

Appel à participation aux Groupes de Travail du consortium DataCite France

Contexte

DataCite est une association à but non lucratif, sous le régime de droit allemand, inscrite au registre des associations sous le nom « DataCite – International Data Citation Initiative ». Elle a pour objectif d'attribuer des identifiants numériques pérennes, appelés DOI (*Digital object identifier*, en français « identifiant numérique d'objet ») aux produits de la recherche dans leur définition la plus large : jeux de données, articles scientifiques, textes, logiciels, OMP (*Outputs Management Plan*, en français « plans de gestion de produits de recherche »), audio, visioconférence, image, etc. Ces DOI servent à identifier, décrire et découvrir les données de recherche de manière univoque (publications, entrepôts et centres de données, etc.). Ils sont utilisés dans de nombreux systèmes d'information des établissements de l'enseignement supérieur, la recherche et l'innovation.

Depuis 2010, le CNRS et par délégation l'INIST est engagé auprès de DataCite pour piloter une adhésion consortiale des institutions françaises. Cette adhésion constitue un des engagements français de l'Open Government Partnership (engagement 18 : Construire un écosystème de la « science ouverte »).

Objectifs du consortium DataCite France

Son objectif est de développer et maintenir une communauté de pratique, de promouvoir l'intégration des identifiants dans les systèmes d'information des établissements et infrastructures de recherche français. La dynamique qui en résultera permettra de renforcer la représentation française et de porter une vision collective au sein du consortium international DataCite.

La gouvernance du consortium DataCite France

Le consortium DataCite France se compose des instances suivantes :

- Une Assemblée générale, organe d'échange et de décision du consortium, qui regroupe l'ensemble des adhérents au consortium ;
- Un Comité exécutif, pour la mise en œuvre du projet et de son plan d'actions, qui rend compte à l'Assemblée générale ;
- Un porteur du consortium (ou *consortium lead*) qui assure la coordination politique et technique entre les membres du consortium DataCite France et le consortium international DataCite.

Les groupes de travail

Les travaux du consortium seront organisés au travers de groupes de travail ; ces derniers se composent de membres du consortium, ainsi que d'un membre du Comité exécutif et d'un membre désigné par le porteur du consortium DataCite France. Des membres extérieurs tels que des experts peuvent également faire partie d'un groupe de travail.

Chaque groupe de travail proposera au Comité exécutif un plan d'action et un livrable. Un groupe de travail peut avoir une durée de vie variable et clôturera son activité lorsque le livrable sera fourni.

Chaque groupe de travail aura pour animateur(s) un ou des membres du Comité exécutif. Les réunions se feront par visioconférence. Il est attendu que chaque groupe de travail se réunisse au moins six fois par an.

Les premiers groupes de travail (GT)

No. GT	Titre du GT	Animation du GT
1	Métadonnées	Davien BLANC Michel JACOBSON
2	Interactions avec ORCID et autres PIDs	Geneviève MICHAUD
3	Modèle économique et périmètres	Pascal AVENTURIER Valérie WARTH (Inist) Geneviève MICHAUD
4	Supports, sensibilisation et formations	Joel SUDRE
5	Communication	Geneviève MICHAUD Claire FRANCOIS

Description des groupes de travail

<p>Titre du GT : Métadonnées</p> <p>Objectifs :</p> <ul style="list-style-type: none"> • Fournir des métadonnées permettant de rendre FAIR les données de recherche ; • Définir la politique à suivre vis-à-vis des vocabulaires contrôlés et les ontologies à utiliser afin d'améliorer l'interopérabilité et la <i>Machine-Actionability</i>. <p>Problématiques à traiter : Les six propriétés obligatoires du schéma de métadonnées DataCite ne sont pas suffisantes pour garantir la FAIRisation minimale des données de la recherche. Ainsi, il sera notamment abordé deux sujets clés. D'une part, définir les champs de métadonnées à recommander pour décrire les jeux de données. D'autre part, fournir des recommandations sur l'attribution de DOI à différents niveaux de granularité allant des jeux de données spécifiques à des collections de données, mais aussi permettant la citation de fragments de données.</p> <p>Pistes de livrables : Guide de bonnes pratiques ; Recommandations sur un noyau de métadonnées à utiliser.</p> <p>N.B. : Interactions avec le groupe de travail sur les métadonnées du consortium international qui a défini le schéma de métadonnées DataCite (Metadata Working Group).</p>

Titre du GT : Interactions avec ORCID et autres PIDs

Objectifs :

- Encourager l'usage des différents identifiants pérennes dans la description des produits de la recherche ;
- Coordonner les actions de DataCite France avec ORCID et toute autre entité gérant un PID utilisable dans le schéma DataCite.

Problématiques à traiter : L'existence de plusieurs types d'identifiants pérennes, tels que les ORCID, les ROR, les DOI, et leur fragmentation rendent difficile l'utilisation de ces derniers. Afin de généraliser leur adoption, les sujets à traiter sont l'analyse des documentations autour de ces identifiants, des politiques mises en place à leur propos, afin d'être en capacité de préparer les recommandations opérationnelles. Enfin, le lien entre les DOI des jeux de données et ceux des publications associées a posteriori sera également abordé.

Pistes de livrables : Liste d'identifiants pérennes recommandés avec les URI de leur schéma. À intégrer dans le guide de bonnes pratiques qui sera défini par le GT Métadonnées.

N.B. : Suivi du comité de coordination sur l'ensemble des PIDs piloté par le MESRI dans le cadre de la politique des données, des algorithmes et des codes sources. Lien avec les travaux de la Task Force EOSC sur ce thème.

Titre du GT : Modèle économique et périmètres

Objectifs :

- Proposer un nouveau modèle économique interne au consortium français en tenant compte des évolutions du modèle économique défini au niveau du consortium international.

Problématiques à traiter : Le modèle économique actuel n'est pas supportable par le porteur (Inist). Pour répondre à cette question pressante, les sujets à aborder concernent la typologie des membres (étrangers, privés, associations à caractère scientifique ?), l'estimation de l'ensemble des coûts (cotisation vers DataCite International, coûts internes, frais liés à l'animation du consortium), le financement (service public ? prix coûtant ?) et la comparaison avec les autres consortiums internationaux.

Pistes de livrables : Scénarios de modèle économique ; Déclinaisons dans les contrats et devis.

Titre du GT : Supports, sensibilisation et formations

Objectifs :

- Sensibiliser sur la complétude des métadonnées ;
- Sensibiliser sur l'utilisation des vocabulaires contrôlés et ontologies ;
- Généraliser l'utilisation des identifiants pérennes associés à des métadonnées de qualité permettant la découverte et la réutilisation des données de recherche ;
- Faire connaître et accompagner à l'usage du guide de bonnes pratiques préparé par le GT Métadonnées.

Problématiques à traiter : Comment accompagner la mise en œuvre des bonnes pratiques ? Pour cela, il est important d'agréger et rendre accessibles les supports et documents de références, de développer un réseau de formateurs pour mettre en place cette sensibilisation. Enfin, il sera nécessaire d'interroger et d'assurer la qualité des formations. Des exemples de thématiques concernent la création de DOI sur la plateforme de gestion de DOI de DataCite (Fabrica) et l'utilisation des API DataCite (construction des requêtes sur ces API).

Pistes de livrables : Formations ; Supports en français ; FAQ ; Traduction en français du schéma de métadonnées DataCite ; Cartographie des supports et formations nécessaires (dans quel cas et pour qui).

N.B. : Collaboration avec les URFIST ; Lien avec les ateliers de la donnée et les centres de ressource de Recherche Data Gov.

Titre du GT : Communication

Objectif :

- Proposer et mettre en œuvre un plan de communication aussi bien interne au consortium DataCite France qu'externe à celui-ci.

Problématiques à traiter :

Communication interne : comment faciliter un bon partage des avancées de chaque GT, entre eux, avec le Comité exécutif et l'ensemble des membres ? Comment faciliter la restitution au consortium DataCite France des décisions, activités et informations provenant de DataCite International ? Comment organiser une préparation de l'AG annuelle DataCite international avec les membres du consortium DataCite France ?

Communication externe : faire connaître le consortium DataCite France et accompagner les actions de sensibilisations et de formations.

Pistes de livrables : Plan de communication et outils associés (site web ; forum de discussion ; organisation de rencontres).

Les réponses sont à envoyer à pid.opidor@inist.fr jusqu'au 10 novembre 2023 en indiquant le nom et prénom, la structure, le poste actuel, les objectifs et le(s) groupe(s) de travail souhaité(s).