Excel-like Data Definition Editor for Real-time Generation of Define.xml and Mapping specs Management

Dylan Song Yinzhen Information Technology Co., Ltd.
John Wang dMed Biopharmaceutical Co., Ltd.

ABSTRACT

Define.xml is mandatory for study data submissions to regulatory agencies. CDISC Mapping specs management is crucial in the process of SDTM and ADaM datasets creation. Users are eager to have an Excel-like Data Definition Editor (DDE) that is specialized for Define.xml generation (Including Analysis Results Metadata). This DDE app is web-based, and an intuitive way of developing SDTM and ADaM mapping specs collaboratively with version control and generates the corresponding Define.xml in real-time, which improve efficiency and quality. Other features include the easy reusability between studies, Snapshots management, Standards customization, export of Excel and Word format of define.xml for the easy review of other stakeholders.

INTRODUCTION

CDISC Define.xml is data specification for submission study data. It consists of dataset level specs, variable level specs, value level specs, controlled terminology and other specs for SDTM and ADaM datasets. Its purpose is to facilitate the reviewer to quickly understand study data.

Define.xml is now mandated document for study data submission to FDA and PMDA. So, it has to be accurate and cover every details. It's an important task for clinical programmers.

Generating Define.xml efficiently and accurately now becomes a hot topic in the industry. People want a handy generating tool/system to increase the efficiency, lower the learning curve, and reduce the error rate.

DDE COVERS THREE MAJOR FEATURES

1. EXCEL LIKE USER INTERFACE (UI)

Many users have the experience of using excel files for mapping specs. DDE uses Excel like UI so that users don't need spent additional time to learn this system. The tabs in DDE are like Excel sheets, and the most of columns are almost same as Excel files. For some Define.xml features such as pop-up windows for selecting controlled terms that Excel file cannot support, DDE enhances the user's experience.

With such familiar UI, users can focus more on the metadata itself and increase the working efficiency.

2. LOWER THE LEARNING CURVE, AND REDUCE THE ERROR RATE IN DAILY OPERATION

Usually users like to use Excel for mapping specs development and use SAS macros or other software like P21 to read in those mapping specs for Define.xml.

SAS macros may need additional debugging time once something goes wrong as there are complex components' relationships inside different metadata tables. It's especially difficult for SAS newbies to get the components' connection right.

Lowering the learning curving and making users at different levels to efficiently generate Define.xml in simple and familiar way can reduce the cost significantly.

3. AUTOMATIC ASSISTING FUNCTION TO INCREASE EFFICIENCY
With respect of users’ habits for the flexibility of excel editing style, DDE uses additional automatic assisting function to increase the efficiency and reduce the operational redundancy.

- One click to preview Define.xml at any time, while doing metadata editing.
- Cross checking for commonly used components between different metadata tables, quickly locate to the error/issue, easy for quick fix and potentially reduce the operating error.
- Bring in the database lock concept, users can lock the metadata for Define.xml, to avoid the misoperation after all metadata ready for publish.
- Support multiple users’ collaboration at same time, easy for tasks assignments and peer review.

DESIGN PRINCIPLE

Normally one clinical study is conducted by a team of professionals with close collaboration. According to the study background and the specific requirements, this system or tool must have following features:

1. Based on Server/Client framework, to ensure the data integrity.
2. The metadata in the Define.xml, need to split into different tables, or formatted storage, in accordance to database design principles, to facilitate different format output, such as XML, Excel, or Word.
3. Easy for project management. Users can create, edit, copy, lock, and delete projects. Also bring in the snapshots function for each metadata table.
4. Support different versions of CDISC standards (Controlled Terms, SDTM and ADaM), convenient for users to manage and add custom standards, Controlled Terms, etc.
5. Simple and intuitive user interface for the operation process, which is easy for newbies to easily understand the business rules, such as the user interface for the metadata of the ADaM analysis results.
6. Support the real-time preview of the Define.xml file (one-click preview), which is convenient for users to check if the input is correct or not, and the overall progress.
7. A variety of rich assisting functions help to improve efficiency and reduce costs.
8. Support teamwork collaboration and support RESTful API to facilitate the data exchange between related systems.
9. Using the latest IT technology, stable and reliable, support in-house server deployment.

TECHNICAL ARCHITECTURE

1. ARCHITECTURAL DESIGN

The system adopts the cutting-edge technology framework, based on Web 2.0 rich client solution, and connects the multi-layer architecture seamlessly.

The user edits, submits and presents information through the presentation layer (UI), and the collected data is submitted to different business processes through the business middle layer, performs data filtering, screening, formatting, etc., and performs secondary processing on the collected information. Then extract accurate information for intelligent analysis and aggregation.

2. THE FRAMEWORK IS ILLUSTRATED AS BELOW:
3. IT SUPPORT SYSTEM REQUIREMENTS

The system makes full use of the current mainstream technology platform and application practices and adopts encrypted security communication and application isolation protection technology to ensure the system is stable and reliable.

DDE supports for mainstream server platforms such as Linux and Windows. It also supports the popular Docker implementation.

FUNCTIONAL REQUIREMENT ANALYSIS

1. VISUAL EXCEL STYLE EDITING USER INTERFACE (UI)
The project structure is simple and straightforward. Metadata is classified into different pieces according to Define.xml. When going into each metadata table, the interface looks like Excel, operates like Excel. Users can do copy, delete, edit and modify operation just like Excel. DDE has built-in intelligent analysis combined with bug correction function, so that users can easily jump right to those locations for bug fix.

### 2. TEAMWORK COLLABORATION

DDE supports project creators and teams to work collaboratively. It has strict access management, to facilitate task list allocation and mutual check and work coordination mechanisms.

The project creator can adjust the team members’ access to view and edit the project at any time. Team members in the project team can dynamically check whether there are users in the team who are online and offline at the same time.
3. DEFINE.XML REAL-TIME GENERATION AND PREVIEW

DDE has one-click preview, supports editing and result preview real-time synchronization, quickly and efficiently renders Define.xml file.

This feature is very helpful for users to view real-time progress. This feature also verifies that all edits to the Define.xml file have been successfully completed.

4. SUPPORT DIFFERENT VERSIONS OF CDISC STANDARDS

DDE supports different versions of CDISC standards (Controlled Terms, SDTM and ADaM). It can re-use metadata in the project, arbitrarily combine various metadata lists according to the selected metadata.
dictionary; Users can re-use the metadata by copying the project, to reduce repetitive operations and improve editing efficiency.

**5. SNAPSHOTS AND VERSION MANAGEMENT**

Users can make snapshots of important metadata information. It's like version control of Excel mapping specs. Users can also export all metadata of the current project to Excel file for version management. Data snapshots make it easy for users to do historical queries and rollback operations.
6. PROJECT METADATA LOCK AND UNLOCK

The current project can be locked or unlocked for the solidification of project data for freezing and archiving.

After the project data is locked, all users in the team can only view the metadata and cannot modify or delete the operation. All the users of the team can continue to modify and edit the metadata only after the project creator unlocks the project.

7. CONCISE ADAM ANALYSIS RESULTS METADATA EDITING

Sometimes users may have difficulty on how to edit the ADaM analysis result metadata. DDE has established a complete editor structure through strict metadata cross-checking control.

The system visualizes and rationalizes the metadata of the ADaM analysis results, making the editing process simple and efficient, and is convenient for user to understand the business and editing operations.
8. EXPORT MAPPING SPECS TO EXCEL FORMAT

User can export all metadata of the current project to the Excel file for version control, as needed at different times and different schedules.

CT controlled terms also support the direct import of excel files, which is more convenient and flexible for users.

9. SUPPORT FOR CUSTOM CDISC DOMAINS AND CT TERMS

The system not only supports the CDISC domains and CT terms in the project, but also supports custom domains and CT terms.

In addition, the system also supports editing CDISC metadata dictionaries, to fit the project needs.
10. MULTI-PROJECT MULTI-TAB DISPLAY AND COMPARISON

The system supports multiple metadata editing interfaces in multiple projects at the same time, which is convenient for users to consult, copy, compare, add, delete and change operations, which is convenient and flexible.

This function is commonly used in the comparative analysis of metadata for similar projects, help users to find similarities and differences, then users can copy and paste cell metadata as needed.

11. AUTOMATIC GENERATION OF SAS CODE FOR GENERATING DATASETS

Each data set in the project can generate relevant SAS code for download or copy to the clipboard at user's convenience.

These SAS codes are automatically generated based on the data set metadata filled out by the user.

Users can download these code files or copy the code directly to the system clipboard for easy copying and pasting.
12. SUPPORT FOR RESTFUL API

DDE supports RESTful API data interface, which can seamlessly exchange related data with other systems.

The RESTful API data interface is a commonly used inter-system service data exchange technology and can also be customized to develop special interfaces according to different requirements of other systems.

13. CHINESE AND ENGLISH LANGUAGES FREE TO SWITCH

The system supports different two sets of user interfaces in Chinese and English, which is convenient for users to choose.

14. CROSS CHECK AND BUG TRACKING
DDE has built-in cross-checking between common related components, through the quick jump operation mode, helps users quickly locate the problem. Users can view the bug tracking list in real time to facilitate timely modification and adjustment.

Along each step of the user's operation, DDE performs cross-checking and bug tracking in real time. Following CDISC’s definition rules for Define.xml, important logical errors will be promptly warned and marked.

![Bug Tracking Image]

**SYSTEM DEPLOYMENT**

According to different user requirements, Standard Edition users can directly access the DDE cloud server, which is convenient and flexible. Enterprise users can deploy the system to an in-house server for easy unified management of applications.

The system supports mainstream server platforms such as Linux and Windows, and the popular Docker containers.

**COMPARISON OF ADVANTAGES AND DISADVANTAGES**

Through the comprehensive comparison with most of the related software for Define.xml generation, overall, DDE has a strong cost-effective advantage, as follows:

1. One-click preview function while in the editing operation.
2. Visual Excel editing style, automated assisting functions, cross-checking and bug tracking.
3. Online editing, unified management, support team collaboration.
4. Enterprise server deployment: support most mainstream server platforms.
5. Support project metadata locking and unlocking, metadata snapshot management.
6. Blank dataset generation SAS code is automatically generated by DDE; supports RESTful API
PRODUCT PLANNING

FUTURE PRODUCT DEVELOPMENT
1. Export to Word format for easy review and comment.
2. Enhance dynamic cross-checking and bug tracking capabilities. The business and validation rules from regulatory agencies will also be implemented step by step.
3. More intelligent assisting features.

CONCLUSION
DDE follows CDISC Define.xml standard strictly and provides effective and easy user interface. It considers the actual user operation experience and metadata contents. It increases the efficiency for generating define.xml and standardizes the operation procedures, while lowering the training cost and learning curve.

REFERENCE
CDISC SDTM, https://www.cdisc.org/standards/foundational/sdtmig
CDISC ADaM, https://www.cdisc.org/standards/foundational/adam
FDA Study Data Standards Resources, https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources
Visual Define-XML Editor, https://github.com/defineEditor/editor

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

  Dylan Song
  Yinzhen Information Technology Co., Ltd.
  Song.xinyan@foxmail.com
  https://linksales.cn

  John Wang
  dMed Biopharmaceutical
  jun.wang@dmedglobal.com
  https://www.dmedglobal.com

Any brand and product names are trademarks of their respective companies.