

Introduction into CDISC BRIDG

Jörg Dillert, Oracle Health Sciences, Potsdam, Germany

ABSTRACT

CDISC BRIDG (Biomedical Research Integrated Domain Group) is the canonical model in the biomedical / clinical research arena. On one end it serves as a communication model between domain users and the technical experts, but can be used further for enterprise interfacing for message-based communications and data exchange. Beside, BRIDG is on the way to get ISO approval and has reached an important step on this way in April 2015.

The challenge with BRIDG is, that is partial overwhelming and stays under the impression to be too complicated.

This article will give a simplified view on BRIDG to start up understanding and takes a simple example on how data from one system / standard can be / are mapped into BRIDG.

Finally I will highlight what's new in Version 4 which is (at the time writing the abstract) out for comments.

INTRODUCTION

A LITTLE BIT OF HISTORY

2003-2004: CDISC AND HL7 RCRIM WG

The idea of building a Domain Analysis Model (DAM) which documented the shared semantics of the protocol-driven-research domain began in September 2003. Following a discussion of the benefits, risks, and effort involved in pursuing the path of building a DAM, the CDISC Board of Directors (BoD) elected to begin funding the development of a CDISC DAM. The process officially started with a scope definition in March 2004 and the resulting DAM becoming the BRIDG Model around one year later.

2004-2005: THE NCI CABIG® (AND CDISC) STRUCTURED PROTOCOL REPRESENTATION TEAMS

In February 2004, NCI began a project to develop a structured model of a clinical trial protocol as part of the caBIG® project. The search of other interested groups naturally led to interactions with representatives from both CDISC and the HL7 RCRIM WG. The first joint modeling sessions for a "CDISC Domain Analysis Model project" involving both members of the CDISC DAM team and the NCI team were held at the end of 2004. This initial effort, the combination of the work that CDISC had done during the first 9 months of CDISC Domain Analysis Model Project and the first attempts to integrate the work of the NCI team, resulted in the first DAM of the domain of protocol-driven research.

2005-2006: MATURATION OF THE BRIDG PROJECT

The project was officially renamed "BRIDG" in the summer of 2005. BRIDG, which stands for "Biomedical Research Integrated Domain Group" was chosen to reflect the project's contribution to "bridge" between the various stakeholder organizations. A more formal organizational governance structure must be defined. Interest Groups like HL7, W3C, and OASIS have been involved. An early version of the BRIDG Model was presented to the HL7 RCRIM WG in late-2005/early-2006 and was subsequently accepted as the single Domain Analysis Model for use by the HL7 RCRIM WG.

HL7 and CDISC decided to reverse engineer all of their standards (HL7: implementation-independent semantics from RCRIM WG specifications – CDISC: all existing CDISC specifications [e.g. SDTM, ODM, etc.]).

All four of the BRIDG stakeholder organizations are where involved with applying "BRIDG in context" in a number of application and/or data interchange projects focused on computable semantic interoperability and created the Board of Directors for additional strategic oversight and a Technical Harmonization Committee to manage the operational details in 2005.

PhUSE Annual Conference 2015

2007 – 2012: BRIDG EVOLUTION

As BRIDG became more visible in the community, the BRIDG SCC started hearing feedback from both domain experts and technologists. Many domain experts said that the model was too large, too technical, and did not represent their semantics in language they could understand. The technologists said that the model was not clear enough to map cleanly to downstream models (e.g., design models, implementation models, etc.). As a result, the BRIDG BoD and SCC designed a BRIDG solution that would present each audience of BRIDG with its own perspective.

In addition, the BRIDG BoD decided in late 2008 to move BRIDG forward as an international standard through the ISO and has been forwarded through CDISC

2013 – PRESENT: BRIDG SCOPE EXPANSION – TRANSLATIONAL RESEARCH

Following a gap in BRIDG project in late 2012 to early 2013, during which the BRIDG BoD was reorganized into the BRIDG Advisory Panel, the leadership of BRIDG determined that an NCI model called the Life Sciences Domain Analysis Model (LS DAM) should be harmonized with BRIDG. The LS DAM harmonization introduced 3 new sub-domains covering concepts related to biospecimens, molecular biology and experiments.

In 2014, a new work group called the HL7 BRIDG Work Group was started. This work group provides a platform for BRIDG users and implementers to discuss BRIDG from an implementation view point and identify areas of enhancement.

Subsequently, in early 2015, the CDISC Pharmacogenetics and Pharmacogenomics (PGx) domains were harmonized with BRIDG to compliment the related genetic/genomic concepts added to the model for LS DAM. The PGx concepts leveraged and helped flesh out some concepts identified earlier by the LS DAM harmonization.

A WORD ABOUT UML

The BRIDG model is documented using the UML: Unified Modeling Language. By definition through the Object Management Group (OMG), UML is “... is a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system. The UML offers a standard way to write a system's blueprints, including conceptual things such as business processes and system functions as well as concrete things such as programming language statements, database schemas, and reusable software components.”

UML contains 13 different diagram types with a package diagram on top to divide the model into logical container. From a simple Use case diagram to a complex class diagram - UML is just a 'language', to 'translate' between different 'interpretations and views' of certain domain topics. Technically it can be used to generate further documentation as well as software – but it still is a communication tool between domain and IT experts.

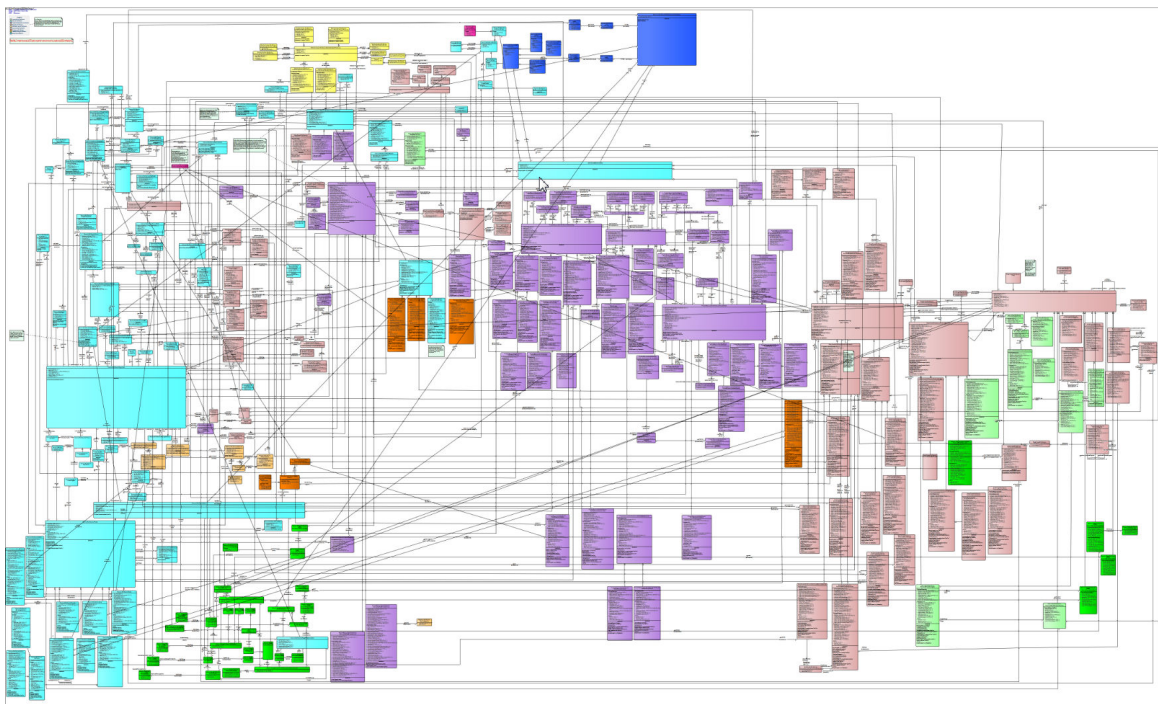
The BRIDG model uses mainly class and activity diagrams.

For a more detailed introduction into UML see the reference section.

LET'S LOOK INTO THE MODEL

If you open up the model the first time, you get overwhelmed with (class) boxes.

PhUSE Annual Conference 2015



And if you look into the documentation, which can be downloaded as part of the package, you will find 1365 pages. Not the best way to start to understand the model.

But if you look into the legend, you will see some color code information about the meaning of certain classes. This is the starting point:

Name: UML-Based Comprehensive BRIDG Model Diagram
Package: BRIDG Domain Analysis Model
Version: 4.0
Author: BRIDG SCC

Legend

- Adverse Event Sub-Domain
- Common Sub-Domain
- Protocol Representation Sub-Domain
- Regulatory Sub-Domain
- Statistical Analysis Sub-Domain
- Study Conduct Sub-Domain
- Biospecimen Sub-Domain
- Molecular Biology Sub-Domain
- Experiment Sub-Domain

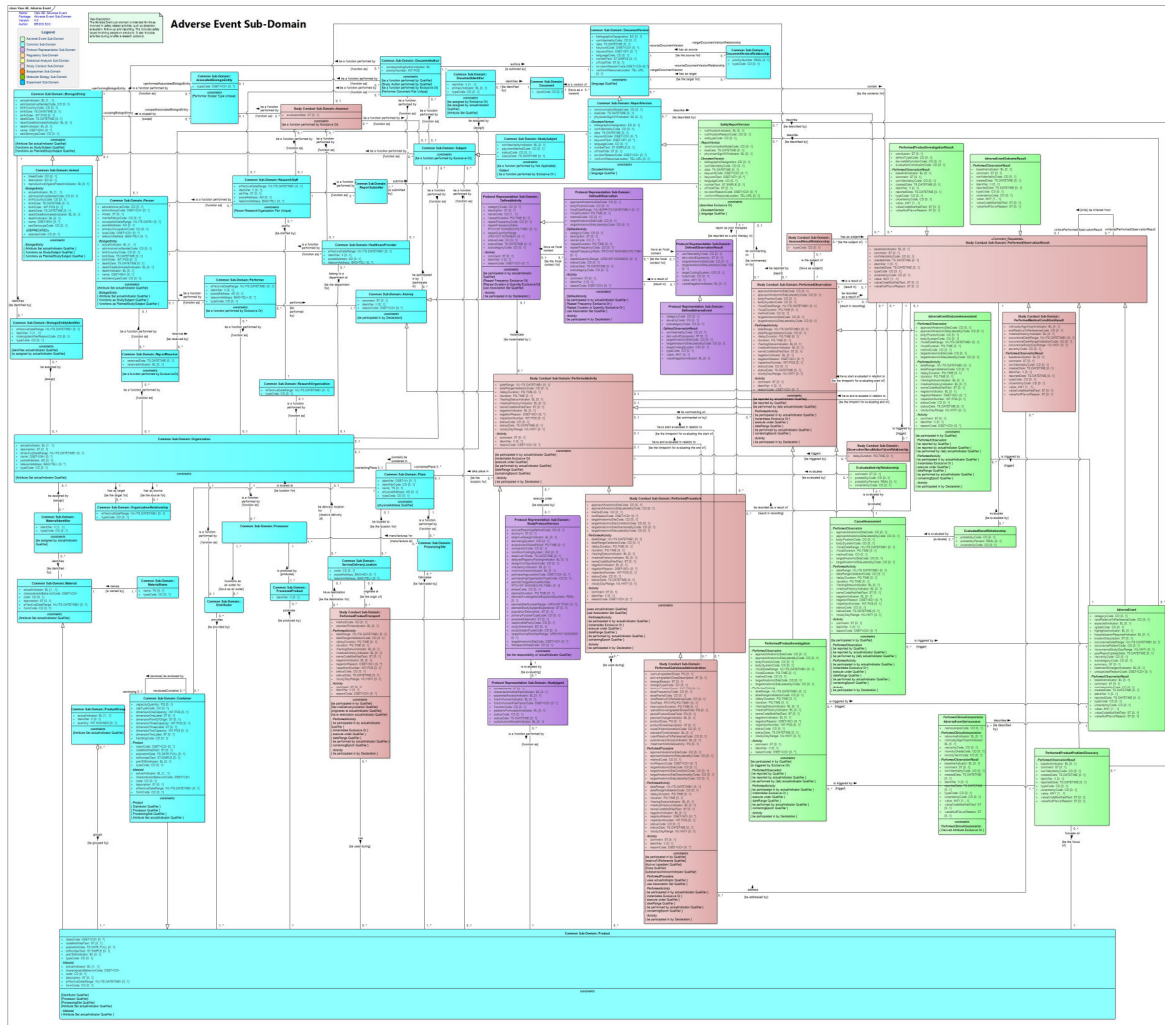
View Description:
This View shows the complete BRIDG Model (current Release) and specifically shows, for each class where it's applicable, the complete set of attributes for the class, partitioning the attributes as to whether they are "local" to the class or inherited from the class' super-type hierarchy.

Also the advice, that ...

Please Note: This comprehensive view of the BRIDG model is not intended for users to learn the model. It is created by the BRIDG modelers to ensure model integrity. It is highly recommended that you look at the sub-domain views to learn and review the concepts and relationships in the BRIDG model.

PhUSE Annual Conference 2015

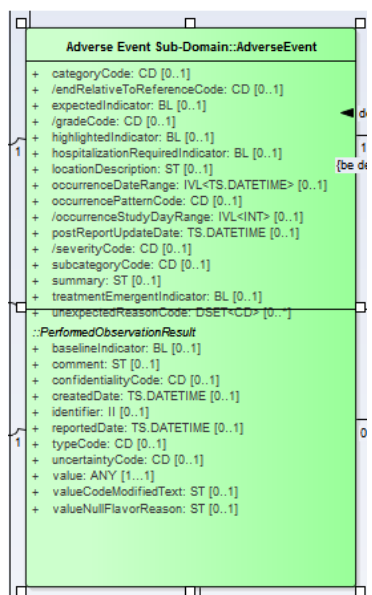
So, means, the model contains sub-views describing , actually 9, sub-domains. One example domain is the Adverse Event (AE) Sub-Domain. This AE sub-view contains all elements and objects defined as 'playing a role in the world of clinical adverse events'. Let's use this sub-view as our beginner's example.



As we can see here, this AE sub-view includes objects (classes). These classes are directly related to the nature of Adverse Events, but include classes and relationships from other models and sub-domains as well. The Adverse Event (AE) sub-domain contains objects from the Common Domain like "Person" and "Animal", from the protocol representation "DefinedAdverseEvent" and "DefinedObservation" for example and from the Study Conduct domain "PerformedActivity" and "PerformedSubstanceAdministration". The Adverse Event sub-domain itself defines the elements of an AE: sure, the class "AdverseEvent" itself but also classes like "AdverseEventSeriousness" and "AdverseEventOutcomeResult". All class names are nearly self-explaining the class meaning.

Every class definition in BRIDG contains the list of attributes. An attribute in a class is defined by its name and the data type. The class "AdverseEvent" contains, for example, an attribute called "reportedDate", which is defined as data type TS.DATETIME (TS=time stamp).

PhUSE Annual Conference 2015

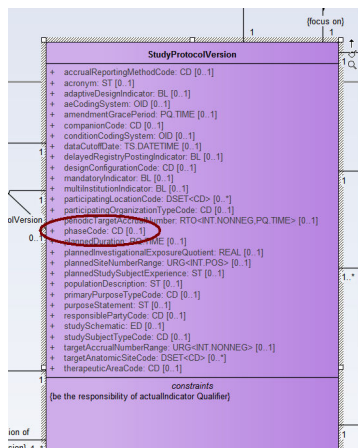


The possible data types are defined by HL7 Abstract Data Types. Examples are:

- ST= String, BL=Boolean, INT=Integer, TS.DATETIME=timestamp / date time
- EN= Entity Name, AD = Address
- CD = Concept Descriptor (code list)
- PQ = Physical Quantity

A Concept Descriptor (CD) is a reference to a concept defined in an external code system, terminology or ontology.

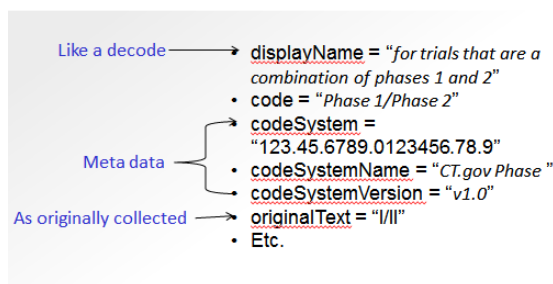
An example of a CD can be seen in the “StudyProtocolVersion” class and the “phaseCode” attribute:



The phaseCode is defined by clinicaltrials.gov and contains some of the following values:

- **N/A:** for trials without phases
- **Phase 0:** exploratory trials, involving very limited human exposure, with no therapeutic or diagnostic intent (e.g., screening studies, microdose studies).
- **Phase 1:** includes initial studies to determine the metabolism and pharmacologic actions of drugs in humans, the side effects associated with increasing doses, and to gain early evidence of effectiveness; may include healthy participants and/or patients
- **Phase 1/Phase 2:** for trials that are a combination of phases 1 and 2
- Etc.

Now, this should be stored in systems as



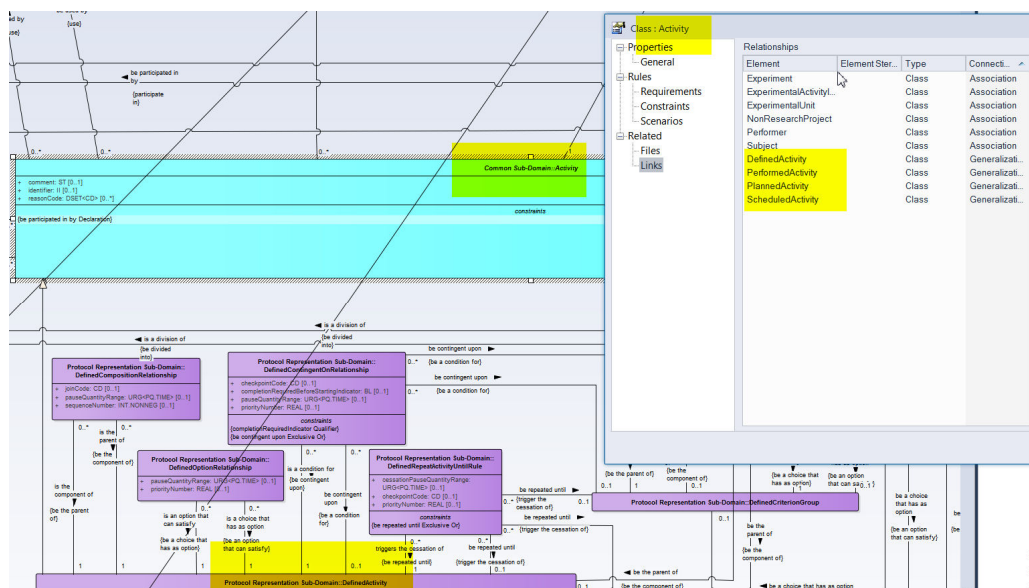
(This example has been taken from the official BRIDG training material).

Another key concept in BRIDG is the concept of 'ACTIVITY'.

An Activity is an action that can, in the context of a study or a post-marketing investigation, be defined, planned, scheduled or performed.

There are 4 activity subtypes:

- Defined activity
 - An activity that frequently occurs in studies (e.g. more than one time in more than one arm) and therefore is called out as a reusable template in a global library of activities outside the context of any particular study, and may be used in the composition of a defined study subject activity group. A defined activity is a "kind of" activity rather than an "instance of" an activity.
- Planned activity
 - An activity that is intended to occur or start at some point in the context of a particular study.
 - For example, Pregnancy tests are planned for StudySubjects who are females of childbearing potential.
- Scheduled activity
 - An activity that is anticipated to occur at some time in the future and has been assigned a time or date when that activity is to be performed.
 - EXAMPLE(S): An X-Ray scheduled for February 15 is in state "Scheduled." If John is unable to have the X-Ray on that date, the X-Ray would either be rescheduled (remain in "Scheduled" state, but "date" attribute would change) or moved to state "Canceled."
- Performed activity
 - An activity that is successfully or unsuccessfully completed.
 - EXAMPLE(S): CBC performed on a specific StudySubject on a given day.



PhUSE Annual Conference 2015

If you are familiar with HL7 'moodCode'(s) then BRDIG activities are following the same idea.

Finally, the relation between classes is classified by constraints and multiplicities (0..1:1, 1:n, m:n etc). Also, some classes are defined as 'abstract', where another class inherits from. An example from the real world would be the abstract class 'car' with its real implementation as a 'passenger car' or a 'lorry'. It inherits all attributes from the abstract class (example: coachwork, steering wheel) and can add their own implementation attributes, an attribute which is just important for this class (example for a passenger car: backdoors)

MAPPING EXAMPLE

To understand how the classes and attributes are mapped you have 2 possibilities: Start in the model to see to which target models the BRDIG classes and attributes has been mapped to, or start in the mapping spreadsheet to start with a target model and get the classes and attributes from BRDIG. The mapping spread sheet is part of the BRDIG package.

BRDIG is mapped against the following models or targets, parts of the 'Biomedical Research' domain:

- LSDAMv2.2.3Plus
- SDTM IGv3.1.3x
- PGxIG r1.0
- CTRPv3.8
- HCTv1.0 2012
- FDA HL7 SD SD DSTU2012
- Statisticsv1.0
- CTRv1.0
- CTRRr3
- NCI CRF Round 3
- NCI CRF Rounds 1&2
- CTSA-HSDBv1.0
- HCTv1.0 Apr2011
- SDTM IGv3.1.2
- CDASHv1.1
- ICSRr2
- C3PRv2.9
- PSCv2.6
- caAERSv2.2
- LabViewer2.2
- CTR&R R2
- COPPA
- Lab
- PSC
- AE
- CTR&R R1 Comprehensive
- SDTM IG
- TDM
- Study Design RMIM
- HL7SP
- RPS1
- C3PR
- CTOM

For every target a single worksheet exists and defines the single classes and attributes.

PhUSE Annual Conference 2015

	A	B	C	D	E	F	G	H	I
1	CTDM 0.53			BRIDG 3.0					
2	ClassNam e	AttributeNam e	Definition	BRIDG Class	BRIDG Attribute	Map:CTDM	Path / Where Clause	ActionItem	Comments
93	ConceptDes criptor	code		Implementation Specific					
94	ConceptDes criptor	codeSystem		Implementation Specific					
95	ConceptDes criptor	codeSystemNa me		Implementation Specific					
96	ConceptDes criptor	codeSystemVer sion		Implementation Specific					
97	ConceptDes criptor	displayText		Implementation Specific					
98									
99	DeathSum mary	autopsiedIndicat or		Derived			Subject > Activity WHERE DefinedActivity.nameCod e = "Autopsy"		
							BiologicEntity.deathDate > Subject > Activity.PerformedActivit y.PerformedObservation > AssessedResult > PerformedObservationRe sult WHERE		

To understand the syntax used in the sheet, a Mapping Path definition is also available.

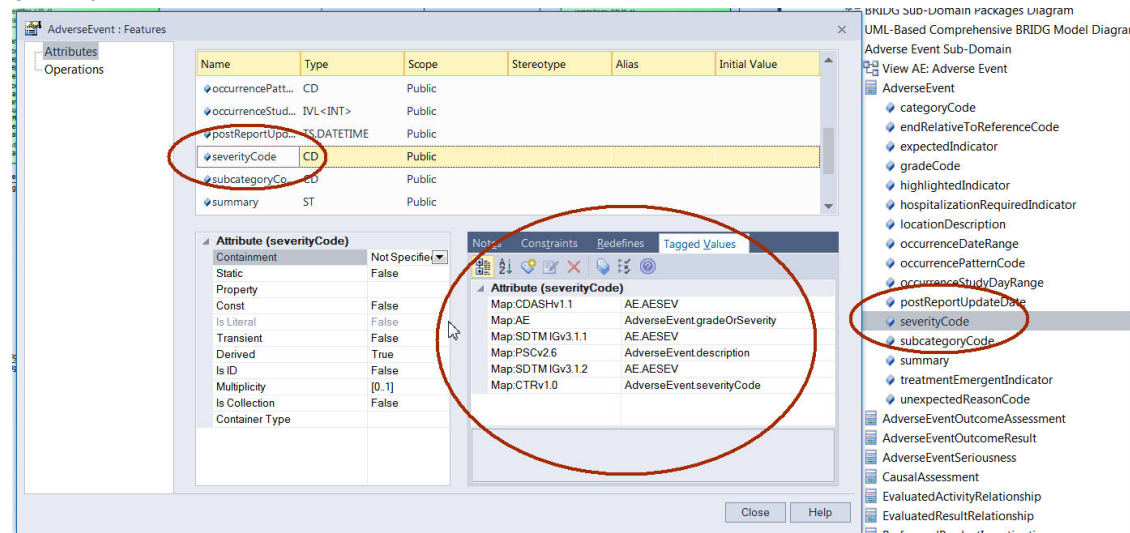
	A	B	C	D	E	F	G	H	I	J	K	L
1	How to Read a Mapping Path											
2	Disclaimer:  Mapping Path conventions have evolved throughout the many versions of BRIDG. Not all mapping paths in all tabs of this spreadsheet may adhere precisely to the conventions as stated.											
3	Mapping Path columns are included in each tab of this spreadsheet. Mapping paths illustrate how the semantics of a specific source model/project element are harmonized with BRIDG. These paths contain the BRIDG classes, attributes, associations and data types that are used in capturing the semantics of the source model/project element. Mapping paths may be straight-forward or more complex depending on the semantic meaning of the element being harmonized.											
4	The starting point of each mapping path is always a BRIDG class. Class names are represented using upper case CamelCase. This starting point (anchor point) represents the context of the source model element. For example, are we talking about a Person, Place or Activity? The end of the mapping path refers to the specific attribute in BRIDG that captures the semantics of the source element. For example, name, date, value, etc. Attribute names are separated from the class name with a "." and are always represented using lower case camelCase. Mappings for all attributes in a given class usually start at the same anchor point to give a consistent context to related data items.											
5	In some cases the mapping path may contain a "WHERE" clause - this means a specific condition applies to the logic of the mapping.											
6	A period (.) is used to separate the following: classes in an inheritance path (i.e., Activity.PerformedActivity), class name from attribute name (i.e., Activity.comment), and attribute from data type (i.e., Activity.comment.ST).											
7	Data type names are in represented in ALL CAPS and have may attributes that map to different data type attributes of a BRIDG class attribute. For example, Dose.amount (from a source model) may map to PerformedSubstanceAdministration.productDose(PQ).value (in BRIDG) and Dose.unit (from a source model) may map to PerformedSubstanceAdministration.productDose(PQ).units (in BRIDG).											
8												
9	Mapping Path Legend:											
10	Symbol	Meaning						Example				
11	" > " (with spaces)	Class to class across an association						Class1 > Class2				
12	" ." (no spaces)	Class to attribute						Class.attribute				
13	" () " (no spaces)	Attribute to data type (DT)						Class.attribute(DT)				
14	" ." (no spaces, outside parentheses)	Data type (DT) to data type attribute (dtAttr)						Class.attribute(DT).dtAttr				
15	" => " (no spaces, inside parentheses)	Data type (DT1) to data type (DT2) constraint						Class.attribute(DT1=>DT2)				
16	" [] " (text goes inside brackets)	Annotation						[subject visit]				

The opposite way is to start in the BRIDG model. For elements which are mapped a TAG has been created in Enterprise Architect. As an example, multiple TAGs have been created for the "severityCode" in Adverse Events

PhUSE Annual Conference 2015

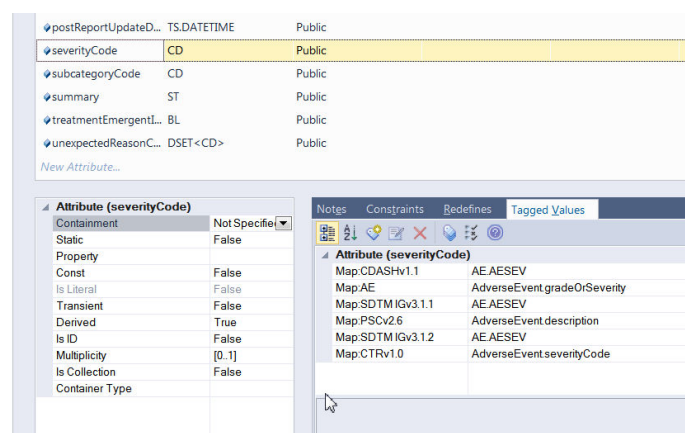
to define the mapping to the target model. In the following screenshot you can see that the “severityCode” has been mapped to:

CDASHv1.1
AE
SDTMIGv3.1.1
PSCv2.6
SDTMIGv3.1.2
CTrv1.0



However, one problem is that some mapping targets are missing. In this particular example we see in the mapping spreadsheet, that “severityCode” is mapped against SDTM IGv3.1.3, but this target version is not listed in the TAG session in Enterprise Architect:

Study Data Tabulation Model Implementation Guide (3.1.3) to BRIDG 4.0 beta (1649) Mapping, 2013-Nov									
Source Specification					Mapping				
Mapped Group Name	Mapped Element Name	Element Type	Card	Definition and Semantics	Custom	Type	Review by	Status	Comments / Issues / Rationale
AE	AESEV	Attrib	Char	0..1 PCOE: Record Qualifier. NOTE:SDSEDESCRIPTION: The severity or intensity of the event. Examples: MILD, MODERATE, SEVERE.	Events AE 26	Supported		Applied	



WHAT'S NEW IN VERSION 4

PhUSE Annual Conference 2015

With the release of BRIDG 4.0, the scope of BRIDG is now officially expressed as “basic, pre-clinical, clinical, and translational research and associated regulatory artifacts. 4.0 is the first release of BRIDG as a translational research model and therefore some of the newer sub-domains are not fully fleshed out as yet.

In Version 4 the LS DAM (Life Sciences Domain Analysis Model) has been added. It is a Hypothesis driven basic and pre-clinical research model and supports discovery sciences as well. Over years it has been developed in parallel to BRIDG. The main stakeholder was the NCI, working closely with the HL7 Clinical Genomics workgroup.

Further new in Version 4 are 3 domains:

- Biospecimen
- Experiment
- Molecular Biology

Version 4 added also related SDTM changes: trial design changes, inclusion of the therapeutic area Oncology and the content published in the SDTM amendment 1. Also, SDTM PGx has been added – Pharmacogenomics and Pharmacogenetics.

CONCLUSION / SUMMARY

BRIDG is the canonical model for the Life Sciences Domain. It uses UML as the ‘language’ to define the model.

For a newbie it can be overwhelming the first time, but diving deeper it is not so complicated to start and understand the model. Once the concept of classes and attributes and the grouping into sub-domains has been understood it can be simple to understand the mappings to different target models like CDISC SDTM and ICSR r2 as well.

With Version 4 the NCI LS DAM has been added and BRIDG is awaiting fully ISO approval and publication over the next few months.

REFERENCES

BRIDG home: <http://bridgmodel.nci.nih.gov/>

BRIDG history: http://bridgmodel.nci.nih.gov/about_bridg/background

HL7: <http://www.hl7.org/special/committees/bridg/index.cfm>

Definition of CD – Study Phases on clinicaltrials.gov: <https://prsinfo.clinicaltrials.gov/definitions.html>

Enterprise Architect, Sparxsystems: <http://www.sparxsystems.com/>

UML Tutorial: http://sparxsystems.com.au/resources/uml2_tutorial/

ISO TC-215:

http://www.iso.org/iso/home/store/catalogue_tc/catalogue_tc_browse.htm?commid=54960&published=on&development=on

GLOSSARY

BRIDG - Biomedical Research Integrated Domain Group

DAM – Domain Analysis Model

UML – Unified Modelling Language

NCI – National Cancer Institute

FDA – US Food and Drug Administration

HL7 – Health Level 7

CDISC – Clinical Data Interchange Standards Consortium

ACKNOWLEDGEMENT

Acknowledgement goes to the CDISC BRDIG team, the NCI engagement as well to HL7.

<http://www.cdisc.org/bridg>

<http://bridgmodel.nci.nih.gov/>

<http://www.hl7.org/special/committees/bridg/index.cfm>

CONTACT INFORMATION

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

Jörg Dillert
Oracle Deutschland B.V. & Co. KG
Schiffbauergasse 14
14467 Potsdam
Work Phone: +49-331 2007388
Email: joerg.dillert@oracle.com
Web: www.oracle.com

Brand and product names are trademarks of their respective companies.

Date of this article: 17th August 2015