PP04

SEND Architecture Facilitates Harmonization and Aggregation of Data from Different Organizations/LIMS

Frederic Mura, Rich Buchanan, Laura Kaufman, Reto Aerni,
Christina Wuermlin, Amelia Bedeoan, Michael Wasko

pds Intuitive Software Insightful Solutions

PDS Preclinical Data Systems Inc., Basel Switzerland and Mt. Arlington, NJ USA

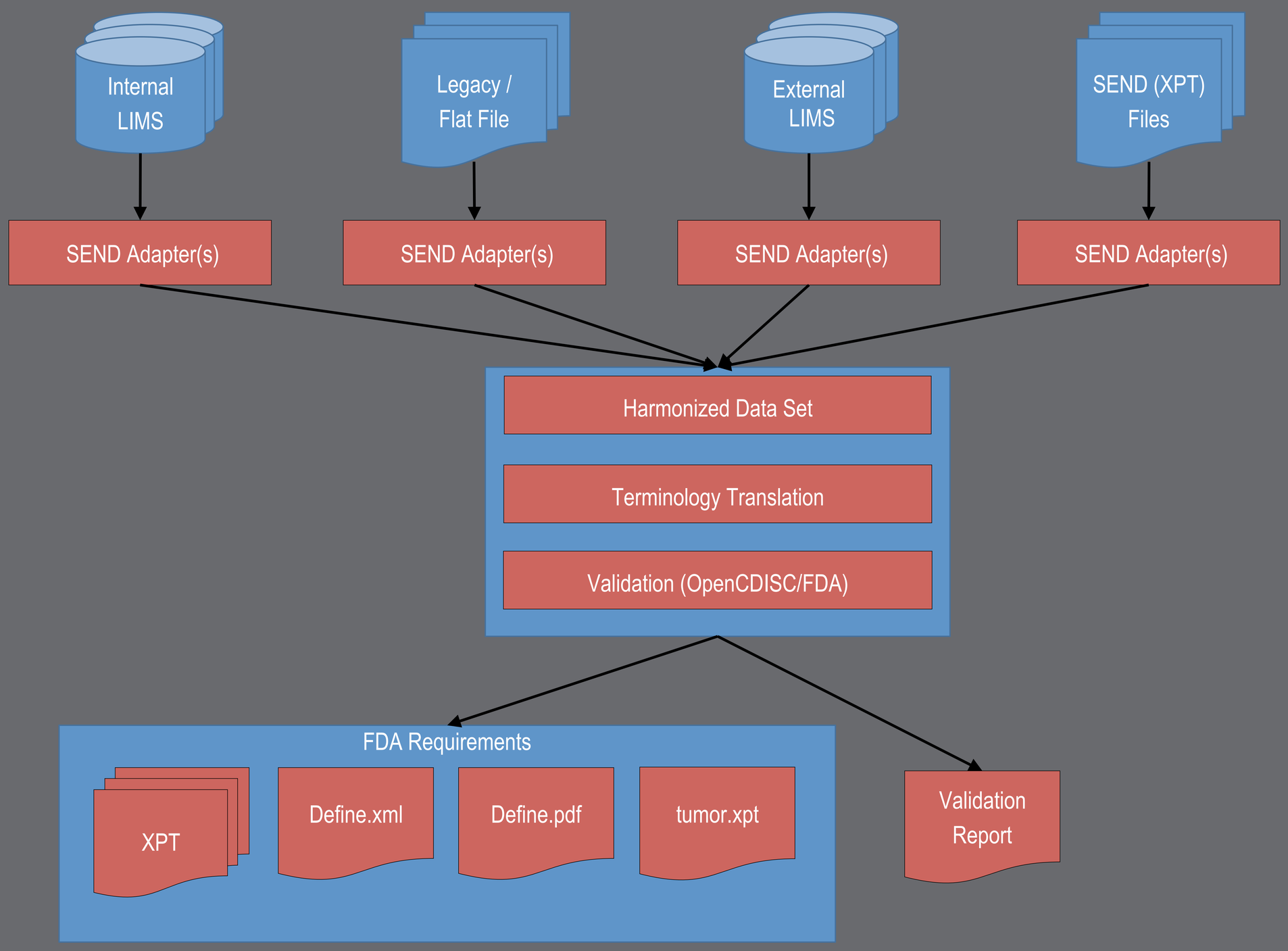
Summary

Toxicology studies in today's business environment involve multiple testing sites, LIMS, and/or departments. Accordingly, SEND solutions need to harmonize potential heterogeneous metadata, file formats, terminology, and other aspects of study documentation. We wish to report a software architecture for SEND that accomplishes the aggregation, harmonization, and translation of data from different sources into one complete dataset including all XPTs, define files, and validation report(s). The architecture involves input from LIMS-specific or file-based adapters, which is then processed through an engine that maps data to appropriate variables and domains; harmonizes data (e.g. animal number, study number) across different domains; consolidates comments (CO); relates records across domains (RELREC) and performs controlled terminology translation. Trials information is provided to the engine through an API. This SEND solution has been used successfully to submit SEND datasets from multiple source systems to FDA.

Integration and Harmonization of SEND Files in Addition to Aggregation

- Harmonization of metadata across all domains is essential
- Toxicology studies require internal correlations for data interpretation. These include:
  - Correlation of macroscopic and microscopic findings
  - Correlation of clinical signs, macroscopic, and microscopic findings when palpable masses are present
  - Correlation of pharmacokinetic calculations with animals and time points from which the calculations were made
  - Correlation of comments with the appropriate finding(s) and animal(s) where appropriate
- The same correlations are also needed in SEND, and can be accomplished in several ways, including RELREC.
- A requirement for SEND-compliant solutions is correlating findings that may have been collected with different LIMS and at different organizations, such as clinical observations and microscopic pathology. The architecture below is platform agnostic and facilitates such required correlations.

Architecture Diagram



Features of SEND Architecture

- Engine supports multiple data sources and formats. Includes direct access to databases as well as flat files and accepts CSV, TXT, XML, and XPT files to facilitate ease on implementation
- *All* files are processed through the engine for harmonization (e.g. animal number, study number); terminology translation and mapping; correlation (RELREC); and validation
- One define file is produced for the entire data set as required by FDA
- One validation report is produced for the entire dataset (option for with or without define.xml)
- The OpenCDISC Validator can be extended to also include additional rules, such as FDA validation rules, increasing FDA acceptance probability
- Easy versioning as SEND evolves with component-based software, e.g. adapters, engine
- Web-based user interface
- Back-end file management

Conclusions

- It is possible to create an architecture that encompasses a complete solution for the creation of a fully compliant SEND dataset.
- A comprehensive SEND solution is critical for FDA compliance.
- The reality of distributed workflow across multiple organizations and LIMS makes data harmonization critical.
- All contributors to preclinical development must have the ability to supply data that can be easily harmonized.
- Simply collecting individual domains from multiple sources will not result in a SEND compliant data set ready for FDA submission.