

Les identifiants pérennes – Un aperçu

Introduction générale

Dans le contexte de la recherche scientifique, on distingue deux grands types d'identifiants pérennes :

- les **identifiants Objet**, pour vos publications et vos données de la recherche,
- les **identifiants Contributeurs**, pour les personnes et les institutions.

A noter qu'on parle souvent de PID pour **Persistent Identifier** en anglais.

Plusieurs systèmes d'identification mettent en place des possibilités de liens entre les auteurs et leurs productions scientifiques, c'est-à-dire leurs publications mais aussi leurs données.

L'utilisation conjointe de ces identifiants présente en effet un triple intérêt. Elle favorise :

- l'agrégation de la production scientifique et son association à son auteur,
- l'augmentation de la visibilité du chercheur et de ses travaux ainsi que la visibilité des établissements et des institutions,
- la recherche documentaire et la découverte de l'information.

Définition d'identifiant pérenne

Un identifiant est une association unique entre un code alphanumérique et une entité ou une ressource.

Dans le monde numérique vaste et ouvert du web, les ressources sont localisées par des URL's. Cependant ces URL's ne sont pas stables. Si la ressource est déplacée et/ou renommée, elle n'est plus accessible. Le navigateur affiche alors le code erreur 404.

Les identifiants pérennes garantissent quant à eux un lien stable à la ressource en ligne. Cette pérennité est obtenue grâce à une gestion active des URLs. Cette gestion est assurée par des organisations reconnues, portées par des infrastructures humaines et techniques. Il s'agit de faire correspondre en permanence l'identité de la ressource à sa localisation sur le web. On y accède à l'aide d'un lien hypertexte en ayant la garantie que ce lien ne sera pas brisé.



Identifiant pérenne Objet

Les identifiants pérennes Objet concernent principalement vos publications et données de recherche.

Quels sont les différents types d'identifiants pérennes Objet ?

Les identifiants pérennes Objet sont assignés aux ressources issues de la production scientifique, par exemple les articles, les jeux de données, les images, les sons, les objets physiques...

Il existe déjà des identifiants tels que l'ISBN pour les livres et l'ISSN pour les revues.

Cependant, avec le déploiement d'internet et la mise en ligne de publications et de données de la recherche, des identifiants mieux adaptés à l'univers numérique ont été mis en place comme :

- Handle
- PURL (persistant URL)
- ARK (Archive Resource Key)
- ePIC (European Persistent Identifier Consortium)
- et le DOI (identifiant numérique d'objet).

A quoi servent-ils ?

- Les identifiants pérennes « objet » permettent d'identifier la ressource et d'y accéder de manière stable sur le long terme.
- Ils font le lien entre les articles publiés et les jeux de données sous-jacentes.
- Ils aident aussi à la découverte, au partage, à la réutilisation et à la citation des résultats de la recherche et de la production scientifique.

Comment les choisir et les obtenir ?

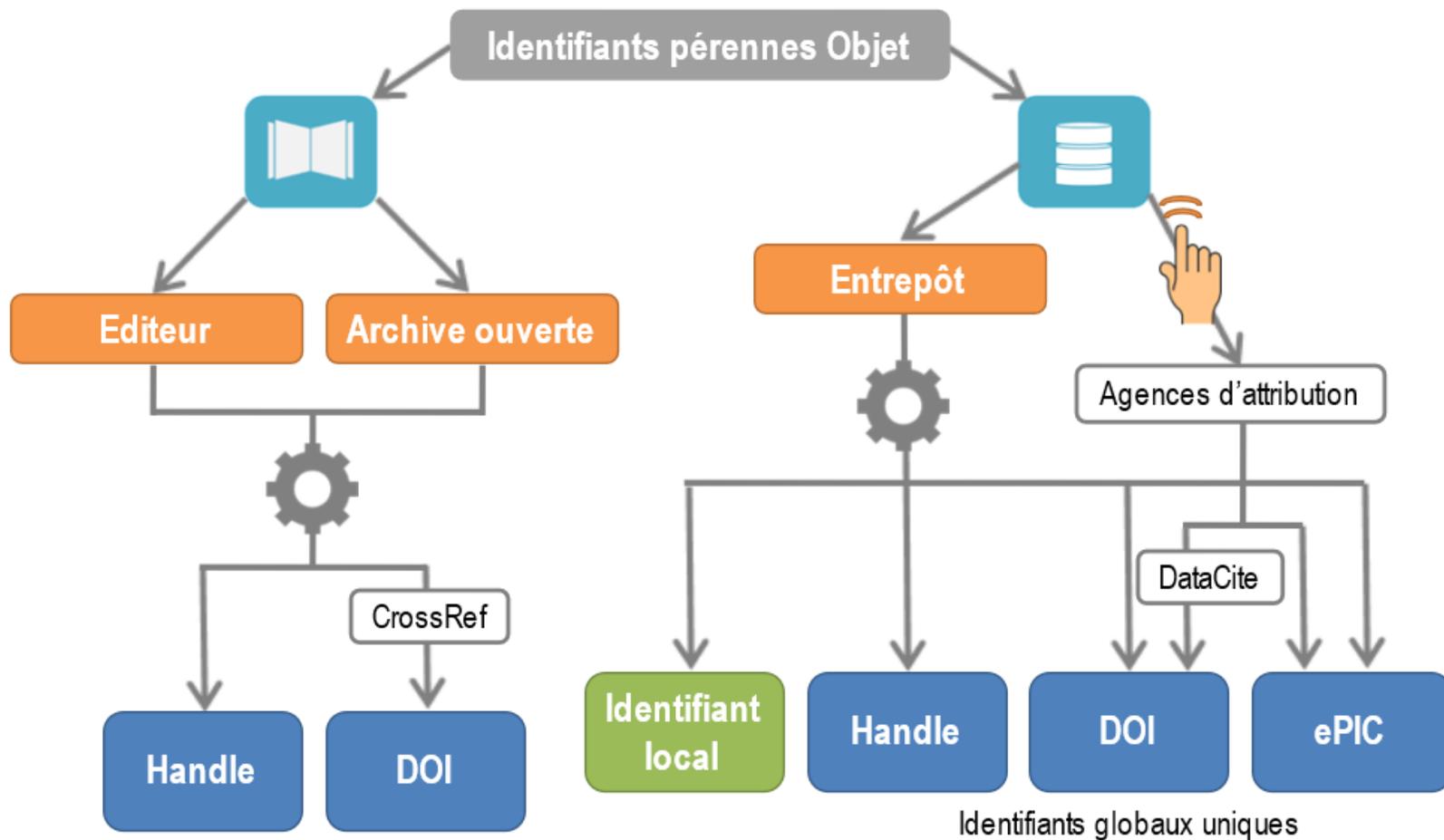
Pour les publications, l'attribution d'un identifiant pérenne est une procédure bien établie et systématisée.

La plupart des éditeurs et des archives ouvertes octroient de manière automatique un identifiant pérenne à chaque article. Il s'agit le plus souvent d'un Handle ou d'un DOI. Ce dernier est attribué par l'intermédiaire de l'agence CrossRef.



Lorsque vos données sont déposées dans un entrepôt, des identifiants leur sont attribués : il peut s'agir d'un identifiant local, ou d'un identifiant global unique, par exemple un Handle, un DOI ou un ePIC.

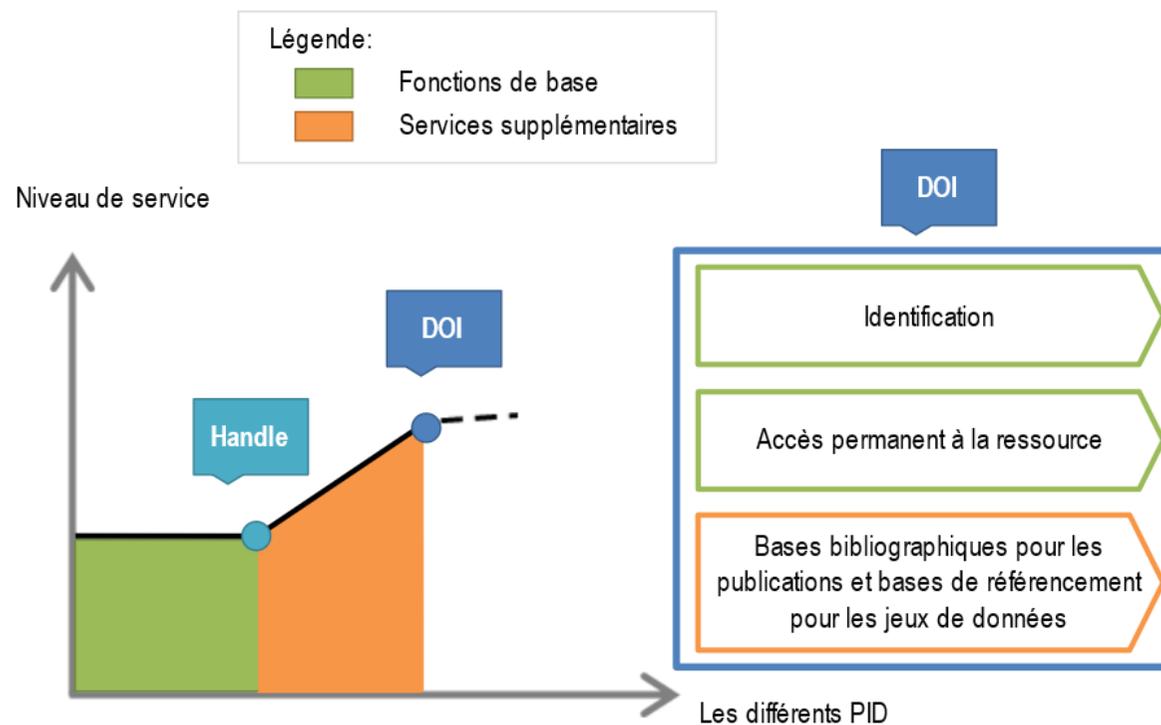
Les institutions et les centres de données peuvent également obtenir des DOI directement en contactant une agence membre de DataCite.



Pour l'accessibilité sur internet et une meilleure interopérabilité, il est recommandé d'utiliser des identifiants globaux uniques basés sur des standards et largement utilisés comme le DOI, ePIC ou Handle.

Il faut savoir que tous les systèmes d'identifiants pérennes offrent les mêmes fonctions de base à savoir l'identification et l'accès permanent à la ressource. Le choix dépend du niveau de service offert par chaque système.

- Le Handle est un système robuste qui offre ces fonctionnalités de base.
- CrossRef et DataCite, autour du DOI, offrent des services supplémentaires, comme la création et le maintien de bases bibliographiques pour les publications et des bases de référencement pour les jeux de données.

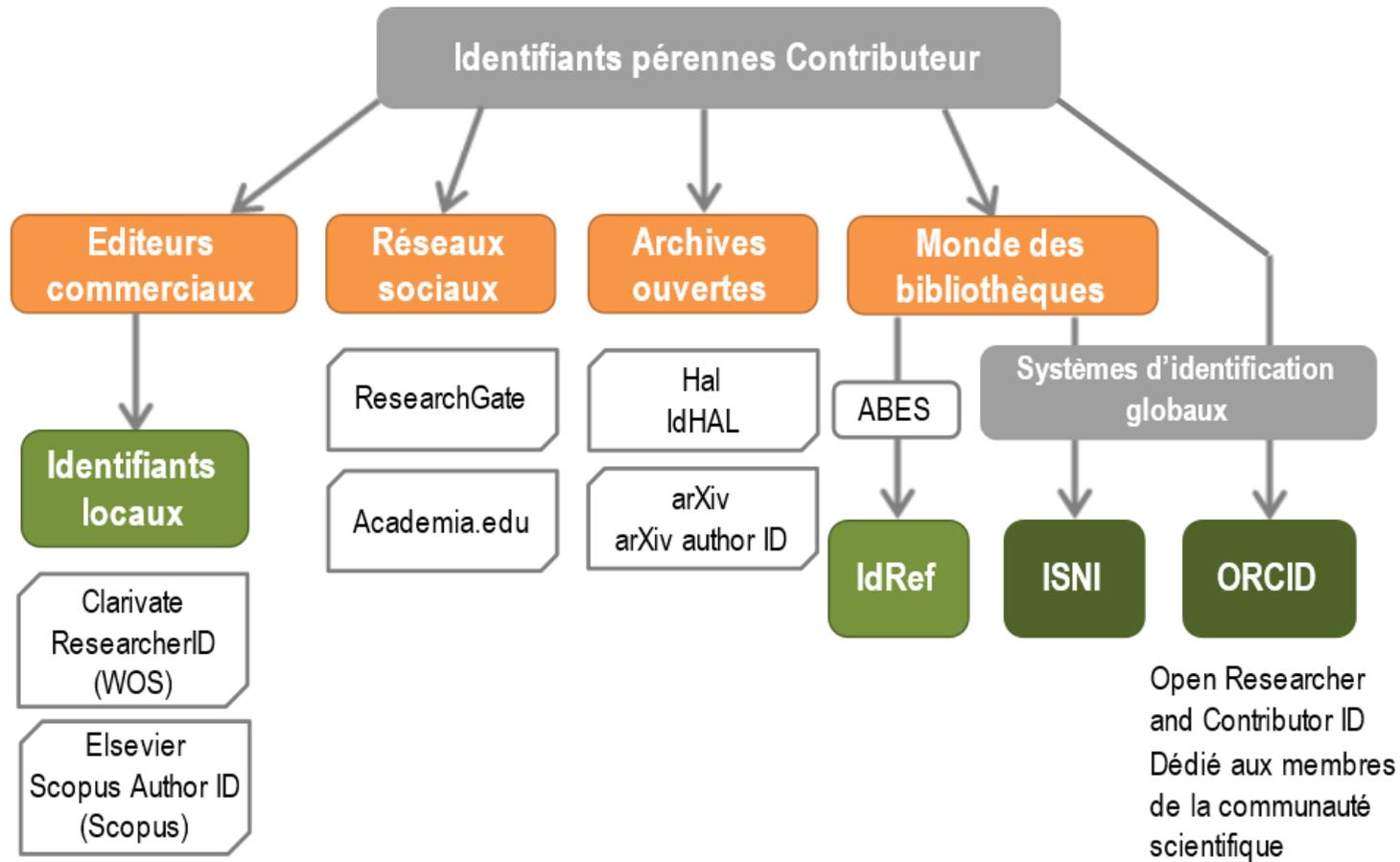


Identifiant pérenne Contributeur

Les identifiants Contributeur sont associés aux auteurs de la production scientifique et aux organismes qui assurent cette production.

Quels sont les différents types d'identifiants pérennes Contributeur?

Il existe plusieurs types d'identifiants dédiés aux auteurs, aux contributeurs et aux institutions impliqués dans la recherche :



- **les éditeurs commerciaux** attribuent des identifiants locaux pour leur base de données : par exemple Clarivate (anciennement Thomson-Reuters) avec l'identifiant ResearcherID ou Elsevier avec Scopus Author ID
- **les réseaux sociaux** comme ResearchGate et Academia.edu attribuent à chaque inscrit son propre identifiant
- **les archives ouvertes** peuvent proposer la création d'un identifiant local, par exemple idHAL pour l'archive ouverte HAL, ou arXiv author ID pour l'archive ouverte Arxiv
- **dans le monde des bibliothèques :**
 - L'ABES, agence bibliographique de l'enseignement supérieur en France, utilise l'identifiant IdRef pour son catalogue SUDOC.
 - L'identifiant ISNI (International Standard Name Identifier) est un identifiant international attribué aux personnes et institutions impliquées dans la production littéraire, artistique et intellectuelle au sens large. Il est défini par une norme ISO.
 - Enfin, ORCID (Open Researcher and Contributor ID) est un système d'identification global, ouvert et gratuit qui s'appuie sur ISNI. Il a été créé par un ensemble d'institutions de recherche, de sociétés savantes, de fondations caritatives et d'éditeurs. ORCID est dédié spécifiquement aux membres de la communauté scientifique.

A quoi servent-ils ?

L'attribution d'identifiants pérennes aux personnes et aux institutions est un moyen efficace de désambiguïser les noms et de résoudre les problèmes d'homonymie.

Exemple :

- l'animateur de télévision Jacques MARTIN (ISNI 0000 0000 7359 228X)
- l'auteur de bande dessinée Jacques MARTIN (ISNI 0000 0001 2033 0240).

D'autre part le chercheur ou l'institution qui a un identifiant pérenne global augmente sa visibilité académique.

Enfin, un identifiant pérenne Contributeur permet d'agréger :

- à un chercheur → ses travaux,
- à un établissement → toutes les publications produites par ses chercheurs,
- à un organisme financeur → les productions scientifiques issues des projets qu'il finance.



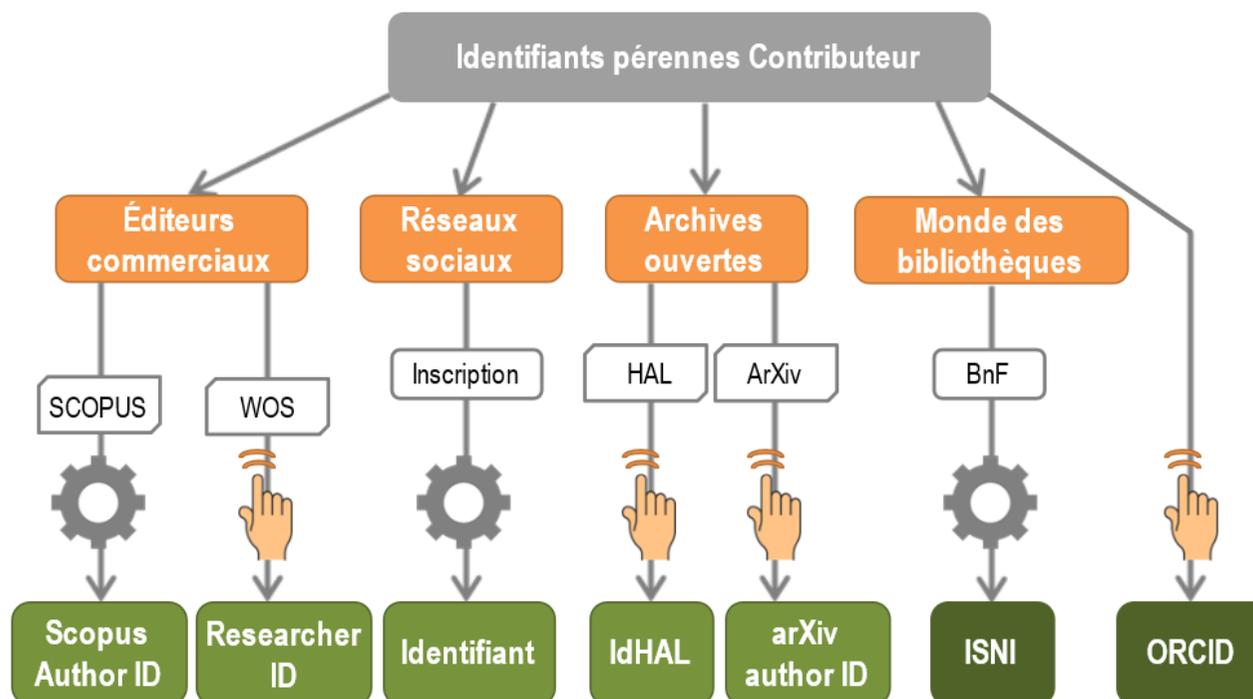
Comment les choisir et les obtenir ?

Souvent, l'obtention d'un identifiant Contributeur se fait automatiquement :

- si vos articles sont référencés dans la base de données Scopus de Elsevier, un identifiant auteur vous est attribué,
- si vous êtes inscrit sur les réseaux sociaux scientifiques, vous avez reçu un identifiant lors de votre inscription,
- si vous êtes l'auteur d'une œuvre référencée par la BnF (bibliothèque nationale de France), un identifiant ISNI vous a été délivré.

En revanche, l'obtention de certains identifiants nécessite une action de votre part. C'est le cas pour les identifiants ResearcherID, IdHAL et Arxiv author ID.

Enfin, si vous voulez un identifiant global, vous pouvez créer votre propre profil en ligne et le gérer sur la plateforme ORCID.



Consignes du CoSo

Le Comité pour la Science ouverte (CoSo) a publié en 2019 une **Note d'orientation** intitulée "**Des identifiants ouverts pour la science ouverte : synthèse**".

<https://www.ouvrirlascience.fr/des-identifiants-ouverts-pour-la-science-ouverte-synthese/>

Voici quelques extraits :

« La mise sur pied du Comité pour la Science ouverte a donné le coup d'envoi à une réflexion conduite par les opérateurs de l'Enseignement supérieur et de la recherche français destinée à améliorer la structuration des identifiants les plus utiles, à accélérer leur adoption par les communautés, et à les rendre plus libres et pérennes. »

Recommandations sur les PID Contributeur :

« ...il n'existe pas encore d'identifiant international standard pour les organisations ni les affiliations. Ces entités sont répertoriées dans des registres nationaux (en France RNSR, AURÉHAL structures) et internationaux (ISNI, GRID, ROR, ...) encore imparfaitement alimentés et connectés. L'un des identifiants d'organisations les plus importants, Ringgold ID, de statut commercial, est utilisé par le monde de l'édition. » (Page 2 de la Note)

« En France plusieurs identifiants sont utilisés, pour les personnes, les affiliations et les publications, notamment ceux proposés par HAL via AURÉHAL (Accès Unifié aux Référentiels HAL), Huma-Num, la BnF et l'Abes. Pour l'Enseignement supérieur et la Recherche, IdRef est un identifiant « universel », au sens où il couvre les personnes, une partie des publications, et les structures, quand l'entité est française ou reliée à d'autres entités utilisées en France. » (Page 2 de la Note)

« Le service de création de PIDs le plus connu aujourd'hui dans le monde de la recherche est ORCID, qui permet à chaque chercheur de créer son (ou ses) identifiant librement et gratuitement. En France, la plupart des services nationaux de dépôt de publications, d'artefacts numériques ou de jeux de données, sont connectés à ORCID. Le référentiel ouvert le plus riche, de plus en plus connu des chercheurs, est à ce jour Wikidata. Wikidata est une initiative relativement récente (2012), dont la gouvernance et le modèle économique ne garantissent pas encore complètement la pérennité. Il est alimenté notamment par les données provenant des principaux référentiels français. » (Page 6 de la Note)

« L'identification des structures de recherche et des affiliations est encore en pleine structuration à l'échelle mondiale. » (Page 9 de la Note)



« Le RNSR semble être à ce jour l’outil national le plus abouti permettant de repérer et d’identifier les structures de recherche. AURÉHAL du CCSD a quant à lui une granularité plus fine et peut identifier les équipes de recherche. » (Page 9 de la Note)

Recommandations sur les PID Objet :

« Du fait de la complexité des données à gérer et des usages, le sujet des identifiants des données de la recherche est encore peu mûr. Divers systèmes cohabitent, principalement basés sur DOI et Handle, mais les besoins des communautés, et les solutions pour y répondre, ne sont pas encore totalement définis. » (Page 11 de la Note)

« Une initiative internationale à laquelle la France participe a développé un système d’identifiant de logiciels prometteur (<https://www.softwareheritage.org>). Les archives institutionnelles et les bases de données elles-mêmes disposent d’un identifiant, l’ISSN (<https://portal.issn.org/>). » (Page 11 de la Note)

Union européenne et Freya :

« L’Union européenne soutient des projets transnationaux visant à développer l’utilisation des identifiants de manière coordonnée par les chercheurs et les institutions des différents pays, parmi lesquels Freya – Connected open identifiers for discovery, access, and use of research resources (<https://www.project-freya.eu>), dans lequel la France est représentée au sein du programme des « ambassadeurs ». Freya semble être amené à jouer un rôle important dans la définition de l’usage des identifiants dans l’European Open Science Cloud (<https://www.eosc-portal.eu>), plateforme européenne qui fédère les outils permettant aux chercheurs de stocker, gérer, analyser et réutiliser de larges quantités de données issues de la recherche.

Depuis 2012 la base Wikidata, issue de la fondation Wikimedia qui porte Wikipedia, est devenue progressivement le point de convergence mondial des identifiants ouverts. Les principaux systèmes d’identifiants décrits ci-dessus sont alignés avec Wikidata ou y versent leurs propres données. » (Pages 2 et 3 de la Note)

Position française et action du CoSo :

« En France, le plan national pour la science ouverte, manifestation d’une volonté politique forte, a accéléré un mouvement concerté autour d’une adhésion nationale à ORCID, un soutien à Open Citations qui s’appuie sur les DOI, ainsi qu’à un soutien aux principes FAIR, qui comprennent l’adoption d’identifiants uniques. Par ailleurs, le plan a créé le Comité pour la science ouverte, qui peut aujourd’hui à la fois proposer des objectifs opérationnels et coordonner un plan d’action national impliquant les acteurs de terrain. » (Extrait du site du CoSo concernant la Note)



« L'année 2019 marque le lancement d'une action coordonnée au niveau national pour le développement en France des identifiants pérennes pour la science ouverte, qui conduira à la mise en oeuvre d'un réel plan d'action. [...]

Cette action nationale coordonnée vise de manière concrète

- le développement de services utiles à partir de cas réels permettant de démontrer les atouts des PIDs
- la communication adaptée vers les chercheurs pour les guider dans leur usage quotidien des identifiants

Compte tenu des besoins identifiés et du niveau de maturité des solutions, elle se décompose en 4 actions distinctes menées en parallèle, classées selon l'urgence à agir et qui concernent :

- Les identifiants de structures et d'organisations, autour du Répertoire national des structures de recherche RNSR
- Les identifiants de chercheurs, en s'appuyant sur une adhésion collective à ORCID
- Les identifiants de publications, en amplifiant les usages du DOI
- Les identifiants de données, logiciels et objets numériques, en s'appuyant sur le groupe RDA France et les collèges et groupes de travail du CoSo. » (Extrait du site du CoSo concernant la Note)

A retenir

Les identifiants pérennes sont des éléments importants du mouvement de l'Open Science. Ils sont essentiels pour la diffusion, le partage et la réutilisation des résultats de la recherche scientifique.

L'association d'identifiants pérennes Contributeur et Objet facilite l'agrégation entre un auteur, son institution, ses articles et les données de recherche sous-jacentes.

Si vous souhaitez être accompagné dans votre choix d'identifiant pérenne, renseignez-vous auprès de votre institution qui dispose probablement d'un service dédié aux données de la recherche.

