



**HAL**  
open science

# Étudier les mécanismes des effets indésirables des médicaments avec l'IA explicable : expériences avec la fouille de graphes de connaissances

Pierre Monnin, Adrien Coulet

## ► To cite this version:

Pierre Monnin, Adrien Coulet. Étudier les mécanismes des effets indésirables des médicaments avec l'IA explicable : expériences avec la fouille de graphes de connaissances. Bulletin de l'Association Française pour l'Intelligence Artificielle, 2022. hal-03669065

**HAL Id: hal-03669065**

**<https://inria.hal.science/hal-03669065v1>**

Submitted on 16 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



## ■ Étudier les mécanismes des effets indésirables des médicaments avec l'IA explicable : expériences avec la fouille de graphes de connaissances

Orange

Equipe HeKA, Inria Paris,  
Centre de Recherche des Cordeliers (Inserm, U. de Paris)  
<https://team.inria.fr/heka/>

**Pierre MONNIN**

Chercheur

[pierre.monnin@orange.com](mailto:pierre.monnin@orange.com)

<https://pmonnin.github.io>

**Adrien COULET**

Chercheur

[adrien.coulet@inria.fr](mailto:adrien.coulet@inria.fr)

[https://team.inria.fr/heka/  
team-members/coulet/](https://team.inria.fr/heka/team-members/coulet/)

### Introduction

Les événements indésirables médicamenteux (EIM) sont mis en évidence et statistiquement caractérisés grâce à des essais cliniques randomisés, puis à la pharmacovigilance après mise sur le marché. Cependant, les mécanismes moléculaires responsables de ces effets restent inconnus dans la plupart des cas. C'est notamment le cas pour les toxicités hépatiques ou cutanées qui sont pourtant particulièrement surveillées durant la phase d'évaluation des médicaments. En marge des essais cliniques, de nombreuses autres connaissances à propos des médicaments et des molécules les constituant sont disponibles. En particulier, des graphes de connaissances en accès libre décrivent leurs propriétés, interactions, et implications dans les réseaux biologiques (*i.e.*, les *pathways*). D'autre part, des classifications expertes ont été établies manuellement et permettent de disposer de listes de médicaments connus pour causer, ou non, différents types d'EIM.

### L'IA explicable pour expliquer les EIM

Dans un récent article [1], nous avons fouillé des graphes de connaissances biomé-

dicaux afin d'identifier les propriétés biomoléculaires des médicaments qui permettent de reproduire automatiquement les classifications expertes, distinguant les médicaments causant ou non un type spécifique d'EIM. Dans la perspective de produire des artefacts capables d'expliquer ces classifications nous avons mis en œuvre des méthodes de classification nativement explicables : les arbres de décision et les règles de classification. En effet, ces méthodes produisent des modèles compréhensibles par les humains qui expliquent la classification de nouveaux exemples [4]. De plus, nous avons évalué l'hypothèse selon laquelle les propriétés associées aux médicaments pourraient aussi fournir aux experts des éléments explicatifs pour les mécanismes moléculaires sous-jacents aux EIM.

Nous avons testé notre approche à partir de deux classifications expertes distinguant les médicaments causant ou non des toxicités hépatiques (ou *DILI*, pour *Drug-Induced Liver Injuries*) et cutanées (ou *SCAR*, pour *Severe Cutaneous Adverse Reactions*). Pour extraire des propriétés associées à ces médicaments, nous les alignons avec un graphe de connaissances que nous fouillons ensuite. Il s'agit ici de PGx-

1. <https://pgxlod.loria.fr>



LOD<sup>1</sup>, un graphe de connaissances biomédicales que nous avons précédemment créé [2] et qui intègre, connecte, et complète des graphes de connaissances publics (e.g., CTD, DisGeNET, DrugBank, etc.). Dans le processus de fouille, les nœuds représentant les médicaments sont appelés “graines” et nous extrayons leurs voisins, les chemins, et les patrons de chemins dont ils sont racines. Par exemple, sur la Figure 1,  $v_4$  est un voisin de  $n_2$ ,  $\frac{p_1}{\rightarrow} v_2 \xrightarrow{p_2} v_3$  est un chemin dont  $n_1$  est racine, et  $\frac{p_1}{\rightarrow} C_1 \xrightarrow{p_2} C_3 \xrightarrow{p_3} v_6$  est un patron de chemins dont  $n_1$  et  $n_2$  sont racines. L'intérêt des patrons de chemins réside dans leur capacité, par généralisation, à être communs à davantage de graines que les chemins. La Figure 1 illustre comment les nœuds de deux chemins distincts peuvent être généralisés et remplacés par les classes que ceux-ci instancient pour créer un patron de chemins unique. Nous avons particulièrement étudié le passage à l'échelle de ce processus de fouille en créant l'algorithme `kgpm` [3]<sup>2</sup>.

Sur la base des propriétés extraites, nous avons entraîné deux classifieurs à distinguer les médicaments causant ou non chacun des deux types d'EIM étudiés. Nous isolons les propriétés qui sont à la fois discriminantes dans la reproduction des classifications expertes et interprétables par des experts (i.e., des termes de la *Gene Ontology*, des molécules cibles de médicaments, ou des noms de réseaux biologiques). Le caractère explicatif des propriétés isolées a ensuite été manuellement évalué par 3 experts en pharmacologie.

## Évaluation quantitative et qualitative

Les propriétés extraites du graphe permettent de reproduire avec un bon niveau de fidélité les classifications expertes des médicaments causant ou non les EIM (*accuracy* de 0,74 et 0,81 pour DILI et SCAR respectivement). Les experts ont évalué unanimement

que 73% (pour DILI) et 38% (pour SCAR) des propriétés discriminantes sont possiblement explicatives des EIM associés. Ils sont partiellement d'accord (2/3) pour dire que 90% (pour DILI) et 77% (pour SCAR) de ces propriétés sont possiblement explicatives.

Par exemple, le patron de chemin  $\frac{\text{interactsWith}}{\rightarrow} \text{Enzyme} \xrightarrow{\text{cellularComponent}}$  **Endoplasmic reticulum** a été unanimement évalué comme potentiellement explicatif pour DILI. En effet, le réticulum endoplasmique est connu, en particulier dans les tissus hépatiques, pour contenir des enzymes de la famille des cytochromes P450 qui sont fortement impliqués dans le métabolisme des médicaments.

## Conclusion et perspectives

Notre étude suggère que le graphe de connaissances utilisé fournit des propriétés suffisamment diverses pour permettre à des modèles simples et explicables de distinguer les médicaments causant ou non des EIM. En plus de permettre de reproduire les classifications expertes, les propriétés discriminantes semblent porter le sens d'une réalité biologique qui en font des candidats à considérer pour étudier de manière plus approfondie les mécanismes biologiques sous-jacents aux EIM.

Dans notre travail, nous nous sommes concentrés sur des classifieurs nativement explicatifs, mais relativement naïfs (arbres de décision et règles de classification). Il semble intéressant de reproduire nos expériences avec des réseaux de neurones profonds adaptés aux graphes comme les *Graph Convolutional Networks* et les *Graph Neural Networks* pour mesurer la perte de performance potentielle associée à notre choix de modèle. Une telle comparaison quantitative devrait toutefois s'accompagner d'une comparaison qualitative prenant en compte la nécessité d'étapes supplémentaires pour expliquer ces modèles profonds. Par

2. <https://github.com/pmonnin/kgpm>

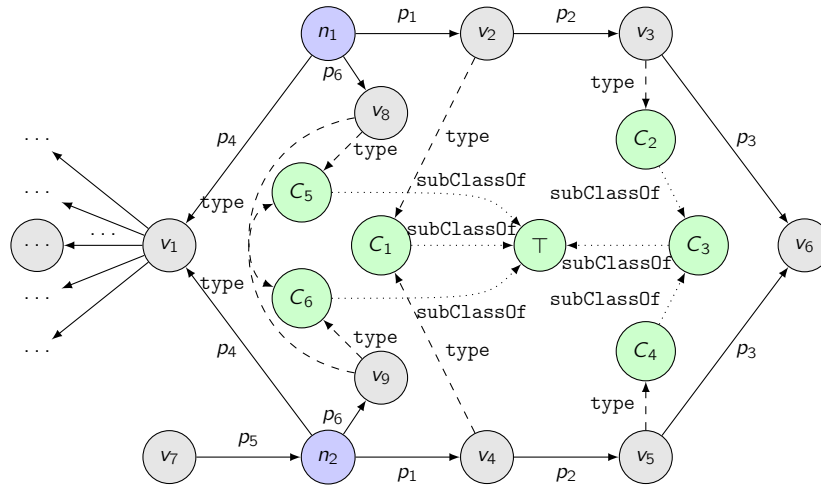


Figure 1 – Exemple de graphe de connaissances.  $n_1$  et  $n_2$  sont les “graines” pour l’extraction de propriétés, les  $v_i$  sont des individus voisins, et les  $C_i$  sont des classes d’ontologie utilisées pour la construction de patrons de chemins. Par exemple,  $n_1$  et  $n_2$  sont caractérisés par le voisin  $v_1$ , le chemin  $\xrightarrow{p_4} v_1$ , et les patrons de chemins  $\xrightarrow{p_6} T$  et  $\xrightarrow{p_1} C_1 \xrightarrow{p_2} C_3 \xrightarrow{p_3} v_6$ . Des paramètres s’appliquent pour limiter le nombre de voisins, chemins, et patrons de chemins extraits (e.g., support, spécificité, longueur) [3].

exemple, l’utilisation de cartes de saillance (*saliency maps*) permet d’obtenir des informations sur la couche ou les neurones activés par certaines instances. Ces informations sont particulièrement intéressantes pour les spécialistes en apprentissage machine mais nécessitent un travail d’interprétation supplémentaire avant d’être compréhensibles par des experts du domaine d’application, parfois non familiers avec les réseaux de neurones. Les valeurs de Shapley sont plus généralement interprétables et pour cette raison, leur adaptation aux méthodes de fouille de graphes mériterait d’être considérée dans la poursuite de ce travail.

## Références

- [1] Emmanuel Bresso et al. Investigating ADR mechanisms with explainable AI : a feasibility study with knowledge graph mining. *BMC Medical Informatics and Decision Making*, 21(1) :171, 2021.
- [2] Pierre Monnin et al. PGxO and PGxLOD : a reconciliation of pharmacogenomic knowledge of various provenances, enabling further comparison. *BMC Bioinformatics*, 20-S(4) :139 :1–139 :16, 2019.
- [3] Pierre Monnin et al. Tackling scalability issues in mining path patterns from knowledge graphs : a preliminary study. In *Proceedings of the First International Conference "Algebras, graphs and ordered sets", ALGOS 2020, Online (originally planned in Nancy, France), August 26-28, 2020*, volume 2925 of *CEUR Workshop Proceedings*, pages 123–137. CEUR-WS.org, 2020.
- [4] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5) :206–215, 2019.