



UNITED
NATIONS

EP

UNEP/MED WG.566/Inf.7



UNITED NATIONS
ENVIRONMENT PROGRAMME
MEDITERRANEAN ACTION PLAN

10 mai 2023
Original: English

Réunion des Points Focaux Nationaux INFO/RAC

Rome, Italie, 7-8 juin 2023

Point 5 de l'ordre du jour. Politique des données du PAM - Annexes des flux de données

Politique de gestion des données du PAM - Implications pour les parties contractantes

Pour des raisons environnementales et de réduction des coûts, ce document est imprimé en nombre limité. Les délégués sont priés d'apporter leurs copies aux réunions et de ne pas demander de copies supplémentaires.

UNEP/MAP
INFO/RAC - Rome, 2023

Politique de gestion des données du PAM – Implications pour les parties contractantes

Le présent document vise à donner un bref aperçu des principaux objectifs et implications de la politique approuvée en matière de données du PAM.

La Politique de gestion des données du PAM a été adoptée par la décision IG.25/10 (UNEP/MED IG.25/27) approuvée lors de la 22^e réunion des parties contractantes à la convention de Barcelone qui s'est tenue à Antalya en décembre 2021, dans le but général d'atteindre un niveau d'interopérabilité de base en établissant des principes, des objectifs et des instruments pour chaque flux de données appartenant au PAM.

La décision IG.25/10 porte sur :

1. l'adoption de la politique du PNUE/PAM en matière de données, telle qu'elle figure à l'annexe I.
2. la demande à l'INFO/RAC de fournir le soutien technique nécessaire aux Parties contractantes et de répondre à tous les besoins identifiés pour mettre en œuvre pleinement la Politique de gestion des données du PNUE/PAM.
3. l'appel aux parties contractantes à prendre des mesures efficaces pour mettre en œuvre la Politique de gestion des données du PNUE/PAM.

La Politique de gestion des données prend en considération deux acteurs principaux :

- L'INFO/RAC, auquel il est demandé d'apporter un soutien technique aux parties contractantes pour l'application de la Politique de gestion des données (par exemple, par le biais de réunions bilatérales, de formations, etc.) et la mise en place d'une infrastructure technologique efficace pour soutenir le processus.
- Les parties contractantes, qui sont appelées à mettre en œuvre la Politique de gestion des données avec l'aide de l'INFO/RAC par le biais, par exemple, de la préparation des données à intégrer dans le système d'information IMAP, l'harmonisation des géodatabases afin d'être correctement reliées par InfoMAPNode, la normalisation des données pour atteindre l'interopérabilité des systèmes.

Les **objectifs** de la Politique de gestion des données sont multiples et peuvent être classés en quatre groupes en fonction de leur objet :

1. Les qualités envisagées pour les éléments. La Politique de gestion des données favorise la disponibilité rapide de la version la plus récente de l'ensemble de données et le maintien de séries de données à long terme ; l'assurance qualité et les procédures d'assurance et de contrôle de la qualité des données ; la protection de l'intégrité, de la transparence et de la traçabilité des données, ainsi que la traçabilité des données environnementales.

2. Les contraintes subies. La Politique de gestion des données vise à respecter la législation nationale pertinente et à reconnaître les droits de propriété intellectuelle dans le partage des données.

3. Les principes qui sous-tendent la politique. La Politique de gestion des données permet l'interopérabilité et l'application de normes, l'accès libre et gratuit aux données du PAM, la réutilisation de données venant de différentes sources, la mise en œuvre des principes INSPIRE, SEIS, Copernicus et GEOSS.

4. Les éléments concernés. La Politique de gestion des données vise à soutenir la gestion des données de recherche, la publication de métadonnées et l'utilisation de données d'externalisation ouverte, telles que les données issues des actions de la science citoyenne.

Les principes qui sous-tendent les prescriptions

Les principes qui guident la Politique de gestion des données sont divers et s'inspirent de principes plus larges promus par des organisations internationales telles que GEOSS (Système mondial des systèmes d'observation de la Terre), INSPIRE (Infrastructure d'information géographique dans la Communauté européenne), SEIS (Système d'information environnementale partagé) et Copernicus. La Politique de gestion des données adapte ces principes généraux au cas spécifique du PNUE-PAM et les traduit en objectifs (groupe 4).

Un principe général sous-tend tous les autres : « *Les données doivent être gérées au plus près de leur source, collectées une seule fois, partagées avec d'autres, et facilement disponibles pour remplir le mandat du PNUE-PAM.* »

Dans cette phrase, il y a des implications :

- Éviter la duplication des données, promouvoir une gestion des données FAIR (faciles à trouver, accessibles, interopérables et réutilisables).
- Promouvoir l'utilisation de services et d'outils interopérables. Lorsque cela est possible, il est juste d'exploiter le travail des autres.
- Promouvoir un changement d'état d'esprit pour passer de la propriété des données à leur gestion (« les données sont un bien public, je comprends leur utilité pour tout le monde, je ferai tout mon possible pour promouvoir le progrès »).

En dehors de ce principe général et de ses implications, qui sont fondamentaux pour la gestion des données, la poursuite de l'interopérabilité est également d'une importance primordiale.

L'interopérabilité signifie pour l'utilisateur final que tout élément d'information peut être partagé entre plusieurs acteurs avec le même niveau de qualité (qualité des instruments, qualité de l'information, qualité de l'élaboration). Pour y parvenir, la voie à sens unique est celle des normes : des règles universellement reconnues pour partager l'information (protocoles). Un exemple simple est le protocole http.

Un autre principe pertinent est celui du **libre accès**. Les données doivent être aussi ouvertes que possible, en respectant les contraintes imposées par la législation locale, la sensibilité des données et les droits d'auteur. En élargissant l'horizon, le libre accès signifie également favoriser la possibilité pour un plus grand nombre de pays d'accéder à la connaissance. De plus, il sous-tend également le principe selon lequel les retombées des données collectées au moyen de fonds publics devraient être disponibles pour le grand public.

Pour formaliser le libre accès, des licences Open Data doivent être appliquées aux données. Parmi toutes les possibilités, le PNUE-PAM a choisi de considérer les licences Creative Commons pour ses données. Il convient de noter que, contrairement au concept commun de « licence » (qui interdit généralement aux utilisateurs d'agir d'une certaine manière), les licences Creative Commons ne se concentrent pas sur ce que l'utilisateur ne peut pas faire avec les données, mais sur les droits transférés de l'auteur (propriétaire des droits d'auteur) à l'utilisateur. Elles réglementent donc les possibilités offertes à l'utilisateur par cette licence spécifique, et ce changement de mentalité est à la base de l'initiative Creative Commons. S'il existe une variété de licences possibles pour les données, de l'absence de droits d'auteur (en vert, figure 1) à la présence de droits d'auteur (en rouge, figure 1), il y a de nombreuses nuances : la possibilité de partager des données en attribuant l'auteur (licence CC-BY), en partageant des données avec la même licence que les données originales (licence CC-BY-SA), en autorisant (ou non) leur exploitation économique (licence CC-BY-NC), en recombinaison des données originales (licence CC-BY-ND) nous pouvons dire que (pour la définition des données ouvertes) les licences de données ouvertes ne sont que 2 : CC-0 et Creative Commons Attribution. Bien que la présente Politique de gestion des données encourage l'utilisation par défaut de la licence CC-BY, d'autres licences plus restrictives pourraient être utilisées éventuellement pour des données sensibles.

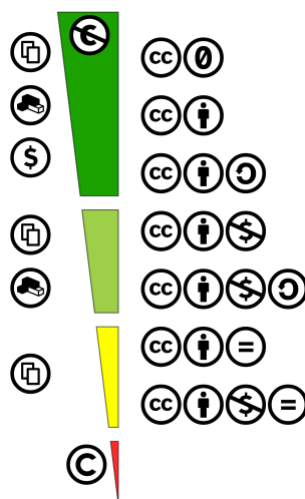


Figure 1 - Les licences Creative Commons réglementent les autorisations des utilisateurs entre les deux extrêmes : absence de droits d'auteur (vert) et présence de droits d'auteur (données en accès restreint, rouge).

Le tableau 1 suivant complète la figure 1 en donnant une explication complète de la licence, de sa signification, de ce qui peut ou ne peut pas être fait avec les données, et en indiquant si la licence peut être classée comme conforme aux données ouvertes ou non.

Symbole	Nom	Signification	Ce qui peut être fait	Données ouvertes ?
	CC-0	Par cette licence, le propriétaire des données renonce à tout droit sur les données, y compris celui d'être reconnu comme l'auteur.	Faire une copie des données, les recombinaison, les utiliser à des fins commerciales	Oui
	CC-BY	Par cette licence, le propriétaire des données est reconnu lorsque les données sont utilisées à quelque fin que ce soit.		Oui
	CC-BY-SA	Par cette licence, le propriétaire des données est reconnu lorsque les données sont utilisées et l'utilisateur est obligé d'utiliser la même licence si les données sont partagées.		Oui
	CC-BY-NC	Par cette licence, le propriétaire des données est reconnu, mais les données ne peuvent être utilisées qu'à des fins non commerciales.	Faire une copie des données, les recombinaison	Non
	CC-BY-NC-SA	Par cette licence, le propriétaire des données est reconnu, mais les données ne peuvent être utilisées qu'à des fins non commerciales, et l'utilisateur est obligé d'utiliser la même licence si les données sont partagées.		Non
	CC-BY-ND	Par cette licence, le propriétaire des données est reconnu, mais les données ne peuvent pas être utilisées (analysées, intégrées,	Faire une copie des données	Non

		recombinées) pour créer de nouveaux produits.		
	CC-BY-NC-ND	Par cette licence, le propriétaire des données est reconnu, mais les données ne peuvent pas être utilisées (analysées, intégrées, recombinaées) pour créer de nouveaux produits, et les données ne peuvent être utilisées qu'à des fins non commerciales.		Non

Tableau 1 - Explication complète des licences Creative Commons.

Le dernier principe pertinent examiné est la promotion de la **réutilisation des données**. L'un des objectifs étant d'éviter la duplication des efforts, les données doivent pouvoir être réutilisées, exploitées et recombinaées à partir de différentes sources dans différents cadres et supports.

La réutilisation des données, des métadonnées et de l'infrastructure n'est possible que si :

- Les méta(données) sont richement décrites par une pluralité d'attributs précis et pertinents
- Les (méta)données sont diffusées avec une licence d'utilisation des données claire et accessible
- Les (méta)données sont associées à une provenance détaillée
- Les (méta)données respectent les normes communautaires pertinentes pour le domaine.

L'implication pour la réutilisation de la provenance est claire : en connaissant l'origine d'une donnée, il est possible de contacter le propriétaire/la personne de contact ou le responsable des données en cas de problèmes (tels que des bugs ou des erreurs) sur les données. Cela augmente considérablement les possibilités de réutilisation des données.

Qui plus est, la présence d'attributs pertinents augmente la possibilité de réutiliser les données : le choix d'attributs codifiés dans des vocabulaires ou des normes augmente fortement la valeur des données en les rendant plus compréhensibles et souhaitables pour l'utilisateur final.

Implications pour les parties contractantes

Deux groupes principaux d'implications concrètes pour les parties contractantes ont été identifiés.

Le premier groupe rassemble les implications s'appliquant de manière générique à toutes les données, le second groupe se concentre sur les implications s'appliquant uniquement aux données restreintes. Toutes les implications (génériques et particulières) doivent être prises en considération par les pays lorsque des données sont partagées et les contraintes éventuelles sur les données doivent être vérifiées AVANT de prendre toute mesure pour le partage des données.

Implications communes pour toutes les données :

- S'il existe des compétences adéquates, les données doivent être traitées aussi près que possible de leur source, en suivant le principe de ne pas dupliquer les données.
- Les données spatiales doivent être en format standard (ISO, OGC et INSPIRE) afin d'être disponibles via une infrastructure interopérable pour les informations spatiales.
- Les données sont mises à disposition gratuitement dans un délai minimum.

- Si les données ont un identifiant unique (URI ou DOI, par exemple), celui-ci doit être mentionné lors de la fourniture ou de la déclaration des données.
 - Les données doivent être fournies avec leur localisation (ou leurs coordonnées Latitude/Longitude) chaque fois que possible pour les données environnementales.

Implications particulières pour les données sensibles :

- Dans le cas de données appartenant aux parties contractantes ou à des tiers, les éventuels obstacles ou limitations au partage des données doivent être vérifiés afin d'évaluer leur conformité avec la Politique de gestion des données avant que toute action ne soit prise sur les données. Egalement, les droits de propriété intellectuelle, les conditions d'utilisation ou de réutilisation, la confidentialité et la déclaration de qualité des données doivent être vérifiés.

En particulier, les restrictions sur lesquelles il convient d'attirer l'attention sont les suivantes :

- Règles contraignantes
- Traités internationaux
- Législation nationale
- Données personnelles (protection des)
- Confidentialité statistique
- Protection des droits de propriété intellectuelle
- Protection de la sécurité nationale
- Défense
- Sécurité publique
- Embargo

Les restrictions demandées doivent être démontrées à l'aide de références spécifiques. Les raisons de la fermeture des données peuvent être principalement liées à quatre natures différentes :

1. Juridique – dans ce cas, il existe des lois (nationales ou internationales) qui réglementent l'utilisation des données. Ce cas comprend également toutes les raisons susmentionnées de défense, de sécurité publique et de propriété intellectuelle.
2. Vie privée – ce cas fait référence à la possible fuite de données personnelles, donc à des données liées d'une certaine manière au Règlement général sur la protection des données RGPD.
3. Recherche – il s'agit ici des données impliquées dans des activités de recherche en cours de développement ou de publication.
4. Conservation – il s'agit de données concernant, par exemple, des espèces menacées, dont la position pourrait être préjudiciable aux espèces elles-mêmes (par exemple, la position exacte des oiseaux migrants).

Si certaines de ces données ne sont pas partageables du tout (les données relatives à la défense ou à la sécurité nationale, par exemple), d'autres peuvent être partiellement ouvertes dans des conditions spécifiques liées à la fois à la licence et à la manipulation des données.

Par exemple, dans la plupart des cas de « conservation » (point 4), la position des données est cruciale pour définir leur nature « sensible ». Une solution possible à la fermeture des données pourrait être le flou des données en ce qui concerne leur position ou l'agrégation spatiale qui fait que les données ne se réfèrent pas à un point spécifique mais, par exemple, à une région ou à un pays.

Dans ces cas, où les données pourraient être partiellement ouvertes, un plan de gestion spécial doit être mis en place entre le propriétaire des données (par exemple la partie contractante), le fournisseur (par exemple le CAR thématique) et le partageur de données (par exemple l'INFO/RAC).

En tout état de cause, si les données sont « restreintes » ou stratégiques pour les parties contractantes, elles ne seront pas ouvertes au grand public, mais ne seront partagées qu'avec les utilisateurs appropriés (autorisés).

Si nécessaire, les données confidentielles ou sensibles peuvent être reclassées ou agrégées en coopérant avec l'INFO/RAC afin d'ouvrir l'ensemble de données au moins partiellement.

Si les données font l'objet d'une publication scientifique ou sont impliquées dans un contrat de consortium ou dans l'enregistrement d'un brevet, elles peuvent être soumises à un embargo. Afin de formaliser l'embargo (qui ne devrait pas durer moins de 24 mois, idéalement), le partenaire doit motiver la demande (d'embargo) et la durée de l'embargo doit être explicitement mentionnée dans les métadonnées.

Conclusions

En conclusion, lorsque des données sont produites et partagées dans le cadre du système PAM, il convient de tenir compte d'un ensemble de règles en matière de Politique de gestion des données. Ces règles sont établies dans le contexte plus large d'un ensemble de principes qui favorisent le libre accès, l'interopérabilité et la gestion FAIR des données.

En général, puisque les données sont un bien public, les retombées de la collecte de données devraient concerner l'ensemble de la communauté, ainsi les données devraient être « aussi ouvertes que possible, aussi fermées que nécessaire ».

Si, pour une raison quelconque, les données doivent être partagées dans des conditions restrictives, il est nécessaire de vérifier si ces nécessités existent réellement et de choisir, avec le personnel spécialisé, la restriction adéquate et la plus appropriée pour ces données. Le propriétaire des données, le fournisseur des données et l'INFO/RAC pourraient également discuter d'une éventuelle agrégation ou reclassification des données sensibles.