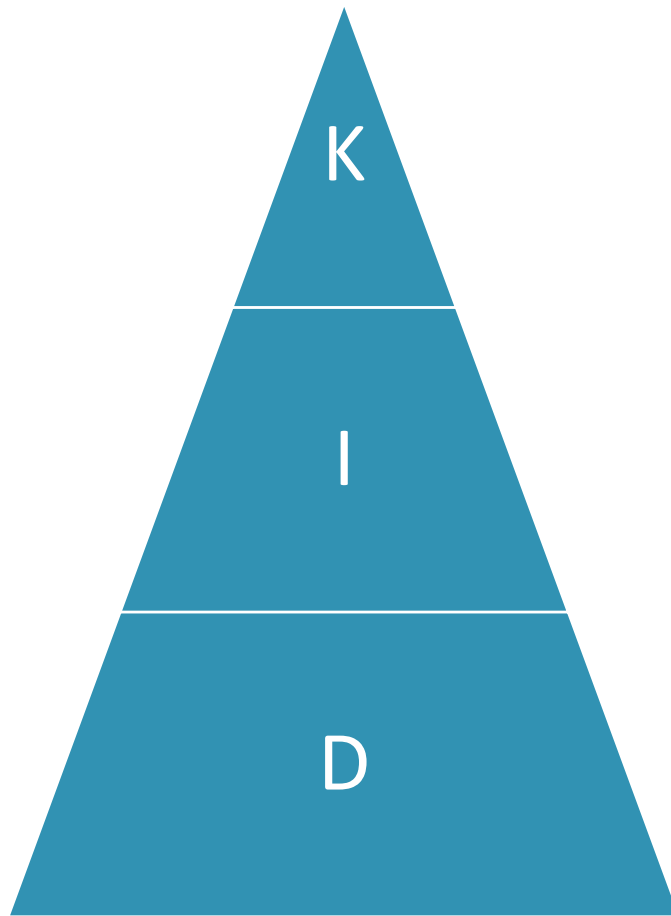


Data, Metadata and the Notion of Holons

Dr Peter Tormay

Capish

Data as the Building Block of Knowledge



Knowledge: actionable insights

Information: Data in context

Data: raw facts

Knowledge Management vs Information Management

Information Management

- Focus on Data and Information
- Structured & Unstructured facts
- Organizing, Analyzing and retrieving
- Codification of explicit knowledge – easy to copy
- Information Systems

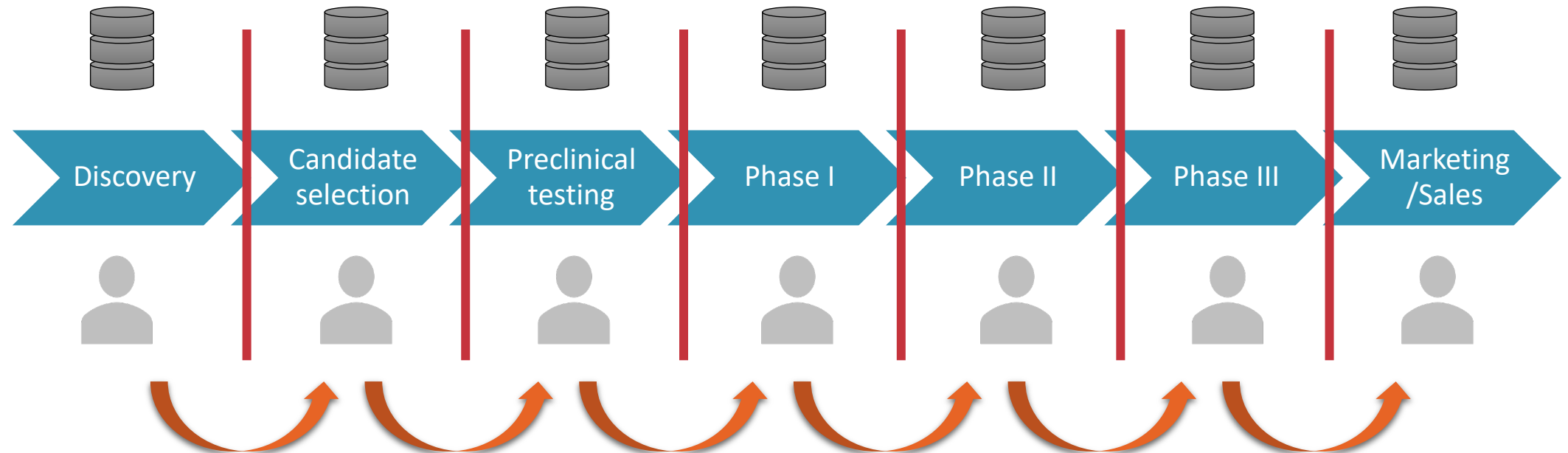
Knowledge Management

- Focus on knowledge & understanding
- Deals with codified (explicit) and uncoded (tacit knowledge)
- Understanding, Enabling, Sharing
- Creation of organisational culture, processes, environments

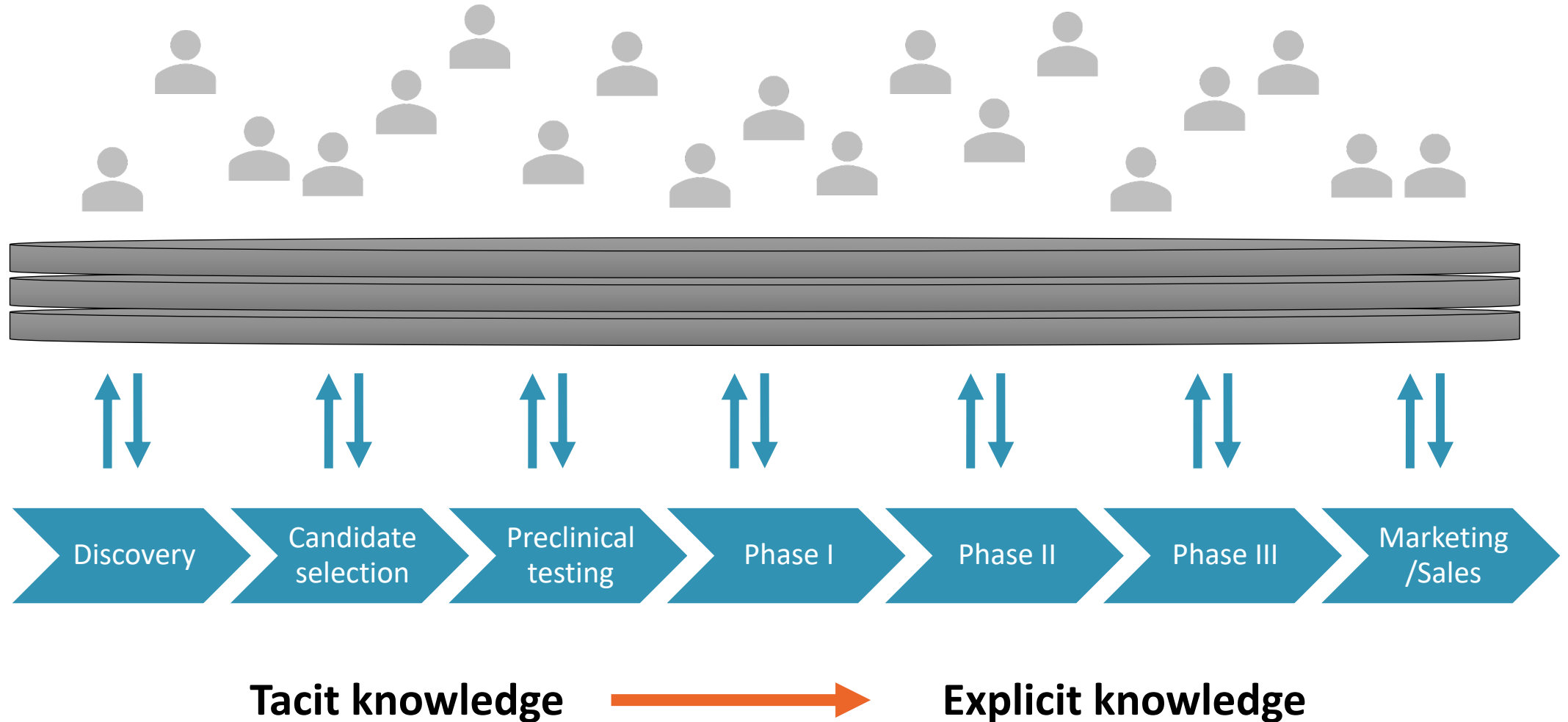


Data Silos

free flow of information between stakeholders
for the purpose of **informed decision making**



Free flow of information

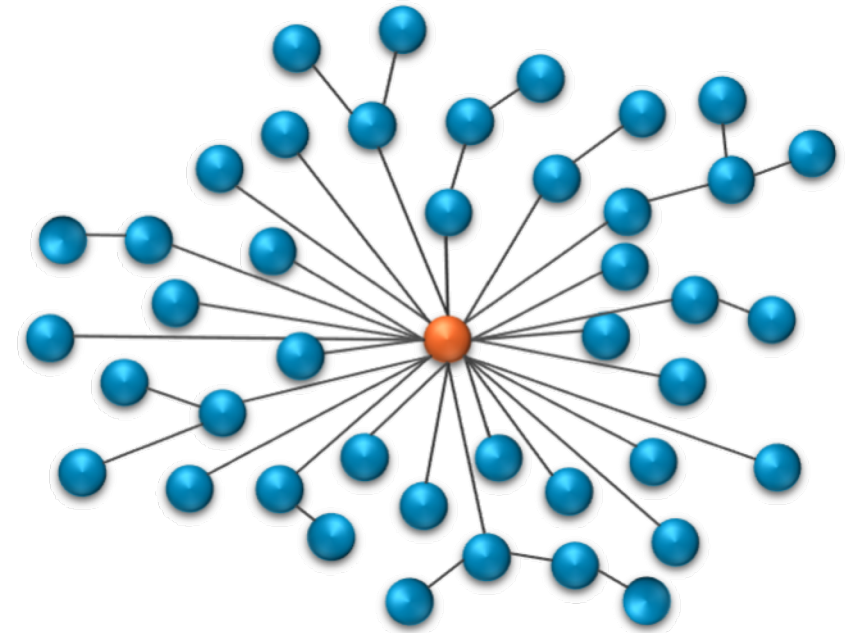


Ontology

Formal representation of a knowledge domain, describing its entities, events and processes and the relationships connecting these entities, events and processes

Benefits

- To share **common understanding** of the structure of information among people or software agents
- To enable **reuse** of domain knowledge
- To make domain assumptions **explicit**
- To **analyse** domain knowledge

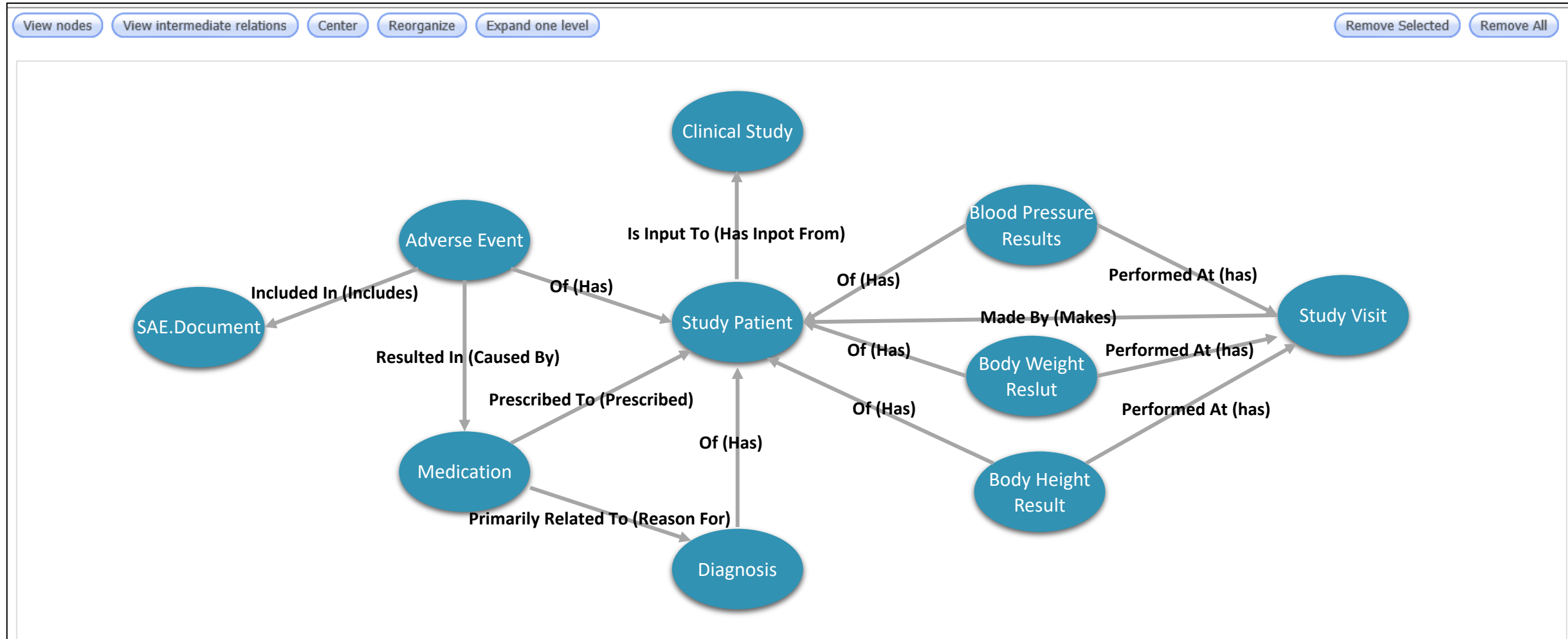


Ontologies in Life Sciences

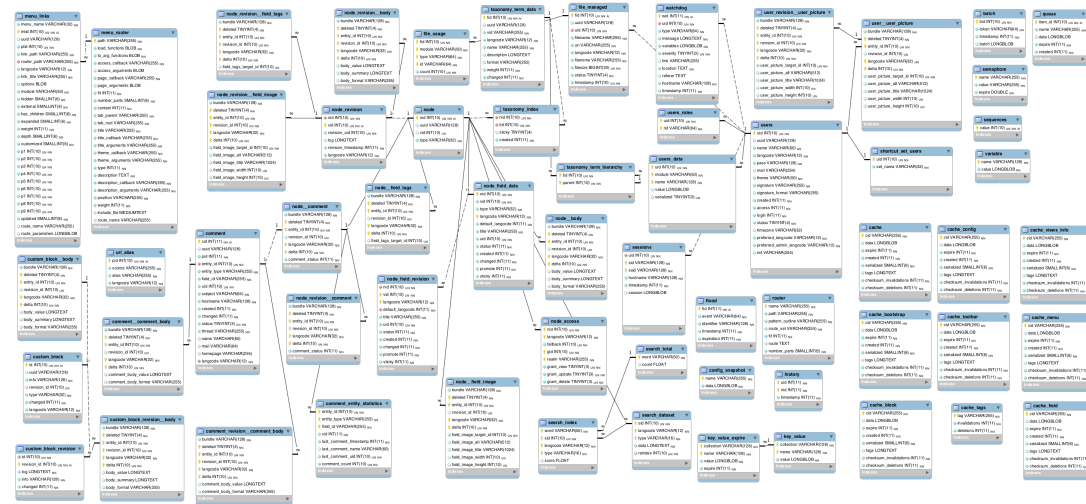
- Snomed CT
 - ICD-09/10
 - MedDRA
- 
- The diagram shows a vertical list of three items: 'Snomed CT', 'ICD-09/10', and 'MedDRA'. To the right of this list is a large closing square bracket ']' that spans the height of the list. A horizontal line extends from the middle of this bracket to the left of the text 'Terminologies – Code lists', which is underlined.
- Terminologies – Code lists

Concerned with the meaning of labels rather than the entity the labels are describing

Real world information modelling



Relational databases – fitting a square peg into a round hole



Graph databases

Triple stores

- Subject – predicate – object
- URI's and primitive data types
- Relationships have strong meaning

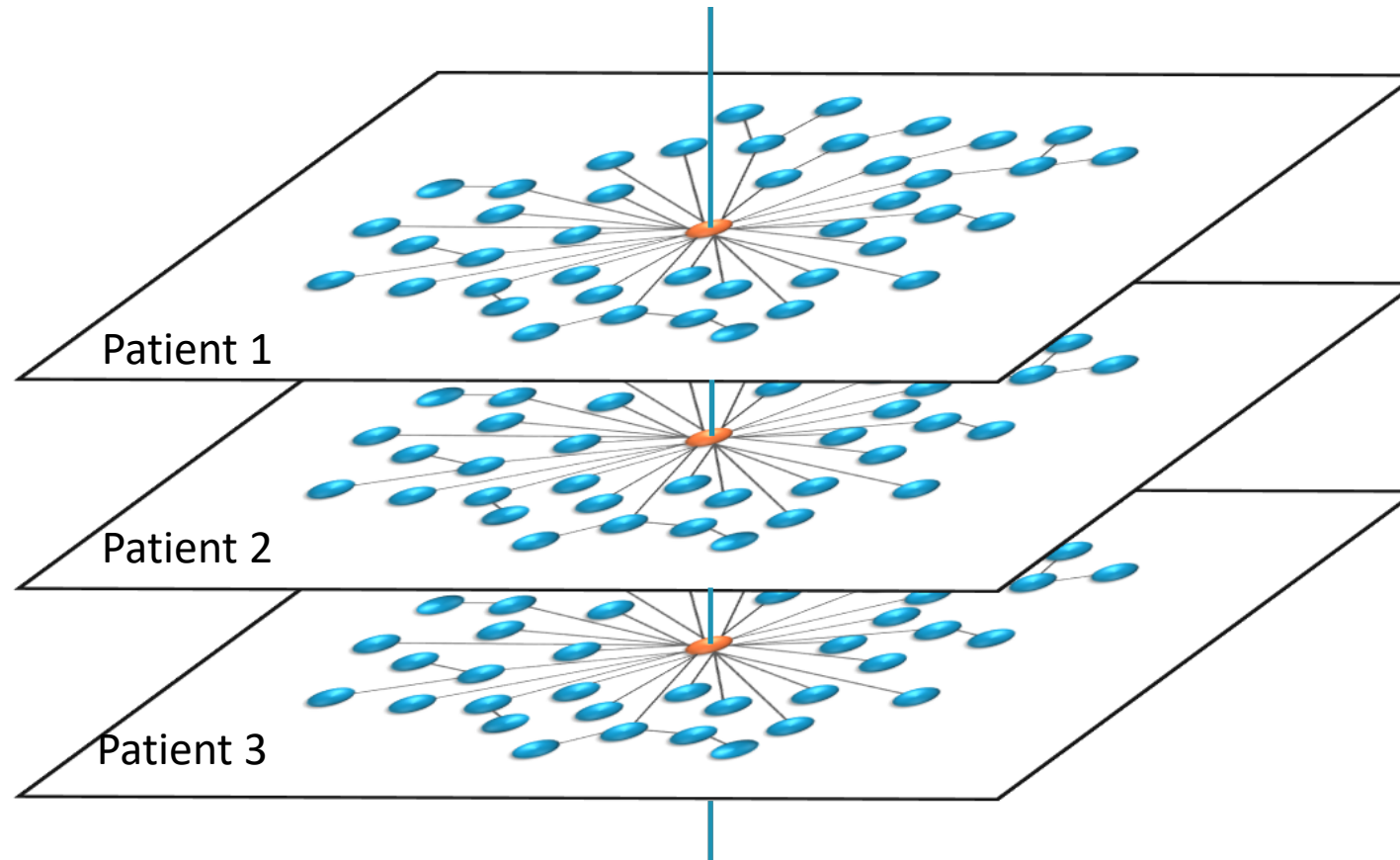
Property graph databases

- Complex objects with properties
- Relationships can/cannot have strong meaning based on the way data is modelled
- Note: Better analytical processing if all information is stored in nodes with little meaning on relationships

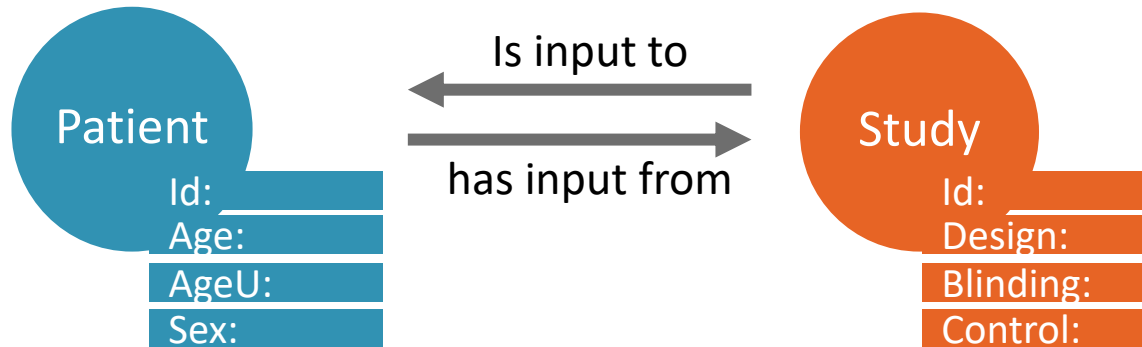
Key elements of an ontology-based data model

- Concerned with the things themselves and NOT with the definition of things (Generalisation of individual instances)
- Based on a taxonomy
 - hierarchy of things
 - Relationships between things in general
- Underpinned by a terminology

Ontology vs Instances



Holons



- A concept that can be interpreted by itself
- Classified according to content
- Contains information
 - Fields, groups and attributes
- Contains relations to other Holons
- Each relation has a specific meaning

Header

```
<o
  holonType= "Body Weight.Results"
  holonConcept="Body Weight"
  dimension="Data"
  domain="616.0754—dc23"
  ontology="(G5.24.1.0).ClinSci.ontology@capish.se"
  usage="Data"
>
<k>
<f>Body Weight.Results Key<v>BodyWeightKey...</v></f>
</k>
```

Message

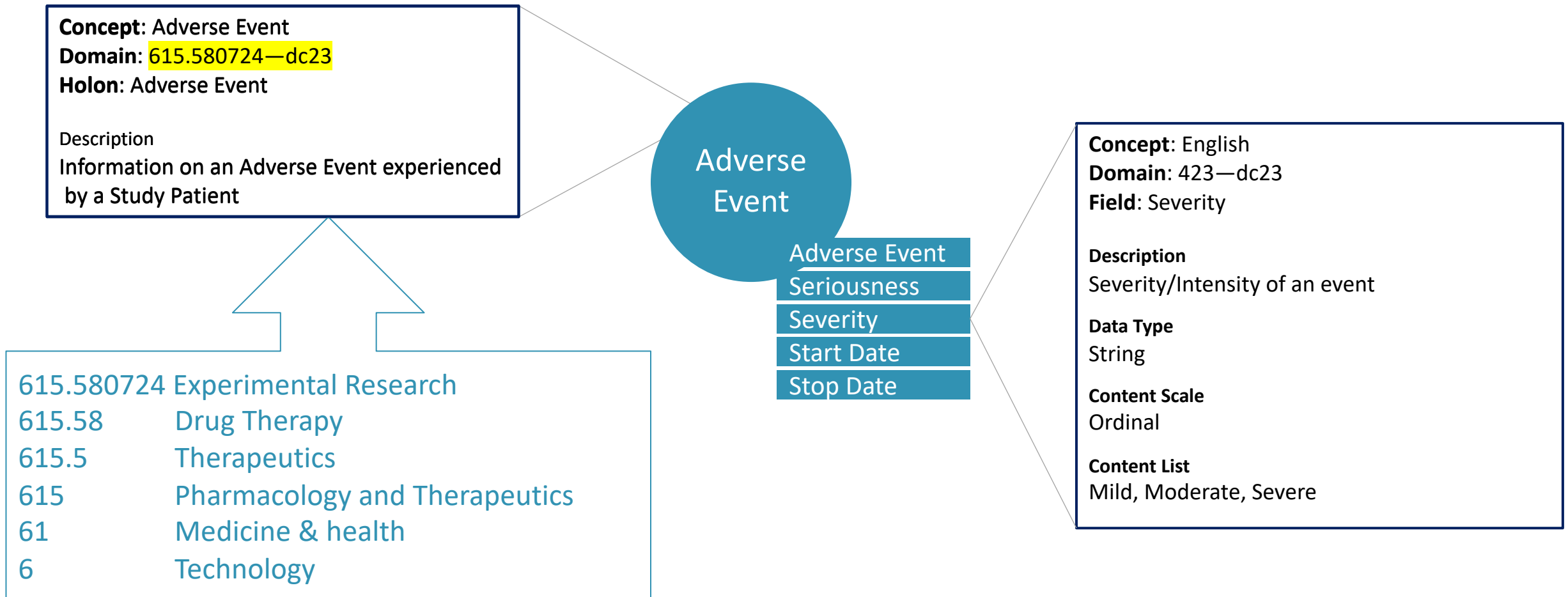
```
<f>Date<v standard="ISO 8610" type="dateTime">2010-01-05</v></f>
<f>Body Weight<v unit="kg">68</v></f>
<f>Body Weight.Original<v unit="pounds">150</v></f>
```

Relation

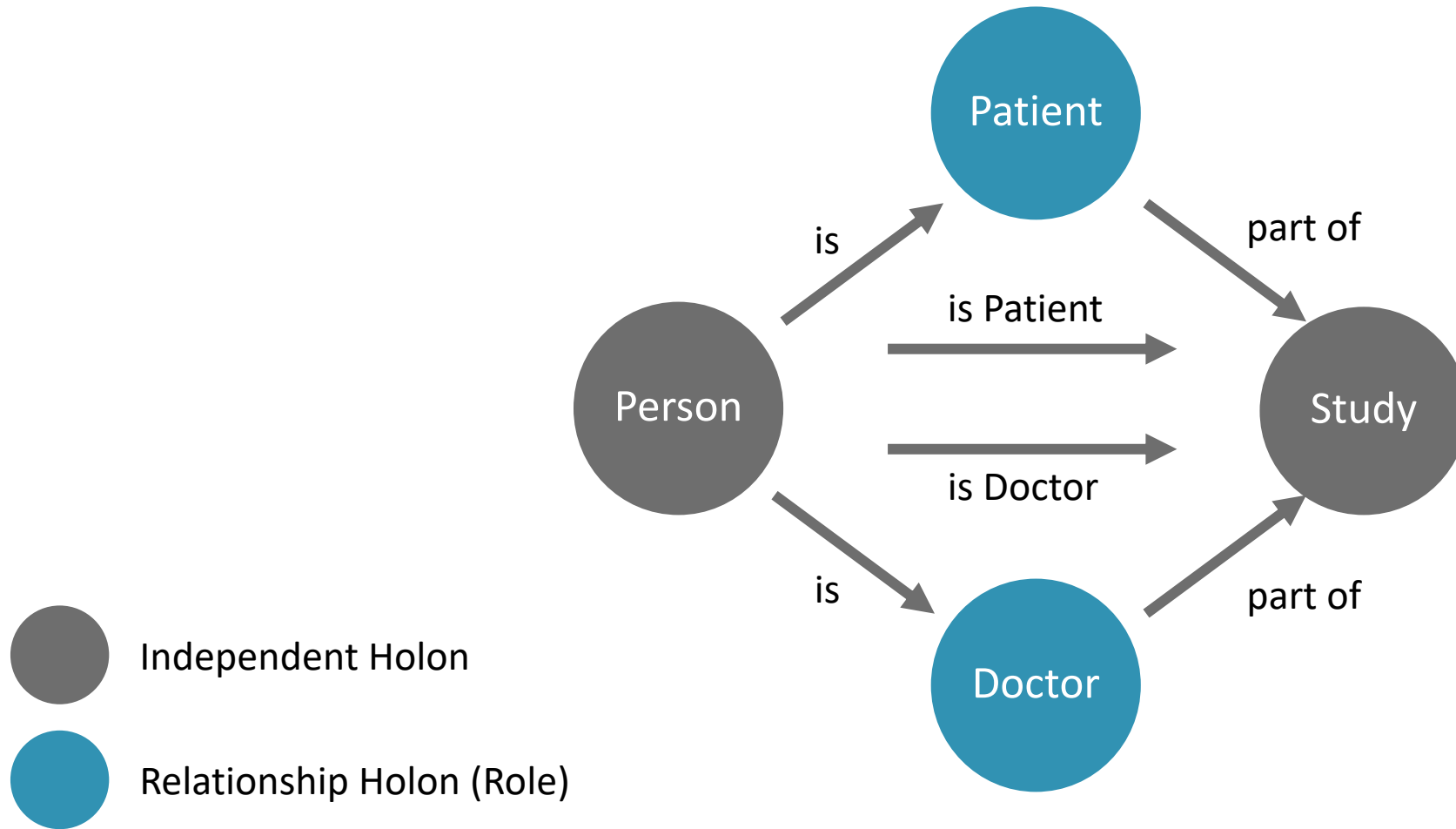
```
<r
  reverse="Has" dimension="Data" holonType="Patient">
  Of
  <f>Patient Key<v>PatientKey...</v></f>
</r>

</o>
```

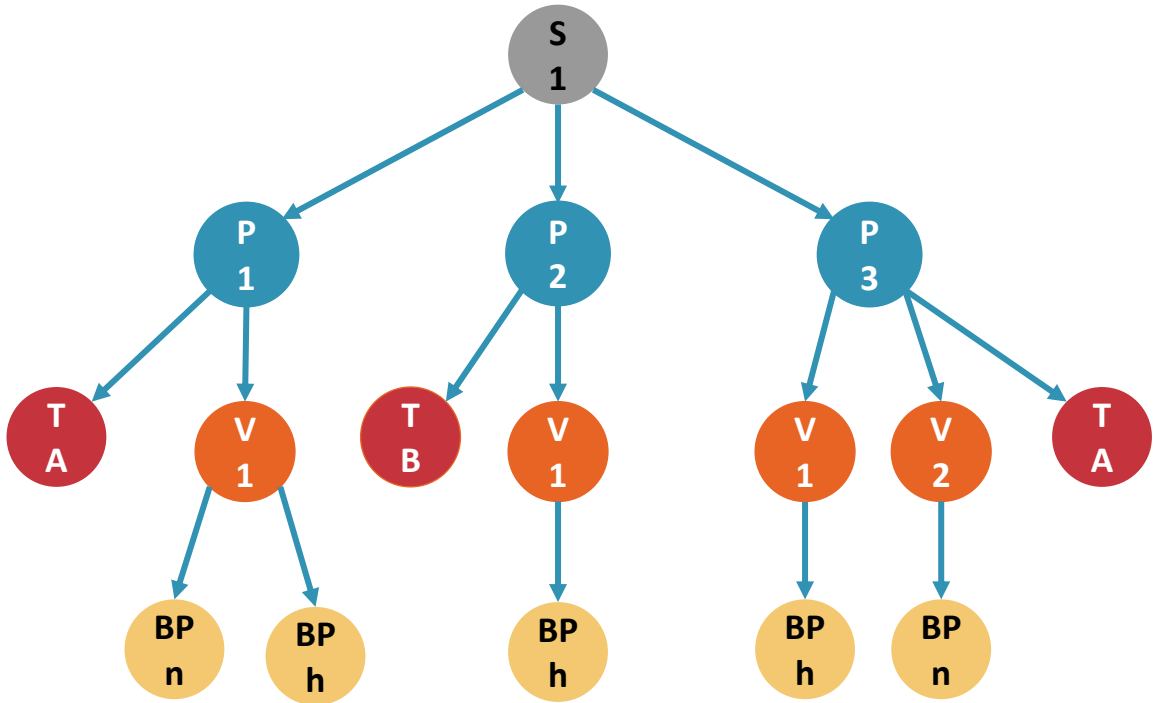
Holons – Data and Metadata



A word on relationships



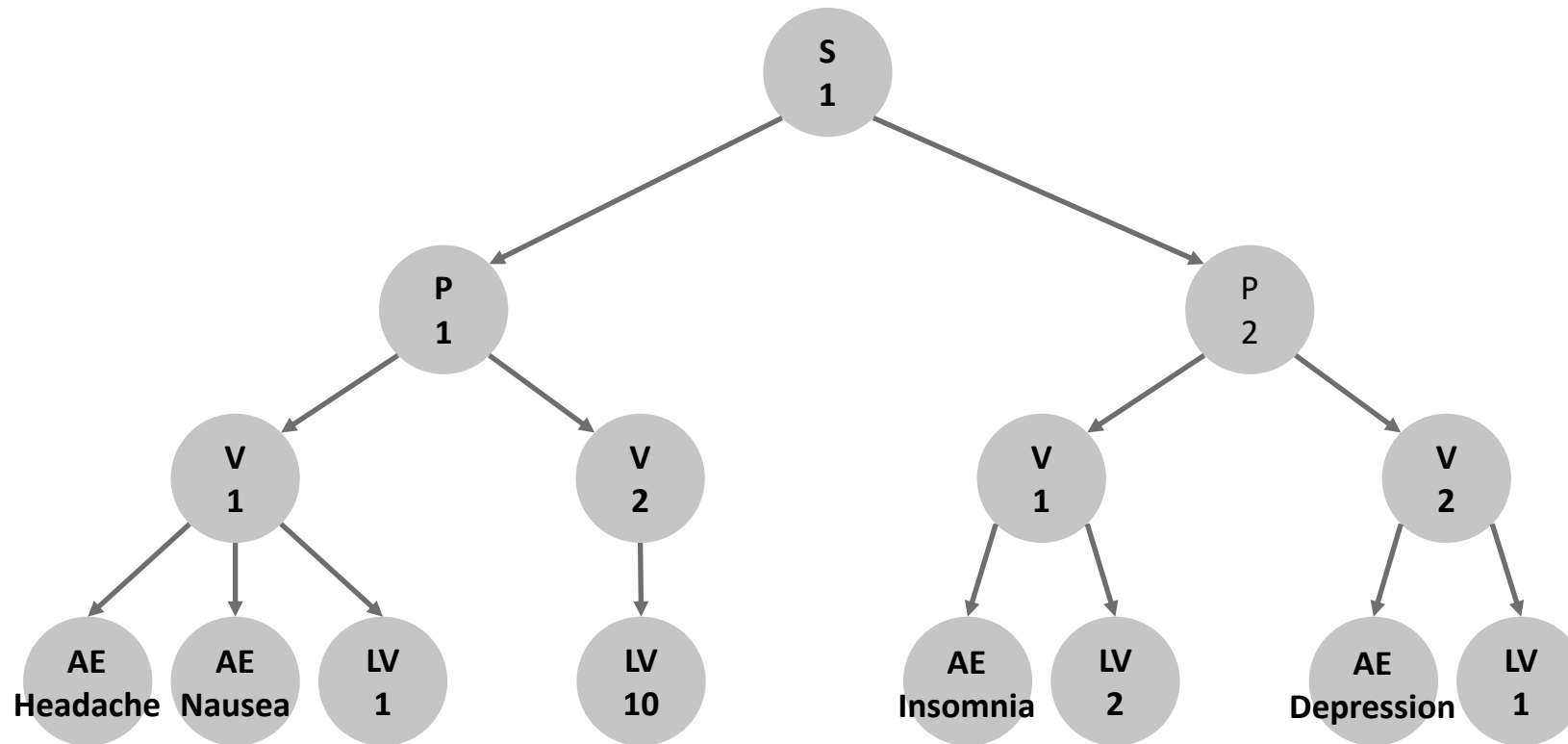
Graph Database Implementation



- individual nodes (identity index)
- node types (node type identity index)
- property values (property index)
- existence of indirect relationships (relation index)

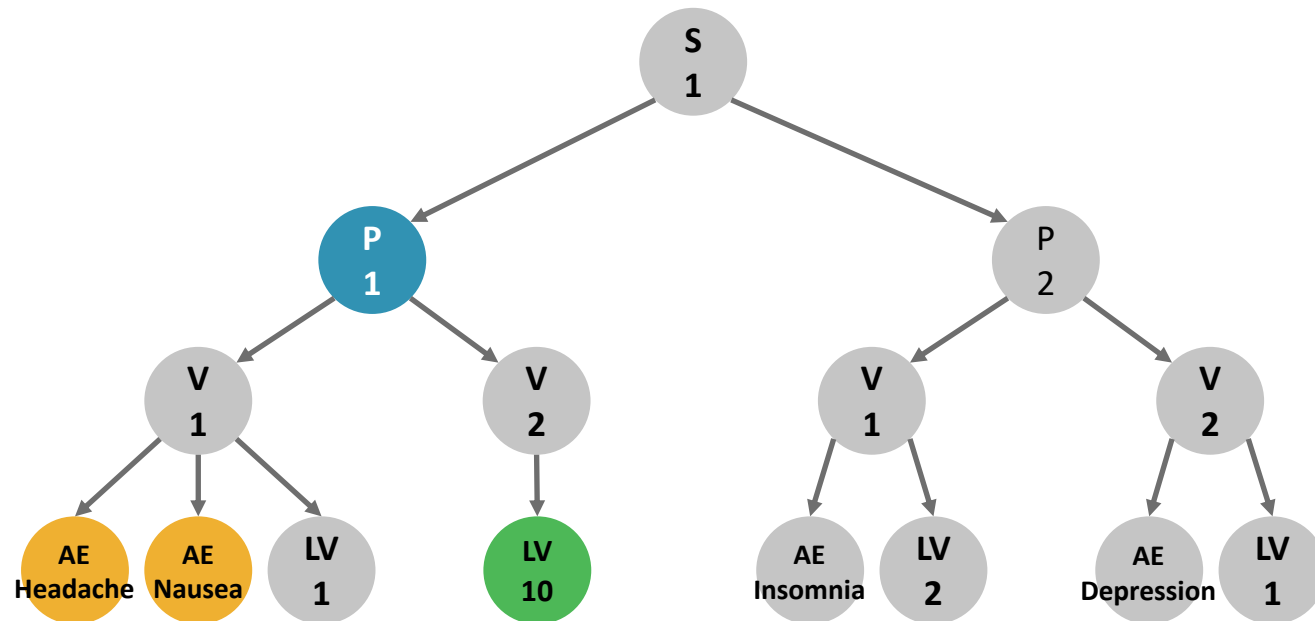
Graph Database Query

What adverse events have been reported for patients with elevated liver values



Graph database query

What adverse events have been reported for patients with elevated liver values



Select data:

Type of Node: "Patient"

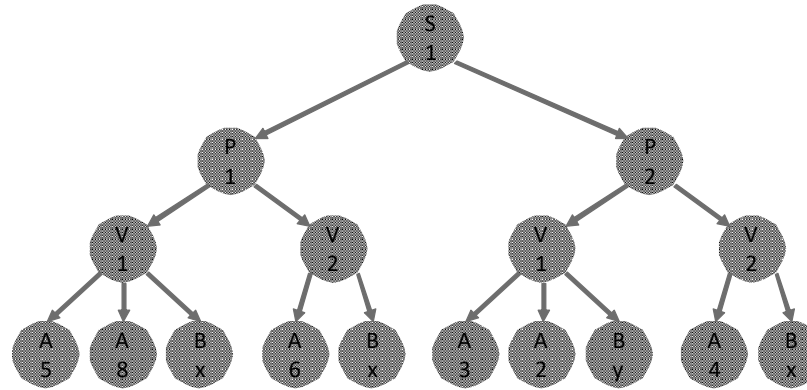
With (Type of Node: "Liver Value", Property: "Value > 5")

Fetch:

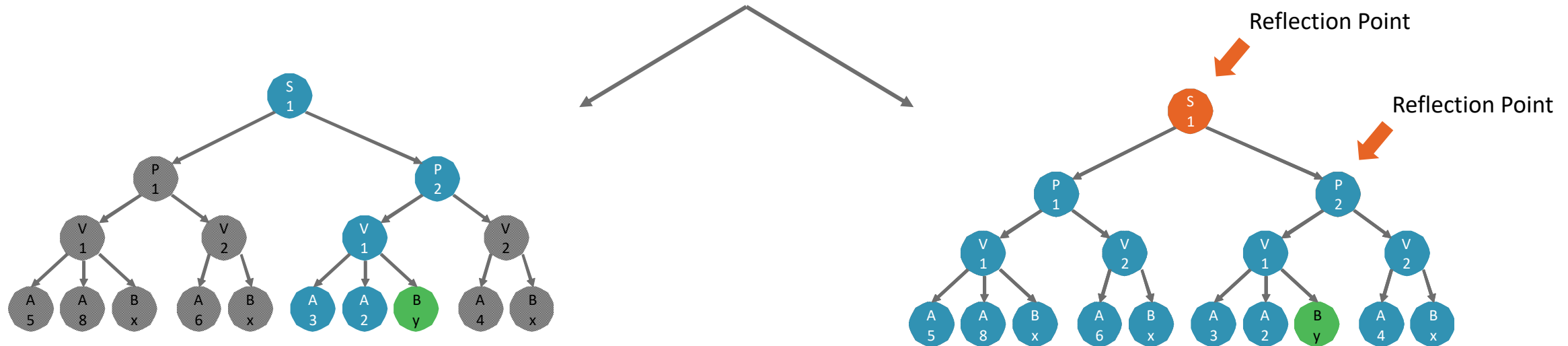
Type of Node: "Adverse Event", property: "Name"

Reflect your data – Building cohorts easily

Without
Reflection

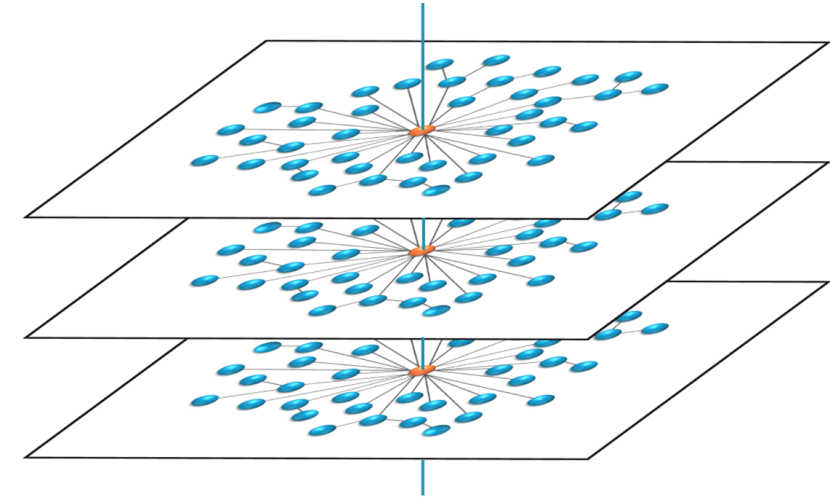


With
Reflection



Conclusions

- Ontologies provide a framework that enables a realistic representation of data and its context
- Ontological entities and relationships can be represented in object based graph databases
- Ontologies allow for data and metadata to be kept together
- Ontology based information models with focus on nodes enable better data structures and search capabilities





Thank you

Peter Tormay

peter.tormay@capish.com

Anna Berg

anna.berg@capish.com