



Suivre et prévenir en vie réelle

Exploitation des réseaux sociaux pour la pharmacovigilance

Marie-Christine Jaulent, DR Inserm

LIMICS – UMR_S1142

Inserm, Sorbonne université & université Paris 13



Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances pour la e-Santé



La pharmacovigilance

- **Surveillance** des médicaments et **prévention** du **risque d'effet indésirable** résultant de leur utilisation, que ce risque soit potentiel ou avéré
 - **Coût économique et clinique importants** des effets indésirables médicamenteux
- Surveillance pour les médicaments avec AMM, ATU ou RTU
 - **Utilisation normale, mésusage, abus, erreur**



Collecter des cas et détecter des signaux

- **Collecte active des données** via les centres de pharmacovigilance ou les laboratoires pharmaceutiques
- **Coder les cas** pour alimenter des bases de données
 - **Activité chronophage**
 - ANSM (France), FAERS (USA), EudraVigilance (EMA), VIGIBASE (OMS, depuis 1968)
 - EN 2019, 136 pays membres et 26 pays associés pour alimenter **VIGIBASE**



Catégorie d'information	Variables collectées
Report ID & Date	Report_id / Receive_date
Patient Data	Sex / Age / Weight
Seriousness	Hospitalization / Death / Death_date / Life_threat / Seriousness_other
Reaction	<p>Couple (Reaction_meddra_pt; Medicinal_product)</p> <p>➤ Pas de codage de la relation de cause à effet</p>
Basic Drug Information	Dosage_text / Drug_start_date / Drug_end_date
Administration Information	Route / Drug_separate_dosage_num / Drug_dosage_form
Report Information (Origin / Duplicates, etc.)	Primary_source_country / Occur_country / Reporter_country / Duplicate / Duplicate_num / Duplicate_source / Drug_recurr_administration / Drug_treatment_duration / Drug_cumulative_dosage_num
Additional Drug Information	Substance_name / Generic_name





Collecter des cas et détecter des signaux

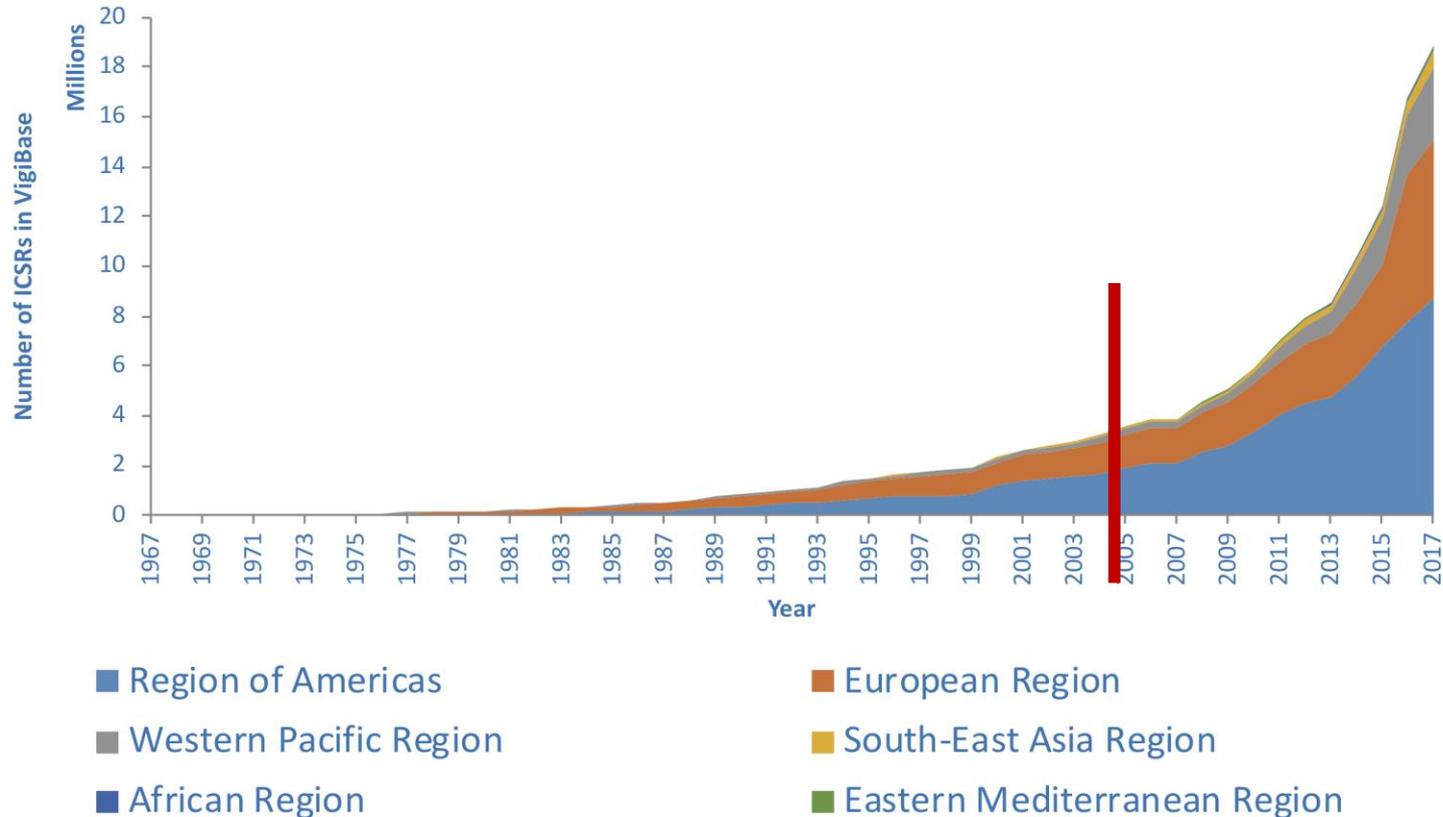
- **Collecte active des données** via les centres de pharmacovigilance ou les laboratoires pharmaceutiques
- **Coder les cas** pour alimenter des bases de données
 - **Activité chronophage**
 - ANSM (France), FAERS (USA), EudraVigilance (EMA), VIGIBASE (OMS, depuis 1968)
 - EN 2019, 136 pays membres et 26 pays associés pour alimenter **VIGIBASE**



VigiBase – Avril 2018



Total Number of Individual Case Safety Reports (ICSRs) in VigiBase





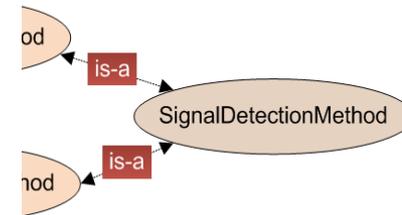
La détection du signal

- **Signal en pharmacovigilance (Définition OMS)**
 - une information sur **une relation possible entre un médicament et un effet indésirable**, la relation étant inconnue ou mal documentée à ce jour.
 - Algorithmes de fouille de données

• A

➤ **Id** Example Classification of Signal Detection Methods

pl
m





Une course contre la montre



La sous-déclaration



Exploitation des réseaux sociaux

- **Collecte passive des données**
 - Les + : volume et disponibilité en temps réel
 - Les - : qualité de la donnée, langage
- **Plusieurs initiatives nationales depuis 2014**
 - Plusieurs acteurs académiques (U1142; U1219 ERIAS; ..) et Industriels (Kappa Santé; Digimind; ...)
 - Plusieurs projets
 - ADR-PRISM (FUI)
 - **Vigi4Med, PHARES (ANSM)**
- Le projet européen **WEB-RADR**





Key points:

van Stekelenborg, J., Ellenius, J., Maskell, S. et al. **Recommendations for the Use of Social Media in Pharmacovigilance: Lessons from IMI WEB-RADR.** *Drug Saf* (2019).

General social media, as exemplified by sample data from Facebook and Twitter, are not recommended for broad statistical signal detection.

Social media channels may provide a useful adjunct to pharmacovigilance activities in specific niche areas such as exposure during pregnancy and abuse/misuse of medicines.

Future enhancement of adverse event recognition algorithms may broaden the scope and utility of social media over time.





Le consortium du projet

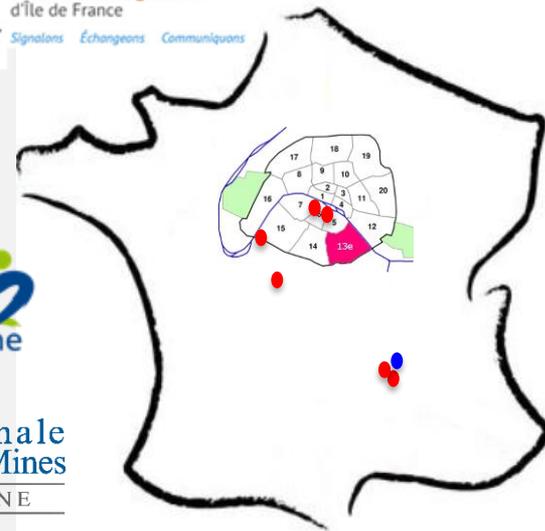


• Deux centres régionaux de pharmacovigilance

- CRPV Paris-HEGP
- CRPV de Saint-Etienne

• Cinq équipes de recherche

- SSPIM, Hôpital Nord (CHU de Saint-Etienne), **coordinateur**
- INSERM U1142 (Paris)
- INSERM U1138-eq22 (Paris)
- CNRS-LIMSI (Orsay)
- Ecole des mines de Saint-Etienne-ISCOD





Etude sur l'intérêt potentiel des messages

Critère présent	Cas des forums N = 1 284	Cas de la BNPV N = 2 521
Age	4,8%	93,3%
Sexe	47,7%	99,5%
Indication	60,7%	37,1%
Durée du traitement	22,6%	54,7%

Comparison

Critère présent	Cas des forums N = 1 284	Cas de la BNPV N = 2 521
Délai d'apparition (a)	19,9%	83,5%
Evolution	6,4%	82,3%

Karapetiantz P, Bellet F, Audeh B, Lardon J, Leprovost D, Aboukhamis R, Morlane-Hondere F, Grouin C, Burgun A, Katsahian S, Jaulent MC, Beyens MN, Lillo-Le Louet A, Bousquet C. Descriptions of Adverse Drug Reactions Are Less Informative in Forums Than in the French Pharmacovigilance Database but Provide More Unexpected Reactions. *FRONTIERS IN PHARMACOLOGY*, MAY 1 2018



Le consortium du projet

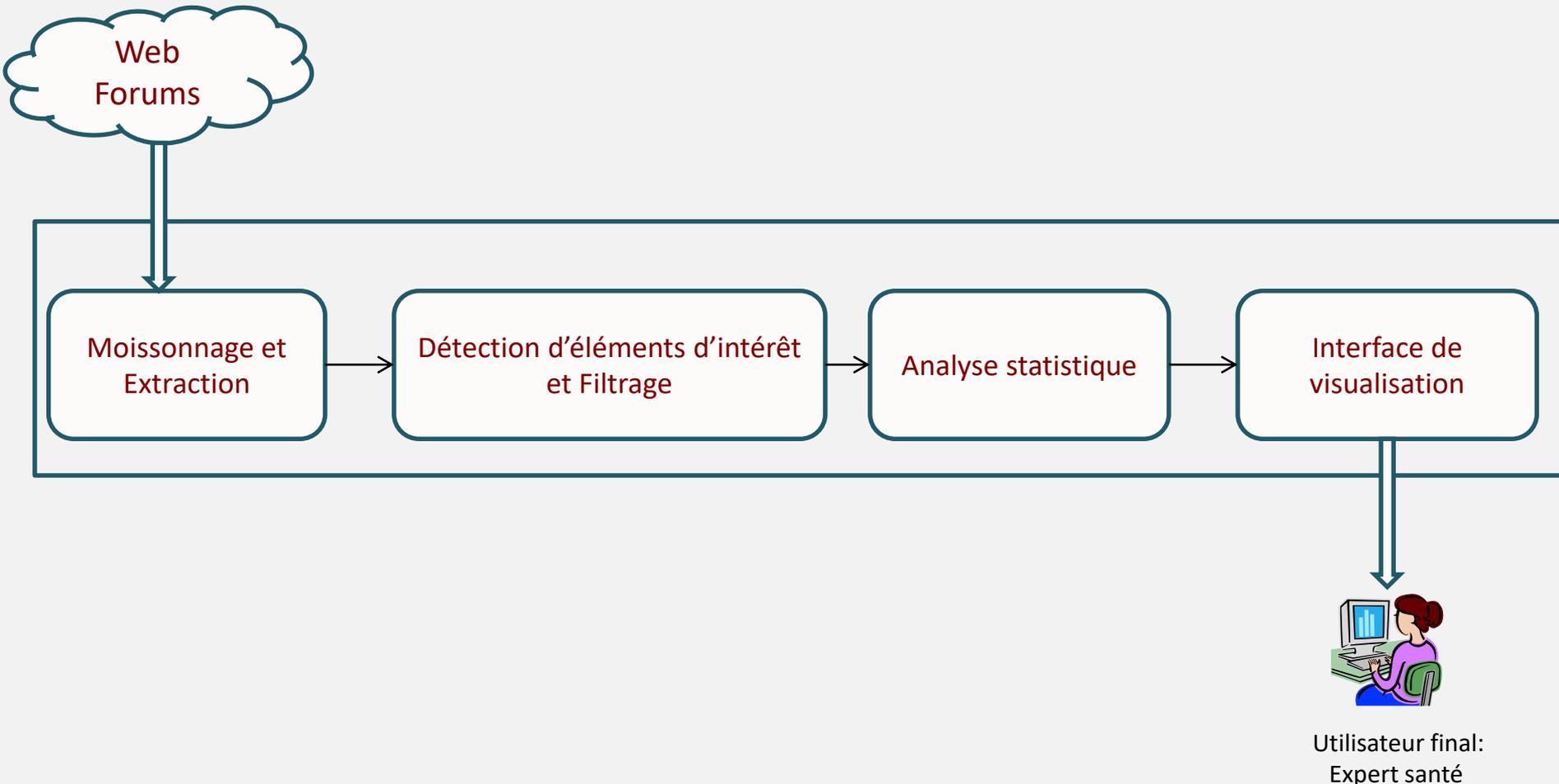


- Un centre régional de pharmacovigilance
 - CRPV Paris-HEGP
- Deux équipes de recherche
 - INSERM U1142 (Paris)
 - CNRS-LIMSI (Orsay)



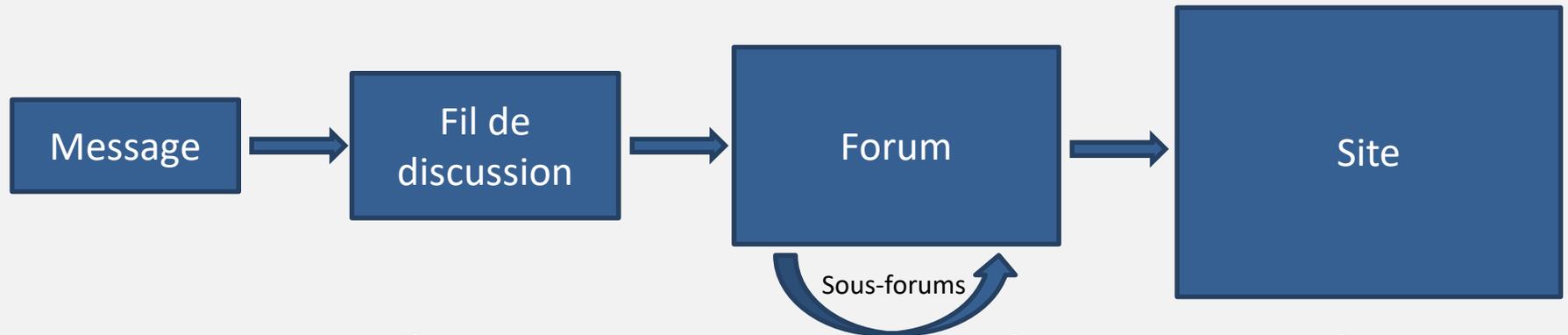


Réalisation d'une suite logicielle





Moissonage : Vigi4Med Scraper



- Moissonne les sites qui respectent la structure des forums de discussions
- Se base sur la structure hiérarchique des pages HTML (DOM: Document Object Model)
- Utilise l'*XPATH* (langage de requête pour localiser une portion d'un document XML)



Exemple: Doctissimo.fr/Santé/Douleur

Fils de discussions:

		s'inquiéter?	franfran /	1	22	iesallies
		mal au genoux		4	67	23/02/2018 à 23:30
		Ses douleurs chroniques disparaissent après un plongeon dans l'eau froide		0	8	Répondre
		AIDEZ MOI douleurs après extraction des 4 dent de sagesse		5	25	23/02/2018 à 15:31

Messages:

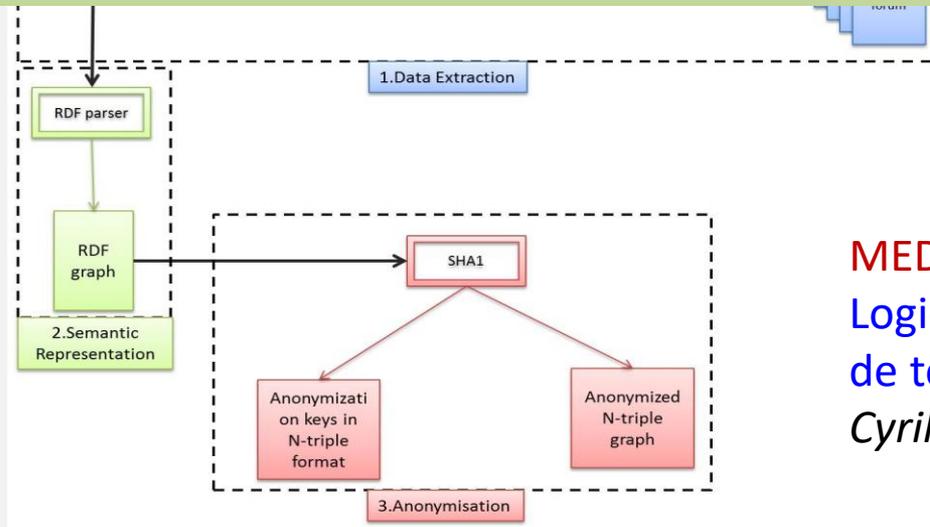
Auteur	Sujet : mal au genoux
Invité	<p>Posté le 25/04/2016 à 12:41:58</p> <p>Bonjour , j'ai ans et j'ai mal a l'extérieur du genoux droit depuis environ 5 mois je suis aller voir le médecin et il ma dit que c'était une tendinite il ma conseiller d'arrêter l'athlétisme c'est ce que j'ai fait et je n'es toujours pas reprit . J'ai aussi demander avis à mon kiné (je vais au kiné car j'ai un scoliose) et il ma d'abord dit que ce n'était pas possible que c'était une tendinite car apparemment la ou j'ai mal il n'y a pas de</p>
Profil : Habitué	<p>Posté le 23/02/2018 à 23:30:14</p> <p>alors t'a fait quoi ? après tout ce temps</p>



Vigi4Med Scraper: Structure



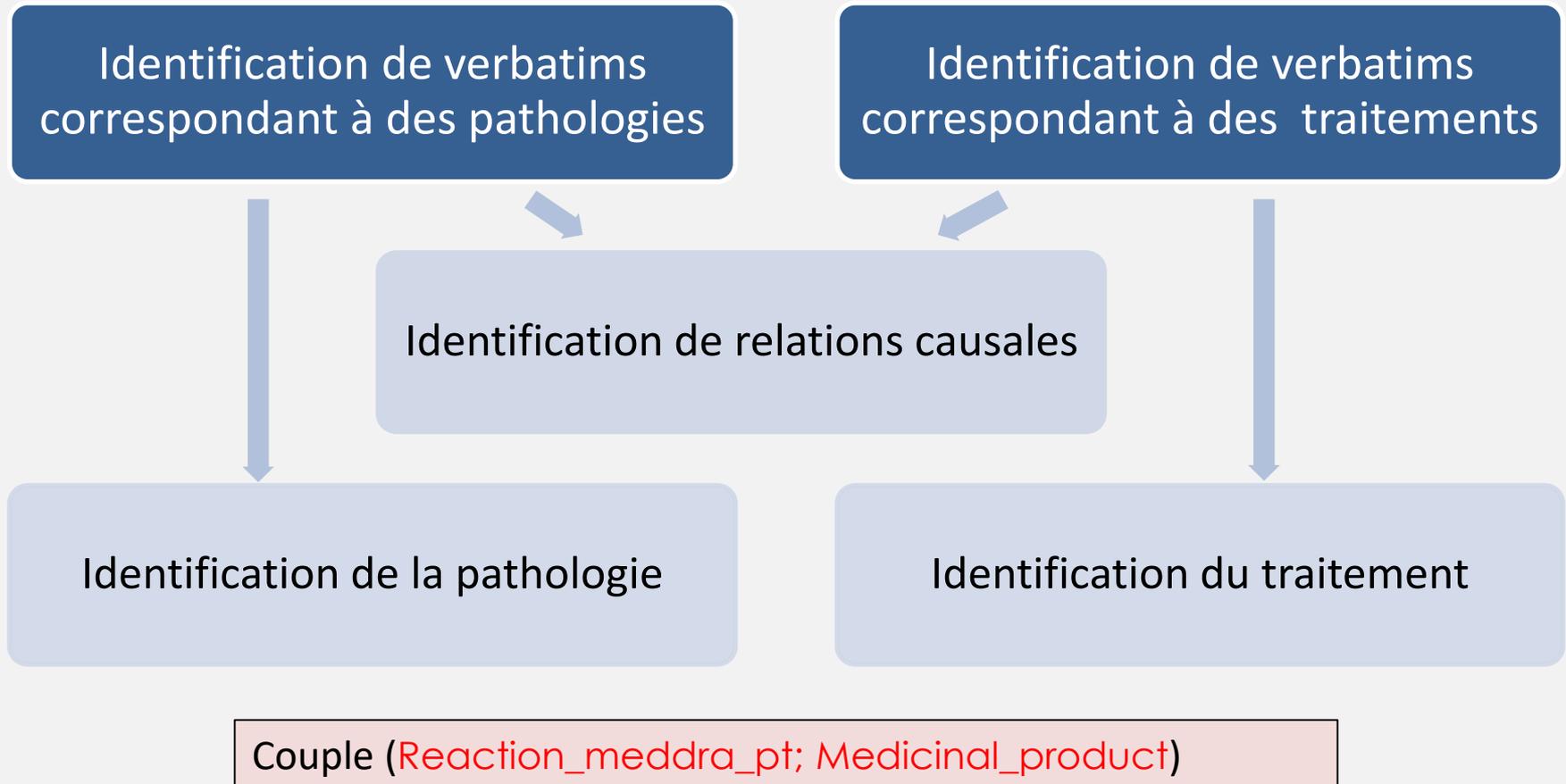
Audeh B, Beigbeder M, Zimmermann A, Jaillon P, Bousquet C. Vigi4Med Scraper: A Framework for Web Forum Structured Data Extraction and Semantic Representation. *PLoS One*. 2017;12(1)



MEDINA
Logiciel d'anonymisation
de textes
Cyril Grouin, LIMSI



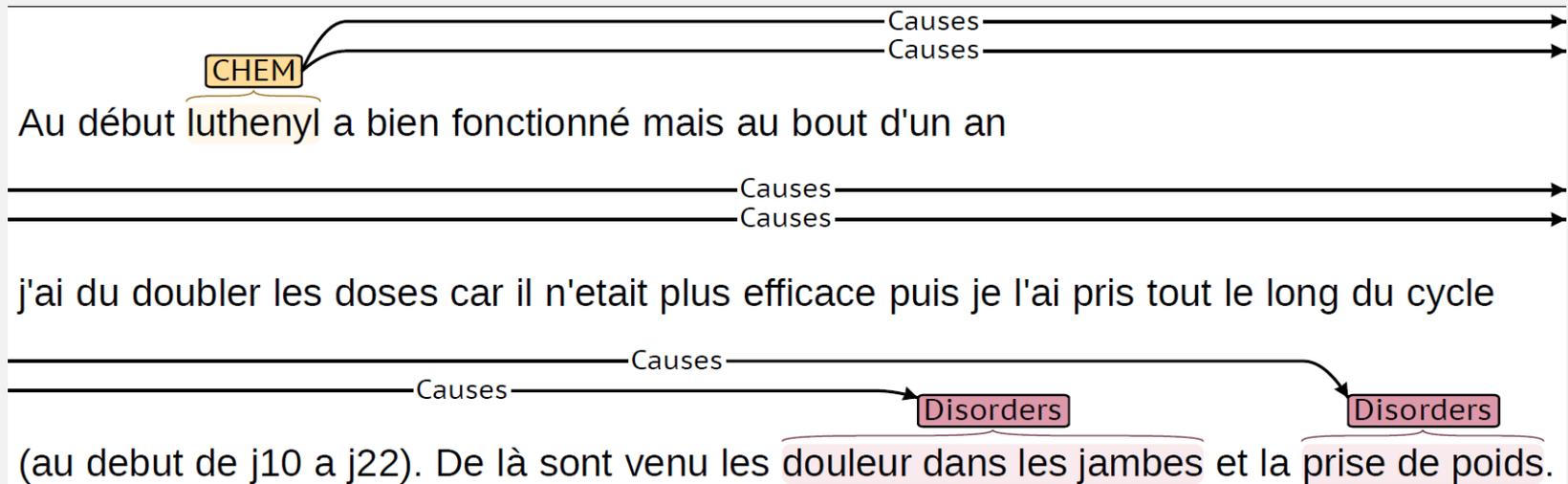
Annotation des messages





Exemple de message annoté

- Traitement automatique du langage (TAL)
 - Détection de médicament (Ex : Lutenyl), effet indésirable (Ex : prise de poids) et relation causale
 - Prise en compte des fautes d'orthographe (Ex : Luthenyl/Lutenyl)





Extraction d'entités annotées

- Fichier en format tabulaire avec les champs suivants :
 - Nom du forum: *Thyroïde*
 - Identifiant du message: *433*
 - Date: *03/21/18 10:52 a. m.*
 - Entité pathologique extraite: *diarrhées*
 - Entité pathologique normalisée (MedDRA PT) (sinon "---- »): *Diarrhée*
 - CUI UMLS de l'entité normalisée (sinon "----") : *C0011991*
 - Entité de type médicament/traitement extraite: *aspirine*
 - Nom normalisé du médicament (sinon "----"): *ASPIRINE*
 - Code MEDDRA de l'entité pathologique extraite: *10023084*
 - Code ATC du médicament extrait: *N02BA01*

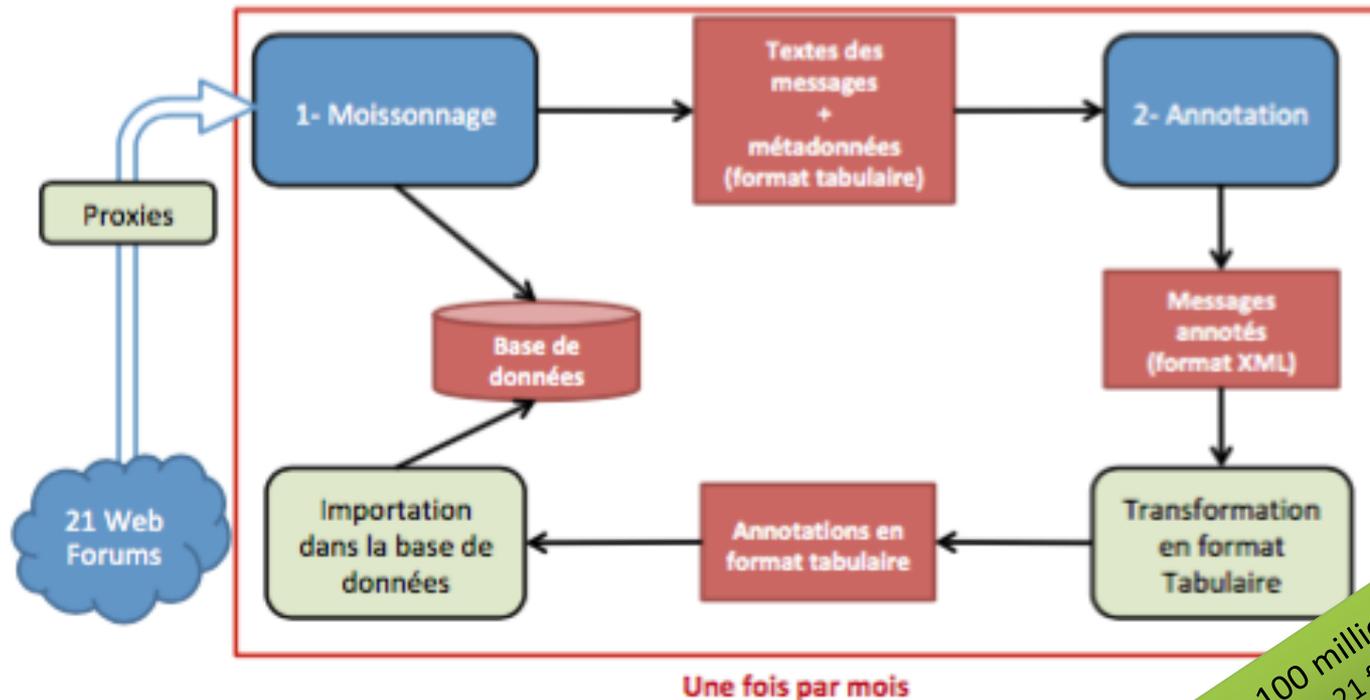


Extraction du contenu de 21 forums de discussion

Site	Taille RDF avant anonymisation	Taille RDF après anonymisation	Nombre de pages	Nombre de fils de discussion	Nombre de messages
http://www.allodocteurs.fr/	84 M	56 M	7 083	6 012	22 476
http://www.atoute.org/	16 G	12 G	294 611	157 267	4773 039
http://www.carenity.com/	4,8 M	4,6 M	149	135	320
http://www.comprendrechoisir.com/	40 M	33 M	3 121	3 013	6 678
http://www.docteurclik.com/	21 M	15 M	5 051	6 828	9 052
http://www.doctissimo.fr (médicament)	7,7 G	5,2 G	2 433 192	212 622	20 953 911
http://www.doctissimo.fr_(santé)	53 G	33 G	780 912	538 818	16 345 311
http://www.doctissimo.fr_(grossesse)	138 G	96 G	1 113 063	1 283 675	34 854 099
http://www.e-sante.fr/	91 M	71 M	1 2321	13 082	35 575
http://www.famili.fr/	10 G	8 G	319 060	63 178	2 558 717
http://sante-az.aufeminin.com/	3,9 G	2,2 G	152 839	140 735	849 655
http://forums.futura-sciences.com/	237 M	193 M	12 001	9 759	76 459
http://sante.journaldesfemmes.com/	58 M	45 M	4 543	4 217	25 844
http://sante-medecine.commentcamarche.net/	2,3 G	1,7 G	219 440	213 156	1 112 084
http://www.vulgaris-medical.com/	343 M	269 M	28 659	27 811	101 912
http://afm-telethon.fr/	8,2 M	6,6 M	1 175	1 012	3 502
http://www.entrepaticients.net/fr	11 M	7,7 M	968	861	3 485
http://www.notrefamille.com/	976 M	586 M	34 909	29 751	273 021
http://www.baclofene.com/index.php	547 M	309 M	4 378	5 377	178 192
http://www.baclofene.fr/portal.php	106 M	79 M	2 819	1 819	36 232
http://www.hepatites.net/	452 M	404 M	16 079	11 783	177 891
http://www.seronet.info/	14,5 M	13,7 M	278	147	175 049



Pipeline : partie automatisée



Plus de 100 millions de messages
extraits de 21 forums (250 Go de
données générées)



Pour quoi faire ?

- Etude sur les médicaments les plus fréquemment mentionnés dans les réseaux sociaux et leur évolution dans le temps
- Analyse du forum **Vivre sans Thyroïde**
- Détection du signal dans les réseaux sociaux : exemple du baclofène





Médicaments les plus mentionnés Vs. les plus prescrits (2015)

Rank	OPEN MEDIC	Web Forum
1	PARACETAMOL	BACLOFENE
2	IBUPROFENE	CLOMIFENE
3	AMOXICILLINE	PARACETAMOL
4	COLECALCIFEROL	PROGESTERONE
5	DICLOFENAC	PHLOROGLUCINOL
6	PREDNISOLONE	DYDROGESTERONE
7	TIXOCORTOL	MAGNESIUM
8	PHLOROGLUCINOL	ESTRADIOL
9	BETAMETHASONE	LEVONORGESTREL
10	OMEPRAZOLE	MORPHINE

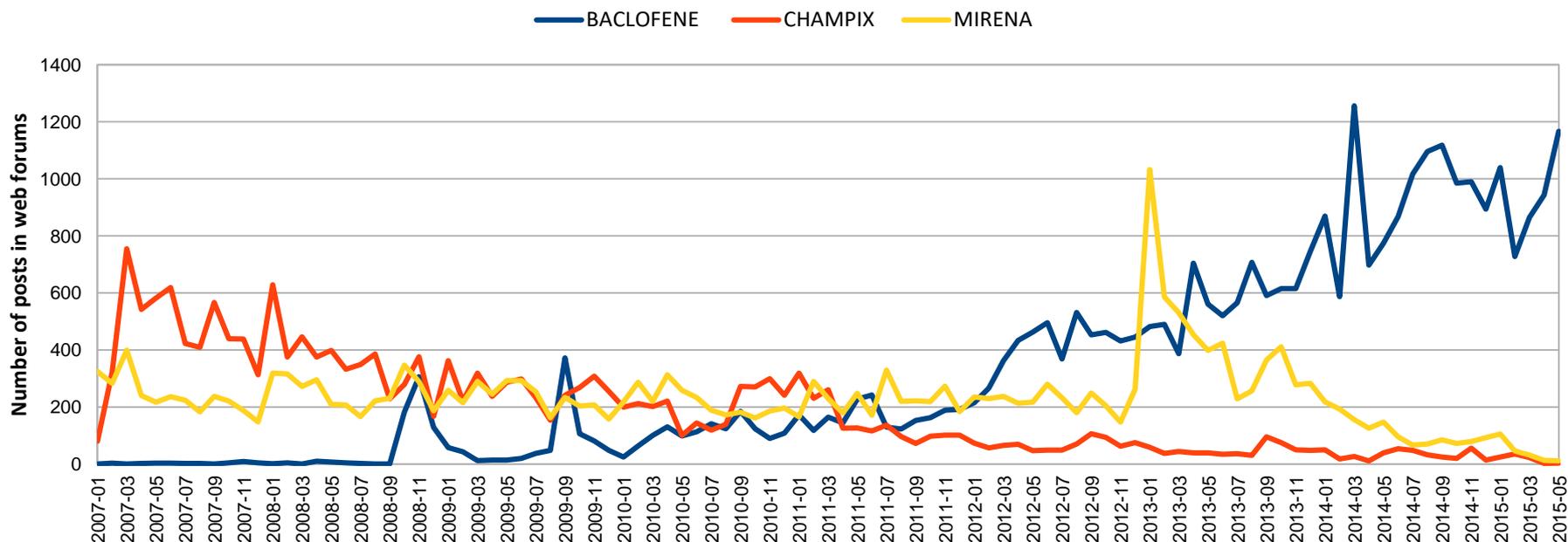
**Corrélation faible
entre les deux listes**

Kendall-tau
corrélation = 0,31
entre les deux listes
(510 principes actifs
en commun)



Evolution dans le temps

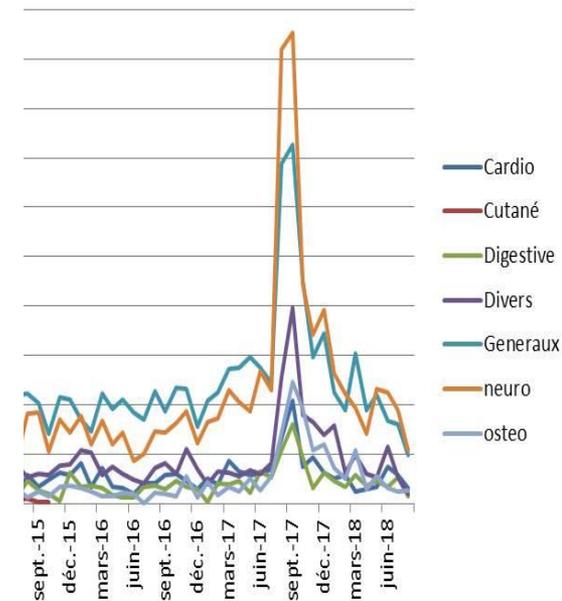
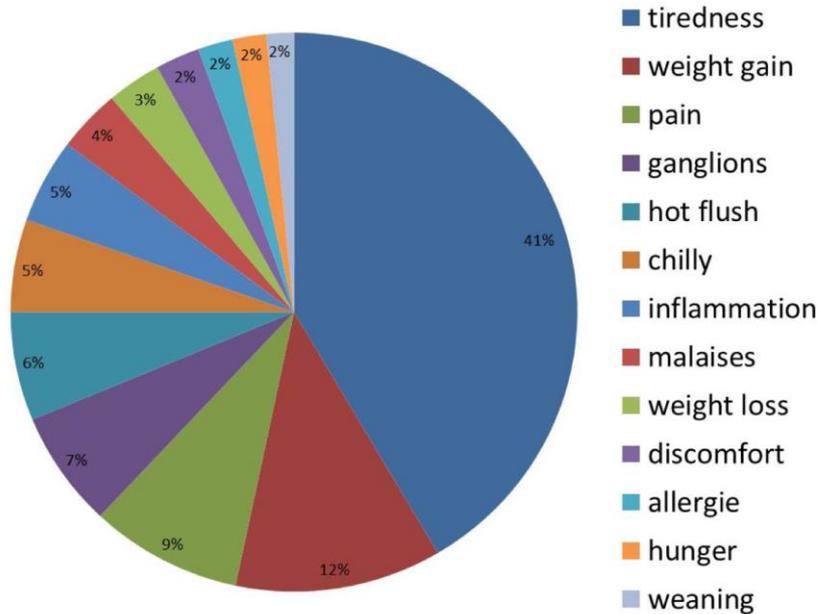
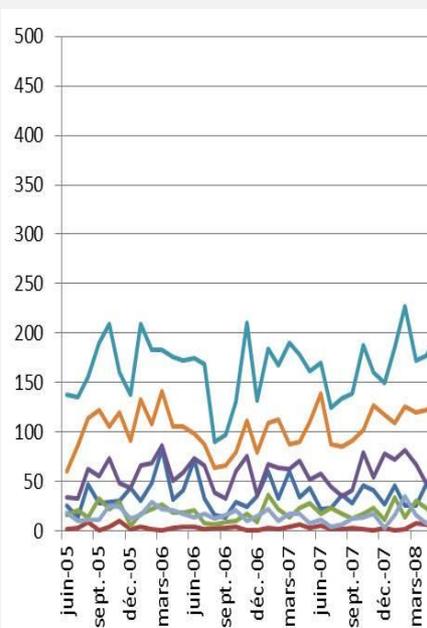
- Champix, Miréna et Baclofène





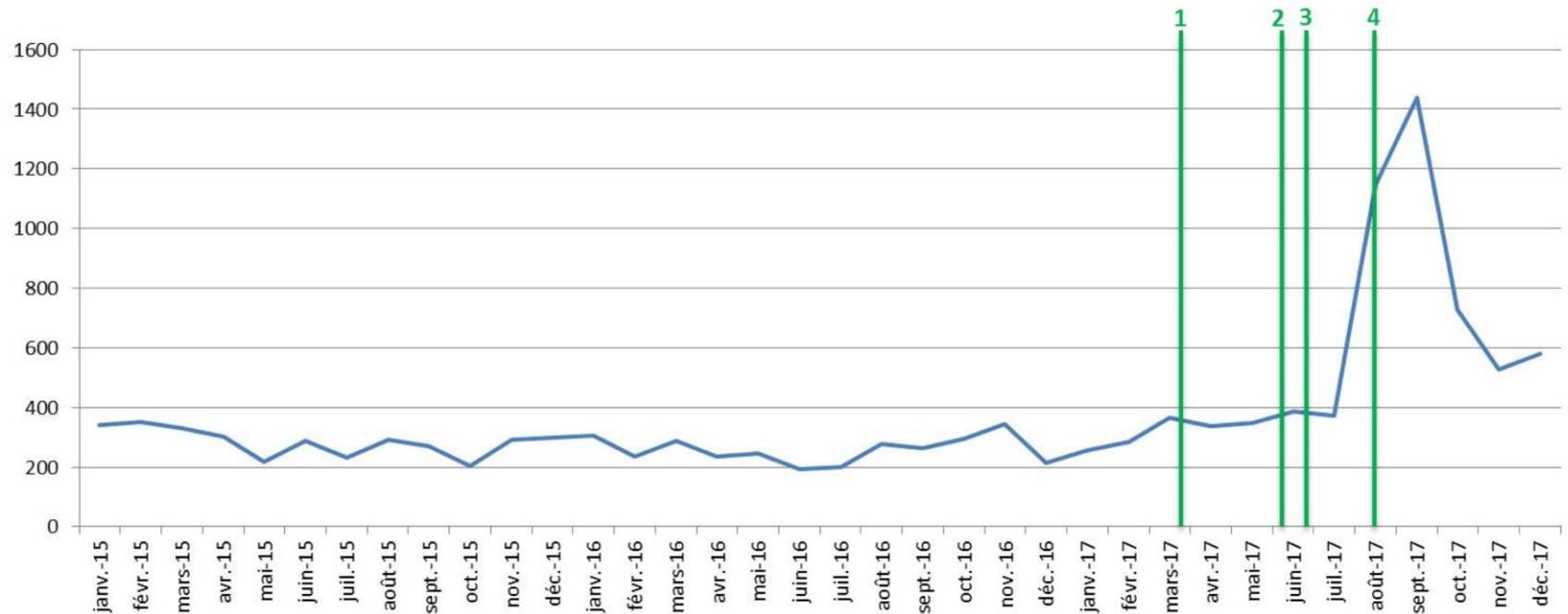
Forum: « vivre sans thyroïde »

Analyse temporelle: effets indésirables par catégorie





El associés au Levothyrox[®] janvier 2015 à Décembre 2017



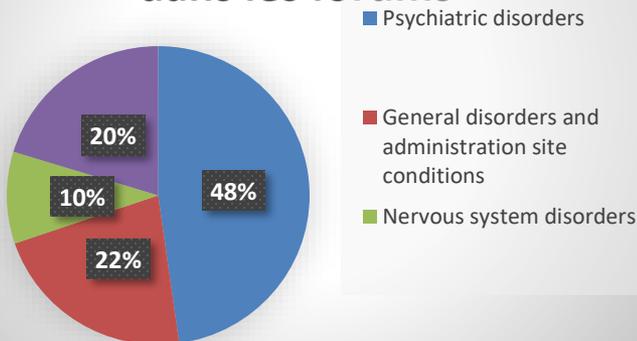
- (1) Marketing of the new formulation of Levothyrox[®]
- (2) First occurrence of an AE associated with the mentions “new formula” and “Levothyrox[®]” in a French web forums
- (3) Online petition against the new formulation
- (4) First article in a national newspaper about the Levothyrox[®] crisis



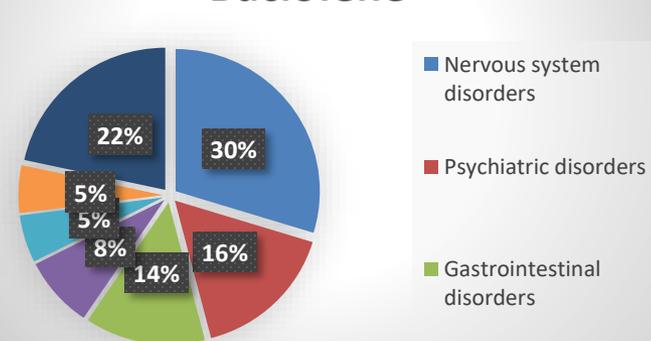
Détection du signal dans les réseaux sociaux : exemple du baclofène

- 2.622.445 couples issus de 747.500 messages (du 02/07/2000 au 02/07/2015)
 - 999 événements indésirables différents retrouvés avec le Baclofène

Distribution des classes d'EIs retrouvés avec le Baclofène dans les forums



Distribution des classes d'EIS retrouvés dans le RCP du Baclofène





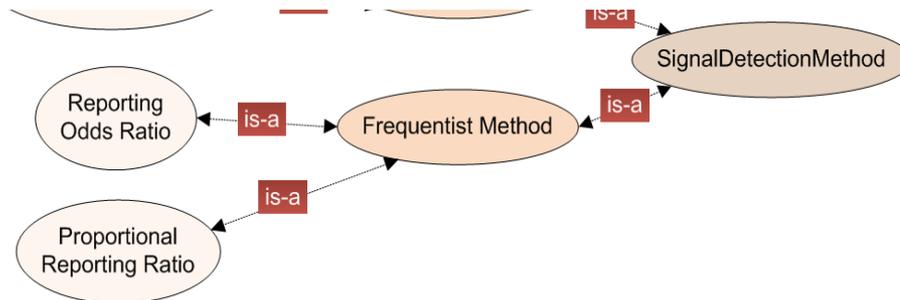
Détection du signal dans les réseaux sociaux : exemple du baclofène

- Détection du signal – Analyse rétrospective
 - Méthodes fréquentistes (PRR, ROR, RFET)
 - Méthodes bayésiennes (BCPNN, GPS)

Table 2. Proportion of AEs in the SPC considered as a signal for each decision rule

Method	PRR1	PRR2	ROR1	ROR2	RFET	midRFET	BCPNN1	BCPNN2	GPS1	GPS2
%SPC signal	50%	64%	64%	64%	64%	64%	61%	64%	36%	64%

Example Classification of
Signal Detection Methods





Détection du signal dans les réseaux sociaux : exemple du baclofène

- Parmi les nouveaux Eis retrouvés comme signaux par les algorithmes, 4 sont confirmés par l'algorithme **CISL** sur toute la base
 - Class-imbalanced subsampling lasso algorithm

EI	N	N couples	N Baclofène	CISL**
Glaucome	236	788	39	Signal
Boulimie	199	2289	268	Signal
SAS*	208	802	127	Signal
Cauchemars	508	6913	286	Signal



Karapetiantz P, Audeh B, Lillo-Le Louët A, Bousquet C. **Signal Detection for Baclofen in Web Forums: A Preliminary Study.** *Stud Health Technol Inform* 2018

Karapetiantz P, Audeh B, Faille J, Lillo-Le Louët A, Bousquet C.
Qualitative and Quantitative Analysis of Web Forums for Adverse Events Detection: "Strontium Ranelate" Case Study. *Stud Health Technol Inform.* 2019

Chen X, Faviez C, Schuck S, Lillo-Le-Louët A, Texier N, Dahamna B, Huot C, Foulquié P, Pereira S, Leroux V, Karapetiantz P, Guenegou-Arnoux A, Katsahian S, Bousquet C, Burgun A. **Mining Patients' Narratives in Social Media for Pharmacovigilance: Adverse Effects and Misuse of Methylphenidate.** *Front Pharmacol.* 2018



Merci de votre attention

Marie-Christine Jaulent
DR INSERM
LIMICS



Cédric Bousquet, PhD
PH CHU St Etienne
LIMICS



Pierre Zweigenbaum
DR CNRS
LIMSI



Pierre Karapetiantz
Statisticien
LIMICS



Bissan Audeh, PhD
Ingénieure Informatique
LIMICS



Agnes Lillo-Le Louet
MD, Paharmacovigilant
HEGP



Cyril Grouin, PhD
IE CNRS
LIMSI

