SDTM domains by query - is it possible?

Johannes Ulander
Standardisation and Harmonisation Specialist
S-Cubed
• What is data?
• Introduction to linked data and graphs
• Try some examples
• Where did my SDTM domain go?
• SDTM by Query
• Summary
What is data?
Introduction to graphs

Triples
Introduction to graphs

Site ID: 1
City: London

Node

Name: Dr. X

Property

Inves8gator

Works at

Relationship

<table>
<thead>
<tr>
<th>Node</th>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site</td>
<td>Site ID</td>
<td>1</td>
</tr>
<tr>
<td>Site</td>
<td>City</td>
<td>London</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Node</th>
<th>Relation</th>
<th>Node</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator</td>
<td>Works at</td>
<td>Site</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Node</th>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator</td>
<td>Name</td>
<td>Dr. X</td>
</tr>
</tbody>
</table>
Introduction to graphs

- **Node**: Site ID: 1
  - City: London
- **Node**: Site ID: 2
  - City: Leeds
- **Property**: Investigator
  - Name: Dr. X
  - Works at Site ID: 1
  - Works at Site ID: 2
  - City: London
  - City: Leeds
- **Property**: Investigator
  - Name: Dr. Who
  - Works at Site ID: 1
  - Works at Site ID: 2
  - City: London
  - City: Leeds
- **Property**: Investigator
  - Name: Dr. Y
  - Works at Site ID: 1
  - Works at Site ID: 2
  - City: London
  - City: Leeds
We recommend offering an option of selecting one or more racial designations or additional subgroup designations. Recommended forms for the instruction accompanying the multiple response questions are “Mark one or more” and “Select one or more.”
The graph can be queried ...

Don’t need to change the question
The Graph is for storing information

Visualization is something different
Medical History

TAUG Parkinson’s Disease
Diagnostic Criteria Example CRF

1. Year of first symptoms as confirmed by history obtained by the physician?
2. Year of Initial Diagnosis?
3. Diagnostic Features/Criteria (as evident on clinical assessment of the patient):
   a. 4-6 Hz Rest Tremor: X Present □ Absent □ Unknown
   b. Bradykinesia: □ Present □ Absent □ Unknown
   c. Rigidity: □ Present □ Absent □ Unknown
   d. Asymmetric Onset: □ Present □ Absent □ Unknown
   e. Substantial Response to Dopaminergic Therapy: □ Present □ Absent □ Unknown
4. Exposure to Genotypic or Diagnostically Related Factors:
   a. Histories (including family history):
      i. Tremor
         - Presence
         - Onset
         - Degree
   b. Idiopathic
      - Presence
      - Onset
      - Degree
   c. Genotypic
      - Presence
      - Onset
      - Degree

Graph representation
Medical History

<table>
<thead>
<tr>
<th>STUDYID</th>
<th>DOMAIN</th>
<th>USUBJID</th>
<th>MHTERM</th>
<th>MHTNM</th>
<th>MHTNT</th>
<th>MHRECQ</th>
<th>MHDECOD</th>
<th>MHCAT</th>
<th>MHRIDYS</th>
<th>MHSTDTIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDISCPLT01</td>
<td>MH</td>
<td>01-701-1015</td>
<td>ALZHEIMER'S DISEASE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PRIMARY DIAGNOSIS</td>
<td></td>
<td>2010-04-30</td>
</tr>
<tr>
<td>CDISCPLT01</td>
<td>MH</td>
<td>01-701-1015</td>
<td>ALZHEIMER'S DISEASE</td>
<td></td>
<td></td>
<td>VERBATIM 0825</td>
<td>TINNITUS</td>
<td>FIRST SYMPTOMS</td>
<td>SIGNIFICANT PRE-EXISTING CONDITION</td>
<td>EAR AND LABYRINTH DISORDERS</td>
</tr>
</tbody>
</table>

| CDISCPLT01 | MH     | 01-701-1148 | ALZHEIMER'S DISEASE |        |       |        |         | PRIMARY DIAGNOSIS |       | 2010-12-12 |
| CDISCPLT01 | MH     | 01-701-1148 | ALZHEIMER'S DISEASE |        |       | VERBATIM 0047 | TINNITUS | FIRST SYMPTOMS | SIGNIFICANT PRE-EXISTING CONDITION | EAR AND LABYRINTH DISORDERS | 2009-06-10 |

Significant pre-existing condition

**Subject**: USUBJID:S01
- **Name: Tinnitus**
  - **Primary Diagnosis**: Date:2008-06-10
  - **First symptoms**: Date:2008-06-10

**Subject**: USUBJID:S02
- **Name: Alzheimer's**
  - **Primary Diagnosis**: Date:2010-04-30
  - **First symptoms**: Date:2010-12-12
Where did my SDTM domain go?
Where did my SDTM domain go?

MATCH (n) WHERE n.DOMAIN = "MH" AND n.USUBJID = "1"
MATCH (s:Subject)-[]->(n),
    (s:Subject)-[]-(study:Study)
WHERE s.USUBJID = '1' AND n.DOMAIN = 'MH'
RETURN study.name as STUDYID,
    s.USUBJID as USUBJID,
    n.DOMAIN as DOMAIN,
    n.name as MHTERM,
    n.STDTC as MHSTDTC
SDTM domains are queries

- Part of the query
- Not duplicated
- Not mapped

First exposure to study treatment
What about Analysis Data?
What about Analysis Data?

Create average Systolic Blood Pressure

MATCH (sbpNode:StandardisedResult {name:"Systolic Blood Pressure"})<-[]-(subjectNode:Subject {USUBJID:"1"})
CREATE (averageNode:Average {name:"Average SYSBP"})
WITH COLLECT(sbpNode) as sbpNodes, averageNode, AVG(sbpNode.STRESN) as theAverageValue, subjectNode
FOREACH (node in sbpNodes | CREATE (node)-[:source:sourceValue]->(averageNode) )
SET averageNode.value = theAverageValue
CREATE (averageNode)<-[:derivedValue]-(subjectNode)
RETURN sbpNodes, averageNode, subjectNode
What about Analysis Data?
MATCH (sbpNode:StandardisedResult {name:"Systolic Blood Pressure"})<-[ ]-(subjectNode:Subject {USUBJID:"1"})
CREATE (averageNode:Average {name:"Average SYSBP"})
WITH COLLECT(sbpNode) as sbpNodes, averageNode, AVG(sbpNode.STRESN) as theAverageValue, subjectNode
FOREACH (node in sbpNodes | CREATE (node) -[:source:sourceValue]-(averageNode) )
SET averageNode.value = theAverageValue
SET averageNode.value = theAverageValue,
    averageNode.definition = "https://en.wikipedia.org/wiki/Average",
    averageNode.method = "AVG() function"
CREATE (averageNode)<-[:derivedValue]-(subjectNode)
RETURN sbpNodes, averageNode, subjectNode
The Galaxy of Clinical data meets the Universe
Summary

• It is possible to create an SDTM domain by query
  • Standardised structures $\rightarrow$ Standardised queries

• In linked data there are no boundaries
  • No mapping

• Linked data excels at handling complex information
  • Solves many of our current problems in relational databases and standards management
Thank you for listening

Johannes Ulander
Standardisation and Harmonisation Specialist
S-Cubed