



# CDISC 2020 Europe Interchange

Virtual Conference

1-2 April 2020





# Beyond Biomedical Concepts: How Study Management Concepts Fill the Gaps

Presented by Dr. Philippe Verplancke  
Head of Business Development, XClinical GmbH

1 April 2020



# Speaker Bio



- Dr. Philippe Verplancke is an electrical engineer and earned a Ph.D. in nuclear fusion research from the university of Ghent in Belgium.

After working at McKinsey and a startup company on mobile patient data collection, he founded XClinical in 2002, an EDC / IWRS / eCOA / eTMF vendor leveraging CDISC XML standards.

He is an early CDISC contributor and a former ODM/define.xml trainer. As XClinical's Head of Business Development, he continues to participate in technical discussions, for example in the CDISC 360 project.

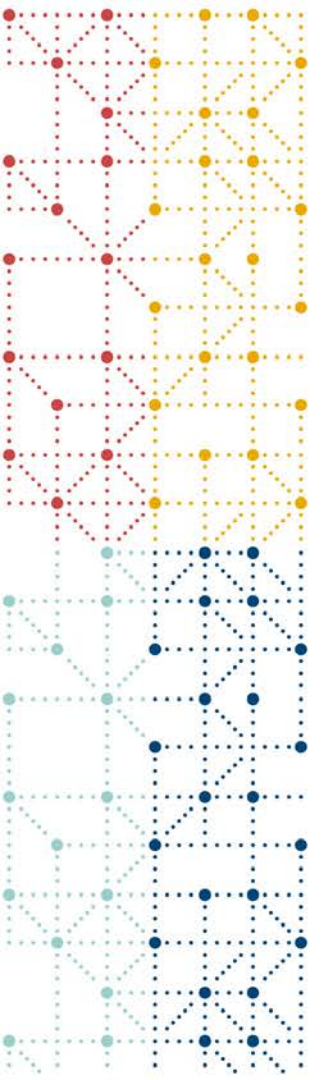
# Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
- *Philippe Verplancke is the founder and Head of Business Development of XClinical, an EDC / IWRS / eCOA / eTMF vendor*



## Agenda

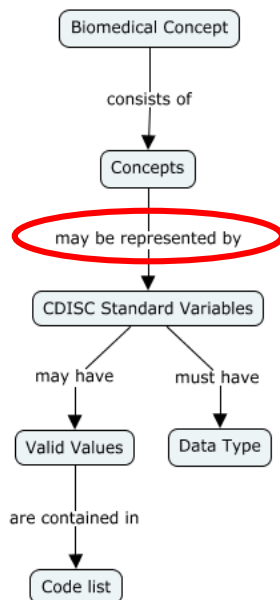
1. The No-Mapping approach with Biomedical Concepts
2. Study Management Concepts
3. One more thing... Version Control !



## The No-Mapping Approach with Biomedical Concepts

# What is a biomedical concept?

## Biomedical Concept Overview



A Biomedical Concept is a unit of knowledge created by a unique combination of characteristics. (ISO 11179)

A Concept is a characteristic of a Biomedical Concept that can be described independently of any particular representation. These concepts have c-codes from NCI EVS.

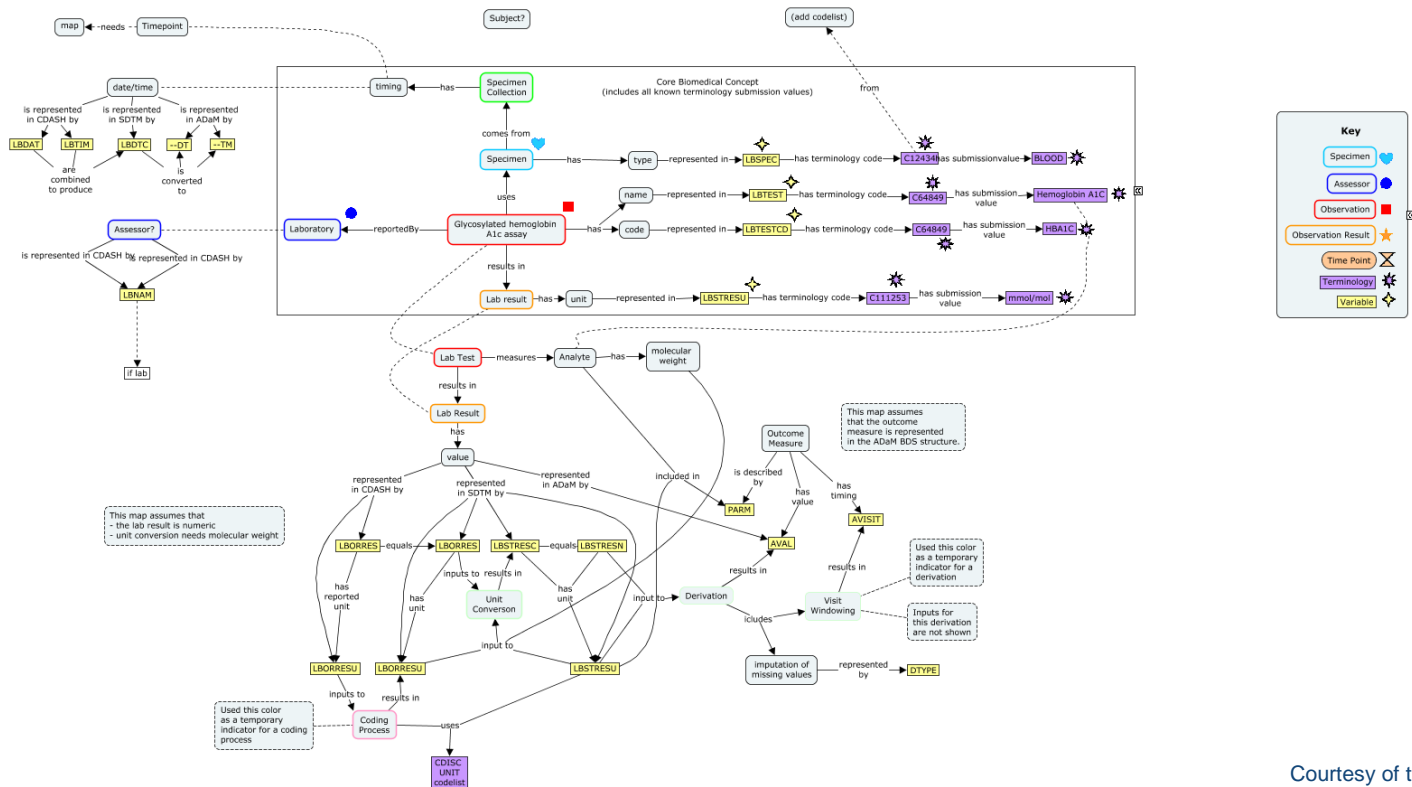
A CDISC Standard Variable is a named data element that contains clinical research data.

Valid values are the permissible codelist items for a variable.  
A data type is a classification identifying one of various types of data, such as real, integer or Boolean.

A code list is a set of valid values for a variable.

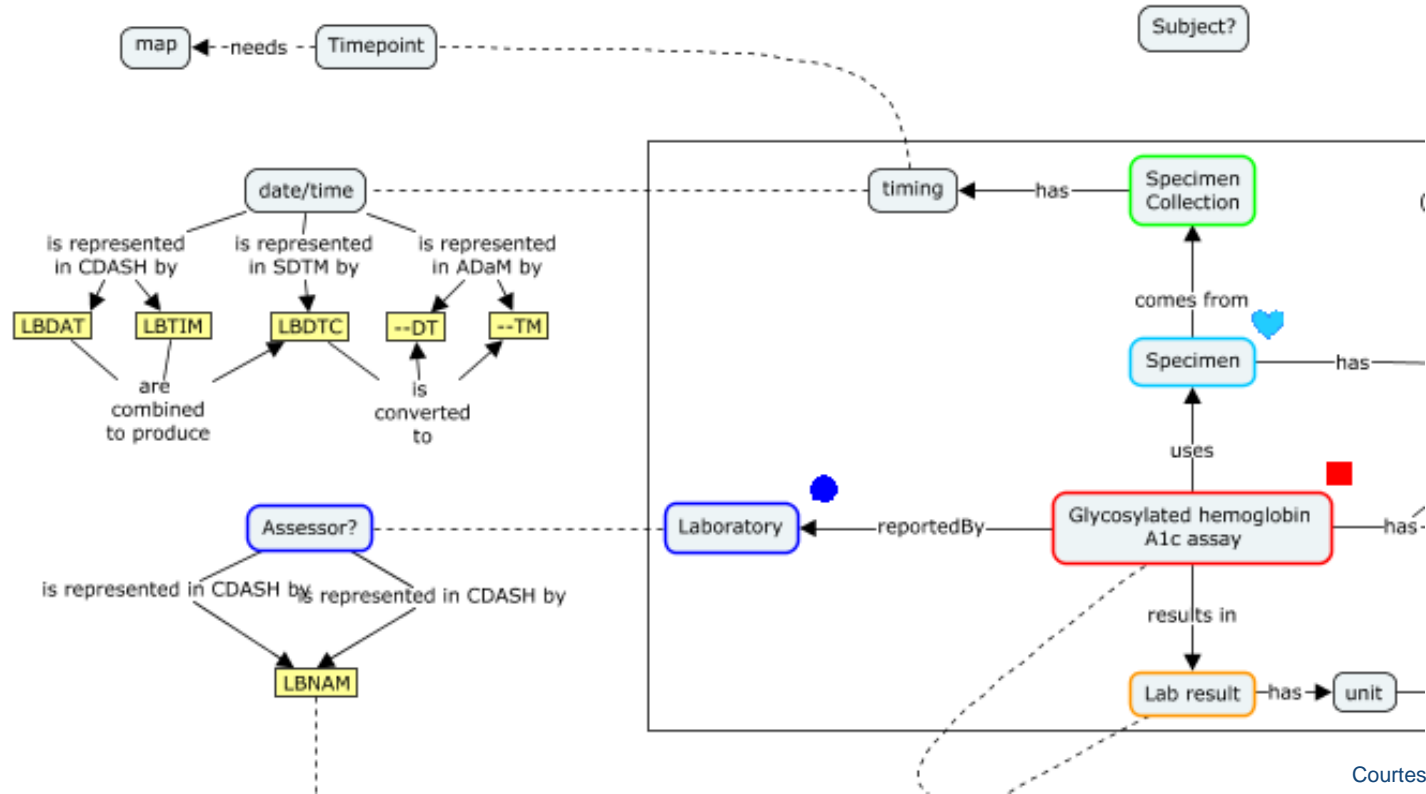
Courtesy of the CDISC 360 team

# Example of a Biomedical Concept including its representations in CDASH, SDTM and ADaM





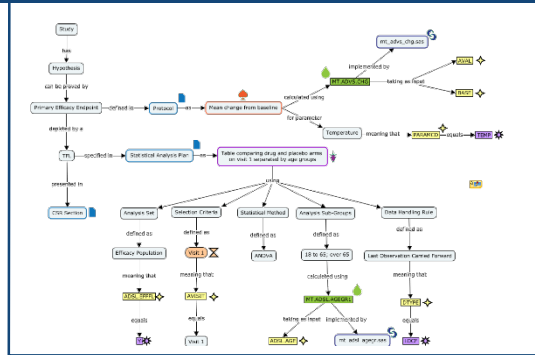
# Zoom in: Date/Time in CDASH, SDTM, ADaM



Courtesy of the CDISC 360 team

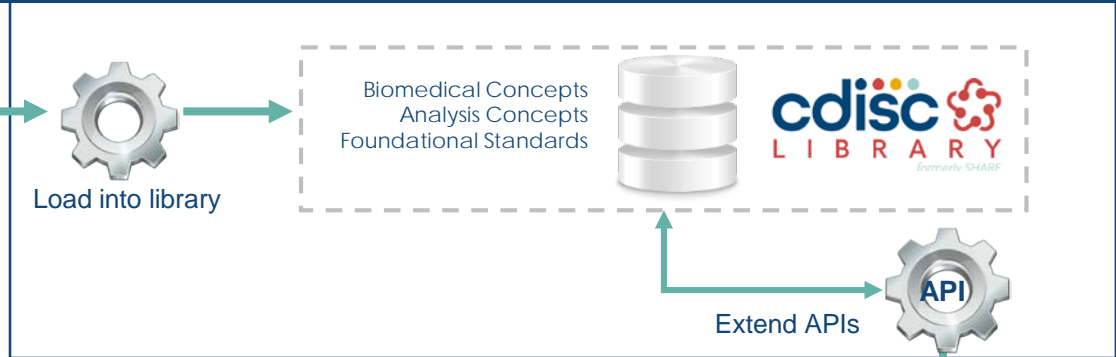
# The Challenge: how to store biomedical concepts in a machine-readable and machine-computable way?

Bio Concepts = *ENHANCED CDISC STANDARDS*



CDISC 360 Workstream 1

New CDISC Library → *PUBLISH STANDARDS IN MACHINE READABLE FORM*



CDISC 360 Workstream 2

1. Transform the Concept Map into a standard meta-model using ISO 11179
2. Transform the meta-model into a formal ontology
3. Validate the concepts in the formal ontology
4. Export the formal ontology for consumption in the CDISC Library

**Any  
Software  
System**



Courtesy of the CDISC 360 team

# Ontology Development 101 (Natalya F. Noy & Deborah L. McGuinness, Stanford University)

1. There is no one correct way to model a domain; there are always viable alternatives. The best solution almost always depends on the application that you have in mind and the extensions that you anticipate.
2. Ontology development is necessarily an iterative process
3. Concepts in the ontology should be close to objects (physical or logical) and relationships in your domain of interest. These are most likely to be nouns (objects) or verbs (relationships) in sentences that describe your domain.

CDISC 360 will need to make decisions that will have a long-term impact

[https://protege.stanford.edu/publications/ontology\\_development/ontology101-noy-mcguinness.html](https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html)

# The No-Mapping Approach with Biomedical Concepts

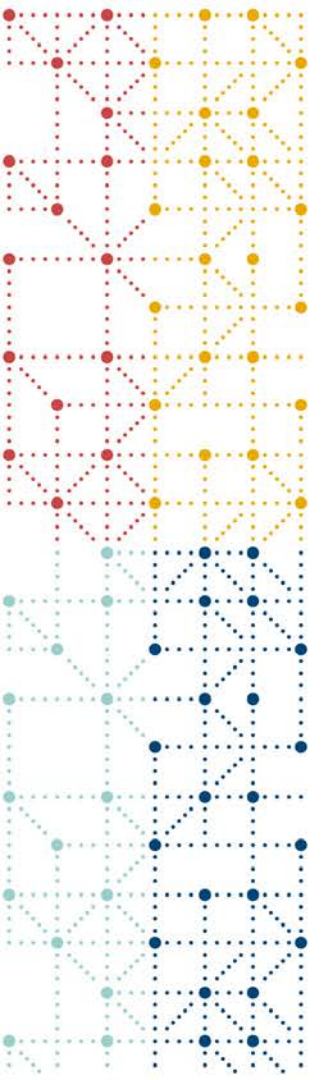
1. Select which biomedical concepts are needed for a study, including analysis concepts and endpoint concepts
2. Collect data
3. Done! 😊

If CDISC is successful  
creating a formal ontology...

- when reading the machine-readable representation of the biomedical concepts, EDC systems will know when\* and how to collect the data, ETL systems will know how to derive\* and tabulate the data, Analysis systems will know how to analyze the data

\* well, not quite....

we need to add Study Management Concepts to the Biomedical Concepts



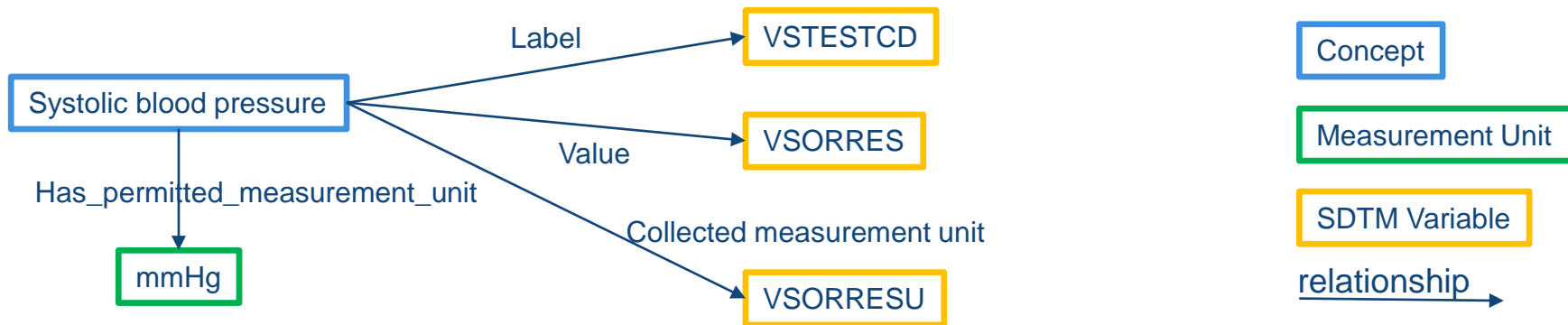
# Study Management Concepts

# Another Disclaimer

- Study Management Concepts are an idea proposed by the author
- This is not an official proposal from CDISC nor from the CDISC 360 team
- If we are serious about the “No-Mapping” approach and if we want to avoid creating yet another set of scripts (just hidden deep inside fancy graph databases), we need to model Study Management Concepts to tell computer systems when to collect data, how to interpret the data, etc.

# Concepts versus Scripts

## an example from the CDISC history books



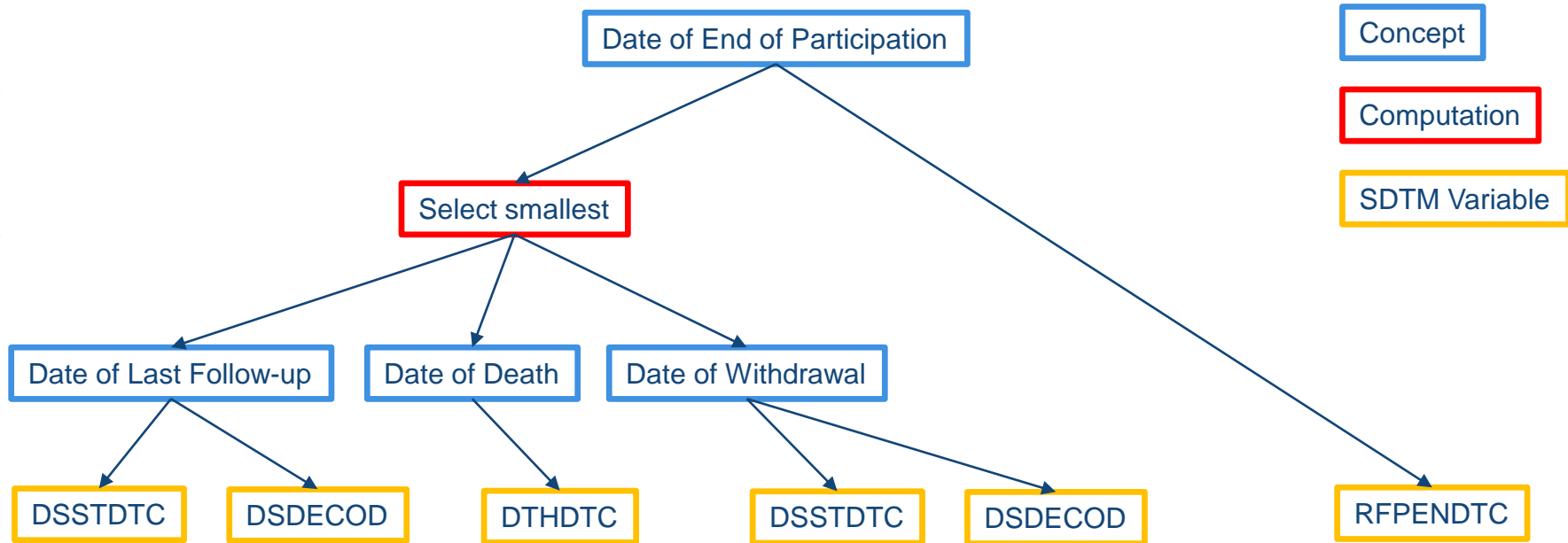
18 years ago, CDISC has already provided a machine-readable representation of these concepts in ODM:

```
<ItemDef OID="VS.SYSBP">  
  <MeasurementUnitRef OID="mmHG">  
</ItemDef>
```

Computer systems can compute SDTM variables by letting the tabulation system and the data collection system point to the same concepts.

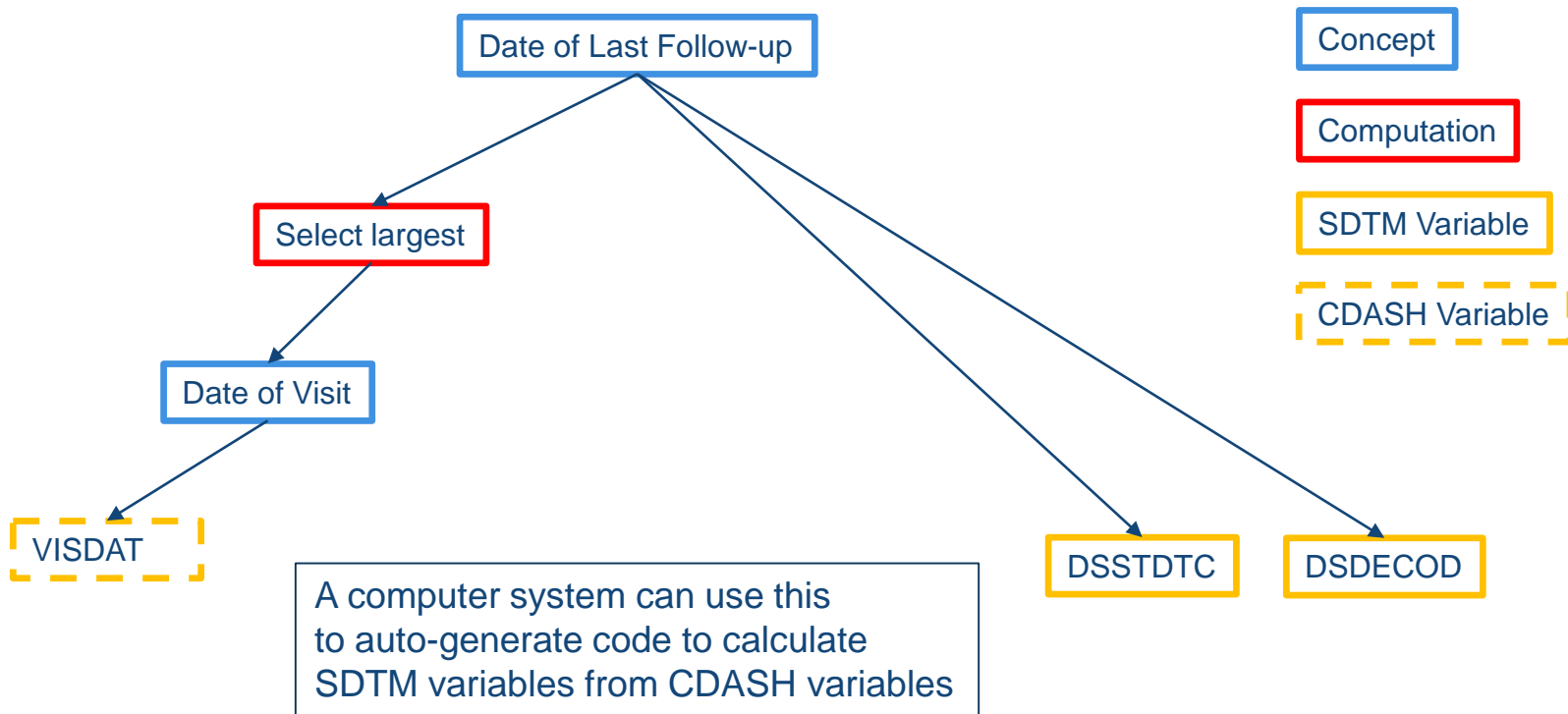
There is no need to provide scripted metadata like in define.xml 2.0 (VSORRESU = "mm.HG" where VSTESTCD="SYSBP")

# Study Mgmt Concept Example 1: Date of End of Participation



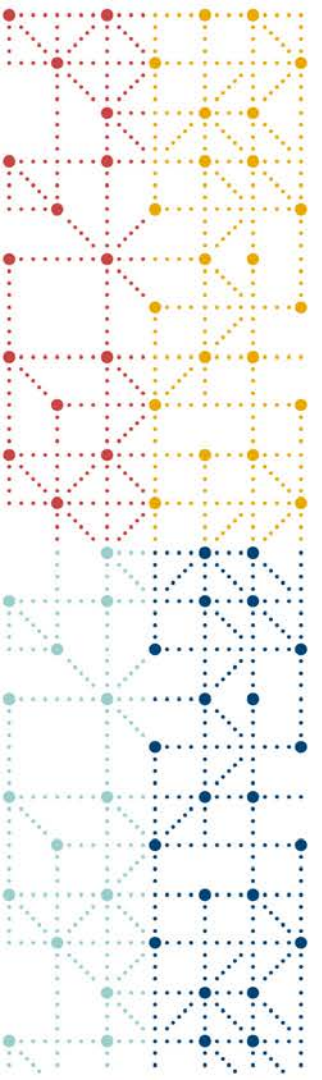


# Study Mgmt Concept Example 2: Date of Last Follow-up



# What's the point? Is this not just mapping again?

- Working with concepts avoids ambiguity
- Working with concepts reduces redundancies
- Pointing all computer systems to the same machine-readable concepts guarantees consistency, avoids gaps, avoids human error
- Working with concepts helps (forces) CDISC as an organization to build machine-readable standards with exact semantics (meaning, using formal ontologies), reducing wiggle-room for interpretation



**One more thing... Version Control !**

# One more thing ... Version Control !

- Each concept, each variable, each codelist item, etc., should have a unique identifier that does not change when the label changes but only when the semantics are different
- When creating a study amendment or when reviewing newer versions of CDISC standards and their impact on the study, a data manager needs to have a transparent overview of the version history of metadata
- CDISC currently only provides version control on a collection of metadata (stored in one file), not on each element of metadata.
- Here is an example of how it could be improved in analogy to software version control:



AE Form version 2 references all the elements with orange color or an orange border.



**Thank You!**

[info@xclinical.com](mailto:info@xclinical.com)

[philippe.verplancke@xclinical.com](mailto:philippe.verplancke@xclinical.com)

Please visit our virtual XClinical booth on Slack!

