Dr Warehouse – A translational Data Warehouse

Nicolas Garcelon
12 juin 2019
A Translational Research Context

Clinical activity
Hospital databases

600 beds
400 pediatric beds

Genetic research
Research databases

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

30 labs
12 platforms
A Translational Research Context

Clinical activity
Hospital databases

600 beds
400 pediatric beds

Genetic research
Research databases

Patient recruitment
and data mining

New Knowledge

Reuse data produced
by clinical teams

Reuse data produced
by research teams
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

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

« 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
Heterogeneous:

- Format
- Transmission (flow, databases)
- Volume
- Technology
A Translational Research Context

Clinical databases

Genetic research
Research databases

Data warehouse

Reuse data produced by clinical teams

New Knowledge

Reuse data produced by research teams

Patient recruitment and data mining
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

Knowledge databases
- GenAtlas
- orphanet
- NCBi
- swissprot

Clinical databases

Search engine
Data mining algorithm

Database

Existing data warehouse

i2b2
(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

- Arthritis, Juvenile
- Diabetes Mellitus, type 1
- Hepatitis, Autoimmune
- Lupus Erythematosus, Systemic
- Addison Disease
- Sjogren’s Syndrome
- Graves Disease
- Arthritis, Rheumatoid
- Dermatitis Herpetiformis
- Thyroiditis, Autoimmune
- Arthritis, Rheumatoid (subclinical)
- Lupus Erythematosus, Systemic
- Multiple Sclerosis
- Polyendocrinopathy, Autoimmune
- Antiphospholipid Syndrome
- Myasthenia Gravis

Comorbidities are only in free text

81 %

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!


A Translational Research Context

Clinical databases

Patient recruitment and data mining

Reuse data produced by clinical teams

DrWarehouse

New Knowledge

Reuse data produced by research teams

Genetic research
Research databases
Dr Warehouse

A data warehouse oriented around the hospital document:

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

3 challenges identified

Dr Warehouse

Find patients for research and clinical studies

Data mining: phenotypic description

Improve Translational Research

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 »
3 challenges identified

- Find patients for research and clinical studies
- Data mining: phenotypic description
- Improve Translational Research
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

1. A search engine on textual and/or coded data

2. A dedicated interface to check the consistency of the results
Dr WareHouse ©imagine
Entrepôt de données

Rechercher des patients
Sur tout l'entrepôt

epidermolysis bullosa

Elider aux synonymes :
+ Avancé - Rechercher le résultat

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

LANCER LA RECHERCHE

341 Patients
2396 Documents

Documents contenant 'epidermolysis bullosa', en excluant les négations
Refaire une recherche sur le résultat :

Export les patients :

Filtrer le résultat ci-dessous :

Echographie abdominale du CRH Hôpital :

Madame [PRENOM_PATIENT], née le [DATE_NAISS_PATIENT], a été hospitalisée en hospitalisation de jour / seance dans le service de [SERVICE] pour la première cure de Vfnour dans le contexte d'epidermolysis bullosa dystrophique recessive avec absence totale de collagène 7 [...] Epidermolysis bullosa dystrophique recessive avec absence totale de collagène 7 [...] Outre la réhydratation de la peau et du tissu conjonctif, des complications cutanées et digestives ont été rencontrées chez l'enfant présentant une epidermolysis bullosa [...]

11 ans

CRH Hôpital :

Madame [PRENOM_PATIENT], née le [DATE_NAISS_PATIENT], a été hospitalisée en hospitalisation de jour / seance dans le service de [SERVICE] pour la première cure de Vfnour dans le contexte d'epidermolysis bullosa dystrophique recessive avec absence totale de collagène 7 [...] Epidermolysis bullosa dystrophique recessive avec absence totale de collagène 7 [...] Outre la réhydratation de la peau et du tissu conjonctif, des complications cutanées et digestives ont été rencontrées chez l'enfant présentant une epidermolysis bullosa [...]

10 ans

630 CRH Dermatologie du :

Madame [PRENOM_PATIENT], née le [DATE_NAISS_PATIENT], a été hospitalisée en hospitalisation de jour / seance dans le service de [SERVICE] pour la première cure de Vfnour dans le contexte d'epidermolysis bullosa dystrophique recessive avec absence totale de collagène VII chez une enfant de [DATE_NAISS_PATIENT] [...] Epidermolysis bullosa dystrophique recessive confirmée sur le plan immunologique (mutation homozygote PARB/13K sur le gène du collagène VII) [...] Epidermolysis bullosa dystrophique recessive chez une jeune fille de [DATE_NAISS_PATIENT] [...]
Dr WareHouse

Rechercher des patients

Sur tout l'entrepôt

 Epidermolyse bulleuse

341 Patients
2396 Documents

Documents contenant 'epidermolyse bulleuse', en excluant les négations

Refaire une recherche sur le résultat :

SAUVER LA REQUETE

Exportation des résultats :

Filtrer les résultats ci-dessous :

Echographie abdominale du 11 ans

RADIOLOGIE PEDIATRIQUE :

INDICATION : 5é Inflammatoire biliaire avec cholestéatome, douleurs abdominales et vomissements chez un enfant présentant une epidermolyse bulleuse [...]
To facilitate the selection of patients for studies
To facilitate the selection of patients for studies
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 »
Reducing noise

False positive patients:
« We can exclude the diagnosis of lupus »
« The father has a crohn’s disease »

We developed a search engine optimization strategy by detecting negation and family history.
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."
### Integration in Dr Warehouse

Dr Warehouse is an electronic medical record system that integrates various data sources to provide comprehensive patient information. The example shown illustrates the search and retrieval of patient data, specifically focusing on cases related to Immunohematology Pediatrics. The interface includes options to search by patient details, diagnosis, treatment, and other medical records. The highlighted section shows a query for patients with specific conditions and treatments, indicating the versatility and depth of data available within the system.
Integration in Dr Warehouse

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

SCID

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

syndrome d'Omenn

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

+ afficher les autres documents

ARTEMIS
Dr Warehouse

3 challenges identified

- Find patients for research and clinical studies
- Data mining: phenotypic description
- Improve Translational Research
Aggregate data visualization tools

Dr Warehouse @imagine
Entrepot de données

Rechercher des patients

Sur tout l'entrepôt

épidermolyse bulleuse 341/341

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

LANDER LA RECHERCHE ?

Requêtes sauvées

Historique requête

Show: 10 entries
Search: }

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

341 Patients
2396 Documents

Afficher le parcours moyen (hospitalisations) des patients
Afficher le parcours moyen complet (hospitalisations et consultations) des patients
Aggregate data visualization tools: Pathcare
Aggregate data visualization tools

Dr. WareHouse
Entrepot de données

Search for patients

Across the entire data warehouse

lupus

104/1041

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

START A SEARCH

Query history

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.
Data warehouse 500,000 patients

For each patient of the data warehouse

Narrative records

UMLS Metathesaurus

Query “Rett syndrome”

N Patients found

Aggregated concepts and ranking

High throughput phenotyping

Phenotypes associated to Rett syndrome:

<table>
<thead>
<tr>
<th>UMLS Concepts</th>
<th>Freq</th>
<th>TF-IDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stéréotypie</td>
<td>71.8</td>
<td>8.06</td>
</tr>
<tr>
<td>Crises</td>
<td>55</td>
<td>8.37</td>
</tr>
<tr>
<td>Scoliose</td>
<td>51.2</td>
<td>6.71</td>
</tr>
<tr>
<td>Epilepsie</td>
<td>50.7</td>
<td>2.96</td>
</tr>
<tr>
<td>Syndrome pyramidal</td>
<td>42.6</td>
<td>2.87</td>
</tr>
<tr>
<td>Ostéoporose</td>
<td>42.6</td>
<td>2.79</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1022 concepts</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

UMLS Concepts | Context | Certainty | Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lupus</td>
<td>Patient</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Anémie</td>
<td>Patient</td>
<td>-1</td>
<td>2</td>
</tr>
<tr>
<td>Insuffisance rénale</td>
<td>Family</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fièvre</td>
<td>Patient</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Asthénie</td>
<td>Patient</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
High throughput phenotyping

Based on concepts extracted from medical reports of identified patients

22 millions of phenotypes extracted
Next generation phenotyping using narrative reports in a rare disease clinical data warehouse

Nicolas Garcelon1,2,3*, Antoine Neuraz2,3, Rémi Salomon4, Nadia Bahi-Buisson5, Jeanne Amiel6,7, Capucine Picard8,9, Nizar Mahiout1,8,10,11, Vincent Berroil1, Anita Burgun5,12 and Bastien Rance1,12

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 (APCDS). 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 (i.e., between 11 and 41 for TF-IDF). Average precision ranges from 0.35 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

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

Example Activated PI3K-delta syndrome (APDS):

Median age of first apparition of the signs
3 challenges identified

Find patients for research and clinical studies

Data mining: phenotypic description

Improve Translational Research

Dr Warehouse

Dr Warehouse

Patient-centered features

Facilitate the exploration of the health records: a search engine for one patient.
Patient-centered features

Visualization of the patient's path over time
Patient-centered features
Visualization of the patient's path over space
Patient-centered features

Automated extraction of concepts from the narrative records
Translational research and similarity

Undiagnosed patients

Search engine
All phenotypic criteria

A diagnosed patient
Complex phenotype

Similarity with index patient
Phenotypic proximity

\[\emptyset\]
Our approach to find similar patients

For each patient of the data warehouse

<table>
<thead>
<tr>
<th>Concepts</th>
<th>Nb</th>
<th>TF-IDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabete</td>
<td>2</td>
<td>0.02</td>
</tr>
<tr>
<td>Cataracte</td>
<td>4</td>
<td>0.2</td>
</tr>
<tr>
<td>Fièvre</td>
<td>5</td>
<td>0.001</td>
</tr>
<tr>
<td>Proteinurie</td>
<td>3</td>
<td>0.02</td>
</tr>
<tr>
<td>Thrombopenie</td>
<td>2</td>
<td>0.03</td>
</tr>
<tr>
<td>Asthenie</td>
<td>1</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Patient index

Concepts

- Syndrome de lowe: 1
- Anemie: -1
- Insuffisance rénale: 1
- Cataracte: 1

Similarity: VSM

Top30 similar patients

Distance

Patient index

X 500,000 patients
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,h,h, Smail Hadji-Rabia a,b,i, Alain Fischer a,h,i,j,k,l, Arnold Munnich a,h,m,n, Anita Burgun c,d,o

1Institut Imagine, Paris Descartes Université Paris Descartes-Sorbonne Paris Cité, Paris, France
2INSERM, Institut Imagine, UMR 1163, Université Paris Descartes, Sorbonne Paris Cité, Paris, France
3INSERM, Centre de Recherche des Cordeliers, UMR 1138 Equipe 22, Université Paris Descartes, Sorbonne Paris Cité, Paris, France
4Département d’Informatique médicale, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, Paris, France
5Service de Néphrologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, Paris, France
6Laboratory of Human Lymphohematopoiesis, INSERM UMR 1163, Paris, France
7Service de Hématologie, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
8Service de néphrologie pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
9Service 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
10Centre de Références Déficits immunitaires Héréditaires, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
11École d’Immunoalérologie-Hématologie et Rhumatologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
12Collège de France, Paris, France
13Département de génétique médicale, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes Sorbonne Paris Cité, France
14Centre de Référence des Maladies Osseuses Constitutionnelles, INSERM UMR 1163, Laboratoire de bases moléculaires et physiopathologiques de l’ostéochondrodystrophies, Paris, France
Integration of the similarity algorithm in Dr Warehouse
Translational research and similarity

A short story

Activated PI3Kδ 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 data warehouse for rare diseases

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

a Institute Imagine, Paris Descartes University 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 Service de Nephrologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, Paris, France
d Département de génétique médicale, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, Paris, France
e Centre de Référence des Maladies Osteopéniques Congénitales, INSERM UMR 1163, Laboratoire de bases moléculaires et physiopathologiques de l’ostéochondrodysplasies, Paris Descartes-Sorbonne Paris Cité University, AP-HP, Institut Imagine, 75015 Paris, France
f Hôpital Européen Georges Pompidou, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, Paris, France

ARTICLE INFO

Keywords: 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 pages, and downloaded 646 documents.
An architecture focused on artificial intelligence for clinicians

Clinical data

**EHR**
PACS, Biology, Genetics [...]

Specific databases

Historical databases

Biomedical Data Warehouse

Knowledge databases: thesaurus, ontologies, literature, annotated corpora etc.

Information retrieval, Natural language processing, Interoperability, FHIR, HPO

Enrichment and quality control

Machine Learning methods

Evaluation of the CDSS

Recruitment of patients

Extraction of data

Mathematics, Artificial intelligence

Application to patient care
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