

Fiche pratique

CIRCULATION DU NIR
AUX FINS D'APPARIEMENT AVEC LE SNDS

CIRCUIT RESPONSABLE DE TRAITEMENT /
CENTRE UNIQUE / eCRF SANS NIR

Produit en collaboration avec le CASD

RT-CU - eCRF sans NIR – Version 1.0 – Juin 2024

1. Introduction

La CNIL publie un ensemble de fiches pratiques, produites avec le CASD, présentant des exemples de circuits d'appariement avec le SNDS, en complément du guide pratique de la CNIL¹ publié en décembre 2020.

Les fiches présentent :

- des schémas fonctionnels détaillés pour chaque étape, dans le même formalisme que ceux du guide ;
- des schémas techniques orientés « tables de données », produits par le CASD.

Ces fiches illustrent en détail des exemples d'implémentation concrète des circuits du guide de 2020, lequel reste valide et se trouve ainsi précisé par les fiches.

Ces exemples respectent les principes issus du guide, qui ont été déclinés pour les fiches pratiques. Vous les trouverez rassemblés dans le document vademécum², comme aide-mémoire et guide de lecture.

La présente fiche concerne une étude monocentrique où le centre investigateur dépend du RT et gère ses données dans un « eCRF », sauf le NIR qui reste stocké localement dans le centre.

Elle correspond à la fusion des cas « RT » et « CU » du guide.

Un tiers de confiance peut être sollicité pour mettre au format et opérer la transmission du NIR à la CNAM.

Présentation du CASD

Le Centre d'accès sécurisé aux données (CASD) est un groupement d'intérêt public (GIP) rassemblant l'État représenté par INSEE, le GENES, le CNRS, l'École polytechnique, HEC Paris et la Banque de France.

Il a été créé par [arrêté interministériel du 29 décembre 2018](#).

Le GIP a pour objet principal d'organiser et de mettre en œuvre des services d'accès sécurisé pour les données confidentielles à des fins non lucratives de recherche, d'étude, d'évaluation ou d'innovation. Il a également pour mission de valoriser la technologie développée pour sécuriser l'accès aux données dans le secteur public et dans le secteur privé.



¹ Guide pratique : Modalités de circulation du NIR pour la recherche en santé aux fins d'appariement avec le SNDS (PDF, 660 ko), CNIL, URL :

https://www.cnil.fr/sites/cnil/files/atoms/files/guide_pratique_circuits_nir_recherche_en_sante.pdf

² Vademécum : circulation du NIR aux fins d'appariement avec le SNDS (PDF, 328 ko), CNIL, URL :

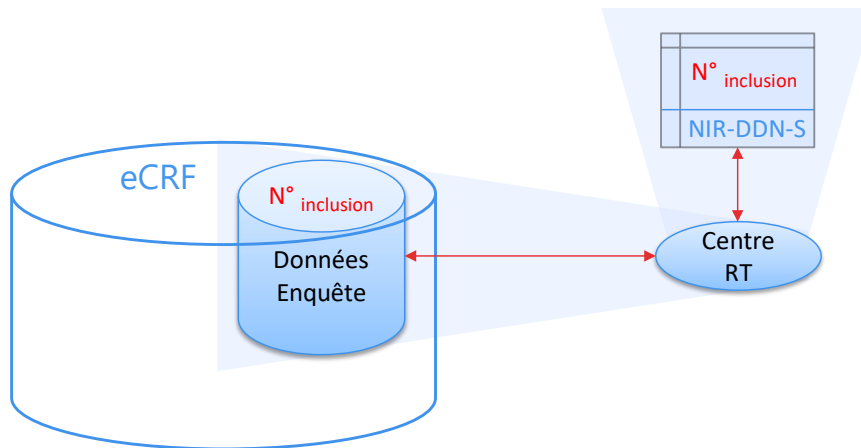
https://www.cnil.fr/sites/cnil/files/2024-06/circuits_nir_vademecum.pdf

2. Implémentation détaillée du circuit RT / centre unique / eCRF sans NIR

Saisie des données d'enquête

Le centre saisit ses données dans un eCRF ne contenant pas le NIR

- La sécurité est assurée par le eCRF selon les règles fixées par le RT.
- Afin de limiter les risques de réidentification, **le eCRF est pseudonymisé et le NIR des patients reste stocké dans le centre, de manière cloisonnée par rapport au numéro d'inclusion** (par exemple, le NIR est chiffré avec une clé spécifique).
- En complément, **une authentification forte est recommandée pour tout accès au eCRF ainsi qu'à la table de correspondance locale** entre le numéro d'inclusion et le NIR, que ce soit en saisie de données ou pour une simple consultation.

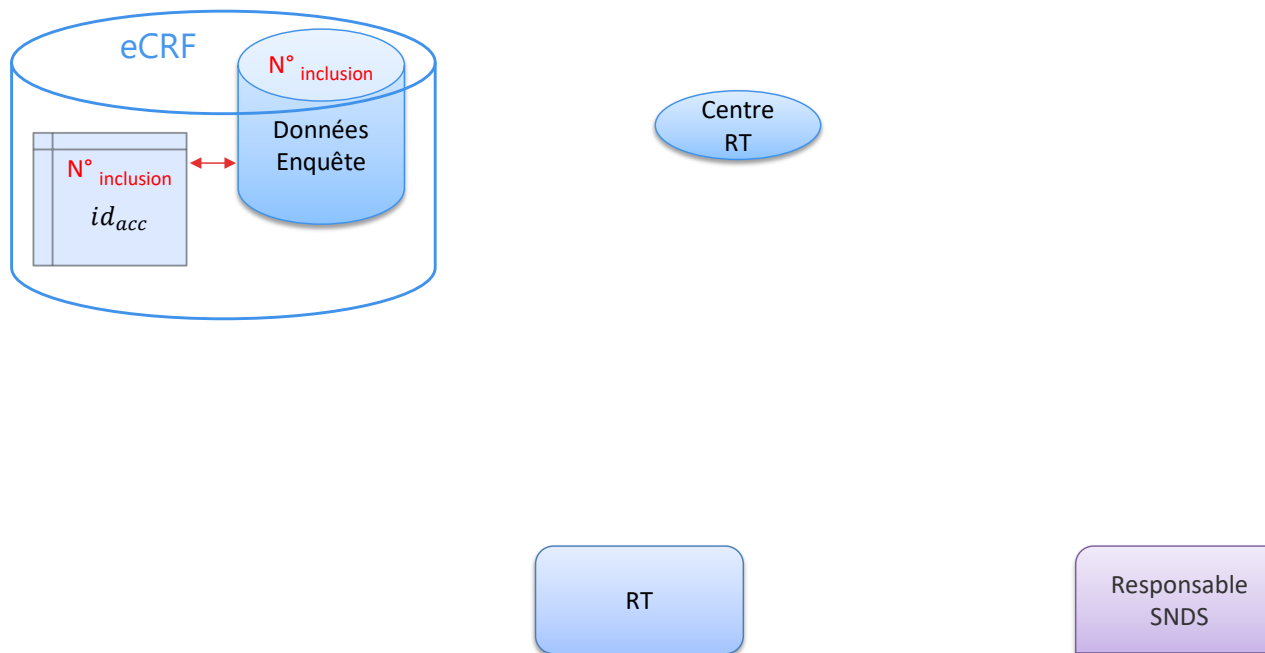


Etape 0 – Génération des identifiants d'accrochage

Le eCRF génère en interne, non visible du centre, une table de correspondance entre le numéro d'inclusion des participants et un identifiant d'accrochage aléatoire et non significatif (id_{acc})

- L'utilisation d'identifiants techniques temporaires (« identifiants d'accrochage ») permet de dissocier le NIR et les données de santé lors des transferts entre acteurs.
- Par principe, **ces numéros sont non significatifs et différents du numéro d'inclusion de la personne dans l'étude**, afin de limiter les risques de réidentification croisée entre le numéro d'inclusion, le NIR et les données de santé (enquête et SNDS).

- De même, **le centre n'a pas d'accès direct à la table de correspondance interne au eCRF** qui fait le lien entre les numéros d'inclusion et les identifiants d'accrochage.
- Si le besoin est dûment justifié (par ex. transmission récurrente, audit des données, alerte des participants), la table de correspondance entre le numéro d'inclusion et l'identifiant d'accrochage peut être conservée dans le eCRF, de manière sécurisée.
- L'identifiant d'accrochage peut être généré par une fonction mathématique aléatoire, mais aussi par une fonction de hachage à clé secrète ; dans le second cas, c'est la clé secrète qui sera conservée au lieu de la table de correspondance.
- Si l'identifiant d'accrochage est généré au fil des inclusions, il doit rester masqué dans l'interface eCRF du centre.

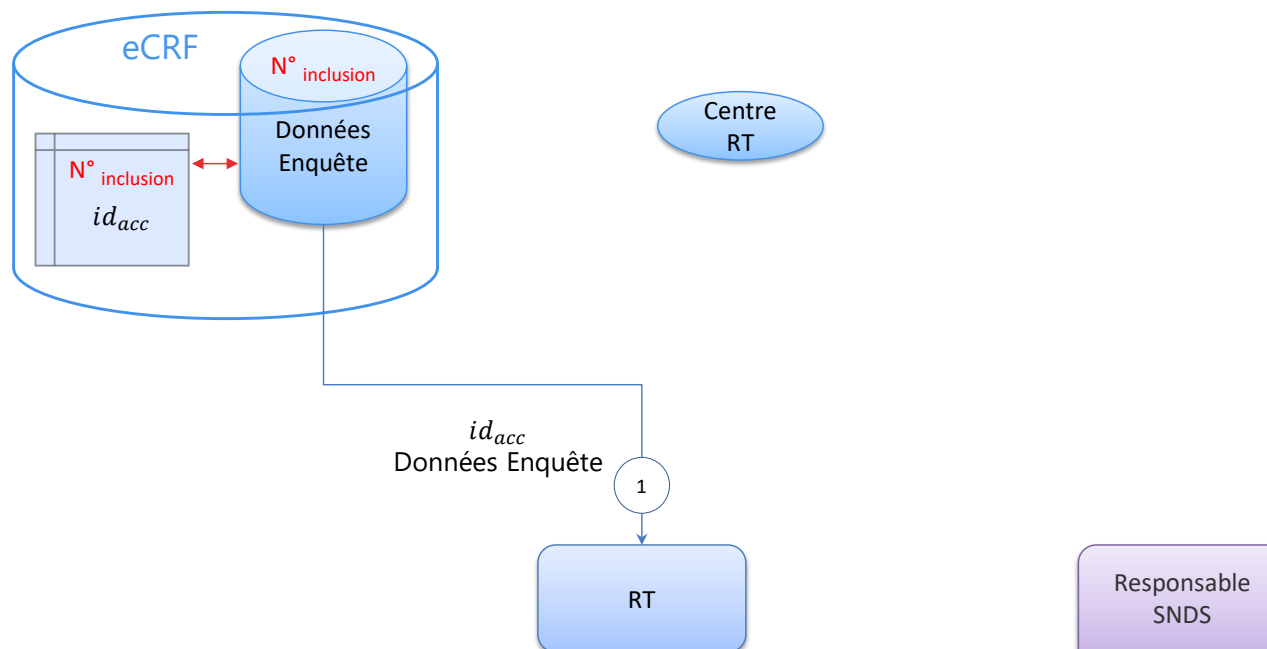


Etape 1 – Extraction des données d'enquête

Les données d'enquête sont extraites du eCRF, associées aux identifiants d'accrochage

- Par principe, afin de limiter les risques de réidentification portant sur les données de l'enquête et sur les données du SNDS auxquelles elles seront appariées, **le numéro d'inclusion n'est pas extrait du eCRF**.
- NB : les risques de réidentification sont à considérer pour la transmission et pour le stockage des données.

- Dès lors, **pour extraire les données strictement nécessaires** (identifiants d'accrochage et données d'enquête), le RT peut demander l'intervention manuelle d'un administrateur habilité ou déclencher une fonction interne au eCRF.
- Dans le cas d'une fonction interne au eCRF, celle-ci pourra également envoyer aux centres les données qui leur sont nécessaires pour le circuit d'appariement (cf. étape 2).

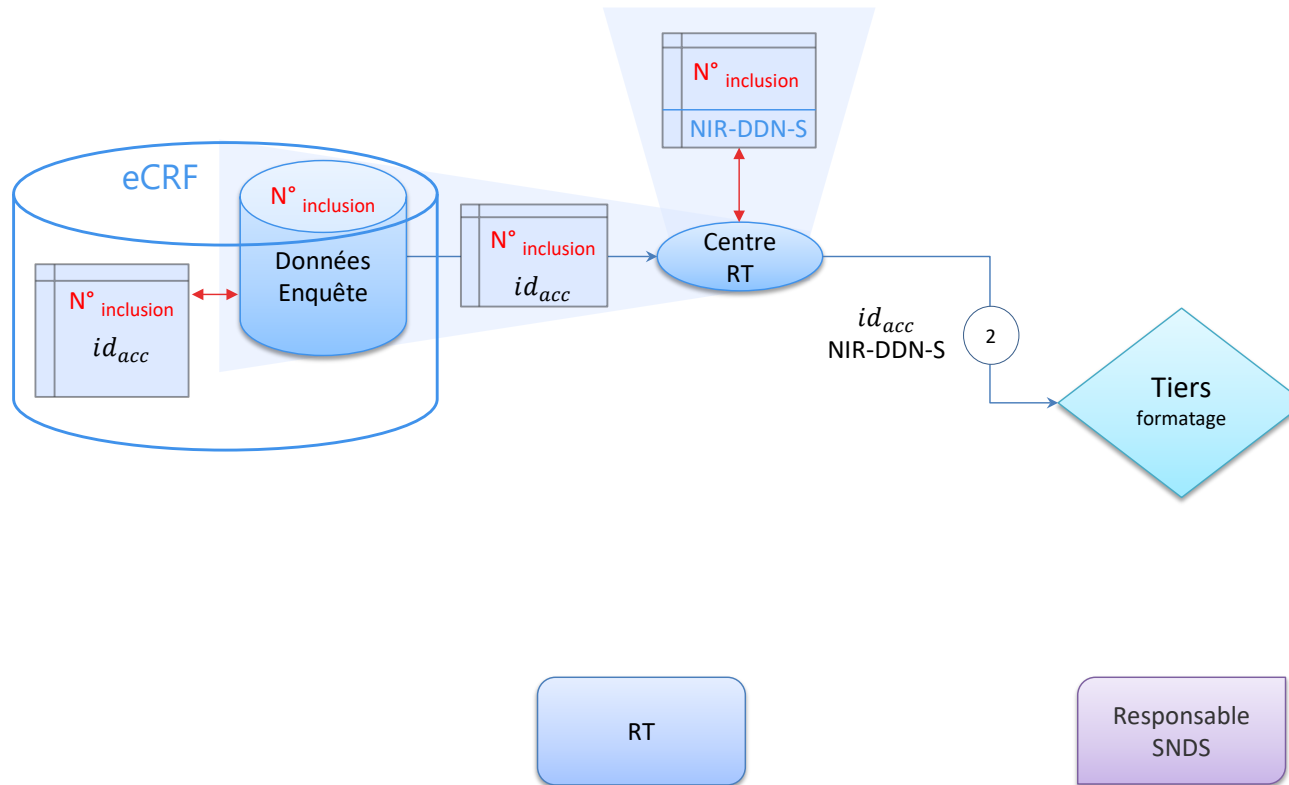


Etape 2 – Envoi des NIR au tiers de formatage

Le centre reçoit la liste des numéros d'inclusion de ses participants, accompagnés des identifiants d'accrochage correspondants

- Par principe, **le centre n'a pas d'accès direct à la table de correspondance du eCRF** entre les numéros d'inclusion et les identifiants d'accrochage. En effet, le centre ne peut être destinataire de cette table qu'au moment de l'appariement.

- L'extraction de cette table va donc être réalisée soit manuellement par un administrateur habilité du eCRF, à la demande du RT, soit par une fonction interne au eCRF qui sera déclenchée par le RT (cf. étape 1), soit par une fonction interne déclenchée par le centre dans son interface eCRF.

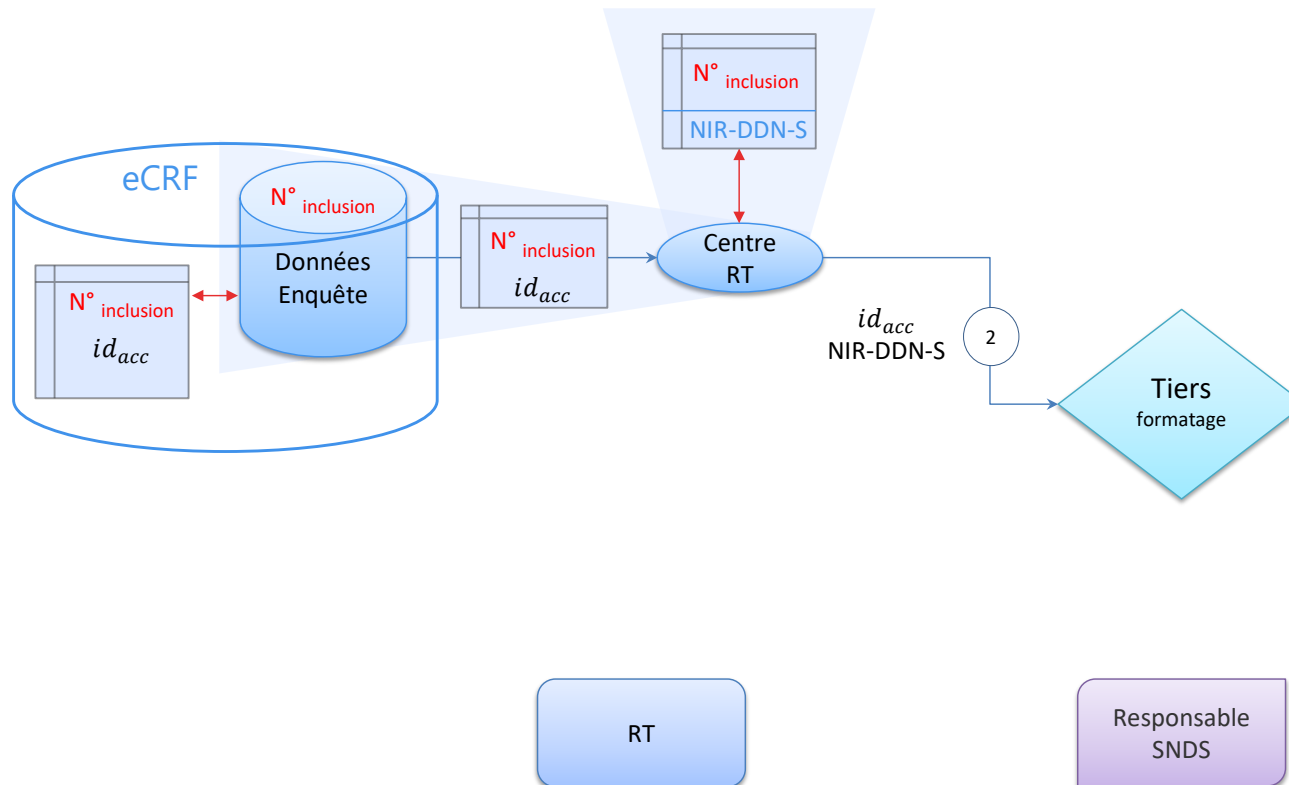


Etape 2 – Envoi des NIR au tiers de formatage (suite)

Le centre transmet au tiers de formatage les identifiants d'accrochage accompagnés des [NIR - Date de Naissance - Sexe] de ses participants

- À partir des numéros d'inclusion reçus et à l'aide de sa table de correspondance locale, le centre va générer un fichier associant les identifiants d'accrochage reçus avec les [NIR - Date de Naissance - Sexe] correspondants. Ce fichier va être transmis au tiers de formatage, comme seules données strictement nécessaires au circuit d'appariement.

- En effet, par principe, **le numéro d'inclusion n'est pas transmis avec le NIR**, ces deux identifiants posant par nature un risque élevé de réidentification.
- De même, **le tiers n'a pas d'accès direct** aux données stockées dans le centre, et aucune autre donnée personnelle liée à l'enquête ne lui est transmise.
- Enfin, **le centre ne conserve pas le fichier contenant les NIR**, qui doit être détruit juste après l'envoi au tiers, **de même que celui qui contenait les numéros d'inclusion**.

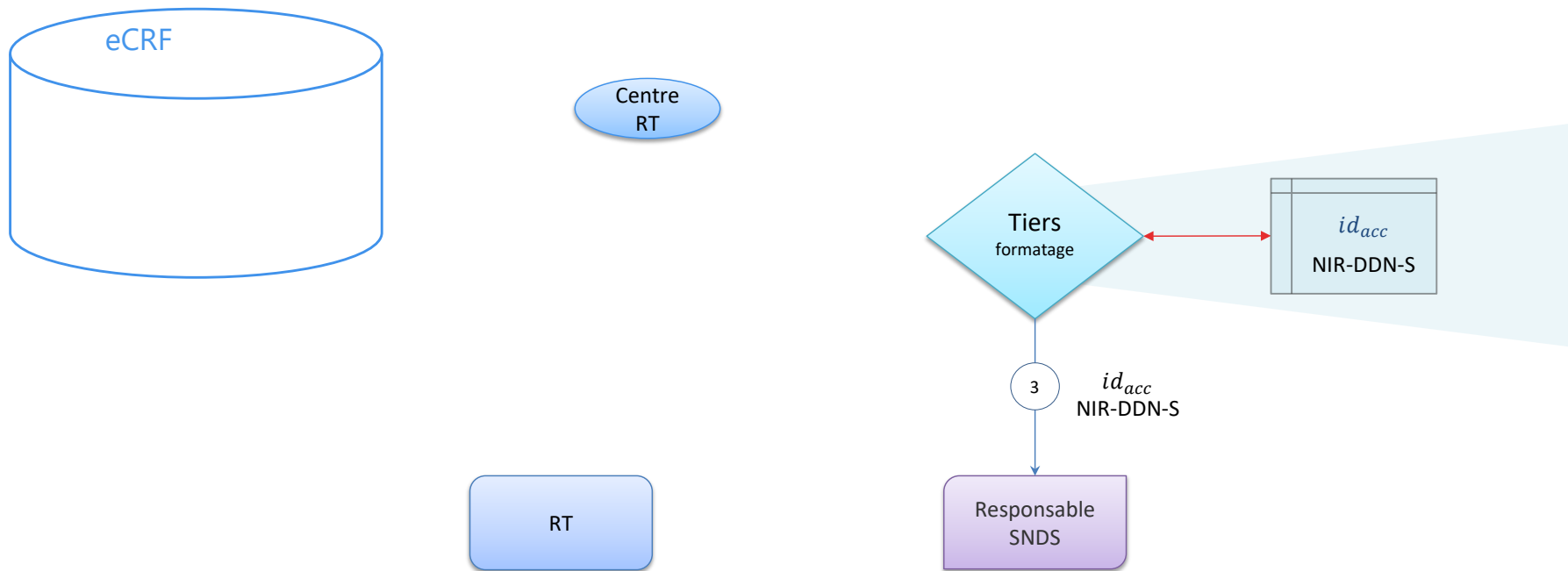


Etape 3 – Envoi des NIR au responsable SNDS

Le tiers de formatage envoie au responsable de la base SNDS les identifiants d'accrochage accompagnés des [NIR - Date de Naissance - Sexe]

- Cet envoi doit se faire à l'aide du téléservice « SAFE » de la CNAM, dans un fichier unique et au format adéquat.

- Le tiers a ici pour rôle de formater les données reçues du centre et d'opérer le téléservice « SAFE ».
- Si le besoin est dûment justifié (par ex. transmission récurrente, audit des données, alerte des participants), le tiers peut conserver le fichier transmis, de manière sécurisée.

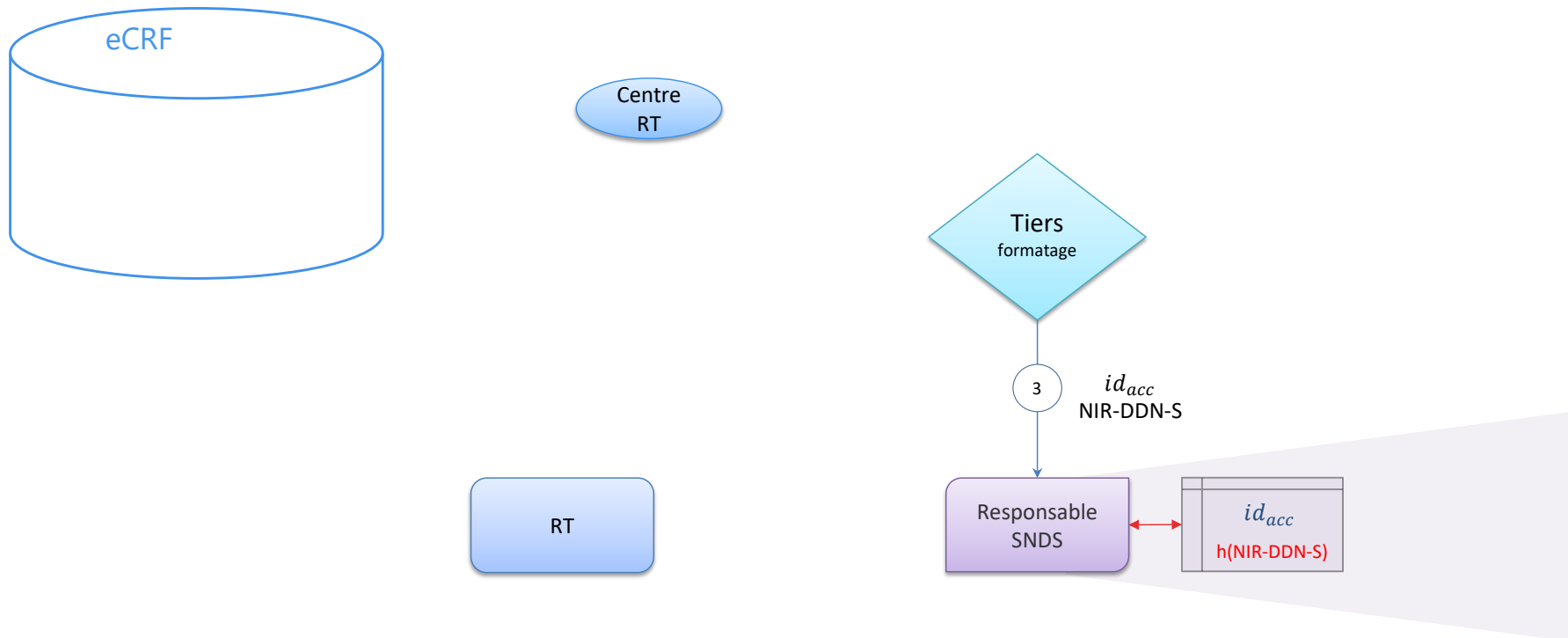


Etape 3bis - Hachage des NIR en entrée

Dès réception, le responsable de la base SNDS procède au « hachage » du triplet [NIR + Date de Naissance + Sexe] pour générer l'identifiant interne du SNDS : $h(\text{NIR-DDN-S})$

- Le hachage désigne ici un calcul cryptographique produisant une pseudonymisation irréversible.

- Dans le cas du SNDS, le NIR est pseudonymisé par plusieurs étapes de hachage successives.
- Par principe, le triplet [NIR + Date de Naissance + Sexe] est remplacé par $h(\text{NIR-DDN-S})$ dès réception des données et il n'est pas conservé, afin de limiter les risques de réidentification des données du SNDS.

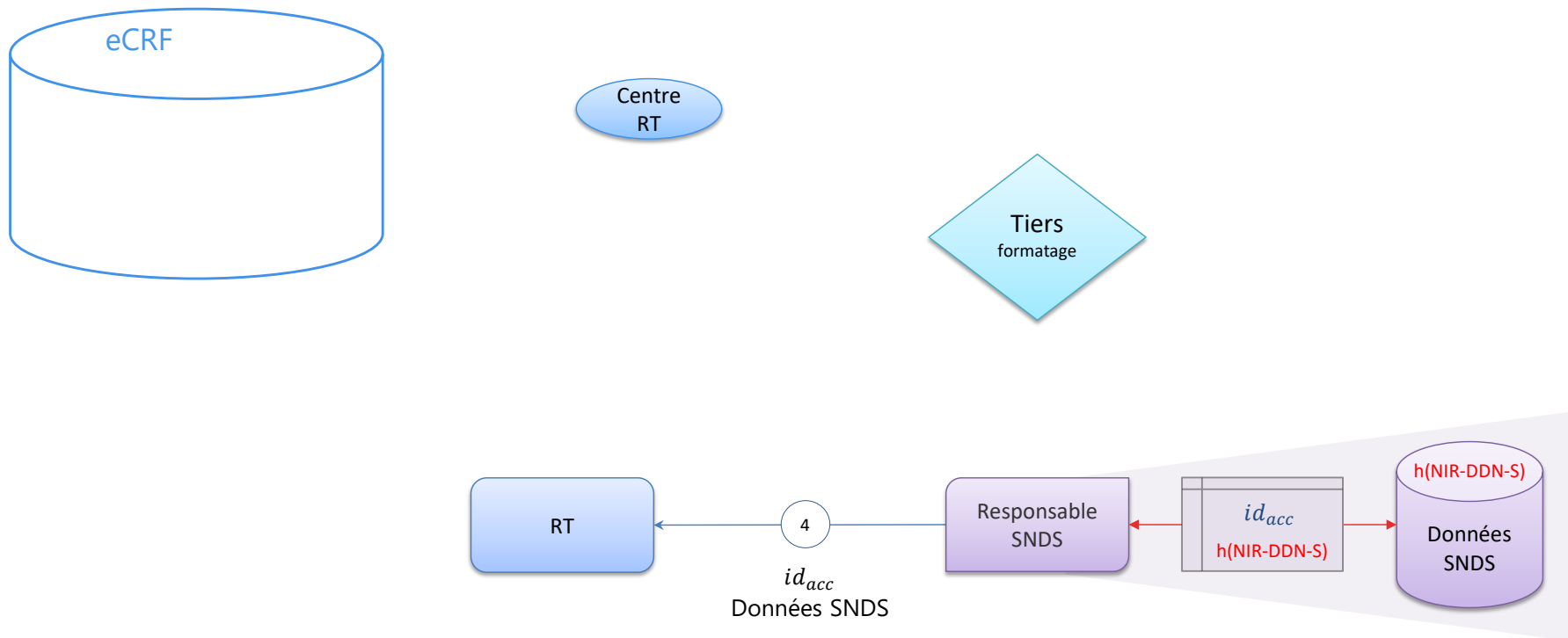


Etape 4 – Extraction et envoi des données SNDS au RT

Le responsable de la base SNDS extrait les données du SNDS correspondant aux $h(\text{NIR-DDN-S})$ des participants

Il transmet les données extraites au responsable de traitement, avec l'identifiant d'accrochage reçu du tiers de formatage

- Afin de limiter les risques de réidentification des données du SNDS, son identifiant interne $h(\text{NIR-DDN-S})$ n'est jamais extrait : **seul l'identifiant d'accrochage est présent avec les données extraites du SNDS.**



Etape 5 – Appariement des données d'enquête avec les données du SNDS, et génération d'un nouvel identifiant pour la base appariée

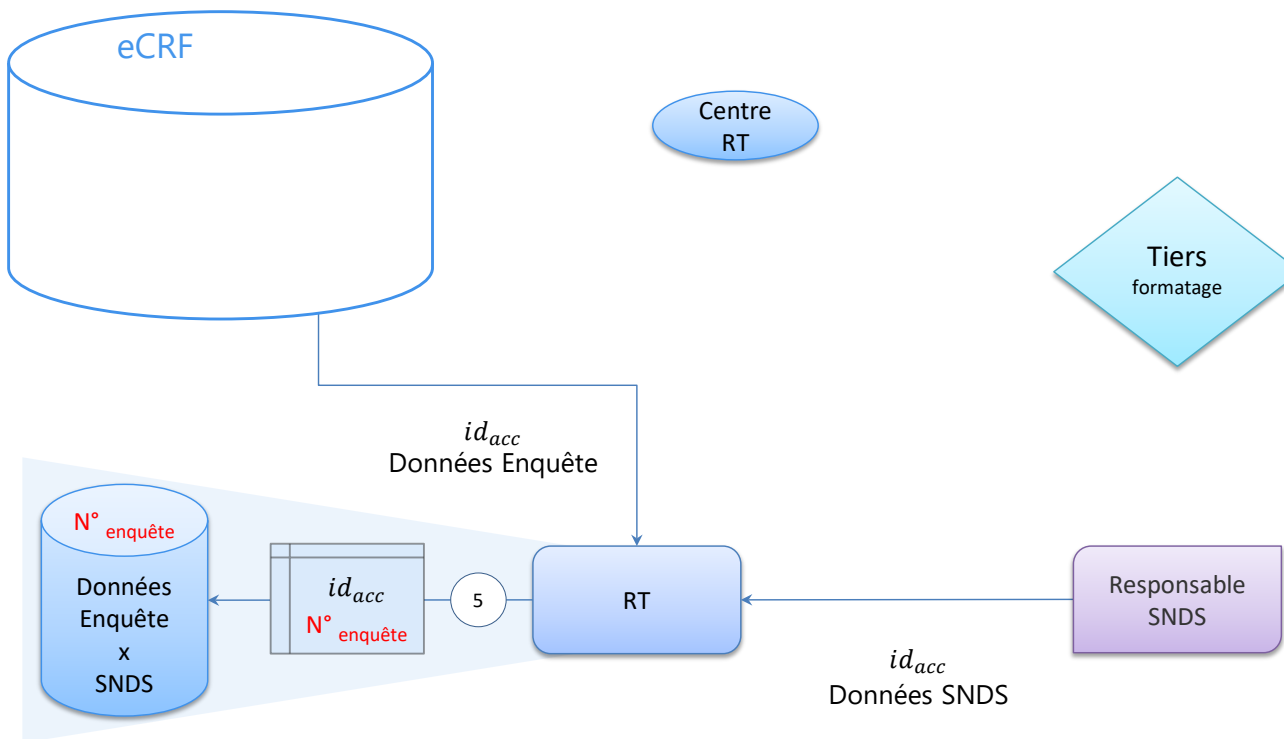
L'appariement doit être effectué sur une plateforme SNDS nationale ou dans un système (« bulle sécurisée ») conforme au [référentiel de sécurité du SNDS](#)

Le responsable de traitement reçoit les données du SNDS et les apparie avec les données d'enquête transmises par le eCRF, à l'aide des identifiants d'accrochage

Après vérification de l'appariement, le responsable de traitement remplace les identifiants d'accrochage par un

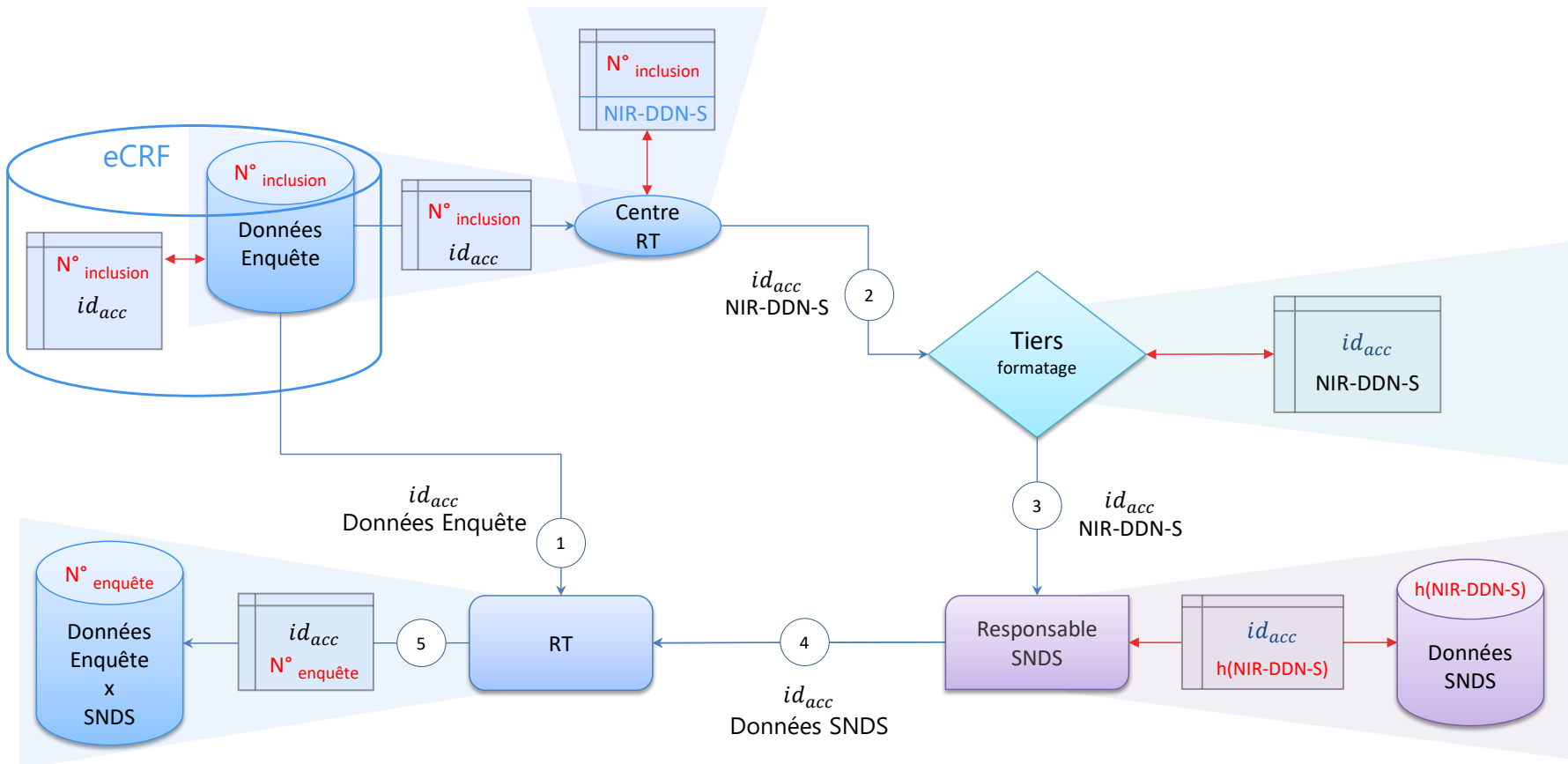
identifiant aléatoire propre à la base des données appariées : $N^{\circ}_{\text{enquête}}$

- Si le besoin est dûment justifié (par ex. transmission récurrente, audit des données, alerte des participants), la table de correspondance entre les identifiants d'accrochage et l'identifiant des données appariées peut être conservée par le responsable de traitement, de manière sécurisée.
- L'identifiant des données appariées peut être généré par une fonction mathématique aléatoire, mais aussi par une fonction de hachage à clé secrète ; dans le second cas, c'est la clé secrète qui sera conservée au lieu de la table de correspondance.



3. Synthèse de l'implémentation du circuit RT / centre unique / eCRF sans NIR

Vue fonctionnelle complète



Vue technique complète

