

LE RENDEZ-VOUS INTERNATIONAL
CONFÉRENCES | ATELIERS | EXPOSITION

THE INTERNATIONAL MEETING
CONFERENCES | WORKSHOPS | EXHIBITION

HPC
SIMULATION
BIG DATA

Forum teratec

Unlock the future!

11 & 12 JUIN
JUNE 11 & 12 2019

ÉCOLE POLYTECHNIQUE
PALAISEAU-FRANCE

PLATINUM SPONSORS



GOLD SPONSORS



SILVER SPONSORS



PARTENAIRE
CAFÉ EUROPÉEN
DE LA RECHERCHE



AVEC LE SOUTIEN DE





Dr Warehouse – A translational Data Warehouse

Nicolas Garcelon

12 juin 2019

A Translational Research Context



Clinical activity
Hospital databases



600 beds
400 pediatric beds



34 Reference centers for rare diseases
25 Competence centers
Including 25 Reference centers integrated to *Imagine*

imagine
INSTITUT DES MALADIES GÉNÉTIQUES

Genetic research
Research databases

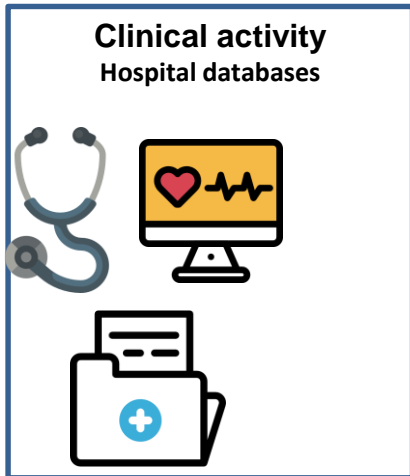


30 labs
12 platforms

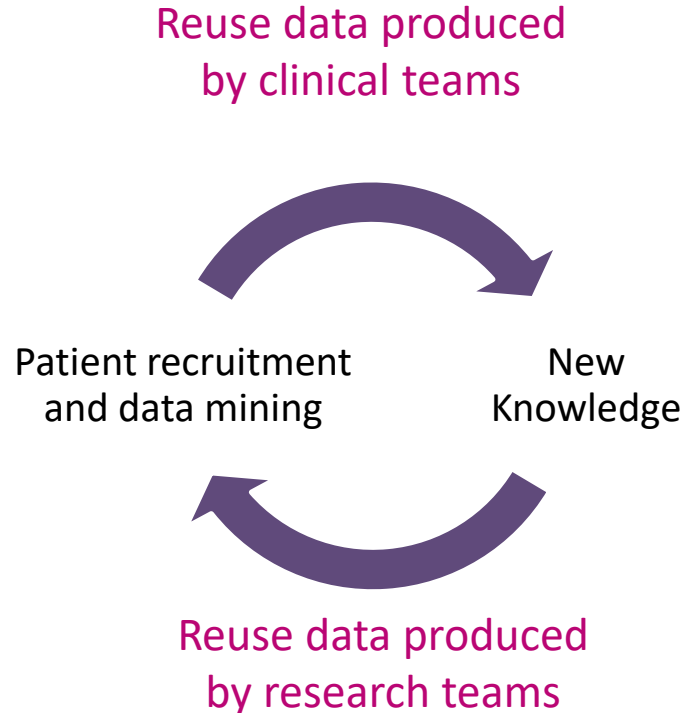
A Translational Research Context



imagine
INSTITUT DES MALADIES GÉNÉTIQUES



600 beds
400 pediatric beds





ELECTRONIC HEALTH RECORDS

=

MINE OF KNOWLEDGE

- ✓ **Clinical research:** inclusion into clinical trial, clustering
- ✓ **Epidemiological research:** cohort
- ✓ **Medico economic :** Quality, reporting
- ✓ **Evaluation of practices:** Study on the quality of patient care
- ✓ **Vigilances :** detection of adverse drug reactions, detection of nosocomial infections
- ✓ Etc.

Evidence-Based Medicine in the EMR Era

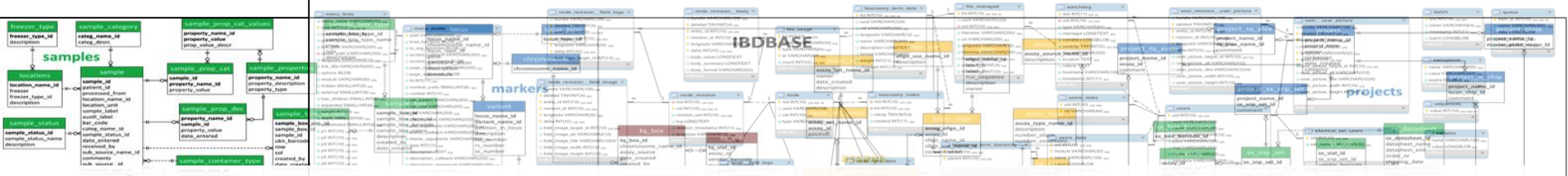
J.Frankovich, C.A. Longhurst, S M. Sutherland.

Stanford, publication NEJM 9 Nov 2011

Results of Electronic Search of Patient Medical Records (for a Cohort of 98 Pediatric Patients with Lupus) Focused on Risk Factors for Thrombosis Relevant to Our 13-Year-Old Patient with Systemic Lupus Erythematosus.*

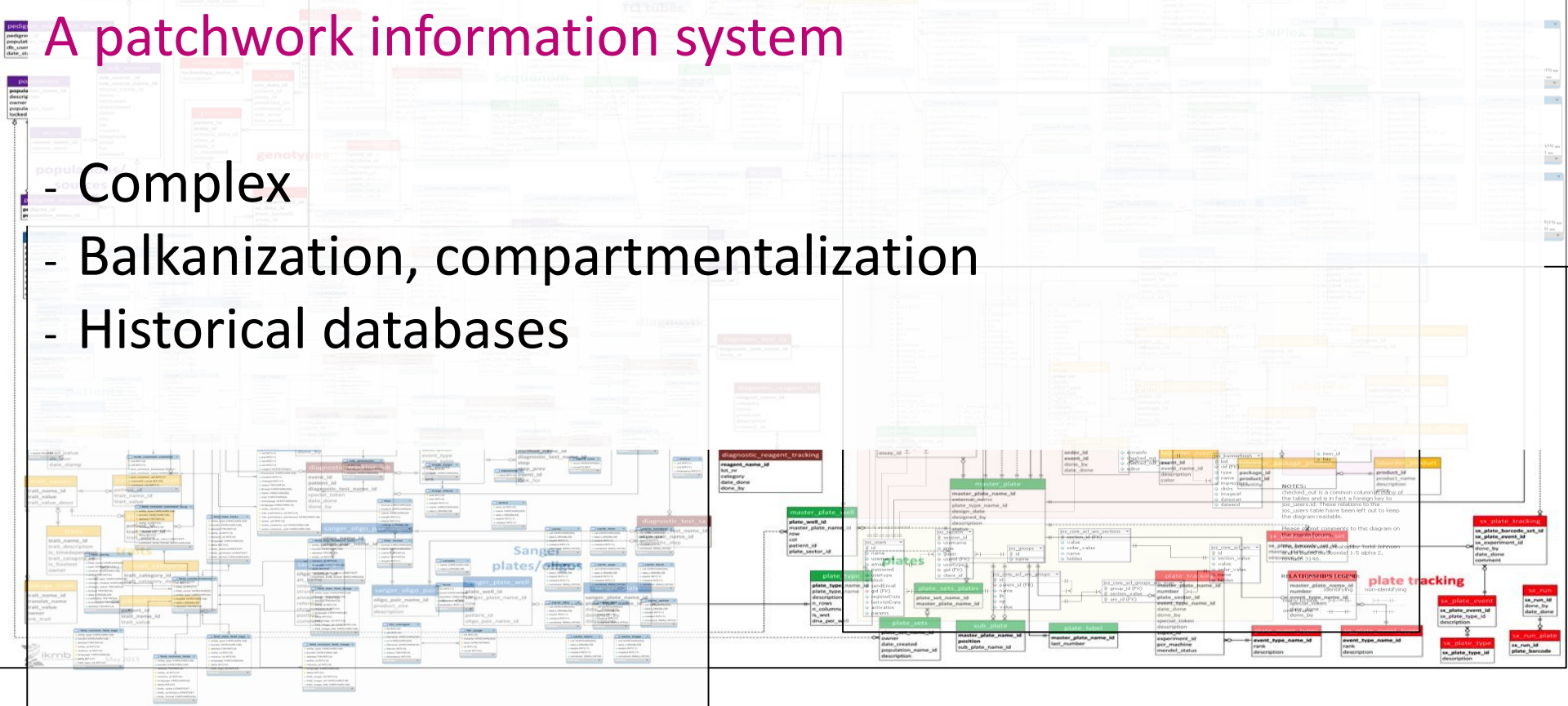
Outcome or Risk Factor	Keywords Used to Conduct Expedited Electronic Search	Prevalence of Thrombosis <i>no./total no (%)</i>	Relative Risk (95% CI)
Outcome — thrombosis	“Thrombus,” “Thrombosis,” “Blood clot”	10/98 (10)	Not applicable
Thrombosis risk factor			
Heavy proteinuria (>2.5 g per deciliter)			
Present at any time	“Nephrosis,” “Nephrotic,” “Proteinuria”	8/36 (22)	7.8 (1.7–50)
Present >60 days	“Urine protein”	7/23 (30)	14.7 (3.3–96)
Pancreatitis	“Pancreatitis,” “Lipase”	5/8 (63)	11.8 (3.8–27)
Antiphospholipid antibodies	“Aspirin”	6/51 (12)	1.0 (0.3–3.7)

« We made the decision on the basis of the best data available [...] in the light of experience as guided by intelligence. »

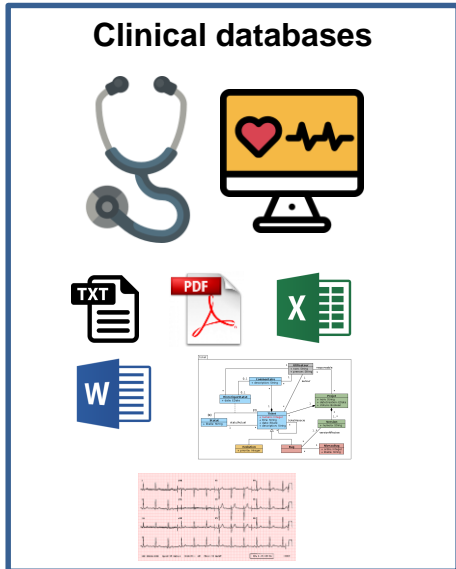


A patchwork information system

- Complex
- Balkanization, compartmentalization
- Historical databases

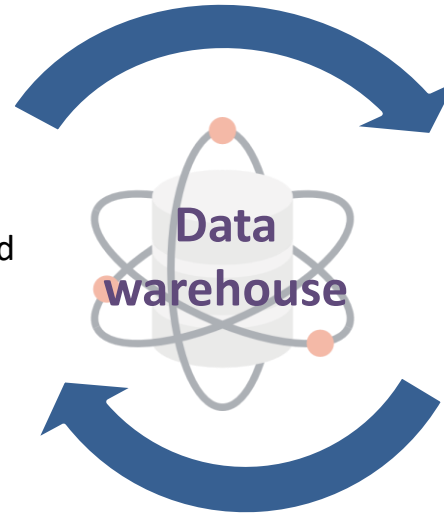


A Translational Research Context



Patient
recruitment and
data mining

Reuse data produced
by clinical teams



New
Knowledge

Reuse data produced
by research teams



A translational data warehouse

Access to the collective hospital memory to:

- Reduce diagnostic wandering
- Reduce therapeutic wandering
- Quickly create cohorts
- Accelerate patient recruitment into clinical trials
- Test hypotheses (drug + signs, sign + sign, etc.)

Biomedical data warehouse

Existing data warehouse

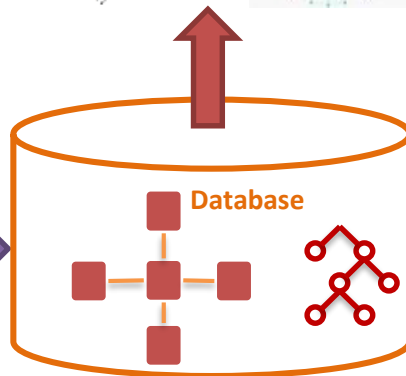
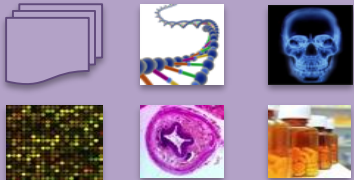
Knowledge databases



Search engine
Data mining algorithm



Clinical databases



(Murphy, 2006)

Stride

(Lowe, 2009)

Vanderbilt

(Danciu, 2014)

Starmaker

(Krasowski, 2015)

Meteor

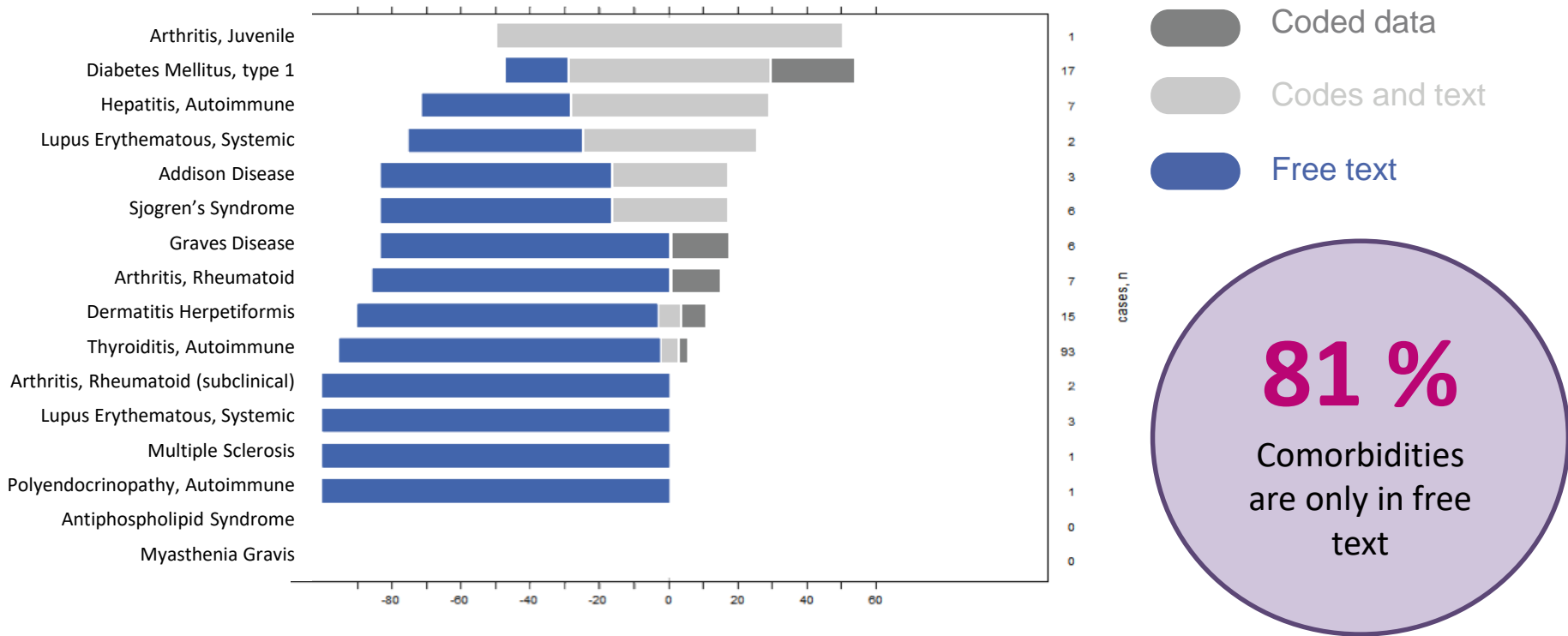
(Puppala, 2015)

OHDSI

(Hripcsak, 2015)

Coded data
(thesaurus)

A case study on autoimmune comorbidities in patients with celiac disease



Narrative records are essential

More detailed phenotypic description^[1,2]:

Description of undiagnosed patients

Expression of doubt and absence of signs

In the context of rare diseases: evolution of knowledge

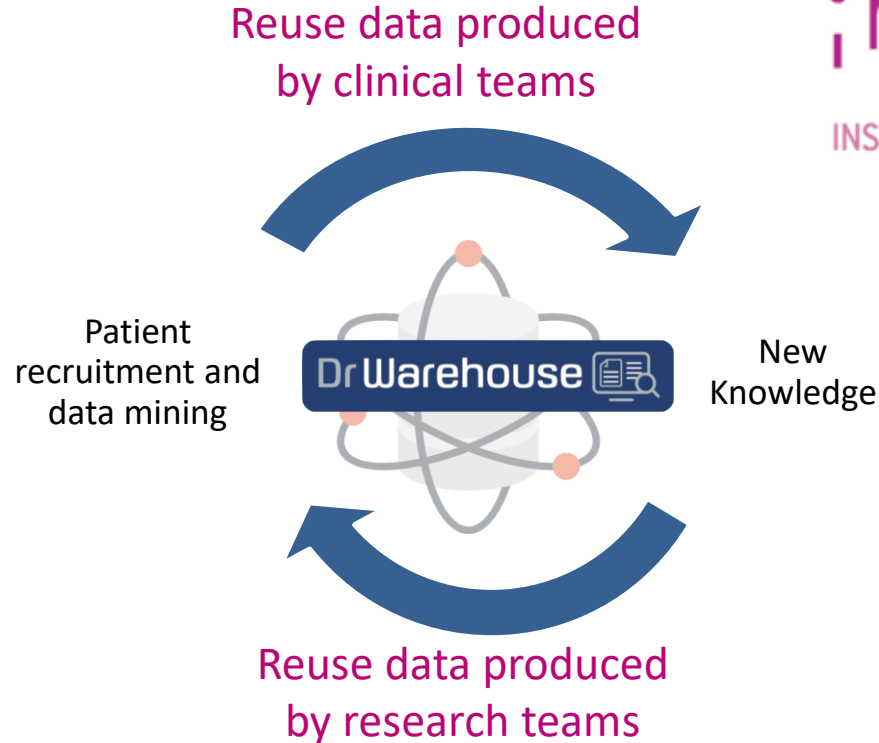
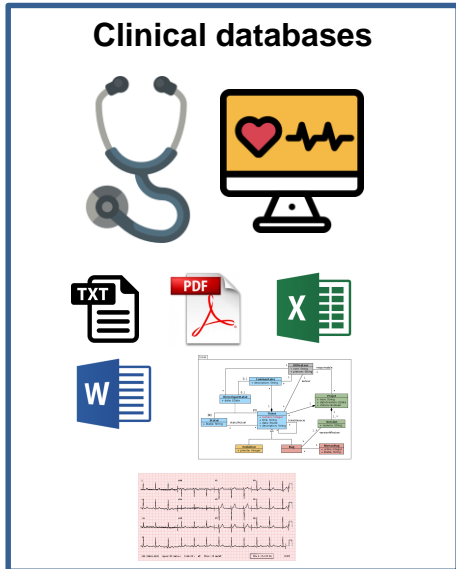
Clinicians have to write!



[1] Hanauer DA. EMERSE: The Electronic Medical Record Search Engine. *AMIA Annu Symp Proc* 2006;**2006**:941.

[2] Raghavan P et al. How essential are unstructured clinical narratives and information fusion to clinical trial recruitment? *AMIA Jt Summits Transl Sci Proc* 2014;**2014**:218–23.

A Translational Research Context

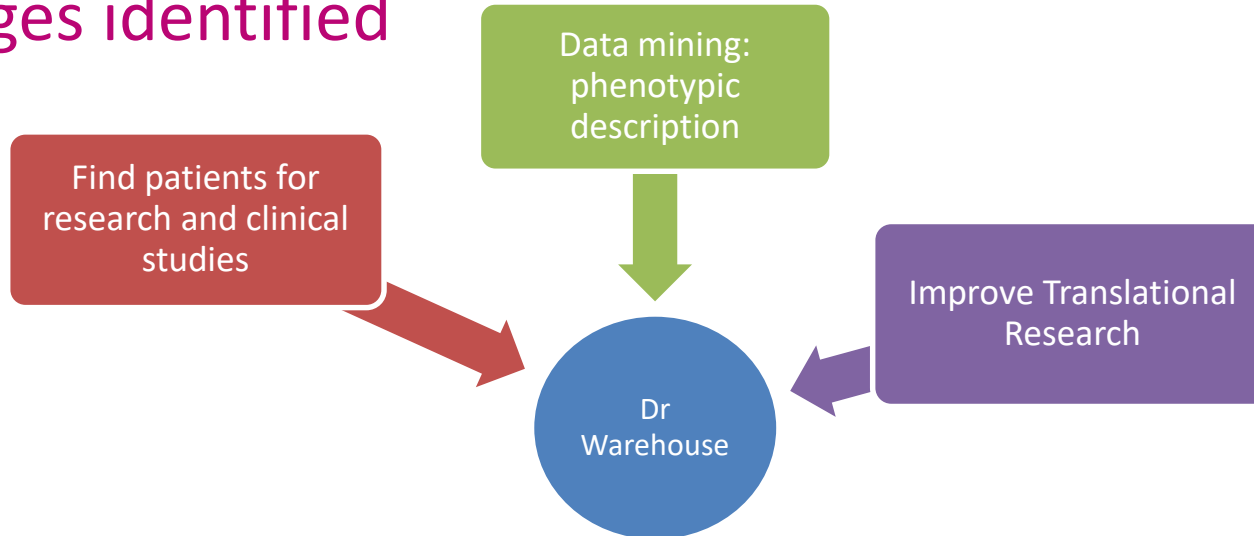


Dr Warehouse

A data warehouse oriented around the hospital document:

- For clinicians & data scientists
- For all types of data

3 challenges identified



Dr Warehouse in Necker

660 000 patients

5.5 M records

26 sources of data

From 1996 to 2019



290 users



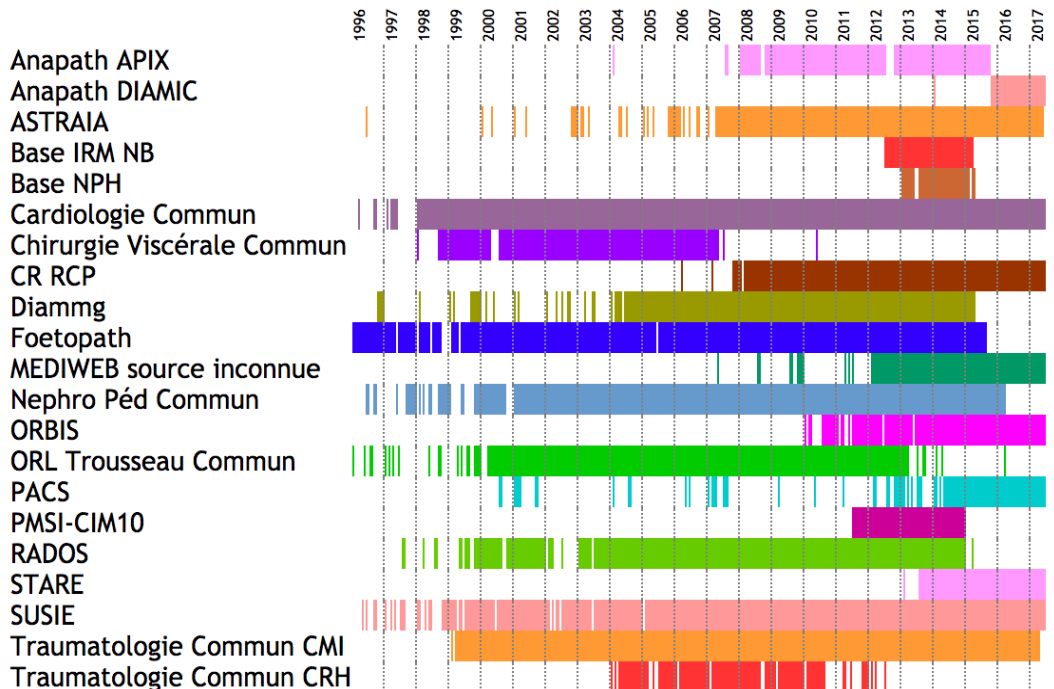
13 000 queries



70 000

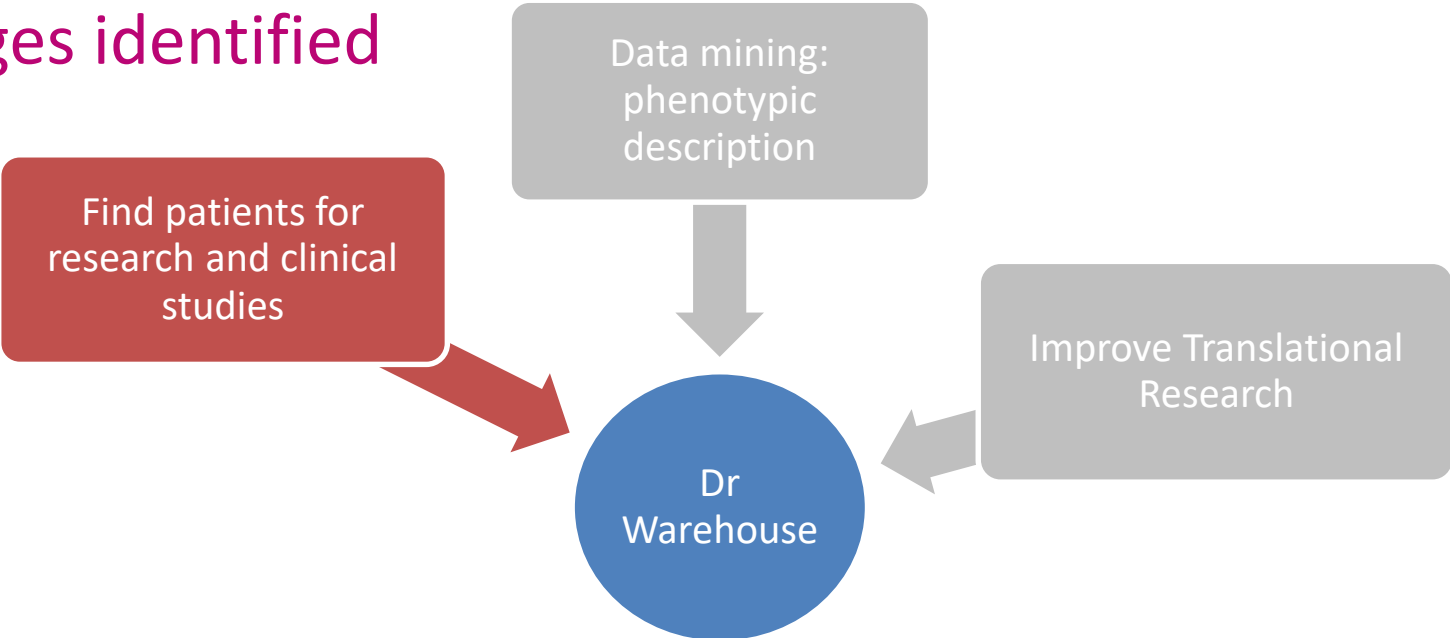


640 « cohorts »



Dr Warehouse

3 challenges identified



Find specific patients

- ✓ To recruit patients into a study

Time saving, and improvement of the inclusion rate

- ✓ To help make a decision on an undiagnosed patient

**From Individual memory to collective memory
organized and sustainable**

(Multidisciplinary Consultation Meeting)



An interface adapted to free text



Rechercher des patients

Sur les patients de vos services

eczema 891
Etendre aux synonymes :
[+ Avancé](#) - [Réécrire la requête](#)

thrombopénie 1157/8172
Etendre aux synonymes :
[+ Avancé](#) - [Réécrire la requête](#)

- + Ajouter un filtre Full text
- + Ajouter un filtre structuré
- + Contraintes temporelles
- + Filtre patient

LANCER LA RECHERCHE ?

Requêtes sauveées

Search:
Titre

amvotrophie spinale

① A search engine on textual and/or coded data

① A dedicated interface to check the consistency of the results

Rechercher des patients

Sur tout l'entrepôt

 341/341
 Etendre aux synonymes :
 + Avancé - Réécrire la requête

- + Ajouter un filtre Full text
- + Ajouter un filtre structuré
- + Filtre patient

LANCER LA RECHERCHE



Requêtes sauveées

- pericardite and calcification
- NETT et MECF2
- Spinaurie
- teratologie
- test mercapto met
- teratologie NEUROLOGIE PED

Historique requête

date	Requêtes
29/10/2015 22:20	Documents contenant 'epidermolyse bulleuse' et excluant les négations
15/10/2015 18:56	Documents contenant 'epidermolyse bulleuse' et excluant les négations
26/08/2015 14:52	Documents contenant 'epidermolyse bulleuse' et excluant les négations

Résultat

Cohortes

Données stat

Concepts

PMSI

Biologie

Genes

Map

Clustering

Datamart

341 Patients
2396 Documents

Documents contenant 'epidermolyse bulleuse', en excluant les négations

SAUVER LA REQUÊTE

Refaire une recherche sur le résultat :

Exporter les patients :

Filtrer le résultat ci-dessous : ok annuler

<p>11 ans</p> <p></p>	<p>Echographie abdominale du 15/10/15 par S. LAMBERT le 15/10/15 - RADIOLOGIE PEDIATRIQUE :</p> <p>INDICATION : Sd inflammatoire bilogique avec fièvre, douleurs abdominales et vomissements chez un enfant présentant une epidermolyse bulleuse [...]</p>
<p>10 ans</p> <p></p>	<p>CRH HdJ du 15/10/15 par S. LAMBERT le 15/10/15 :</p> <p>Madame [PRENOM_PATIENT], nee le [DATE_NAIS_PATIENT], a ete hospitalisee en hopital de jour / seance dans le service de 15/10/15 au 15/10/15 pour cure de Venofar dans le contexte de epidermolyse bulleuse dystrophique recessive avec absence de collagene 7 [...]. Epidermolyse bulleuse dystrophique recessive avec absence totale de collagene 7 [...]</p> <p>Ostéodensitométrie du corps entier- Enfant du 15/10/15 (PACS) par 15/10/15 le 15/10/15 :</p> <p>Premier Examen dans le cadre d'une epidermolyse bulleuse [...]</p> <p>630 CRH Dermatologie du 15/10/15 par S. LAMBERT le 15/10/15 - DERMATOLOGIE :</p> <p>Motif d'hospitalisation : reevaluation d'une epidermolyse bulleuse dystrophique recessive avec absence totale de collagene VII chez une enfant de 15 ans [...]. Epidermolyse bulleuse dystrophique recessive confirmée sur le plan moleculaire (mutation homozygote PARG613X sur l'exon 14 du collagene 7) [...]. Epidermolyse bulleuse dystrophique recessive chez une jeune fille de 15 ans [...]</p> <p>15/10/15 du 15/10/15 par S. LAMBERT le 15/10/15 :</p> <p>Code principal : Q812 epidermolyse bulleuse dystrophique ; Codes significatifs : D509 anemie par carence en fer, sans precision ; R53+2 Fatigue [asthenie] ; K9380 Atteintes de la cavite buccale au cours de</p>
<p>17 ans</p> <p></p>	<p>+ afficher les autres documents</p> <p>302 COURRIER du 15/10/15 par S. LAMBERT le 15/10/15 - METABOLISME-NEUROLOGIE PEDIATRIQUE :</p> <p>J'ai recupere le bilan endocrinien effectue a [PRENOM_PATIENT] que tu suis pour une epidermolyse bulleuse dystrophique severe depuis sa naissance [...]</p> <p>CRH de pédiatrie du 15/10/15 par S. LAMBERT le 15/10/15 :</p> <p>Dilatation oesophagienne chez un patient atteint d'une epidermolyse bulleuse dystrophique recessive [PRENOM_PATIENT] [...]. Epidermolyse bulleuse dystrophique recessive [...]. [PRENOM_PATIENT] de 17 ans [PRENOM_PATIENT]</p> <p>DIAG PMSI du 15/10/15 par S. LAMBERT le 15/10/15 :</p> <p>Code principal : Q812 epidermolyse bulleuse dystrophique ; Codes significatifs : D509 anemie par carence en fer, sans precision ; R53+2 Fatigue [asthenie] ; K9380 Atteintes de la cavite buccale au cours de</p>

Rechercher des patients

Sur tout l'entrepôt

epidermolyse bulleuse

341/341

Etendre aux synonymes : [+ Avancé](#) - [Réécrire la requête](#)[+ Ajouter un filtre Full text](#)[+ Ajouter un filtre structuré](#)[+ Filtre patient](#)

LANCER LA RECHERCHE

?

Requêtes sauées

gastroentérite aigüe et colite hémorragique

NETT et NECPE

dysplasie

hépatite

test hépatique

hépatite NEUROLOGIE PED

Historique requête

Show 10 entries

Search:

▼ date

Requêtes

- 29/10/2015 22:20
Document contenant 'Epidermolyse bulleuse' et 'anémie par carence en fer, sans précision', au cours de...
- 15/10/2015 18:56
Document contenant 'Epidermolyse bulleuse' et 'anémie par carence en fer, sans précision', au cours de...
- 26/08/2015 14:52
Document contenant 'Epidermolyse bulleuse' et 'anémie par carence en fer, sans précision', au cours de...

Résultat

Cohortes

Données stat

Concepts

PMSI

Biologie

Genes

Map

Clustering

Datamart

341 Patients
2396 Documents

Documents contenant 'epidermolyse bulleuse', en excluant les négations

SAUVER LA REQUÊTE

Refaire une recherche sur le résultat : Exporter les patients : Filtrer le résultat ci-dessous : ok annuler

11 ans	Echographie abdominale du ANASSIMED HASSANALY Age 28/11/2005 F 10 ans - RADIOLOGIE PEDIATRIQUE : INDICATION : Sd inflammatoire bilogique avec fièvre, douleurs abdominales et vomissements chez un enfant présentant une epidermolyse bulleuse [...]
10 ans	CRH HdJ du ANASSIMED HASSANALY Age 28/11/2005 F 10 ans Madame [PRENOM_PATIENT] cure de Venofer dans le co Ostéodensitométrie du co Premier Examen dans le ca 630 CRH Dermatologie du Motif d'hospitalisation : ree confirmée sur le plan molec
17 ans	1/3 CRH HDJ NCK DERMATOLOGIE PED, Motif de l'hospitalisation MOTIF D'HOSPITALISATION PARIS CEDEX 15, References : CPH/CPH Compte-rendu d'hospitalisation de jour du jour / seance dans le service de NCK HJ3 DERMATOLOGIE - HJ CLINIQUE ROBERT DEBRE RDC le pour cure de Venofer dans le contexte d' epidermolyse bulleuse dystrophique récessive avec absence de collagene 7.. HISTOIRE DE LA MALADIE TRAITEMENT a L'ENTREe Antecedents medicaux : Antecedents chirurgicaux : - Epidermolyse bulleuse dystrophique recessive avec absence totale de collagene 7.

Search for patients

Throughout the warehouse

 1041/1041

Extending to synonyms :

[+Advanced - Rewrite the query](#)

+ Add a filter Full text

+ Add a structured filter

+ Time constraints

+ Patient filter

START A RESEARCH

?

Saved queries

Search:

title

No data available in table

History query

Search:

queries

28/07/2017
09:07Filtered 1 : Documents containing 'lupus',
Excluding negations

Result

Cohort : Test ⁰x

Stats data

Concept

Biology

Map

clustering

D

1041 patients
5972 Documents

Filtered 1 : Documents containing 'lupus', Excluding negations

SAVE QUERY

SEARCH ON RESULT

EXPORT PATIENTS TO EXCEL

FIT THE RESULT BELOW

SHARE YOUR QUERY

FEEDING A COHORT

+ Create a cohort

Or Select an existing cohort to feed it to patients :

Open Cohort : Test ^x
 include all of the following patients in the cohort

 Exclude all patients below from the cohort

 In doubt all patients below

 Import all the patients below into the cohort

 Do not display patients already included or excluded
Patient F,
42 years

Lettre médicale du [redacted] - NEPHROLOGIE ADULTE :

Pour ce qui est du **lupus**, les articulations font mal comme depuis toujours, pas de lesion cutanee en rapport [...]

280 HDJ NEPHROLOGIE ADULTE du [redacted], par [redacted] - NEPHROLOGIE ADULTE :

LUPUS ERYTHEMATEUX DISSEMINÉ [...] Lésions cutanées brûlées a l'azote liquide, decrites par la patiente comme « cicatrices de **lupus** » [...] 2/ sur le p
lupus, bilan immunologique en attente [...] [redacted] qui avait conclu a l'absence d'arguments pour un neuro-**lupus** [...][+ View other documents](#)To facilitate the selection of
patients for studies

Search for patients

Throughout the warehouse

 1041/1041

Extending to synonyms :

+Advanced - Rewrite the query

+ Add a filter Full text

+ Add a structured filter

+ Time constraints

+ Patient filter

START A RESEARCH

?

Saved queries

Search:

title

No data available in table

History query

Search:

queries

Dated	queries
28/07/2017 09:07	Filtered 1 : Documents containing 'lupus', Excluding negations
27/07/2017 16:31	Filtered 1 : Documents containing 'infection% and eczema and thrombopenie', Excluding negations
27/07/2017 16:28	Filtered 1 : Documents containing 'MECP2', Excluding negations
27/07/2017 15:06	Filtered 1 : Documents containing 'lupus and cicatrice', Excluding negations

Result

Cohort : Test ¹x

Stats data

Concept

Biology

Map

clustering

Datamart

1041 patients
5972 Documents

Filtered 1 : Documents containing 'lupus', Excluding negations

SAVE QUERY

SEARCH ON RESULT

EXPORT PATIENTS TO EXCEL

FIT THE RESULT BELOW

SHARE YOUR QUERY

FEEDING A COHORT

+ Create a cohort

Or Select an existing cohort to feed it to patients :

Choose a Cohort

Open Cohort : Test ^x

Include all of the following patients in the cohort

Exclude all patients below from the cohort

In doubt all patients below

Import all the patients below into the cohort

Do not display patients already included or excluded

Patient F, 42 years	Lettre médicale du [redacted] - NEPHROLOGIE ADULTE : Pour ce qui est du lupus , les articulations font mal comme depuis toujours, pas de lesion cutanee en rapport [...] 280 HDJ NEPHROLOGIE ADULTE du [redacted] par [redacted] - NEPHROLOGIE ADULTE : LUPUS ERYTHEMATEUX DISSEMINÉ [...] Lésions cutanees brulees a l'azote liquide, decrites par la patiente comme « cicatrices de lupus » [...] 2/ sur le plan du lupus : [...] Absence d'argum lupus , bilan immunologique en attente [...] [redacted] qui avait conclu a l'absence d'arguments pour un neuro- lupus [...]
Patient F, 14 years	090 CION RHUMATOLOGIE du [redacted] par [redacted] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : L'histoire familiale revele un LUPUS avec atteinte renale chez la maman [...] Lupus aenticoagulant et anticardiolipine [...]
Patient F, 37 years	373 CRH HEMATOLOGIE ADULTE du [redacted] par [redacted] - HEMATOLOGIE CLINIQUE ADULTE : - Lupus cutane et hematologique [redacted], puis glomerulopathie lupique stade V [...] [...] HEMATOLOGIE ADULTE SCID du 16/08/14 (XPOS) par SUAREZ Felipe - HEMATOLOGIE CLINIQUE ADULTE :

To facilitate the selection of patients for studies

Search engine optimization : NLP

Reducing noise

3 queries on Necker's data warehouse:

"crohn and diabetes"

"lupus and diarrhea"

"nphp1"

Of the 500,000 patients, the search engine returns a total of 350 patients:

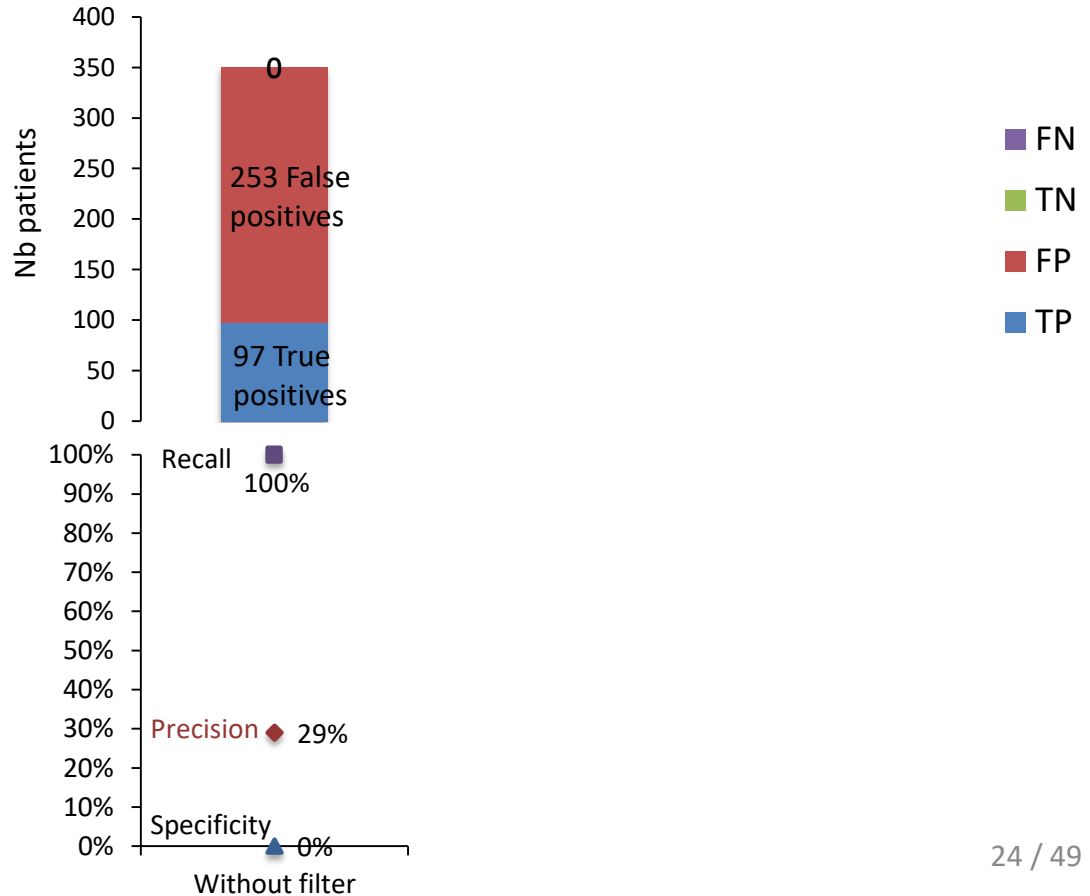
Recall: **1**

Precision: **0.29**

False positive patients:

« We can exclude the diagnosis of lupus »

« The father has a crohn's disease »



Detection of the expression of negation and family history

3 queries on Necker's data warehouse:

"crohn and diabetes"

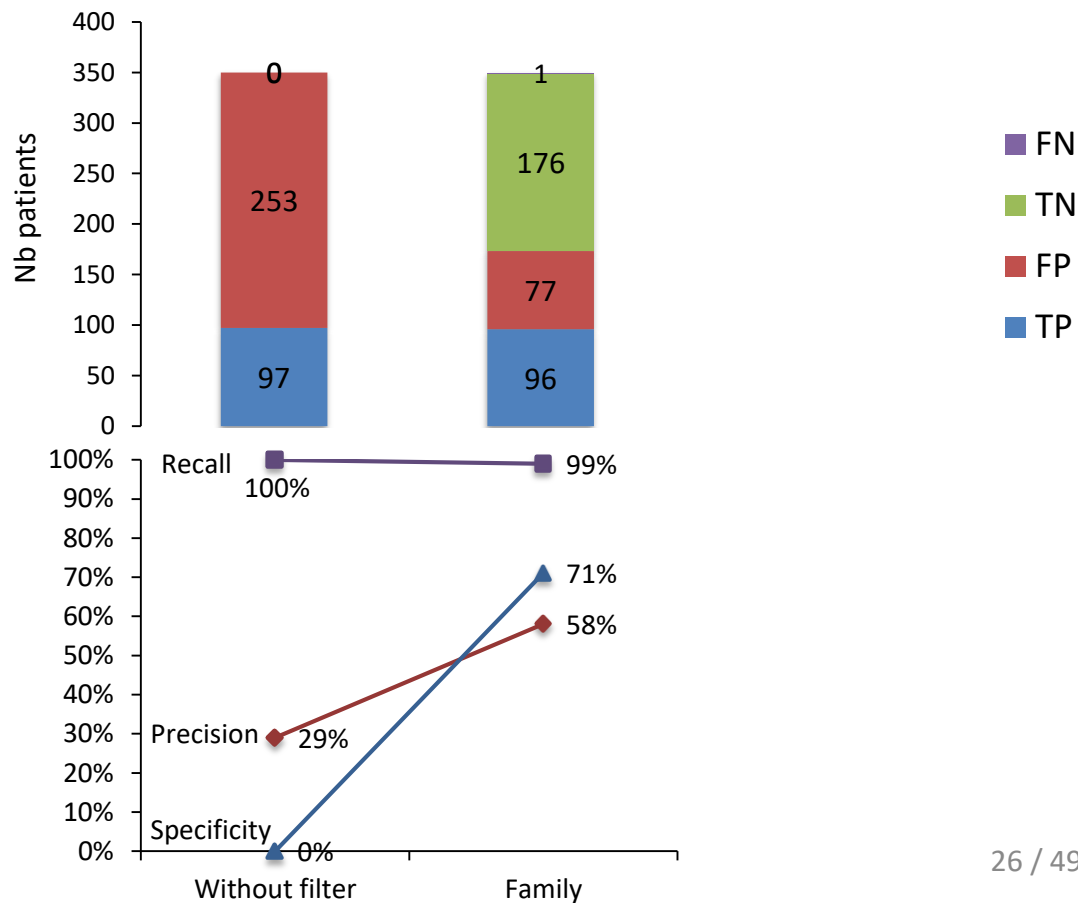
"lupus and diarrhea"

"nphp1"

Of the 500,000 patients, the search engine returns a total of 185 patients:

Recall: **0.99**

Precision: **0.58**



Detection of the expression of negation and family history

3 queries on Necker's data warehouse:

"crohn and diabetes"

"lupus and diarrhea"

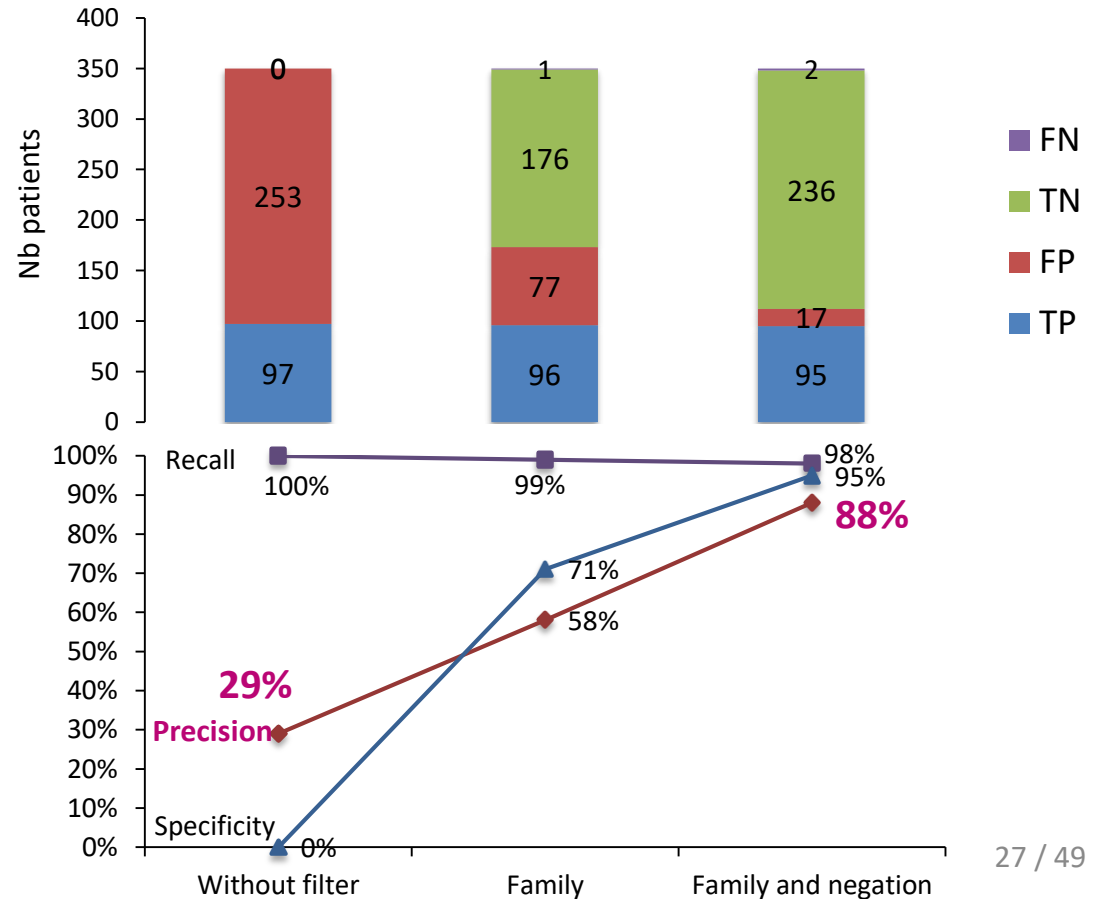
"nphp1"

Of the 500,000 patients, the search engine returns a total of 112 patients:

Recall: **0.98**

Precision: **0.88**

N. Garcelon, A. Neuraz, V. Benoit, R. Salomon, A. Burgun, Improving a full-text search engine: the importance of negation detection and family history context to identify cases in a biomedical data warehouse, J Am Med Inform Assoc. (2016).



Reducing silence

If I'm looking for "SCID," I expect to find the reports that contain "Omenn's syndrome."

Narrative record

HÔTEL - DIEU

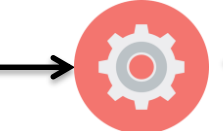
CONCLUSION: Patient masculin âgé de 65 ans souffrant d'insuffisance rénale chronique de stade 4. Il existe une diabète sucré de type 2. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.

REPERMISE: Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.

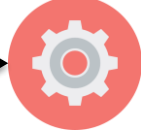
CLINIQUE: Malade à l'admission, dans un contexte d'aggravation de son diabète et de son insuffisance rénale.

Text	Certainty	Context	Enriched Text
<p>CONCLUSION: Patient masculin âgé de 65 ans souffrant d'insuffisance rénale chronique de stade 4. Il existe une diabète sucré de type 2. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>REPERMISE: Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>CLINIQUE: Malade à l'admission, dans un contexte d'aggravation de son diabète et de son insuffisance rénale.</p>	Affirmation	Patient	<p>CONCLUSION: Patient masculin âgé de 65 ans souffrant d'insuffisance rénale chronique de stade 4. Il existe une diabète sucré de type 2. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>REPERMISE: Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>CLINIQUE: Malade à l'admission, dans un contexte d'aggravation de son diabète et de son insuffisance rénale.</p>
<p>CONCLUSION: Patient masculin âgé de 65 ans souffrant d'insuffisance rénale chronique de stade 4. Il existe une diabète sucré de type 2. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>REPERMISE: Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>CLINIQUE: Malade à l'admission, dans un contexte d'aggravation de son diabète et de son insuffisance rénale.</p>	Negation	Patient	<p>CONCLUSION: Patient masculin âgé de 65 ans souffrant d'insuffisance rénale chronique de stade 4. Il existe une diabète sucré de type 2. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>REPERMISE: Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>CLINIQUE: Malade à l'admission, dans un contexte d'aggravation de son diabète et de son insuffisance rénale.</p>
<p>CONCLUSION: Patient masculin âgé de 65 ans souffrant d'insuffisance rénale chronique de stade 4. Il existe une diabète sucré de type 2. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>REPERMISE: Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>CLINIQUE: Malade à l'admission, dans un contexte d'aggravation de son diabète et de son insuffisance rénale.</p>	Affirmation	Famille	<p>CONCLUSION: Patient masculin âgé de 65 ans souffrant d'insuffisance rénale chronique de stade 4. Il existe une diabète sucré de type 2. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>REPERMISE: Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4. Les analyses de laboratoire sont compatibles avec une insuffisance rénale chronique de stade 4.</p> <p>CLINIQUE: Malade à l'admission, dans un contexte d'aggravation de son diabète et de son insuffisance rénale.</p>

For each UMLS term



Regular expression



Exact match

Integration of terms, synonyms and ascendants UMLS-based Hierarchy

Concept	Certainty	Context
DNID	1	Patient
SLE	-1	Patient
End-stage renal failure	1	Family

Integration in Dr Warehouse

The screenshot displays the Dr Warehouse interface. At the top, the header includes the logo 'Dr Warehouse' and the tagline 'Entrepôt de données'. Below the header is a navigation bar with tabs for 'Accueil', 'Moteur de recherche', 'Mes requêtes', 'Mes Cohortes', 'Outils', 'Patient Norm / DPP', 'Admin', 'Notifications', 'Nicolas Garcelon', and 'deconnexion'. The main content area is titled 'Rechercher des patients' and shows a search for 'SCID' with 146 patients and 883 documents. A filter is applied: 'Documents contenant 'scid' Etendu aux synonymes, en excluant les négations, dans les services : IMMUNO-HEMATOLOGIE PEDIATRIQUE'. Below the filter are buttons for 'ALIMENTER UNE COHORTE', 'SAUVER LA REQUÊTE', 'RECHERCHER SUR LE RÉSULTAT', 'EXPORTER LES PATIENTS SUR EXCEL', 'FILTRER LE RÉSULTAT CI DESSOUS', and 'PARTAGER VOTRE REQUÊTE'. The search results are listed in a table with columns for patient information and document details. Two results are highlighted with red boxes: one for 'syndrome d'Omenn' and another for 'ARTEMIS'. On the left side, there is a 'Historique requête' section showing a list of previous search filters and their dates.

Dr Warehouse © imagine
Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes Cohortes | Outils | Patient Norm / DPP | Admin | Notifications | Nicolas Garcelon | deconnexion

Rechercher des patients

Sur tout l'entrepôt

SCID 146/146

Etendre aux synonymes :
+Avancé - Réécrire la requête

+ Ajouter un filtre Full text
+ Ajouter un filtre structuré
+ Contraintes temporelles
+ Filtre patient

LANCER LA RECHERCHE ?

Historique requête

Date	Requêtes
17/11/2017 17:01	17/11/2017 Filtre 1 : Documents contenant 'syndrome de rett', en excluant les négations
17/11/2017 17:00	17/11/2017 Filtre 1 : Documents contenant 'lupus', en excluant les négations
07/11/2017 16:58	07/11/2017 Filtre 1 : Documents contenant 'rachis', datés du 01/01/2017 au 30/06/2017, en excluant les négations dans les services : CHIRURGIE ORTHOPEDIQUE PEDIATRIQUE.
26/10/2017 17:04	26/10/2017 Filtre 1 : Documents contenant 'myocardite', en excluant les négations Filtre 2 : Documents contenant 'greffe', en excluant les négations, dans les services : CARDIOLOGIE, CARDIOLOGIE PEDIATRIQUE, CHIRURGIE CARDIAQUE PEDIATRIQUE.
26/10/2017 16:57	26/10/2017 Filtre 1 : Documents contenant 'mutisme', datés du 01/09/2011, en excluant les négations Filtre 2 : Documents contenant 'perfusion asi or Perfusion par arterial spin labeling or perfusion par asi or near(perfusion, asi),2, false', en excluant les négations
26/10/2017 16:47	26/10/2017 Filtre 1 : Documents contenant 'syndrome de rett', en excluant les négations
26/10/2017 16:25	26/10/2017 Filtre 1 : Documents contenant 'mecp2', en excluant les négations
26/10/2017 16:02	26/10/2017 Filtre 1 : Documents contenant 'pneumopathie', en excluant les négations
25/10/2017 14:26	25/10/2017 Filtre 1 : Documents contenant 'syndrome hyper igm', en excluant les négations

Résultat | Données stat | Concept | Biologie | Gènes | Map | Clustering | Export

146 patients
883 documents

Filtre 1 : Documents contenant 'scid' Etendu aux synonymes, en excluant les négations, dans les services : IMMUNO-HEMATOLOGIE PEDIATRIQUE,

ALIMENTER UNE COHORTE | SAUVER LA REQUÊTE | RECHERCHER SUR LE RÉSULTAT | EXPORTER LES PATIENTS SUR EXCEL | FILTRER LE RÉSULTAT CI DESSOUS | PARTAGER VOTRE REQUÊTE

Patient	Document
Patient F, décédé à 6 mois	091 CRH SC IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>SCID</u>
Patient M, 34 ans	090 CION IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>SCID</u>
Patient F, 19 ans	090 CION IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>SCID</u>
Patient M, 6 ans	131 CRH REA MEDICO CHIR PEDIATRIQUE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>SCID</u>
Patient F, 3 ans	091 CRH SC IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>syndrome d'Omenn</u>
Patient M, 7 ans	090 CION IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>ARTEMIS</u>
Patient M, 15 ans	090 CION IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>SCID</u>
Patient F, 6 ans	131 CRH REA MEDICO CHIR PEDIATRIQUE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>SCID</u>
Patient F,	090 CION IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE : <u>SCID</u>

29 / 49

Integration in Dr Warehouse

Dr Warehouse © imagine
Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes Cohortes | Outils | Patient Norm / DPP | Admin | Notifications | Nicolas Garcelon | deconnexion

Rechercher des patients

Sur tout l'entrepôt

SCID

146/146

Etendre aux synonymes : ✓

Avancé - Réécrire la requête

+ Ajouter un filtre Full text

+ Ajouter un filtre structuré

+ Contraintes temporelles

Résultat

Données stat

Concept

Biologie

Gènes

Map

Clustering

Export

146 patients
883 documents

Filtre 1 : Documents contenant 'scid' Etendu aux synonymes , en excluant les négations , dans les services : IMMUNO-HEMATOLOGIE PEDIATRIQUE,

ALIMENTER UNE COHORTE

SAUVER LA REQUÊTE

RECHERCHER SUR LE RÉSULTAT

EXPORTER LES PATIENTS SUR EXCEL

FILTRE LE RÉSULTAT CI DESSOUS

PARTAGER VOTRE REQUÊTE

Patient F, 091 CRH SC IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :
décédé à 6 mois

131 CRH REA MEDICO CHIR PEDIATRIQUE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :

SCID

SCID

SCID

091 CRH SC IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :

syndrome d'Omenn

+ afficher les autres documents

090 CION IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :

ARTEMIS

+ afficher les autres documents

- 26/10/2017 16:47 négations Filtre 1 : Documents contenant 'syndrome de rett' , en excluant les négations
- 26/10/2017 16:25 Filtre 1 : Documents contenant 'mecp2' , en excluant les négations
- 26/10/2017 16:02 Filtre 1 : Documents contenant 'pneumopathie' , en excluant les négations
- 25/10/2017 14:26 Filtre 1 : Documents contenant 'syndrome hyper igm' , en excluant les négations

Patient F, 6 ans

Patient F,

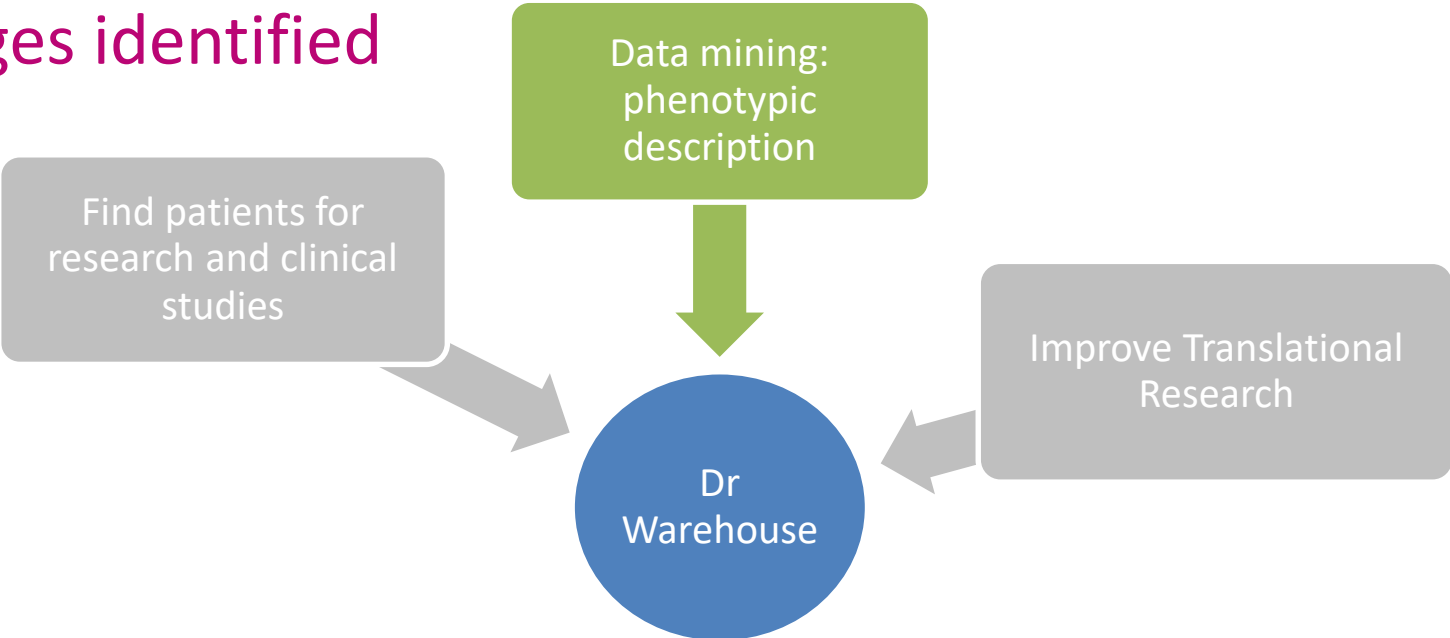
+ afficher les autres documents

131 CRH REA MEDICO CHIR PEDIATRIQUE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :

090 CION IMMUNO HEMATOLOGIE du [DATE] (SUSIE) par [AUTHOR] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :

Dr Warehouse

3 challenges identified



Aggregate data visualization tools

Dr WareHouse @Imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Outils | Contact | Patient Nom / IPP | Admin

Nicolas Garcelon Déconnexion

Rechercher des patients

Sur tout l'entrepôt

Résultat

Cohortes

Données stat

Concepts

PMSI

Biologie

Genes

Map

Clustering

Datamart

epidermolyse bulleuse 341/341

Etendre aux synonymes :

+ Avancé - Réécrire la requête

+ Ajouter un filtre Full text

+ Ajouter un filtre structuré

+ Filtre patient

LANCER LA RECHERCHE

?

Requêtes sauveées

pericardite and calcification

NETT et MCEP

apoplexie

hémorragie

test mercap nit

hémorragie NEUROLOGIE MD

Historique requête

Show 10 entries

Search:

date

Requêtes

29/10/2015 22:20

15/10/2015

Sur tout l'hôpital :

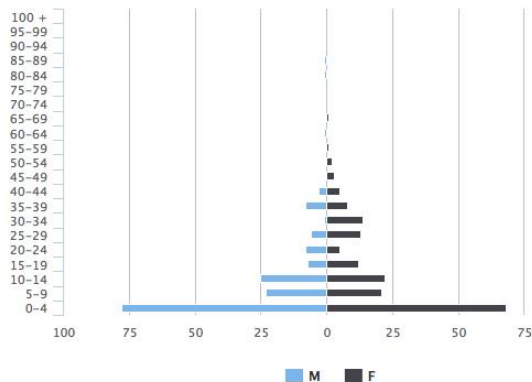
341 Patients

2396 Documents

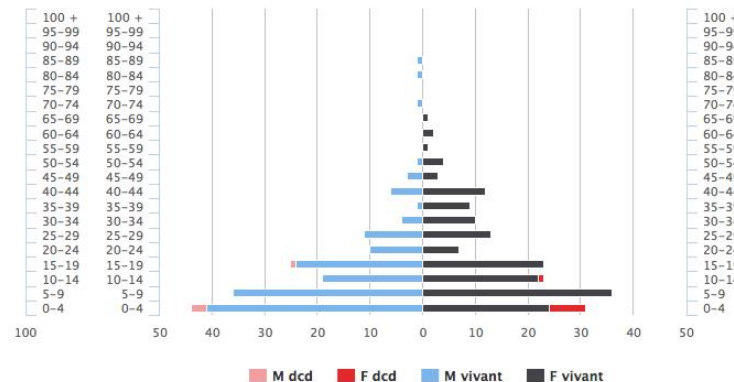
[Afficher le parcours moyen \(hospitalisations\) des patients](#)

[Afficher le parcours moyen complet \(hospitalisations et consultations\) des patients](#)

Pyramide des âges au 1er document trouvé



Pyramide des âges aujourd'hui



Pyramide des âges < 20 ans au 1er document trouvé

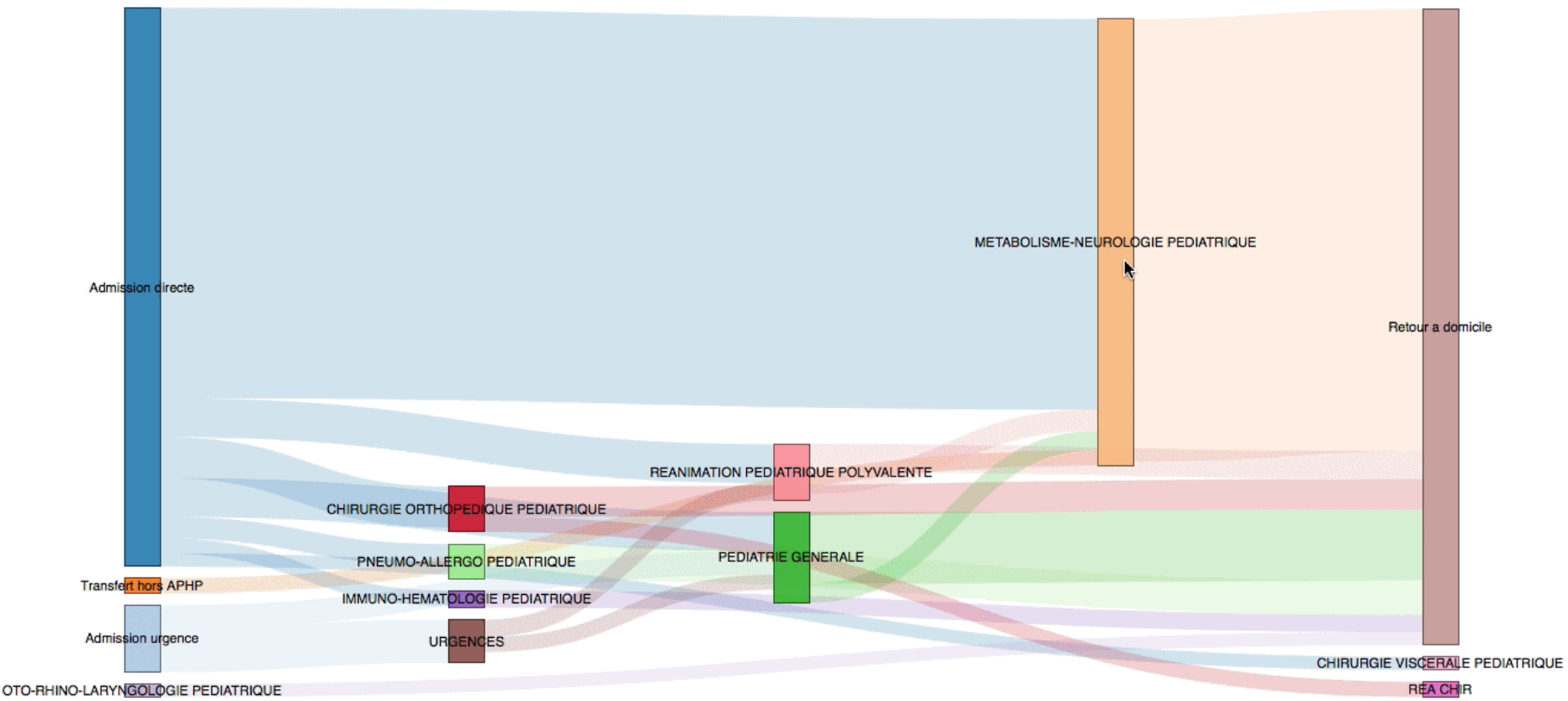


Pyramide des âges < 20 ans aujourd'hui



37 ans 49

Aggregate data visualization tools: Pathcare



Aggregate data visualization tools

Dr. Warehouse ©Imagine
Entrepôt de données

Home | Search engine | My queries | My Cohorts | Tools | Patient | Notifications | Nicolas Garcelon | Log out

Search for patients

Across the entire data warehouse

1041/1041

Extend to synonyms :

+Advanced - Rewrite the query

- + Add a full text filter
- + Add a structured filter
- + Time constraints
- + Patient filter

START A SEARCH ?

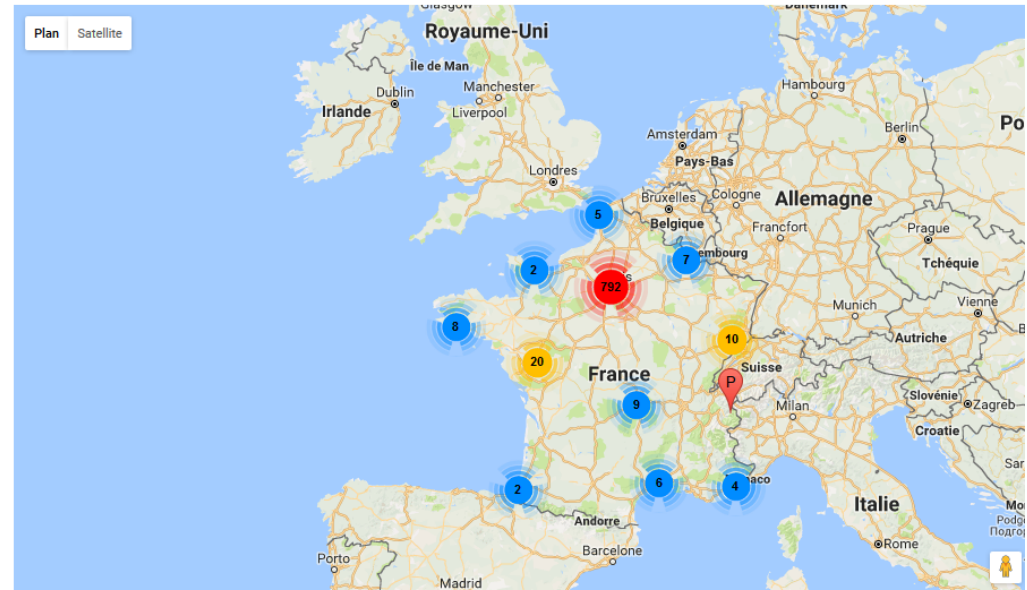
Query history

Search:

Date	Queries
29/08/2017 10:09	Filtered 1 : Documents containing 'lupus', Excluding negations
18/10/2016 13:37	Filtered 1 : Documents containing 'LIPA', Excluding negations
18/10/2016 13:35	Filtered 1 : Documents containing 'mecip2', Excluding negations
18/10/2016 13:31	Filtered 1 : Documents containing 'lupus', Excluding negations
12/10/2016 15:30	Filtered 1 : Documents containing 'syndrome de lowe', Excluding negations

- Result | Stats data | Concept | Biology | Genes | Map | Clustering

Across the entire data warehouse :
1041 patients
5972 Documents



High throughput phenotyping

Objectives:

Automated phenotypic description of a group of patients.

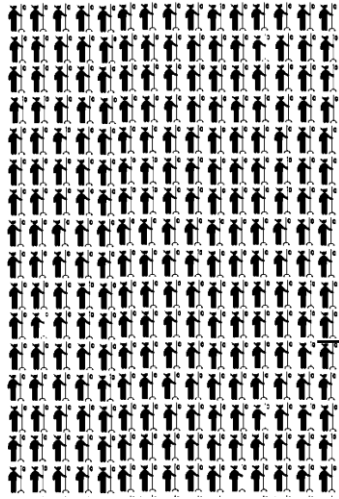
Enrichment of knowledge bases on rare diseases.

Our approach:

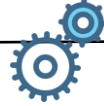
Reuse hospital reports, in French, without manual annotation phase, to carry out a high throughput phenotyping of rare diseases.

High throughput phenotyping

Data warehouse
500,000 patients



Query
"Rett syndrome"



N Patients found



Aggregated
concepts and
ranking

Phenotypes associated to Rett syndrome:

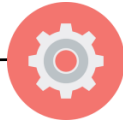
UMLS Concepts	Freq	TF-IDF
Stéréotypie	71.8	8.06
Crises	55	8.37
Scoliose	51.2	6.71
Epilepsie	50.7	2.96
Syndrome pyramidal	42.6	2.87
Ostéoporose	42.6	2.79
...		
1022 concepts		

For each patient
of the data
warehouse

Narrative records



UMLS Metathesaurus



UMLS Concepts	Context	Certainty	Frequency
Lupus	Patient	1	3
Anémie	Patient	-1	2
Insuffisance rénale	Family	1	1
Fièvre	Patient	1	5
Asthénie	Patient	1	4

RESEARCH

Open Access



Next generation phenotyping using narrative reports in a rare disease clinical data warehouse

Nicolas Garcelon^{1,2,13*} , Antoine Neuraz^{2,3}, Rémi Salomon^{1,4}, Nadia Bahi-Buisson^{1,5}, Jeanne Amiel^{1,6,7}, Capucine Picard^{1,8,9}, Nizar Mahlaoui^{1,8,10,11}, Vincent Benoit¹, Anita Burgun^{2,3,12} and Bastien Rance^{2,12}

Abstract

Background: Secondary use of data collected in Electronic Health Records opens perspectives for increasing our knowledge of rare diseases. The clinical data warehouse (named Dr. Warehouse) at the Necker-Enfants Malades Children's Hospital contains data collected during normal care for thousands of patients. Dr. Warehouse is oriented toward the exploration of clinical narratives. In this study, we present our method to find phenotypes associated with diseases of interest.

Methods: We leveraged the frequency and TF-IDF to explore the association between clinical phenotypes and rare diseases. We applied our method in six use cases: phenotypes associated with the Rett, Lowe, Silver Russell, Bardet-Biedl syndromes, DOCK8 deficiency and Activated PI3-kinase Delta Syndrome (APDS). We asked domain experts to evaluate the relevance of the top-50 (for frequency and TF-IDF) phenotypes identified by Dr. Warehouse and computed the average precision and mean average precision.

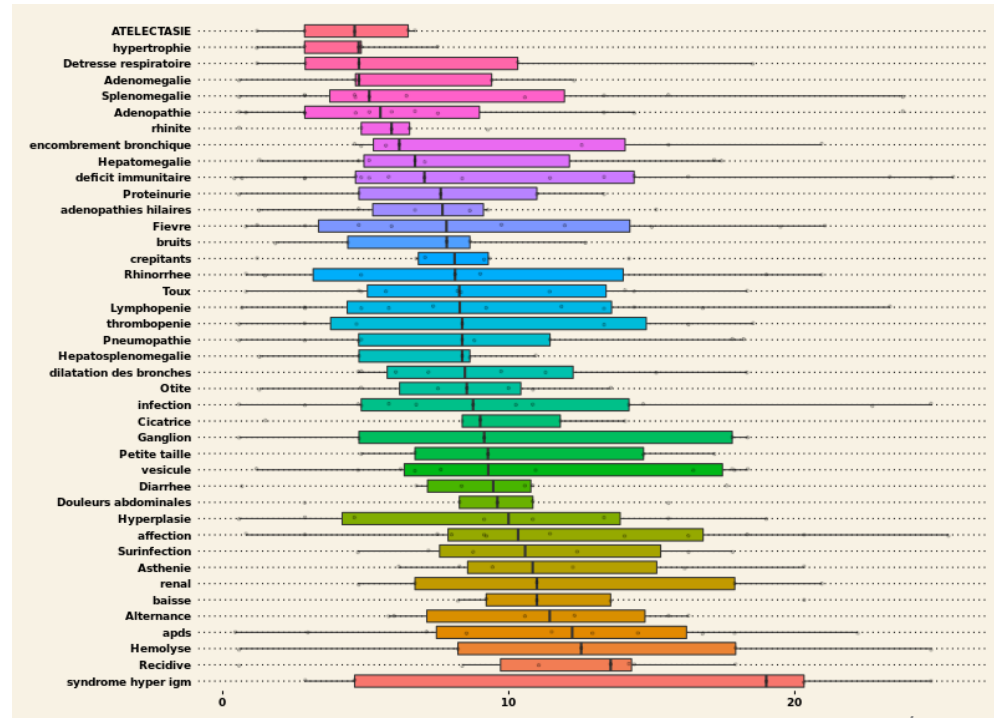
Results: Experts concluded that between 16 and 39 phenotypes could be considered as relevant in the top-50 phenotypes ranked by descending frequency discovered by Dr. Warehouse (resp. between 11 and 41 for TF-IDF). Average precision ranges from 0.55 to 0.91 for frequency and 0.52 to 0.95 for TF-IDF. Mean average precision was 0.79. Our study suggests that phenotypes identified in clinical narratives stored in Electronic Health Record can provide rare disease specialists with candidate phenotypes that can be used in addition to the literature.

Conclusions: Clinical Data Warehouses can be used to perform Next Generation Phenotyping, especially in the context of rare diseases. We have developed a method to detect phenotypes associated with a group of patients using medical concepts extracted from free-text clinical narratives.

Keywords: Data warehouse, Next generation phenotyping, Data mining, Rare diseases, Natural language processing

High throughput phenotyping

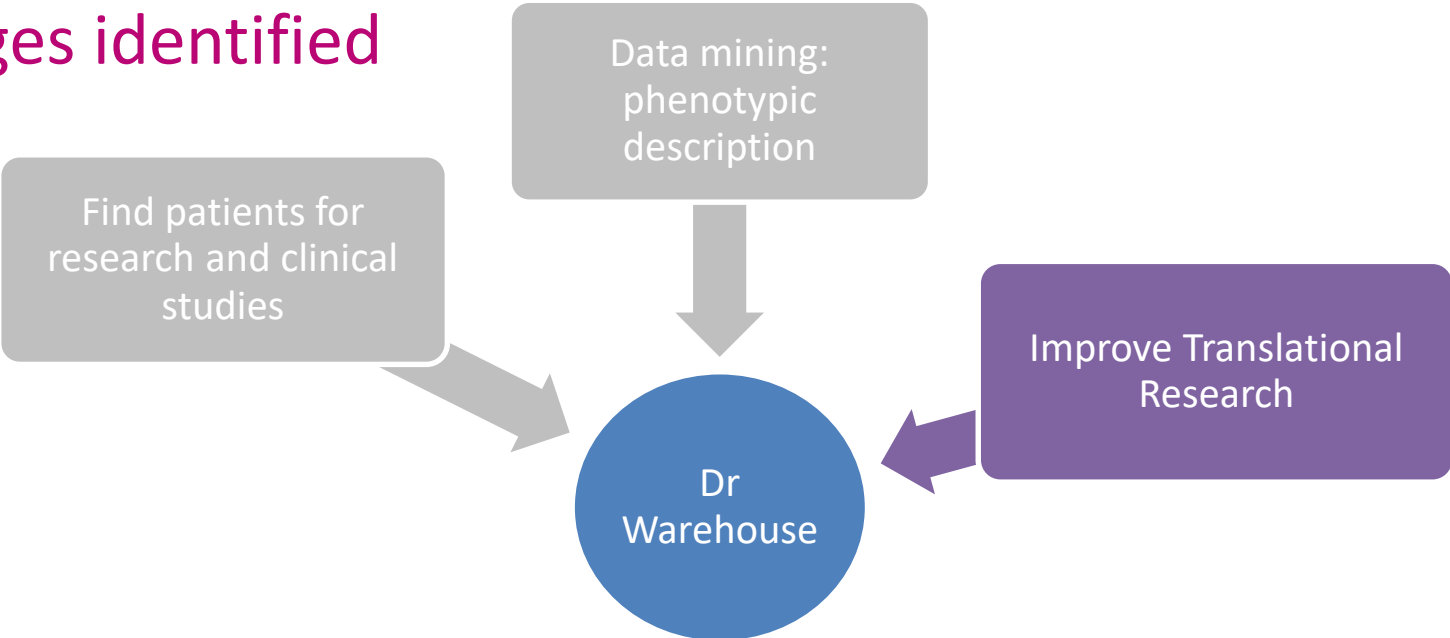
Example Activated PI3K-delta syndrome (APDS) :



To take into account the age of the first apparition of signs

Dr Warehouse

3 challenges identified



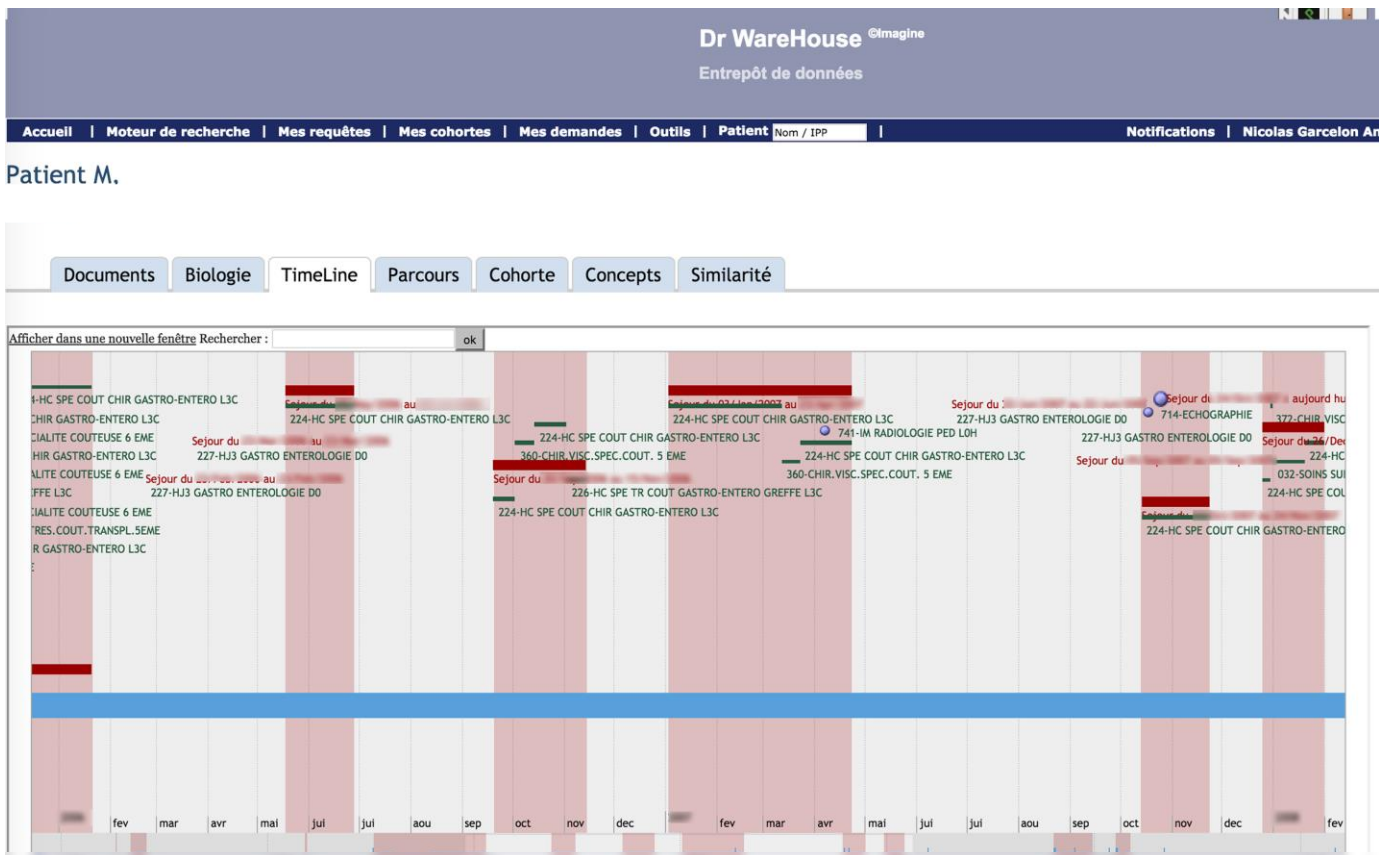
Patient-centered features

Facilitate the exploration of the health records: a search engine for one patient

The screenshot displays the Dr Warehouse web application interface. At the top, the header includes the logo 'Dr Warehouse @Imagine' and the tagline 'Entrepôt de données'. A navigation bar contains links for 'Accueil', 'Moteur de recherche', 'Mes requêtes', 'Mes cohortes', 'Mes demandes', 'Outils', 'Patient', 'Notifications', and 'Nicolas Garcelon An'. Below this, the user is identified as 'Patient M,'. A secondary navigation bar offers tabs for 'Documents', 'Biologie', 'TimeLine', 'Parcours', 'Cohorte', 'Concepts', and 'Similarité'. A search bar contains the text 'biopsie' and a 'RECHERCHER' button. The results section shows '38 documents trouvés' and a list of search results. The first result is a document titled 'Biopsie cutanée avec colo spé pour diag d'affection non carcinologique, pa...', with a snippet mentioning 'ORBIS CR - courrier' and 'Histologie de la biopsie realisee le ... : lésions de [...]'. To the right of the search results, a detailed view of a document is shown, including the title 'Biopsie cutanée avec colo spé pour diag d'affection non carcinologique, pa...', the service name 'SERVICE D'ANATOMIE ET DE CYTOLOGIE PATHOLOGIQUES', the address 'Groupe Hospitalier Necker-Enfants Malades, 149 rue de Sevres 75743 PARIS CEDEX 15', and contact information for the 'Chef de Service' and 'Secretariat'.

Patient-centered features

Visualization of the patient's path over time



Patient-centered features

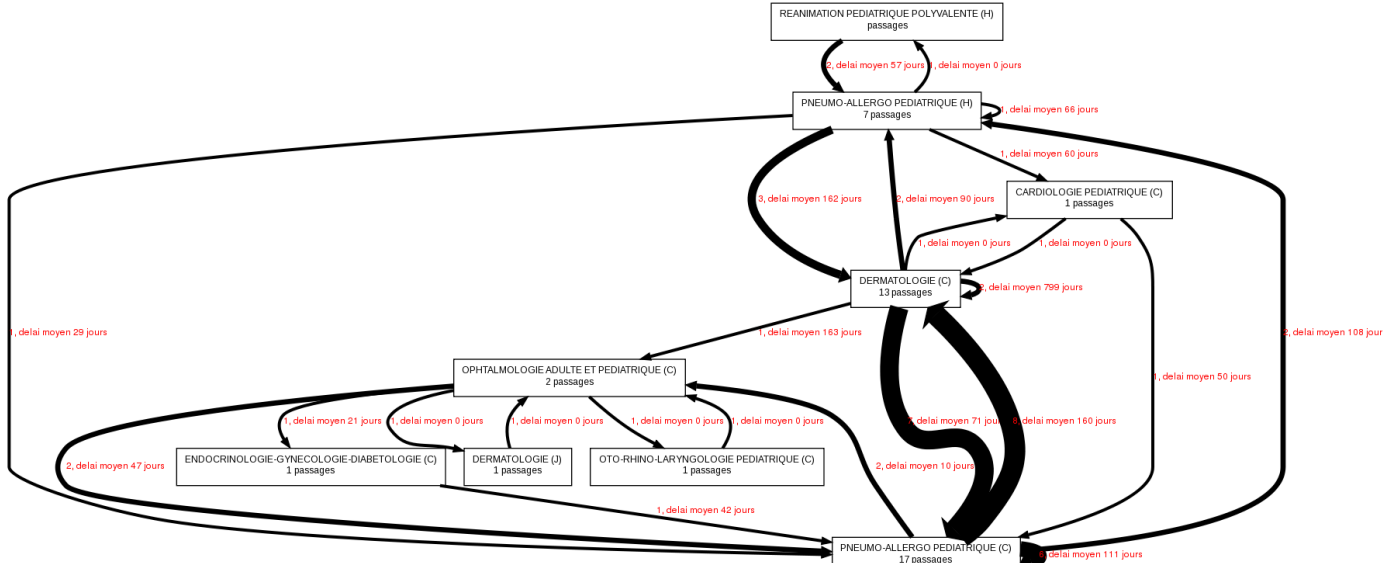
Visualization of the patient's path over space

Dr Warehouse ^{©Imagine}
Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes Cohortes | Outils | Patient | Admin |

Patient M,
18 ans

documents | Biologie | TimeLine | Parcours | Cohorte | Concepts | Similarité



Patient-centered features

Dr Warehouse ^{©Imagine}
Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes Cohortes | Outils | Patient Nom / IPP | Admin | Notifications | Nicolas Garcelon

Patient M,
18 ans

- documents
- Biologie
- TimeLine
- Parcours
- Cohorte
- Concepts
- Similarité

Concepts

Affichage simplifié - Affichage pour validation

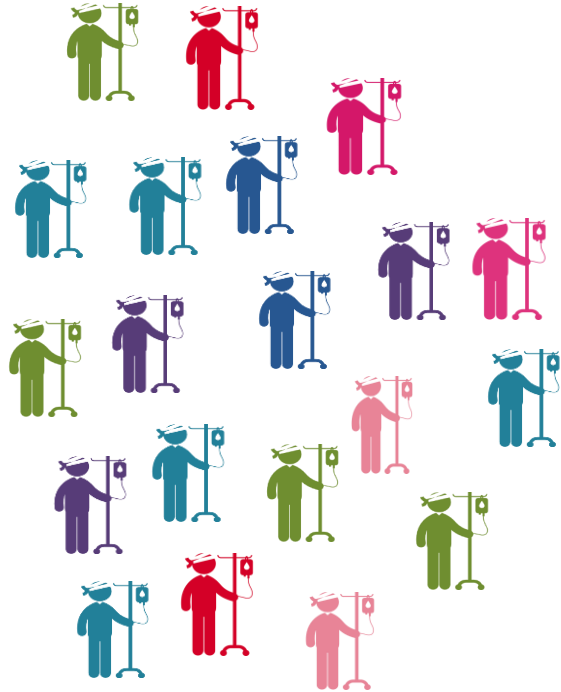
Search:

Concepts	Nb documents	TF-IDF doc	TF-IDF patient
Asthme	27	2.26	0.33
Hamartome	21	5.13	0.83
Surpoids	9	2.41	0.15
Rhinite allergique	9	2.44	0.26
Obstruction	8	2.24	0.11
hyperkeratose	7	1.08	0.22
engorgement bronchique	7	0.85	0.1
hypertrophie	6	1.06	0.08
Obstruction nasale	5	0.75	0.07
Wheezing	5	0.77	0.08
dyspnee	5	0.77	0.05
rhinite	5	1.08	0.08
Bronchopathie	4	2.15	0.08
Kyste	4	2.92	0.07
Crise d'asthme	4	1.91	0.05
Dermatophytie	3	0.95	0.07
cgh	3	2.85	0.07
Naevus epidermique	3	1.19	0.09
troubles du developpement	3	0.93	0.09
complications	3	0.39	0.03
Keratodermie	3	0.39	0.07
Gene respiratoire	3	2.21	0.03
Hematome	2	3.56	0.02
discontinue	2	0.55	0.03

Automated extraction of concepts from the narrative records

Translational research and similarity

Undiagnosed patients



A diagnosed patient
Complex phenotype

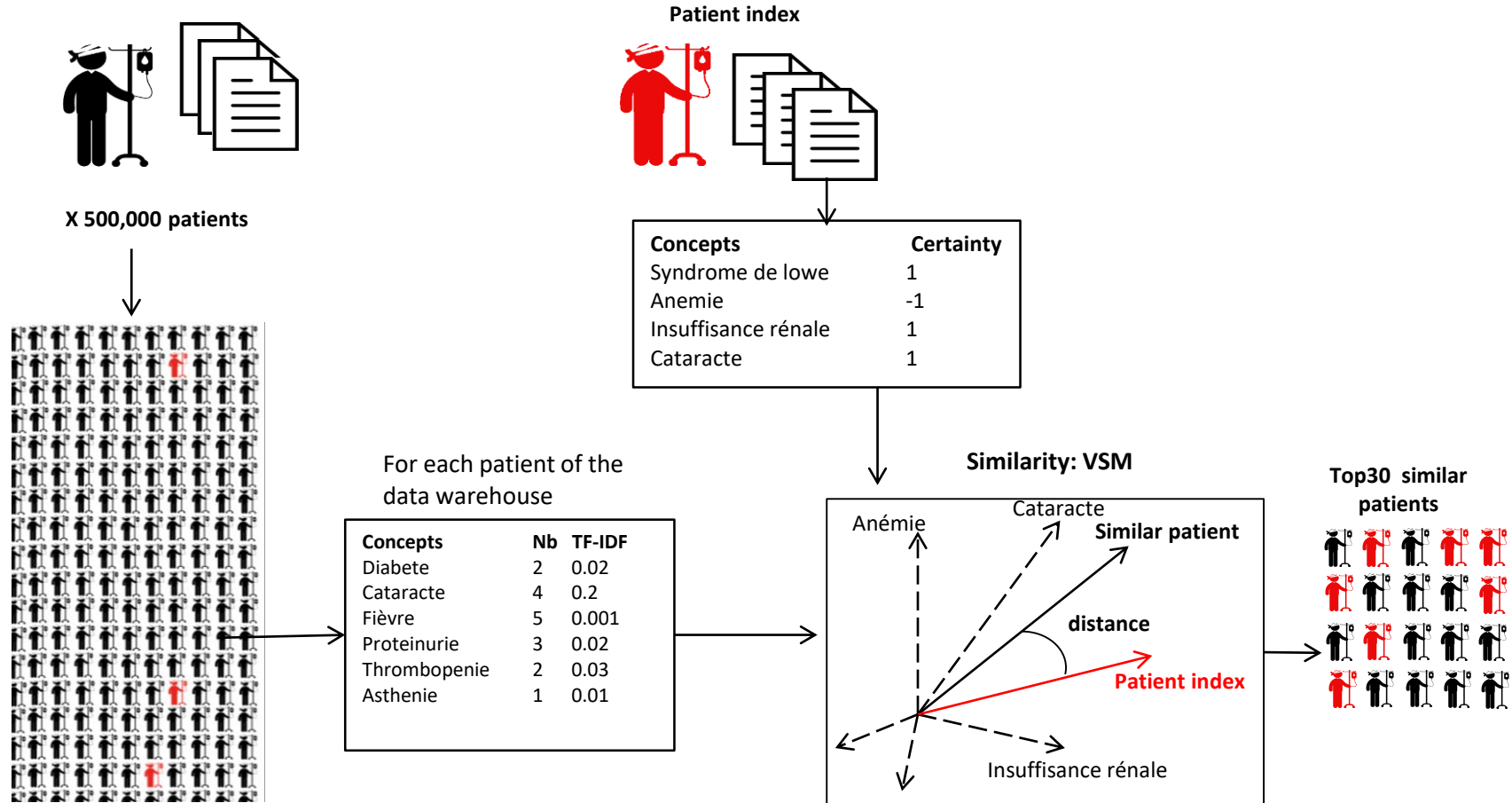
Search engine
All phenotypic criteria



Similarity with
index patient
Phenotypic proximity



Our approach to find similar patients



Translational research and similarity

Journal of Biomedical Informatics 73 (2017) 51–61



Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin



Finding patients using similarity measures in a rare diseases-oriented clinical data warehouse: Dr. Warehouse and the needle in the needle stack



Nicolas Garcelon^{a,b,c,*}, Antoine Neuraz^{c,d}, Vincent Benoit^{a,b}, Rémi Salomon^{a,b,e}, Sven Kracker^{a,b,f}, Felipe Suarez^{a,b,g}, Nadia Bahi-Buisson^{a,b,h}, Smail Hadj-Rabia^{a,b,i}, Alain Fischer^{a,b,j,k,l}, Arnold Munnich^{a,b,m,n}, Anita Burgun^{c,d,o}

^a Institut Imagine, Paris Descartes Université Paris Descartes-Sorbonne Paris Cité, Paris, France

^b INSERM, Institut Imagine, UMR 1163, Université Paris Descartes, Sorbonne Paris Cité, Paris, France

^c INSERM, Centre de Recherche des Cordeliers, UMR 1138 Equipe 22, Université Paris Descartes, Sorbonne Paris Cité, Paris, France

^d Département d'informatique médicale, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

^e Service de Néphrologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

^f Laboratory of Human Lymphohematopoiesis, INSERM UMR 1163, Paris, France

^g Service de Hématologie, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

^h Service de neurologie pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

ⁱ Service de Dermatologie, Centre de Références maladies Génétiques à Expression Cutanée (MAGEC), Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

^j Centre de Référence Déficiences Immunitaires Héritaires, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

^k Unité d'Immunologie-Hématologie et Rhumatologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

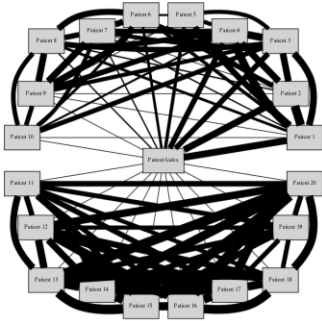
^l Collège de France, Paris, France

^m Département de génétique médicale, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

ⁿ Centre de Référence des Maladies Osseuses Constitutionnelles, INSERM UMR 1163, Laboratoire de bases moléculaires et physiopathologiques de l'ostéochondrodysplasie, Paris

Nom Prénom né le 01/01/2000

Documents | Biologie | TimeLine | Parcours | Cohorte | Concepts | Similarité

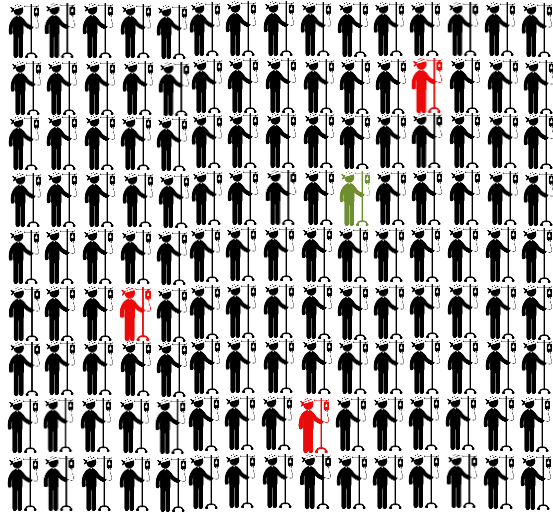


Integration of the similarity algorithm in Dr Warehouse

Patient	Similarité	Concepts communs	Concepts patient
Patient 1 (statut cohorte APDS : 1) (statut cohorte APDS2 : 1) (statut cohorte DIH inclus CEREDIH : 1) Afficher le comparateur	81	Conjonctivite, Toux, DYSPHAGIE, dyspnee, Fievre, Hyperplasie, deficit immunitaire, Lymphopenie, mycose, Otite, Douleur, Sinusite, syndrome d'apnee du sommeil, Surinfection, Sinusite chronique, larmolement, Rhinorrhoe purulente, syndrome hyper igm, Adenopathie, Intervention chirurgicale, Gene respiratoire, rhinosinusite, Cicatrice, Recidive	Vegetations, stenose pharyngee , Gaz, Epanchement, Retrait, hypertrophie des amygdales, suppuration, pi3k, Thrombus, deficit immunitaire humoral, Ptosis, Croute, septicemie, dacryocystite, hypertrophie, Observation, obstruction laryngee, Bouton, Respiration, Aphte, tor, Complication, Thrombose, Hernie inguinale bilaterale, obstruction chronique des voies aeriennes, nutrition, Stenose laryngee, Oedeme, Hypopnee, vesicule, obstruction nasale bilaterale, Deficit en IgG, Apnee, Obstruction, Lésions muqueuses, Nodule, SAOS, Circulation collaterale, Douleurs abdominales, adherence, hypertrophie des vegetations adenoides, apds, appareil respiratoire
Patient 2 (statut cohorte APDS : 1) (statut cohorte APDS2 : 1) (statut cohorte DIH inclus CEREDIH : 1) Afficher le comparateur	80	Anorexie, Asthenie, dilatation des bronches, rhinite, infection, Conjonctivite, Toux, Fievre, gastroenterite, cephalées, Hyperplasie, deficit immunitaire, Lymphopenie, Otite, Douleur, syndrome d'apnee du sommeil, Surinfection, baisse, Sinusite chronique, Adenopathie cervicale, Tolerance, Ecoulement, Rhinorrhoe purulente, syndrome hyper igm, Adenopathie, Pansinusite, Rhinorrhoe, Cicatrice, Recidive	Enuresie, bruits, Aphte, Hypopnee, 3 6, bronchite chronique, Ptosis, Toux nocturne, Hypogammaglobulinemie, somnolence diurne, deviation de la cloison nasale, Gastro-enterite aigue, ANEMIE MICROCYTAIRE, Diphterie, Apnee, 3 4, Anemie, Injection intraveineuse, baisse de l'acuite visuelle, Conjonctivite chronique , Ronchi, crepitations, deficit de l'immunte humorate, Conjonctivite purulente, pai, lymphoproliferation, erytheme localise, infections respiratoires, PIK3R1, cauchemars, ronflement, Obstruction, Enuresie nocturne, Gynecomastie, hypertrophie, Urgences, endormissement, Epistaxis, Rhinite chronique, pi3k.apds , Irritable, bronchorrhee, Agitation, Prise de poids, Toux chronique, Taux d'anticorps, cavite, vomissement, Sueurs nocturnes, SAOS, Frequence cardiaque maximale, sommeil agile, Respiration buccale, Toux persistante
Patient 3 (statut cohorte DIH inclus CEREDIH : 1) (statut cohorte APDS : 1)		douleur abdominale, Asthme, ATELECTASIE, dilatation des bronches, rhinite, infection, Toux, Diarhee,	vesicule, Decouverte fortuite, Toux chronique, Douleurs abdominales, Rhinite chronique, regression, bruits, PIK3CD, HPA3, cavite, PIK3, Vegetation, distension, Diarhee chronique, hypertrophie, Retrecissement, CD31, Purpura, apds , ANEMIE MICROCYTAIRE, consideration, eruption, fenetre aorto-pulmonaire, adenopathies hilaires, Sequestration splenique, positive, Hematome, douleur dentaire, Hypersplenisme, Cytolyse hepatique, infection virale, Toux nocturne, Moelle riche, EMPHYSEME, Hypogammaglobulinemie, selles glaireuses, malformation pulmonaire, Respiration, Sueurs nocturnes, 70%, Pancytopenie, pi3k, Detresse

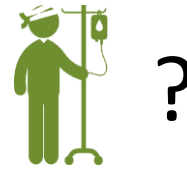
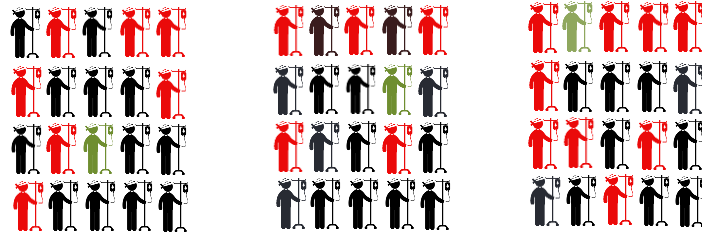
Translational research and similarity

A short story



Activated PI3K δ syndrome (APDS) : 2 genes
PIK3CD (Agnulo et al, 2013)
PIK3R1 (Deau et al, 2014)

Evaluation on the APDS cohort



Last visit 2012
Unclassified immune deficiency



PIK3R1 deletion

Dr Warehouse : A translational datawarehouse for rare diseases

Journal of Biomedical Informatics 80 (2018) 52–63



Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin



Dr Warehouse published in February 2018

Dr Warehouse® is open source

A clinician friendly data warehouse oriented toward narrative reports: Dr. Warehouse



Nicolas Garcelon^{a,b,*}, Antoine Neuraz^{b,c}, Rémi Salomon^{a,d}, Hassan Faour^a, Vincent Benoit^a, Arthur Delapalme^a, Arnold Munnich^{a,e,f}, Anita Burgun^{b,c}, Bastien Rance^{b,g}

^a Institut Imagine, Paris Descartes Université Paris Descartes-Sorbonne Paris Cité, Paris, France
^b INSERM, Centre de Recherche des Cordeliers, UMR 1138 Equipe 22, Université Paris Descartes, Sorbonne Paris Cité, Paris, France
^c Department of Medical informatics, Hôpital Necker-Enfant Malades, Assistance Publique des Hôpitaux de Paris, Paris, France
^d Service de Néphrologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
^e Département de génétique médicale, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
^f Centre de Référence des Maladies Osseuses Constitutionnelles, INSERM UMR 1163, Laboratoire de bases moléculaires et physiopathologiques de l'ostéochondrodysplasie, Paris Descartes-Sorbonne Paris Cité University, AP-HP, Institut Imagine, 75015 Paris, France
^g Hôpital Européen Georges Pompidou, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

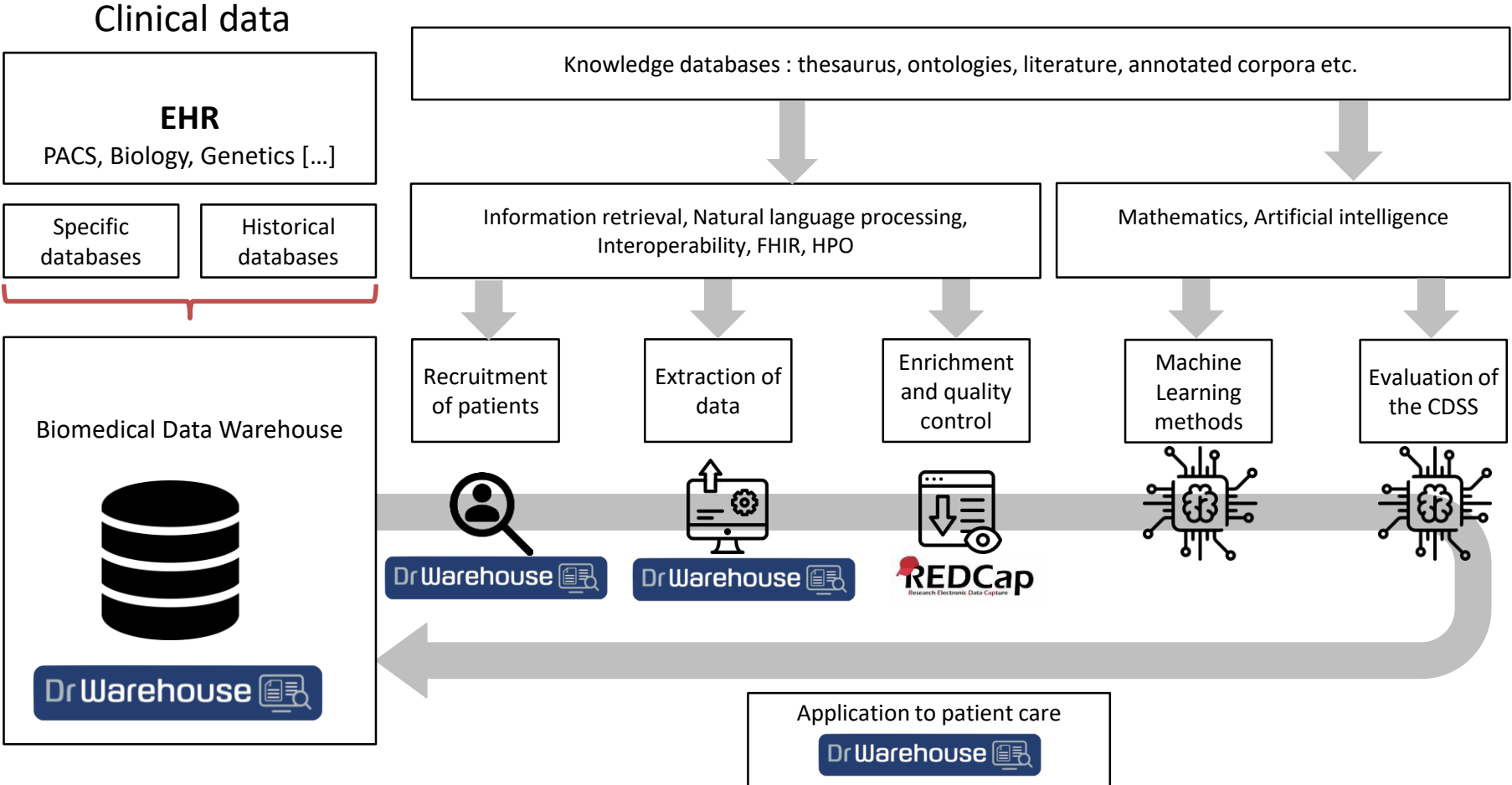
ARTICLE INFO

Keywords:
Software
Computational biology
Method
Data warehouse
Rare diseases
Electronic health records
Information storage and retrieval
Text-mining

ABSTRACT

Introduction: Clinical data warehouses are often oriented toward integration and exploration of coded data. However narrative reports are of crucial importance for translational research. This paper describes Dr. Warehouse*, an open source data warehouse oriented toward clinical narrative reports and designed to support clinicians' day-to-day use.
Method: Dr. Warehouse relies on an original database model to focus on documents in addition to facts. Besides classical querying functionalities, the system provides an advanced search engine and Graphical User Interfaces adapted to the exploration of text. Dr. Warehouse is dedicated to translational research with cohort recruitment capabilities, high throughput phenotyping and patient centric views (including similarity metrics among patients). These features leverage Natural Language Processing based on the extraction of UMLS* concepts, as well as negation and family history detection.
Results: A survey conducted after 6 months of use at the Necker Children's Hospital shows a high rate of satisfaction among the users (96.6%). During this period, 122 users performed 2837 queries, accessed 4,267 pa-

An architecture focused on artificial intelligence for clinicians



Conclusion

Accelerate the reuse of **collective memory** for both research and patient care.

Put the tools of data science back into **the hands of doctors** to enable them to explore their own data in their daily work.

Break down data from the clinic and research.

Accelerate translational research

Creation of a start up : codoc

Dr Warehouse is in 4 hospitals : Necker, HEGP, Foch, Sainte Anne

Demonstration available here : <http://www.drwarehouse.org/>

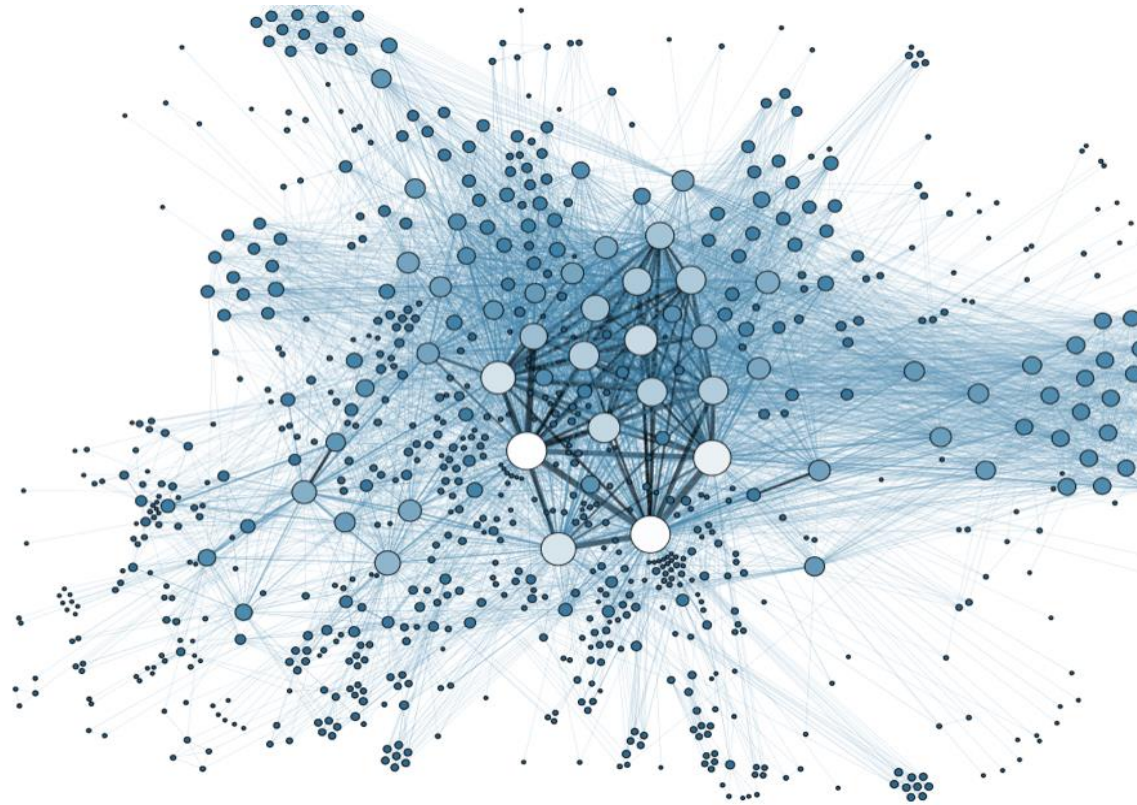
Imagine – Data science Platform Team :

Nicolas Garcelon

Vincent Benoit

Hassan Faour

Marc Vincent



nicolas.garcelon@institutimagine.org