



MARCHE ORDINAIRE DE PRESTATIONS INTELLECTUELLES

CAHIER DES CLAUSES TECHNIQUE PARTICULIÈRES

Mise en conformité, construction et exploitation d'un Entrepôt de
Données de Santé (EDS) et appariement avec le Système
National des Données de Santé (SNDS)

REFERENCE

N° 2026H001

ETABLISSEMENTS DU GHT ALPES DAUPHINE
concerné par le présent contrat



Centre Hospitalier Universitaire
Grenoble Alpes



ARTICLE 1 – OBJET ET CONTEXTE DU MARCHÉ

1.1 Objet du marché

Le présent Cahier des Clauses Techniques Particulières (CCTP) définit les prestations attendues dans le cadre du chaînage d'un Entrepôt de Données de Santé (EDS) dont le CHU Grenoble Alpes est le responsable de traitement avec le Système National des Données de Santé (SNDS) pour constituer une EDS « fils ».

Ce marché s'inscrit dans la continuité d'une démarche engagée par le CHUGA. L'établissement dispose d'ores et déjà d'un EDS conforme au référentiel CNIL. L'EDS MARS (Multimorbidity Apnea Respiratoire failure Sleep database), qui repose sur la base légale du consentement des patients, est constitué des données de recherche prospective et de suivi courant des patients atteints de maladies du sommeil. A noter qu'un premier prestataire avait été sollicité pour un dépôt dans le cadre du Health Data Hub (HDH) mais qu'il est apparu que notre cas d'usage relevait davantage de la création d'un EDS « fils » que d'un projet à soumettre au HDH.

Ce partenaire s'est alors désengagé du projet mais il avait déjà produit une Analyse d'Impact relative à la Protection des Données (AIPD) et un protocole de recherche partiels. Le présent marché vise à sélectionner un nouveau partenaire capable de reprendre ces travaux et de mener à terme la constitution de l'EDS « fils », depuis la finalisation du dossier CNIL jusqu'à la mise en exploitation opérationnelle.

Ce marché s'articule autour de six Phases fonctionnels, couvrant l'ensemble du cycle de vie du projet :

- Phase 1 – Mise en conformité réglementaire de l'EDS (durée estimée : 4 mois)
- Phase 2 – Construction de l'infrastructure technique de l'EDS (6 mois, mutualisé avec le Phase 3)
- Phase 3 – Appariement avec le SNDS (6 mois, mutualisé avec le Phase 2)
- Phase 4 – Préparation et simplification des bases SNDS (durée variable selon projets)
- Phase 5 – Hébergement et infogérance de l'EDS
- Phase 6 – Mise à disposition d'un espace projet sécurisé (« bulle analytique »)

1.2 Contexte réglementaire et institutionnel

Les données de santé constituent une ressource stratégique pour la recherche médicale, l'épidémiologie et l'amélioration des soins. Leur exploitation est néanmoins soumise à un cadre juridique strict, notamment :

- Le Règlement Général sur la Protection des Données (RGPD – UE 2016/679) ;
- La loi Informatique et Libertés modifiée (loi n° 78-17 du 6 janvier 1978) ;
- Le référentiel EDS de la Commission Nationale de l'Informatique et des Libertés (CNIL) ;
- Les conditions d'accès au SNDS définies par la Caisse nationale d'Assurance Maladie (CNAM) ;
- La politique de sécurité des systèmes d'information de santé (PGSSI-S) de l'ANS (Agence du Numérique en Santé).

Le titulaire devra démontrer une maîtrise approfondie de l'ensemble de ces référentiels et justifier d'expériences similaires dans le secteur de la santé numérique.

1.3 Périmètre fonctionnel global

Rappel :

Le SNDS est géré par la CNAM et rassemble les données de remboursement de l'Assurance Maladie, les données hospitalières (PMSI), les données relatives aux causes médicales de décès, et les données médico-sociales. Son accès est strictement encadré et requiert une autorisation de la CNIL.

L'EDS de l'établissement acheteur vise à constituer une base de données secondaire à partir des données de soins de l'EDS déjà constitué enrichie par le chaînage avec le SNDS, afin de permettre la conduite d'études médico-économiques, épidémiologiques et de recherche clinique dans un environnement conforme au cadre réglementaire. Dans le cadre d'une précédente collaboration, une partie des documents nécessaire au chaînage a déjà été rédigée.

ARTICLE 2 – PHASE 1 : MISE EN CONFORMITÉ DE L'EDS

2.1 Objectifs

Le prestataire accompagne le pouvoir adjudicateur dans l'obtention de l'autorisation CNIL nécessaire à l'exploitation de l'EDS « fils », en conformité avec le référentiel EDS publié par la CNIL. Cette phase conditionne l'ensemble des Phases suivants et doit être menée avec rigueur méthodologique.

2.2 Analyse d'écart

Le titulaire réalise une analyse d'écart exhaustive entre la situation actuelle et les exigences du référentiel EDS. Cette analyse porte notamment sur :

- La cartographie des données traitées, leur nature, leur sensibilité et leur cycle de vie ;
- Les flux de collecte, de pseudonymisation et de stockage existants ;
- Les mesures de sécurité techniques et organisationnelles déjà en place ;
- La gouvernance des données et les rôles des acteurs identifiés (DPO, RSSI, équipe data) ;
- Les écarts entre l'architecture actuelle et l'architecture cible attendue par la CNIL.

Un rapport de l'analyse d'écart est remis à l'établissement dans les quatre semaines suivant le démarrage de la mission.

2.3 Montage du dossier CNIL et gestion des échanges

Le prestataire accompagne l'établissement dans la construction complète du dossier d'autorisation CNIL. Cet accompagnement inclut :

- La rédaction de l'ensemble des pièces du dossier CNIL, incluant notamment la description de la finalité du traitement, la base légale, le périmètre des données traitées et les mesures de sécurité ;
- La description de l'architecture technique de l'EDS (schéma de déploiement, flux de données, modalités de pseudonymisation, cloisonnement des environnements) ;
- La rédaction de l'Expression des Besoins relative à l'appariement SNDS, document opposable qui devra préciser les finalités de recherche, les données nécessaires et les modalités techniques du chaînage ;
- La gestion active des allers-retours (A/R) avec la CNIL, incluant la réponse aux demandes de compléments et la participation aux échanges techniques avec les services instructeurs ;
- La coordination avec le Délégué à la Protection des Données (DPO) de l'établissement, qui devra être associé à toutes les étapes clés.

2.4 Analyse d'Impact relative à la Protection des Données (AIPD)

Le titulaire conduit, ou corédige avec le DPO, l'Analyse d'Impact relative à la Protection des Données (AIPD) conformément à l'article 35 du RGPD. L'AIPD devra couvrir :

- La description systématique et la finalité du traitement ;
- L'évaluation de la nécessité et de la proportionnalité des traitements ;
- L'évaluation des risques pesant sur les droits et libertés des personnes concernées ;
- Les mesures envisagées pour faire face aux risques, y compris les garanties, les mesures de sécurité et les mécanismes visant à assurer la protection des données personnelles.

2.5 Cadrage de la gouvernance de l'EDS

Au-delà des aspects réglementaires, le titulaire accompagne l'établissement dans la structuration de la gouvernance opérationnelle de l'EDS. Cela comprend :

- La définition des rôles et responsabilités des parties prenantes internes (DPO, RSSI, direction médicale, équipes de recherche) ;
- La proposition d'un cadre de gouvernance des accès aux données (procédure de demande d'accès, comité d'accès, traçabilité des utilisations) ;
- L'anticipation de l'architecture technique cible de l'EDS, en lien avec les contraintes exprimées dans le dossier CNIL.

2.6 Livrables du Phase 1

Livrable	Description	Échéance
L1.1	Rapport de l'analyse d'écart	M+1
L1.2	Dossier complet d'autorisation CNIL (avec architecture EDS)	M+3
L1.3	Expression des besoins – appariement SNDS	M+2
L1.4	AIPD finalisée et validée par le DPO	M+3
L1.5	Document de gouvernance de l'EDS	M+4
L1.6	Accusé de réception / décision CNIL	M+4 (dépend CNIL)

ARTICLE 3 – PHASE 2 : CONSTRUCTION DE L'EDS

3.1 Objectifs

Ce Phase couvre la mise en place de l'infrastructure technique sécurisée de l'EDS, conforme aux exigences du référentiel CNIL et aux standards de sécurité du secteur de la santé. La durée de réalisation est de six mois, en parallèle du Phase 3.

3.2 Construction de l'infrastructure

Le titulaire est chargé de la conception, du déploiement et de la configuration de l'infrastructure hébergeant l'EDS. Les exigences techniques minimales sont les suivantes :

3.2.1 Configuration des machines virtuelles (VM)

- Déploiement d'un environnement virtualisé dédié à l'EDS, isolé du reste du système d'information de l'établissement ;
- Configuration des systèmes d'exploitation dans le respect des bonnes pratiques de durcissement (hardening) ;
- Définition et mise en œuvre des politiques de gestion des droits d'accès (contrôle d'accès basé sur les rôles – RBAC) ;
- Paramétrage des ressources de calcul (CPU, RAM, stockage) en adéquation avec les volumes de données attendus et les besoins analytiques.

3.2.2 Mesures de sécurité

Les solutions de sécurité à déployer comprennent au minimum :

- Enregistrement de session (session recording) : journalisation complète des actions réalisées par les utilisateurs disposant de droits d'accès aux données sensibles ;
- Collecte et centralisation des logs : mise en place d'une solution de SIEM (Security Information and Event Management) ou a minima d'un collecteur de logs centralisé, avec politique de conservation conforme aux exigences CNIL ;

- Analyse du trafic réseau : déploiement d'outils de détection d'intrusion (IDS/IPS) et de supervision des flux réseau afin de détecter toute activité anormale ou tentative d'exfiltration de données.

L'ensemble des mesures de sécurité implémentées sera documenté dans un dossier d'exploitation remis à l'établissement et mis à jour à chaque évolution significative.

3.2.3 Environnement analytique

L'EDS doit disposer d'un environnement permettant la réalisation d'analyses statistiques avancées et le développement de modèles d'intelligence artificielle. Le titulaire met en place :

- L'installation et la configuration du langage Python (version LTS recommandée) avec les principales bibliothèques d'analyse de données (pandas, numpy, scikit-learn, matplotlib, etc.) et de machine learning (TensorFlow, PyTorch, ou équivalent) ;
- L'installation et la configuration du langage R avec les packages nécessaires à la bonne réalisation des statistiques en recherche clinique.
- La mise en place d'un environnement de développement interactif (JupyterHub, RStudio Server ou équivalent) accessible par les utilisateurs habilités depuis la bulle analytique.

3.3 Documentation d'exploitation

Le titulaire produit une documentation d'exploitation complète incluant :

- L'architecture technique détaillée (schémas de déploiement, flux de données, matrice de flux réseau) ;
- Les procédures d'administration courante (gestion des comptes, mises à jour, sauvegardes) ;
- Les procédures de gestion des incidents et de reprise d'activité (PRA/PCA) ;
- Le registre des mesures de sécurité mises en place.

3.4 Livrables du Phase 2

Livrable	Description	Échéance
P 2.1	Dossier d'architecture technique détaillée	M+2
P 2.2	Infrastructure EDS déployée et testée	M+5
P 2.3	Environnement analytique (Python, R, notebooks) opérationnel	M+5
P 2.4	Documentation d'exploitation complète	M+6
P 2.5	Rapport de recette technique	M+6

ARTICLE 4 – PHASE 3 : APPARIEMENT SNDS

4.1 Objectifs

Ce Phase couvre l'ensemble des opérations nécessaires à la réalisation du chaînage probabiliste combinatoire entre les données de l'EDS de l'établissement et les données du SNDS. Il est mené en parallèle du Phase 2, sur une durée de six mois à compter de la réception de l'autorisation CNIL.

4.2 Suivi de la convention d'appariement

Le titulaire assiste l'établissement dans la négociation et le suivi de la convention d'appariement à conclure avec la CNAM et/ou le Health Data Hub. Cela comprend :

- La vérification de la complétude et de la cohérence des documents contractuels ;
- Le suivi des jalons administratifs jusqu'à la signature de la convention ;
- La coordination avec les équipes juridiques de l'établissement.

4.3 Réception et sécurisation des données SNDS

À réception des données SNDS dans l'environnement sécurisé de l'EDS, le titulaire effectue :

- La vérification de l'intégrité et de la complétude des fichiers transmis ;
- La mise en place des mesures de sécurité spécifiques requises pour le stockage des données SNDS (cloisonnement, traçabilité des accès, durée de conservation) ;
- La documentation des opérations de réception.

4.4 Chaînage probabiliste combinatoire

Le chaînage entre les données de l'EDS et le SNDS est réalisé selon la méthode du chaînage probabiliste combinatoire, conformément aux procédures définies par la CNAM. Ce processus comprend :

4.4.1 Construction des bases de correspondance

- Création d'une base de données de correspondance à partir des données du SNDS, intégrant les variables identifiantes pseudonymisées nécessaires au chaînage ;
- Création d'une base de données de correspondance à partir des données de la base MARS permettant d'établir le lien entre les identifiants locaux et les identifiants SNDS.

4.4.2 Processus d'appariement

- Lancement de la procédure d'appariement selon l'algorithme validé, intégrant les règles de gestion des doublons, des cas limites et des non-appariés ;
- Calcul et documentation des taux d'appariement obtenus, avec analyse des causes de non-appariement ;

- Mise à disposition des tables chaînées dans l'environnement analytique de l'EDS.

4.5 Contrôle qualité et traçabilité

Toutes les étapes du processus d'appariement sont documentées et tracées. Le titulaire fournit un rapport d'appariement détaillé incluant les taux d'appariement, les statistiques de qualité et les éventuelles anomalies détectées.

4.6 Livrables du Phase 3

Livrable	Description	Échéance
P 3.1	Convention d'appariement signée (avec assistance du titulaire)	M+2
P 3.2	Rapport de réception et contrôle qualité des données SNDS	M+4
P 3.3	Bases de correspondance SNDS et MARS constituées	M+5
P 3.4	Tables chaînées disponibles dans l'EDS	M+6
P 3.5	Rapport d'appariement (taux, qualité, anomalies)	M+6

ARTICLE 5 – PHASE 4 : PRÉPARATION ET SIMPLIFICATION DES BASES SNDS

5.1 Contexte et périmètre

Les données brutes du SNDS nécessitent un travail important de data management avant d'être exploitables dans le cadre d'études. Ce Phase, dont la durée et le périmètre précis dépendent des projets de recherche à venir, couvre l'ensemble des opérations de préparation, de structuration et de simplification des bases SNDS mises à disposition dans l'EDS.

Note importante :

Ce Phase est conditionné aux projets de recherche identifiés conjointement entre l'établissement et le titulaire. Un ordre de service spécifique sera émis pour chaque projet, précisant le périmètre des données à traiter, la méthodologie retenue et les livrables attendus.

5.2 Prestations de data management

Les prestations comprennent notamment :

- Analyse et documentation des tables SNDS disponibles dans l'EDS ;

- Nettoyage et contrôle qualité des données : identification et traitement des valeurs manquantes, des incohérences et des doublons ;
- Recodage et harmonisation des variables selon les nomenclatures de référence (CIM-10, ATC) ;
- Construction de variables dérivées et d'indicateurs utiles aux projets de recherche (comorbidités, scores, trajectoires de soins) ;
- Documentation des algorithmes de construction des variables et des choix méthodologiques (data dictionaries, manuels de codage) ;
- Mise à disposition de tables analytiques simplifiées et documentées à destination des équipes de recherche.

5.3 Livrables du Phase 4 (par projet)

Livrable	Description	Échéance
P 4.1	Plan de data management par projet	J+15 après OS
P 4.2	Rapport de qualité des données	Selon planning projet
P 4.3	Tables analytiques simplifiées et documentées	Selon planning projet
P 4.4	Dictionnaire des variables et manuel de codage	Selon planning projet

ARTICLE 6 – PHASE 5 : HÉBERGEMENT ET INFOGÉRANCE DE L'EDS

6.1 Périmètre de l'hébergement

Le titulaire assure l'hébergement de l'EDS dans un environnement certifié Hébergeur de Données de Santé (HDS) conformément à la loi du 13 octobre 2016 et au référentiel de certification ISO 27001 / HDS. L'ensemble des données traitées dans l'EDS, y compris les données SNDS, sont hébergées exclusivement sur le territoire de l'Union européenne.

6.2 Prestations d'infogérance

Les prestations d'infogérance couvrent :

- La maintenance préventive et corrective de l'infrastructure (systèmes d'exploitation, middleware, bases de données) ;
- La supervision continue de la disponibilité et des performances (monitoring 24h/24, 7j/7) avec astreinte en cas d'incident critique ;
- La gestion des sauvegardes et des tests de restauration selon une politique définie conjointement (RPO, RTO) ;

- La gestion des mises à jour de sécurité (patch management) dans le respect des délais préconisés par l'ANSSI ;
- La gestion des capacités et l'évolutivité de l'infrastructure (scalabilité) en fonction de la croissance des données ;
- Le support technique aux utilisateurs de l'EDS (niveau 2 et 3).

6.3 Niveaux de service (SLA)

Le titulaire s'engage sur les niveaux de service minimaux suivants, précisés dans le contrat :

- Disponibilité de l'environnement EDS : 99,5 % minimum (hors maintenance planifiée) ;
- Délai de traitement des incidents critiques (P1) : prise en charge sous 1 heure, résolution sous 4 heures ;
- Délai de traitement des incidents majeurs (P2) : prise en charge sous 4 heures, résolution sous 8 heures ouvrées ;
- Fréquence des sauvegardes : quotidienne, avec test de restauration mensuel documenté.

ARTICLE 7 – PHASE 6 : ESPACE PROJET SÉCURISÉ (BULLE ANALYTIQUE)

7.1 Objectifs

Le titulaire met à disposition des équipes de recherche un espace de travail sécurisé et isolé (communément désigné « bulle analytique » ou « espace projet »), permettant la réalisation d'études sur les données de l'EDS, y compris les données appariées SNDS, dans un cadre conforme aux exigences réglementaires.

7.2 Caractéristiques techniques de la bulle analytique

La bulle analytique présente les caractéristiques suivantes :

- Isolation complète du réseau extérieur : aucun flux sortant non autorisé, liste blanche des URL et services autorisés ;
- Authentification forte (MFA) pour tout accès à l'espace projet ;
- Traçabilité exhaustive des accès et des actions réalisées sur les données ;
- Impossibilité d'exporter des données brutes en dehors de l'espace sécurisé ; seuls les résultats statistiques agrégés, validés par le gestionnaire de l'EDS, peuvent être extraits ;
- Accès aux outils analytiques (Python, R, SQL, notebooks) installés et maintenus par le titulaire ;
- Gestion des droits par projet et par utilisateur, avec révocation immédiate possible.

7.3 Gestion opérationnelle des projets

Pour chaque étude ou projet de recherche accédant à la bulle, le titulaire assure :

- La création et la configuration de l'espace projet (répertoires, droits d'accès, ressources dédiées) ;
- L'ouverture et la fermeture des accès utilisateurs selon les habilitations validées ;
- Le support technique aux chercheurs dans l'utilisation de l'environnement analytique ;
- L'archivage sécurisé des travaux à l'issue de chaque projet.

7.4 Livrables du Phase 6

Livrable	Description	Échéance
P 6.1	Spécifications techniques de la bulle analytique	M+4
P 6.2	Environnement bulle déployé et testé	M+6
P 6.3	Procédure d'ouverture/fermeture de projets	M+6
P 6.4	Guide utilisateur de la bulle analytique	M+6

ARTICLE 8 – EXIGENCES GÉNÉRALES ET TRANSVERSALES

8.1 Certification et qualifications requises

Le titulaire devra justifier des qualifications suivantes :

- Certification ISO 27001 en cours de validité ;
- Certification Hébergeur de Données de Santé (HDS) en cours de validité pour les activités d'hébergement ;
- Références démontrables dans la mise en œuvre d'EDS en environnement hospitalier ou de santé publique ;
- Expertise documentée dans les procédures d'accès au SNDS et le chaînage de données de santé.

8.2 Organisation de projet

Le titulaire désigne un chef de projet dédié, interlocuteur unique de l'établissement, disposant d'une expérience d'au moins cinq ans dans des projets similaires. Un comité de piPhaseage mensuel réunissant les représentants de l'établissement et du titulaire est mis en place pour le suivi de l'ensemble des Phases.

8.3 Confidentialité et protection des données

Le titulaire s'engage à respecter les obligations issues du RGPD en sa qualité de sous-traitant. Un contrat de sous-traitance conforme à l'article 28 du RGPD est annexé au marché. Le personnel du titulaire amené à accéder aux données de santé signe un engagement de confidentialité individuel.

8.4 Réversibilité

Le titulaire garantit la réversibilité totale des prestations à l'issue du marché. Un plan de réversibilité est fourni dans les trois mois suivant la notification du marché, décrivant les modalités de restitution des données, des documentations et des droits d'accès à l'établissement ou à tout tiers désigné.

ARTICLE 9 – PLANNING PRÉVISIONNEL

Le tableau ci-dessous présente le planning prévisionnel global du projet. Les délais sont exprimés en mois à compter de la date de notification du marché.

Phase / Prestation	M1	M2	M3	M4	M5–6	M6+
Phase 1 – Mise en conformité EDS						
Phase 2 – Construction EDS						
Phase 3 – Appariement SNDS						
Phase 4 – Préparation bases SNDS						Selon projets
Phase 5 – Hébergement / Infogérance						Continu
Phase 6 – Bulle analytique						Continu