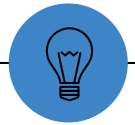
SNOMED CT Translation User Group Meeting
(28.06.2022)

Definitions in SNOMED CT: the role of Terminology and Knowledge Organisation



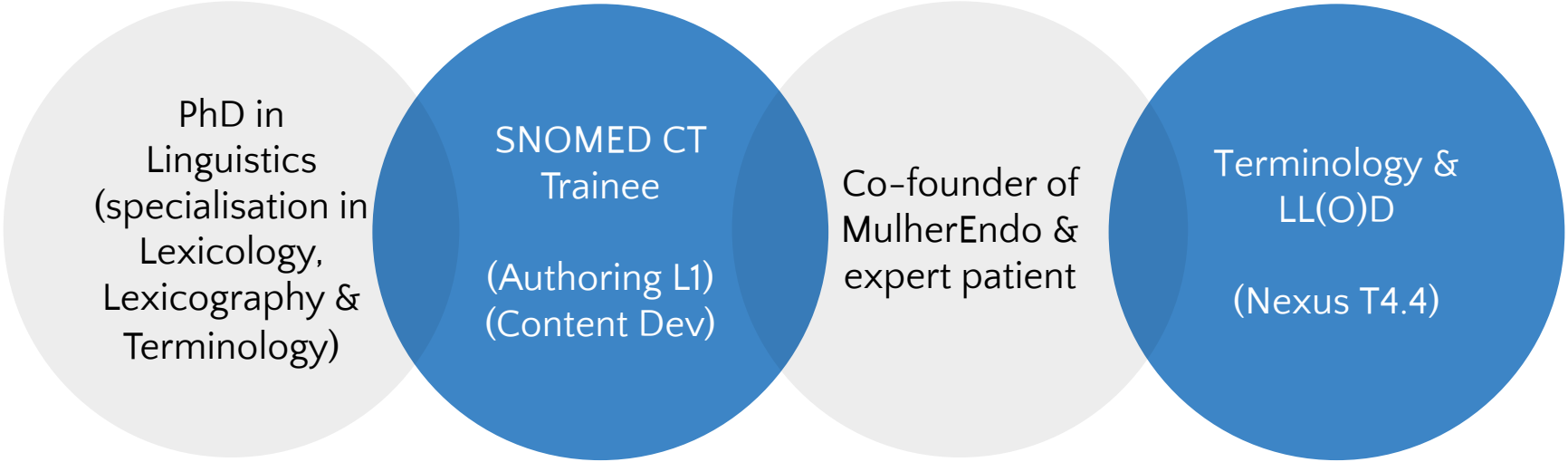
Sara Carvalho
(NOVA CLUNL & CLLC-UA)

CLUNL
CENTRO DE LINGUÍSTICA DA UNIVERSIDADE NOVA DE LISBOA

universidade de aveiro

cllc centro de línguas, literaturas e culturas

My research interest in healthcare



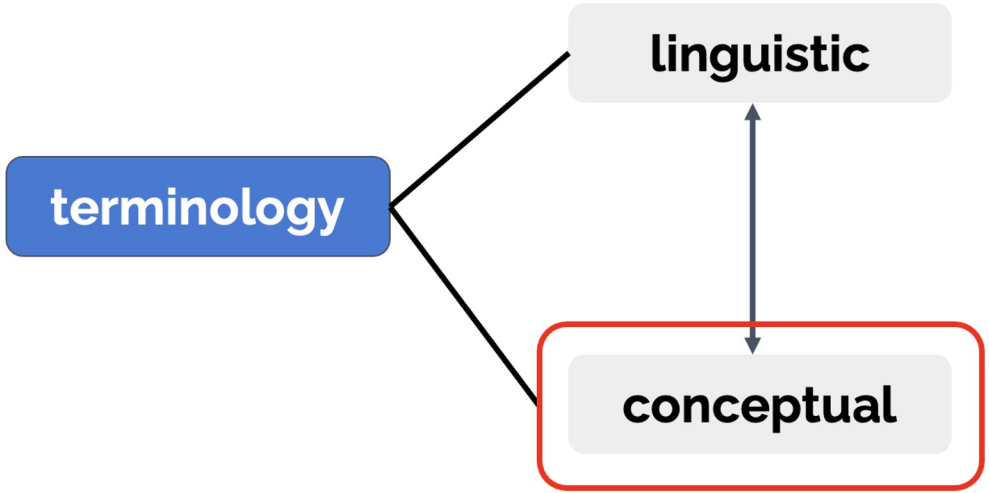
Outline

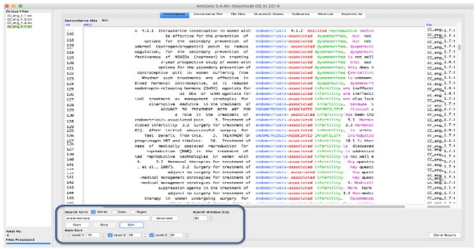
- **Background**
  - Theoretical and methodological framework
  - Definition(s) in Terminology
- **Definition(s) in SNOMED CT**
  - A (very) brief state of the art
  - Analysing examples
- **Definition(s) in the EndoTerm project**
- **Concluding remarks (and prompts for discussion)**

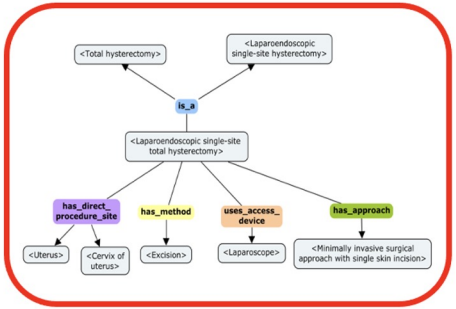
1 Background

(Costa, 2013; Santos & Costa, 2015; Carvalho, 2018; Costa et al., 2019)

Double-dimensional approach to Terminology

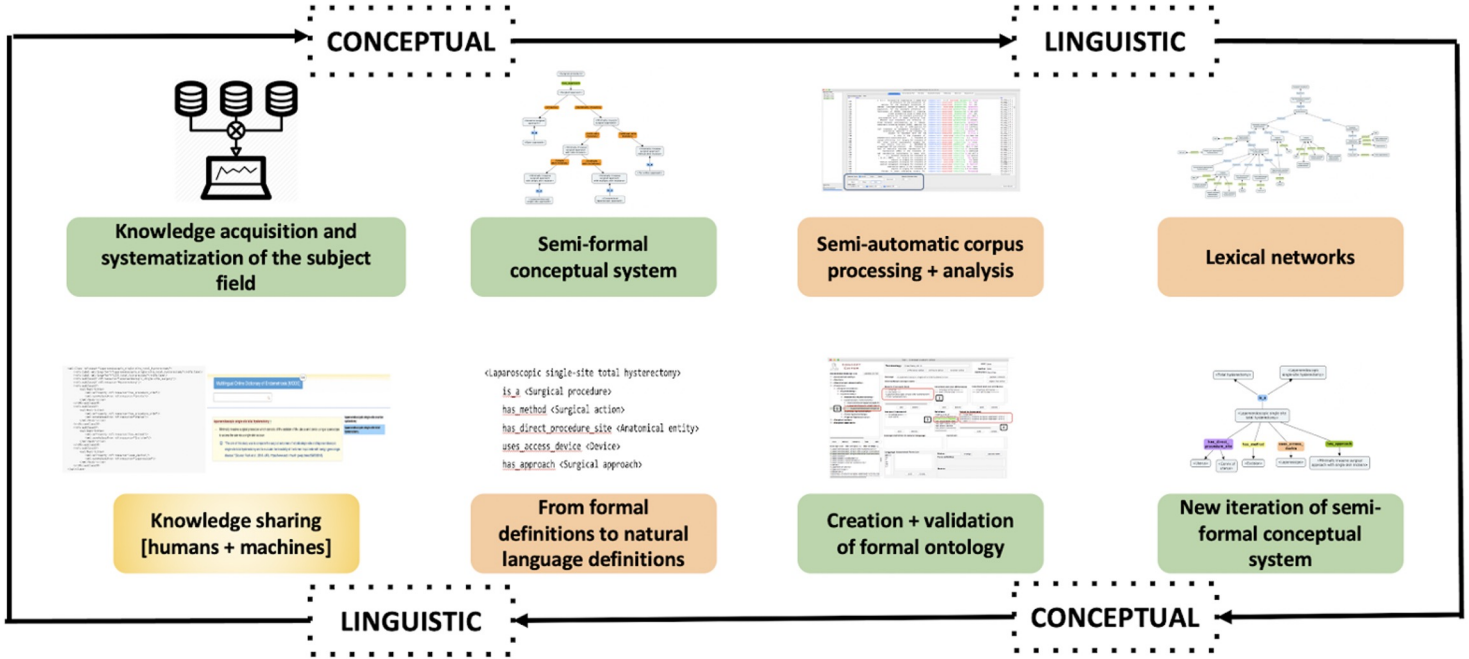






(Santos & Costa, 2015; Carvalho, 2018)

Mixed approach in terminology work



Approaches to definition(s) in Terminology

- Aristotelian tradition (Smith, 2018)
  - *definiendum = genus proximum + differentia specifica*
- ISO 1087: 2019 & ISO 704: 2009
  - **Definition**: "representation of a concept by an expression that describes it and differentiates it from related concepts"
  - **Intensional definition**: "conveys the intension of a concept" > "immediate generic concept" + "delimiting characteristics"
  - **Guidelines** for drafting intensional definitions (e.g. Felber, 1984; Sager, 1990; Suonuuti, 2001; Löckinger et al., 2015)

Approaches to definition(s) in Terminology

- Going beyond hierarchically-based definitions > the role of **non-hierarchical conceptual relations** (cf. the work by Nuopponen, 2010, 2011, 2014, 2018) in the drafting of NL definitions
- **Definitional templates > can "guide formulation of definitions"** (Martin, 1998)
  - Frame-Based Terminology (Faber, 2012, 2015)
    - Definitions as "mini-knowledge representations" (Faber, 2002)
    - Several domains: healthcare (García de Quesada et al., 2002); environment (San Martín & León-Araúz, 2013); tourism (Durán-Muñoz, 2016)
- "Despite the centrality of [natural language] definitions, the activity of creating definitions must still be realized **manually**; this is **time-consuming**, **costly**, requires **uncommon expertise**, and is prone to all kinds of **inconsistencies.**" (Seppälä, 2015: 24)

2

Definition(s) in SNOMED CT

A (very) brief state of the art

- Survey of SNOMED CT direct users (Elhanan et al., 2011): **63% claim textual definitions would be useful** > The **US version** of SNOMED CT contains only **4,372 text definitions**, covering 2,608 diseases, corresponding to **1.3% of all SNOMED CT concepts** and **3.3% of disease concepts**." (Lyudovyk & Weng, 2019)
- **Textual definition**:
  - "textual description applied to **some** SNOMED CT concepts that provides **additional information** about the intended meaning or usage of the concept. Definitions are **not mandated** and are **considered for addition on a case-by-case basis**, and **if required, to differentiate a concept from its related concepts**. Adding a definition to a concept **provides additional clarity on its context of use**." (SNOMED CT Editorial Guide, 2022) [highlights are my responsibility]
  - It **enhances** the definition provided by the modeled relationships but it **should neither be confused with the formal logic definitions** of concepts expressed using OWL axioms or defining relationships **nor be contradictory** to the modelling. (adapted from SNOMED CT Ed. Guide & Glossary)

A (very) brief state of the art

1. From texts to formal definitions

- Ma & Distel (2013)
- Petrova et al. (2015)
- Prokhorov et al. (2019)

Automatic generation of formal concept definitions from textual ones (SNOMED CT as a 'model' for formal definitions)

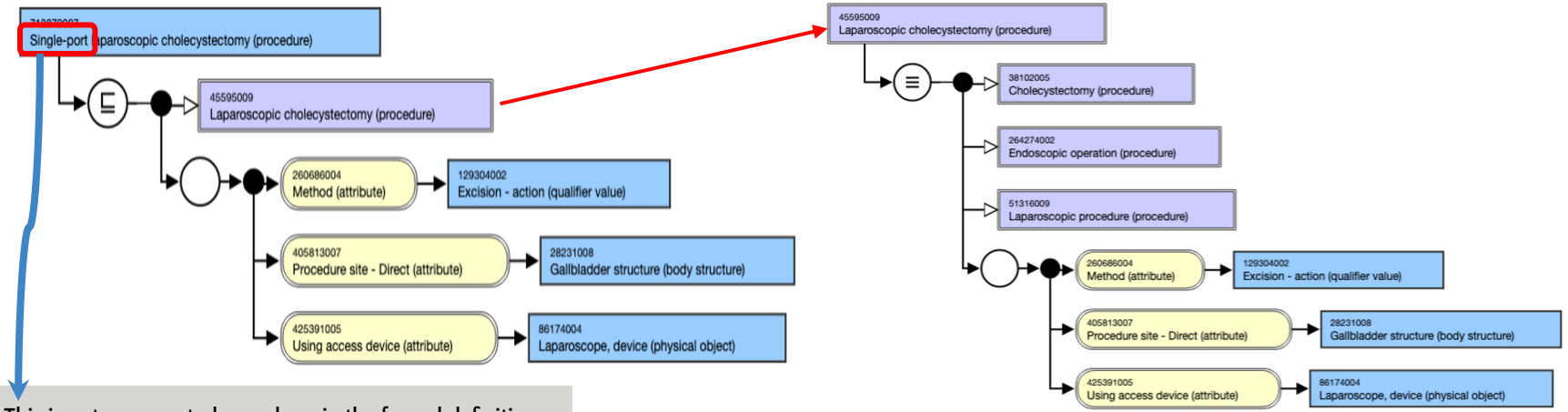
2. From formal definitions to texts

- Jiang et al. (2013) > basis for **ICD-11** textual definitions
- Liang et al. (2013) > **OntoVerbal** (SNOMED CT) > Protégé plugin - not available in recent versions
- Lyudovyk & Weng (2019) > **SNOMEDtxt** (https://sno2eng.shinyapps.io/sno2Eng/)

Spelling inconsistencies: cf. single-port vs. single port
(714031000 |Single port laparoscopic nephrectomy (procedure)|

Analysing examples

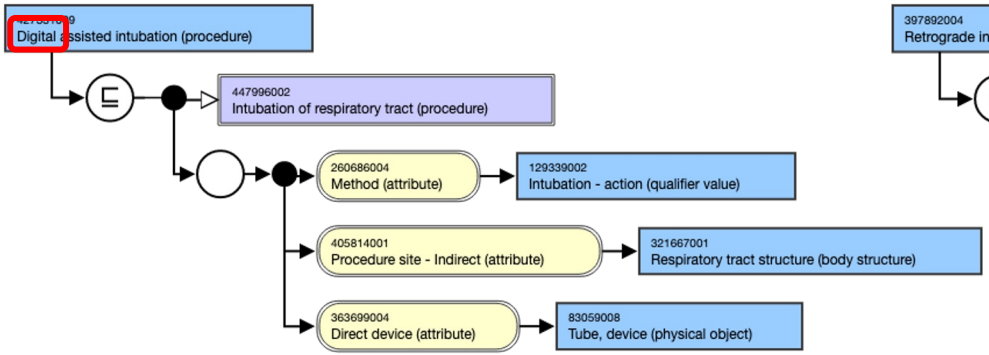
- **Case Study 1:** same formal definition so technically there is no difference between the 2 concepts (yet, there is > delimiting characteristic does not appear anywhere; no NL definition
  - **713872007 |Single-port laparoscopic cholecystectomy (procedure)|** > **inferred** view; Int. Ed. **2022-05-31**

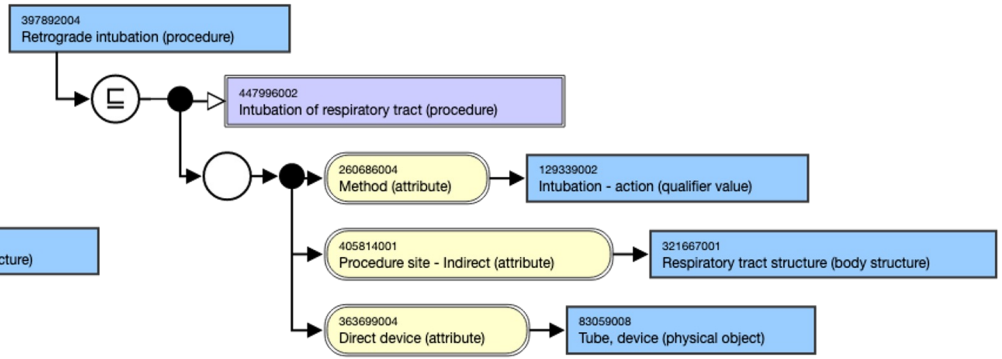


This is not represented anywhere in the formal definition > challenges will arise when drafting the NL definition

Analysing examples

- **Case Study 2:** same formal definition so technically there is no difference between the 2 concepts (or is there?); NL definition in one of them BUT does not match the logical definition
  - **427331009 |Digital assisted intubation (procedure)| vs. 397892004 |Retrograde intubation (procedure)|** (**inferred** view; Int. Ed. **2022-05-31)**



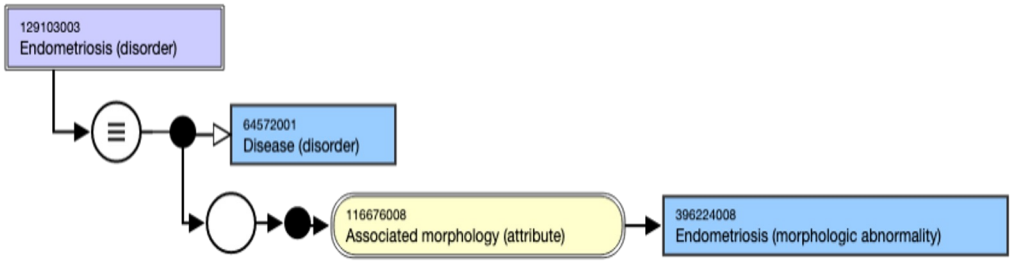


D ★ Digital intubation uses the index and middle fingers to blindly direct the ET tube into the larynx. (added this year)

- What is an ET tube?
- What is the difference (if any) between one concept and the other after all? What does retrograde entail, then?

Analysing examples

- **Case Study 3:** formal definition needs to be readjusted to incorporate recent advances in the area; no NL definition
  - **129103003 |Endometriosis (disorder)| (inferred** view; Int. Ed. **2022-05-31)**



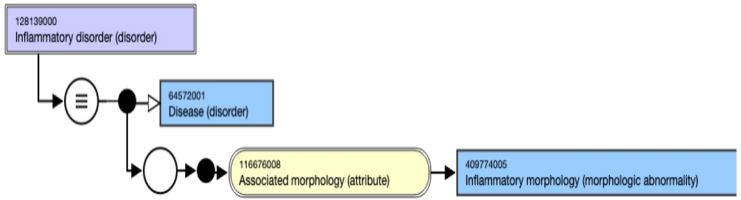
Suggestion:

=== **128139000 |Inflammatory disorder (disorder)|** :

{ 116676008 |Associated morphology (attribute)| = 396224008 |Endometriosis (morphologic abnormality)| }

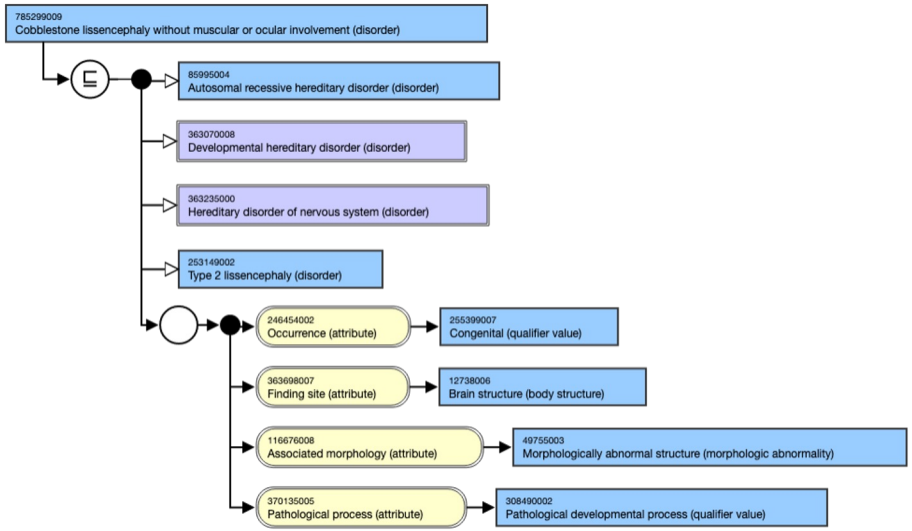
Need to incorporate <Inflammation>
(cf. Consensus Paper "An International Terminology for Endometriosis", Nov 2021)

"A disease characterised by the presence of endometrium-like epithelium and/or stroma outside the endometrium and myometrium, usually with an associated inflammatory process"



Analysing examples

- **Case Study 4:** primitive concept (formal definition may not be stable); NL definition exists but does not match formal one
  - **785299009 |Cobblestone lissencephaly without muscular or ocular involvement (disorder)|** (**inferred** view; Int. Ed. **2022-05-31**)



A rare genetic cobblestone lissencephaly disease with characteristics of the presence of a constellation of brain malformations, including cortical gyral and sulcus anomalies, white matter signal abnormalities, cerebellar dysplasia and brainstem hypoplasia, existing alone or in conjunction with minimal muscular and ocular abnormalities, typically manifesting with severe developmental delay, increased head circumference, hydrocephalus and seizures. There is evidence the disease is caused by homozygous or compound heterozygous mutation in the LAMB1 gene on chromosome 7q31.

Analysing examples

- **Case Study 4:** primitive concept (formal definition may not be stable); NL definition exists but does not match formal one
  - **785299009 |Cobblestone lissencephaly without muscular or ocular involvement (disorder)|** (**inferred** view; Int. Ed. 2022-05-31)

A rare genetic cobblestone lissencephaly disease with characteristics of the presence of a constellation of brain malformations, including cortical gyral and sulcus anomalies, white matter signal abnormalities, cerebellar dysplasia and brainstem hypoplasia, existing alone or in conjunction with minimal muscular and ocular abnormalities, typically manifesting with severe developmental delay, increased head circumference, hydrocephalus and seizures. There is evidence the disease is caused by homozygous or compound heterozygous mutation in the LAMB1 gene on chromosome 7q31.

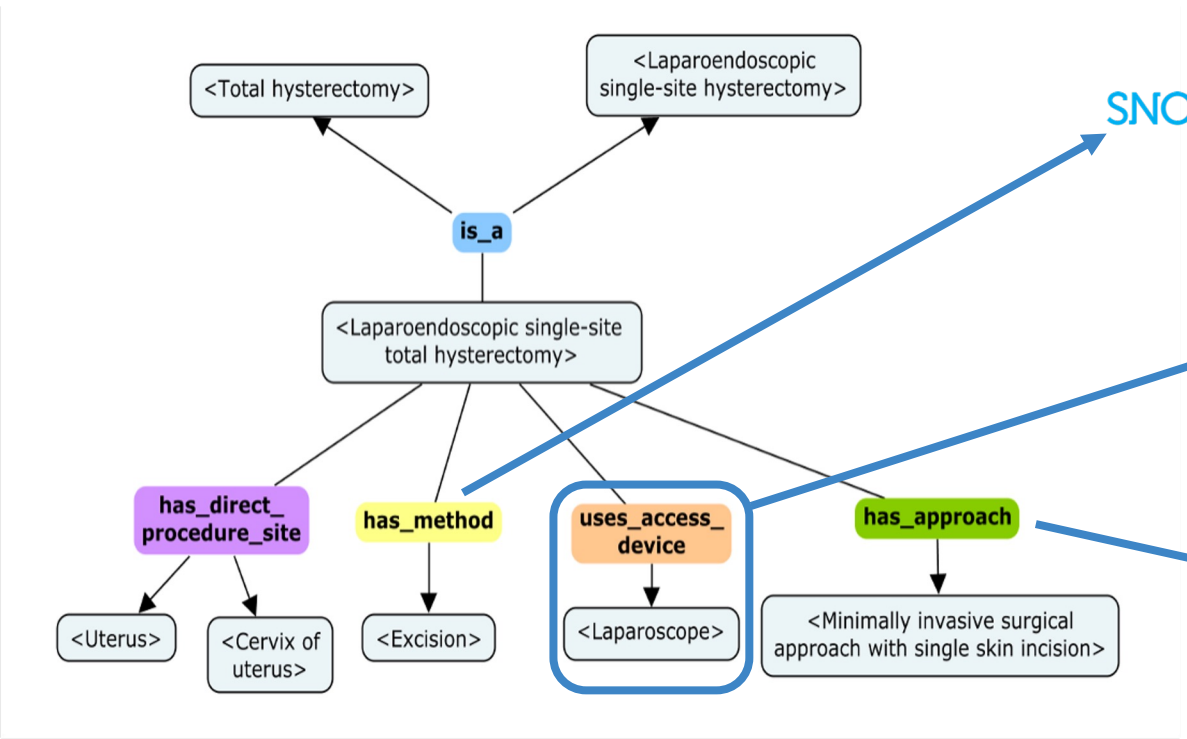
- The characteristic rare is not in the formal definition.
- The types of malformations are not in the formal definition.
- If it is in conjunction with minimal abnormalities, why does the FSN state without muscular or ocular involvement?
- Includes signs/symptoms (not in the formal definition)
- The causes are much more detailed in the NL definition than in the formal definition > source?

3

Definition(s) in the EndoTerm project

Terminology and ontologies

An example from EndoTerm (Carvalho, 2018)



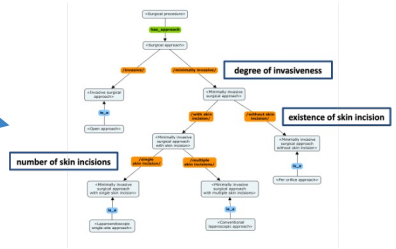
SNOMED CT

Unified Medical Language System ®

<Laparoscopic hysterectomy> (procedure)

has_procedure_site

<Laparoscope> (physical object)



```
<owl:Class rdf:about="Laparoendoscopic_single-site_total_hysterectomy">
    <rdfs:label xml:lang="en">"laparoendoscopic_single-site_total_hysterectomy"</rdfs:label>
    <rdfs:label xml:lang="en">"LESS_total_hysterectomy"</rdfs:label>
    <rdfs:subClassOf rdf:resource="Laparoendoscopic_single-site_surgery"/>
    <rdfs:subClassOf rdf:resource="Hysterectomy"/>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="has_procedure_site"/>
            <owl:someValuesFrom rdf:resource="Cervix"/>
        </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="has_procedure_site"/>
            <owl:someValuesFrom rdf:resource="Uterus"/>
        </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="has_method"/>
            <owl:someValuesFrom rdf:resource="Excision"/>
        </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="uses_device"/>
            <owl:someValuesFrom rdf:resource="Laparoscope"/>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
```

Formal concept definition (RDF/OWL)

Natural language definition

```
<Laparoscopic single-site total hysterectomy>
    is_a <Surgical procedure>
    has_method <Surgical action>
    has_direct_procedure_site <Anatomical entity>
    uses_access_device <Device>
    has_approach <Surgical approach>
```

Proposal: minimally invasive surgical procedure which consists of the excision of the uterus and cervix using a laparoscope to access the site via a single skin incision.

4 Concluding remarks

Some final remarks + prompts for discussion

- In SNOMED CT, “concepts are defined in 3 different ways” (cf. SNOMED CT Translation Guidelines):
  - the **Fully Specified Name** (FSN) > human-readable form
  - the **formal concept definition** > computer-processable representation
  - **textual definition** > describe the meaning of a concept in natural language
- Ideally, **the 3 must align** > but the examples have shown challenges in all of them:
  - **Ambiguity** in the FSN
  - Formal concept definition **lacking one or more delimiting characteristic**
  - Textual definitions containing **extra** information that is nowhere to be found in the formal definition

Some final remarks + prompts for discussion

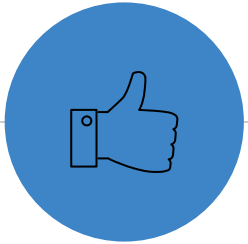
- What's **next** for **textual definitions** in SNOMED CT?
  - Are they a priority for SNOMED CT at the moment? > MeSH, ICD-11, Disease Ontology and other resources have them… > If so, for which **target group(s)**?
  - What would be the **scope**? Would the parts of a given textual definition **match** the parts of the corresponding logical definition (for consistency purposes) or would it go beyond that (i.e. would extra information be included and based on what criteria)?
  - For which concepts? **Fully defined** concepts only or also **primitive** concepts (but how, if the concept's formal definition is not 100% stable)?
  - **Who** would draft these textual definitions and what would be the **workflow**? > importance of methodological framework + **guidelines** (cf. OBO Foundry's example - Seppälä et al., 2017) + **automation**?

Some final remarks + prompts for discussion

- Added value of the synergies involving **Terminology** – with its double-dimensional nature – and **ontologies** in what concerns knowledge representation, organization and sharing:
  - **Logical consistency** (term formation + drafting of natural language definitions – cf. Carvalho, 2018 and ongoing work)
  - More **effective communication** among the various stakeholders (HC providers but also terminologists, translators… and what about **patients**?)
  - Foundation for **interoperability** – both at the conceptual and linguistic level

References (selection)

- Carvalho, S. (2018). *A terminological approach to knowledge organization within the scope of endometriosis: the EndoTerm project*. PhD. Lisboa: FCSH/UNL. https://run.unl.pt/handle/10362/49745
- Durán-Muñoz, I. (2016). "Producing frame-based definitions". *Terminology* 22:2 (2016). Amsterdam: John Benjamins, pp. 224-250.
- Faber, P. (ed.) (2012). *A Cognitive Linguistics View of Terminology and Specialized Language.* Berlin, Boston: De Gruyter Mouton.
- Lyudovyk, O. and Weng, C. (2019). "SNOMEDtxt: Natural Language Generation from SNOMED Ontology". In *Stud Health Technol Inform*. 2019 August 21; 264: 1263–1267. doi:10.3233/SHTI190429.
- Santos, C. and Costa, R. (2015). "Domain specificity: Semasiological and onomasiological knowledge representation". In: Hendrik J. Kockaert & Frieda Steurs (eds.), *Handbook of Terminology*, Volume 1. Amsterdam / Philadelphia: John Benjamins Publishing Company, pp. 153-179.
- Seppälä, S., Ruttenberg, A. and Smith, B. (2017). "Guidelines for writing definitions in ontologies". *Ci.Inf.,* Brasília, DF, v.46 n.1, p.73-88, jan./abr. 2017, pp. 73-88.



Thank you!

Feel free to contact me:

- sara.carvalho@ua.pt
- https://orcid.org/0000-0002-7501-5405
- https://www.cienciavitae.pt//en/D518-A669-DE57