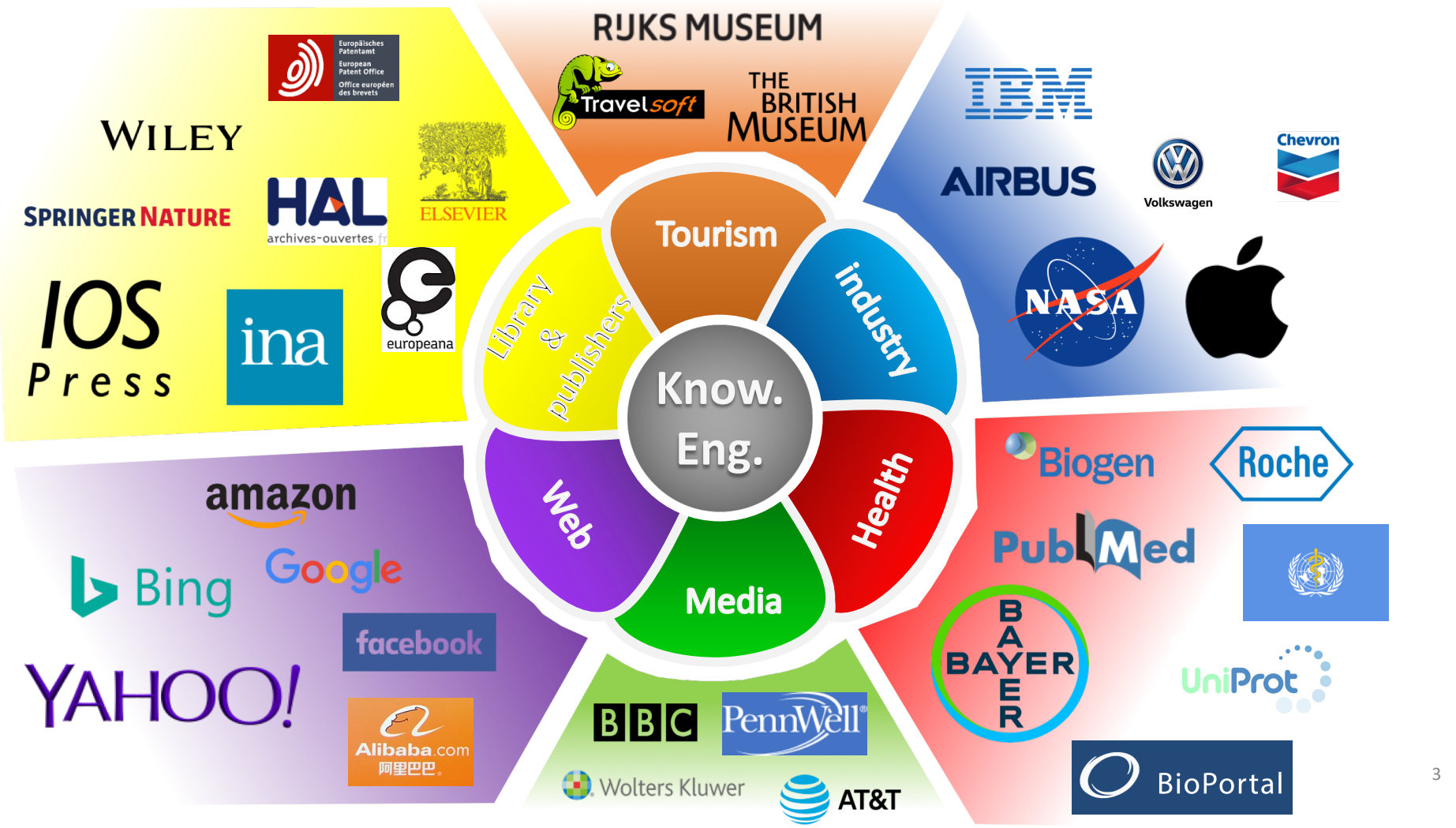


Dynaccurate

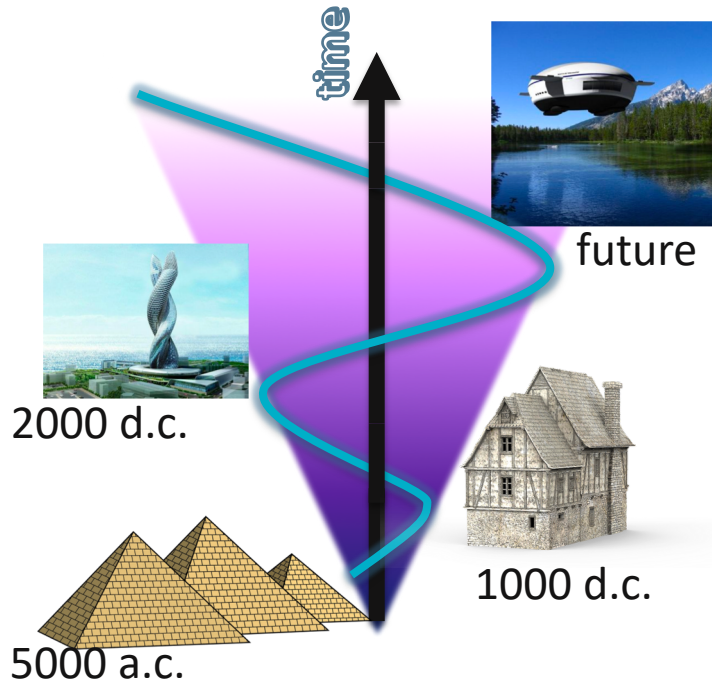
AI to keep ontology mappings and semantic annotations up to date

Marcos Da Silveira
marcos.dasilveira@list.lu

Cédric Pruski
cedric.prusk@list.lu



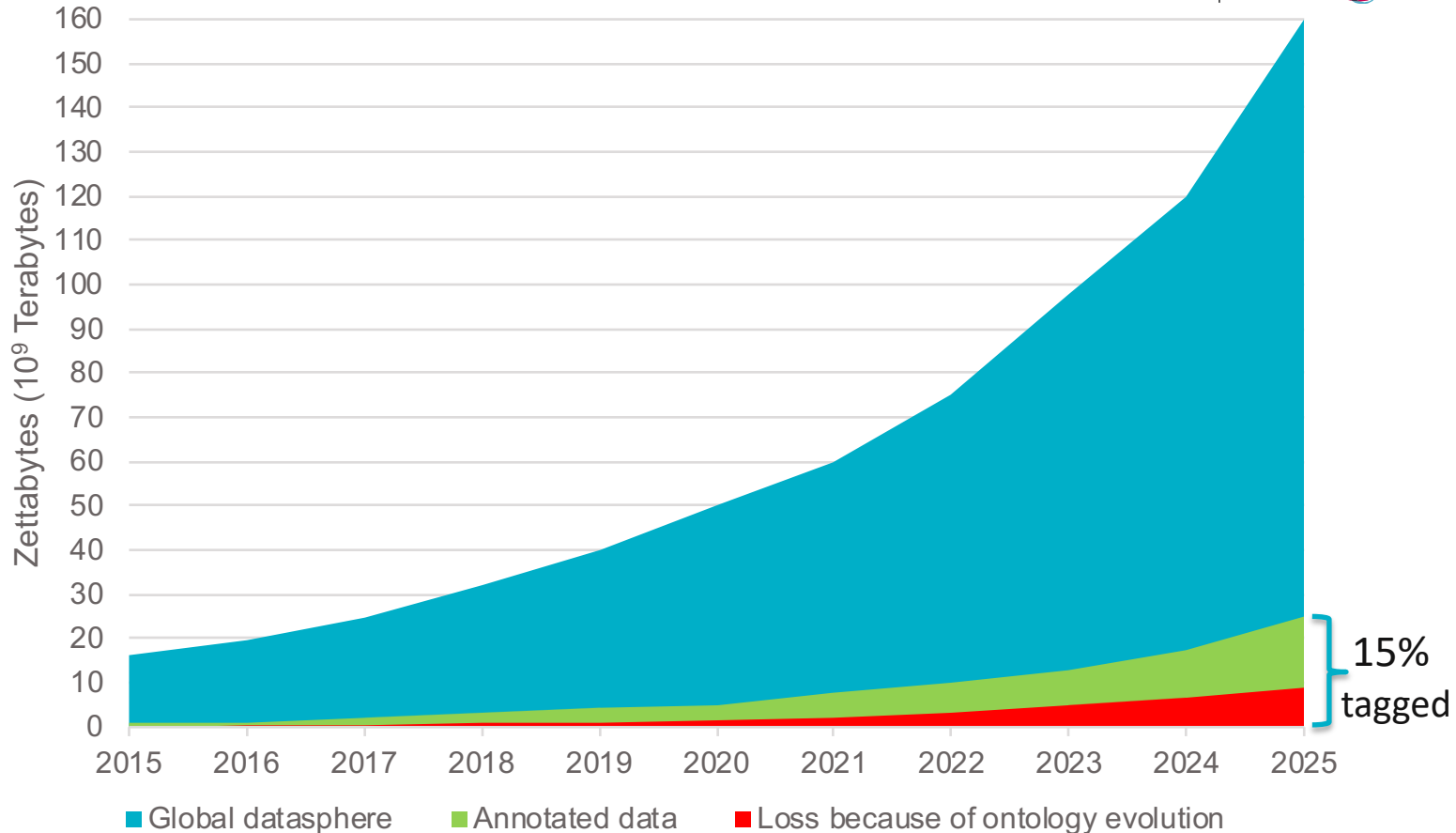
Management of the evolutions



- Knowledge evolves
- Models evolve
- Data evolve

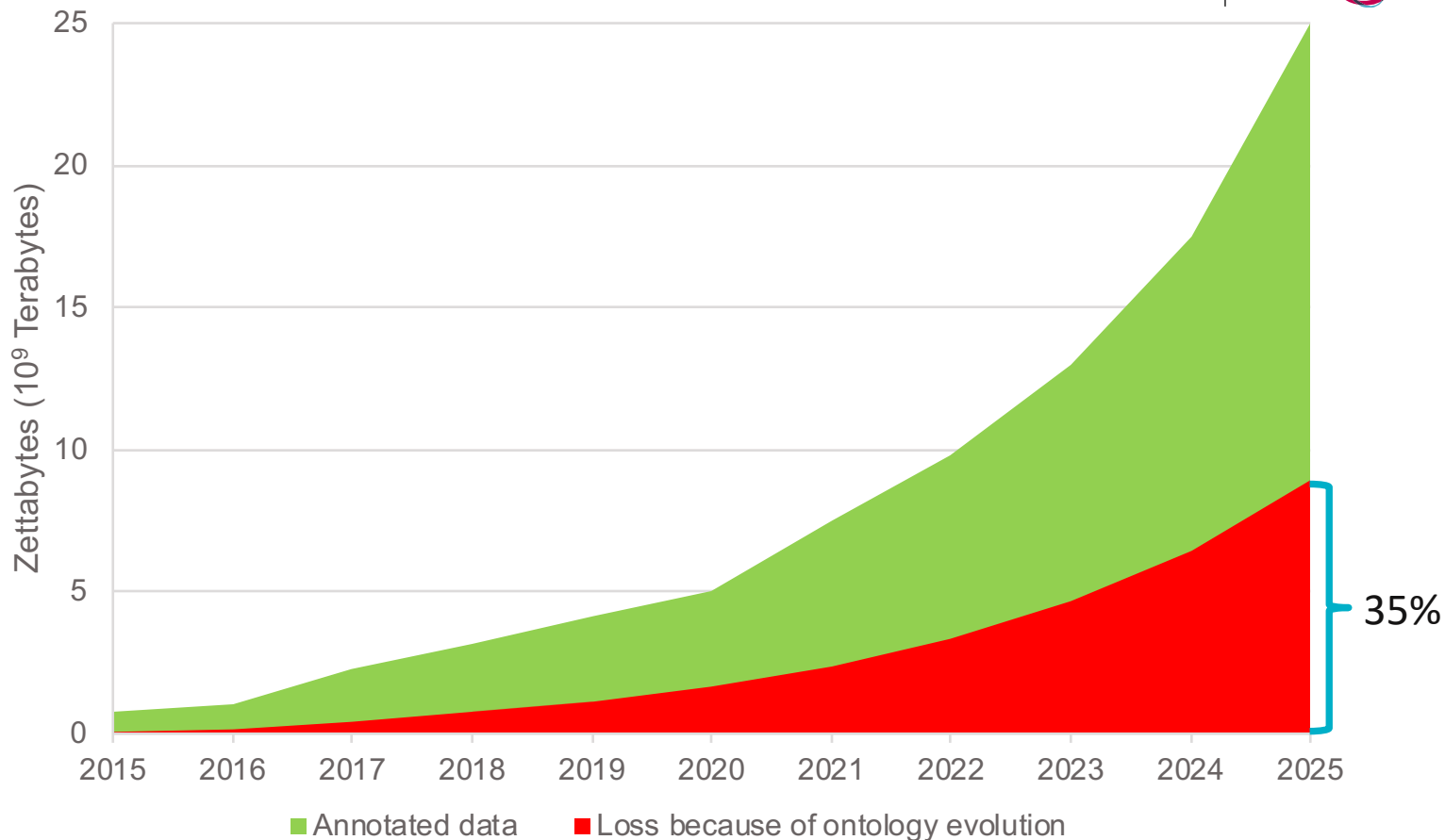
Users should profit of it

WHY MGT OF EVOLUTION IS A PROBLEM?



Source: IDC's Data Age 2025 study/Seagate

IF YOU DO NOTHING, YOU WILL LOOSE ALL



WHAT WE ARE DOING?

- Fixing mappings
- Fixing semantic annotations
- Retrieving past knowledge

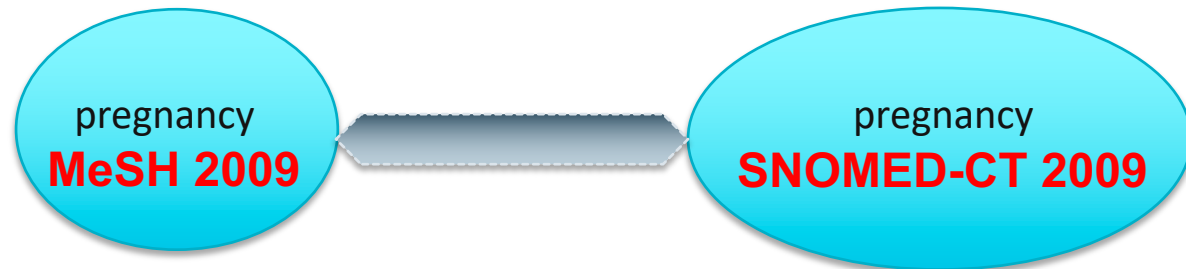


WHAT ARE ANNOTATIONS AND MAPPINGS?

Annotation

Diabetes mellitus and pregnancy-induced hypertension .

Mappings



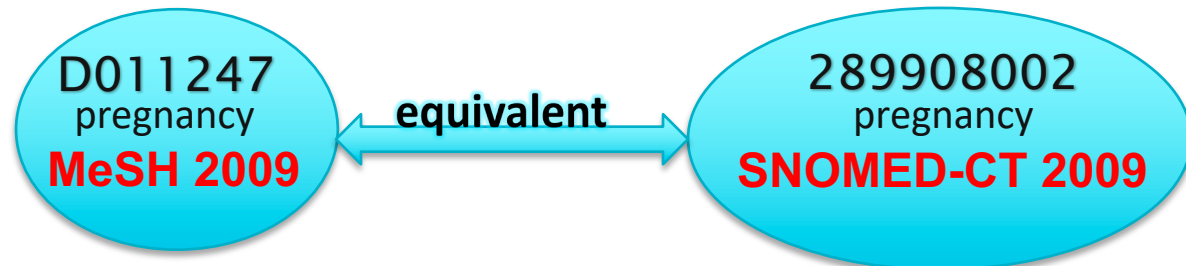
WHAT ARE ANNOTATIONS AND MAPPINGS?

Annotation

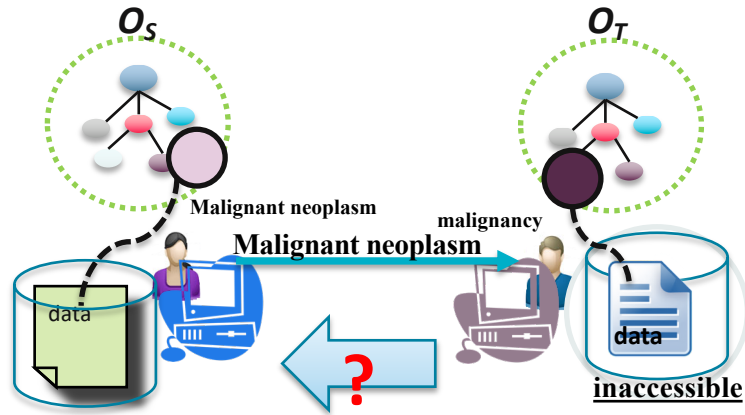


Diabetes mellitus and pregnancy-induced hypertension.

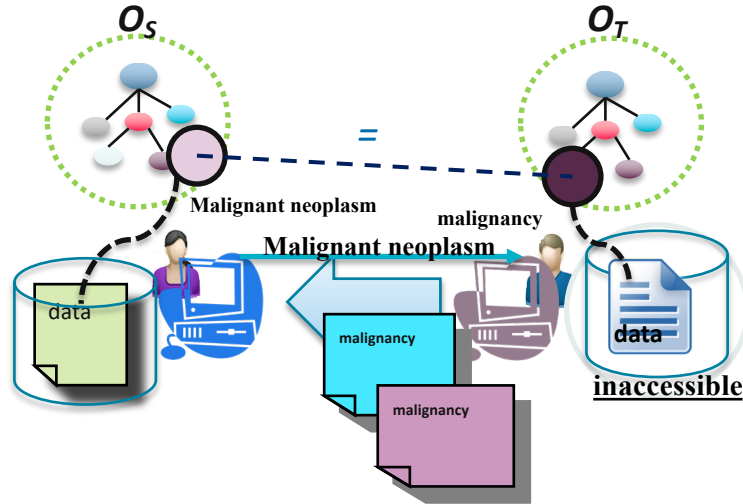
Mappings



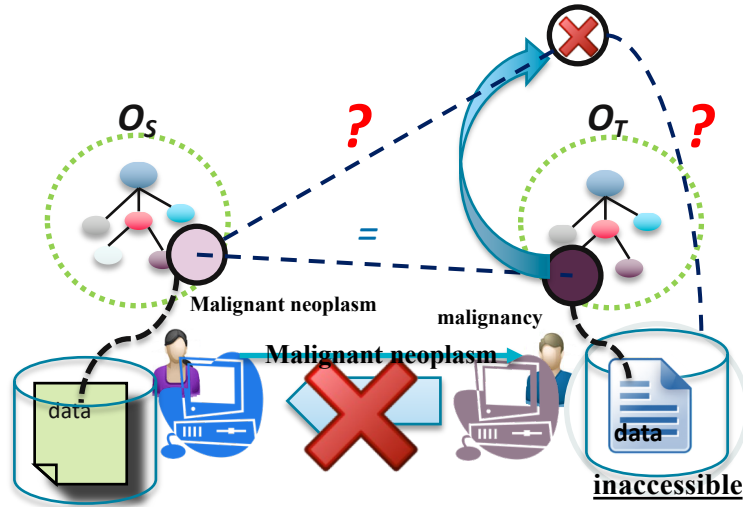
IMPACT OF KNOWLEDGE EVOLUTION IN INFORMATION SYSTEMS



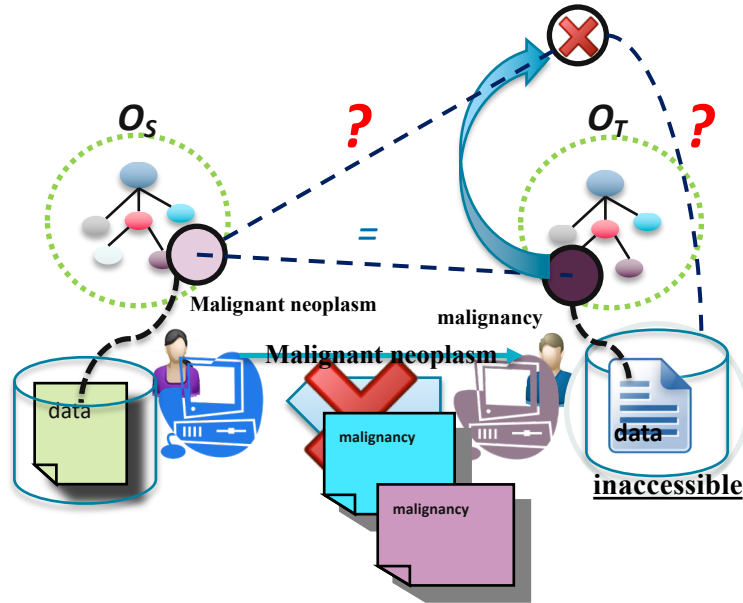
IMPACT OF KNOWLEDGE EVOLUTION IN INFORMATION SYSTEMS



IMPACT OF KNOWLEDGE EVOLUTION IN INFORMATION SYSTEMS



IMPACT OF KNOWLEDGE EVOLUTION IN INFORMATION SYSTEMS



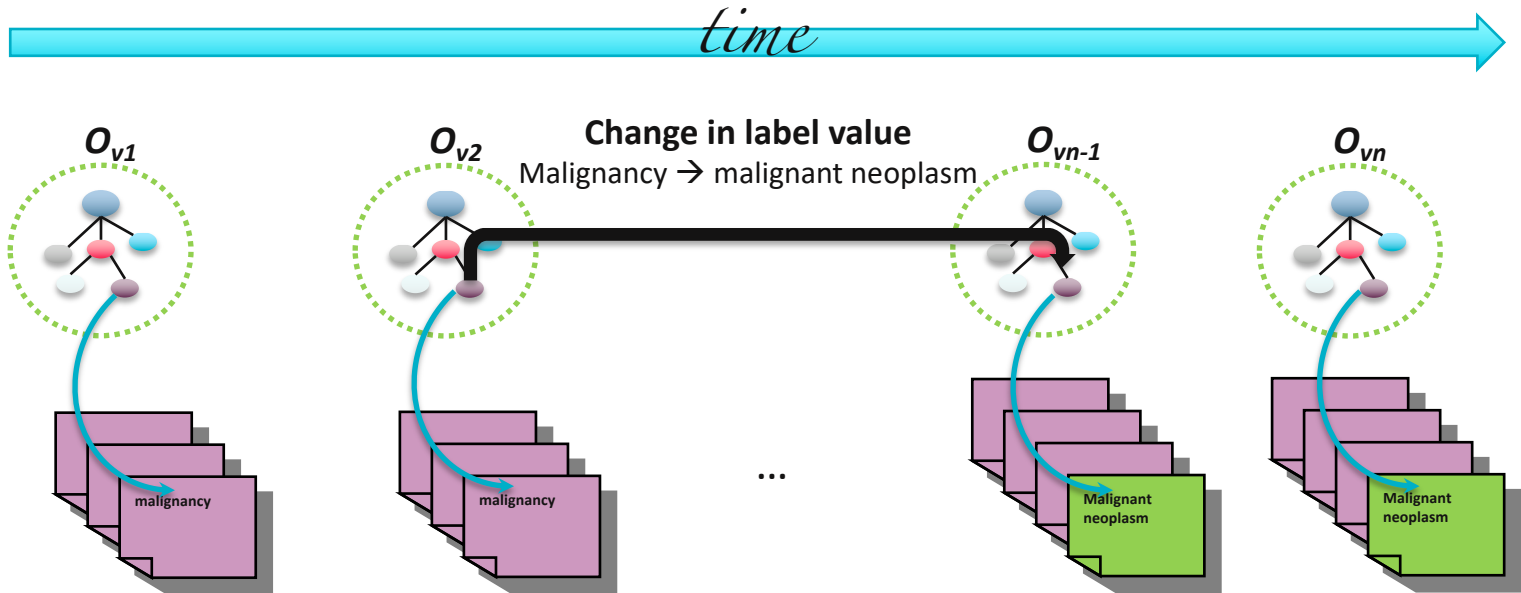
Options:

✓ 1-) If you are aware, you can update the mappings by hand

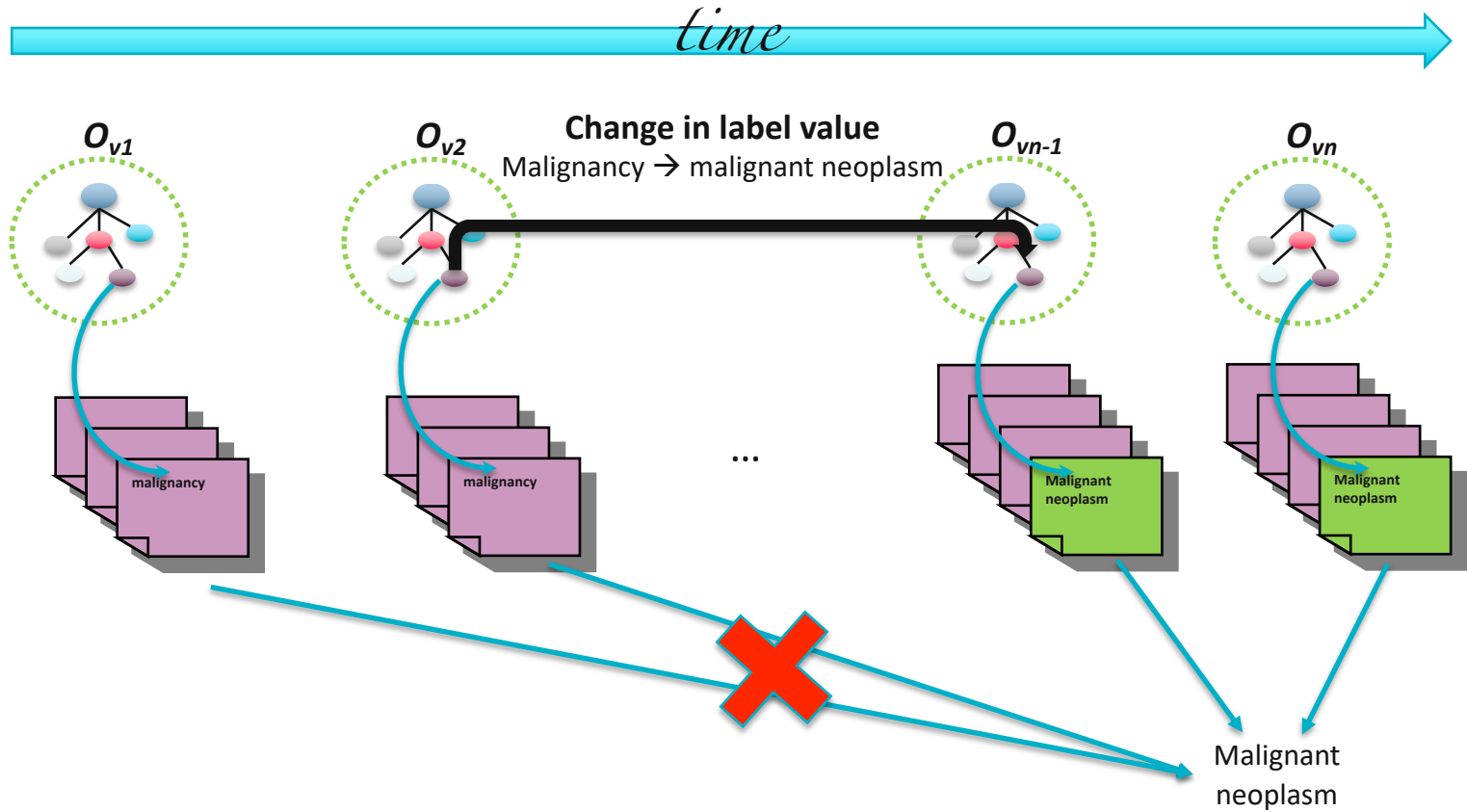
👍 2-) You can use Dynaccurate

⚠ 3-) You will lose information

IMPACT OF KNOWLEDGE EVOLUTION IN INFORMATION SYSTEMS

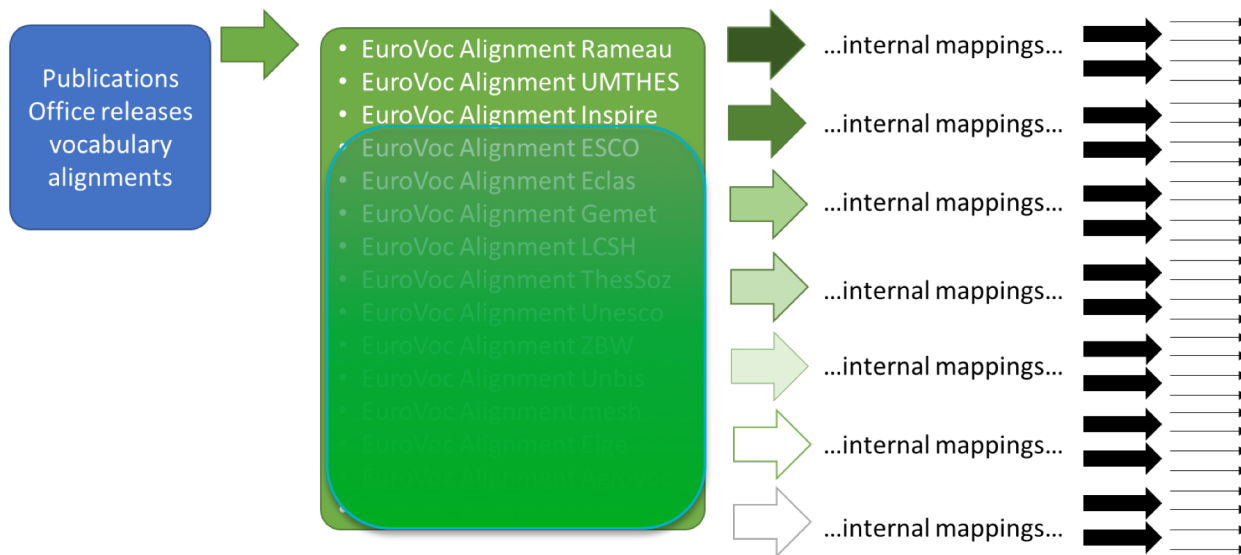


IMPACT OF KNOWLEDGE EVOLUTION IN INFORMATION SYSTEMS



EXAMPLE OF EUROVOC MAPPINGS

Public Authority | Releases to stakeholders | Downstream impacts



- Current: Resources must be invested **at each step** of this process
- This is common context where Dynaccurate can have added value
 - Eurovoc new release OR external terminologies new releases

WHEN USING DYNACCURATE

- Dynaccurate can identify what has changed in the ontologies
 - What concepts have changed between versions
- Dynaccurate can identify the type of change
 - Basic changes: addition, removal, change attribute values ...
 - Complex changes: split, merge, move
- This can provide the information for the release notes
- Dynaccurate will also facilitate impact assessments for any remaining manual changes
- Dynaccurate's output is also readable at machine level, allowing for further automation opportunities in the future

Let's talk about technical stuffs



WHAT WE ARE DOING?

Fixing mappings

Fixing semantic annotations

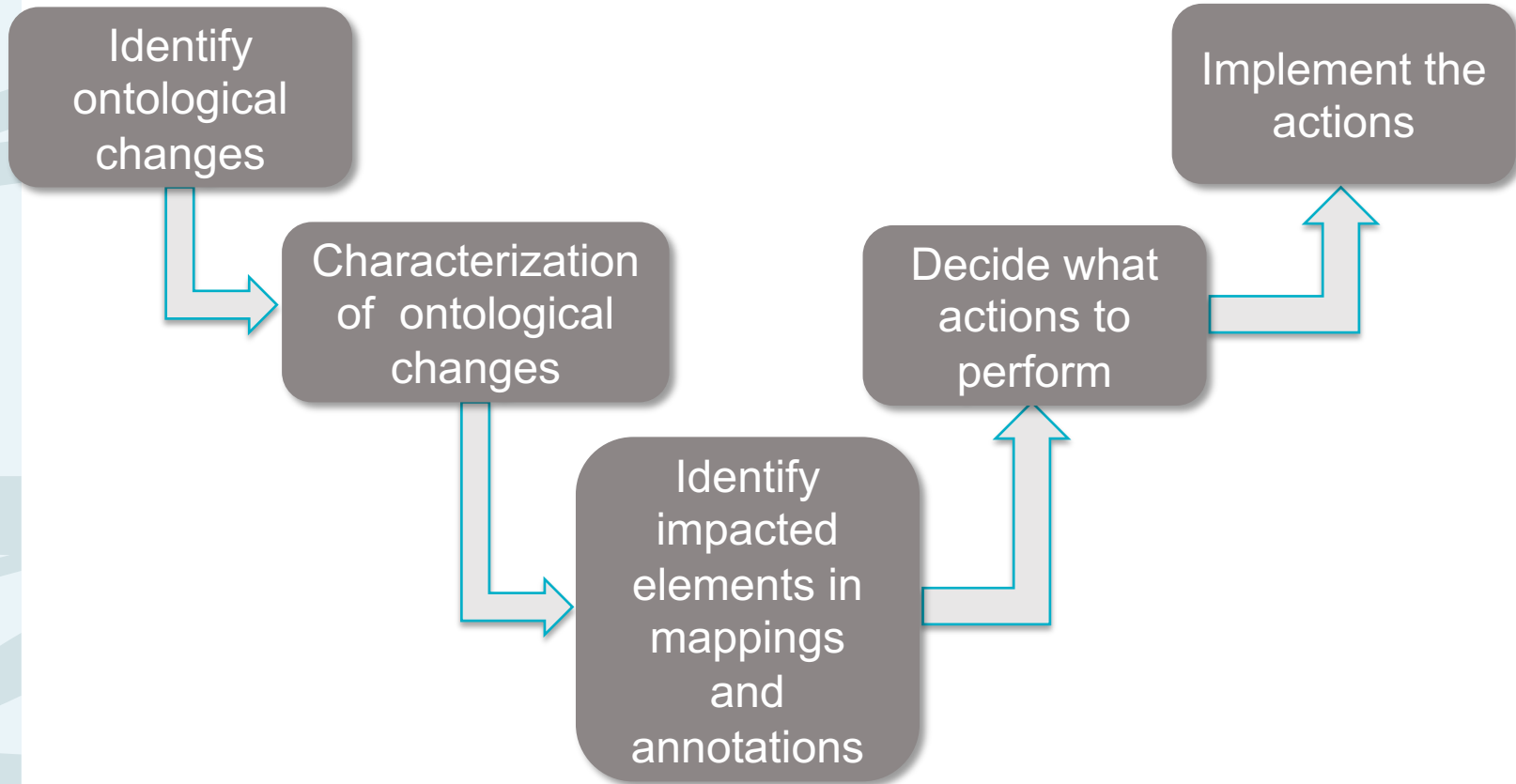
Retrieving past knowledge



FIXING MAPPINGS



DYNACCURATE - GENERAL APPROACH



1-) Change identification

- Use of **COnto-Diff** to identify changes in concept involved in mappings
 - Basic changes (addition, removal, change attribute values ...)
 - Complex changes (split, merge, move)

2-) Change characterization

- Definition of **change patterns** to identify specific changes:
 - Lexical change patterns
 - Total copy, partial copy, total transfer, partial transfer
 - Semantic change patterns
 - Equivalent, more specific, more general, partial match

3-) Identify the impact

- Use the outcomes of previous phase to identify the **set of mappings** that requires updating

4-) Select and implement the maintenance actions

- Evaluate the context and decide what action is necessary to implement
 - Mappings: add, remove, move, derive, change relation

The **scope is reduced** to the set to impacted mappings only

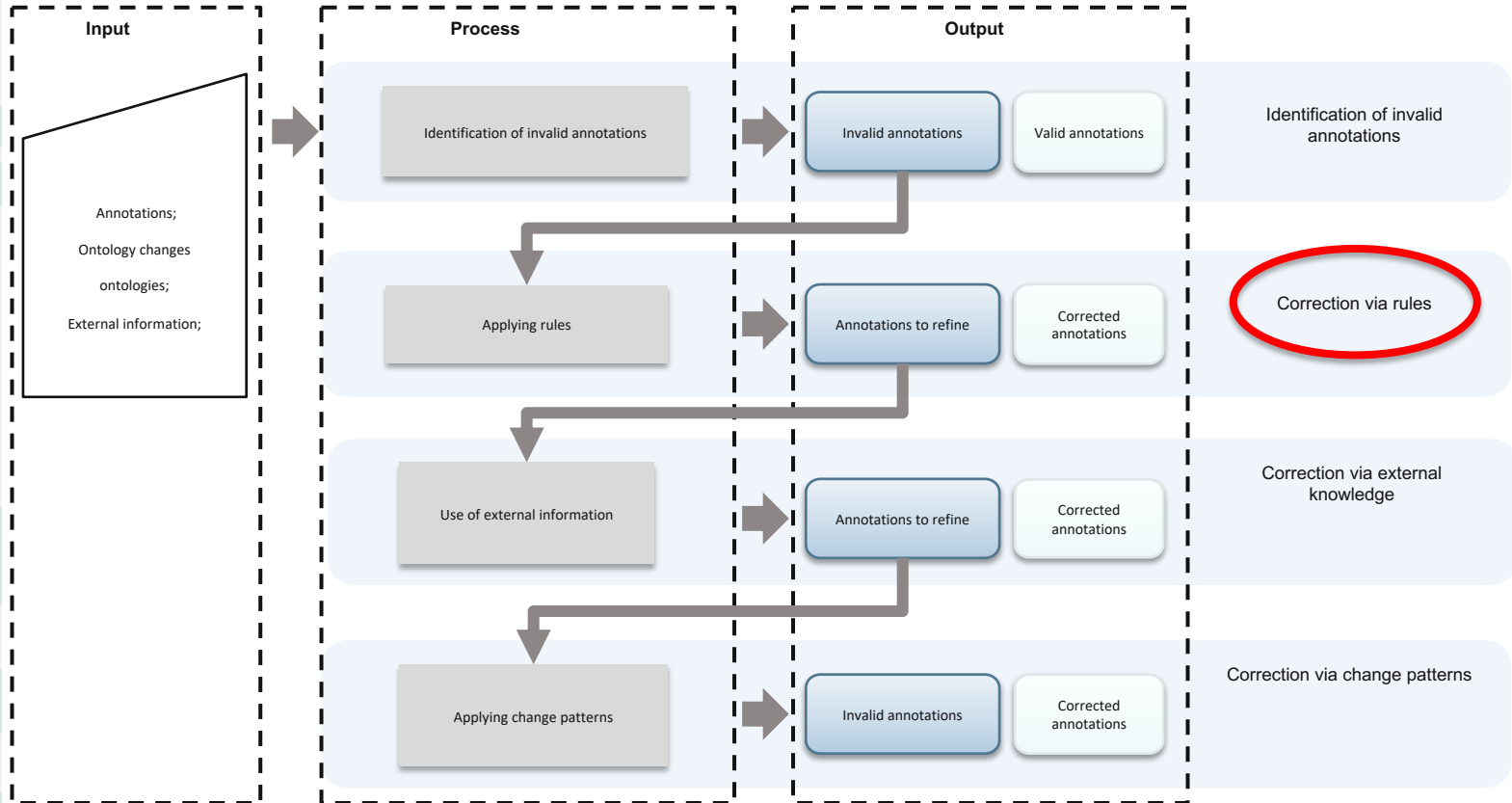
- Reduce the maintenance execution **time**
- Reduce the **work** of experts for validation tasks

FIXING SEMANTIC ANNOTATIONS



DECIDE WHAT ACTIONS TO PERFORM

Rule-based approach for semantic annotations adaptation



DECIDE WHAT ACTIONS TO PERFORM

Rule-based approach for semantic annotations adaptation

- *MergeAnnot*: Merge of two annotations
- *IncreaseAnnot*: Increase the amount of information to annotate
- *ResurrectAnnot*: Reconsideration of deleted annotations
- *PluralAnnot*: Consideration of plurals
- *ChangeConceptAnnot*: Change of concept ID in the annotations
- *SplitAnnot*: Split of annotation
- *SuperClassAnnot*: Use of super concept to annotate

The application of these rules is governed by a well-accepted guideline¹

1. Dogan et. al. "NCBI disease corpus: a resource for disease name recognition and concept normalization."

EXAMPLE OF RULE

MergeAnnot

MeSH 2009

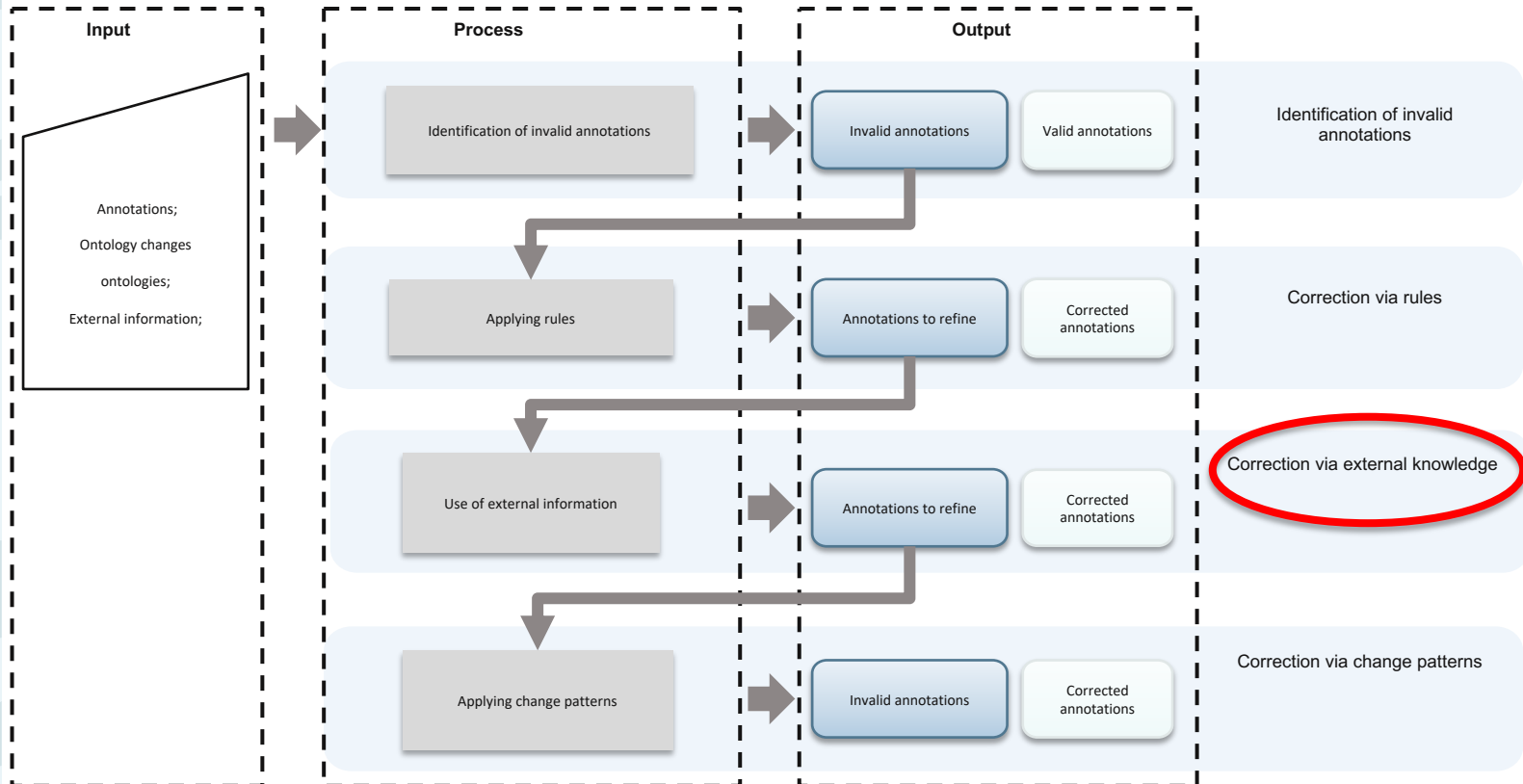
diabetes mellitus and ^{D011247} pregnancy-induced ^{D006973} hypertension .

diabetes mellitus and pregnancy-induced hypertension .

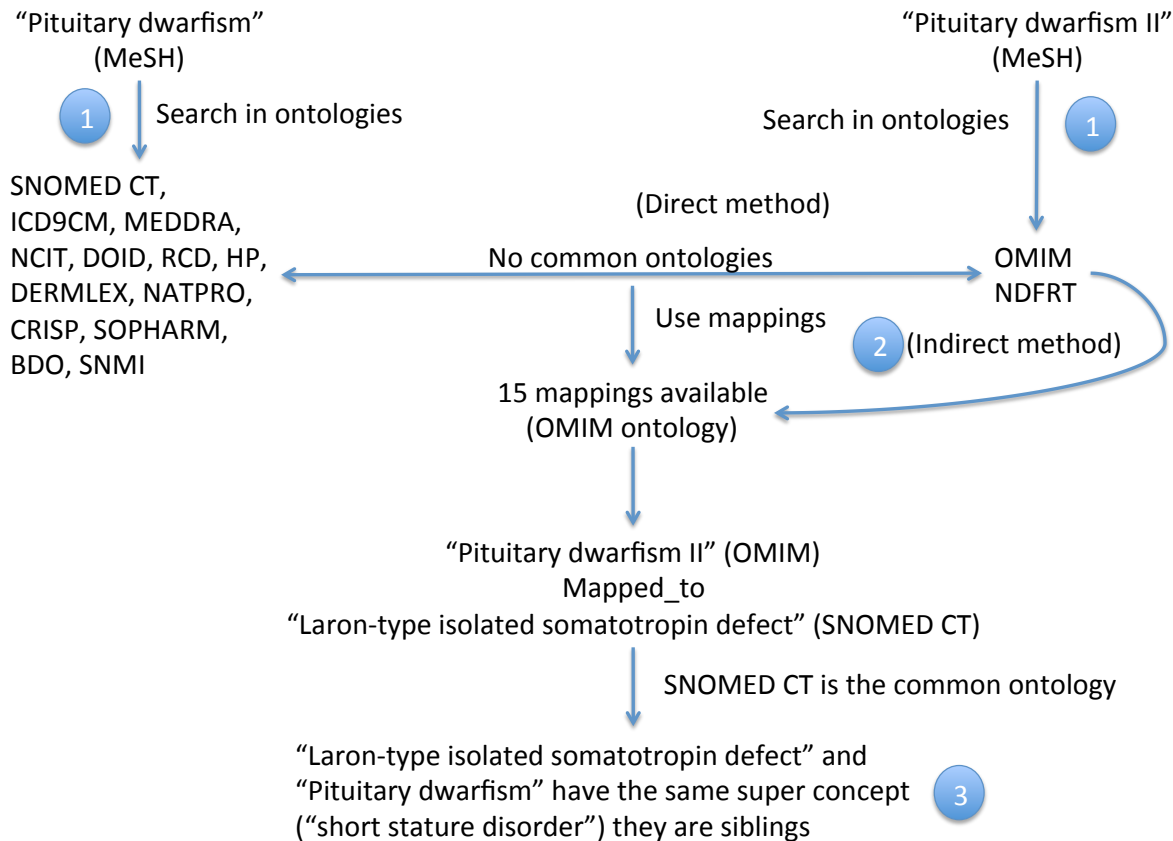
D046110

MeSH 2010

DIRECT MAINTENANCE



BACKGROUND KNOWLEDGE



WHAT WE ARE DOING?

Fixing mappings

Fixing semantic annotations

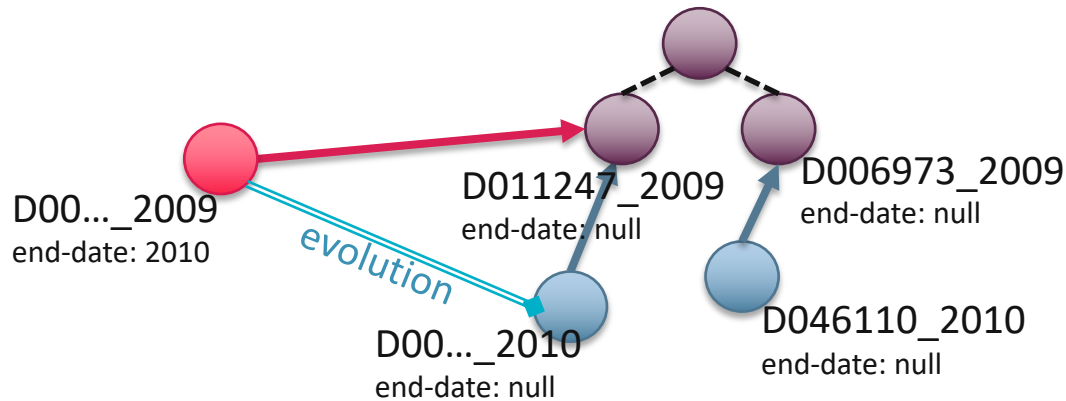
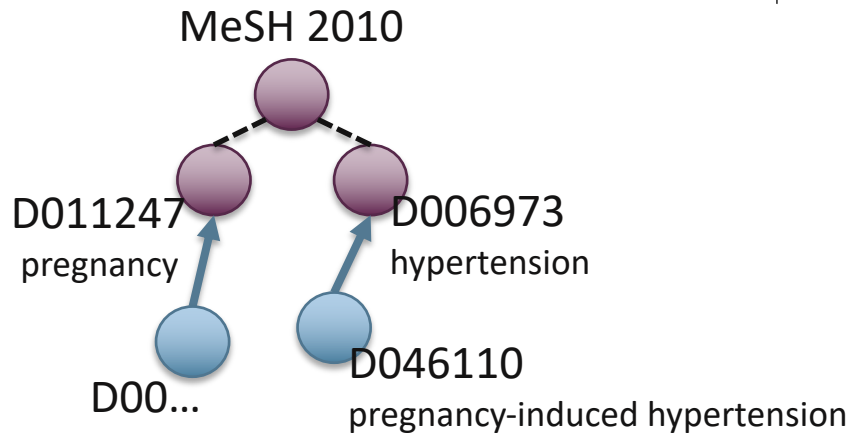
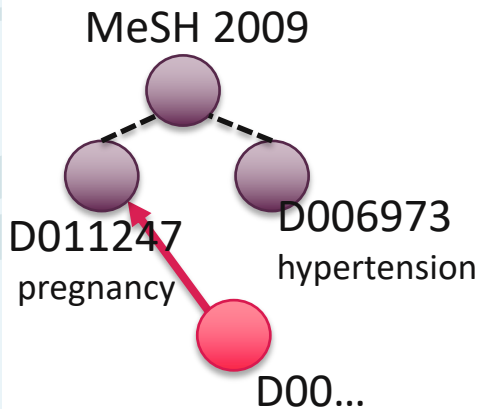
Retrieving past knowledge



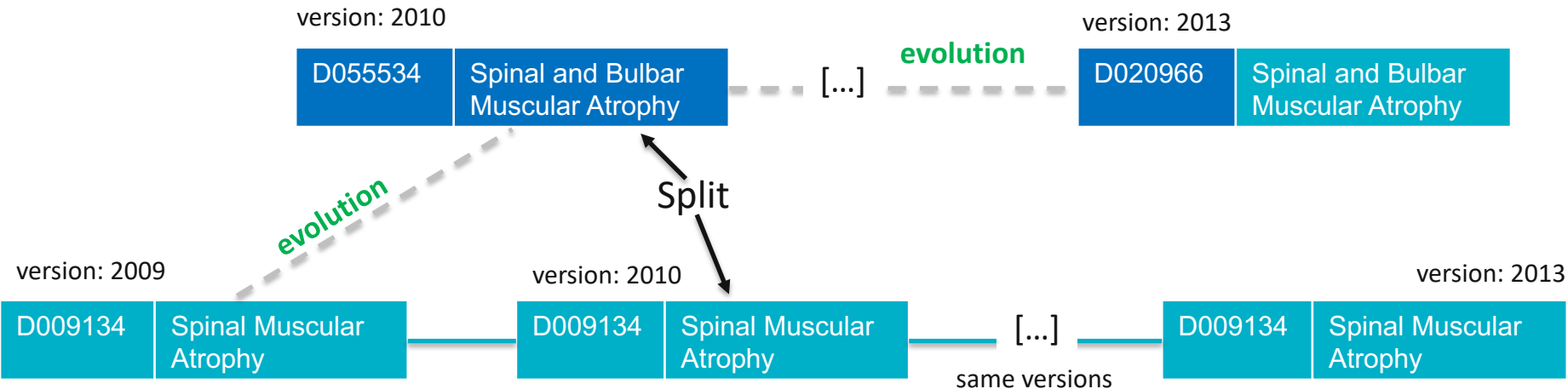
KNOWLEDGE GRAPH RETRIEVING PAST INFORMATION

- Problem:
 - In many cases metadata cannot be modified (data privacy)
 - Ex: Patient's data
- Objectives:
 - Keep annotated data searchable even if the annotation cannot be modified according to KOS changes
 - Avoid mismatch between KOS version used to annotate data and the one used to query it
- Build a historical knowledge graph containing the history of ontological changes in a compact way that can be queried to gain information about concept evolution

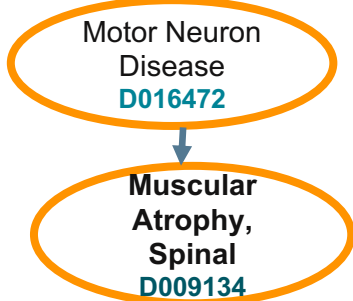
HISTORICAL KNOWLEDGE GRAPH CREATION



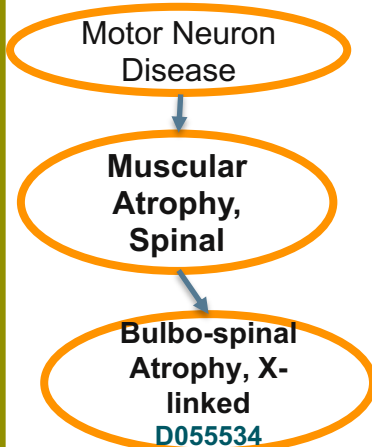
HISTORICAL KNOWLEDGE GRAPH, EXAMPLE



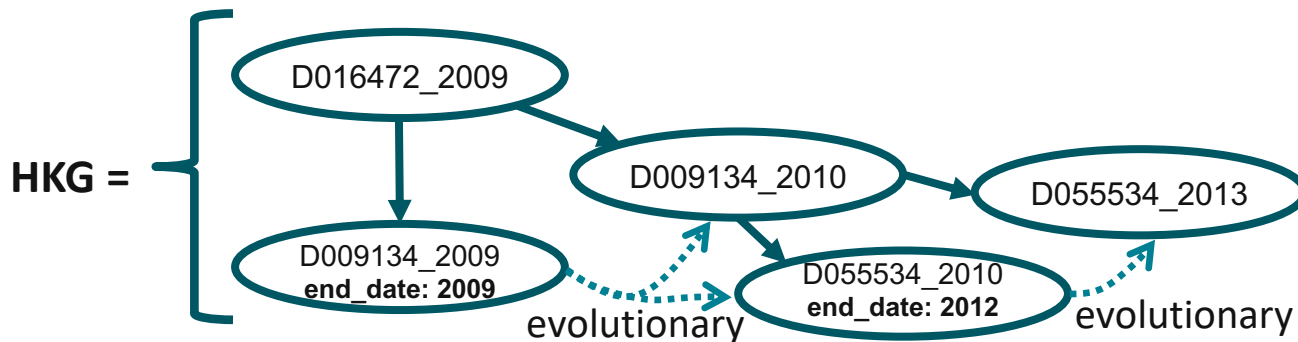
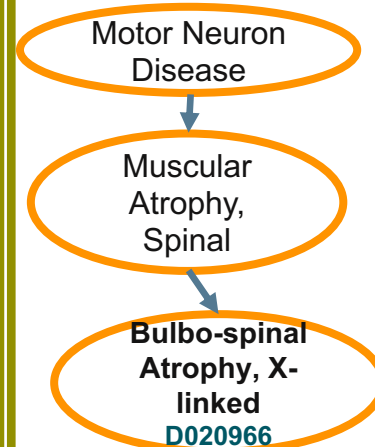
Version: 2009



Version: 2010



Version: 2013



- Adding explanation to recommendations and to evaluate trends and impacts
- Adding evaluation of quality and trends
- Improving visualization to facilitate analysis and validation tasks

VISIT US AT

<https://www.elisa-project.lu/>



Cedric.Pruski@list.lu

Marcos.dasilveira@list.lu

<https://marcao02.github.io/>