Assisting Data Retrieval With a Drug Knowledge Graph

Romain LELONG[a,b,1], Badisse DAHAMNA[a,b], Romain LEGUILLON[a], Julien GROSJEAN[a,b], Catherine LETORD[a,b], Stéfan J.DARMONI[a,b], Lina F. SOUALMIA[b,c]

[a] *CHU Rouen, Department of Biomedical Informatics, F-76000 Rouen, France*
[b] *LIMICS U1142, Sorbonne Université, Paris, France*
[c] *Normandie Univ, UNIROUEN, TIBS-LITIS EA 4108, F-76000 Rouen, France*

12 octobre 2021

1. romain.lelong@gmail.com

Prescription orders integration

1,452,616 prescription orders
6,978,586 atomic prescriptions
2011 $\longrightarrow$ August 2021

Information retrieval ?

Rouen University Hospital (RUH)

RUH Health Data Warehouse
EDSaN

Prescriptions metadata & Information Retrieval

Patient information
id, age, gender, birthdate

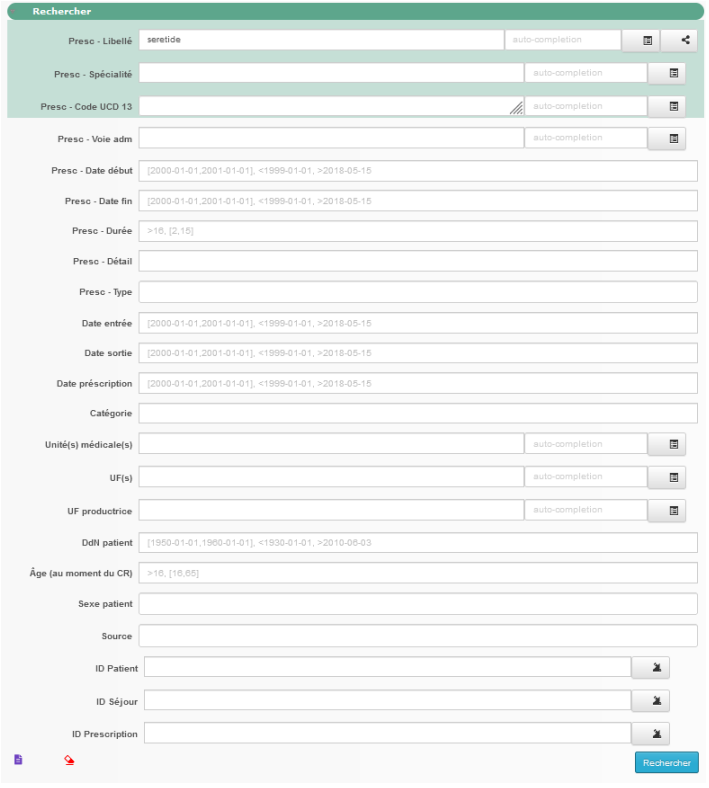
Stay information
id, entry date, leaving date, units

prescription information
id, date

List of prescribed drugs
drug label, details, UCD codes

APACHE LUCENE

Lucene indexes ⇒ allow the search by metadata



Prescriptions metadata & Information Retrieval

Search for prescription orders based on what drug were prescribed ?

Patient information
id, age, gender, birthdate

Stay information
id, entry date, leaving date, units

prescription information
id, date

List of prescribed drugs
drug label, details, UCD codes

UCD = Common Dispensing Unit

$\Rightarrow$ Encoding and billing of administrated drugs in France

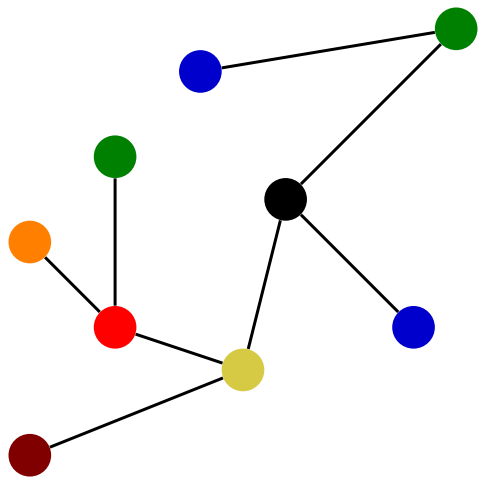
$\Rightarrow$ Good quality of data

Selecting drugs
=
Selecting UCD codes

Drug Knowledge Graph

Medical Subject Heading descriptors
Anatomical Therapeutic Chemical (ATC) concepts
International Non proprietary Names (INNs)
Virtual drugs from the MédicaBase
Pharmacological roots
Pharmacological specialties
Drug components
Drug components groups
UCD codes

131,277 vertices
703,807 edges

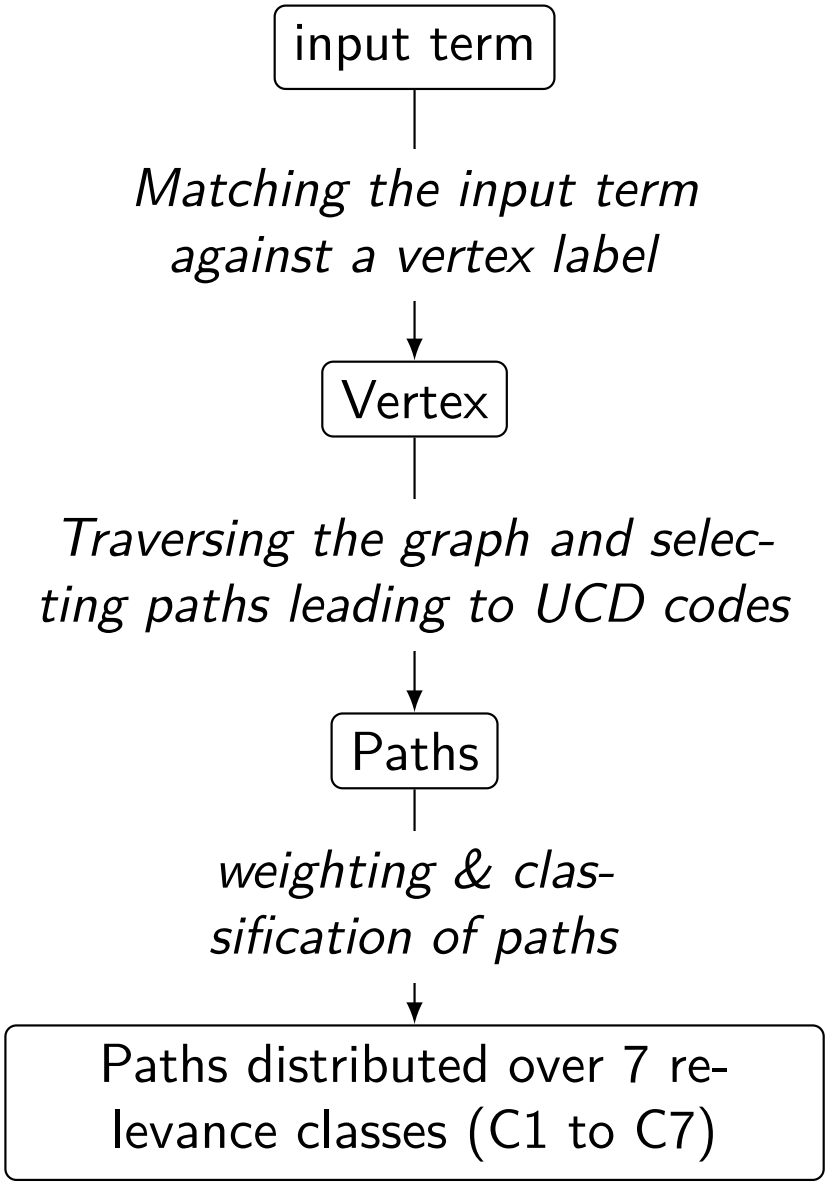


Health Terminology/Ontology Portal
$>$ 90 terminologies or ontologies
$>$ 2 million concepts

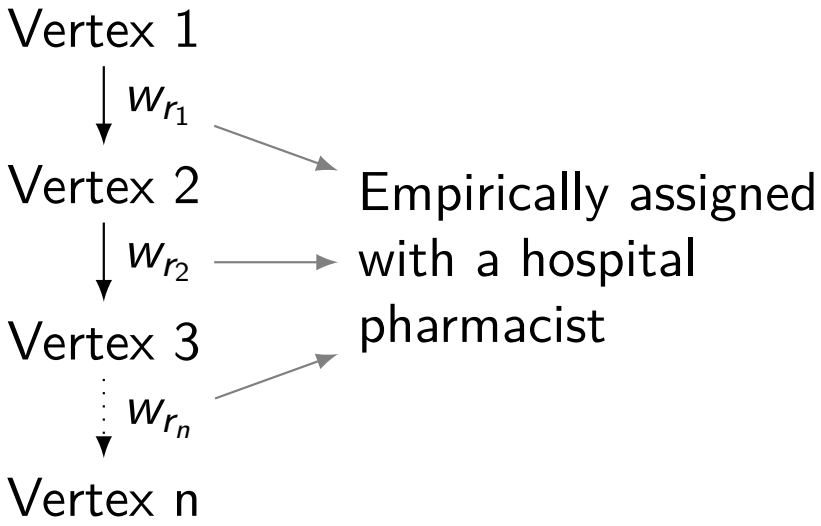
Conceptual graph of drug-specific concept

2. https ://hetop.eu

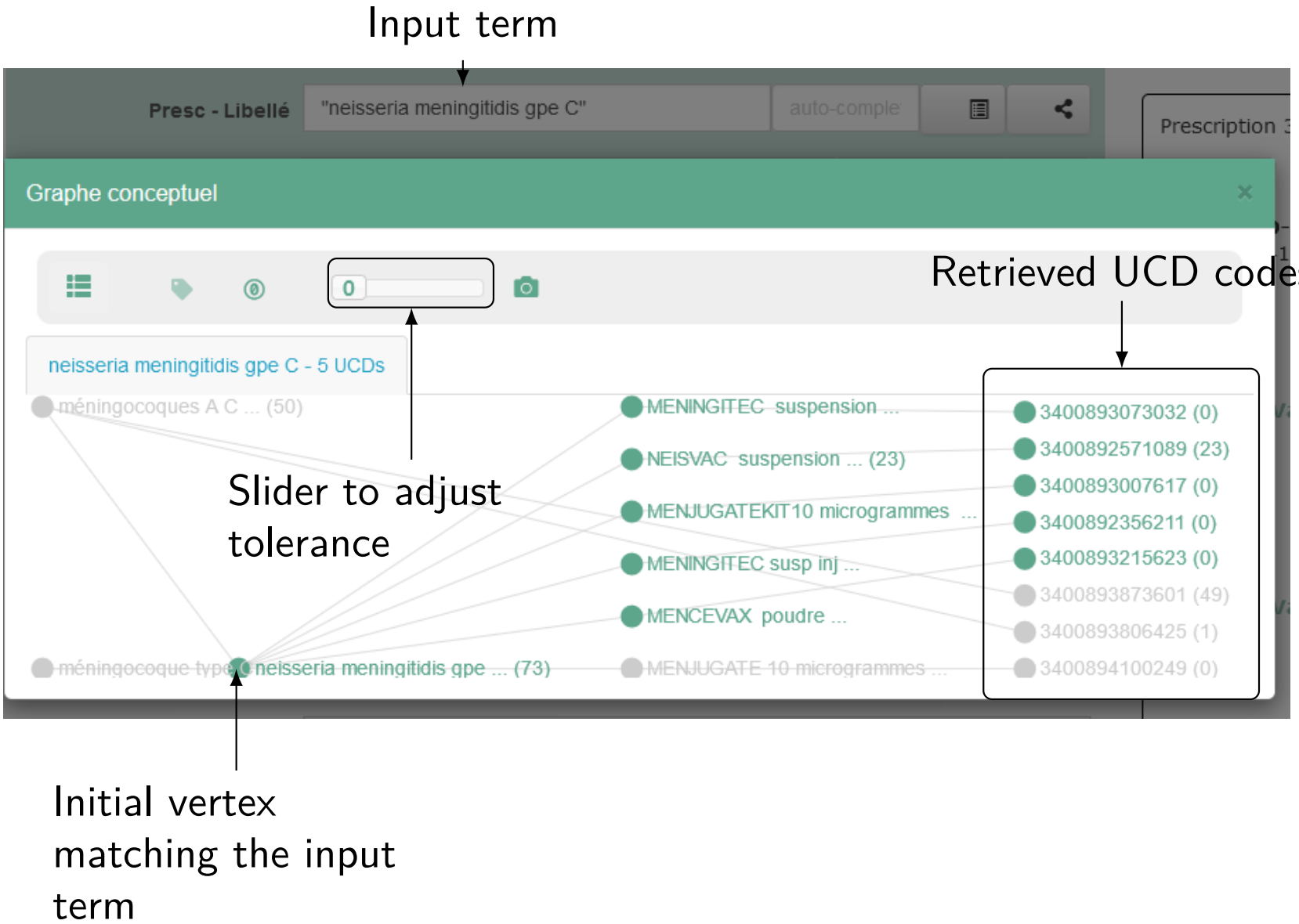
Exploitation of the knowledge graph



Paths weighting



$$w_{path} = w_{r_1} + w_{r_2} + \cdots + w_{r_n}$$


Input term
Presc - Libellé
"neisseria meningitidis gpe C"
auto-comple
Prescription 3
Graphe conceptuel
0
Retrieved UCD codes
neisseria meningitidis gpe C - 5 UCDs
méningocoques A C ... (50)
MENINGITEC suspension ...
NEISVAC suspension ... (23)
MENJUGATEKIT10 microgrammes ...
MENINGITEC susp inj ...
MENCEVAX poudre ...
méningocoque type C
neisseria meningitidis gpe ... (73)
MENJUGATE 10 microgrammes ...
3400893073032 (0)
3400892571089 (23)
3400893007617 (0)
3400892356211 (0)
3400893215623 (0)
3400893873601 (49)
3400893806425 (1)
3400894100249 (0)
Slider to adjust tolerance
Initial vertex matching the input term

Evaluation

88 terms randomly drawn among the possible types of vertex.

System

Only the paths/UCD codes from the first three non-empty relevance classes were considered. More than 100,000 paths were retrieved.

Hospital pharmacist evaluation

Attribution of a score to each path/UCD code from 1 to 3 :
1 unsatisfactory
2 could be improved
3 consistent result.

Result & Discussion

Type of term	C_1	C_2	C_3	C_4	C_5
Drug composition	2.71	1.08	1.04		
Drug composition groups	-	3.00	1.00	1.00	
MeSH Descriptors	-	2.94	1.49	1.05	
INN	-	-	1.95	1.01	1.00
virtual drugs	3.00	1.23	1.28		
Medical indications	-	-	3.00	2.82	1.52
ATC Code	2.25	-	3.00	1.00	
pharmacological specialties	2.50	1.59	1.89		
pharmacological roots	3.00	1.84	1.05		

1. Overall, average scores of paths tends to decrease with level of relevance but exceptions can be found ;
2. Highest score of 3 (or close to 3) reached for several types of terms ;
3. Variability of the best possible relevance class.

Conclusion

Does the drug knowledge graph assist the user in selecting drugs of interest ?

Yes : The ranking of resulting UCDs codes were overall congruent with the expert judgment.

But : Some inconsistencies remain.

$\Rightarrow$ A refinement of the weights assigned to the edges of the graph is necessary. !

Thank you for your attention

Questions ?