Analyse d'impact du Projet "Dante"

(Données d'Accès précoces Normalisées, suivies par Traitement automatisé du langage en vue de leur Evaluation)

Version du 20 mai 2025, réalisée avec le logiciel « PIA » de la Commission Nationale de l'Informatique et des Libertés (CNIL).

Historique des versions et évolutions

Version 1.0

- Conception et évaluation : Elisabete de Carvalho
- Validation : Marco Fiorini
- Envoi à la CNIL le 30 mai 2025
- Dépôt le 16 septembre 2025

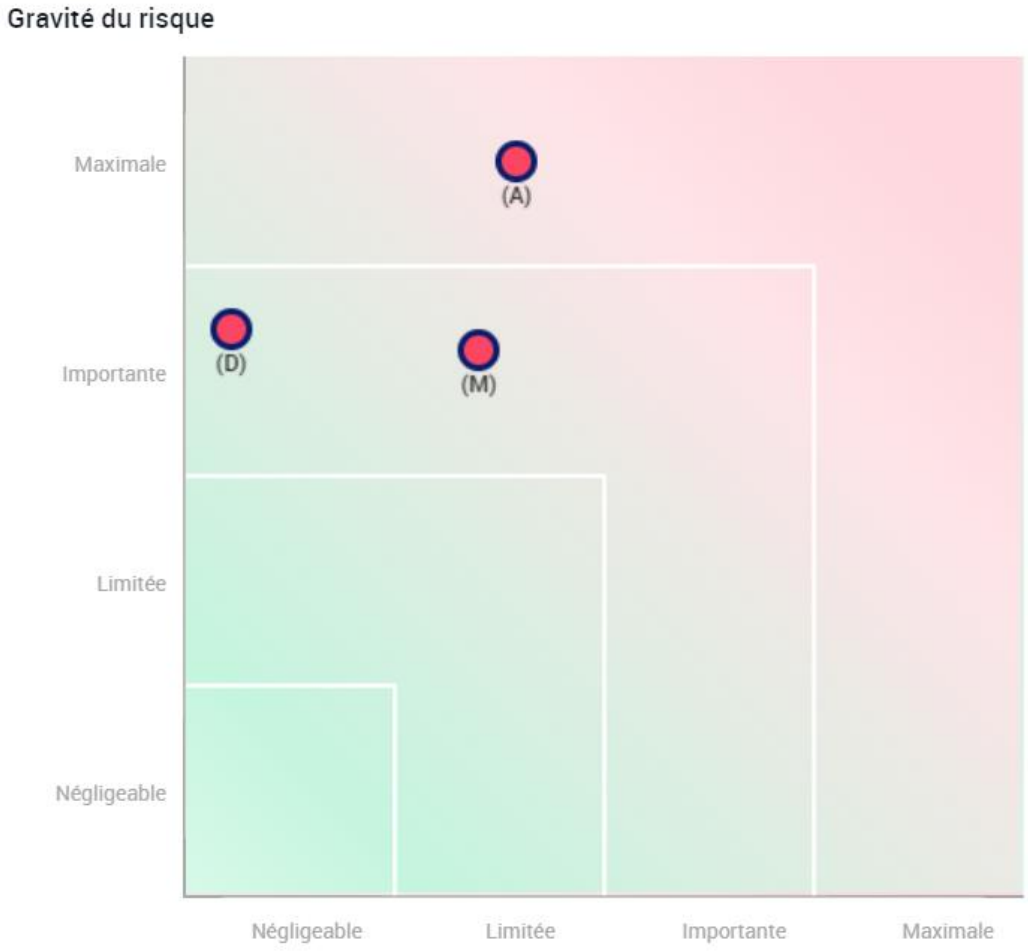
Version 1.1

- Conception et évaluation : Anne Haziza
    - Modifications :
        - 1. Base légale (art. 6) — Remplacement de l'article 6.1.e (mission d'intérêt public) par l'article 6.1.f (intérêt légitime), avec ajout du test de mise en balance correspondant.
        - 2. Portabilité — Retrait du droit à la portabilité (art. 20), inapplicable sous intérêt légitime comme il l'était sous mission d'intérêt public.
        - 3. Clarification des deux phases de l'étude : la phase 1 (constitution du bras IA par extraction automatisée des DPI, sans intervention humaine) relève de la méthodologie de référence MR-004 ; la phase 2 (constitution du bras de référence par un personnel de recherche accédant au DPI source, sur site ou à distance) relève d'une autorisation spécifique de la CNIL.
- Validation : Marco Fiorini
- Envoi projeté à la CNIL le 22 juin 2026
- Dépôt projeté le 1er juillet 2026

Table des matières

1 Vision Synthétique

1.1 Cartographie des risques



- Mesures prévues ou existantes
- Avec les mesures correctives mises en oeuvre
- (A)ccès illégitime à des données
- (M)odification non désirée de données
- (D)isparition de données

09/05/2025

2 Validation

2.1 Avis du DPO et des personnes concernées

2.1.1 Nom du DPO

Marco Fiorini

2.1.2 Avis du DPO

Le traitement pourrait être mis en œuvre.

2.1.3 Opinion détaillée du DPO

Au regard des éléments fournis, le traitement envisagé dans le cadre du projet « DANTE » peut recevoir un avis favorable à sa mise en œuvre, dans les conditions suivantes :

- Conformité au RGPD : Le contrat intègre les **clauses contractuelles types prévues par la décision d'exécution (UE) 2021/915**, garantissant le respect des articles 28 et suivants du Règlement Général sur la Protection des Données (RGPD).

- Finalités légitimes et définies : **Le traitement est mis en place à des fins scientifiques et méthodologiques clairement identifiées (comparaison de méthodes de collecte de données)**, dans un cadre limité et proportionné.

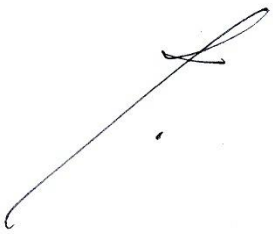
- Données pseudonymisées et sécurisées : Le traitement prévoit la **pseudonymisation des données de santé, leur hébergement exclusivement en France auprès de prestataires certifiés HDS**, ainsi que des mesures robustes de sécurité technique et organisationnelle (chiffrement, contrôle d'accès, etc.).

- **Encadrement contractuel des sous-traitants** : Les obligations des sous-traitants sont détaillées, les responsabilités bien réparties, et un droit de regard du Responsable de traitement est prévu sur les éventuels sous-traitants ultérieurs.

- **Droits des personnes et transparence** : Une procédure de coopération avec le DPO est prévue, ainsi que les modalités d'exercice des droits des personnes concernées.

En conséquence, le projet tel que cadré contractuellement présente des garanties suffisantes pour le respect des droits et libertés des personnes concernées, conformément au RGPD. Cet avis favorable est toutefois assorti d'une condition : au regard de la diversité des solutions effectivement opérationnelles sur le terrain, la phase 2 (création d'un bras de référence), qui suppose la consultation du DPI source par un personnel de recherche, ne pourra être mise en œuvre qu'après obtention de l'autorisation de la CNIL propre à cette phase, la phase 1 (bras IA) demeurant pour sa part couverte par la MR-004.

Le DPO



2.1.4 Recherche de l'avis des personnes concernées

L'avis des personnes concernées n'a pas été sollicité en amont, conformément aux principes de conformité « by design » à la Méthodologie de Référence MR-004 de la CNIL.

En effet, le projet « DANTE » (Données d'Accès précoces Normalisées, suivies par Traitement automatisé du langage en vue de leur Evaluation) s'inscrit dans un cadre de recherche n'impliquant pas le recueil du consentement, mais reposant sur une base légale d'intérêt légitime (article 6.1.f du RGPD), avec des traitements de données de santé pseudonymisées, proportionnés et réalisés sous la responsabilité d'un promoteur institutionnel respectant les obligations de sécurité, d'information et de documentation prévues par la MR-004.

3 Contexte

3.1 Vue d'ensemble

3.1.1 Quel est le traitement qui fait l'objet de l'étude ?

Le projet DANTE est une étude à visée méthodologique portée par la Filière IA & Cancers. Cette association à but non lucratif regroupe des partenaires publics (Institut National du Cancer, Health Data Hub, Agence de l'Innovation en Santé) et privés (industriels et association d'acteurs de santé). Elle est le responsable du traitement des données de cette étude.

L'étude est menée pour mieux comprendre comment améliorer la collecte des données médicales liées aux traitements innovants reçus en accès précoce. Actuellement, ces données sont souvent saisies manuellement par les professionnels de santé, ce qui prend du temps, peut générer des erreurs ou des oublis, et aboutit à des dossiers incomplets.

Le traitement compare cette méthode traditionnelle avec une méthode automatisée utilisant des technologies d'intelligence artificielle (IA), capables de lire les informations présentes dans les comptes rendus médicaux et de les structurer automatiquement. Le but de l'étude est de vérifier si cette nouvelle méthode permettrait de produire des données plus fiables, plus complètes et plus faciles à exploiter, tout en allégeant le travail des équipes médicales.

L'étude se déroule en deux phases, encadrées par deux régimes juridiques distincts. La phase 1 constitue le « bras IA » : les données sont extraites automatiquement des dossiers patients informatisés (DPI) par une solution d'IA, sans intervention humaine et avec pseudonymisation locale avant tout transfert ; cette phase relève de la méthodologie de référence MR-004 de la CNIL. La phase 2 constitue le « bras de référence » (gold standard) : un personnel de recherche (attachés de recherche clinique) saisit les mêmes variables en consultant directement le DPI source, sur site ou à distance ; cet accès humain à des données potentiellement identifiantes ne relève

pas de la MR-004 et fait l'objet d'une demande d'autorisation spécifique auprès de la CNIL. La phase 2 ne peut débuter avant l'obtention de cette autorisation.

3.1.2 Quelles sont les responsabilités liées au traitement ?

Dans le cadre de l'étude DANTE, les responsabilités de traitement des données personnelles peuvent être synthétisées comme suit :

3.1.2.1 Responsable de traitement

La Filière IA & Cancers, en tant qu'initiateur et pilote de l'étude, agit en tant que responsable de traitement pour la finalité de comparaison des méthodes de collecte (manuelle vs automatisée) dans les accès précoces. Elle détermine les finalités de traitement (comparaison, qualité des données, impact des méthodes) et s'assure de la conformité globale au RGPD.

3.1.2.2 Sous-traitants ou mandataires du responsable

Plusieurs entités agissent pour le compte de la Filière, avec des missions spécifiques :

- Les centres de soins participants en direct (CHU, CLCC, CHG, ESPIC…) et des PME, partenaires innovateurs IA. Tous interviennent pour collecter automatiquement les données médicales, à partir de documents source, via des solutions de traitement automatisé du langage (TAL).

- Ces partenaires doivent :
  - effectuer la pseudonymisation des données avant leur transmission,
  - gérer une clé de correspondance pseudonyme dans des conditions sécurisées.

3.1.2.3 Tiers de confiance

Un tiers de confiance indépendant, l'entreprise RCTs, reçoit les données pseudonymisées issues des collectes automatisées et manuelles. Il constitue également le bras de référence (phase 2) : il fait saisir les variables de l'étude par des attachés de recherche clinique accédant directement aux DPI, sur site ou à distance, cet accès étant subordonné à l'autorisation de la CNIL propre à cette phase.

Il est chargé :

- de fusionner les données,
- de les analyser conformément au protocole scientifique,
- de documenter la comparaison des méthodes selon des indicateurs de qualité, complétude et fiabilité pur aboutir à un « score d'exactitude ».

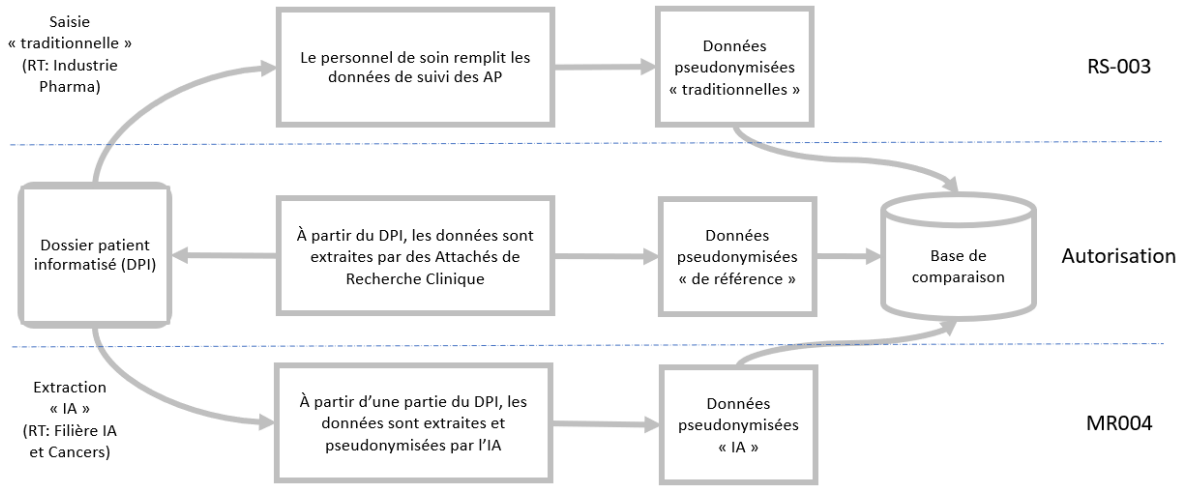
3.1.3 Quels sont les référentiels applicables ?

La collecte « manuelle » des données nécessaires à l'étude est encadrée par la **référence méthodologique RS-003**, applicable aux traitements réalisés dans le cadre des accès précoces. Cette partie est sous la responsabilité de traitement des diverses entreprises pharmaceutiques qui veulent bien participer à Dante à travers les accès précoces. Pour bien comprendre cette répartition, il faut distinguer deux choses : cette collecte « manuelle » n'est pas créée par DANTE. Elle correspond au recueil de données que chaque laboratoire doit déjà réaliser, indépendamment de l'étude, au titre de ses obligations réglementaires d'accès précoce — via le protocole d'utilisation thérapeutique et de recueil de données (PUT-RD) défini par les autorités de santé. DANTE ne fait que réutiliser ces données déjà existantes, sans modifier la façon dont elles sont saisies. C'est la raison pour laquelle chaque laboratoire demeure responsable de traitement pour la base « manuelle » qui le concerne, la Filière IA & Cancers n'intervenant aucunement dans cette saisie.

Leur **réutilisation à des fins de recherche et d'analyse**, à des fins de recherche scientifique d'intérêt général, est quant à elle réalisée conformément à la **méthodologie de référence MR-004** de la CNIL pour la phase 1 (bras IA), qui encadre l'usage secondaire de données de santé sans intervention directe auprès des personnes ni consultation humaine des documents sources.

En revanche, la phase 2 (bras de référence) implique la consultation directe du DPI source par un personnel de recherche, sur site ou à distance, et donc un accès humain à des données potentiellement identifiantes. Ce traitement sort du champ de la MR-004 ; il fait l'objet d'une demande d'autorisation auprès de la CNIL, instruite par le Health Data Hub et soumise à l'avis préalable du CESREES. La base de licité demeure l'article 6(1)(f) du RGPD et la condition de l'article 9(2)(j) ; seule la formalité préalable change (autorisation au lieu d'engagement de conformité à une méthodologie de référence). La phase 2 ne pourra démarrer qu'après obtention de cette autorisation.

Voici un schéma d'ensemble :



3.2 Données, processus et supports

3.2.1 Quelles sont les données traitées ?

Les données utilisées dans cette étude sont issues de la prise en charge des patients sous accès précoces.

Elles permettent de retracer les différentes étapes du suivi : la demande d'accès au traitement, son instauration, les visites de suivi, ainsi que l'arrêt éventuel du traitement.

Ces données concernent uniquement les éléments nécessaires à l'analyse. Elles sont regroupées en plusieurs catégories détaillées ci-dessous :

- **Données démographiques et type d'établissement prescripteur** ; ces informations permettent de caractériser le profil des patients inclus dans l'étude :
    - Date de naissance (mois et année uniquement, au format MM/AAAA)
    - Sexe
    - Catégorie de l'établissement ayant prescrit ou initié le traitement (par exemple : hôpital universitaire (CHU), centre privé type (ESPIC), Centre de Lutte contre le Cancer (CLCC)...
- **Données cliniques au moment de la demande** ; ces données permettent d'évaluer l'état général du patient au moment de la demande d'accès au traitement :
    - Statut « ECOG », (i.e. un indicateur couramment utilisé en oncologie pour mesurer l'autonomie du patient).
- **Données liées à la première administration du traitement** : ces informations retracent les étapes initiales de la prise en charge :
    - Date de la demande d'accès au traitement
    - Date de la première administration ou d'instauration du traitement
    - Détails de la posologie initiale :
    - Dose administrée
    - Voie d'administration (par exemple orale, intraveineuse)
    - Effets indésirables observés ou situations particulières (si renseignés par les soignants)
- **Données issues du suivi du traitement** (visites successives) : le traitement et son suivi pouvant s'étendre sur plusieurs semaines ou mois, plusieurs occurrences de visites sont possibles. Pour chaque visite, les données suivantes peuvent être collectées :
    - Date de la visite
    - Posologie actualisée : Dose

- Voie d'administration
- Fréquence d'administration (nombre de prises par semaine ou par jour)
- Fréquence exprimée en jours ou en semaines
- Évaluation de la qualité de vie, si disponible, à l'aide du questionnaire standardisé publié par l'EORTC et appelé « QLQ-C30 ». Le QLQ-C30[1] (pour Quality of Life Questionnaire - Core 30) est un questionnaire standardisé conçu pour mesurer la qualité de vie des patients atteints de cancer.

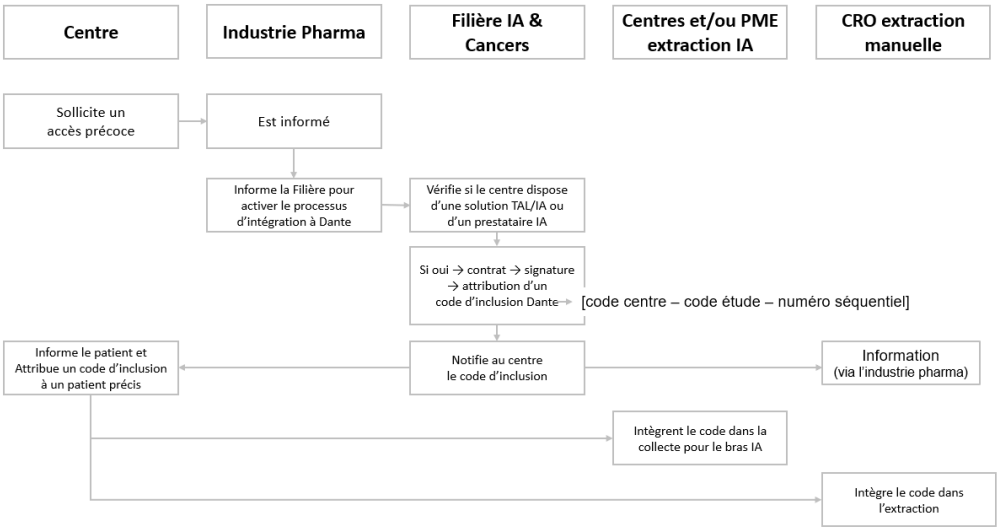
- **Données liées à l'interruption ou l'arrêt du traitement** : si le traitement a été interrompu ou arrêté définitivement, les données suivantes pourront être analysées :
  - Date d'interruption temporaire du traitement, le cas échéant
  - Motif de cette interruption
  - Date de reprise du traitement si applicable
  - Date d'arrêt définitif du traitement
  - Raisons de l'arrêt, telles que documentées par l'équipe soignante

3.2.2 Comment le cycle de vie des données se déroule-t-il (description fonctionnelle) ?

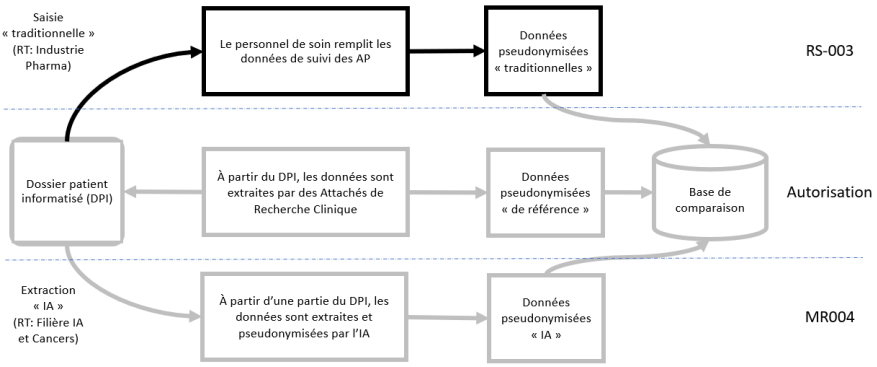
Le cycle de vie des données dans le cadre de l'étude DANTE se déploie selon plusieurs étapes successives, intégrant à la fois les processus de collecte, pseudonymisation, structuration, transfert, analyse et archivage. En voici une description complète :

- **Identification des patients et inclusion**
  - Les établissements de santé identifient les patients éligibles et remplissent une fiche d'accès au traitement dans le cadre d'un protocole d'utilisation thérapeutique et de recueil de données (PUT-RD) validé par la HAS. Cette première étape déclenche l'inclusion du patient, et marque le début de la génération des données exploitables pour l'étude.

[1] Le QLQ-C30 a été développé par l'EORTC (*European Organisation for Research and Treatment of Cancer*), un organisme européen de référence en recherche clinique en oncologie.



- **Collecte des données cliniques – triple modalité (manuelle, IA, référence)**
  - Les données sont collectées selon trois modalités distinctes :
    - Par les CRO « traditionnellement » mandatées par les laboratoires industriels partenaires ; en noir ci-dessous :

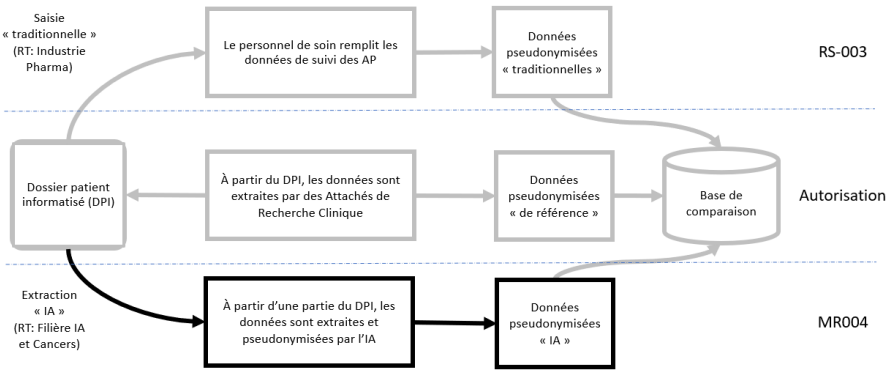


NB : Il convient de distinguer deux opérations. La collecte « traditionnelle » des données n'est pas placée sous la responsabilité de traitement de la Filière : indépendante du projet DANTE, elle correspond au recueil que chaque industriel partenaire réalise déjà au titre de ses propres obligations d'accès précoce, et dont il demeure le responsable de traitement. Cette opération est régie par le référentiel RS-003 (cf. schéma ci-dessus) et n'entre pas, en tant que telle, dans le périmètre de la présente analyse.

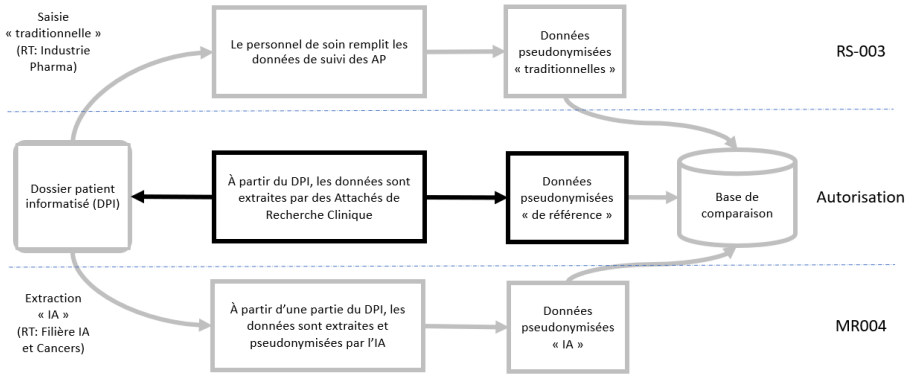
En revanche, la transmission des données pseudonymisées ainsi recueillies vers la base de comparaison (cf. § « Transfert vers une base de comparaison pour analyse comparative » ci-dessous) est, elle, spécifique au projet DANTE : effectuée pour le compte

de la Filière aux fins de l'étude comparative, elle relève de la présente analyse d'impact. Pour cette seule opération, l'entité qui y procède agit en qualité de sous-traitant de la Filière et se trouve tenue des mêmes obligations que celles prévues à l'article 28 du RGPD et reportées sur l'ensemble des sous-traitants du projet (cf. §4.2.5) — à savoir le respect de standards de sécurité minimaux, une obligation de dialogue et de coopération, voire d'audit en cas de difficulté. Ces obligations sont intégralement reportées par voie contractuelle sur les sous-traitants concernés.

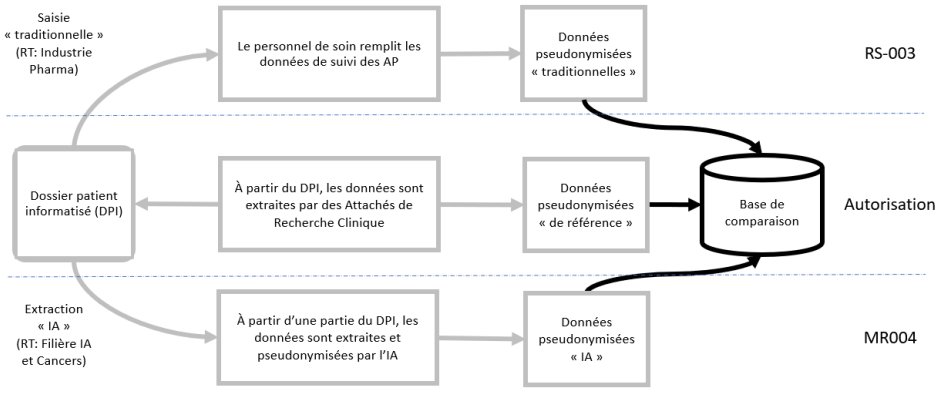
- Dans une première phase (phase 1), automatiquement, via des solutions d'IA de traitement du langage naturel (NLP) déjà installées dans les établissements de santé partenaires par leurs propres soins ou via des PME innovantes. Ces outils extraient et pseudonymisent directement les informations à partir des documents cliniques du patient (comptes rendus, prescriptions, etc.), sans intervention humaine, en noir ci-dessous :



- Par un personnel de recherche (attachés de recherche clinique), pour constituer le bras de référence (phase 2) : la saisie est réalisée en consultant directement le DPI source, sur site ou à distance. Cette modalité suppose un accès humain à des données potentiellement identifiantes et ne sera mise en œuvre qu'après autorisation de la CNIL, en noir ci-dessous :



- **Transfert vers une base de comparaison pour analyse comparative**
    - Les données pseudonymisées issues des deux modalités (manuelle et automatique) sont transmises par une CRO centrale, indépendante, RCTs en charge de fusionner et chaîner les trois bases de données pour permettre l'étude comparative entre les deux méthodes de collecte, comme décrit en noir ci-dessous :



    - Analyse comparative : la CRO centrale applique les méthodes définies dans le protocole d'étude pour comparer :
        - La complétude
        - L'exactitude
        - La cohérence des données collectées par IA vs manuellement
        - Cette analyse vise à évaluer l'intérêt des technologies d'IA pour fiabiliser et simplifier la collecte de données d'accès précoce en vie réelle.

- **Archivage et destruction**
    - Les données pseudonymisées sont conservées pendant 2 ans après la fin de l'étude pour les besoins d'analyse.
    - Elles sont ensuite archivées pendant 5 ans maximum, à des fins de traçabilité scientifique. À l'issue de cette période, elles sont supprimées définitivement.

3.2.3 Quels sont les supports des données ?

Dans le cadre de l'étude DANTE, **seuls des supports HDS (Hébergeurs de Données de Santé certifiés) sont utilisés** pour héberger et sécuriser les données pseudonymisées collectées par intelligence artificielle.

Modalités de recours à un hébergeur HDS :

- Les données pseudonymisées sont stockées sur un espace sécurisé.
- Cet espace est hébergé en France, exclusivement par un hébergeur certifié HDS.
- L'hébergement HDS est mentionné comme une exigence de sécurité pour toutes les données collectées automatiquement dans le cadre du projet.

Concrètement, les prestataires impliqués dans la collecte automatisée des données via IA doivent :

- Garantir que le stockage des données pseudonymisées se fait dans un environnement certifié HDS,
- Assurer la pseudonymisation préalable dans l'établissement d'origine,
- Et transmettre ces données, via une chaîne sécurisée, au tiers de confiance pour analyse.

4 Principes fondamentaux

4.1 Proportionnalité et nécessité

4.1.1 Les finalités du traitement sont-elles déterminées, explicites et légitimes ?

les finalités du traitement sont bien déterminées, explicites et légitimes, en conformité avec les exigences du RGPD (article 5.1.b) et de la méthodologie de référence MR-004.

4.1.1.1 Finalités déterminées et explicites

Le projet DANTE a une finalité clairement définie : comparer la qualité des données collectées manuellement par les professionnels de santé avec celles extraites automatiquement par des technologies d'intelligence artificielle (IA) dans le contexte des traitements administrés en accès précoce**.**

Cette comparaison vise :

- à **évaluer la complétude, la fiabilité et la cohérence** des données issues des deux méthodes,
- à identifier si l'IA peut **améliorer la collecte de données en vie réelle** tout en réduisant la charge pour les professionnels,
- à contribuer à des **usages d'intérêt public**, comme l'optimisation des recommandations thérapeutiques et l'évaluation en vie réelle des innovations.

4.1.1.2 Finalités légitimes

La finalité de l'étude est directement liée à :

- l'**amélioration du système de santé**,
- la **qualité des données d'usage en vie réelle**,
- et la **transparence scientifique** autour des accès précoces.

Elle repose sur un **intérêt légitime** (santé publique, innovation thérapeutique, équité d'accès), ce qui répond aux exigences de la **base légale de l'article 6(1)(f)** du RGPD.

4.1.1.3 Synthèse

Le projet DANTE vise un objectif **méthodologique, scientifique et d'intérêt général**, sans détournement de finalité. Les données sont utilisées **uniquement dans le cadre de cette étude**, avec pseudonymisation, traçabilité et limitation dans le temps. La description est suffisamment précise pour informer les personnes concernées, comme l'exige la CNIL.

4.1.2 Quels sont les fondements qui rendent ce traitement licite ?

4.1.2.1 L'intérêt légitime du responsable de traitement

Base légale : article 6(1)(f) du RGPD ; **le traitement relève de l'intérêt légitime**, en l'occurrence :

- L'évaluation et l'amélioration des méthodes de collecte de données en vie réelle,
- La contribution à la qualité et à la sécurité des soins,
- L'appui à la régulation des traitements innovants (accès précoces).

4.1.2.2 Traitement de données sensibles autorisé

Base complémentaire : article 9(2)(j) du RGPD ; le traitement porte sur des données de santé (catégorie particulière) et est **autorisé à des fins de recherche scientifique**, sous réserve de garanties appropriées (pseudonymisation, sécurité, durée limitée…).

4.1.2.3 Articulation des deux régimes

Les bases de licité exposées ci-dessus — l'article 6(1)(f) et la condition de l'article 9(2)(j) — sont communes à l'ensemble de l'étude. La formalité préalable diffère toutefois selon la phase :

- **Phase 1 — bras IA :** l'extraction automatisée des DPI, sans intervention humaine et avec pseudonymisation locale avant tout transfert, relève de la méthodologie de référence MR-004. Elle est mise en œuvre sur le fondement d'un engagement de conformité, sans autorisation préalable de la CNIL.
- **Phase 2 — bras de référence :** la saisie par un personnel de recherche accédant directement au DPI source (sur site ou à distance) implique un accès humain à des données potentiellement identifiantes. Ce traitement sort du champ de la MR-004 et requiert une autorisation spécifique de la CNIL, instruite par le Health Data Hub et soumise à l'avis préalable du CESREES. La phase 2 ne peut débuter avant l'obtention de cette autorisation.

4.1.2.4 Appréciation de l'intérêt légitime (test de mise en balance)

Le recours à l'article 6(1)(f) impose d'apprécier l'intérêt légitime poursuivi au regard des intérêts et des droits et libertés fondamentaux des personnes concernées. Cette appréciation est conduite en trois temps.

1. Test de finalité — existence d'un intérêt légitime

La Filière IA & Cancers, association à but non lucratif, poursuit un intérêt légitime consistant à évaluer et à améliorer les méthodes de collecte des données générées dans le cadre des accès précoces, dans une finalité méthodologique et scientifique d'intérêt général (fiabilisation des données en vie réelle, allègement de la charge des équipes soignantes, appui à la régulation des traitements innovants). Cet intérêt est réel, présent et clairement défini ; **il ne poursuit aucune finalité commerciale, ni aucune visée individuelle à l'égard des patients. Il est précisé aux finalités exposées supra** (§ « Finalités déterminées, explicites et légitimes »).

2. Test de nécessité — proportionnalité du traitement

Le traitement est nécessaire à la réalisation de cet intérêt et aucun moyen moins intrusif ne permettrait de l'atteindre. La nécessité se vérifie par plusieurs caractéristiques : le traitement porte sur un ensemble strictement limité de variables (set prédéfini de 22 variables), à l'exclusion de toute collecte exhaustive ou indifférenciée ; il repose sur la réutilisation de données déjà produites dans le cadre du soin, sans sollicitation ni examen supplémentaires des patients ; et il ne mobilise que des données pseudonymisées, les données directement identifiantes demeurant au sein des établissements. Le principe de minimisation est donc respecté (cf. § minimisation).

3. Test de mise en balance proprement dit

Mis en regard des droits et libertés des personnes concernées, l'intérêt légitime poursuivi n'apparaît pas surpassé par ceux-ci, compte tenu des garanties suivantes :

- *Attentes raisonnables des personnes* : la réutilisation à des fins de recherche d'intérêt public s'inscrit dans le prolongement direct de la prise en charge en accès précoce et du dispositif de recueil de données qui lui est inhérent ; elle n'est donc pas de nature à surprendre les personnes concernées, d'autant qu'elle fait l'objet d'une information individuelle préalable.

- *Nature des données et niveau de risque* : bien que des données de santé soient en cause, l'impact sur les personnes est limité par la pseudonymisation réalisée localement avant tout transfert, la conservation de la clé de correspondance au sein de l'établissement, l'hébergement HDS en France, le chiffrement (AES-256, TLS 1.2), la journalisation et le recours à un tiers de confiance indépendant. Aucun profilage ni décision individuelle automatisée au sens de l'article 22 du RGPD n'est mis en œuvre.

- *Encadrement réglementaire* : le traitement s'inscrit dans la méthodologie de référence MR-004 et le référentiel RS-003, qui définissent un socle de garanties préétabli par la CNIL.

- *Droit d'opposition* : les personnes peuvent s'opposer à tout moment, sans justification et sans condition, à l'utilisation de leurs données, sans aucune incidence sur la qualité ou la continuité de leur prise en charge. Ce droit, porté à leur connaissance constitue la contrepartie déterminante de l'absence de recueil du consentement et assure l'équilibre du traitement.

Au regard de ces éléments, l'intérêt légitime de la Filière IA & Cancers n'est pas méconnu par les intérêts ou les droits et libertés des personnes concernées, sous réserve du maintien effectif des garanties précitées.

4.1.3 Les données collectées sont-elles adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données) ?

Oui, au vu du **cahier des charges** du projet DANTE, les **données collectées respectent bien le principe de minimisation** tel que défini par l'article 5-1-c du RGPD. Voici pourquoi :

- **Finalité clairement définie** : le projet vise à comparer la qualité et la complétude des données issues de la saisie manuelle versus celles extraites automatiquement par IA, dans le contexte du suivi des accès précoces. La collecte sert **exclusivement à cette comparaison**, et s'inscrit dans une finalité de santé publique et d'amélioration de la collecte en vie réelle.

- **Ensemble limité et structuré de données** : le set de données collectées est strictement défini à partir des **quatre fiches officielles du PUT-RD** validées par la HAS (accès, instauration, suivi, arrêt du traitement). Il comprend **22 variables** essentielles à l'analyse, certaines récurrentes selon la chronologie du traitement.
- **Aucune donnée inutile ou hors périmètre** : les données ne sont pas collectées de façon exhaustive ou indifférenciée. Elles sont choisies pour leur **utilité directe dans l'évaluation de la fiabilité des méthodes de collecte**, dans le cadre des obligations de suivi réglementaire des accès précoces.
- **Exclusion des données directement identifiantes** : seules des données **pseudonymisées** sont utilisées pour l'analyse, avec un cloisonnement clair entre les identifiants et les données traitées, ce qui réduit le risque d'exposition de données inutiles.

Le projet DANTE applique rigoureusement le principe de minimisation des données : **les informations collectées sont nécessaires, proportionnées et directement liées à l'objectif scientifique du projet**, sans excès ni traitement superflu. Il est donc **conforme** aux exigences du RGPD sur ce point.

4.1.4 Les données sont-elles exactes et tenues à jour ?

Non, les données ne sont **pas garanties comme exactes ou tenues à jour de façon systématique** dans l'état initial, car l'étude DANTE a justement pour objectif de **tester si l'IA permettrait d'améliorer la qualité, la cohérence et l'actualisation des données** par rapport aux pratiques actuelles, souvent manuelles, incomplètes et variables selon les établissements.

4.1.5 Quelle est la durée de conservation des données ?

Dans l'étude DANTE, les **données pseudonymisées** sont :

- **Conservées pendant deux ans** à compter de la fin de l'étude, afin de permettre aux équipes scientifiques de finaliser les analyses et, le cas échéant, de publier les résultats ;
- Puis **archivées de manière sécurisée pendant cinq années supplémentaires**, à des fins de traçabilité et de contrôle réglementaire ;
- Au terme de ces cinq années, **elles sont supprimées définitivement**, dans le respect des normes de sécurité en vigueur.

4.2 Mesures protectrices des droits

4.2.1 Comment les personnes concernées sont-elles informées du traitement ?

Dans le cadre de l'étude **DANTE**, les personnes concernées (les patients) sont informées du traitement de leurs données via une **note d'information**. Cette note informe les patients, de manière claire et accessible, des deux modalités de traitement de leurs données : (i) en phase 1, de l'extraction automatisée de leurs données issues de leur dossier patient informatisé par une solution d'intelligence artificielle, sans intervention humaine ; et (ii) en phase 2, de la consultation de leur dossier patient informatisé par un personnel de recherche, sur site ou à distance, pour la constitution du bras de référence. Elle les informe également de leur faculté d'opposition, qui peut être exercée à tout moment et porte aussi bien sur la phase 1 que sur la phase 2. Elle contient les éléments suivants :

- **Nom et finalité de l'étude** (DANTE, visant à comparer deux méthodes de collecte de données)
- **Identité du responsable de traitement** (Filière IA & Cancers)
- **Base légale** du traitement (intérêt légitime, art. 6.1.f du RGPD)
- **Données traitées** et leur origine (issues du dossier médical dans le cadre d'un accès précoce)
- **Destinataires** des données et modalités de pseudonymisation
- **Durée de conservation**
- **Droits** des personnes (accès, opposition, rectification, effacement, limitation)
- **Modalités d'exercice de ces droits** (contact DPO)
- **Référence à la CNIL** en cas de réclamation

4.2.2 Comment les personnes concernées peuvent-elles exercer leurs droits d'accès ?

Dans l'étude DANTE, les personnes concernées peuvent exercer leur **droit d'accès** de manière claire et encadrée, conformément au **Règlement général sur la protection des données (RGPD)**.

4.2.2.1 Droit d'accès

Chaque personne peut :

- **Demander à consulter les données** pseudonymisées utilisées dans le cadre de l'étude,
- **Obtenir une copie** des informations la concernant, dans un format compréhensible.

4.2.2.2 Comment exercer ces droits ?

Les personnes concernées peuvent formuler leur demande :

- **Par e-mail** : à l'adresse du Délégué à la protection des données (DPO) : dpo@filiere-ia.fr
- **Par courrier postal** : Délégué à la protection des données Filière IA & Cancers, PariSanté Campus, 2-10 rue d'Oradour-sur-Glane 75015 Paris

L'exercice de ces droits se fait **sans justification** et **sans impact sur la qualité de la prise en charge médicale**.

4.2.3 Comment les personnes concernées peuvent-elles exercer leurs droits de rectification et à l'effacement (droit à l'oubli) ?

Nota bene : ces droits s'exercent dans la limite de l'objet de l'étude, qui est précisément de mesurer l'exactitude des données produites, et dans les limites de la pseudonymisation et du moment, où une personne demande l'accès à ses données ; leur rectification ou leur effacement peut ne plus être possible techniquement.

Au-delà et si les personnes concernées peuvent exercer leur **droit de rectification** et leur **droit à l'effacement** (ou « droit à l'oubli ») conformément aux règles suivantes.

4.2.3.1 Droit de rectification

Les personnes concernées peuvent demander la **correction de données inexactes** ou incomplètes, si elles constatent qu'une information traitée dans le cadre de l'étude est erronée.

4.2.3.2 Droit à l'effacement

Elles peuvent demander que les **données soient effacées**, notamment si :

- **Elles s'opposent au traitement** (voir droit d'opposition),
- Les données ne sont **plus nécessaires** aux finalités de l'étude,
- Ou si leur traitement est **non conforme à la réglementation**.

Les modalités d'exercice sont les mêmes que celles mentionnées ci-dessus au § « Comment exercer ces droits ? » au 4.2.2.2 ci-dessus ci-dessus

4.2.4 Comment les personnes concernées peuvent-elles exercer leurs droits de limitation et d'opposition ?

Dans la note d'information pour l'étude DANTE, les modalités d'exercice du **droit de limitation** et du **droit d'opposition** sont clairement détaillées. Les personnes concernées peuvent limiter (suspendre temporairement) ou s'opposer à tout moment à l'utilisation de leurs données dans le cadre de l'étude DANTE, y compris à la consultation de leur dossier patient informatisé par un personnel de recherche au titre

du bras de référence (phase 2), sans incidence sur la qualité ou la continuité de leur prise en charge.

S'agissant de l'opposition, pour des raisons techniques, il est recommandé aux patients, sans obligation, dans la note d'information dédiée, d'exercer ce droit **dans un délai de 30 jours** suivant la réception de la note.

Les modalités d'exercice sont les mêmes que celles mentionnées ci-dessus au § « Comment exercer ces droits ? » au 4.2.2.2 ci-dessus en page 21

4.2.5 Les obligations des sous-traitants sont-elles clairement définies et contractualisées ?

Oui, les obligations des sous-traitants sont **clairement définies et contractualisées** dans le contrat de collecte des données. Voici une synthèse structurée des éléments principaux :

4.2.5.1 Cadre légal clair et conforme

Le contrat intègre les **clauses contractuelles types de la Décision d'exécution (UE) 2021/915**, en application de l'article 28 du RGPD. Cela garantit que toutes les parties (Responsable du traitement et Sous-traitants) sont engagées contractuellement à respecter les normes européennes sur la protection des données.

4.2.5.2 Engagement explicite sur l'analyse d'impact (AIPD)

L'article *Assistance au Responsable du traitement* (Section 3) stipule explicitement que le Sous-traitant :

- **Aide à la réalisation d'une analyse d'impact sur la protection des données** dans les cas où un type de traitement est susceptible de présenter un risque élevé pour les droits et libertés des personnes physiques.
- **Assiste le Responsable du traitement en cas de consultation préalable de l'autorité de contrôle**, si l'analyse révèle un risque élevé non atténué.

4.2.5.3 Contractualisation formelle

L'ensemble de ces engagements sont **insérés dans le contrat**, avec des conséquences en cas de non-respect :

- Le Responsable peut **suspendre ou résilier le contrat** en cas de manquement du Sous-traitant aux obligations RGPD (Article « Non-respect des clauses et résiliation »).
- Le Sous-traitant doit **documenter sa conformité**, permettre des audits et collaborer en cas d'incident (violation de données, demande d'un patient, etc.).

4.2.5.4 Mesures techniques et organisationnelles

Le contrat décrit aussi de manière détaillée les **mesures concrètes à mettre en œuvre**, telles que :

- **Chiffrement des données** (en transit et au repos, avec AES-256, TLS 1.2).
- **Hébergement HDS** (données uniquement en France).

4.2.5.5 Documentation, audits et vérifiabilité

- Le Sous-traitant doit tenir à disposition un **registre des traitements et incidents de sécurité**.
- Le Responsable du traitement est autorisé à **réaliser des audits documentaires** pour s'assurer de la conformité.
- En cas de changement de Sous-traitant ultérieur, une procédure d'information et d'approbation par le Responsable est définie.

5 Mesures de prévention des risques

5.1 Sécurité des données conservées (au repos)

5.1.1 Chiffrement des données

- Toutes les **données de santé sont chiffrées** dans leur intégralité avec des **algorithmes robustes** : **AES-256** est un standard minimum.

5.1.2 Hébergement sécurisé

- Les données sont **exclusivement stockées chez des prestataires certifiés HDS** (Hébergeurs de Données de Santé) en France, conformément à l'article L1111-8 du Code de la santé publique. Cela garantit le respect des standards français les plus stricts en matière de confidentialité et de sécurité.

5.1.3 Contrôle d'accès

- Des **politiques strictes de contrôle d'accès** sont en place :
  - Accès selon le **principe du moindre privilège** ;
  - **Révision régulière des droits d'accès** ;
  - Seules les personnes formées à la sécurité des données de santé peuvent accéder aux données.

5.1.4 Sauvegardes sécurisées

- Des **sauvegardes régulières et chiffrées** sont prévues.
- Ces sauvegardes incluent des **mesures de sécurité contre la perte, l'altération ou la destruction accidentelle**.

5.1.5 Gestion des clés de chiffrement

- Les systèmes utilisés doivent garantir que les clés de chiffrement sont **protégées contre l'accès non autorisé**, ce qui est une obligation implicite dans le cadre d'un hébergement HDS et d'un usage de chiffrement AES-256.

- En cas de **violation de sécurité ou suspicion de compromission**, le Sous-traitant doit :
    - **Notifier immédiatement** le Responsable du traitement ;
    - Fournir **les mesures correctrices envisagées**.

5.2 Sécurité des flux de données (en transit)

Le contrat de sous-traitance impose l'usage des **protocoles sécurisés suivants** pour tout échange de données :

- **TLS** (Transport Layer Security), minimum version **1.2**.
- **SFTP** (Secure File Transfer Protocol) ;
- **HTTPS** avec authentification renforcée ;
- **Authentification forte** : par exemple, **authentification multifacteur (MFA)** pour les utilisateurs accédant aux systèmes traitant les données.

5.3 Traçabilité et journalisation

Dans le cadre du projet **DANTE**, les politiques de **journalisation des événements** et de **conservation des traces** sont décrites avec précision dans le contrat, conformément aux exigences du **Règlement Général sur la Protection des Données (RGPD)** et aux standards applicables aux données de santé. Voici un exposé des principales mesures :

Le contrat exige que les prestataires mettent en œuvre un **journal d'audit** pour chaque action réalisée sur les données. Cela inclut :

- **Horodatage précis** de chaque opération ;
- **Identification de l'utilisateur ou du système** ayant réalisé l'action ;
- **Nature exacte de l'opération** (lecture, extraction, transfert, suppression, etc.)
- **Origine et destination des flux**, en cas de transmission de données.

Par ailleurs, tous les systèmes d'information impliqués (bases de données, plateformes de traitement IA, modules d'accès) doivent intégrer un mécanisme de **journalisation natif**, sans possibilité de désactivation. Cela concerne aussi bien :

- Les accès en lecture ou écriture aux données ;
- Les connexions administratives ou techniques ;
- Les transferts vers le tiers de confiance (CRO centrale).

Le Responsable du traitement (Filière IA & Cancers) peut **exiger la transmission des journaux d'audit** dans le cadre :

- D'**audits documentaires réguliers** ;
- De **contrôle de conformité RGPD** ;
- D'**investigations sur des incidents ou violations de données**.

Mise à disposition aux autorités

- En cas de demande de la **CNIL ou d'une autorité de contrôle compétente**, les journaux doivent être mis à disposition **immédiatement**, incluant :
  - Les traces techniques complètes ;
  - La documentation du dispositif de journalisation.

5.4 Application du principe de cloisonnement

5.4.1 Double modalité de collecte : manuelle vs automatisée

- Les traitements manuels (effectués par les CROs « traditionnelles ») sont **totalement séparés** des traitements automatisés (réalisés par les prestataires IA).
- La **CRO centrale** agit comme **tiers de confiance** pour regrouper les deux bases, mais selon un protocole clairement défini et encadré.

5.4.2 Interdiction d'accès aux documents sources

- Aucune intervention humaine ou consultation de documents originaux n'est autorisée **dans le cadre du traitement automatisé** : Cela garantit que la solution d'IA fonctionne **sans biais ni interférence humaine**, et sans croisement non justifié entre les sources de données.
- Cette interdiction vaut exclusivement pour la phase 1 (bras IA). La phase 2 (bras de référence) repose au contraire, par construction méthodologique, sur une consultation humaine du DPI source par un personnel de recherche, sur site ou à distance. Cet accès constitue une exception délibérée et circonscrite, subordonnée à l'autorisation de la CNIL et assortie de garanties propres (habilitation nominative, canal sécurisé, traçabilité des accès, absence de copie locale en dehors de l'environnement sécurisé, engagement de confidentialité). Le personnel de référence saisit en aveugle, sans accès aux données du bras IA, de sorte que le cloisonnement entre les bras est préservé.

5.4.3 Restrictions contractuelles et géographiques

- **Interdiction des transferts hors France** : tous les traitements sont hébergés **exclusivement en France**, interdisant explicitement tout **transfert de données en dehors du territoire national**, même au sein de l'UE.

NB : À noter par ailleurs que le projet prévoit une **répartition équilibrée des patients entre établissements** (aucun centre ne devant représenter plus de 15 %

des patients inclus – i.e. 22 patients maximum par centre), limitant la concentration excessive de données dans un même flux ou point d'entrée.

5.5 Désignation d'un Délégué à la Protection des Données (DPO)

5.5.1 DPO officiellement déclaré auprès de la CNIL

Monsieur Marco Fiorini est **Délégué à la Protection des Données (DPO)**, avec effet au **10 mai 2025**. La désignation est enregistrée sous le numéro **DPO-160814** et est **valide auprès de la CNIL**, avec publication des coordonnées publiques comme l'exige le RGPD.

5.5.2 Coordonnées officielles du DPO

- **Nom** : Marco Fiorini
- **Adresse** : PariSanté Campus, 2-10 rue d'Oradour-sur-Glane, 75015 Paris
- **Email de contact public** : dpo@filiere-ia.fr

5.5.3 Rôle du DPO dans le projet DANTE

Dans le contrat du projet DANTE, le DPO est explicitement mentionné comme **interlocuteur de référence** pour :

- Les **audits de conformité** ;
- La **coopération avec les sous-traitants** ;
- La **gestion des incidents de sécurité** ;
- L'**exercice des droits des personnes concernées**.

5.5.4 Autres structures de gouvernance

5.5.4.1 Responsable de traitement clairement identifié

La **Filière IA & Cancers** agit comme **Responsable de traitement** dans le projet Dante, ce qui renforce son obligation de piloter la gouvernance RGPD.

5.5.4.2 Comité de suivi RGPD

Un **comité dédié à la protection des données** (type comité de pilotage ou comité éthique) sera constitué pour être une gouvernance structurée conforme au RGPD qui suive le projet, au-delà de son lancement.

6 Evaluation des risques

L'introduction de la phase 2 (bras de référence) crée une surface de risque spécifique, distincte de celle de la phase 1 : un accès humain à des données de santé potentiellement identifiantes du DPI source, le cas échéant à distance. Ce risque accru est traité par des mesures dédiées — habilitation nominative, authentification renforcée, canal sécurisé, journalisation des accès, absence de copie locale hors environnement sécurisé, engagement de confidentialité du personnel de recherche — ainsi que par la subordination de cette phase à l'autorisation préalable de la CNIL et à l'avis du CESREES. Les fiches de risque ci-dessous s'appliquent aux deux phases ; les impacts et vraisemblances doivent être lus en tenant compte de cet accès humain propre à la phase 2.

6.1 Accès illégitime à des données

6.1.1 Quels pourraient être les principaux impacts sur les personnes concernées si le risque se produisait ?

- **Atteinte à la vie privée** : Réidentification possible via recoupement, exposition de traitements ou pathologies sensibles,
- **Discrimination / stigmatisation** : Si des données de santé sont exposées (cancer, traitement expérimental), risque de discrimination à l'emploi, assurance ou logement,
- **Perte de confiance** : exercice du droit d'opposition, non-participation à d'autres études, tension dans la relation médecin-patient,
- **Préjudice moral / psychologique** si une fuite est rendue publique ou exploitée malveillamment,
- **Impact juridique / indemnisation** : Actions collectives ou individuelles si préjudice prouvé (cf. jurisprudence européenne)

6.1.2 Quelles sont les principales menaces qui pourraient permettre la réalisation du risque ?

- **Menace humaine interne** : Un membre du personnel ou un prestataire autorisé peut, par erreur ou malveillance, accéder, modifier ou divulguer des données sensibles hors cadre, compromettant la confidentialité des patients,
- **Menace humaine externe** : Un acteur malveillant extérieur, tel qu'un cybercriminel ou un concurrent, peut tenter d'infiltrer les systèmes pour voler, bloquer ou exploiter les données à des fins frauduleuses
- **Menace technique** : Une faille logicielle, une configuration inadéquate ou un malware peut altérer l'intégrité des données, en perturber le traitement ou en permettre la fuite
- **Menace environnementale** : Un événement physique tel qu'un incendie, une inondation ou une panne matérielle peut entraîner une perte définitive des

données ou une interruption du service sans possibilité immédiate de restauration.

6.1.3 Quelles sources de risques pourraient-elles en être à l'origine ?

- **Erreur humaine** : Une personne autorisée peut, par inadvertance, mal manipuler, divulguer ou modifier des données sensibles,
- **Malveillance interne** : Un collaborateur ou sous-traitant mal intentionné peut exploiter son accès pour exfiltrer ou compromettre des données,
- **Vulnérabilité technique** : Un système non mis à jour ou mal configuré peut être exploité par un attaquant pour accéder illégalement aux données,
- **Intrusion par logiciel malveillant** : Un virus, ransomware ou spyware peut infecter le système et compromettre les données sans ciblage spécifique,
- **Gouvernance insuffisante** : Une mauvaise répartition des responsabilités ou un manque de pilotage RGPD fragilise le dispositif de conformité,
- **Incident physique ou environnemental** : Un incendie, une inondation ou une coupure électrique peut entraîner la perte ou la destruction de données.

6.1.4 Quelles sont les mesures initiales, parmi celles identifiées, qui contribuent à traiter le risque ?

Chiffrement, Cloisonnement, Journalisation, Organisation de la politique de protection de la vie privée

6.1.5 Comment estimez-vous la gravité du risque, notamment en fonction des impacts potentiels et des mesures prévues ?

Maximale, **Dans le projet DANTE**, les **données de santé**, même pseudonymisées, représentent un **impact potentiellement élevé à critique** si exposées, notamment en cas de réidentification croisée.

6.1.6 Comment estimez-vous la vraisemblance du risque, notamment au regard des menaces, des sources de risques et des mesures prévues ?

Limitée, **La vraisemblance du risque est modérée** : les menaces sont crédibles, les sources de risque réelles, mais les **mesures de sécurité mises en œuvre dans DANTE sont de nature à limiter significativement la probabilité d'occurrence**.

6.2 Modifications non désirées de données

6.2.1 Quels pourraient être les principaux impacts sur les personnes concernées si le risque se produisait ?

- **Impact scientifique** : Une donnée modifiée pourrait fausser les résultats comparatifs entre saisie manuelle et automatisée, compromettant la validité de l'étude,
- **Impact institutionnel** : L'exploitation de données altérées pourrait discréditer la Filière ou ses partenaires et remettre en cause la confiance des parties prenantes,
- **Impact juridique** : L'utilisation ou la diffusion de données inexactes expose le responsable de traitement à des sanctions de la CNIL et à des recours potentiels des personnes concernées.

6.2.2 Quelles sont les principales menaces qui pourraient permettre la réalisation du risque ?

- **Malveillance interne** : Un collaborateur ou sous-traitant mal intentionné peut volontairement falsifier ou corrompre des données pour influencer les résultats,
- **Défaillance logicielle** : Un bug ou une mauvaise configuration dans l'outil d'extraction automatique peut entraîner la modification erronée de nombreuses données,
- **Erreur d'appariement des bases** : Une mauvaise correspondance entre les données issues de sources manuelles et automatisées peut provoquer une altération ou une incohérence des dossiers patients.

6.2.3 Quelles sources de risques pourraient-elles en être à l'origine ?

- **Erreur humaine** : Une personne autorisée peut, par inadvertance, mal manipuler, divulguer ou modifier des données sensibles,
- **Gouvernance insuffisante** : Une mauvaise répartition des responsabilités ou un manque de pilotage RGPD fragilise le dispositif de conformité,
- **Intrusion par logiciel malveillant** : Un virus, ransomware ou spyware peut infecter le système et compromettre les données sans ciblage spécifique,
- **Malveillance interne** : Un collaborateur ou sous-traitant mal intentionné peut exploiter son accès pour exfiltrer ou compromettre des données,
- **Incident physique ou environnemental** : Un incendie, une inondation ou une coupure électrique peut entraîner la perte ou la destruction de données,
- **Vulnérabilité technique** : Un système non mis à jour ou mal configuré peut être exploité par un attaquant pour accéder illégalement aux données.

6.2.4 Quelles sont les mesures, parmi celles identifiées, qui contribuent à traiter le risque ?

Journalisation, Organisation de la politique de protection de la vie privée

6.2.5 Comment estimez-vous la gravité du risque, notamment en fonction des impacts potentiels et des mesures prévues ?

Importante, La gravité du risque est considérée comme élevée en raison des conséquences potentiellement majeures sur la validité scientifique de l'étude, la conformité réglementaire (notamment au regard du RGPD) et la protection des droits des personnes concernées, les données traitées étant à la fois sensibles et utilisées à des fins d'évaluation comparative critiques.

6.2.6 Comment estimez-vous la vraisemblance du risque, notamment au regard des menaces, des sources de risques et des mesures prévues ?

La vraisemblance du risque de modification non désirée de données peut être évaluée comme **négligeable**, grâce aux mesures contractuelles et techniques mises en œuvre. Toutefois, elle reste **"Limitée"** dans certains cas spécifiques, notamment en lien avec les éventuelles erreurs algorithmiques dans l'IA (non reliées à une attaque).

6.3 Disparition de données

6.3.1 Quels pourraient être les principaux impacts sur les personnes concernées si le risque se produisait ?

- **Impact institutionnel** : L'exploitation de données altérées pourrait discréditer la Filière ou ses partenaires et remettre en cause la confiance des parties prenantes,
- **Perte de confiance** : exercice du droit d'opposition, non-participation à d'autres études, tension dans la relation médecin-patient,
- **Préjudice moral / psychologique** si une fuite est rendue publique ou exploitée malveillamment,
- **Traçabilité scientifique compromise** : La disparition des données empêcherait toute analyse comparative entre méthodes de collecte, compromettant la valeur scientifique du projet,
- **Perte de bénéfice collectif** : L'absence de données limiterait les retombées positives du projet pour les futurs patients en accès précoce.

6.3.2 Quelles sont les principales menaces qui pourraient permettre la réalisation du risque ?

- **Défaillance logicielle** : Un bug ou une mauvaise configuration dans l'outil d'extraction automatique peut entraîner la modification erronée de nombreuses données,
- **Malveillance interne** : Un collaborateur ou sous-traitant mal intentionné peut volontairement falsifier ou corrompre des données pour influencer les résultats,
- **Menace environnementale** : Un événement physique tel qu'un incendie, une inondation ou une panne matérielle peut entraîner une perte définitive des données ou une interruption du service sans possibilité immédiate de restauration,

- **Menace humaine externe** : Un acteur malveillant extérieur, tel qu'un cybercriminel ou un concurrent, peut tenter d'infiltrer les systèmes pour voler, bloquer ou exploiter les données à des fins frauduleuses,
- **Menace technique** : Une faille logicielle, une configuration inadéquate ou un malware peut altérer l'intégrité des données, en perturber le traitement ou en permettre la fuite,
- **Erreur d'appariement des bases** : Une mauvaise correspondance entre les données issues de sources manuelles et automatisées peut provoquer une altération ou une incohérence des dossiers patients,
- **Menace humaine interne** : Un membre du personnel ou un prestataire autorisé peut, par erreur ou malveillance, accéder, modifier ou divulguer des données sensibles hors cadre, compromettant la confidentialité des patients.

6.3.3 Quelles sources de risques pourraient-elles en être à l'origine ?

- **Erreur humaine** : Une personne autorisée peut, par inadvertance, mal manipuler, divulguer ou modifier des données sensibles,
- **Gouvernance insuffisante** : Une mauvaise répartition des responsabilités ou un manque de pilotage RGPD fragilise le dispositif de conformité,
- **Intrusion par logiciel malveillant** : Un virus, ransomware ou spyware peut infecter le système et compromettre les données sans ciblage spécifique,
- **Malveillance interne** : Un collaborateur ou sous-traitant mal intentionné peut exploiter son accès pour exfiltrer ou compromettre des données,
- **Vulnérabilité technique** : Un système non mis à jour ou mal configuré peut être exploité par un attaquant pour accéder illégalement aux données,
- **Incident physique ou environnemental** : Un incendie, une inondation ou une coupure électrique peut entraîner la perte ou la destruction de données.

6.3.4 Quelles sont les mesures, parmi celles identifiées, qui contribuent à traiter le risque ?

Organisation de la politique de protection de la vie privée, Cloisonnement

6.3.5 Comment estimez-vous la gravité du risque, notamment en fonction des impacts potentiels et des mesures prévues ?

Importante, dans le cadre du projet Dante, le risque serait grave, car une perte de données de santé pseudonymisées compromettant à la fois la validité scientifique de l'étude, les droits des patients et la confiance des partenaires.

6.3.6 Comment estimez-vous la vraisemblance du risque, notamment au regard des menaces, des sources de risques et des mesures prévues ?

Négligeable, Malgré des menaces réelles, la probabilité de concrétisation de ce risque est jugée **faible**, grâce à un **cadre contractuel rigoureux**, des **exigences techniques précises**, et une **traçabilité complète** du traitement.

6.4 Vue d'ensemble