Introduction to Knowledge Graphs and Semantic Web technologies for integrative bioinformatics

Alban Gaignard, CNRS, institut du thorax, Nantes

DUBii - 2021

Knowledge Graphs, quick introduction

Google

pasteur

Tous | Maps | Actualités | Images | Vidéos | Plus | Paramètres | Outils

Environ 71300000 résultats (0,59 secondes)

Institut Pasteur | Pour la recherche, pour la santé, pour demain
https://www.pasteur.fr/fr
Faire un pour la recherche - Institut **Pasteur**. Fermer. EN · FR · Accueil. Saisissez vos mots-clés. L'Institut **Pasteur**. Retour L'Institut **Pasteur** · Notre histoire.
Notre histoire · Centre médical Institut Pasteur · Institut Pasteur · Dons - pasteur

Histoire de Louis Pasteur et de l'Institut Pasteur
https://www.pasteur.fr/fr/institut-pasteur/notre-histoire
Découvrez l'histoire de l'Institut **Pasteur** depuis 1888 et les diverses actions de Louis **Pasteur**, son combat pour la vaccination et la recherche médicale.

Pasteur — Wikipédia
https://fr.wikipedia.org/wiki/Pasteur
Le mot **pasteur** peut avoir plusieurs significations. Sommaire. 1 Fonction; 2 Saints chrétiens; 3 Patronyme; 4 Toponyme; 5 Spectacles. 5.1 Films; 5.2 Pièce de ...
Louis Pasteur · Pasteur (christianisme) · Institut Pasteur

Louis Pasteur — Wikipédia
https://fr.wikipedia.org/wiki/Louis_Pasteur
Louis **Pasteur**, né à Dole (Jura) le 27 décembre 1822 et mort à Marnes-la- Coquette (Hauts-de-Seine, à cette époque en Seine-et-Oise) le 28 septembre 1895 ...
Renommé pour: Vaccin contre la rage **Domaines**: Chimie, microbiologie
Nationalité: Française **Étudiants en thèse**: Charles Friedel
Catégorie:Louis Pasteur · Pasteur Vallery-Radot · Maison de Louis Pasteur à ...

Note · Horaires

Centre Pasteur Dermatologie
2,6 ★★★★★ (54) · Clinique dermatologique
6bis Boulevard Pasteur · 02 51 84 06 06
Ouvert · Ferme à 12:30
SITE WEB ITINÉRAIRE

Maison de Retraite Bon Pasteur
5,0 ★★★★★ (1) · Maison de retraite
11 Rue du Haut Moreau · 02 40 74 37 31
ITINÉRAIRE

Louis Pasteur

Scientifique

Louis Pasteur, né à Dole le 27 décembre 1822 et mort à Marnes-la-Coquette le 28 septembre 1895, est un scientifique français, chimiste et physicien de formation. Pionnier de la microbiologie, il connut, de son vivant même, une grande notoriété pour avoir mis au point un vaccin contre la rage. Wikipédia

Date et lieu de naissance : 27 décembre 1822, Dole

Date et lieu de décès : 28 septembre 1895, Marnes-la-Coquette

Enfants : Jean-Baptiste Pasteur, Marie-Louise Pasteur, Jeanne Pasteur, Cécile Pasteur, Camille Pasteur

Invention : Pasteurisation

Enseignement : École Normale Supérieure (1847), PLUS

Livres

Voir d'autres éléments (plus de 20)

Studies on fermentati... 1876 · Germ Theory and Its Applic... · Écrits scientifiques et médica... · Oeuvres de Pasteur · Germ Theory and Its Applic...

Recherches associées

Voir d'autres éléments (plus de 15)

René Descartes (Tendances) · Robert Koch · Joseph Lister · Alexander Fleming · Edward Jenner

Commentaires

Google

pasteur

Tous | Maps | Actualités | Images | Vidéos | Plus | Paramètres | Outils

Environ 71300000 résultats (0,59 secondes)

Institut Pasteur | Pour la recherche, pour la santé, pour demain
https://www.pasteur.fr/fr
Faire un pour la recherche - Institut **Pasteur**. Fermer. EN · FR · Accueil. Saisissez vos mots-clés. L'Institut **Pasteur**. Retour L'Institut **Pasteur** · Notre histoire.
Notre histoire · Centre médical Institut Pasteur · Institut Pasteur · Dons - pasteur

Histoire de Louis Pasteur et de l'Institut Pasteur
https://www.pasteur.fr/fr/institut-pasteur/notre-histoire
Découvrez l'histoire de l'Institut **Pasteur** depuis 1888 et les diverses actions de Louis **Pasteur**, son combat pour la vaccination et la recherche médicale.

Pasteur — Wikipédia
https://fr.wikipedia.org/wiki/Pasteur
Le mot **pasteur** peut avoir plusieurs significations. Sommaire. 1 Fonction; 2 Saints chrétiens; 3 Patronyme; 4 Toponyme; 5 Spectacles. 5.1 Films; 5.2 Pièce de ...
Louis Pasteur · Pasteur (christianisme) · Institut Pasteur

Louis Pasteur — Wikipédia
https://fr.wikipedia.org/wiki/Louis_Pasteur
Louis **Pasteur**, né à Dole (Jura) le 27 décembre 1822 et mort à Marnes-la- Coquette (Hauts-de-Seine, à cette époque en Seine-et-Oise) le 28 septembre 1895 ...
Renommé pour: Vaccin contre la rage **Domaines**: Chimie, microbiologie
Nationalité: Française **Étudiants en thèse**: Charles Friedel
Catégorie:Louis Pasteur · Pasteur Vallery-Radot · Maison de Louis Pasteur à ...

Note · Horaires

Centre Pasteur Dermatologie
2,6 ★★★★★ (54) · Clinique dermatologique
6bis Boulevard Pasteur · 02 51 84 06 06
Ouvert · Ferme à 12:30
SITE WEB ITINÉRAIRE

Maison de Retraite Bon Pasteur
5,0 ★★★★★ (1) · Maison de retraite
11 Rue du Haut Moreau · 02 40 74 37 31
ITINÉRAIRE

Louis Pasteur

Scientifique

Louis Pasteur, né à Dole le 27 décembre 1822 et mort à Marnes-la-Coquette le 28 septembre 1895, est un scientifique français, chimiste et physicien de formation. Pionnier de la microbiologie, il connut, de son vivant même, une grande notoriété pour avoir mis au point un vaccin contre la rage. Wikipédia

Date et lieu de naissance : 27 décembre 1822, Dole

Date et lieu de décès : 28 septembre 1895, Marnes-la-Coquette

Enfants : Jean-Baptiste Pasteur, Marie-Louise Pasteur, Jeanne Pasteur, Cécile Pasteur, Camille Pasteur

Invention : Pasteurisation

Enseignement : École Normale Supérieure (1847), PLUS

Livres

Voir d'autres éléments (plus de 20)

Studies on fermentati... 1876 · Germ Theory and Its Applic... · Écrits scientifiques et médica... · Oeuvres de Pasteur · Germ Theory and Its Applic...

Recherches associées

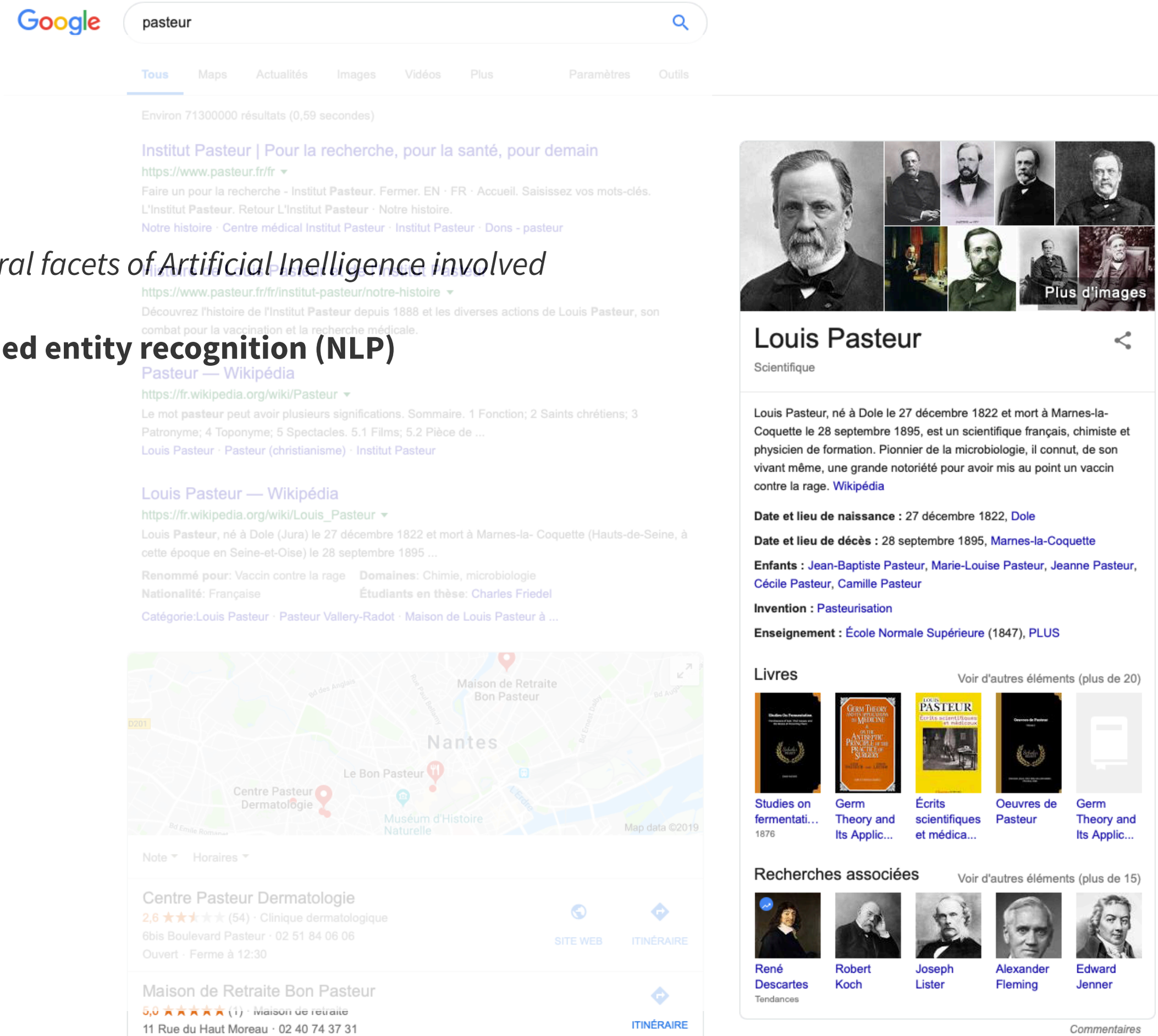
Voir d'autres éléments (plus de 15)

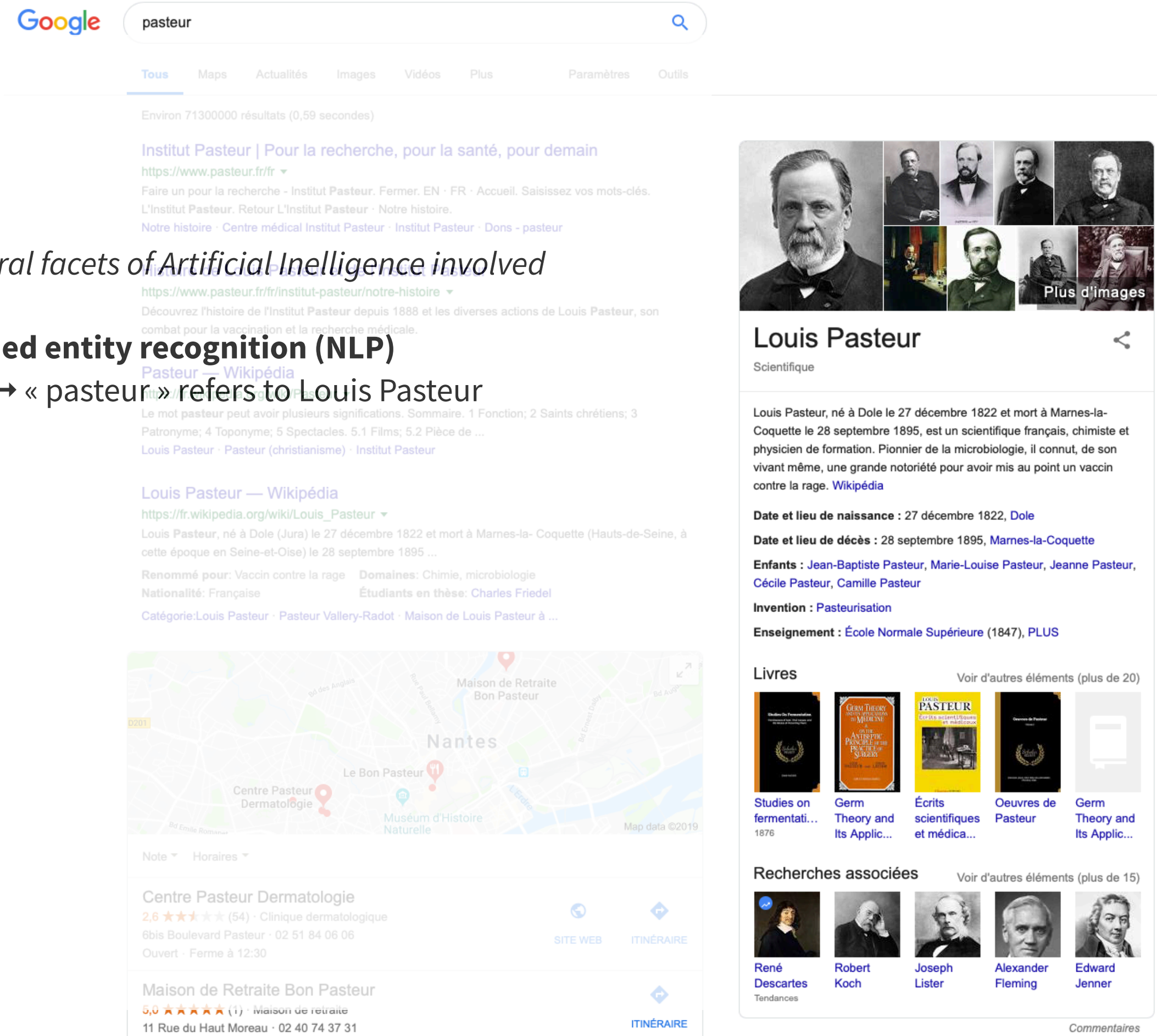
René Descartes (Tendances) · Robert Koch · Joseph Lister · Alexander Fleming · Edward Jenner

Commentaires

Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

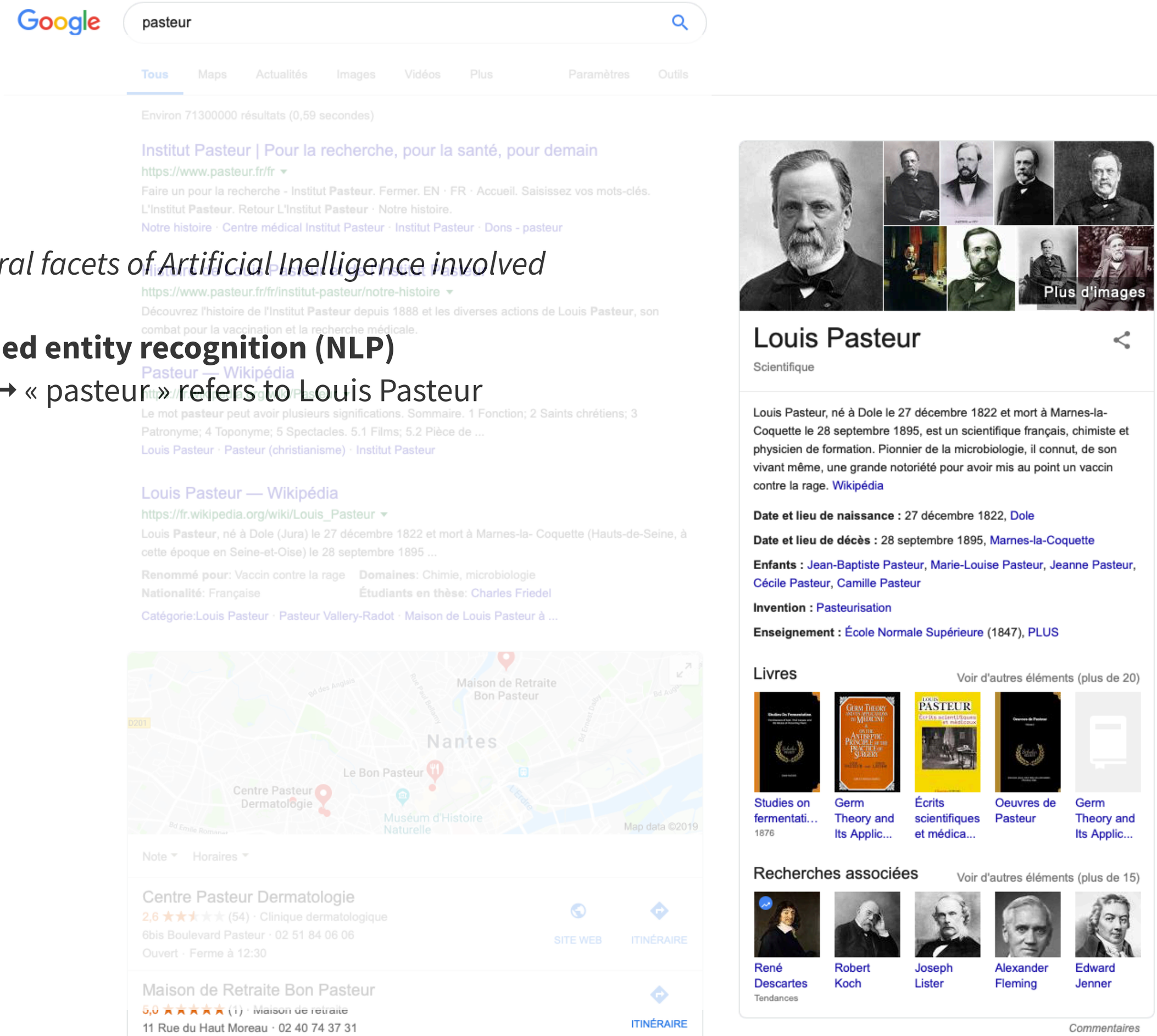




Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

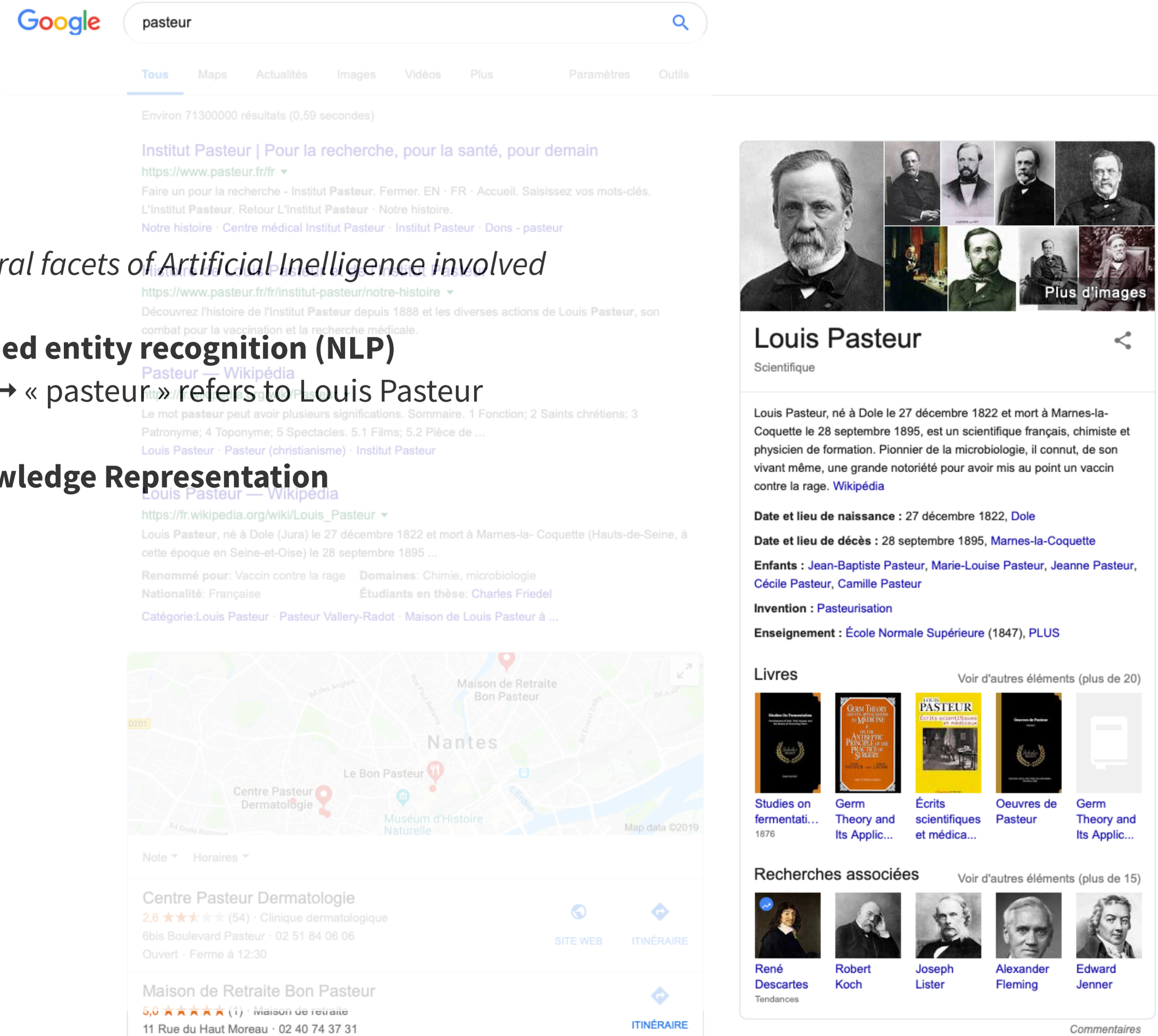
→ « pasteur » refers to Louis Pasteur



Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

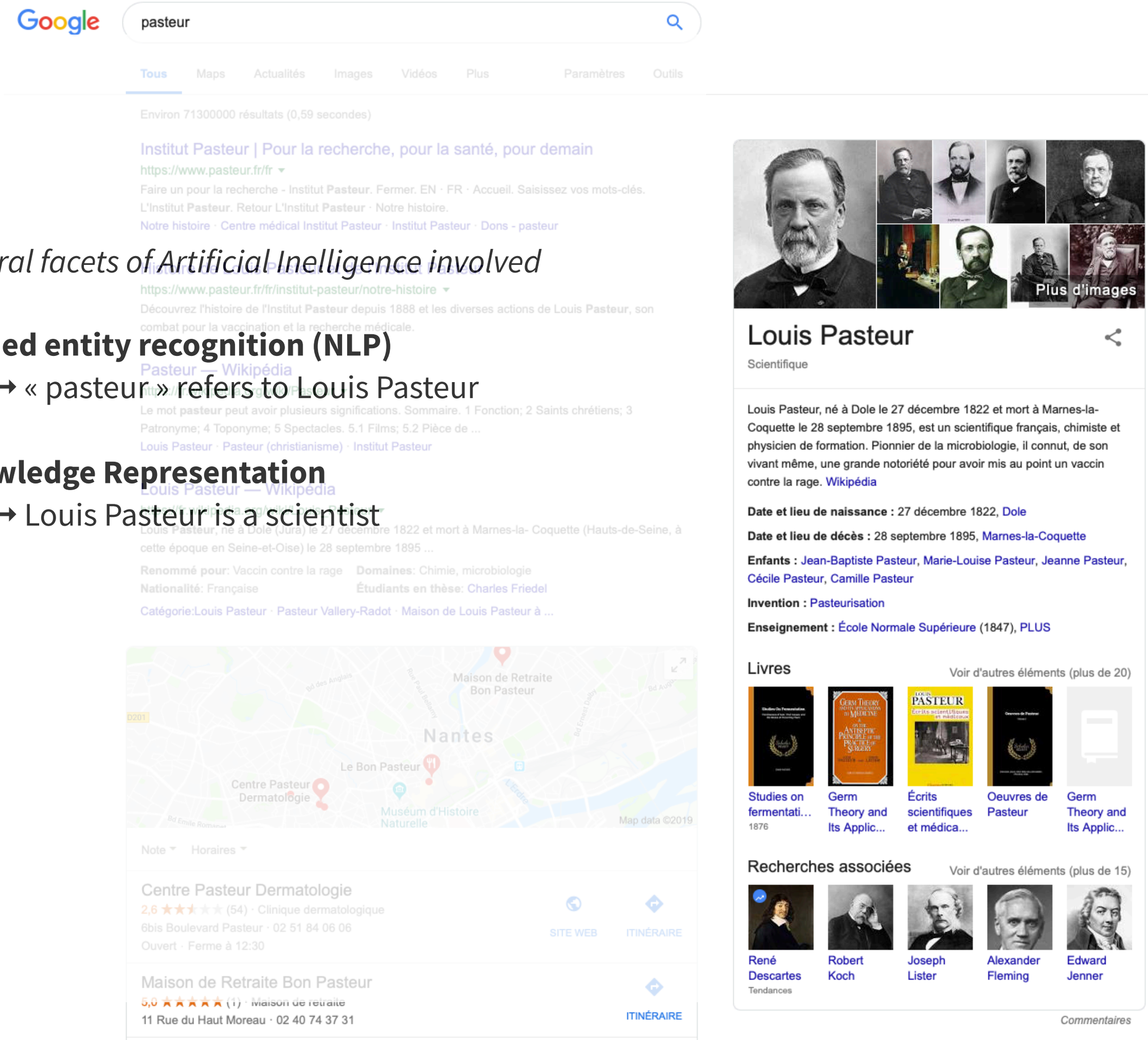


Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation



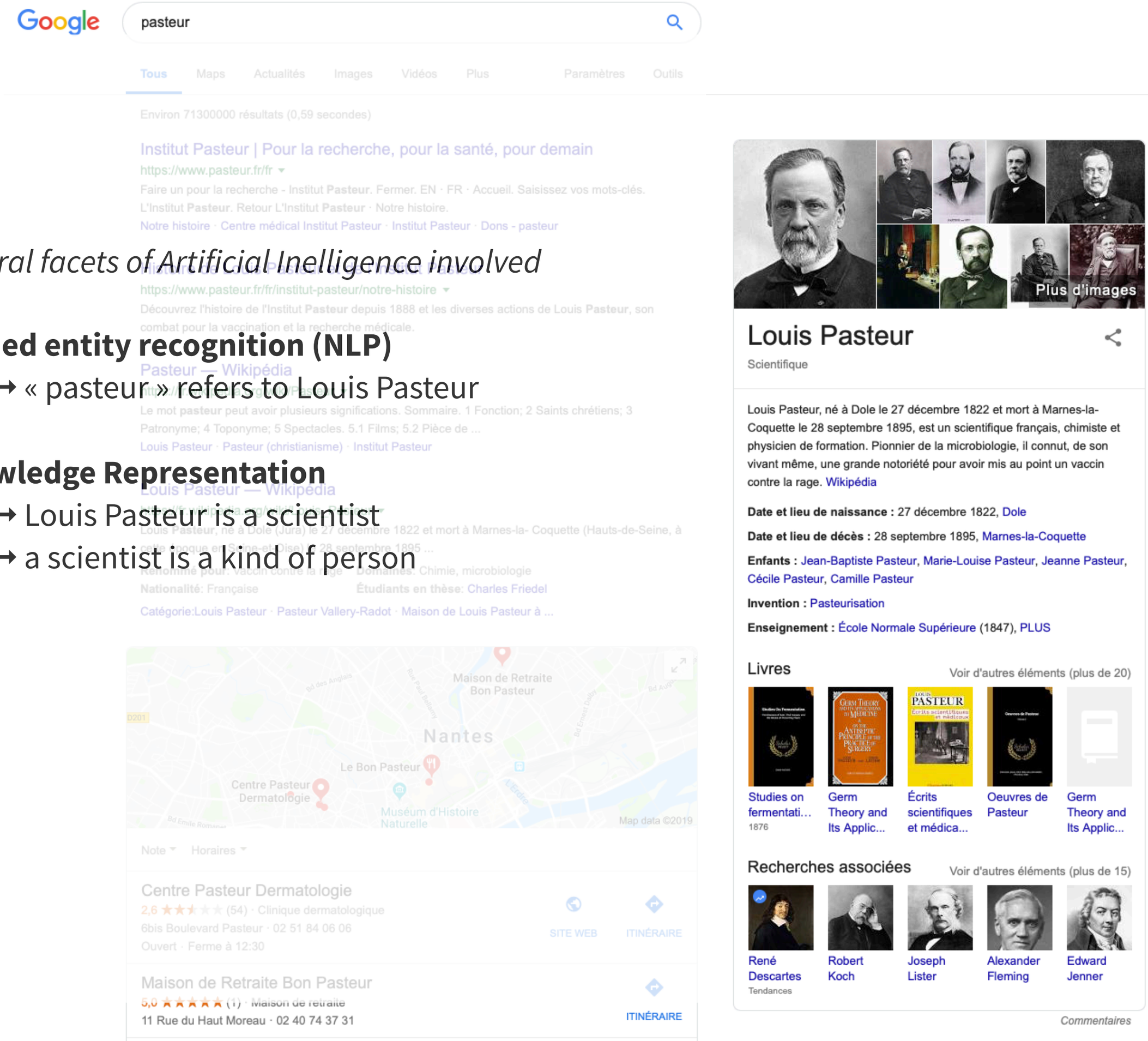
Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist



Several facets of Artificial Inelligence involved

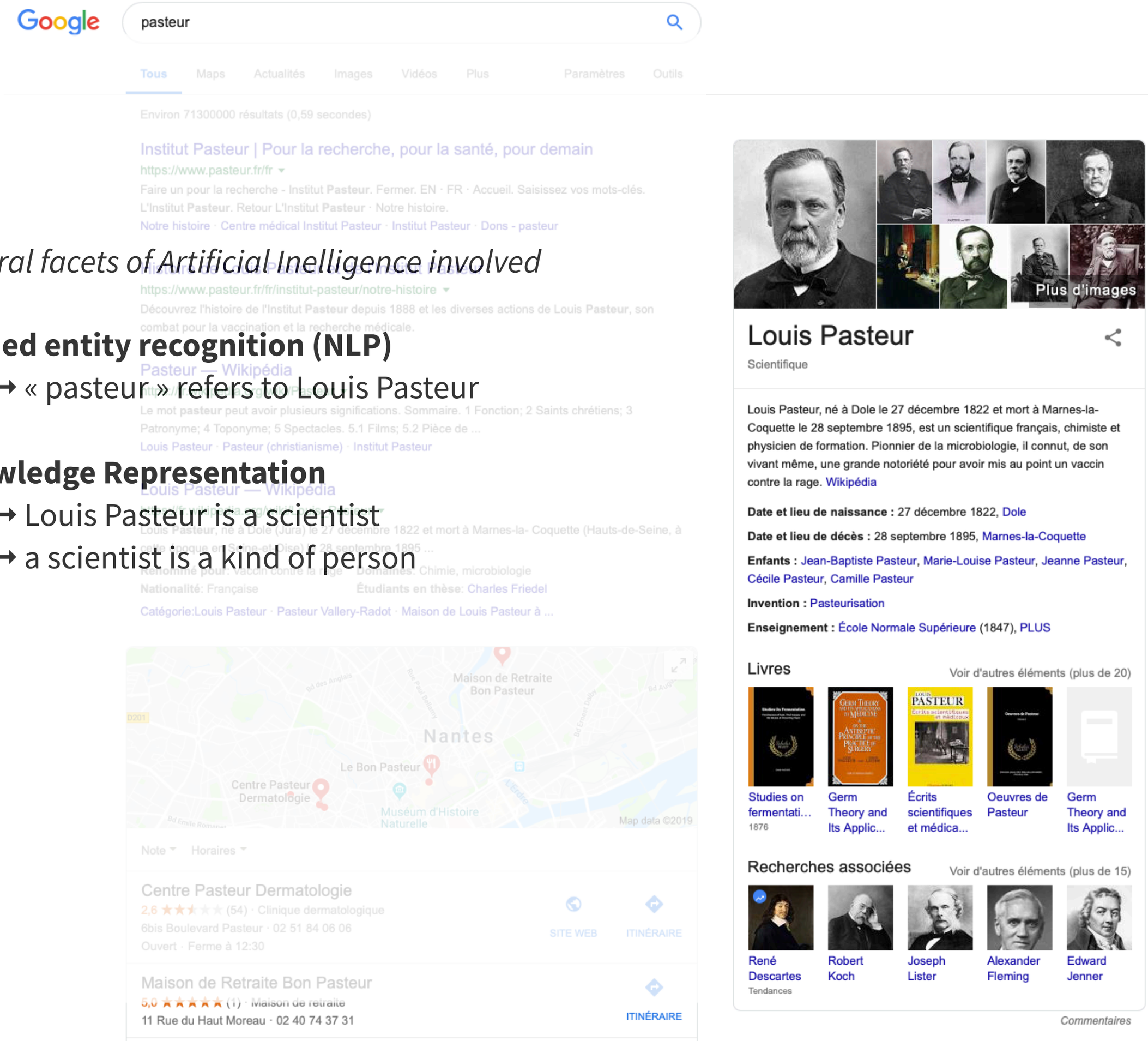
Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist

→ a scientist is a kind of person



Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist

→ a scientist is a kind of person

Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist

→ a scientist is a kind of person

Reasoning

Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist

→ a scientist is a kind of person

Reasoning

→ « if an entity is a scientists then its also a person »

Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist

→ a scientist is a kind of person

Reasoning

→ « if an entity is a scientists then its also a person »

→ « If a web resource is a person then display his/her portrait »

Louis Pasteur

Scientifique

Louis Pasteur, né à Dole le 27 décembre 1822 et mort à Marnes-la-Coquette le 28 septembre 1895, est un scientifique français, chimiste et physicien de formation. Pionnier de la microbiologie, il connut, de son vivant même, une grande notoriété pour avoir mis au point un vaccin contre la rage. Wikipédia

Date et lieu de naissance : 27 décembre 1822, Dole

Date et lieu de décès : 28 septembre 1895, Marnes-la-Coquette

Enfants : Jean-Baptiste Pasteur, Marie-Louise Pasteur, Jeanne Pasteur, Cécile Pasteur, Camille Pasteur

Invention : Pasteurisation

Enseignement : École Normale Supérieure (1847), PLUS

Livres

Voir d'autres éléments (plus de 20)

Studies on fermentati... 1876 | Germ Theory and Its Applic... | Écrits scientifiques et médica... | Oeuvres de Pasteur | Germ Theory and Its Applic...

Recherches associées

Voir d'autres éléments (plus de 15)

René Descartes (Tendances) | Robert Koch | Joseph Lister | Alexander Fleming | Edward Jenner

Commentaires

Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist

→ a scientist is a kind of person

Reasoning

→ « if an entity is a scientists then its also a person »

→ « If a web resource is a person then display his/her portrait »

Louis Pasteur

Scientifique

Louis Pasteur, né à Dole le 27 décembre 1822 et mort à Marnes-la-Coquette le 28 septembre 1895, est un scientifique français, chimiste et physicien de formation. Pionnier de la microbiologie, il connut, de son vivant même, une grande notoriété pour avoir mis au point un vaccin contre la rage. Wikipédia

Date et lieu de naissance : 27 décembre 1822, Dole

Date et lieu de décès : 28 septembre 1895, Marnes-la-Coquette

Enfants : Jean-Baptiste Pasteur, Marie-Louise Pasteur, Jeanne Pasteur, Cécile Pasteur, Camille Pasteur

Invention : Pasteurisation

Enseignement : École Normale Supérieure (1847), PLUS

Livres

Voir d'autres éléments (plus de 20)

Studies on fermentati... 1876 | Germ Theory and Its Applic... | Écrits scientifiques et médica... | Oeuvres de Pasteur | Germ Theory and Its Applic...

Recherches associées

Voir d'autres éléments (plus de 15)

René Descartes (Tendances) | Robert Koch | Joseph Lister | Alexander Fleming | Edward Jenner

Commentaires

Several facets of Artificial Inelligence involved

Named entity recognition (NLP)

→ « pasteur » refers to Louis Pasteur

Knowledge Representation

→ Louis Pasteur is a scientist

→ a scientist is a kind of person

Reasoning

→ « if an entity is a scientists then its also a person »

→ « If a web resource is a person then display his/her portrait »

… and many other such as Computer Vision, Machine Learning …

Knowledge representation and reasoning : still needed for question answering systems !

Travailler chez Apple

À propos d'Apple | Équipes | Apple Retail | Étudiants | Profil | Se connecter | Rechercher

Siri - Software Engineer - Knowledge Graph

Santa Clara Valley (Cupertino), California, United States
Machine Learning and AI

Envoyer un CV

‹ Retour aux résultats de recherche

Summary

Posted: **May 29, 2019**
Weekly Hours: **40**
Role Number: **200039719**

The Knowledge Graph team is looking for outstanding engineers to build the next- generation of knowledge graph and data infrastructure at Apple to power features including Siri and Spotlight. If you are interested in building a world-class

Key Qualifications

- Extensive systems programming experience in either Python or Java. Solid system development skills in UNIX-type OS (e.g. Linux, Mac OS)
- Experience working with large data sets and pipelines, ideally using the Apache
- software stack (e.g. Spark, HBase)
- Excellent problem-solving and analytic skills
- You are self-motivated and able to quickly learn new domains
- You have good attention to detail
- Broad knowledge of computer science and systems
- Excellent communication and collaborative skills; Able to work as part of a small, focused team and give your best effort

Description

The Siri Knowledge Graph team is building groundbreaking technology in the areas of question answering, knowledge base construction and machine learning. We aim to be a "know-it-all" question answering system, capable of answering questions from hundreds of millions of users about nearly anything. The question answering system is backed by a knowledge graph that was automatically constructed from a vast number of data sources including natural language text, HTML tables, and many others. You will have exciting opportunities to working on rapidly building a more complete and accurate knowledge graph with impact across all of Apple. This knowledge graph also enables many other features across Apple besides the question answering feature at Siri.
The problems we pursue include

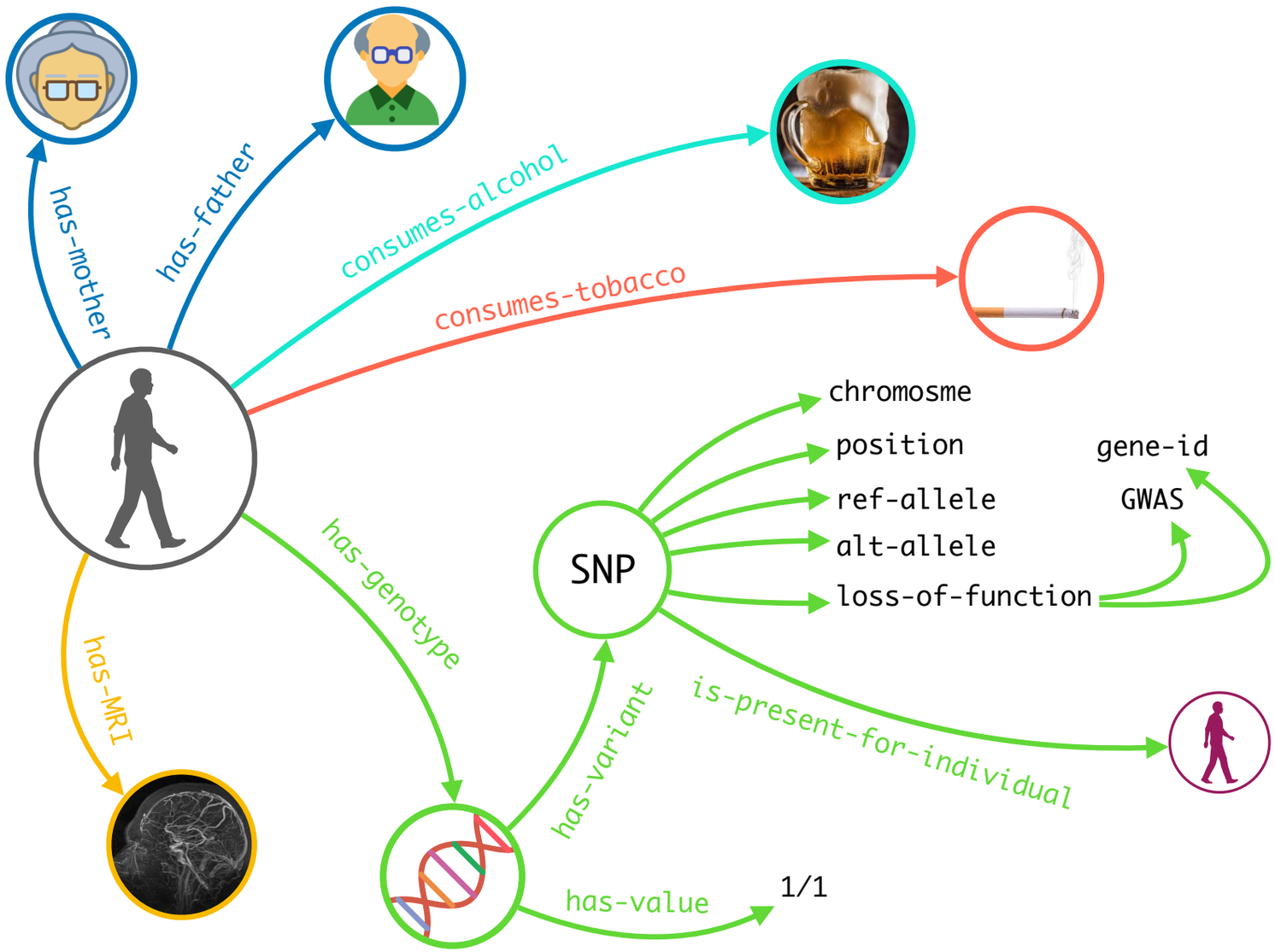
* Information Extraction from natural language text and semi-structured data such as HTML tables
* Data Integration (e.g., Entity Resolution and Knowledge Fusion)
* Knowledge graph reasoning and inference
* Named Entity Linking

What is a Knowledge Graph ?

« ***a collection of interlinked descriptions of things***
(real-word objects, abstract concepts, events, etc.) »

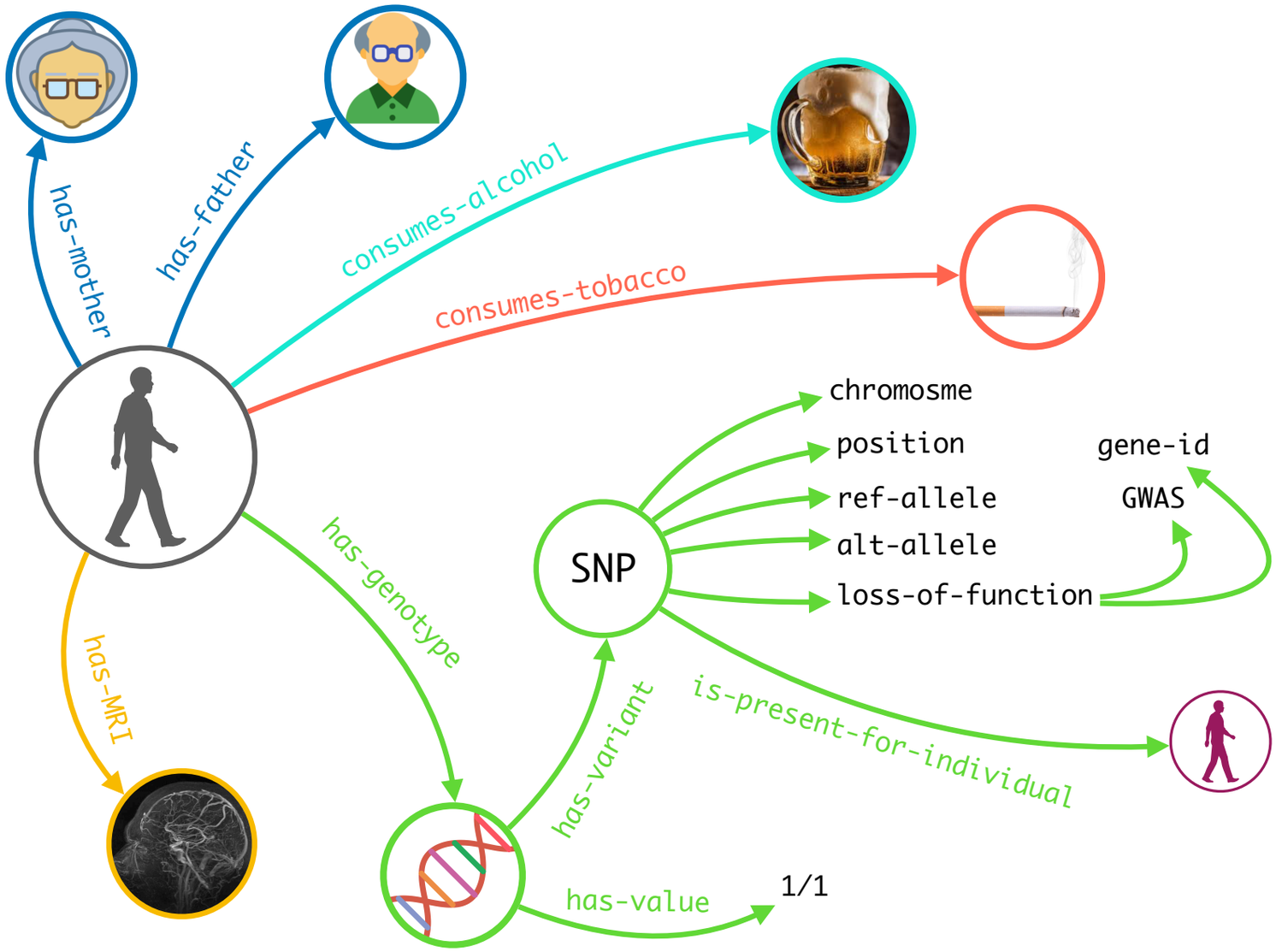
What is a Knowledge Graph ?

« ***a collection of interlinked descriptions of things***
(real-word objects, abstract concepts, events, etc.) »



What is a Knowledge Graph ?

« ***a collection of interlinked descriptions of things***
(real-word objects, abstract concepts, events, etc.) »

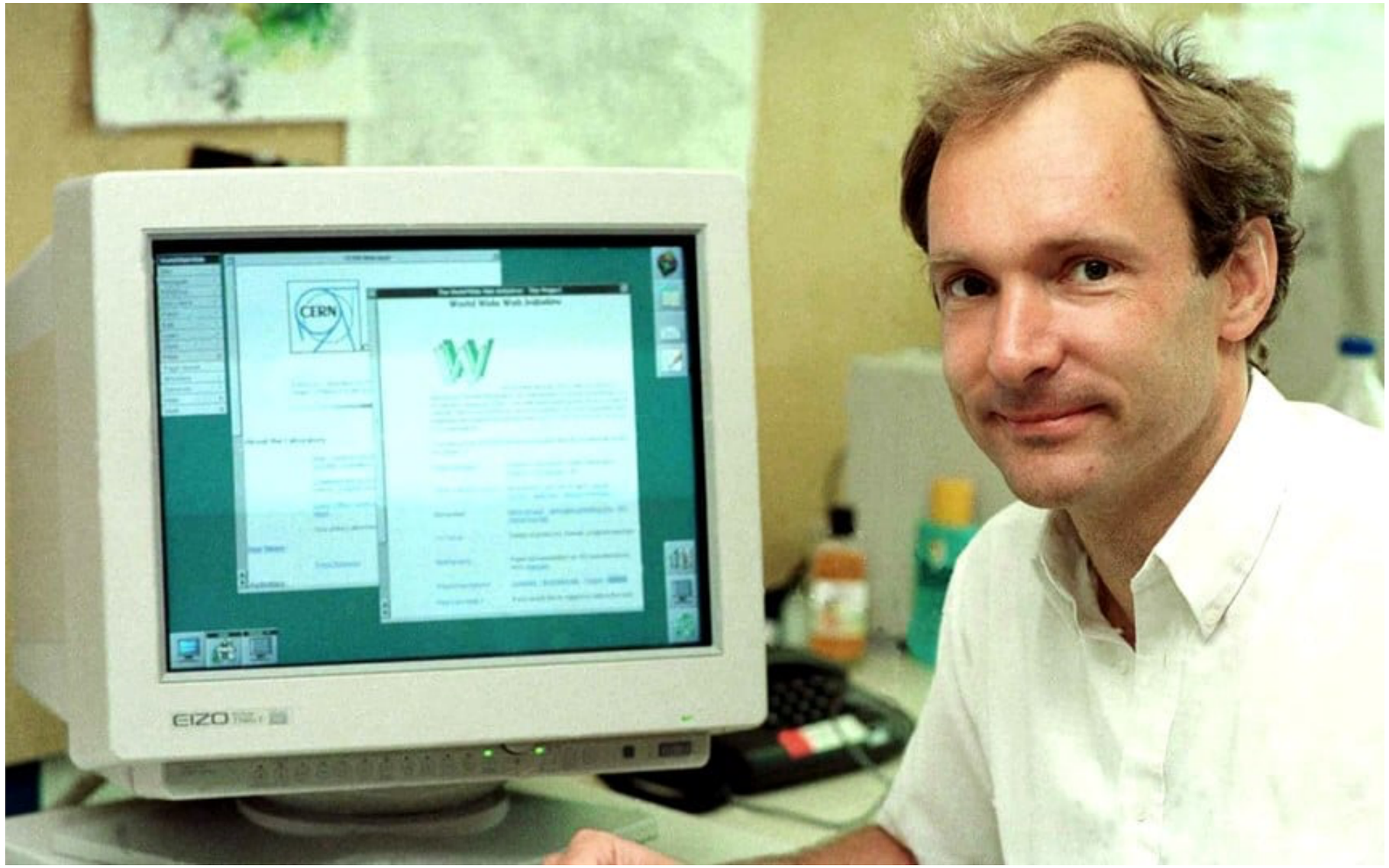


- a **database**: information storage / extraction
- a **graph**: network analysis
- a **knowledge base** with formal semantics :
  + logical facts
  + logical inferences

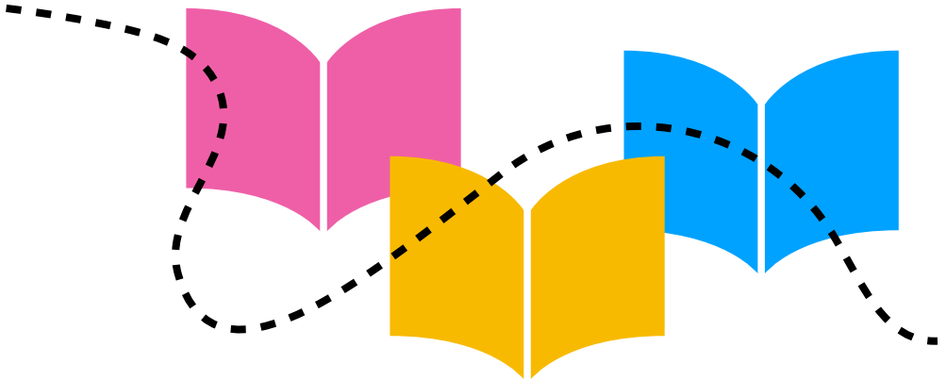
From linked documents (**Web**) …

… to linked data (**Semantic Web**)

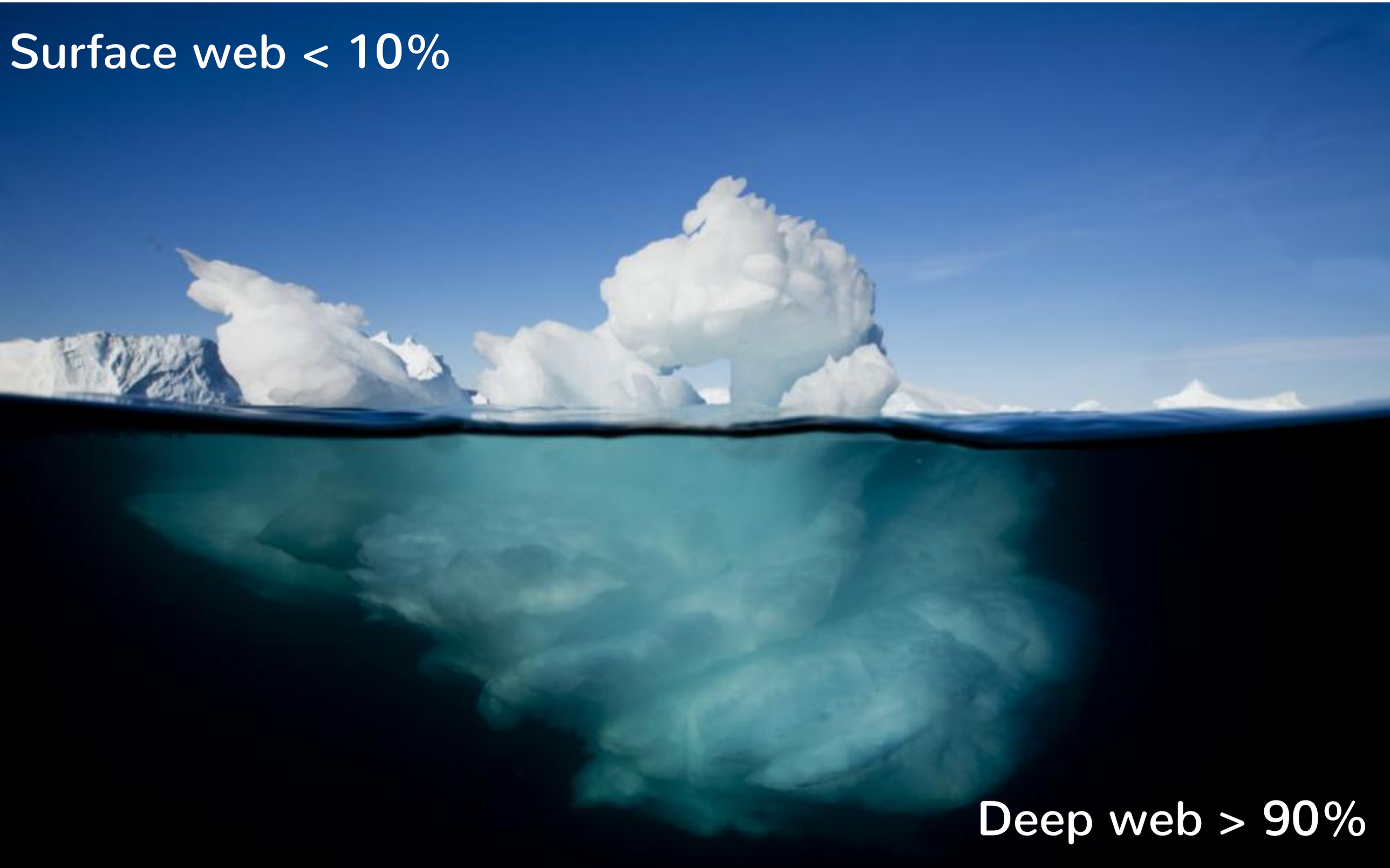
90’s: web of documents



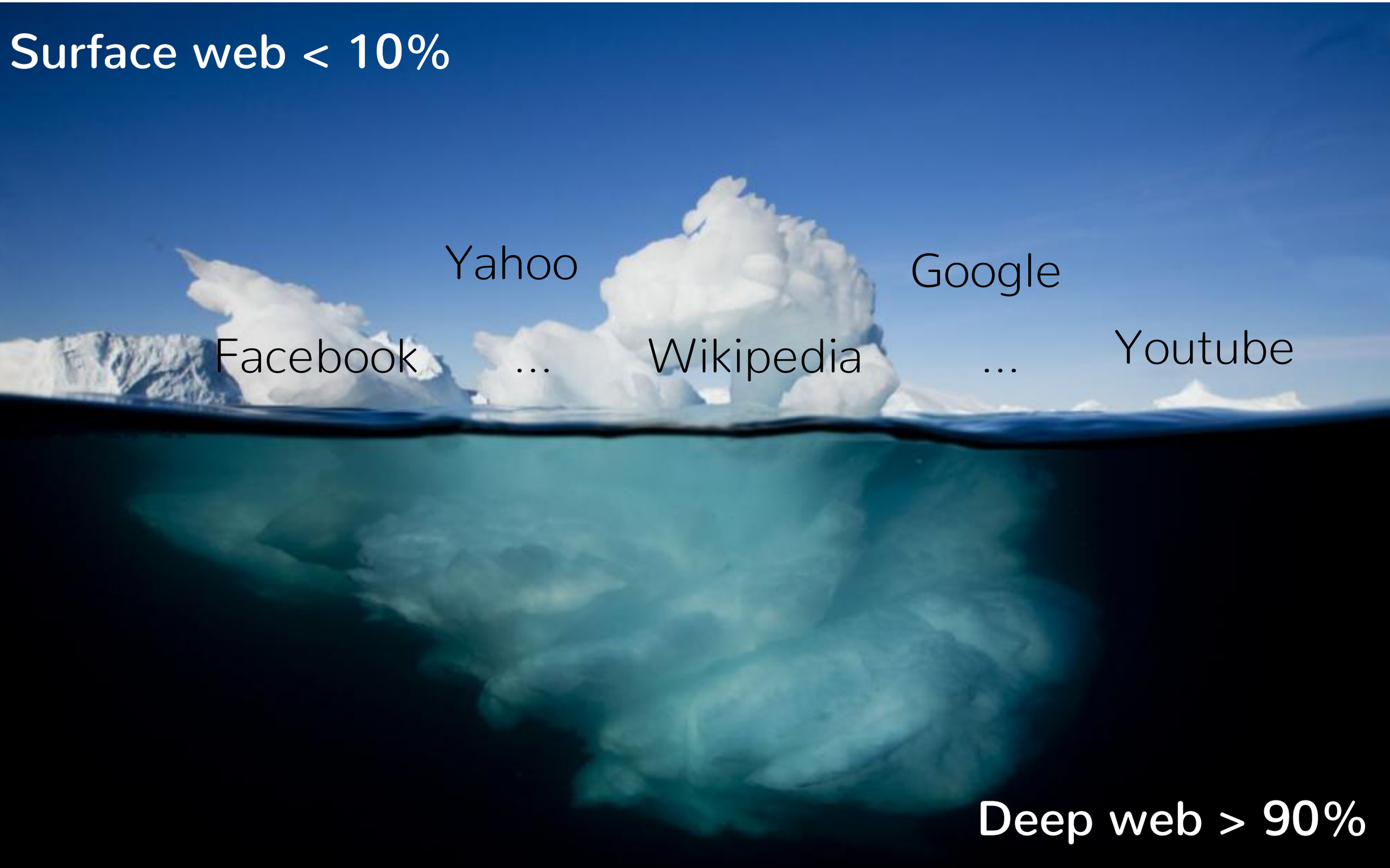
A distributed system of Hypertext documents based on URL, HTTP, and HTML standards (World Wide Web consortium, W3C).



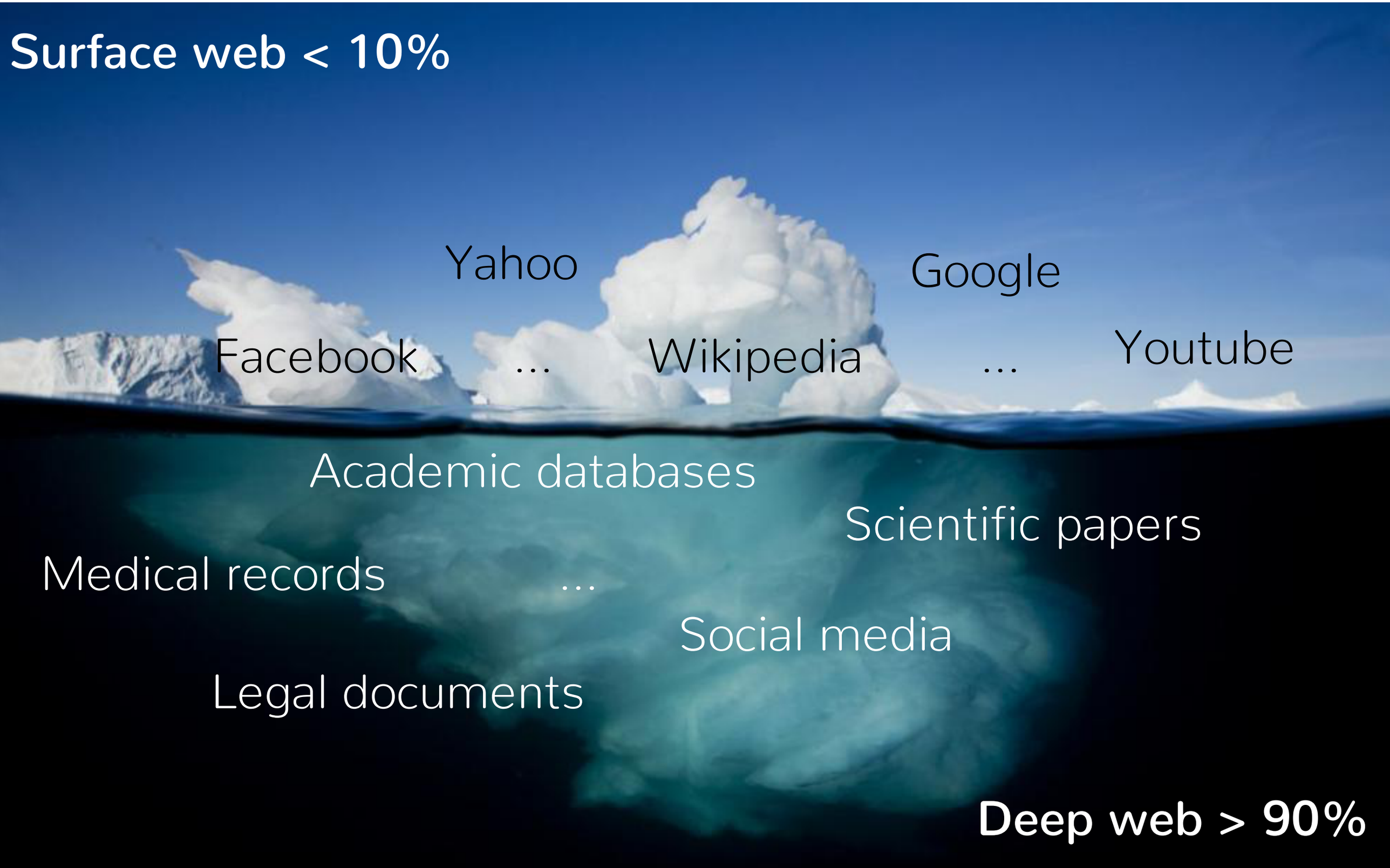
Today: web of data



Today: web of data



Today: web of data



Diverse data, not easily searchable !

Diverse data, not easily searchable !

```
@HWI-ST534_129:2:24:20503:16510:CGATGT
CTGAGAGCCGGGGAAGCCGGCGGAGCCGGGGGACTGGCGAGCCGGAACAT
+
HHHHHHHHHHEFDDGDDFBFGG>7D4<9;<&?:;<DC>CCDD@?=?A###
@HWI-ST534_129:2:42:2118:9580:CGATGT
GGCGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGA
+
GEECGGGBGIDF6FFFFEF=IDEFBEE8E8E?EEB@6=9B#########
@HWI-ST534_129:2:2:12654:80229:CGATGT
CGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAG
+
GGEGFCDCBBAEEEEGGFGFG;EGEEGFFBBEDEBDFGFCFF;DF2D<DD
@HWI-ST534_129:2:48:12356:179714:CGATGT
GAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAA
+
E=GHFHEGHHBCGDDBEEBBCBDDDE@EGBD=ABDCB?EC;@@8@EEB;E
@HWI-ST534_129:2:44:8225:39540:CGATGT
GGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAACTATG
+
HHHEHHHHHHHHHGHHHHFHHHHHHHHDFDHHBHFDDFEFEFF>G<CCCE
```

Diverse data, not easily searchable !

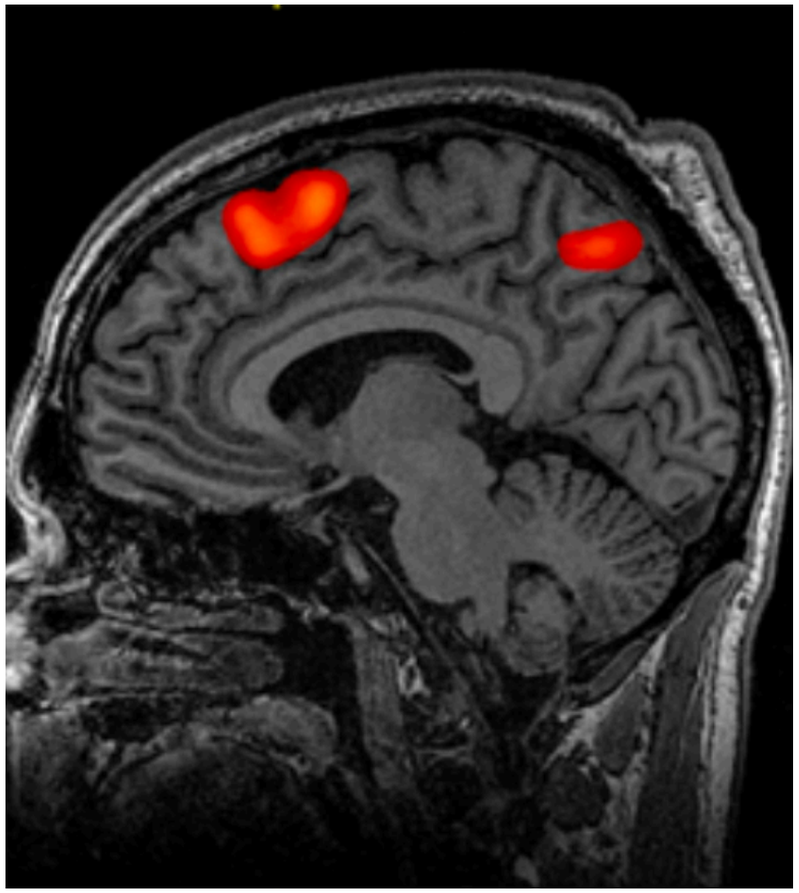
```
@HWI-ST534_129:2:24:20503:16510:CGATGT
CTGAGAGCCGGGGAAGCCGGCGGAGCCGGGGGACTGGCGAGCCGGAACAT
+
HHHHHHHHHHEFDDGDDFBFGG>7D4<9;<&?:;<DC>CCDD@?=?A###
@HWI-ST534_129:2:42:2118:9580:CGATGT
GGCGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGA
+
GEECGGGBGIDF6FFFFEF=IDEFBEE8E8E?EEB@6=9B#########
@HWI-ST534_129:2:2:12654:80229:CGATGT
CGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAG
+
GGEGFCDCBBAEEEEGGFGFG;EGEEGFFBBEDEBDFGFCFF;DF2D<DD
@HWI-ST534_129:2:48:12356:179714:CGATGT
GAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAA
+
E=GHFHEGHHBCGDDBEEBBCBDDDE@EGBD=ABDCB?EC;@@8@EEB;E
@HWI-ST534_129:2:44:8225:39540:CGATGT
GGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAACTATG
+
HHHEHHHHHHHHHGHHHHFHHHHHHHHDFDHHBHFDDFEFEFF>G<CCCE
```

Genetic sequences

- 1st line = label
- 2nd line = raw sequence (A,T,C,G)
- 4th line = quality score for each base

Diverse data, not easily searchable !

```
@HWI-ST534_129:2:24:20503:16510:CGATGT
CTGAGAGCCGGGGAAGCCGGCGGAGCCGGGGGACTGGCGAGCCGGAACAT
+
HHHHHHHHHHEFDDGDDFBFGG>7D4<9;<&?:;<DC>CCDD@?=?A###
@HWI-ST534_129:2:42:2118:9580:CGATGT
GGCGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGA
+
GEECGGGBGIDF6FFFFEF=IDEFBEE8E8E?EEB@6=9B#########
@HWI-ST534_129:2:2:12654:80229:CGATGT
CGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAG
+
GGEGFCDCBBAEEEEGGFGFG;EGEEGFFBBEDEBDFGFCFF;DF2D<DD
@HWI-ST534_129:2:48:12356:179714:CGATGT
GAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAA
+
E=GHFHEGHHBCGDDBEEBBCBDDDE@EGBD=ABDCB?EC;@@8@EEB;E
@HWI-ST534_129:2:44:8225:39540:CGATGT
GGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAACTATG
+
HHHEHHHHHHHHHGHHHHFHHHHHHHHDFDHHBHFDDFEFEFF>G<CCCE
```

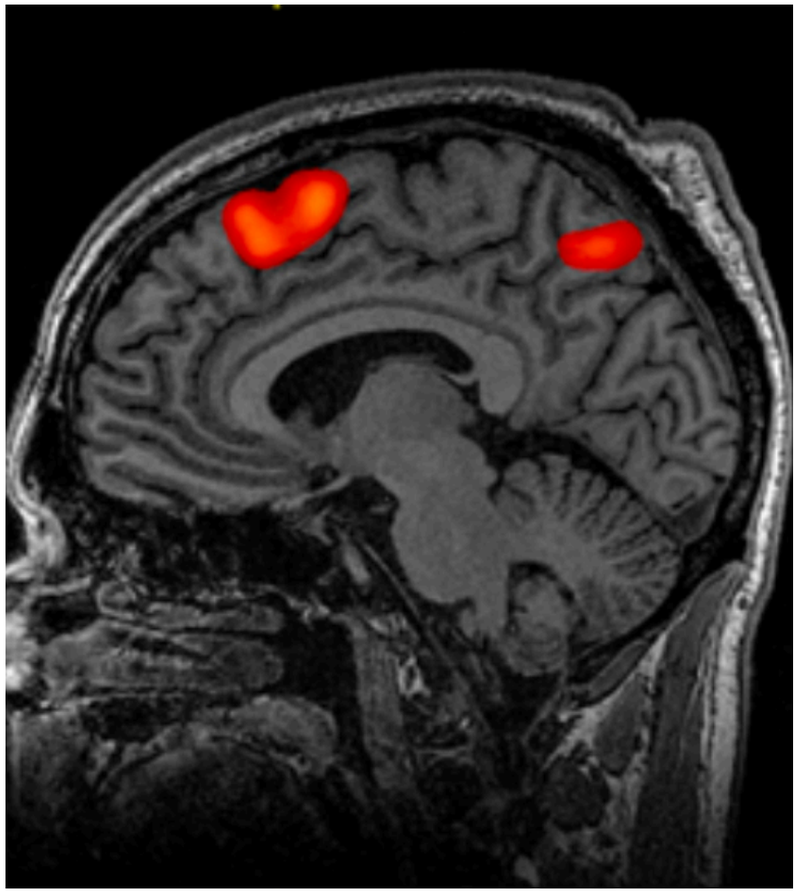


Genetic sequences

- 1st line = label
- 2nd line = raw sequence (A,T,C,G)
- 4th line = quality score for each base

Diverse data, not easily searchable !

```
@HWI-ST534_129:2:24:20503:16510:CGATGT
CTGAGAGCCGGGGAAGCCGGCGGAGCCGGGGGACTGGCGAGCCGGAACAT
+
HHHHHHHHHHEFDDGDDFBFGG>7D4<9;<&?:;<DC>CCDD@?=?A###
@HWI-ST534_129:2:42:2118:9580:CGATGT
GGCGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGA
+
GEECGGGBGIDF6FFFFEF=IDEFBEE8E8E?EEB@6=9B#########
@HWI-ST534_129:2:2:12654:80229:CGATGT
CGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAG
+
GGEGFCDCBBAEEEEGGFGFG;EGEEGFFBBEDEBDFGFCFF;DF2D<DD
@HWI-ST534_129:2:48:12356:179714:CGATGT
GAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAA
+
E=GHFHEGHHBCGDDBEEBBCBDDDE@EGBD=ABDCB?EC;@@8@EEB;E
@HWI-ST534_129:2:44:8225:39540:CGATGT
GGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAACTATG
+
HHHEHHHHHHHHHGHHHHFHHHHHHHHDFDHHBHFDDFEFEFF>G<CCCE
```



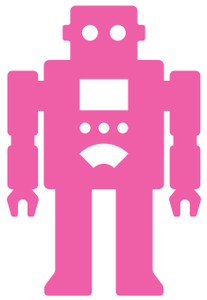
Genetic sequences

- 1st line = label
- 2nd line = raw sequence (A,T,C,G)
- 4th line = quality score for each base

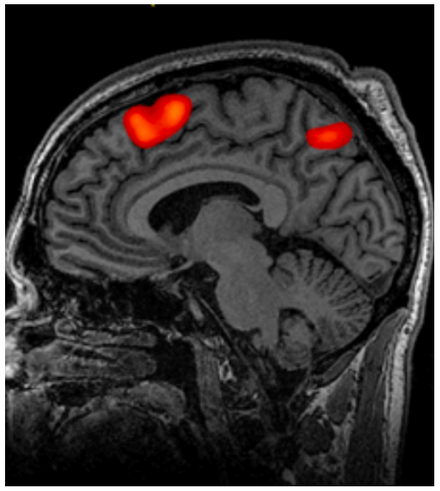
Medical image

- MRI as imaging modality
- brain as imaging target
- labeled data (intensity/volume ?)

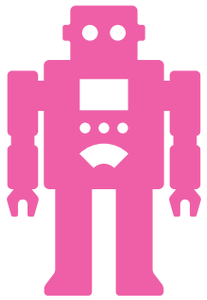
Semantic web

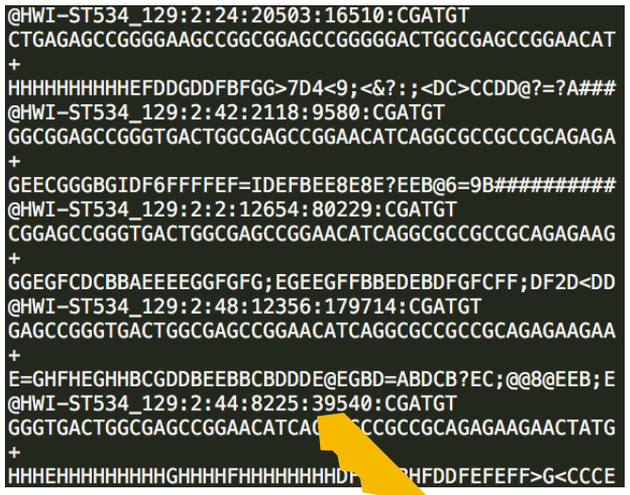
make  and better interpret (reason) and exchange ≠ data

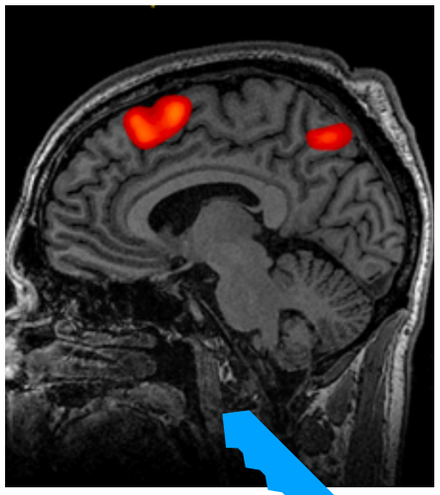
```
@HWI-ST534_129:2:24:20503:16510:CGATGT
CTGAGAGCCGGGGAAGCCGGCGGAGCCGGGGGACTGGCGAGCCGGAACAT
+
HHHHHHHHHHHEFDDGDDFBFGG>7D4<9;<&?:;<DC>CCDD@?=?A###
@HWI-ST534_129:2:42:2118:9580:CGATGT
GGCGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGA
+
GEECGGGBGIDF6FFFFEF=IDEFBEE8E8E?EEB@6=9B#########
@HWI-ST534_129:2:2:12654:80229:CGATGT
CGGAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAG
+
GGEGFCDCBBAEEEEGGFGFG;EGEEGFFBBEDEBDFGFCFF;DF2D<DD
@HWI-ST534_129:2:48:12356:179714:CGATGT
GAGCCGGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAA
+
E=GHFHEGHHBCGDDBEEBBCBDDDE@EGBD=ABDCB?EC;@@8@EEB;E
@HWI-ST534_129:2:44:8225:39540:CGATGT
GGGTGACTGGCGAGCCGGAACATCAGGCGCCGCCGCAGAGAAGAACTATG
+
HHHEHHHHHHHHHGHHHHFHHHHHHHHDFDHHBHFDDFEFEFF>G<CCCE
```



Semantic web

make  and better interpret (reason) and exchange ≠ data





→ build « intelligent » applications

(one of the many facets of AI)

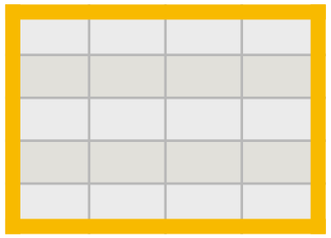
W3C recommendations

W3C recommendations

http://dbpedia.org/resource/RAC1

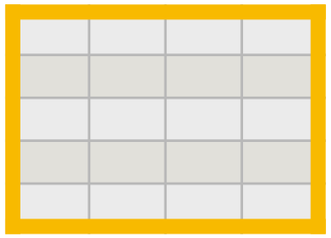
W3C recommendations

http://dbpedia.org/resource/RAC1



W3C recommendations

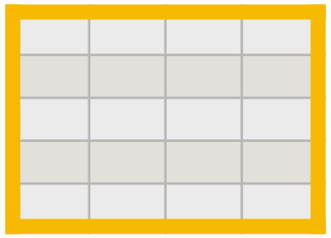
http://dbpedia.org/resource/RAC1



.csv

W3C recommendations

http://dbpedia.org/resource/RAC1



.csv

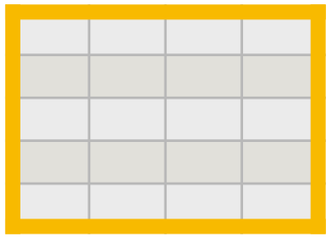
abstract

Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions. (en)

RAC1

W3C recommendations

http://dbpedia.org/resource/RAC1



.csv

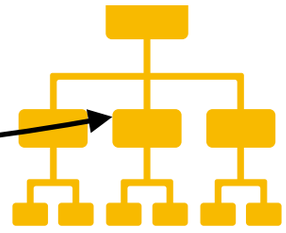
abstract

Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions. (en)

RAC1

is a

Gene Ontology



Wikipedia … for humans

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

Languages
日本語
Русский
Українська
Edit links

Article Talk Read Edit View history Search Wikipedia

RAC1

From Wikipedia, the free encyclopedia

"Rac1" redirects here. For the first game in the Ratchet & Clank series, see Ratchet & Clank.

Rac1, also known as **Ras-related C3 botulinum toxin substrate 1**, is a protein found in human cells. It is encoded by the *RAC1* gene.[5][6] This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions.[7]

Contents [hide]

1 Function
2 Role in cancer
3 Role in glucose transport
4 Clinical significance
5 Interactions
6 References
7 Further reading
8 External links

Function [edit]

Rac1 is a small (~21 kDa) signaling G protein (more specifically a GTPase), and is a member of the Rac subfamily of the family Rho family of GTPases. Members of this superfamily appear to regulate a diverse array of cellular events, including the control of GLUT4[8][9] translocation to glucose uptake, cell growth, cytoskeletal reorganization, antimicrobial cytotoxicity,[10] and the activation of protein kinases.[11]

Rac1 is a pleiotropic regulator of many cellular processes, including the cell cycle, cell-cell adhesion, motility (through the actin network), and of epithelial differentiation (proposed to be necessary for maintaining epidermal stem cells).

Role in cancer [edit]

Along with other subfamily of Rac and Rho proteins, they exert an important regulatory role specifically in cell motility and cell growth. Rac1 has ubiquitous tissue expression, and drives cell motility by formation of lamellipodia.[12] In order for cancer cells to grow and invade local and distant tissues, deregulation of cell motility is one of the hallmark events in cancer cell invasion and metastasis.[13] Overexpression of a constitutively active Rac1 V12 in mice caused a tumor that's phenotypically indistinguishable form human Kaposi's sarcoma.[14] Activating or gain-of-function mutations of Rac1 are shown to play active roles in promoting mesenchymal-type of cell movement assisted by NEDD9 and DOCK3 protein complex.[15] Such abnormal cell motility may result in epithelial mesenchymal transition (EMT) – a driving mechanism for tumor metastasis as well as drug-resistant tumor relapse.[16][17]

Role in glucose transport [edit]

Rac1 is expressed in significant amounts in insulin sensitive tissues, such as adipose tissue and skeletal muscle. Here Rac1 regulated the translocation of glucose transporting GLUT4 vesicles from intracellular compartments to the plasma membrane.[9][18][19] In response to insulin, this allows for blood glucose to enter the cell to lower blood glucose. In conditions of obesity and type 2 diabetes, Rac1 signaling in skeletal muscle is dysfunctional, suggesting that Rac1 contributes to the progression of the disease. Rac1 protein is also necessary for glucose uptake in skeletal muscle activated by exercise[8][20] and muscle stretching[21]

Clinical significance [edit]

Activating mutations in Rac1 have been recently discovered in large-scale genomic studies involving melanoma [22][23][24] and non-small cell lung cancer.[25] As a result, Rac1 is considered a therapeutic target for many of these diseases.[26]

A few recent studies have also exploited targeted therapy to suppress tumor growth by pharmacological inhibition of Rac1 activity in metastatic melanoma and liver cancer as well as in human breast cancer.[27][28][29] For example, Rac1-dependent pathway inhibition resulted in the reversal of tumor cell phenotypes, suggesting Rac1 as a predictive marker and therapeutic target for trastuzumab-resistant breast cancer.[28] However, given Rac1's role in glucose transport, drugs that inhibits Rac1 could potentially be harmful to glucose homeostasis

RAC1

Available structures	
PDB	Ortholog search: PDBe RCSB
	List of PDB id codes [show]

Identifiers	
Aliases	RAC1, MIG5, Rac-1, TC-25, p21-Rac1, ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1), Rac family small GTPase 1
External IDs	OMIM: 602048 MGI: 97845 HomoloGene: 69035 GeneCards: RAC1

Gene location (Human) [hide]

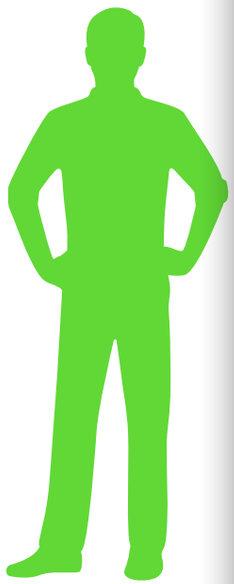
Chr. Chromosome 7 (human)[1]

Band 7p22.1 Start 6,374,523 bp[1] End 6,403,977 bp[1]

Gene location (Mouse) [show]

RNA expression pattern [hide]

Wikipedia … for humans



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

Languages
日本語
Русский
Українська
Edit links

Article Talk Read Edit View history Search Wikipedia

RAC1

From Wikipedia, the free encyclopedia

"Rac1" redirects here. For the first game in the Ratchet & Clank series, see Ratchet & Clank.

Rac1, also known as **Ras-related C3 botulinum toxin substrate 1**, is a protein found in human cells. It is encoded by the *RAC1* gene.[5][6] This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions.[7]

Contents [hide]

1 Function
2 Role in cancer
3 Role in glucose transport
4 Clinical significance
5 Interactions
6 References
7 Further reading
8 External links

Function [edit]

Rac1 is a small (~21 kDa) signaling G protein (more specifically a GTPase), and is a member of the Rac subfamily of the family Rho family of GTPases. Members of this superfamily appear to regulate a diverse array of cellular events, including the control of GLUT4[8][9] translocation to glucose uptake, cell growth, cytoskeletal reorganization, antimicrobial cytotoxicity,[10] and the activation of protein kinases.[11]

Rac1 is a pleiotropic regulator of many cellular processes, including the cell cycle, cell-cell adhesion, motility (through the actin network), and of epithelial differentiation (proposed to be necessary for maintaining epidermal stem cells).

Role in cancer [edit]

Along with other subfamily of Rac and Rho proteins, they exert an important regulatory role specifically in cell motility and cell growth. Rac1 has ubiquitous tissue expression, and drives cell motility by formation of lamellipodia.[12] In order for cancer cells to grow and invade local and distant tissues, deregulation of cell motility is one of the hallmark events in cancer cell invasion and metastasis.[13] Overexpression of a constitutively active Rac1 V12 in mice caused a tumor that's phenotypically indistinguishable form human Kaposi's sarcoma.[14] Activating or gain-of-function mutations of Rac1 are shown to play active roles in promoting mesenchymal-type of cell movement assisted by NEDD9 and DOCK3 protein complex.[15] Such abnormal cell motility may result in epithelial mesenchymal transition (EMT) – a driving mechanism for tumor metastasis as well as drug-resistant tumor relapse.[16][17]

Role in glucose transport [edit]

Rac1 is expressed in significant amounts in insulin sensitive tissues, such as adipose tissue and skeletal muscle. Here Rac1 regulated the translocation of glucose transporting GLUT4 vesicles from intracellular compartments to the plasma membrane.[9][18][19] In response to insulin, this allows for blood glucose to enter the cell to lower blood glucose. In conditions of obesity and type 2 diabetes, Rac1 signaling in skeletal muscle is dysfunctional, suggesting that Rac1 contributes to the progression of the disease. Rac1 protein is also necessary for glucose uptake in skeletal muscle activated by exercise[8][20] and muscle stretching[21]

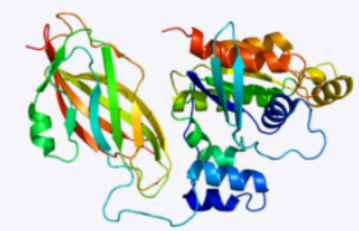
Clinical significance [edit]

Activating mutations in Rac1 have been recently discovered in large-scale genomic studies involving melanoma [22][23][24] and non-small cell lung cancer.[25] As a result, Rac1 is considered a therapeutic target for many of these diseases.[26]

A few recent studies have also exploited targeted therapy to suppress tumor growth by pharmacological inhibition of Rac1 activity in metastatic melanoma and liver cancer as well as in human breast cancer.[27][28][29] For example, Rac1-dependent pathway inhibition resulted in the reversal of tumor cell phenotypes, suggesting Rac1 as a predictive marker and therapeutic target for trastuzumab-resistant breast cancer.[28] However, given Rac1's role in glucose transport, drugs that inhibits Rac1 could potentially be harmful to glucose homeostasis

"Infobox"

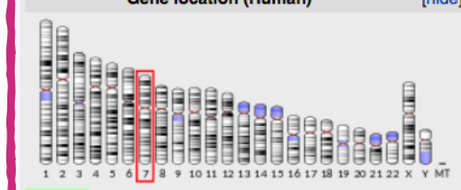
RAC1



Available structures	
PDB	Ortholog search: PDBe RCSB
	List of PDB id codes [show]

Identifiers	
Aliases	RAC1, MIG5, Rac-1, TC-25, p21-Rac1, ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1), Rac family small GTPase 1
External IDs	OMIM: 602048 MGI: 97845 HomoloGene: 69035 GeneCards: RAC1

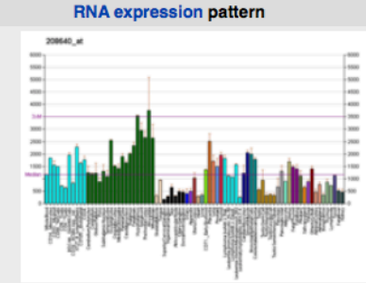
Gene location (Human) [hide]



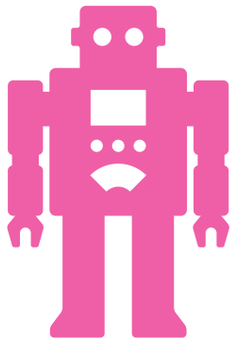
Chr.	Chromosome 7 (human)[1]		
Band	7p22.1	Start	6,374,523 bp[1]
		End	6,403,977 bp[1]

Gene location (Mouse) [show]

RNA expression pattern [hide]



Wikipedia … for machines (DBpedia)



DBpedia Browse using Formats Faceted Browser Sparql Endpoint

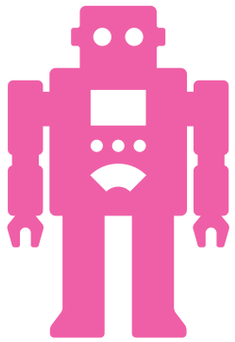
About: RAC1

An Entity of Type : Biomolecule, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions.

Property	Value
dbo:abstract	▪ Rac1 (RAS-related C3 botulinus toxin substrate 1) は、ヒト細胞に存在するタンパク質であり、RAC1遺伝子によりコードされている。RAC1は選択的スプライシングにより異なる機能を持ったいくつかのタンパク質を生成しており、このうちの1つがRac1である。Rac1は、悪性黒色腫や肺非小細胞癌 を含むさまざまな癌の発生において、重要な役割を果たしていると考えられている。そのため、現在これらの疾患に対する治療標的と考えられている。 (ja) ▪ Rac1 (англ. Ras-related C3 botulinum toxin substrate 1) — внутриклеточный белок из суперсемейства ГТФаз, относится к «малым» G-белкам. Находится в двух состояниях: активном ГТФ-связанном и неактивном ГДФ-связанном состоянии. В своей активной форме Rac1 связывается в клетке с целым рядом эффекторных белков и приводит к регулировке многих клеточных процессов, таких как секреция, фагоцитоз апоптозных клеток, поляризация эпителиальных клеток и индуцированное факторами роста образование мембранных складок и выростов (англ. membrane ruffles). (ru) ▪ Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions. (en)
dbo:entrezgene	▪ 5879
dbo:wikiPageExternalLink	▪ http://cmkb.cellmigration.org/report.cgi?report=orth_overview&gene_id=5879 ▪ http://www.cellmigration.org/index.shtml
dbo:wikiPageID	▪ 13562705 (xsd:integer)
dbo:wikiPageRevisionID	▪ 703055305 (xsd:integer)
rdf:type	▪ owl:Thing ▪ wikidata:Q206229 ▪ wikidata:Q8054 ▪ dbo:Biomolecule ▪ dbo:Protein

Wikipedia … for machines (DBpedia)



Give me all gene IDs described with « toxin »

DBpedia | Browse using | Formats | Faceted Browser | Sparql Endpoint

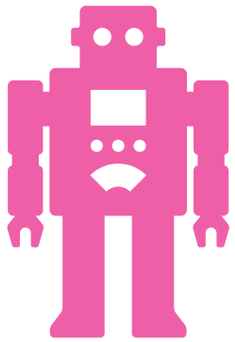
About: RAC1

An Entity of Type : Biomolecule, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions.

Property	Value
dbo:abstract	▪ Rac1 (RAS-related C3 botulinus toxin substrate 1) は、ヒト細胞に存在するタンパク質であり、RAC1遺伝子によりコードされている。RAC1は選択的スプライシングにより異なる機能を持ったいくつかのタンパク質を生成しており、このうちの1つがRac1である。Rac1は、悪性黒色腫や肺非小細胞癌 を含むさまざまな癌の発生において、重要な役割を果たしていると考えられている。そのため、現在これらの疾患に対する治療標的と考えられている。 (ja) ▪ Rac1 (англ. Ras-related C3 botulinum toxin substrate 1) — внутриклеточный белок из суперсемейства ГТФаз, относится к «малым» G-белкам. Находится в двух состояниях: активном ГТФ-связанном и неактивном ГДФ-связанном состоянии. В своей активной форме Rac1 связывается в клетке с целым рядом эффекторных белков и приводит к регулировке многих клеточных процессов, таких как секреция, фагоцитоз апоптозных клеток, поляризация эпителиальных клеток и индуцированное факторами роста образование мембранных складок и выростов (англ. membrane ruffles). (ru) ▪ Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions. (en)
dbo:entrezgene	▪ 5879
dbo:wikiPageExternalLink	▪ http://cmkb.cellmigration.org/report.cgi?report=orth_overview&gene_id=5879 ▪ http://www.cellmigration.org/index.shtml
dbo:wikiPageID	▪ 13562705 (xsd:integer)
dbo:wikiPageRevisionID	▪ 703055305 (xsd:integer)
rdf:type	▪ owl:Thing ▪ wikidata:Q206229 ▪ wikidata:Q8054 ▪ dbo:Biomolecule ▪ dbo:Protein

Wikipedia … for machines (DBpedia)



Give me all gene IDs described with « toxin »

DBpedia | Browse using | Formats | Faceted Browser | Sparql Endpoint

About: RAC1

An Entity of Type : Biomolecule, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions.

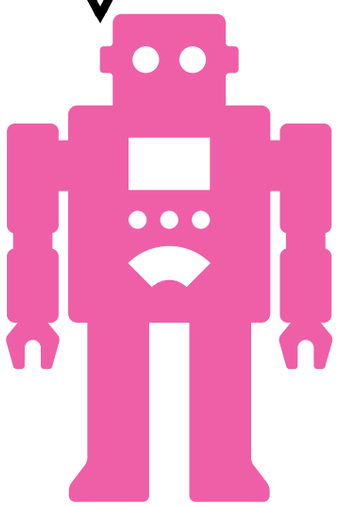
abstract

entrezgene

Property	Value
dbo:abstract	▪ Rac1 (RAS-related C3 botulinus toxin substrate 1) は、ヒト細胞に存在するタンパク質であり、RAC1遺伝子によりコードされている。RAC1は選択的スプライシングにより異なる機能を持ったいくつかのタンパク質を生成しており、このうちの1つがRac1である。Rac1は、悪性黒色腫や肺非小細胞癌 を含むさまざまな癌の発生において、重要な役割を果たしていると考えられている。そのため、現在これらの疾患に対する治療標的と考えられている。 (ja) ▪ Rac1 (англ. Ras-related C3 botulinum toxin substrate 1) — внутриклеточный белок из суперсемейства ГТФаз, относится к «малым» G-белкам. Находится в двух состояниях: активном ГТФ-связанном и неактивном ГДФ-связанном состоянии. В своей активной форме Rac1 связывается в клетке с целым рядом эффекторных белков и приводит к регулировке многих клеточных процессов, таких как секреция, фагоцитоз апоптозных клеток, поляризация эпителиальных клеток и индуцированное факторами роста образование мембранных складок и выростов (англ. membrane ruffles). (ru) ▪ Rac1, also known as Ras-related C3 botulinum toxin substrate 1, is a protein found in human cells. It is encoded by the RAC1 gene. This gene can produce a variety of alternatively spliced versions of the Rac1 protein, which appear to carry out different functions. (en)
dbo:entrezgene	▪ 5879
dbo:wikiPageExternalLink	▪ http://cmkb.cellmigration.org/report.cgi?report=orth_overview&gene_id=5879 ▪ http://www.cellmigration.org/index.shtml
dbo:wikiPageID	▪ 13562705 (xsd:integer)
dbo:wikiPageRevisionID	▪ 703055305 (xsd:integer)
rdf:type	▪ owl:Thing ▪ wikidata:Q206229 ▪ wikidata:Q8054 ▪ dbo:Biomolecule ▪ dbo:Protein

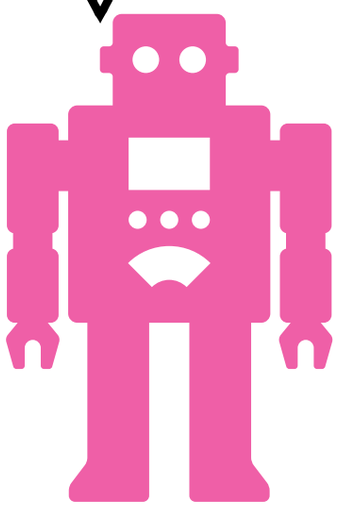
Querying DBpedia with SPARQL

```
SELECT DISTINCT ?gene ?entrez_id ?uniprot_id WHERE {
    ?gene dbo:abstract ?abstract .
    FILTER (regex(?abstract, "toxin")).
    ?gene dbo:entrezgene ?entrez_id .
    OPTIONAL {?gene dbo:uniprot ?uniprot_id} .
}
```

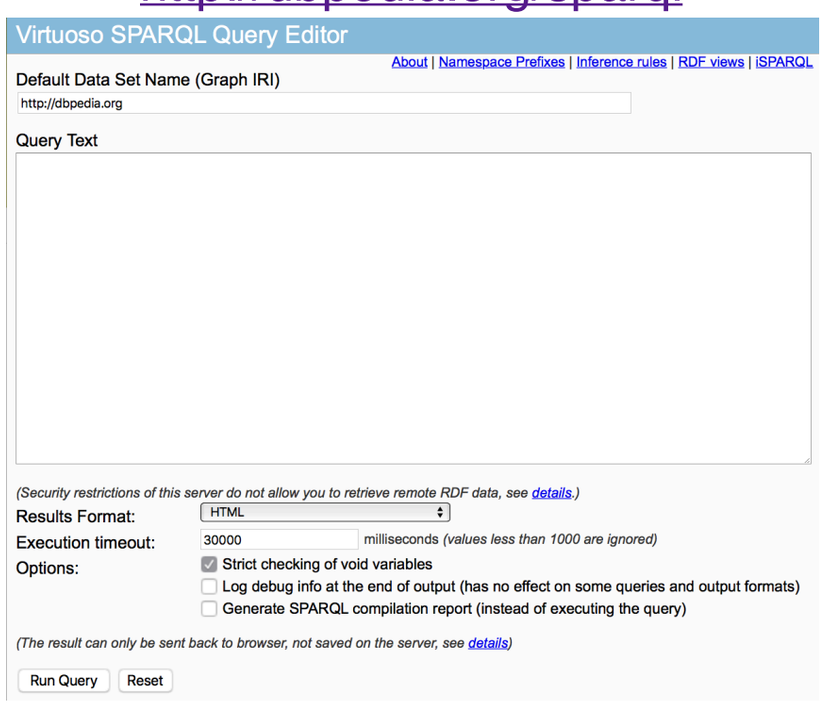


Querying DBpedia with SPARQL

```
SELECT DISTINCT ?gene ?entrez_id ?uniprot_id WHERE {
    ?gene dbo:abstract ?abstract .
    FILTER (regex(?abstract, "toxin")).
    ?gene dbo:entrezgene ?entrez_id .
    OPTIONAL {?gene dbo:uniprot ?uniprot_id} .
}
```

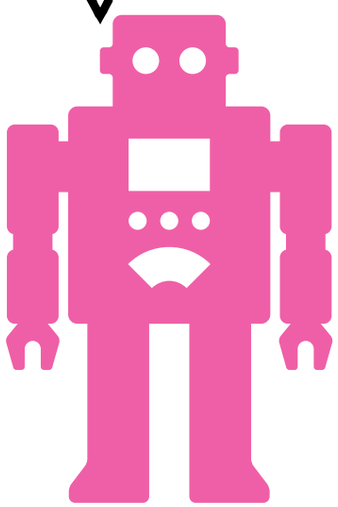


http://dbpedia.org/sparql

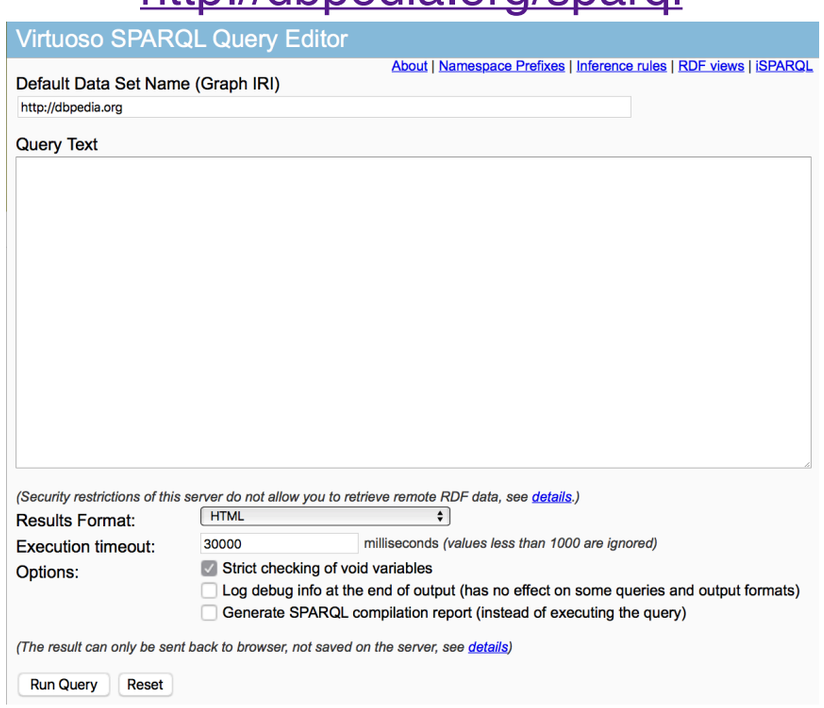


Querying DBpedia with SPARQL

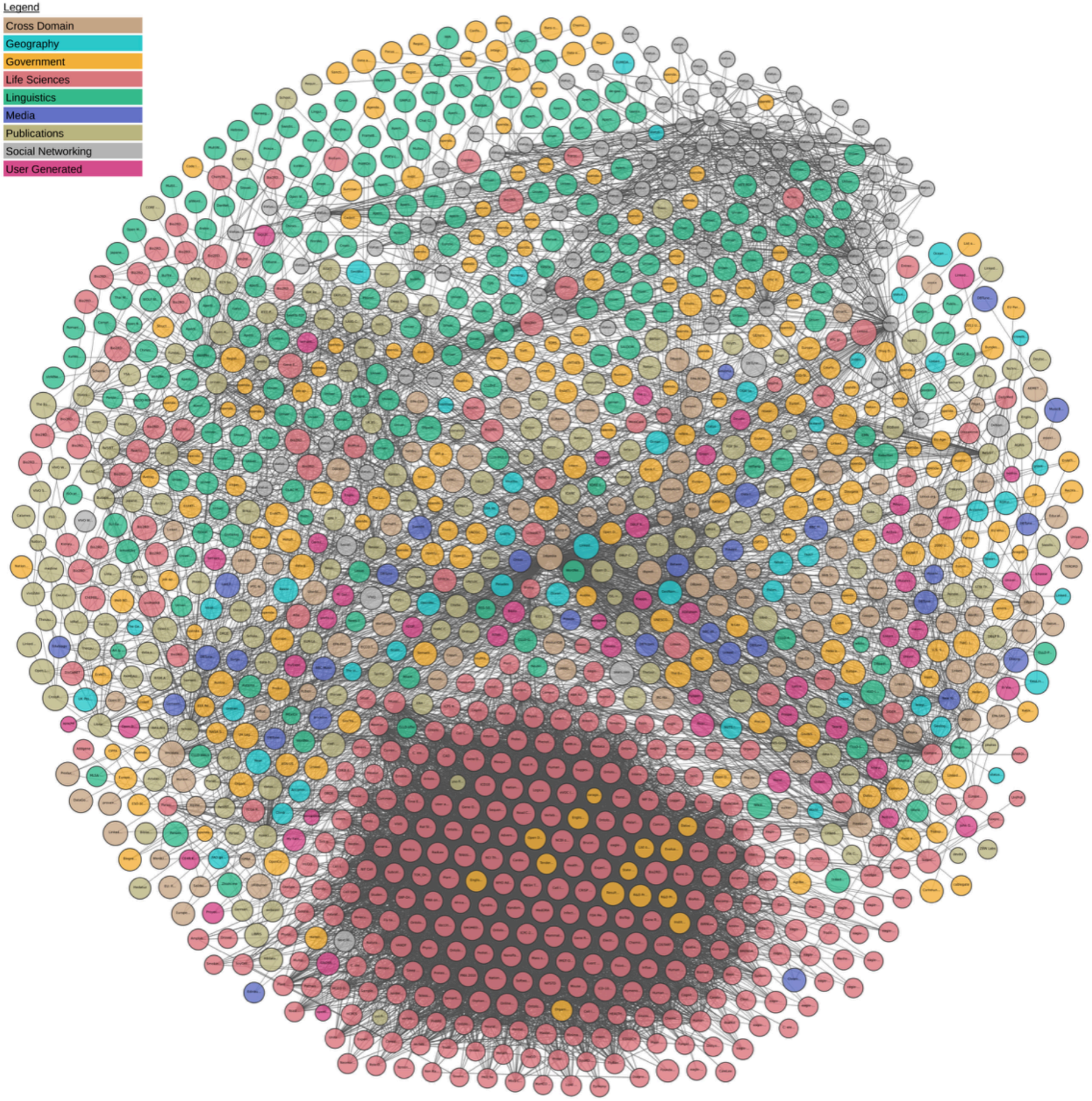
```
SELECT DISTINCT ?gene ?entrez_id ?uniprot_id WHERE {
    ?gene dbo:abstract ?abstract .
    FILTER (regex(?abstract, "toxin")).
    ?gene dbo:entrezgene ?entrez_id .
    OPTIONAL {?gene dbo:uniprot ?uniprot_id} .
}
```



http://dbpedia.org/sparql

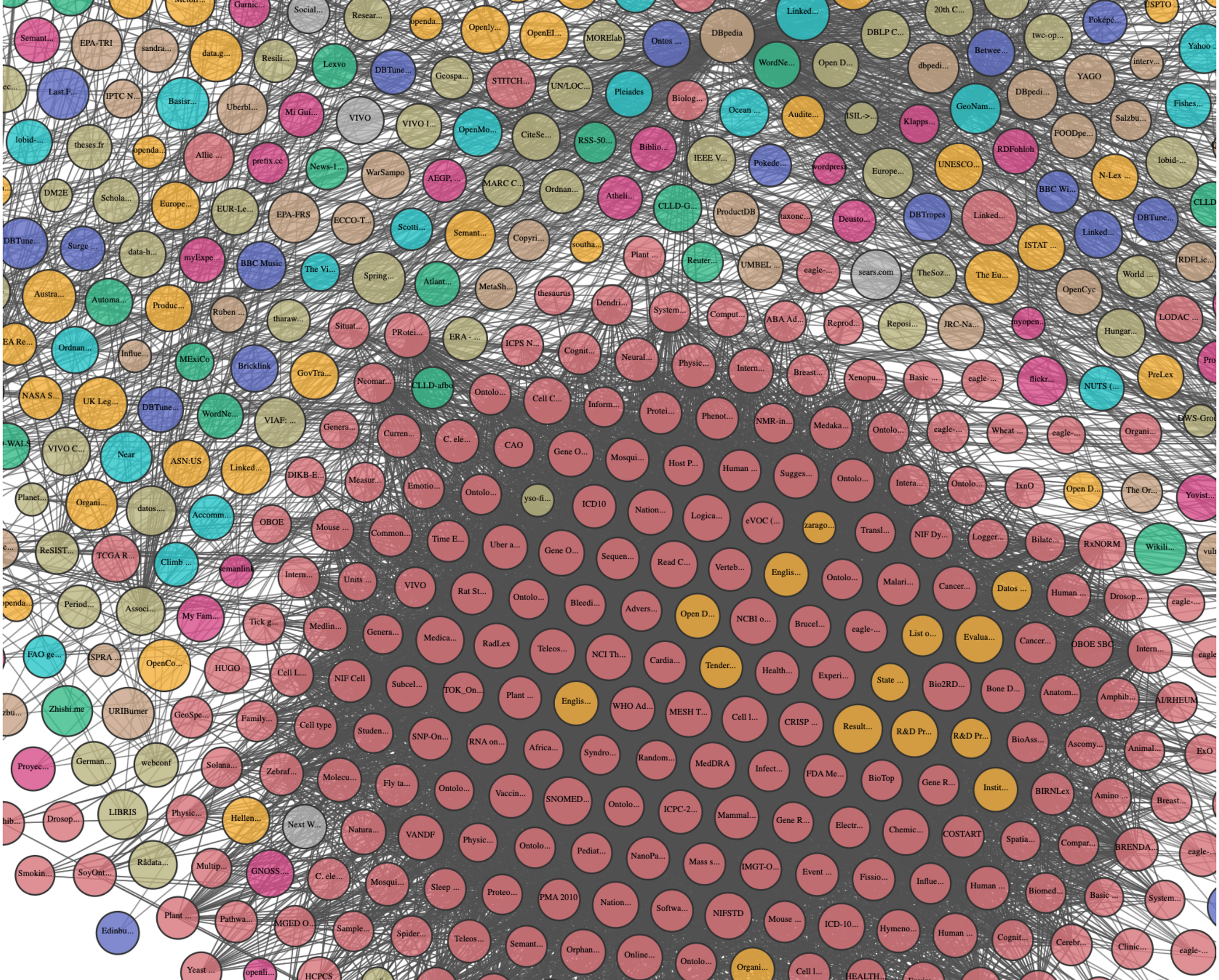


gene	entrez_id	uniprot_id
http://dbpedia.org/resource/DsbA	"948353"	"P0AEG4"
http://dbpedia.org/resource/Cholinesterase	"590"	"P06276"
http://dbpedia.org/resource/Cholinesterase	"590"	"P22303"
http://dbpedia.org/resource/Cholinesterase	"43"	"P06276"
http://dbpedia.org/resource/Cholinesterase	"43"	"P22303"
http://dbpedia.org/resource/Clostridium_perfringens_alpha_toxin	"988262"	
http://dbpedia.org/resource/Lymphotoxin	"4049"	"P01374"
http://dbpedia.org/resource/Lymphotoxin	"4049"	"Q06643"
http://dbpedia.org/resource/Lymphotoxin	"4050"	"P01374"
http://dbpedia.org/resource/Lymphotoxin	"4050"	"Q06643"
http://dbpedia.org/resource/Casein_kinase_2	"1457"	"P19784"
http://dbpedia.org/resource/Casein_kinase_2	"1457"	"P67870"
http://dbpedia.org/resource/Casein_kinase_2	"1457"	"P68400"
http://dbpedia.org/resource/Casein_kinase_2	"1460"	"P19784"
http://dbpedia.org/resource/Casein_kinase_2	"1460"	"P67870"
http://dbpedia.org/resource/Casein_kinase_2	"1460"	"P68400"
http://dbpedia.org/resource/Casein_kinase_2	"1459"	"P19784"
http://dbpedia.org/resource/Casein_kinase_2	"1459"	"P67870"
http://dbpedia.org/resource/Casein_kinase_2	"1459"	"P68400"
http://dbpedia.org/resource/Collagenase	"4317"	"P03956"
http://dbpedia.org/resource/Collagenase	"4317"	"P22894"
http://dbpedia.org/resource/Collagenase	"4312"	"P03956"
http://dbpedia.org/resource/Collagenase	"4312"	"P22894"
http://dbpedia.org/resource/Guanylin	"2980"	"Q02747"
http://dbpedia.org/resource/Macrophage_inflammatory_protein	"6348"	"P10147"
http://dbpedia.org/resource/Macrophage_inflammatory_protein	"6348"	"P13236"
http://dbpedia.org/resource/Macrophage_inflammatory_protein	"6351"	"P10147"
http://dbpedia.org/resource/Macrophage_inflammatory_protein	"6351"	"P13236"



The Linked Open Data Cloud from lod-cloud.net

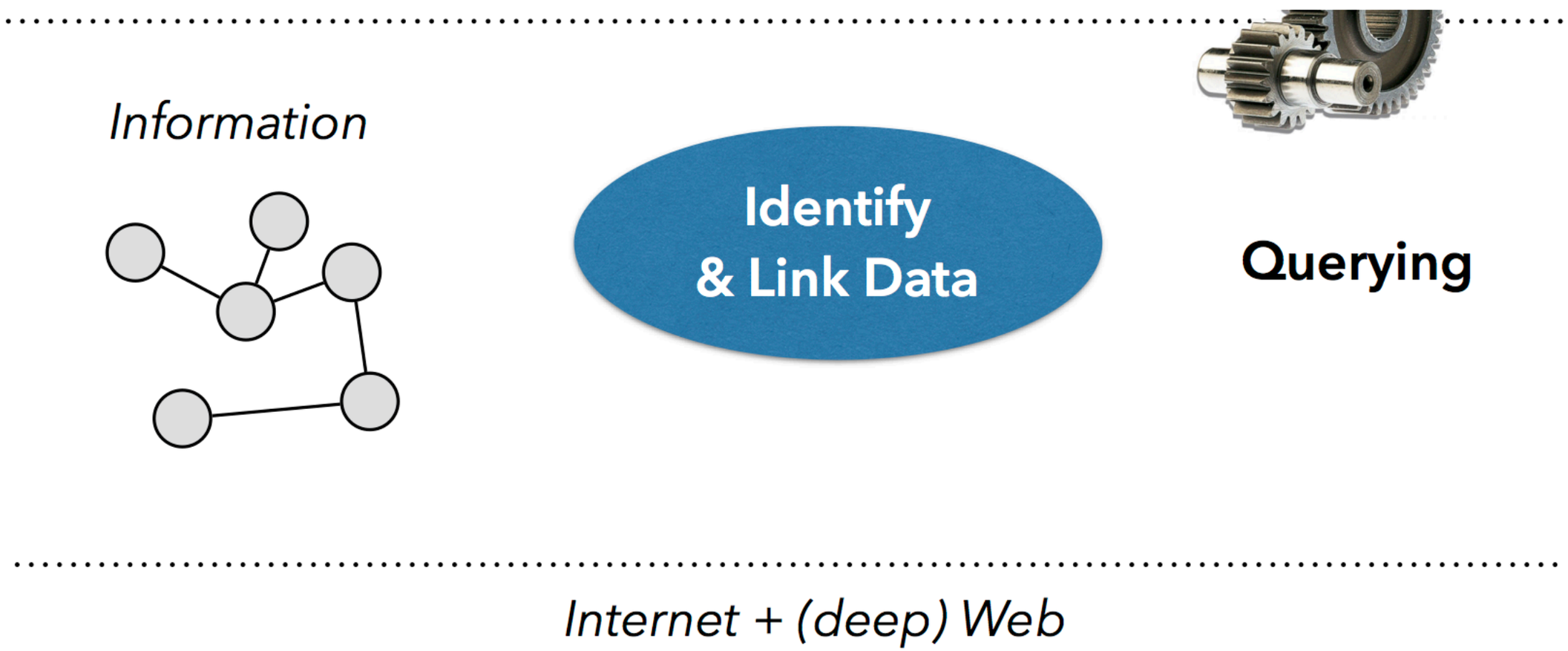
CC BY


Semant...
EPA-TRI
sandra...
data.g...
Resili...
Lexvo
DBTune...
Geospa...
STITCH...
UN/LOC...
Pleiades
Biolog...
Ocean ...
Audite...
ISIL->...
Klapps...
GeoNam...
DBpedi...
YAGO
Fishes...
Salzbu...
FOODpe...
RDFohloh
UNESCO...
Europe...
wordpres...
Pokede...
IEEE V...
Biblio...
RSS-50...
CiteSe...
OpenMo...
VIVO I...
VIVO
Mi Guí...
Uberbl...
Basisr...
IPTC N...
Last.F...
lobid-...
theses.fr
openda...
Allie ...
prefix.cc
News-1...
WarSampo
AEGP, ...
MARC C...
Ordnan...
Atheli...
CLLD-G...
ProductDB
taxonc...
Deusto...
DBTropes
Linked...
BBC Wi...
N-Lex ...
DBTune...
Linked...
ISTAT ...
The Eu...
TheSoz...
sears.com
eagle-...
UMBEL ...
Reuter...
Plant ...
southa...
Copyri...
Semant...
Scotti...
ECCO-T...
EPA-FRS
EUR-Le...
Europe...
Schola...
DM2E
DBTune...
Surge ...
data-h...
myExpe...
BBC Music
The Vi...
Spring...
Atlant...
MetaSh...
thesaurus
Dendri...
System...
Comput...
ABA Ad...
Reprod...
Reposi...
JRC-Na...
myopen...
OpenCyc
World ...
RDFLic...
LODAC ...
Hungar...
Austra...
Automa...
Produc...
Ruben ...
tharaw...
Situat...
PRotei...
ERA - ...
ICPS N...
Cognit...
Neural...
Physic...
Intern...
Breast...
Xenopu...
Basic ...
eagle-...
flickr...
NUTS (...
PreLex
EA Re...
Ordnan...
Influe...
MExiCo
Bricklink
GovTra...
Neomar...
CLLD-afbo
Ontolo...
Cell C...
Inform...
Protei...
Phenot...
NMR-in...
Medaka...
Ontolo...
eagle-...
Wheat ...
eagle-...
Organi...
NASA S...
UK Leg...
DBTune...
WordNe...
VIAF: ...
Genera...
Curren...
C. ele...
CAO
Gene O...
Mosqui...
Host P...
Human ...
Sugges...
Ontolo...
Intera...
Ontolo...
IxnO
Open D...
The Or...
Yovist...
VIVO C...
Near
ASN:US
Linked...
DIKB-E...
Measur...
Emotio...
Ontolo...
yso-fi...
ICD10
Nation...
Logica...
eVOC (...
zarago...
Transl...
NIF Dy...
Logger...
Bilate...
RxNORM
Wikili...
Planet...
Organi...
datos....
Accomm...
OBOE
Mouse ...
Common...
Time E...
Uber a...
Gene O...
Sequen...
Read C...
Verteb...
Englis...
Ontolo...
Malari...
Cancer...
Datos ...
Human ...
Drosop...
eagle-...
ReSIST...
TCGA R...
Climb ...
Intern...
Units ...
VIVO
Rat St...
Ontolo...
Bleedi...
Advers...
Open D...
NCBI o...
Brucel...
eagle-...
List o...
Evalua...
Cancer...
OBOE SBC
Intern...
Period...
Associ...
My Fam...
Tick g...
Medlin...
Genera...
Medica...
RadLex
Teleos...
NCI Th...
Cardia...
Tender...
Health...
Experi...
State ...
Bio2RD...
Bone D...
Anatom...
Amphib...
FAO ge...
SPRA ...
OpenCo...
HUGO
Cell L...
NIF Cell
Subcel...
TOK_On...
Plant ...
Englis...
WHO Ad...
MESH T...
Cell l...
CRISP ...
Result...
R&D Pr...
R&D Pr...
BioAss...
Ascomy...
Animal...
ExO
Zhishi.me
URIBurner
GeoSpe...
Family...
Cell type
Studen...
SNP-On...
RNA on...
Africa...
Syndro...
Random...
MedDRA
Infect...
FDA Me...
BioTop
Gene R...
Instit...
BIRNLex
Amino ...
Breast...
Proyec...
German...
webconf
Solana...
Zebraf...
Molecu...
Fly ta...
Ontolo...
Vaccin...
SNOMED...
Ontolo...
ICPC-2...
Mammal...
Gene R...
Electr...
Chemic...
COSTART
Spatia...
Compar...
BRENDA...
Drosop...
LIBRIS
Physic...
Hellen...
Next W...
Natura...
VANDF
Physic...
Ontolo...
Pediat...
NanoPa...
Mass s...
IMGT-O...
Event ...
Fissio...
Influe...
Human ...
Biomed...
Basic ...
System...
Smokin...
SoyOnt...
Rådata...
Multip...
GNOSS...
C. ele...
Mosqui...
Sleep ...
Proteo...
PMA 2010
Nation...
Softwa...
NIFSTD
Mouse ...
ICD-10...
Hymeno...
Human ...
Cognit...
Cerebr...
Clinic...
Edinbu...
Plant ...
Pathwa...
MGED O...
Sample...
Spider...
Teleos...
Semant...
Orphan...
Online...
Ontolo...
Organi...
Cell l...
HEALTH...
Yeast ...
openli...
HCPCS

Not only DBpedia !
31 billion triples in 2011 ...

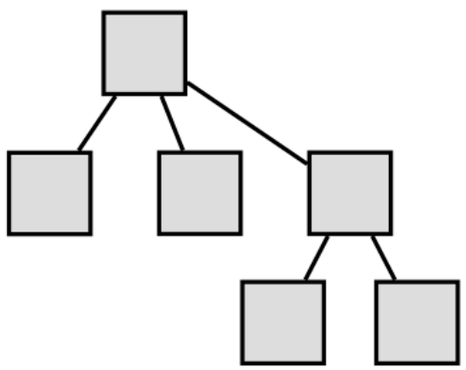
Summary

Summary


Information
Identify
& Link Data
Querying
Internet + (deep) Web

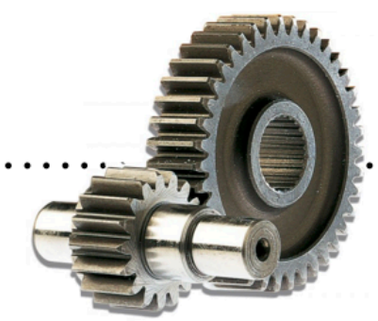
Summary

Knowledge

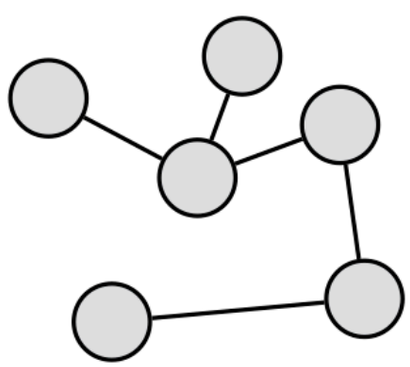


Vocabularies
Ontologies

Reasoning



Information



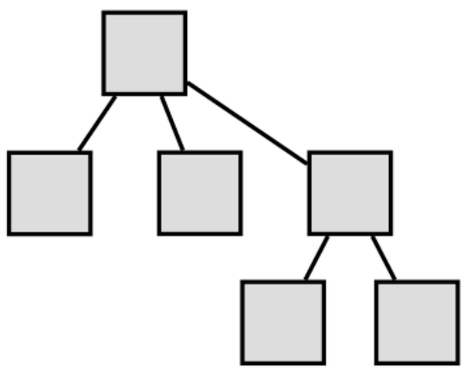
Identify
& Link Data

Querying

Internet + (deep) Web

Summary

Knowledge



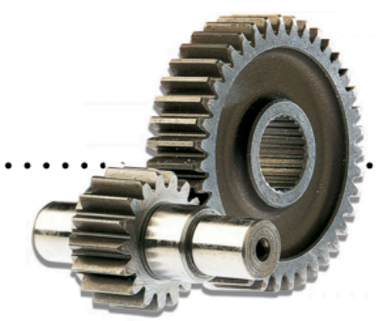
RDF-S

OWL

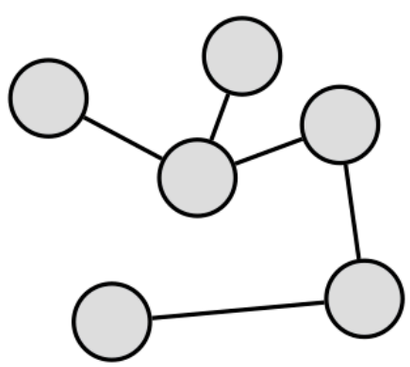
SWRL

Vocabularies
Ontologies

Reasoning



Information



Identify
& Link Data

URIs

RDF

Querying

SPARQL

Internet + (deep) Web

RDF graphs

RDF to link data

RDF to link data

Definitions

(1) An RDF statement expresses a **relationship** between two resources

(2) The **subject** and the **object** represent the two resources being related ; the **predicate** represents the nature of their relationship

(3) The relationship is phrased in a **directional** way (from subject to object) and is called in RDF a **property**

(4) Because RDF statements consist of three elements they are called **triples**

RDF to link data

Definitions

(1) An RDF statement expresses a **relationship** between two resources

(2) The **subject** and the **object** represent the two resources being related ; the **predicate** represents the nature of their relationship

(3) The relationship is phrased in a **directional** way (from subject to object) and is called in RDF a **property**

(4) Because RDF statements consist of three elements they are called **triples**

Sample RDF triples :

```
<RAC1> <is a> <human gene> .
<seq1> <is a variant of> <RAC1> .
```

RDF graphs

RDF graphs

Definitions

(1) A **graph** structure is formed with a set of **nodes** (resources) and **edges**(relationships between resources)

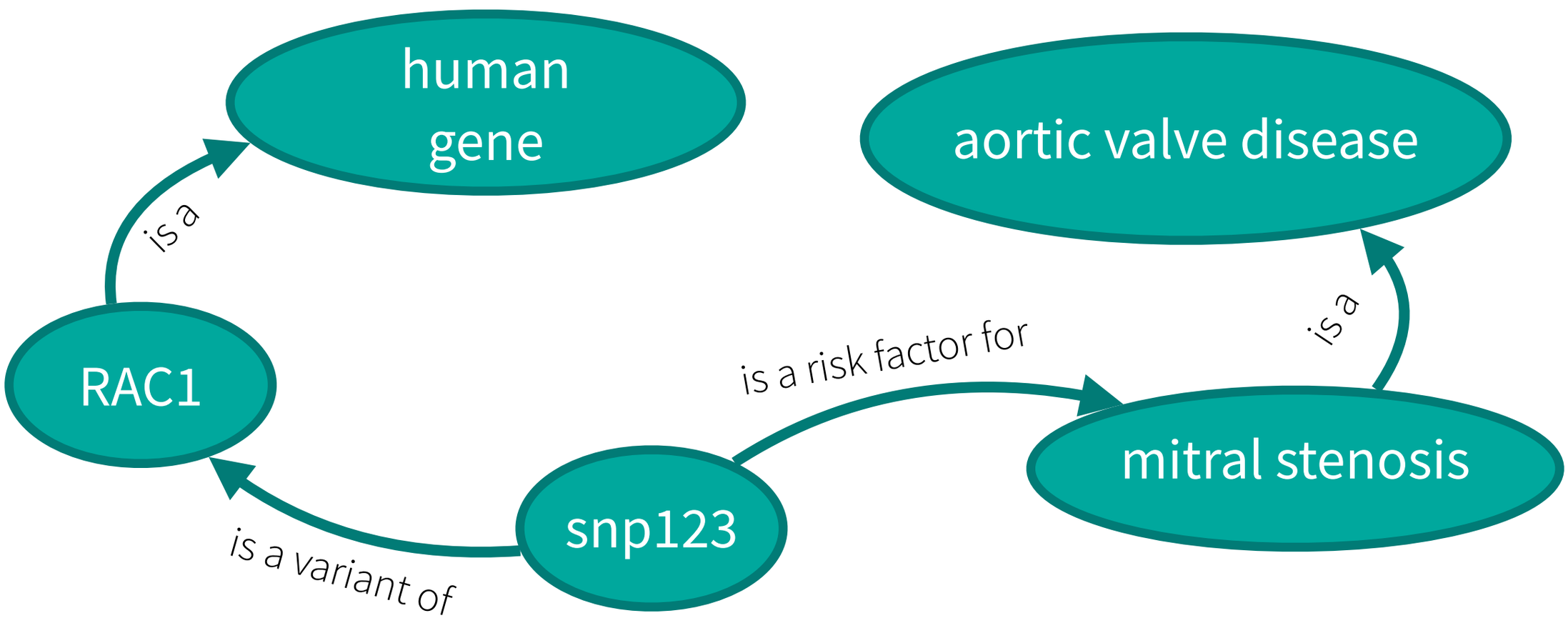
(2) A set of RDF triples is called an RDF graph. RDF is a **directed**, **labeled graph** data format for representing information in the Web.

RDF graphs

Definitions

(1) A **graph** structure is formed with a set of **nodes** (resources) and **edges**(relationships between resources)

(2) A set of RDF triples is called an RDF graph. RDF is a **directed**, **labeled graph** data format for representing information in the Web.



RDF-S to represent knowledge graphs

RDF-Schema aims at providing a simple vocabulary to **organize domain-specific knowledge** through classes (**concepts**) and properties (**relationships**).

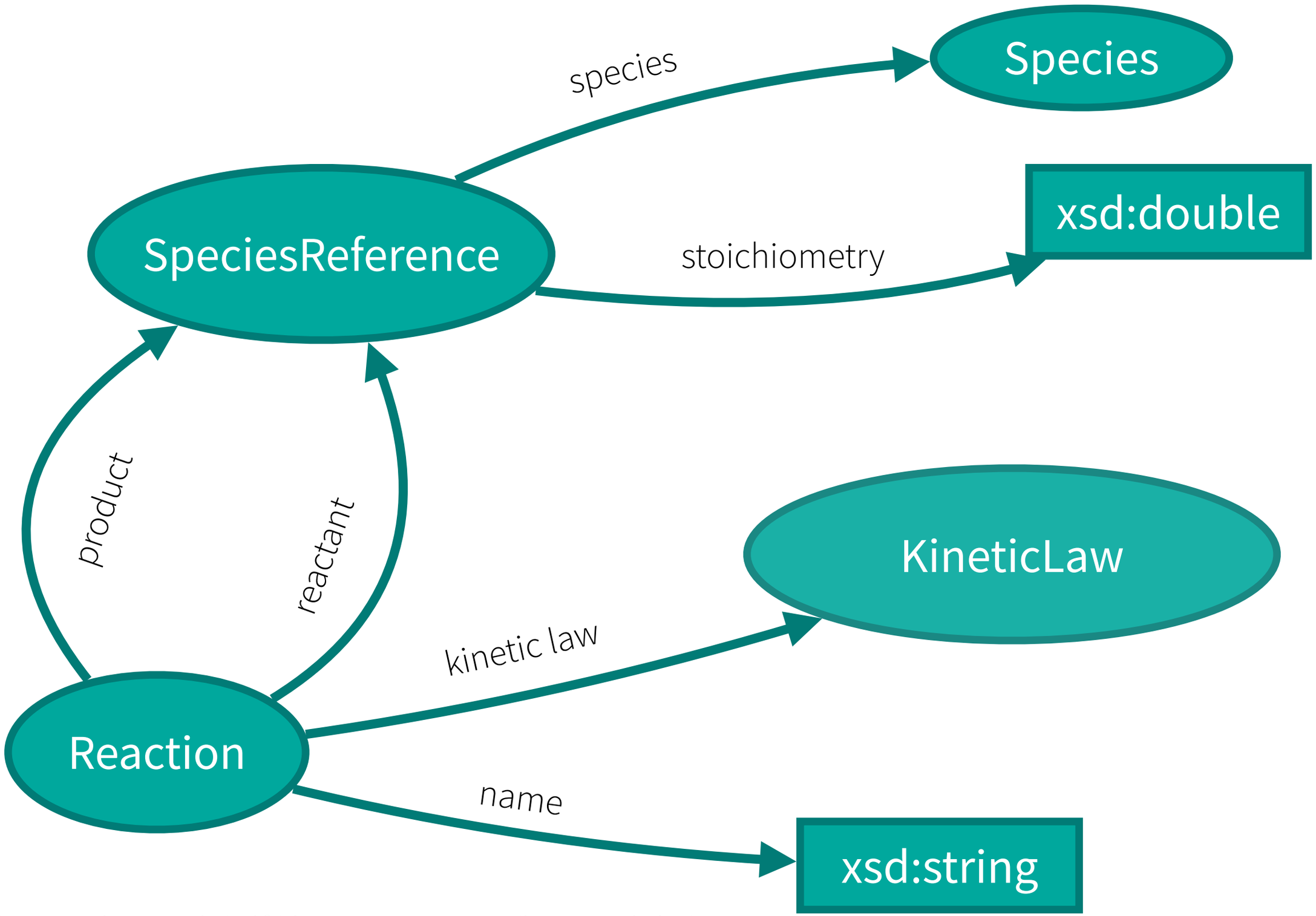
Classes

Resources may be divided into groups called **classes**. The members of a class are known as **instances** of the class. The rdf:**type** property may be used to state that a resource is an instance of a class (« is a » relation).

Properties

- rdfs:**range**: a constraint on the class membership(s) for values of this property
- rdfs:**domain**: a constraint on the class membership(s) for resources having this property
- rdfs:**label**, rdfs:**comment**
- rdf:**type**: to state that a resource is an instance of a class
- rdfs:**subClassOf**: to state that all the instances of one class are instances of another
- rdfs:**subPropertyOf**: to state that all resources related by one property are also related by another

Sample RDF-S vocabulary



https://www.ebi.ac.uk/rdf/documentation/biomodels/

Defining a domain-specific RDF property

```
my_ns:stoichiometry   rdf:type      rdfs:Property .
my_ns:stoichiometry   rdfs:domain   my_ns:SpeciesReference .
my_ns:stoichiometry   rdfs:range    xsd:double .
my_ns:stoichiometry   rdfs:label    "stoichiometry" .
```

Intuition for **logical inferences** (reasoning):

"Each thing designated by the stoichiometry property should be numerical value."

"*Each thing at the origin of a stoichiometry property should be a SpeciesReference object.*"

To go further: https://www.w3.org/TR/rdf11-mt/#rdf-entailment section 9.2.1

webprotege.stanford.edu/#projects/5ea9c833-134e-496c-a844-6619a3f7b9d5/edit/Properties?selection=Dat

Home | Mini-ontology-M2-BBRT | Project | Share | alban.gaignard | Help

Classes | Properties | Individuals | Comments | Changes by Entity | History | OWL Classes | Add tab

Property Hierarchy

Create Delete

- owl:topObjectProperty
  - hasFather
  - hasMother
  - hasGenotype
  - hasVariant
  - isPresentForIndividual
  - hasLoF
- owl:topDataProperty
  - hasChromosome
  - hasPosition
  - hasRefAllele
  - hasAltAllele
  - hasGwas
  - hasGeneId
  - hasValue
- Annotation properties

Property: hasGeneId

IRI

http://webprotege.stanford.edu/RDSpwGBLBluQRRMaZZVt4Yf

Annotations

rdfs:label | hasGeneId | lang

Enter property | Enter value | lang

☐ **Functional** (Single valued)

Domain

LossOfFunction

Range

xsd:string

Comments: hasGeneId

Start new thread

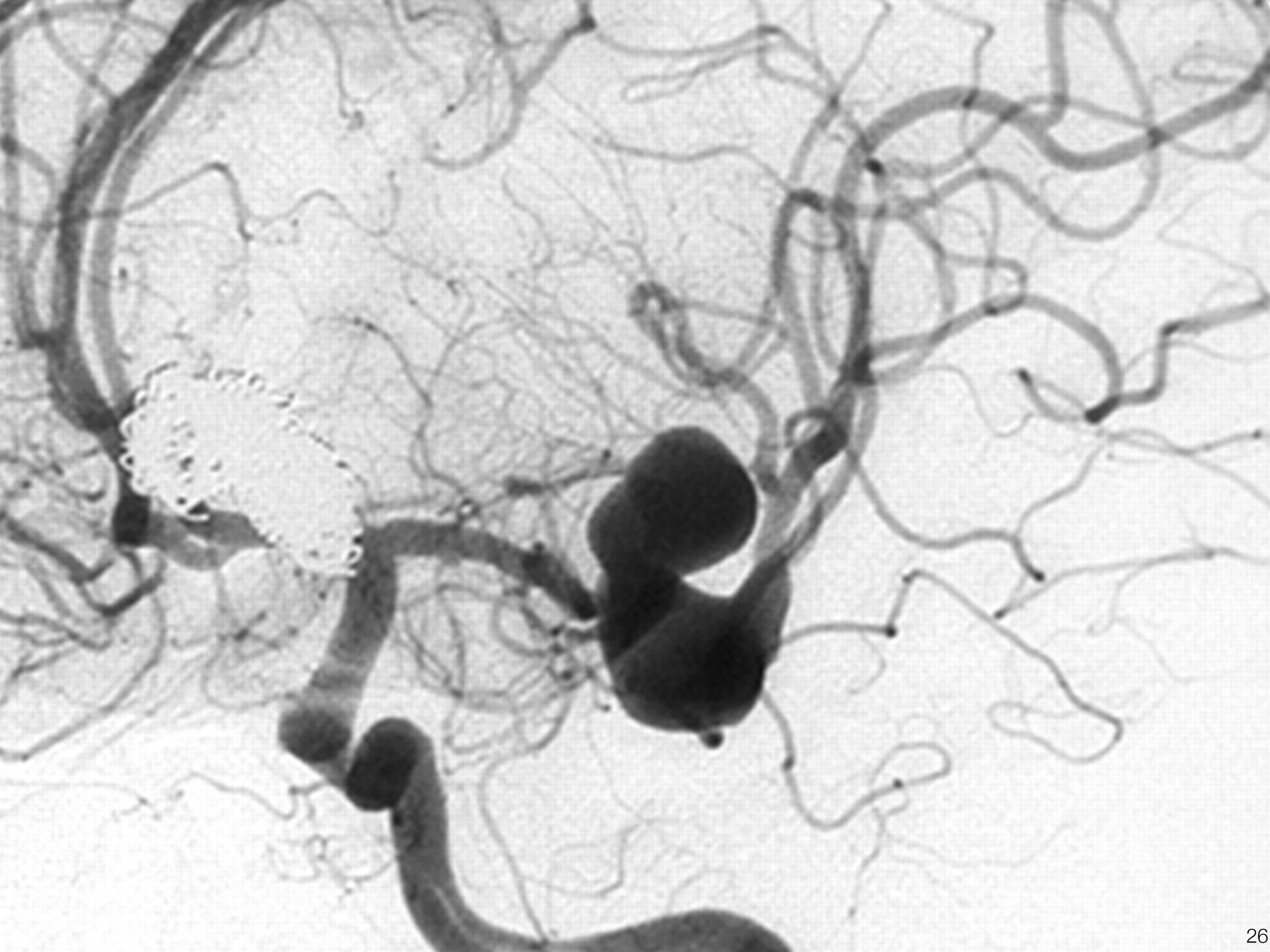
Project Feed

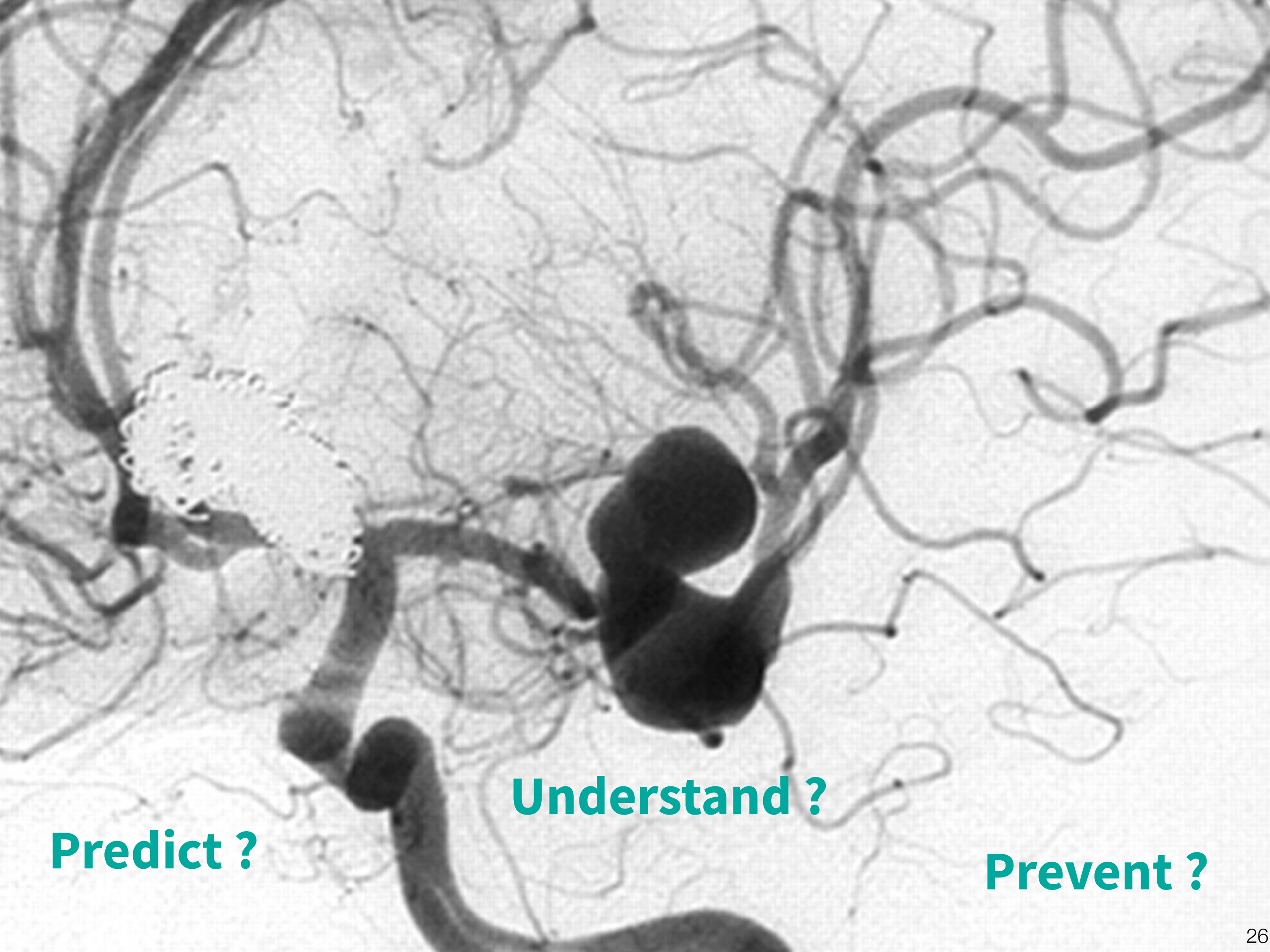
AG alban.gaignard Added xsd:string to the range of hasValue
DataProperty: hasValue
Less than one minute ago

AG alban.gaignard Added Genotype to the domain of hasValue
DataProperty: hasValue
Less than one minute ago

AG alban.gaignard Created data property hasValue as a sub-property of owl:topDataProperty

Knowledge Graphs, applied to intracranial aneurisms




Understand ?
Predict ?
Prevent ?

ICAN cohort : 34 univ. hospitals / 3000 subjects

3000 MRIs
1000 genotypes *(2000 more expected)*
93 exomes *(whole genomes expected)*

Understand ?

Predict ?

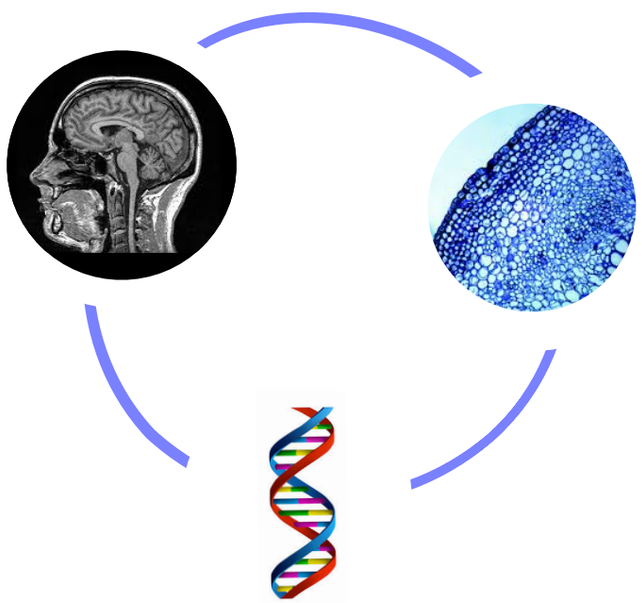
Prevent ?

Multiscale observations

predict risk of rupture
better **understand** mecanisms, causes
hopefully **prevent**

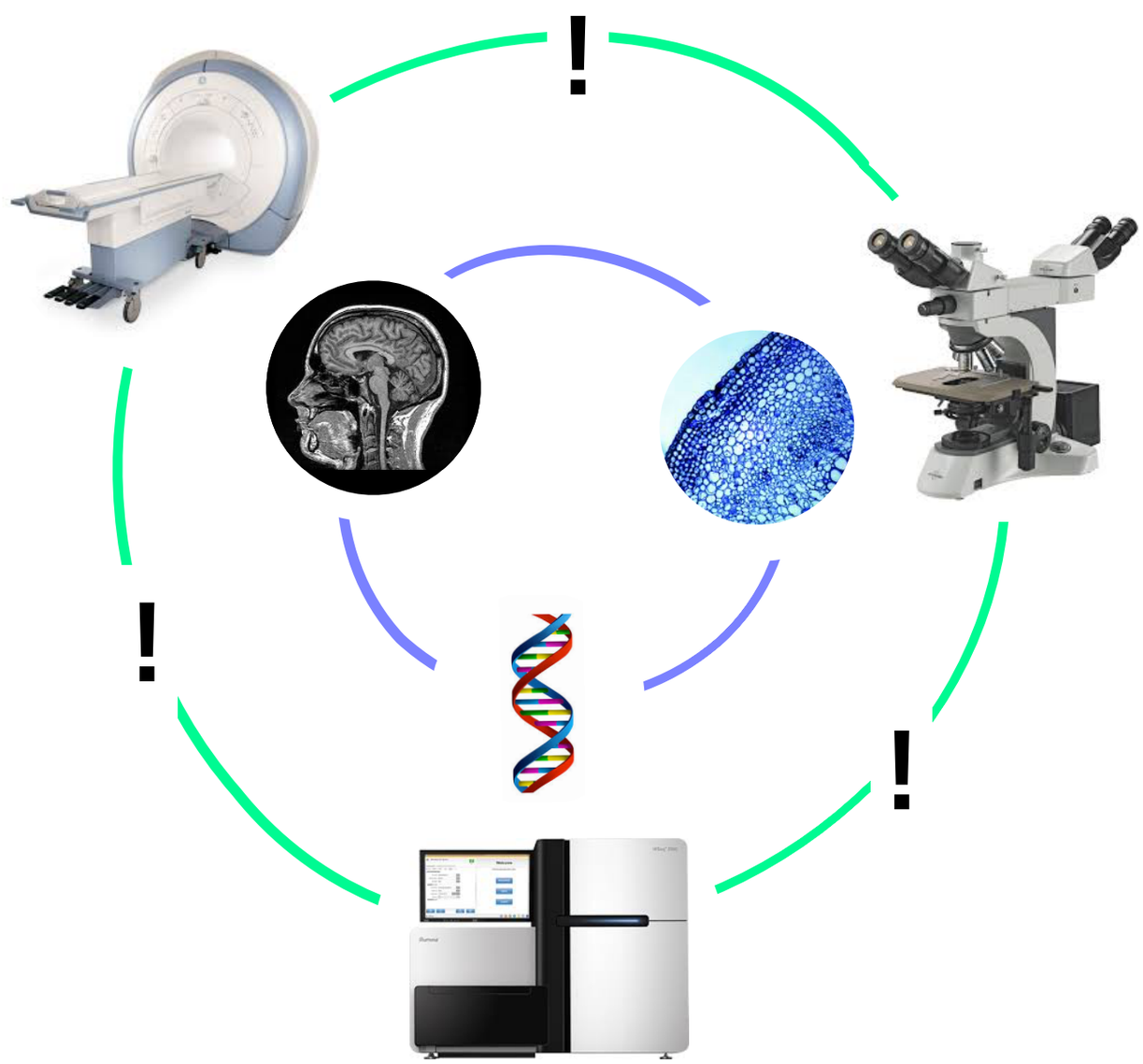
Multiscale observations

predict risk of rupture
better **understand** mecanisms, causes
hopefully **prevent**



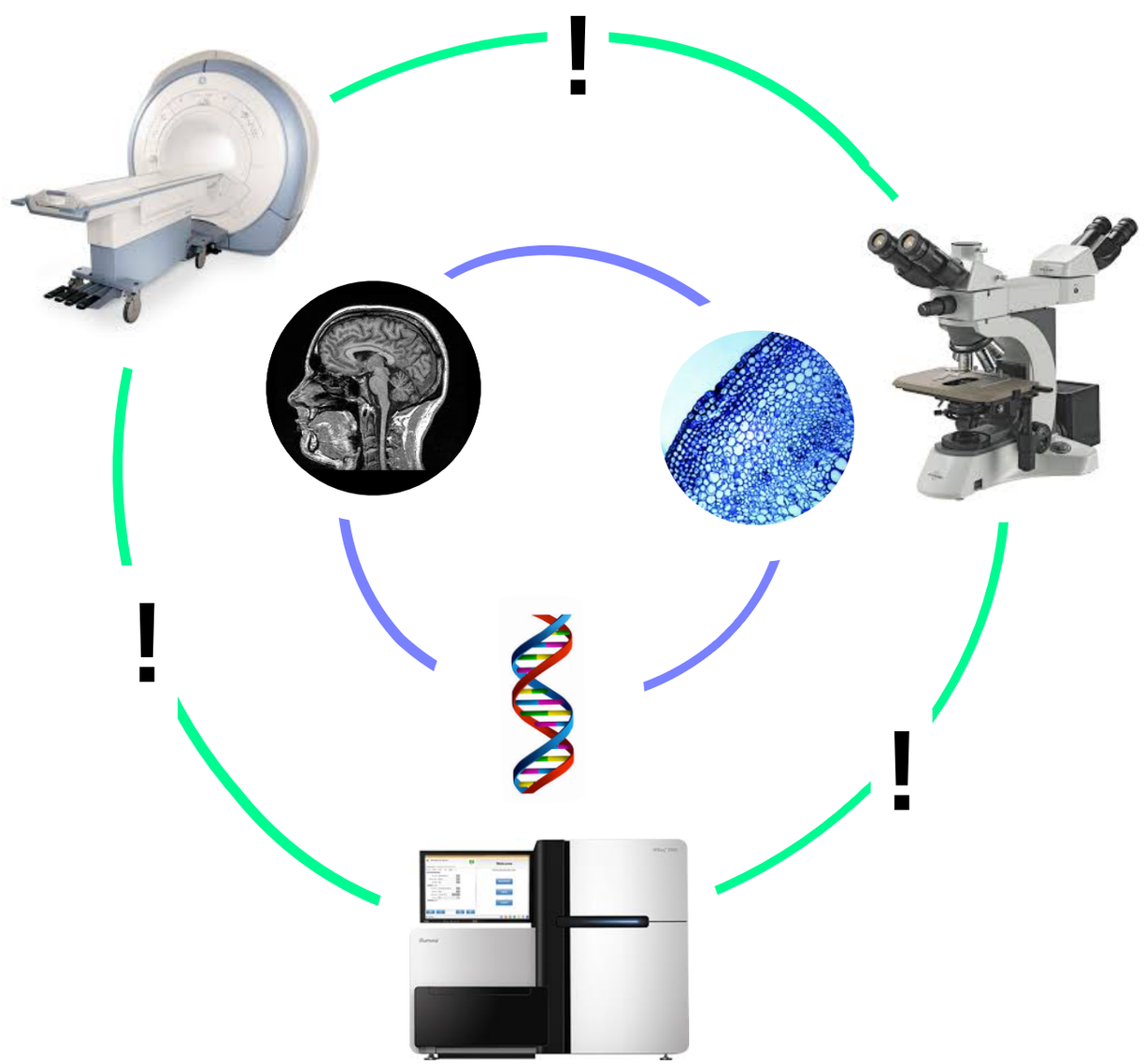
Multiscale observations

predict risk of rupture
better **understand** mecanisms, causes
hopefully **prevent**



Multiscale observations

predict risk of rupture
better **understand** mecanisms, causes
hopefully **prevent**



Challenges

- data silos
- interoperability
- data diversiy
- ≠ vocabularies / semantics
- variables >> samples (N)
- decentralization

Maxime Folschette

Kirsley Chennen

Bridging imaging-omics-clinical data: **INEX-MED**

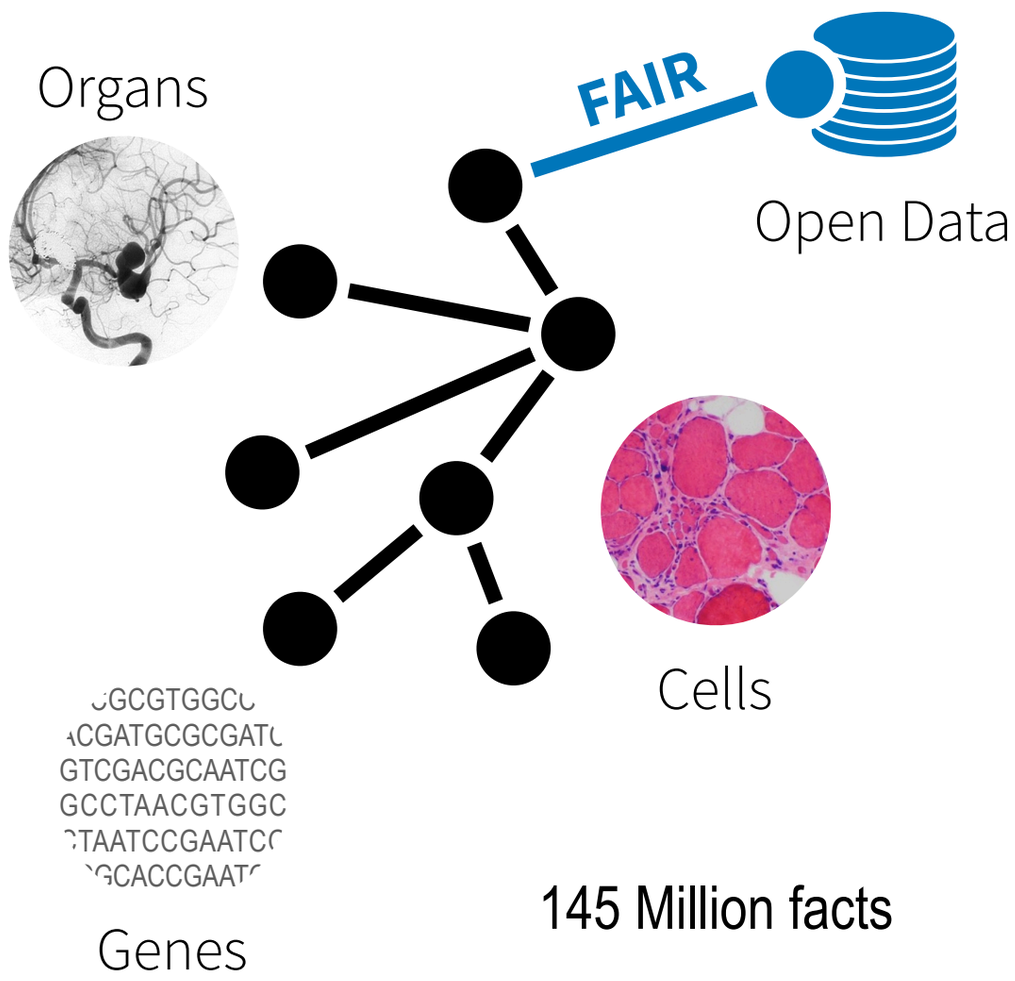
INEX-MED

Intracranial aneurysms (3.2 % world population)
Congenital myopathies (rare diseases)

Association between imaging phenotypes - « omics » signatures ?
Patients with high/low aneurysm rupture risk ?

INEX-MED

Intracranial aneurysms (3.2 % world population)
Congenital myopathies (rare diseases)

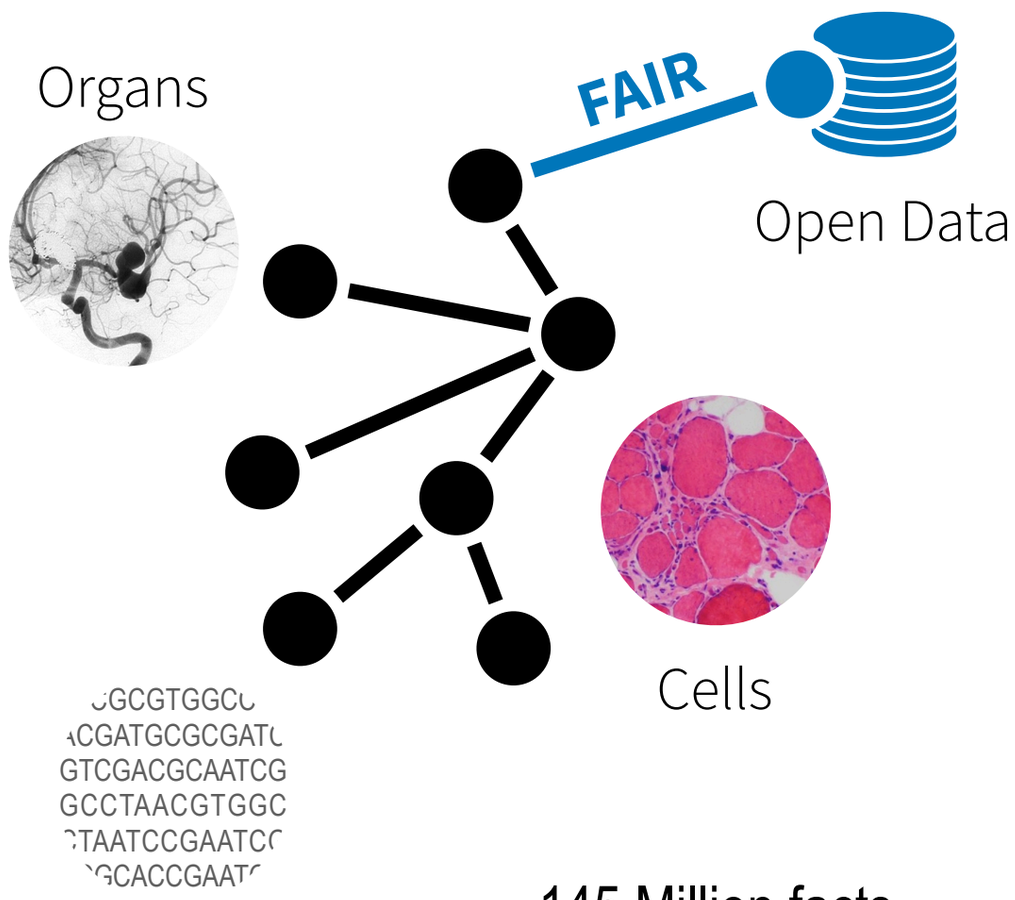


Multi-scale « Knowledge Graph »

Association between imaging phenotypes - « omics » signatures ?
Patients with high/low aneurysm rupture risk ?

INEX-MED

Intracranial aneurysms (3.2 % world population)
Congenital myopathies (rare diseases)



145 Million facts

Multi-scale « Knowledge Graph »

Select
Combine
Reuse
(Share)

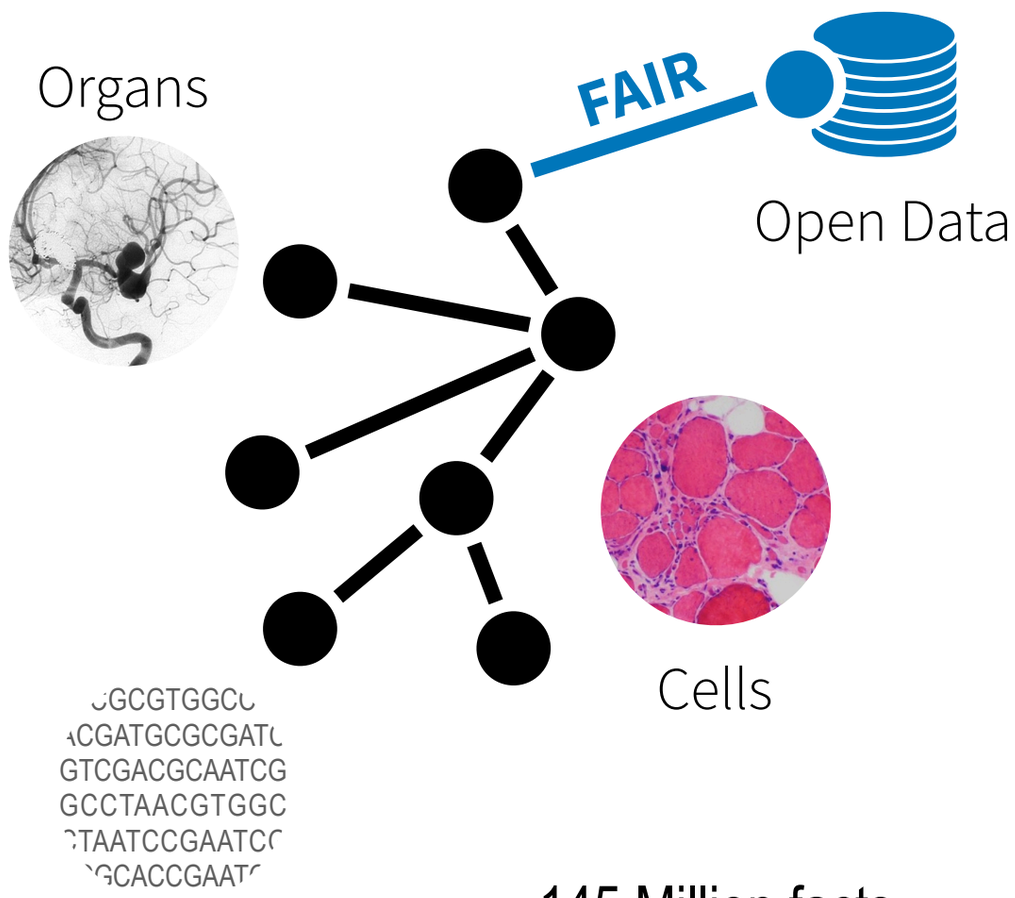
Ontologies
SPARQL queries

@Nantes

Association between imaging phenotypes - « omics » signatures ?
Patients with high/low aneurysm rupture risk ?

INEX-MED

Intracranial aneurysms (3.2 % world population)
Congenital myopathies (rare diseases)



145 Million facts

Multi-scale « Knowledge Graph »

Select
Combine
Reuse
(Share)

Ontologies
SPARQL queries

Predict
Explain

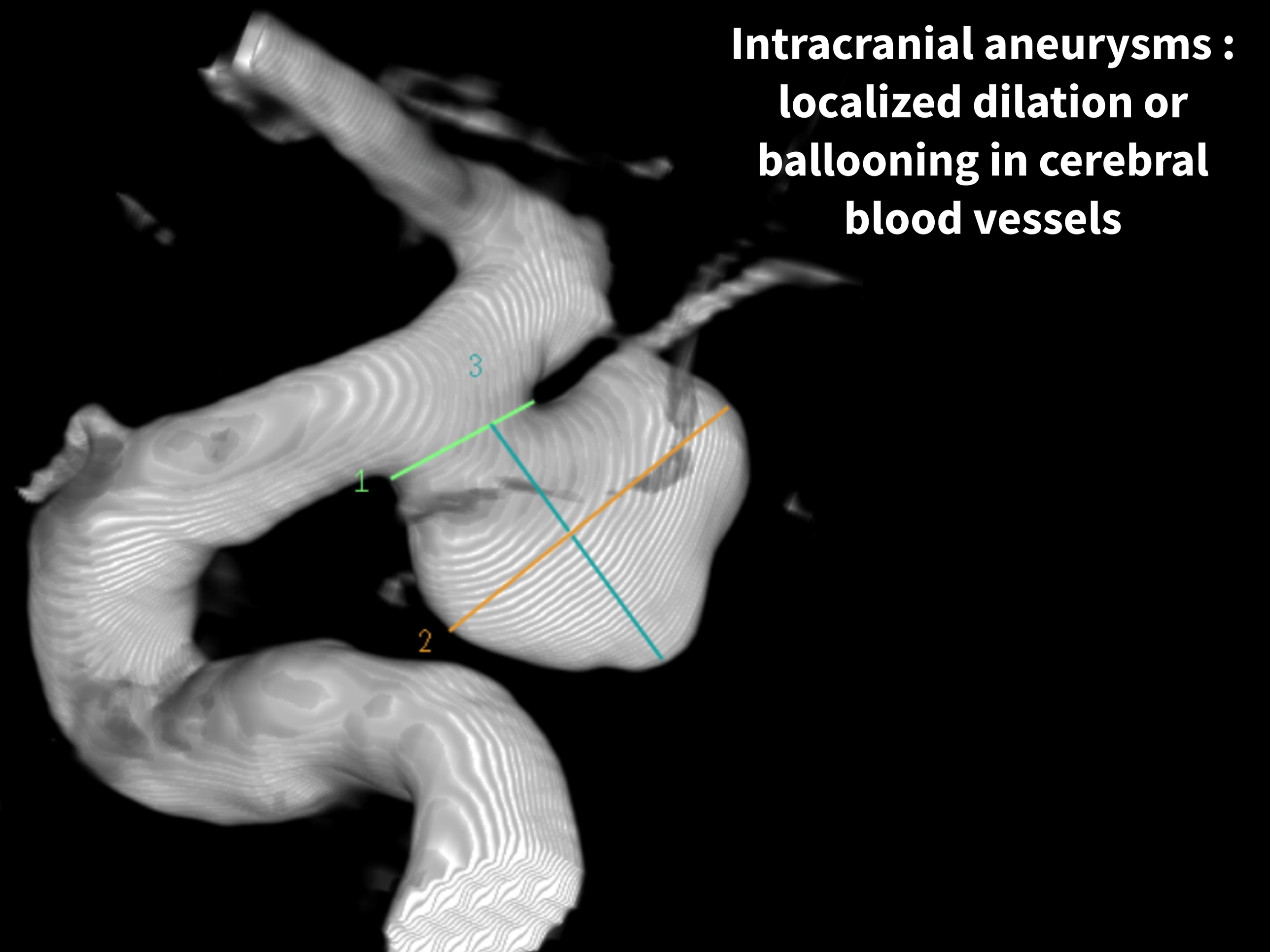
Statistics
Reasoning
Machine learning

@Nantes

@Strasbourg

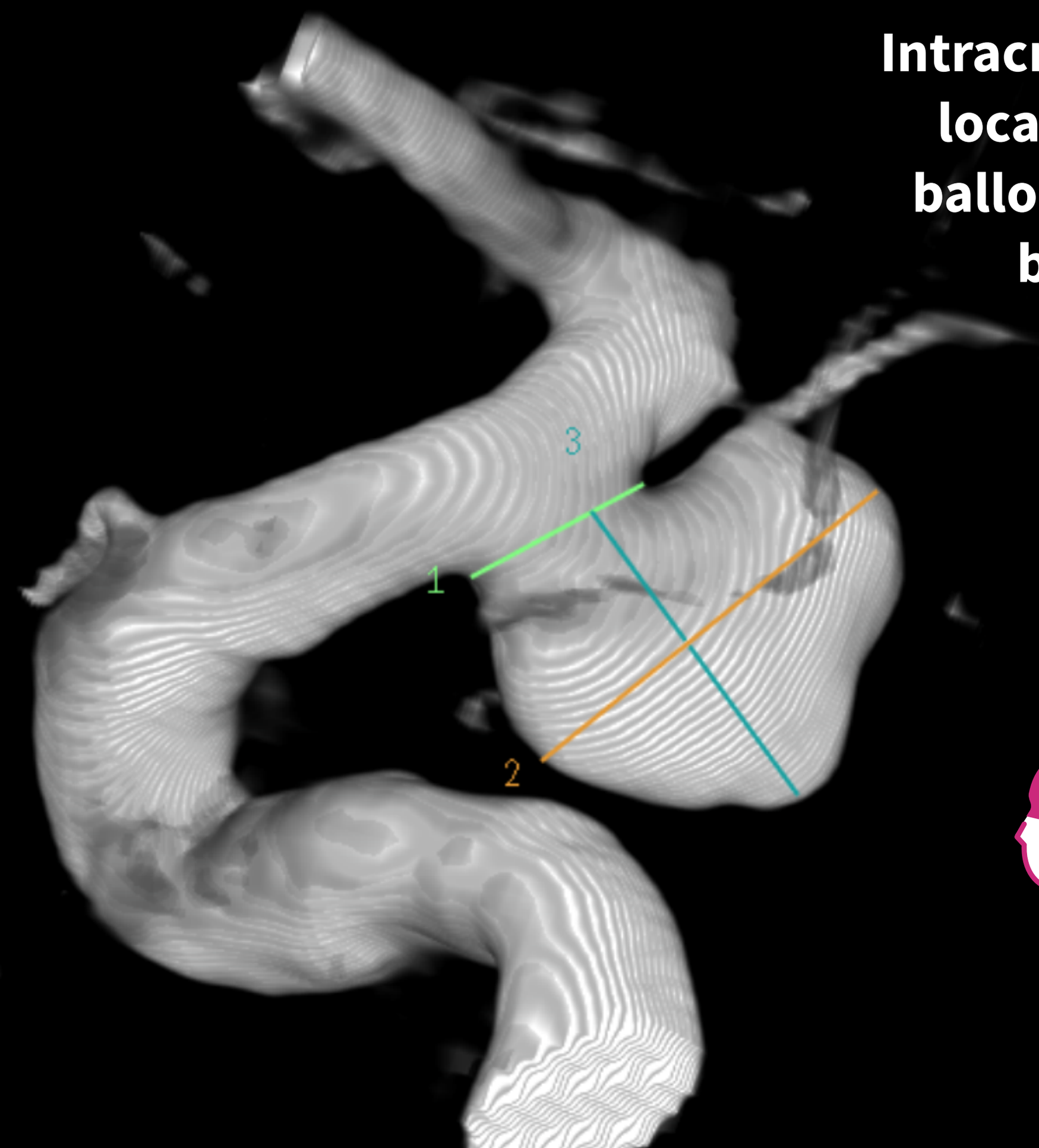
Association between imaging phenotypes - « omics » signatures ?
Patients with high/low aneurysm rupture risk ?

Imaging Data

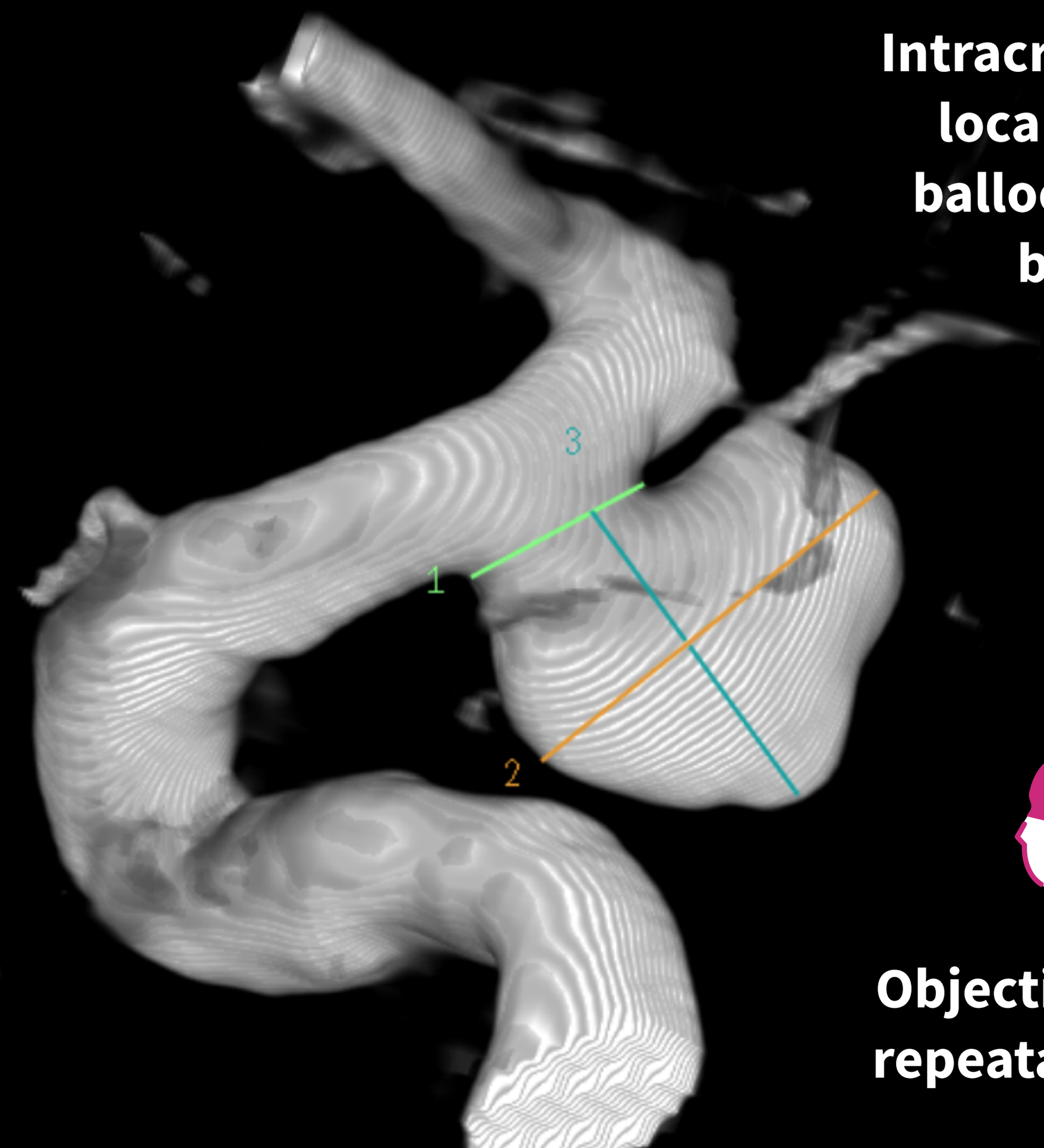


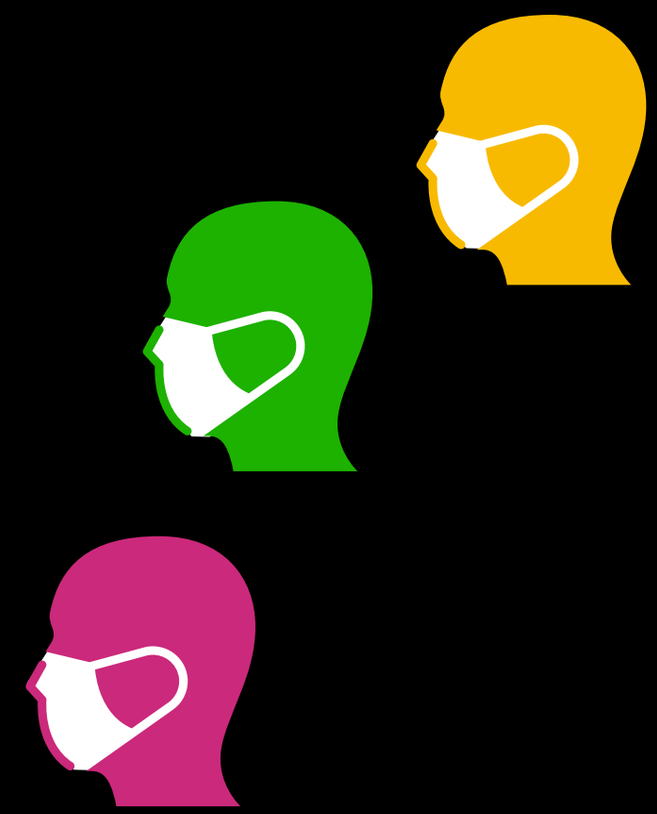
Intracranial aneurysms : localized dilation or ballooning in cerebral blood vessels

Intracranial aneurysms : localized dilation or ballooning in cerebral blood vessels



Intracranial aneurysms : localized dilation or ballooning in cerebral blood vessels

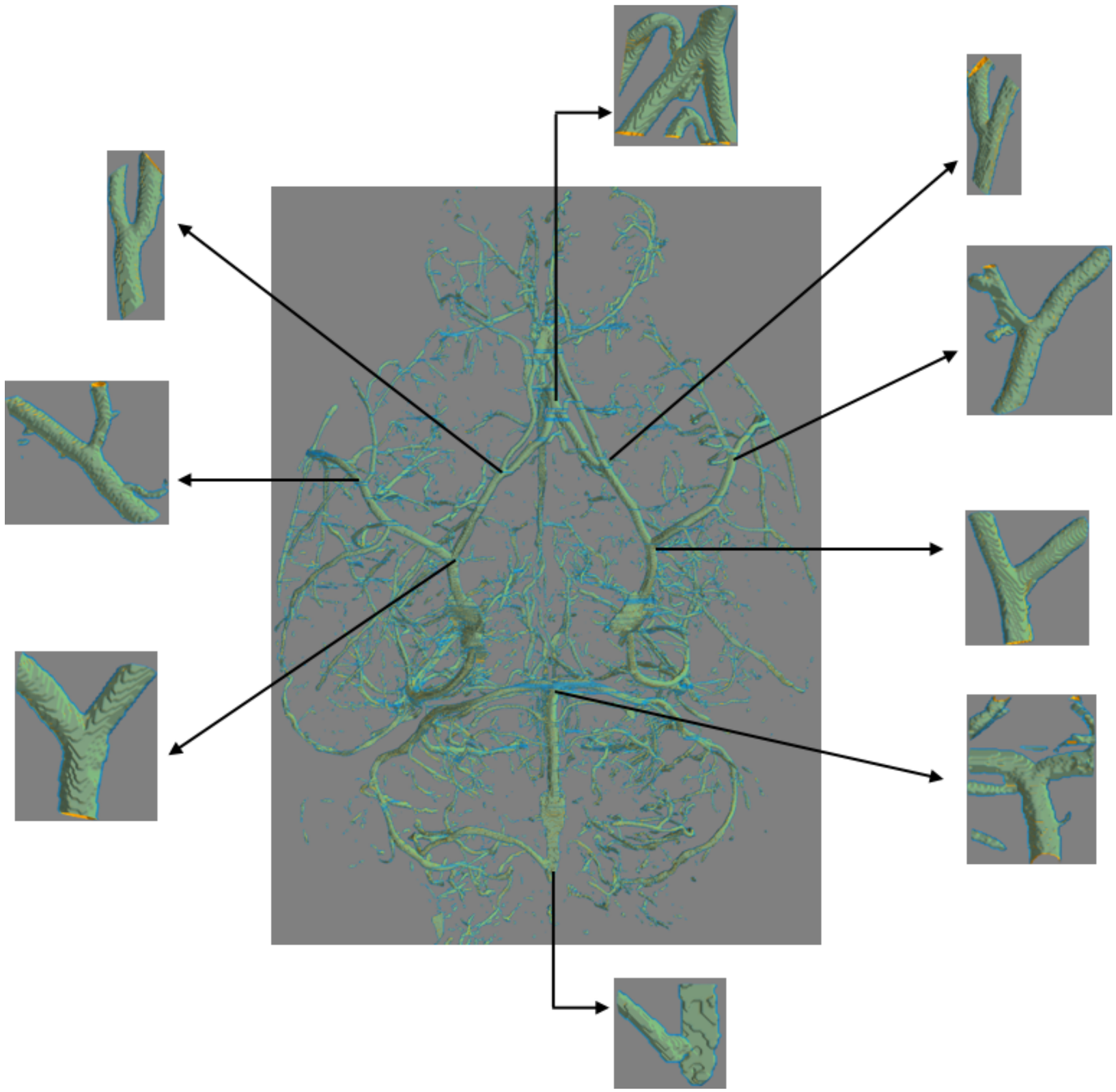


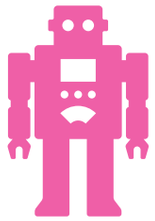


Objective, systematic, repeatable measures ?

Quantifications of cerebral artery bifurcations

Anass Nouri, Florent Autrusseau





- distances ?
- angles ?
- radius ?
- tortuosity ?

Anass Nouri

Quantifications of cerebral artery bifurcations

Anass Nouri, Florent Autrusseau

X coord	Y coord	Z coord	centrers bifs ID	CS (to neighbor1)	CS (to neighbor2)	CS (to neighbor3)	CS (to neighbor4)	neighbor1 ID	neighbor2 ID	neighbor3 ID	neighbor4 ID	FB1	FB2	FB3	Dist. to FB1	Dist. to FB2	Dist. to FB3
176	241	0	3	2.82842712474619	2.9428090415820	4.93223042511211	-1	1	2	4	-1	-1	-1	1	-1	-1	58
212	275	0	4	4.47213595499958	4.8670420531638	5.21895141649746	-1	0	8	3	-1	-1	3	-1	-1	36	-1
186	299	0	7	2.55228474983079	2.276142374915 4	2.94280904158206	-1	24	22	6	-1	-1	-1	-1	-1	-1	-1
195	273	4	8	-1	6	3.21895141649746	-1	10	4	5	-1	4	-1	-1	5	-1	-1
163.66666	230.66666	8.6666666	10	-1	-1	-1	-1	8	17	12	15	3	-1	-1	5	-1	-1
122	240	10	12	-1	-1	3.97874585671244	-1	25	10	18	-1	10	-1	-1	5	-1	-1
50	162	15	15	2.82842712474619	-1	-1	-1	9	10	17	-1	-1	4	-1	-1	3	-1
33.666666	139.66666	25	17	-1	-1	-1	-1	10	26	15	-1	4	-1	-1	4	-1	-1
127	223	28	18	3.82136720504592	6.4944382578492	-1	-1	12	13	30	-1	5	-1	-1	23	-1	-1
111	228	28	19	-1	2	4.15737865166653	-1	26	11	28	-1	11	-1	-1	3	-1	-1
132.5	222.5	34	25	4	4.157378651666 5	-1	-1	20	26	12	-1	-1	11	-1	-1	20	-1
171	225	34	26	-1	-1	4	-1	17	19	25	-1	7	-1	-1	4	-1	-1
134.33333	218.33333	36.333333	28	3.31207919004578	2.9428090415820	4.15737865166653	-1	16	41	19	-1	-1	-1	9	-1	-1	78
124	176	38	29	5.63299316185545	3.8856180831641	4	-1	49	38	31	-1	30	-1	-1	18	-1	-1
173	180	38	30	-1	4.3147573033330	4.15737865166653	-1	18	37	38	-1	8	-1	-1	9	-1	-1
128.66666	225	38.333333	31	4.70966340149732	4.157378651666 5	6.60262336462821	-1	35	29	14	-1	19	-1	-1	27	-1	-1
202	295	38	32	4.61441713185517	4.157378651666 5	5.00932323001944	-1	40	34	35	-1	23	-1	-1	10	-1	-1
138	226	39	33	2.76417624662798	-1	-1	-1	36	37	39	-1	20	-1	-1	29	-1	-1
113	200	41	34	4.15737865166653	2.276142374915 4	-1	-1	32	40	42	-1	16	-1	-1	22	-1	-1
147	229	41	35	4.45703848018865	4.8240453183331	2.2761423749154	-1	32	31	23	-1	16	-1	-1	33	-1	-1
114	239	41	36	2.82842712474619	2.552284749830 7	-1	-1	33	46	39	-1	17	-1	-1	29	-1	-1
84	62	42	37	-1	3.218951416497 4	5.10456949966159	-1	33	60	30	-1	17	-1	-1	7	-1	-1
48	119	42	39	-1	-1	4.47213595499958	-1	33	36	47	-1	17	-1	-1	9	-1	-1
214	185	42	40	3.97874585671244	2.276142374915 4	4.70966340149732	-1	32	34	55	-1	16	-1	-1	10	-1	-1
8	144	43	42	-1	-1	3.64273441009184	-1	34	44	53	45	18	-1	-1	5	-1	-1
134	216	43	43	-1	-1	-1	-1	57	50	38	-1	37	-1	-1	5	-1	-1
143	228	43	44	-1	-1	-1	-1	48	42	55	-1	29	-1	-1	8	-1	-1

Angles1	Angles2	Angles3	Angles4	Angles5	Angles6	Tortuosity to neighbor1	Tortuosity to neighbor2	Tortuosity to neighbor3	Tortuosity to neighbor 4
59.34616592	61.2194009107	59.437684484785	59.3461659263	61.2194009107 3	59.437684484785	0.446666666666667	0.429347826086956	0.316666666666667	-1
96.98249728	36.6992252004	46.9772860610453	96.9824972879	36.69922520048	46.9772860610453	0.447115384615385	0.405637254901961	0.310909090909091	-1
25.52460190	65.5755794768	89.7700811909635	25.5246019011	65.57557947681	89.7700811909635	0.434782608695652	0.381118881118881	0.478835978835979	-1
39.61656823	75.1211354659	72.0797868606077	39.6165682339	75.12113546596	72.0797868606077	-1	0.405637254901961	0.386904761904762	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	0.471875	-1
-1	-1	-1	-1	-1	-1	0.406666666666667	-1	-1	-1
99.81105020	67.3460505207	18.2911379083032	99.8110502024	67.34605052073	18.2911379083032	-1	-1	-1	-1
83.73604728	26.0845415520	74.498640433063	83.7360472800	26.08454155204	74.498640433063	0.471875	0.378787878787879	-1	-1
-1	-1	-1	-1	-1	-1	-1	0.439393939393939	0.401269841269841	-1
61.56310985	83.0175027120	35.8609692198911	61.5631098511	83.01750271208	35.8609692198911	0.328502415458937	0.364705882352941	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	0.364705882352941	-1
56.78908923	66.4017912129	58.2485925838408	56.7890892391	66.40179121297	58.2485925838408	0.425925925925926	0.506060606060606	0.401269841269841	-1
22.57459595	79.8973882083	84.8372987563639	22.5745959504	79.89738820831	84.8372987563639	0.341666666666667	0.25	0.355555555555556	-1

Quantifications of cerebral artery bifurcations

Anass Nouri, Florent Autrusseau

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	X coord	Y coord	Z coord	centrers bifs ID	CS (to neighbor1)	CS (to neighbor2)	CS (to neighbor3)	CS (to neighbor4)	neighbor1 ID	neighbor2 ID	neighbor3 ID	neighbor4 ID	FB1	FB2	FB3	Dist. to FB1	Dist. to FB2	Dist. to FB3
2	176	241	0	3	2.82842712474619	2.9428090415820	4.93223042511211	-1	1	2	4	-1	-1	-1	1	-1	-1	58
3	212	275	0	4	4.47213595499958	4.8670420531638	5.21895141649746	-1	0	8	3	-1	-1	3	-1	-1	36	-1
4	186	299	0	7	2.55228474983079	2.2761423749154	2.94280904158206	-1	24	22	6	-1	-1	-1	-1	-1	-1	-1
5	195	273	4	8	-1	6	3.21895141649746	-1	10	4	5	-1	4	-1	-1	5	-1	-1
6	163.66666	230.66666	8.6666666	10	-1	-1	-1	-1	8	17	12	15	3	-1	-1	5	-1	-1
7	122	240	10	12	-1	-1	3.97874585671244	-1	25	10	18	-1	10	-1	-1	5	-1	-1
8	50	162	15	15	2.82842712474619	-1	-1	-1	9	10	17	-1	-1	4	-1	-1	3	-1
9	33.666666	139.66666	25	17	-1	-1	-1	-1	10	26	15	-1	4	-1	-1	4	-1	-1
10	127	223	28	18	3.82136720504592	6.4944382578492	-1	-1	12	13	30	-1	5	-1	-1	23	-1	-1
11	111	228	28	19	-1	2	4.15737865166653	-1	26	11	28	-1	11	-1	-1	3	-1	-1
12	132.5	222.5	34	25	4	4.1573786516665	-1	-1	20	26	12	-1	-1	11	-1	-1	20	-1
13	171	225	34	26	-1	-1	4	-1	17	19	25	-1	7	-1	-1	4	-1	-1
14	134.33333	218.33333	36.333333	28	3.31207919004578	2.9428090415820	4.15737865166653	-1	16	41	19	-1	-1	-1	9	-1	-1	78
15	124	176	38	29	5.63299316185545	3.8856180831641	4	-1	49	38	31	-1	30	-1	-1	18	-1	-1
16	173	180	38	30	-1	4.3147573033330	4.15737865166653	-1	18	37	38	-1	8	-1	-1	9	-1	-1
17	128.66666	225	38.333333	31	4.70966340149732	4.1573786516665	6.60262336462821	-1	35	29	14	-1	19	-1	-1	27	-1	-1
18	202	295	28	32	4.61441713185517	4.1573786516	5.0093253001944	-1	40	34	35	-1	23	-1	-1	10	-1	-1
19	138	226	19	[illegible]	[illegible]	[illegible]	[illegible]	-1	36	37	39	-1	20	-1	-1	29	-1	-1
20	113	200	41	[illegible]	[illegible]	[illegible]	[illegible]	-1	32	40	42	-1	16	-1	-1	22	-1	-1
21	147	229	41	35	4.4570384[illegible]18865	4.8240453183331	2.2761423749154	-1	32	31	23	-1	16	-1	-1	33	-1	-1
22	114	239	41	36	2.82842712474619	2.5522847498307	-1	-1	33	46	39	-1	17	-1	-1	29	-1	-1
23	84	62	42	37	-1	3.2189514164974	5.10456949966159	-1	33	60	30	-1	17	-1	-1	7	-1	-1
24	48	119	42	39	-1	[illegible]	[illegible]	-1	32	36	47	-1	17	-1	-1	9	-1	-1
25	214	185	42	40	3.9787458[illegible]7124	2.2[illegible]	[illegible]	32	32	3[illegible]	55	-1	16	-1	-1	10	-1	-1
26	8	144	43	42	-1	-1	3[illegible]	34	34	44	53	45	18	-1	-1	5	-1	-1
27	134	216	43	43	-1	-1	-1	-1	57	50	38	-1	37	-1	-1	5	-1	-1
28	143	228	43	44	-1	-1	-1	-1	48	42	55	-1	29	-1	-1	8	-1	-1

Angles1	Angles2	Angles3	Angles4	Angles5	Angles6	Tortuosity to neighbor1	Tortuosity to neighbor2	Tortuosity to neighbor3	Tortuosity to neighbor 4
59.34616592	61.2194009107	59.437684484785	59.3461659263	61.2194009107	[illegible]	[illegible]666667	0.429347826086956	0.316666666666667	-1
96.98249728	36.6992252004	46.9772860610453	96.9824972879	36.6992252004	46.9772860610453	0.447118[illegible]5385	0.405637254901961	0.310909090909091	-1
25.52460190	65.5755794768	89.7700811909635	25.5246019011	65.5755794768	89.7700811909635	0.434782608695652	0.381118881118881	0.478835978835979	-1
39.61656823	75.1211354659	72.0797868606077	39.6165682339	75.12113546596	72.0797868606077	-1	0.405637254901961	0.386904761904762	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	0.471875	-1
-1	-1	-1	-1	-1	-1	0.406666666666667	-1	-1	-1
99.81105020	67.3460505207	18.2911379083032	99.8110502024	67.3460505207	18.2911379083032	-1	-1	-1	-1
83.73604728	26.0845415520	74.498640433063	83.7360472800	26.08454155204	74.498640433063	0.471875	0.378787878787879	-1	-1
-1	-1	-1	-1	-1	-1	-1	0.439393939393939	0.401269841269841	-1
61.56310985	83.0175027120	35.8609692198911	61.5631098511	83.01750271208	35.8609692198911	0.328502415458937	0.364705882352941	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	0.364705882352941	-1
56.78908923	66.4017912129	58.2485925838408	56.7890892391	66.40179121297	58.2485925838408	0.425925925925926	0.506060606060606	0.401269841269841	-1
22.57459595	79.8973882083	84.8372987563639	22.5745959504	79.89738820831	84.8372987563639	0.341666666666667	0.25	0.355555555555556	-1

interpretating these values ?

comparing them ?

sharing?

J Digit Imaging (2013) 26:630–641
DOI 10.1007/s10278-013-9599-2

Quantitative Imaging Biomarker Ontology (QIBO) for Knowledge Representation of Biomedical Imaging Biomarkers

Andrew J. Buckler · M. Ouellette · J. Danagoulian · G. Wernsing · Tiffany Ting Liu · Erica Savig · Baris E. Suzek · Daniel L. Rubin · David Paik

Published online: 16 April 2013
© Society for Imaging Informatics in Medicine 2013

J Digit Imaging (2013) 26:630–641
DOI 10.1007/s10278-013-9599-2

Quantitative Imaging Biomarker Ontology (QIBO) for Knowledge Representation of Biomedical Imaging Biomarkers

Andrew J. Buckler · M. Ouellette · J. Danagoulian · G. Wernsing · Tiffany Ting Liu · Erica Savig · Baris E. Suzek · Daniel L. Rubin · David Paik

Published online: 16 April 2013
© Society for Imaging Informatics in Medicine 2013

better **interpret** imaging measures

better **reuse** imaging measures

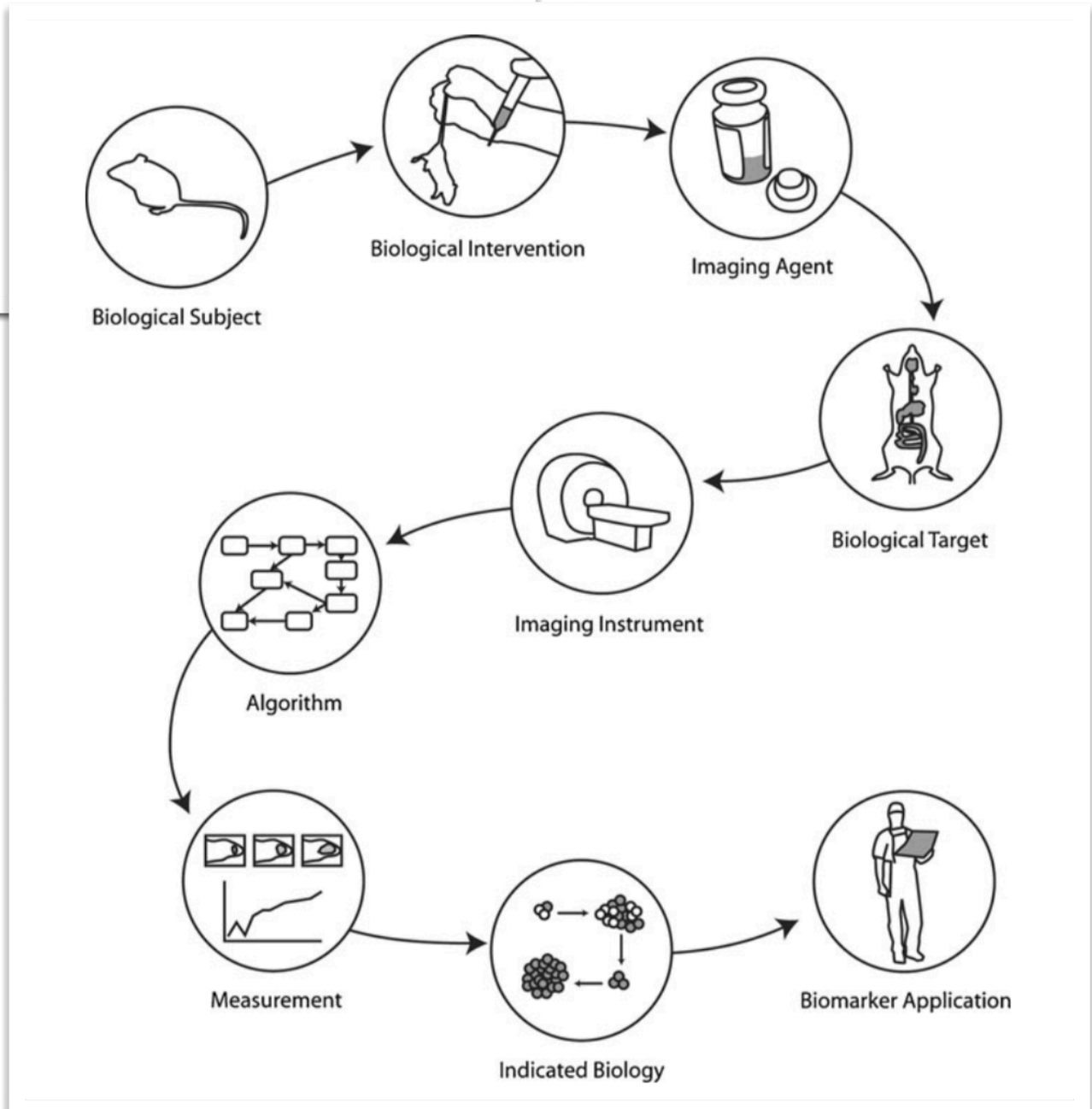
cross-link imaging measures

J Digit Imaging (2013) 26:630–641
DOI 10.1007/s10278-013-9599-2

Quantitative Imaging Biomarker Ontology (QIBO) for Knowledge Representation of Biomedical Imaging Biomarkers

Andrew J. Buckler · M. Ouellette · J. Danagoulian · G. Wernsing · Tiffany Ting Liu · Erica Savig · Baris E. Suzek · Daniel L. Rubin · David Paik

Published online: 16 April 2013
© Society for Imaging Informatics in Medicine 2013



better **interpret** imaging measures

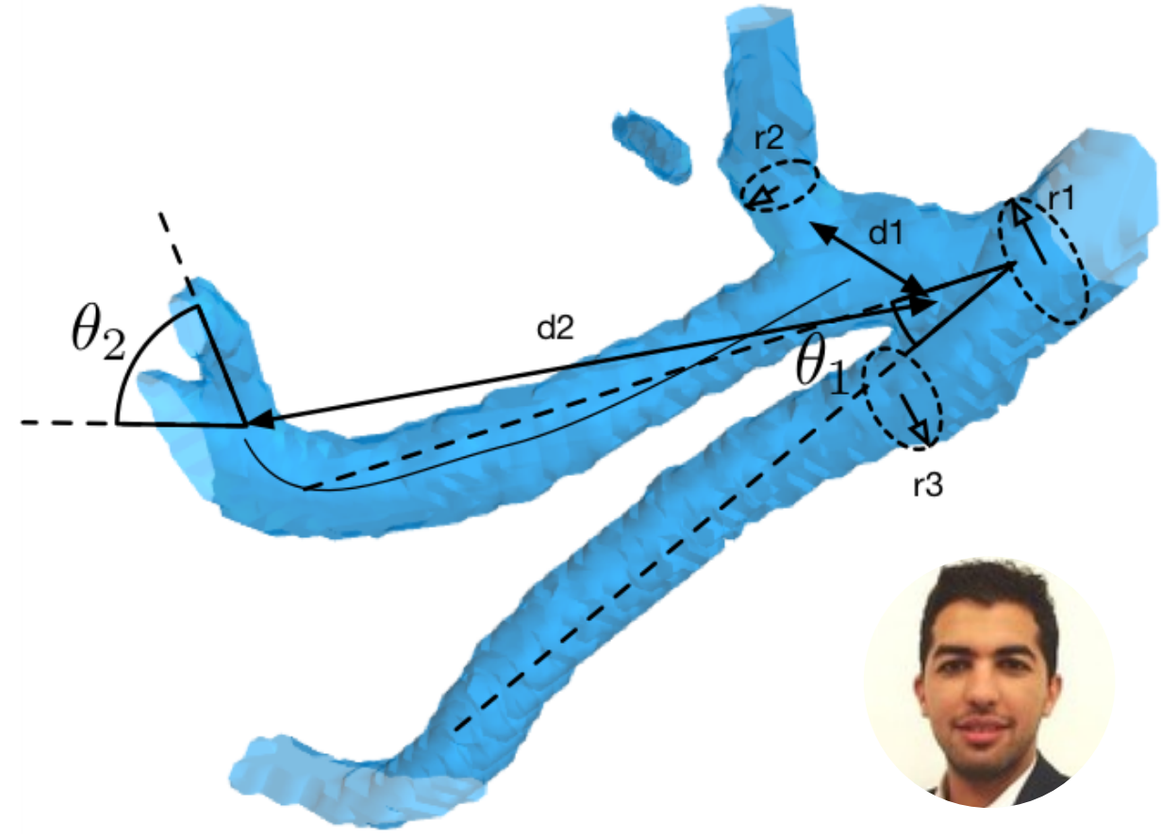
better **reuse** imaging measures

cross-link imaging measures

d_1, d_2 : **distances** between **bifurcations**

r_1, r_2, r_3 : **radius** of vessels

θ_1, θ_2 : **angles** of bifurcations

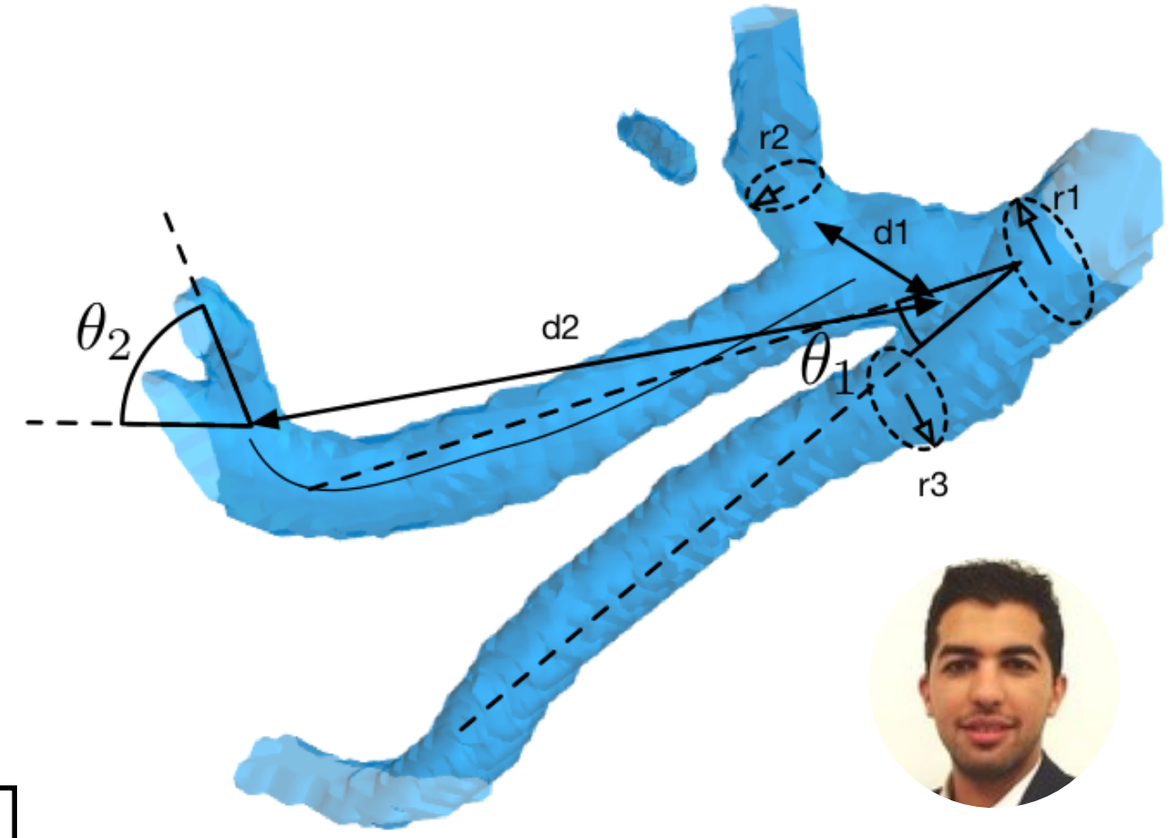


Anass Nouri

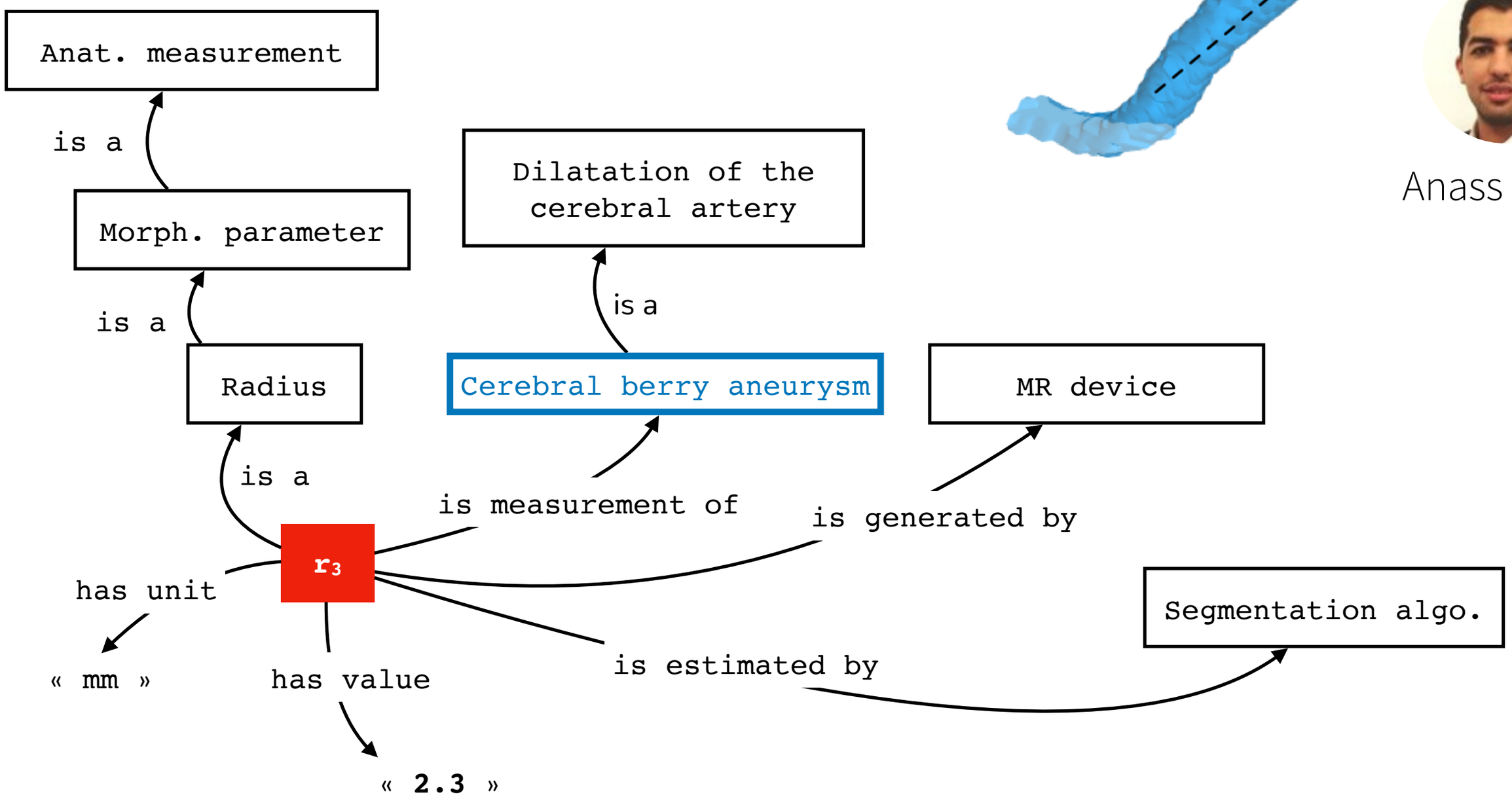
d_1, d_2 : **distances** between **bifurcations**

r_1, r_2, r_3 : **radius** of vessels

θ_1, θ_2 : **angles** of bifurcations

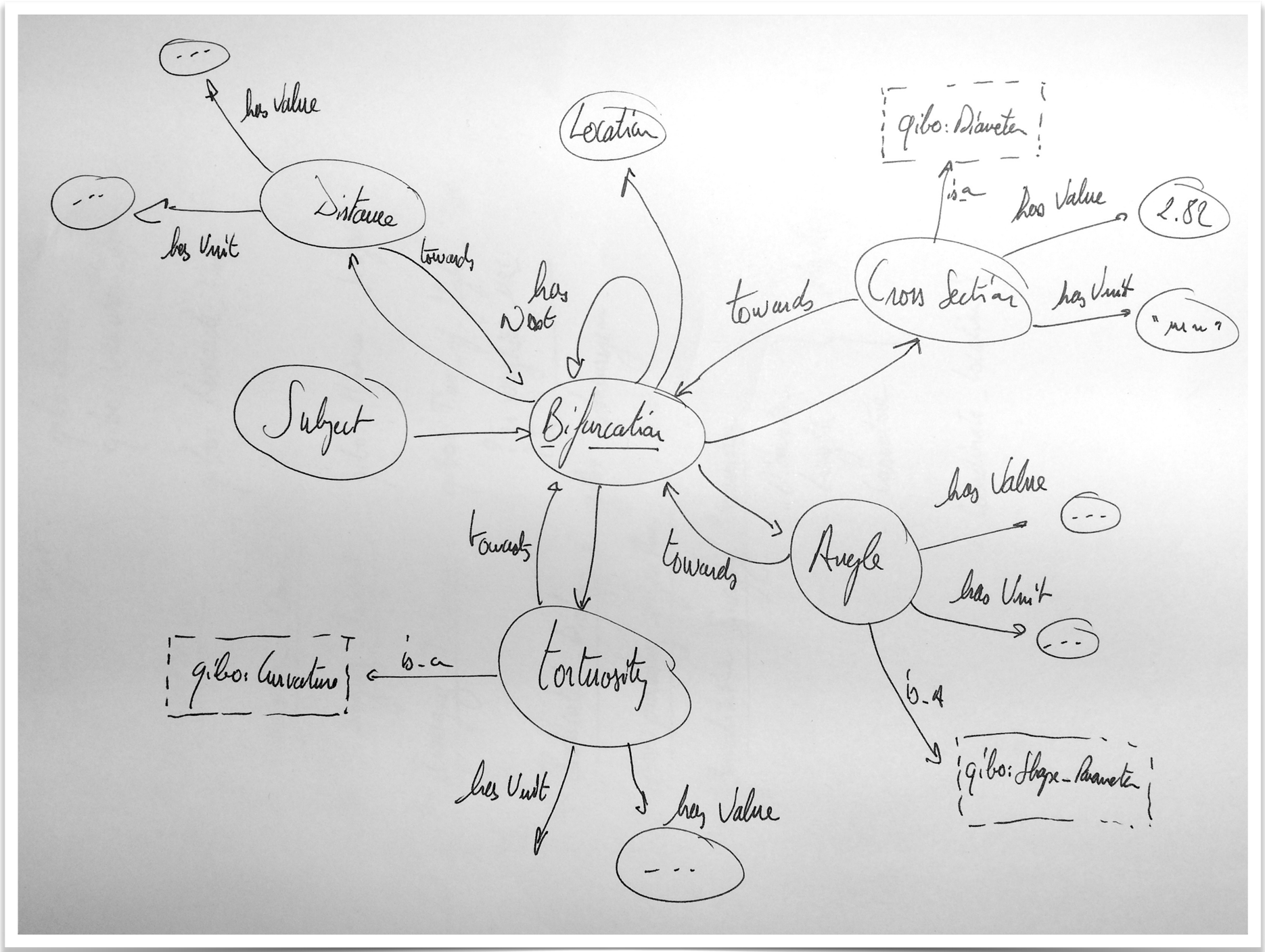


Anass Nouri

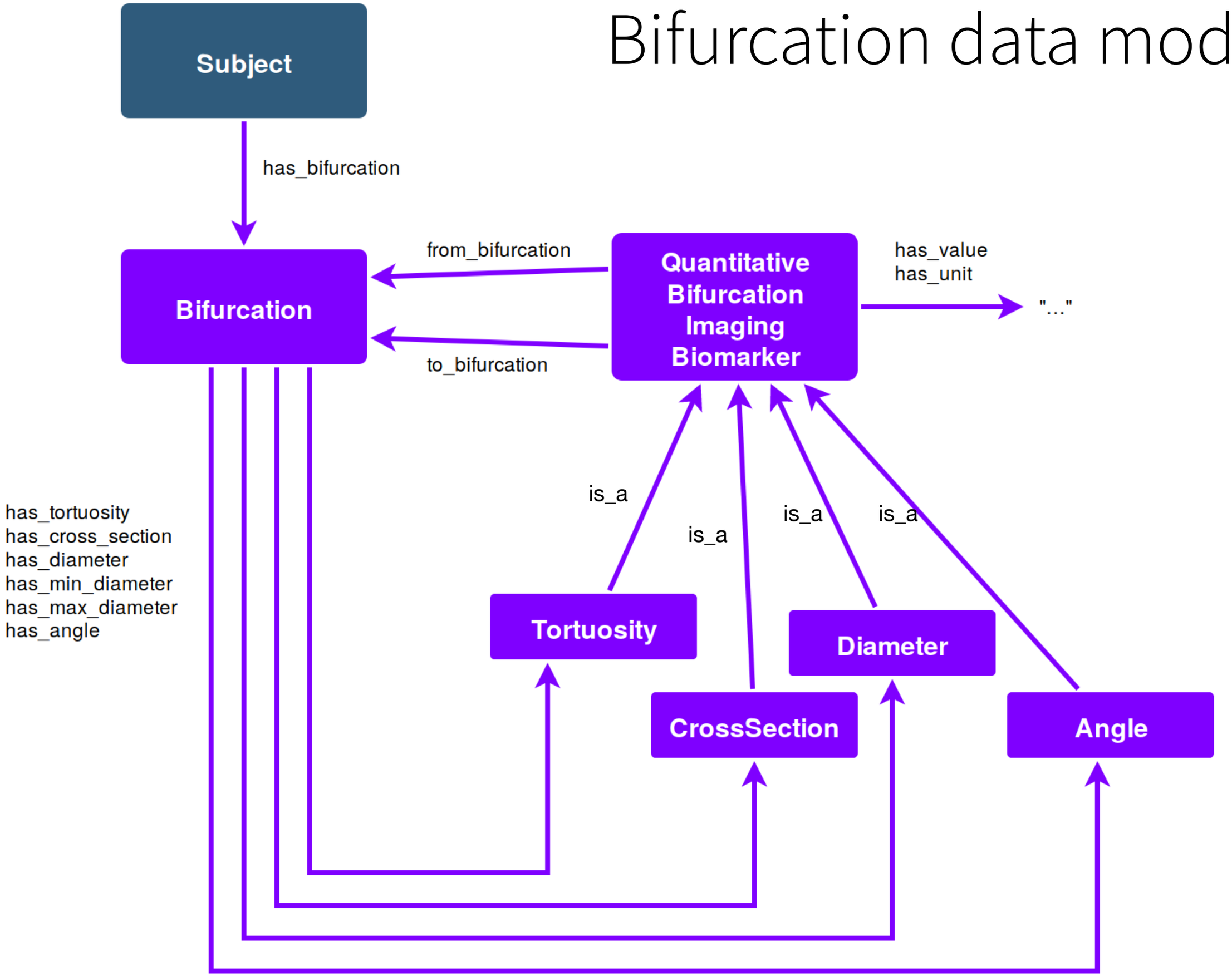


BrainHack project

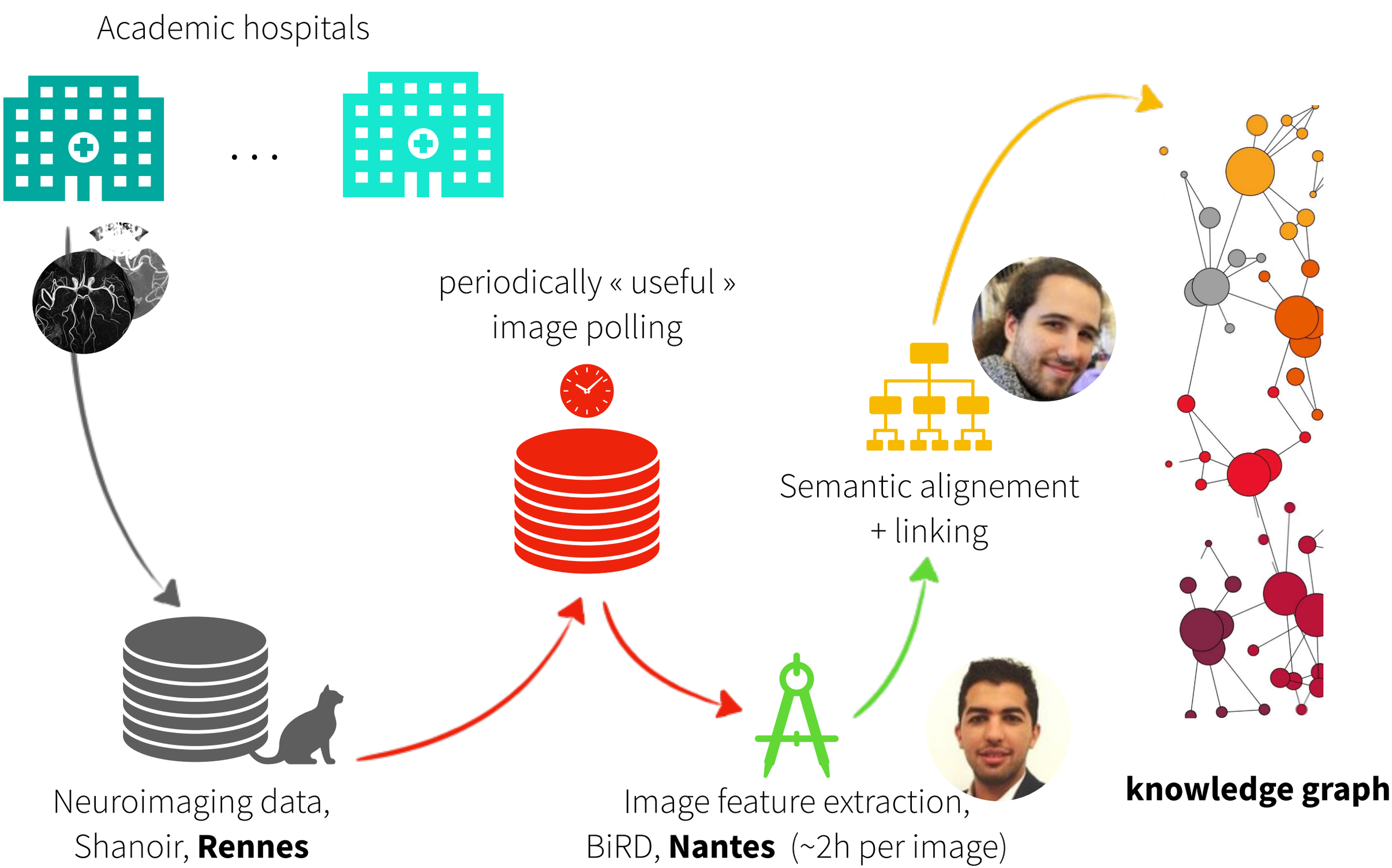
« From MS excel sheets to semantic bioimaging markers : representing and querying cerebral vascular measures with the QIBO ontology »



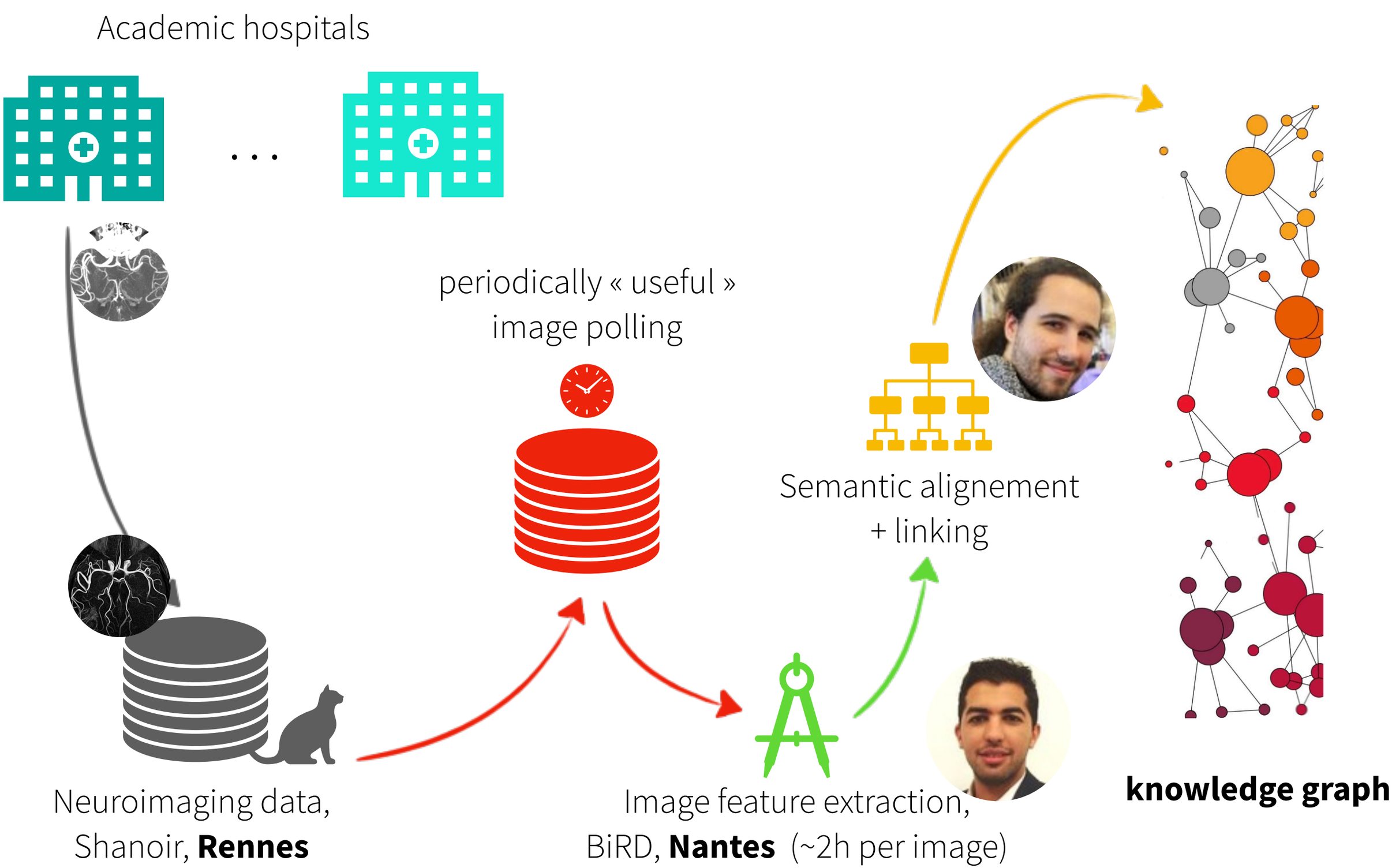
Bifurcation data model



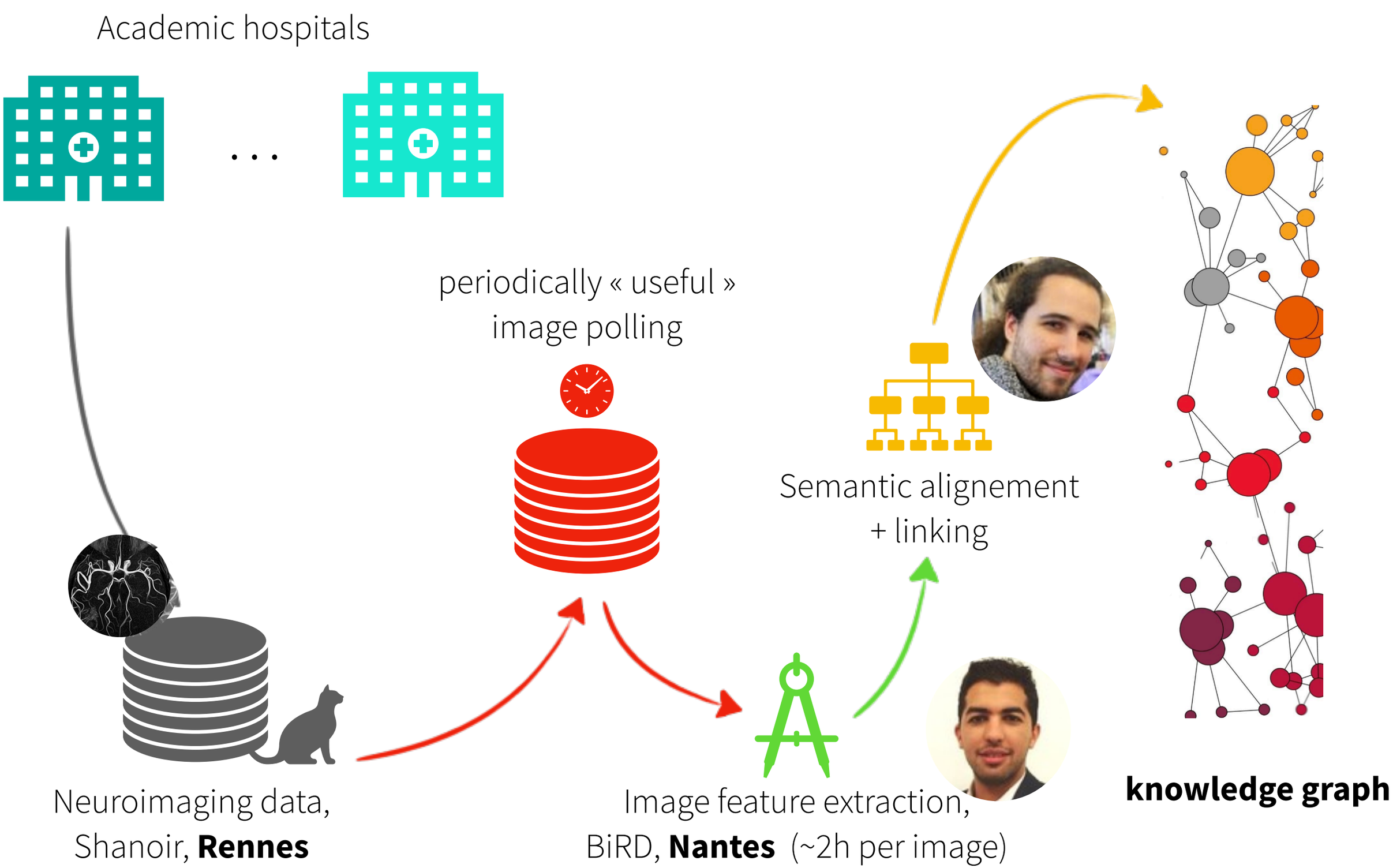
From Images to Knowledge Graph



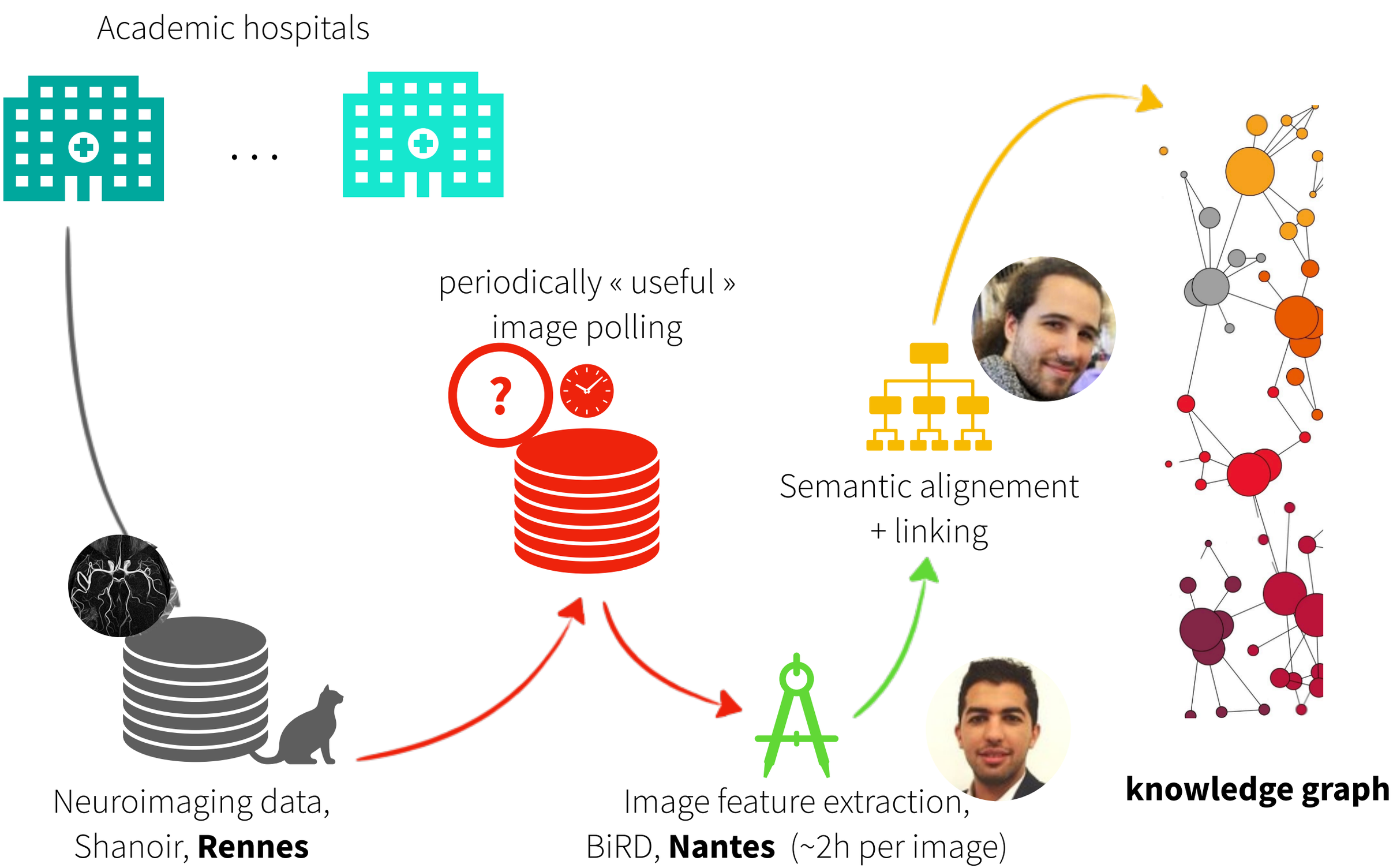
From Images to Knowledge Graph



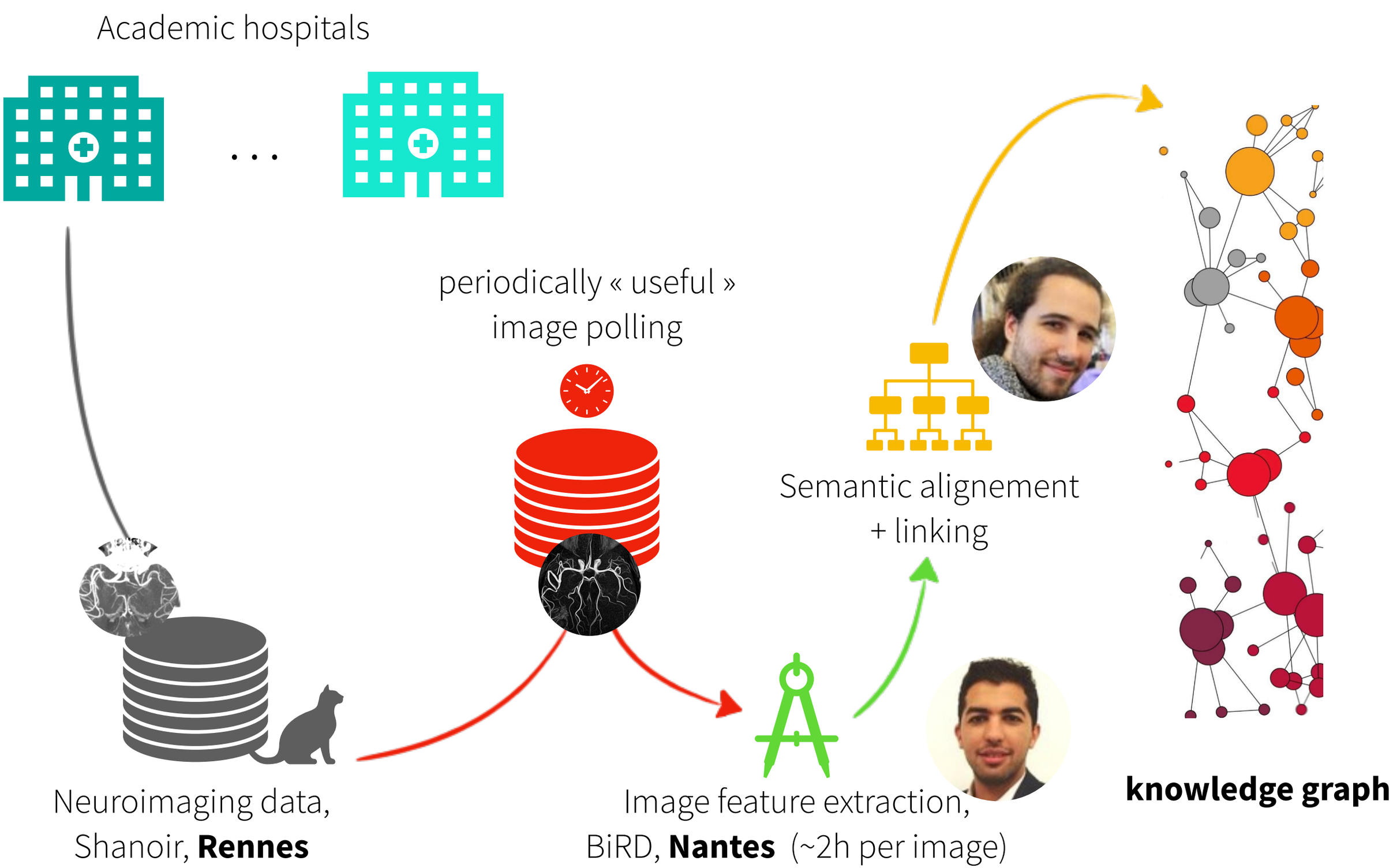
From Images to Knowledge Graph



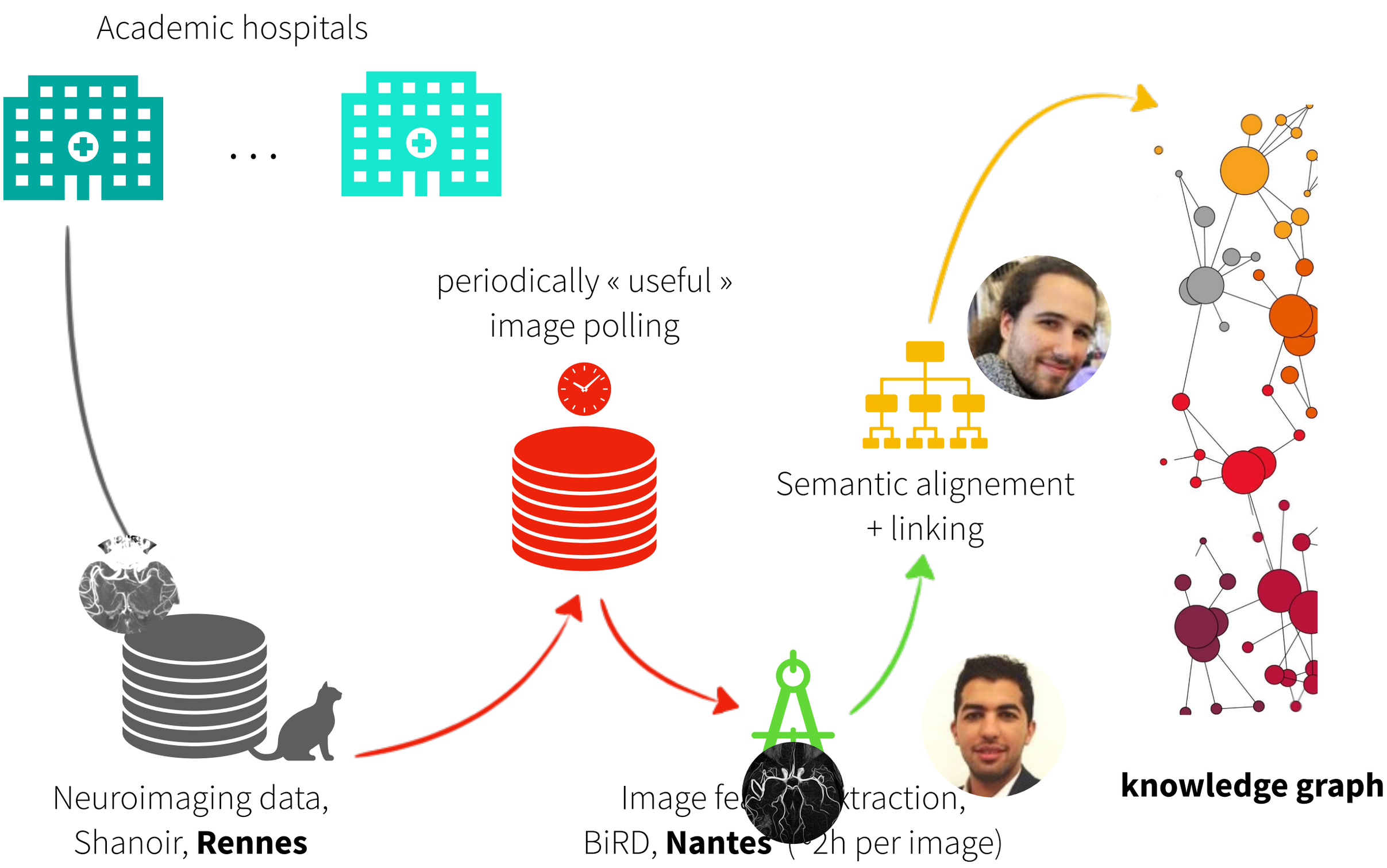
From Images to Knowledge Graph



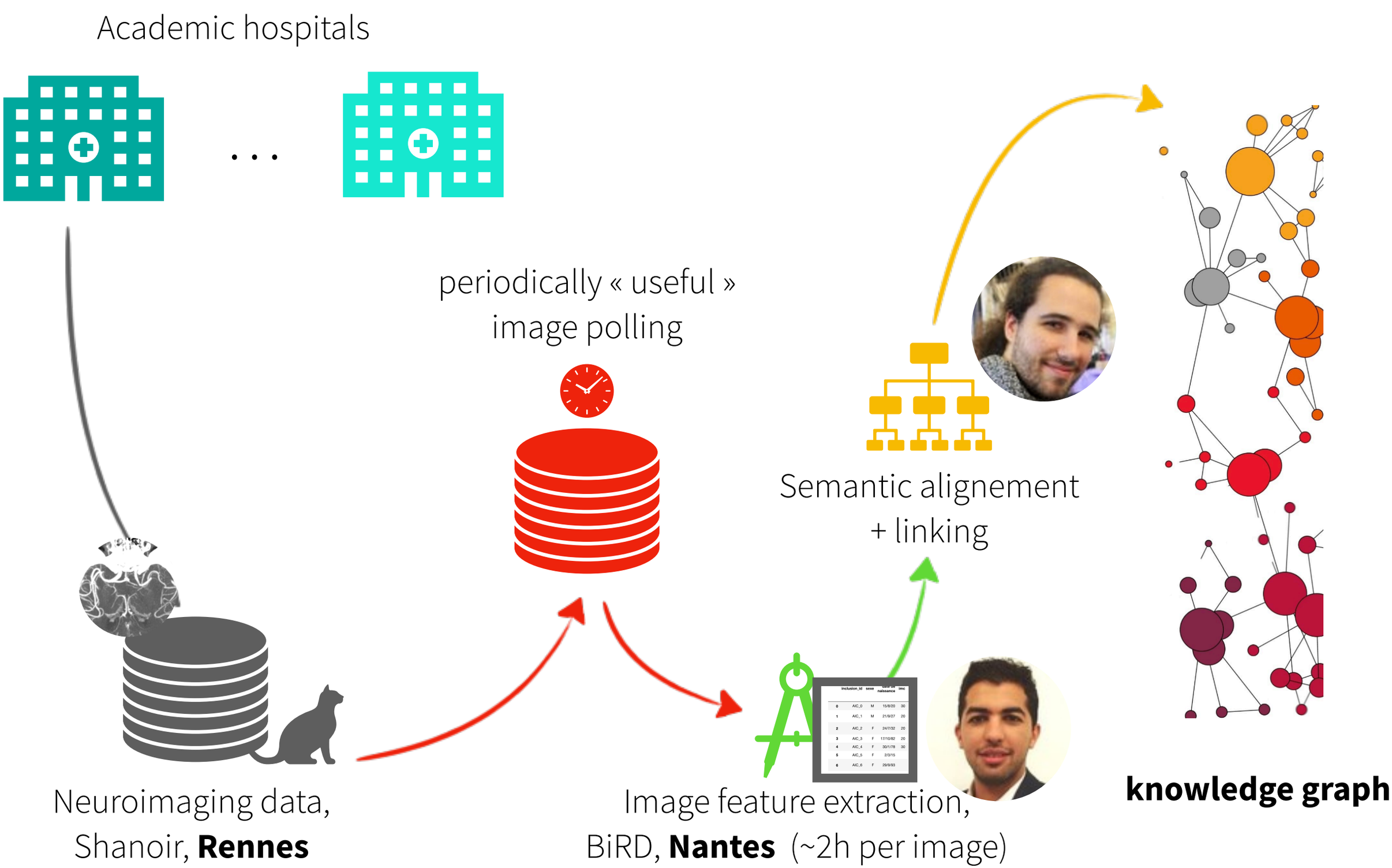
From Images to Knowledge Graph



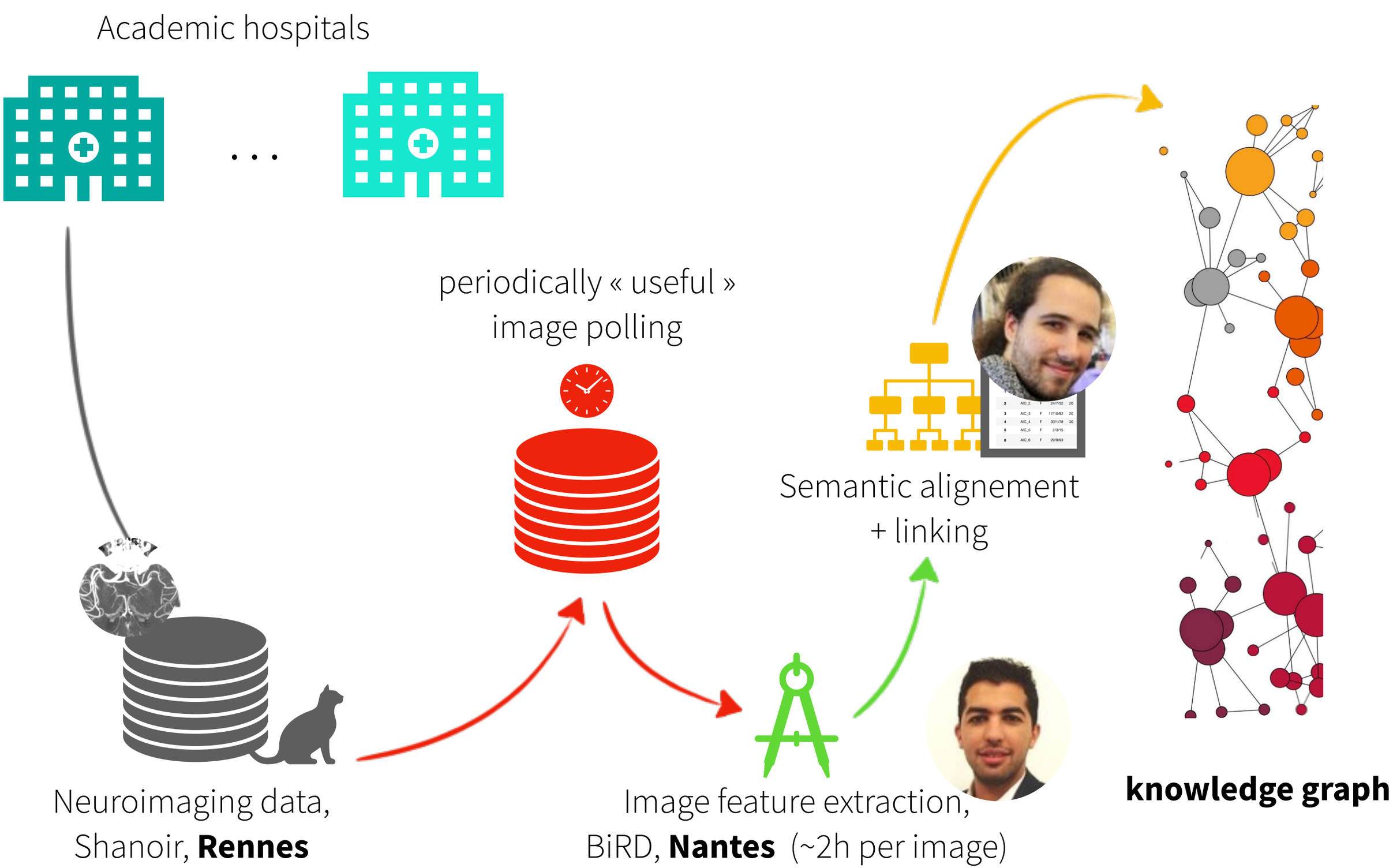
From Images to Knowledge Graph



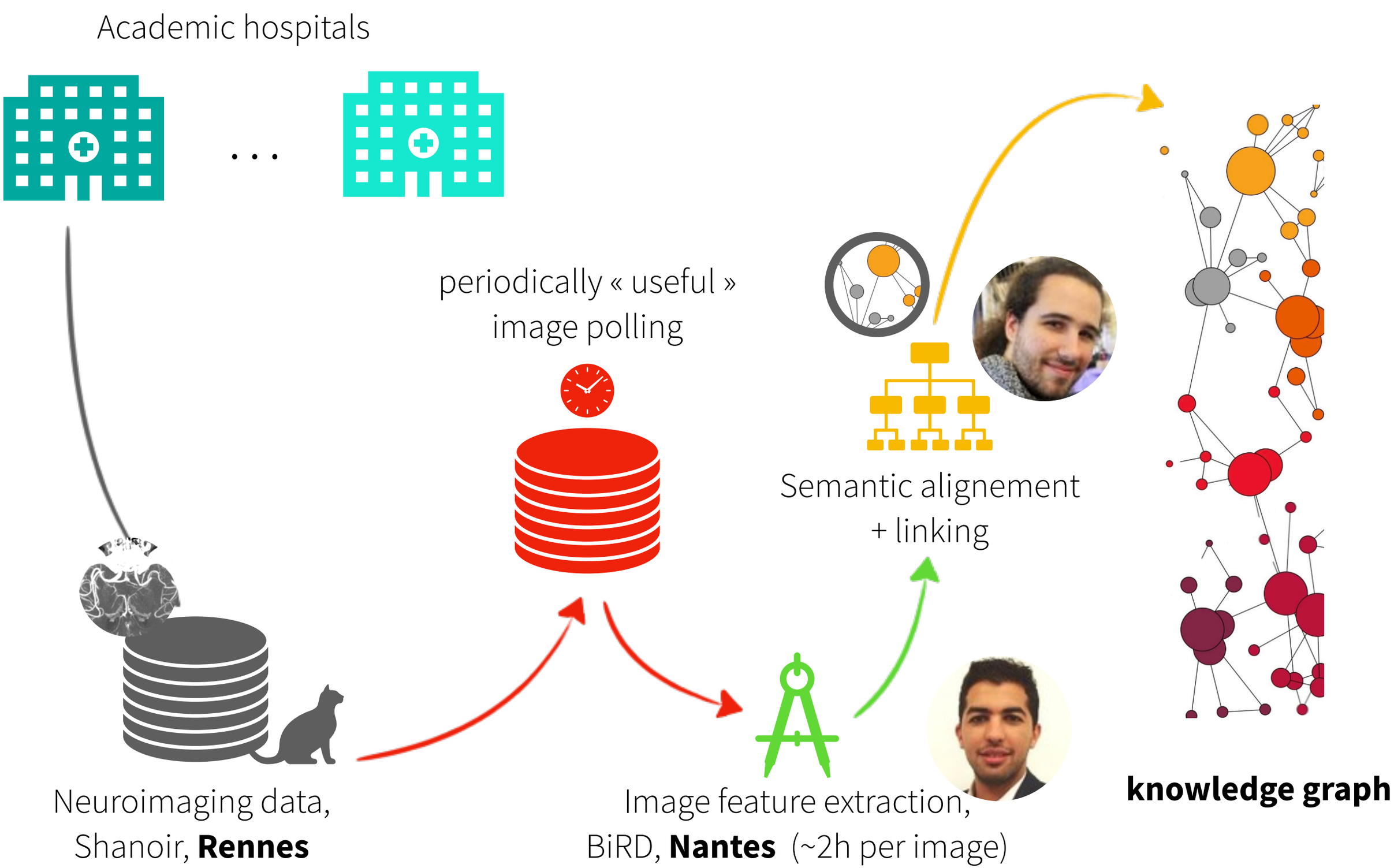
From Images to Knowledge Graph



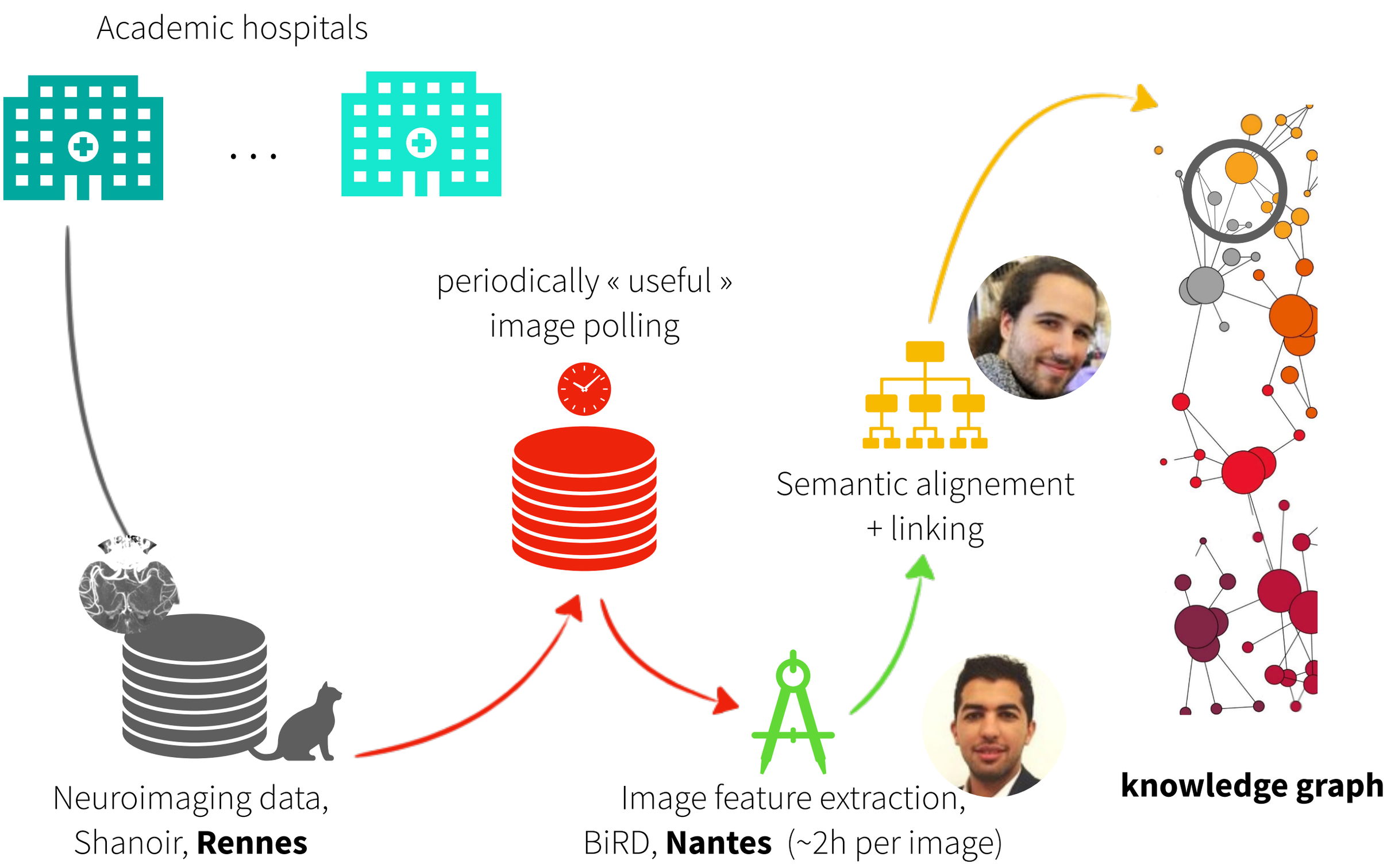
From Images to Knowledge Graph



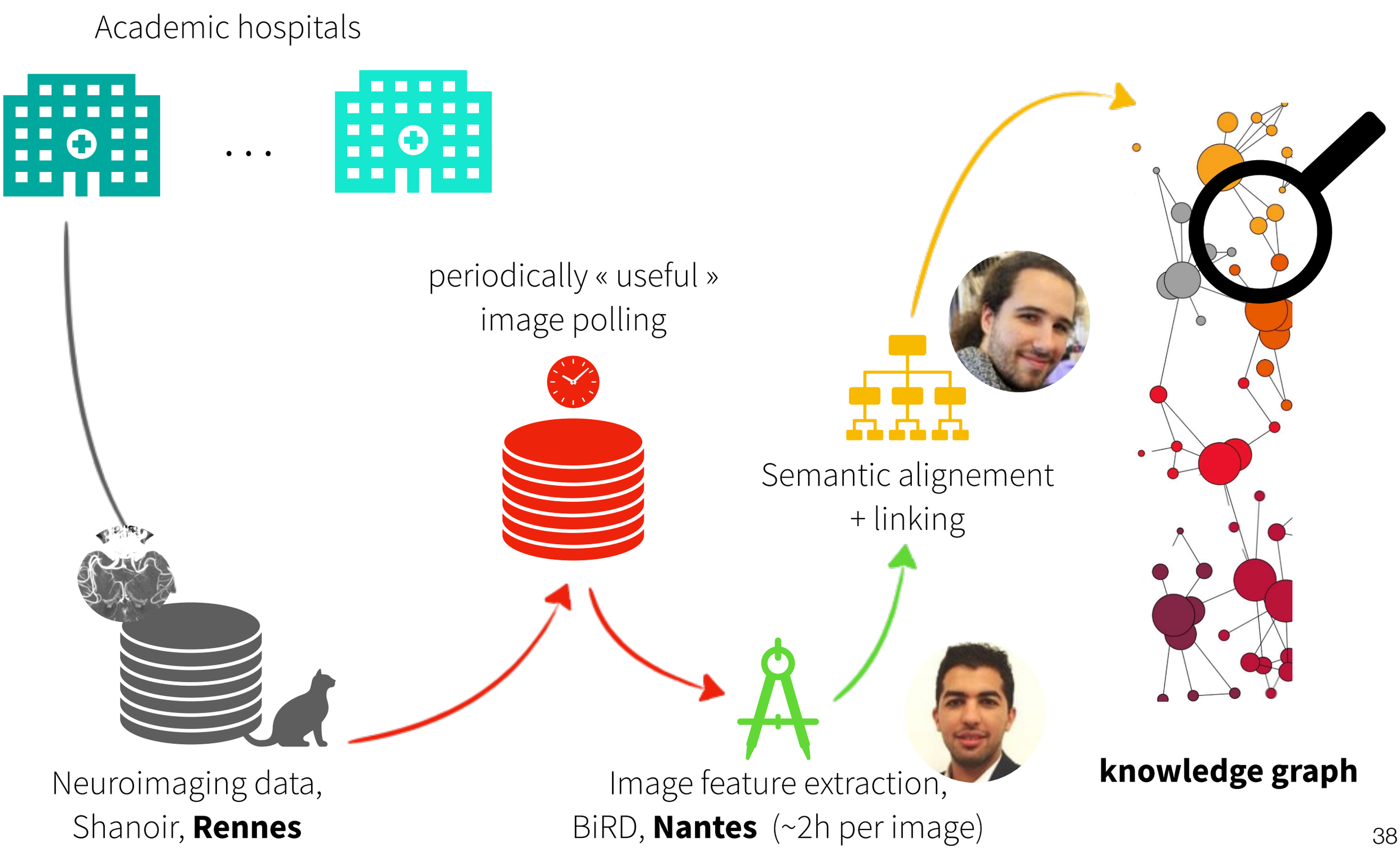
From Images to Knowledge Graph



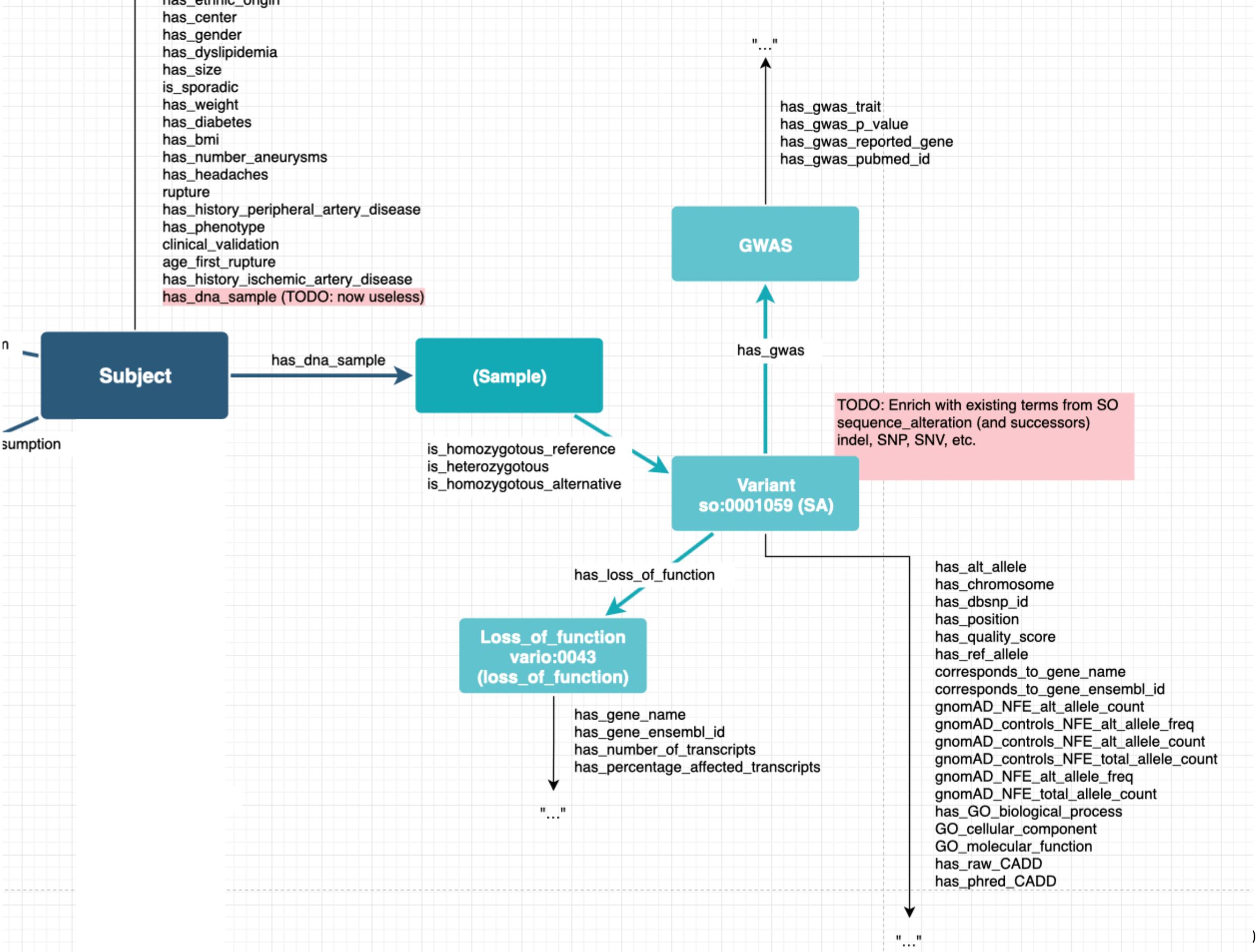
From Images to Knowledge Graph

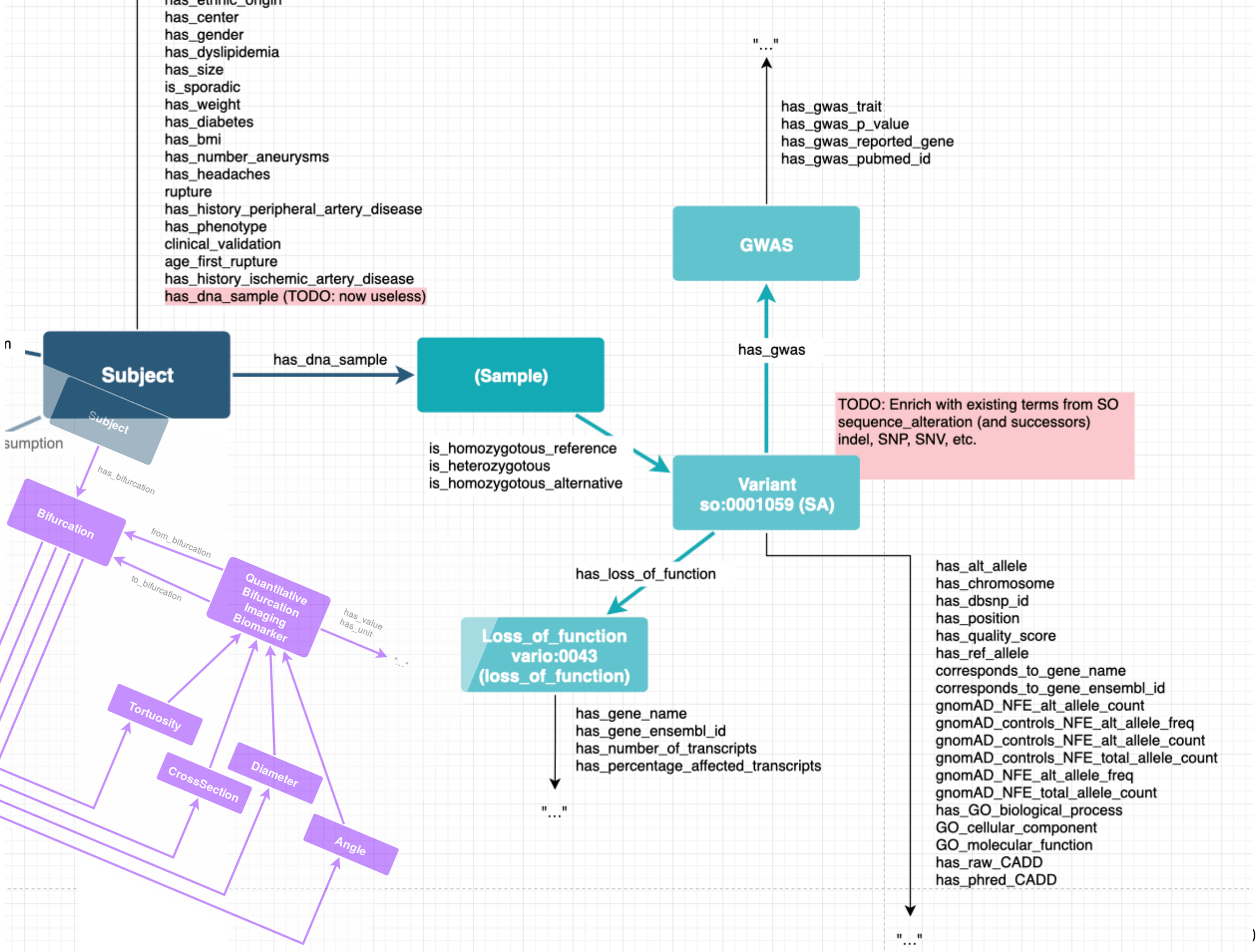


From Images to Knowledge Graph



Genomic Data


has_ethnic_origin
has_center
has_gender
has_dyslipidemia
has_size
is_sporadic
has_weight
has_diabetes
has_bmi
has_number_aneurysms
has_headaches
rupture
has_history_peripheral_artery_disease
has_phenotype
clinical_validation
age_first_rupture
has_history_ischemic_artery_disease
has_dna_sample (TODO: now useless)
Subject
has_dna_sample
(Sample)
sumption
is_homozygotous_reference
is_heterozygotous
is_homozygotous_alternative
Variant
so:0001059 (SA)
"..."
has_gwas_trait
has_gwas_p_value
has_gwas_reported_gene
has_gwas_pubmed_id
GWAS
has_gwas
TODO: Enrich with existing terms from SO
sequence_alteration (and successors)
indel, SNP, SNV, etc.
has_loss_of_function
Loss_of_function
vario:0043
(loss_of_function)
has_gene_name
has_gene_ensembl_id
has_number_of_transcripts
has_percentage_affected_transcripts
"..."
has_alt_allele
has_chromosome
has_dbsnp_id
has_position
has_quality_score
has_ref_allele
corresponds_to_gene_name
corresponds_to_gene_ensembl_id
gnomAD_NFE_alt_allele_count
gnomAD_controls_NFE_alt_allele_freq
gnomAD_controls_NFE_alt_allele_count
gnomAD_controls_NFE_total_allele_count
gnomAD_NFE_alt_allele_freq
gnomAD_NFE_total_allele_count
has_GO_biological_process
GO_cellular_component
GO_molecular_function
has_raw_CADD
has_phred_CADD
"..."


has_ethnic_origin
has_center
has_gender
has_dyslipidemia
has_size
is_sporadic
has_weight
has_diabetes
has_bmi
has_number_aneurysms
has_headaches
rupture
has_history_peripheral_artery_disease
has_phenotype
clinical_validation
age_first_rupture
has_history_ischemic_artery_disease
has_dna_sample (TODO: now useless)
Subject
has_dna_sample
(Sample)
"..."
has_gwas_trait
has_gwas_p_value
has_gwas_reported_gene
has_gwas_pubmed_id
GWAS
has_gwas
TODO: Enrich with existing terms from SO
sequence_alteration (and successors)
indel, SNP, SNV, etc.
is_homozygotous_reference
is_heterozygotous
is_homozygotous_alternative
Variant
so:0001059 (SA)
has_loss_of_function
Loss_of_function
vario:0043
(loss_of_function)
has_gene_name
has_gene_ensembl_id
has_number_of_transcripts
has_percentage_affected_transcripts
"..."
has_alt_allele
has_chromosome
has_dbsnp_id
has_position
has_quality_score
has_ref_allele
corresponds_to_gene_name
corresponds_to_gene_ensembl_id
gnomAD_NFE_alt_allele_count
gnomAD_controls_NFE_alt_allele_freq
gnomAD_controls_NFE_alt_allele_count
gnomAD_controls_NFE_total_allele_count
gnomAD_NFE_alt_allele_freq
gnomAD_NFE_total_allele_count
has_GO_biological_process
GO_cellular_component
GO_molecular_function
has_raw_CADD
has_phred_CADD
"..."
Subject
has_bifurcation
Bifurcation
from_bifurcation
to_bifurcation
Quantitative Bifurcation Imaging Biomarker
has_value
has_unit
"..."
Tortuosity
CrossSection
Diameter
Angle

Querying and **Reasoning** on this Knowledge Graph

Queries with graph patterns

```
SELECT * WHERE {
  # Search variants
  ?subject ican:has_dna_sample ?sample .
  ?sample ican:is_heterozygotous ?variant .
  ?variant ican:has_chromosome "19" ;
           ican:has_position ?position ;
           ican:has_alt_allele ?alt_allele ;
           ican:has_ref_allele ?ref_allele .
  FILTER (?position > 10092337 && ?position < 10106407)
  # dbSNP identifier
  OPTIONAL { ?variant ican:has_dbsnp_id ?dbSNPid . }
  # First aneurysm rupture: age and localisation
  OPTIONAL { ?subject ican:age_first_rupture ?ageRupt . }
  OPTIONAL { ?subject ican:loc_first_rupture ?locRupt . }
}
```

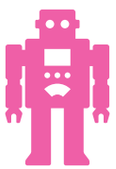
Reasoning with ontologies

Quels sont les patients atteints de **HP:0002813 (“Abnormality of limb bone morphology”)** ?

	Subject	HPO term	Label
0	r18	HP_0010442	Polydactyly
1	r18	HP_0001830	Postaxial foot polydactyly
2	r18	HP_0200055	Small hand
3	r18	HP_0004209	Clinodactyly of the 5th finger
4	r18	HP_0001162	Postaxial hand polydactyly
5	r9	HP_0001162	Postaxial hand polydactyly
6	r9	HP_0010442	Polydactyly
7	r9	HP_0001830	Postaxial foot polydactyly
8	r11	HP_0001830	Postaxial foot polydactyly
9	r11	HP_0001162	Postaxial hand polydactyly
10	r11	HP_0010442	Polydactyly

Recap

Knowledge Graphs & Linked Data, **why ?**

- the gaps
- Something in the middle, between 
- « Artificial intelligence » (knowledge representation, reasoning)
- Make sense out of data diversity ➞ better understand aneurysms

how ?

- **RDF** to represent statements (knowledge graphs)
- **Ontologies** as shared vocabularies / conceptuals models
  - **QIBO**: quantitative imaging biomarkers ontology
  - **Bioportal**: inter-linked life science ontologies

Next episode ?

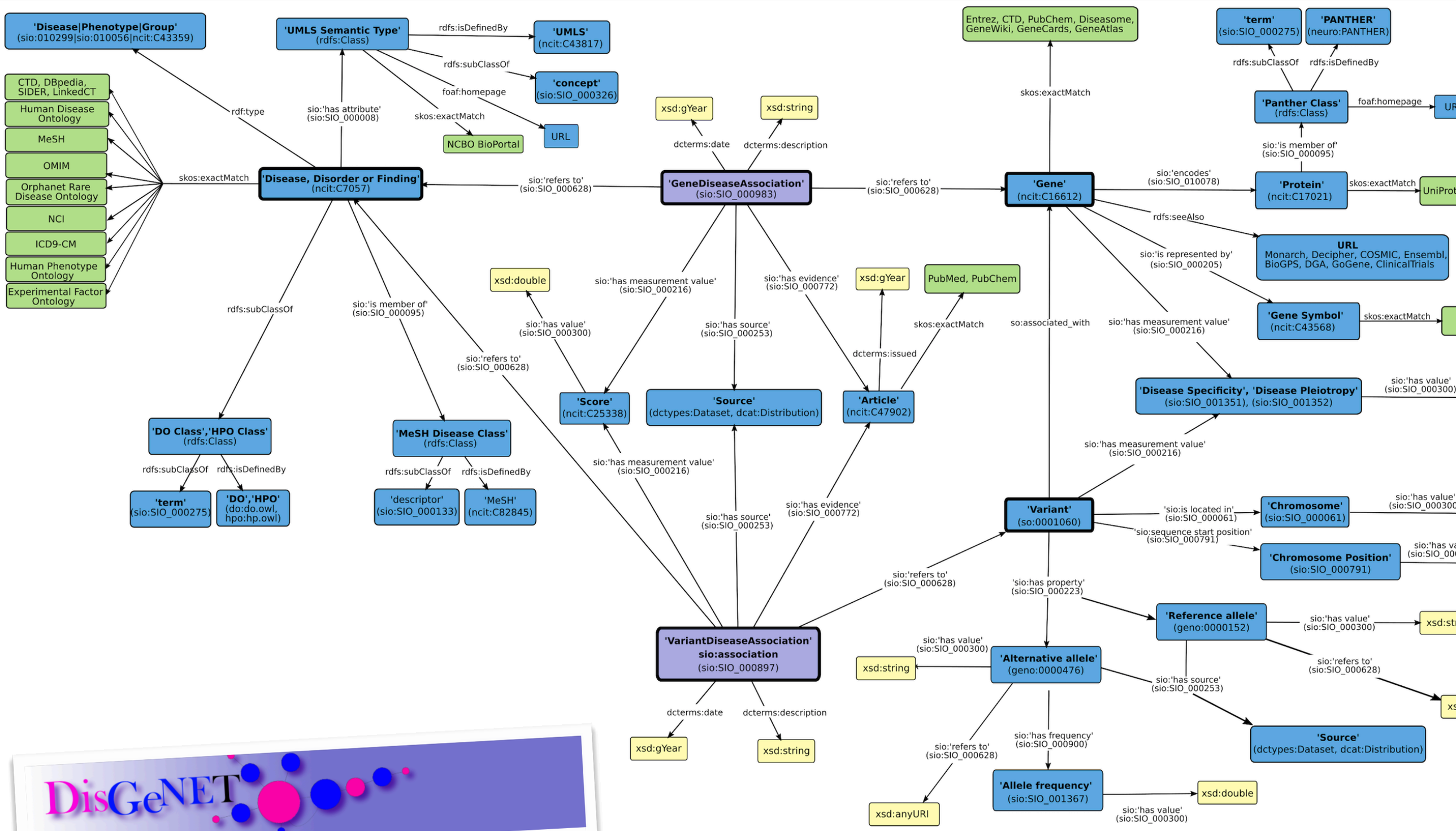
- RDFization of your data (towards FAIRification)
- Hosting your data through a SPARQL endpoint
- Advanced SPARQL querying (property path expressions, nested queries, aggregates)
- Designing controlled vocabularies
- Designing formal ontologies
- Logical reasoning (Inference rules, RDFS entailments, OWL entailments)
- etc.

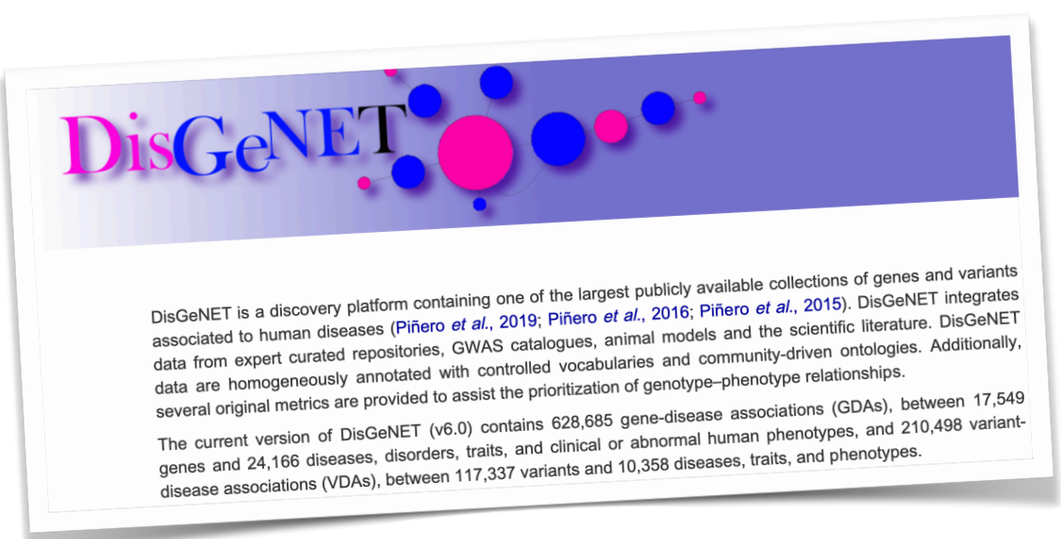
Web resources

- **Standards**
  - https://www.w3.org/TR/rdf11-primer/
  - https://www.w3.org/TR/sparql11-query/
  - https://www.slideshare.net/LeeFeigenbaum/sparql-cheat-sheet
- **Applications** to Life Sciences
  - http://www.biomedbridges.eu/trainings/knowledge-exchange-workshop-resource-description-framework-rdf
  - http://www.wikipathways.org/index.php/Help:WikiPathways_Sparql_queries
- **Coding**
  - Java: Corese (https://project.inria.fr/corese/), Apache Jena (https://jena.apache.org)
  - Python: RDFlib (https://rdflib.readthedocs.io)

Thanks you for your attention

contact: alban.gaignard@univ-nantes.fr





DisGeNET is a discovery platform containing one of the largest publicly available collections of genes and variants associated to human diseases (Piñero *et al.*, 2019; Piñero *et al.*, 2016; Piñero *et al.*, 2015). DisGeNET integrates data from expert curated repositories, GWAS catalogues, animal models and the scientific literature. DisGeNET data are homogeneously annotated with controlled vocabularies and community-driven ontologies. Additionally, several original metrics are provided to assist the prioritization of genotype–phenotype relationships.

The current version of DisGeNET (v6.0) contains 628,685 gene-disease associations (GDAs), between 17,549 genes and 24,166 diseases, disorders, traits, and clinical or abnormal human phenotypes, and 210,498 variant-disease associations (VDAs), between 117,337 variants and 10,358 diseases, traits, and phenotypes.