

Analyse des parcours de soins vus par le SNIIRAM

**Enjeux et défis de
la recherche en informatique**

Thomas GUYET

LACODAM
IRISA / INRIA

Erwan DREZEN

Plateforme PEPS/ANSM
UPRES EA 7449 – REPERES

Épidémiologie numérique

▶ **Données numériques relatives à la santé : de plus en plus facilement exploitables**

- Diminution des problèmes de volumétrie
- Meilleure structuration / uniformisation
- Vecteur d'amélioration des systèmes de soins et des soins

▶ **Un concept important : le parcours de soins**

- « Ensemble des événements de soins associés à un patient »
- Informations temporelles, taxonomiques, ...

▶ **Le cas du SNIIRAM**

- Comment l'utiliser dans l'étude des parcours de soins ?

SNIIRAM & Parcours de soins

► Que nous offre le SNIIRAM ?

- Dimension temporelle

Dates de délivrance de médicaments, de séjours hospitaliers, ...

- Dimension taxonomique

ATC, CIP, CIM10, CCAM, NABM, ...

- Données multi-dimensionnelles

1 délivrance = 1 médicament, 1 prescripteur, 1 pharmacien

► En pratique, comment y étudier des parcours ?

- Écart sémantique entre l'information disponible (remboursements) et la connaissance médicale à en extraire
- Incomplétude (médicament non remboursé p.e.) / imprécision (erreur codage)
- Objets complexes par rapport à des infos tabulaires

SNIIRAM & Parcours de soins

- ▶ **Ex: récupérer dans le SNIIRAM la séquence des médicaments délivrés à un patient**

```
SELECT
  EXE_SOI_DTD, PHA_PRS_IDE, PFS_PRE_NUM, PFS_EXE_NUM,
  PSE_SPE_COD, PSP_SPE_COD, PSE_ACT_NAT, PSP_ACT_NAT
FROM
  ER_PRS_F as prs,
  ER_PHA_F as pha
WHERE
  (prs.FLX_DIS_DTD = pha.FLX_DIS_DTD) AND (prs.FLX_TRT_DTD = pha.FLX_TRT_DTD) AND
  (prs.FLX_EMT_TYP = pha.FLX_EMT_TYP) AND (prs.FLX_EMT_NUM = pha.FLX_EMT_NUM) AND
  (prs.FLX_EMT_ORD = pha.FLX_EMT_ORD) AND (prs.ORG_CLE_NUM = pha.ORG_CLE_NUM) AND
  (prs.DCT_ORD_NUM = pha.DCT_ORD_NUM) AND (prs.PRS_ORD_NUM = pha.PRS_ORD_NUM) AND
  (prs.REM_TYP_AFF = pha.REM_TYP_AFF) AND
  (BEN_NIR_PSA = '123456789')
ORDER BY
  EXE_SOI_DTD,
  PHA_PRS_IDE;
```

SNIIRAM & Parcours de soins

- ▶ Ex: récupérer dans le SNIIRAM la séquence des médicaments délivrés à un patient

EXE_SOI_DTD	PHA_PRS_IDE	PFS_PRE_NUM	PFS_EXE_NUM	PSE_SPE_COD	PSP_SPE_COD	PSE_ACT_NAT	PSP_ACT_NAT
2013-01-05	3200805	A30339679	A30435923	0	1	50	0
2013-01-05	3474419	A30339679	A30435923	0	1	50	0
2013-01-05	3531413	A30339679	A30435923	0	1	50	0
2013-01-05	3555537	A30339679	A30435923	0	1	50	0
2013-01-05	3599397	A30339679	A30435923	0	1	50	0
2013-01-05	3890891	A30339679	A30435923	0	1	50	0
2013-01-15	3230901	A30339679	A30435923	0	1	50	0
2013-01-15	3404459	A30339679	A30435923	0	1	50	0
2013-02-11	3368663	A30339679	A30435984	0	1	50	0
2013-02-11	3436850	A30339679	A30435984	0	1	50	0
2013-02-11	3474419	A30339679	A30435984	0	1	50	0

SNIIRAM & Parcours de soins

- ▶ **Ex: récupérer dans le SNIIRAM la séquence des médicaments délivrés à un patient**

```
SELECT
  EXE_SOI_DTD, PHA_PRS_IDE, PFS_PRE_NUM,
  PSE_SPE_COD, PSP_SPE_COD, PSE_ACT_NAT,
FROM
  ER_PRS_F as prs,
  ER_PHA_F as pha
WHERE
  (prs.FLX_DIS_DTD = pha.FLX_DIS_DTD) AND
  (prs.FLX_EMT_TYP = pha.FLX_EMT_TYP) AND
  (prs.FLX_EMT_ORD = pha.FLX_EMT_ORD) AND
  (prs.DCT_ORD_NUM = pha.DCT_ORD_NUM) AND
  (prs.REM_TYP_AFF = pha.REM_TYP_AFF) AND
  (BEN_NIR_PSA = '123456789')
ORDER BY
  EXE_SOI_DTD,
  PHA_PRS_IDE;
```

EXE_SOI_DTD	PHA_PRS_IDE	PFS_PRE_NUM	PFS_EXE_NUM
2013-01-05	3200805	A30339679	A30435923
2013-01-05	3474419	A30339679	A30435923
2013-01-05	3531413	A30339679	A30435923
2013-01-05	3555537	A30339679	A30435923
2013-01-05	3599397	A30339679	A30435923
2013-01-05	3890891	A30339679	A30435923
2013-01-15	3230901	A30339679	A30435923
2013-01-15	3404459	A30339679	A30435923
2013-02-11	3368663	A30339679	A30435984
2013-02-11	3436850	A30339679	A30435984
2013-02-11	3474419	A30339679	A30435984

- ▶ **Constats**

- **Technicité importante (SQL)**
- **Connaissance fine requise des tables**
- **Phase d'apprentissage importante**



Le chercheur face au SNIIRAM

► Aujourd'hui

- Interroger des données : requêtes SQL
- Vérifier des hypothèses : tests statistiques (SAS, R)

► Demain ? ... des besoins complémentaires

- Visualiser des parcours de soins : outils de visualisation
- Interroger des parcours de soins : requêtes « médicales »
- Extraire des connaissances : fouille de données

► Besoin d'outils spécifiques pour le SNIIRAM

- Notamment à cause de cette notion de parcours de soins
- Des défis pour la recherche en sciences des données

Modèle de données PEPS

► Vision « parcours de soins » du SNIIRAM

- Vocabulaire technique commun entre les acteurs de PEPS
- Conserve la structure de la base de données SNIIRAM
- Transformation flexible à la volée des données

► Concepts explicites pour l'épidémiologie

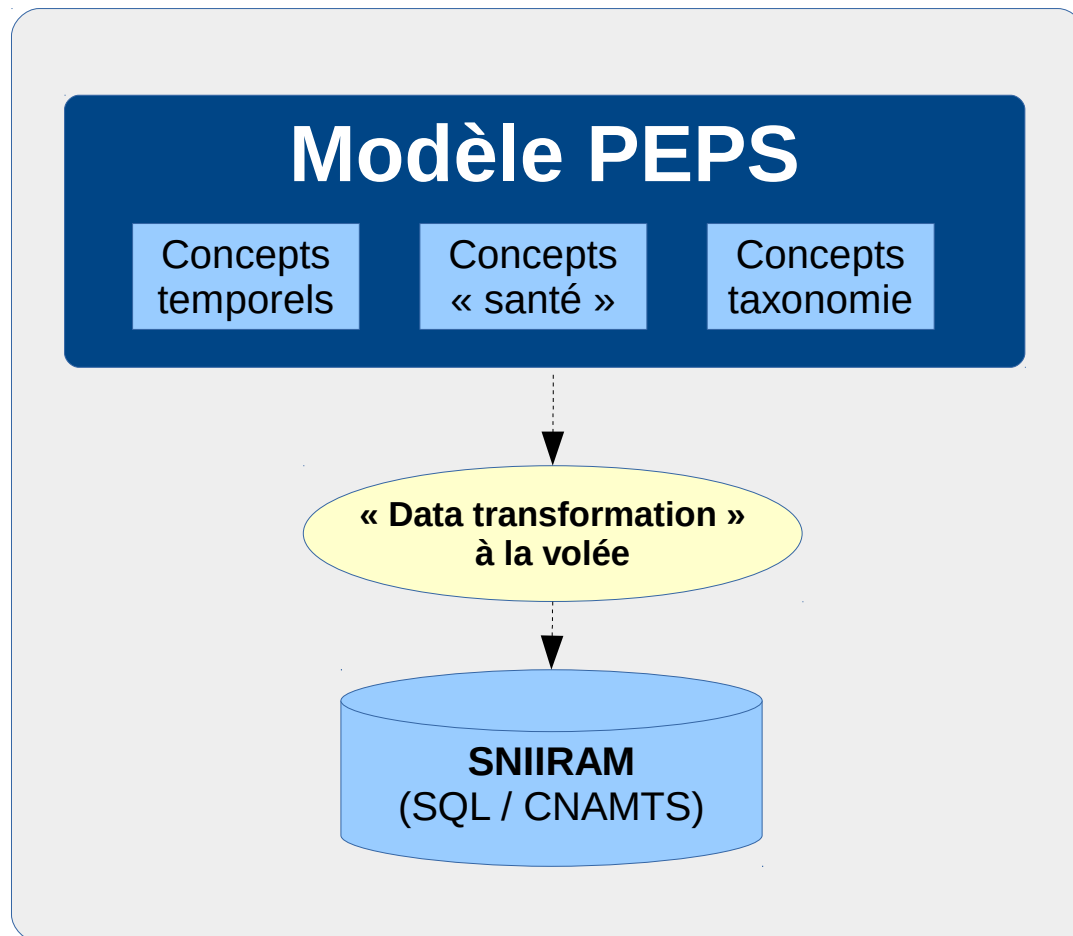
- Concepts médicaux : patient, médecin, séjour, médicament
- Concept de séquences temporelles : notion de « parcours »

► Implémenté sous forme d'un composant logiciel

- Brique de base pour le développement d'outils logiciels
- Utilisable via différentes technologies (java, python, webservice)

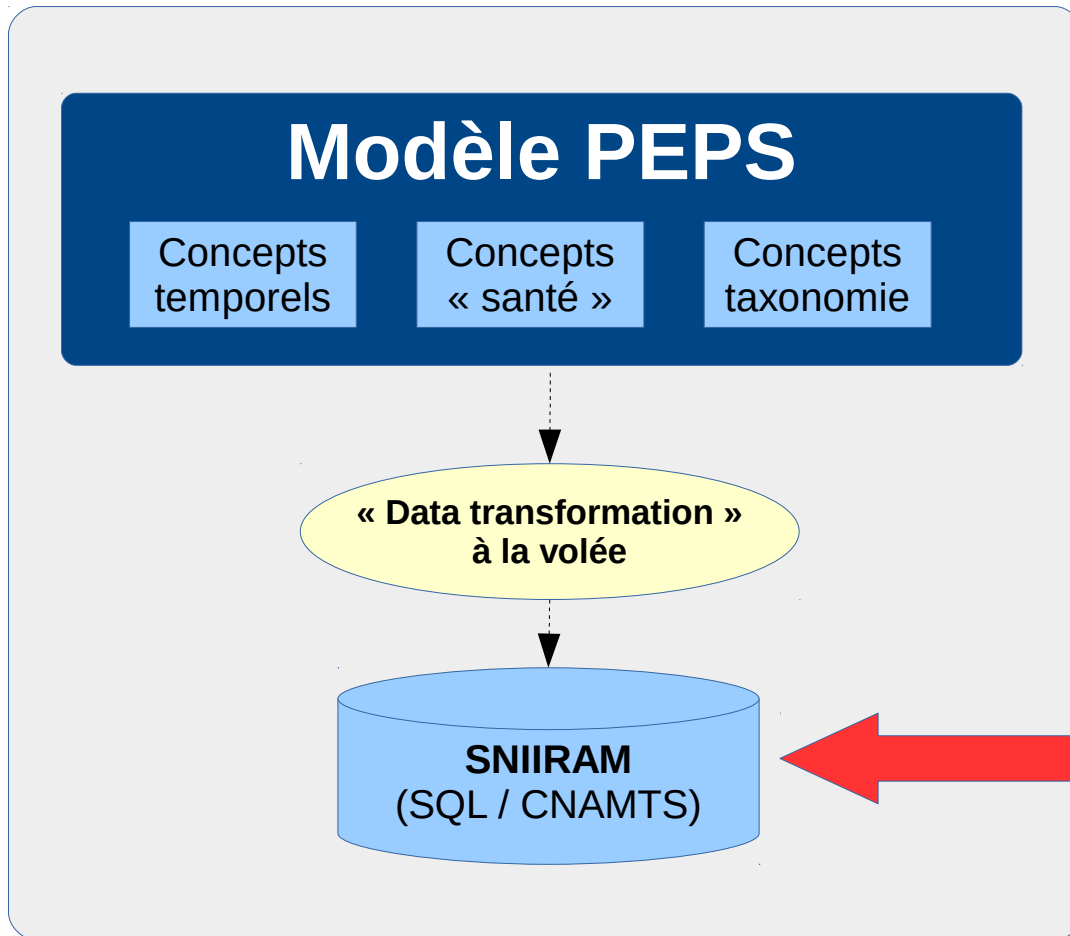
Modèle de données PEPS

1. Modéliser



Modèle de données PEPS

1. Modéliser

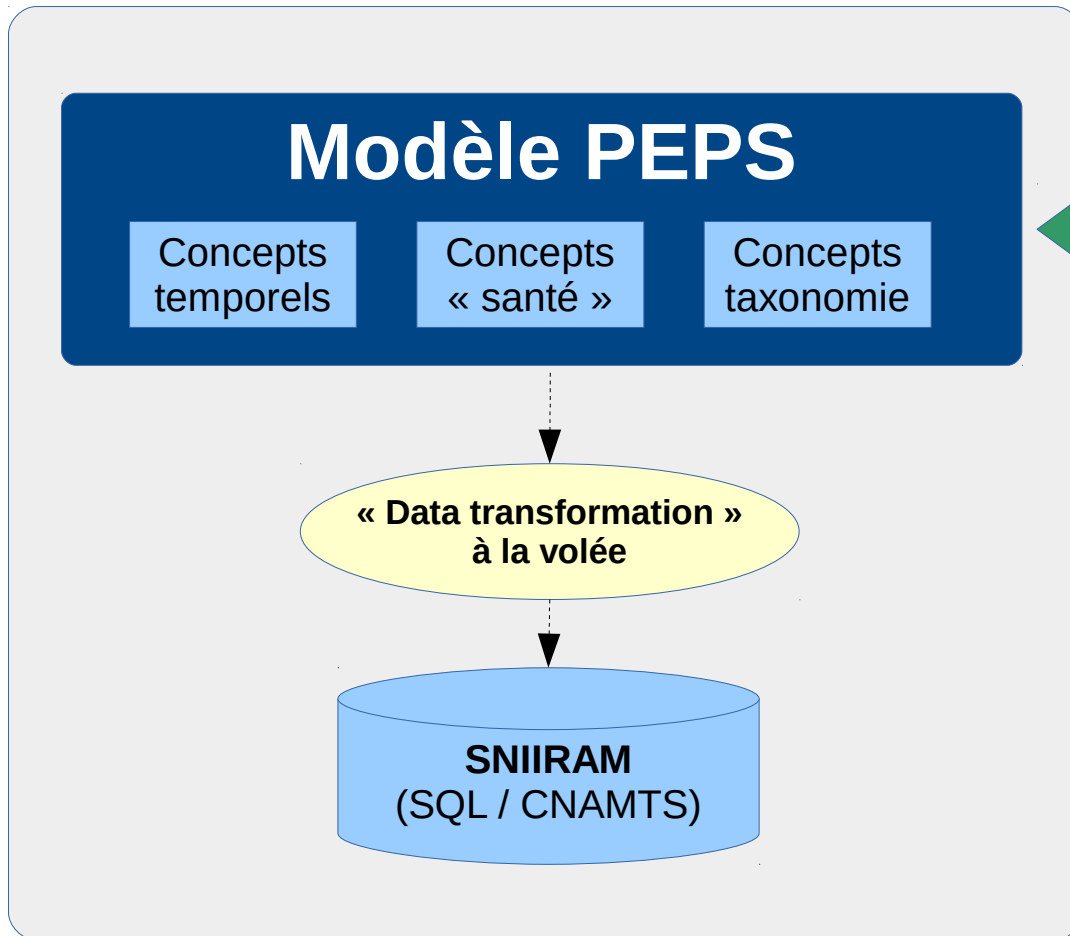


Avant...

```
SELECT
  EXE_SOI_DTD, PHA_PRS_IDE, PFS_PRE_NUM, PFS_EXE_NUM,
  PSE_SPE_COD, PSP_SPE_COD, PSE_ACT_NAT, PSP_ACT_NAT
FROM
  ER_PRS_F as prs,
  ER_PHA_F as pha
WHERE
  (prs.FLX_DIS_DTD = pha.FLX_DIS_DTD) AND
  (prs.FLX_TRT_DTD = pha.FLX_TRT_DTD) AND
  (prs.FLX_EMT_TYP = pha.FLX_EMT_TYP) AND
  (prs.FLX_EMT_NUM = pha.FLX_EMT_NUM) AND
  (prs.FLX_EMT_ORD = pha.FLX_EMT_ORD) AND
  (prs.ORG_CLE_NUM = pha.ORG_CLE_NUM) AND
  (prs.DCT_ORD_NUM = pha.DCT_ORD_NUM) AND
  (prs.PRS_ORD_NUM = pha.PRS_ORD_NUM) AND
  (prs.REM_TYP_AFF = pha.REM_TYP_AFF) AND
  (BEN_NIR_PSA = '123456789')
ORDER BY
  EXE_SOI_DTD,
  PHA_PRS_IDE;
```

Modèle de données PEPS

1. Modéliser



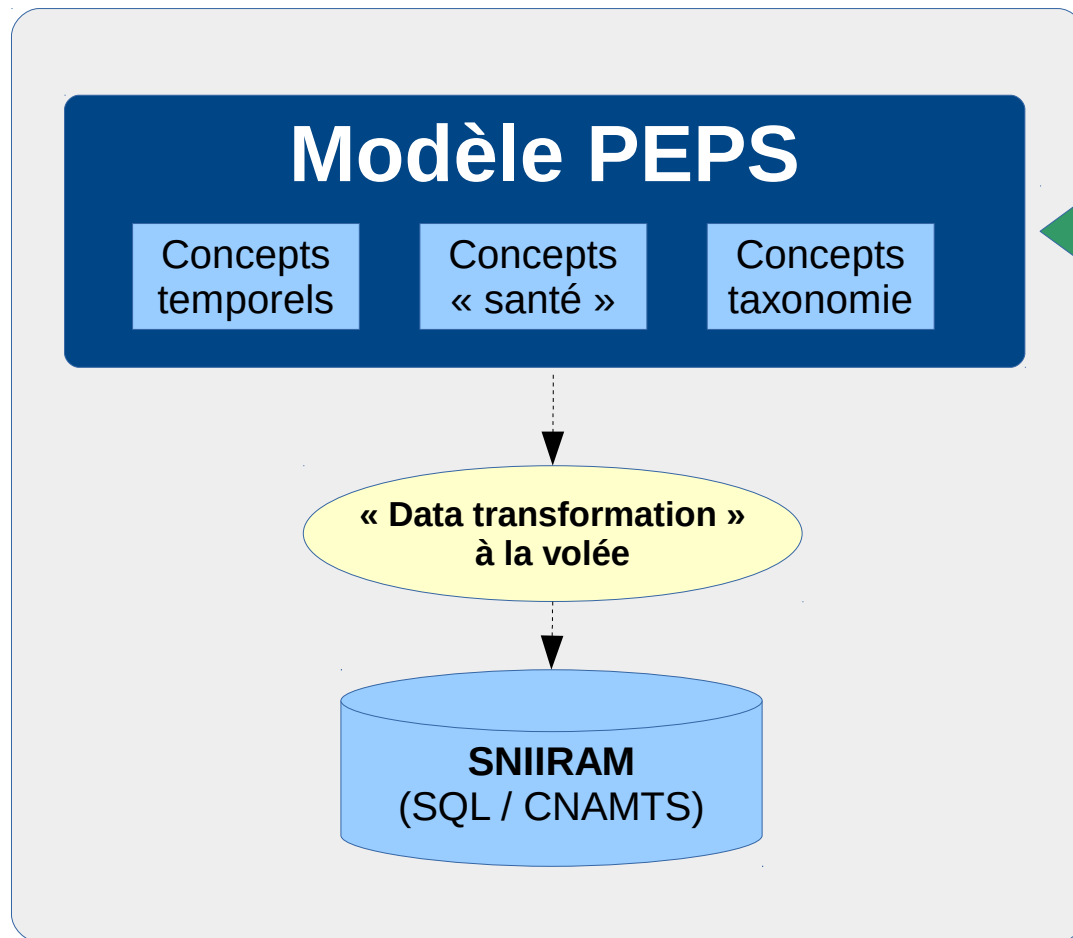
Maintenant...

```
for p in patient.prestaMedicaments:  
    print p.begin  
  
    print p.prescriber.type  
  
    print p.medicament.cip  
    print p.medicament.atc
```

```
SELECT  
    EXE_SOI_DTD, PHA_PRS_IDE, PFS_PRE_NUM, PFS_EXE_NUM,  
    PSE_SPE_COD, PSP_SPE_COD, PSE_ACT_NAT, PSP_ACT_NAT  
FROM  
    ER_PRS_F as prs,  
    ER_PHA_F as pha  
WHERE  
    (prs.FLX_DIS_DTD = pha.FLX_DIS_DTD) AND  
    (prs.FLX_TRT_DTD = pha.FLX_TRT_DTD) AND  
    (prs.FLX_EMT_TYP = pha.FLX_EMT_TYP) AND  
    (prs.FLX_EMT_NUM = pha.FLX_EMT_NUM) AND  
    (prs.FLX_EMT_ORD = pha.FLX_EMT_ORD) AND  
    (prs.ORG_CLE_NUM = pha.ORG_CLE_NUM) AND  
    (prs.DCT_ORD_NUM = pha.DCT_ORD_NUM) AND  
    (prs.PRS_ORD_NUM = pha.PRS_ORD_NUM) AND  
    (prs.REM_TYP_AFF = pha.REM_TYP_AFF) AND  
    (BEN_NIR_PSA = '123456789')  
ORDER BY  
    EXE_SOI_DTD,  
    PHA_PRS_IDE;
```

Modèle de données PEPS

1. Modéliser



Maintenant...

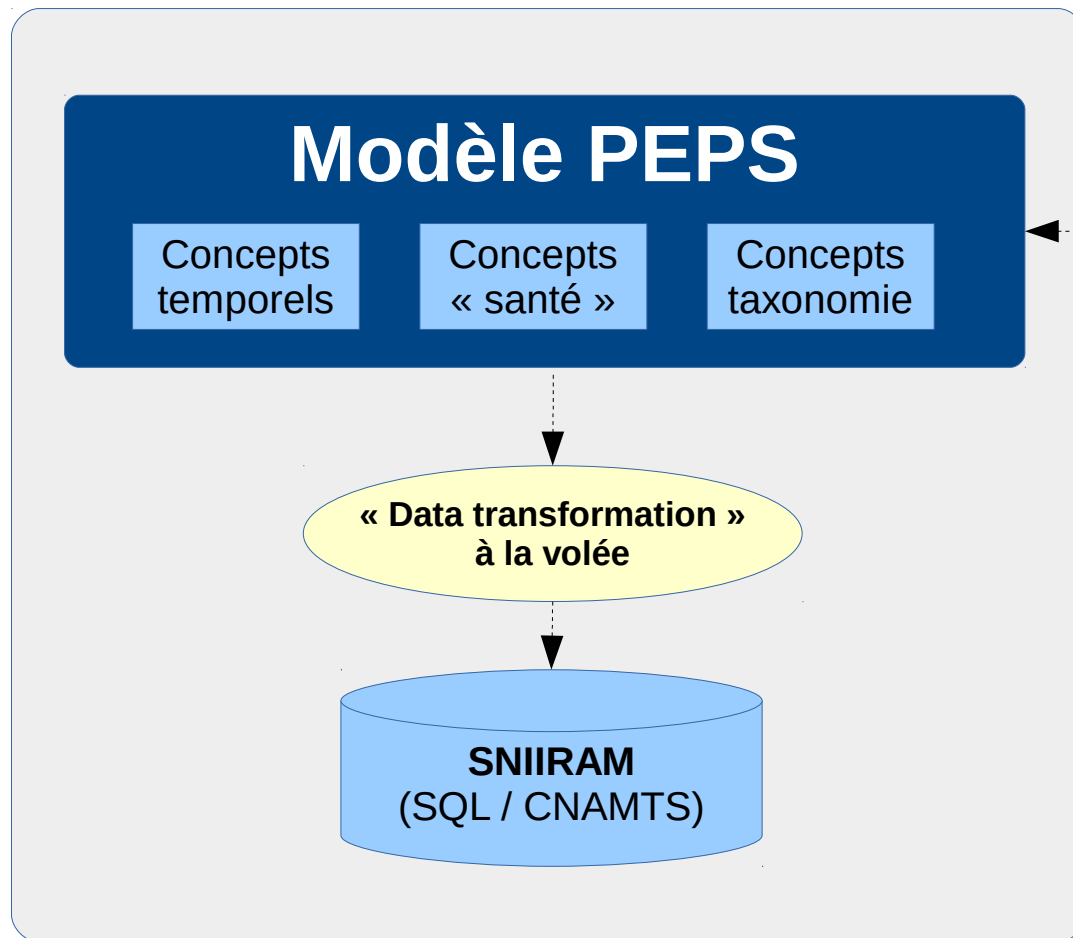
```
for p in patient.prestaMedicaments:  
    print p.begin  
  
    print p.prescriber.type  
  
    print p.medicament.cip  
    print p.medicament.atc
```

« Data Transformation »

- Effectue les requêtes SQL sur le SNIIRAM
- Peut s'adapter aux évolutions du SNIIRAM
- Modèle PEPS et outils indépendants du SNIIRAM

Modèle de données PEPS

1. Modéliser



2. Manipuler



« Data Transformation »

- Effectue les requêtes SQL sur le SNIIRAM
- Peut s'adapter aux évolutions du SNIIRAM
- Modèle PEPS et outils indépendants du SNIIRAM

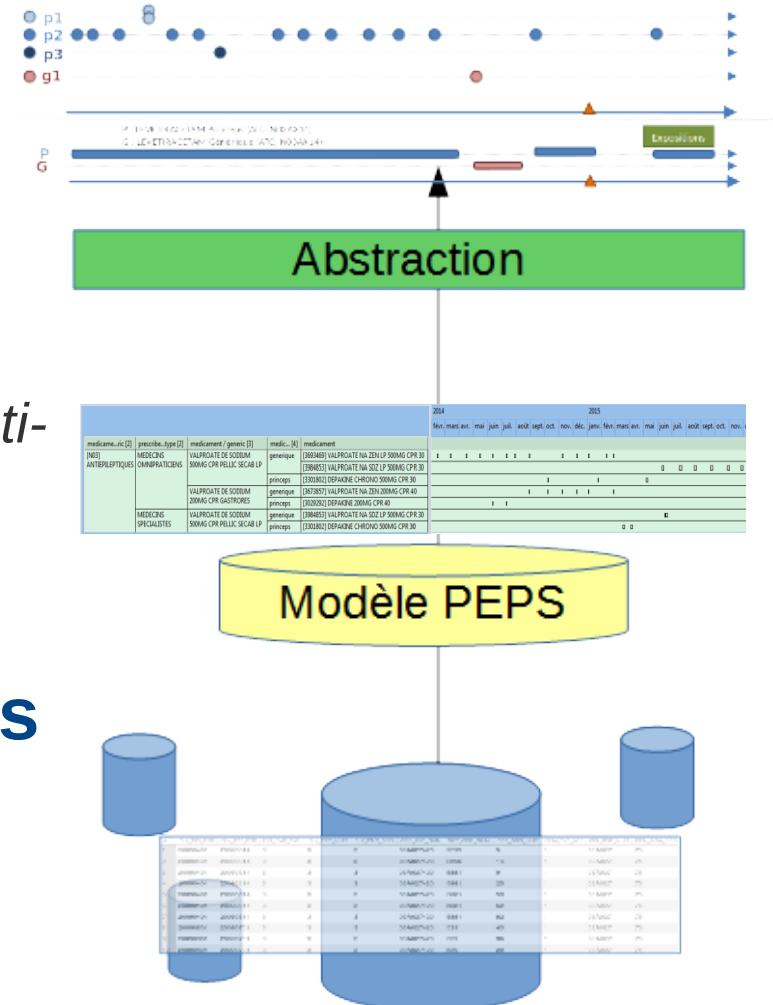
Abstraire des parcours de soins

► Créer des vues personnalisées et adaptées aux études

- Modèle PEPS : une vue médicale des données
- Abstraction des parcours : choix des variables, brutes ou calculées, qui sont utilisées pour l'analyse
- Exemple de proxy : **expositions aux anti-épileptiques**

► Objectif : améliorer la capacité d'analyse des épidémiologistes

- Flexibilité : répondre aux besoins spécifiques d'un grand nombre d'études
- Expressivité : permettre des vues complexes



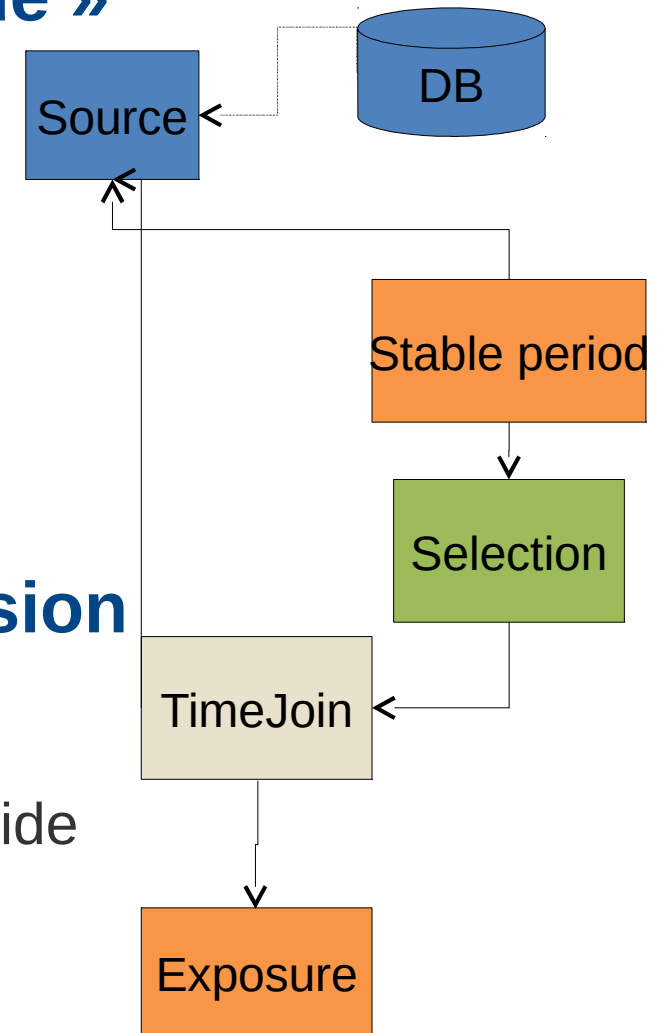
Abstraire des parcours de soins

► Outil de conception des « vues-étude » d'un parcours de soins

- Faciliter la conception des vues abstraites
- Difficultés additionnelles : dimensions temporelles et ontologiques

► Idée directrice : conception d'un langage « SQL-like » pour l'expression des abstractions

- « Langage » : formalisation informatique solide
- Syntaxe adaptée à l'expression de proxies sémantiquement riches
- Automatisation de l'exécution des requêtes



Analyser des parcours de soins

► Vers des analyses de parcours de soins

- Pour **quels patients** observe-t-on une séquence de soins ?
- Quelles sont les **séquences de soins communes** aux parcours d'un groupe de patients ?
- Existe-t-il des groupes de patients aux **parcours similaires** ?

► Objectif : induire de nouvelles connaissances à partir des « parcours de soins »

- Éviter une nouvelle simplification des données pour conserver la richesse sémantique
- Exploration (semi-)automatiquement

► Besoin d'outils d'analyse spécifiques

Outils PEPS : vue générale

► Différentes tâches de traitement des parcours de soins

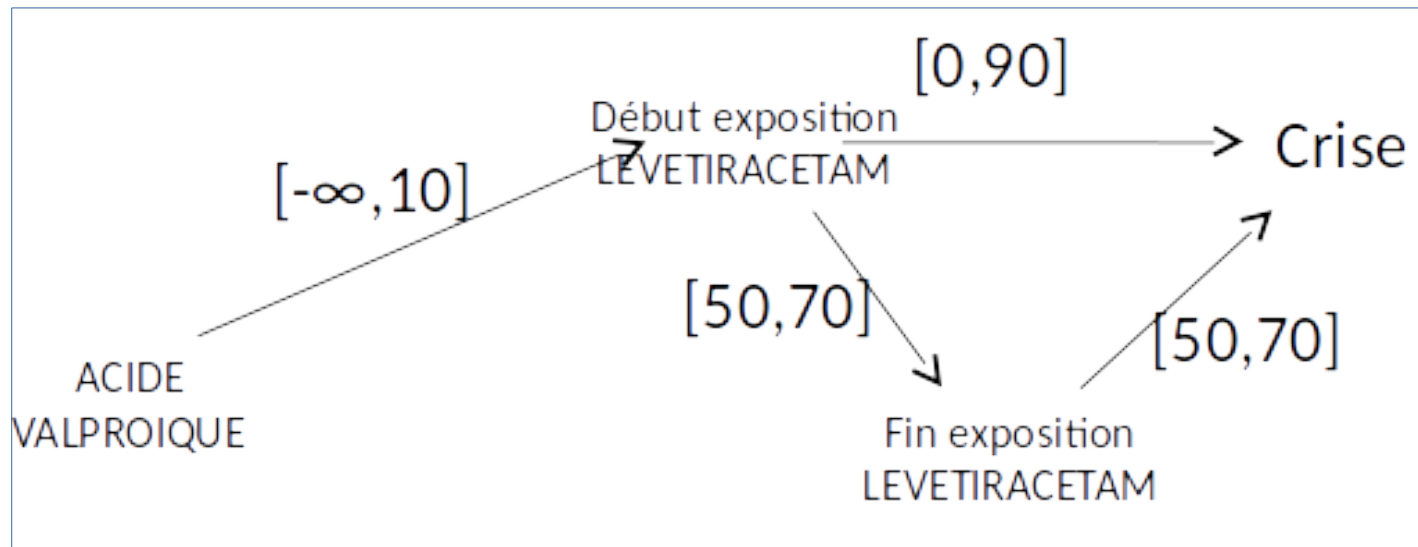
- Interrogation de parcours de soins
- Comparaison de parcours de soins
- Fouilles de données séquentielles

► Différentes approches

- Approches inspirées des méthodes du web sémantique (SPARQL)
- Approches inspirées de la programmation logique (ASP)
- Approches algorithmiques : chroniques

Analyser des parcours de soins

- **Modèle de chroniques pour l'interrogation et la fouille de séquences de soins**



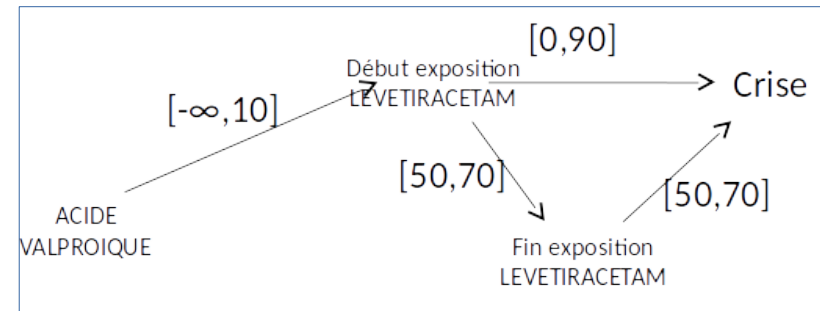
Exemple de chronique

(contexte d'étude GenEpi – patients épileptiques)

Interroger les parcours de soins

► Identifier des parcours de soins conformement à une séquence

- Sélectionner des patients
- Nouveau proxy : induction de nouveaux éléments (cf abstraction)



► Modèle de séquences de soins

- Propriétés recherchées
 - Expressif : dans les dimensions temporelles et ontologiques
 - Efficace : capable de traiter efficacement de longs parcours de soins
 - Facilement exprimable / interprétable
- De nombreux modèles existants et à explorer dans la littérature informatique
- Modèle de chronique répond à ces exigences

Découvrir des séquences de soins

► Abduction automatique de chroniques dans de larges ensembles de parcours de soins

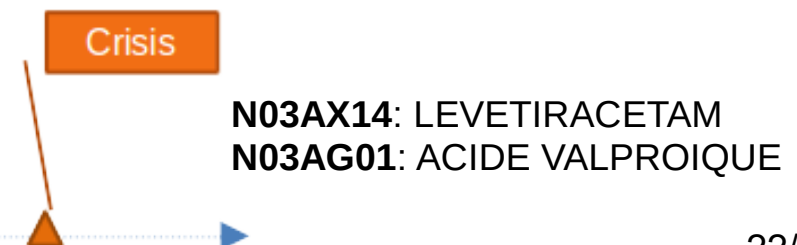
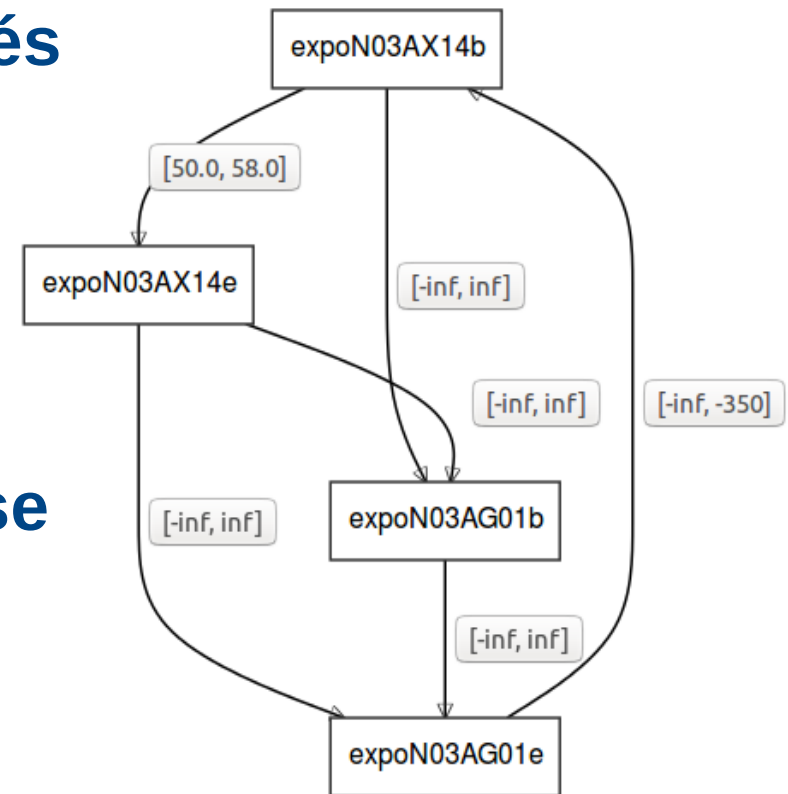
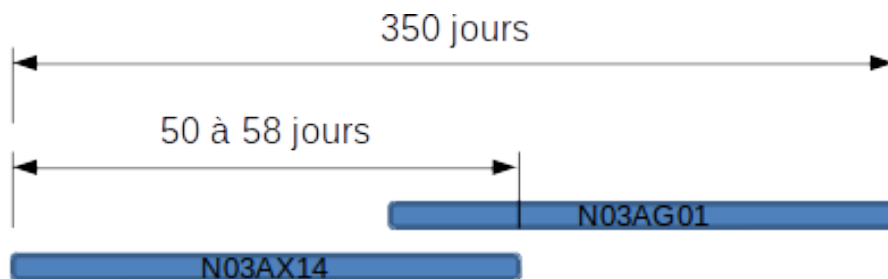
- Principe de la « fouille de données » : découvrir les chroniques qui sont « cachées » dans les données
- Développement d'algorithmes de découverte de
 - **Chroniques fréquentes** : séquences de soins qui apparaissent régulièrement dans un ensemble de parcours
 - **Chroniques discriminantes** : séquences de soins plus spécifiques à un groupe de parcours qu'à un autre

► Découvrir des chroniques caractéristiques dans les données

- Raffiner la connaissance d'une séquence de soins à partir des données
- Suggérer l'existence de séquences cachées

Exemple pratique

- ▶ **Extraction de chroniques discriminant les patients ayant eu une crise parmi ceux exposés aux antiépileptiques**
- ▶ **Avec reconstruction des expositions aux médicaments**
- ▶ **Chronique = Nouvelle hypothèse à explorer**
 - Un type de séquence de soins ?
 - Un groupe de patients « similaires » ?



- ▶ **Le SNIIRAM : une opportunité majeure pour étudier les parcours de soins**
 - Données longitudinales finalement assez riches (++) avec appariements)
- ▶ **Mais tâche non triviale !**
 - Matière brute complexe (remboursements)
 - Sémantique peu adaptée des outils traditionnels des chercheurs (SQL, SAS / R)
 - Besoin réel de nouveaux outils pour l'étude des parcours de soins
 - Voie PEPS R&D (modélisation, visualisation, chroniques, fouille de données...)
- ▶ **Des défis informatiques encore d'actualité pour traiter la notion complexe de parcours de soins**
 - Défis « sémantiques »
 - Défis « computationnels »

**Merci de
votre attention.**