

POLITIQUE SOLEIL DE GESTION DES DONNEES

Date de Diffusion	Rédacteur	Vérificateurs	Approbateur	Modifications	Indice
02/10/2018	B.Gagey	Direction SOLEIL Juriste SOLEIL Représentants des Utilisateurs	J.Daillant		0

PUBLIC

Diffusion : Pour action / Pour information

Destinataires : Utilisateurs SOLEIL, Personnel SOLEIL

POLITIQUE SOLEIL DE GESTION DES DONNEES

TABLE DES MATIERES

1. PRINCIPES GENERAUX	3
2. DEFINITIONS.....	4
3. PROPRIETE	6
4. ROLES ET RESPONSABILITES	6
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 SOLEIL de Gestion des Données

2 octobre 2018

Conscient que « la recherche moderne se construit sur un dialogue scientifique approfondi et progresse en s'appuyant sur les travaux antérieurs »¹, SOLEIL encourage ses Utilisateurs et ses Chercheurs à produire des données répondant aux principes FAIR : des données Faciles à trouver, Accessibles, Interopérables et Réutilisables ("Findable, Accessible, Interoperable, Reusable" en anglais). SOLEIL vise à faciliter cela grâce à des services de gestion de données pour stocker et récupérer les données expérimentales et les métadonnées associées qui ont été collectées et/ou stockées à SOLEIL. 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. Elle est basée sur un modèle établi dans le cadre du projet PaNdata², co-financé par la Commission européenne via le septième Programme-cadre, ainsi que sur son application dans les autres installations dédiées à la production et à l'utilisation de faisceaux de photons ou de neutrons. Ce modèle a été adapté à SOLEIL avec l'aide d'experts en données de la recherche du CNRS. La présente politique de gestion des données sera mise en œuvre à partir de 2018.

1. PRINCIPES GENERAUX

- 1.1. La présente politique de gestion des données porte sur la propriété et la curation des données expérimentales, et des métadonnées associées, qui sont collectées, réduites, traitées et/ou stockées à SOLEIL, ainsi que sur l'accès à ces données et métadonnées.
- 1.2. SOLEIL 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 de faisceau.
- 1.4. Les Utilisateurs ne doivent pas tenter d'accéder, d'exploiter ou de distribuer des données ou des métadonnées, à moins qu'ils n'y soient autorisés par les clauses de cette politique.
- 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 de faisceau à SOLEIL.
- 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. Les Utilisateurs de l'installation SOLEIL doivent être enregistrés personnellement via le portail dédié ([SUN set](#)). Chaque Utilisateur réalisant une expérience à SOLEIL doit accepter et respecter les règles de la Charte Utilisateurs de SOLEIL.

¹ Cf. [Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020](#), Commission européenne, mars 2017, et [Traduction INIST-CNRS](#)

² Cf. [PaNdata Europe](#). Common Data Policy Framework on Scientific Data, décembre 2010

2. DEFINITIONS

Aux fins de la présente politique de gestion des données :

- 2.1. Le terme **données expérimentales** (voir la Figure 1) désigne les données collectées à partir des expériences réalisées sur des instruments de l'installation SOLEIL. 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 dans le but de faciliter l'analyse ultérieure des données expérimentales.
- 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 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 instruments de l'installation SOLEIL, y compris (mais sans s'y restreindre) les informations sur le contexte de l'expérience, l'équipe expérimentale, les conditions expérimentales et autres informations logistiques.

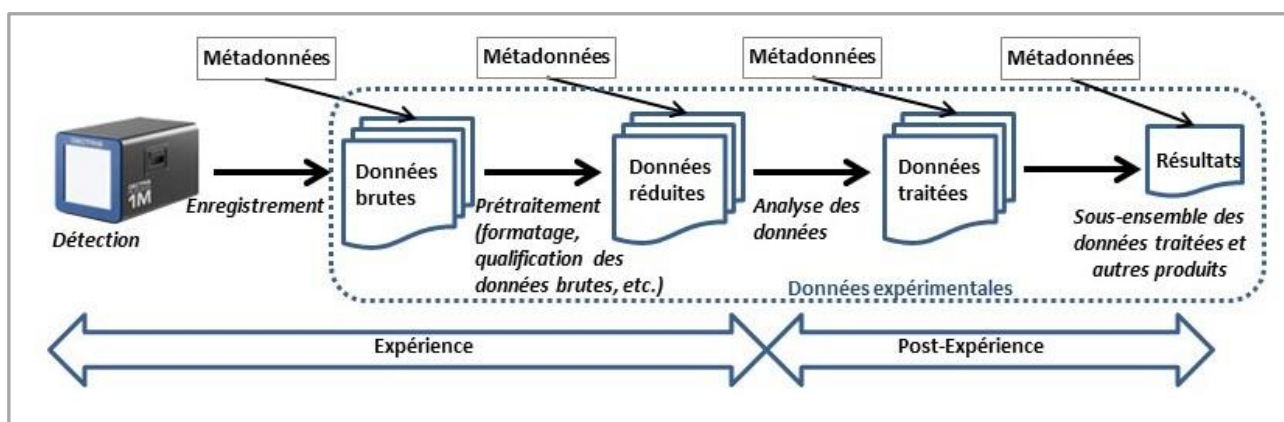


Figure 1 : Illustration simplifiée des classes de données

- 2.7. Le terme **temps de faisceau** signifie la période temporelle allouée par SOLEIL à l'équipe expérimentale pour la réalisation d'expériences scientifiques sur un instrument de l'installation SOLEIL. Chaque temps de faisceau est associé à une date de début, une date de fin et une équipe expérimentale bien définies.
- 2.8. Le terme **cycle d'allocation de temps de faisceau** désigne la période temporelle entre un appel à projets et le suivant, soit typiquement six mois.
- 2.9. Le terme **projet standard** désigne une proposition d'expérience s'étendant sur un cycle d'allocation de temps de faisceau.
- 2.10. Le terme **projet à long terme** désigne une proposition d'expérience s'étendant sur plusieurs cycles d'allocation de temps de faisceau.

- 2.11. Le terme **projet BAG** ("Block Allocation Group") désigne une proposition de plusieurs expériences soumise par des consortiums locaux, nationaux ou internationaux et s'étendant sur au moins deux cycles d'allocation de temps de faisceau. Chaque BAG est constitué d'un coordinateur ou "BAG coordinator" (désigné parmi les membres des équipes du BAG) et de plusieurs équipes, chacune dirigée par un **Principal Investigateur** ou "**Principal Investigator**" (PI).
- 2.12. Le terme **Responsable du Projet** ou "**Main Proposer**" (MP) désigne le responsable identifié sur un projet standard ou à long terme, ou le coordinateur identifié sur un projet BAG. Le MP est la personne qui soumet la proposition et en gère tous les aspects. Il ou elle est aussi la personne à contacter par SOLEIL pour toute question administrative relative à la proposition.
- 2.13. Le terme **équipe expérimentale** inclut le MP et l'ensemble des co-proposants d'une proposition, ainsi que l'ensemble des participants à l'expérience ou aux expériences exécutées pour la proposition (y compris le ou les référents locaux SOLEIL, dits "local contact"), et enfin toutes les autres personnes auxquelles le MP accorde le droit d'accéder aux données expérimentales résultantes et aux métadonnées associées.
Dans le cas d'un BAG, chaque équipe expérimentale comprend le PI et l'ensemble des co-proposants de l'expérience, ainsi que l'ensemble des participants à cette expérience (y compris le ou les référents locaux SOLEIL, dits "local contact"), et enfin toutes les autres personnes auxquelles le PI octroie le droit d'accéder aux données expérimentales résultantes et aux métadonnées associées.
- 2.14. Le terme **recherche propriétaire** désigne la recherche réalisée via un accès payé (commercial).
- 2.15. Le terme **recherche publique** désigne la recherche réalisée via un accès examiné par des pairs (notamment les comités de programme SOLEIL pour l'évaluation des propositions), ou via un temps de faisceau pour la recherche interne SOLEIL ou pour des accès rapides, mais pas via une recherche propriétaire ou des accords spécifiques.
- 2.16. Le terme **catalogue en ligne** désigne une base de données informatique de métadonnées, qui contient des liens vers des fichiers de données brutes, accessibles par diverses méthodes telles que (mais sans s'y restreindre) les navigateurs Web.
- 2.17. Un **utilisateur enregistré** désigne une personne à qui a été attribué un compte personnel [SUN set](#), ce dernier pouvant être demandé via la page Web de connexion du portail des Utilisateurs de l'installation SOLEIL.
- 2.18. Le terme **dépositaire** désigne l'institut qui stocke et conserve les données brutes, les métadonnées et les résultats, et donne accès à ceux-ci.
- 2.19. Le terme **à long terme** signifie jusqu'à cinq ans (SOLEIL visant 10 ans) à compter de la fin du temps de faisceau correspondant. La durée précise dépendra de la ligne de lumière utilisée, du type et du volume des données concernées, et des conséquences économiques associées au stockage à long terme de ces données. Ainsi, SOLEIL 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 ligne de lumière ou station d'expérience seront communiquées sur le portail des Utilisateurs de l'installation SOLEIL ([SUN set](#)), et rappelées au MP lors de l'acceptation de la proposition et avant la fin de la durée à long terme.
- 2.20. Le terme **accès ouvert** signifie appartenant à la communauté dans son ensemble, non protégé par un copyright ou un brevet, et susceptible d'appropriation par quiconque. Les données stockées à long terme seront mises à disposition sous la licence CC-BY ³.

³ Cf. (<https://creativecommons.org/licenses/by/4.0/>).

- 2.21. Le terme **période d'embargo** signifie la période, à compter de la fin du temps de faisceau correspondant, pendant laquelle l'accès aux données brutes collectées et aux métadonnées associées est limité à l'équipe expérimentale. Cette période sera de trois (3) années par défaut, avec la possibilité de l'étendre à cinq (5) ans, conformément à l'article 6.5.
- 2.22. Le terme **Plan de Gestion de Données**, ou DMP ("Data Management Plan"), fait référence au document décrivant le cycle de vie de gestion des données pour les 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). »⁴. 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 les métadonnées associées obtenues via des recherches publiques menées à SOLEIL seront en accès ouvert après une période d'embargo.
- 3.2. Toutes les données brutes et les métadonnées associées obtenues via des recherches propriétaires menées à SOLEIL 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 de SOLEIL après les expériences, sauf accord contraire avec la direction de SOLEIL avant le début de l'expérience.
- 3.3. Dans certains cas spécifiques (tels que les expériences en lien avec la sécurité nationale, les expériences classifiées réalisées par un laboratoire ZRR⁵, etc.), SOLEIL 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 de SOLEIL dès l'acceptation de cette proposition.
- 3.1. 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.

4. ROLES ET RESPONSABILITES

- 4.1. Sauf indication contraire, SOLEIL agira comme dépositaire des données brutes et des métadonnées associées.
- 4.2. SOLEIL vise à fournir des moyens de réduction et/ou de traitement des données brutes.
- 4.3. Le MP a l'obligation d'établir un Plan de Gestion des Données. SOLEIL peut apporter une aide en fournissant un modèle ou en envoyant au MP les informations disponibles si le bailleur de fonds, l'institut de recherche, ou toute autre organisation concernée, impose des exigences particulières en matière de gestion des données.
- 4.4. Il relève de la responsabilité du MP (et/ou de chaque PI d'un BAG) de s'assurer que le numéro d'expérience est saisi correctement dans les métadonnées correspondant à chaque jeu de

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

⁵ ZRR = « Zone à Régime Restrictif », cf. <http://www.sgdsn.gouv.fr/missions/protection-du-potentiel-scientifique-et-technique-de-la-nation/>

données afin d'associer correctement ces jeux de données au MP et, le cas échéant, au PI. À défaut, l'équipe expérimentale ne pourra 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 (ou chaque PI d'un BAG) 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 des données à long terme.
- 4.7. SOLEIL 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 (et/ou de chaque PI d'un BAG) 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. SOLEIL n'est pas responsable de ladite conformité avec la législation relative à la protection des données. SOLEIL peut apporter une aide en fournissant les informations en 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 (et/ou de chaque PI d'un BAG) 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 (et/ou chaque PI d'un BAG) a la possibilité de transférer ses droits (ou d'en octroyer une partie) pendant la période d'embargo à une autre personne enregistrée.
- 4.12. Le MP (ou chaque PI d'un BAG) a l'obligation de récupérer et de sauvegarder toutes les données provenant de l'expérience au cours de la durée stipulée à l'acceptation de la proposition. SOLEIL peut lui prodiguer des conseils pour trouver un dépôt de données destiné à la conservation à très long terme.
- 4.13. Si un MP (ou un PI) n'est plus en capacité d'assumer son rôle, il/elle devra en informer la direction de SOLEIL et l'équipe expérimentale, qui trouveront ensemble un remplaçant, si possible.
- 4.14. SOLEIL 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. SOLEIL 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.
- 4.15. SOLEIL et le MP (ou chaque PI d'un BAG) 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.
- 4.16. SOLEIL et le MP (ou chaque PI d'un BAG) ne peuvent être tenus responsables en cas d'indisponibilité ou de perte des logiciels d'analyse de données.
- 4.17. SOLEIL ne peut être tenu responsable des conséquences de toute interprétation des données.

- 4.18. À titre de précaution (et sans préjudice de la question de propriété), tous les membres de l'équipe expérimentale octroient à SOLEIL 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 seront préservées dans des formats bien définis, pour lesquels les moyens de lire les données seront mis à disposition par SOLEIL.
- 5.2. Les métadonnées, qu'elles soient capturées automatiquement par des instruments ou enregistrées manuellement avec des outils SOLEIL, seront préservées soit à l'intérieur des fichiers de données brutes, soit dans un catalogue en ligne associé, ou encore par ces deux méthodes.
- 5.3. Les données brutes et les métadonnées resteront en lecture seule pendant la totalité de leur durée de vie.
- 5.4. Les données brutes et les métadonnées seront migrées ou copiées sur des infrastructures de stockage à long terme (qui sont susceptibles d'être sous-traitées) en vue de leur préservation à long terme.
- 5.5. Il est prévu que chaque expérience et chaque jeu de données soient associés à un identifiant persistant unique permettant de les attribuer de manière unique.
- 5.6. Sauf mention contraire, les données réduites et les métadonnées associées ne seront pas préservées à long terme par SOLEIL. Les outils permettant de les régénérer à partir des données brutes seront mis à disposition par SOLEIL 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 préservées à long terme par SOLEIL.
- 5.8. Les résultats provenant d'analyses effectuées sur les données brutes et les métadonnées avec des moyens fournis par SOLEIL seront stockés à long terme par SOLEIL s'il s'agit d'une recherche interne SOLEIL, et seulement dans la mesure du possible pour les autres cas. SOLEIL ne pourra être tenu responsable de la préservation complète de ces données, par exemple de garantir la sémantique de ces données ou la disponibilité des logiciels nécessaires à leur lecture ou manipulation.

6. ACCES AUX DONNEES ET REUTILISATION DE CELLES-CI

- 6.1. L'accès aux données brutes, aux métadonnées et aux résultats stockés par SOLEIL sera donné via un catalogue en ligne consultable dans les conditions décrites ci-dessous.
- 6.2. L'accès au catalogue en ligne sera restreint aux utilisateurs enregistrés.
- 6.3. Le téléchargement des données pourra être soumis à des restrictions en matière de volume que SOLEIL pourra fournir.
Le téléchargement des données sera enregistré dans un journal et ces informations seront mises à disposition du MP (ou PI dans le cas d'un BAG), dans les limites prévues par la législation.
- 6.4. Pendant la période d'embargo, l'accès aux données brutes 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 (ou PI dans le cas d'un BAG).

Si des personnes extérieures à l'équipe expérimentale désignent l'accès aux données, SOLEIL transférera cette demande au MP (ou PI dans le cas d'un BAG), qui l'examinera en coordination avec l'équipe expérimentale.

- 6.5. Au terme de la période d'embargo, toutes les données brutes ainsi que les métadonnées associées seront mises en accès ouvert.
SOLEIL rappellera la date limite 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 trois (3) ans (par exemple pour s'aligner avec les pratiques des disciplines scientifiques en lien avec la proposition) aura cette possibilité de façon annuelle, soumis à justification (via l'outil [SUN set](#)), dans la limite de deux (2) années de prolongation. Pour les prolongations plus longues, une demande précisant les raisons de la prolongation devra être soumise aux directeurs scientifiques de SOLEIL, qui seront décisionnaires quant à cette demande. Dans des circonstances exceptionnelles, les données pourront être mises en accès ouvert avant le terme de la période d'embargo initiale si le MP en fait la demande auprès de SOLEIL.
- 6.6. L'accès aux résultats provenant d'analyses réalisées sur des données brutes et aux métadonnées est limité à la ou aux personnes réalisant l'analyse, sauf demande contraire de ces personnes. Toutefois, si l'accès aux données brutes analysées est encore restreint, l'accès aux résultats devra être octroyé par le MP (ou PI dans le cas d'un BAG) sur demande.
- 6.7. Pour des raisons liées au fonctionnement de l'installation, le personnel autorisé de SOLEIL (par exemple les scientifiques exploitant les instruments ou le personnel informatique) a accès à toutes les données ou métadonnées conservées. SOLEIL s'engage à préserver la confidentialité des données dont l'accès est limité.
- 6.8. Le catalogue en ligne permettra de relier les données expérimentales aux propositions d'expériences. L'accès aux propositions sera seulement accordé à l'équipe expérimentale et au personnel pertinent de SOLEIL, sauf demande contraire du MP (ou PI dans le cas d'un BAG).
- 6.9. Un sous-ensemble restreint de métadonnées sera rendu public dès la fin des expériences. Ces informations seront mises à disposition sur le Web via le lien associé à l'identifiant persistant unique.
- 6.10. SOLEIL recommande fortement que les chercheurs qui souhaitent réaliser des analyses de données brutes et de métadonnées en accès ouvert prennent contact, dans la mesure du possible, avec le MP (ou PI dans le cas d'un BAG) pour les en informer, suggérer des collaborations et leur proposer de co-signer toute publication ultérieure découlant des données, le cas échéant.
- 6.11. Toute personne publiant des résultats basés sur des données en accès ouvert devra citer la source des données avec leur identifiant persistant unique, ainsi que toutes les éventuelles publications associées aux mêmes données brutes ou encore la provenance des données brutes. SOLEIL encourage en outre la mise en accès ouvert de tels résultats.
- 6.12. L'allocation de temps de faisceau SOLEIL engage le MP (et/ou chaque PI d'un BAG) à répondre aux demandes concernant les données obtenues par l'expérience après leur mise en accès ouvert.

7. INFORMATIONS RELATIVES AUX PUBLICATIONS

- 7.1. Les publications relatives aux données des expériences réalisées à SOLEIL devront citer l'identifiant persistant unique de l'expérience et des données.
- 7.2. Les références des publications relatives aux expériences réalisées à SOLEIL devront être déposées dans la base de données des publications (fonction « Registrer Publications » du

menu Users du [SUN_set](#)) dès la publication de l'article. Tout manquement à fournir les références d'une publication pourra conduire au refus de nouvelles allocations de temps de faisceau.

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

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