Version for Website/Registration (Submitted Jan 2019)

**Let's Make a Knowledge Graph! An Interactive, Hands-on Workshop.**

**– Tim Williams, UCB**

*Knowledge Graph* is a term that is gaining popularity to describe multi-dimensional graph databases that use a reasoner to infer knowledge from data. It sounds complex, but at its core is a very simple way to join data together using meaningful relationships. F.A.I.R. data (<https://www.go-fair.org/fair-principles/>) is built on these Linked Data concepts and it can provide future-proof data for pharma, breaking down traditional data silos in a highly inter-connected, extensible way.

This is an updated version of previous workshops at the CSS and EUConnect conferences We invite attendees who have not participated in the previous workshops to experience Linked Data in this interactive session. You will use a web application to diagram relationships for clinical trial processes and data, then convert your white board drawing to Resource Description Framework (RDF). You will then query the data, using an ontology to infer values and relationships not in your original content. As a last step you will seamlessly merge your study with data from all the other attendees.

This introductory workshop provides the background you need to launch your own exploration of this technology or participate in a PhUSE project. We welcome attendees with no previous experience with Linked Data.

Preregistration is required. You must bring a laptop with remote desktop capability and attend a preparatory webinar in the days preceding the workshop.

Version for Brochure (To Submit: May 14 2018) : ~350 Words

Linked Data has the potential to solve many of the challenges currently facing the pharmaceutical industry. Clinical trials are increasingly sophisticated, with the need to integrate information from diverse sources including devices, genomics, Real World Evidence, and even social media. Our current standards exist across multiple documents, are prone to inconsistent interpretation, and new versions are released regularly. It is difficult to determine if data created in one version of a standard can be merged with data created using a different version.

A new approach is needed where we model clinical trial processes and data using highly interconnected, extensible data with precise meaning (semantics).  When the clinical trial process is an integral part of the data model, a positive impact can be realized across the entire lifecycle, from design, through conduct and data collection, all the way to submission and publication. It also can fuel the surging interest in Natural Language Processing, Machine Learning, and Artificial Intelligence.

Linked Data represents values, concepts, and knowledge at an atomic level. These simple building blocks are explicitly identified, with meaningful relationships between them that are both human and machine readable.  This allows construction of flexible, extensible data models that can be built incrementally and facilitate the breakdown of traditional data silos.

In this hands-on workshop you will use a web application to diagram relationships for clinical trial processes and data, then convert your white board drawing to Linked Data as Resource Description Framework (RDF). You will then query the data, using an ontology to infer values and relationships not in your original content. In a last step you will seamlessly merge your study with data from all of the other attendees.

This introductory workshop provides the background you need to launch your own exploration of this technology or participate in a PhUSE project. No previous experience is assumed.

Preregistration is required. You must bring a laptop with remote desktop capability and attend a preparatory webinar in advance of the conference.