



**HAL**  
open science

## Traitement automatique des déclarations des Erreurs Médicamenteuses avec un modèle Transformer GPT

Touati Hanae, Thabet Rafika, Fontanili Franck, Pruski Marc, Elyes Lamine

### ► To cite this version:

Touati Hanae, Thabet Rafika, Fontanili Franck, Pruski Marc, Elyes Lamine. Traitement automatique des déclarations des Erreurs Médicamenteuses avec un modèle Transformer GPT. Alass-Giseh 2024 - 1er congrès de l'Association latine pour l'analyse des systèmes de santé et de la communauté de Gestion et Ingénierie des systèmes hospitaliers, Jul 2024, Liège (B), Belgique. pp.390-397. hal-04684078

**HAL Id: hal-04684078**

**<https://imt-mines-albi.hal.science/hal-04684078>**

Submitted on 2 Sep 2024

**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.

# Traitement automatique des déclarations des Erreurs Médicamenteuses avec un modèle Transformer GPT

Touati Hanae <sup>1</sup>, Thabet Rafika <sup>1,2</sup>, Fontanili Franck <sup>1</sup>, Pruski Marc <sup>3</sup>, Lamine Elyes <sup>1,4</sup>

<sup>1</sup> IMT Mines Albi, Allée des Sciences, 81000 Albi, {hanae.touati, franck.fontanili, rafika.thabet}@mines-albi.fr

<sup>2</sup> ISITCom, Laboratoire MARS, LR17ES05, Sousse, Tunisie

<sup>3</sup> Atout Majeur Concept, Rue du Château de Ribaute, 31130 Quint-Fonsegrives

<sup>4</sup> ISIS – Université Champollion, Campus Universitaire, Rue Firmin Oulès, 81104 Castres, elyes.lamine@univ-jfc.fr

## Résumé. Français

Les Erreurs médicamenteuses (EM) constituent un défi majeur pour les établissements de santé, en raison de leurs conséquences potentiellement graves qui peuvent aller jusqu'à mettre en jeu le pronostic vital des patients. Afin de mieux gérer ces incidents, les professionnels de santé suivent un processus de gestion qui s'appuie sur le signalement puis l'analyse des déclarations renseignées par les professionnels de santé via différents outils de déclaration, tant numériques que papier. Bien que ces outils soient personnalisables et demandent de renseigner des informations spécifiques à travers des questions à choix multiples, des listes déroulantes ou des descriptions textuelles, ils partagent tous la caractéristique commune d'inclure un champ textuel libre. Ainsi, l'analyse du texte est cruciale pour une compréhension approfondie et une analyse efficace des erreurs médicamenteuses. Étant donné le volume important des déclarations à traiter dans des délais restreints, il est indispensable d'automatiser ce traitement pour aider les professionnels de santé à prioriser les EM à traiter. Dans ce contexte, le traitement automatique de langage naturel à travers des modèles Transformers comme GPT se présente comme une solution pertinente. Le modèle GPT est capable d'extraire rapidement les informations clés de la déclaration et permet la structuration des descriptions textuelles des EM.

**Mots clés :** Gestion des risques pour les patients, Erreur médicamenteuse, Traitement automatique du langage.

## Resumen. Español

Los errores de medicación (EM) representan un desafío mayor para las instituciones de salud debido a sus consecuencias potencialmente graves que pueden llegar a comprometer el pronóstico vital de los pacientes. Para manejar mejor estos incidentes, los profesionales de la salud siguen un proceso de gestión que se basa principalmente en el reporte y luego en el análisis de las declaraciones proporcionadas por los profesionales de la salud a través de diferentes herramientas de reporte, tanto digitales como en papel. Aunque estas herramientas son personalizables y requieren que se introduzcan informaciones específicas a través de preguntas de opción múltiple, listas desplegables o descripciones textuales, todas comparten la característica común de incluir un campo de texto libre. Por lo tanto, el análisis de texto es crucial para una comprensión profunda y un análisis efectivo de los errores de medicación. Dado el volumen significativo de declaraciones a procesar en plazos limitados, es indispensable automatizar este procesamiento para ayudar a los profesionales de la salud a priorizar los EM a tratar. En este contexto, el procesamiento automático del lenguaje natural a través de modelos Transformadores como GPT se presenta como una solución relevante. El modelo GPT es capaz de extraer rápidamente la información clave de la declaración y permite la estructuración de las descripciones textuales de los EM.

**Palabras claves :** Gestión de riesgos para los pacientes, Error de medicación, Procesamiento automático del lenguaje.

# 1 Introduction

Dans le domaine médical et dans le parcours de soins, les erreurs médicamenteuses (EM) figurent parmi les risques critiques affectant la sécurité des patients. Elles englobent toute erreur non intentionnelle d'un professionnel de santé, d'un patient ou d'un tiers, selon le cas, survenue au cours du processus de soin impliquant un médicament lors de la prise en charge médicamenteuse (PECM) de la prescription à l'administration et à la surveillance du patient [Schmitt et al., 2006]. Elles peuvent avoir des conséquences ayant différents niveaux de gravité compromettant l'efficacité des soins. Ainsi, la gestion ces risques liés à la PECM est cruciale pour améliorer la sécurité des patients et optimiser les résultats thérapeutiques.

Pour ce faire, la gestion est organisée autour de 3 niveaux [HAS, 2022] : (1) le niveau interne aux établissements de santé (ES), (2) le niveau régional orchestrés par les Agences Régionales de Santé (ARS) et les Structures Régionales d'Appui (SRA) et (3) le niveau national représenté par la Haute Autorité de Santé (HAS). Chacun de ces niveaux joue un rôle prépondérant dans l'analyse des erreurs, la mise en place de plan d'actions, l'élaboration de cadres, de normes et de recommandations et bonnes pratiques réglementaires visant à minimiser les risques d'EM. Cependant, malgré tous les efforts mis en place pour les diminuer et corriger, il existe encore et toujours des erreurs fréquentes et non gérées au niveau interne des ES. Ainsi, nous nous intéressons dans ce travail à la gestion des EM dans ces ES (publics ou privés).

En interne, les ES ont instauré un processus de gestion qui comporte la déclaration des EM, l'analyse de leurs causes et effets et finalement la mise en place et le suivi du plan d'action correctif et préventif. Ce processus est destiné à prévenir et à corriger les incidents liés à la PECM dont les EM. Cependant, l'application de cette procédure se heurte à des obstacles majeurs, notamment l'inadéquation des démarches existantes pour gérer l'ensemble des EM survenant dans les divers services, et dans les délais imposés par le règlement interne de chaque établissement. Cette réalité met en lumière la nécessité urgente de trouver de nouvelles approches pour soutenir le comité de retour d'expérience dans sa gestion. En vue des démarches actuelles, souvent manuelles et chronophages, il est quasi impossible de couvrir efficacement et rapidement toutes les EM déclarées pour une période et survenant dans les différents services lors d'une même réunion. Il faut noter qu'il n'y a que deux à trois réunions d'analyse programmées par année selon le règlement intérieur de chaque établissement. Cette situation soulève une problématique principale qui consiste à chercher comment assister au mieux les professionnels de santé dans la gestion des EM en interne à l'hôpital ? Parmi les solutions envisagées, l'automatisation de quelques étapes du processus de gestion des EM, notamment lors du choix des EM à analyser, émerge comme une voie prometteuse. Elle offre l'espoir d'une réponse plus rapide et plus fiable à cette problématique, pour garantir une sécurité optimale des patients au sein des établissements de santé.

Ce papier est structuré comme suit : dans la Section 2, nous présentons le contexte général de notre travail, à savoir le processus de gestion des EM en interne à l'hôpital, qui repose sur l'analyse des déclarations qui sont sous forme de texte descriptif ce qui pose une première problématique à l'automatisation de ce processus. Pour y remédier, nous présenterons les techniques de l'IA qui ont émergé dans le domaine médical et qui peuvent aider à structurer les déclarations des EM. Finalement, la section se termine par la présentation des modèles Transformers de langage. La Section 3 présente l'application du modèle GPT3.5 sur des déclarations des EM pour la compréhension et structuration de l'EM. Enfin, nous présentons nos conclusions et perspectives dans la Section 4.

## 2 Contexte général

### 2.1 Processus de gestion des EM en interne à l'hôpital

La gestion des EM dans les systèmes hospitaliers, suit un processus composé essentiellement de trois étapes interdépendantes comme le présente la Figure 1. Cependant, il est confronté, dès sa première étape de déclaration, à des défis majeurs compromettant l'efficacité de l'analyse et celle de la correction et la prévention des incidents futurs. La gestion s'articule autour des déclarations faites dans des systèmes de déclaration visant à collecter des informations précises caractérisant chaque EM. Ces systèmes requièrent souvent que le déclarant (professionnel de santé) remplisse des formulaires détaillant l'erreur, en combinant des descriptions textuelles avec des éléments structurés comme des cases à cocher, des listes déroulantes ou des champs préremplis paramétrés différemment selon les établissements de santé et leurs spécialités. Ces descriptions permettent d'expliquer le contexte et les circonstances de l'erreur, tandis que les éléments structurés facilitent la compréhension, le classement et l'analyse statistique des EM. Néanmoins, ce processus peut souffrir de sous-déclaration due à la charge de travail du personnel, la peur de la sanction, ou la non-réception de réponse suite à de précédents signalement. Par ailleurs, l'efficacité de l'analyse des déclarations peut être limitée par la qualité des informations fournies, qui varie en fonction de la capacité du déclarant à communiquer efficacement les détails de l'erreur, le manque de temps ou tout autre facteur.

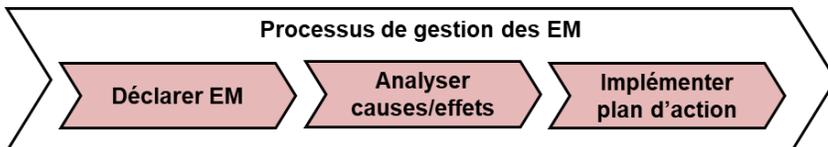


Figure 1. Processus de gestion des EM en interne à l'hôpital

Après une étude détaillée de l'existant et une comparaison des pratiques de diverses hôpitaux en France, nous avons constaté que son obstacle majeur réside dans l'hétérogénéité des outils de déclaration utilisés au sein des établissements de santé, qui varient considérablement d'un établissement à un autre. Cette diversité, résultant de la personnalisation des systèmes de déclaration numériques ou papier selon les préférences de chaque institution, entraîne un manque de standardisation des informations recueillies. De plus, la tendance à privilégier la saisie de texte libre et l'absence de champs obligatoires dans ces outils engendrent des déclarations non structurées et souvent incomplètes, limitant ainsi la précision, la profondeur de l'analyse des erreurs déclarées et le potentiel d'utilisation des techniques de science de données pour une analyse automatisée.

La programmation préférentielle des outils de déclaration avec l'existence majeure du texte descriptif libre conduit quant à elle à une variabilité significative dans la qualité et la quantité des informations fournies dans les déclarations des erreurs médicamenteuses. Le texte libre en déclaration, sans directive claire ou exigence pour la saisie des informations spécifiques, rend difficile l'extraction de données cohérentes et pertinentes pour l'analyse. En conséquence, le comité de retour d'expérience (CREX), chargé de l'analyse des causes profondes des erreurs médicamenteuses et la mise en place d'un plan d'actions correctif et préventif, doit souvent se référer à des documents complémentaires, tels que les dossiers des patients ou les fiches de service et rapport pour obtenir une vue d'ensemble des circonstances entourant chaque erreur. Cette approche fragmentée et laborieuse ne favorise pas une compréhension complète des EM, entravant ainsi les efforts et le temps pour améliorer la sécurité des patients et la qualité des soins. Pour remédier à cet obstacle, la solution est de structurer la phase de déclaration à l'aide d'un formulaire type, dans le but d'harmoniser la collecte des données essentielles à la fois pour la déclaration et l'analyse des EM. Mais avant qu'un tel formulaire ne soit validé et utilisé, il existe un large volume de déclarations des EM à exploiter afin d'améliorer le processus de gestion. Le point commun entre toutes ces déclarations se résume dans le texte descriptif. Ainsi, cette amélioration reposera principalement sur le traitement et l'extraction des informations pertinentes de ce texte descriptif.

## **2.2 Place de l'IA dans le traitement du langage appliqué en domaine médical**

L'intégration de l'intelligence artificielle (IA) dans le domaine de la santé représente une avancée majeure, promettant d'améliorer significativement la qualité et l'efficacité des soins médicaux [Damiani et al., 2023], en termes de diagnostic, de prévention, de résultats. L'application de l'IA dans le traitement des données de santé a démontré d'importantes performances, notamment dans l'analyse d'imagerie médicale, la prédiction des maladies, et la personnalisation des traitements [Subrahmanya et al, 2022]. Cependant, c'est dans le traitement automatique du langage naturel (TAL) que l'IA pourra afficher un potentiel révolutionnaire pour le secteur médical [Alowais et al, 2023], notamment dans la gestion des EM. Bien que les EM constituent une préoccupation majeure pour la sécurité des patients, impliquant des risques significatifs pour leur santé. La gestion efficace de ces erreurs nécessite une analyse minutieuse des données recueillies, à savoir les déclarations des EM qui sont souvent sous forme de textes descriptifs complexes. Ici, l'IA, et plus précisément le TAL, se révèle être un outil puissant. En automatisant l'analyse de ces déclarations descriptives d'EM, le TAL peut aider à la compréhension du contexte et l'identification rapide des caractéristiques des EM présentes dans la déclaration qui seront capables d'aider le professionnel de santé dans sa gestion. Alimentés par des modèles d'apprentissage profond et des algorithmes avancés, les systèmes de TAL, sont capables de comprendre le contexte, d'extraire les informations pertinentes, et de traiter le langage naturel de manière à refléter la complexité des communications humaines dans toutes les langues [Chowdhary, 2020]. Ils permettent la classification du texte, l'analyse des sentiments, la reconnaissance des entités nommées, génération de texte..., et ce dans toutes les langues et dans différents contextes même ceux les plus spécifiques comme le domaine médical. La capacité à interpréter le langage médical spécifique permet non seulement d'analyser les données médicales plus efficacement, mais aussi d'offrir des solutions pour une meilleure amélioration. En outre, dans le cas des erreurs médicamenteuses, l'application de l'IA dans le traitement du langage naturel ouvre la voie à des systèmes de déclaration et de suivi des EM plus intuitifs et accessibles. L'analyse des descriptions textuelles, avec du TAL, permet la compréhension des déclarations et l'extraction des informations caractéristiques de l'EM ce qui va faciliter la structuration de la déclaration dans le futur facilitant la saisie des données par le déclarant. Ces systèmes peuvent contribuer à une meilleure capture et à une déclaration plus précise des EM, réduisant ainsi le risque des erreurs non signalées. Les promesses du TAL dans le domaine médical vont au-delà de la gestion des erreurs médicamenteuses. Elles englobent l'amélioration de la documentation clinique, l'optimisation de la communication entre les professionnels de santé et les patients, et le soutien à la recherche médicale [Locke et al, 2021]. En permettant une analyse rapide et précise de vastes volumes de texte, le TAL aide à libérer le potentiel des données de santé non structurées, ouvrant de nouvelles perspectives pour les soins basés sur les données probantes et la médecine personnalisée.

## **2.3 Modèles Transformers pour les tâches de TAL**

Les modèles « Transformers » ont révolutionné le traitement du langage naturel grâce à leur architecture unique, capable d'apprendre à partir d'un volume important de données non labélisées dans différentes langues et d'utiliser cet apprentissage pour performer dans différentes tâches de TAL en prenant compte du contexte du texte étudié. Appliqués aux descriptions des EM, ces modèles peuvent analyser le langage complexe et technique propre au domaine médical, afin d'extraire des informations clés caractérisant les EM à partir des textes descriptifs non structurés. Le niveau de connaissance qu'ils ont pu acquérir durant leur apprentissage leur permet de détecter des corrélations et des tendances qui échapperaient aux analyses manuelles ou aux méthodes statistiques traditionnelles. Ainsi, les modèles Transformers peuvent aider à structurer les déclarations textuelles des EM en identifiant les personnes impliquées, les phases du processus de la Prise en charge Médicamenteuse, les médicaments sujet de l'EM, les causes et conséquences de l'EM ainsi que toutes autres informations favorisant la compréhension de l'incident. Ceci contribuera significativement à l'amélioration de la sécurité des patients. Leur mise en œuvre dans les systèmes de gestion des EM promet donc non seulement d'accroître l'efficacité des processus d'analyse mais également de fournir de nouvelles pistes actionnables pour la prévention des erreurs. L'avènement des modèles Transformers a marqué une révolution dans le domaine du traitement

automatique du langage naturel (TAL), posant de nouvelles bases pour la compréhension et la génération du langage par les machines. Ces modèles se distinguent par leur capacité à traiter des séquences de données en parallèle et à capturer les dépendances complexes entre les mots, grâce à leur mécanisme d'attention qui permet de pondérer l'importance relative des différents mots dans une phrase [Rahali et Akhloufi, 2023]. Depuis l'introduction de BERT [Devlin et al, 2020], les modèles Transformers ont connu une évolution rapide, chaque nouveau modèle apportant des améliorations significatives en termes de compréhension et de génération de texte. Le Tableau 1 présente une comparaison entre les différents modèles Transformers les plus connus et utilisés depuis leur première apparition. La comparaison du tableau comporte l'architecture du modèle suivi, s'il suit une structure d'encodeur, décodeur ou encodeur et décodeur ainsi que les tâches de TAL qu'il permet de réaliser (classification du texte ou/et des tokens, le résumé ou traduction des documents, la réponse aux questions, la génération du texte ou la complétion des textes incomplets...). Aujourd'hui, il existe plusieurs modèles Transformers pré-entraînés pour exécuter des tâches spécifiques de TAL dans une langues spécifiques. Généralement, ils sont prêts à être utilisés sur des données en langue anglaises, mais peuvent être réajustés et entraînés pour la langue souhaitée. Par exemple, l'étude faite par [Frei et Kramer, 2023] a affiné GPT pour une reconnaissance d'entités nommées à partir des données cliniques en langue allemande, [Jarrar et al., 2022] ont utilisé AraBERT [Antoum et al, 2020], un nouveau modèle découlant de BERT, pour une tâche de reconnaissance d'entités nommées dans des textes en dialecte arabe.

**Tableau 1. Tableau comparatif des modèles Transformers**

| Modèle                     | Architecture du modèle |          |                   | Tâches de NLP  |        |            |     |                     |           |
|----------------------------|------------------------|----------|-------------------|----------------|--------|------------|-----|---------------------|-----------|
|                            | Encodeur               | Décodeur | Encodeur-décodeur | Classification | Résumé | Traduction | Q/R | Génération de texte | Fill Mask |
| BERT [Devlin et al., 2020] | X                      |          |                   | Texte/Token    |        |            | X   |                     | X         |
| GPT [Radford et al., 2018] |                        | X        |                   | Texte/Token    | X      | X          | X   | X                   | X         |
| ALBERT [Lan et al., 2019]  | X                      |          |                   | Texte/Token    |        |            | X   |                     | X         |
| RoBERTa [Liu et al., 2019] | X                      |          |                   | Texte/Token    |        |            | X   |                     | X         |
| T5 [Raffel et al., ]       |                        |          | X                 | Texte/Token    | X      | X          | X   | X                   | X         |
| BART [Lewis et al., 2019]  |                        |          | X                 | Texte          | X      | X          | ~X  |                     | ~X        |
| XLNet [Yang et al., 2019]  |                        | X        |                   | Texte/Token    |        |            | X   |                     | X         |

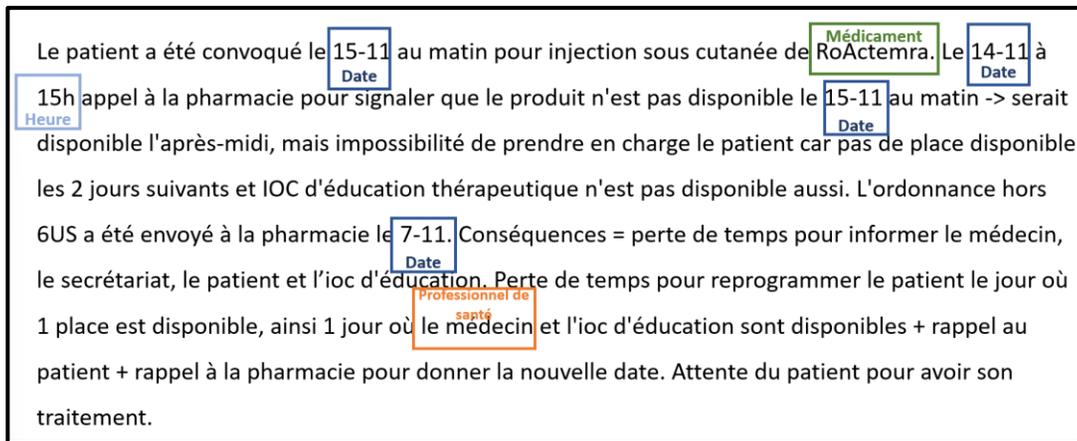
### 3 Méthodologie de traitement des déclarations des erreurs médicamenteuses et résultats obtenus

#### 3.1 Collecte des données

La collecte des données pour cette étude s'est appuyée sur des déclarations d'erreurs médicamenteuses provenant de différents établissements de santé. Ces déclarations ont été recueillies à l'aide de divers outils, y compris des fiches papier et des plateformes numériques telles que BlueKango<sup>1</sup>, Sesame<sup>2</sup>, et Ennov<sup>3</sup>. Chaque outil propose un guide de déclaration qui incite à répondre à des questions à choix multiples, ciblant les caractéristiques spécifiques de l'erreur médicamenteuse. Malgré la variété des formats, tous ces outils partagent un élément commun : ils permettent d'inclure une description textuelle libre de l'incident. Cette uniformité dans la collecte des données textuelles ouvre la voie à une analyse cohérente, malgré la diversité des sources et des formats de déclaration.

### 3.2 Résultats

Face au défi de traiter et d'analyser les vastes quantités de textes descriptifs libres issues des déclarations d'erreurs médicamenteuses, l'adoption d'une approche basée sur l'intelligence artificielle s'est avérée cruciale. En particulier, l'utilisation d'un modèle Transformer, tel que GPT-3.5 Turbo, a été identifiée comme la méthode la plus efficace pour extraire des informations pertinentes à partir de ces descriptions complexes sans annotation préalable. Dans notre étude, nous avons exploité le potentiel du modèle GPT-3.5 Turbo, avec sa capacité à comprendre et à générer du texte naturel, pour analyser des exemples de déclaration des EM et en extraire des informations clés, selon la reconnaissance d'entités nommées (NER). Cette approche permet d'identifier et de classer automatiquement les informations clés contenues dans les textes, telles que les types de médicaments impliqués, les étapes du processus de soin où l'erreur s'est produite, et les conséquences observées, selon les informations citées dans une déclaration. Cette exploration s'est déroulée selon deux phases distinctes :



Le patient a été convoqué le 15-11 au matin pour injection sous cutanée de RoActemra. Le 14-11 à 15h appel à la pharmacie pour signaler que le produit n'est pas disponible le 15-11 au matin -> serait disponible l'après-midi, mais impossibilité de prendre en charge le patient car pas de place disponible les 2 jours suivants et IOC d'éducation thérapeutique n'est pas disponible aussi. L'ordonnance hors 6US a été envoyé à la pharmacie le 7-11. Conséquences = perte de temps pour informer le médecin, le secrétariat, le patient et l'IOC d'éducation. Perte de temps pour reprogrammer le patient le jour où 1 place est disponible, ainsi 1 jour où le médecin et l'IOC d'éducation sont disponibles + rappel au patient + rappel à la pharmacie pour donner la nouvelle date. Attente du patient pour avoir son traitement.

Figure 2. Exemple d'extraction non dirigée des informations importantes dans une déclaration d'EM

#### Phase 1 : Extraction non supervisée

Lors de la première phase, nous avons utilisé le modèle pour effectuer une reconnaissance d'entités nommées sans spécifier au préalable les labels à rechercher. Le modèle a démontré une capacité remarquable à identifier plusieurs types d'entités au sein des textes, tels que la date, le nom du médicament, les conséquences, ainsi que les identités des professionnels de santé et/ou des patients impliqués. Il est important de noter que la nature et la quantité d'informations extraites variaient considérablement d'une déclaration à l'autre, reflétant la diversité et l'inconsistance des données fournies dans chaque rapport. La Figure 2 présente le résultat de l'extraction non dirigée sur un exemple de déclaration.

#### Phase 2 : Extraction supervisée

Lors la deuxième phase, nous avons affiné notre approche en spécifiant explicitement les labels recherchés : médicament, date, temps, professionnel de santé, patient, cause, et conséquence. Cette précision a servi de guide pour le modèle lors de l'extraction d'informations correspondant à ces catégories, améliorant ainsi la pertinence et la profondeur des données recueillies. Par exemple, dans le premier cas testé Figure 3.a, le modèle a réussi à identifier la date, le médicament, l'heure, les professionnels de santé impliqués, le patient, les conséquences, et même une autre personne impliquée. Cependant, dans un second cas Figure 3.b, où la déclaration était moins détaillée, seules les entités concernant les personnes impliquées (patient, professionnel de santé, et autre) et le médicament ont été reconnues, illustrant encore une fois la dépendance de l'efficacité de l'extraction aux informations présentes initialement dans les déclarations. Il est à noter que, malgré la présence des causes comme entités à extraire, le modèle n'en a extrait aucune car elles n'étaient pas précisées dans les déclarations. Cela souligne une limitation inhérente à l'approche qui dépend strictement des données d'entrée fournies.

Le **Patient** a été convoqué le **15-11** au **matin** pour injection sous cutanée de **Médicament RoActemra**. Le **14-11** à **15h** **Heure** appel à la pharmacie pour signaler que le produit n'est pas disponible le **15-11** au **matin** -> serait disponible l'après-midi, mais impossibilité de prendre en charge le patient car pas de place disponible les 2 jours suivants et **Professionnel de santé** **IOC d'éducation thérapeutique** n'est pas disponible aussi. L'ordonnance hors 6US a été envoyé à la pharmacie le **7-11**. **Conséquences = perte de temps** pour informer le médecin, **P. Impliquée** le secrétariat, le patient et l'ioc d'éducation. **Perte de temps** pour reprogrammer le patient le jour où 1 place est disponible, ainsi 1 jour où **Professionnel de santé** **le médecin** et l'ioc d'éducation sont disponibles + rappel au patient + rappel à la pharmacie pour donner la nouvelle date. Attente du patient pour avoir son traitement.

Figure 3.a Exemple 1 d'extraction dirigée

Lors de la répartition des traitements par **Professionnel de santé** **l'infirmier Diplômé d'État (IDE)** aidée par **Professionnel de santé** **Aide Soignant** (traitement vérifié, mis dans pots étiquetés au nom des résidents par IDE). L'IDE donne le pot de **traitement de Mme C.** a l'Aide Soignant **nommant son plateau repas** et le met sur le plateau de Mme N. **Médicament** **Les traitements (AVK)** ont été administré a **Patient** **Mme N.** par sa **P. Impliquée** **petite fille** qui après avoir administré un cp, réalise que le pot est marqué du nom de **Patient** **Mme C.** et le signale a l'IDE.

Figure 3.b Exemple 2 d'extraction dirigée

Figure 3. Extraction des informations caractéristiques de deux exemples de déclaration d'EM

Les résultats obtenus grâce à l'application de GPT-3.5 Turbo ont été significatifs. Le modèle a démontré sa capacité à discerner une variété de labels pertinents à partir des textes des déclarations, permettant ainsi un étiquetage fin des déclarations pouvant servir pour la classification des EM. Bien que l'exactitude ne soit pas parfaite à 100% et que les extractions nécessitent une vérification ou une complétion manuelle, le modèle permet d'établir une base solide pour l'étiquetage automatique des déclarations. Ces premiers succès indiquent que l'utilisation des modèles Transformers pour l'analyse du texte dans le cas des déclarations des EM est une solution prometteuse.

En automatisant l'extraction des informations cruciales, cette approche ouvre la voie à des interventions plus précises et fondées sur des données probantes pour prévenir les erreurs médicamenteuses dû à la préparation des déclarations des EM pour être exploitable. Ceci œuvre vers l'amélioration du processus de gestion des EM à l'hôpital.

## 4 Conclusion

En conclusion, cet article a mis en lumière la complexité et les enjeux associés à la gestion des erreurs médicamenteuses (EM) dans les établissements de santé, soulignant l'importance cruciale d'un traitement efficace et rapide des déclarations fournies par le personnel soignant. À travers l'examen du processus actuel, qui repose largement sur l'analyse de descriptions textuelles libres recueillies via divers outils, nous avons identifié le besoin pressant d'innovation pour surmonter les défis posés par le volume important de déclarations et les contraintes de temps.

L'automatisation du traitement de ces déclarations, grâce au traitement automatique du langage naturel (TAL) et spécifiquement à l'emploi de modèles Transformers comme GPT, apparaît comme une solution prometteuse. Cette approche offre la possibilité d'extraire rapidement les informations clés nécessaires à la priorisation et à la gestion des EM, et elle facilite également une structuration efficace des données, permettant une analyse plus approfondie et une meilleure compréhension des incidents rapportés.

## Remerciement

Nous tenons à remercier la région Occitanie et la société Atout Majeur Concept (AMC) qui nous accompagnent financièrement pour la réalisation de ces travaux de recherche. Nous tenons également à exprimer notre gratitude au Centre Hospitalier Intercommunal de Castres-Mazamet et plus particulièrement à Dr. Marie-Hélène CLEOSTRATE, Pharmacien hospitalier au CHIC et Chef du pôle « Médico-Technique », et à Dr. Marie-Noëlle CUIFI, Médecin gériatre et Chef du pôle « Médecine », pour leur soutien et implication dans ce projet de recherche.

## Références

- Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A. et al. (2023), 'Revolutionizing healthcare: the role of artificial intelligence in clinical practice', *BMC medical education* 23(1), 689.
- Antoun, W., Baly, F. & Hajj, H. (2020), 'Arabert : Transformer-based model for arabic language understanding'.
- Chowdhary, K. & Chowdhary, K. (2020), 'Natural language processing', *Fundamentals of artificial intelligence* pp. 603–649.
- Damiani, G., Altamura, G., Zedda, M., Nurchis, M. C., Aulino, G., Alizadeh, A. H., Cazzato, F., Della Morte, G., Caputo, M., Grassi, S. et al. (2023), 'Potentiality of algorithms and artificial intelligence adoption to improve medication management in primary care: a systematic review', *BMJ open* 13(3), e065301.
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018), 'Bert : Pre-training of deep bidirectional transformers for language understanding'.
- Frei, J. & Kramer, F. (2023), 'Annotated dataset creation through large language models for non-english medical nlp', *Journal of Biomedical Informatics* 145, 104478.
- Jarrar, M., Khalilia, M. & Ghanem, S. (2022), 'Wojood : Nested arabic named entity corpus and recognition using bert'.
- Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P. & Soricut, R. (2019), 'Albert : A lite bert for self-supervised learning of language representations'.
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V. & Zettlemoyer, L. (2019), 'Bart : Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension', *arXiv preprint arXiv :1910.13461*
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. & Stoyanov, V. (2019), 'Roberta : A robustly optimized bert pretraining approach'.
- Locke, S., Bashall, A., Al-Adely, S., Moore, J., Wilson, A. & Kitchen, G. B. (2021), 'Natural language processing in medicine : a review', *Trends in Anaesthesia and Critical Care* 38, 4–9.
- HAS (2022), 'Haute Autorité de Santé : Déclarer les événements indésirables graves associés aux soins (EIGS)'.
- Radford, A., Narasimhan, K., Salimans, T., Sutskever, I. et al. (2018), 'Improving language understanding by generative pre-training'.
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. & Liu, P. J. (2020), 'Exploring the limits of transfer learning with a unified text-to-text transformer', *Journal of machine learning research* 21(140), 1–67.
- Rahali, A. & Akhloufi, M. A. (2023), 'End-to-end transformer-based models in textual-based nlp', *AI* 4(1), 54–110.
- Schmitt, E., Antier, D., Bernheim, C., Dufay, E., Husson, M. C. & Tissot, E. (2006), 'Dictionnaire français de l'erreur médicamenteuse', *Société française de pharmacie clinique*.
- Subrahmanya, S. V. G., Shetty, D. K., Patil, V., Hameed, B. Z., Paul, R., Smriti, K., Naik, N. & Somani, B. K. (2022), 'The role of data science in healthcare advancements: applications, benefits, and future prospects', *Irish Journal of Medical Science (1971-)* 191(4), 1473–1483.
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R. & Le, Q. V. (2019), 'Xlnet : Generalized autoregressive pretraining for language understanding', *Advances in neural information processing systems* 32.