

## POLITIQUE INFRANALYTICS DE GESTION DES DONNEES

Date de la version	Rédacteurs	Vérificateurs	Approbateur	Modifications	Numéro de version
05/05/2021	S. Lecocq G. Ohanessian	Comité de direction d'Infranalytcs Romain Magré GT Politique de données			1

Ce document est une adaptation de la politique de données du synchrotron Soleil aux conditions particulières d'Infranalytcs. Soleil ne peut pas être tenu responsable du contenu final de ce document.

Nous remercions chaleureusement Brigitte Gagey, directrice Stratégie des Systèmes d'Information et Qualité du synchrotron Soleil, pour son aide précieuse dans l'élaboration de cette politique de données.

Diffusion : Pour action / Pour information

Destinataires : Utilisateurs Infranalytcs, Personnel Infranalytcs

# POLITIQUE INFRANALYTICS DE GESTION DES DONNEES

## TABLE DES MATIERES

1. PRINCIPES GENERAUX .....	3
2. DEFINITIONS .....	4
3. PROPRIETE .....	6
4. ROLES ET RESPONSABILITES .....	7
5. STOCKAGE ET CURATION DES DONNEES .....	8
6. ACCES AUX DONNEES ET REUTILISATION DE CELLES-CI .....	8
7. INFORMATIONS RELATIVES AUX PUBLICATIONS .....	9
8. FIN DE LA PRESERVATION DES DONNEES OU DU CATALOGUE EN LIGNE.....	10

# Politique Infranalytcs de gestion des données

05 mai 2021

Infranalytcs (<https://infranalytcs.cnrs.fr/>) est la fusion de trois infrastructures de recherche : RMN-THC (<https://www.ir-rmn.fr/>), Renard (<https://renard.univ-lille.fr/>) et FT-ICR (<http://www.fticr.org/>), mettant à disposition d'une large communauté d'utilisateurs des moyens expérimentaux au meilleur niveau international pour l'analyse de composés chimiques, biologiques, environnementaux, minéraux, etc., sur la base de propositions de recherche évaluées et sélectionnées par des comités d'experts. Elles fournissent également un support de compétences pour l'utilisation des dispositifs expérimentaux et des logiciels associés, pour l'analyse des données produites, ainsi que pour le stockage et la sauvegarde des données acquises.

La Loi pour une République numérique<sup>1</sup> de 2016 impose à l'ensemble des acteurs de la recherche financée par des fonds publics français d'ouvrir les données produites en particulier en répondant aux principes FAIR : des données Faciles à trouver, Accessibles, Interopérables et Réutilisables ("Findable, Accessible, Interoperable, Reusable" en anglais). La publication du [Plan National pour la Science Ouverte](#) par la ministre de l'Enseignement supérieur, de la Recherche et de l'Innovation en 2018<sup>2</sup> fournit aux acteurs du domaine un plan d'action incluant en particulier des mesures sur la gestion des données de la recherche. En novembre 2019, le CNRS s'est doté d'une [feuille de route pour la science ouverte](#)<sup>3</sup> qui, pour les données, comporte une action visant à « soutenir et accompagner les infrastructures de recherche, productrices de données, dans la définition et la mise en œuvre de politiques de données ».

Infranalytcs vise à atteindre les objectifs FAIR en mettant en place des services de gestion de données pour récupérer et stocker les données expérimentales et les métadonnées associées qui ont été collectées et/ou stockées sur l'un des sites expérimentaux d'Infranalytcs. Ceci implique également de définir un cadre d'accès à ces données et métadonnées.

C'est l'objet de la présente politique de gestion des données, qui définit un cadre contractuel entre Infranalytcs et ses utilisateurs. Elle sera progressivement mise en œuvre à partir de 2021 avec une période d'implémentation de trois ans et pourra faire l'objet d'adaptations ultérieures en fonction des évolutions des pratiques scientifiques et de gestion des données, ainsi que des éventuelles évolutions de la loi.

## 1. PRINCIPES GENERAUX

- 1.1. La présente politique de gestion des données porte sur la définition, la propriété et la curation des données expérimentales et des métadonnées associées, qui sont acquises, réduites, traitées et/ou stockées sur les équipements d'Infranalytcs ainsi que sur l'accès à ces données et métadonnées.

---

<sup>1</sup> [Loi pour une république numérique](#), 2016.

<sup>2</sup> [Plan National pour la Science Ouverte](#), 4 juillet 2018.

<sup>3</sup> [Feuille de route du CNRS pour la science ouverte](#), 18 novembre 2019.

- 1.2. Infranalytics pourra recourir à des tiers sous-traitants pour remplir ses obligations dans le cadre de cette politique de gestion des données. Ces tiers joueront le rôle de « sous-traitants de données » si des données personnelles sont concernées.
- 1.3. L'acceptation de cette politique est une condition nécessaire à l'allocation de temps d'expérience.
- 1.4. Les Utilisateurs des données ne doivent pas modifier les données et métadonnées sur le support de stockage mis en place par Infranalytics.
- 1.5. Les infractions délibérées à cette politique peuvent conduire à l'interdiction d'accéder à des données ou métadonnées, et/ou au rejet des demandes ultérieures de temps d'expérience sur les équipements d'Infranalytics.
- 1.6. Si (et dans la mesure où) les données et métadonnées contiennent des données personnelles, les législations de la France et de l'Union européenne relatives à la protection des données s'appliqueront respectivement. Cette politique de gestion des données sera régie et interprétée dans le respect de la loi française.
- 1.7. Le Responsable du projet, tel que défini à l'article 2.8, doit être enregistré personnellement via le portail dédié.
- 1.8. Chaque membre de l'équipe expérimentale réalisant une expérience à Infranalytics doit accepter et respecter les règles du règlement intérieur d'Infranalytics.

## 2. DEFINITIONS

Aux fins de la présente politique de gestion des données :

- 2.1. Le terme **données expérimentales** (voir la Figure 1) désigne l'ensemble des données collectées à partir des expériences réalisées sur les équipements d'Infranalytics. Cette définition comprend (mais sans s'y restreindre) les données créées automatiquement ou manuellement par des logiciels spécifiques de l'installation et/ou grâce à l'expertise de son personnel.
- 2.2. Le terme **données brutes** (voir la Figure 1) désigne les données expérimentales produites par les systèmes de détection et enregistrées pendant les expériences, et qui ne peuvent être dérivées d'autres données persistantes.
- 2.3. Le terme **données réduites** (voir la Figure 1) désigne les données expérimentales dérivées des données brutes par un prétraitement pendant les expériences, y compris (mais sans s'y restreindre) les prétraitements de formatage et de qualification des données brutes, ainsi que d'aide à la décision de poursuivre l'expérience.
- 2.4. Le terme **données traitées** (voir la Figure 1) désigne les données expérimentales dérivées des données brutes ou réduites pendant les étapes d'analyse.
- 2.5. Le terme **résultats** (voir la Figure 1) désigne un sous-ensemble des données traitées et des autres produits de l'analyse des données expérimentales, à l'exception des publications basées sur de telles analyses et des droits de Propriété Intellectuelle.
- 2.6. Le terme **métadonnées** (voir la Figure 1) décrit les informations portant sur les données collectées par des équipements d'Infranalytics, ainsi que (mais sans s'y restreindre) les informations sur l'échantillon, le contexte de l'expérience, l'équipe expérimentale, les conditions expérimentales et les informations administratives et logistiques.

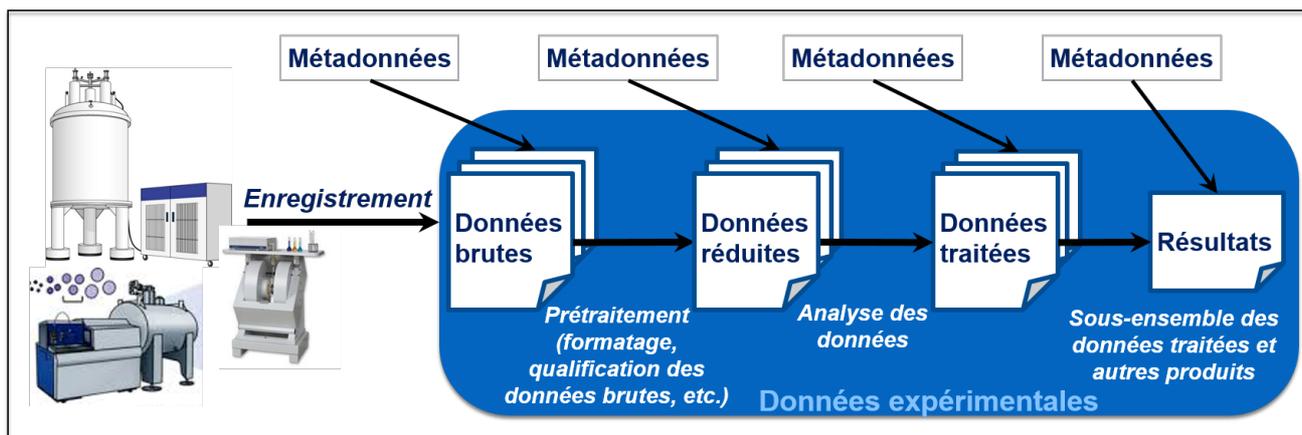


Figure 1 : Illustration simplifiée des classes de données

- 2.7. Le terme **temps d'expérience** signifie la période temporelle allouée par Infranalytics à l'équipe expérimentale pour la réalisation d'expériences scientifiques sur un des équipements d'Infranalytics. Chaque temps d'expérience est associé à une équipe expérimentale, un site et une installation bien définis.
- 2.8. Le terme **responsable du projet** ou "**Main Proposer**" (MP) désigne le responsable identifié du projet soumis. Le MP est la personne qui soumet le projet et en gère tous les aspects. Il ou elle est aussi la personne à contacter par Infranalytics pour toute question administrative relative au projet. Le MP est garant pour l'équipe expérimentale du respect des règles figurant dans cette politique de données. Le MP doit être enregistré personnellement via le portail dédié.
- 2.9. Le terme **équipe expérimentale** inclut le MP et l'ensemble des co-proposants d'un projet, ainsi que l'ensemble des participants à l'expérience ou aux expériences exécutées pour le projet (y compris l'équipe d'accueil du site Infranalytics), et enfin toutes les autres personnes auxquelles le MP accorde le droit d'accéder aux données expérimentales résultantes et aux métadonnées associées.
- 2.10. Le terme **utilisateur des données** désigne toute personne désirant accéder aux données produites sur les équipements d'Infranalytics.
- 2.11. Le terme **projet** désigne la description scientifique et les modalités de l'expérience menée par l'équipe expérimentale.
- 2.12. Le terme **recherche propriétaire** désigne la recherche réalisée via un accès payé à finalité commerciale, non soumis à évaluation par les pairs.
- 2.13. Le terme **recherche publique** désigne la recherche, financée en majorité par des fonds publics, réalisée via un accès examiné par des pairs (notamment les comités d'experts d'Infranalytics pour l'évaluation des propositions), ou via un temps d'expérience pour la recherche interne Infranalytics, mais pas via une recherche propriétaire ou des accords spécifiques.
- 2.14. Le terme **catalogue en ligne** désigne une base de données informatique de métadonnées, qui contient des indications permettant d'accéder aux fichiers de données brutes ou réduites.
- 2.15. Le terme **dépositaire** désigne l'entité qui assure le stockage et la conservation des données brutes, des métadonnées et des résultats, et donne accès à ceux-ci.
- 2.16. Le terme **stockage à long terme** signifie stockage jusqu'à cinq ans (Infranalytics visant 10 ans) à compter de la fin du temps d'expérience correspondant. La durée précise dépendra de l'équipement utilisé, du type et du volume des données concernées, et des conséquences économiques associées au stockage à long terme de ces données. Ainsi, Infranalytics se réserve le droit de restreindre les périodes de stockage ou les jeux de données, en consultant

les communautés concernées, pour les instruments à haut débit de données. Les durées à long terme applicables à chaque type d'instrument seront communiquées sur le portail de dépôt de projet, et rappelées au MP lors de l'acceptation du projet.

- 2.17. Le terme **période d'embargo** signifie la période, à compter de la clôture du projet, pendant laquelle l'accès aux données brutes et réduites collectées et aux métadonnées associées est limité à l'équipe expérimentale.
- 2.18. Le terme **accès ouvert** signifie appartenant à la communauté dans son ensemble, non protégé par un copyright ou un brevet, et susceptible de réutilisation par quiconque. Les données stockées à long terme seront mises à disposition sous la licence CC-BY<sup>4</sup> à l'issue de la période d'embargo.
- 2.19. Le terme **Plan de Gestion de Données**, ou DMP ("Data Management Plan"), fait référence au document décrivant le cycle de vie des données qui seront collectées, traitées et/ou générées par un projet de recherche : « la gestion des données de recherche pendant le projet et après la fin de ce dernier ; quelles données seront collectées, traitées et/ou générées ; quelle méthodologie et quelles normes seront appliquées ; ces données seront-elles partagées / en accès ouvert ; et comment ces données seront-elles conservées et préservées (y compris après la fin du projet) »<sup>5</sup>. Un tel document peut être exigé par les bailleurs de fonds externes, les instituts de recherche et d'autres organisations.

### 3. PROPRIETE

- 3.1. Toutes les données brutes et/ou réduites et les métadonnées associées obtenues via des recherches publiques menées à Infranalytics seront en accès ouvert après une période d'embargo.
- 3.2. Toutes les données brutes et/ou réduites et les métadonnées associées obtenues via des recherches propriétaires menées à Infranalytics seront la propriété exclusive du client qui a acheté l'accès ; ces données ne sont pas couvertes par la présente politique de gestion des données. Les données provenant de recherches propriétaires seront supprimées des systèmes de stockage d'Infranalytics après les expériences, sauf accord contraire avec la direction d'Infranalytics avant le début de l'expérience.
- 3.3. Dans certains cas spécifiques (tels que les expériences réalisées par un laboratoire ZRR<sup>6</sup>, etc.), Infranalytics et l'organisme portant l'expérience peuvent convenir d'une politique de gestion des données particulière. Le MP est alors invité à l'indiquer dans sa proposition et à contacter la direction d'Infranalytics dès l'acceptation de cette proposition. Les cas d'exemption à l'ouverture de données engendrées grâce à des fonds publics sont listés de manière exhaustive dans la [Loi pour une République numérique](#)<sup>7</sup>.
- 3.4. La propriété (y compris tous les éventuels droits de Propriété Intellectuelle qui en découlent) de l'ensemble des résultats de l'analyse des données expérimentales est déterminée par les obligations contractuelles de la ou des personnes réalisant l'analyse des données et/ou l'interprétation des données, conformément au droit applicable.

---

<sup>4</sup> Cf. (<https://creativecommons.org/licenses/by/4.0/>).

<sup>5</sup> Cf. [Guidelines on FAIR Data Management in Horizon 2020](#), Commission européenne, juillet 2016, p. 4

<sup>6</sup> ZRR = « Zone à Régime Restrictif », cf. <http://www.sqdsn.gouv.fr/missions/protection-du-potentiel-scientifique-et-technique-de-la-nation/>

<sup>7</sup> [Loi pour une République numérique](#), 2016. On pourra aussi se reporter au guide d'application établi par le Comité pour la Science Ouverte : <https://www.ouvrirlascience.fr/ouverture-des-donnees-de-recherche-guide-danalyse-du-cadre-juridique-en-france-v2/>

## 4. ROLES ET RESPONSABILITES

- 4.1. Sauf indication contraire, Infranalytics agira comme dépositaire des données brutes et/ou réduites et des métadonnées associées.
- 4.2. Infranalytics fournit dans la limite de ses possibilités des moyens de réduction et/ou de traitement des données brutes.
- 4.3. Le MP est invité à établir un Plan de Gestion des Données associé au projet déposé. Si le bailleur de fonds, l'institut de recherche, ou toute autre organisation concernée impose des exigences particulières en matière de gestion des données, le DMP correspondant sera accepté et mis en œuvre par Infranalytics dans la limite de la présente politique de données.
- 4.4. Il relève de la responsabilité du MP de s'assurer que le numéro d'expérience est saisi correctement dans les métadonnées correspondant à chaque jeu de données afin d'associer correctement ces jeux de données au MP. À défaut, le MP ne pourra pas accéder aux données via le catalogue en ligne (6.1), ou bien d'autres utilisateurs risqueront d'avoir par erreur des droits d'accès à ces données. Ceci peut normalement être réalisé en saisissant le numéro d'expérience dans le logiciel d'acquisition de données.
- 4.5. Le MP doit s'assurer que les métadonnées contiennent les descriptions des échantillons.
- 4.6. L'équipe expérimentale est encouragée à s'assurer que les métadonnées sont aussi complètes que possible, car ceci facilitera pour tous la recherche, la récupération et l'interprétation à long terme des données.
- 4.7. Infranalytics s'efforce de fournir des moyens de capture des éléments de métadonnées qui ne sont pas capturés automatiquement par les instruments, y compris le Plan de Gestion de Données, afin de faciliter l'enregistrement de la description la plus complète possible des données brutes.
- 4.8. Il relève de la responsabilité du MP de s'assurer que toutes les informations (y compris les données personnelles) ajoutées aux métadonnées par l'équipe expérimentale sont conformes à la réglementation française et européenne relative à la protection des données. Infranalytics n'est pas responsable de ladite conformité avec la législation relative à la protection des données. Infranalytics peut apporter une aide en fournissant les informations à sa disposition.
- 4.9. Il relève de la responsabilité du MP de s'assurer que l'équipe expérimentale consent à toutes les modifications éventuellement apportées à la période d'embargo (6.5).
- 4.10. Il relève de la responsabilité du MP de gérer les éventuelles autorisations d'accès aux données octroyées aux personnes extérieures à l'équipe expérimentale désignée pendant cette période d'embargo (6.4) et en conformité avec la législation.
- 4.11. Le MP a la possibilité de transférer ses droits (ou d'en octroyer une partie) pendant la période d'embargo à un autre membre de l'équipe expérimentale.
- 4.12. Infranalytics fera de son mieux pour assurer le stockage et la curation adéquats des données expérimentales et des métadonnées, ainsi que l'accès ininterrompu à celles-ci, pendant la durée stipulée à l'acceptation de la proposition. Cependant, des défaillances causées par des erreurs techniques ou humaines ne peuvent pas être exclues en ce qui concerne le traitement de données. Infranalytics ne peut donc garantir un stockage et une curation entièrement adéquats. D'autre part, l'accès peut être temporairement limité ou impossible, en particulier du fait des nécessaires interventions de maintenance ou de rénovation des services, ou en cas de défaillance de prestataires de services tiers.
- 4.13. Infranalytics et le MP ne peuvent être tenus responsables en cas de perte ou de défaillance de données, de métadonnées ou de résultats, ni en cas de limitation ou d'indisponibilité de l'accès.

- 4.14. Si un MP n'est plus en capacité d'assumer son rôle, il/elle devra en informer la direction d'Infranalytics et l'équipe expérimentale, qui trouveront ensemble un remplaçant, si possible.
- 4.15. Infranalytics et le MP ne peuvent être tenus responsables en cas d'indisponibilité ou de perte des logiciels d'analyse de données.
- 4.16. Infranalytics ne peut être tenu responsable des conséquences de toute interprétation des données.
- 4.17. Il relève de la responsabilité du MP de récupérer et de sauvegarder toutes les données provenant de l'expérience au cours de la durée stipulée à l'acceptation de la proposition.
- 4.18. À titre de précaution (et sans préjudice de la question de propriété), tous les membres de l'équipe expérimentale octroient à Infranalytics le droit d'utilisation illimité et absolu des données expérimentales et des métadonnées, dans la mesure nécessaire à leur préservation et à leur mise à disposition en conformité avec cette politique de gestion des données.

## **5. STOCKAGE ET CURATION DES DONNEES**

- 5.1. Toutes les données brutes et/ou réduites seront préservées dans des formats bien définis, pour lesquels les moyens de lire les données seront mis à disposition par Infranalytics dans la mesure du possible. Ces formats pourront être différents pour les différents types d'équipements.
- 5.2. Les métadonnées, qu'elles soient capturées automatiquement par des instruments ou enregistrées manuellement avec des outils mis en place par Infranalytics, seront préservées soit associées aux fichiers de données brutes, soit dans un catalogue en ligne associé, ou encore par ces deux méthodes.
- 5.3. Les données brutes et/ou réduites et les métadonnées associées resteront en lecture seule pendant la totalité de leur durée de vie.
- 5.4. Les données brutes et/ou réduites et les métadonnées associées seront migrées ou copiées sur des infrastructures de stockage sécurisé (qui sont susceptibles d'être sous-traitées) en vue de leur préservation à long terme (objectif jusqu'à 10 ans). À l'issue de la réalisation de leur expérience, les MP seront invités à réfléchir à la valeur de long terme des données de leurs projets, en utilisant une « Data value checklist » proposée par Infranalytics.
- 5.5. Infranalytics assurera l'attribution à chaque projet d'un identifiant persistant unique de type DOI (Digital Object Identifier).
- 5.6. Sauf mention contraire, un seul jeu de données brutes et/ou réduites et les métadonnées associées sera préservé à long terme. Les autres jeux de données réduites et métadonnées associées ne seront pas préservés à long terme par Infranalytics. Les outils permettant de les régénérer à partir du jeu conservé seront mis à disposition par Infranalytics dans la mesure du possible (via des versions récentes des outils supportés fonctionnant sur des systèmes d'exploitation récents).
- 5.7. Sauf mention contraire, les données traitées provenant des étapes intermédiaires d'analyse ainsi que les métadonnées associées ne seront pas préservés à long terme par Infranalytics.

## **6. ACCES AUX DONNEES ET REUTILISATION DE CELLES-CI**

- 6.1. L'accès aux données brutes et/ou réduites et aux métadonnées stockées par Infranalytics sera à terme donné via un catalogue en ligne consultable dans les conditions décrites ci-dessous.

- 6.2. Le téléchargement des données pourra être soumis à des restrictions en matière de volume qu'Infranalytics pourra fournir. Le téléchargement des données sera enregistré dans un journal et ces informations seront mises à disposition du MP, dans les limites prévues par la législation.
- 6.3. La période d'embargo sera de trente-six (36) mois par défaut, prolongeable une fois de douze (12) mois sur justification. Elle s'appliquera aux données brutes et/ou réduites et métadonnées associées issues de la recherche publique. Le MP pourra lever l'embargo à tout moment.
- 6.4. Pendant la période d'embargo, l'accès aux données brutes et/ou réduites et aux métadonnées associées obtenues par une expérience est limité à l'équipe expérimentale, hormis les cas spécifiques de transfert ou d'octroi de droits par le MP.  
Si des personnes extérieures à l'équipe expérimentale désignée sollicitent l'accès aux données, Infranalytics transférera cette demande au MP, qui l'examinera en coordination avec l'équipe expérimentale.
- 6.5. Au terme de la période d'embargo, toutes les données brutes et/ou réduites ainsi que les métadonnées associées seront mises en accès ouvert.  
Infranalytics rappellera la date limite ainsi que les conditions d'une éventuelle prolongation au MP au moins deux mois avant la fin de la période d'embargo.  
Tout MP souhaitant prolonger les restrictions d'accès aux données au-delà de trente-six (36) mois (par exemple pour s'aligner avec les pratiques des disciplines scientifiques en lien avec la proposition) aura cette possibilité, soumise à justification, dans la limite de douze (12) mois de prolongation.
- 6.6. Les Utilisateurs des données ne doivent pas tenter d'accéder à des données sous période d'embargo ou à des données et métadonnées fermées pour des raisons légales, stockées par Infranalytics, à moins qu'ils n'y soient autorisés par les clauses de cette politique.
- 6.7. L'accès aux résultats provenant d'analyses réalisées sur des données brutes et/ou réduites et aux métadonnées associées est limité à la ou aux personnes réalisant l'analyse, sauf demande contraire de ces personnes.
- 6.8. Pour des raisons liées au fonctionnement de l'installation, le personnel autorisé d'Infranalytics (par exemple les scientifiques exploitant les instruments ou le personnel informatique) a accès à toutes les données ou métadonnées conservées.
- 6.9. Infranalytics s'engage à préserver la confidentialité des données et métadonnées dont l'accès est limité.
- 6.10. Le catalogue en ligne permettra de relier les données expérimentales aux projets. L'accès aux projets sera seulement accordé à l'équipe expérimentale et au personnel pertinent d'Infranalytics pendant la période d'embargo, sauf demande contraire du MP.
- 6.11. Un sous-ensemble restreint de métadonnées sera rendu public dès la fin des expériences. Ces informations seront mises en ligne dans le catalogue via le lien associé à l'identifiant persistant unique.
- 6.12. Toute personne publiant des résultats basés sur des données en accès ouvert devra citer la source des données avec leur identifiant persistant unique, en accord avec la licence CC-BY, ainsi que toutes les éventuelles publications associées aux mêmes données brutes et/ou réduites ou encore la provenance des données brutes et/ou réduites. Infranalytics encourage en outre la mise en accès ouvert de tels résultats.

## **7. INFORMATIONS RELATIVES AUX PUBLICATIONS**

- 7.1. Les publications relatives aux données des expériences réalisées à Infranalytics devront impérativement citer l'identifiant persistant unique du projet.

- 7.2. Les références des publications relatives aux expériences réalisées à Infranalytics devront être communiquées par le MP au site d'Infranalytics qui a conduit l'expérience pour publication dans le catalogue, dès la publication de l'article. Tout manquement à fournir les références d'une publication pourra conduire au refus de nouvelles allocations de temps d'expérience.

## **8. FIN DE LA PRESERVATION DES DONNEES OU DU CATALOGUE EN LIGNE**

- 8.1. Si Infranalytics décide de cesser d'agir comme dépositaire et/ou d'entretenir et de mettre à disposition le catalogue en ligne, Infranalytics informera en temps opportun les MP concernés et leur fournira des moyens effectifs pour faire une copie des données brutes et/ou réduites, des métadonnées et des résultats correspondants, à condition qu'Infranalytics connaisse les adresses e-mail correctes des MP à cette date.