

SYNODOS

Système de **N**ormalisation et d'**O**rganisation de **D**onnées médicales textuelles pour l'**O**bservation en **S**anté

18 ans de CISMéF
5 Février 2013, Rouen



Contexte - Justification

- **Le dossier patient informatisé (DPI)**
 - aide à la décision médicale
 - surveillance épidémiologique
 - data et semantic mining
- **Toutefois très peu de données sont structurées et codées dans les DPI pour permettre ce type d'utilisation**
 - A l'heure actuelle, seules les données nécessaires à la tarification à l'activité sont systématiquement structurées et codées dans les dossiers médicaux hospitaliers.
- **1^{ère} étude de faisabilité réalisée dans le cadre du projet ALADIN (ANR TecSan 2008)**
 - bonnes performances globales de détection des infections nosocomiales
 - mise en évidence de certains verrous technologiques à lever pour un déploiement industriel

Objectifs scientifiques

- **Développer une solution générique**
 - d'extraction sémantique de données médicales
 - organisant les données afin de les rendre exploitable à des fins épidémiologiques ou d'aide à la décision
 - Caractère générique du domaine d'application médicale assuré en permettant à l'utilisateur d'écrire lui-même ses règles métier
- **Evaluer la qualité des données extraites par la solution**
 - 2 domaines d'application : infections associées aux soins, cancer

Objectifs technologiques

- **Au niveau des expressions temporelles, enrichir les règles de détection des expressions temporelles**
- **Au niveau terminologique, interfacier l'analyseur sémantique et le serveur multi-terminologique en amont pendant la phase d'extraction**
- **Interfacier l'analyseur sémantique et le système de représentation des connaissances**
- **Générer de règles expert indépendamment des règles linguistiques**
- **Interfacier ces différents outils pour constituer une solution opérationnelle**

Le consortium SYNODOS

- **Coordonnateur**
 - Lyon 1 (LBBE) : MH Metzger
- **Partenariats universitaires et industriels**
 - Industriels : Objet Direct, CELI
 - Universitaires : CiSMeF, Lyon 1
- **Durée du projet**
 - 1^{er} octobre 2012 –
30 septembre 2015
- **Financement**
 - Coût Global = 2.0 M€
 - Aide allouée = 0.785 M€



Grandes Etapes du projet

- 1 Définition de l'architecture SYNODOS
- 2 Développement d'un outil d'anonymisation
- 3 Développement d'un outil d'aide à l'annotation : MedIndex
- 4 Identification et intégration au serveur terminologique CISMeF des terminologies pertinentes pour les 2 domaines d'application
- 5 Annotation du corpus pour organiser la base de connaissances et préparer la base d'évaluation
- 6 Développement de l'outil de traitement sémantique des données textuelles médicales
- 7 Interfaçage de l'outil sémantique et du serveur terminologique
- 8 Développement d'un système générateur de règles expert
- 9 Intégration des différents outils (sémantique, terminologie, règles expert, interface SIH)
- 10 Evaluation de la solution SYNODOS (qualité des données extraites)



Projet SYNODOS

SYstème de Normalisation et d'Organisation de Données
médicales textuelles pour l'Observation en Santé (ANR –
TecSan 2012)

CiSMeF Catalogue et Index des Sites
Médicaux de langue Française



Utilisation de l'outil terminologique

CiSM^eF

- Connexion du serveur sémantique (CELI) au serveur terminologique pour repérer les concepts médicaux
- Annotations manuelles : Appel à l'outil terminologique pour faire une proposition d'indexation automatique des dossiers patients servant à l'apprentissage du système (LBBE)



Utilisation des outils sémantiques

CISMeF

- Connexion du système expert (Objet Direct) au serveur terminologique pour exploiter les relations entre concepts (intra ou extra terminologiques)



Import de terminologie

CISMeF

- Import et traduction de l'US National Cancer Institute thesaurus (NCIt)
- Création d'une ontologie formelle adaptée aux infections associées aux soins à partir de terminologies existantes



Transmission des offset

CISMeF

- Réécriture du système d'indexation
 - Possibilité d'indiquer les mots qui ont permis de trouver un concept



MH Metzger
Q Gicquel
S Chariout



L Dini
M Rühlmann
A Bittar

*Ce projet a obtenu le financement de
l'Agence Nationale de Recherche
(programme TecSan 2012)*



SJ Darmoni
I Kergourlay



F Segond
P Arnod-Prin
C Lopez