

Table des matières

I. Préambule	2
A. Présentation IHU-ICAN	2
B. Enjeux	2
C. Présentation et objectifs de la plateforme ICAN I/O Data Sciences	2
D. Objectifs de la charte	2
E. Localisation	3
II. Organisation	3
III. Fonctionnement	4
A. Modalités d'utilisation	4
B. Prestations / offres de service	4
C. Facturation	8
D. Responsabilités	8
E. Formation	8
IV. Sécurité	8
V. Partenaires	8
VI. Publication (selected)	8
VIII. Conditions générales d'exécution	11
A. Modalités de citation	11
B. Sanction du non-respect de la charte	11

I. Préambule

A. Présentation IHU-ICAN

Situé à l'Hôpital de la Pitié-Salpêtrière à Paris, la Fondation de Cardiométabolisme et Nutrition (IHU-ICAN) a été officiellement inauguré en novembre 2011 parmi les six centres d'excellence scientifique en France. Le financement de cet Institut a été accordé dans le cadre du programme d'investissements d'avenir initié par le gouvernement français pour faire face aux enjeux de santé en Europe et dans le monde.

L'IHU-ICAN s'appuie sur les forces et l'expertise des unités de recherche médicale et scientifique de l'INSERM et de Sorbonne Université et des équipes médicales de l'AP-HP pour développer la médecine du futur dans le domaine du cardiométabolisme et de la nutrition.

La communauté ICAN rassemble des chercheurs et des cliniciens du cœur et du métabolisme ainsi que des paramédicaux (infirmiers, diététiciens, psychologues, etc.).

B. Enjeux

L'IHU-ICAN vise à lutter contre l'obésité, le diabète, les maladies cardiovasculaires, la NASH (stéatose hépatique non alcoolique) et les dyslipidémies en traduisant les découvertes engendrées par la recherche centrée sur le patient en innovations thérapeutiques et diagnostiques. Pour cela, IHU-ICAN met en œuvre une approche pluridisciplinaire.

Les grands objectifs d'IHU-ICAN sont de :

- Développer une recherche translationnelle dans le domaine des maladies cardiométaboliques à l'échelle internationale.
- Développer une médecine personnalisée dans laquelle les innovations sont traduites en soins.
- Former les futurs professionnels de santé.
- Valoriser la recherche via des partenariats public/privé.
- Disséminer la connaissance scientifique au sein des professionnels et du grand public.

C. Présentation et objectifs de la plateforme ICAN I/O Data Sciences

ICAN I/O Data Sciences est une plateforme d'expertise axée sur l'analyse et l'intégration de données multi-omiques et cliniques. La plateforme va de pair avec l'évolution des technologies modernes et l'utilisation d'algorithmes robustes pour répondre aux objectifs de recherche grâce à une approche collaborative et holistique intégrant les technologies multi-omiques et l'analyse statistique.

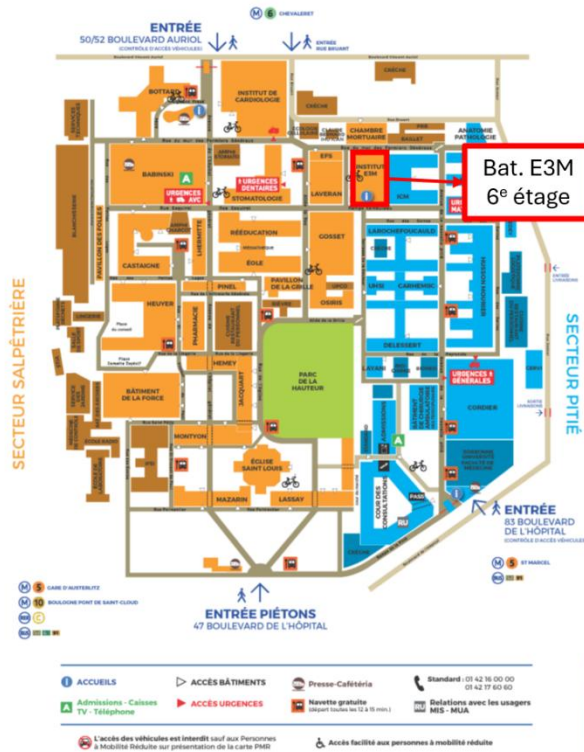
L'équipe comprend des data managers et des data scientists qui possèdent tous une grande expérience acquise en travaillant sur des projets nationaux et internationaux, en collaboration avec d'éminents scientifiques, médecins et chercheurs sur de nombreux sujets scientifiques.

D. Objectifs de la charte

La présente charte a pour vocation de régir les rapports entre l'IHU-ICAN et les utilisateurs de la plateforme ICAN IO-Data Sciences.

Les termes de la présente Charte s'imposeront à toute personne accédant aux services de la plateforme ICAN IO-Data Sciences.

E. Localisation



Adresse de la Plateforme ICAN I/O Data Sciences :

Hôpital de la Pitié Salpêtrière
 Bâtiment E3M 6^{ème} étage
 75013 Paris

II. Organisation

Le pôle est composé des personnels suivants :

Responsable ICAN I/O Data Sciences : Dr. Maharajah Ponnaiah

Data Scientists : Dr. Maharajah Ponnaiah, Dr. Pauline Stephen Raj, Mme. Rose Tchala Sare (M1 Alternance)

Data Managers : Dr. Pauline Stephen Raj, Dr Mehdi Menai

Coordinateur des plateformes : Dr. Ludovic Le Chat

Comité Scientifique :

Dr. Antonio GALLO, MD, PhD

Maître de Conférence de Universités - Praticien Hospitalier, Unité de Lipidologie et Prévention Cardiovasculaire.

Service de Nutrition (Pr Jean-Michel Oppert)

DMU ARCHIMEDE

Hôpital Pitié-Salpêtrière, APHP,

83 bd de l'hôpital, 75013 PARIS, France

Pr. Matthieu Schmidt, MD, PhD

PU-PH; Service de Médecine Intensive Réanimation Institut de Cardiologie, Hôpital Pitié-Salpêtrière ICAN, Sorbonne Université, Groupe de recherche clinique RESPIRE

47-83 bd de l'Hôpital, 75013 PARIS, France

Dr. Wilfried Le Goff

Head of the SLIM research team –

Delegate Vice-Dean Research of the Faculty of Medicine

Inserm | Sorbonne Université UMR_S1166 ICAN

91 boulevard de l'Hôpital | 75013 PARIS, France

Dr. Xavier Fresquet

Deputy Director

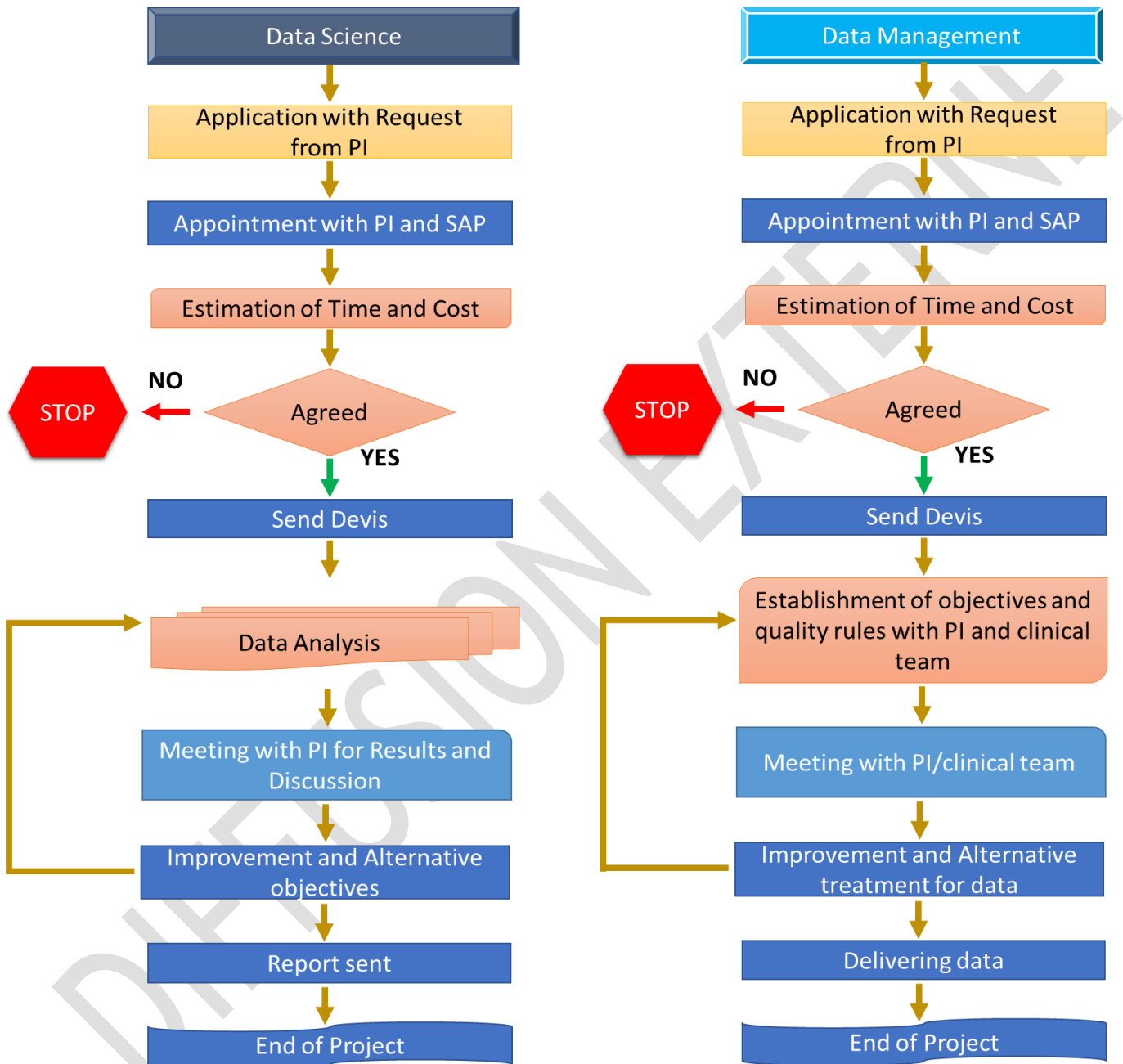
Sorbonne Center for Artificial Intelligence (SCAI)

4, Place Jussieu, 75005 PARIS, France

III. Fonctionnement

A. Modalités d'utilisation

La plateforme ICAN I/O Data Sciences est accessible à toutes les équipes de recherche membres de l'IHU-ICAN ainsi qu'aux équipes extérieures (académiques ou entreprises privées). Les différentes étapes de la soumission à la réalisation d'un projet sont exposées ci-contre :



B. Prestations / offres de service

Gestion et curation des bases de données

- **Gestion du CRF papier :**
 - ✓ **Lecture du protocole :** Examen du protocole pour aligner la conception du CRF en conséquence.
 - ✓ **Révision du CRF papier :** Ajustements nécessaires du CRF papier selon les exigences du protocole et les retours.
 - ✓ **Validation du CRF papier :** Confirmation que le CRF papier répond à toutes les exigences réglementaires et de l'étude.
- **Création de l'eCRF :**
 - ✓ **Liste de variables :** Identification de toutes les variables à inclure dans le Formulaire de Rapport de Cas Électronique (eCRF).
 - ✓ **Implémentation dans REDCap :** Configuration de l'eCRF au sein du système REDCap.
 - ✓ **Test et validation du CRF :** Assurer le bon fonctionnement de l'eCRF et la collecte des données requises.
 - ✓ **Création des comptes utilisateurs :** Mise en place des comptes pour tous les utilisateurs qui accéderont à l'eCRF.
 - ✓ **Mise en production :** Lancement actif de l'eCRF finalisé pour l'utilisation dans l'étude.
 - ✓ **Edit Checks :** Mise en place de contrôles logiques dans l'eCRF pour valider la saisie des données en temps réel.
 - ✓ **Validation des données et contrôle de qualité**
 - ✓ **Création et validation du Plan de Validation des Données :** Développement et approbation d'un plan complet pour garantir la qualité des données tout au long de l'étude.
 - ✓ **Sélection des champs obligatoires—Développement, Test, Validation :** Identification, test et validation des champs essentiels dans l'eCRF nécessitant une saisie de données.
- **Extraction des données :**
 - ✓ **Extraction des données / Création des datasets :** Extraction des données du système et structuration en datasets pour l'analyse.
- **Documentation :**
 - ✓ **Guide d'utilisation REDCap :** Instructions sur l'utilisation du système REDCap.
 - ✓ **Guide de saisie de l'étude :** Guide détaillé pour la saisie des données spécifique à l'étude.
 - ✓ **Plan de Gestion des Données de l'étude :** Description des processus et responsabilités pour la gestion des données de l'étude.
- **Intégration et données externes**
 - ✓ **Lien avec les données externes :** Intégration des données provenant de sources externes avec les données de l'étude maintenues dans l'eCRF.
- **Maintenance de l'eCRF et contrôle qualité des données**
 - ✓ **Correction des problèmes :** Adresse et résolution de tous les problèmes survenant avec l'eCRF.

- ✓ **Nouvelle version / Amendement** : Mise à jour de l'eCRF pour s'adapter aux modifications du protocole de l'étude ou à d'autres exigences.
- ✓ **Gestion des utilisateurs** : Supervision de l'accès à l'eCRF et des rôles des utilisateurs.
- ✓ **Établissement de requêtes** : Mise en place de requêtes pour aborder les incohérences de données identifiées lors de la surveillance des données.
- **Monitoring**
 - ✓ **Mise à disposition du moniteur d'une fraction de données à monitorer** : Fourniture au moniteur de l'accès à un sous-ensemble des données pour le suivi de la qualité et de l'intégrité des données de l'étude.

Data Analysis

Analyse Statistique et Machine Learning

- **Plan d'Analyse Statistique (SAP)** : Un plan détaillé et prédéfini qui décrit la méthodologie statistique, les procédures et les analyses à réaliser pour répondre aux objectifs de recherche d'une étude.

Analyse Descriptive

- **Analyse Descriptive** : Résume les caractéristiques de base des données en fournissant des résumés simples sur l'échantillon et les mesures. Inclut généralement des métriques statistiques telles que la moyenne, la médiane, le mode et l'écart type.

Analyse Univariate

- **Paramétrique / Non – Paramétrique** : Repose sur des hypothèses concernant la distribution des données (généralement une distribution normale).

Intégration Multiomique

- **L'intégration multiomique** : est une approche qui combine les données issues de différentes études omiques comme la génomique, la transcriptomique, la protéomique, la métabolomique et la lipidomique. Elle vise à fournir une compréhension complète des systèmes biologiques en analysant les interactions complexes entre divers types de données biomoléculaires, permettant ainsi une exploration plus approfondie des mécanismes biologiques et des phénotypes.

Outils de Visualisation

- **Les outils de visualisation personnalisés** sont conçus pour améliorer la présentation et l'interprétation des résultats de recherche. Ces visualisations peuvent inclure des graphiques, des cartes de chaleur, des diagrammes réseau, et bien d'autres formats adaptés aux besoins spécifiques de l'étude

Analyse Multivariable

- **Analyse Multivariable** : Analyse simultanément plusieurs variables pour déterminer leur effet et leurs interactions sur la variable dépendante.

Modèles de Régression

- **Modèles de Régression Logistique / Multinomiale / Cox** :
 - ✓ **Régression Logistique** : Utilisée pour prédire le résultat d'une variable dépendante catégorielle basée sur une ou plusieurs variables prédictives.

- ✓ **Régression Multinomiale** : Extension de la régression logistique qui gère les variables dépendantes décrivant plusieurs classes.
- ✓ **Régression de Cox** : Un modèle de survie utilisé pour explorer la relation entre le temps de survie des sujets et une ou plusieurs variables prédictives.

Analyse Longitudinale et de Séries Temporelles

- **Analyse de Séries Temporelles Longitudinales** : Analyse des données collectées au fil du temps pour évaluer la variation temporelle des données et la causalité potentielle.
- **Modèles GLM / Mixtes** :
 - ✓ **Modèles Linéaires Généralisés (GLM)** : Étendent les modèles de régression linéaire à des données non normales.
 - ✓ **Modèles Mixtes** : Intègrent à la fois des effets fixes et aléatoires, utiles dans les données où les observations ne sont pas indépendantes.

Modélisation Prédictive

- **Prédictions de Risque** : Prédit la probabilité qu'un événement se produise en utilisant des algorithmes statistiques, souvent utilisés dans les secteurs de la finance, de l'assurance et de la santé.
- **Analyse Discriminante** : Détermine quelles variables discriminent entre deux ou plusieurs groupes naturellement existants.

Calibration et Validation de Modèles

- **Calibration de Modèle** : Ajuste les prédictions du modèle pour mieux correspondre aux données du monde réel.
- **Validation de Modèle** : Processus de confirmation que les sorties d'un modèle statistique sont acceptables par rapport aux données réelles.

Techniques de Classification

- **Classification – Clustering / Analyse ROC** :
 - ✓ **Clustering** : Un type d'apprentissage non supervisé utilisé pour regrouper des ensembles d'objets de manière à ce que les objets dans le même groupe soient plus similaires entre eux qu'avec ceux d'autres groupes.
 - ✓ **Analyse ROC** : Utilisée pour évaluer la performance d'un modèle de classification, en traçant le taux de vrais positifs contre le taux de faux positifs.

Techniques Statistiques Avancées

- **Appariement selon le Score de Propension** : Une technique de jumelage statistique qui tente d'estimer l'effet d'un traitement en tenant compte des covariables qui prédisent la réception du traitement.
- **Analyse des Risques Concurrents** : Utilisée dans l'analyse de survie pour déterminer la probabilité de différents événements concurrents.

Techniques de Survie et de Préviation

- **Analyse de Survie – KM (Kaplan-Meier)** : Statistique non paramétrique utilisée pour estimer la probabilité de survie à partir des temps de survie observés.

- **Prévision de Séries Temporelles** : Utilise des données historiques pour prédire les résultats futurs en identifiant les motifs dans les données horodatées.

Imputation de Données et Analyse de Sensibilité

- **Imputation Multiple suivie d'une Analyse de Sensibilité** : Adresse les données manquantes en imputant les valeurs manquantes plusieurs fois pour

C. Facturation

La tarification des services pour chacune des plateformes est dépendante des caractéristiques de l'étude. L'utilisateur devra se rapprocher du responsable de la plateforme pour le chiffrage de son projet. Celui-ci s'appuiera sur une grille tarifaire basée sur une tarification différentielle : tarif interne, tarif institutionnel/académique et tarif industriel. De plus, le chiffrage étant basé sur une estimation de temps nécessaire, la plateforme se réserve le droit de facturer tout dépassement d'horaires indépendant de sa volonté.

D. Responsabilités

Toute utilisation ou interprétation de manière partielle, inappropriée ou déformée des conclusions des rapports émis par la plateforme ne saurait engager la responsabilité de la plateforme et de l'IHU-ICAN.

E. Formation

Les ressources humaines et matérielles sur les plateformes peuvent être mises à profit pour la formation d'étudiants (stagiaires, doctorants, etc.), d'ingénieurs et techniciens et pour des programmes d'animation et de formation scientifique et technique (séminaires, ateliers, sessions d'études, etc.) selon l'accord et la disponibilité du responsable de la plateforme.

IV. Sécurité

La plateforme ICAN I/O met en œuvre des mesures de sécurité rigoureuses pour protéger les données sensibles qu'elle traite. L'accès est strictement réservé aux membres autorisés du personnel. Les données sont stockées sur un serveur sécurisé appartenant à Sorbonne Université, bénéficiant ainsi des protocoles de sécurité avancés mis en place par l'institution. Ces mesures incluent un chiffrement des données, ainsi que des contrôles d'accès basés sur des identifiants uniques. En outre, des audits de sécurité sont effectués pour garantir la conformité aux normes de sécurité les plus strictes et pour détecter et corriger rapidement toute vulnérabilité potentielle. Notre engagement envers la sécurité des données assure la confidentialité et l'intégrité des données traitées.

V. Partenaires



VI. Publication (selected)

- ✓ Denimal D, Ponnaiah M, Jeannin AC, et al. Non-alcoholic fatty liver disease biomarkers estimate cardiovascular risk based on coronary artery calcium score in type 2 diabetes: a cross-sectional study with

two independent cohorts. **Cardiovasc Diabetol.** 2024;23(1):69. Published 2024 Feb 13. doi:10.1186/s12933-024-02161-x

- ✓ Huang F, Ammirati E, Ponnaiah M, et al. Fulminant myocarditis proven by early biopsy and outcomes [published correction appears in Eur Heart J. 2024 Feb 1;45(5):375]. **Eur Heart J.** 2023;44(48):5110-5124. doi:10.1093/eurheartj/ehad707
- ✓ Ha Van G, Schuldiner S, Sultan A, et al. Use of the SINBAD score as a predicting tool for major adverse foot events in patients with diabetic foot ulcer: A French multicentre study. **Diabetes Metab Res Rev.** 2023;39(8):e3705. doi:10.1002/dmrr.3705
- ✓ Urbain F, Ponnaiah M, Ichou F, et al. Impaired metabolism predicts coronary artery calcification in women with systemic lupus erythematosus. **EBioMedicine.** 2023;96:104802. doi:10.1016/j.ebiom.2023.104802
- ✓ Danial P, Olivier ME, Bréchet N, et al. Association Between Shock Etiology and 5-Year Outcomes After Venoarterial Extracorporeal Membrane Oxygenation. **J Am Coll Cardiol.** 2023;81(9):897-909. doi:10.1016/j.jacc.2022.12.018
- ✓ Lebreton G, Schmidt M, Ponnaiah M, et al. Extracorporeal membrane oxygenation network organisation and clinical outcomes during the COVID-19 pandemic in Greater Paris, France: a multicentre cohort study **Lancet Respir Med.** 2021;9(8):851-862. doi:10.1016/S2213-2600(21)00096-5
- ✓ Mack-Inocentio D, Menai M, Doré E, et al. Large-Scale Assessment of Health-Related Physical Fitness in French Older Adults: Feasibility and Validity. **Front Public Health.** 2020;8:487308. Published 2020 Dec 17. doi:10.3389/fpubh.2020.487308

VII. Normes éthiques et légales relatives à la collecte et le traitement de données à caractère personnel

Chaque Partie se conformera à ses obligations en vertu de toute législation pertinente en matière de protection des données personnelles, notamment la loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, modifiée (« LIL »), et du Règlement UE 2016/679 du parlement européen et du conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (« RGPD ») et s'engage à traiter, en conformité avec l'accord des parties, les données à caractère personnel.

Dans le cadre des prestations de service réalisées par la plateforme ICAN I/O Data Sciences la plateforme se conformera à ses obligations en vertu de toute législation pertinente en matière de protection des données personnelles, notamment la LIL et le RGPD.

Les partenaires demeurent les seuls responsables pour déterminer les données à caractère personnel qui seront traitées par la plateforme ICAN I/O Data Sciences, en proportionnalité avec les finalités du traitement et l'usage qui en sera fait.

Le partenaire est le responsable de traitement des données personnelles relatif à l'étude objet du recours à la plateforme.

L'IHU-ICAN à travers la plateforme ICAN I/O Data Sciences est le sous-traitant du traitement de données personnelles demandé par le partenaire dans le cadre de ladite étude.

Tout transfert de données entre le partenaire et la plateforme ICAN I/O Data Sciences doit se faire avec des données pseudonymisées ou anonymisées, au sens du RGPD.

Pour la réalisation de l'étude, la plateforme et le partenaire ont défini avant tout transfert, la liste des catégories des données échangées avec la plateforme.

Tout envoi de fichiers ou documents ne répondant pas à cette liste et qui viendrait violer les dispositions du RGPD, sera signalé à l'utilisateur et le fichier effacé du système d'information de l'IHU-ICAN. Cette violation sera inscrite au registre de traitement sous-traitant de l'IHU-ICAN. Il sera alors possible pour le partenaire de renvoyer le fichier conforme à la liste prédéfinie ne contenant que les données nécessaires à leur utilisation par la plateforme.

Pour les besoins de la prestation, et sur la base de l'article 6.1e (intérêt public) et 6.1f (intérêt légitime) du RGPD, l'IHU-ICAN réalise, en tant que sous-traitant, le traitement des données personnelles nécessaire à la réalisation de la prestation, tel que précisé au contrat et/ou au devis.

Pour les besoins de la prestation, et sur la base de l'article 6.1f) du RGPD, l'IHU-ICAN réalise le traitement des données personnelles du partenaire suivantes : nom, prénom, numéro de téléphone, courriel, profession.

Conformément aux articles 48 et suivants de la Loi Informatique et Libertés, toute demande d'opposition, d'accès et de rectification doit être adressée à mesdonnees@ihuican.org.

La politique d'utilisation des données personnelles de l'IHU-ICAN est consultable sur le site : www.ihuican.org.

VIII. Conditions générales d'exécution

A. Modalités de citation

Tous les articles qui utilisent les ressources de la plateforme ICAN I/O Data Sciences - quelles qu'elles soient (images, ressources logicielles, espace de travail, etc.) doivent avoir un auteur directement affilié à l'IHU-ICAN selon la condition suivante :

- Lorsque ICAN I/O Data Sciences fournit une contribution de collaboration significative sur le plan scientifique (implication dans la conception et l'élaboration de l'étude clinique et/ou fondamentale, etc.), technique (développement des modèles ou technique, d'analyse d'images et analyse et interprétation des résultats, mise en place des procédures d'analyse des données, encadrement des étudiants, etc.) ou financier (contribution à l'achat de matériel, participation au financement de l'étude), les membres de l'équipe ICAN I/O Data Sciences - qui ont contribué sont listés comme coauteurs et la position de co-auteurs sera en fonction de leur contribution.
- Lorsque ICAN I/O Data Sciences fournit une contribution de plateforme sous forme de prestation simple (analyse des données), un co-auteur de ICAN I/O Data Sciences n'est pas requis. Pour assurer néanmoins la visibilité de l'ICAN, nous demandons aux auteurs de rajouter IHU ICAN dans la partie remerciements (acknowledgments).

L'affiliation est la suivante :

Fondation Cardiométabolisme et Nutrition (IHU-ICAN), Paris, France.

Les financements suivants doivent être également listés dans les remerciements :

This study was supported by Agence Nationale de la Recherche, 'Investissements d'avenir' [ANR-10-IAHU-05].

Merci de nous consulter avant soumission en envoyant les articles concernés au Responsable de Plateforme.

B. Mesures du non-respect de la charte

L'acceptation de tout devis ou la signature de tout contrat relatif aux prestations réalisées par la plateforme constitue une acceptation de la présente charte, sans réserve. En cas de non-respect majeur et répété de la présente charte, les responsables de la plateforme se réservent le droit de refuser tout nouveau projet.