#### **PhUSE /FDA Project Request**

1. Project Title:

|  |
| --- |
| **Study Data Validation and Submission Conformance** |

1. Workgroup Category:

**Emerging Trends & Technologies**

**Optimizing the Use of Data Standards**

**Standard Analyses and Code Sharing**

**Nonclinical Topics**

**Educating for the Future**

**Data Transparency**

1. Affected Stakeholders(s):

|  |
| --- |
| Industry and regulatory agency stakeholders responsible for creating, validating, reviewing, and/or consuming both clincial and nonclinical data. Potential direct collaboration with FDA and academia. |

1. Project Scope:

|  |
| --- |
| This project will replace its predecessor "Going Translational With Linked Data (GoTWLD)" by extending existing work in both SDTM (CTDasRDF project) SEND Conformance ("SENDConform sub-project of GoTWLD) by further developing the data models and instance data conversion. The project will include modeling FDA Technical Rejection Criteria to faclitate submission of data to the FDA.  **Value Proposition**   * Conformance errors for study data submissions to the FDA can be largely decreased using an ontology-based, Linked Data model, including validation using Shapes Constraint Language (SHACL). * The project will demonstrate creation of highly-compliant, high-quality preclinical and clinical study data for submissions, using a much more automated process than is currently available. Costs for data review, validation and re-work will be greatly reduced. * Separation of the results (instance) data from the standards data and metadata results in a version-free graph data structure for nonclinical studies and clinical trials results. CDISC-compliant data for submissions will be created by mapping the standards data to the results data. Costs for recoding between CDISC versions will be drastically reduced. * Metadata for submissions packages to the FDA can be standardized, validated, and semi-automated. * The project provides a gateway for Knowledge Graph technology to support the FDA's Technology and Modernization Action Plan ([TMAP](https://www.fda.gov/media/130883/download))   F.A.I.R. Principles (<https://www.go-fair.org/fair-principles/>) will be followed with all work made available on Github, including a comprehensive website for documentation, explanation, and resources (<https://phuse-org.github.io/SENDConform/> - currently under early construction, URL subject to change). Web Protégé (<https://protegewiki.stanford.edu/wiki/WebProtege>) will be used for collaborative ontology development. |

1. Project Dependencies:

|  |
| --- |
| * Participants with expertise in RDF creation and querying, including ontologies, SHACL, and related tools. * CDISC SDTM and SEND terminology and domain knowledge. * Participation from the FDA; specifically a sample or dummy eCTD file to support RDF modeling   Team members actively working on the GoTWLD project will roll-over into the new project, ensuring adequate staffing and consistency of approach. By inviting FDA participation (and potentially academia) the project hopes to avoid the lack of staff resources which have hampered past initiatives. |

1. Project Objectives and Timeline:

The following deliverables are planned for February 2021.

* Supporting Ontologies
* Techincal Rejection Criteria Proof of Concept
* Submission Metadata Collection Proof of Concept
* Documentation and Resources (website)
* Project Conclusion

1. Project Lead(s):

|  |
| --- |
| * Co-lead: Tim Williams, UCB Biosciences * Co-lead: Armando Oliva, Semantica LLC |

1. Project Requestor(s):

|  |
| --- |
| Tim Williams, Armando Oliva |