Modularity Meets Forgetting a case study with the SNOMED CT

Jieying Chen, Ghadah Alghamdi, Renate Schmidt,
Dirk Walther, Yongsheng Gao

University of Manchester, UK[1]
IHTSDO
DNV GL, Norway

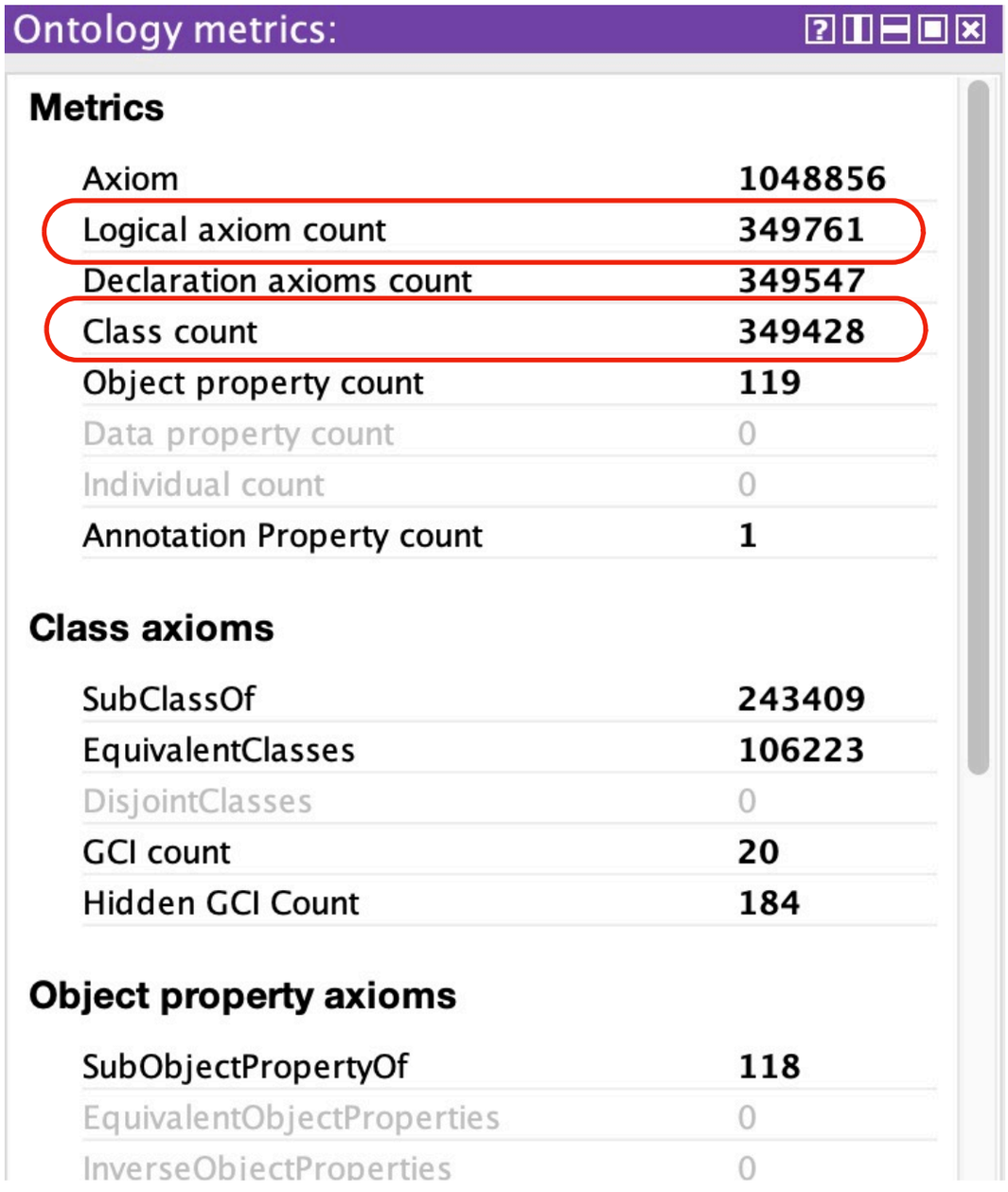
SNOMED CT Expo 2019, Kuala Lumpur, 31 Oct. 2019

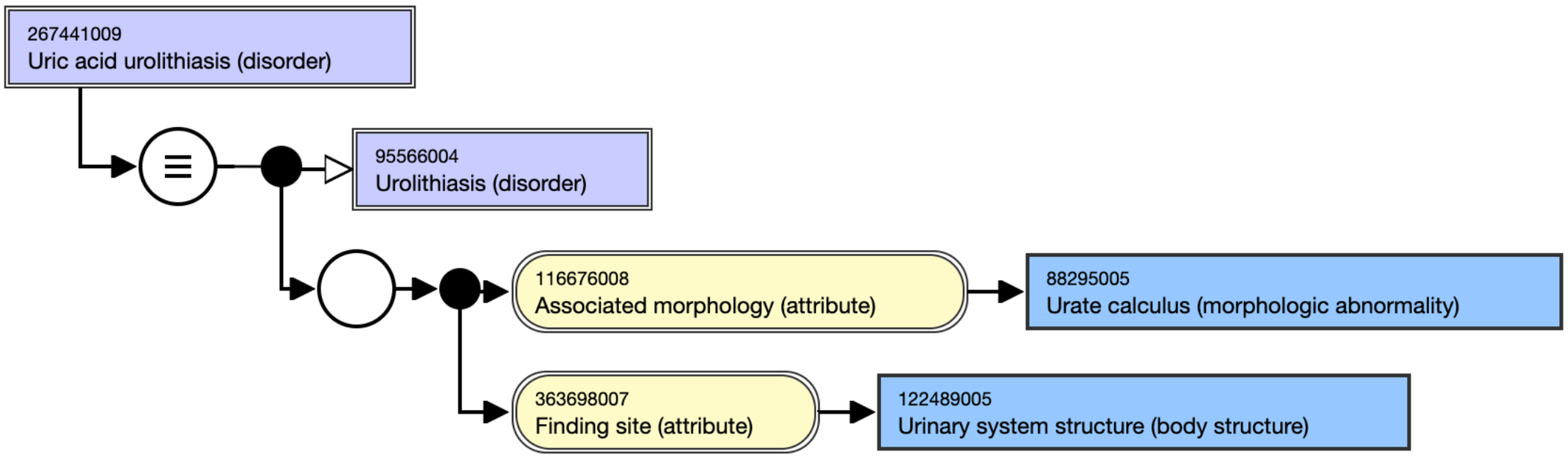
---

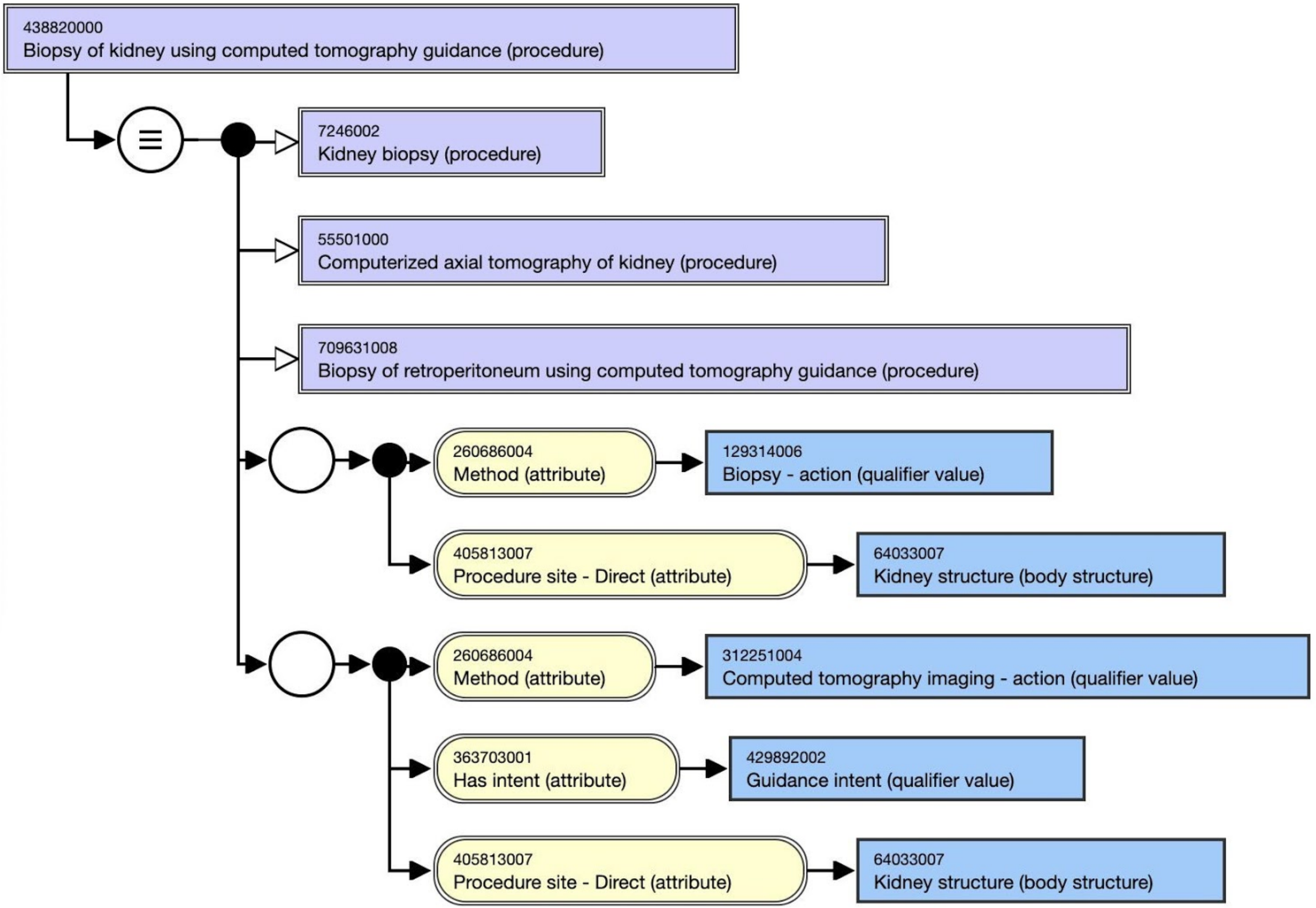
1. This work is funded by EPSRC IAA and IHTSDO, Snomed International.
2. Thanks Yizheng Zhao, Patrick Koopmann, Boris Konev.

Motivation

- Snomed CT
  - > 340 000 classes
  - > 340 000 logical axioms




267441009
Uric acid urolithiasis (disorder)
95566004
Urolithiasis (disorder)
116676008
Associated morphology (attribute)
88295005
Urate calculus (morphologic abnormality)
363698007
Finding site (attribute)
122489005
Urinary system structure (body structure)

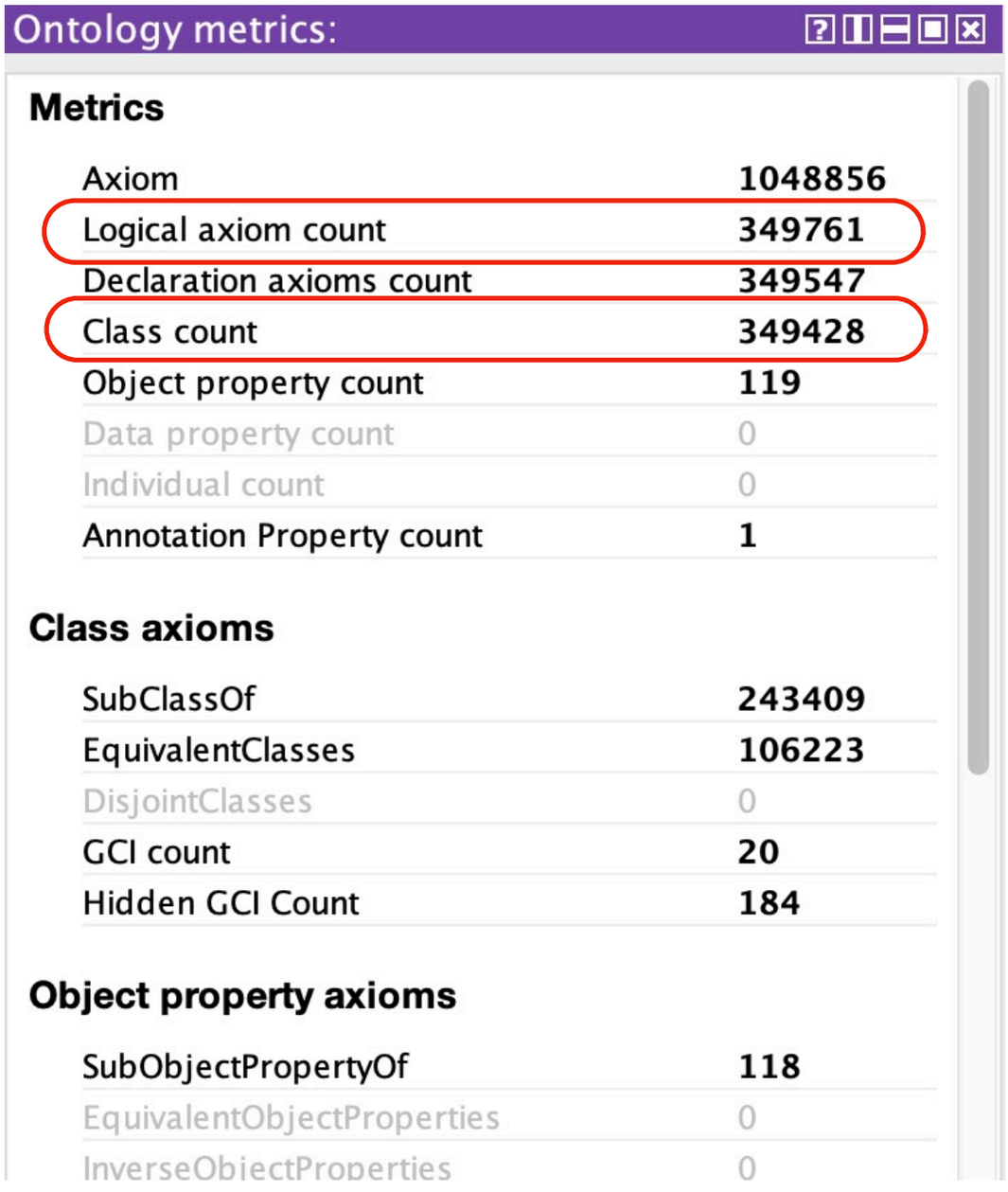

438820000
Biopsy of kidney using computed tomography guidance (procedure)
7246002
Kidney biopsy (procedure)
55501000
Computerized axial tomography of kidney (procedure)
709631008
Biopsy of retroperitoneum using computed tomography guidance (procedure)
260686004
Method (attribute)
129314006
Biopsy - action (qualifier value)
405813007
Procedure site - Direct (attribute)
64033007
Kidney structure (body structure)
260686004
Method (attribute)
312251004
Computed tomography imaging - action (qualifier value)
363703001
Has intent (attribute)
429892002
Guidance intent (qualifier value)
405813007
Procedure site - Direct (attribute)
64033007
Kidney structure (body structure)

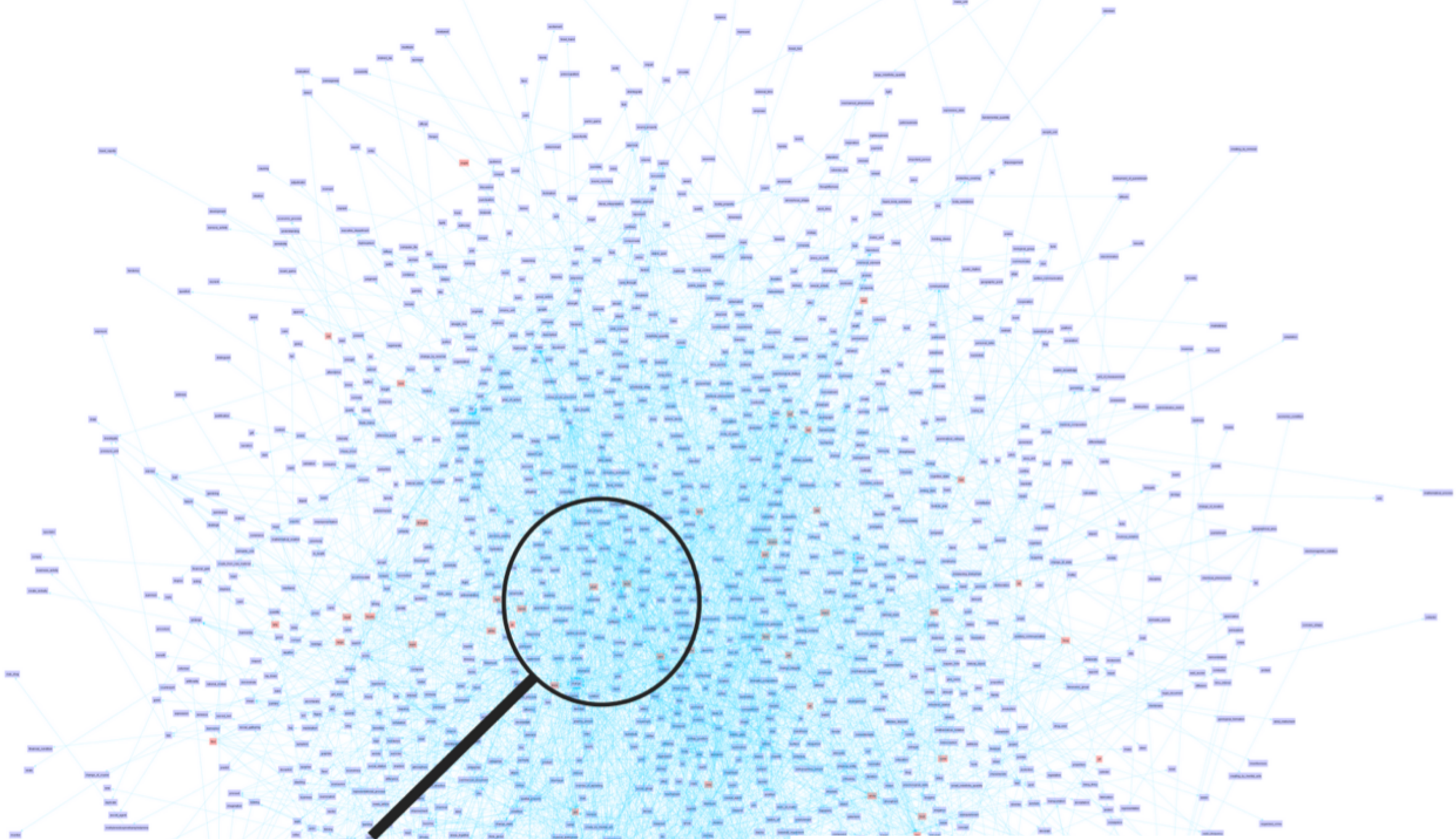
Motivation

- Snomed CT
  - > 340 000 classes
  - > 340 000 logical axioms

- Goal

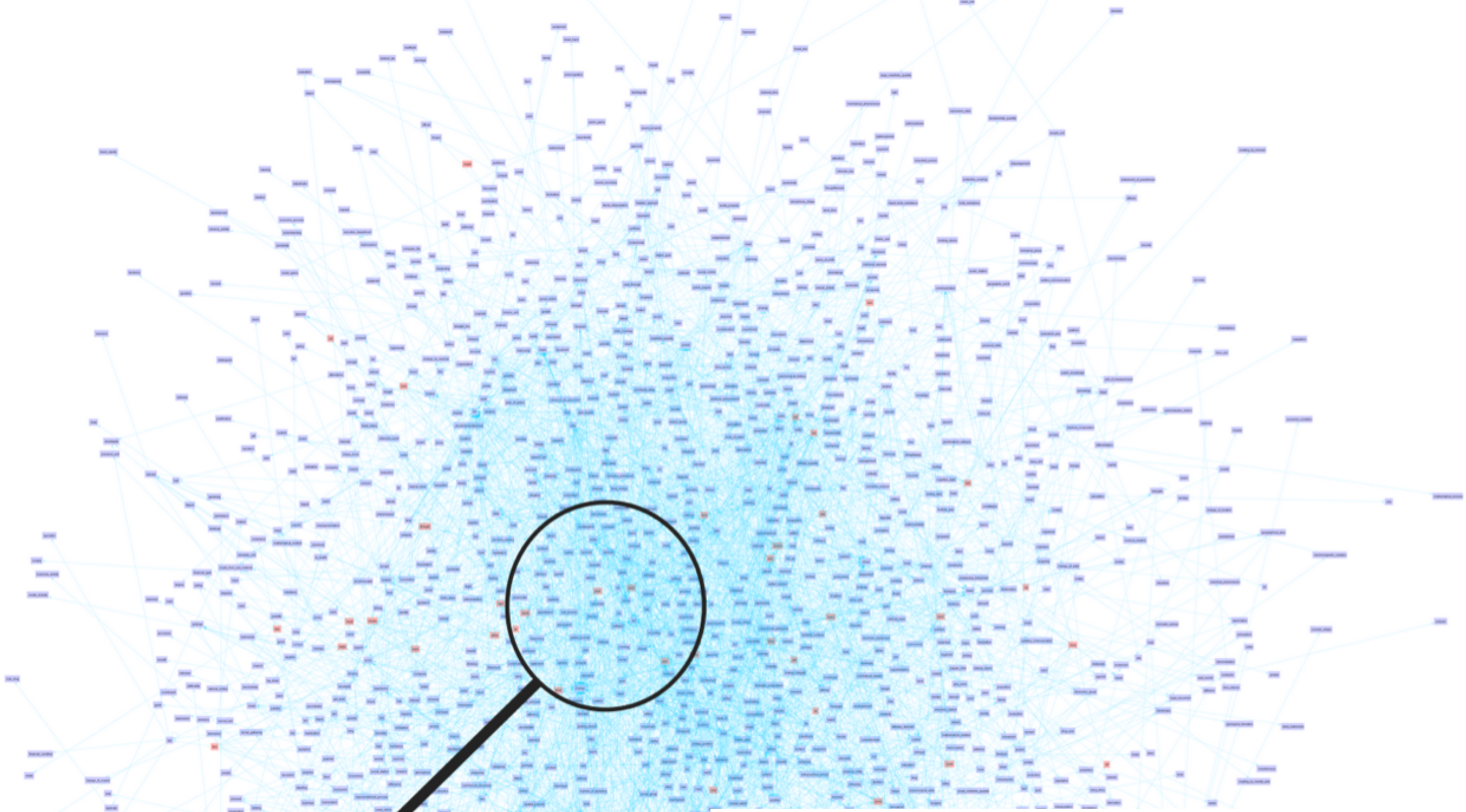
  Compute a concise extract





Use cases

- Ontology abstraction / summarisation
- Ontology reuse
- Ontology versioning
- Information hiding



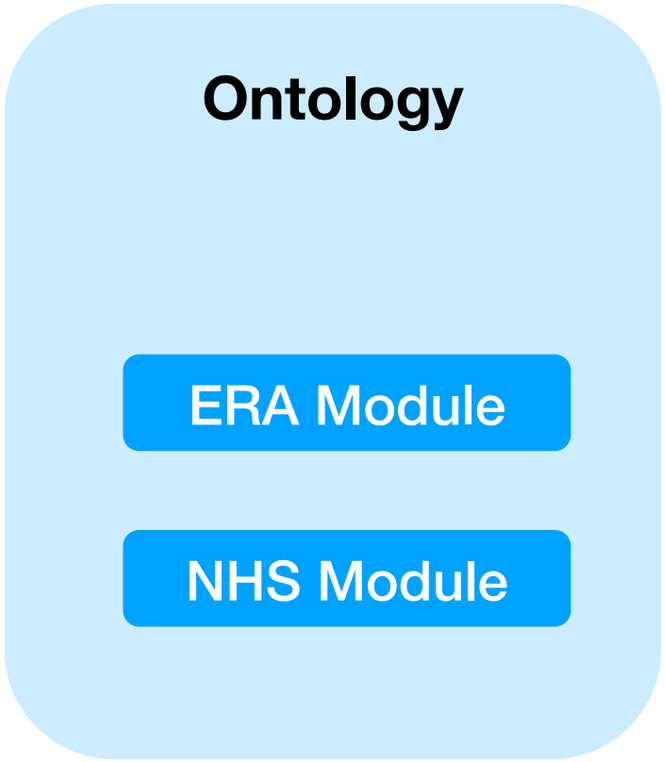
Approaches:

Computing extracts of Snomed CT

- ontology modularity
- forgetting / uniform interpolation

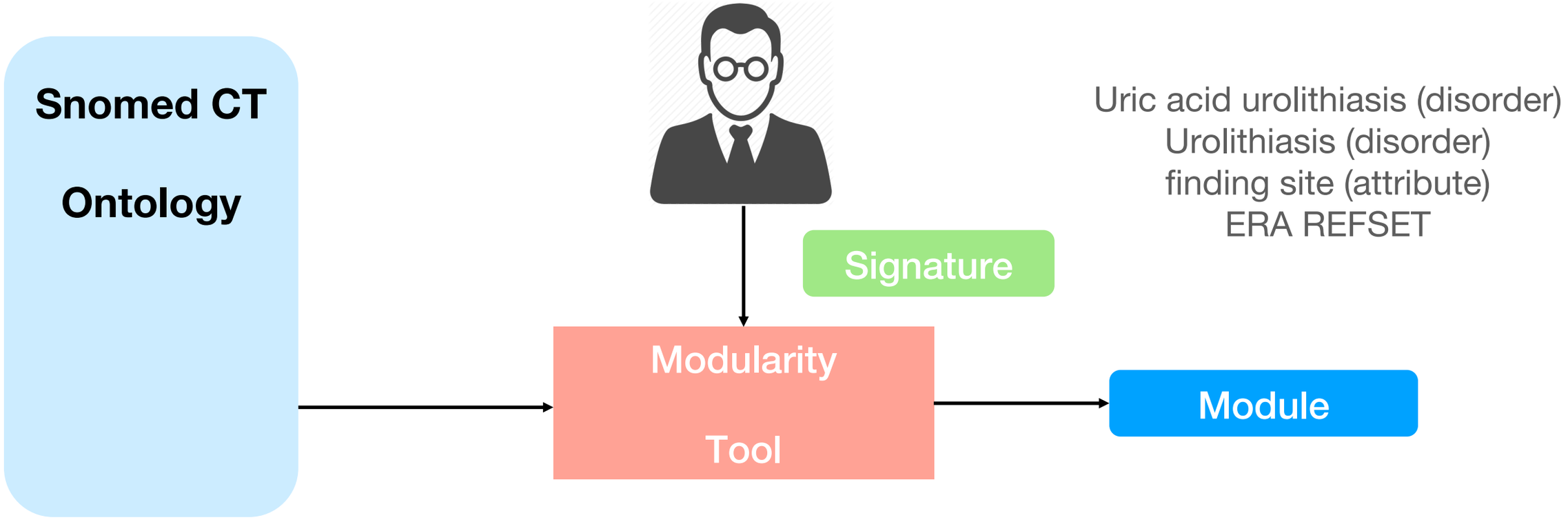
Ontology Modularity

- Ontology is a set of axioms.
- A Module is a subset of an ontology.
- We are interested in modules that preserve answers to queries about specified classes and properties, e.g., ERA REFSET, NHS REFSETs, ...



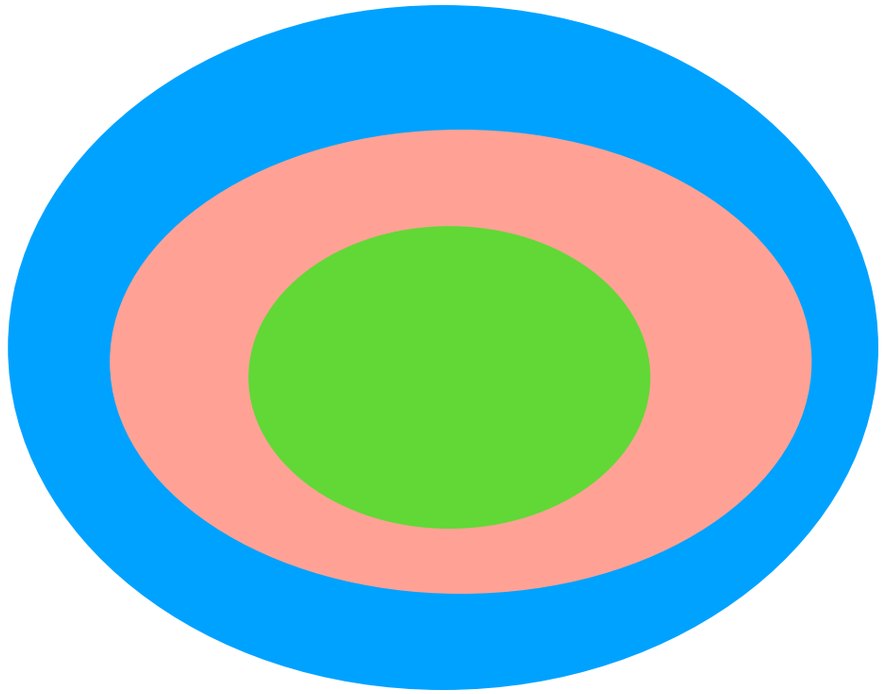
Ontology Modularity

- Input
  - Ontology
  - Signature: set of specified classes and properties



Tools for Ontology Modularity

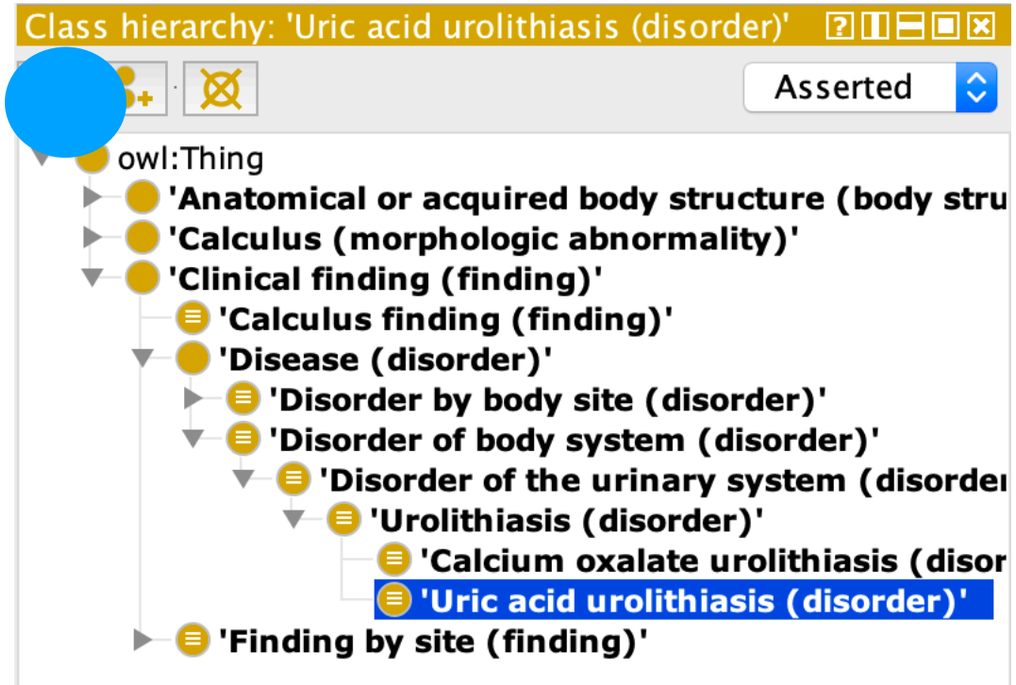
- OWLAPI: Locality-based module[1]
- MEX: Minimal Semantic module [2]
- Minimal Subsumption Module [3]



1. U. Sattler, T. Schneider, M. Zakharyaschev: Which Kind of Module Should I Extract? In Proceedings of DL'09.
2. B.Konev, C.Lutz, D.Walther, F.Wolter: Model-theoretic inseparability and modularity of description logic ontologies. JAI'13.
3. J.Chen, M. Ludwig, Y. Ma, D. Walther: Zooming in on Ontologies: Minimal Modules and Best Excerpts. In Proceedings of ISWC'17.

Example: Ontology Modularity

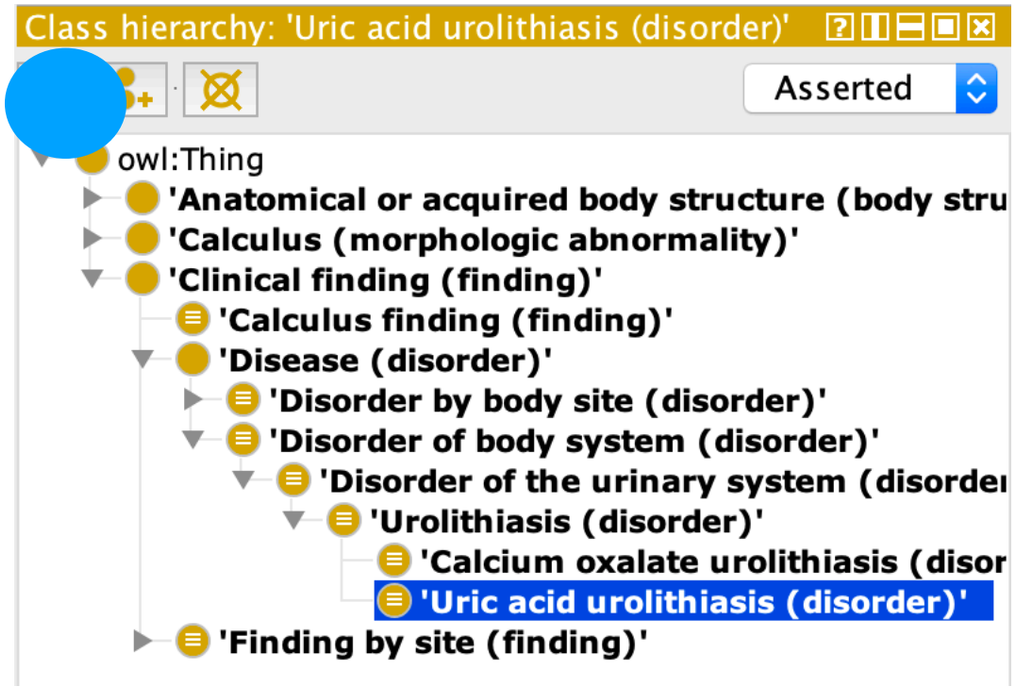
Σ:={ Urolithiasis (disorder), Calcium oxalate urolithiasis (disorder)
Uric acid urolithiasis (disorder) }

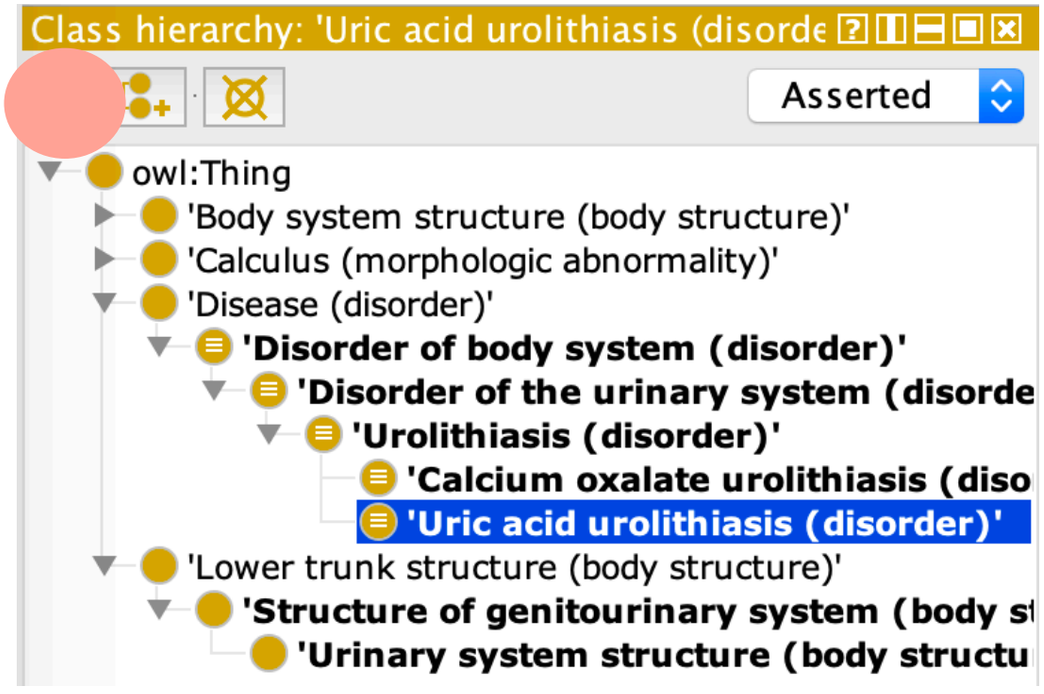


...

Example: Ontology Modularity

Σ:={ Urolithiasis (disorder), Calcium oxalate urolithiasis (disorder)
Uric acid urolithiasis (disorder)}

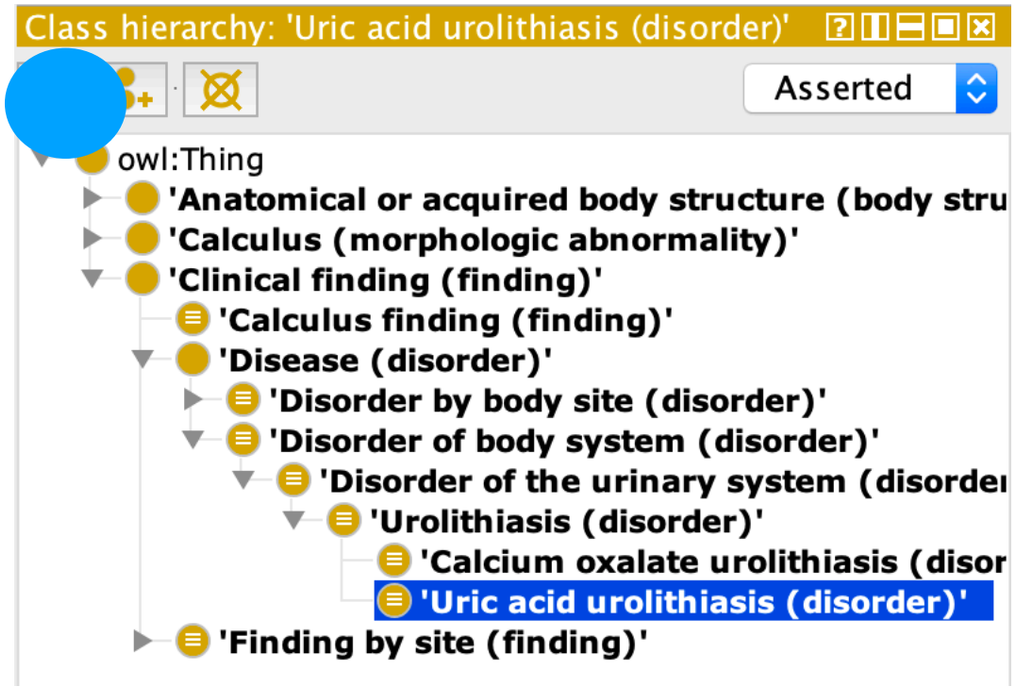


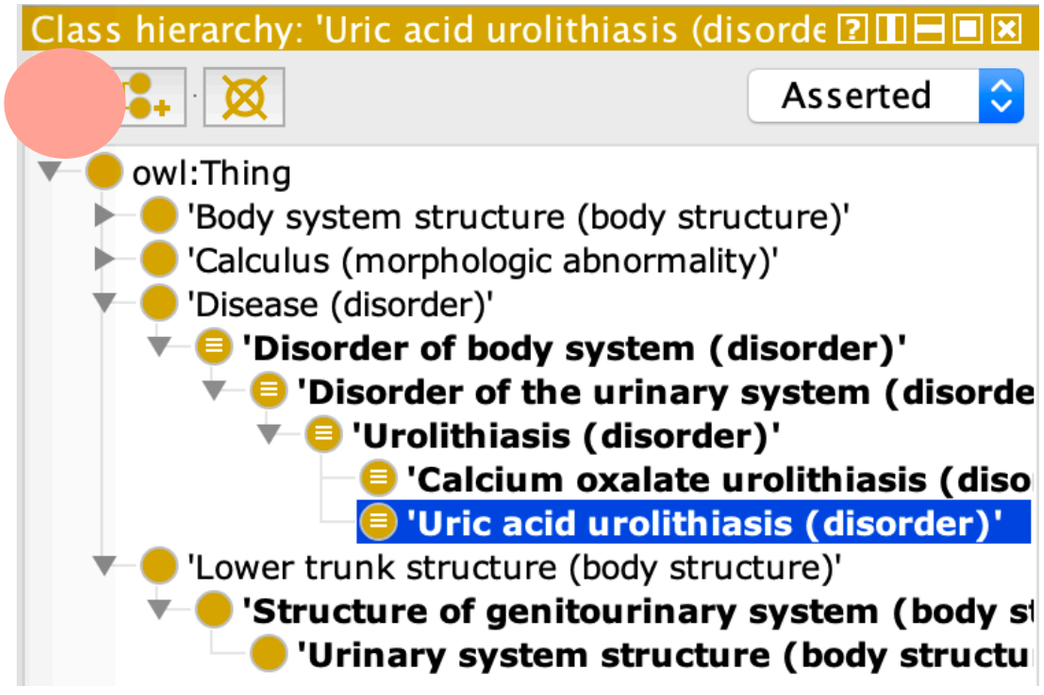


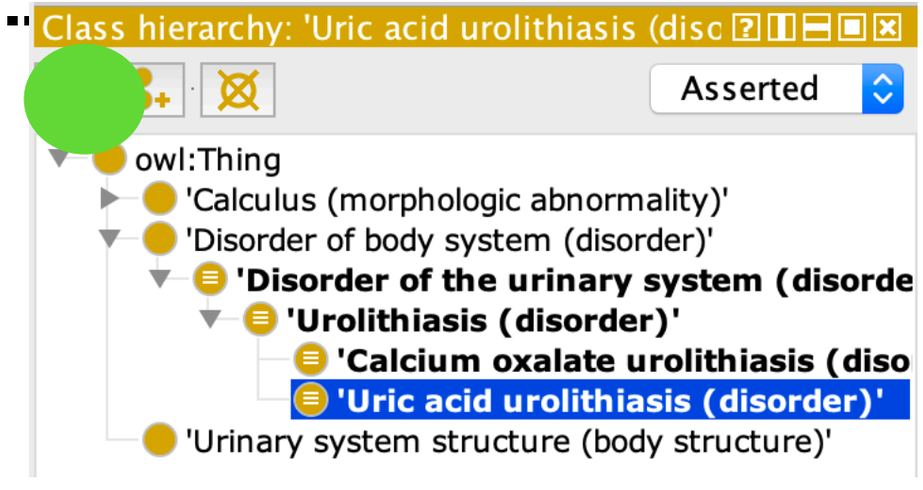
...

Example: Ontology Modularity

Σ:={ Urolithiasis (disorder), Calcium oxalate urolithiasis (disorder)
Uric acid urolithiasis (disorder) }

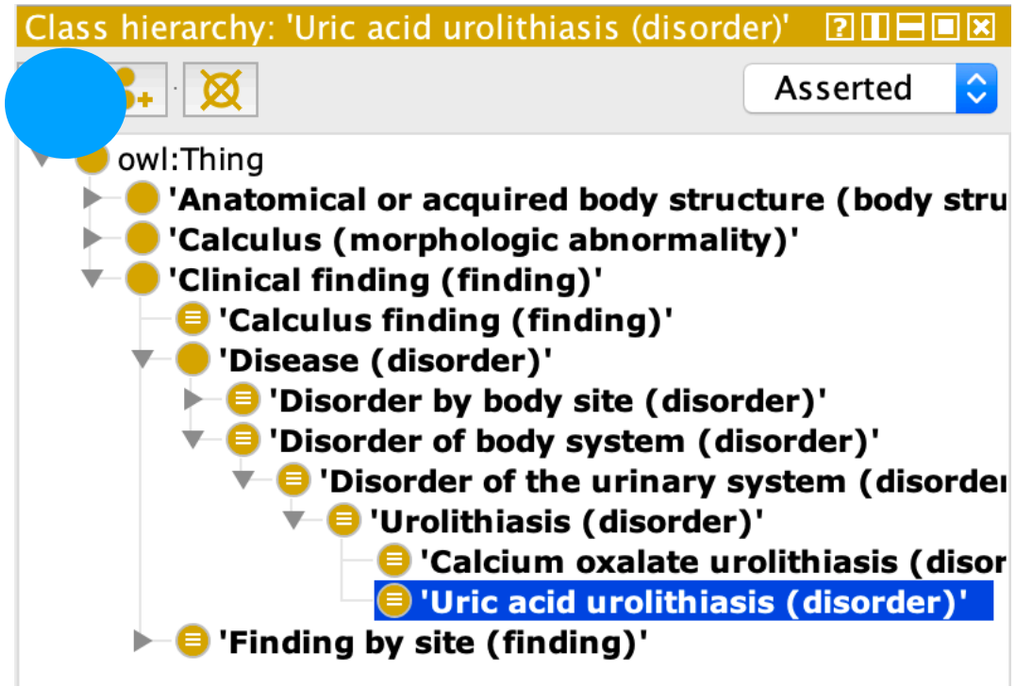


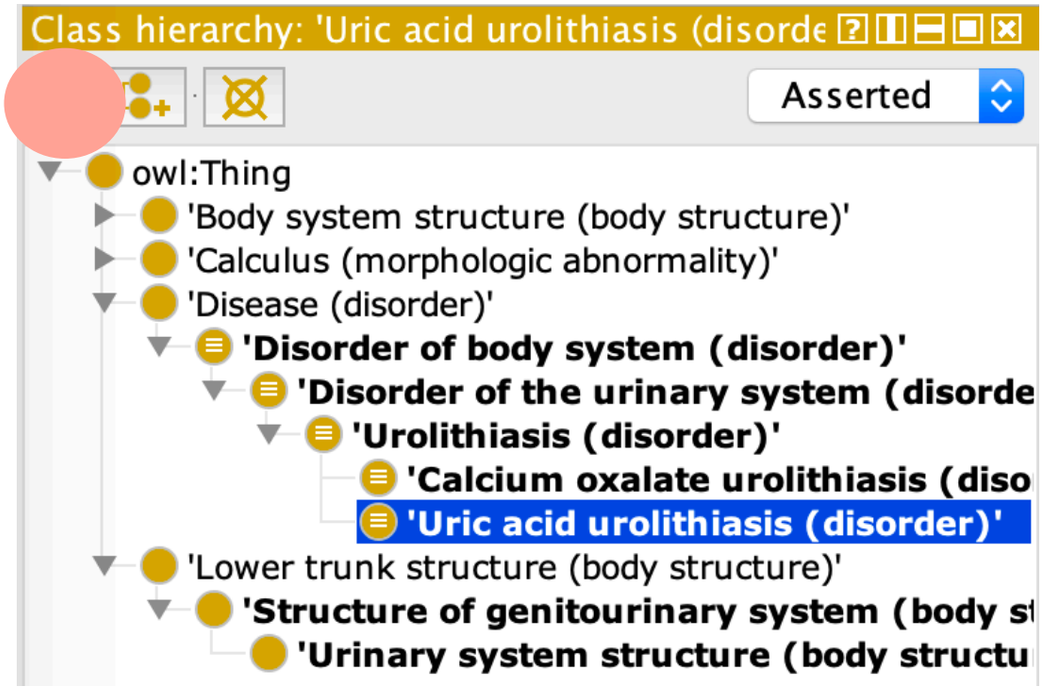


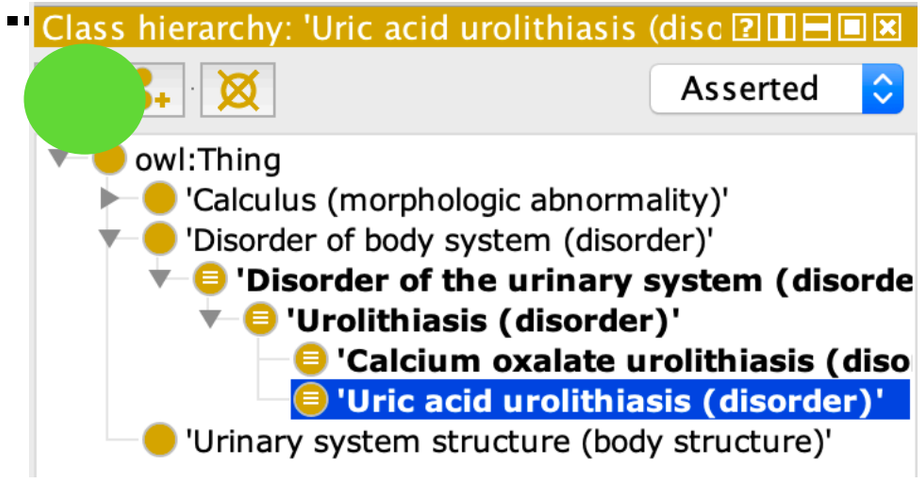


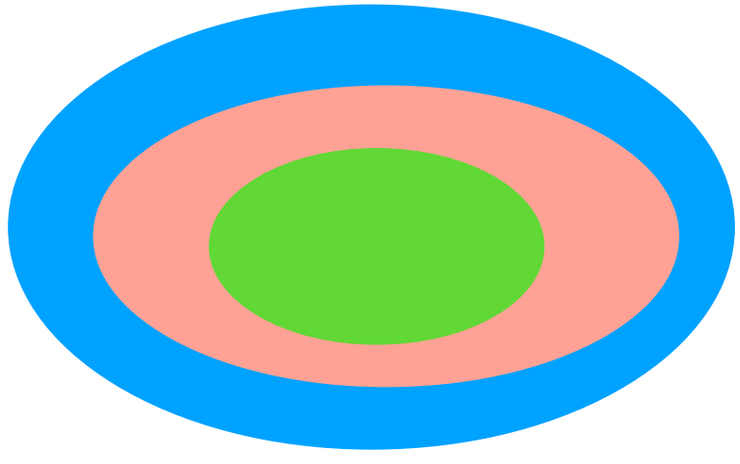
Example: Ontology Modularity

Σ:={ Urolithiasis (disorder), Calcium oxalate urolithiasis (disorder)
Uric acid urolithiasis (disorder),... }



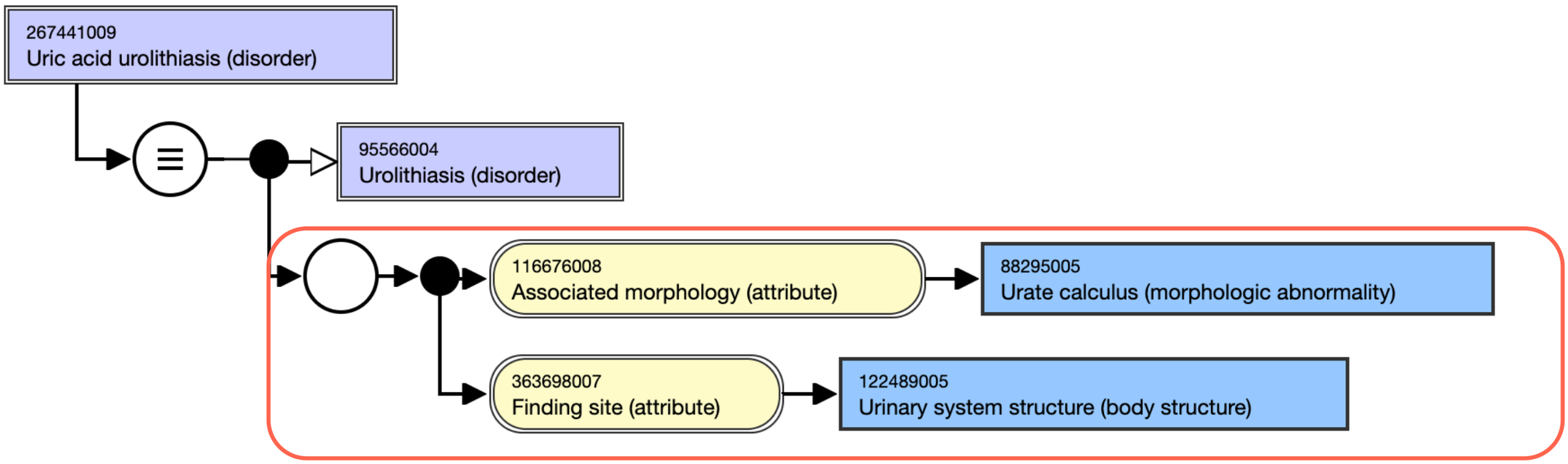




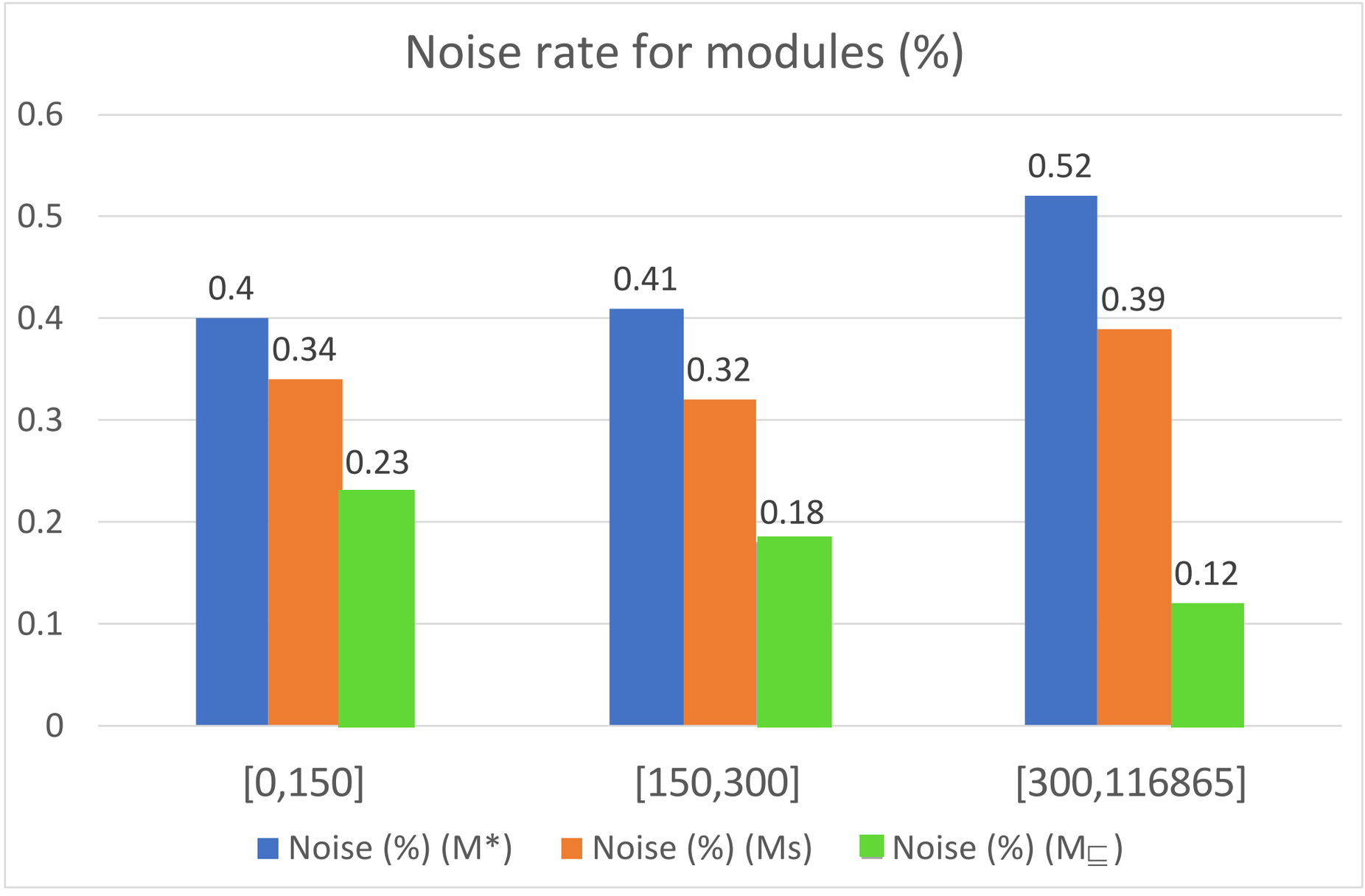


Example: Ontology Modularity

Σ:={ Urolithiasis (disorder), Calcium oxalate urolithiasis (disorder)
Uric acid urolithiasis (disorder) }



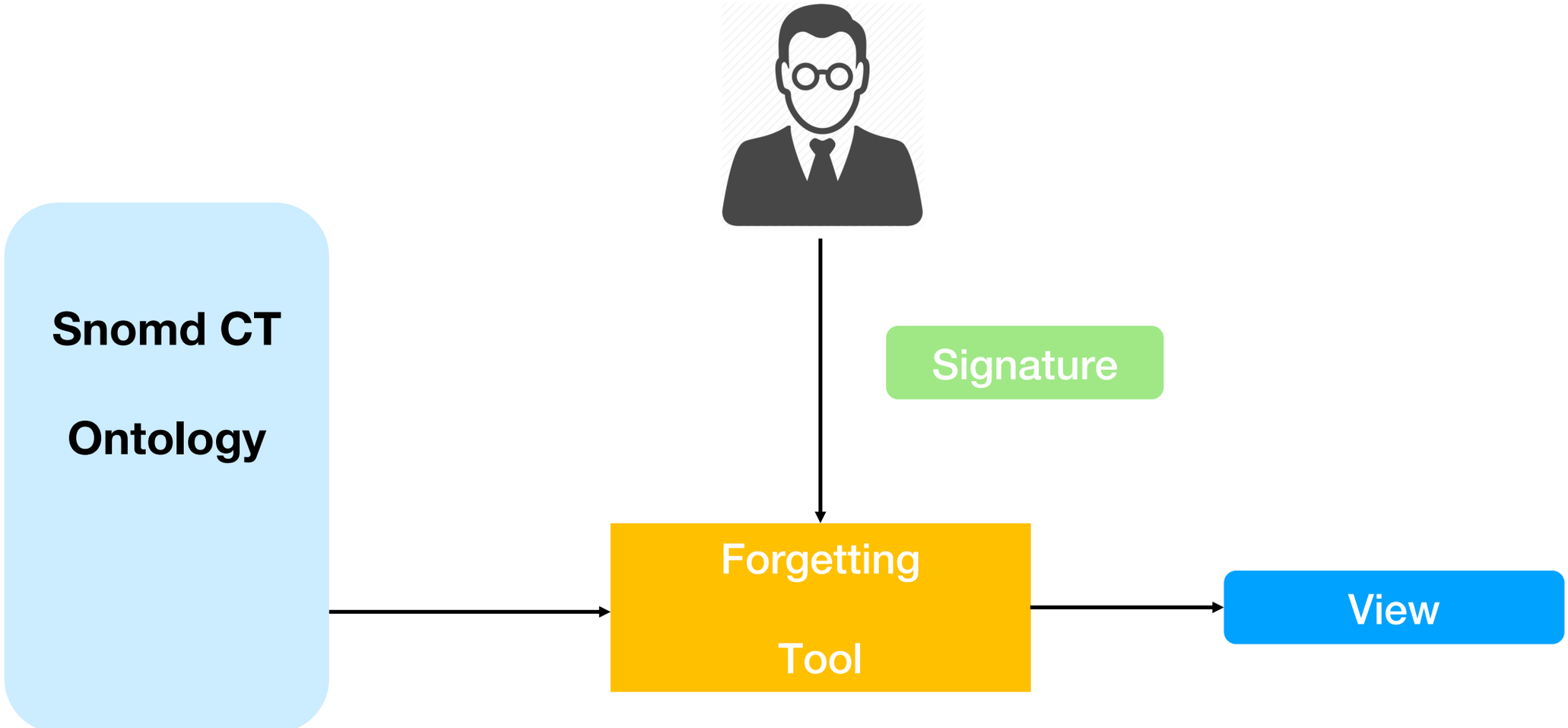
“Noise”



Modules contain noise: classes and properties that are not in signatures.

Forgetting

- a set of axioms over signature, specified classes and properties
- preserve answers to queries about signature
- no noise: not a subset of the ontology



➢ requires reasoning, hard to compute

Forgetting Tools

- NUI[1]
- LETHE[2]
- FAME[3]

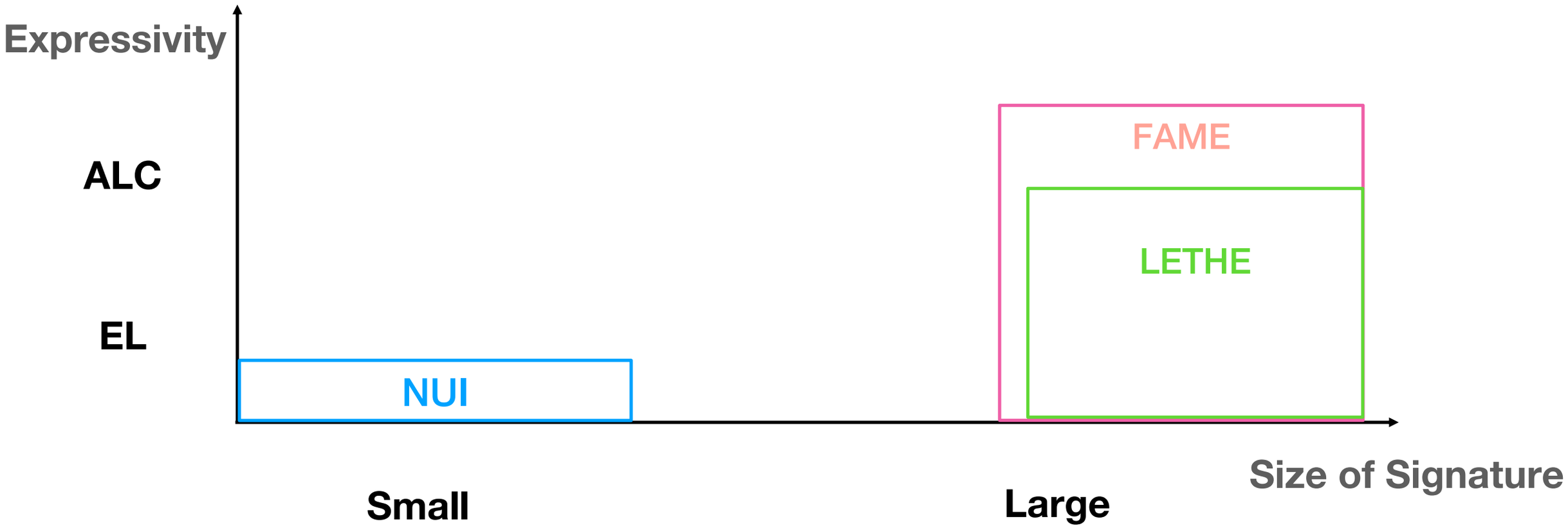
1. B.Konev, D.Walther, F.Wolter: Forgetting and Uniform Interpolation in Large-Scale Description Logic Terminologies In Proceedings of IJCAI'09.

2. P.Koopmann, R.Schimidt: https://lat.inf.tu-dresden.de/~koopmann/LETHE/

3. Y.Zhao, R.Schimidt: http://www.cs.man.ac.uk/~schmidt/sf-fame/

Forgetting Tools

- NUI[1]
- LETHE [2]
- FAME[3]

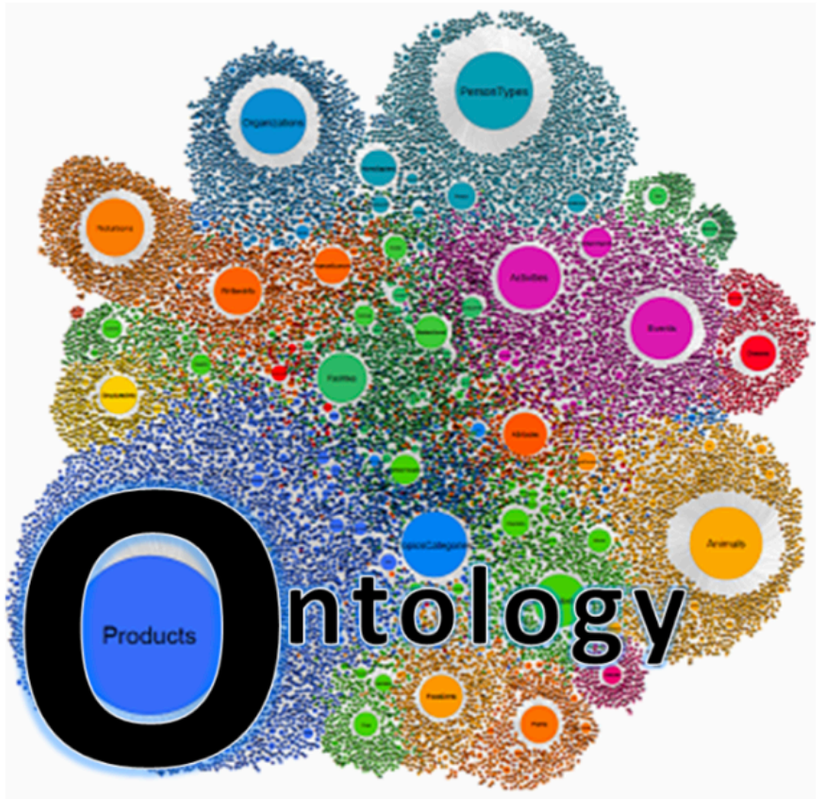


Signature Adjustment

- Assumption of the tools: the signature is given by the user

Motivation of Signature Adjustment

SNOMED CT
The global language of healthcare



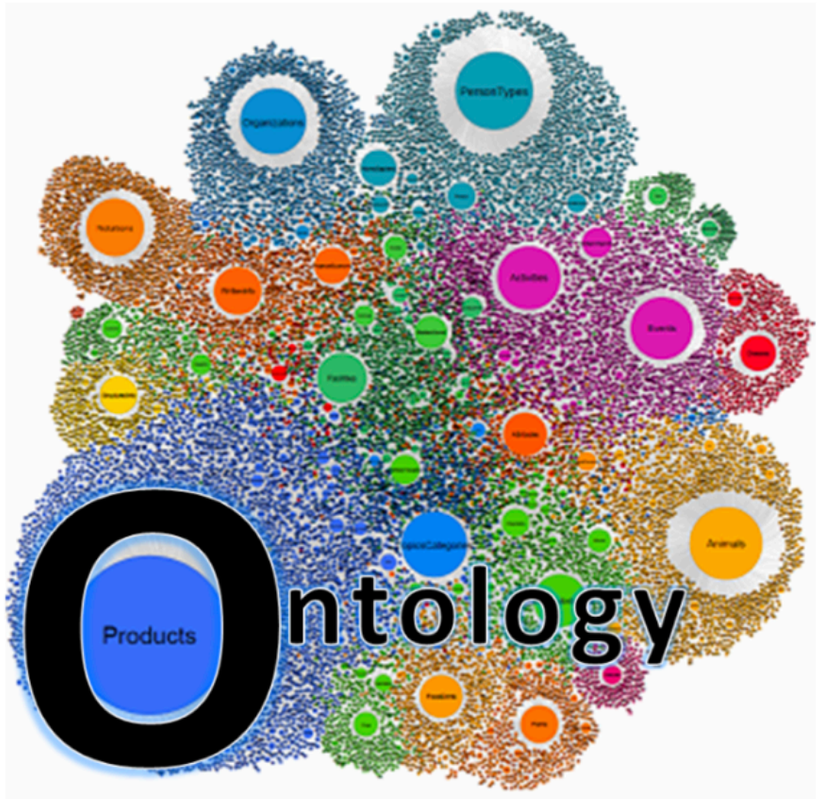
Disease

Heart structure

Heart tissue (body structure)

Motivation of Signature Adjustment

SNOMED CT
The global language of healthcare



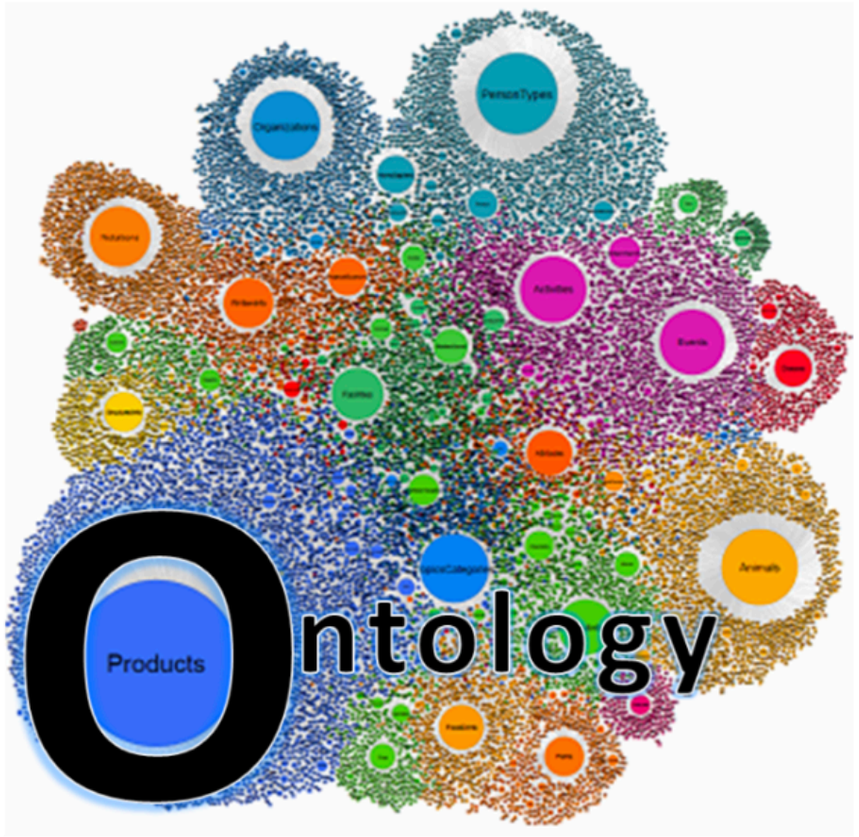
Disease

Heart structure

Heart tissue (body structure)

View is empty!

Motivation of Signature Adjustment



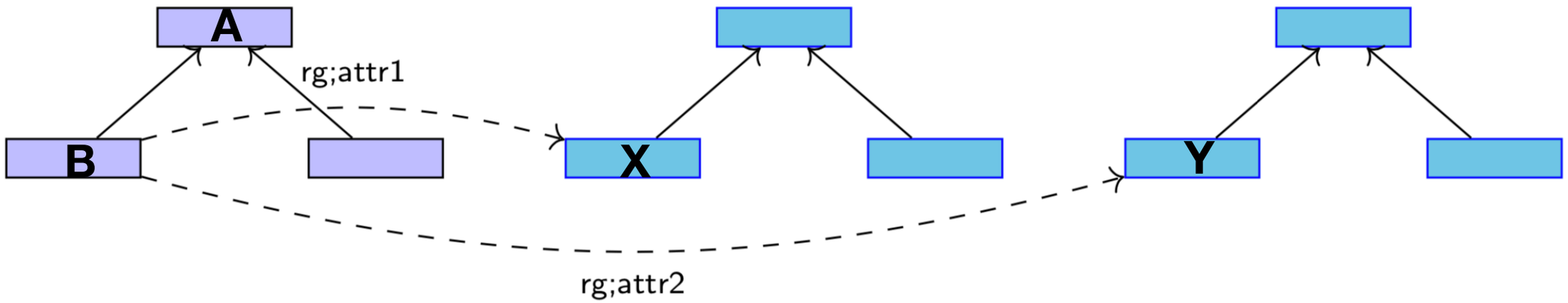

```
refcompid
140004
281004
297009
330007
368009
563001
568005
602001
792004
805002
811004
815008
965003
967006
1023001
1085006
1145003
1201005
1261007
1335005
```

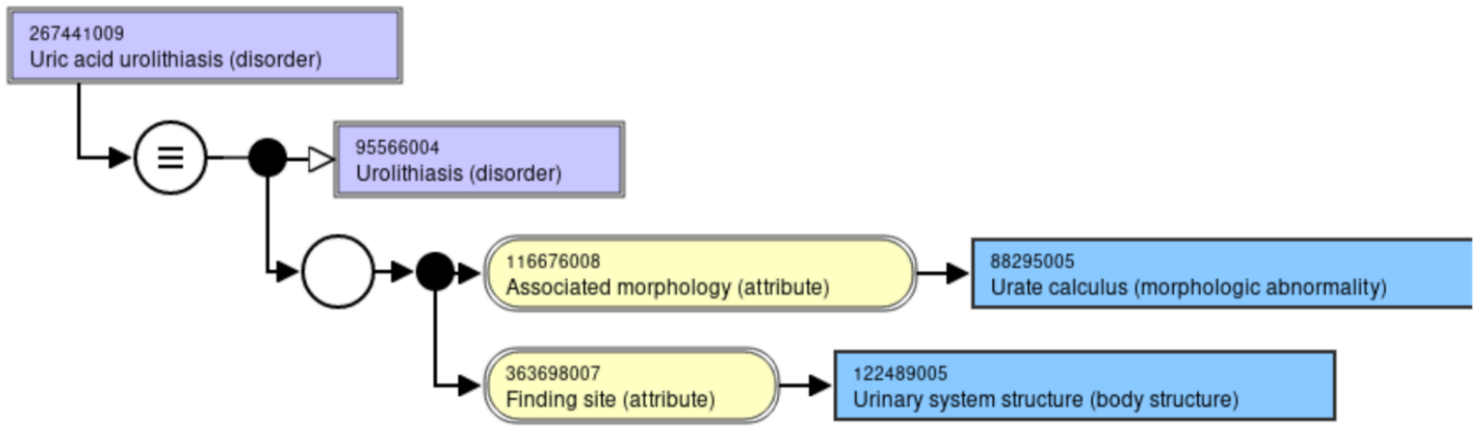
ERA refset

View is not really useful!

Signature Adjustment

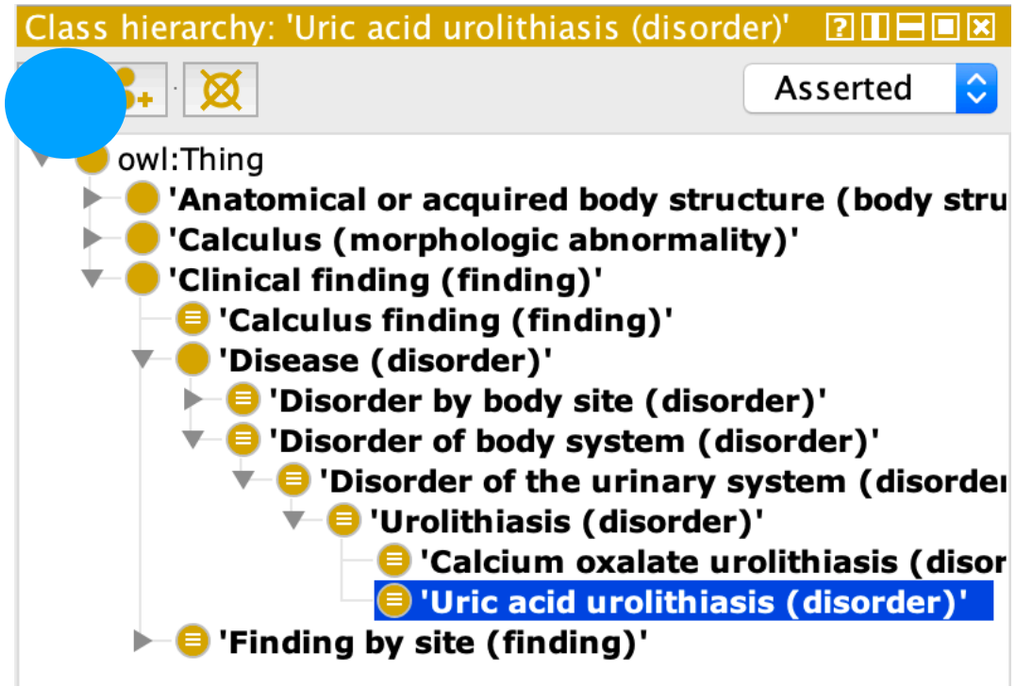
- Assumption of the tools: the signature is given by the users
- **Solution:** utilise the ontology to adjust signature

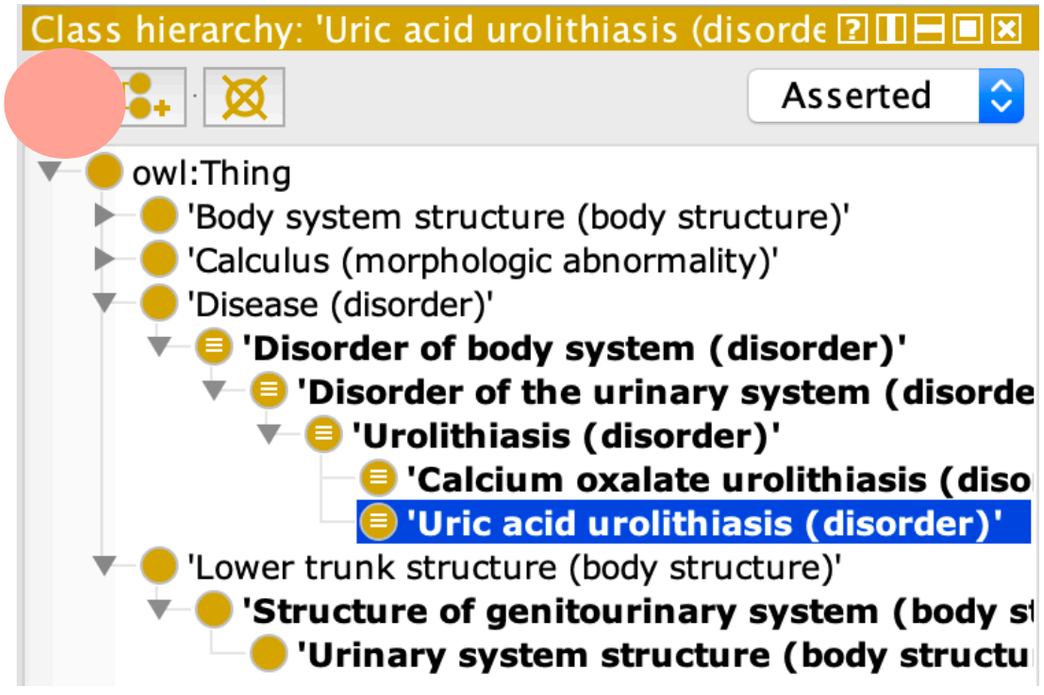


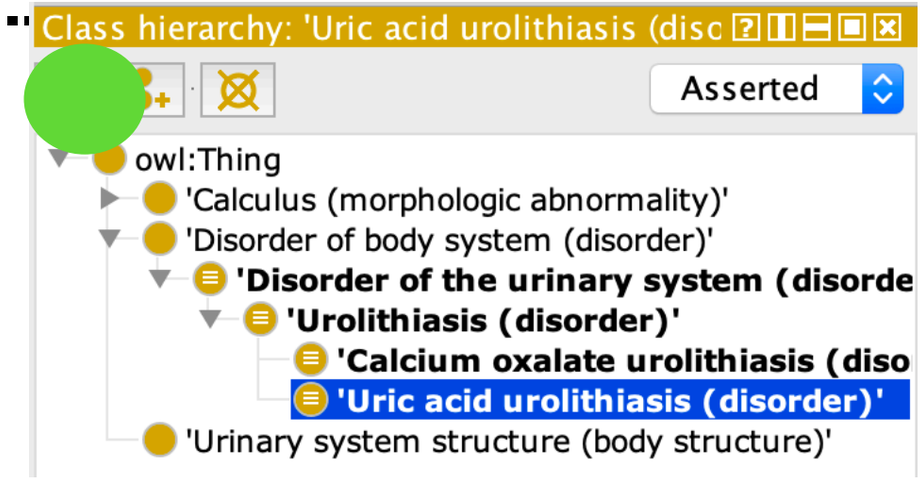


Example: Forgetting

Σ*:={ Urolithiasis (disorder), Calcium oxalate urolithiasis (disorder)
Uric acid urolithiasis (disorder), Calculus (morphologic abnormality)
Urinary system structure (body structure),... }

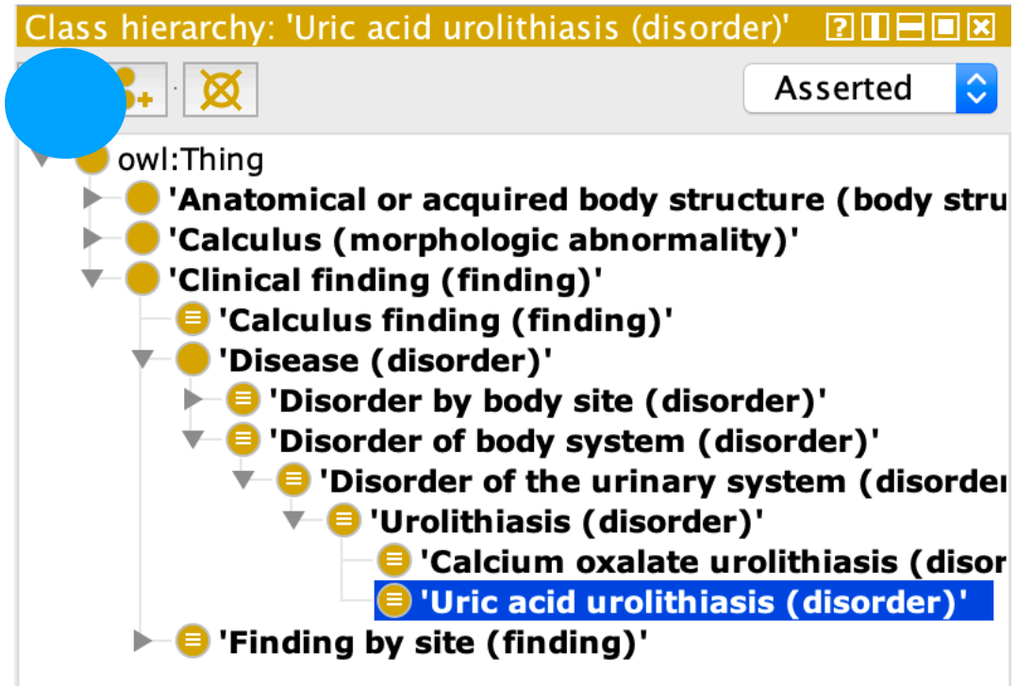


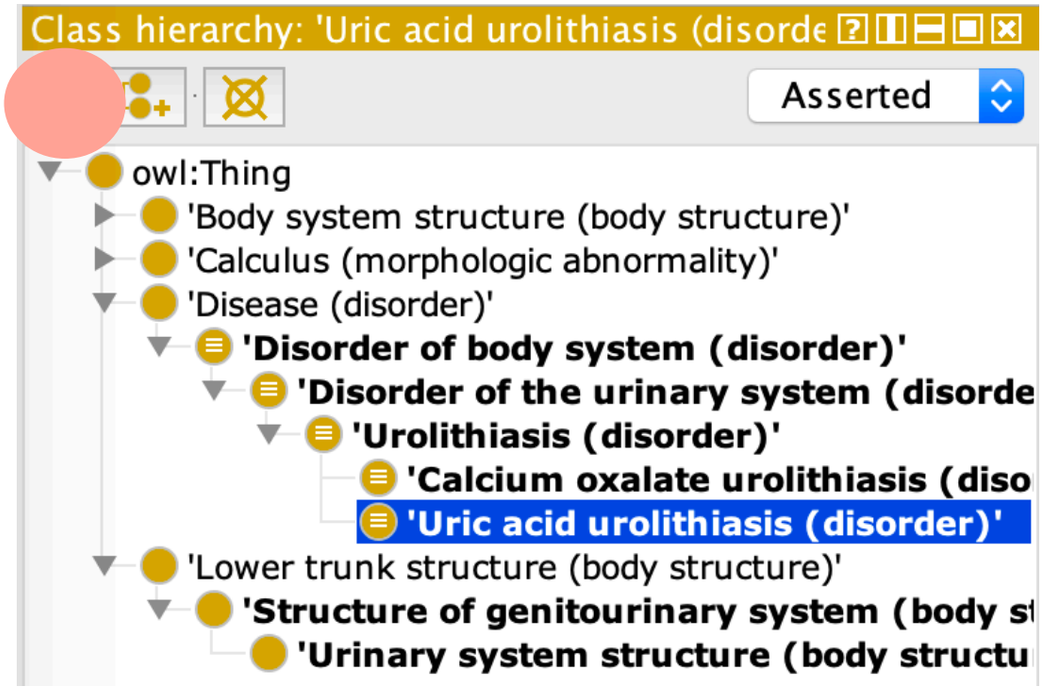


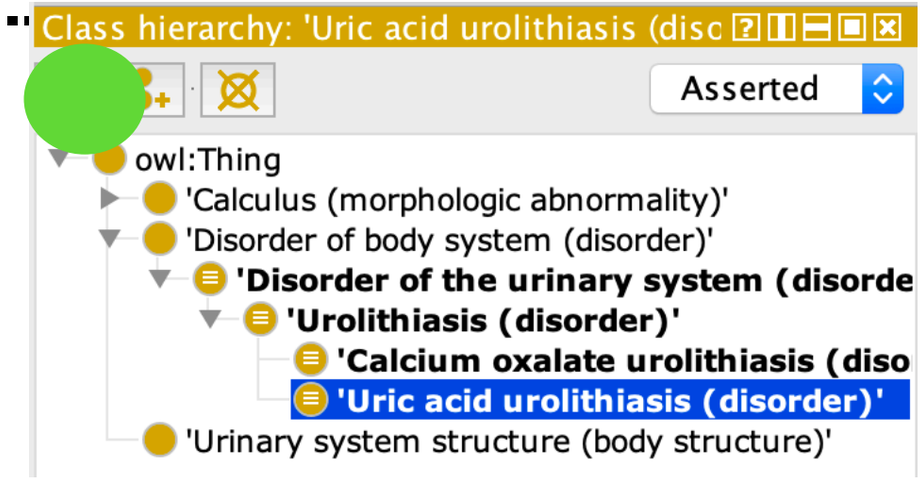


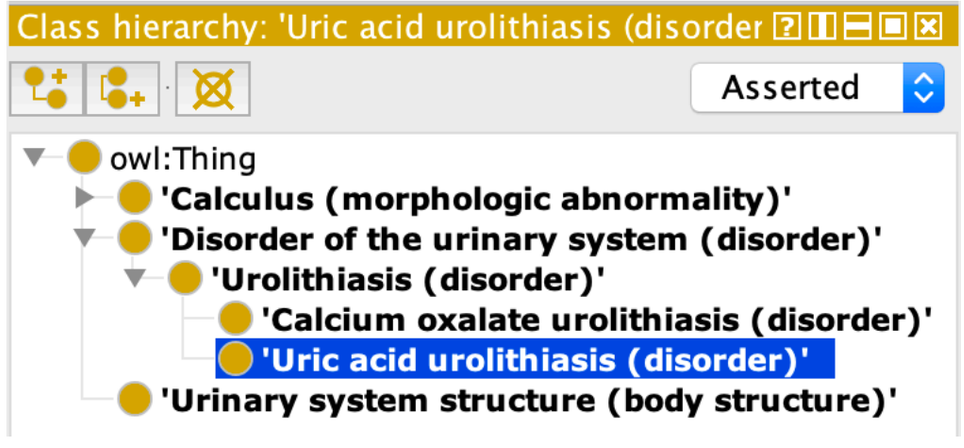
Example: Forgetting

Σ*:={ Urolithiasis (disorder), Calcium oxalate urolithiasis (disorder)
Uric acid urolithiasis (disorder), Calculus (morphologic abnormality)
Urinary system structure (body structure),... }



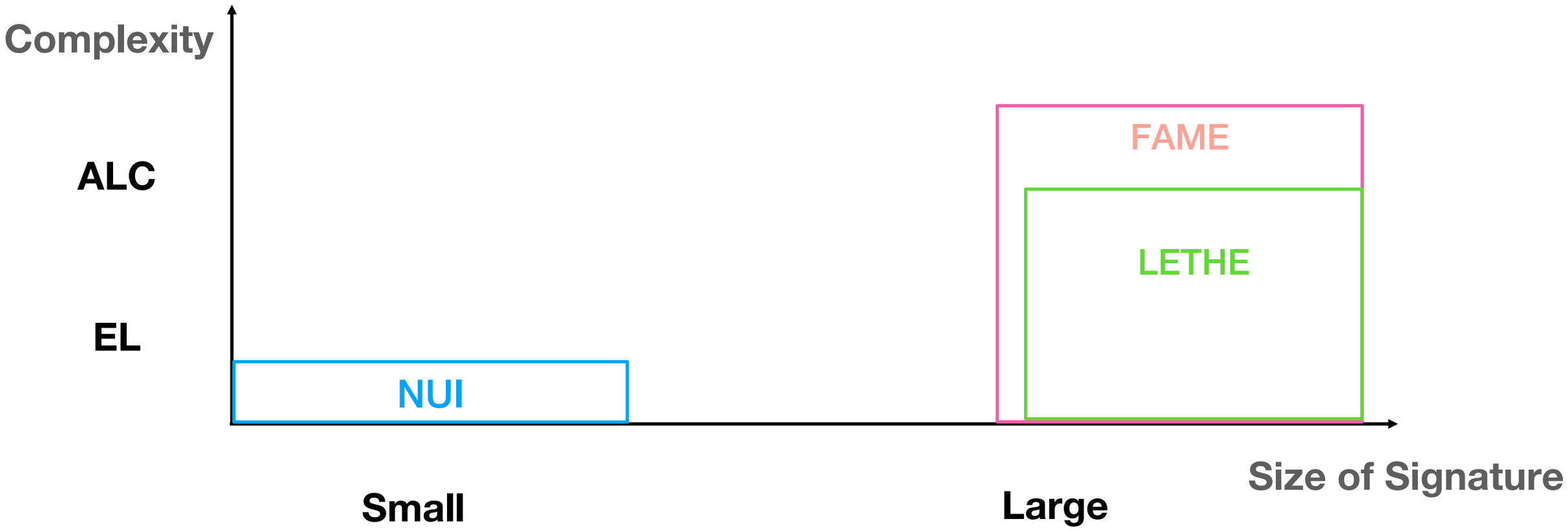






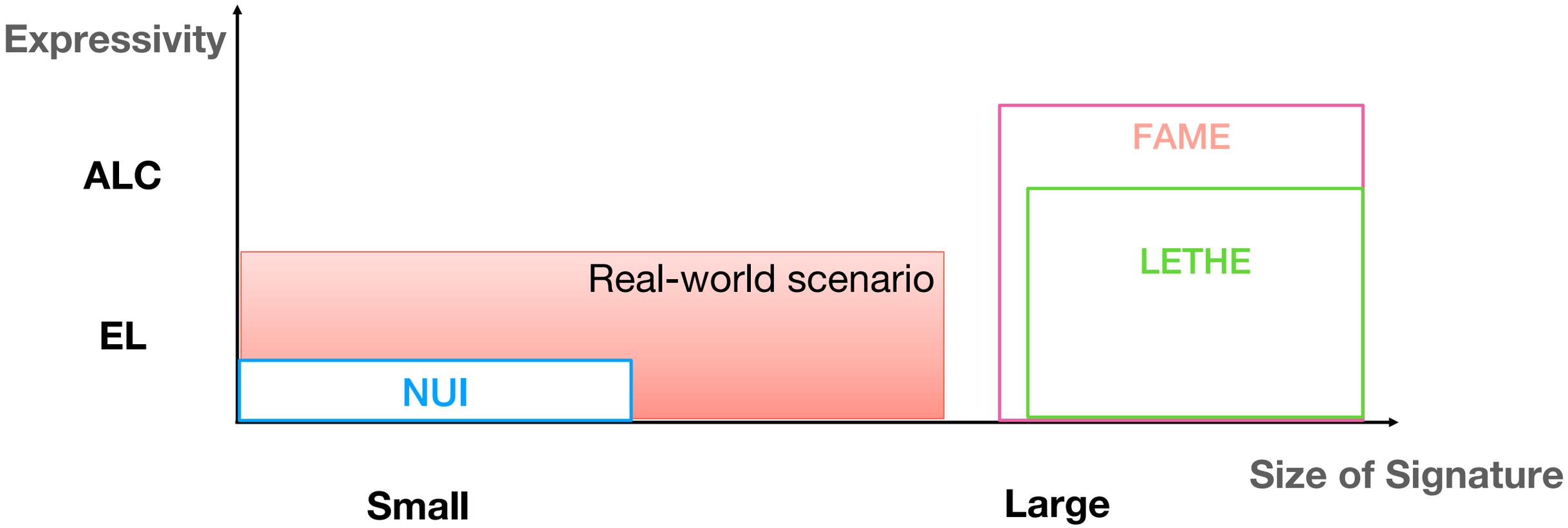
Forgetting Tools

- NUI[1]
- LETHE [2]
- FAME[3]



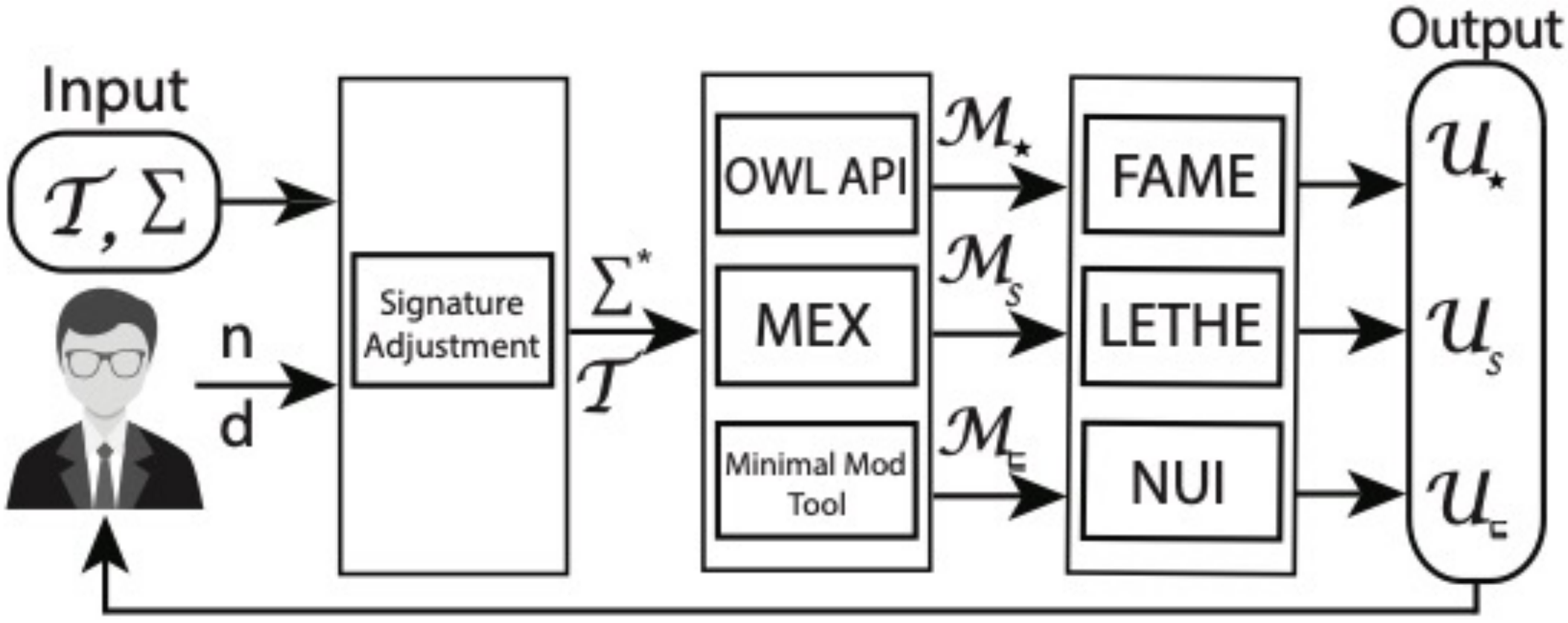
Forgetting Tools

- NUI[1]
- LETHE [2]
- FAME[3]



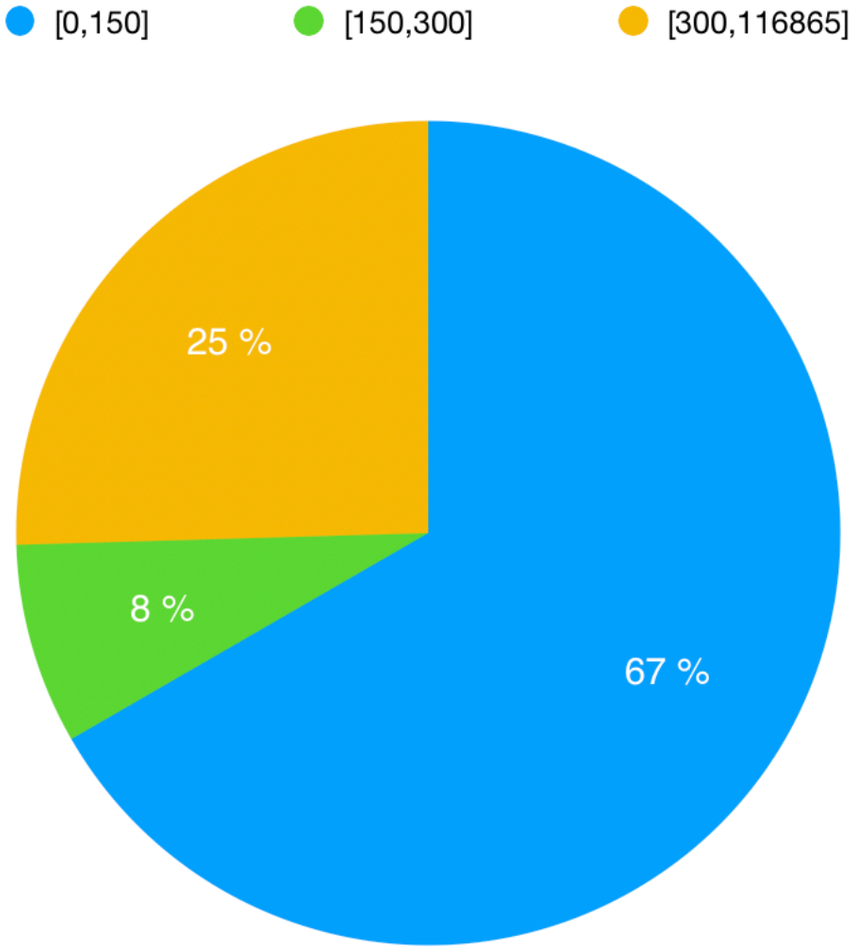
Workflow

Signature Adjustment + **Ontology Modularity** + **Forgetting**



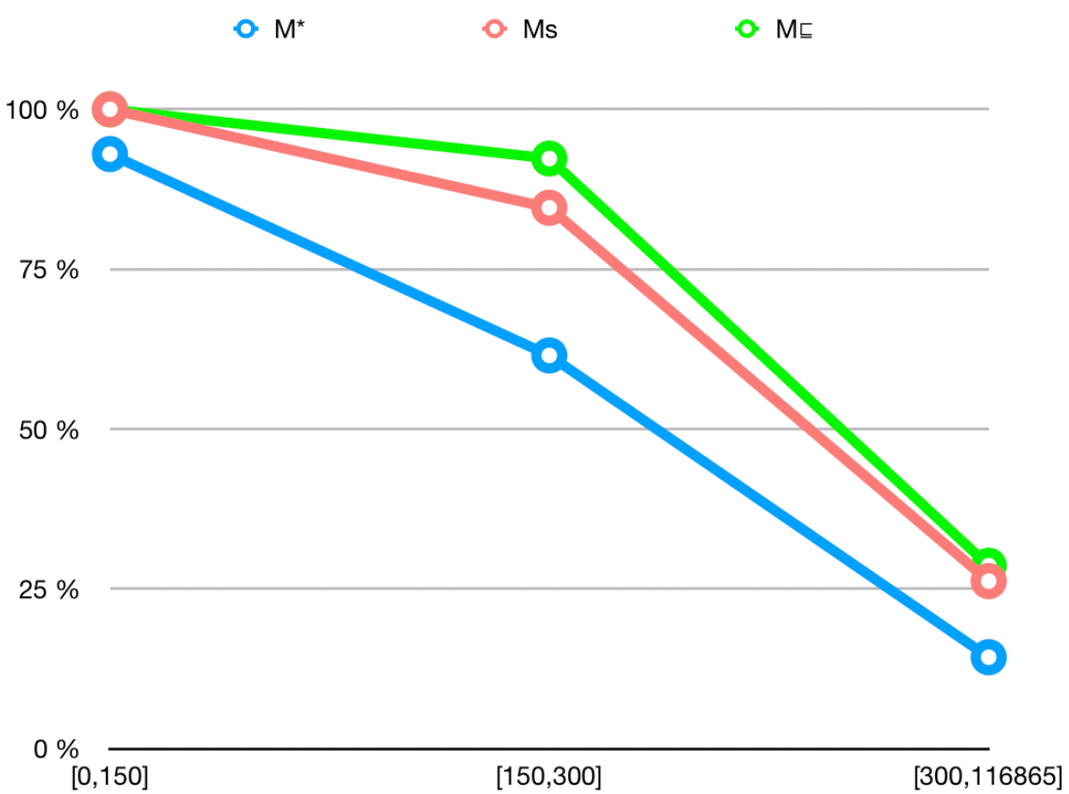
Evaluation: NHS Refsets

- In total: 165 signatures
- 0-33 properties
- 2-116 865 classes

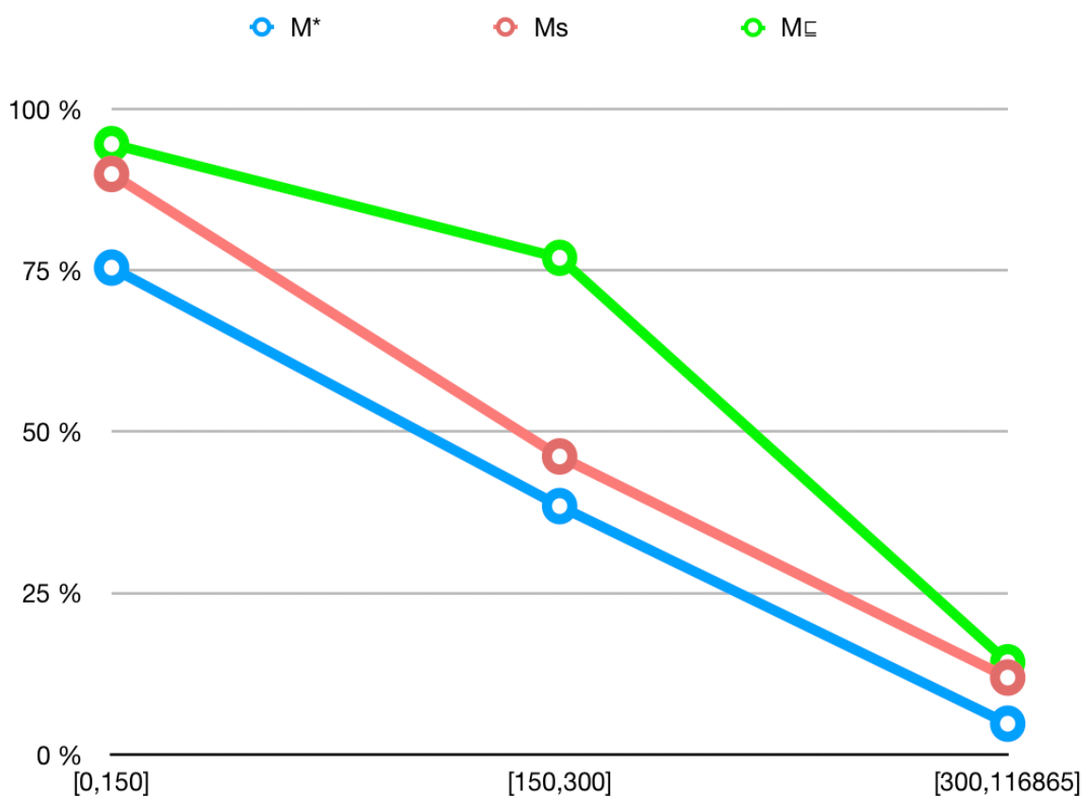


Evaluation: NHS Refsets

- **Success Rate**



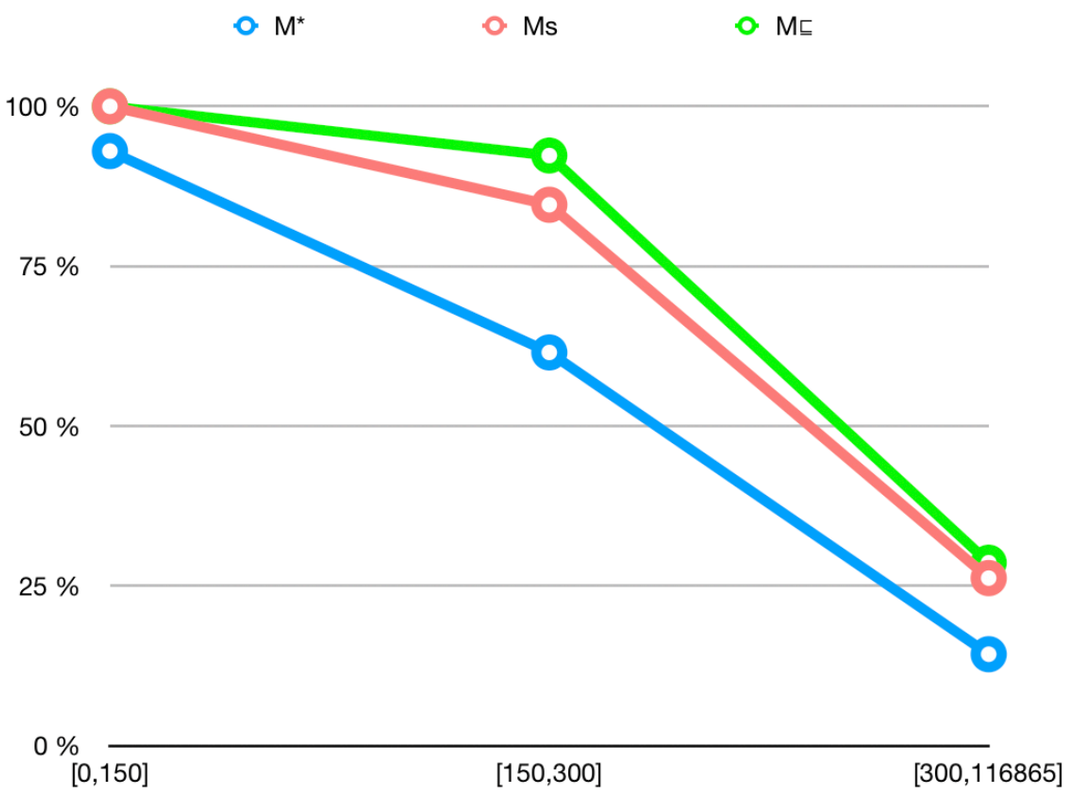
LETHE



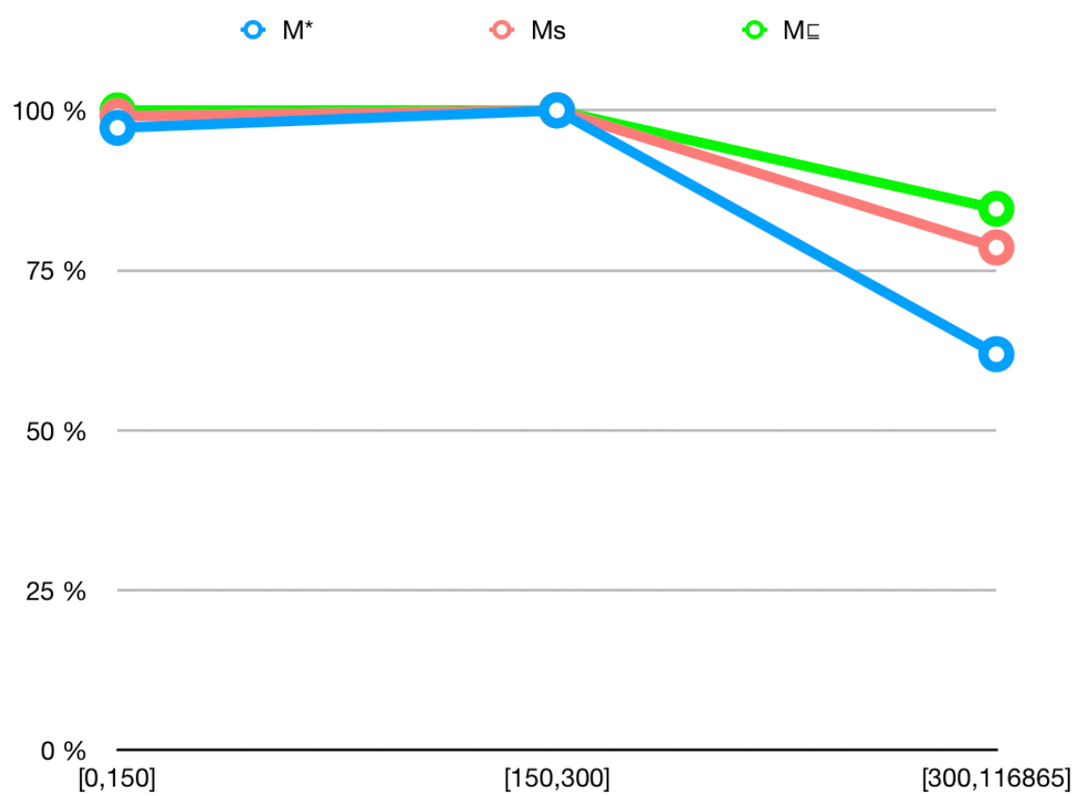
FAME

Evaluation: NHS Refsets

- **Success Rate**



LETHE

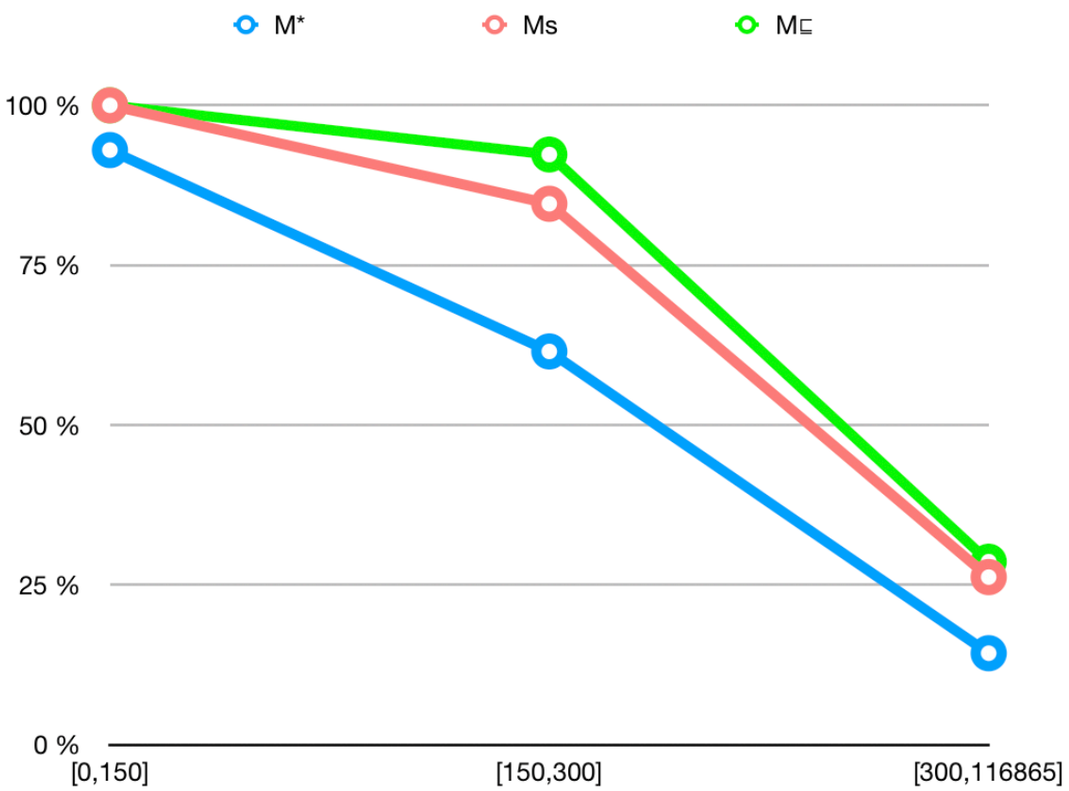


NUI

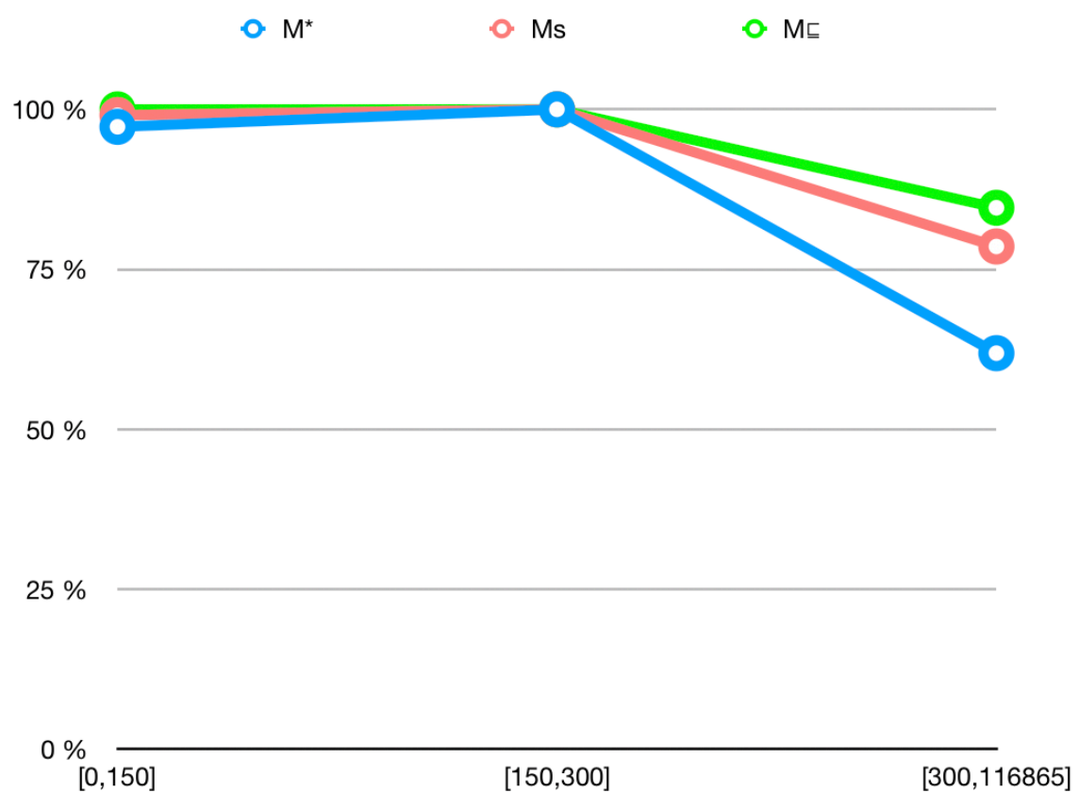
Evaluation: NHS Refsets

- **Success Rate**

Ontology Modularity improves performance of forgetting tools.

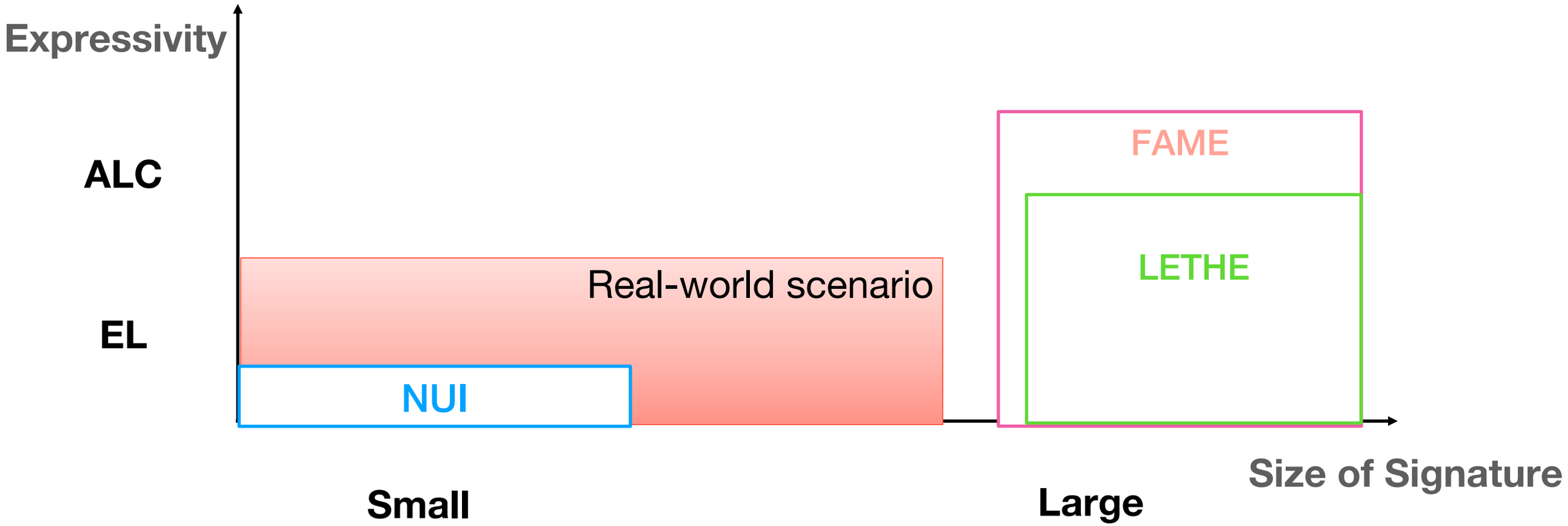


LETHE

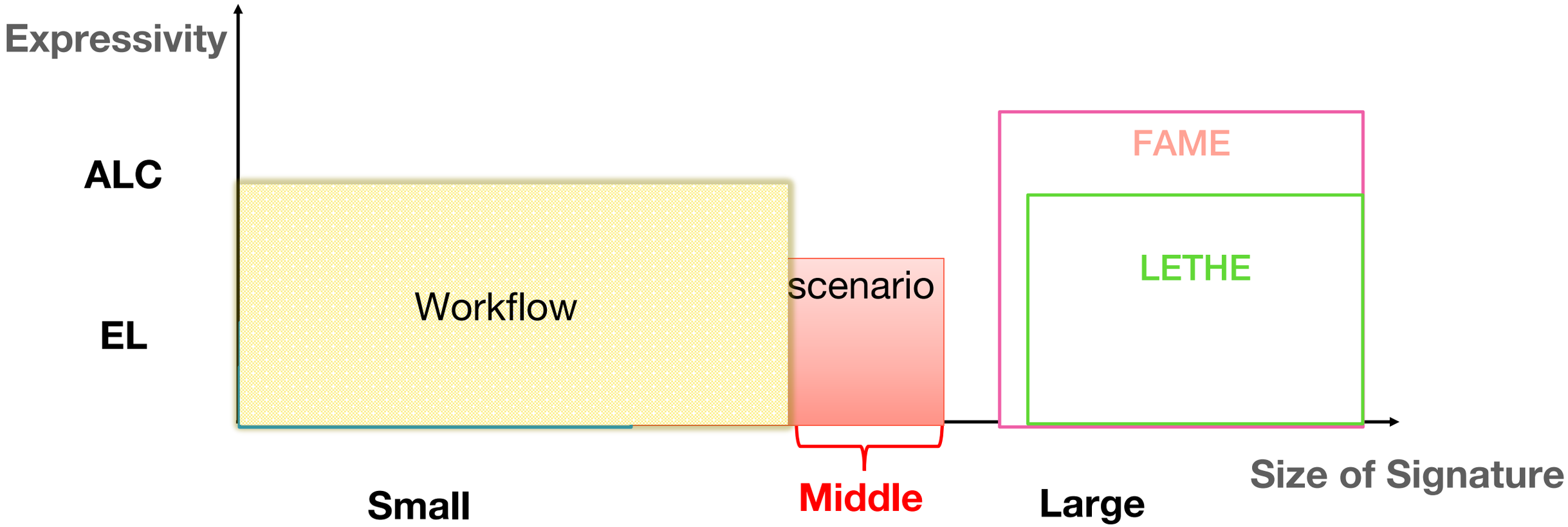


NUI

Challenge


Expressivity
ALC
EL
Real-world scenario
NUI
FAME
LETHE
Small
Large
Size of Signature

Challenge

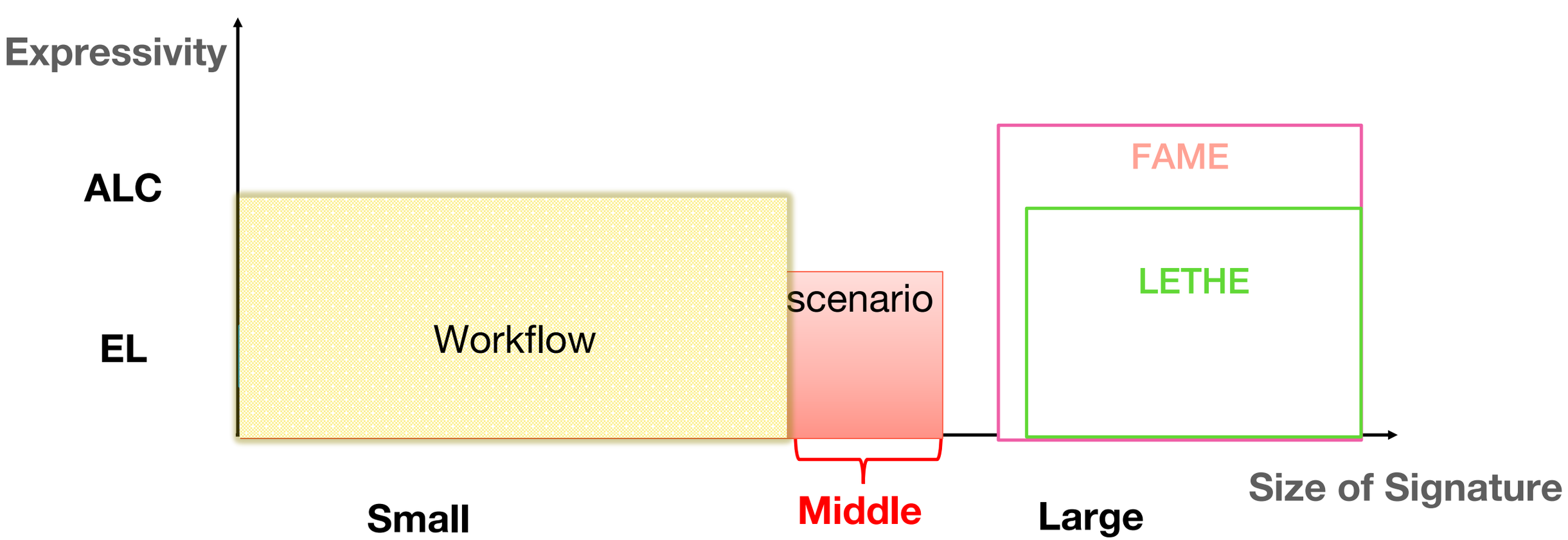

Expressivity
ALC
EL
Workflow
scenario
FAME
LETHE
Small
Middle
Large
Size of Signature

Conclusion

- **Workflow allows feasible computation of ontology extracts of SNOMED CT**
- **Signature adjustment techniques**

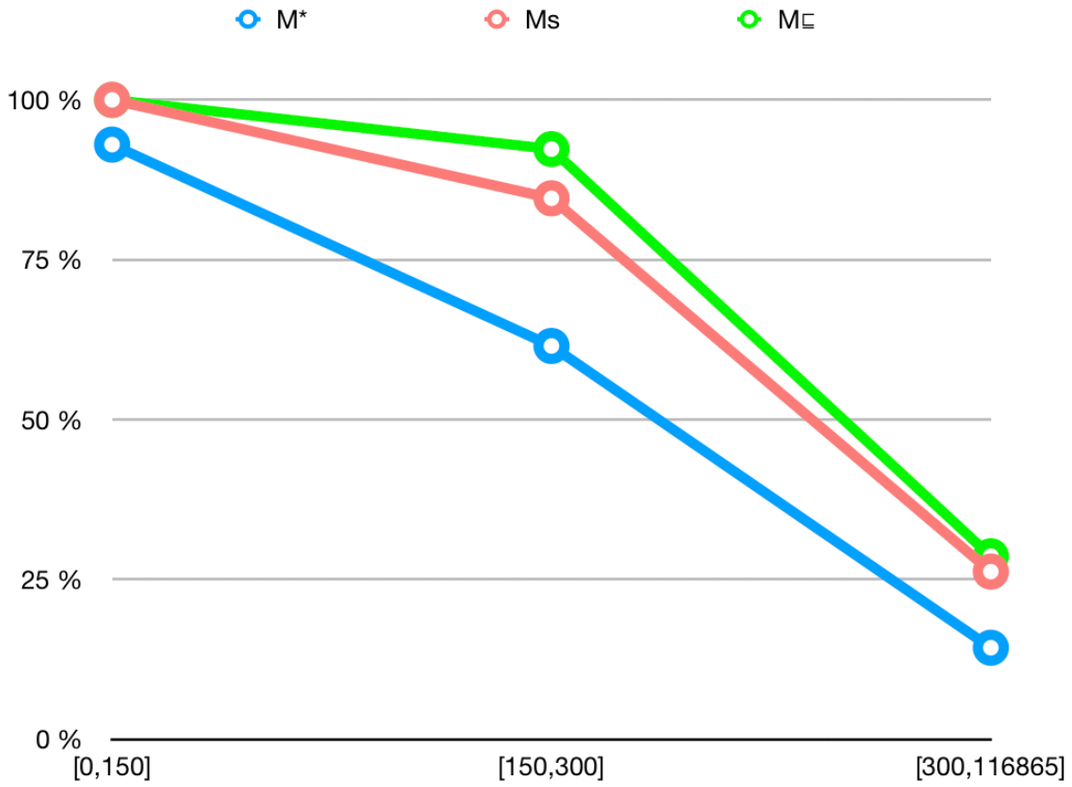
i. allow user to specify and enhance quality of extracts

ii. small sets of key concepts

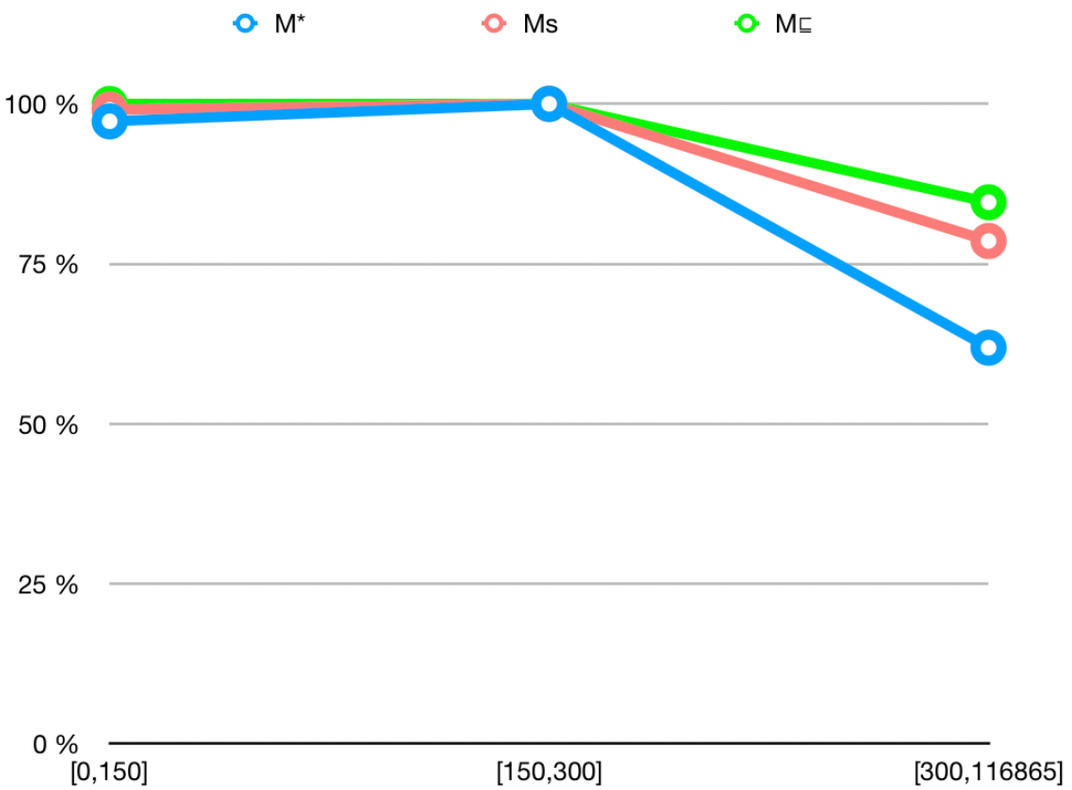


Future Work

- Extend algorithm for computing minimal subsumption modules to handle new features of SNOMED CT
- Continue to improve our forgetting tools



LETHE



NUI


рахмат
danke
謝謝
ngiyabonga
teşekkür ederim
spas
Баярлалаа
mersi
kia ora
barka
welalin
tack
mahalo
tapadh leat
спасибо
faafetai lava
vinaka
спасибі
blagodaram
dank je
misaotra
matondo
paldies
grazzi
thank you
gracias
хвала
kiitos
dankie
nanni
nandri
dhanyavad
maururu
köszönöm
asante
manana
obrigada
tenki
bedankt
bayarlalaa
gràcie
hvala
akun
dankon
aciū
djiere dieuf
tau
mochchakkeram
murakoze
enkosi
dziękuję
дякую
mamnun
chokrane
chnorakaloutioun
gratias ago
gràcies
sulpáy
go raibh maith agat
sobodi
dėkuji
sagolun
sukriya
kop khun krap
taiku
obrigado
mèsi
didi madloba
ありがとう
arigatô
takk
dakujem
trugarez
najis tuke
tanemirt
grazie
terima kasih
rahmet
kam sah hamnida
rahmat
diolch
dhanyavadagalu
shukriya
mercé
мерси
তোমাকে ধন্যবাদ
xièxie
ευχαριστώ
감사합니다
merci