

# Linking Clinical Data Standards

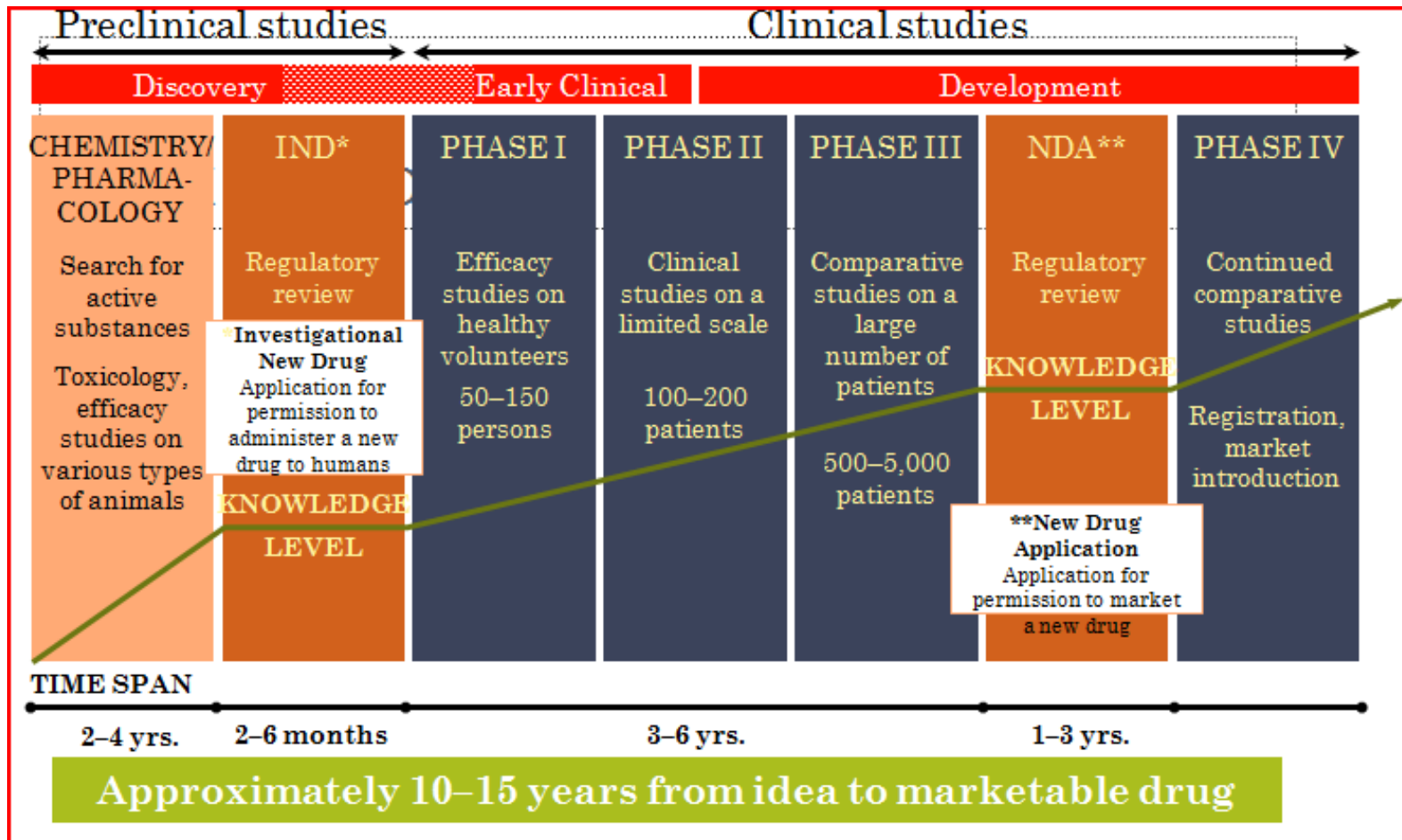
Kerstin Forsberg

AstraZeneca, R&D Information

[kerstin.l.forsberg@astrazeneca.com](mailto:kerstin.l.forsberg@astrazeneca.com)

**kerfors** on Twitter, Google+, LinkedIn, Slideshare, Blogspot

# Pharmaceutical Research and Development



## Clinical Studies And The Road To Linked Data

By Jennifer Zaino on February 11, 2013 8:00 AM

Clinical studies aren't what they used to be. In the past, the process was one-off: You conducted a study, gathered a lot of data, analyzed it, wrote a report, and submitted it to the authorities. But, says long-time Linked Data advocate Kerstin Forsberg, an information architect at [AstraZeneca](#), that's all changed in the last few years.



"A study is not a study on its own," says Forsberg. Today, the goal is to do meta-analysis across many studies, so parties ranging from pharmaceuticals companies to contract research organizations to government authorities all are 'customers' of clinical data, so to speak. Data from various studies must be shared among all these parties. "It

puts a new context around clinical trial data, that it must be easy to link data together, to link across several different studies," she says.

The case is there to use modern information standards, like semantic web standards and Linked Data principles, to address this need. It's why Forsberg is one of the individuals spearheading a volunteer effort to create RDF and OWL representations of the standards published by the [Clinical Data Interchange Standards Consortium \(CDISC\)](#) an international, non-profit organization that develops and supports global data standards for medical research.

The problem is that CDISC publishes its standards in huge PDF documents, with Excel matrices and some first-generation XML implementations, she notes. "If I want to know how to exchange lab data with a number of different measurements, I must pick up the PDF document to see how to structure the data sets, what elements are in them," she says. That's fine for a human to do, but better yet would be to represent the standards data in RDF format so that it would be machine-readable and thus seamless to refer to a single element in a full data set of labs data.

[http://semanticweb.com/clinical-studies-and-the-road-to-linked-data\\_b35252](http://semanticweb.com/clinical-studies-and-the-road-to-linked-data_b35252)

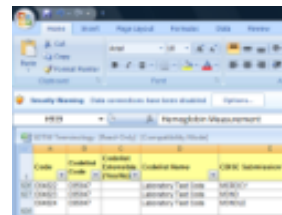
# Data standard organisation



CLINICAL DATA INTERCHANGE STANDARDS CONSORTIUM

*The CDISC vision is to inform patient care & safety  
through higher quality medical research.*

# Human readable standards



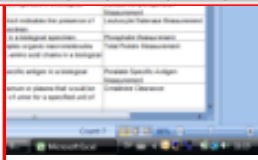
Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	CDISC Notes	Core	References
AESCAT	Subcategory for Adverse Event	Char	*	Sponsor Defined	Grouping Qualifier	A further categorization of adverse event. Example: NEUROLOGIC.	Perm	SDTMG 4.1.2.6
AEOCCUR	Adverse Event Occurrence	Char	**Y, N or Null	CRF or Sponsor Defined	Record Qualifier	Used when the occurrence of specific adverse events is solicited to indicate whether an adverse event occurred or not. Values are null for spontaneously reported events. Note: Use of this variable is under evaluation by the SDS Team.	Perm	
AEBODSYS	Body System or Organ Class	Char	**	CRF or Derived	Record Qualifier	Body system or organ class (Primary SOC) that is involved as an event or assessment from the standard hierarchy (e.g., MedDRA).	Exp	SDTMG 4.1.3.3
AELOC	Location of the Reaction	Char	*	CRF or Derived	Record Qualifier	Describes anatomical location relevant for the event. (e.g., LEFT ARM for skin rash).	Perm	
AESEV	Severity/Intensity	Char	*	CRF or Derived	Record Qualifier	The severity or intensity of the event. Examples: MILD, MODERATE, SEVERE.	Perm	
AESEV	Serious Event	Char	**Y, N	CRF or Derived	Record Qualifier	Is this a serious event?	Exp	
AEACN	Action Taken with Study Treatment	Char	*	CRF	Record Qualifier	Describes changes to the study treatment as a result of the event. Examples include ICH E1B values: DRUG WITHDRAWN, DOSE REDUCED, DOSE INCREASED, DOSE NOT CHANGED, UNKNOWN or NOT APPLICABLE.	Exp	
AEACNOTH	Other Action Taken	Char	*	CRF	Record Qualifier	Describes other actions taken as a result of the event. Usually reported as free text. Example: "Treatment unblinded. Primary care physician notified."	Perm	
AEREL	Causality	Char	*	CRF	Record Qualifier	Records the investigator's opinion as to the causality of the event to the treatment. Examples: NOT RELATED, UNLIKELY RELATED, POSSIBLY RELATED, RELATED.	Exp	
AERELNST	Relationship to Non-Study Treatment	Char	*	CRF	Record Qualifier	Records the investigator's opinion as to whether the event may have been due to a treatment other than study drug. Reported as free text. Example: "Most likely related to aspirin use".	Perm	
AEPAIT	Pattern of Adverse Event	Char	*	CRF	Record Qualifier	Used to indicate the pattern of the event over time. Examples: INTERMITTENT, CONTINUOUS, SINGLE EVENT.	Perm	

**CDISC**  
**Study Data Tabulation  
 Implementation Guide  
 Human Clinical Trials**  
 Prepared by the  
**CDISC Submission Data Standards Team**

- Notes to Readers**
- This is the implementation guide for Version 1.1 of the CDISC Study Data Tabulation Model, posted for comment.
  - This Implementation Guide comprises version 3.1.1 of the CDISC Submission Data Standards and Domain Models.
  - See CDISC notes and assumptions regarding use of -OCCUR variable in CM, SU, AE, MH domains in Section 6.

**Revision History**

Date	Version	Summary of Changes
2004-07-14	3.1	Released version reflecting all changes and corrections identified during comment periods.
2005-08-26	3.1.1 Final	Released version reflecting all changes and corrections identified during comment period.



# Machine computable and queryable standards

The image shows three overlapping screenshots of the TopBraid Composer ontology editor. The top-left screenshot shows a 'Resource Form' for 'C66767.C49501' with 'Annotations' and 'Other Properties' sections. The middle-left screenshot shows a table of instances for 'C66767.C49501' with columns for 'rdf:type' and 'mms:PermissibleValue'. The rightmost screenshot shows a detailed 'Resource Form' for 'Column.AE.AEACN' with various properties like 'mms:broader', 'mms:context', 'mms:dataElementDescription', and 'mms:dataElementDomain'.

**Resource Form**  
Name: C66767.C49501

**Annotations**

**Other Properties**

ctsnc:Definition  
An indication that a medication schedule was modified by temporarily terminating a prescribed regimen of medication. (NCI)

ctsnc:SubmissionValue  
DRUG INTERRUPTED

ctsnc:Synonyms

ctsnc:Code  
C49501

ctsnc:PreferredTerm  
Drug Interrupted

mms:context  
Drug Interrupted

mms:valueDomain  
C66767

mms:ordinal  
mms:PermissibleValue

**Incoming References**

Resource	rdf:type	rdfs:label
C66741.C81255	mms:PermissibleValue	
C66767.C49501	mms:PermissibleValue	

**Resource Form**  
Name: Column.AE.AEACN

**Annotations**

**Other Properties**

mms:broader  
sdm:DE.Event.--ACN

mms:context  
Table.AE

mms:dataElementDescription  
Describes changes to the study treatment as a result of the event. AEACN is specifically for the relationship to study treatment. AEACNOTH is for actions unrelated to dose adjustments of study treatment. Examples of AEACN values include JCH E2B values: DRUG WITHDRAWN, DOSE REDUCED, DOSE INCREASED, DOSE NOT CHANGED, UNKNOWN or NOT APPLICABLE

mms:dataElementDomain  
Action Taken with Study Treatment

mms:dataElementLabel  
AEACN

mms:dataElementName  
AEACN

mms:dataElementValueDomain  
xsd:string

mms:ordinal  
sdmct:C66767

sdmts:dataElementCompliance  
sdmts:Classifier.ExpectedVariable

sdmts:dataElementRole  
sdmts:Classifier.RecordQualifier

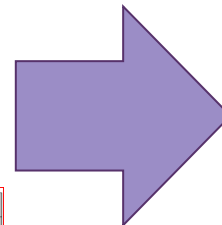
sdmts:dataElementRole  
sdmts:Classifier.Character

sdmts:qualifies

Screenshots from the ontology tool:  
TopBraid Composer

We want to push back to CDISC, and other public and internal standard groups, and show in practice how to: *“Use (semantic web) standards for standards”*

Human readable standards



Machine computable and queryable standards



Variable Name	Variable Label	Type	Controlled Terminology Format	Origin	Role	CDISC Note	Core	Reference
AECA1	Subcategory for Adverse Event	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Is further categorization of adverse event. Example: NEUROLOGIC.	Yes	SDTM304.1.1.1
AECCUR	Adverse Event Occurrence	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Event when the occurrence of specific adverse events is indicated to indicate whether an adverse event occurred at all. Values are used for spontaneously reported events. Note: Use of this variable is under evaluation by the SDTM Team.	Yes	
AEBOVYS	Body System or Organ Class	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Body system or organ class (Primary body part is involved in an event or compartment from the medical literature (e.g., SKIN/DA).	Yes	SDTM304.1.1.1
AELC	Location of the Reaction	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Describes anatomical location where the event occurred (e.g., LEFT ARM or chest only).	Yes	
AELV	Severity Intensity	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	The severity or intensity of the event. Examples: MILD, MODERATE, SEVERE.	Yes	
AELR	Adverse Event	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Is this a serious event?	Yes	
AEACN	Action Taken with Study Treatment	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Describes changes to the study treatment as a result of the event. Examples: WITHDREW STUDY TREATMENT, DRUG WITHDRAWN, DOSE REDUCED, DOSE INCREASED, DOSE NOT CHANGED, UNKNOWN or NOT APPLICABLE.	Yes	
AEACNTH	Other Action Taken	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Describes other actions taken as a result of the event. Usually reported as free text. Example: Treatment withheld. Primary care physician consulted.	Yes	
AELRST	Causality	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Records the investigator's opinion as to the causality of the event in the treatment. Examples: NOT RELATED, UNRELATED, POSSIBLY RELATED, POSSIBLY RELATED - RELATED.	Yes	
AELRSTNT	Relationship to Study Treatment	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Records the investigator's opinion as to whether the event may have been due to a treatment other than study drug. Reported as free text. Example: "More likely related to surgery."	Yes	
AEPATY	Pattern of Adverse Event	Char	CDISC or NCI	Source/Qualifier	Record/Qualifier	Used to track the pattern of the event over time. Examples: INTERMITTENT, CONTINUOUS, SINGLE EVENT.	Yes	

**CDISC**

**Study Data Tabulation Model Implementation Guide: Human Clinical Trials**

Prepared by the CDISC Submission Data Standards Team

---

**Notes to Readers**

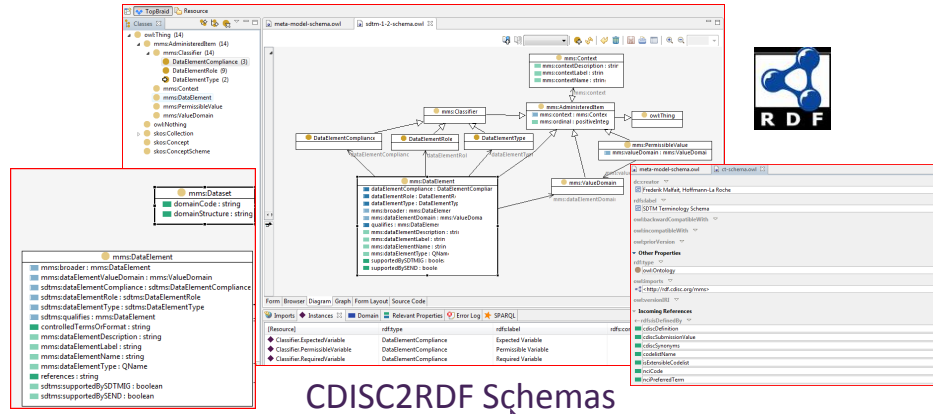
- This is the implementation guide for Version 1.1 of the CDISC Study Data Tabulation Model, revised for comment.
- This Implementation Guide corresponds to version 3.1.1 of the CDISC Submission Data Standards and Domain Models.
- See CDISC notes and assumptions regarding use of \_OCCUR variable in CM, SU, AE, MH domains in Section 6.

**Revision History**

Date	Version	Summary of Changes
2004-07-14	3.1	Released version reflecting 24 changes and corrections identified during comment period.
2007-08-20	3.1.1 Final	Released version reflecting 10 changes and corrections identified during comment period.

# CDISC2RDF

We want to push back to CDISC, and other public and internal standard groups, and show in practice how to: "Use (semantic web) standards for standards"



Human readable standards

Machine computable and queryable standards

Variable Name	Variable Label	Type	Controlled Term or Format	Origin	Role	CDISC Note	Code	Reference
RECAT	Subcategory for Adverse Event	Char	None	Source: SDR	Record Qualifier	A 3-letter composition of adverse event. Example: NEUROLOGIC.	None	SDR550.4.1.1.1
REOCCUR	Adverse Event Occurrence	Char	PV, N, or U	Record Qualifier	Record Qualifier	Used when the occurrence of specific adverse events is isolated to indicate whether an adverse event occurred at all. Values are null for spontaneously reported events. Note: Use of this variable is under evaluation by the SDG Team.	None	
AEBOOVS	Body System or Organ Class	Char	None	SDR or Source	Record Qualifier	Body system or organ class (Primary BODY, not as specified as an event or subevent from the related intensity of a SDR(S)).	None	SDR550.4.1.1.1
AELOC	Location of the Reaction	Char	None	Record Qualifier	Record Qualifier	Describe regional location relevant for the event (e.g., LEFT ARM or distal tibia).	None	
AELRV	Severity Intensity	Char	None	Record Qualifier	Record Qualifier	The severity or intensity of the event. Examples: MILD, MODERATE, SEVERE.	None	
AELER	Adverse Event	Char	PV, N	Record Qualifier	Record Qualifier	Is this a serious event?	None	
AEACN	Action Taken with Study Treatment	Char	None	Record Qualifier	Record Qualifier	Describe changes to the study treatment as a result of the event. Examples: DRUG WITHDRAWN, DOSE REDUCED, DOSE INCREASED, DOSE NOT CHANGED, UNKNOWN or NOT APPLICABLE.	None	
AEACNTH	Action Taken	Char	None	Record Qualifier	Record Qualifier	Describe other actions taken as a result of the event. Usually reported as Yes/No. Example: Treatment withheld. Primary use electronic health.	None	
AELREL	Relationship	Char	None	Record Qualifier	Record Qualifier	Record the relationship's response as to the severity of the event to this treatment. Examples: NOT RELATED, CLOSELY RELATED, POSSIBLY RELATED, RELATED.	None	
AELRELST	Relationship to Study Treatment	Char	None	Record Qualifier	Record Qualifier	Record the relationship's response as to whether the event may have been due to a treatment other than study drug. Reported as Yes/No. Example: "Most likely related to study drug."	None	
REPATY	Pattern of Adverse Event	Char	None	Record Qualifier	Record Qualifier	Must track over the pattern of the event over time. Examples: INTERMITTENT, CONTINUOUS, SINGLE EVENT.	None	

**CDISC**

**Study Data Tabulation Model Implementation Guide: Human Clinical Trials**

Prepared by the CDISC Submission Data Standards Team

**Notes to Readers**

- This is the implementation guide for Version 1.1 of the CDISC Study Data Tabulation Model, revised for comment.
- This Implementation Guide corresponds to version 3.1.1 of the CDISC Submission Data Standards and Domain Models.
- See CDISC notes and assumptions regarding use of \_OCCUR variable in CM, SU, AI, MH domains in Section 6.

**Revision History**

Date	Version	Summary of Changes
2024-07-14	3.1	Released version reflecting all changes and corrections identified during comment period.
2022-09-26	3.1.1 Final	Released version reflecting all changes and corrections identified during comment period.

Blog: <http://cdisc2rdf.com/>  
Google Code: <https://code.google.com/p/cdisc2rdf/> (under Source)

Diagram showing the CDISC2RDF Instance View. It includes a resource form for `Column.AEACN` with fields like `description`, `intensity`, `location`, and `actionTaken`. A sidebar on the right shows a form view for `Column.AEACN` with fields like `intensity`, `location`, and `actionTaken`. A large purple arrow points from the 'Machine computable and queryable standards' section to the 'Human readable standards' section.



# Next step

- A general approach to the use of multiple standards that collectively support standardized data/meta-data exchange and re-purposing
  - Standards-as-is
  - Standards-in-context
  - Interoperability across standards and the data collected using them



## Overview to the Semantic Web Technologies

**Mitra Rocca**  
**Medical Informatician**  
**FDA/CDER/OTS**

**28 January 2013**

1

## Partnerships for Innovation ...

### Research Project Description

**Semantic Web Technologies Fellowship**  
**Office of Translation Sciences**  
**Center for Drug Evaluation and Research**  
**U.S. Food and Drug Administration**  
**Silver Spring, MD**

**FDA-CDER-2013-0009**

#### **Project Description:**

A postdoctoral fellowship is available within the Office of Translation Sciences (OTS) in the Center for Drug Evaluation and Research (CDER) of the U.S. Food and Drug Administration (FDA). The selected participant will be involved with efforts to use Semantic Web technologies to develop clinical data standards for various therapeutic areas at CDER.

The participant may be involved in one or more of the following aspects of the project:

- Reviewing artifacts, such as case report forms (CRF)
- Sampling data
- Gathering requirements from subject matter experts/reviewers at CDER
- Conducting research utilizing Semantic Web technologies for data warehousing and re-use of Electronic Health Record (EHR) for clinical research.

This project aims to improve the efficiency of human drug review through required electronic submissions and standardization of electronic drug application data utilizing Semantic Web technologies.