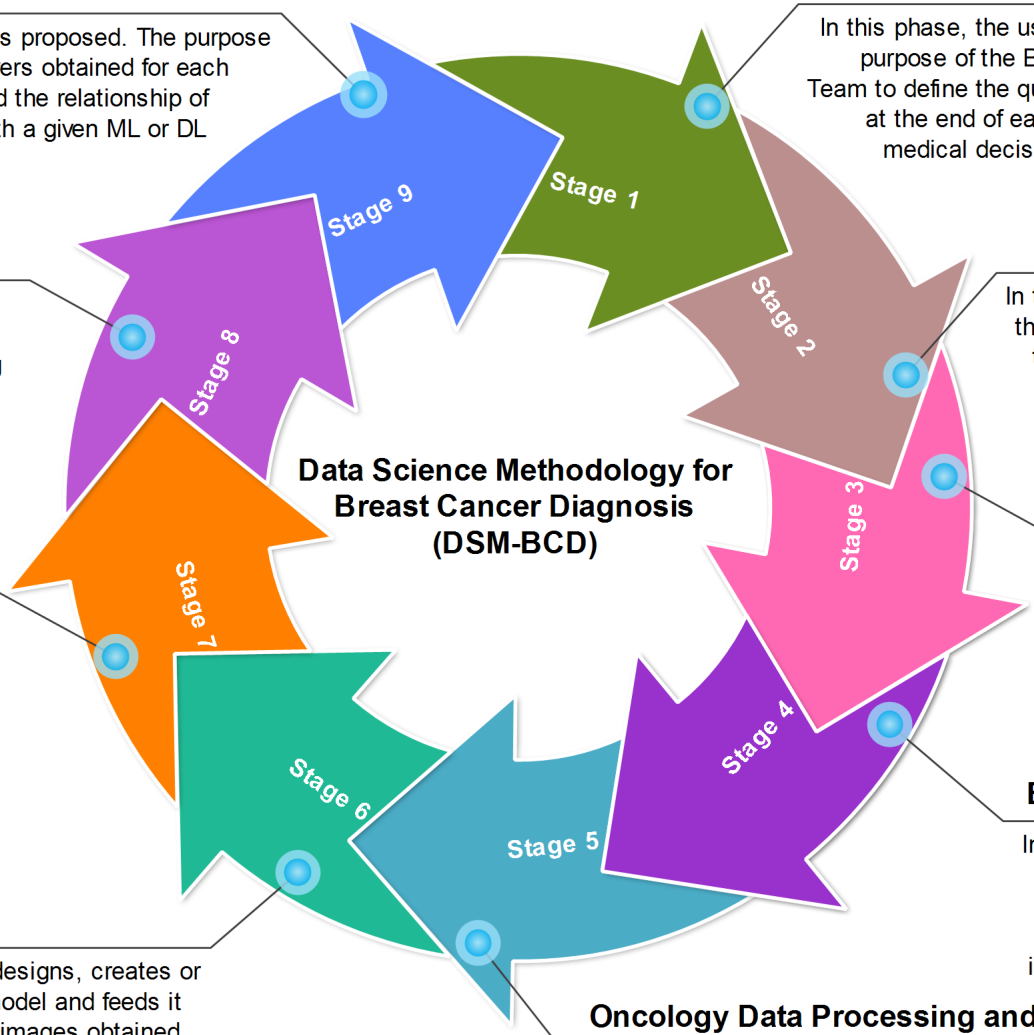

Breast Cancer Diagnostic Logbook (BCDL)

In this phase, the use of *BCDL* is proposed. The purpose of the BCDL is to store the answers obtained for each question posed in the BCQM and the relationship of these questions and answers with a given ML or DL model.

Breast Cancer Question Map (BCQM)

In this phase, the use of *BCQM* is proposed. The purpose of the BCQM is for the Data Analysis Team to define the questions that will be answered at the end of each Release and that will allow medical decisions to be made regarding the diagnosis of this disease.

Medical feedback

In this phase, the medical oncology expert determines whether the results generated by the ML or DL model succeeded in answering the questions posed in the BCQM and whether the new information obtained is sufficient to diagnose breast cancer.

Activity planning

In this phase the Data Analysis Team, based on the questions asked in the BCQM, analyses all the tasks to be carried out, estimates them in time and distributes them among the people who will carry them out during the Release

Evaluation and Interpretation

In this phase, the Data Analysis Team evaluates the model to understand its quality and ensure that it addresses the questions generated in the BCQM adequately and completely. It is necessary that specialized measures based on model performance, sensitivity and specificity are used to perform the evaluation.

Oncology data acquisition

In this phase, based on the tasks performed in the activity planning, the medical oncology expert together with the engineer and the data scientist identify and gather the available data resources (structured, unstructured and semi-structured) and relevant to solve the questions posed in the BCQM.

Exploratory Oncology Data Analysis

In this phase, the scientist obtains the set of data or images that were previously organized by the engineer and performs an exploratory data analysis to discover general patterns in the information generated. The variables and patterns are verified by the oncology expert physician.

Modeling and Execution

In this phase, the data scientist designs, creates or uses a predictive or descriptive model and feeds it with the version of the dataset or images obtained. scientist must select the type of learning and the determined technique according to the questions posed in the BCQM.

Oncology Data Processing and Transformation

This phase encompasses all activities to build the dataset or images that will be used in the next stage of modeling and execution. Among the activities of oncology data processing and transformation are data cleaning, combining data from multiple sources, and transforming the data into variables of value.