

## COMMUNIQUÉ DE PRESSE

Paris, le 25 mars 2026

### Publication d'un corpus inédit de comptes-rendus médicaux de patients fictifs en open data pour accélérer l'IA en santé

Dans un cadre réglementaire européen exigeant, garantissant un accès et un partage sécurisés des données de santé, le projet PARTAGES propose une réponse concrète aux défis qu'il implique. Mobilisant 32 partenaires français et coordonné par la Plateforme des données de santé (Health Data Hub), PARTAGES publie en open data un corpus inédit de plus de 6 000 comptes-rendus médicaux de patients fictifs. Rédigés et validés par des médecins, ces documents constituent une ressource stratégique pour entraîner et évaluer des outils d'intelligence artificielle (IA) au service des professionnels de santé.

Le développement d'outils d'intelligence artificielle appliqués aux comptes rendus médicaux s'inscrit dans un cadre structurant et protecteur. La réglementation encadre strictement l'accès aux données cliniques, garantissant leur usage sécurisé, tout en posant des **défis en matière de disponibilité de corpus ouverts et d'évaluation des modèles**. De plus, les systèmes entraînés sur ces données cliniques peuvent mémoriser des informations des données d'entraînement les rendant eux-mêmes sensibles, ce qui complique leur partage et contribue à un écosystème fragmenté.

Afin de répondre à ces enjeux, un appel a été diffusé auprès des associations d'internes et des centres hospitalo-universitaires partenaires, dans le cadre du projet PARTAGES, lancé en avril 2025. Il a permis de constituer un panel de **120 médecins** représentatif de la **diversité des spécialités médicales et des territoires** pour rédiger un corpus de compte-rendus de patients fictifs.

Fruit de cette mobilisation, le corpus réunit plus de **6 000 comptes-rendus médicaux de patients fictifs** rédigés par ces praticiens. Élaboré à partir de situations cliniques fictives, le corpus de comptes-rendus médicaux ne contient aucune donnée issue de patients réels. Il peut ainsi être mobilisé librement, **sans aucune contrainte réglementaire**, tout en conservant un haut niveau de réalisme et de pertinence clinique.

Par son ampleur et son niveau d'exigence méthodologique, il constitue une ressource inédite en France pour **l'entraînement, l'évaluation et la comparaison de modèles d'intelligence artificielle** au service de l'innovation et des professionnels de santé. Sa publication a été officiellement annoncée le 25 mars, à l'occasion du comité stratégique de mi-parcours du projet PARTAGES.

Ce livrable résulte d'un travail conjoint entre la Plateforme des données de santé et plusieurs partenaires du projet : l'**AP-HP**, le **LISN**<sup>1</sup> et le **LIMICS**<sup>2</sup>, et a fait l'objet d'une publication scientifique : [\*PARHAF, a human-authored corpus of clinical reports for fictitious patients in French.\*](#)

<sup>1</sup> Laboratoire interdisciplinaire des sciences du numérique (CNRS/Université Paris-Saclay)

<sup>2</sup> Laboratoire de recherche en informatique pour la santé (Inserm/Sorbonne Université/Université Sorbonne Paris-Nord)

## **PARTAGES : un projet national au service des communs numériques**

Lauréat de l'appel à projets "Communs numériques pour l'intelligence artificielle générative", PARTAGES bénéficie d'un soutien de l'État dans le cadre du **plan France 2030** opéré par Bpifrance. Le projet est doté d'un budget total de 9,4 millions d'euros pour une durée de deux ans (2025-2027).

Il réunit un consortium inédit de **32 partenaires**, composé d'équipes de recherche, de groupes d'établissements de santé publics et privés, ainsi que d'acteurs *deeptech* spécialisés en intelligence artificielle. Son ambition : structurer des ressources ouvertes pour favoriser l'émergence de solutions d'IA générative en santé, au bénéfice de l'écosystème académique, hospitalier et industriel.

En complément du corpus, PARTAGES met à disposition [plusieurs livrables](#) :

- un guide méthodologique encadrant la production et la relecture des comptes-rendus,
- des sous-ensembles annotés du corpus,
- quatre guides d'annotation dédiés aux cas d'usage du projet,
- un corpus de données médicales ouvertes (articles scientifiques, notices de médicaments, cas cliniques) ayant servi à l'entraînement de modèles de fondation.

## **Cap sur 2027 : modèles spécialisés et évaluation en conditions réelles**

Dans la continuité de ces travaux, le consortium de PARTAGES développera d'ici 2027 **sept modèles d'IA spécialisés** pour des cas d'usage à fort impact : codage d'informations médicales, résumés automatiques de comptes-rendus, détection automatique en infectiologie, analyse de la réponse aux traitements en oncologie, ... Les premiers modèles entraînés ainsi que leur protocole d'évaluation seront publiés prochainement sur le site de la Plateforme des données de santé.

En parallèle, une **plateforme nationale d'évaluation fédérée** est en cours de développement. Déployée dans 20 hôpitaux et entrepôts de données de santé hospitaliers, elle permettra d'évaluer des algorithmes sur des données réelles dans un cadre réglementaire sécurisé.

[En savoir plus sur PARTAGES](#)

### **À PROPOS DE LA PLATEFORME DES DONNÉES DE SANTÉ**

La Plateforme des données de santé (Health Data Hub) est un groupement d'intérêt public créé par la Loi du 24 juillet 2019 relative à l'organisation et la transformation du système de santé. Elle associe 56 parties prenantes, issues de la sphère publique (Etat, représentants de la recherche et de l'enseignement, les caisses et les organismes d'assurance maladie, établissements et professionnels de santé, les agences publiques) mais aussi les représentants des industriels, France Assos Santé... et met en œuvre les grandes orientations stratégiques relatives au Système National des Données de Santé fixées par l'Etat et notamment le ministère de la santé et de l'accès aux soins. C'est un service à destination de l'écosystème de santé, des acteurs à l'origine de la collecte de données, des porteurs de projets d'intérêt général et de la société civile. En ce sens, la PDS promeut l'innovation en santé et l'accessibilité des données et des connaissances par le biais, entre autres, d'événements fédérateurs comme l'organisation de data challenge et d'appels à projets.

**Contact presse :**  
[presse@health-data-hub.fr](mailto:presse@health-data-hub.fr) 06 29 55 33 71

Abonnez-vous à l'infolettre sur le [site internet](#)  
Suivez la PDS sur [LinkedIn](#), [X](#) et [Bluesky](#)