

Ontologies en médecine

De l'Ingénierie des connaissances au *Big Data*

Jean Charlet^{1,2,3}

¹INSERM UMR_S 1142 ; ²AP-HP ; ³Université Pierre et Marie Curie

Jean.Charlet@upmc.fr

Atelier IN-OVIVE, Montpellier – 07 juin 2016



Diapositives réutilisables sous licence *Creative Commons*
« Attribution, partage dans les mêmes conditions 3.0 France »

<http://creativecommons.org/licenses/by-sa/3.0/fr/>



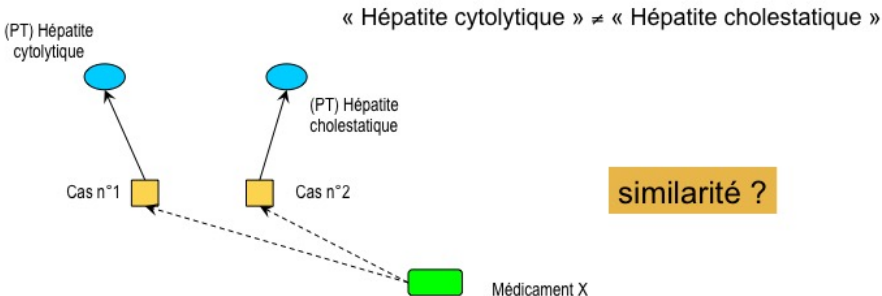
- 1 Introduction
 - Des ontologies
- 2 Des modèles
 - Plusieurs modèles?
 - Big Data
- 3 Usages des ontologies
 - Des usages
 - R.I.
 - Intégration de données
- 4 Quelques éléments de réflexion
 - Les ontologies
 - Les standards
 - Le Big Data
 - Aide à la décision
 - Et les SHS ?
- 5 Conclusion
 - Discussion
- 6 Annexes

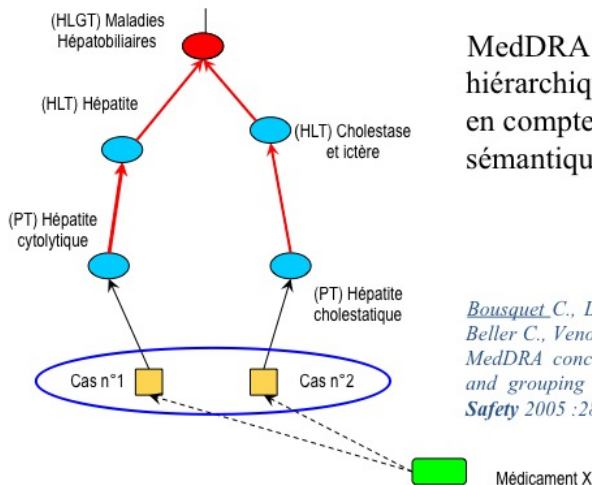
shrink=50

Dans mon Système de déclaration de ...

... pharmacovigilance, j'ai beaucoup de mal à regrouper des cas similaires

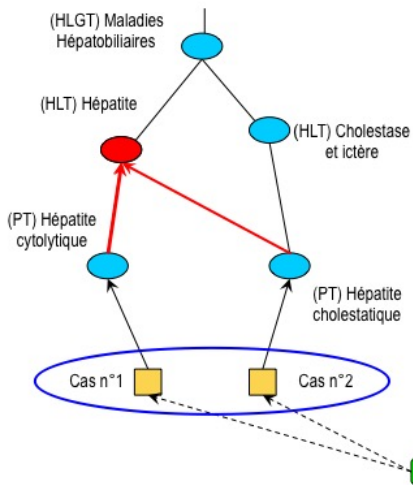
Pourquoi ? A cause de cela :





MedDRA : structure hiérarchique qui ne prend pas en compte la dimension sémantique des termes

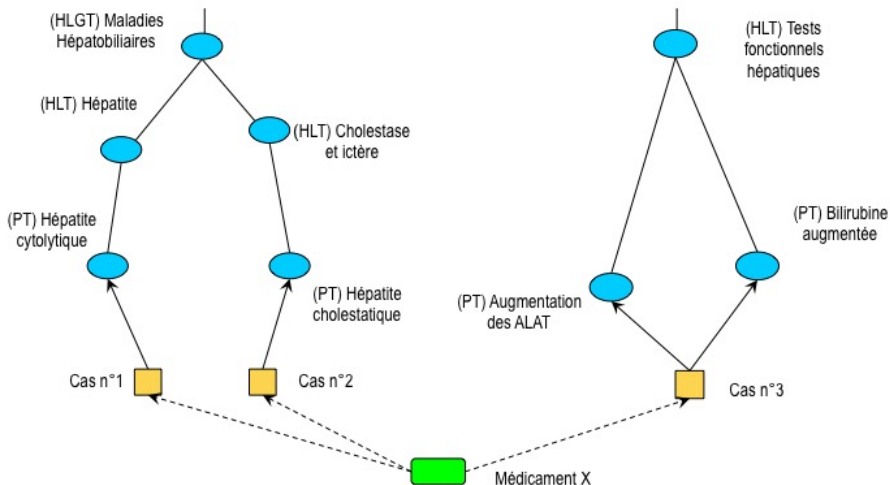
*Bousquet C., Lagier G., Lillo-Le Louët A., Le Beller C., Venot A., Jaulent M.-C. Appraisal of MedDRA conceptual structure for describing and grouping adverse drug reactions. **Drug Safety** 2005 ;28(1) :19-34.*

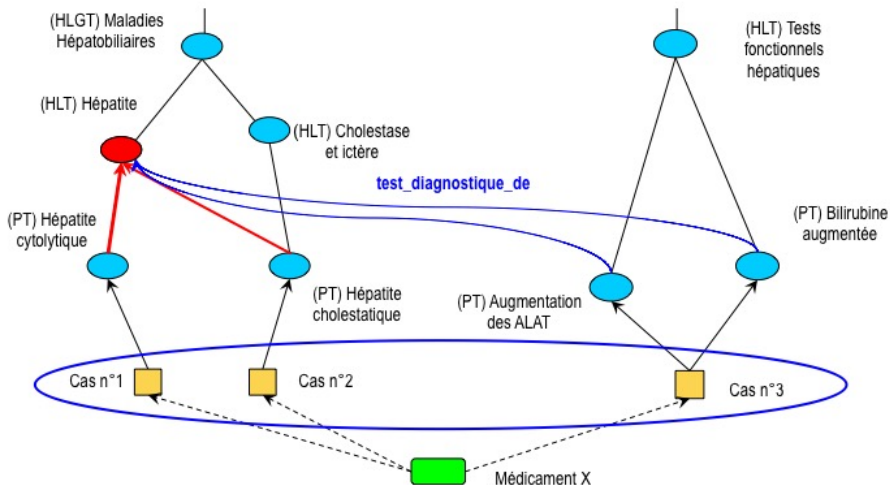


- **Problème à résoudre**

Représenter le contenu sémantique des termes MedDRA pour permettre le regroupement automatique des cas → **modélisation ontologique**

- **Hepatitis** = $(isStructuralDisorderOf\ SOME\ HepaticStructuralLevels)\ AND(hasTypeOfDisorder\ SOME\ InflammationTypes)$
- **Hepatitis_cholestatic** = $\{(isStructuralDisorderOf\ SOME\ HepaticStructuralLevels)\ AND\ (hasTypeOfDisorder\ SOME\ InflammationTypes)\}\ AND\ \{(isFunctionalDisorderOf\ SOME\ (Excretion\ AND\ [isFunctionOf\ SOME\ HepaticStructuralLevels]))\}$





Ressource sémantique

Trouver une ressource telle que le sens des concepts médicaux et les relations sémantiques qui existent entre eux puisse être représentés formellement pour pouvoir être exploités sans ambiguïté par un système informatique

Cette ressource qui a à la fois la propriété d'être compréhensible par un être humain et "par" une machine est une *ontologie informatique*

Ressource sémantique

Trouver une ressource telle que le sens des concepts médicaux et les relations sémantiques qui existent entre eux puisse être représentés formellement pour pouvoir être exploités sans ambiguïté par un système informatique

Cette ressource qui a à la fois la propriété d'être compréhensible par un être humain et "par" une machine est une *ontologie informatique*

Une ontologie

est un modèle muni de deux propriétés

Sémantique interprétative

Interprétable par les être humains (partagée)

Sémantique formelle

“Calculable” par un ordinateur

*Une ontologie est une **spécification partagée** d'une **conceptualisation***

(by Tom Gruber)

Une ontologie

est un modèle muni de deux propriétés

Sémantique interprétative

Interprétable par les être humains (partagée)

Sémantique formelle

“Calculable” par un ordinateur

*Une ontologie est une **spécification partagée** d'une **conceptualisation***

(by Tom Gruber)

Une ontologie

est un modèle muni de deux propriétés

Sémantique interprétative

Interprétable par les être humains (partagée)

Sémantique formelle

“Calculable” par un ordinateur

*Une ontologie est une **spécification partagée** d'une **conceptualisation***

(by Tom Gruber)

Une ontologie

est un modèle muni de deux propriétés

Sémantique interprétative

Interprétable par les être humains (partagée)

Sémantique formelle

“Calculable” par un ordinateur

*Une ontologie est une **spécification partagée** d'une **conceptualisation***

(by Tom Gruber)

- 1 Introduction
 - Des ontologies
- 2 Des modèles
 - Plusieurs modèles?
 - Big Data
- 3 Usages des ontologies
 - Des usages
 - R.I.
 - Intégration de données
- 4 Quelques éléments de réflexion
 - Les ontologies
 - Les standards
 - Le Big Data
 - Aide à la décision
 - Et les SHS ?
- 5 Conclusion
 - Discussion
- 6 Annexes

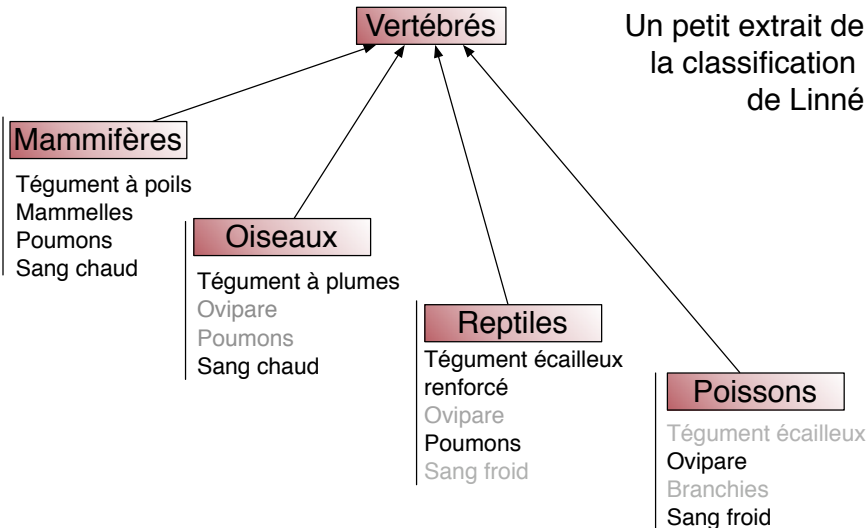
Une ontologie I/V



Image extraite du mnhn

Une ontologie II/V

Un petit extrait de
la classification
de Linné



Une ontologie III/V

The screenshot displays a web-based ontology editor interface. The browser address bar shows the URL: `http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23`. The page title is "untitled-ontology-23". A search bar is present with the text "Search for entity".

The main interface is divided into several panels:

- Active Ontology:** A navigation bar with tabs for "Active Ontology", "Entities", "Classes", "Object Properties", "Data Properties", "Individuals", "SKOS view", "OWLviz", and "DL Query".
- Class hierarchy:** A tree view showing the hierarchy of classes. The root is "Thing", which branches into "Animal", "Systeme", and "Attributs". "Animal" further branches into "Vertebre", "Amphibien", "Mammifere", "Oiseau", "Poisson", and "Reptile". "Vertebre" branches into "Amphibien" and "Mammifere". "Systeme" branches into "SystemeRespiratoire", "SystemeAPoumons", and "SystemeABranchies". "Attributs" branches into "Caracteristique" and "TypeTegument". "Caracteristique" branches into "CaracteristiqueVertebre", "CaracteristiqueNourritureBebe", "CaracteristiqueSangChaudFroid", "CaracteristiqueSangChaud", "CaracteristiqueSangFroid", "CaracteristiqueSortieEmbryon", "Ovipare", "Vivipare", and "TypeTegument". "TypeTegument" branches into "TegumentAPlumes", "TegumentAPOils", "TegumentEcailleux", and "TegumentEcailleuxRenforce".
- Annotations: Mammifere:** A panel showing the annotations for the class "Mammifere". It currently displays "Annotations +".
- Description: Mammifere:** A panel showing the description for the class "Mammifere". It displays "Equivalent To +", "SubClass Of +", and "SubClass Of (Anonymous Ancestor)".

The "Description: Mammifere" panel shows the following information:

- Equivalent To:** A plus sign (+) indicating no equivalents are currently listed.
- SubClass Of:** A plus sign (+) indicating no subclasses are currently listed.
- SubClass Of (Anonymous Ancestor):** A list of subclasses with their relationships to "Mammifere":
 - Vertebre** (some)
 - aPourCaracteristique some AnimalAMammelles** (some)
 - aPourCaracteristique some CaracteristiqueSangChaud** (some)
 - aPourCaracteristique some TegumentAPOils** (some)
 - aPourCaracteristique some Vivipare** (some)
 - aPourSystemeRespiratoire some SystemeAPoumons** (some)

Une ontologie IV/V

untitled-ontology-23 (http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23) : [Users/jc/Documents/GlobDjin/txt/communications/2012ITS...]

untitled-ontology-23 (http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23) Search for entity

Active Ontology Entities Classes Object Properties Data Properties Individuals SKOS view OWLViz DL Query

Annotations Selected entailments

Ontology header:

Ontology IRI

Ontology Version IRI

Annotations +

Ontology metrics:

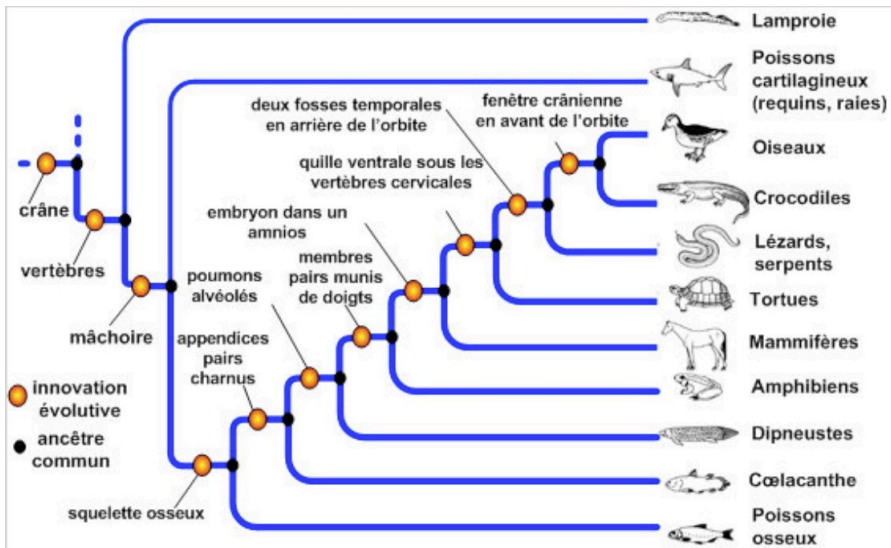
Metrics	
Axiom	78
Logical axiom count	49
Class count	27
Object property count	2
Data property count	0
Individual count	0
Blank node count	0

Ontology imports General axioms RDF/XML rendering OWL/XML rendering OWL functional syntax rendering

RDF/XML rendering:

```
<!-- http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#Mammifere -->
<owl:Class rdf:about="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#Mammifere">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#Vertebre" />
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#aFourCaracteristique" />
      <owl:someValuesFrom rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#Vivipare" />
    </owl:Restriction>
  </rdfs:subClassOf>
</rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#aPourSystemeRespiratoire" />
    <owl:someValuesFrom rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#SystemeAPoumons" />
  </owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#aPourCaracteristique" />
    <owl:someValuesFrom rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#TegumentAPoils" />
  </owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#aPourCaracteristique" />
    <owl:someValuesFrom rdf:resource="http://www.semanticweb.org/ontologies/2012/9/untitled-ontology-23#CaracteristiqueSangChaud" />
  </owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf>
```


Une ontologie V/V



<http://www.cosmovisions.com/mammiferesClassification.htm>

Les 5V du Big Data

Volumetry,
Velocity,
Variety,
Veracity,
Value

- 1 Introduction
 - Des ontologies
- 2 Des modèles
 - Plusieurs modèles?
 - Big Data
- 3 Usages des ontologies**
 - Des usages
 - R.I.
 - Intégration de données
- 4 Quelques éléments de réflexion
 - Les ontologies
 - Les standards
 - Le Big Data
 - Aide à la décision
 - Et les SHS ?
- 5 Conclusion
 - Discussion
- 6 Annexes

Usages des ontologies en médecine

- Corriger des systèmes d'informations cliniques opérationnels mais fonctionnant en silos non interopérables
- *Compléter les informations contenues dans les S.I. par des connaissances médicales*
- Améliorer les systèmes d'indexation ou *de codage (Indexation médicale vs T2A)*
- Faire de la *recherche d'information dans les DPI/DMP*
- Fournir des modèles jusqu'au plus haut niveau de formalisation (dictionnaires, classifications, taxonomies, ontologies, signaux) pour divers SIAD
- *Trouver un modèle pivot unique et standardisé pour unifier des modèles de données différents* dans les entrepôts de données cliniques qui sont alors dits *sémantiques* (EDCS)

LeRUDI : Une ontologie pour la R.I.

Pour de la R.I.

- Répertorier les concepts dans les documents médicaux
- À travers les termes les dénotant
- Dans le cadre du projet LeRUDI

Pour faire des raisonnements sur les structures trouvées

- À travers la relations de subsomption
- Pour remonter des disciplines médicales dans un nuage de mots
- Pour tenir compte des processus et des localisations des maladies

L'œdème de la moelle épinière dans ONTOLURGENCES

The screenshot displays the OntolURGENCES web interface. The browser address bar shows the URL: `http://doe-generated-ontology.com/UrgencesDMP`. The search bar contains the text `infar`. The main content area is divided into two panes. The left pane, titled "Class hierarchy", shows a tree structure of classes. The right pane, titled "Annotations: OedemeDeLaMoelleEpinriere", displays various annotations for the selected class.

Class hierarchy: OedemeDeLaMoelleEpinriere

- ▶ DegenerescenceCorticoBasale
- ▶ MaladieCerebroVasculaire
- ▶ MaladieCorticale
- ▶ MaladieDemyelinisanteDuSystemeNerveuxCentral
- ▶ MaladieDesNoyauxGrisCentraux
- ▼ MaladieMedullaire
 - AbcesIntrarachidien
 - AmyotrophieSpinale
 - AtaxieSpinale
 - CompressionMedullaire
 - Myelopathies
 - MyelopathiePostRadiotherapie
 - ▼ MyelopathieVasculaire
 - InfarctusAiguDeLaMoelleEpinriere
 - **OedemeDeLaMoelleEpinriere**
 - ThromboseArterielleDeLaMoelleEpinriere
 - Neuromyelite
 - Syringomyelie
 - PoliomyeliteAigue
 - ScleroseLateraleAmyotrophique
 - ▼ MaladieMeninges
 - AbcesSousDural
 - Arachnoidite
 - Ependymite
 - HemorragieSousDuraleNonTraumatique
 - Leptomeningite
 - Meningisme
 - ▶ Meningite
 - MeningiteLymphocytaireChroniqueBenigne
 - Pneumoencephale
 - ▶ MaladieSousCorticale
 - ▶ MigraineSP

Annotations: OedemeDeLaMoelleEpinriere

Annotations

- altLabel [language: fr] oedème rachidien
- hiddenLabel [language: fr] OedemeDeLaMoelleEpinriere
- prefLabel [language: fr] oedème de la moelle épinière
- prefLabel [language: en] edema of spinal cord

Description: OedemeDeLaMoelleEpinriere

Equivalent To

SubClass Of

- MyelopathieVasculaire

General class axioms

SubClass Of (Anonymous Ancestor)

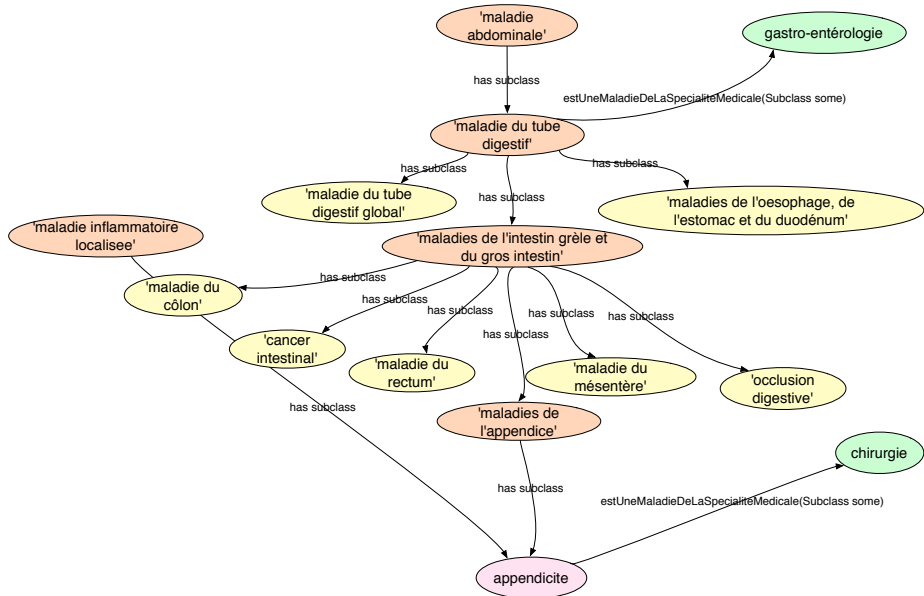
- estUneMaladieDeLaSpecialiteMedicale some Neurologie
- peutAvoirUneEvolution some Chronique

Instances

Target for Key

To use the reasoner click Reasoner → Start reasoner Show Inferences

LeRUDI : extrait de l'ontologie



LeRUDI : Écran de synthèse

Dominique Gilles



DATE NAISSANCE - AGE
01/04/1970 - 40 ans

TÉLÉPHONE
06 06 06 06 06

ADRESSE
9 rue Georges Pitard - 69000 LYON

Documents récents

Ensemble des documents récents (66)

Artériopathie oblitérante du MI droit ... En conclusion, douleur thoracique chez un patient polyvasculaire. Le bilan biologique est normal...	04/03/2008	Voir ce doc Voir les docs d'URO (17)	RCP	Service d'urologie - Unité... ...Au total, cancer de la prostate traité par résection endo-urétrale. Bilan à 3 ans normal...	04/03/2008	Voir ce doc Voir le de RC
Cardiologie - Unité de soins intensifs ... Au total, patient admis pour décompensation cardiaque grave dans les suites d'un infarctus du myocarde antérieur. Il sort stabilisé, pour un séjour...	04/03/2008	Voir ce doc Voir les docs de REA-SI(17)	CONSULT	Chirurgie vasculaire -En conclusion, ulcération interdigitale surinfectée ayant donnée une lymphangite de tout le pied...	04/03/2008	Voir ce doc Voir le de COI (17)
Service de pneumologie ... Au total: décompensation cardiaque globale chez un patient avec de lourds antécédents cardio-vasculaires d'évolution favorable...	04/03/2008	Voir ce doc Voir les docs d'HO3PI (17)	BIO			

Pathologies		Actes médicaux		Traitements	
Ensemble des pathologies (6)					
Voir l'ensemble des actes (4)					
Prescription hospitalière la plus récente : 17/01/2010					
Artériopathie oblitérante du MI droit	22 docs	Voir tous les docs	Pontage fémopoplité droit	10 docs	Voir tous les docs
Diabète de type 1	17 docs	Voir tous les docs	Pose de pacemaker	7 docs	Voir tous les docs
Insuffisance cardiaque grave	13 docs	Voir tous les docs	Pose de stent	4 docs	Voir tous les docs
Troubles du rythme ventriculaire chroniques	9 docs	Voir tous les docs	Résection prostatique endo-urétrale	3 docs	Voir tous les docs
Cancer de la prostate	4 docs	Voir tous les docs			
					Inexium 20
					Kardegic 75
					Amlor 5
					Movicor LP 20
					Sectral 200

LeRUDI : Écran de recherche

Dominique Gilles

DATE NAISSANCE - AGE
01/04/1970 - 40 ans

TÉLÉPHONE
06 06 06 06 06

ADRESSE
9 rue Georges Pitard - 69000 LYON

Visualiser le dossier : B747897

 Lancer

Filtre de services : Tous ▼

 Réinitialiser le formulaire

Supprimer un critère :

 Mot 1

 Mot 2



Résultats de votre recherche :

Tri et personnalisation des résultats: ⌵

Documents correspondants : 25 documents trouvés

Page(s) : « Précédent | 1 2 3 |

Menu des documents

Pathologies - 14 docs

Pathologies neurologiques (7)



Asthme

Symptômes - 14 docs

Signes cliniques (7)

Actes - 14 docs

Contenus acte (7)

Service	Unité fonctionnelle	Date
 HOSPI	Accueil et urgences Saint Julien ... Un sixième décompensation cardiaque illud autem non dubitatur quod cum esset aliquando virtutum omnium domicilium Roma, ingenuos advenas plerique nobilium, ut Homerici bacarum suavitare...	01/01/2010
 HOSPI	Département de médecine polyclinique Saint Julien ... ud autem non dubitatur quod cum esset aliquando virtutum omnium domicilium Roma, ingenuos advenas plerique nobilium, ut Homerici bacarum suavitare sixième décompensation cardiaque illud autem non dubitatur quod cum esset aliquando virtutum omnium domicilium Roma, ingenuos advenas plerique nobilium, ut Homerici bacarum suavitare...	01/10/2009

[Voir ce doc](#)

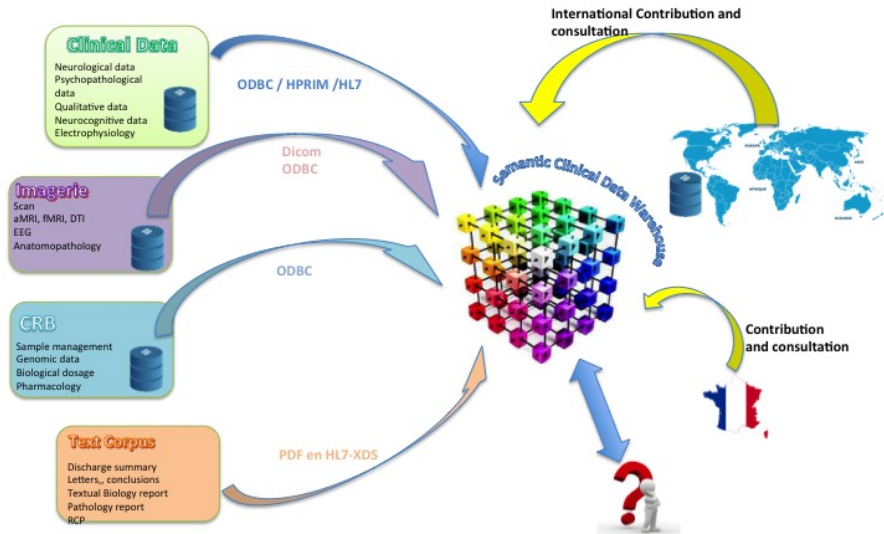
[Voir ce doc](#)

Intégration de données

- **Que ce soit** dans le contexte du cancer où le séquençage du génome ouvre des voies de compréhension et, encore plus vite, de thérapeutiques. . .
- **Que ce soit** pour les MR pour 80% mendéliennes mais dont on n'a souvent que le phénotype. . .
- **Que ce soit** en psychiatrie où les troubles sont principalement décrits par des syndromes cliniques et où l'on a besoin de mieux caractériser le phénomène (e.g. EMA, oculomotricité) et le croiser avec des variants génomiques. . .
- **Que ce soit**. . .

*Volumetry, Velocity, **Variety**, Veracity*

EDCS



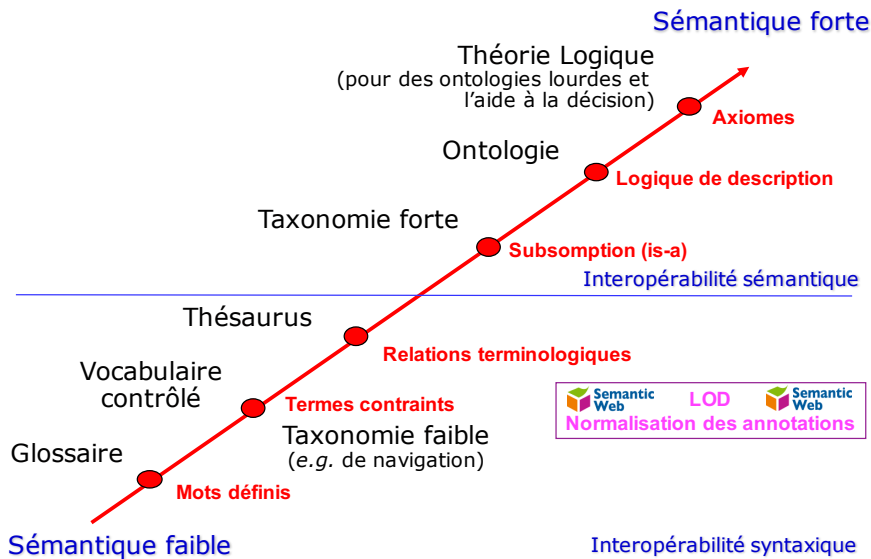
- 1 Introduction
 - Des ontologies
- 2 Des modèles
 - Plusieurs modèles?
 - Big Data
- 3 Usages des ontologies
 - Des usages
 - R.I.
 - Intégration de données
- 4 Quelques éléments de réflexion**
 - Les ontologies
 - Les standards
 - Le Big Data
 - Aide à la décision
 - Et les SHS ?
- 5 Conclusion
 - Discussion
- 6 Annexes

Quelques bonnes propriétés des ontologies

- Possibilité d'associer aux concepts des termes dans **différentes langues** ou des termes du **langage médical** vs du **langage patient**
- Selon les besoins et les points de vues, les personnes ne feront pas les mêmes ontologies du même domaine mais
- Des outils `<http://sourceforge.net/projects/onagui/>` et des langages (SKOS) permettent d'**aligner** les ontologies
- Chaque concept a un identifiant unique (IRI) dans le monde Internet (aussi vrai d'un thésaurus bien modélisé)

<http://www.limics.fr/ontologies/OntolUrgences#Appendicite>

Le spectre des SOC/ KOS



Disponibilité et standards I

- Le problèmes des ontologies de l'IA (raisonnements) et du Web de données ont des échelles différentes : *big* ontologies et relativement peu de données (10^4 et 10^6) vs petites ontologies et *big data* ($\leq 10^2$ et $\geq 10^8$)
- De nombreuses ontologies (et terminologies) médicales sont disponibles sur Internet. Il est nécessaire de les mettre à disposition (voir <http://bioportal.bioontology.org>) et les aligner (voir <http://www.hetop.eu/hetop/>)
- Des services professionnels de type aide au codage ou fouille de données sémantiques passent par des applications de faible champ mais avec des développements sémantiques forts (et coûteux)

Disponibilité et standards II

- Dans tous les cas, il faut assurer l'interopérabilité syntaxique via les **standards** (langages de représentation des connaissances, etc.)
- Et la SNOMED-CT ? Et si c'était un problème de gouvernance ?
- Il faut d'urgence un "lieu" où sont les SOC importants de la médecine, dans les formats du WS, dans toutes les versions nécessaires (à ce jour, à plusieurs endroits, dans n'importe quel format, non versionnés, etc.)

Le big data ou la question de la réutilisation des données I

Volumetry, Velocity, Variety, Veracity

De grosses variations dans l'interprétation a posteriori du codage PMSI avec l'affirmation que ça pourrait s'arranger ¹, qu'on manque de données ², que les données sont biaisées (à 10%) ³, qu'on loupe 20% des cas dans une méta-analyse sur le diabète ⁴, 25% des cas dans une méta-analyse sur l'insuffisance cardiaque ⁵

1. Benchimol *et al.*, 2011

2. Shahian *et al.*, 2007

3. Townsend *et al.*, 2012

4. Leong *et al.*, 2013

5. McComick *et al.*, 2014

Le big data ou la question de la réutilisation des données II

- Les données sont codées pour des **but**s précis, avec des **règles de codage précises**, avec une **réduction du sens imposée** et un **retour difficile** (vers le texte ou, plus généralement, la ressource d'origine)
- Des problèmes analogues avec les données des capteurs
 - *Prendre les données pour ce qu'elles sont, en tenant compte du contexte de captation, comme point d'entrée pour d'autres études plus ciblées si nécessaire*

Le big data ou la question de la réutilisation des données III

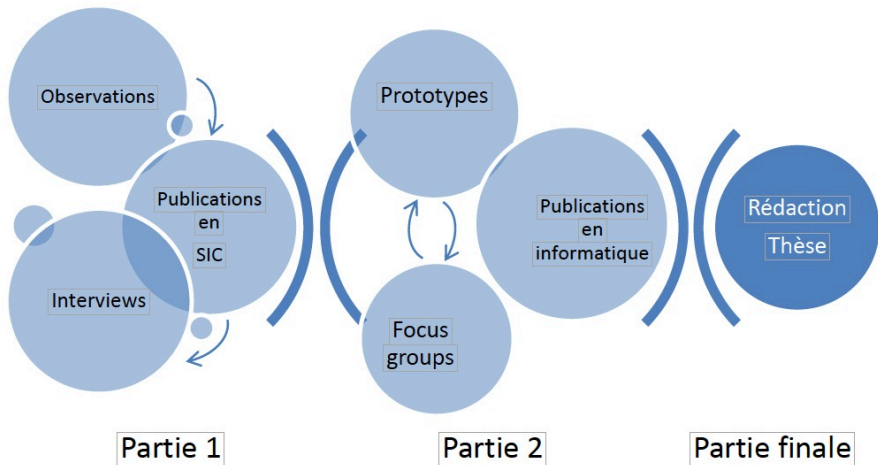
- L'annotation des textes⁶ ($\approx 80\%$ de l'inf. méd.) fondée sur des terminologies/ontologies est un outil majeur de l'informatique médicale
- Dans de nombreux pays, des outils sont commercialement proposés pour **indexer** des bases textuelles
 - Pour **améliorer, vérifier le codage médical**
 - Pour, souvent, in fine, **améliorer la détection (automatique) des patients pour de potentiels essais cliniques**
- Nous travaillons sur ces outils pour analyser les trajectoires des patients atteints de SLA ou pour déclarer des cas de fœtopathologie dans le but de faire du RàPC

CDSS : la convergence

- Des promesses de l'IA non tenues jusque là
- Après 20 ans de séparation, les modèles ontologies et les CDSS sont arrivés à une maturité qui permet de voir se développer
- Des CDSS qui font des raisonnements formels fondés sur des ontologies et modélisés avec le couple OWL/SWRL

Une étude SHS du dossier médical

- Un **croisement de démarches** ethnographiques et informatiques pour analyser un dossier patient en cancérologie
- Une explicitation des **tensions organisationnelles** dans les services de soins
- Description du **processus de pratique annotative** et explicitation d'un certain nombre d'heuristiques de conception d'un logiciel outillant ce processus



Equipement informatique des Annotations et des Pratiques d'écriture professionnelles :
une étude ancrée pour l'organisation des soins en cancérologie,
Thèse, Philippe Marrast

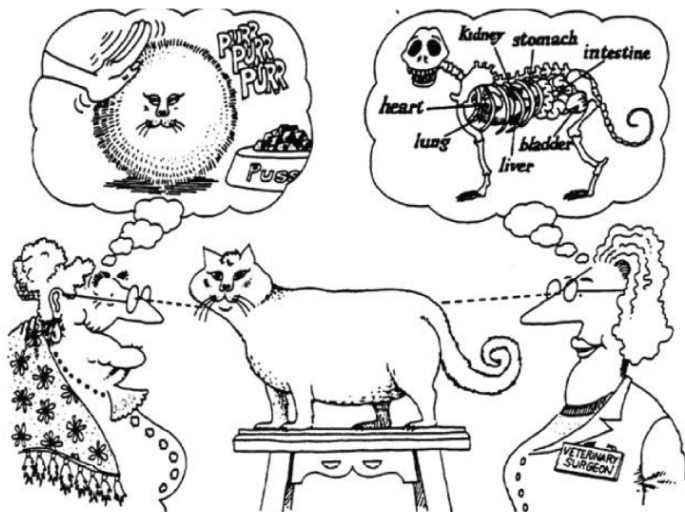
- 1 Introduction
 - Des ontologies
- 2 Des modèles
 - Plusieurs modèles?
 - Big Data
- 3 Usages des ontologies
 - Des usages
 - R.I.
 - Intégration de données
- 4 Quelques éléments de réflexion
 - Les ontologies
 - Les standards
 - Le Big Data
 - Aide à la décision
 - Et les SHS ?
- 5 **Conclusion**
 - **Discussion**
- 6 Annexes
 - Web sémantique
 - intégrer des données : DebugIT
 - Les médicament diurétiques dans OntoMédicaments

Discussion

- Des données au service de la santé pour peu qu'on en **maîtrise les modèles** et qu'on développe les applications qui s'en servent
- Dans le contexte du Big Data, de l'intégration des données de toutes natures en étant capable de caractériser **la nature, la provenance, la fiabilité, le pourquoi** de la captation **des données**
- Quels enjeux éthiques ?
 - Au niveau réglementaire/CNIL : traitement/++ mais délais/—
 - Au niveau des acteurs : public ou privé ?

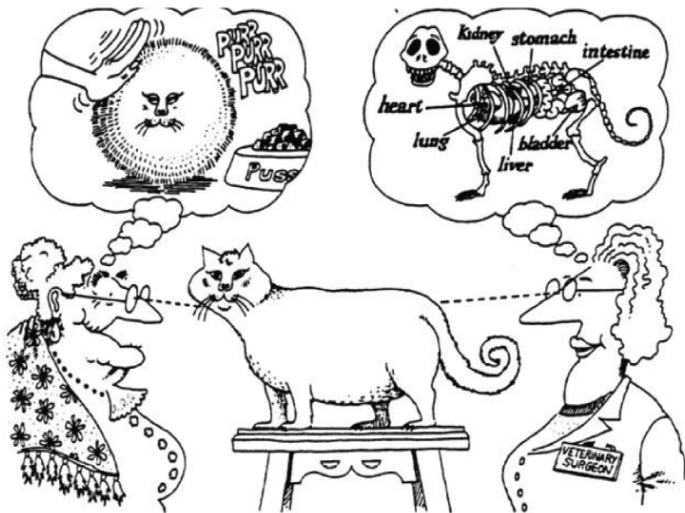
<https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/>
http://affordance.typepad.com//mon_weblog/2016/05/un-rein-vaut-mieux-que-deux-ia.html

Un chat est un chat !



Abstraction focuses upon the essential characteristics of some object, relative to the perspective of the viewer.

Un chat est un chat !



Abstraction focuses upon the essential characteristics of some object, relative to the perspective of the viewer.

Je vous remercie de votre attention.

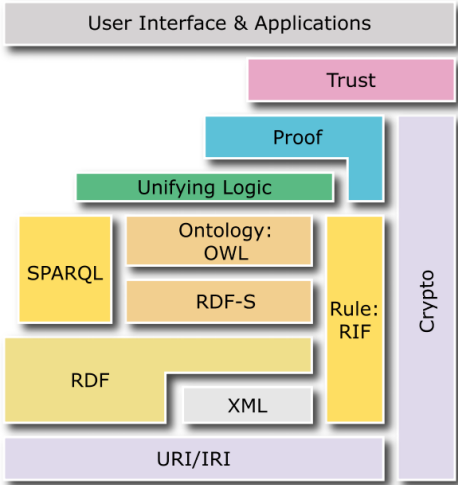
- 1 Introduction
 - Des ontologies
- 2 Des modèles
 - Plusieurs modèles?
 - Big Data
- 3 Usages des ontologies
 - Des usages
 - R.I.
 - Intégration de données
- 4 Quelques éléments de réflexion
 - Les ontologies
 - Les standards
 - Le Big Data
 - Aide à la décision
 - Et les SHS ?
- 5 Conclusion
 - Discussion
- 6 Annexes
 - Web sémantique
 - intégrer des données : DebugIT
 - Les médicament diurétiques dans OntoMénégas

Les standards

- ☞ L'interopérabilité doit d'abord être **syntaxique**
- ☞ L'**interopérabilité sémantique** se fonde sur des ontologies
- ☞ Les machines ne traitent la sémantique que parce que l'on respecte la syntaxe ☞ On parle de **sémantique formelle**

Un seul mot d'ordre, respecter les standards syntaxiques

Le « cake » du Web de données (2007)



RDF

RDF est un modèle de triplets :
tout élément de connaissance se réduit
à (**sujet** , **prédictat** , **objet**)



(by Thomas Francart, SPARNA)

RDF

L'appendicite est une sous-classe des maladies de l'appendice et des maladies inflammatoires localisées

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:urgences="http://www.limics.fr/ontologies/ontolurgences#"
  >

  <owl:Class rdf:about="&Urgences;Appendicite">
    <rdfs:subClassOf rdf:resource=
      "urgences;MaladieInflammatoireLocalisee"/>
    <rdfs:subClassOf rdf:resource=
      "urgences;MaladiesDeLAppendice"/>
  </owl:Class>

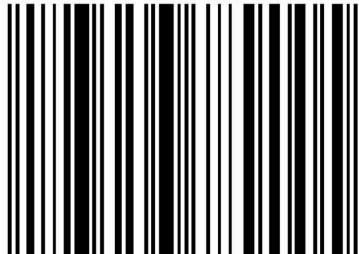
</rdf:RDF>

```

RDF

en **RDF**

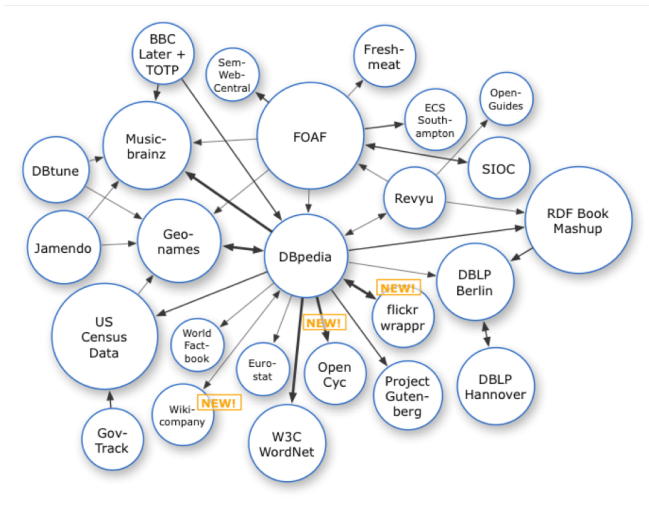
les ressources
et les propriétés
sont identifiés par
des **URIs**.



<http://mydomain.org/mypath/myresource>

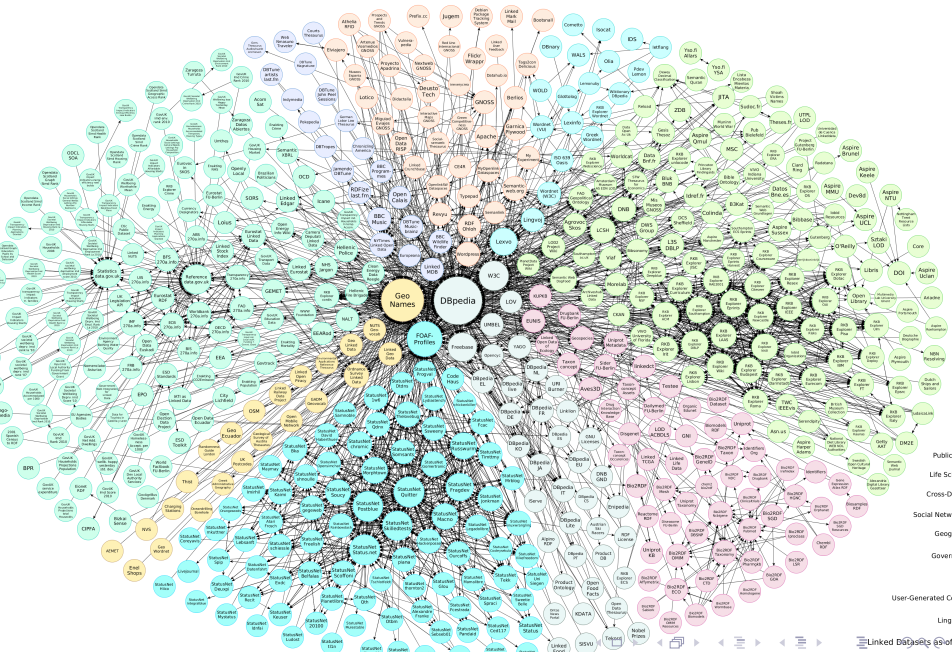
<http://www.limics.fr/ontologies/OntoUrgences#Appendicite>

Linking Open Data Cloud Diagram 1



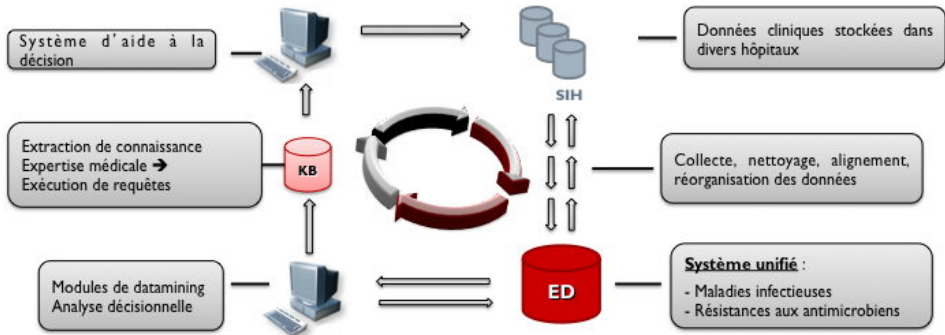
2007-10-08

Linking Open Data Cloud Diagram 4



Public
Life Sci
Cross-D
Social Netw
Govern
User-Generated Co
Lingu

Le projet européen DebugIT II



- 1^{er} Champ d'application : les infections urinaires

- Exécution de requêtes :

Liste des patients ayant une infection spécifique (ex : infection urinaire) par un agent pathogène donné (ex : E. coli) résistante à un antibiotique donné (ex : amoxicilline)

Comment répondre à des questions “complexes”

ID	Bacterium	Antibiotic	Result	Value
1	E. Coli	Amoxicilline	Resistant	25
2	Escherichia coli	Cefoperazone	Sensitive	28
3	Pseudomonas aeruginosa	Ofloxacin	Intermediate	10

ID	Bactérie	Antibiotique	R	V
10	Escherichia.Coli	Cefpirome	R	25
20	Hafnia alvei	Quinolones	S	28
30	Pseudomonas aeruginosa	Ofloxacin	I	10

Sources de données

Question : Quels sont les résultats d'antibiogrammes du test de l'E. Coli résistante aux B-lactamines ?

Comment répondre à des questions "complexes"

ID	Bacterium	Antibiotic	Result	Value
1	E. Coli	Amoxicilline	Resistant	25
2	Escherichia coli	Cefoperazone	Sensitive	28
3	Pseudomonas aeruginosa	Ofloxacin	Intermediate	10

ID	Bactérie	Antibiotique	R	V
10	Escherichia.Coli	Cefpirome	R	25
20	Hafnia alvei	Quinolones	S	28
30	Pseudomonas aeruginosa	Ofloxacin	I	10

Sources de données

Question : Quels sont les résultats d'antibiogrammes du test de l'E. Coli résistante aux B-lactamines ?

Ce qu'on ne sait pas :

- ✗ Les champs « bacterium » & « bactérie » font référence au même concept « bactérie »
- ✗ Amoxicilline est une B-Lactamine
- ✗ E. Coli et Escherichia coli font référence à la même bactérie

Comment répondre à des questions “complexes”

ID	Bacterium	Antibiotic	Result	Value
1	E. Coli	Amoxicilline	Resistant	25
2	Escherichia coli	Cefoperazone	Sensitive	28
3	Pseudomonas aeruginosa	Ofloxacin	Intermediate	10

ID	Bactérie	Antibiotique	R	V
10	Escherichia.Coli	Cefpirome	R	25
20	Hafnia alvei	Quinolones	S	28
30	Pseudomonas aeruginosa	Ofloxacin	I	10

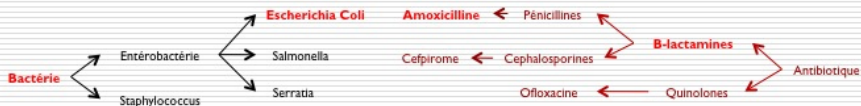
Hypothèse : Utiliser une ontologie pour enrichir les données cliniques

- Technologies du web sémantique
- Grâce à cette combinaison de source de données, on saura que pour les antibiogrammes 1 et 10, E. Coli est résistante aux B-lactamines, et particulièrement à l'amoxicilline et à la Cefpirome.

Comment répondre à des questions "complexes"

ID	Bacterium	Antibiotic	Result	Value
1	E. Coli	Amoxicilline	Resistant	25
2	Escherichia coli	Cefoperazone	Sensitive	28
3	Pseudomonas aeruginosa	Ofloxacin	Intermediate	10

ID	Bactérie	Antibiotique	R	V
10	Escherichia.Coli	Cefpirome	R	25
20	Hafnia alvei	Quinolones	S	28
30	Pseudomonas aeruginosa	Ofloxacin	I	10



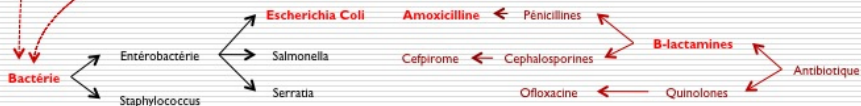
Hypothèse : Utiliser une ontologie pour enrichir les données cliniques

- Technologies du web sémantique
- Grâce à cette combinaison de source de données, on saura que pour les antibiogrammes 1 et 10, E. Coli est résistante aux B-lactamines, et particulièrement à l'amoxicilline et à la Cefpirome.

Comment répondre à des questions "complexes"

ID	Bacterium	Antibiotic	Result	Value
1	E. Coli	Amoxicilline	Resistant	25
2	Escherichia coli	Cefoperazone	Sensitive	28
3	Pseudomonas aeruginosa	Ofloxacin	Intermediate	10

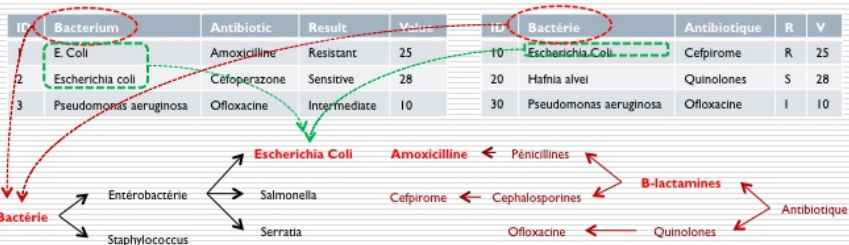
ID	Bactérie	Antibiotique	R	V
10	Escherichia.Coli	Cefpirome	R	25
20	Hafnia alvei	Quinolones	S	28
30	Pseudomonas aeruginosa	Ofloxacin	I	10



Hypothèse : Utiliser une ontologie pour enrichir les données cliniques

- Technologies du web sémantique
- Grâce à cette combinaison de source de données, on saura que pour les antibiogrammes 1 et 10, E. Coli est résistante aux B-lactamines, et particulièrement à l'amoxicilline et à la Cefpirome.

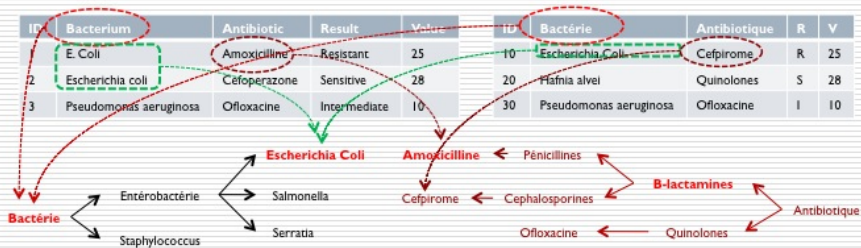
Comment répondre à des questions "complexes"



Hypothèse : Utiliser une ontologie pour enrichir les données cliniques

- Technologies du web sémantique
- Grâce à cette combinaison de source de données, on saura que pour les antibiogrammes 1 et 10, E. Coli est résistante aux B-lactamines, et particulièrement à l'amoxicilline et à la Cefpirome.

Comment répondre à des questions "complexes"



Hypothèse : Utiliser une ontologie pour enrichir les données cliniques

- Technologies du web sémantique
- Grâce à cette combinaison de source de données, on saura que pour les antibiogrammes 1 et 10, E. Coli est résistante aux B-lactamines, et particulièrement à l'amoxicilline et à la Cefpirome.

La mécanique de l'intégration dans DebugIT

Résultats

antibiotic1	bacteria	rl	susceptibility1	antibiogram
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51506337>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51506308>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51546026>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51545998>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51546143>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51546063>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51524644>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51524616>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51529294>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51529265>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51529635>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51529606>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51545236>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51545208>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51545339>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51545310>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51573721>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51573692>
<http://debugit1.spin.jussieu.fr/resource/antibiotic/60>	<http://debugit1.spin.jussieu.fr/resource/bacterie/65>	<http://debugit1.spin.jussieu.fr/resource/Vocab/Sensitive>	<http://debugit1.spin.jussieu.fr/resource/test/51573430>	<http://debugit1.spin.jussieu.fr/resource/antibiogram/51573401>

Résultats

Données initiales:

- **238 623** résultats d'antibiogrammes
- Effectués sur **61** antibiotiques (12 pénicillines, 13 céphalosporines soit 21 bêta-lactamines) et **165** bactéries distinctes

Réponse à la requête SPARQL:

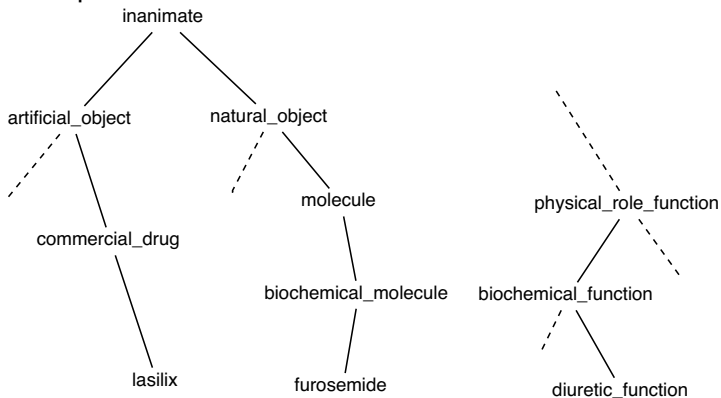
- Par couplage de l'ontologie avec la base de données
- ✗ On obtient **4 001** résultats de résistance de l'E. Coli aux Bêta-lactamines. (25 673 sensibles, 2 716 Intermédiaires)

Quelques remarques

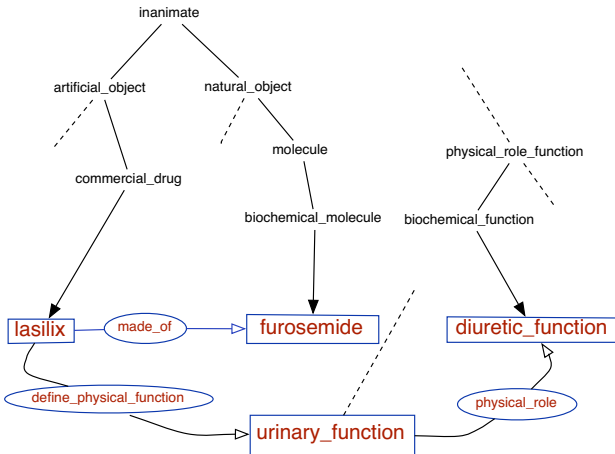
- Cela suppose la présence de la prescription connectée
- Les résultats obtenus ne pourront être validés qu'en supposant que les données d'origine sont riches et *de qualité*
- Les vocabulaires de qualité en machine, pour des données de qualité en machine, pour produire d'autres données de qualité

② La normalisation sémantique

Normalisation sémantique. Selon le contexte et la tâche ➡ arbre de concepts

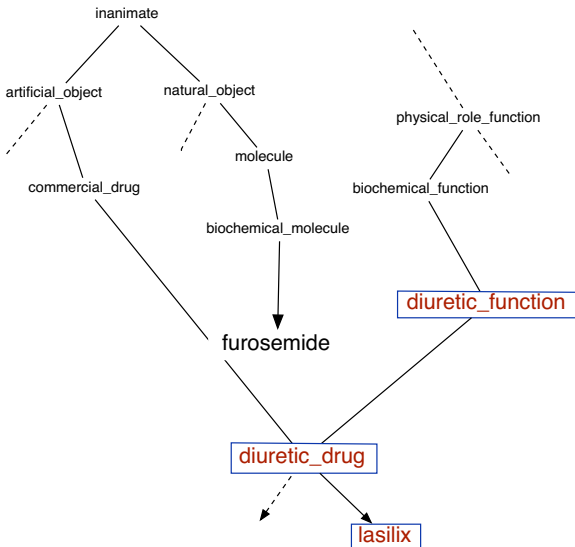


③ Engagement ontologique



Construire (ou pas)
des *concepts définis*
par CN&S...

③ Engagement ontologique (4)



...qui peuvent hériter des attributs de leurs pères et des valeurs de ceux-ci et qui permettent de re-classifier automatiquement certains concepts dessous

La puissance des raisonnements mais à manier avec méthode... Surtout s'il y en a beaucoup

Les médicaments diurétiques II

The screenshot shows a web-based ontology editor interface. The browser address bar displays the URL: `http://doe-generated-ontology.com/Top_ontologie_de_la_medecine`. The main navigation bar includes tabs for **Active Ontology**, **Entities**, **Classes** (selected), **Object Properties**, **Data Properties**, **Individuals**, **OWL Viz**, and **DL Query**. A search bar on the right contains the text `Q: diu`.

The interface is divided into two main panels:

- Left Panel: Class hierarchy: lasilix**
 - Class hierarchy (inferred)
 - Class hierarchy: lasilix
 - artificial_object
 - buildings
 - commercial_drug
 - actilyse
 - adalate
 - aldomet
 - aspegic
 - befizial
 - bitildiem
 - calciparine
 - captopril
 - cebutid
 - cipralan
 - coragine
 - cordarone
 - cordipatch_5
 - cordium
 - cortancyl
 - corvasal
 - daonil
 - diafusor
 - diamicron
 - dilzem
 - diuretic_drug
 - fraxiparine
 - glucophage
 - isoket
 - isoptine
 - lasilix**
 - lexomil

- Right Panel: Annotations: lasilix**
- Annotations: lasilix
- Annotations +
- definition: "CG Representation : [lasilix: _x]"@en
- hiddenLabel: "lasilix"@en
- hiddenLabel: "lasilix"@fr
- Bottom Panel: Description: lasilix**
- Description: lasilix
- Equivalent classes +
- Superclasses +
 - commercial_drug and (defines_physical_function some (urinary_function and (physical_role some diuretic_function)))
- Inherited anonymous classes
- Members +
- Keys +

At the bottom of the interface, there is a status bar with the text: **To use the reasoner click Reasoner->Start Reasoner** and a checked checkbox for **Show Inferences**.

Les médicaments diurétiques III

Top_ontologie_de_la_medecine (http://doe-generated-ontology.com/Top_ontologie_de_la_medecine) - [~/Volumes/Home/Users/jc/Applications/java/fichierstravail/owl/Ont...

Top_ontologie_de_la_medecine (http://doe-generated-ontology.com/Top_ontologie_de_la_medecine) Q diu

Active Ontology Entities **Classes** Object Properties Data Properties Individuals OWLViz DL Query

Class hierarchy Class hierarchy (inferred)

Class hierarchy (inferred): lasilix

- cebutid
- cipralan
- coragine
- cordarone
- cordipatch_5
- cordium
- cortancyl
- corvasal
- daonil
- diafusor
- diamicon
- dilzem
- ▼ ● diuretic_drug
 - **lasilix**
- fraxiparine
- glucophage
- isoket
- isoptine
- lexomil
- lipanthyl
- logirene
- lopirine
- loproresor_oros
- loproresor
- lopril
- lovenox
- maxepa
- panorex
- persantine
- reniten
- risordan
- sartel

Annotations: lasilix

Annotations +

definition

"CG Representation : [lasilix: _x]"@en

hiddenLabel

"lasilix"@en

hiddenLabel

"lasilix"@fr

Description: lasilix

Superclasses +

- commercial_drug
 - and (defines_physical_function some (urinary_function and (physical_role some diuretic_function)))
- diuretic_drug

Inherited anonymous classes

- commercial_drug
 - and (defines_physical_function some (urinary_function and (physical_role some diuretic_function)))

Reasoner active Show Inferences