Paper TT17

Transforming Clinical Trials with Linked Data

Armando Oliva, Semantica LLC, Fort Lauderdale, USA

Tim Williams, UCB Biosciences Inc., Raleigh, USA

# Abstract

The pharmaceutical industry continues to be plagued by data integration and management challenges across the clinical trial data life cycle. Considerable progress has been made in recent years with the implementation of CDISC standards. Historically, standards focused on distinct segments of the clinical trial process: study design, submission, publication. To provide a future-proof solution, these standards must be adapted and integrated holistically and consistently across all use cases.

Linked Data provides a potential solution by representing clinical trial concepts at their atomic level, then leveraging ontological classification and rules integration. This paper reports results from the PhUSE project "Clinical Trials Data as RDF." SDTM data was converted to Linked Data based on CDISC and custom ontologies, then reassembled into high-quality, submission-ready data sets. The approach has several advantages, including the inextricable representation of data and their meaning in ways not possible in traditional approaches.

# Introduction

Intro text ....

Points to Cover:

* Current standards are themselves silos
* Why model the clinical trial
* Scope [could be A Main Section? or under a section called Project Approach or similar?

# A Main Section

Next section text...

# Another Main Section.

Another section text.

# A Subsection

Subsection text as needed.......

# Another subsection

Subsection text as needed.......

# Data Conversion [TW]

Points to cover:

* original method (rrdf/rrdflibs, redland.. )
* SMS (1) method: what it is, compatible with R2RML (1)
* Development and validation tools in SPARQL, RShiny. Future: SHACL?
* Future: Why convert at all? Virtual graphs and mapping to source data

Conclusion

Conclusion text...

References [to be updated based on paper content]

1. **Stardog.** Stardog User Manual. *Stardog Mapping Syntax (SMS).* [Online] [Cited: 04 08, 2018.] https://www.stardog.com/docs/#\_stardog\_mapping\_syntax.

2. **W3C.** R2RML: RDB to RDF Mapping Language. *W3C Recommendation.* [Online] [Cited: 04 08, 2018.] https://www.w3.org/TR/r2rml/.

3. **PhUSE Emerging Trends and Technologies.** *Transport for the Next Generation.* s.l. : PhUSE, 2017.

4. *Technical Rejection Criteria for Study Data - Preliminary Findings.* **Allard, Crystal.** Boston : PhUSE Single Day Event (SDE), 2017.

5. *Managing Study Workflow Using the Resource Description Framework (RDF).* **Oliva, Armando.** Endinburgh : PhUSE Annual Conference, 2017.

6. **CDISC.** About CDISC. *CDISC Website.* [Online] [Cited: 05 01, 2017.] https://www.cdisc.org/about.

7. *State of the Union: The Crossroads of CDISC Standards and SAS' Supporting Role.* **Decker, Chris.** Las Vegas, Nevada : SAS Global Forum 2011, 2011.

ACKNOWLEDGEMENTS

The authors are indebted to the contributions from the "CTDasRDF" PhUSE project team members.

This paper is largely based on free, open-source software and the efforts of volunteers in PhUSE working groups. Please support those who donate their time and expertise through your own collaboration, participation, and promotion of these activities.

Contact Information

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

|  |  |
| --- | --- |
| Armando Oliva, M.D.  Semantica LLC  Fort Lauderdale, FL, USA  aoliva@semanticallc.com  C:\_sandbox\sas\Conferences\PhUSE-SDE-CARY2015\images\Twitter_logo_blue.png @nomini  https://www.linkedin.com/in/aolivamd | Tim Williams  UCB BioSciences, Inc  Raleigh, NC, USA  tim.williams@ucb.com  C:\_sandbox\sas\Conferences\PhUSE-SDE-CARY2015\images\Twitter_logo_blue.png @NovasTaylor  <https://www.linkedin.com/in/timpwilliams> |

All project files, data, and this paper are available from the project's Github repository: <https://github.com/phuse-org/CTDasRDF>. Study instance data: <https://raw.githubusercontent.com/phuse-org/ctdasrdf/master/data/rdf/cdiscpilot01.ttl>

Brand and product names are trademarks of their respective companies.