

Mapping Programs Written by MDR

Jörn Bilow, Entimo AG, Berlin, Germany

Abstract

As Metadata Repositories become a widely used tool supporting standards implementation in the pharmaceutical industry, they evolve from being pure tools for curation, governance and provision of industry standards and metadata. Making metadata operational is the next step forward. Metadata driven data transformation from one data model to another, e.g. CDASH to SDTM, is often a challenging task. Mapping specifications to guide either in-house or external subject matter experts through the process can be provided by advanced MDRs in multiple formats. Any missing link between data models can be detected and reported in a comprehensive way. To further assist with data transformation, the latest release of entimICE MDR can even generate validated and readily executable SAS programs that will run data transformation either within or outside the entimICE environment.

1 Introduction

Cooperation between pharmaceutical organizations and outsourcing partners, mainly CROs, is relying heavily on a standardized and reliable knowledge transfer. Specifications for services executed by CROs should be easy to generate and to understand. entimICE MDR provides numerous tools to help with these tasks. Structure metadata and algorithms stored in the MDR can be used to generate specifications for data transformation tasks, e.g. raw data to SDTM, or SDTM to ADaM with just a few mouse clicks. entimICE MDR can even be used to automatically generate validated SAS® programs in an automated way from information stored in the repository. These programs can be executed within or outside of the entimICE Statistical Computing Environment. On delivery of the generated output, the system provides checking capabilities to ensure conformance with target data structures.

2 Mapping Specifications

Specifications for data transformation tasks are essential to ensure consistency and conformance with SDTM or ADaM sponsor structures. entimICE MDR can generate specification documents in two widely used data formats from information stored in the repository. Reusable global standards can be adapted to specific projects and studies, benefitting from the entimICE inheritance concept.

2.1 Base technology

The entimICE MDR provides a model-independent standards and metadata repository, handling CDISC data standards, codelists, computational algorithms, value level metadata as well as customer-specific data models throughout all stages of the drug development process. Every configured model has an extendible sets of attributes, configurable lifecycles and workflows along with an integrated change management system. The system supports company-specific data models on global, therapeutic, project and study levels utilizing an inheritance concept for easy standards maintenance and optimized performance.

2.2 Excel or define-like mapping specifications

From the information stored about source and target data models, along with algorithms describing the relationship between individual attributes across the models, the system can generate specifications describing all necessary steps to take towards data transformation. The system user can choose between an Excel file output and an xml-file related to the define.xml. Both formats are well known in the pharmaceutical industry and provide an easy way of ensuring smooth and transparent communication either within an organization, or between a sponsor and CRO, guiding statistical programmers through the process. All information stored in the MDR is reusable for minimum effort during study design.

2.3 Relationship reports across data models

System users can generate a report describing relationships across various data models, for example from raw or EDC data to SDTM to ADaM. Hyperlinks guide the user through these dependencies, and color coded highlights indicate any missing links or information in the transition towards submission ready data. The report can be generated with minimum effort and guides data managers and programmers through the transformation processes.

2.4 Program generation

Data transformation algorithms can easily be transferred into the entimICE Mapping Generator. This tool, driven entirely by metadata, not only visualizes relations between source and target variables. It also enables programmers to review the data transformation in comprehensive reports, refine it with different algorithms or custom code, and then in a final step automatically generates executable SAS code. The generated programs consists of checking and transformation steps, and can be executed within the entimICE environment or outside. entimICE MDR thus provides a comprehensive and reusable way of using metadata to facilitate data transformation downstream in the drug development process, whether it is executed by the sponsor or by a CRO.

New data transformation program code can be generated easily for changes in the data models, because all information leading to the program generation is stored in a reusable and transparent way.

Contact Information

Your comments and questions are valued and encouraged. Contact the author at:

Jörn Bilow
Entimo AG | Stralauer Platz 33-34 | 10243 Berlin | Germany
Email: bil [at] entimo.de
Work Phone: +49 30 520 024 100
Web: www.entimo.com

Brand and product names are trademarks of their respective companies.