



Entrepôts de Données de Santé

Objectifs pédagogiques

- Connaître la définition d'un EDS
- Connaître les objectifs principaux des EDS
- Appréhender la façon de construire et d'exploiter un EDS
- Appréhender les avantages et limites des EDS

EDS - Qu'est ce que c'est

- Les Entrepôts de Données de Santé (EDS) sont des outils informatiques permettant la collection, l'intégration puis le traitement des données de santé provenant d'un grand nombre de sources d'information clinique (dossier patient informatisé, système d'information des laboratoires et d'imagerie, prescription informatisée, dossier infirmier...).
- Agrégation d'un **maximum** d'informations disponibles sur les patients quelle que soit l'application source
- ➔ Permet de croiser ces informations et donc de sélectionner finement des patients/données

Objectifs des EDS

- Améliorer la recherche interventionnelle grâce aux études de faisabilité d'essais cliniques et l'optimisation des inclusions (*pré-screening*) ;
- Améliorer la recherche observationnelle (registres et cohortes) ;
- Détecter des profils de santé particuliers pour la prévention (primaire ou secondaire) ;
- Créer et maintenir des outils de détection d'événements liés à la vigilance (par exemple, les infections nosocomiales, pharmacovigilance...) ;
- Développer des indicateurs d'activité plus précis et variés ;
- Renforcer la prise en compte des indicateurs de qualité ;
- Améliorer le codage PMSI, par la détection semi-automatique d'atypie entre le codage réalisé et les données issues de l'EDS ;
- Participer à la pédagogie médicale ;
- Développer et évaluer les algorithmes d'aide à la décision.

Aujourd'hui dans les CH

- Large utilisation du Programme de Médicalisation des Systèmes d'Information (PMSI) via CIM-10 / CCAM
- Pour la recherche : « usage détourné » d'un système médico-économique
- Très pratique et fiable dans certains cas...
- ...mais beaucoup de biais et approximations

État des lieux des EDS

Dans le monde :

- i2b2  (Harvard MS)
- ...

En France :

-  (CHU de Rennes)
- **Dr Warehouse**  (Necker, Foch – Paris)
-  (Centre de Lutte Contre le Cancer)
Continuum Soins Recherche
- EDSaN ...

État des lieux des EDS hospitaliers en France



Types d'EDS

- Hospitaliers
- Médecine libérale
- CLCC
- Pharmacie

- Publiques (CHU...)
- Privés (IQVIA, THIN...)

Approche rouennaise pour l'EDS

- Sémantique :
 - Apport d'un lexique en français extrêmement riche (HeTOP)
 - RI améliorée (subsomption/explosion)
- Technologique :
 - Temps de réponses plus bas possible pour une utilisation en temps réel/interactive
 - Annotateur sémantique ECMT (BOW)
 - 6 ms/doc de santé vs. 134 ms/doc Quick UMLS vs. plusieurs secondes pour Metamap
 - 13 M de documents de santé en 22 h => possibilité de lancer ECMT chaque semaine
 - Moteur de recherche
 - I2B2 assez lent ; pas du tout temps réel
 - EDSaN : temps de réponse < 2s ; possibilité de tester de nombreuses hypothèses de requêtes lors d'une session de 2 heures

EDS Rouen

- Projet diligenté par la Direction Générale du CHU en octobre 2017
- **EDSaN (Entrepôt de Données de Santé Normand)**
- **Accord CNIL : octobre 2020**
- Publication princeps :

*Pressat-Laffouilhère, Thibaut; Balayé, Pierre; Dahamna, Badisse; Lelong, Romain; Billey, Kévin; Darmoni, Stéfan J. & Grosjean, Julien. [Evaluation of Doc'EDS: a French semantic search tool to query health documents from a clinical data warehouse](#). **BMC medical informatics and decision making**, February, Volume 22, Number 1, Pages 34, 2022.*

EDSaN au CHU de Rouen

- Volumétrie

Patients	2,1 millions
Séjours (hospitalisations, séances, consultations)	21,9 millions
Documents (CR, notes infirmières, lettres, ordonnances...)	23,3 millions
Administrations médicamenteuses (en hospitalisation)	2,2 millions
Résultats de biologie (hématologie, biochimie...)	193 millions
Dispositifs médicaux (prothèses...)	193 000
Diagnostics (CIM-10)	12,0 millions
Actes (imagerie, chirurgie...) (CCAM)	11,1 millions

- Antériorité : patients venus au moins une fois au CHU de Rouen les 20 dernières années

EDSaN au CHU de Rouen

- Exploitation+++ du texte (80% information clinique pertinente)
 - gestion négation, hypothèses/suspensions, antécédents familiaux...
- Dossiers de spécialité non intégrés nativement **mais exploitables** via le « lac de données »
 - ICCA
 - LOISABI
 - Softalmo
 - Urqual
 - ...

EDS Rouen – couverture fonctionnelle

Gestion des référentiels

Annuaire des structures

Référentiels des circuits de biologie, imagerie, chirurgie

Annuaire des Professionnels de santé

Prise en charge du dossier administratif

Identité Patient

Prise en charge du patient
Mouvements

Coordination et planification

Gestion des rendez-vous

Gestion de la planification des blocs opératoires

Production de soins et médico-techniques

Soins

Production des documents cliniques

Prescription multi modale

Soins infirmiers

Urgences

Dossiers de spécialités

Cancérologie DCC

Dossier social

Médico-technique

Radiothérapie

Chimiothérapie

CR Imagerie

CR Anatomopathologie

Imagerie

Moniteurs de réanimation

Biologie conventionnelle

Bactériologie / Virologie

Dispositifs médicaux

Génétique

Biobanque

Gestion médico-économique

Codage des actes

Codage des diagnostics

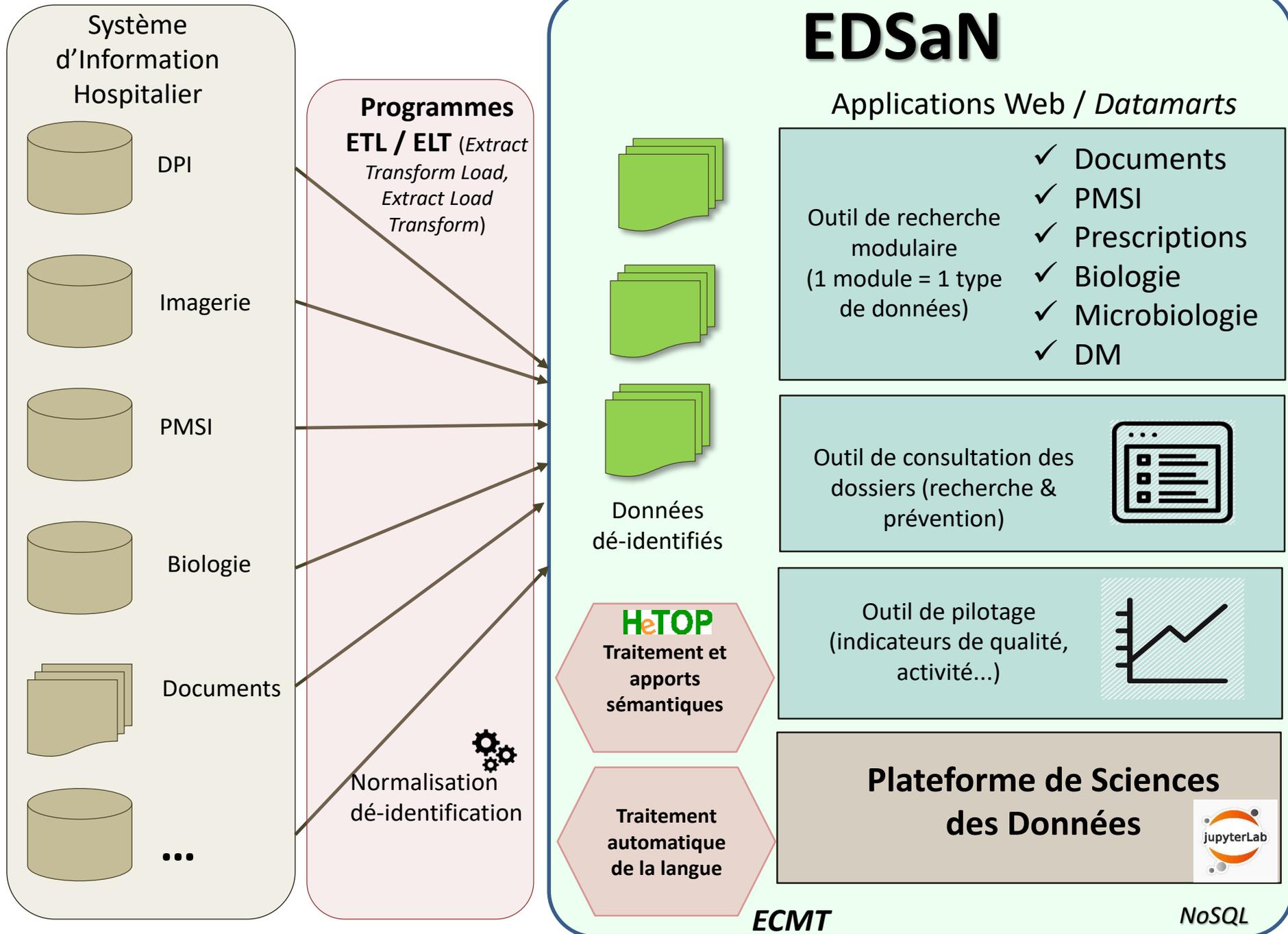
Groupage PMSI

 Domaines métiers présents dans EDS

 Domaines métiers hors périmètres EDS

Schéma adapté de l'APHP

Aspects méthodologiques et techniques



Architecture technique EDSaN

ENVIRONNEMENT DE PRODUCTION EDS

HEBERGEMENT SECURISE

SERVEUR APPLICATIF (HTTPS) EDS

- AUCUNE DONNEE NOMINATIVE
- IDS PATIENTS INTERNES EDS
- PARTITION CHIFFREE
- CR ANONYMISES



IMDG

SERVEUR DE PERSISTENCE DONNEES MEDICALES



POSTGRESQL 10

- AUCUNE DONNEE NOMINATIVE
- IDS PATIENTS INTERNES EDS
- PARTITION CHIFFREE
- CR ANONYMISES

SERVEUR DE PERSISTENCE DONNEES NOMINATIVES CHIFFREES



POSTGRESQL 10

- DONNEES NOMINATIVES SEULES
- CORRESPONDANCE CHIFFREE (SEL)
IDS PATIENTS EDS - NIP
- PARTITION CHIFFREE
- DONNEES SENSIBLES
(NOM, PRENOM,...) CHIFFREES
DANS LA BASE DE DONNEES

RESTRICTION AUX IP AUTORISEES DANS LE CHU

UTILISATEURS

REFERENTS AUTORISES



DEPOT SECURISE (TEMPORAIRE)
ARCHIVE CHIFFREE



- Séparation physique des données nominatives et des données cliniques (dont documents textuels dé-identifiés)
- Chiffrement des données (tables de correspondances identifiant « public » / identifiant « privé ») ; réversibilité via procédure particulière et par le personnel agréé

Procédures

Politique d'accès à EDSaN

<https://edsan.chu-rouen.fr/edsan/acces-aux-donnees/procedures/>

1. Étude de chaque demande par le Comité Scientifique et Éthique de l'établissement
2. Rendez-vous « physique » avec un membre du D2IM pour créer la (ou les) requête(s)
 - Raisons réglementaire et éthique : l'outil de requête donne accès à toutes les données
 - Raison scientifique : biais des données, biais des outils...
 - Accompagnement pour « délester » le chercheur
3. Accès aux données délimitées par la (ou les) requête(s) dans un outil dédié (accès nominatif)
4. Exports possibles par le Comité d'Export (personnels agréés du D2IM)

Résultats & Outils

Plateforme EDSaN



Consulter vos jeux de données

Cet environnement est consacré à la sélection des données ou des patients afin de créer des cohortes ou des tableaux de données à partir de l'EDSaN. Pour cela, une ou plusieurs requêtes doivent avoir été élaborées et exécutées pour pouvoir rendre accessible les données à explorer.

 [Accéder à votre espace personnel](#)

Créer ou ouvrir des requêtes

Uniquement accessible par les administrateurs EDSaN, cet environnement permet de créer des requêtes pour quelles soient notamment disponibles à la consultation par les chercheurs.

 [Créer une requête](#)

Outil modulaire de recherche

- 1 module = 1 type de données
- Disponible à ce jour :
 - Documents textuels (80% information clinique pertinente¹)
 - PMSI (mouvements, codages diagnostics et actes médicaux)
 - Prescriptions/administrations médicamenteuses dans l'établissement
 - Dispositifs médicaux
 - Biologie « conventionnelle »

The screenshot shows the EDSaN web application interface. At the top, there is a navigation bar with the EDSaN logo, a 'Projet' dropdown menu, a 'Sélection' dropdown menu, and an 'Outils' dropdown menu. On the right side of the navigation bar, there are logos for 'D2IM' (Département d'Informatique et d'Information Médicales) and 'CHU ROUEN NORMANDIE', along with a share icon. Below the navigation bar, there are three search modules, each with a circular icon and a title:

- Doc'EDS** (purple icon): Module permettant la recherche sémantique au sein de l'ensemble des documents cliniques du système d'information hospitalier. Il offre des fonctionnalités d'assistance à la recherche textuelle en exploitant des techniques de traitement automatique de la langue, et les concepts issus de 80 terminologies de santé. Ce module recense actuellement 18 437 736 documents.
- EDSaN PMSI** (red icon): Module dévolu à la recherche au sein de l'information relative à la T2A (Tarification À l'Activité), notamment les données réglementaires du DIM, soit : les diagnostics, actes et informations inhérentes aux séjours... Ce module recense actuellement 8 042 521 séjours.
- EDSaN MEDICAMENTS** (green icon): Module offrant des possibilités de recherche au sein des données médicamenteuses du SIH : notamment les prescriptions des patients hospitalisés. Il exploite un modèle du médicament innovant afin d'assister l'utilisateur dans sa recherche. Ce module recense actuellement 304 459 prescriptions.

¹ Raghavan P, Chen JL, Fosler-Lussier E, Lai AM. How essential are unstructured clinical narratives and information fusion to clinical trial recruitment? AMIA Jt Summits Transl Sci Proc. 2014 Apr 7;2014:218-23. PMID: 25717416; PMCID: PMC4333685.

Les fonctions techniques des EDS

- 4 outils complémentaires :
 - Effectuer des requêtes
 - Consulter et trier des dossiers patients
 - Tableaux de bord
 - Plateforme de traitement de données (JupyterLab) : gestion des données, extractions automatiques de données, transformations, statistiques, algorithmes d'IA...
- Utilisation d'assistants sémantiques (outil d'amont)
- Statistiques / exports / visualisations / *data mining* / extraction de valeurs (*feature extraction*) (outils d'aval)

Analyses textuels (module Doc'EDS)

Doc'EDS (module document)

- 80% de l'information clinique se trouve dans les documents textuels (comptes rendus, lettres, transmissions infirmières...)
- Il faut prendre en compte la structure du document et le discours (Traitement Automatique du Langage Naturel ou *NLP*) :
 - Négation
 - Condition/suspicion/hypothèse/futur
 - Antécédents familiaux
 - Segmentation du document

Partie « requête »

Partie « résultats »

sigma AND fracture

Texte

décédé OR décés

Texte (atcd familiaux)

Date doc. [2000-01-01,2001-01-01], <1999-01-01, >2018-05-15

Type doc. auto-compl

Unité(s) médicale(s) auto-compl

UF(s)

DdN patient [1950-01-01,1960-01-01], <1930-01-01, >2010-06-03

Âge (au moment du CR) >16, [16,65]

Sexe patient

ID Patient

ID Séjour

ID Document

Rechercher

12 patients

4 de 16

Document 768941148

Patient 214354096 (Masculin, 66 ans) né(e) en 1951

Du 2018-01-27 en CHIR.ORTHOPEDIQUE ET TRAUMATOLOGIQUE (ORTH) (UF 6112), Évènement 262658880

Malade sous anesthésie générale decubitus dorsal.
Voie d'abord antéro-médiale. Luxation de la patella. Patelloplastie.
A l'aide de l'ancillaire DEPUY SIGMA TC3 recoupe fémorale première puis distale proximale.
Recoupe fémorale 4 en 1.
Prothèse d'essai qui valide une bonne course avec une correction du varus.
On effectue une radiographie per opératoire avec les implants d'essai. Celle-ci retrouve une correction satisfaisante du varus qui maintenant est compris entre 5 et 10°.
On a la persistance d'une pente antérieure mais qui *ne sera pas réduite en raison d'un risque de conflit mécanique entre* la tige et la corticale.
Mise en place des implants finaux TC3 fémur taille 4 tibia taille 3. Insert taille 4. PE 12.5.
Quille d'extension fémorale et tibiale 13x30.
Protocole RAAC per opératoire. Fermeture de la capsule sur genou fléchi. Fermeture de la sous peau et de la peau par points séparés et drain de redon aspiratif

EVOLUTION

Suites simples
Pas de déficit sensitivomoteur
Douleurs contrôlées par les antalgiques usuels
Pas de signe de thrombose veineuse profonde
Bilan biologique *sans particularité*
Retour à domicile autorisé
Traitement prescrit dans le service

Traitement habituel
Antalgiques usuels
EVA à la sortie : 0/10

CONSIGNES DE SORTIE

- Réfection du pansement tous les 2 jours jusqu'à cicatrisation complète par IDE à domicile
- Ablation des sutures à J15
- Rééducation en flexion 0-40° immédiate selon protocole fourni au patient
- Appui autorisé protégé par 2 cannes pendant 6 semaines

RV DE CONSULTATION :
Prochain RV dans 3 mois avec le Pr [DOCTOR]
Radio (hon fait)

Visualisation des documents afin d'affiner sa recherche

sigma AND fracture

Texte

décédé OR décès

Texte (atcd familiaux)

Date doc. [2000-01-01,2001-01-01], <1999-01-01, >2018-05-15

Type doc. auto-compl

Unité(s) médicale(s) auto-compl

UF(s) auto-compl

DdN patient [1950-01-01,1960-01-01], <1930-01-01, >2010-06-03

Âge (au moment du CR) >16, [16,65]

Sexe patient

ID Patient

ID Séjour

ID Document

Rechercher

Opérateurs booléens disponibles

Assistant sémantique

Sélectionner les segments du texte

Les métadonnées disponibles pour affiner la requête (dates, âge...)

KAMIQUOL (ORJ) 5 mg, un par jour
FUROSEMIDE 20 mg, un par jour
SELOKEN 200 mg, 1 par jour
SINVASTATINE 40 mg, un par jour
INNOHEP 14000, une injection par jour
IMOVANE 7.5 mg, un au coucher
VESICARE 5 mg, un par jour
UROREC 8 mg, un par jour

MODE DE VIE

Marié (femme aidante)

HISTOIRE DE LA MALADIE :

Patient de ● ans hospitalisé le matin ●/04 à Becquerel pour une cure de radiothérapie dans le cadre de la prise en charge d'un carcinome épidermoïde de l'oesophage suivi au CHU en gastro-entérologie (médecin référent : [DOCTOR]).

Le patient décrit des expectorations mêlées de sang survenues en deuxième moitié de nuit (nuit du ● au ●/04). Un second épisode de crachats sanglants après effort de toux au cours de l'évaluation pré radiothérapie. Dans le contexte de tumeur de l'oesophage, le patient est transféré en unité de soins intensifs de gastro-entérologie au CHU pour une probable hématomèse.

EXAMEN CLINIQUE :

Taille : 176 cm : 76 kg. IMC : 24.5 kg/m². Température : 36.4°C Pouls : 77 par mn. TA : 109/70. EVA : 0. Hb : 12.6 g/dl. Bonne tolérance hémodynamique de l'hémorragie. pas de déglobulisation

Dénutrition. OMS 2-3.

Conjonctives colorées. Absence d'ictère.

Cardiovasculaire : BDC irrégulières. Pas de souffle. Pas d'OMI Pas de Turgescence des jugulaires.

Respiratoire : Râles crépitants à la base pulmonaire gauche. Toux importantes suivies d'expectorations teintées de sang.

Digestif : présence d'une sonde naso-gastrique d'alimentation entérale. Abdomen souple, indolore. BHA présents et normaux. Absence d'argument en faveur d'une ascite clinique. Toucher rectal : selles dures de couleur normale (pas de méléna ni de rectorragies). Prostate de taille augmentée avec disparition du sillon médian.

Neurologique : bonne orientation temporo-spatiale. Pas de déficit sensitivo-moteur.

BIOLOGIE :

Hémoglobine à 12,3 g/dl. Sérologie et antigénurie aspergillaire en cours.

RADIOLOGIE :

Radio de thorax :

TDM du ●/02/20● : Stabilité de l'épaississement circonférentiel oesophagien avec multiples adénomégalies médiastinales. Apparition de plusieurs nodules lobaires inférieures gauches :

Anonymisation
(patients et
médecins)

Détection des
suspensions/hypothèses/doutes/futur

Détection des
négations/infirmations

SCHEMA DE TRAITEMENT : Cacic (Campto + Cisplatine) + radiothérapie.

CIBLES MESURABLES :

Tumeur primitive : **oesophage**.

Adénopathies : oui.

ANAMNESE (depuis la précédente hospitalisation) :

Antécédents : chirurgie sur le ménisque droit. Chirurgie sur varices superficielles des membres inférieurs. Terrain atopique : eczéma.

Frère décédé d'un **cancer** du poumon il y a 30 ans.

Oncle décédé d'un **cancer** de l'**oesophage** il y a 16 ans.

EXAMEN CLINIQUE :

Indice de performance : grade OMS 0. Poids : 89 kg. Surf. corp.: 2 m2.

EXAMENS COMPLEMENTAIRES :

Biologie : GB : 6600 G/L PN : 3890 G/L Plaquettes : 272000 G/L Hb : 13,5 g/dl. TP : 99 %.

Urée : 0,24. Créat : 7,3 mg/l. TCA : 1,18.

TRAITEMENT RECU :

Chimiothérapie :

- Campto 60 mg/m2/j soit 120 mg/j (soit 100 % de la dose standard).
- Cisplatine 30 mg/m2/j soit 60 mg/j (soit 100 % de la dose standard).
- Traitement anti-émétique préventif par Zophren.
- Traitement anti-émétique curatif par Solumédrol.

Radiothérapie : à prévoir.

« Segmentation »
du document

Détection des
antécédents familiaux

Différents bilans

Statistiques & bilans

Stats démographiques

Date doc.

Type doc.

Unité(s) médicale(s)

UF(s)

DdN patient

Age (au moment du CR)

Sexe patient

Code(s) acte(s)

Code(s) diag(s)

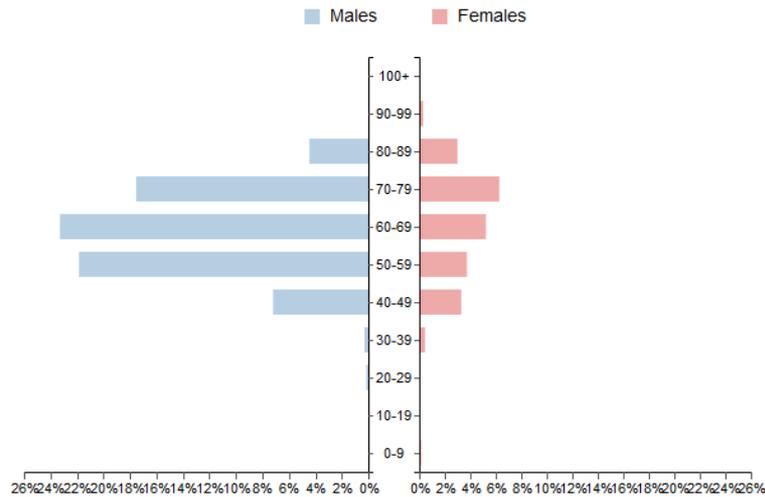


150 femmes (22,4%)

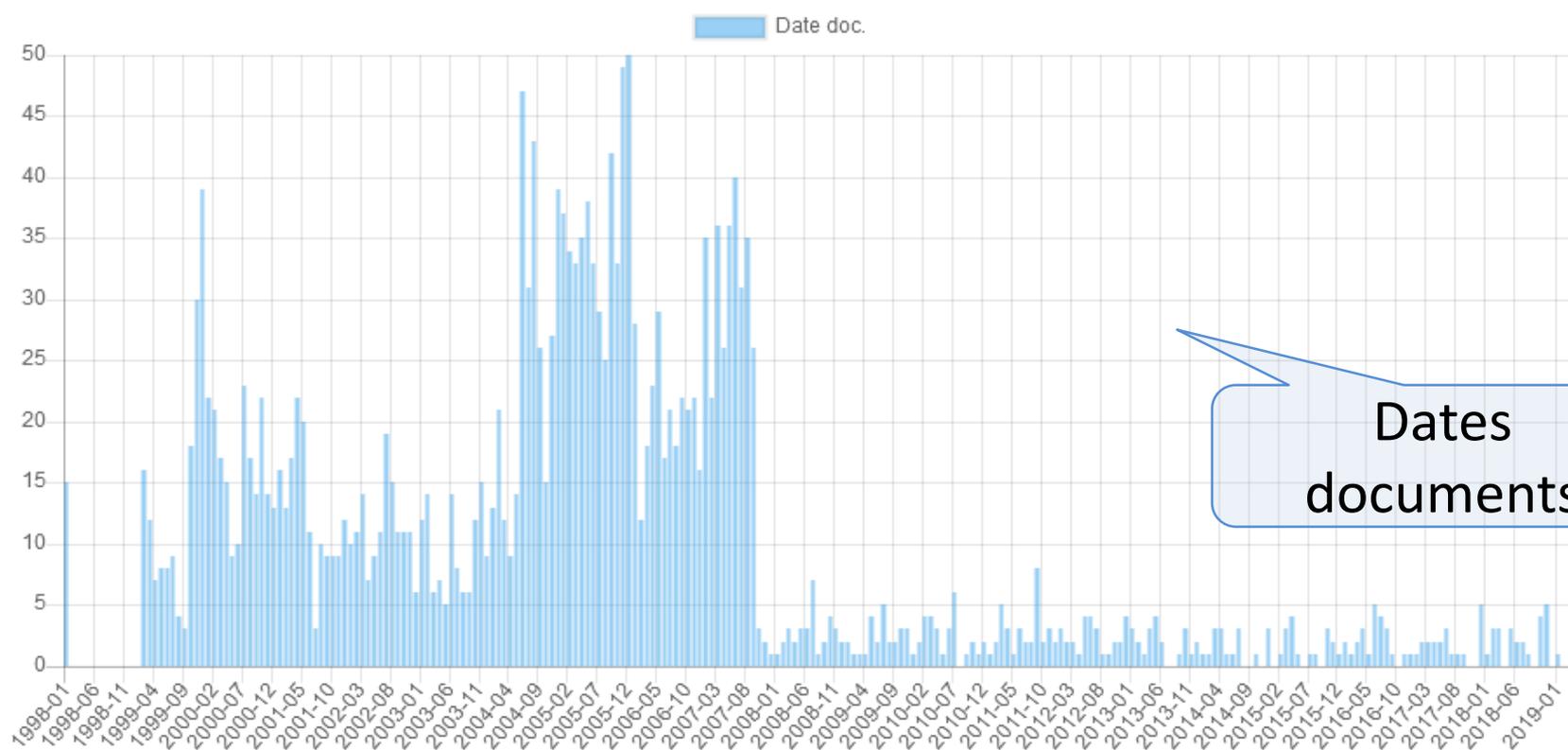
520 hommes (77,6%)

Statistiques des âges (au moment du CR)

Moyenne	63,3
Écart type	11,5
Minimum	7
Maximum	93
Q1	54
Q2	63
Q3	72



Population



Dates documents

SEJUM	#	%
DIGE HEPATO GASTRO ENTEROLOGIE NUTRITION	1685	61,9%
CGCD CHIRURGIE GENERALE ET DIGESTIVE	235	8,6%
RADI IMAGERIE CENTRALE	149	5,5%
URGE URGENCES	128	4,7%
PHIE PHARMACIE	93	3,4%
ORLO O.R.L ADULTES	61	2,2%
PNM1 CLINIQUE PNEUMOLOGIQUE HCN	59	2,2%
PHYS PHYSIOLOGIE DIGESTIVE	34	1,2%
CARD CARDIOLOGIE	25	0,9%
REAC REANIMATION CHIRURGICALE	23	0,8%

Répartition dans les unités médicales / services

lobstein

Type de discours

Contenu avéré

Restreindre à

Slope (-) 3

Synonymes

Expansion sémantique

	Terme	# doc
Equivalent <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> "lobstein"	1775
Libellé <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> "ostéogenèse imparfaite"~3	2394
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> "OI"	1543
Equivalent <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> "fragilité"	35169
	<input checked="" type="checkbox"/> "os de verre"~3	344
	<input checked="" type="checkbox"/> "maladie des os de verre"~3	225
	<input checked="" type="checkbox"/> "fragilité osseuse"~3	1227
	<input checked="" type="checkbox"/> "maladie de Porak et Durante"~3	4
	<input checked="" type="checkbox"/> "osteogenesis imperfecta"~3	3
	<input checked="" type="checkbox"/> "fragilité osseuse héréditaire"~3	2
	<input checked="" type="checkbox"/> "fragilité osseuse constitutionnelle"~3	84
	<input checked="" type="checkbox"/> "maladie de lobstein"~3	591
Plus précis <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> "syndrome de Bruck"~3	4
	<input checked="" type="checkbox"/> "ostéogenèse imparfaite type IV"~3	19
	<input checked="" type="checkbox"/> "ostéogenèse imparfaite légère"~3	7
	<input checked="" type="checkbox"/> "ostéogenèse imparfaite létale"~3	11
	<input checked="" type="checkbox"/> "ostéogenèse imparfaite avec sclérotiques bleues et dent	2
	<input type="text"/>	

Assistant sémantique

Assistant sémantique ✕

lobstein

Type de discours: Contenu avéré

Restreindre à:

Slope (~): 3

Synonymes | **Expansion sémantique**

	Terme	# doc
Prescription hors AMM <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> "FOSAMAX"	15996
	<input checked="" type="checkbox"/> "STEOVESS"	30
	<input type="text"/>	

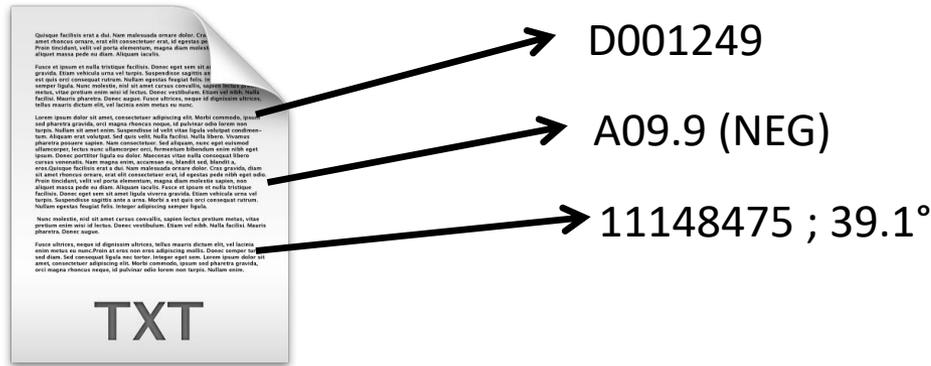
Ajouter un autre concept

Sous-requête	# doc
"lobstein" OR "ostéogenèse imparfaite"~3 OR "OI" OR "fr..."	56728

Finaliser

ECMT : annotation automatique des documents textuels

- Extracteur de Concepts Multi-Terminologique



- Utilisation en aval (documents d'un groupe de patients)
- Utile pour le *data mining*, observer des tendances ou structurer des informations précises (médicaments...)

Filtrer (maladies, procédures, ...) : ici médicaments

Afficher éléments

Filtrer : PHA_RAC

Catégorie	Concept	Identifiant	Terminologie	Nb documents
Substance pharmacologique; produit chimique inorganique;	cisplatine	PHA_RAC_1826	PHA	1150
chimie organique; Substance pharmacologique;	zophren	PHA_RAC_8895	PHA	1015
chimie organique; Substance pharmacologique;	campto	PHA_RAC_1470	PHA	517
chimie organique; Substance pharmacologique;	solumédrol	PHA_RAC_7444	PHA	401
médicament clinique;	solupred	PHA_RAC_7445	PHA	306

Affichage de l'élément 1 à 5 sur 332 éléments (filtré de 3,710 éléments au total)

Précédent

1

2

3

4

5

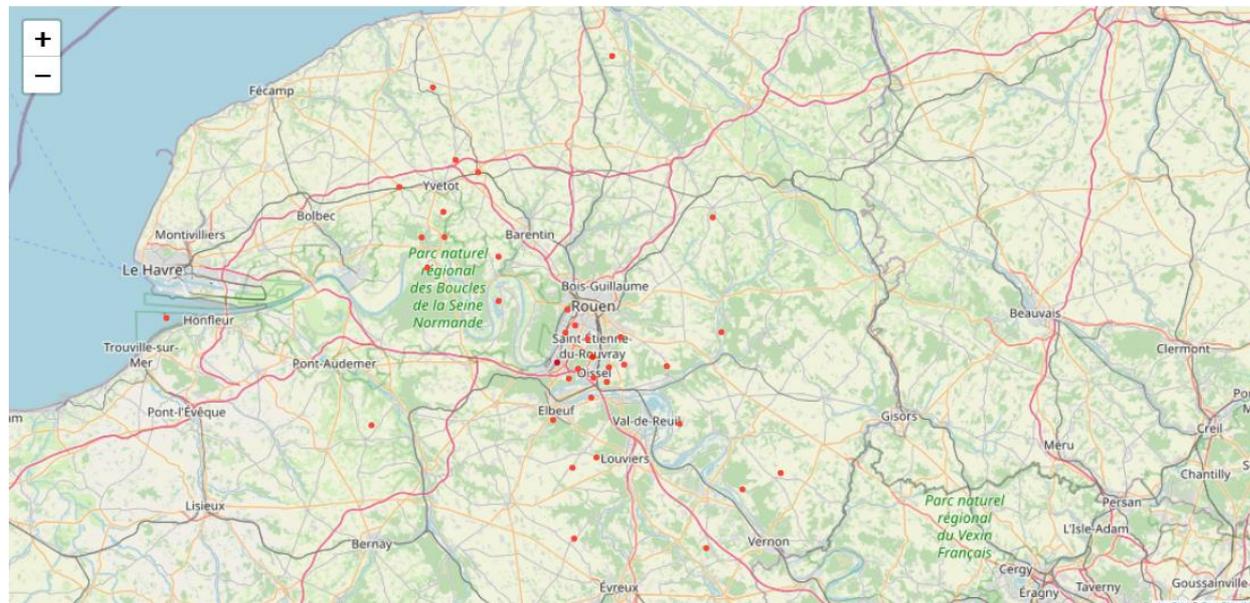
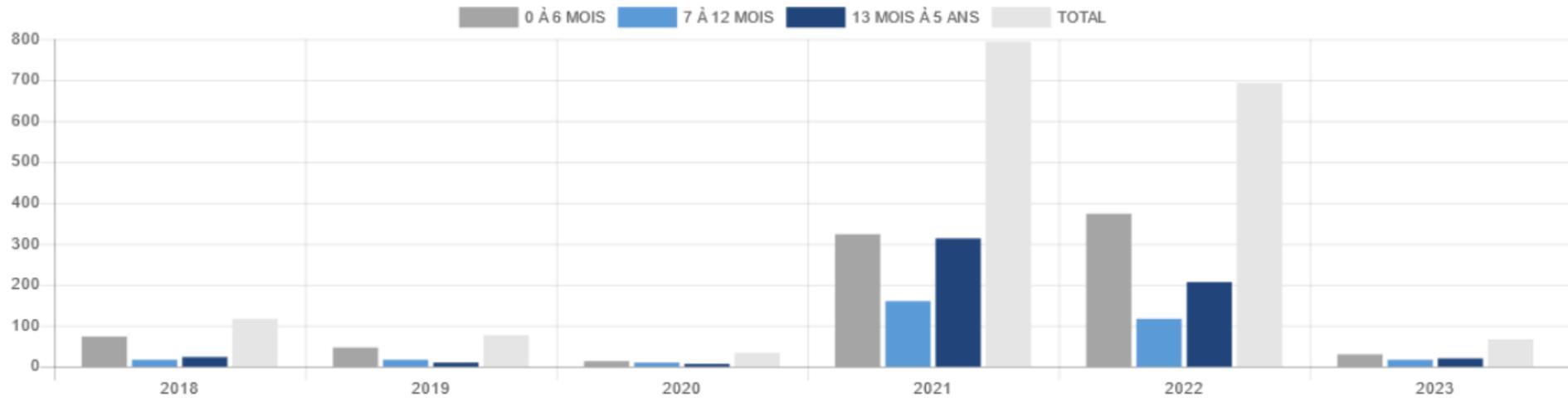
...

67

Suivant

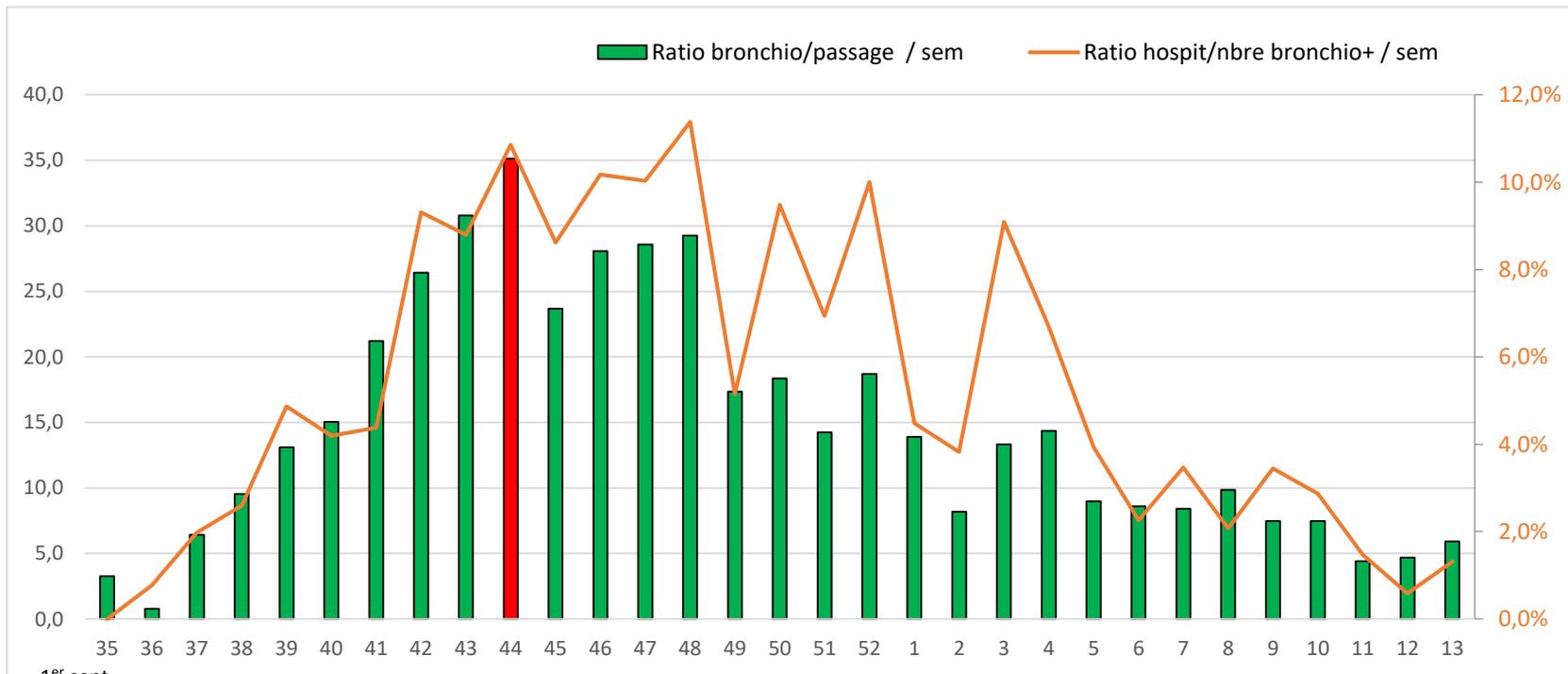
Tableaux de bord

Nb de patients VRS+ par tranche d'âge sur les 5 dernières années, CHU Rouen



Saison des bronchiolites en 2022/23

aux urgences du CHU de Rouen



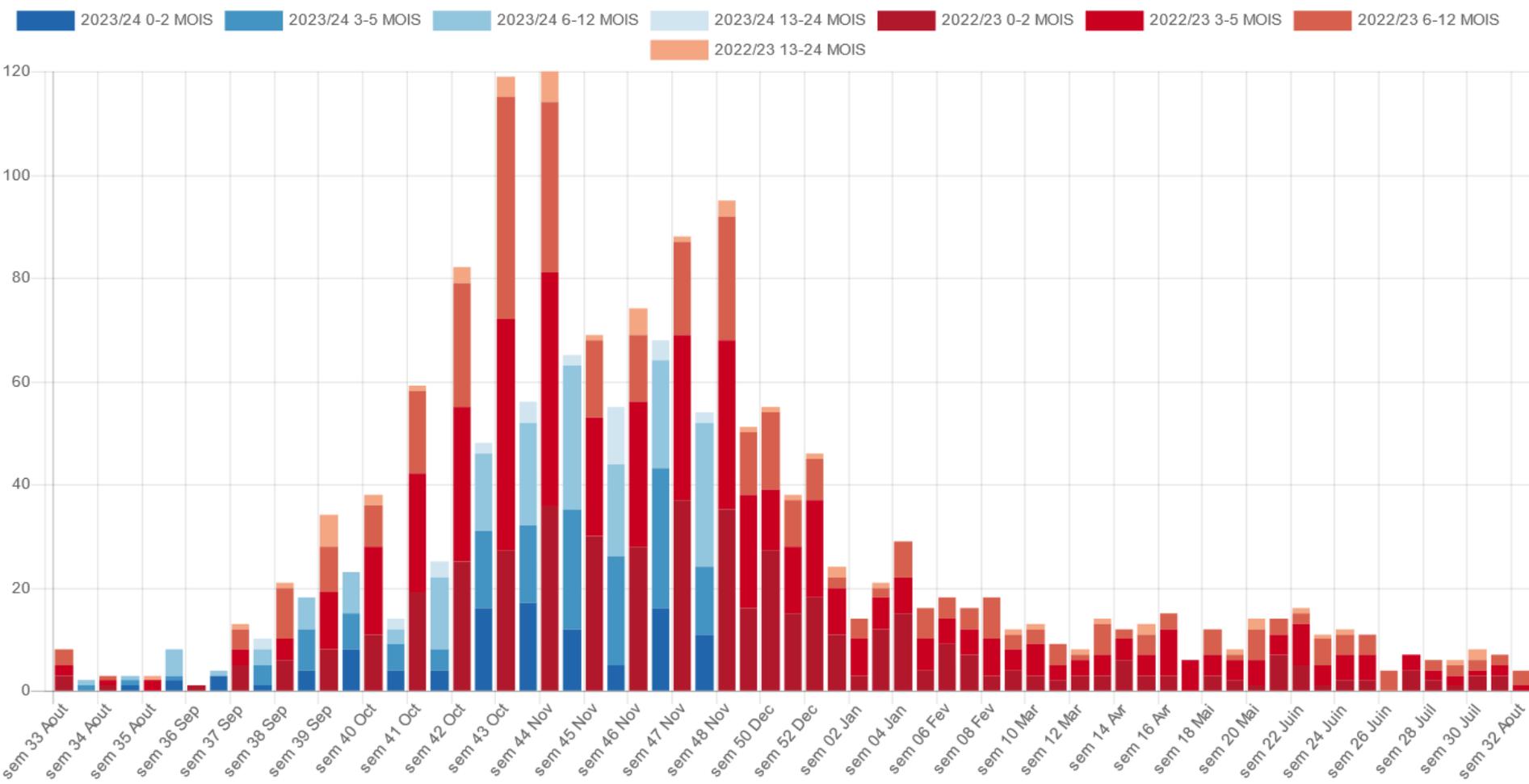
1^{er} sept.

**La semaine 44 en 2023 débutera
le lundi 30 octobre et se terminera le
dimanche 5 novembre**

Observations non publiées issues de l'entrepôt de données de santé EDSAN

Saisons des bronchiolites (tranches d'âges)

aux urgences du CHU de Rouen



Plateforme de traitements de données

EDSaNCoh : cohortes automatiques

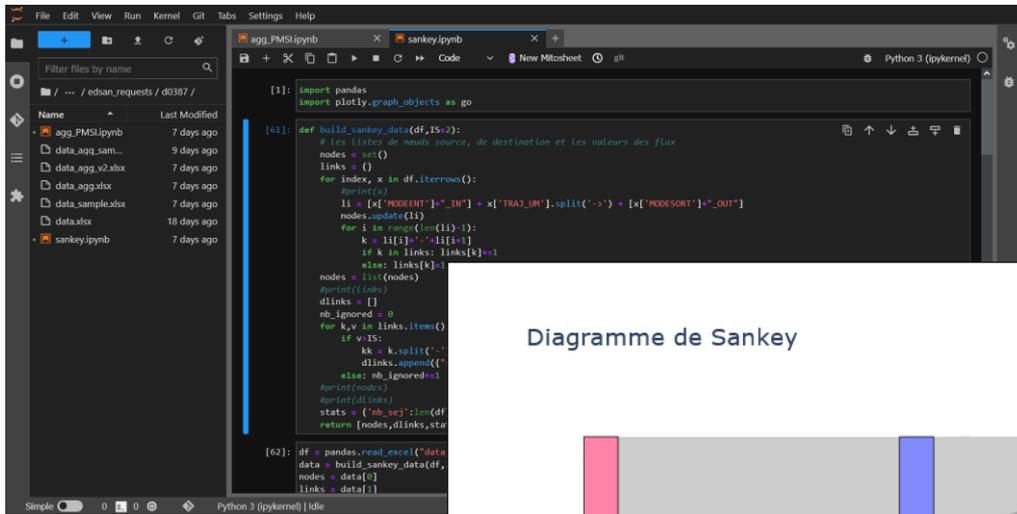
- En fonction des variables : utilisation du PMSI, des documents textuels ou de la biologie
- Performance globale : 94% d'exactitude
- Possibilité d'effectuer des améliorations itératives

Table : FP=false positive; FN=false negative; IDK="I don't know" (no match neither for the positive or negative query); "?"=unintelligible (match for both the positive and negative query), □ = extracted via EDSaN queries; □ = extracted via regular expressions. Between brackets, the improvements respect to the previous automatic evaluation.

	MATCHs	MISMATCHEs	FP	FN	"IDKs" OR "?"
Physical activity	57(+2)	43	12	2	29
Depression	83(+3)	17			17
Anxiety	85(+5)	15	1		14
Smoking/vaping	89(+3)	11	3		8
Diabetes	90	10			10
Death of a beloved one	92	8			8
History of anxiety	93(+4)	7			7
History of depression	94	6			6
Suicidal thoughts/self-harm attempts	95(+1)	5			5
Fat-free mass (kg)	97(+2)	3			3
Fat mass (kg)	97(+1)	3			3
ED (infancy or adolescence)	97	3			3
Trauma (infancy or adolescence)	97	3			3
Overweight/obesity (infancy or adolescence)	98	2			2
Bone tissue mass (kg)	99(+3)	1			1
BMI (kg/m ²)	99	1			1
Body weight (kg)	99	1			1
Body height (m)	99	1			1
History of physical/verbal abuse	99	1			1
BMD Femur Z score (z)	99(+3)	1			1
BMD Spine Z score (z)	100(+2)	0			0
Endometriosis	100	0			0
Lean mass (kg)	100(+4)	0			0
Autistic disorder	100	0			0
TOTAL	2258 (94.08%)	142 (5.92%)			
	2400				

Traitements en aval

- Gestion des données : mise en qualité, structuration...
- « Traitements complexes » : temporalités...
- Traitements statistiques
- Graphiques avancés spécifiques
- Algorithmes de phénotypage
- Algorithmes d'IA
- ...

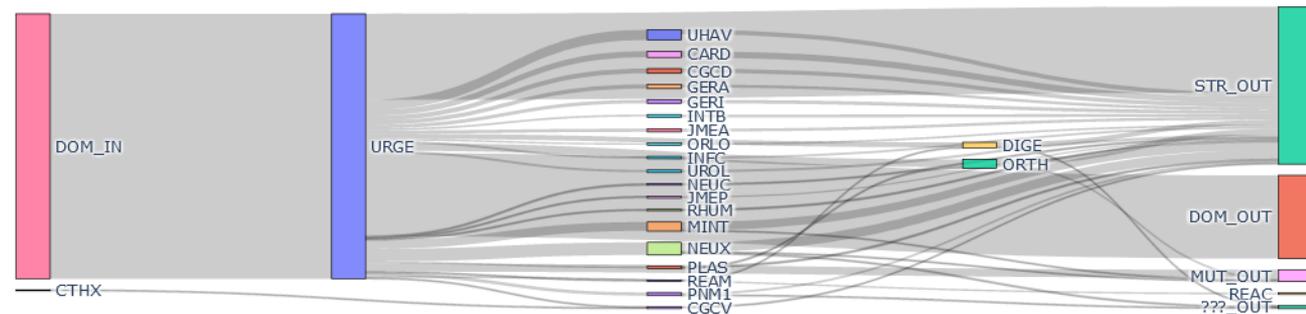


```
[1]: import pandas
import plotly_graph_objects as go

[61]: def build_sankey_data(df, IS=2):
    # Les listes de maus source, de destination et les valeurs des flux
    nodes = set()
    links = {}
    for index, x in df.iterrows():
        #print(x)
        li = [x['MODEENT']+"_IN"+ x['TRAJ_UN']+'.split(">") + [x['MODESORT']+"_OUT"]
        nodes.update(li)
        for i in range(len(li)-1):
            k = li[i]+">"+li[i+1]
            if k in links: links[k]=1
            else: links[k]=1
        nodes = list(nodes)
        #print(li)
        dlinks = []
        nb_ignored = 0
        for k,v in links.items():
            if v>IS:
                kk = k.split(">")
                dlinks.append((">".join(kk), v))
            else: nb_ignored+=1
        #print(nodes)
        #print(dlinks)
        stats = {'nb_seg': len(df)
        return (nodes, dlinks, stats)

[62]: df = pandas.read_excel("data
data = build_sankey_data(df)
nodes = data[0]
links = data[1]
```

Diagramme de Sankey

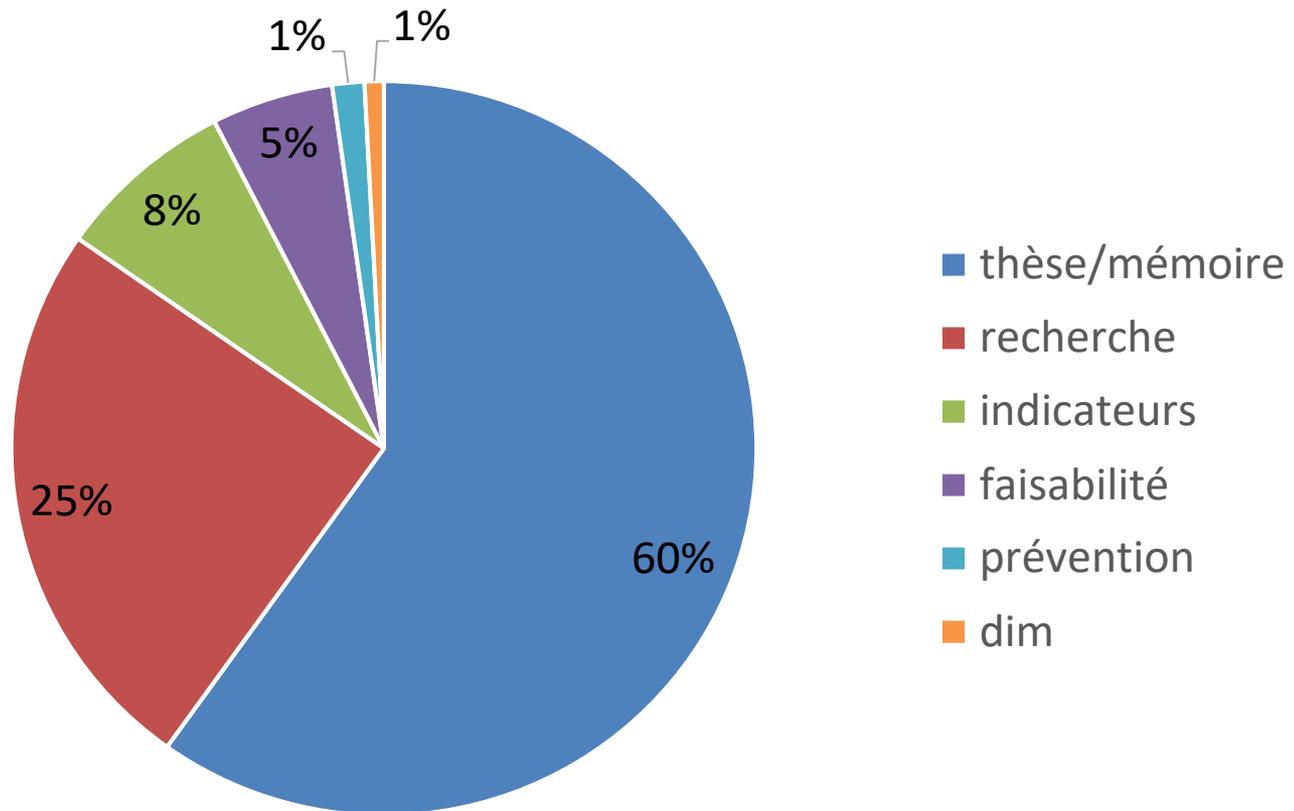


Expérience EDSaN : bilan novembre 2023

État des lieux

- Opérationnel depuis début 2018 pour des cas de tests et prototypages
- Accord CNIL obtenu en octobre 2020
- Au 29 septembre 2023 : **402 demandes traitées** (recensées dans un registre, conformément à la RGPD)

Finalités des demandes



Exemples cas d'usage COVID-19 :

- fourniture DG/CME des documents produits par semaine depuis pandémie COVID-19 (reflet des activités)
- surveillance de pathologies particulières (AVC, infarctus... pendant les confinements)
- élaboration d'une liste de patients « à très haut risque COVID » pour priorisation de la vaccination

Et maintenant...

- Lien EDS hôpital – hôpital (interopérabilité...)
- Lien EDS hôpital et EDS ville (médecine générale)
- Comparaison avec des **bases nationales** (SNDS...)

UN PROJET MULTI-PARTENARIAL

