FZI & KIT Karlsruhe

Process Report for

Report for Process: Reference_Architecture_CDA

Prepared for the Community

By

Martin Trat & Matthes Elstermann

September 13, 2024

Table of Contents:

Subject Interaction Diagram: 3

Subject Short Description 4

Messages 5

Subjects and Their Internal Behavior 7

Concept Drift Detection Unit 7

Concept Drift Adaptation Unit 10

Inference Unit 13

Evaluation Unit 16

Model (Re-) Training Unit 19

Concept Drift Understanding Unit 22

Interface Subjects & Groups 25

Maintenance Process System 25

CDA Core 25

Inference Process System 25

Streaming Data Source 26

Consuming Entity 26

Data Scientist 26

CDA System Data Buffer 27

Delayed Label Source 27

Subject Interaction Diagram:

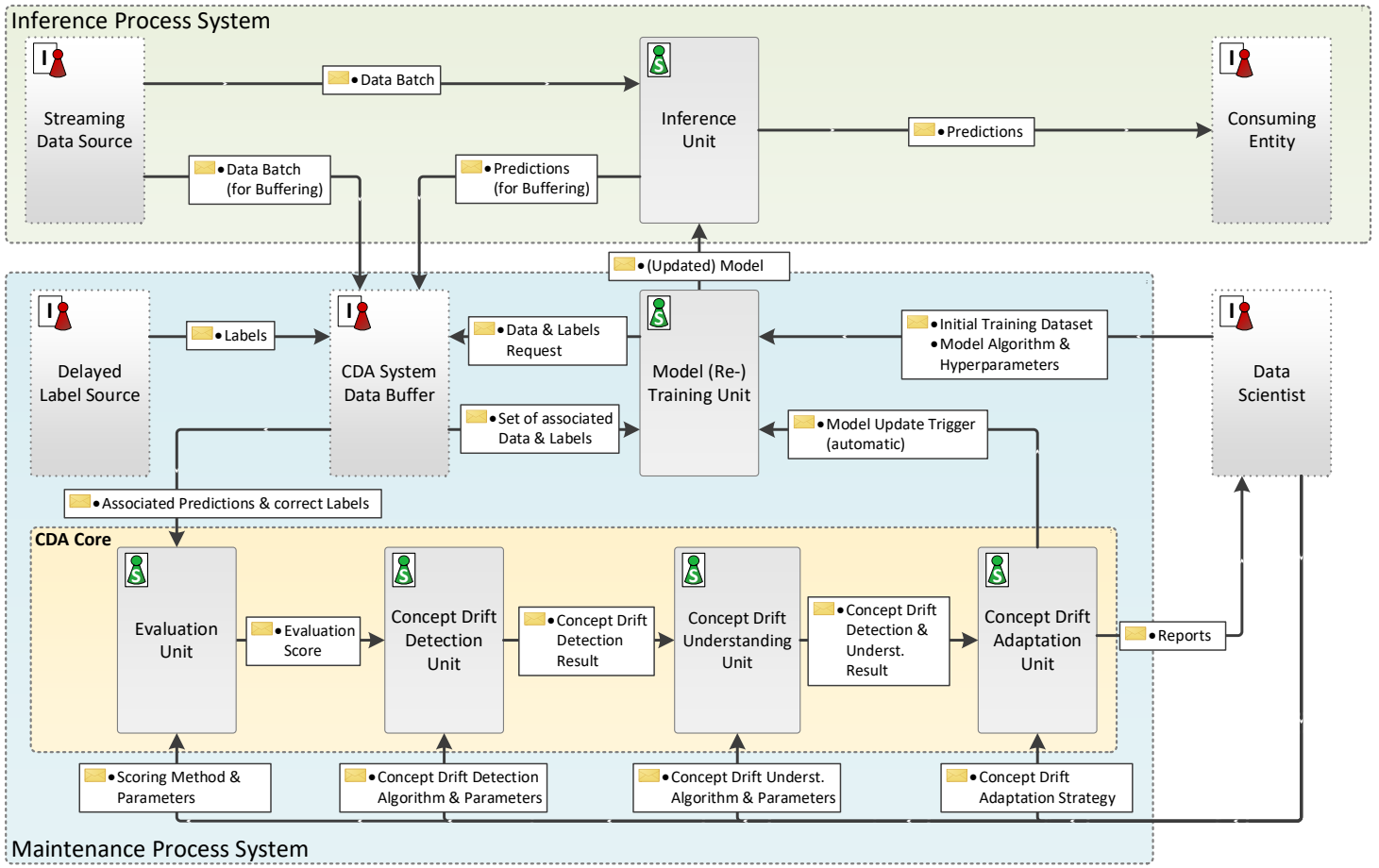

Figure 1: Subject-Oriented Referential Process Model of Report for Process: Reference_Architecture_CDA

Subject Short Description

The areas of activity described in the model represent the principal aspects necessary for the holistic execution of strategic product planning.

There are **14 subjects** being part of :

Maintenance Process System

CDA Core

Inference Process System

Streaming Data Source

Comment: One or a set of sources streaming data within a real-world scenario.

Concept Drift Detection Unit

Comment: It is tasked with evaluating concept drift evidence and alerting in case of detections.

Concept Drift Adaptation Unit

Comment: It receives alerts from the Concept Drift Detection and optionally from the Understanding Unit and orchestrates associated responses sustaining the robustness of employed inference models.

Inference Unit

Comment: The center of any AI-based system executing inference processes. It pools compute resources and routes input data from their data stream sources through preprocessing pipelines to the inference model.

Consuming Entity

Comment: The System or human that is regularly provided with the inference predictions.

Evaluation Unit

Comment: It is provided with associated pairs of model predictions and labels to compute performance scores.

Model (Re-) Training Unit

Comment: It can request a range of labeled data samples to update estimation models when triggered. It pools compute resources for such updates and supplies the Inference Unit with the resulting model updates, i.e. sequential versions.

Data Scientist

Comment: She is is involved to parametrize various other system units and can receive reports on CD understanding and adaptation measures, e.g. containing concept drift timing and treatment details.

CDA System Data Buffer

Comment: Is a vital element of CDA systems in practice as it takes into consideration the delay of ground truth often occurring in real-world scenarios, which are provided by a Delayed Label Source.

The buffer's main purpose is to store original predictions and match them with the associated input data samples and ground truth.

Delayed Label Source

Comment: The source of correct labels for incoming data batches (post factum). Provision occurs after the predictions were made. It needs to be configured before system start.

Concept Drift Understanding Unit

Comment: It analyzes concept drift evidence e.g. regarding the type, location or magnitude of concept drifts.

Messages

Depending on the organization size and chosen level of formalism, the possible forms of messages are broad. They may be elaborated formal reports spanning many pages of written text, conveyed or stored in an organization's internal information management systems. Equally possible is a scenario where the same information, if at all, is formulated and conveyed purely informal on a personal level in form of verbal instructions and decisions or in meetings.

There are **20 messages** being sent in :

Evaluation Score (SID_1_MessageSpecification_482)

Comment: Batch size >= 1

Data Batch (SID_1_MessageSpecification_489)

Comment: Batch size >= 1

Concept Drift Detection Result (SID_1_MessageSpecification_496)

Predictions (SID_1_MessageSpecification_622)

Predictions (for Buffering) (SID_1_MessageSpecification_676)

Model Update Trigger (automatic) (SID_1_MessageSpecification_789)

Model Algorithm & Hyperparameters (SID_1_MessageSpecification_887)

Initial Training Dataset (SID_1_MessageSpecification_895)

Concept Drift Adaptation Strategy (SID_1_MessageSpecification_901)

Comment: defines what will be done upon drift detection, (e.g. trigger model update, inform responsible persons) etc.

Data Batch (for Buffering) (SID_1_MessageSpecification_952)

Labels (SID_1_MessageSpecification_1092)

Comment: the correct ground truth/static descriptions - the correct descriptions for the buffered data batches describing what the predictions should have been!

Associated Predictions & correct Labels (SID_1_MessageSpecification_1098)

Comment: for comparison what was predicted and what it should have been

Data & Labels Request (SID_1_MessageSpecification_1155)

Set of associated Data & Labels (SID_1_MessageSpecification_1263)

Scoring Method & Parameters (SID_1_MessageSpecification_1292)

Concept Drift Detection Algorithm & Parameters (SID_1_MessageSpecification_1297)

Reports (SID_1_MessageSpecification_1302)

(Updated) Model (SID_1_MessageSpecification_564)

Comment: including Training Strategy

Concept Drift Underst. Algorithm & Parameters (SID_1_MessageSpecification_1356)

Concept Drift Detection & Underst. Result (SID_1_MessageSpecification_1362)

Subjects and Their Internal Behavior

In this section, each subject from the subject interaction diagram is individually presented.

Concept Drift Detection Unit

Figure 2: Concept Drift Detection Unit

Subject behavior:

This subject is connected with a Subject Behavior Diagram. The SBD is displayed on a following page.

Comment:

It is tasked with evaluating concept drift evidence and alerting in case of detections.

Functions States:

Execute Concept Drift Detection Algorithm

Comment: The acutal work is done here

Preprocess Evaluation

Initialize/Update Detection Algorithm

Receiving States:

Await Data Evaluation Input

Sending States:

Send Concept Drift Detection Result

Outgoing messages:

"Concept Drift Detection Result" **To** Concept Drift Understanding Unit

Receiving messages:

"Evaluation Score" **From** Evaluation Unit

"Concept Drift Detection Algorithm & Parameters" **From** Data Scientist

Complete Subject Behavior Diagram SBD_4_SID_1_FullySpecifiedSubject_47

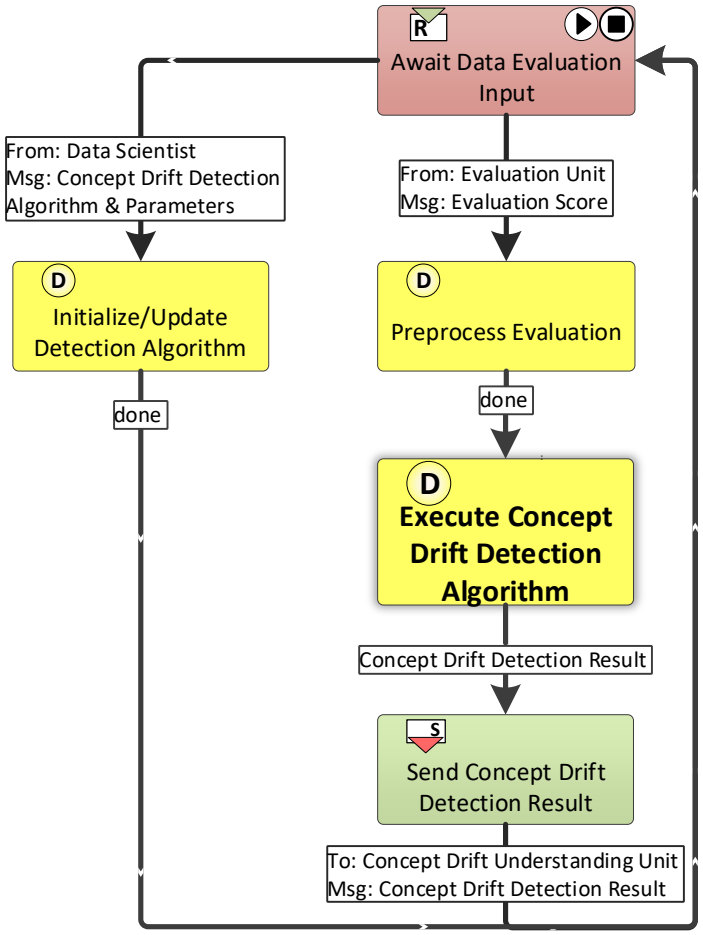

Figure 3: Subject Behavior for Concept Drift Detection Unit

Concept Drift Adaptation Unit

Figure 4: Concept Drift Adaptation Unit

Subject behavior:

This subject is connected with a Subject Behavior Diagram. The SBD is displayed on a following page.

Comment:

It receives alerts from the Concept Drift Detection and optionally from the Understanding Unit and orchestrates associated responses sustaining the robustness of employed inference models.

Functions States:

Set Concept Drift Adaptation Strategy

Evaluate Concept Drift Detection & Understanding Result

Receiving States:

Await Input

Sending States:

Trigger Retraining of Classifier Unit Model

Inform Data Scientist

Outgoing messages:

"Model Update Trigger (automatic)" **To** Model (Re-) Training Unit

"Reports" **To** Data Scientist

Receiving messages:

"Concept Drift Adaptation Strategy" **From** Data Scientist

"Concept Drift Detection & Underst. Result" **From** Concept Drift Understanding Unit

Complete Subject Behavior Diagram SBD_6_SID_1_FullySpecifiedSubject_102

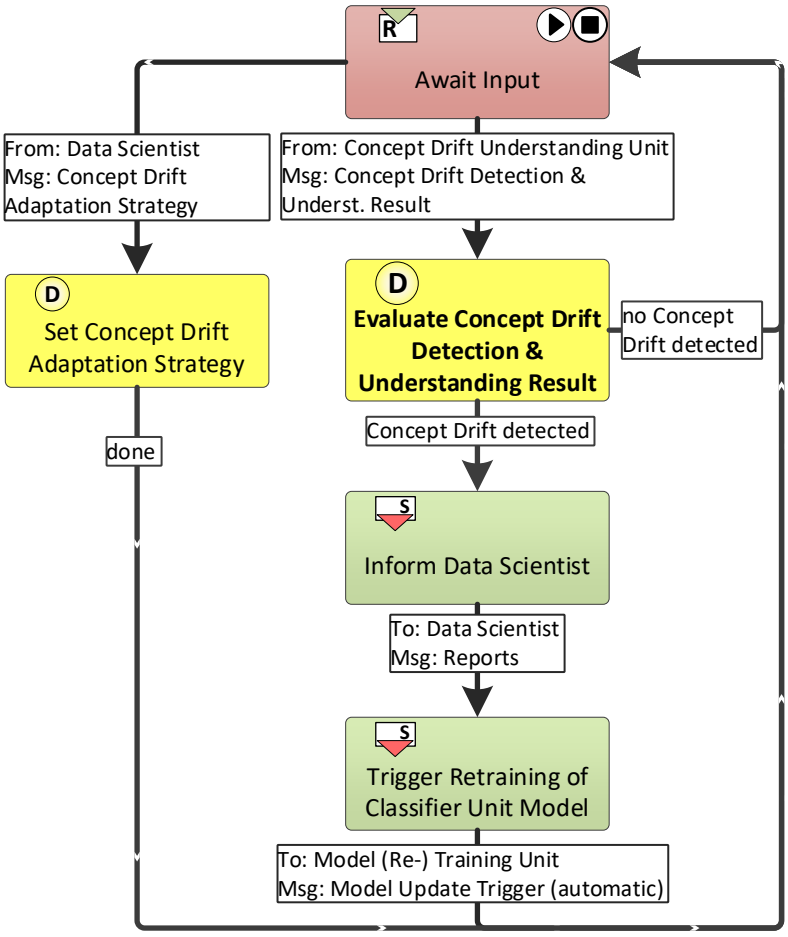

Figure 5: Subject Behavior for Concept Drift Adaptation Unit

Inference Unit

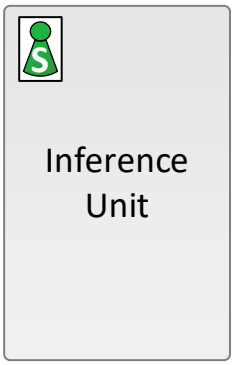

Figure 6: Inference Unit

Subject behavior:

This subject is connected with a Subject Behavior Diagram. The SBD is displayed on a following page.

Comment:

The center of any AI-based system executing inference processes. It pools compute resources and routes input data from their data stream sources through preprocessing pipelines to the inference model.

Functions States:

Preprocess Data Batch

Replace Estimation Model

Conduct Inference

Receiving States:

Await Input

Sending States:

Publish Inference Result to Consumer

Forward Result to CDA System

Outgoing messages:

"Predictions" **To** Consuming Entity

"Predictions (for Buffering)" **To** CDA System Data Buffer

Receiving messages:

"Data Batch" **From** Streaming Data Source

"(Updated) Model" **From** Model (Re-) Training Unit

Complete Subject Behavior Diagram SBD_9_SID_1_FullySpecifiedSubject_295

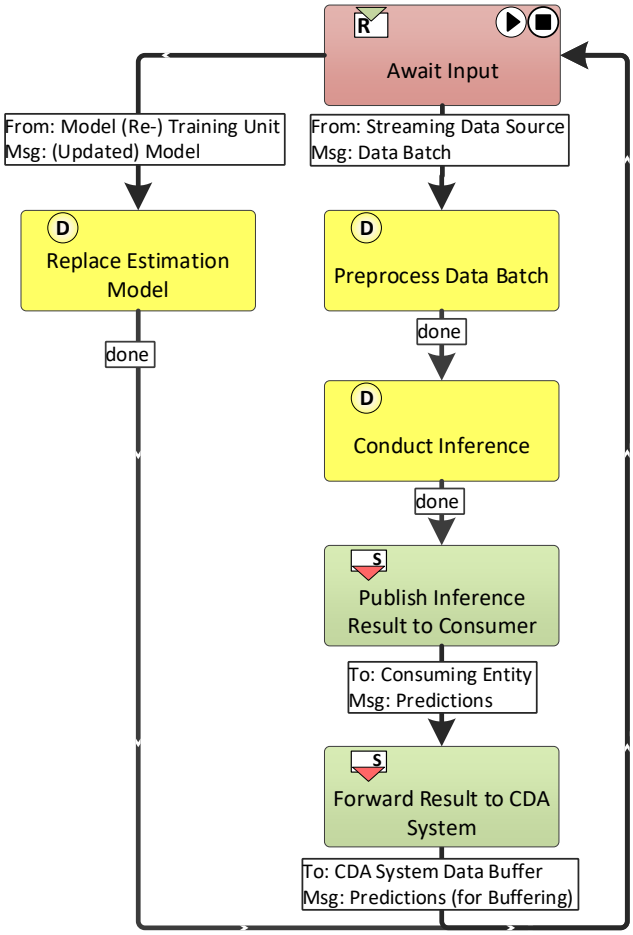

Figure 7: Subject Behavior for Inference Unit

Evaluation Unit

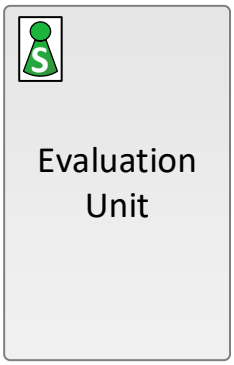

Figure 8: Evaluation Unit

Subject behavior:

This subject is connected with a Subject Behavior Diagram. The SBD is displayed on a following page.

Comment:

It is provided with associated pairs of model predictions and labels to compute performance scores.

Functions States:

Compute Evaluation Score

Comment: The score is also very important for CDA details can be found:

Initialize/Update Scoring Method

Receiving States:

Await Input

Sending States:

Send Evaluation Score

Outgoing messages:

"Evaluation Score" **To** Concept Drift Detection Unit

Receiving messages:

"Associated Predictions & correct Labels" **From** CDA System Data Buffer

"Scoring Method & Parameters" **From** Data Scientist

Complete Subject Behavior Diagram SBD_15_SID_1_FullySpecifiedSubject_625

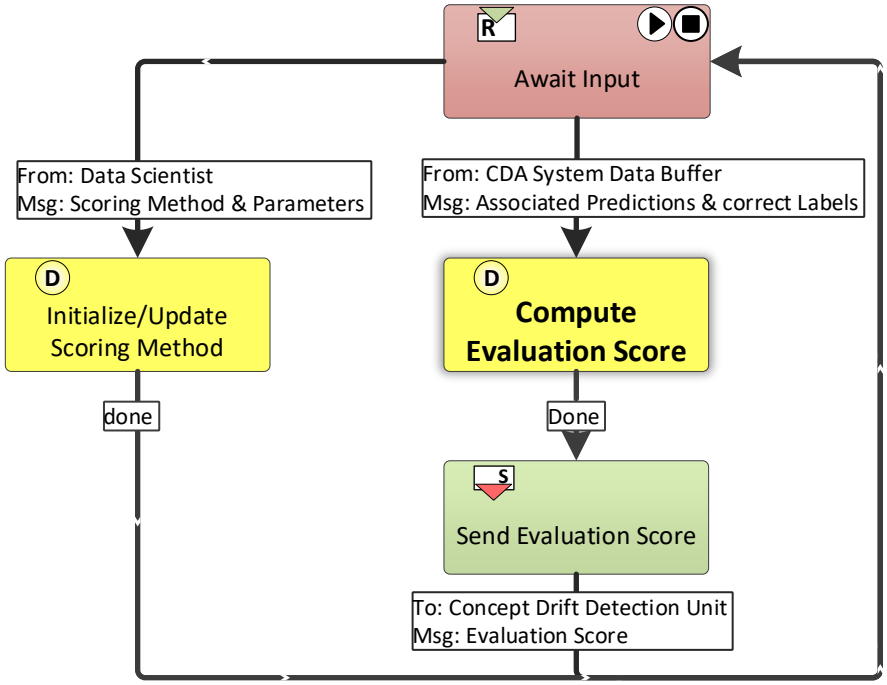

Figure 9: Subject Behavior for Evaluation Unit

Model (Re-) Training Unit

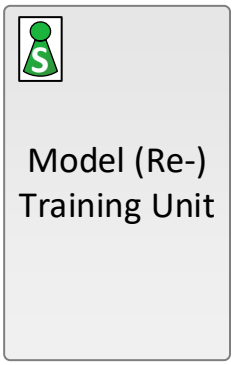

Figure 10: Model (Re-) Training Unit

Subject behavior:

This subject is connected with a Subject Behavior Diagram. The SBD is displayed on a following page.

Comment:

It can request a range of labeled data samples to update estimation models when triggered. It pools compute resources for such updates and supplies the Inference Unit with the resulting model updates, i.e. sequential versions.

Functions States:

Set Algorithm and Hyperparameters

Conduct Model Training

Receiving States:

Initially Set-Up Training Unit

Get Parameters for Model

Await Input/Updates

Receive (new) Training Data

Sending States:

Request Data & Labels for Incremental Training

Publish new Model Version

Outgoing messages:

"Data & Labels Request" **To** CDA System Data Buffer

"(Updated) Model" **To** Inference Unit

Receiving messages:

"Model Update Trigger (automatic)" **From** Concept Drift Adaptation Unit

"Model Algorithm & Hyperparameters" **From** Data Scientist

"Initial Training Dataset" **From** Data Scientist

"Set of associated Data & Labels" **From** CDA System Data Buffer

Complete Subject Behavior Diagram SBD_19_SID_1_FullySpecifiedSubject_738

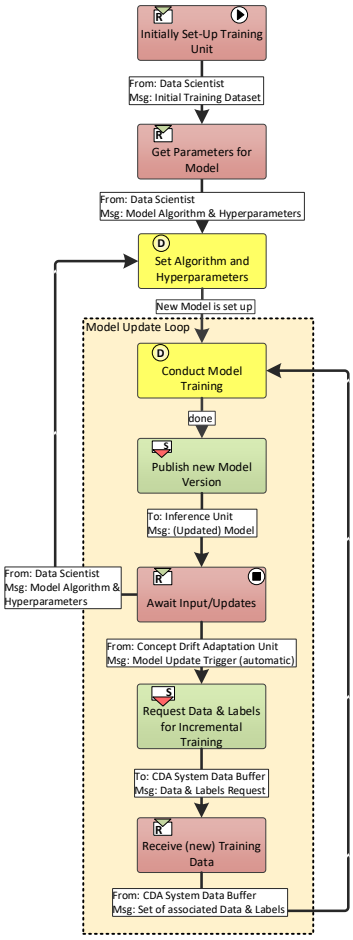

Figure 11: Subject Behavior for Model (Re-) Training Unit

Concept Drift Understanding Unit

Figure 12: Concept Drift Understanding Unit

Subject behavior:

This subject is connected with a Subject Behavior Diagram. The SBD is displayed on a following page.

Comment:

It analyzes concept drift evidence e.g. regarding the type, location or magnitude of concept drifts.

Functions States:

Preprocess Input

Execute Concept Drift Understanding Algorithm

Initialize/Update Understaning Algorithm

Receiving States:

Await Input

Sending States:

Send Concept Drift Detection & Understanding Result

Outgoing messages:

"Concept Drift Detection & Underst. Result" **To** Concept Drift Adaptation Unit

Receiving messages:

"Concept Drift Detection Result" **From** Concept Drift Detection Unit

"Concept Drift Underst. Algorithm & Parameters" **From** Data Scientist

Complete Subject Behavior Diagram SBD_22_SID_1_FullySpecifiedSubject_1305

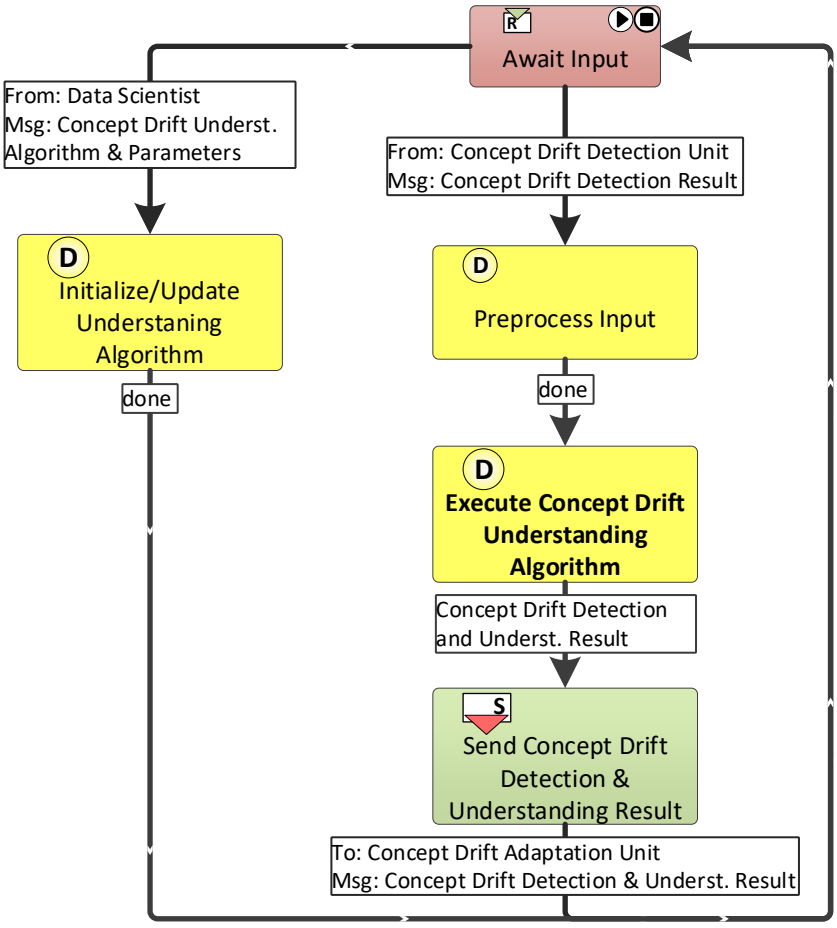

Figure 13: Subject Behavior for Concept Drift Understanding Unit

Interface Subjects & Groups

Reminder: interface subjects are actors without a defined behavior within the bounds of a given model. Defining no behavior diagram may have various reasons: the behavior may be defined in another process model, the subject maybe a technical system only reacting towards request but without a complex process flow, or simply because it is not important or impossible to describe how message are created within the model.

Maintenance Process System

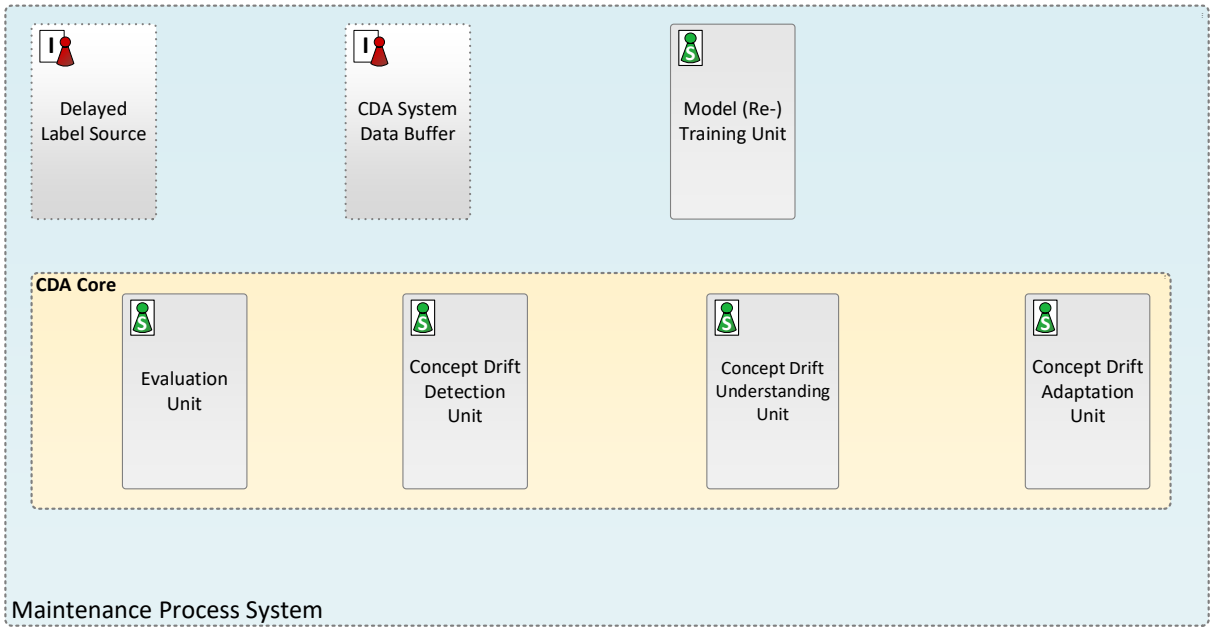

Figure 14: SID_0_SubjectGroup_1271

CDA Core

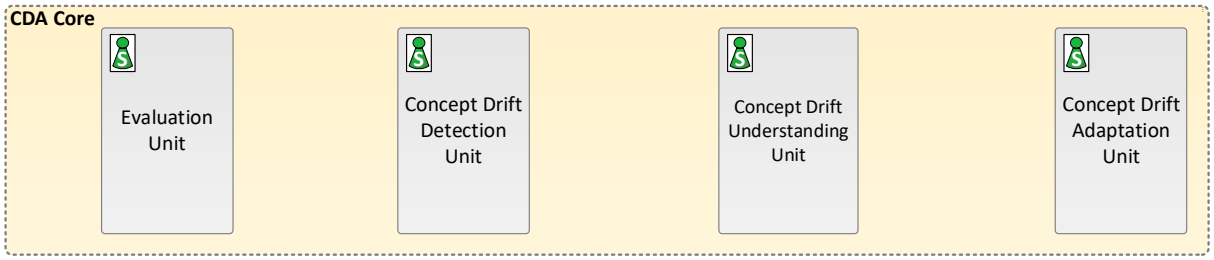

Figure 15: SID_0_SubjectGroup_1279

Inference Process System

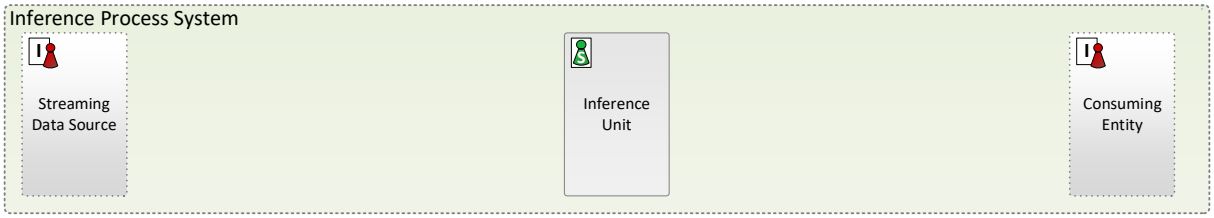

Figure 16: SID_0_SubjectGroup_1266

Streaming Data Source

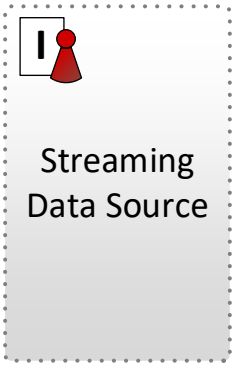

Figure 17: SID_1_InterfaceSubject_2

Comment: One or a set of sources streaming data within a real-world scenario.

Consuming Entity

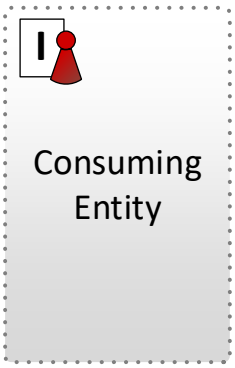

Figure 18: SID_1_InterfaceSubject_573

Comment: The System or human that is regularly provided with the inference predictions.

Data Scientist

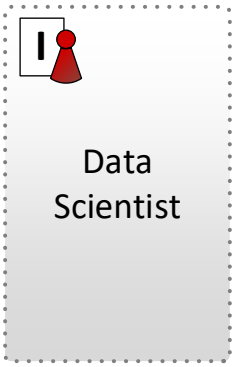

Figure 19: SID_1_InterfaceSubject_838

Comment: She is is involved to parametrize various other system units and can receive reports on CD understanding and adaptation measures, e.g. containing concept drift timing and treatment details.

CDA System Data Buffer

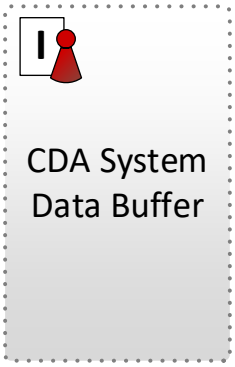

Figure 20: SID_1_InterfaceSubject_903

Comment: Is a vital element of CDA systems in practice as it takes into consideration the delay of ground truth often occurring in real-world scenarios, which are provided by a Delayed Label Source. The buffer's main purpose is to store original predictions and match them with the associated input data samples and ground truth.

Delayed Label Source

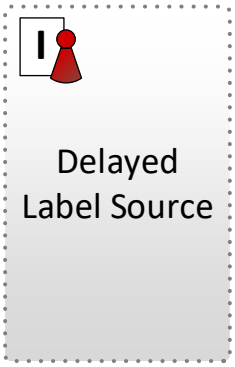

Figure 21: SID_1_InterfaceSubject_1043

Comment: The source of correct labels for incoming data batches (post factum). Provision occurs after the predictions were made. It needs to be configured before system start.