

Question du mois de SEPTEMBRE 2025

« BIG DATA » : Quel intérêt d'un entrepôt de données en médecine d'Urgence ?



Credit photo : <https://www.robots-et-compagnie.com/intelligence-artificielle-bientot-plus-forte-homme/>

Auteurs Violeau M¹, Forgeot C¹, Maillard L¹ .

¹ Observatoire régionale des urgences de Nouvelle aquitaine (ORUNA, 23 Quai de Paludate, 33800 Bordeaux

Mais pourquoi cette question ? Quel intérêt pour moi clinicien ? Pourquoi dois-je m'intéresser aux données médicales ? Je ne suis ni statisticien, ni médecin en santé publique ni informaticien.. Laissez-moi donc tranquille au lit du malade avec ma sonde d'échographie. Et bien non !! Que tu sois médecin au lit du malade, urgentiste en ville ou en rase campagne, médecin en cabinet libéral, cadre, infirmier, assistant de régulation, étudiant, interne, chirurgien ... Tu devrais t'y intéresser ... Peut-être même sans le savoir, tu génères tous les jours, tout un arsenal de données médicales qui sont ensuite utilisées pour développer les connaissances, améliorer les prises en charge ou optimiser les organisations. Or, pour faire ça, on a besoin de toutes les données mais aussi des opérateurs qui les collectent et en permettent l'accès. Ces opérateurs ont des profils différents et les modalités d'accès sont assez complexes. Ainsi, exploiter ces données médicales multiples revient parfois à un sacerdoce. C'est là que les entrepôts de données de santé entrent en jeu....



Commençons par la base , c'est quoi une donnée de santé ?

Les données de santé sont des informations relatives à la santé physique ou mentale, passée, présente ou future, d'une personne physique, d'un patient ou d'un malade qui révèlent des informations sur l'état de santé de cette personne. Ces données sont particulièrement sensibles et protégées par différentes instances et réglementations (règlement européen sur la protection des données personnelles, loi Informatique et Libertés, code de la santé publique, etc.) afin de garantir le respect de la vie privée des personnes. La CNIL identifie 3 catégories de données [1]:

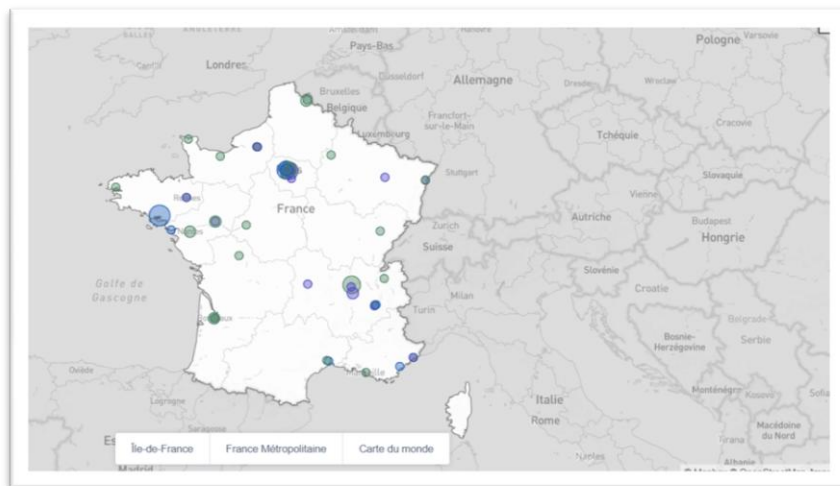
1/ par leur nature : antécédents médicaux, maladies, prestations de soins réalisés, résultats d'examens, traitements, handicap, ...

2/ par croisement avec d'autres données, deviennent des données de santé en ce qu'elles permettent de tirer une conclusion sur l'état de santé ou le risque pour la santé d'une personne : croisement d'une mesure de poids avec d'autres données (nombre de pas, mesure des apports caloriques...), croisement de la tension avec la mesure de l'effort, etc....

3/ celles qui deviennent des données de santé en raison de leur destination, c'est-à-dire de l'utilisation qui en est faite au plan médical.

Pourquoi mettre en place des entrepôts de données dans le paysage de la santé ?

Un **entrepôt de données de santé (EDS)** est un système de stockage numérique centralisé qui permet de **collecter**, **intégrer** et **organiser** de grandes quantités de données en santé à plus ou moins long terme provenant de diverses sources internes et externes (dossiers patients, imagerie, biologie, prescriptions, données de capteurs. L'objectif est de permettre l'exploitation secondaire **des données (recherche, épidémiologie, qualité)** permettant ainsi aux utilisateurs de prendre des décisions éclairées et stratégiques selon une réglementation bien définie [2] Contrairement aux bases de données opérationnelles utilisées pour les transactions en temps réel, un entrepôt de données est spécialement conçu pour **stocker** des données historiques et actuelles dans un format cohérent, **optimisé** pour les requêtes analytiques et les agrégations. L'AFM téléthon a par exemple créé un EDS en 2023 pour accélérer les connaissances en matière de pathologies neuromusculaires mais surtout favoriser le déploiement de nouveaux traitements. Depuis 2017, 100 entrepôts publics (comme des centres hospitaliers, des instituts de recherche, etc.) et privés (associations, centres de lutte contre le cancer, etc) et 22 du secteur privé (entreprises, cliniques, etc.) ont été créés et mis en œuvre par 88 acteurs différents. Ces derniers peuvent gérer un ou plusieurs entrepôts, voire se mettre en groupement pour gérer un même entrepôt [3].



Crédit photo : d'après la CNIL

Répartition des entrepôts de santé (EDS) et des acteurs responsables sur le territoire français [3]

Quelle différence entre un entrepôt de données, un collecteur de données et un hébergeur de données ?

L'**entrepôt** est une infrastructure logicielle et organisationnelle où l'on centralise d'importantes quantités de données issues de nombreuses sources. Le **collecteur** de données c'est un dispositif ou un service qui capte recueille et transmet les données pour faire remonter les informations depuis les patients, les professionnels ou les systèmes (ça ne vous dit rien le Dossier Médical Partagé (DMP), lancé en mars 2024...), capteurs, objets connectés, logiciels. Il ne stocke pas mais c'est la porte d'entrée des données. L'**hébergeur** appelé Hébergeur des données en santé (HDS) c'est une société, un prestataire certifié qui assure la conservation, la sécurité et à la disponibilité des données

Pourquoi est-ce un vrai enjeu pour le domaine de la santé ? Particulièrement aux urgences ?

Les données médicales (ou données en santé) sont devenues un enjeu majeur et particulièrement sensibles car elles représentent à la fois la richesse scientifique, un levier d'innovation et un sujet à haute valeur ajoutée d'un domaine, d'un pays, d'une entreprise, et donc d'un Hôpital. Ces données médicales sont particulièrement stratégiques pour plusieurs raisons dans le domaine de la santé et donc de l'Urgence :



- Pour améliorer la **prise en charge des patients** (suivi personnalisé, coordination multi-professionnelles, traçabilité, évènements indésirables ou erreurs médicales)
- Pour accélérer **la recherche et l'innovation** ainsi favoriser les grandes cohortes ou les essais multicentriques, identifier les patients éligibles
- Pour développer plus efficacement le lancement **de nouvelles molécules ou nouveaux traitements**, de nouveaux protocoles, de nouveaux examens
- Pour entraîner **de nouveaux modèles algorithmiques numériques** (l'IA vous connaissez...) ainsi détecter plus tôt, plus efficacement certains marqueurs ou signaux faibles de pathologies du XXI ème siècle (cancer, maladies cardio-vasculaires, maladies rares)
- Pour **organiser, piloter, gérer une filière** ou un réseau : par exemple pour anticiper le besoin de lits (urgences, réanimation, épidémies, ...)
- Pour mieux adapter **les politiques publiques et les stratégies d'orientation budgétaire**
- Pour mesurer **la qualité des soins** et favoriser l'amélioration continue

A l'échelle d'un service d'Urgence, c'est aussi valoriser la dynamique d'une équipe, mieux organiser et gérer son service, comprendre les dysfonctionnements, fixer des objectifs, favoriser le développement de projets, optimiser les prises en charge, améliorer les conditions de travail ... (dans une entreprise, on pourrait parler d'augmentation de la productivité en optimisant les ressources ☺), assurer la qualité des soins et prévenir les risques, les défaillances mais aussi comprendre les flux et anticiper la réponse.

Avec la numérisation de notre société, notre dépendance à Internet et aux objets connectés, la donnée en santé devient précieuse et en plus, elle est rare. Et comme la ruée vers l'or au XIXème siècle, les Etats Unis ont un coup d'avance avec les GAFAM qui voient ce nouvel or numérique comme un vrai business et le rêve d'un nouvel eldorado [4]. ... **D'ailleurs un pays qui ne maîtrise pas sa donnée perd en souveraineté et en compétitivité pour développer sa stratégie. Il s'agit donc d'un vrai enjeu politique et stratégique ...**



Crédit photo: Getty Images / Hulton Archive

Avons-nous un retard dans l'exploitation des données en santé en France ?

*Vous avez sûrement entendu parler de la création le 1er juillet 2025 d'un comité stratégique exceptionnel dans le cadre de la stratégie nationale en matière d'intelligence artificielle et d'utilisation secondaire des données de santé. Face à l'essor mondial de l'intelligence artificielle (IA) et la nécessité de renforcer la souveraineté sanitaire et numérique de la France, le gouvernement vient de publier sa **stratégie nationale « Intelligence artificielle et données de santé » pour la période 2025-2028**. Ce plan vise à tirer le plein potentiel du patrimoine unique que constituent les données de santé françaises tout en garantissant la confiance des citoyens et la sécurité de leurs informations médicales. La stratégie comprend deux volets indissociables. D'une part, construire un patrimoine national solide de données de santé, et d'autre part, développer une IA de confiance au service de la médecine. [5]*

Mais l'histoire des données en santé ne date pas d'hier, c'est en 2016, que le sujet suscite un regain d'intérêt avec la création d'un Système National des Données de Santé (SNDS), une infrastructure unique au monde, qui centralise les données en santé pour la recherche, enrichie depuis 2019 grâce à la Plateforme des Données de Santé (Health Data Hub) pour faciliter le partage et l'accès.

2020-2021, vous vous en souvenez ? ...coup de projecteur dans la sphère publique avec le COVID qui a permis de souligner l'importance de la donnée pour surveiller, modéliser, prévenir et traiter les populations[6]. La Haute Autorité de santé (HAS) s'empare du dossier et lance sa stratégie nationale 2021-2024 pour structurer les EDS, mieux exploiter la data et renforcer l'interopérabilité [7]. Depuis 2016, le cadre réglementaire avait déjà été posé par l'Europe. En France, ça s'appelle le RGPD (règlement général sur la protection des données). En 2022, l'Europe va plus loin et pose les esquisses



de l'European Health Data Space (EHDS) pour le partage de données entre les pays membres de l'UE avec une entrée en vigueur en mars 2025 et mise en application en 2029 (presque maintenant donc !). Ça paraît complètement fou, mais c'est maintenant !' Bref, c'est un nouveau niveau débloqué pour l'exploitation des données de santé et croyez-nous, les chercheurs sont dans les starting blocks ! [8]

Finalelement, comment ça marche aux Urgences ?

Si tout ça reste encore un peu loin de vous pour le moment, nous allons essayer de le rendre plus concret dans ce paragraphe.

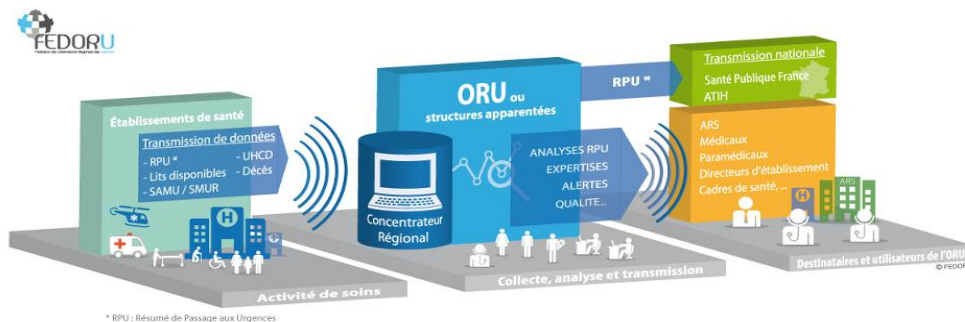
*Vous le savez sûrement, les données que vous renseignez chaque jour dans vos services d'urgences alimentent les **résumés de passage aux urgences (RPU)**. Un passage génère un résumé. Ces données sont envoyées chaque jour par votre établissement vers un collecteur régional. Cette transmission est rendue obligatoire par l'Arrêté du 24 juillet 2013, relatif au recueil et au traitement des données d'activité médicale produites par les établissements de santé publics ou privés ayant une activité de médecine d'urgence (...) [9], qui fixe les destinataires ainsi que le format des données transmises.*

Ces collecteurs régionaux (on parle aussi de concentrateurs), selon les régions sont gérés par les GRADeS (Groupement régional d'appui à la e-Santé) ou les ORU (Observatoires régionaux des urgences), mais toujours sous l'égide des ARS. Les données collectées, elles, sont exploitées par les ORU et restituée aux ARS et aux établissements pour le pilotage opérationnel, la définition des politiques publiques à l'échelle régionale et territoriale. Elles peuvent parfois être également mobilisées par les établissements, dans le cadre de projet d'études ou projets plus spécifiques (à condition d'avoir les autorisations CNIL *ad hoc* – RGPD, vous vous souvenez ?).

Depuis ces concentrateurs, elles sont également transmises à 2 agences nationales : mensuellement à l'Agence technique de l'information hospitalière (ATIH) et quotidiennement, à Santé publique France, qui disposent toutes deux de leur propre base nationale, exploitée à des fins médico-économiques pour la 1^{ère}, et de veille et surveillance sanitaire pour la 2^{nde} d'où la nécessité de disposer de données précoces. Bien que ces données ne soient pas disponibles en open source (Eh oui ! Elles contiennent des données individuelles donc sensibles ! RGPD encore !), elles peuvent être mises à disposition par l'un de ces deux organismes, sur demande, dans le cadre d'un projet précis accompagné de l'encadrement CNIL adéquat (RGP... vous avez compris). Si ça vous intéresse :

- Plateforme Syrius de l'ATIH
- [\[Accès aux documents administratifs et aux données scientifiques | Santé publique France\]](#)

Evidemment, les données d'urgences ne se résument pas aux données RPU ! Elles comprennent également, les données enregistrées dans le cadre de l'activité pré-hospitalière, par les centres de régulation des SAMU (Résumé de régulation - RDR) et dans le cadre des sorties SMUR (Résumé patient d'intervention SMUR – RPIS). Cependant, à l'heure actuelle, ces dernières ne font l'objet d'aucune réglementation nationale. Ainsi, si certaines régions sont en avance de phase avec une remontée de données pré-hospitalières fonctionnelle, dans d'autres, les formats de données pré-hospitalières restent très hétérogènes et aucune remontée d'information n'est organisée. Parmi les autres données d'urgences, on peut également citer les données relatives aux disponibilités en lits, les tensions en établissements de santé etc.



Crédit photo: FEDORU

Schématisation des processus liés à la collecte et à la transmission des données

Les ORU : des acteurs clés dans le champ des données d'urgence...

Alors ? Les ORU, c'est quoi ? Les ORU sont les observatoires régionaux des urgences. Comme leur nom l'indique, ils sont positionnés à l'échelle régionale, auprès des ARS et travaillent à la collecte des données d'urgences et à leur exploitation. Les équipes qui les composent sont variables d'un ORU à l'autre mais comprennent systématiquement des médecins urgentistes (qui poursuivent leur activité clinique) et des profils DATA (statisticien, data scientists...). Les ORU sont donc détenteurs d'une expertise technique et surtout (!) d'une expertise métier et de terrain permettant une exploitation juste et pertinente des données. Au-delà des analyses, les ORU ont également à charge la restitution d'information sur l'activité des urgences aux établissements, au moyen de bulletins, tableaux de bord voire d'outils BI permettant la visualisation d'indicateurs de l'activité des urgences en temps réel. Ils travaillent également au plus près avec les établissements à la qualité de la donnée et en particulier, l'exhaustivité et la qualité du codage des données RPU. En effet, un codage de qualité est un des



ingrédients indispensables pour produire des analyses... de qualité. C'est là que vous entrez en jeu ! Ce n'est pas toujours évident de coder correctement. On n'a pas toujours le temps de prendre le temps. On ne sait pas toujours quoi coder (combien de codes déjà dans la CIM10 ?). L'interface n'est pas ergonomique et peut être même carrément rebutant ! Mais...on a besoin de vous !! Vous savez à présent que ces données vont être utilisées, notamment par d'autres médecins urgentistes, pour tout un tas de travaux, qui ont pour objectifs d'améliorer la prise en charges aux urgences, l'organisation des services, etc.

Si maintenant, on reprend un peu de hauteur, les ORU sont des opérateurs techniques et métier qui collectent, exploitent les données d'urgences et font en sorte de les améliorer voir de les faire évoluer. Pour mutualiser les compétences, les idées, capitaliser sur les projets et porter leur voix au niveau national, ils se sont constitués en Fédération : la Fedoru (Fédération... ORU, vous l'avez ?) [10].

C'est bien beau tout ça mais en 2025, bien qu'il y ait un ORU dans quasiment toutes les régions métropolitaines, la pluralité des opérateurs complexifie l'accès à ces dernières et constitue bien souvent une limite dans la capacité à analyser. Dans ce contexte, les entrepôts de données de santé pour la médecine d'urgence pourraient être le point de convergence de l'ensemble de ces données. Leur mise en œuvre permettrait un regroupement de ces dernières au sein d'un opérateur régional unique, un encadrement des remontées et une uniformisation des formats.

Les ORU, appuyés par les GRADeS, pourraient offrir le socle technique et d'expertise nécessaire à la mise en place de telles infrastructures et à une exploitation croisée et pertinente des données disponibles au sein de ses EDS. Ils offriraient ainsi, un accès simultané à une grande variété de données et, en les croisant, permettraient d'affiner les analyses, d'approfondir les capacités en termes de recherche, de travailler sur les parcours de soins dans leur globalité, voire de rendre possible le chainage de données qui reste, encore aujourd'hui un véritable obstacle. Ils permettraient de servir des objectifs plus ambitieux. *Une application qui vous dirait, en temps réel vers quel hôpital vous devez aller pour le patient grand brûlé pris en charge en SMUR ? Quel est le service le plus adapté pour ce patient polypathologique ? Comment améliorer les délais de prise en charge ? Combien de ressources sont nécessaires pour faire face à l'afflux de l'hiver prochain ou face à un événement imprévu dans votre service, ça vous dirait ?*



Enfin, avec l'avènement de l'IA, les perspectives en termes de collecte de données évoluent. Ainsi, demain, le périmètre des EDS pourrait s'élargir à l'ensemble des informations textuelles des dossiers patients, et notamment les observations médicales, ouvrant ainsi de nouveaux champs d'exploration tels que : l'analyse des filières et parcours de soins, l'étude des problématiques d'accidentologie (domestique, voie publique, etc.), le suivi et l'anticipation des tensions structurelles...

Conclusion

Les EDS ne sont pas une finalité en soi : ce sont des outils. Pour les données de médecine d'urgence, leur utilité dépendra évidemment de la qualité de ces données et de la pertinence des travaux menés à partir d'elles. De nombreux points devront être considérés pour la mise en place de ces EDS : gouvernance, environnement technique, conformité RGPD, bénéficiaires des données... Ils devront être pensés dès le départ, avec la conscience de leurs limites, et encadrés par des professionnels disposant de compétences techniques scientifiques adaptées, notamment au sein des ORU et des sociétés savantes comme la SFMU.



Référentiels bibliographiques

[1] La Commission Nationale de l'Informatique et des Libertés (CNIL) (2018) Qu'est-ce qu'une donnée de santé ? <https://www.cnil.fr/fr/quest-ce-que-une-donnee-de-sante> (Dernier accès le 25 septembre 2025)

[2] République Française (2019), Arrêté du 29 novembre 2019 portant approbation d'un avenant à la convention constitutive du groupement d'intérêt public « Institut national des données de santé » portant création du groupement d'intérêt public « Plateforme des données de santé ». JORF n°0278



du 30 novembre 2019. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000039433105> (dernier accès le 25 septembre 2025)

[3] La Commission Nationale de l'Informatique et des Libertés (CNIL) (2024) Cartographie des entrepôts de données de santé et des acteurs responsables des EDS en France. <https://carto-eds.beta.cnil.fr/> (dernier accès le 25/09/2025)

[4] Mandon R, Bellit S (2021) Vos données valent-elles de l'or ? L'Internet industriel des objets à l'épreuve du réel. <https://www.la-fabrique.fr/fr/publication/vos-donnees-valent-elles-de-lor-linternet-industriel-des-objets/> (dernier accès le 1/10/2025)

[5] Ministère de la santé (2025) Stratégie intelligence artificielle et données de santé. <https://sante.gouv.fr/actualites-presse/presse/communiqués-de-presse/article/intelligence-artificielle-et-donnees-de-sante-une-strategie-nationale-pour> (dernier accès le 1/10/2025)

[6] Zins M, Cuggia M, Goldberg M (2010) Les données de santé en France - Abondantes mais complexes. *Med Sci* 37:179-184

[7] Haute Autorité de santé (2021) Stratégie de données de santé 2021-2024. https://www.has-sante.fr/jcms/p_3272150/fr/donnees-en-sante-et-social-savoir-les-exploiter-pour-mieux-suivre-les-pratiques (dernier accès le 1/10/2025)

[8] European Association for Data and Cloud AISBL (2025) <https://gaia-x.eu/> (dernier accès le 1/10/2025)

[9] République Française (2013) Arrêté du 24 juillet 2013 relatif au recueil et au traitement des données d'activité médicale produites par les établissements de santé publics ou privés ayant une activité de médecine d'urgence et à la transmission d'informations issues de ce traitement dans les conditions définies à l'article L. 6113-8 du code de la santé publique et dans un but de veille et de sécurité sanitaires. JORF n°0185 du 10 août 2013.

<https://www.legifrance.gouv.fr/loda/id/JORFTEXT000027825549> (dernier accès le 25/09/2025)

[10] Fedoru (2013) Les Observatoires Régionaux des Urgences (ORU) ou structures apparentées <https://www.oruna.fr/> (dernier accès 25/09/2025)

Contact:

Docteur Violeau Mathieu

40 Avenue Charles de Gaulle 79021 Niort Cedex, France

mathieu.violeau@oruna.fr / innovation@sfmu.org