

Serveur terminologique inter-lingue : vers des applications pratiques

Stefan J. Darmoni, MD, PhD

*TIBS, LITIS Lab
Rouen University Hospital & Rouen Medical School, Normandy University, France
LIMICS, INSERM U1142*

Email: Stefan.Darmoni@chu-rouen.fr



Pôle de Santé Publique

JF. Gehanno, V. Mangot, MH. Roux

CISMED

Enseignement et
recherche (sous-
section CNU 46-02)

Enseignement et recherche (sous-section CNU
46-04)

Enseignement et recherche (sous-section CNU 46-01)

JF. Gehanno

Département de
santé au travail et
pathologie
professionnelle
(DSTPF)

Unité de santé au
travail des
personnels
hospitaliers
(USTPH)

Unité de pathologie
professionnelle
(UPF)

SJ. Darmoni

Département
d'information et
d'informatique
médicales (DIIM)

Unité d'information
médicale (UIM 1)

Unité
d'informatique
médicale (UIM 2)

J. Benichou

Département
d'appui à la
recherche clinique
(DARC)

Centre
d'investigation
clinique (CIC)

Centre de ressources
biologiques (CRB)

Unité de
biostatistiques (UB)

J. Ladner

Département de
santé publique et
épidémiologie
(DSPE)

Unité de santé
publique et
d'épidémiologie
clinique (USPEC)

Unité transversale
d'éducation
thérapeutique du
Patient (UTEP)

I. Maréchal

Département
d'appui à la qualité
et à la sécurité des
soins (DAQSS)

Unité de prévention
des risques associés
aux soins (UPRAS)

Unité de sécurité
transfusionnelle et
hémovigilance
(USTH)

Coordination des
vigilances sanitaires
(CVS)

(+ activité d'identito-
vigilance)

V. Merle

Département
d'hygiène
hospitalière (DHH)

Unité de prévention
des infections
nosocomiales (UPIN)

Laboratoire
d'hygiène
hospitalière (LHH)
(Département
Microbiologie - Pôle
Biologie)



Department of Biomedical Informatics

MDs	Engineers	University	Librarians	Secretaries
Stéfan DARMONI (Prof.), head	Badisse DAHAMNA	Lina Soualmia (Senior lecturer, computer science)	Benoit THIRION (Head of the Medical Library)	Annie-Claude LANCELEVEE
Jean-Philippe LEROY	Ivan KERGOURLAY	Chloé CABOT (PhD student)	Catherine LETORD (Pharmacist, librarian)	Sandrine VOURIOT
Nicolas GRIFFON	Julien GROSJEAN (PostDoc)	Wiem CHEBIL (PhD student)	Gaétan KERDELHUE (librarian)	
Matthieu SCHUERS (GP, PhD student)	Romain LELONG (PhD student)	Melissa Mary (PhD student, Grant CIFRE BioMérieux)	Léa SEGAS (librarian)	
2 residents in Public Health				
1 ou 2 residents in GP				
Philippe MASSARI (retired)				
	Financement CHU	Financement Université Rouen	Financement Conseil Régional Normandie	Financement sur projets de recherche



HeTOP content

- HeTOP is a repository dedicated to (European) health professionals and students. URL: www.hetop.eu; version NoSQL (beta-test): cispro.chu-rouen.fr/hetop
- HeTOP provides access to 69 health terminologies and ontology (T/O) available mainly in French or in English, but also German, Italian and Dutch (European languages) but also with no Latin alphabet (Greek, Russian) and more recently outside Europe (Japanese, Mandarin, Arabic & Hebrew) (**32 different languages**).
- HeTOP can be used by humans and by computers via Web services.
- The main objective of HeTOP is to provide an access to terminologies and ontology, allowing dynamic browsing and navigation.
 - Free portal for over 20 T/O: e.g. MeSH, CISMeF, ICD10, & CCAM; extended access restricted by ID/pwd for academic use only
 - To be included next (Lausanne collaboration): extended ICD10 + CHOP



HeTOP content

- HeTOP provides the usual data for each concept: preferred terms, original code, synonyms, definitions and other attributes, relations and hierarchies.
- Double (matricial) navigation:
 - among T/O
 - among languages
- Time consuming task > 20 man-years (to develop) + 2 man-years per year to maintain (integration & maintenance of T/O + mappings)
- Time consuming task to translate terminologies +++
- Several services on demand
 - access to other resources on the Internet (PubMed, CISMeF, etc.) through a French InfoButton (InfoRoute)
 - access to mappings tools (integrated in a beta version)
 - access to automatic indexing tool (ECMT)



Methods

To integrate terminologies and ontology into EHTOP, three steps are necessary:

- (1) designing a meta-model into which each terminology and ontology can be integrated,
- (2) developing a process to include terminologies into EHTOP,
- (3) building and integrating existing and new inter & intra-terminology semantic harmonization into EHTOP.



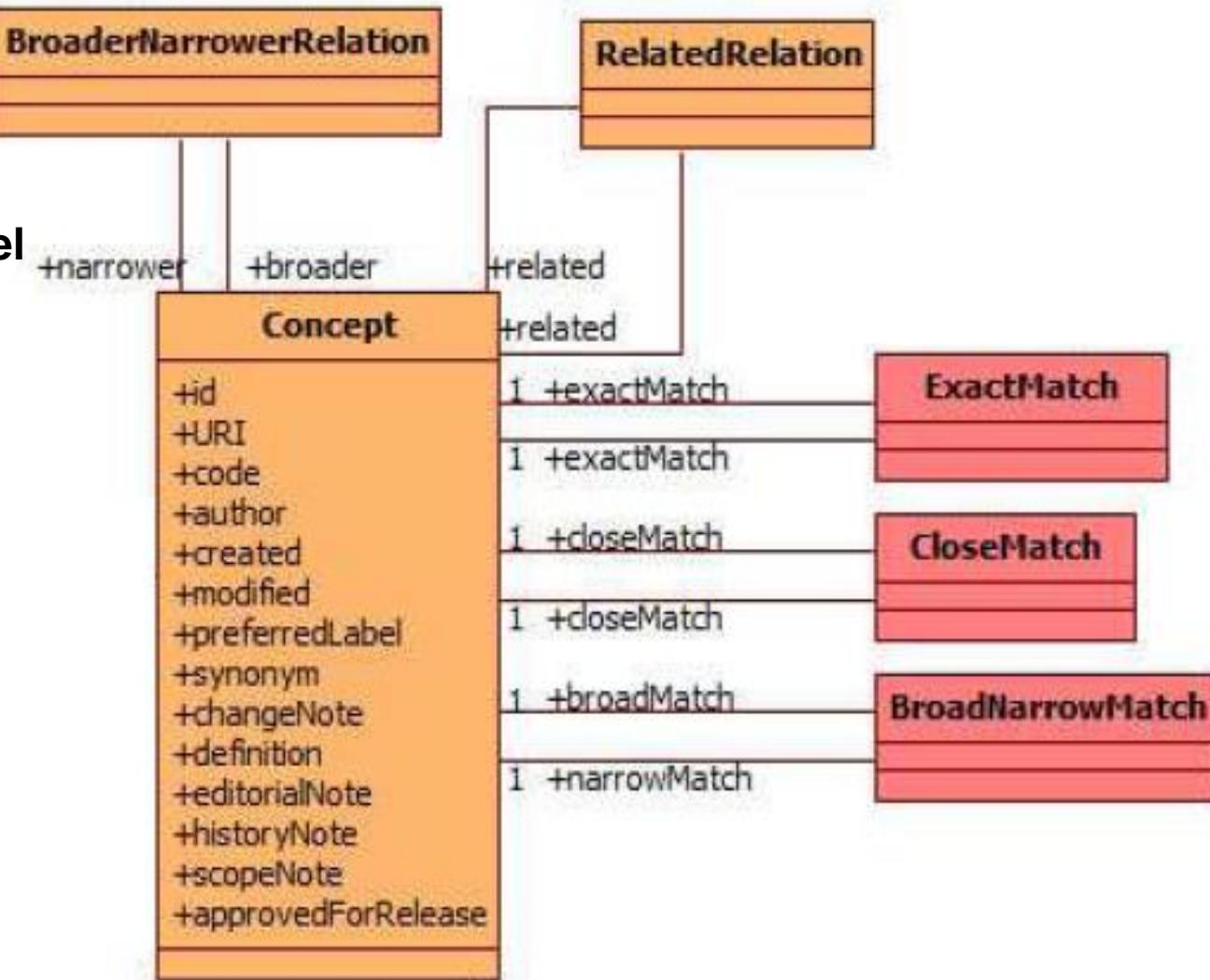
HeTOP generic model

Compliant with
ISO Terminology model
More simple

Caption

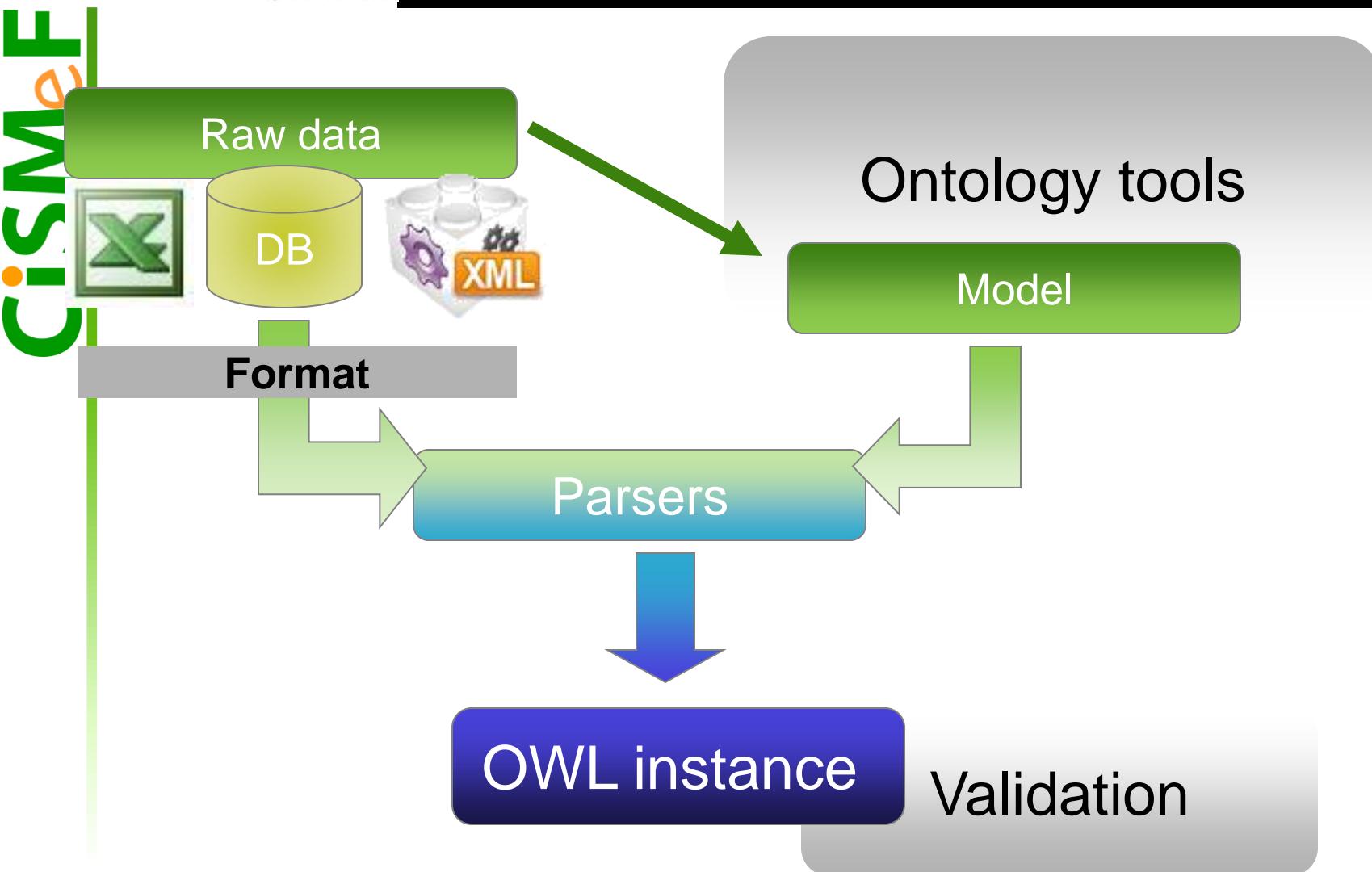
Alignment part

Conceptual part



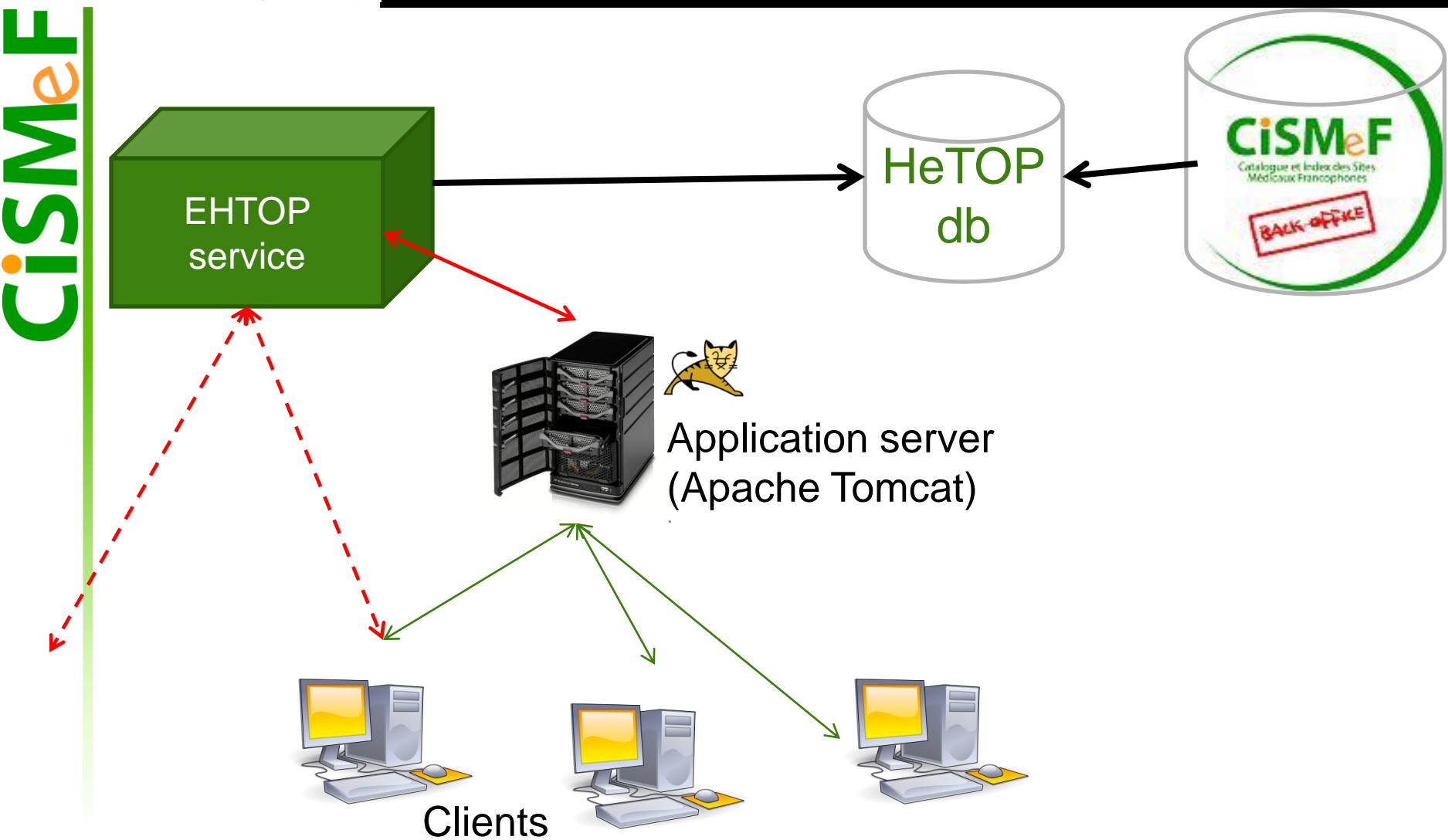


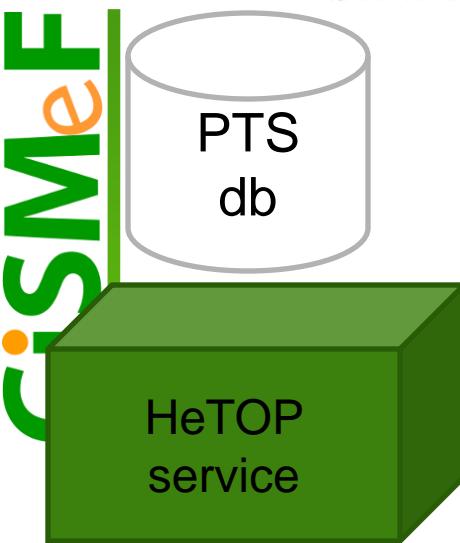
Integration: OWL instances





Methods & technologies (1)





Methods & technologies (2)

- Oracle 11.1g (optimizations & partitionning) => NoSQL in 2015
- Java J2EE
- CISMeF APIs
- Apache Tomcat
- Infinispan cache layer
- Cross-browser (Vaadin framework)
=> new framework in 2016 (INSA Rouen Engineering School)



Croslingual Health Multi-Terminology/Ontology Portal

- First version before HeTOP (French & English)
- URL: <http://pts.chu-rouen.fr/>
- Access for humans and computers (Web services)
 - Since September 2010, daily used by CISMeF team to index manually and automatically Web resources
 - Since January 2011, MeSH is freely available (500 unique users per working day)
 - Teaching tool: Rouen Medical School (since Sept. 2010) to teach anatomy and rare diseases
 - Terminology auditing: HPO/Orphanet
 - T/O translations into French: FMA, HPO, SNOMED CT, MEDLINEplus
- Restricted access to the other terminologies (2,250 registered)
- Cooperation with BioPortal: Clement Jonquet & Mark Musen (ANR Jeunes Chercheurs: project SIFR)

94 entrées trouvées en 2,39 s ★★★

Sélection terminologies

Vos recherches (2)

Historique des consultations (4)

Résultats

MeSH (21)

Descripteur MeSH (8)

- antiasthmatiques
- asthme
- asthme à l'effort
- asthme induit par l'aspirine
- asthme professionnel
- dyspnée paroxystique
- état de mal asthmatique
- remodelage des voies aériennes

MeSH Concept Supplémentaire (4)

- asthme et polypes nasaux
- asthme, petite taille et IgA élevés
- asthme, polypes nasaux et intolérance à l'aspi
- dermatite atopique, type 3

MeSH Concept (9)

- Asthme
- Asthme à l'effort
- Asthme cardiaque
- asthme et polypes nasaux
- Asthme induit par l'aspirine
- Asthme professionnel
- asthme, petite taille et IgA élevés
- asthme, polypes nasaux et intolérance à l'aspi
- Remodelage des voies aériennes dans l'asth...

CIM-9 (8)

CIM-10 (18)

DRC (4)

HPO (3)

MedlinePlus (2)

NCIt (2)

Concept NCIt (2)

- asthme
- Asthme chronique obstructif

OMIM (2)

RADLEX (1)

SNOMED int. (30)

WHO-ART (3)

Description Hiérarchies Relations PubMed / Doc'CISMef

Asthme (Descripteur MeSH) 

Ressource consultée 1675 fois

Afficher toutes les langues

 BioPortal  NLM  Inserm

Libellé préféré

 asthme
 asthma

Identifiant d'origine

D001249

Definition du MeSH

 Forme de maladie bronchique présentant une obstruction des voies respiratoires, marquée par des attaques récurrentes de dyspnée paroxysmale avec sifflements dûs à la contraction spasmotique des bronches. [Traduction effectuée avant 2008]
 A form of bronchial disorder with three distinct components: airway hyper-responsiveness (RESPIRATORY HYPERSENSITIVITY), airway INFLAMMATION, and intermittent AIRWAY OBSTRUCTION. It is characterized by spasmotic contraction of airway smooth muscle, WHEEZING, and dyspnea (DYSPNEA, PAROXYSMAL).

Synonyme CISMeF

 asthmas, bronchial  bronchial asthmas

Synonyme MeSH

 Asthme bronchique  asthmas  bronchial asthma

CUI UMLS

C0004096

VIDAL

 Maladie caractérisée par une difficulté à respirer, se traduisant souvent par des sifflements. L'asthme, permanent ou survenant par crise, est dû à un spasme et à une inflammation des bronches.

url (CISMef)

CISMef

CISMef

Annotation MeSH

 note specifics; ASTHMA, CARDIAC see DYSPNEA, PAROXYSMAL is also available; in historical literature consider indexing "phthisic" here; "phthisis" probably goes under TUBERCULOSIS, PULMONARY

Icone VCM hiérarchique

94 entrées trouvées en 2,39 s ★★★

Sélection terminologies

Vos recherches (2)

Historique des consultations (4)

Résultats

MeSH (21)

- Descriptor MeSH (8)
 - antiasthmatiques
 - **asthme**
 - asthme à l'effort
 - asthme induit par l'aspirine
 - asthme professionnel
 - dyspnée paroxystique
 - état de mal asthmatique
 - remodelage des voies aériennes

MeSH Concept Supplémentaire (4)

- asthme et polypes nasaux
- asthme, petite taille et IgA élevée
- asthme, polypes nasaux et intolérance à l'aspi
- dermatite atopique, type 3

MeSH Concept (9)

- Asthme
- Asthme à l'effort
- Asthme cardiaque
- asthme et polypes nasaux
- Asthme induit par l'aspirine
- Asthme professionnel
- asthme, petite taille et IgA élevée
- asthme, polypes nasaux et intolérance à l'aspi
- Remodelage des voies aériennes dans l'asthme

CIM-9 (8)

CIM-10 (18)

DRC (4)

HPO (3)

MedlinePlus (2)

NCl (2)

Concept NCl (2)

- asthme
- Asthme chronique obstructif

OMIM (2)

RADLEX (1)

SNOMED int. (30)

WHO-ART (3)

Description Hiérarchies Relations PubMed / Doc'CISMeF

Asthme (Descripteur MeSH)

Ressource consultée 1675 fois

BioPortal NLM Inserm

Libellé préféré

- France asthme
- Anglais asthma
- Allemand astma
- Espagnol asthma
- Italien Astma
- Danois astma
- Portugais Asma
- Néerlandais Asthma
- Norvégien Astma
- Russe АСТМА БРОНХИАЛЬНАЯ
- Allemand ACTMA БРОНХИАЛЬНАЯ
- Anglais Astma
- Chinois 哮喘

Identifiant d'origine

D001249

Definition du MeSH

- Forme de maladie bronchique présentant une obstruction des voies respiratoires, marquée par des attaques récurrentes de dyspnée paroxysmale avec sifflements dûs à la contraction spastique des bronches. [Traduction effectuée avant 2008]
A form of bronchial disorder with three distinct components: airway hyper-responsiveness (RESPIRATORY HYPERSENSITIVITY), airway INFLAMMATION, and intermittent AIRWAY OBSTRUCTION. It is characterized by spasmotic contraction of airway smooth muscle, WHEEZING, and dyspnea (DYSPNEA, PAROXYSMAL).

Synonyme CISMeF

- Anglais asthmas, bronchial
Anglais Alternative helseforskringssystemer

Synonyme MeSH

- Anglais Asthme bronchique
Anglais asthma, bronchial

Anglais asthmas

Anglais bronchial asthma

CUI UMLS

94 entrées trouvées en 2,39 s ★★★

Sélection terminologies

Vos recherches (2)

Historique des consultations (4)

Résultats

MeSH (21)

Descripteur MeSH (8)

- antiasthmatisques
- asthme
- asthme à l'effort
- asthme induit par l'aspirine
- asthme professionnel
- dyspnée paroxystique
- état de mal asthmatique
- remodelage des voies aériennes

MeSH Concept Supplémentaire (4)

- asthme et polyopes nasaux
- asthme, petite taille et IgA élevés
- asthme, polyopes nasaux et intolérance à l'aspirine
- dermatite atopique, type 3

MeSH Concept (9)

- Asthme
- Asthme à l'effort
- Asthme cardiaque
- asthme et polyopes nasaux
- Asthme induit par l'aspirine
- Asthme professionnel
- asthme, petite taille et IgA élevés
- asthme, polyopes nasaux et intolérance à l'aspirine
- Remodelage des voies aériennes dans l'asthme

CIM-9 (8)

CIM-10 (18)

DRC (4)

HPO (3)

MedlinePlus (2)

NCIt (2)

Concept NCIt (2)

- asthme
- Asthme chronique obstructif

OMIM (2)

RADLEX (1)

SNOMED int. (30)

WHO-ART (3)

Description Hiérarchies Relations PubMed / DocCISMéF

Asthme (Descripteur MeSH) ↴

Intra-terminologiques Inter-terminologiques

Liste des qualificatifs affiliables (37)

Voir aussi (2)

Type(s) sémantique(s) (1)

maladie ou syndrome

Type sémantique

Concepts Supplémentaires MeSH en relation (4)

Concept(s) lié(s) au record (1)

Métaforme(s) (2)

Informations(s) d'indexation (20)

Topic(s) MedlinePlus (2)

asthme

Topic MedlinePlus

asthme chez l'enfant

Topic MedlinePlus

Alignements manuels CISMeF (1/2)

asthme chez l'enfant

Topic MedlinePlus

Alignements automatiques CISMeF supervisés (9/14)

0804493 - asthme

Code CIM-9

asthme

Topic MedlinePlus

asthme

Notion SNOMED

ASTHME

Résultat de consultation DRC

asthme

Terme préféré WHO-ART

Asthme

RCE DRC

asthme

Terme HPO

J45

Concept NCIt

J45 - asthme

Catégorie CIM-10

Correspondances UMLS (même concept) (7/17)

C0004096

493.9 - asthme, sai

Code CIM-9

C0004096

asthme

Notion SNOMED

C0004096

ASTHME

Terme préféré WHO-ART

C0004096

asthme

Topic MedlinePlus

C0004096

asthme

Concept NCIt

C0004096

J45 - asthme

Catégorie CIM-10

C0004096

J45.9 - asthme, sans précision

Sous Catégorie CIM-10

Alignements automatiques supervisés en BTNT (1/2)

493.9 - asthme, sai

Code CIM-9

Alignements automatiques exacts (par équipe CISMeF) (3/9)

493.9 - asthme, sai

Code CIM-9

Asthme

Concept Radlex

prédisposition à l'asmthe

Phénotype OMIM

Alignements automatiques faux (1/3)

94 entrées trouvées en 2,39 s ★★★

Sélection terminologies

Vos recherches (2)

Historique des consultations (4)

Résultats

MeSH (21)

Descripteur MeSH (8)

- antiasthmatiques
- **asthme**
- asthme à l'effort
- asthme induit par l'aspirine
- asthme professionnel
- dyspnée paroxystique
- état de mal asthmatique
- remodelage des voies aériennes

MeSH Concept Supplémentaire (4)

- asthme et polypes nasaux
- asthme, petite taille et IgA élevée
- asthme, polypes nasaux et intolérance à l'aspi
- dermatite atopique, type 3

MeSH Concept (9)

- Asthme
- Asthme à l'effort
- Asthme cardiaque
- asthme et polypes nasaux
- Asthme induit par l'aspirine
- Asthme professionnel
- asthme, petite taille et IgA élevée
- asthme, polypes nasaux et intolérance à l'aspi
- Remodelage des voies aériennes dans l'asthme

CIM-9 (8)

CIM-10 (18)

DRC (4)

HPO (3)

MedlinePlus (2)

NCIt (2)

Concept NCIt (2)

- asthme
- Asthme chronique obstructif

OMIM (2)

RADLEX (1)

SNOMED int. (30)

WHO-ART (3)

Description

Hierarchies

Relations

PubMed / Doc'CISMef

Asthme (Descripteur MeSH) 

Qualificatif(s) utilisable(s) pour ce mot clé :

Afficher la liste alphabétique des qualificatifs

analyse

- liquide céphalorachidien
- sang
- urine

anatomie et histologie

cytologie

- anatomie pathologique

embryologie

classification

diagnostic

- anatomie pathologique

échographie

radiographie

scintigraphie

étiologie

complications

congénital

embryologie

génétique

immunologie

induit chimiquement

microbiologie

virologie

parasitologie

histoire

médecine vétérinaire

organisation et administration

économie

physiologie

génétique

immunologie

métabolisme

enzymologie

liquide céphalorachidien

sang

urine

physiopathologie

psychologie

statistiques et données numériques

épidémiologie

éthnologie

mortalité

thérapie

chirurgie

diétothérapie

prévention et contrôle

radiothérapie

rééducation et réadaptation

soins infirmiers

traitement médicamenteux

toutes les ressources

tous les types

seulement les principales

recommandations professionnelles

sans explosion

documents concernant l'enseignement

documents concernant les patients

documents SIGAPS A

documents SIGAPS A ou B

documents SIGAPS A ou B



Three main terminology servers in health

UMLS

- NIH, Bethesda (USA)
- Plus de 150 T/O
- Essentiellement en anglais
- La référence internationale pour la diffusion, mais pas pour la consultation

BioPortal*

- NCBO, Stanford (USA)
- Plus de 400 T/O (beaucoup en biologie, avec peu de concepts)
- Essentiellement en anglais (silo pour les autres langues)
- La référence pour poster une ontologie

HeTOP*

- SIBM, Rouen (France)
- 70 T/O en 32 langues
- La référence inter-lingue (navigation entre les langues) et dans le monde francophone

*Grosjean J et coll. An Approach to Compare Bio-Ontologies Portals. **Stud Health Technol Inform**, 2014;205:1008-1012.



HeTOP: main figures

May 2010

Terminologies & ontologies	Concepts	Synonymes	Définitions	Relations & hiérarchies
25	> 580 000	> 840 000	> 220 000	> 1 200 000

May 2011

Terminologies	Concepts	Synonymes	Définitions	Relations
32	> 980 000	> 2 300 000	> 220 000	> 4 000 000

April 2013

Terminologies	Concepts	Synonymes	Définitions	Relations
45	≈ 1 620 000	≈ 3 700 000	≈ 220 000	≈ 5 500 000

February 2016

Terminologies	Concepts in English	Concepts in French	Terms	Attributes	Relations
69 (17 UMLS)	2,358,544	1,171,340	9,134,636	15,348,985	9,697,000



Main figures

Registered users

> 2 600

traffic

15 000 hits/day
(6 to 700 users per working day)



Terminologies in French AND in UMLS (2015AA)

CiSMeF

Source vocabulary (Translator) [N lang.]	Number of Strings (PT + syn. + acro.)	Number of CUIs (% Translation in French)
MeSH Fr* [14]	105,758	41,229
MedDRA Fr [9]	73,860	73,608
WHO-ART Fr [5]	3,631	3,091
MTHMST Fr	1,833	1,636
ICPC2 [19]	702	722 (100%)
*MeSH Fr HeTOP		
Descriptors (INSERM) [16]	105,274	27,747 (100%)
Supplementary Concepts (CISMeF) [2]	58,514	42,693 (18.24%)
Concepts (INSERM & CISMeF) [2]	99,184	94,052 (26.20%)



Terminologies in UMLS NOT in French

Source vocabulary (Translator) [N lang.]	Number of Strings	Number of CUIs (% of translation Fr)
ICD10 Fr (WHO) [12]	26,337	12,143 (100%)
ICD10 PCS (CISMeF) [2]	7,297	7,297 (5%)
ICD9 Fr (WHO) [1]	10,716	7,356 (100%)
ICDO Fr (WHO+CISMeF) [3]	1,462	1,362 (100%)
ICF (WHO) [2]	1,496	1,495
FMA Fr (U. of Washington)	4,564	4,452
FMA Fr in HeTOP (CISMeF) [7]	16,631	16,369 (20.20%)
ICNP [2]	2,811	1,158 (92%)
SNOMED Int. [2]	139,792	96,756 (94.51%)
ATC (WHO) [3]	5,834	5,757 (100%)
SNOMED CT		126,492 (41.85%)



Terminologies in UMLS NOT in French

CISMeF

Source vocabulary (Translator) [N lang.]	Number of Strings Fr/En/% of translations	Number of CUIs via CISMeF mappings
WHO-ICPS [2]	424	105
RADLEX (CISMeF) [2]	9,465/42,313/22.30	240
LOINC (APHP & SFIL) [2]	58,950	58,500 (60.71%)
MEDLINEplus Fr (CISMeF & LIMSI) [2]	849	846 (100%)
NCIT Fr (CISMeF) [2]	57,827	56,727 (60.40%)



Terminologies in French that are not included in UMLS

CiSMeF

Source vocabulary (Translator) [N lang.]	Number of Strings	Number of CUIs
OMIM Fr (CISMeF) [2]	7,770	6,904 (88.84%)
HRDO (Orphanet) [2]	13,535/id/100	4,943
HPO (CISMeF) [2]	11,127/11,908/93.44	1,541
CCAM (procedure) [1]	10,121	0
BNPC (toxicology) [2]	91,751	11,539
Q-Codes	184	
... including interface terminologies	in biology and imaging	

Overall, number of distinct CUI with at least one French translation in HeTOP
≈ 333,239 vs. ≈ 88,000 in UMLS (x3.79)

108 millions of RDF triplets (big data in health) in 2014



HeTOP relationships (examples & numbers)

CiSMeF

	Source Term (Terminology)	Target Term (Terminology)	Number of relations in HeTOP
UMLSalignment	<i>Myocardial Infarction</i> (MeSH)	<i>Myocardial infarction,</i> <i>NOS</i> (SNOMED Int)	537,978
CISMeFmanual	<i>Riedel thyroiditis</i> (HRDO)	<i>Riedel's thyroiditis</i> (MedDRA)	56,265
CISMeFexact	<i>appetite stimulants</i> (ATC)	<i>Appetite stimulated</i> (WHOART)	653,709
CISMeFSupervised	<i>Gonadotropin releasing hormone</i> (MeSH)	<i>Luteotropin-releasing factor</i> (FMA)	286,561



HeTOP limits

- Formal representation of complex clinical data structures = none
- Formal representation of physiological models = none
- Temporal relations = none
- Data quality = based on T/O quality and point of view
- Formalism & reasoning capabilities = none
- Collaborative editing/searching/sharing tools = collaboration with BioPortal to share tools (Clement Jonquet)
- T/O versioning = not yet provided by HeTOP
- Semantic resources distribution/dissemination processes = 56 T/O available in OWL format (latest version)/SKOS/RDF in several languages



Other tools integrated HeTOP

- ECMT Extracteur de Concepts Multi Terminologiques
 - Able to extract health concepts from any text; e.g. discharge summary in $\frac{1}{2}$ second (NoSQL)
 - Valorization with Alicante SME
 - Used in daily practice in the Catholic University Hospital of Lille, France; Dr. Arnaud Hansske; around one million discharge summaries indexed with ECMT
- InfoRoute, a French InfoButton
 - URL: inforoute.chu-rouen.fr
 - Access to a contextualized knowledge based on semantic expansion based on manual & supervised mapping among terminologies
- MT@HeTOP, tool to perform automatic mappings & translations
- Generic semantic search engine
 - Doc'CISMeF (URL: doccismef.chu-rouen.fr) on grey literature about health in French on the Internet (10^5 resources)
 - LISSA (URL : www.lissa.fr), a PubMed in French ($0,7 \times 10^6$ citations d'articles)
 - RIDOPI, search engine in EHR (8×10^6 discharge summaries in Rouen; around 10^9 health concepts in these summaries; 10^8 numerical data in Rouen)

Examens paracliniques : * Biologie : globule plaquettes 245 000 mm³, TP 88 , ionogramme 3,8 mmol l, ASAT 23, ALAT 25, phosphatase Radiographie pulmonaire : récidive d'un pneumonie régulier et sinusal sans trouble de la repolarisation supra aortiques et trans- crânien (23 04): thrombose interne gauche dès son origine, sténose de l'irrigation du bulbe ancienne calcifiée, évaluation artères vertébrales droite et gauche. Pas d'au niveau du polygone de Willis et notamment postérieures. * Angioscanneur des TSA (12 04) : présence d'une infiltration athéromateuse au niveau de la crosse aortique. Infiltration athéromateuse calcifiée au niveau de l'artère carotide commune. Au niveau du bulbe, infiltration athéromateuse plus importante avec présence d'une plaque calcifiée quasi circonférentielle. Thrombose complète de la carotide interne sur tout son trajet depuis le bulbe jusqu'à sa portion terminale. A droite, infiltration athéromateuse prédominant également au niveau du bulbe avec présence de plaque calcifiée sans sténose significative retrouvée. Calcifications au niveau de l'artère carotide interne dans sa portion intra caverneuse. Sténose moyennement serrée de l'origine de la vertébrale droite. Plusieurs boucles vasculaires sur le trajet de la vertébrale droite. Artère vertébrale gauche est le siège de boucle vasculaire. Pas de sténose significative retrouvée. Au niveau du polygone de Willis, celui-ci est bien opacifié avec présence d'une artère communicante antérieure et des deux artères communicantes postérieures. * Scanner thoracique sans et avec injection : plusieurs adénopathies centimétriques connues stables, épanchement péricardique gauche, dont l'épaisseur maximale mesure 16,4 mm. Bonne opacification des gros vaisseaux médiastinaux. On retrouve la lésion bourgeonnante au sein de la lumière de la bronche souche gauche mesurée ce jour à 22 x 33 mm, responsable d'une atélectasie complète du poumon gauche. Présence d'un emphysème sous cutané thoracique antérieur gauche et axillaire gauche en rapport avec le drainage thoracique. * Echographie de stress : non réalisable en raison du pneumothorax. Evolution dans le service : Sur le plan respiratoire : le patient bénéficie de la pose d'un drain thoracique à gauche pour évacuation de l'épanchement aérique. La radiographie pulmonaire après mise en place montre un décollement persistant ainsi qu'une atélectasie du lobe inférieur gauche, avec ascension de la coupole diaphragmatique. Le drain est donc mis en aspiration sans amélioration de la radiographie thoracique. Une fibroscopie bronchique est réalisée et montre une obstruction par un caillot frais au niveau du tronc souche gauche expliquant la non réexpansion du poumon. Compte tenu de la taille du caillot, nous programmons une bronchoscopie au tube rigide sous anesthésie générale le 16 04 2009 qui permet la thermocoagulation de la lésion qui apparaît comme un caillot, puis extraction de celui-ci. Après extraction, il existe un saignement régulier issu de B6 qui inonde l'arbre bronchique gauche. On procède à un nettoyage au fibroscope souple de tout le lobe supérieur qui apparaît sain, sans anomalie. La lobaire inférieure est totalement obstruée par un caillot. La lobaire inférieure est cathétérisable mais l'orifice d'entrée est totalement déformé, refoulé par une tumeur issue de B6 saignant facilement au contact. On termine le geste en thermocoagulant B6. Des biopsies de carène, épicondyle lobaire supérieur et inférieur gauche sont réalisées et sont négatives.

artère carotide commune ([MSH_D_017536](#))

- catégorie(s) : médecine et chirurgie vasculaire; cardiologie; partie du corps, organe ou composant d'un organe; système d'organisme;
- ascendant(s) : arborescence MeSH ([MSH_D_ARBO](#)); artères carotides ([MSH_D_002339](#)); système cardiovasculaire ([MSH_D_002319](#)); artères ([MSH_D_001158](#)); Anatomie ([MSH_D_A](#)); vaisseaux sanguins ([MSH_D_001808](#));
- descendant(s) : artère carotide interne ([MSH_D_002343](#)); artère carotide externe ([MSH_D_002342](#));
- relié(s) : artère carotide primitive, saï ([SNO_NO_T-45100](#));

ancien

anesthésie

anesthésie générale

anesthésie générale

anesthésie pour bronchoscopie

anomalie

antérieur

aorte, saï

appareil respiratoire, saï

arbre

artère carotide commune

artère carotide interne

artère cérébrale postérieure

artère communicante antérieure

artère communicante postérieure

artère vertébrale

artère vertébrale droite

artère vertébrale gauche

artères

aspartate aminotransferases

aspiration

aspiration (technique)

atélectasie

atélectasie complète

atélectasie pulmonaire

athérome

B05BC02 - urée

biopsie

biopsie

bon

bronche souche gauche

bronche, saï

bronches

Ter. Code	CUI	Cond. Ctxt.
SNO G-C410	C0332310	
SNO G-C470	C1331475	
SNO F-06305	C0231301	
MSH D000124	C0001072 non	
SNO T-D8100	C0004454	
MSH D000410	C0001899	25,
SNO G-A463	C0580836	
MED T140	C0002903	16
SNO P1-C0010	C0002915	16
MSH D000768	C0002915	16
SNO P1-C2A04	C0198953	16
SNO M-01130	C1457869 pas	
SNO G-A105	C1704448	
SNO T-42000	C0003483	
SNO T-20000	C0521346	
SNO L-D0012	C0040811	
MSH D017536	C0162859	
MSH D002343	C0007276	
MSH D020769	C0149576	
SNO T-45530	C0149562	
SNO T-45320	C0149559	
MSH D014711	C0042559	
SNO T-45710	C0226230	
SNO T-45720	C0226231	
MSH D001158	C0003842	
MSH D001219	C0004002	23,
SNO F-29200	C0700198	
MSH D013396	C0038638	
SNO D2-60300	C0004144	
SNO D2-60306	C0264497	
MSH D001261	C0004144	
SNO M-52100	C0264956	
ATC B05BC02	C0041942	3,8
SNO P1-03100	C0005558	
MSH D001706	C0005558	
SNO G-A223	C0205170	
SNO T-26500	C0225630	
SNO T-26000	C0205039	
MSH D001980	C0006255	



Metrics of ECMT (April 2015)

CLEF eHealth in French

	TP	FP	FN	P	R	F1
Anatomy	142	149	54	0.4880	0.7245	0.5832
Chemistry	153	38	108	0.8010	0.5862	0.6670
Devices	13	12	6	0.5200	0.6842	0.5909
Disorders	375	96	209	0.7962	0.6421	0.7109
Geography	14	4	7	0.7778	0.6667	0.7179
Live Beings	125	38	31	0.7669	0.8013	0.7837
Objects	3	16	28	0.1579	0.0968	0.1200
Phenotype	14	35	17	0.2857	0.4516	0.3500
Physiology	60	33	74	0.6452	0.4478	0.5286
Procedure	195	105	109	0.6500	0.6414	0.6457
Overall	1094	526	643	0.6753	0.6298	0.6518



Results obtained in 2015 and 2016 from the training set (832 documents) provided by CLEF e-Health

	Exact match			Inexact match		
	Precision	Recall	F-measure	Precision	Recall	F-measure
2015 results	0,2660	0,1285	0,1732	0,6848	0,5113	0,5855
2016 results	0,4643	0,6154	0,5293	0,6886	0,7835	0,7330
Variation	+74%	+379%	+206%	+0,55%	+53%	+25%

- 2015 : participation in the CLEF e-Health challenge
(<https://sites.google.com/site/clefhealth2015/>)
- Addresses clinical named entity recognition in languages other than English
- 2016 : Deduplication of identical terms from multiples terminologies
Reevaluation in the 2016 CLEF e-Health challenge



Detailed 2016 results obtained from the training set (832 documents) provided by CLEF e-Health

2016	Entity Recognition : exact match						Entity Recognition : inexact match					
	TP	FP	FN	Precision	Recall	F1	TP	FP	FN	Precision	Recall	F1
ANAT	151	340	64	0,3075	0,7023	0,4278	216	275	46	0,4399	0,8244	0,5737
CHEM	172	172	106	0,5000	0,6187	0,5531	257	87	63	0,7471	0,8031	0,7741
DEVI	17	22	9	0,4359	0,6538	0,5231	23	16	5	0,5897	0,8214	0,6866
DISO	465	485	267	0,4895	0,6352	0,5529	789	161	134	0,8305	0,8548	0,8425
GEOG	26	7	12	0,7879	0,6842	0,7324	28	5	11	0,8485	0,7179	0,7778
LIVB	170	124	94	0,5782	0,6439	0,6093	223	71	67	0,7585	0,7690	0,7637
OBJC	8	19	36	0,2963	0,1818	0,2254	8	19	36	0,2963	0,1818	0,2254
PHEN	13	47	41	0,2167	0,2407	0,2281	15	45	40	0,2500	0,2727	0,2609
PHYS	81	79	74	0,5063	0,5226	0,5143	102	58	67	0,6375	0,6036	0,6201
PROC	276	296	159	0,4825	0,6345	0,5482	384	188	96	0,6713	0,8000	0,7300
Overall	1 379	1 591	862	0,4643	0,6154	0,5293	2 045	925	565	0,6886	0,7835	0,7330

- Disparities between UMLS semantic groups :
 - ANAT (anatomy), DISO (disorders) : ICD10, MeSH, FMA...
 - OBJC (objects, food), PHEN (natural, environmental phenomenon or process) : few terminological resources

2015	Entity Recognition : exact match						Entity Recognition : inexact match					
	TP	FP	FN	Precision	Recall	F1	TP	FP	FN	Precision	Recall	F1
ANAT	90	401	498	0,1833	0,1531	0,1668	225	266	155	0,4582	0,5921	0,5166
CHEM	98	246	1 064	0,2849	0,0843	0,1301	250	94	211	0,7267	0,5423	0,6211
DEVI	7	32	56	0,1795	0,1111	0,1373	22	17	29	0,5641	0,4314	0,4889
DISO	276	674	1 782	0,2905	0,1341	0,1835	784	166	537	0,8253	0,5935	0,6904
GEOG	23	10	31	0,6970	0,4259	0,5287	28	5	25	0,8485	0,5283	0,6512
LIVB	94	200	381	0,3197	0,1979	0,2445	196	98	153	0,6667	0,5616	0,6096
OBJC	4	23	117	0,1481	0,0331	0,0541	12	15	105	0,4444	0,1026	0,1667
PHEN	11	49	130	0,1833	0,0780	0,1095	19	41	117	0,3167	0,1397	0,1939
PHYS	41	119	441	0,2563	0,0851	0,1277	93	67	305	0,5813	0,2337	0,3333
PROC	146	426	860	0,2552	0,1451	0,1850	405	167	307	0,7080	0,5688	0,6308
Overall	790	2 180	5 360	0,2660	0,1285	0,1732	2 034	936	1 944	0,6848	0,5113	0,5855
2016	Entity Recognition : exact match						Entity Recognition : inexact match					
	TP	FP	FN	Precision	Recall	F1	TP	FP	FN	Precision	Recall	F1
ANAT	151	340	64	0,3075	0,7023	0,4278	216	275	46	0,4399	0,8244	0,5737
CHEM	172	172	106	0,5000	0,6187	0,5531	257	87	63	0,7471	0,8031	0,7741
DEVI	17	22	9	0,4359	0,6538	0,5231	23	16	5	0,5897	0,8214	0,6866
DISO	465	485	267	0,4895	0,6352	0,5529	789	161	134	0,8305	0,8548	0,8425
GEOG	26	7	12	0,7879	0,6842	0,7324	28	5	11	0,8485	0,7179	0,7778
LIVB	170	124	94	0,5782	0,6439	0,6093	223	71	67	0,7585	0,7690	0,7637
OBJC	8	19	36	0,2963	0,1818	0,2254	8	19	36	0,2963	0,1818	0,2254
PHEN	13	47	41	0,2167	0,2407	0,2281	15	45	40	0,2500	0,2727	0,2609
PHYS	81	79	74	0,5063	0,5226	0,5143	102	58	67	0,6375	0,6036	0,6201
PROC	276	296	159	0,4825	0,6345	0,5482	384	188	96	0,6713	0,8000	0,7300
Overall	1 379	1 591	862	0,4643	0,6154	0,5293	2 045	925	565	0,6886	0,7835	0,7330



Semantic harmonization: mapping, alignment

Three methods employed

URL: http://cispro.chu-rouen.fr/MT_EHTOP/

- Conceptual
 - Same CUI
 - Other relations: close match, BT-NT, NT-BT (SKOS)
 - On UMLS (n=17 included in HeTOP)
- NLP
 - More or less same algorithm of automatic indexing
 - Bag of words
 - on $n^*(n-1)/2$ T/O (included in the HeTOP)
- Statistical
 - Co-occurrence matrix
 - CCAM-ICD10; CCAM-LPP

Terminologies selection

Your queries (1)

Results

MeSH (10)

MeSH Descriptor (7)

anti-asthmatic agents

asthma

asthma, Aspirin-Induced

asthma, exercise-induced

asthma, occupational

conyza

status asthmaticus

MeSH Supplementary Concept (3)

asthmalgine

asthmasedine

NPSR1 protein, human

Record concept(s) (1)

Automatic exact mapping(s) (from CISMef team) (3)

 Asthma
 Asthma (disorder)
 Asthma finding (finding)

Validated automatic narrower mappings (3)

 493.9 - asthme, sai
 Asthma NOS
 asthme bronchique

Curated CISMef NLP mapping (13)

 .0804493 - asthme
 asthma
 asthma
 asthma
 Asthma
 asthma
 Asthma
 Asthma (disorder)
 asthma, nos
 Asthme
 asthme
 asthme
 asthme

False automatic mappings (3)

UMLS correspondence (same concept) (9)

 asthma
 Asthma
 asthma
 asthma
 asthma
 Asthma
 Asthma (disorder)
 asthma, nos
 Asthma, unspecified

examples: asthma, asthma, D001249.nu, asth,

 NClt concept
 SNOMED CT concept
 SNOMED CT concept

 T_DESC_CIM9CODE
 MedDRA Preferred Term
 TUV Concept

 T_DESC_CIM9CODE
 ICPC-2 Descriptor
 MedDRA Preferred Term
 MedlinePlus Topic
 ICD-10 category
 WHO-ART Preferred Term
 HPO term
 SNOMED CT concept
 SNOMED Notion
 TUV Term
 TUV Concept
 DRC Consultation result
 DRC RCE

 ICPC-2 Descriptor
 ICD-10 category
 MedDRA Preferred Term
 WHO-ART Preferred Term
 MedlinePlus Topic
 NClt concept
 SNOMED CT concept
 SNOMED Notion
 ICD-10 Sub-category

NLP

Bag of words

 Conceptual
 Same CUI

Système d'information du SIBM

2007

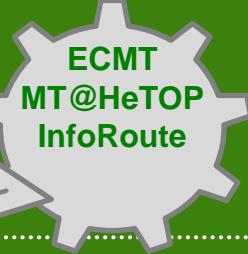


Alignement et traduction



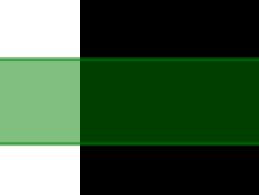
Terminologies de santé
N = 68
2,3 millions concepts

SIBM



indexation

1995



Ressources gratuites francophones
N = 110 000



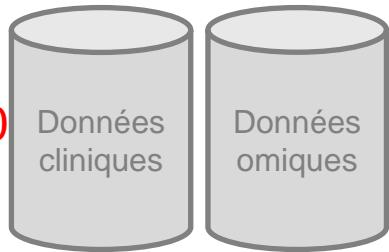
2012



Références d'articles scientifiques francophones
N = 900 000



2010



Données du dossier patient
N = 10^9 concepts médicaux





Many thanks

Questions

Email: Stefan.Darmoni@chu-rouen.fr

