



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

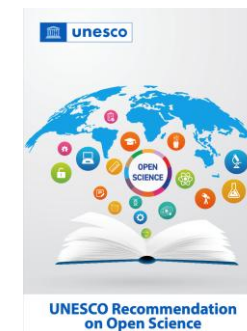
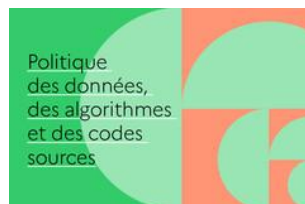


Les engagements de l'ANR en faveur de la science ouverte – Le Plan de Gestion des Données

Zoé Ancion – responsable pôle science ouverte
ANR



Un contexte national, européen et international



Une approche progressive

La politique science ouverte de l'ANR en **10 étapes clés**



Mise en œuvre de la politique science ouverte dans les appels à projets

Depuis 2019, lors du conventionnement d'un projet, les engagements des bénéficiaires portent sur deux volets :

- Le volet publications scientifiques
- Le volet données de la recherche
- + une recommandation sur le code source et logiciel (2023)
- + une recommandation sur l'accès ouvert pour les monographies (2024)

Les données de la recherche

Faciliter le partage et la réutilisation des données de la recherche en adoptant une démarche FAIR (Facile à trouver, accessible, interopérable, réutilisable).

Le coordinateur ou la coordinatrice du projet s'engage à fournir un Plan de Gestion des Données (PGD) qui devra être transmis à l'ANR à T0+6 et à la fin du projet.

- Il est recommandé de mettre à jour son PGD régulièrement.
- Il s'agit d'un livrable avec plusieurs versions.
- La réception de ces livrables conditionne le dernier versement de l'aide.
- L'ANR recommande l'utilisation de son [modèle de PGD structuré](#), disponible sur [l'outil DMP OPIDOR](#) de l'INIST pour une saisie en ligne

Le PGD prépare au partage et à l'ouverture potentielle des données de la recherche dans le respect du principe « aussi ouvert que possible aussi fermé que nécessaire ».

Code source et logiciel

Conformément au 2ème Plan national pour la science ouverte, l'ANR recommande que les logiciels développés pendant le projet de recherche soient mis à disposition sous une licence libre et que les codes sources soient archivés dans [Software Heritage](#) et décrits dans HAL en indiquant la référence (code décision) du projet ANR.

Les coûts éligibles

Les données de la recherche

- Toutes les dépenses liées à la gestion des données sont éligibles durant le projet : acquisition, collecte, stockage, personnel dédié à la gestion des données.....
- Le coût de stockage des données, en cas de recours à un tiers, est éligible, [jusqu'à 5 ans](#) après la date de fin scientifique du projet, sous réserve que le contrat avec ce tiers soit conclu avant la fin du projet.

Une approche concertée

Au niveau national

- Le Comité pour la science ouverte
- DMP OPIDoR (Inist – CNRS)
- Le réseau science ouverte entre les agences de financement françaises – ADEME, Anses, ANRS-MIE, FRM, INCa

Aux niveaux européen et international

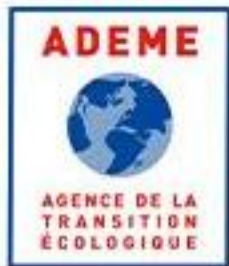
- RDA Funders Forum
- Science Europe Open Science Working Group



Le réseau science ouverte entre les agences de financement françaises

- ✓ [Déclaration conjointe](#) signée en 2020, animé par l'ANR
- ✓ Définir une approche concertée en faveur de la science ouverte et harmoniser nos pratiques
- ✓ Des engagements communs en faveur de la science ouverte pour les chercheurs financés par les 6 agences du réseau

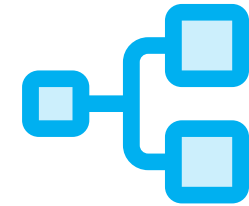
« Déployer un modèle commun de plan de gestion des données structuré. Afin de préparer les données à leur partage et diffusion éventuelle, les agences s'engagent à la mise en œuvre progressive d'un plan de gestion des données exploitable par les machines. Une première étape consistera au déploiement de leur modèle commun de PGD structuré, en s'appuyant sur l'outil [DMP OPIDoR](#). »



Un modèle commun à privilégier modèle Science Europe structuré

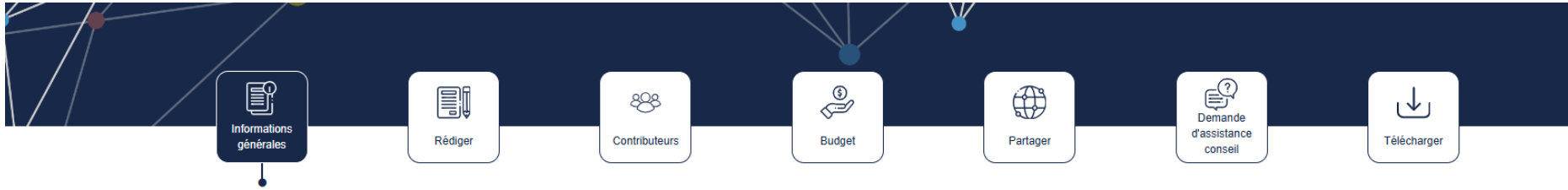
Adopté entre autres en France par l'ANR, le CNRS, le réseau science ouverte entre les agences de financement françaises sous la forme « ANR - Modèle de PGD structuré (français) »

- Organisé selon le cycle de vie des données
- Constitué de formulaires pour répondre aux questions et vous guider
- Facilite la saisie et la réutilisation d'informations
- Pour bénéficier de toutes les fonctionnalités développées



Encourage les bonnes pratiques pour une gestion des données facilitant leur partage dans le respect des principes FAIR

Récupérer automatiquement les informations administratives des projets financés par l'ANR



DMP du projet "Mettre le plan de gestion des données automatisé dans les mains des biologistes"

Cliquez ici si vous avez un projet financé

Si votre projet est financé par l'un des financeurs de la liste, vous pouvez récupérer automatiquement les informations administratives que vous avez déjà saisies lors de la demande de subvention.

Veuillez sélectionner un finaceur

Agence Nationale de la Recherche (ANR)

Puis sélectionnez l'acronyme, le titre ou l'identifiant du projet

maDMP4LS - Mettre le plan de gestion des données automatisé dans les mains des biologistes [ANR-19-DATA-0017]

Enregistrer

**Vous pourrez récupérer le titre,
acronyme, code décision,
résumé, nom des institutions
partenaires, nom du
coordinateur du projet**

*Disponible aussi pour les projets Commission
Européenne (Horizon Europe)*

Renseignements sur le projet

● Titre du projet

Mettre le plan de gestion des données automatisé dans les mains des biologistes

● Acronyme du projet

maDMP4LS

● Résumé du projet

B

I

U

A

L'Institut Français de Bioinformatique (IFB), avec ses deux infrastructures informatiques et ses 30 plates-formes, est une structure essentielle pour les Sciences de la Vie et ses applications (santé, agriculture, environnement), fournissant un environnement de production, d'analyse et de gestion de données pour les communautés d'utilisateurs. Bien que ces communautés puissent être considérées comme bien dotées d'outils de structuration et de gestion des données, des problèmes subsistent pour mettre pleinement en œuvre les bonnes pratiques et gérer les données, de leur production à leur préservation, en assurant leur accessibilité, leur reproductibilité et leur réutilisation. Le Plan de Gestion des Données (PGD ou DMP pour Data Management Plan) est considéré comme un élément-clé qui facilitera la mise en œuvre des principes FAIR (Findable, Accessible, Interoperable, Reusable). Le DMP contient les éléments qui guident le processus de gestion des données tout au long de leur cycle de vie et fournissent également des traces de la provenance des données (inclus tous les résultats de recherche dérivés des données brutes). L'Inst-CNRS propose l'outil Web DMP-OPIDoR qui facilite la rédaction des DMP et contribue également à l'harmonisation des bonnes pratiques en fournissant conseils, exemples et modèles spécifiques de certaines institutions. L'évolution du DMP OPIDoR vers un DMP automatisable (« machine-actionable DMP ») est actuellement à l'étude sur la base des retours d'expérience des utilisateurs, des cas collectés et des travaux de RDA sur le DMP actif. L'outil DMP OPIDoR « machine actionable » sera conforme au modèle commun RDA DMP mais inclura également une extension afin de servir les différents acteurs et besoins disciplinaires validés par un large groupe d'utilisateurs.

Recherche.data.gouv

Il est à présent possible de relier automatiquement ses jeux de données déposés dans Recherche Data Gouv à son projet ANR. Comment ? Suivez le guide !



Au moment du dépôt de données dans l'entrepôt Recherche Data Gouv, des métadonnées de description doivent être saisies pour contextualiser au mieux le jeu de données : un titre, un résumé, des auteurs, mais aussi le type de données ou encore la date de collecte...

Toutes ces informations sont indispensables pour comprendre à quoi correspond le jeu de données : qui l'a produit ? À quelle période ? De quoi est-il constitué ? Peut-il être réutilisé ? Comment s'en servir ?

Dans le cadre de données issues de projets de recherche financés par subvention, il est aussi important de renseigner l'agence de financement ainsi que l'identifiant du projet. Cette étape permet de lier ses données à ce projet, ce qui permet une traçabilité des produits de recherche créés pendant le projet. Une étape identique est aussi nécessaire pour le dépôt des publications scientifiques dans HAL.

Depuis janvier 2024, cette étape est grandement facilitée grâce à un remplissage automatique des métadonnées de subvention, de titre et d'acronyme du projet, dès lors que l'utilisateur renseigne l'identifiant de son projet ANR.

Une fonctionnalité qui sans nul doute, fera gagner un temps précieux aux déposants et permettra de suivre plus précisément les jeux de données issus de projets financés par l'ANR.



RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité

recherche.data.gouv.fr

Premier point de contact : Les ateliers de la donnée

*Expertise généraliste en proximité des équipes de
recherche pour toute question relative à la donnée*

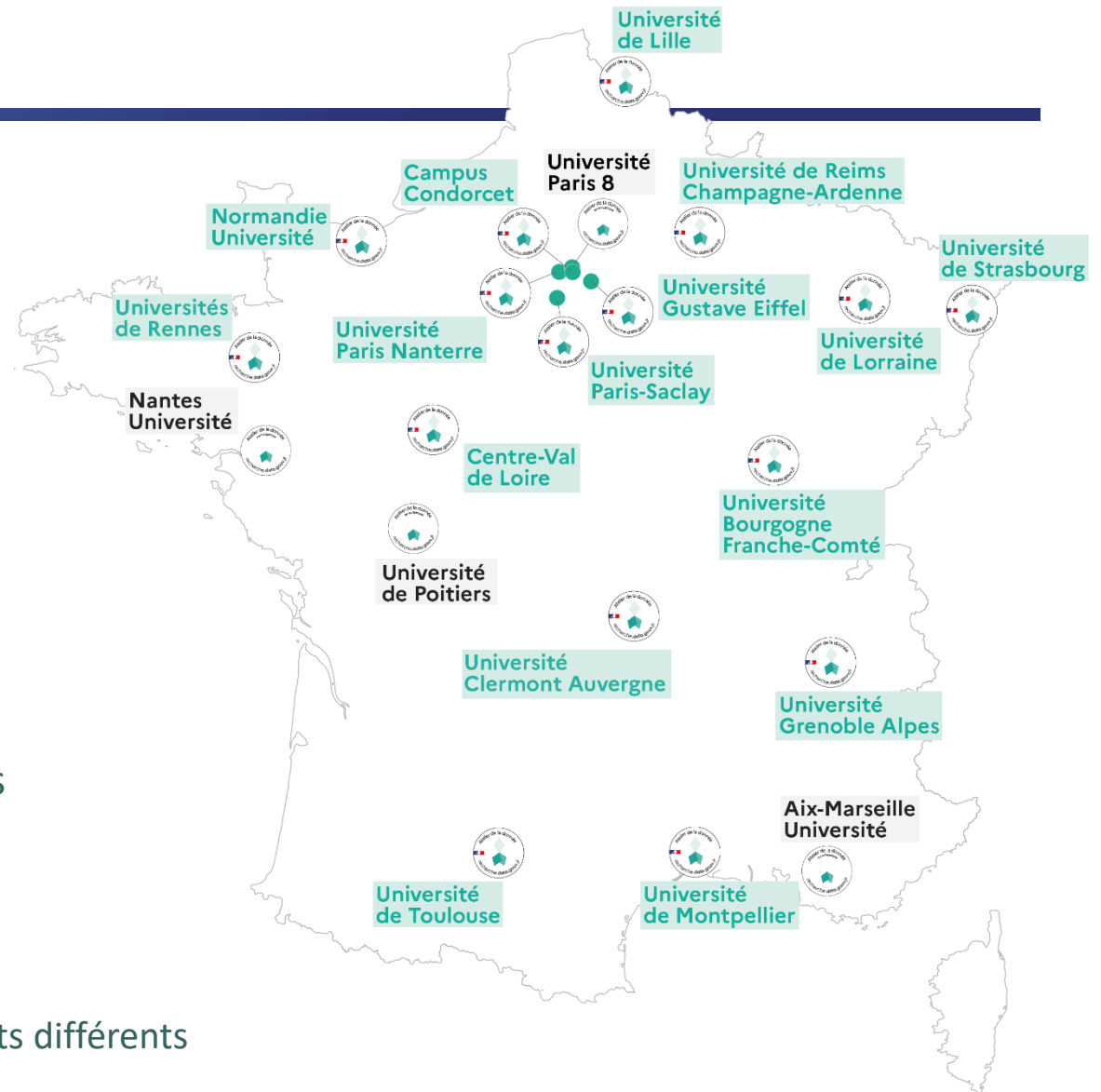
Regroupement d'expertises complémentaires de différents
établissements

Un réseau d'entraide, de mutualisation et de capitalisation des
services actifs

20 ateliers de la donnée opérationnels

95+ ETP
370+ personnes
80+ établissements différents

Plusieurs ateliers en projet



Entrepôts de confiance

Sur cette page, Recherche Data Gouv se fait le relai des entrepôts thématiques de confiance proposés par le Collège des Données de la recherche du Comité pour la science ouverte.

Le partage des données issues des activités de recherche implique de les mettre à disposition dans des entrepôts reconnus par les communautés scientifiques, thématiques, institutionnels ou généralistes.



Retour en haut

Pour faciliter le travail des équipes de recherche, le collège des Données de la recherche propose une méthode d'identification des entrepôts thématiques de confiance, ainsi qu'une **première liste d'entrepôts découlant de la grille de critères d'analyse retenus à découvrir ci-dessous**.

Cette liste est évolutive et a vocation à être progressivement complétée par le collège des Données de la Recherche. **Elle n'est par ailleurs pas exhaustive** : ainsi, ne pas figurer dans la liste n'implique pas que l'entrepôt concerné n'est pas de confiance.

⚠ N'oubliez pas, la dernière version à jour de la liste de référence est celle publiée sur le [site Ouvrir la Science](#).

📧 Pour toute question relative aux entrepôts de confiance, contactez le [Collège des Données de la recherche du Comité pour la science ouverte](#).

Archéologie	+
Astronomie	+
Biologie	+
Chimie	+
Physique	+
Sciences de l'environnement	+
Sciences humaines et sociales	+

Les entrepôts de stockage mentionnés dans les PGD – premiers éléments d'analyse

- **Appel à projets générique édition 2019** : 1001 PGD analysés dont 804 avec des questions du modèle Science Europe et 754 avec des éléments aux questions 9 et/ou 10.
- **Appel à projets générique édition 2023** : 604 PGD analysés dont 515 avec des questions du modèle Science Europe dont 276 avec des éléments de réponse aux questions 9 et/ou 10

Modèle Science Europe

Question 0 : "Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?"

Question 1 : "Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?"

Question 2 : "Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?"

Question 3 : "Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?"

Question 4 : "Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?"

Question 5 : "Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?"

Question 6 : "Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?"

Question 7 : "Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?"

Question 8 : "Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?"

Question 9 : "Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?"

Question 10 : "Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?"

Question 11 : "Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?"

Question 12 : "Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?"

Question 13 : "Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?"

Question 14 : "Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?"

Comparaison des tendances entre AAPG 2019 et 2023

Pour les données :

AAPG2019 → AAPG2023

- | | |
|---------------|------------------------|
| 1) Zenodo | 1) RDG ↑ |
| 2) Hal | 2) Zenodo ↓ |
| 3) Github | 3) Hal ↓ |
| 4) PDB | 4) Nakala ↑ |
| 5) Dryad | 5) Github ↓ |
| 6) NCBI geo | 6) PDB ↓ |
| 7) EMBL pride | 7) Genebank ↑ |
| 8) RDG | 8) EMBL pride ↓ |
| 9) Nakala | 9) Software Heritage ↑ |
| 10) EMBL ena | 10) NCBI geo ↓ |

Pour les publications :

AAPG2019 → AAPG2023

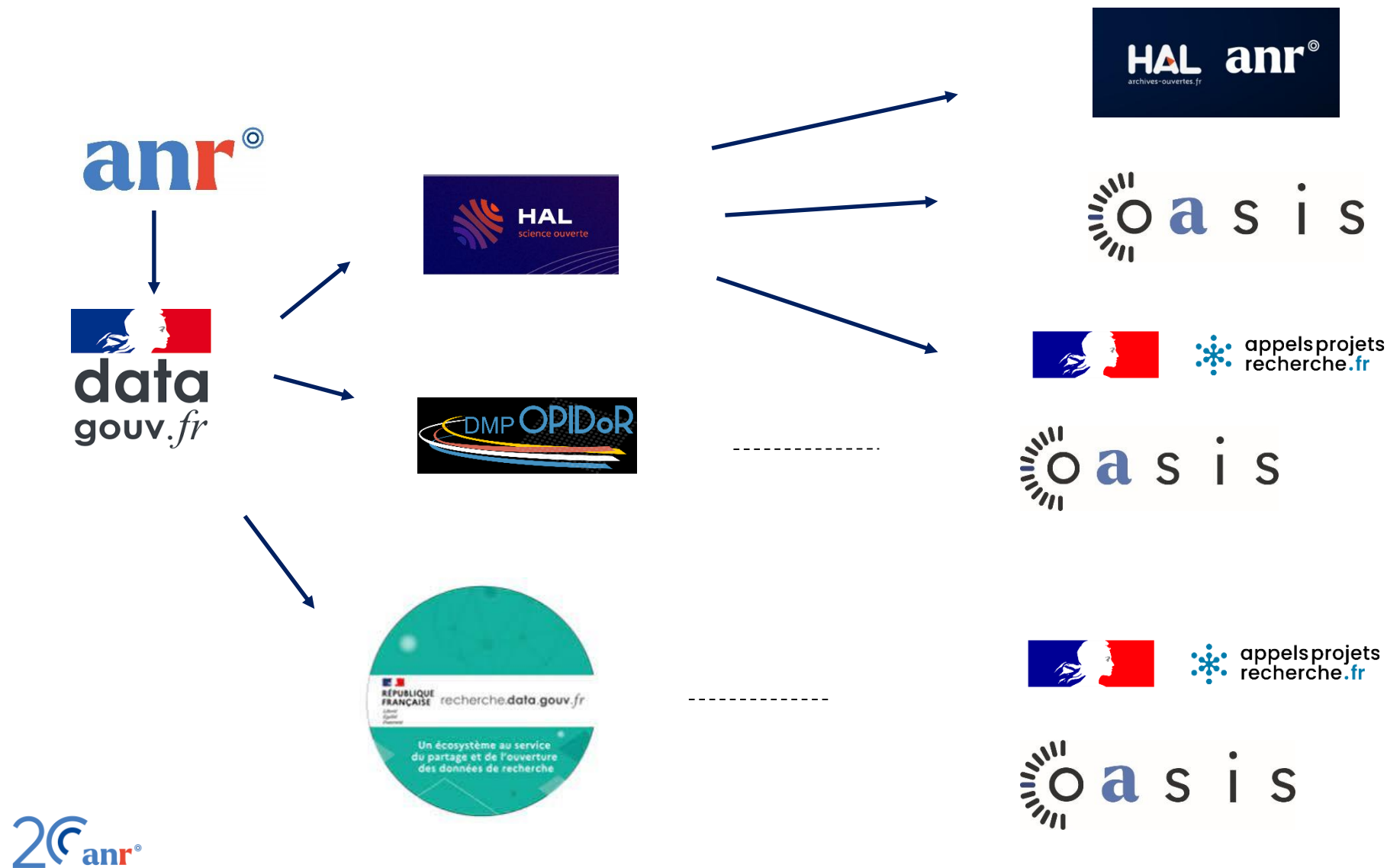
- | | |
|--------|--------|
| 1) Hal | 1) Hal |
|--------|--------|

Pour les logiciels :

AAPG2019 → AAPG2023

- | | |
|----------------------|------------------------|
| 1) Github | 1) Software Heritage ↑ |
| 2) Zenodo | 2) Github ↓ |
| 3) Hal | 3) Hal |
| 4) Software Heritage | 4) Zenodo ↓ |
| 5) Dryad | |

Dites-le nous une fois!



Pour la suite

- ✓ Poursuivre l'adoption du DMP structuré comme un levier de simplification pour tous les acteurs (chercheurs, organismes et établissements, agences de financement)
- ✓ Prise en compte des aspects financiers (dont coût de stockage)
- ✓ Promouvoir la reconnaissance de l'ensemble des produits de la recherche y compris les jeux de données, les codes sources et les logiciels



BARCELONA DECLARATION ON OPEN RESEARCH INFORMATION

Because equitable
decision-making
requires inclusive
data

We
signed the
declaration!



**Plus d'information sur :
Site de l'ANR**

<https://anr.fr/>

Abonnement à la Newsletter mensuelle

<https://anr.fr/fr/newsletter/>

Portail national commun des appels à projets

<https://www.appelsprojetsrecherche.fr/>