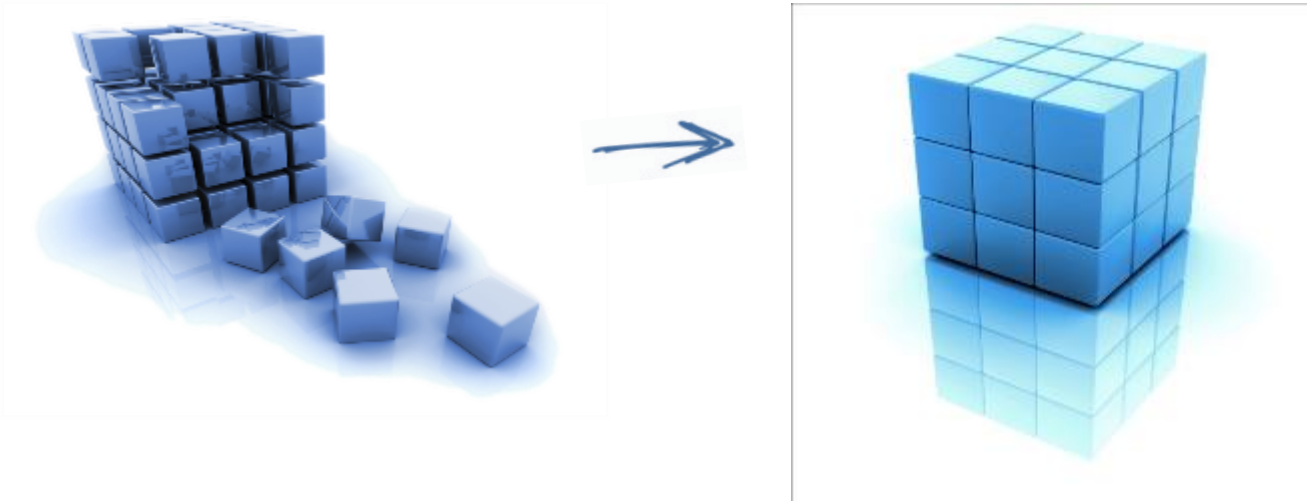
Accelerating Consistent and Quality Nonclinical SEND Study Submissions with Automation and Machine Learning

Suresh Madhavan, Kurien Abraham



Nonclinical eData (SEND) had a poor start

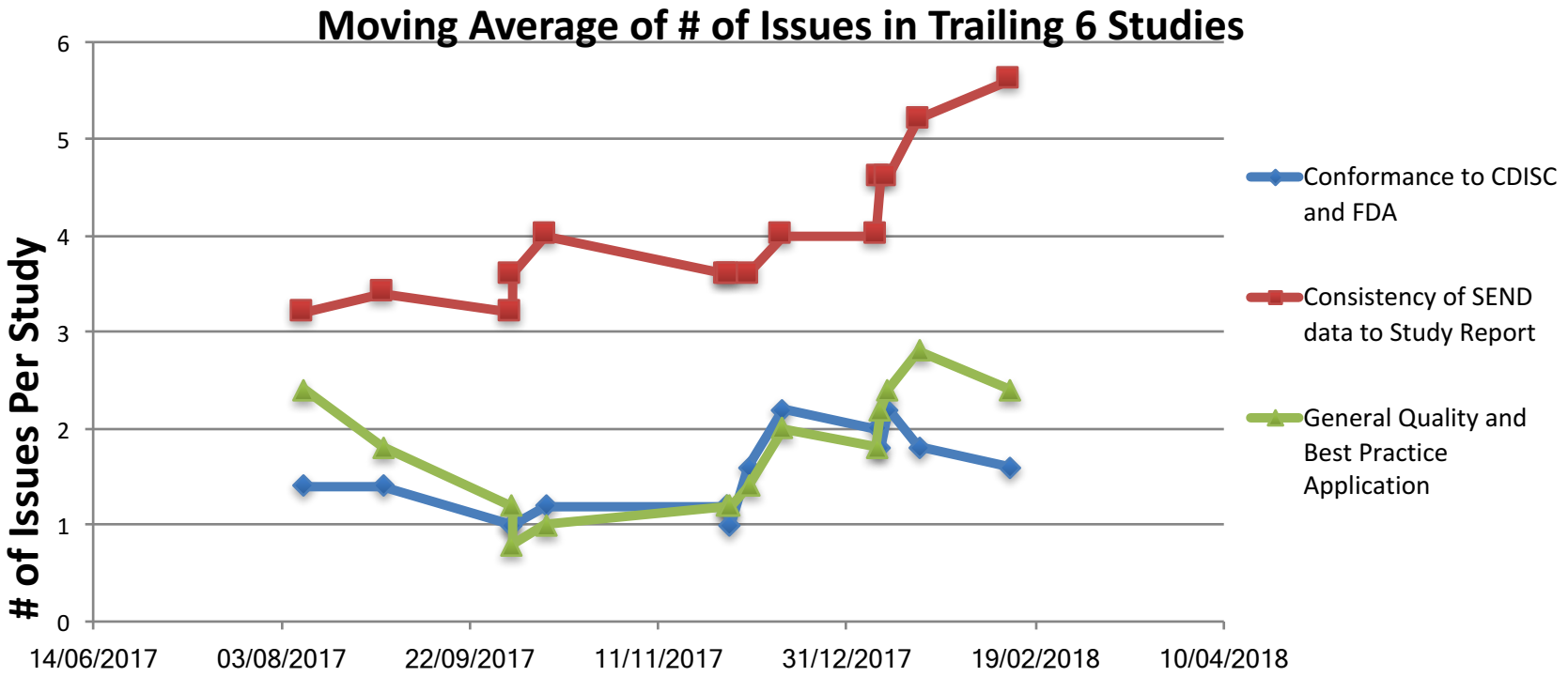
- In the run up to the Dec 2017 mandate hardly 15 nonclinical SEND studies, out of an annual 5,000 typically expected, were submitted to the FDA. More than 75% of the clinical studies were compliant 2 years ahead of their mandate in Dec 2016

- Readiness means studies prepared to the digital eData (CDISC SEND) standards are:
  - **Technically Conformant**:
    - To the CDISC SDTM or SEND standards
    - To the FDA business rules
  - **Consistent** with the GLP (or GCP) audited, signed Study Reports
    - Don't give reviewers a reason to doubt the SEND data
    - Forcing them to go to the Study Report defeats the purpose of eData
  - Meets the **Quality** under best practices of eData production

- As skills and experience grows, Technical Conformance and General Quality is expected to improve. But Consistency is getting worse. Accelerating submissions will require attention to improving consistency

What does the data show?

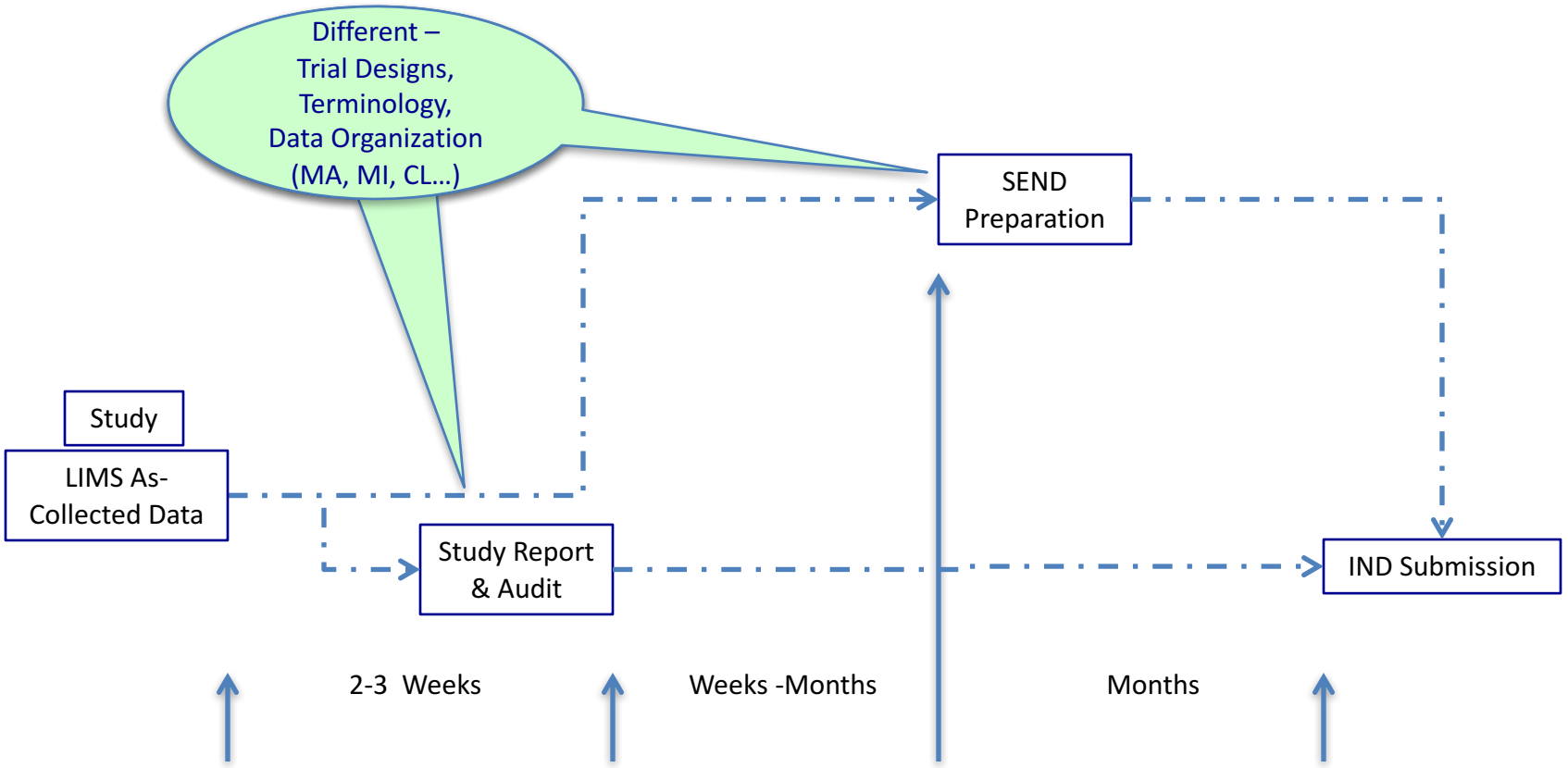
- SEND data sets "consistently inconsistent" with the Study Report



- SEND preparers are still struggling just to create technically conformant SEND data sets

- This was a topic at the FDA presentation at the October '17 SEND Face to Face at Silver Spring, MD and eData quality and consistency is an issue.

Timelines for a typical Nonclinical Safety Study



- Study Close
- **Study Audit**
- **Report Preparation** by Study Director's Team
- SEND Preparation begins by a separate team
- Submissions (IND, studies to NDA)
- SEND IG? CT? Different for each study or unify to same?

Reasons why SEND data sets are deviating from Study Reports

- Study Report Team or Director rarely gets involved in the SEND preparation process
- SEND terminology and structures are still foreign to most toxicologists
- SEND teams tend to be standards and data management savvy and distant from the actual study, its design or the findings until they begin the SEND process
- No restoring forces that ensure SEND data is regressed to Study Report – other than manual checking
- Although both start from the same source, many factors cause SEND teams and their work products to militate away from the Study and the Study Report

How to Ensure Consistency with Study Report

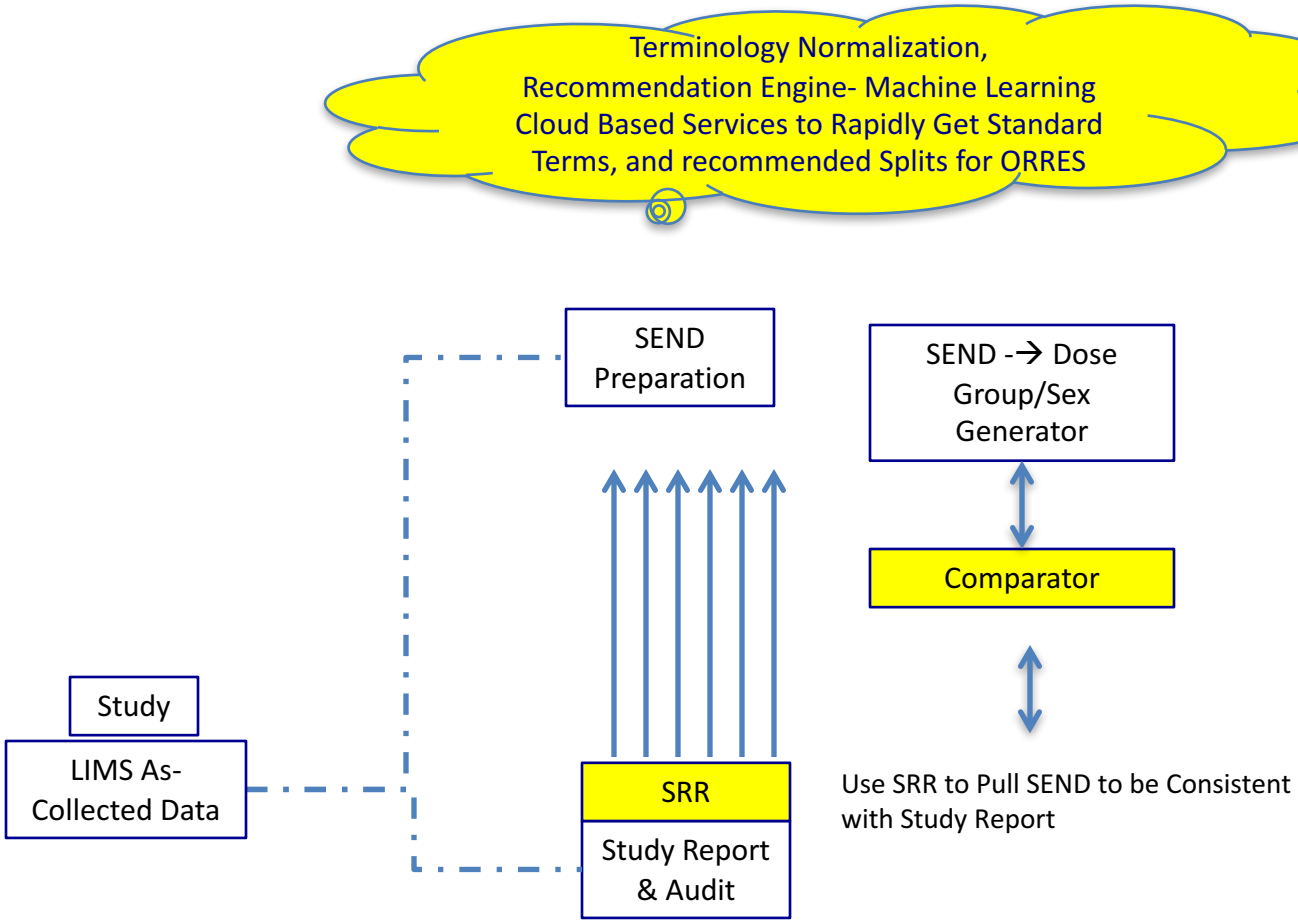
- Begin with digitizing key parts of Study Report
  - Field values that fill Trial Summary (TSVAL)
  - Group summary tables of quantitative data – means, std-dev, counts
  - Group summary tables of qualitative observations – incidence counts - clinical observations, organ macroscopy, tissue microscopy, tumor findings…
  - List of organs observed
  - List of visit days for findings
  - List of test measurement
  - List of derived parameters

- Generate a columnar representation of extracted data from Study Report with provenance

- Use the Study Report Reference (SRR) file (in columnar format) to guide preparation or consistency checks

Generating SEND with SRR as a Guide

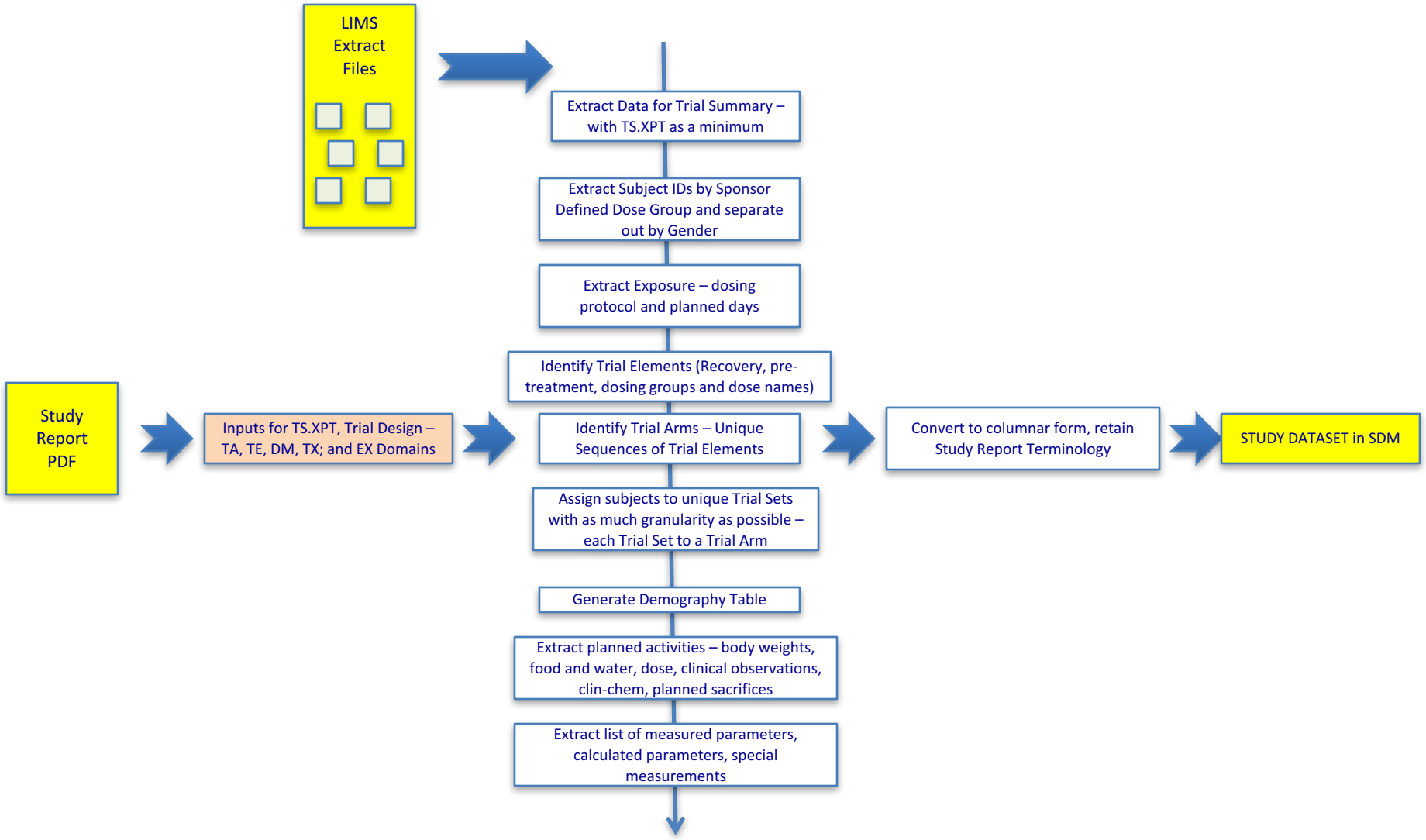
- Using the Study Report (SRR) as a guiding reference ensures that SEND is always aligned with the Study Report and true to the as-collected GLP Source



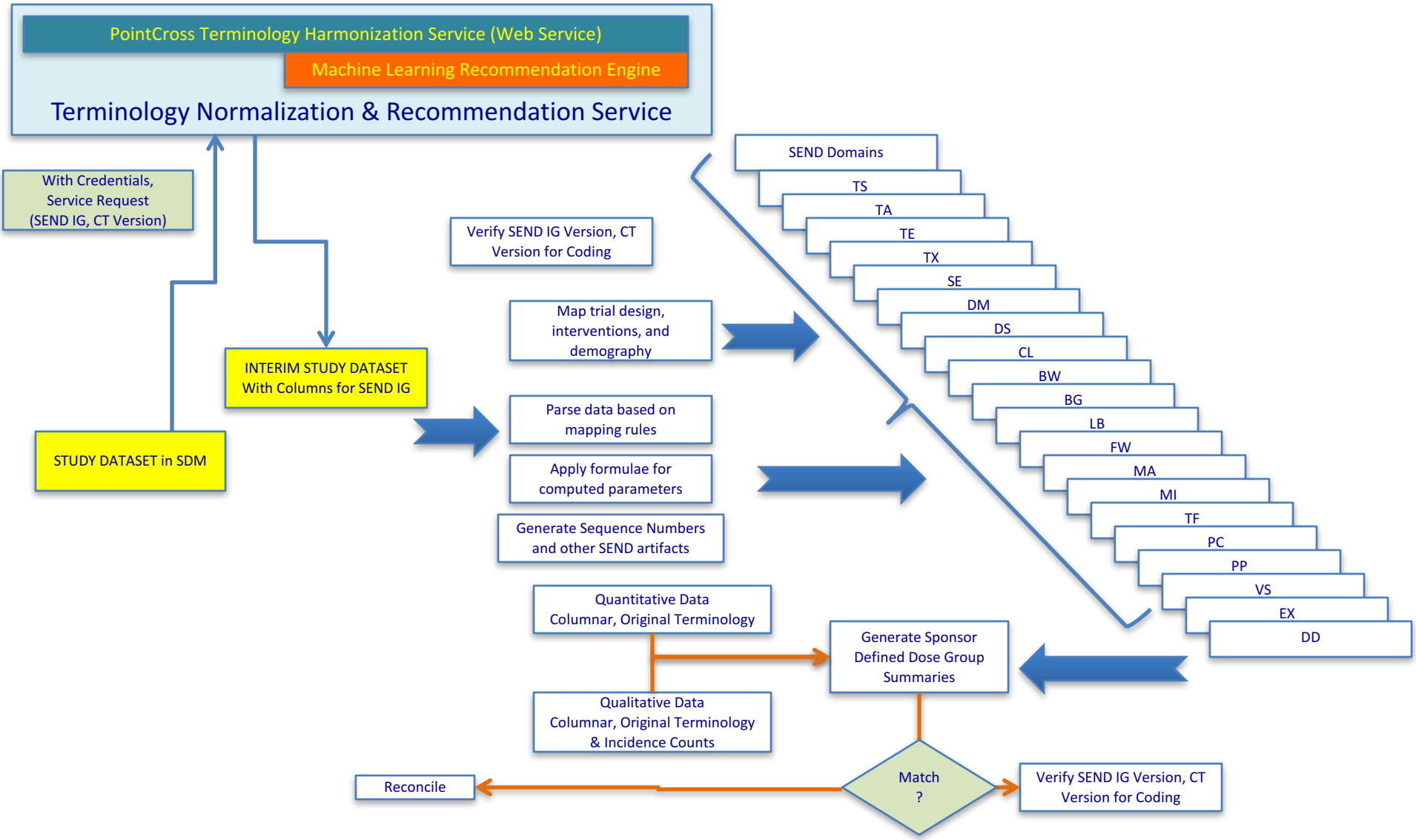
Auto-comparison Analytics

- 100% comparison checks of SEND against Study Report is essential
- Compute the Sponsor defined group summaries using the sponsor dose group flags in the SEND data – means, std-dev, counts
- Normalize the Test Names and Specimen Types in SEND back to the Globally Acceptable standardized terms or the SEND CT that the SRR was coded
- Use the semantic equalizing tools to automatically re-generate the qualitative findings from the SEND components (STRESC, QVAL, Modifier(s), ANTREG, LAT, SEV and others) to match the SRR's findings extracted from the Study Report
  - Use the recommended findings splits to confirm this matching
- Align the SEND data against the SRR records
- Generate comparison of the statistics and flag those that indicate mismatch.

Stage 1 of the SEND Preparation



Final Stage of SEND Preparation



Process for SEND Preparation and Quality Assurance

Cloud based Terminology Normalization

- Drag & Drop Study Report (PDF)
- Generate a Study Report Reference File (SRR) using the SEND Process tools

PDF

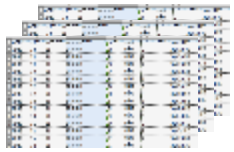
Study Report Reference (SRR) File

Auto-Generate SEND to IG and CT Version

- Ensure Source data for SEND is complete and consistent with SRR
- Auto-Generate SEND to required IG and CT Version
- Generate Reports for updating Define.xml and nSDRG

SEND Process tools

- Drag & Drop SEND Datasets
- Drag & Drop LIMS Extracts



Auto-Compare SRR and Normalized LIMS Data: (Consistency Assurance Tool)

- 100% Cross Check of Quantitative Data
- 100% Check of Qualitative Observation Parsing and Incidence Counts
- Compare Trial Summary Field Values from Study Report to SEND
- Gaps in data in Study Report or LIMS Data

Generate:

- Trial Design for SEND (TE, TA, TX) – SEND Trial Design Generator Tool
- TS.XPT field values
- Group Summary Statistics for Automated Consistency Check of SEND Data Sets
- Unique List of Measured, Derived Test Names, and Qualitative Observations

SEND Validator and Visualization

- Validate SEND against CDISC and FDA rules
- Data Tabulation of all domains
- Data Visualization and Interactive Graphics for SEND data

Improvements take time

- Immediate focus for 2018, 2019 should be to:
  - Generate conformant data sets that pass SEND validation without errors
  - Become better at generating SEND Trial Designs that meet the intent behind Trial Arms and Elements and granular Trial Sets
  - Ensure that the SEND data set is consistent with the Study Report and build the confidence of the reviewers
  - Help the toxicologists and Study Directors cross the barrier of arcane and unfamiliar CDISC and SEND terminology by making SEND data easier to see in their terms

- Longer term focus for late 2019 onwards
  - Exploiting the ability to represent the study digitally for strategic and drug safety purposes
  - Make SEND the normal way to report GLP and non-GLP studies