

TERMINOLOGY AND KNOWLEDGE ORGANISATION: THE ROLE OF ONTOLOGIES IN DIGITAL HUMANITIES

Margarida Ramos | NOVA FCSH @ ROSSIO WORKSHOP



Theoretical framework

Terminology

- Naturally linked to domain-specific knowledge
- Primarily concerned with the relationship between concepts, although it also focuses on the relationships found between their designations – the terms. This is essential for achieving quality in multilingual communication contexts
- The relationships identified between concepts are the core organising principle of terminology work and are commonly reflected in the chosen environment to represent the knowledge organisation



cf. <http://www.computing.surrey.ac.uk/ai/pointer/report/section1.html>

Theoretical framework

semasiological approach

onomasiological approach

TERMINOLOGY

LINGUISTIC

CONCEPTUAL

has dimension

complementary

inferring concepts from text analysis



...vinho (o "Comprador")
...características do vinho e do
...deve ter em conta:
...separada:
...contorno, a funcionalidade da rolha de cortiça.
...Este conhecimento permitirá ao Comprador, juntamente com a consulta ao Vendedor de Rolhas (o "Vendedor"), a seleção de rolhas do tipo, e onde aplicável, classe visual, correntes e das apropriadas características físicas, químicas e microbiológicas. Enquanto que a maioria destas últimas características se relacionam com aspectos de desempenho das rolhas, detalhadas na Seção 8, a classificação das rolhas está maioritariamente dependente de características visuais externas da cortiça. As características visuais da cortiça podem ser avaliadas por pessoas treinadas e/ou por máquinas.

7.1 Dimensões da rolha

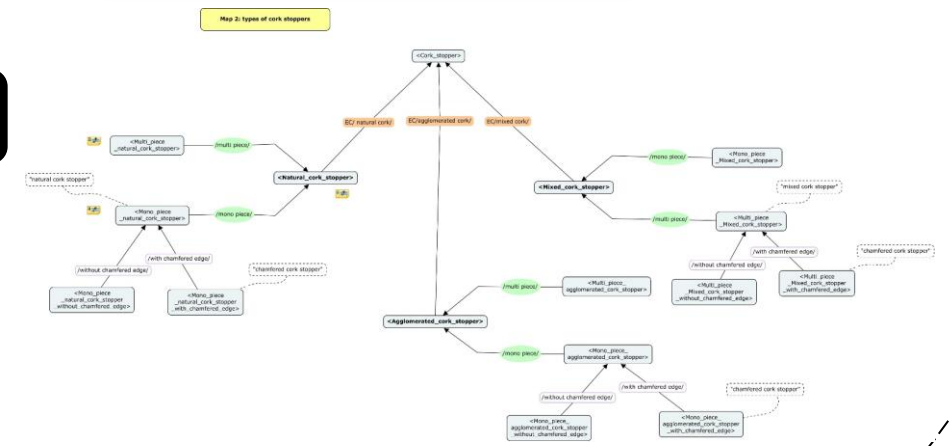
Para determinar as dimensões das rolhas a usar recomenda-se um estudo do perfil interno da garrafa, em conjunto com o conhecimento das condições de engarrafamento e as características do vinho.

7.1.1 Comprimento da rolha

O comprimento da rolha selecionada deve estar de acordo com o nível de enchimento da garrafa. Se o comprimento da rolha for demasiado grande, o espaço de cabeça pode ser comprimido, forçando o vinho a sair da garrafa, fazendo com a que rolha levante após inserção na garrafa. De qual forma, a rolha não deve ser tão curta que cause um espaço de cabeça excessivo, o qual, em determinadas condições, pode contribuir para o desenvolvimento de características sensoriais desfavoráveis.

7.1.2 Diâmetro da rolha

Deve-se ter o cuidado de adequar o diâmetro da rolha ao perfil interno do gargalo da garrafa.



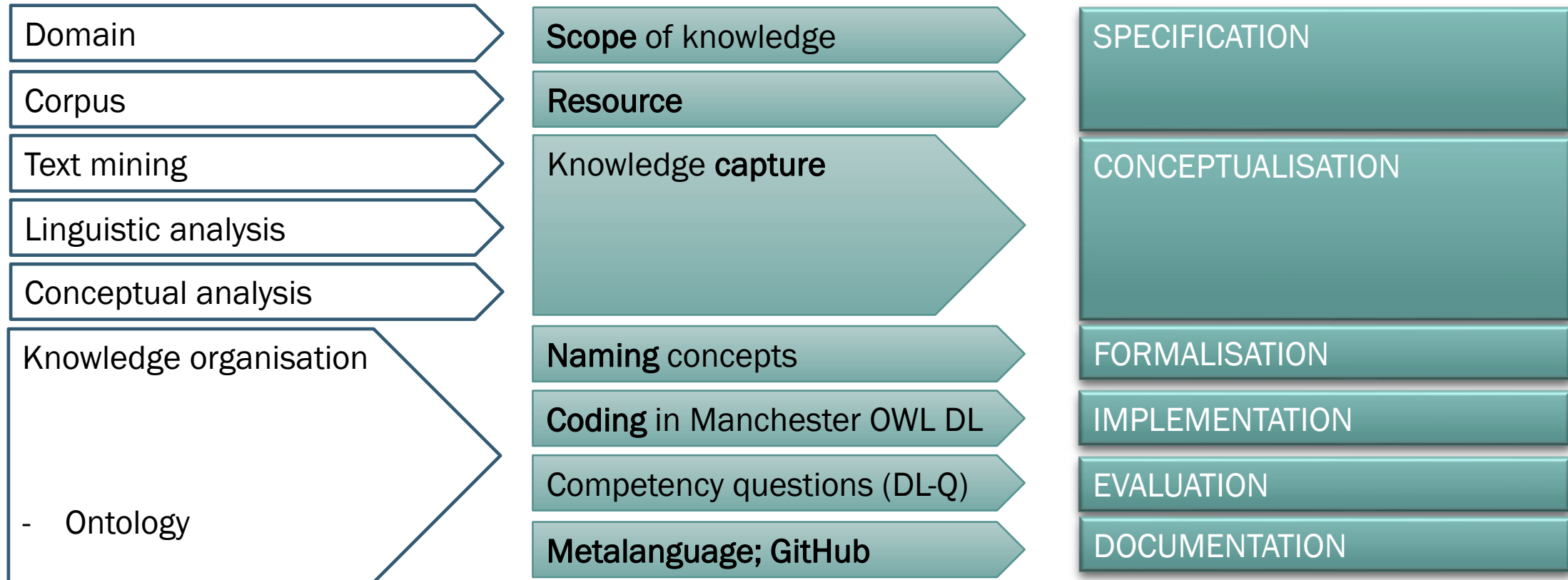
term
verbal designation of a
concept in a specific
subject field (ISO 1087-1)

concept
unit of knowledge created by a
unique combination of
characteristics (ISO 1087-1)

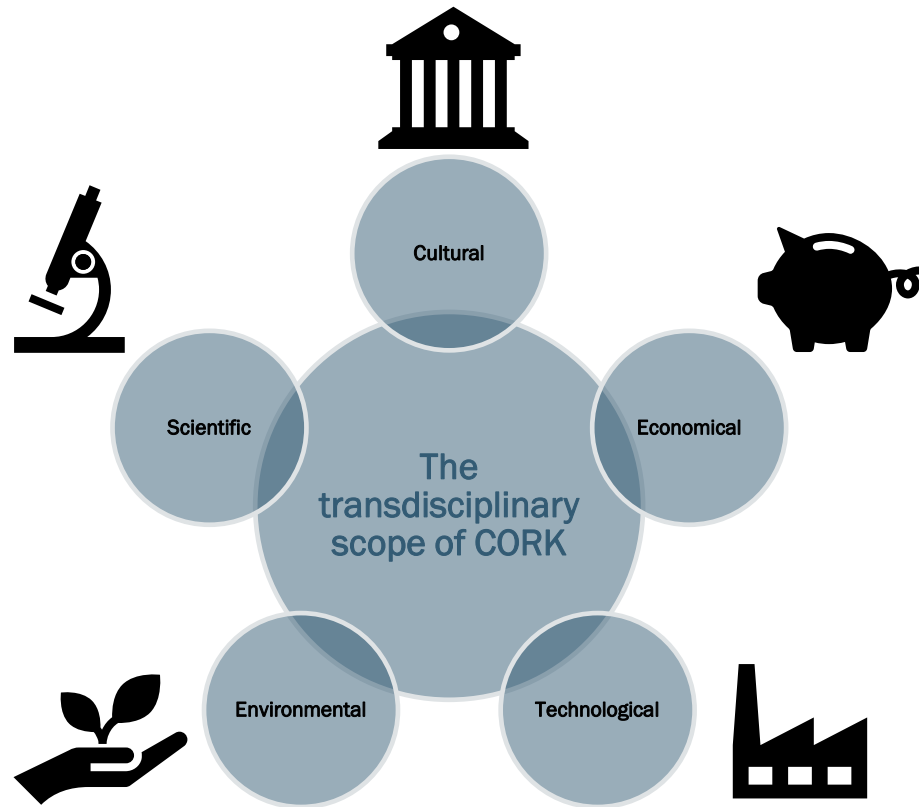
designates

is designated

ROADMAP OF TERMCORK FOLLOWING THE CLASSIC CYCLE TO BUILD AN ONTOLOGY



DOMAIN: CORK, A MULTIFACETED SCOPE OF INTERESTS

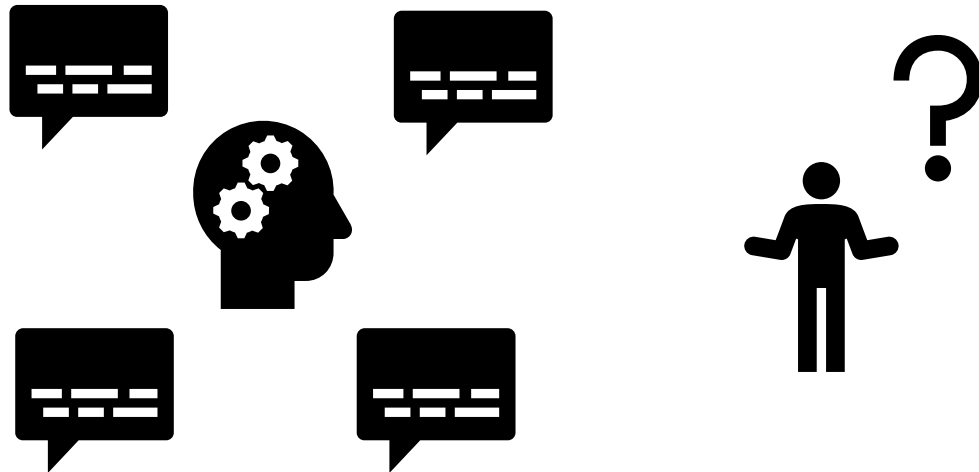
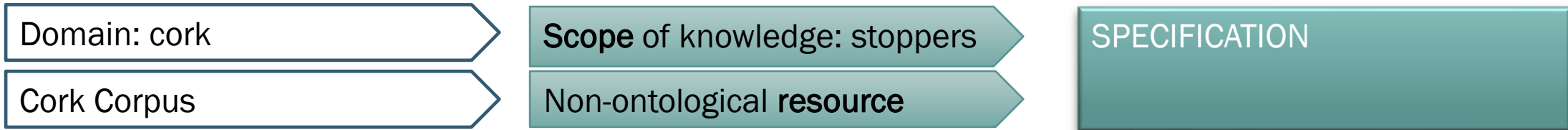


Cork oak forests have a high economic, social and environmental value in Portugal

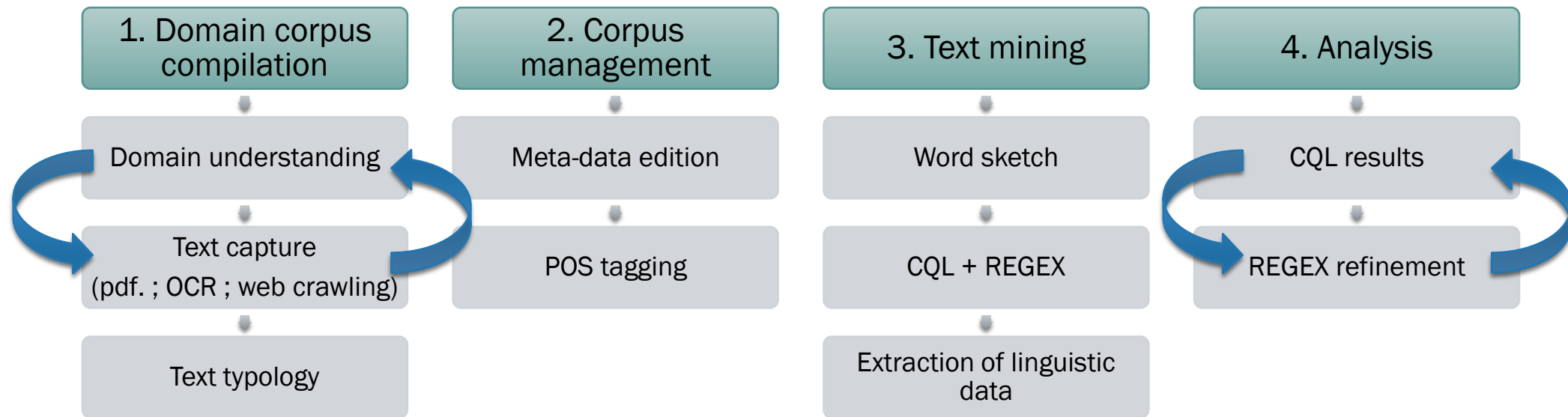
- ✓ Production (“montados”)
- ✓ Biodiversity and CO² capture
- ✓ **Transformation** (industry)
- ✓ Leader in the world ranking of international market shares

→ An endless field of terminological study

A CORPUS BUILT FROM SCRATCH AS A RESOURCE TO GRASP EXPERTS' CONCEPTUALISATIONS



CORPUS BUILDING AND PROCESSING: OVERVIEW



Ramos & Costa, Toth 2019

LINGUISTIC ANALYSIS: THE 2ND TASK OF KNOWLEDGE CAPTURE



4 DEFINITIONS SELECTED TO DEMONSTRATE OUR METHODOLOGY

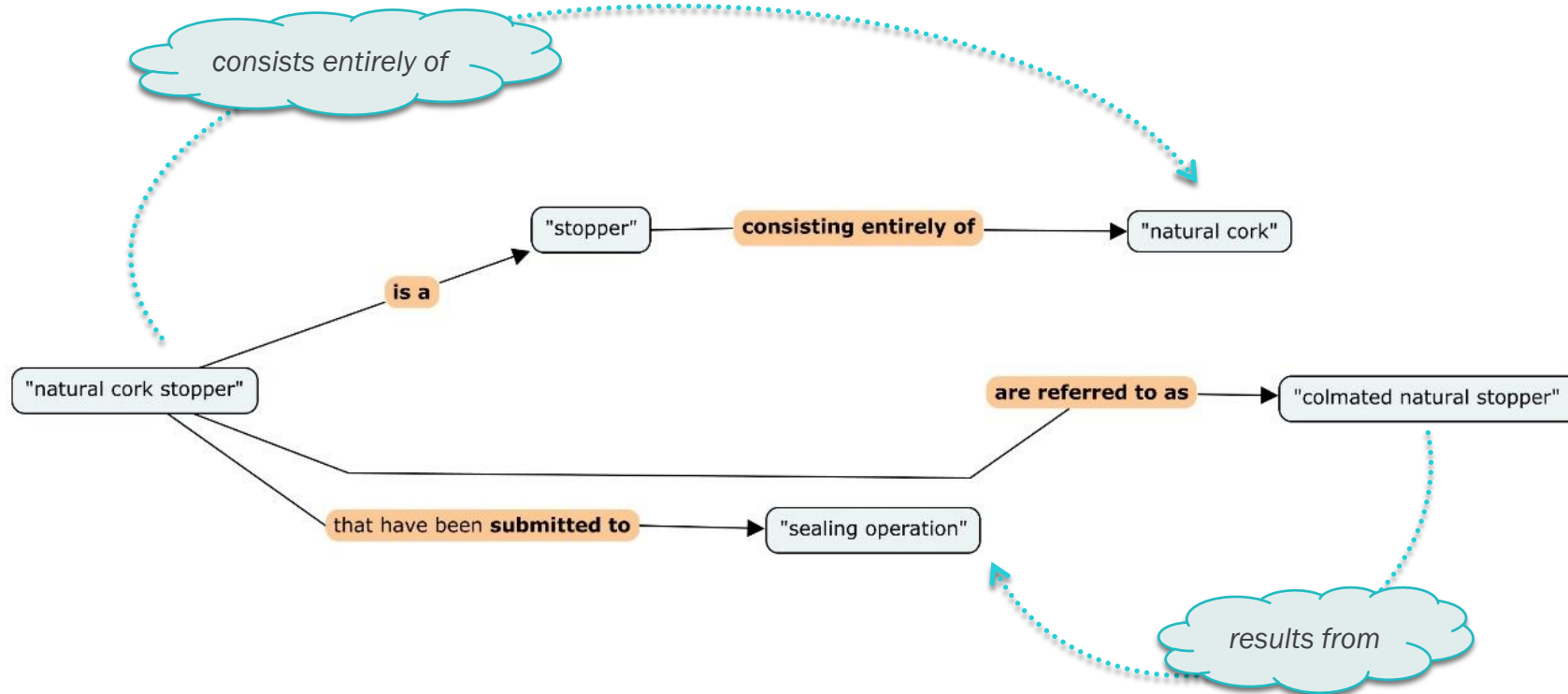
4 definitions (literal translations from pt)	4 definitions (pt) extracted from the Cork corpus
<p>stopper</p> <p>Product obtained from natural cork and / or agglomerated cork, consisting of one or more pieces, intended to seal bottles or other containers and to preserve their contents. (5.1 - NORM)</p>	<p>rolha</p> <p>Produto obtido da cortiça natural e / ou de cortiça aglomerada, constituído por uma ou mais peças, destinado a vedar garrafas ou outros recipientes e a preservar o seu conteúdo. (5.1 - NORM)</p>
<p>STOPPER</p> <p>piece of cork, usually cylindrical, conical or prismatic quadrangular, sometimes with rounded or chamfered lateral edges, consisting of one or several glued elements and intended to seal the containers or contribute to their water tightness. (7.8 - TECH)</p>	<p>ROLHA</p> <p>peça de cortiça, em geral cilíndrica, troncocónica ou prismática quadrangular, por vezes de arestas laterais boleadas ou chanfradas, constituída por um ou vários elementos colados e destinada a vedar os recipientes ou a contribuir para a sua *estanquicidade (7.8 - TECH)</p>
<p>natural cork stopper</p> <p>Stopper consisting entirely of natural cork</p> <p>Note: Natural cork stoppers that have been submitted to the sealing operation (see 6.5.5) are commonly referred to as colmated natural stoppers. (5.5 - NORM)</p>	<p>rolha de cortiça natural</p> <p>Rolha totalmente constituída por cortiça natural.</p> <p>Nota: As rolhas naturais que tenham sido submetidas à operação de colmatagem (ver 6.5.5) são comumente designadas por rolhas naturais colmatadas. (5.5 - NORM)</p>
<p>colmated natural cork stopper</p> <p>The colmated natural cork stopper is a stopper made of natural cork in which its lenticels are filled with a mixture of glues and cork powder from the dimensional finishing processes of natural cork stoppers. (6.1 - REP)</p>	<p>rolha de cortiça natural colmatada</p> <p>A rolha de cortiça natural colmatada é uma rolha feita de cortiça natural em que são obturadas as suas lenticelas com uma mistura de colas e pó de cortiça proveniente dos acabamentos dimensionais das rolhas de cortiça natural. (6.1 - REP)</p>

Lexical Map 3 - Representation of Definition 3 :

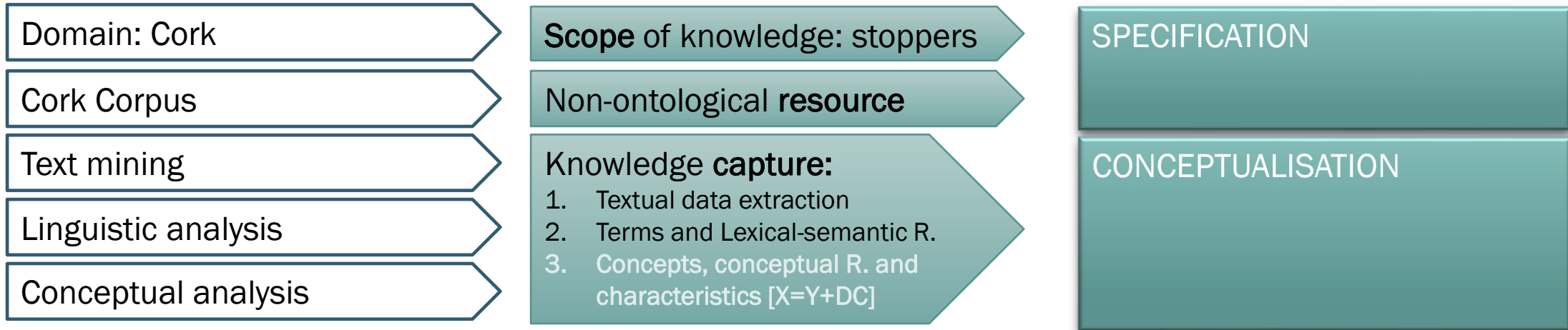
natural Cork Stopper

stopper consisting entirely of natural cork

Note: Natural cork stoppers that have been submitted to the sealing operation are commonly referred to as colmated natural stoppers



CONCEPTUAL ANALYSIS: THE 3RD TASK OF KNOWLEDGE CAPTURE



CONCEPTUAL ANALYSIS: $X=Y+DC$, A MECHANISM TO INFER KNOWLEDGE FROM TEXTS

Aristotelian formula X [SPECIES] = Y [GENUS] + DC [DIFFERENTIAL CHARACTERISTICS]

1.

We can systematically infer:

- **Characteristics**

colmated natural stopper [SPECIES] = natural stopper [GENUS]+ colmated [DC]

- **Concept's place**

- *proximum* genus
- species

colmated natural cork stopper [SPECIES] = natural cork stopper [GENUS] + sealing operation [DC]

2.

Propose **conceptual relations identifiers** to mirror how concepts relate

- **Conceptual relations**






- Subsumption
- Associative
- Partitive

has_process [corresponds to LM 'submitted to']

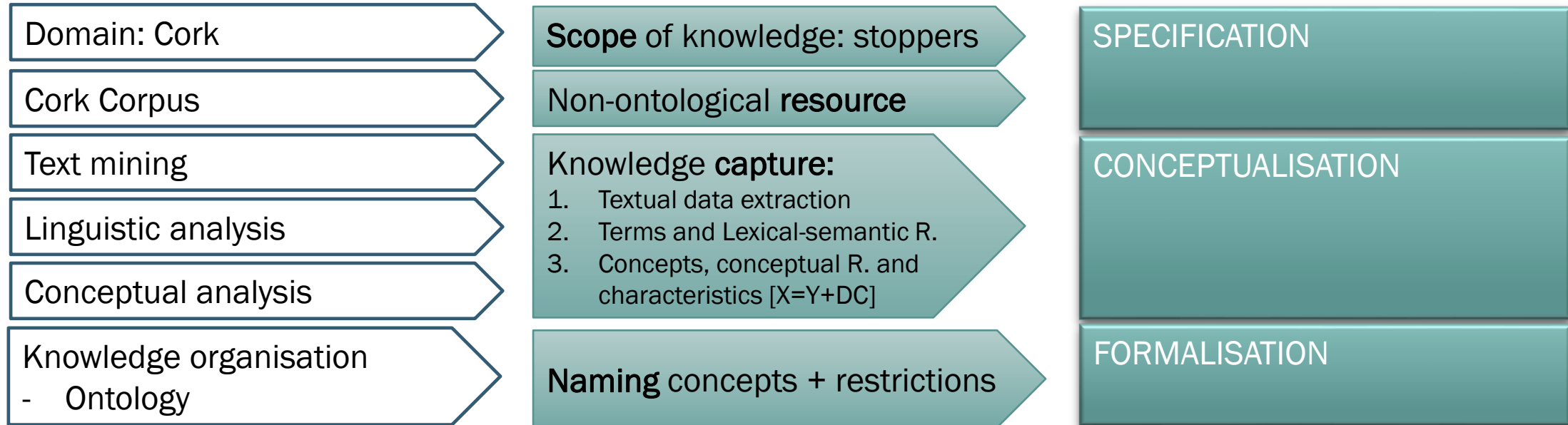
ASSOCIATIVE relation [PROCESS-RESULT]

The starting point to name concepts and domain descriptive relations to build the ontology

CONCEPTUAL ANALYSIS: FINDING AXES OF ANALYSIS TO BUILD AN ONTOLOGY

Lexical marker (en)	Lexical-semantic relation	Conceptual relation identifier	Conceptual relation	Axis of analysis
'usually'	HYPERNYMY - HYPONYMY	has_shape	ASSOCIATIVE object-shape	Shape 
'sometimes with'	HYPERNYMY - HYPONYMY	has_process	ASSOCIATIVE process-result	Finishing Process 
'commonly referred to as'	HYPERNYMY - HYPONYMY	is_a	SUBSUMPTION	
'consisting of'	HOLONYMY-MERONYMY object-components	has_part	PARTITIVE	Parts 
'obtained from'	HOLONYMY-MERONYMY object-stuff	has_raw_material	ASSOCIATIVE product-raw material	Substance 
'consisting entirely of'	HOLONYMY-MERONYMY object-stuff	has_substance	ASSOCIATIVE matter/substance - property	Substance
'have been submitted to'	HOLONYMY-MERONYMY activity-feature	has_process	ASSOCIATIVE process-result	Finishing Process
'is made of'	HOLONYMY-MERONYMY object-stuff	has_substance	ASSOCIATIVE product-raw material	Substance
'are filled with'	HOLONYMY-MERONYMY activity-feature	has_process	ASSOCIATIVE process-result	Finishing Process
'results from the'	HOLONYMY-MERONYMY activity-feature	has_process	ASSOCIATIVE process-result	Finishing Process
'intended to'	---	has_function	ASSOCIATIVE object-function	Function 

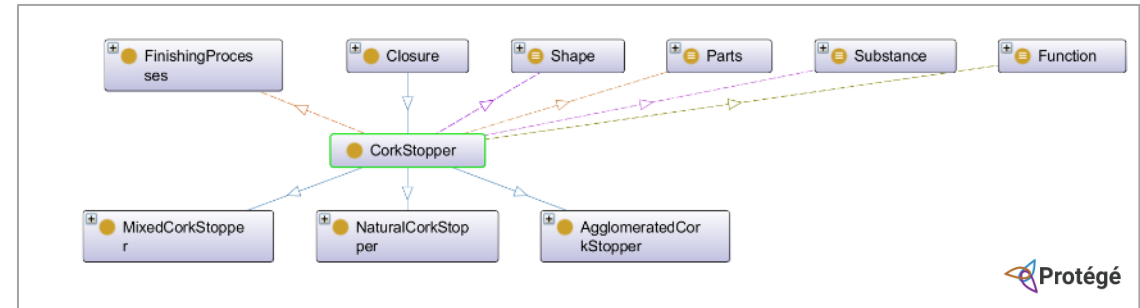
KNOWLEDGE ORGANISATION: NAMING CONCEPTS AND RESTRICTIONS TO BUILD THE ONTOLOGY



KNOWLEDGE ORGANISATION: NAMING RESTRICTIONS BASED ON THE 5 AXES OF ANALYSIS

Conceptual relation identifiers
has_shape
has_process
has_part
has_raw_material
has_substance
has_function

Domain description relations = 5 axes of analysis
hasShape
hasFinishingProcess
hasStructure
IsMadeOf
hasFunction



Protégé

Equivalent To +

SubClass Of +

- (not (FinishingProcesses)) or (hasFinishingProcess some FinishingProcesses)
- Closure
- hasFunction some Function
- hasShape some Shape
- hasStructure some Parts
- isMadeOf some Substance

General class axioms +

SubClass Of (Anonymous Ancestor)

instances +

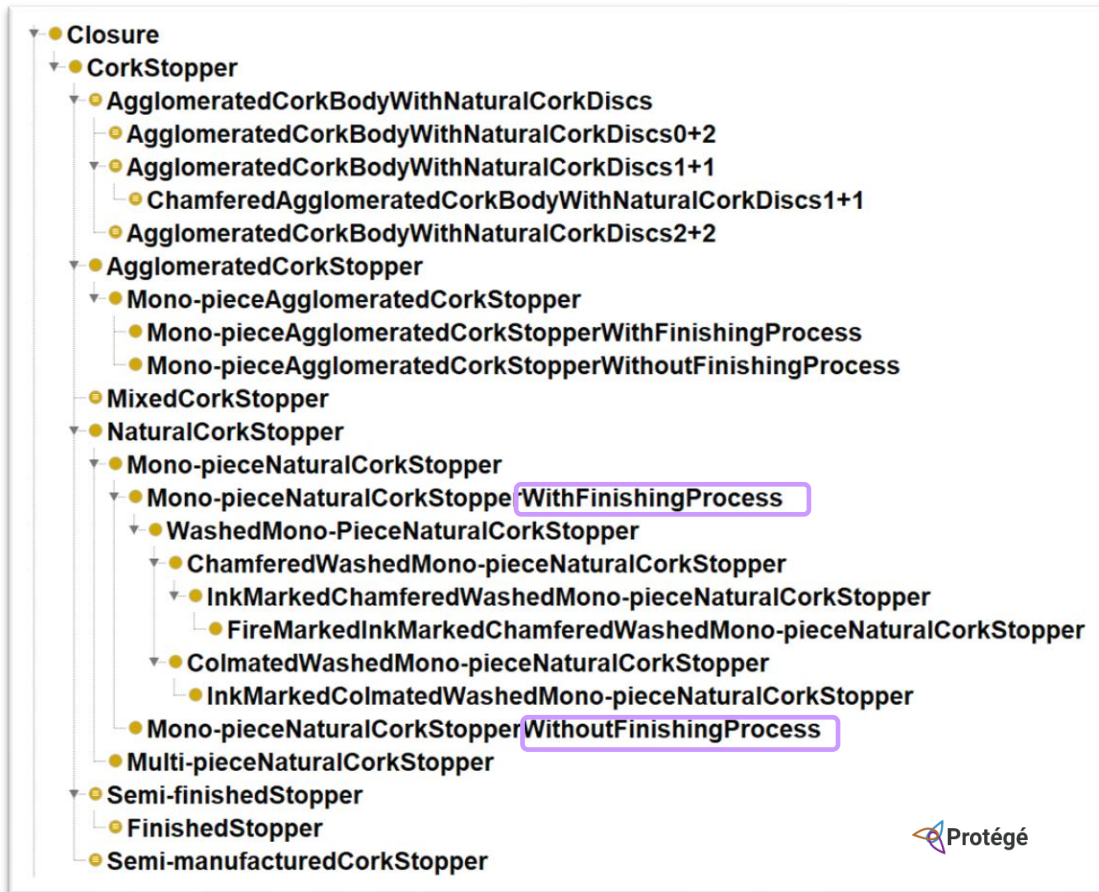
Target for Key +

Disjoint With +

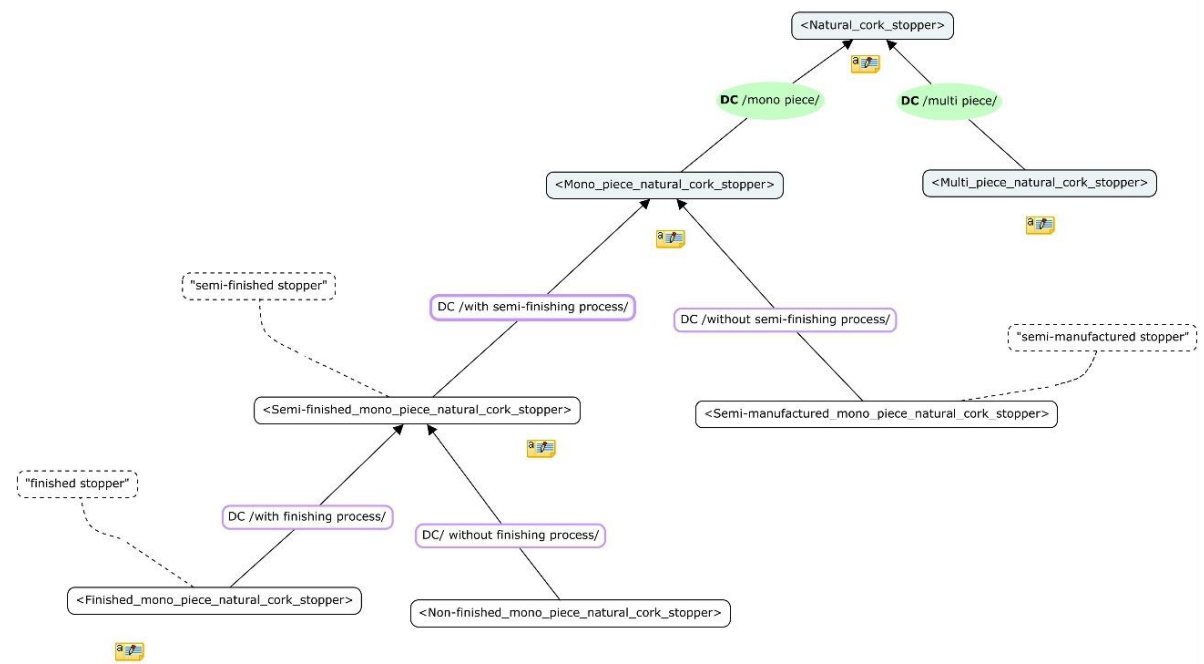
Disjoint Union Of +

Protégé

KNOWLEDGE ORGANISATION: NAMING CONCEPTS ACCORDING TO THE (I) WHOLE SET OF CHARACTERISTICS OR (II) THE STAGE IN THE MANUFACTURE PROCESS



Map 5: natural cork stopper with / without finishing processes



KNOWLEDGE ORGANISATION: NAMING CONCEPTS ACCORDING TO THEIR PURPOSE

Conceptual Map of <Finishing_processes> : the starting point for naming the associative relations [PROCESS-RESULT]

Genus-differentia :
Final Finishing vs. Semi-finishing operations

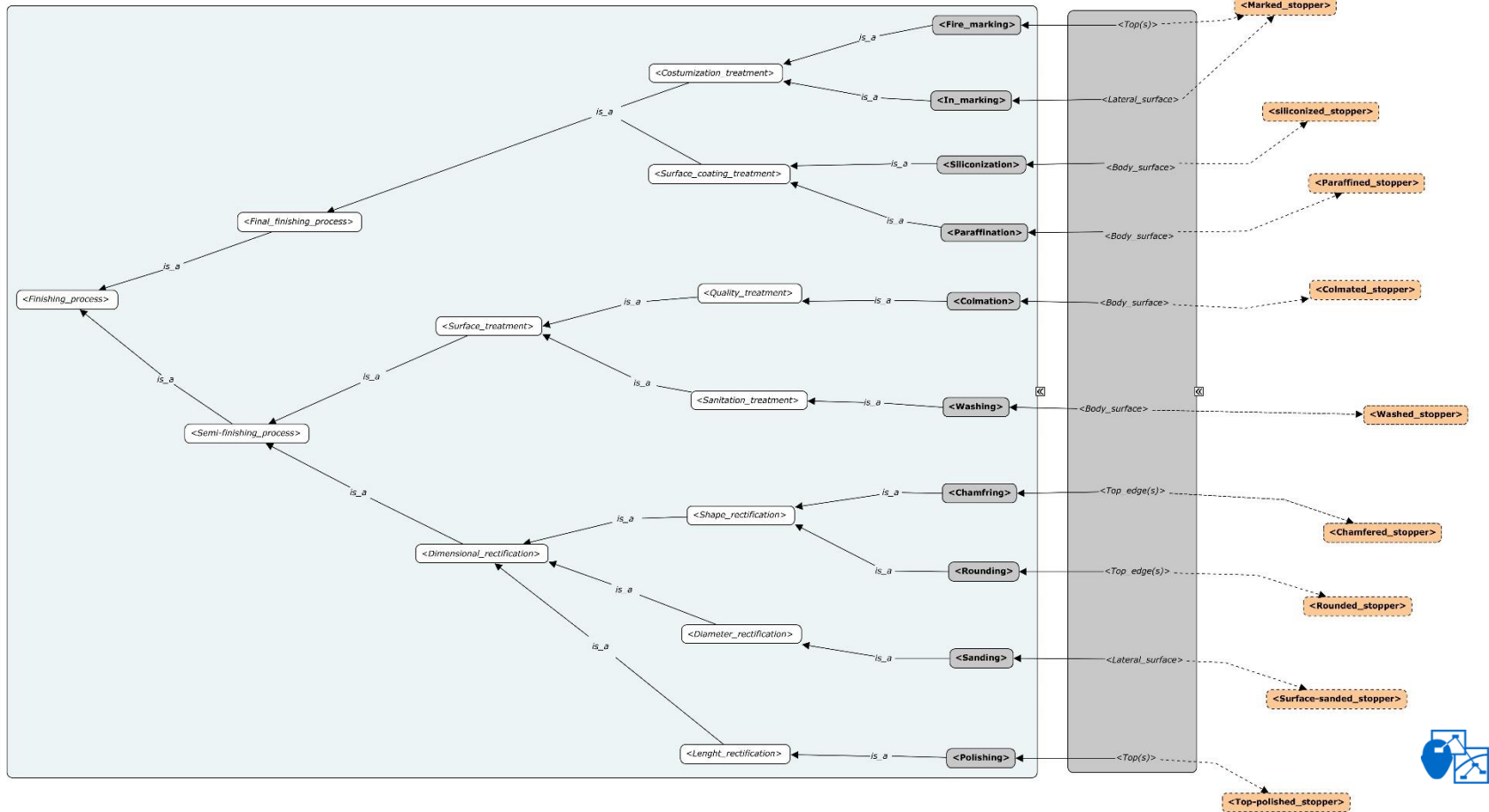
(i) Genus-differentia : Surface treatment operations vs. Dimensional rectification operations
(ii) Differential characteristic (DC) : Purpose

<Operation>

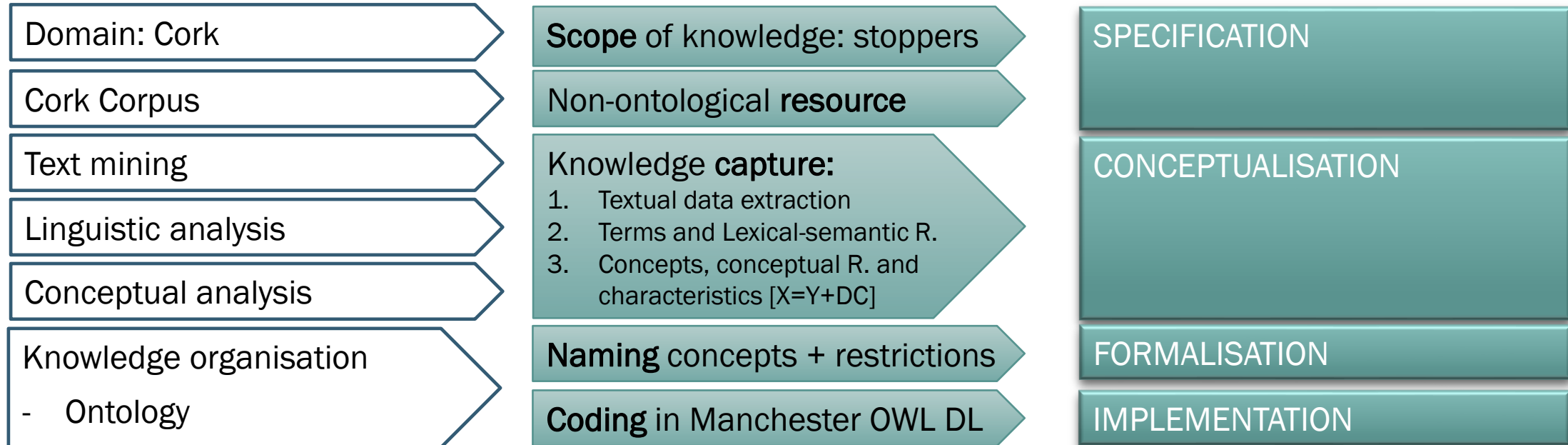
<Location>

<Cork_stopper>

- **FinishingProcesses**
 - **FinalFinishingProcesses**
 - CustomizationTreatment
 - FireMarkinOperation
 - InkMarkingOperation
 - SurfaceCoatingTreatment
 - SurfaceParaffination
 - SurfaceSiliconization
 - **Semi-finishingProcesses**
 - **DimensionalRectification**
 - DiameterRectification
 - SurfaceSandingOperation
 - LengthRectification
 - TopPolishingOperation
 - ShapeRectification
 - EdgeChamferingOperation
 - EdgeRoundingOperation
 - **SurfaceTreatment**
 - QualityTreatment
 - LenticelsColmation
 - SanitationTreatment
 - SurfaceWashing



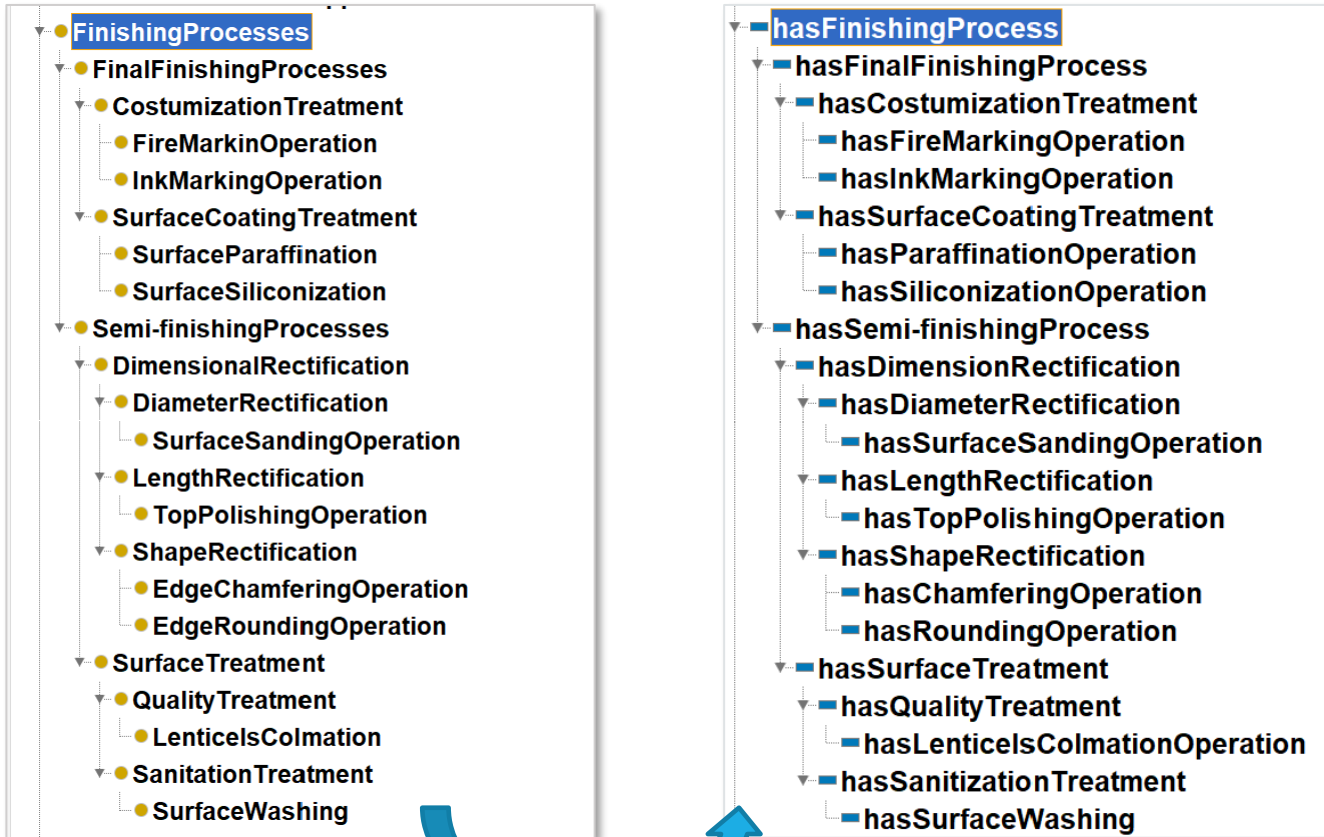
KNOWLEDGE ORGANISATION CODING IN MANCHESTER OWL DL



KNOWLEDGE ORGANISATION: CODING IN MANCHESTER OWL

hasFinishingProcesses

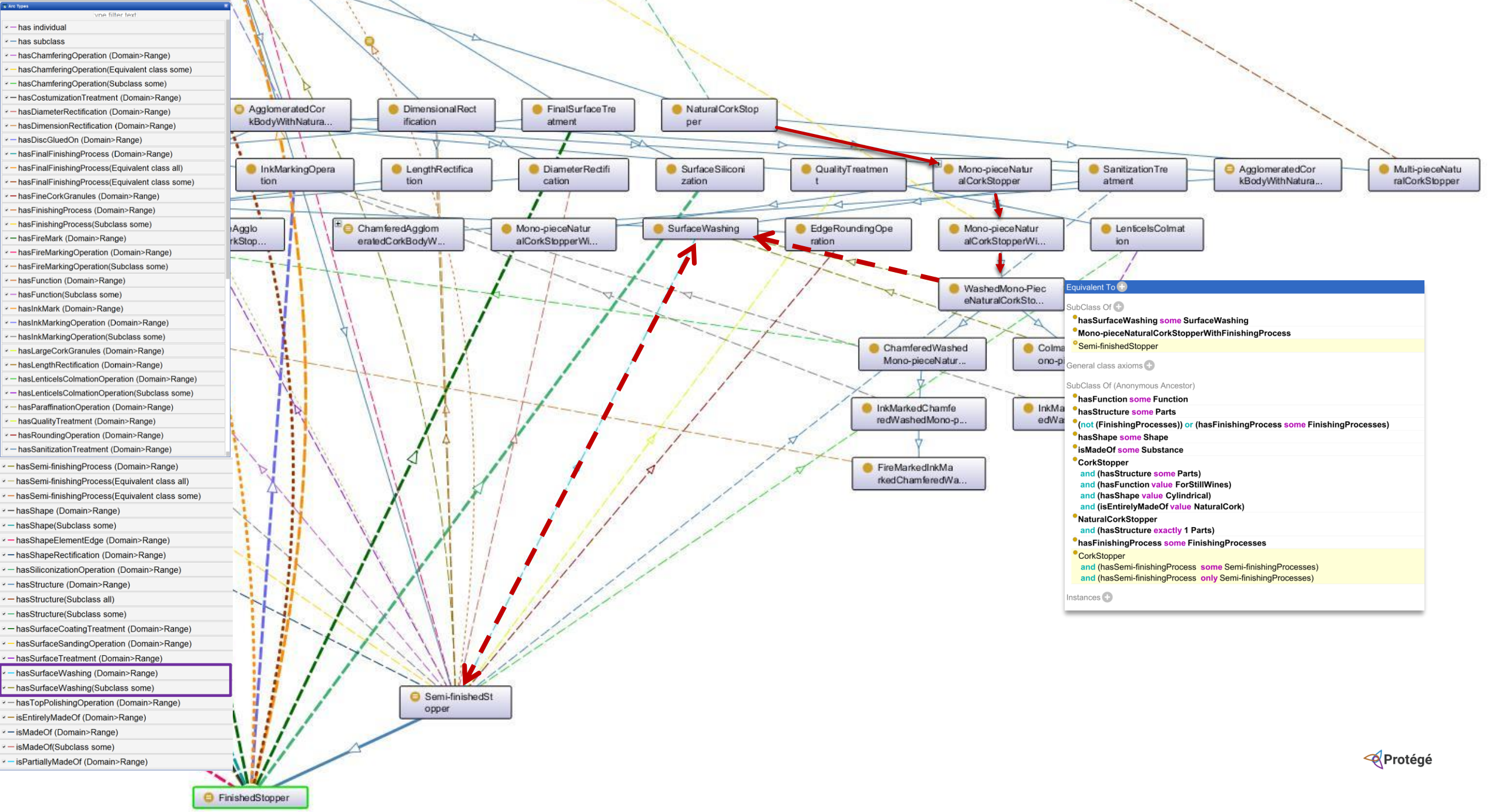
is a restriction used to express the conceptual relation [PROCESS-RESULT]



owl:domain and owl:range restrictions dictate the classification of concepts in the manufacturing process

The screenshot shows the Protégé interface for the class `hasSurfaceWashing`. It displays various axioms and restrictions:

- Equivalent To:** `WashedMono-PieceNaturalCorkStopper` (highlighted in a blue callout box).
- SubProperty Of:** `hasSanitizationTreatment`.
- Inverse Of:** (None listed).
- Domains (intersection):** `Semi-finishedStopper`.
- Ranges (intersection):** `SurfaceWashing`.
- SubClass Of:**
 - `hasSurfaceWashing some SurfaceWashing`
 - `Mono-pieceNaturalCorkStopperWithFinishingProcess`
 - `Semi-finishedStopper` (highlighted in yellow)
- General class axioms:**
 - `hasFunction some Function`
 - `hasStructure some Parts`
 - `(not (FinishingProcesses)) or (hasFinishingProcess some FinishingProcesses)`
 - `hasShape some Shape`
 - `isMadeOf some Substance`
 - `CorkStopper`
 - `and (hasStructure some Parts)`
 - `and (hasFunction value ForStillWines)`
 - `and (hasShape value Cylindrical)`
 - `and (isEntirelyMadeOf value NaturalCork)`
 - `NaturalCorkStopper`
 - `and (hasStructure exactly 1 Parts)`
 - `hasFinishingProcess some FinishingProcesses`
 - `CorkStopper`
 - `and (hasSemi-finishingProcess some Semi-finishingProcesses)`
 - `and (hasSemi-finishingProcess only Semi-finishingProcesses)`



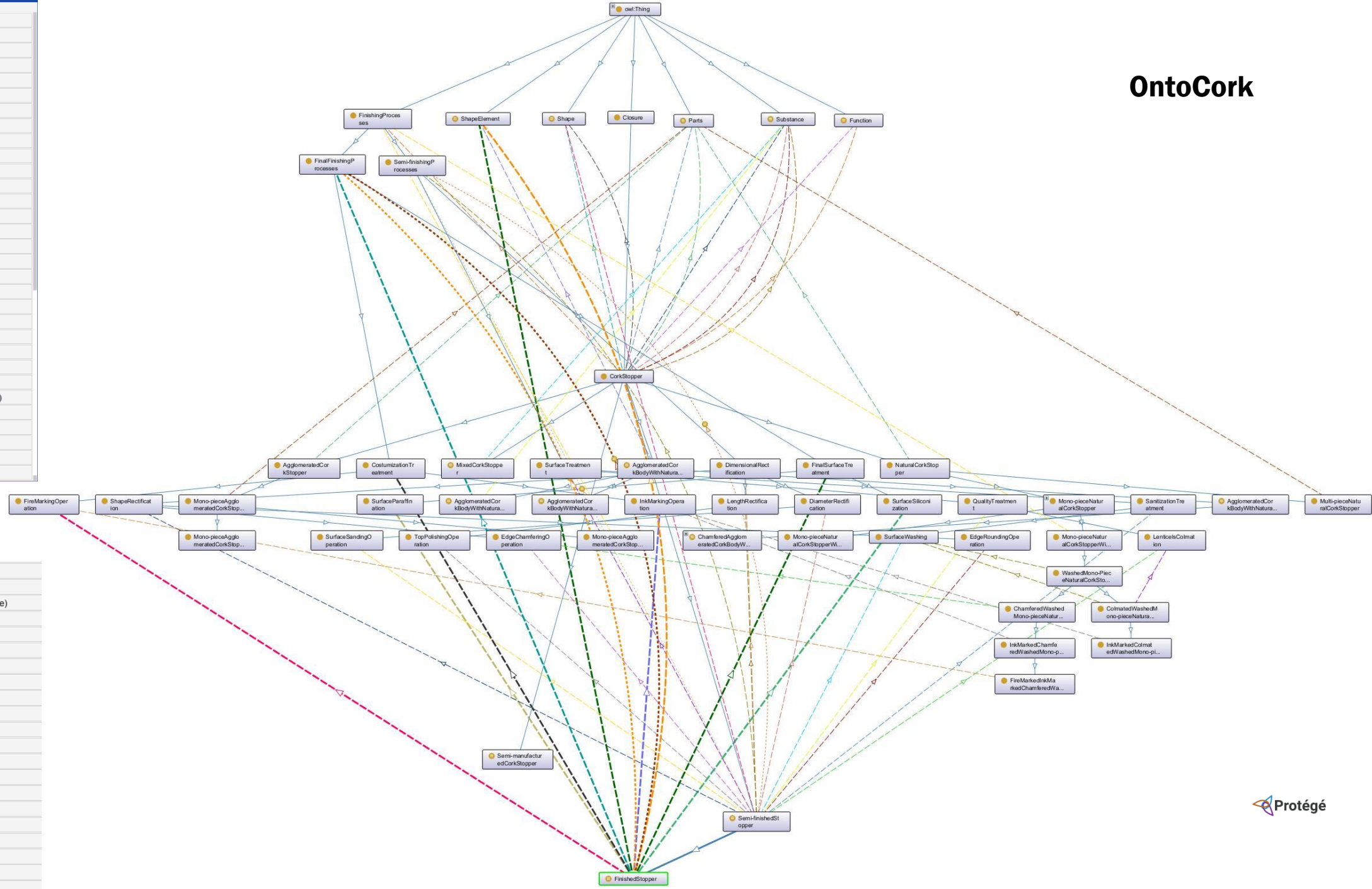
- vne filter text
- ✓ has individual
 - ✓ has subclass
 - ✓ hasChamferingOperation (Domain>Range)
 - ✓ hasChamferingOperation(Equivalent class some)
 - ✓ hasChamferingOperation(Subclass some)
 - ✓ hasCustomizationTreatment (Domain>Range)
 - ✓ hasDiameterRectification (Domain>Range)
 - ✓ hasDimensionRectification (Domain>Range)
 - ✓ hasDiscGluedOn (Domain>Range)
 - ✓ hasFinalFinishingProcess (Domain>Range)
 - ✓ hasFinalFinishingProcess(Equivalent class all)
 - ✓ hasFinalFinishingProcess(Equivalent class some)
 - ✓ hasFineCorkGranules (Domain>Range)
 - ✓ hasFinishingProcess (Domain>Range)
 - ✓ hasFinishingProcess(Subclass some)
 - ✓ hasFireMark (Domain>Range)
 - ✓ hasFireMarkingOperation (Domain>Range)
 - ✓ hasFireMarkingOperation(Subclass some)
 - ✓ hasFunction (Domain>Range)
 - ✓ hasFunction(Subclass some)
 - ✓ hasInkMark (Domain>Range)
 - ✓ hasInkMarkingOperation (Domain>Range)
 - ✓ hasInkMarkingOperation(Subclass some)
 - ✓ hasLargeCorkGranules (Domain>Range)
 - ✓ hasLengthRectification (Domain>Range)
 - ✓ hasLenticelsColmationOperation (Subclass some)
 - ✓ hasParaffinationOperation (Domain>Range)
 - ✓ hasQualityTreatment (Domain>Range)
 - ✓ hasRoundingOperation (Domain>Range)
 - ✓ hasSanitizationTreatment (Domain>Range)
 - ✓ hasSemi-finishingProcess (Domain>Range)
 - ✓ hasSemi-finishingProcess(Equivalent class all)
 - ✓ hasSemi-finishingProcess(Equivalent class some)
 - ✓ hasShape (Domain>Range)
 - ✓ hasShape(Subclass some)
 - ✓ hasShapeElementEdge (Domain>Range)
 - ✓ hasShapeRectification (Domain>Range)
 - ✓ hasSiliconizationOperation (Domain>Range)
 - ✓ hasStructure (Domain>Range)
 - ✓ hasStructure(Subclass all)
 - ✓ hasStructure(Subclass some)
 - ✓ hasSurfaceCoatingTreatment (Domain>Range)
 - ✓ hasSurfaceSandingOperation (Domain>Range)
 - ✓ hasSurfaceTreatment (Domain>Range)
 - ✓ hasSurfaceWashing (Domain>Range)
 - ✓ hasSurfaceWashing(Subclass some)
 - ✓ hasTopPolishingOperation (Domain>Range)
 - ✓ isEntirelyMadeOf (Domain>Range)
 - ✓ isMadeOf (Domain>Range)
 - ✓ isMadeOf(Subclass some)
 - ✓ isPartiallyMadeOf (Domain>Range)

Equivalent To +

- SubClass Of +
 - hasSurfaceWashing some SurfaceWashing
 - Mono-pieceNaturalCorkStopperWithFinishingProcess
 - Semi-finishedStopper
- General class axioms +
- SubClass Of (Anonymous Ancestor)
 - hasFunction some Function
 - hasStructure some Parts
 - (not (FinishingProcesses)) or (hasFinishingProcess some FinishingProcesses)
 - hasShape some Shape
 - isMadeOf some Substance
 - CorkStopper
 - and (hasStructure some Parts)
 - and (hasFunction value ForStillWines)
 - and (hasShape value Cylindrical)
 - and (isEntirelyMadeOf value NaturalCork)
 - NaturalCorkStopper
 - and (hasStructure exactly 1 Parts)
 - hasFinishingProcess some FinishingProcesses
 - CorkStopper
 - and (hasSemi-finishingProcess some Semi-finishingProcesses)
 - and (hasSemi-finishingProcess only Semi-finishingProcesses)
- Instances +

OntoCork

- vne filter text
- has individual
 - has subclass
 - hasChamferingOperation (Domain>Range)
 - hasChamferingOperation(Equivalent class some)
 - hasChamferingOperation(Subclass some)
 - hasCustomizationTreatment (Domain>Range)
 - hasDiameterRectification (Domain>Range)
 - hasDimensionRectification (Domain>Range)
 - hasDiscGluedOn (Domain>Range)
 - hasFinalFinishingProcess (Domain>Range)
 - hasFinalFinishingProcess(Equivalent class all)
 - hasFinalFinishingProcess(Equivalent class some)
 - hasFineCorkGranules (Domain>Range)
 - hasFinishingProcess (Domain>Range)
 - hasFinishingProcess(Subclass some)
 - hasFireMark (Domain>Range)
 - hasFireMarkingOperation (Domain>Range)
 - hasFireMarkingOperation(Subclass some)
 - hasFunction (Domain>Range)
 - hasFunction(Subclass some)
 - hasInkMark (Domain>Range)
 - hasInkMarkingOperation (Domain>Range)
 - hasInkMarkingOperation(Subclass some)
 - hasLargeCorkGranules (Domain>Range)
 - hasLengthRectification (Domain>Range)
 - hasLenticelsColmationOperation (Domain>Range)
 - hasParaffinationOperation (Domain>Range)
 - hasQualityTreatment (Domain>Range)
 - hasRoundingOperation (Domain>Range)
 - hasSanitizationTreatment (Domain>Range)



- hasSemi-finishingProcess (Domain>Range)
- hasSemi-finishingProcess(Equivalent class all)
- hasSemi-finishingProcess(Equivalent class some)
- hasShape (Domain>Range)
- hasShape(Subclass some)
- hasShapeElementEdge (Domain>Range)
- hasShapeRectification (Domain>Range)
- hasSilicizationOperation (Domain>Range)
- hasStructure (Domain>Range)
- hasStructure(Subclass all)
- hasStructure(Subclass some)
- hasSurfaceCoatingTreatment (Domain>Range)
- hasSurfaceSandingOperation (Domain>Range)
- hasSurfaceTreatment (Domain>Range)
- hasSurfaceWashing (Domain>Range)
- hasSurfaceWashing(Subclass some)
- hasTopPolishingOperation (Domain>Range)
- isEntirelyMadeOf (Domain>Range)
- isMadeOf (Domain>Range)
- isMadeOf(Subclass some)
- isPartiallyMadeOf (Domain>Range)

FROM FORMAL OWL TO A LESS FORMAL MODEL: SKOS

For triples involving the `rdf:type` property, the RDF/XML syntax allows a shortened form to model multilingual SKOS labels and link resources to the concept

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">
```

```
<skos:Concept
  rdf:about="http://www.clunl.fcsh.unl.pt/OntoCork#ColmatedWashedMono-pieceNaturalCorkStopper">
```

```
<skos:prefLabel xml:lang="pt">rolha de cortiça natural colmatada</skos:prefLabel>
<skos:prefLabel xml:lang="en">colmated natural cork stopper</skos:prefLabel>
<skos:prefLabel xml:lang="fr">bouchon en liège naturel colmaté</skos:prefLabel>
<skos:altLabel xml:lang="pt">rolha colmatada</skos:altLabel>
```

```
<skos:definition xml:lang="pt">rolha de cortiça natural submetida a operação de colmatagem</skos:definition>
```

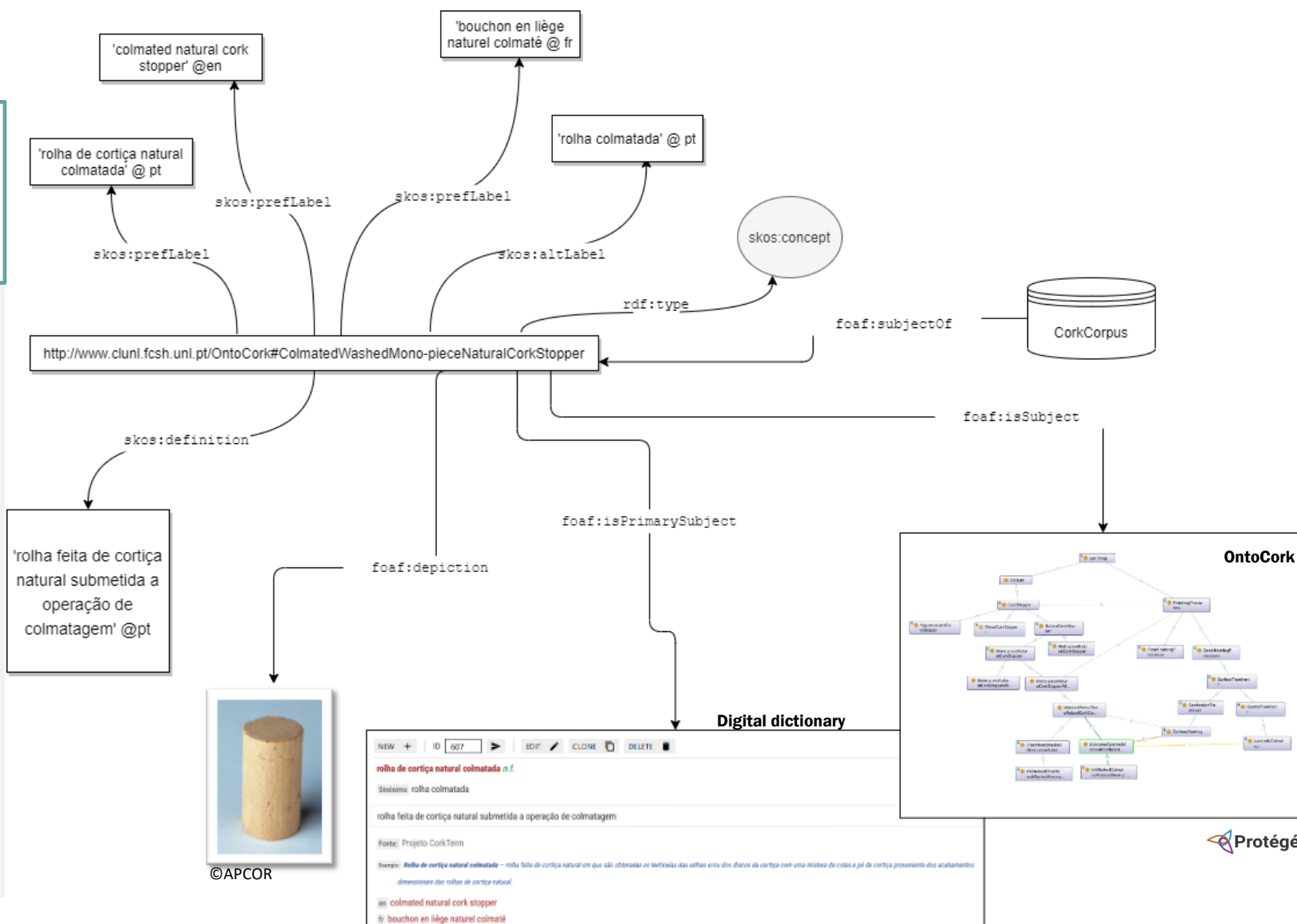
```
<foaf:depiction rdf:resource="https://www.apcor.pt/wp-content/uploads/2015/09/colmatada.jpg">
```

```
<foaf:isPrimarySubject
  rdf:resource="https://www.lexonomy.eu/k4ysn6um/edit/entry/">
```

```
<foaf:isSubject rdf:resource="http://www.clunl.fcsh.unl.pt/OntoCork/">
```

```
<foaf:SubjectOf rdf:resource="http://www.clunl.fcsh.unl.pt/CorkCorpus/">
```

```
</skos:Concept>
</rdf:RDF>
```



FINAL REMARKS

- Linking data and metadata, not only within the same lexicographic resource but also between different resources, is increasingly relevant in the Web of Data, which is based on the use of RDF, URIs and others, which allows users to search and retrieve information.
- In this line of thought, ontologies are suitable for such purposes since they provide several useful features for intelligent systems, as well as for knowledge representation in general and the knowledge engineering process.
- An ontology provides a vocabulary and a machine-processable common understanding of the concepts that terms denote. The meanings of terms in an ontology can be unambiguously communicated between users and applications, as the semantics used are independent of reader and context.



ACKNOWLEDGMENTS

OntoCork was developed within the PhD thesis project, funded by the FCT – Fundação para a Ciência e a Tecnologia, Portugal – through the PhD scholarship PD/BD/113972/2015.