Data Deduplication and Profiling

APP 243 – Course Outline

Version 9.1

AUTHOR: Stibo Systems Academy
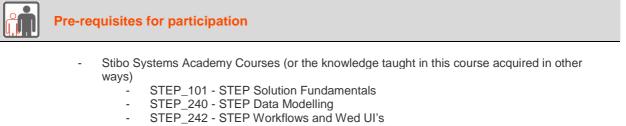
CONFIDENTIALITY LEVEL:

Public

Target group

This course is targeted towards application consultants or other roles responsible for transforming the requirements of data quality maintenance into a functional configuration.

Pre-requisites for participation

- Stibo Systems Academy Courses (or the knowledge taught in this course acquired in other ways)
    - STEP_101 - STEP Solution Fundamentals
    - STEP_240 - STEP Data Modelling
    - STEP_242 - STEP Workflows and Wed UI's

- Some knowledge of XML and JavaScript will be helpful.

- At least one year of application consultancy experience and/or system implementation.

Course duration

The course duration is 3.5 days.

Training form

☒ Instructor led class room training.
☐ E-learning
☐ Online training
☐ Self-study

The course activities will be will be a combination of lectures, case studies in teams, problem solving exercises and presentations by the trainees.

Course purpose

In this course, you will gain the knowledge and skills to configure the STEP solution to enable a customer to effectively monitor product and customer data, troubleshoot issues with data quality, and configure STEP components that will locate and manage duplicate objects.

Course objectives

Upon completion of this course, you will be able to:

1. Identify which STEP data quality components would best address business requirements for product and customer master data.

2. Build a data quality solution that meets the specified business requirements.

3. Maintain implemented data quality components over time as business requirements evolve.

4. Demonstrate how configured data quality components meet business requirements.

Detailed learning objectives

The objectives below describe precisely what is taught during the training: (please note that the learning objectives can belong to more than one course objective).

Course objective 1 – Identify which STEP data quality components would best address business requirements for product and customer master data.

Learning objectives:

- Utilize available dashboard widgets to meet specified requirements.

- Explain the differences of the available Match Criterion (including Levenshtein, Damerau Levenshtein, Javascript, Number Distance, Decision Tables, etc).

- Differentiate between Match Codes and Matching Algorithms.

- Articulate when Look Up Tables are applied to Data Quality components.

- Summarize the differences between 'Product' and 'Customer' data model implementations.

Course objective 2 – Build a data quality solution that meets the specified business requirements.

Learning objectives:

- Apply an event processor to a matching algorithm according to customer specified requirements.

- Apply match codes for multiple requirements and data domains.

- Formulate survivorship rules for the creation and maintenance of Golden Records.

- Implement data normalization using available STEP Quality components, including JavaScript, Lookup Tables, Regular Expressions, etc.

- Implement 'Matching and Linking' component model setup using specified elements (including object types, justification attributes, source attributes, duplicate reference types and non-duplicate reference types).

- Implement the 'Identify Duplicates' Match Action and explain how it differs from the Golden Records process.

- Utilize available Decision Table components to address required specifications.

Course objective 3 – Maintain implemented data quality components over time as business requirements evolve.

Learning objectives:

- Adapt the configured dashboard components to new customer requirements or changes to existing requirements.

- Manage the priority of attributes and references using completeness scores.

- Operate Match Tuning setup to tune matching algorithms.

- Examine the setup and configuration options of the CMDM matching and merging starter package.

Course objective 4 – Demonstrate how configured data quality components meet business requirements

Learning objectives:

- Comprehend dashboard options as a business user (including the generation of bulk updates, searches, and collections) and filtering those results.

- Operate configured data quality components in the Web UI.

- Use the Import Manager Profile to manage data prior to importing.

About Stibo Systems

Stibo Systems provides global organisations with a leading multi-domain Master Data Management (MDM) solution. Stibo Systems enables its customers to better manage enterprise intelligence on a global scale, improve sales, and quickly adjust to changes in business requirements. Stibo Systems' STEP technology is a flexible MDM solution that provides a single trusted source of operational information for the entire enterprise. Stibo Systems offers industry-specific solutions, engineered and supported to meet the strategic information needs of global customers including: GE, Sears, Siemens, Target and Thule. Stibo Systems is a subsidiary of the privately held Stibo A/S group, originally founded in 1794 with corporate headquarters in Aarhus, Denmark.

For more information, visit www.stibosystems.com

For more information:

Stibo Systems Academy
training@stibosystems.com