



FONDATION
DE L'AP-HP

CHAIRE
AI-RACLÈS

Inria

Journée Santé et IA

Conceptions de phénotypes computationnels
pour la recherche en épidémiologie

[Pegdwendé N. Sawadogo](#), Thomas Guyet, Etienne Audureau

28 juin 2022

Plan

- 1 Introduction
- 2 Notion de phénotype computationnel
- 3 Définition multi-dimensionnelle
- 4 Comparaison de systèmes
- 5 Perspectives

Digitalisation des systèmes médicaux

De nouveaux services

- Aide au diagnostic
- Implants/prothèses sur mesure
- Suivi des traitements



Digitalisation des systèmes médicaux

De nouveaux services

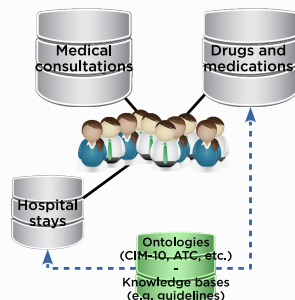
- Aide au diagnostic
- Implants/prothèses sur mesure
- Suivi des traitements



nicematin.com

De nouveaux outils

- Imprimantes 3D/capteurs
- Techniques d'IA
- Entrepôts de données de santé



Vers un usage secondaire des données de santé

Utilisation primaire

- Partage/exploitation des infos sur le patient (EDS)
- Suivi/remboursement des actes médicaux



<https://www.huderf.be>

Usage en épidémiologie

- Surveillance des effets secondaires
- Etude des interactions médicamenteuses



<https://santetudiant.com>

Problématique d'interrogation des EDS

Nécessité de constituer des cohortes

- Etape primordiale des études épidémiologiques
- Identifier des patients par des critères divers
- Formaliser le processus

Problèmes

- EDS conçus pour un usage différent
- Combinaison de critères complexes
- Gap sémantique entre infos stockées et recherchées



Solution :

Formalisation de phénotypes computationnels

Plan

- 1 Introduction
- 2 Notion de phénotype computationnel**
- 3 Définition multi-dimensionnelle
- 4 Comparaison de systèmes
- 5 Perspectives

Notion de phénotype

(Johannsen, 1909; Loebe et al., 2012; Uciteli et al., 2020)

Un phénotype est un ensemble de caractéristiques observables permettant de distinguer des organismes.

C'est une combinaison de **caractéristiques...**

- **morphologiques** : couleur des yeux, taille
- **bio-chimiques** : gènes, groupes sanguins
- **cliniques** : symptômes

Phénotypes computationnels

Un phénotype computationnel est...

(De Moor et al., 2014; Ahmad et al., 2020)

la matérialisation d'un phénotype sous la forme d'une **requête exécutable** sur un EDS

(Mo et al., 2015)

un **workflow** dédié à l'interrogation de données cliniques

(Sharma et al., 2019)

un **modèle d'apprentissage** supervisé pour constituer une cohorte de patients en fonction d'un ensemble de critères

Exemple - Patients épileptiques dans le SNDS

« *Patients ayant eu au moins 10 délivrances d'un médicament anti-épileptique sur 1 an.* »

Discussion

De nombreuses contradictions et insuffisances...

- **Besoin d'un outil pour concevoir/utiliser des phénotypes computationnels sur l'EDS de l'AP-HP**
- **≠ visions de la matérialisation d'un phénotype computationnel**
 - Script?
 - Requête?
 - Workflow?
 - Modèle d'apprentissage?
- **Définitions partielles**
 - Spécification des **entrées** (De Moor et al., 2014; Uciteli et al., 2020)
 - Spécification des **sorties** (Sharma et al., 2019)

Outline

- 1 Introduction
- 2 Notion de phénotype computationnel
- 3 Définition multi-dimensionnelle**
- 4 Comparaison de systèmes
- 5 Perspectives

Vision multi-dimensionnelle

Définition

Un phénotype computationnel (\mathcal{P}) est caractérisable par un ensemble de 5 dimensions que sont :

- ses **interactions** avec l'EDS (\mathcal{I}),
- ses **entrées** (\mathcal{E}),
- ses **sorties** (\mathcal{S}),
- sa **matérialisation** (\mathcal{M}),
- sa **complétude** (\mathcal{C})

$$\mathcal{P} = \langle \mathcal{I}, \mathcal{E}, \mathcal{S}, \mathcal{M}, \mathcal{C} \rangle$$

- Définition plus générique et plus complète
- Outil de comparaison de systèmes de phénotypage

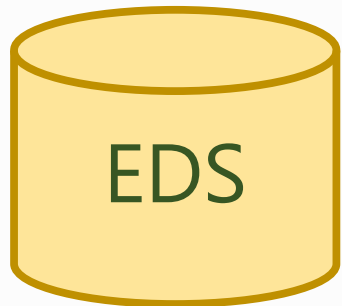
Interactions avec l'EDS $\langle I, E, S, M, C \rangle$

Génération via l'EDS

- Caractéristique souhaitable, mais optionnelle

Exécution sur l'EDS

- Caractéristique strictement nécessaire



Conception

A yellow arrow pointing from the EDS cylinder towards the code icon.



Exécution

A blue arrow pointing from the code icon back towards the EDS cylinder.

Phénotype
computationnel

Entrées $\langle \mathcal{I}, \mathbf{E}, \mathcal{P}, \mathcal{M}, \mathcal{C} \rangle$

Faits

- Caractéristiques du patient
- Évolutifs ou permanents
- Ex : sexe, groupe sanguin, taille



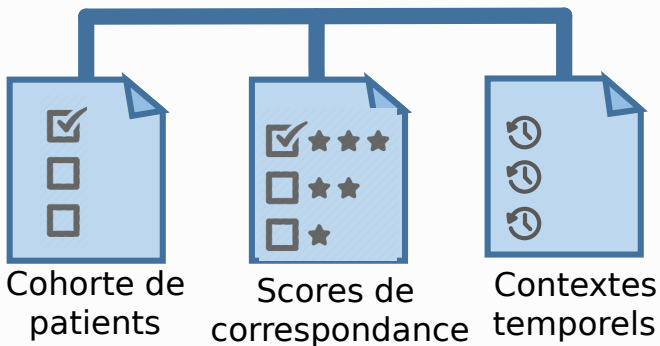
<https://publicdomainvectors.org/>

Évènements

- Actions médicales horodatées
- Ponctuels ou continus
- Ex : Consultation, hospitalisation



Sorties $\langle \mathcal{I}, \mathcal{E}, \mathcal{S}, \mathcal{M}, \mathcal{C} \rangle$



Matérialisation $\langle \mathcal{I}, \mathcal{E}, \mathcal{P}, \mathbf{M}, \mathcal{L} \rangle$

1/2

Modèle apprenable

- Instance d'un algo d'apprentissage machine
- Totalement inadapté aux non-informaticiens



Matérialisation $\langle \mathcal{I}, \mathcal{E}, \mathcal{P}, \mathcal{M}, \mathcal{C} \rangle$

1/2

Modèle apprenable

- Instance d'un algo d'apprentissage machine
- Totalement inadapté aux non-informaticiens



Script

- Séquence d'instruction dans un langage de programmation
- Inadapté aux non-informaticiens



Matérialisation $\langle \mathcal{I}, \mathcal{E}, \mathcal{P}, \mathcal{M}, \mathcal{C} \rangle$

2/2

Requête

- Restrictions exprimées via un langage de requête
- Plutôt inadapté aux non-informaticiens

```
SELECT EmployeeID, FirstName, LastName, HireDate, Ci
FROM Employees
WHERE HireDate BETWEEN '1-june-2012' AND '15-decemb
```

```
SELECT EmployeeID, FirstName, LastName, HireDate, City
WHERE City IN ('Seattle', 'Tacoma', 'Redmond')
```

```
SELECT EmployeeID, FirstName, LastName, HireDate, Ci
FROM Employees
WHERE HireDate NOT BETWEEN '1-june-2012' AND '15-de
```

Matérialisation $\langle \mathcal{I}, \mathcal{E}, \mathcal{P}, \mathcal{M}, \mathcal{C} \rangle$

2/2

Requête

- Restrictions exprimées via un langage de requête
- Plutôt inadapté aux non-informaticiens

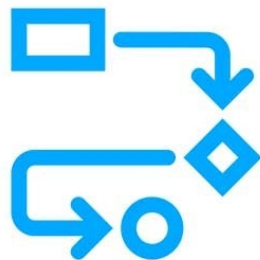
```
SELECT EmployeeID, FirstName, LastName, HireDate, Ci
FROM Employees
WHERE HireDate BETWEEN '1-june-2012' AND '15-decem

SELECT EmployeeID, FirstName, LastName, HireDate, City
WHERE City IN ('Seattle', 'Tacoma', 'Redmond')

SELECT EmployeeID, FirstName, LastName, HireDate, Ci
FROM Employees
WHERE HireDate NOT BETWEEN '1-june-2012' AND '15-de
```

Workflow

- Séquence d'opérations sous formes graphiques
- Adapté aux non-informaticiens



Critères de Complétude $\langle \mathcal{I}, \mathcal{E}, \mathcal{P}, \mathcal{M}, \mathbf{C} \rangle$

- Contraintes temporelles
- Entrées hétérogènes
- Plusieurs modèles de données
- Moindre technicité d'utilisation
- Extension sémantique

Outline

- 1 Introduction
- 2 Notion de phénotype computationnel
- 3 Définition multi-dimensionnelle
- 4 Comparaison de systèmes**
- 5 Perspectives

Comparaison de quelques systèmes

Système	Interactions	Entrées	Sorties	Matérialisation	Complétude
ECLECTIC	Gen, Exc	Fai, Evn	Coh	Req	Tmp, Sem
Jiang <i>et al.</i> (2017)	Exc	Fai, Evn	Coh	Req	Tec
Uciteli <i>et al.</i> (2020)	Gen, Exc	Fai, Evn	Coh	Req	Tec, Sem
PhenoFlow	Exc	Fai, Evn	Coh	Wkf	Tec
SCALPEL	Exc	Fai, Evn	Coh	Scr	Tmp, Sem
Bakalara <i>et al.</i> (2021)	Exc	Evn	Coh	Req	Tec, Tmp, Sem

Comparaison de quelques systèmes

Système	Interactions	Entrées	Sorties	Matérialisation	Complétude
ECLECTIC	Gen, Exc	Fai, Evn	Coh	Req	Tmp, Sem
Jiang <i>et al.</i> (2017)	Exc	Fai, Evn	Coh	Req	Tec
Uciteli <i>et al.</i> (2020)	Gen, Exc	Fai, Evn	Coh	Req	Tec, Sem
PhenoFlow	Exc	Fai, Evn	Coh	Wkf	Tec
SCALPEL	Exc	Fai, Evn	Coh	Scr	Tmp, Sem
Bakalara <i>et al.</i> (2021)	Exc	Evn	Coh	Req	Tec, Tmp, Sem

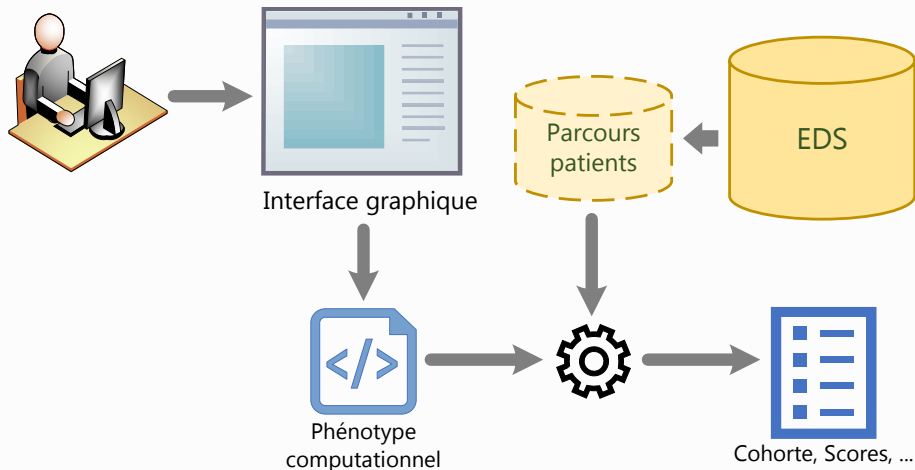
- le **phénotype computationnel** est toujours **exécuté** sur un EDS
- seule la **sortie** sous forme de **cohorte de patients** est proposée
- la **matérialisation** sous forme de requête est **prédominante**

Outline

- 1 Introduction
- 2 Notion de phénotype computationnel
- 3 Définition multi-dimensionnelle
- 4 Comparaison de systèmes
- 5 Perspectives**

Perspectives

Un système plus complet adapté aux non-informaticiens





FONDATION
DE L'AP-HP

CHAIRE
AI-RACLÈS

Inria

Journée Santé et IA

Conceptions de phénotypes computationnels
pour la recherche en épidémiologie

[Pegdwendé N. Sawadogo](#), Thomas Guyet, Etienne Audureau

28 juin 2022