



# CDISC Italian User Network 2020

Milan, Italy | 7 October 2020





S-CUBED

Decomposing SDTM – changing perspective

Johannes Ulander  
Italian CDISC User Network  
Milano  
2020-10-07

# Agenda

- Decompose SDTM
- Perspectives on data

# Person and cat



Name: John  
Age: 32  
Gender: Male

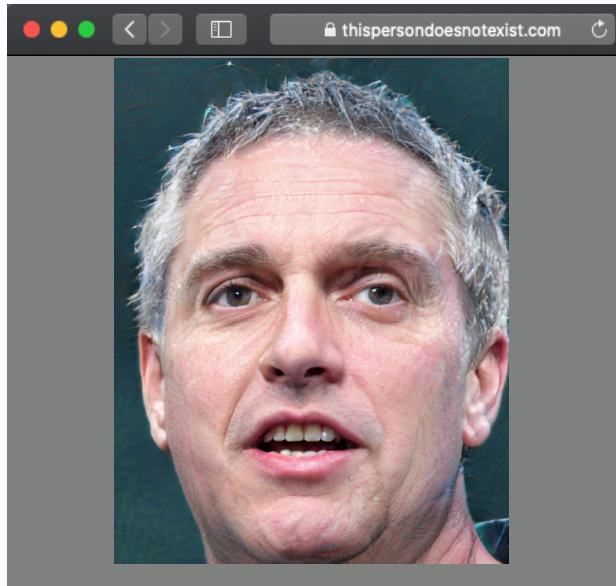
Name: Jazzy  
Age: 6  
Gender: Female





# Person and cat – or is it?

<https://thispersondoesnotexist.com>



Name:  
Age:  
Gender:

Name: Jazzy  
Age: 6  
Gender: Female



Yes, this is our cat

Imagined by a GAN ([generative adversarial network](#))  
[StyleGAN](#) (Dec 2018) - [Karras et al.](#) and Nvidia  
[Original GAN](#) (2014) - [Goodfellow et al.](#)  
Don't panic. Learn about [how it works](#).  
Help me figure out what was learned [here](#).  
[Help](#) this AI continue to dream  
[Another](#) | [Save](#) • [Cats](#) | [Articles](#) | [TV Friends](#) - [Office](#) | [x](#)

<https://arxiv.org/abs/1406.2661>

<https://arxiv.org/abs/1812.04948>

36

42

<i>AAGE</i>
36
42

## DM – Specification for the Demographics Domain Model

dm.xpt, Demographics — Version 3.2. One record per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify submissions involving the primary compound identifier formed
INVID	Investigator Identifier	Char		Record Qualifier	An identifier to describe the SITEID. Not needed if SITEID
AGE	Age	Num		Record Qualifier	Age expressed in AGEU. May not be available if BIRTHDTC
AGEU	Age Units	Char	(AGEU)	Variable	Units associated with AGE.

SDTM IG

STUDYID	USUBJID	INVID	AGE
STUDY1	S01	I01	36

# Study Data Reviewer's Guide

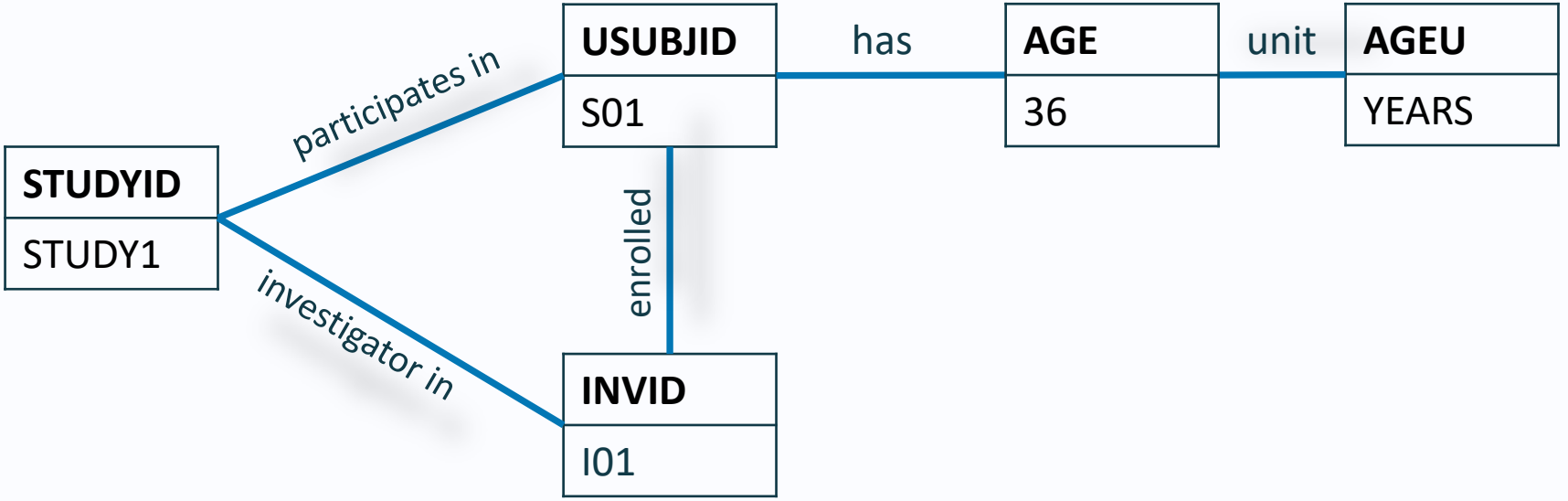
SDRG, Inc.

Study SDRG-001A

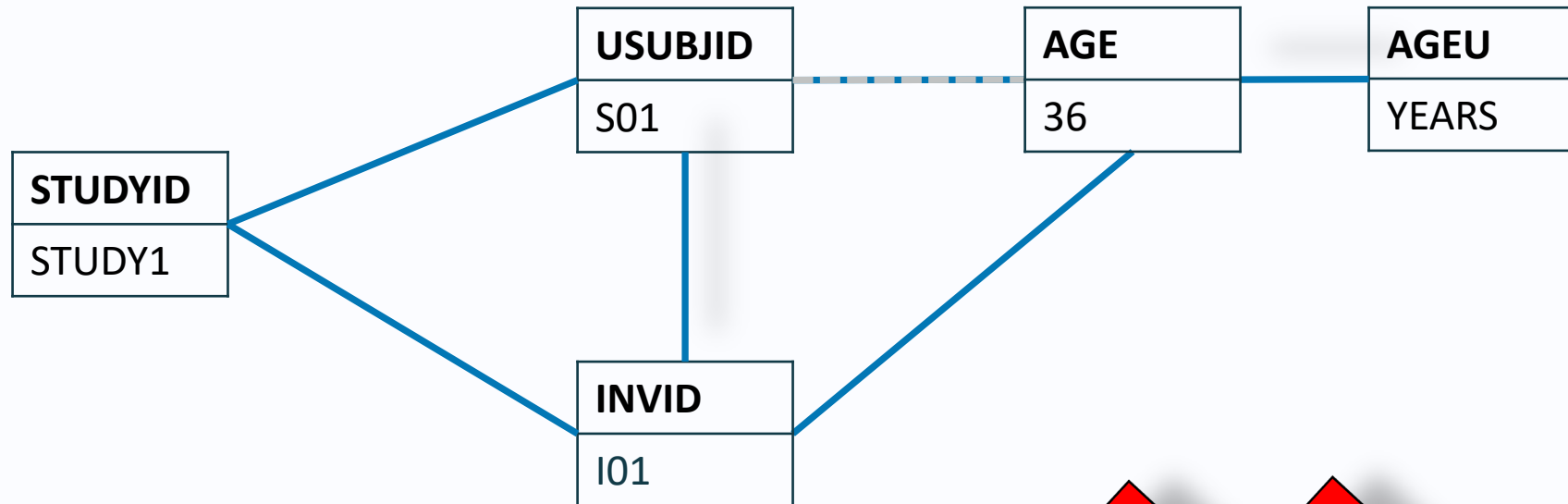
Version 2014-08-06

### 3.3.1. DM – Demographics

To make sure that subjects did not get biased by the investigators age, variable AGE is used for the age of the investigator.

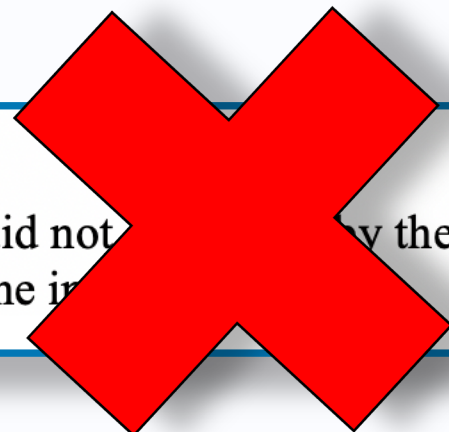


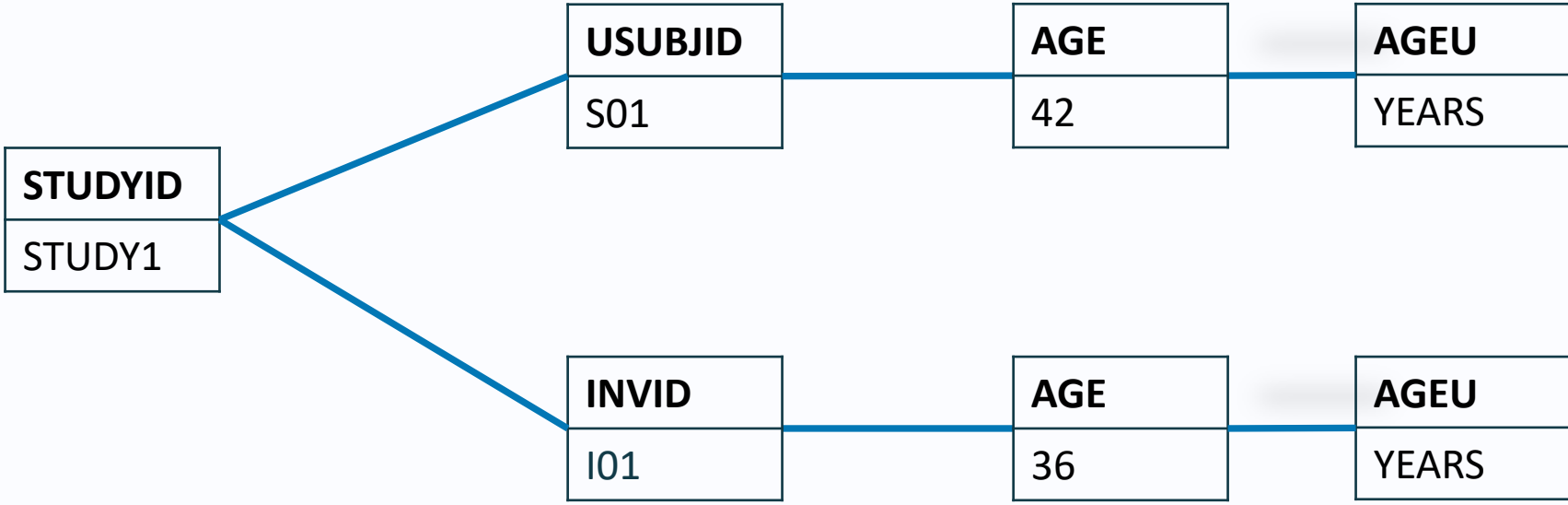




### 3.3.1. DM – Demographics

To make sure that subjects did not know the investigators age, variable AGE is used for the age of the investigator. Variable AGEU is used for the investigator's age in years.





**CDER Common Data Standards Issues Document**  
(Version 1.0/May 2011)

**CDER Common Data Standards Issues Document**  
(Version 1.1/December 2011)

**STUDY DATA  
TECHNICAL CONFORMANCE GUIDE**

*Technical Specifications Document*

**February 2014**

**December 2014**

**October 2015**

**March 2016**

**July 2016**

**October 2016**

**November 2016**

**March 2017**

**October 2017**

**October**

**March 2018**

# Our data is incomplete

Scattered

Versions

Lack precision

SDTM IG

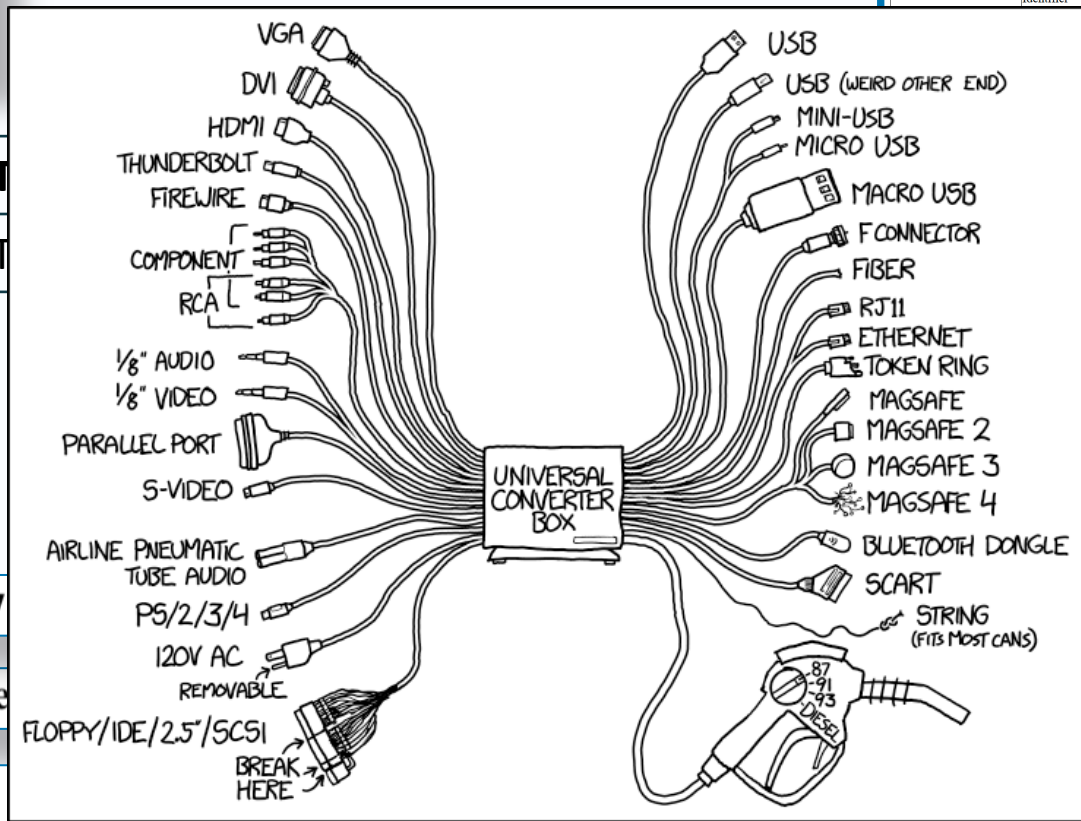
**DM - Specification for the D**

dm.xpt, Demographics — Version 3.2. One

Variable Name	Variable Label	Type
USUBJID	Unique Subject Identifier	Char
INVID	Investigator Identifier	Char

Version	Related
SDTMIG v3.3 20 November 2018	SDTM v1.7 SDTMIG for Medical Devices v1.1 SDTMIG-PGx v1.0 SDTMIG-AP v1.0 Conformance Rules v1.1 for SDTMIG v3.2 and v3.3
SDTMIG v3.2 26 November 2013	SDTM v1.4 SDTMIG for Medical Devices v1.0 SDTMIG-AP v1.0 Conformance Rules v1.1 for SDTMIG v3.2 and v3.3
SDTMIG for Medical Devices v1.1 1 February 2019	SDTM v1.7
SDTMIG for Medical Devices v1.0 23 January 2012	SDTM v1.4
SDTMIG-PGx v1.0 1 June 2015	SDTM v1.5
SDTMIG-AP v1.0 12 December 2013	SDTM v1.4 Conformance Rules v1.1 for SDTMIG v3.2 and v3.3

	Core
Applications or data could be a UBJID.	Req
In addition to	Perm
FIDIC, but concerns	Exp
	Exp

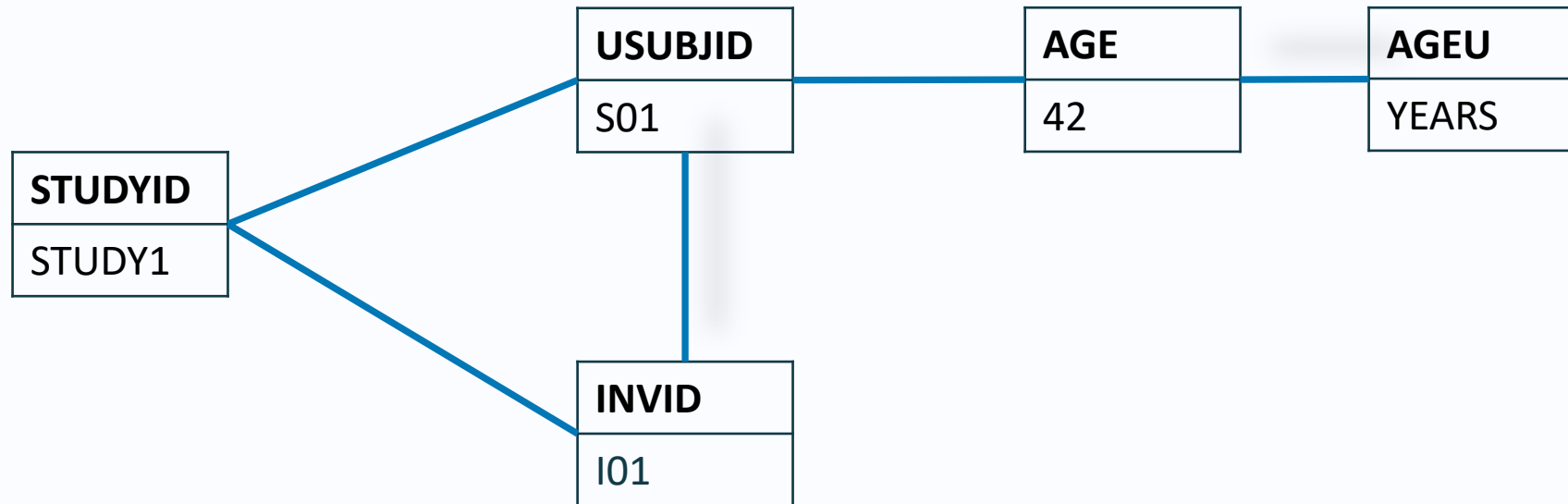


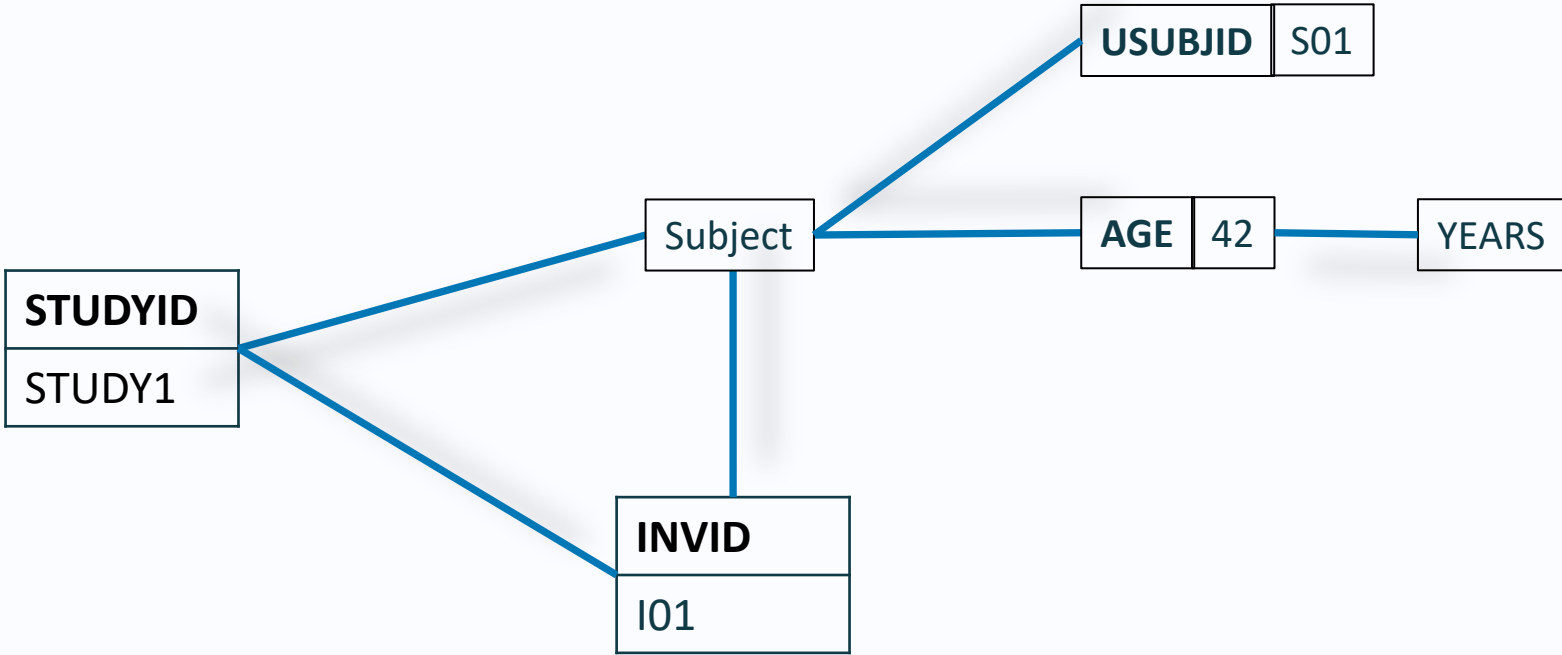
**Study Data Reviewer's Guide**

SDRG, Inc.  
Study SDRG-001A  
Version 2014-08-06

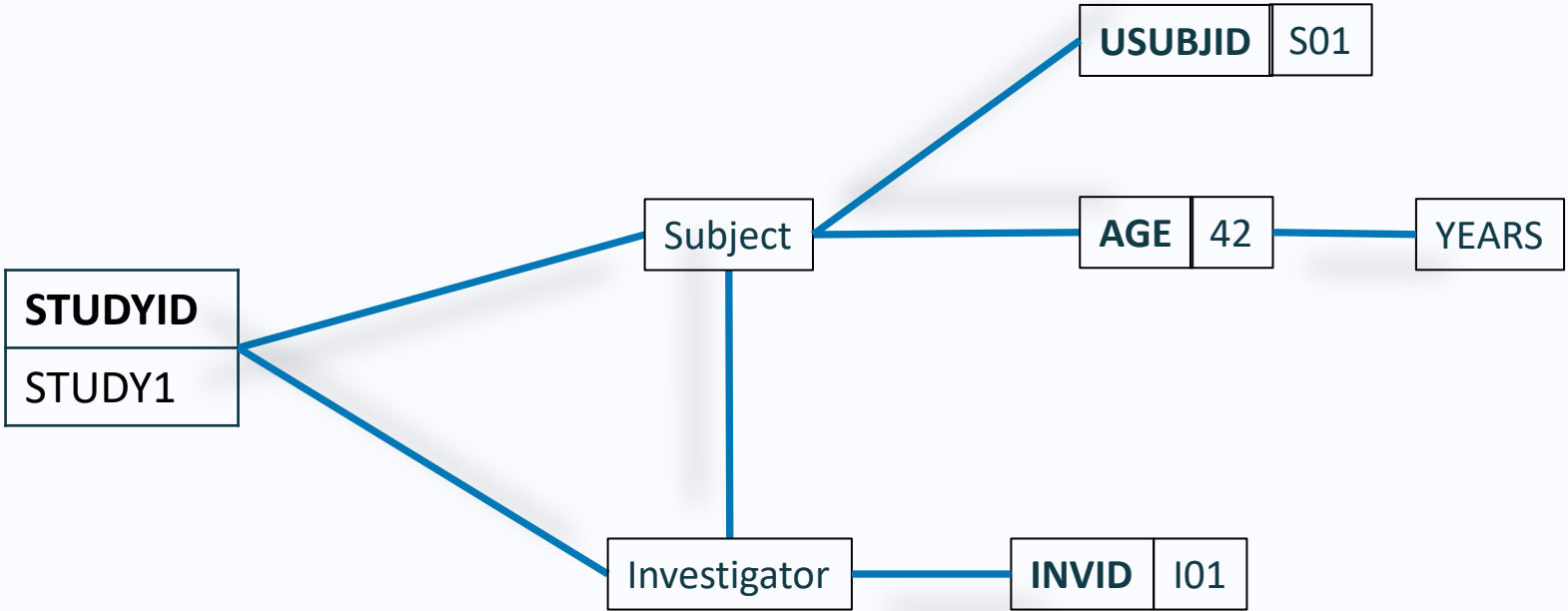
Source: <https://xkcd.com/1406/>

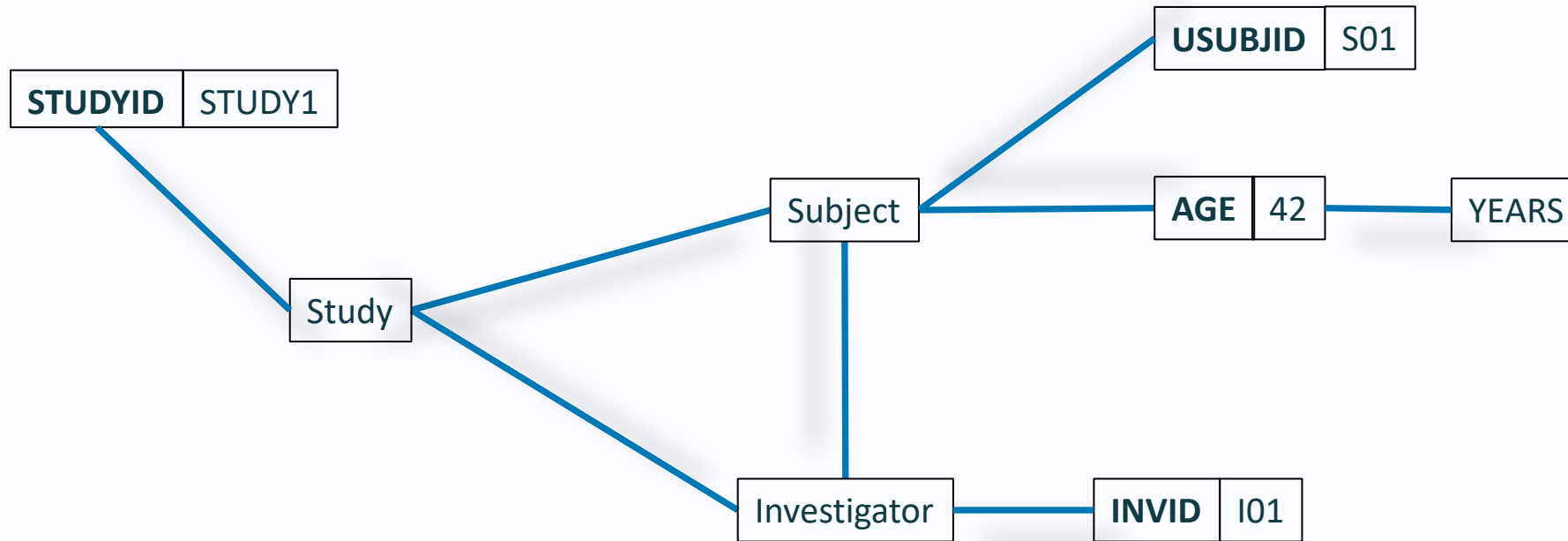
# I want to decompose some more

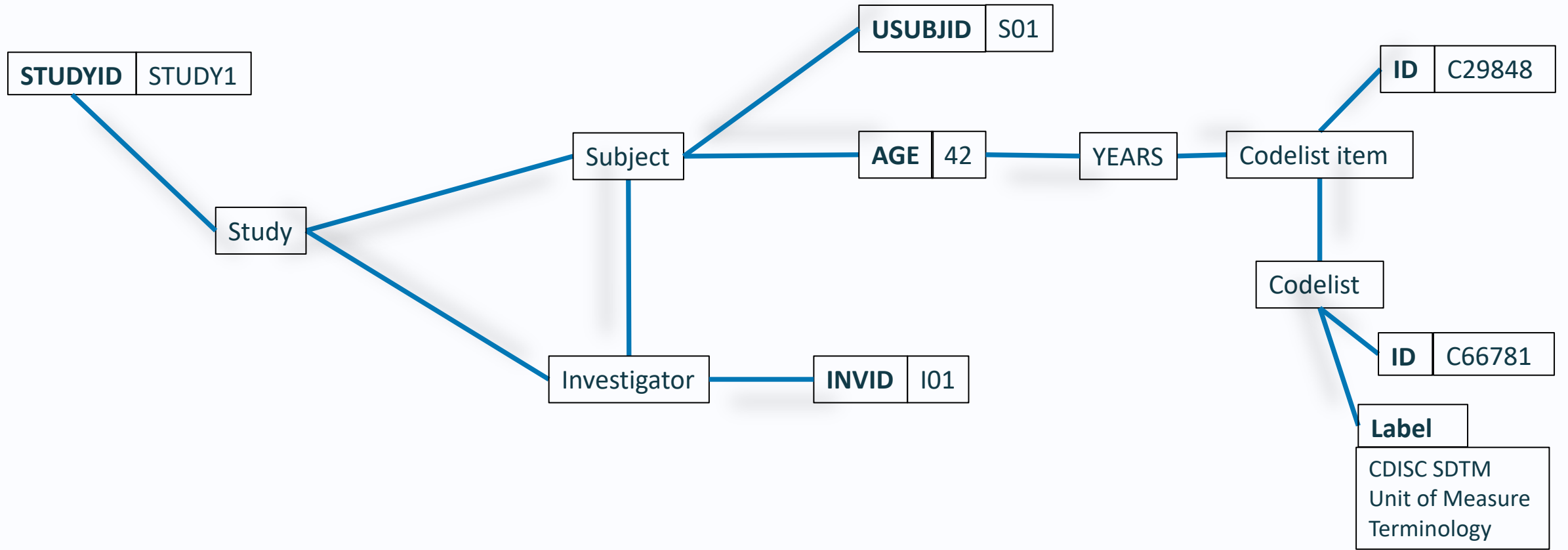


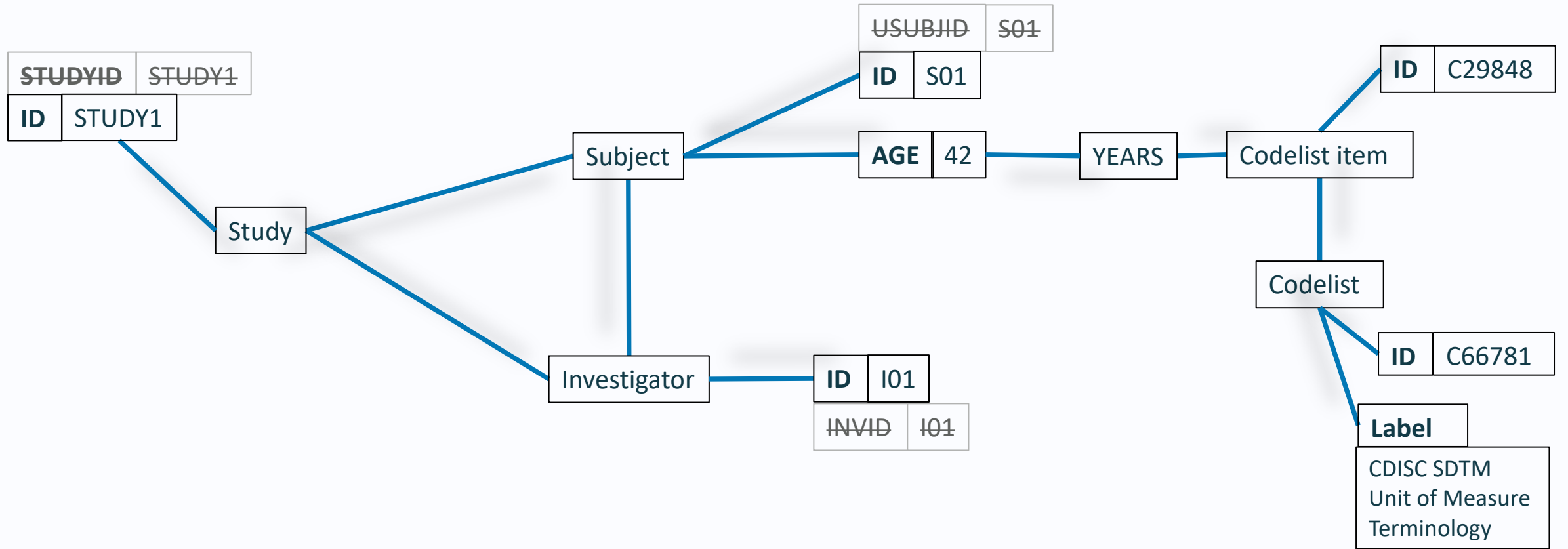












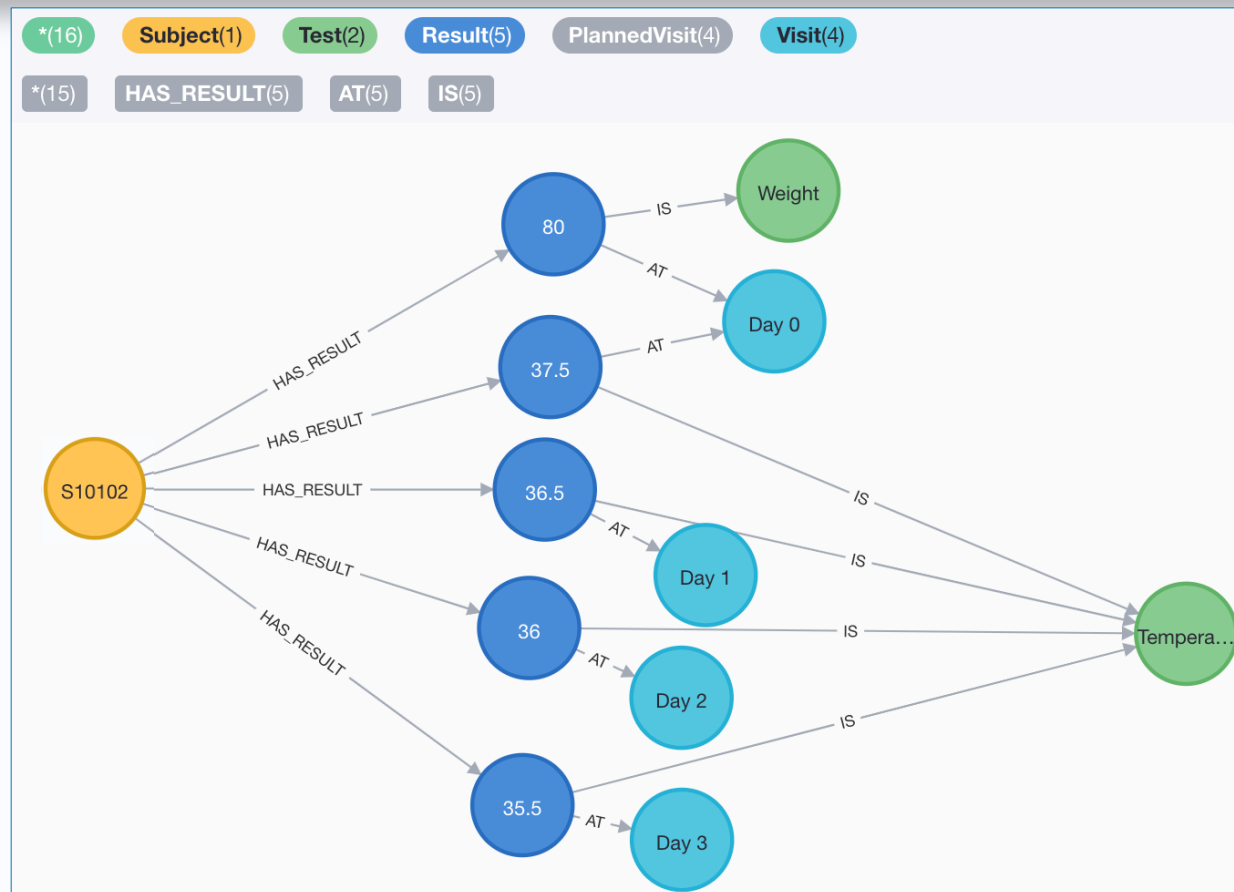
# Why do this?

- Better representation of data
  - More precise
  - Easier to understand
  - Less validation and documentation 😊
- 
- I'm going to show you some data you've never seen



# Data with relationships

```
1 MATCH (s:Subject)-[]→(r:Result)-[]→(t:Test)
2 WHERE s.id="S10102" and t.testcd in ["TEMP","WEIGHT"]
3 MATCH (r)-[:AT]→(v)
4 RETURN s, t, r, v
```



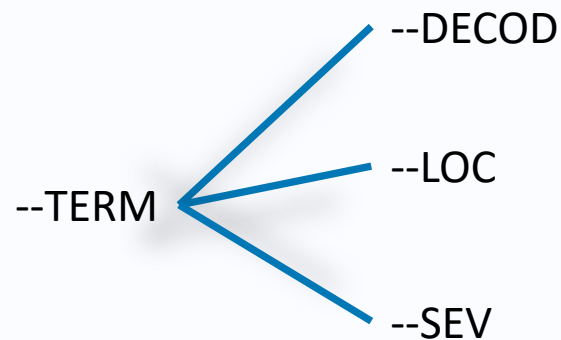
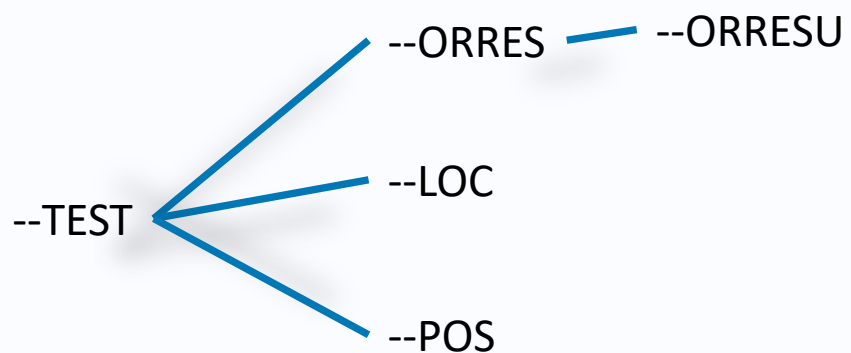
# Data with relationships – different views

```
MATCH (n:Subject)-[]→(r:Result)-[]→(t:Test)
WHERE n.id="S10102" and t.testcd in ["TEMP","WEIGHT"]
MATCH (r)-[:AT]→(v)
RETURN n.id AS USUBJID, t.testcd as TESTCD, r.value as ORRES, v.name as VISIT
```

USUBJID	TESTCD	ORRES	VISIT
"S10102"	"TEMP"	"37.5"	"Day 0"
"S10102"	"WEIGHT"	"80"	"Day 0"
"S10102"	"TEMP"	"36.5"	"Day 1"
"S10102"	"TEMP"	"36"	"Day 2"
"S10102"	"TEMP"	"35.5"	"Day 3"

# Data as Patterns/Shapes

Finding

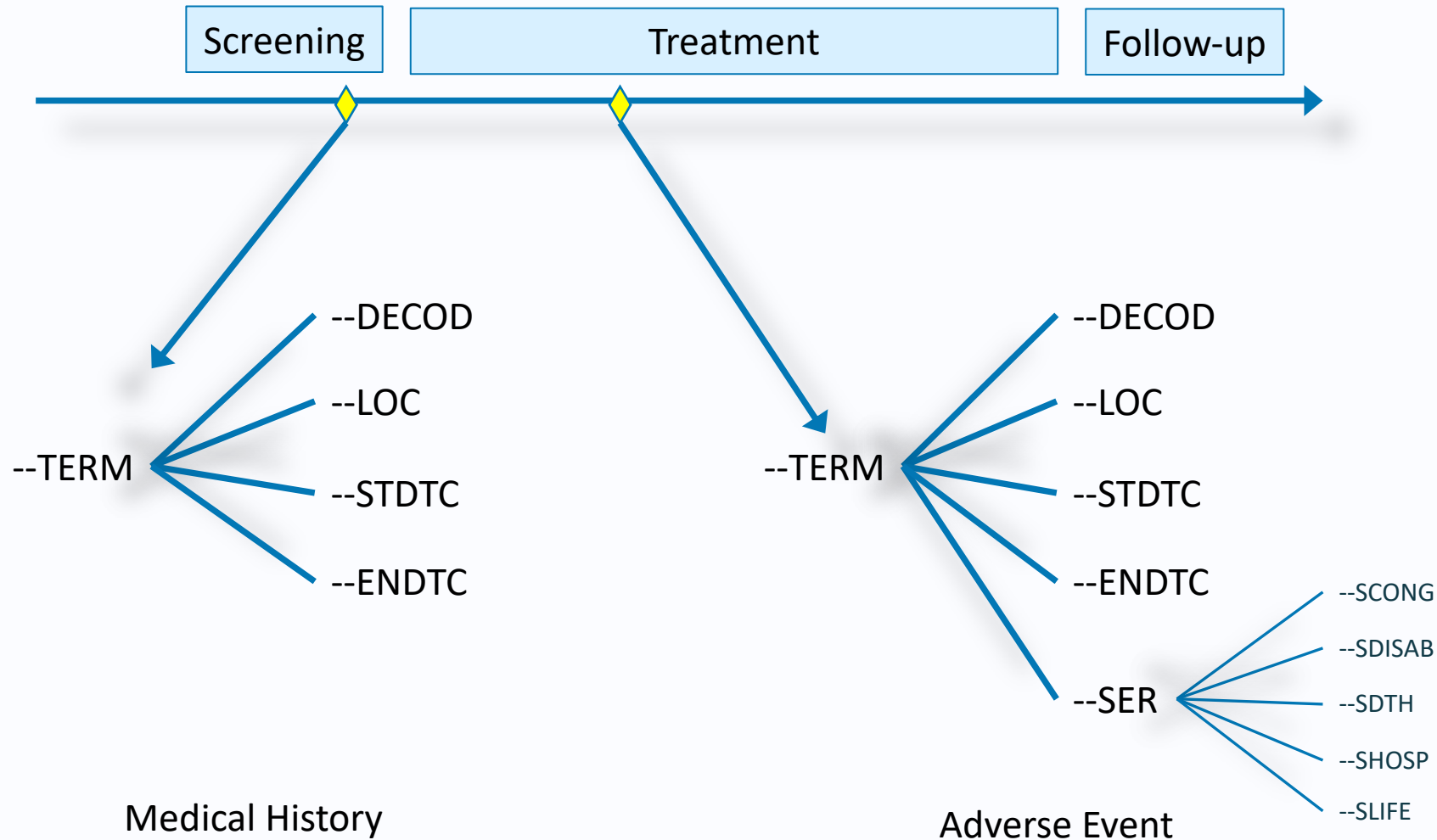


Event

Intervention

Biomedical Concepts

# Data as Patterns/Shapes



# Study setup - timeline

Study Build

Content Design and Interventions Objectives Endpoints **Timeline** SoA

### Data Collection Timeline

Arms/Epochs	Screening	Treatment
High Dose	Screen	High Dose
Low Dose	Screen	Low Dose
Placebo	Screen	Placebo

High Dose

Screening Treatment

Day 0 Week 0 Week 8 Week 12 Week 16 Week 24 Day 174

Visits: BL, Wk8, Wk12, Wk16, Wk24

Time Point

Week 16

Offset (days): 112

Epoch: Treatment

Arm: High Dose

Visit: Week 16

Unit: Day Week Month

Attached Assessments: (click to remove)

- Clinician's Interview-Based Impression of Change (CIBIC+) (Pilot) (F CIBIC) v0.1.0
- Disability Assessment For Dementia (DAD) (Pilot) (F DAD) v0.1.0

+ Add

**A Timepoint can be linked to multiple Assessments, Forms and BCs**








# Study setup – Schedule of Assessments

Study Build

Content Design and Interventions Objectives Endpoints Timeline **SoA**

### Schedule of Assessments

Type	Assessment / Visit	BL	Wk8	Wk12	Wk16	Wk24
	ASS DAD Disability Assessment for Dementia (DAD)	●	●			
	F DAD Disability Assessment For Dementia (DAD) (Pilot)	●	●		●	●
	F LAB SAMPLES Laboratory Samples	●		●		
	F ECG ECG Measurements	●	●			
	F CIBIC Clinician's Interview-Based Impression of Change (CIBIC+) (Pilot)	●	●		●	●


The relationships among the data objects are reflected in the generated Schedule of Assessments.

# Terminology releases

Community A3 MDR (v2.39.1) Johannes  
Community Reader

Terminology / CDISC, CT

Item History View Changes View Submission value changes Return

 **Controlled Terminology**  
Owner: CDISC | Identifier: CT

**History**

Show 15 entries Search:

Version	Last Change Date	Owner	Identifier	Label	Version Label	State
65.0.0	2020-06-26 02:00	CDISC	CT	Controlled Terminology	2020-06-26 Release	Standard
64.0.0	2020-05-08 02:00	CDISC	CT	Controlled Terminology	2020-05-08 Release	Standard
63.0.0	2020-03-27 01:00	CDISC	CT	Controlled Terminology	2020-03-27 Release	Standard
62.0.0	2019-12-20 01:00	CDISC	CT	Controlled Terminology	2019-12-20 Release	Standard
61.0.0	2019-09-27 02:00	CDISC	CT	Controlled Terminology	2019-09-27 Release	Standard
60.0.0	2019-06-28 02:00	CDISC	CT	Controlled Terminology	2019-06-28 Release	Standard
59.0.0	2019-03-29 01:00	CDISC	CT	Controlled Terminology	2019-03-29 Release	Standard

65 releases

# Terminology versions

Community A3 MDR (v2.39.1) Johannes Community Reader

Terminology / CDISC, C66767 / V56.0.0

CL

**ACN**

Preferred term: CDISC SDTM Action Taken with Study Treatment Terminology

Synonym(s): Action Taken with Study Treatment

Definition: Terminology specifying changes to the study treatment as a result of an adverse event.

Not extensible

Show more

PDF Report

Return

Differences ⓘ

Show 15 entries Search:

Date	Identifier	Submission Value	Preferred Term	Synonym	Definition
2007-04-20	C66767	ACN	CDISC SDTM Action Taken with Study Treatment Terminology	ACN; Action Taken with Study Treatment	The reason that changes were made to the therapy under study. (NCI)
2007-05-31	↓	↓	↓	ACN; Action Taken with Study Treatment	↓
2009-10-06	↓	↓	↓	↓	The reason that changes were made to the therapy under study. (NCI)
2011-12-09	↓	↓	↓	↓	↓
2013-06-28	↓	↓	↓	↓	Action Taken with Study Treatment
2018-06-29	↓	↓	↓	↓	Terminology specifying changes to the study treatment as a result of an adverse event.

Showing 1 to 6 of 6 entries Previous 1 Next

65 releases – not the same as 65 versions

# Summary

- Moving to linked data/graph databases to make things easier
  - Use different views of the data, not transformations (mapping)
- SDTM more important than SDTMIG
  - The thinking is correct, but the technology wasn't available
- Biomedical Concepts, think patterns/shapes
  - We have an better future ahead
  - Less documentation
  - More automation

Thank you for listening

[www.s-cubed-global.com](http://www.s-cubed-global.com)

Johannes Ulander

[ju@s-cubed.dk](mailto:ju@s-cubed.dk)