

**ACHATS CENTRAUX
HOTELIERS, ALIMENTAIRES
ET TECHNOLOGIQUES**

Hôpital Bicêtre
78, rue du Général Leclerc
94270 Le Kremlin Bicêtre
Tél. : 01 53 14 69 00

**CAHIER DES CLAUSES
TECHNIQUES
PARTICULIERES****n° 26-087**

Procédure : marché passé en procédure adaptée.

Objet : Mise en place d'une plateforme d'hébergement de données de santé destinée à l'organisation de compétitions de data challenges en santé.

Pour une période allant de la notification au 31 décembre 2027 et reconductible expressément deux fois pour une durée de six (par reconduction). Le marché est éventuellement résiliable sans indemnités à la seule initiative de l'Assistance Publique – Hôpitaux de Paris, six mois avant la date de fin du marché.

Ce document comprend 9 pages et est associé au Cahier des Clauses Administratives Particulières

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	1 / 9

SOMMAIRE

ARTICLE 1 : OBJET	3
ARTICLE 2: DECOMPOSITION EN LOT	3
ARTICLE 3: SPECIFICATIONS TECHNIQUES	3
3.1 Présentation du besoin.....	3
3.2 Caractéristiques de la plateforme	3
3.3 Phases des projets.....	5
3.4 Equipe projet du Titulaire.....	9
3.5 Calendrier prévisionnel des prestations.....	9

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	2 / 9

ARTICLE 1 : OBJET

Le marché a pour objet la mise en place d'une plateforme d'hébergement de données de santé destinée à l'organisation de compétitions de data challenges en santé, nécessaire aux besoins de la Direction de la Recherche Clinique et de l'Innovation de l'Assistance Publique - Hôpitaux de Paris.

ARTICLE 2: DECOMPOSITION EN LOT

Les prestations sont réparties en un lot unique au motif que l'allotissement rendrait techniquement plus difficile et complexe l'exécution du marché et le suivi des prestations.

ARTICLE 3: SPECIFICATIONS TECHNIQUES

3.1 Présentation du besoin

Le présent marché porte sur la mise à disposition d'un accompagnement et d'une plateforme de gestion et d'hébergement pour des compétitions d'envergure internationale se basant sur l'analyse d'un large jeu de données de santé anonymisées. L'objectif de ces compétitions est de développer une solution algorithmique innovante, notamment par de l'apprentissage automatique, capable de répondre à une problématique précise.

La plateforme mise en place vise à fournir un environnement en ligne sécurisé où les compétiteurs peuvent s'inscrire, accéder à la description et au règlement de la compétition, accéder aux données, soumettre des modèles prédictifs et travailler de façon collaborative pour résoudre la problématique posée.

La compétition se déroule en plusieurs phases successives :

- L'entraînement des algorithmes sur un premier jeu de données d'apprentissage ;
- Une phase de test intermédiaire où les compétiteurs peuvent tester leurs algorithmes sur un nouveau jeu de données et permettant l'affichage d'un classement intermédiaire des modèles ;
- Une évaluation finale des algorithmes sur un dernier jeu de données ;
- La récompense des compétiteurs ayant développé les modèles prédictifs les plus performants.

Chaque compétition est rattachée à un projet dit de « Data challenges en santé » porté par une équipe au sein de l'Assistance Publique – Hôpitaux de Paris (AP-HP).

Trois projets intègrent le périmètre du marché. Ces derniers sont nommés : AID-ORAL, MYOCARDIA et DATACESAR.

3.2 Caractéristiques de la plateforme

Le Titulaire doit mettre à disposition une plateforme pouvant accueillir les challenges. Cette plateforme doit disposer des fonctionnalités suivantes :

- Portail composé de plusieurs pages contenant :
 - Description détaillée du challenge : présentation du contexte et des enjeux autour du challenge, de la problématique précise à laquelle on cherche à répondre.
 - Description des données mises à disposition : type, format, conditions d'accès, terminologie, labels, etc.
 - Calendrier et étapes du challenge.
 - Informations relatives à la politique de remise des prix.
 - Informations relatives aux organisateurs de la compétition, aux porteurs du projet et aux sponsors.
 - Informations relatives à l'inscription et formulaire d'inscription.

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	3 / 9

- Conditions Générales d'Utilisation (CGU) de la plateforme.
- Règlement intérieur de la compétition.
- Instructions et espaces dédiés pour la soumission et l'évaluation des modèles prédictifs : format attendu, métrique d'évaluation, boîte de dépôt des modèles prédictifs, limites et autorisations pour les dépôts.
- Ressources d'aide au bon déroulement du challenge : forum de discussion, page d'aide, bibliothèques disponibles, webinar. Le chat / forum est spécifique à chaque projet il permet les échanges entre les participants et les mentors du Challenge pour permettre de présenter les questions médicales ainsi que le set de données d'apprentissage.
- Affichage d'un classement en temps réel des participants sur un leaderboard (public et privé) selon leurs scores de performance.
- Annonce et mise en valeur des résultats et des lauréats à la fin du challenge.
- Inscription des compétiteurs via un formulaire d'inscription.
- Accès à la plateforme par les utilisateurs avant le début du challenge via une landing page, dédiée au Data Challenge sur laquelle les potentiels compétiteurs peuvent retrouver les informations clés et pouvoir se pré-enregistrer ou s'inscrire. Un accès à la plateforme, pendant toute la durée de la compétition pour y participer, et également après la fin du challenge afin de consulter les résultats.
- Assurer la mise à disposition d'une infrastructure d'hébergement sécurisée pour stocker les données de manière fiable, garantissant la disponibilité et la confidentialité des informations tout au long du déroulement du Data Challenge.
- Intégration d'une infrastructure de calcul afin de fournir les ressources de calcul nécessaires pour traiter les données de manière efficace, permettant ainsi une exécution rapide des algorithmes d'apprentissage automatique.
- Test de bout en bout de la plateforme : soumission test, évaluation, affichage leaderboard.
- Monitoring de la compétition permettant de détecter les comportements anormaux, les tentatives de fraude.
- Support par l'équipe de la plateforme, qui s'engage à répondre aux organisateurs et aux compétiteurs dans un délai maximal de 48 heures ouvrables. De plus, l'équipe de la plateforme veillera à ce qu'aucun problème technique ne persiste plus de 24 heures ouvrables.
- Pour chaque projet, le nombre d'inscrits estimé devrait être estimé à plusieurs centaines de personnes regroupés par équipe. Le nombre de compétiteurs par équipe peut varier selon le projet.
- Disposer d'une plateforme ayant une puissance de calcul et de stockage suffisante pour supporter le(s) challenge(s) et pouvant être adapté si besoin.
- Plateforme et espace d'hébergement des jeux de données sécurisés devant se situer en Union Européenne et de préférence en France.
- Accès sécurisé et journalisé (logs traçant l'identité de l'utilisateur, la date, et les actions effectuées).

Précisions complémentaires :

- Les compétiteurs peuvent utiliser leur propre environnement ou l'environnement mis à disposition par le Titulaire ;
- Certains projets impliquent le traitement d'images médicales pouvant nécessiter des ressources de calcul GPU (processeurs graphiques) pour l'entraînement de modèles de deep learning.
- Le machine learning doit être compatible avec les bibliothèques Python standards.

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	4 / 9

3.3 Phases des projets

La mise en place de chaque compétition se fait par phase. Ces dernières sont les suivantes.

Propos liminaires : ci-après est détaillé le besoin pour les 3 projets de data challenges précités. Les dispositions en noires sont générales et s'appliquent à l'ensemble des projets. Toutefois, certains projets répondent à des spécificités d'exécution, ces dernières sont présentées dans un encadré bleu.

3.2.1 Phase 1 : Initialisation et préparation de l'environnement de compétition

En amont de la compétition, les prestations suivantes sont réalisées par le Titulaire seul ou concomitamment avec l'équipe projet de chaque data challenge :

- Une réunion de lancement est organisée entre le Titulaire et l'équipe de chaque projet. Cette équipe est notamment composée de médecin porteur du projet, chef de projet, chargé de suivi administratif et financier.

Cette réunion a pour objectif de définir de manière précise et détaillé le calendrier de chaque projet, leurs jalons ainsi que les responsabilités respectives de chacune des parties. Au cours de cette réunion est également présenté, coté Titulaire, le chef de projet.

Cette réunion se tient dans le mois suivant la notification du marché ou dans le moins suivant l'information au Titulaire de la mise en place d'un nouveau projet. Cette réunion se tient en distanciel.

A la suite de la réunion, et dans les 3 jours ouvrés qui suivent, le Titulaire rédige et envoie à l'équipe projet de chaque data challenge le compte-rendu de cette réunion ainsi que le calendrier détaillé de l'ensemble des phases. L'équipe projet valide ces derniers sous un délai de 5 jours ouvrés. Le cas échéant, le Titulaire retravaille les documents selon les recommandations ou demandes de ladite équipe.

Projets AID-ORAL et MYOCARDIA : la réunion de lancement de chaque projet a lieu dans le mois suivant la notification.

Projet DATACESAR : la réunion de lancement a lieu dans le mois suivant l'information au Titulaire de la mise en place du nouveau projet.

- Consultation et accompagnement de chaque équipe projet sur la formulation de la problématique et la transposition de la question médicale vers une problématique qui puisse être résolue par une solution d'apprentissage automatique.
- Accompagnement dans la définition de la métrique d'évaluation pour qu'elle soit la plus pertinente et appropriée pour évaluer les performances des algorithmes développés lors de la compétition.
- Point sur les données de l'AP-HP mises à disposition dans le cadre des compétitions. Le Titulaire effectue un point sur la collecte des données de l'AP-HP qui seront mises à disposition des candidats lors de la compétition.

Pour faire ce lien, le Titulaire :

- Effectue un état des lieux de ces données : type, volume, taille ;
- Discute conjointement avec l'AP-HP du format de données pour l'intégration dans la plateforme (formats d'export disponibles côté AP-HP, formats d'imports supportés par le prestataire, éventuelles conversions nécessaires). Si une conversion est nécessaire, les deux parties conviendront ensemble de la solution technique la plus adaptée ;

L'équipe projet valide conjointement avec le Titulaire le format d'intégration des données. Cette validation s'effectue sur un délai de 10 jours ouvrés.

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	5 / 9

- Réalisation de tests préliminaires afin d'évaluer la faisabilité du Challenge, en vérifiant la disponibilité des données, la fonctionnalité de la plateforme, et en assurant la capacité à gérer le flux des participants et des soumissions de modèles avant le lancement officiel de la compétition.
- Accompagnement dans la rédaction des éléments de contexte à destination des compétiteurs.
- Communication préliminaire auprès de potentiels compétiteurs et de la communauté de data scientists de la plateforme, dans le but d'attirer un maximum de participants.
- Gestion et enregistrement des inscriptions. Les compétiteurs se regroupent en équipe de 3 à 5 membres avant d'être invités sur la plateforme. Le cas échéant si les inscrits ne parviennent pas à former des groupes, le Titulaire pourrait être sollicité en ce sens.
Le Titulaire doit également réaliser des reportings réguliers (par mail chaque vendredi pendant cette phase) à l'AP-HP sur l'état des inscriptions : nombre d'inscrits, profils des participants (académiques, industriels), provenance géographique, afin de suivre la dynamique de participation et d'ajuster la stratégie de communication si nécessaire.
- Conception du contenu de la plateforme selon les caractéristiques précitées.
- Communication sur les challenges et appel à concourir :
 - Préparation des supports de communication et d'appel à participation à destination des futurs participants : landing page dédiée au Data Challenge, visuels et bannières promotionnelles pour les réseaux sociaux et sites partenaires, textes de présentation synthétiques adaptés aux différents canaux de diffusion (newsletters, posts, mails d'invitation), vidéo ou webinar de présentation du challenge (contexte médical, données, règles de participation) ;
 - Diffusion des supports et des informations relatives aux compétitions. Cette stratégie vise à maximiser le nombre de participants. A titre d'exemple, ces supports sont : communauté propre du prestataire (base de data scientists inscrits sur la plateforme), réseaux académiques et de recherche (universités, laboratoires, écoles d'ingénieurs), communautés spécialisées en IA, machine learning et data science (forums, groupes, meetups), canaux de communication de l'AP-HP et du Health Data, réseaux sociaux professionnels et scientifiques (LinkedIn, X/Twitter), relais auprès de sociétés savantes et congrès médicaux pertinent.
La landing page dédiée à chaque data challenge doit être mise en ligne huit semaines avant la date prévue pour le lancement du challenge.

Le contenu de la landing page et des supports de communication sont expressément validés par voie écrite par l'équipe projet AP-HP avant toute diffusion. Cette validation se fait sous un délai de 10 jours ouvrés.

Cette campagne de communication doit durer 3 mois pour donner de la visibilité aux challenges.

- Élaboration et publication d'un règlement intérieur selon un modèle de Titulaire. Ce règlement doit préciser les modalités de participation, les critères d'évaluation et métrique d'évaluation, les conditions de remise des prix, les règles de propriété intellectuelle, les mentions relatives à la protection des données personnelles et la conformité légale pour un concours international (aspects transfrontaliers, versement d'un prix à l'étranger, etc.)
Après sa rédaction, le règlement est communiqué aux équipes projet pour validation expresse avant toute publication et utilisation. Après envoi, l'AP-HP valide ce dernier sous un délai de 10 jours ouvrés.
- Transmission des jeux de données au Titulaire. Les jeux de données sont transférés sur la plateforme du Titulaire par l'AP-HP. Pour ce faire, le Titulaire doit mettre en place un canal sécurisé. Ce transfert s'opère via un protocole de transfert chiffré, type SFTP ou équivalent.

Ce transfert des données réalisé par l'AP-HP se déroule en deux étapes :

1. Transfert d'un jeu de données test (échantillon réduit) permettant de valider la compatibilité technique avec la plateforme du prestataire (formats, intégration, bon fonctionnement du pipeline d'évaluation) ;
2. Transfert du jeu de données complet une fois la compatibilité technique confirmée.

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	6 / 9

Une vérification conjointe de l'intégrité des données est réalisée après chaque transfert. Cette vérification comporte :

- Un contrôle de la complétude : nombre de fichiers reçus conforme au nombre de fichiers envoyés ;
- Un contrôle de l'intégrité : vérification que les fichiers n'ont pas été corrompus ou altérés pendant le transfert ;
- Un contrôle de la lisibilité : ouverture et lecture d'un échantillon de fichiers pour s'assurer qu'ils sont exploitables sur la plateforme.

Cette vérification se fait conjointement dans les 5 jours ouvrés après chaque transfert effectué.

Avant le transfert du jeu de données global, celui-ci est divisé par l'AP-HP en trois sous-ensembles :

- Jeu d'entraînement (train) : données avec leurs labels (réponses connues), mises à disposition des participants pour développer et entraîner leurs algorithmes ;
- Jeu de test public : données sans labels. Les participants soumettent leurs prédictions sur ce jeu pendant la compétition, et obtiennent un score affiché sur le leaderboard ;
- Jeu de test privé : données sans labels, inconnues des participants pendant toute la compétition. Le classement final est calculé sur ce jeu à la clôture de la compétition, afin de garantir que les algorithmes généralisent bien sur des données jamais vues.

La constitution de ces trois jeux (choix de la répartition, sélection des données) est réalisée par l'AP-HP. Le prestataire est chargé de les intégrer dans sa plateforme et de configurer le système d'évaluation automatique associé.

- Préparation d'un starter kit / notebook de démonstration pour les participants.
 - Réalisation d'un bon fonctionnement de la plateforme par l'équipe projet AP-HP avant ouverture de chaque compétition. Cela a pour objectif d'effectuer : la vérification du bon fonctionnement du pipeline de soumission, de l'évaluation automatique, du leaderboard et du starter kit. Cette validation se fait dans un délai de 10 jours ouvrés.
- L'ouverture de la compétition est conditionnée à cette validation.

3.2.2 Phase 2 : Compétition

Pendant la compétition, les prestations suivantes sont réalisées par le Titulaire seul ou conjointement avec l'équipe projet de chaque data challenge :

- Mise en ligne du portail dédié à chaque data challenge après autorisation écrite de l'AP-HP et après test du bon fonctionnement de la plateforme pour le challenge concerné (comme indiqué ci-dessus) ;
- Mise à disposition des compétiteurs du premier jeu d'entraînement aux participants inscrits (téléchargement autorisé mais traçable) ;
- Gestion du système de soumission des prédictions ;
- Gestion du chat / forum intégré à la plateforme ;
- Organisation d'un webinar pour expliquer le challenge, présenter le set de données du challenge et permettre un matchmaking pour compléter les équipes (40 au maximum) ;
- Mise à disposition d'un espace de dépôt des algorithmes intermédiaires pour une évaluation sur un jeu de données de test, qui est différent de la base de données d'apprentissage.
- Évaluation des performances des algorithmes produits selon une métrique d'évaluation prédéfinie à la fin de la compétition.
- Évaluation automatique et réalisation d'un classement intermédiaire en temps réel, tout au long de la phase d'entraînement et publication et mise à jour de ce classement en temps réel (leaderboard).
- Surveillance constante pour garantir le bon déroulement de la compétition, détecter les problèmes techniques et prévenir la triche.
- Réalisation d'un classement final qui n'est rendu public qu'après la clôture de la compétition.
- Gestion et mise à disposition du support technique.
- Signalisation de tout incident à l'équipe projet AP-HP sous 4 heures ouvrées. Tout changement de règles ou de fonctionnement en cours de compétition nécessite une validation de l'équipe projet AP-HP se faisant sous 2 jours ouvrés.

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	7 / 9

Clôture de la compétition : fermeture des soumissions, gel du leaderboard privé, calcul du classement final et communication des résultats préliminaires aux compétiteurs. Validation par l'équipe projet AP-HP de la date de clôture et de la procédure de gel du leaderboard, sous 5 jours ouvrés avant la date prévue

3.2.3 Phase 3 : Analyse des résultats et remise des prix

A la suite de la compétition, le Titulaire doit procéder en deux temps.

Dans un premier temps, il effectue l'ensemble des prestations inhérentes à l'analyse des résultats. Cela comporte un(e) :

- Exportation et transmission au jury AP-HP de l'ensemble des résultats (classements, scores par métrique, historique des soumissions)
 - Accès aux prédictions des meilleures équipes pour analyse qualitative
- L'équipe de chaque projet procède à la validation de la complétude et de l'exactitude des résultats transmis (classements, scores, historique), sous 10 jours ouvrés.
- Evaluation des algorithmes proposés par les compétiteurs. L'évaluation se fait le jeu de test privé (de taille au moins équivalente au set du jeu d'entraînement) ce qui permettra au jury du Data Challenge d'apprécier :
 - l'adéquation de la Méthodologie I.A à la problématique (vs Data et vs Objectif)
 - la cohérence du modèle final (vs Résultats Cliniques finaux observés) et sa performance prédictive
 - la simplicité et la vitesse d'exécution de l'algorithme sur le plan pratique.

Projet AID-ORAL : Les performances évaluées sont :

- La capacité à discriminer une LOPM d'un CECO
- La capacité à relier photographie de LOPM avec une catégorie diagnostique
- La capacité à relier photographie de CECO avec le diagnostic
- La capacité à relier photographie de LOPM avec un risque de transformation maligne

- Analyse des performances : comparaison des approches et synthèse méthodologique.
- Rédaction d'un rapport de synthèse de la compétition (nombre de participants, évolution des scores, approches utilisées, revue de code, détection de fraude, synthèse méthodologie). Le contenu de ce rapport pourra évoluer selon les demandes des équipes projet. Le rapport doit faire l'objet d'une validation de la part de ces équipes sous 10 jours ouvrés.
- Vérification et validation des algorithmes en tête du classement (top 3 des participants) afin de garantir l'intégrité, l'explicabilité et la fiabilité des résultats, en intégrant des procédures rigoureuses de détection des anomalies et de violations des règles établies (notamment le plagiat).
- Classement définitif des résultats. Ce classement est validé par l'équipe projet avant toute communication publique. Cette validation se fait sous 5 jours ouvrés. Aucune annonce des résultats ne peut être effectuée avant validation écrite de l'équipe projet.

Dans un second temps, il effectue la notification des résultats puis la remise de prix :

- Communication officielle des résultats et des lauréats de la compétition.
- Collecte des informations nécessaires au versement des prix des 3 lauréats.
- Maintien d'une page de résultats accessible sur la plateforme après la clôture du challenge (classement final, description des approches gagnantes, liens vers les publications et les algorithmes en open source le cas échéant)
- Remise des prix de récompense aux lauréats à condition que ces derniers acceptent la publication de leur algorithme en Open Source le mettre à disposition de l'ensemble de la communauté.

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	8 / 9

3.4 Equipe projet du Titulaire

Le Titulaire dimensionne son équipe en nombre suffisant pour permettre la bonne exécution du marché dans les délais impartis.

Cette équipe est notamment composée d'un chef de projet qui assure la planification, la coordination et l'exécution globale du Data Challenge.

3.5 Calendrier prévisionnel des prestations

Phase	Durée estimée
Phase 1 : Initialisation et préparation de l'environnement de compétition	2 à 3/4 mois
Phase 2 : Compétition	3 mois
Phase 3 : Analyse des résultats et remis des prix	3 mois
Durée totale des prestations	8 à 10 mois

A titre informatif uniquement (et non contractuel), la notification du marché interviendrait en mi-juillet 2026.

Sur la base de cette date prévisionnelle de notification, les grandes étapes des Data challenge AID-ORAL et MYOCARDIA seraient :

- Phase 1 : Juillet à Septembre 2026 ;
- Phase 2 : Octobre 2026
- Phase 3 : Janvier à Mars 2026.

Comme présenté à l'article 3.2.1, ce calendrier sera acté lors de la réunion de lancement de chaque projet. Ces phases pourraient potentiellement être mises à jour afin de répondre à la nécessité de chaque projet.

Enfin, pour le projet DATACESAR, les phases ci-dessus débiteront à compter de l'information au Titulaire de la mise en place de ce nouveau projet.

Annexe n°1 : Cadre de réponse technique (CRT)

Document joint au DCE

A.P.-H.P.	Consultation n° 26-087	ACHAT
CCTP.2 18/03/2026	Dernière mise à jour du 15/06/2026	9 / 9