La gestion des données à l'EFA

- Pourquoi verser mes données ?
  - Un constat de départ
  - Enjeux techniques
  - Enjeux éthiques et légaux
  - Enjeux d'exploitation des données scientifiques
  - Enjeux de valorisation de mes données
- Le parcours de la donnée à l'EFA
  - Avant la mission
  - Pendant la mission
  - Après la mission
- Quels référentiels utilise-t-on à l'EFA ?
  - Qu'est-ce qu'un référentiel ?
  - Référentiels extérieurs communs à l'enseignement supérieur et la recherche
  - Référentiels extérieurs communs à une discipline
  - Référentiels internes à l'EFA
  - Quels usages pour ces référentiels ?
- Que deviennent mes données quand je les verse à l'EFA ?
  - Du versement au signalement
  - Quelle licence pour les données ?
- Comment valoriser mes données produites à l'EFA ?

    - Les données versées dans Archimage
    - Les données déposées dans un entrepôt externe
    - Comment valoriser les données à l'EFA

- Comment ouvrir mes données ?
    - Les prérequis
    - Ouverture à l'initiative du producteur
    - Ouverture à l'initiative de l'EFA
    - Accéder au formulaire

- Quel est le rôle d'un référent archives ?

Pourquoi verser mes données ?

Un constat de départ

- 90% des chercheurs disent « effectuer de manière individuelle le stockage, l'archivage ou la transmission de leurs données. » (European Commission, Directorate-General for Research and Innovation, O'Carroll, C., Hyllseth, B., Berg, R., et al., *Providing researchers with the skills and competencies they need to practise Open Science*, Publications Office, 2017, https://data.europa.eu/doi/10.2777/121253. Cité par DATACC, "Gestion des données : une nouvelle exigence, de nouvelles compétences », disponible sur le Web : https://www.datacc.org/bonnes-pratiques/adopter-un-plan-de-gestion-des-donnees/gestion-des-donnees-une-nouvelle-exigence-de-nouvelles-competences/#note. Consulté le 26/04/2023.)
- En 2016, on estimait que 80 % des données produites durant les 20 dernières années étaient « perdues faute de politique de sauvegarde coordonnée », tous domaines de recherche confondus. (Exposé des motifs de la loi pour une République numérique, consulté sur Legifrance. Cité par DATACC, « Gestion des données : une nouvelle exigence, de nouvelles compétences », disponible sur le Web : https://www.datacc.org/bonnes-pratiques/adopter-un-plan-de-gestion-des-donnees/gestion-des-donnees-une-nouvelle-exigence-de-nouvelles-competences/#note. Consulté le 26/04/2023.)

Le dépôt des données répond à des enjeux éthiques et légaux, techniques, mais aussi d'exploitation et de valorisation de la production scientifique.

Enjeux techniques

- L’étude et la publication des données se fait sur le temps long alors que les supports de stockage (disques durs et clés USB) sont fragiles et soumis aux contingences matérielles. Si les données d’une mission de 2021 sont stockées sur un disque dur mais non déposées au service des archives, et que celui-ci est remporté sur le terrain l’année suivante et égaré, volé ou endommagé, l’ensemble des données sera perdu.
- Certains supports de stockage sont concernés par une obsolescence rapide. Qui plus est l’érosion de ces supports est invisible : ils peuvent ne plus fonctionner du jour au lendemain. Les CD-Rom versés il y a quelques années sont d’ores et déjà devenus difficiles à lire.
- Les logiciels et les formats deviennent rapidement obsolètes. Les fichiers déposés aux service des archives sont régulièrement migrés afin qu’il soit toujours possible de les ouvrir.
- Les serveurs des archives de l’EFA sont répliqués et sauvegardés. Ils sont donc mieux protégés qu’un ordinateur individuel.

Enjeux éthiques et légaux

- Les données produites dans le cadre des missions de l'EFA sont des données publiques même si le chercheur en est l'auteur. Pour en savoir plus, voir la fiche : *Le statut des données*.
- Elles s'inscrivent donc dans le cadre de gestion des archives publiques. La gestion, et parfois le dépôt des données, est une exigence des financeurs. Pour en savoir plus, voir la fiche : *Les aspects législatifs*.
- Verser mes données permet de me prémunir contre le plagiat : le dépôt authentifie le producteur et la date de dépôt. La base Archimage indique ainsi de manière explicite la date de découverte, l'auteur, le statut (inédit ou publié) de chaque donnée.
- Le versement des données n'est pas synonyme de diffusion immédiate. Pour en savoir plus, voir la fiche : *Que deviennent mes données déposées à l'EFA ?*

Enjeux d'exploitation des données scientifiques

- Le dépôt des données me permet de retrouver et de retravailler mes données, alors que la masse de données est de plus en plus importante.
- Les missions nécessitent de pouvoir exploiter les données précédemment produites. Une étude est impossible si les données précédemment versées au retour d'une mission de terrain n'ont pas été versées et ne sont pas correctement construites et décrites.
- Le dépôt des données permet d'exploiter dans un nouveau contexte de recherche des outils précédemment produits, une base de données par exemple, permettant ainsi à la recherche de gagner en efficacité.

→ Pour être utilisables, les données déposées doivent donc être documentées. Pour en savoir plus, voir les fiches : *Qu'est-ce qu'une donnée de qualité ?* ; *Quels référentiels utilise-t-on à l'EFA ?* ; *Quelles métadonnées utilise-t-on à l'EFA ?*

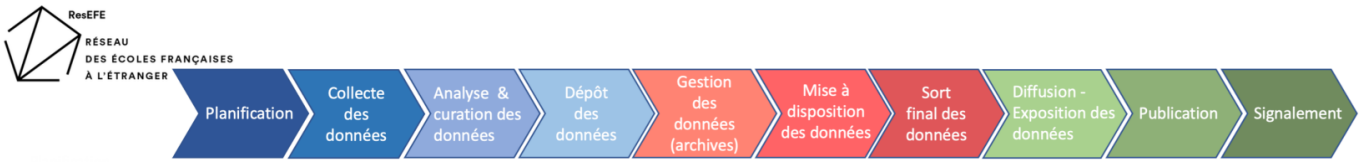
Enjeux de valorisation de mes données

- Déposer mes données permet de les rendre visibles, ou du moins leur existence visible : le dépôt des données constitue un second point d’accès vers mes travaux en sus de mes publications, grâce à leur signalement sur différentes plateformes.
- Déposer mes données me permet d’être facilement identifié comme leur inventeur : mes travaux pourront être facilement cités. Il est aussi plus facile de citer d’autres travaux. Ce peut être aussi un moyen de susciter de nouvelles collaborations scientifiques. Pour en savoir plus : Données et identité numérique du chercheur.
- Mes données déposées peuvent être utilisées pour accompagner des publications, comme dans les récents articles du BAEFE. Pour en savoir plus : Comment valoriser mes données produites à l’EFA ?

Le parcours de la donnée à l'EFA

Ces pages détaillent dans l'ordre les étapes de la vie d'une donnée dans le cadre d'une mission de l'EFA. Les acteurs sont pour chacune d'entre elles spécifiés : les tâches relevant du responsable de la mission sont surlignées en vert et celles du référent archives en bleu.

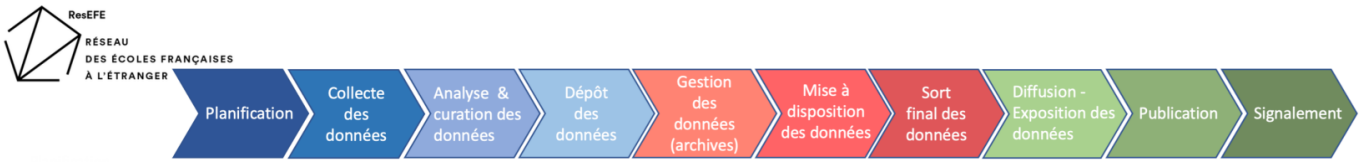
Avant la mission



Cette phase concerne l'étape de planification, antérieure au début de la mission.

1. **Le chef de mission** désigne un référent archives au sein de l'équipe et transmet ses coordonnées au service des archives
2. **Le référent archives** prend contact avec la responsable du service des archives (marie.stahl@efa.gr) & la responsable de la photothèque/planothèque (calliopi.christophi@efa.gr)
3. **La responsable du service des archives** donne les "bonnes pratiques" en vue du versement global
4. **La responsable photothèque/planothèque** donne les "bonnes pratiques" en vue du versement des photographies & communique un numéro d'enregistrement photothèque : un pour les photographies de fouille, et un autre pour chaque type de données composites comportant des photos (photogrammétrie, RTI, etc.)

Pendant la mission



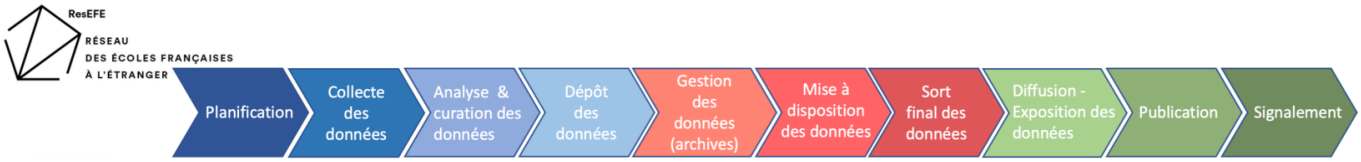
Cette phase concerne les étapes allant de la collecte à l'analyse et la curation des données.

1. **Le référent archives** contrôle la conformité et la sauvegarde des documents papier et numérique & contrôle la conformité et la sauvegarde des données liées aux documents
2. **Les membres de la mission** scannent les documents produits, si les conditions le permettent

Vers la fin de la mission :

1. **Le chef de mission** sélectionne les documents à destination de l'Éphorie
2. **Le référent archives** récupère les documents (papier et numérique) et les données liées et prépare le versement

Après la mission



Cette phase concerne toutes les étapes à partir du dépôt au service des archives jusqu'à leur sort final.

1. **Le chef de mission** verse au plus tard le 31 octobre les documents papier et dépose les documents numériques soit sur place, au retour de la mission, soit via le carnet numérique de la mission ou celui du service des archives
2. **La responsable du service des archives** récupère les documents et données destinés aux archives manuscrites & vérifie la conformité du versement
3. **La responsable photothèque/planothèque** récupère les documents et données destinés à la photothèque & vérifie la conformité du versement
4. **La responsable du service des archives** transmet à la direction les documents destinés à l'Éphorie
5. **La direction** envoie les documents destinés à l'Éphorie
6. **La responsable du service des archives** rédige le bordereau de versement et le transmet au chef de mission avec copie au référent archives
7. **Le chef de mission** signe le bordereau et le transmet scanné au responsable du service des archives (marie.stahl@efa.gr)
8. **Les chargés de numérisation** numérisent les documents papier
9. **La responsable du service des archives** intègre les documents et données liées dans les bases archives manuscrites
10. **La responsable photothèque/planothèque** intègre les documents et données liées dans la base photothèque
11. **La responsable du service des archives** transmet les documents scannés au chef de mission

> Attention, à la fin du programme pluri-annuel, le responsable de mission doit prévoir un dernier versement des documents qui n'auraient pas pu être versés au cours des années passées

> Pour vous renseigner sur la site du chemin de la donnée une fois le versement effectué par les membres de la mission, voir la fiche pratique *Que deviennent mes données quand je les*

verse à l'EFA

Quels référentiels utilise-t-on à l'EFA ?

Qu'est-ce qu'un référentiel ?

Un référentiel est un répertoire d’autorités permettant d’identifier sans ambiguïté une entité nommée (personne, objet, etc.), géographique ou chronologique. Cela passe par l’attribution d’un identifiant pérenne à l’entité (de type DOI, ARK, ou autre) ainsi selon les cas qu’une description, localisation, étendue.

Quelques liens :

- https://doranum.fr/identifiants-perennes-pid/ La fiche synthétique et les ressources de DoRANum sur la question.
- https://isidore.science/vocabularies Le portail développé par la TGIR HumaNum, Isidore, recense et décrit les référentiels qu’il interroge pour apporter des données de bonne qualité à ses utilisateurs.

Référentiels extérieurs communs à l'enseignement supérieur et la recherche

Servent à toutes les disciplines et leurs chercheurs afin d'identifier sans ambiguïté une personne ou sa production. Ils attribuent un PID (identifiant pérenne) aux entités et fournissent un lien permettant de s'y référer.

- IdRef : mis à disposition par l'Abes (pour Agence bibliographiques de l'enseignement supérieur, qui a une mission de signalement notamment avec le catalogue SUDOC et de gestion des métadonnées), fournit une page d'autorité pour toute entité. On y trouve par exemple les personnes ayant une activité de recherche et des publications dans l'enseignement supérieur et la recherche, avec des ressources extérieures vers le catalogue général de la BnF, VIAF ou ISNI. À L'EFA, les secteurs sont aussi alignés avec IdRef, ce qui permet de fournir des traductions et synonymes d'un nom. Les alignements avec d'autres référentiels permettent d'enrichir les informations et de créer un réseau de données et métadonnées pérennes très utilisé pour faire des croisements.
- Orcid : organisation internationale permettant d'identifier un auteur et ses publications. La page d'un chercheur donne des renseignements sur ses activités et un moyen unique de s'y référer.
- HAL : archive ouverte du CNRS pour le dépôt des publications dans leurs versions antérieures à l'édition, ou tout autre support de communication. Grâce à son référentiel lié, elle permet également d'identifier un auteur, auquel est attribué un **IDHAL**. Elle a des déclinaisons pour les SHS, les thèses ou les sciences médicales par exemple, et une institution peut créer et administrer sa propre instance d'archive ouverte.
- Persée : plateforme de numérisation et publication des revues et monographies de l'enseignement supérieur, met à disposition une page d'autorité pour les auteurs et leurs contributions recensées sur le site, grâce à son référentiel **Data Persée**.
- ROR : Ce *Research organisation registry* propose d'identifier de manière pérenne et sans ambiguïté des institutions de recherche et d'enseignement supérieur avec des informations hiérarchisées (départements et laboratoires) et géolocalisation. À l'EFA, il est notamment employé par la *Chronique des fouilles en ligne*.

Référentiels extérieurs communs à une discipline

Spécialisés en archéologie et/ou sciences humaines et sociales, ces référentiels sont disponibles librement.

- Pactols : thésaurus développé par la fédération de recherche autour des sciences de l'antiquité (Frantiq) pour les sujets et lieux étudiés en archéologie et sciences humaines et sociales. Multilingue et multi-hiérarchique, le thésaurus attribue à chaque concept un URI et crée des relations entre concepts associés.
- Pleiades : référentiel géographique spécialisé dans les lieux anciens, permettant de les identifier et de s'y référer en fonction des nomenclatures en vigueur. Il localise les sites présentant un intérêt archéologique avec leurs noms anciens mais aussi dans certains cas leur nom moderne.
- GeoNames : référentiel géographique pour les lieux modernes et contemporains.
- PeriodO : référentiel chronologique, dont un pan a été construit par l'EFA, permettant d'identifier et de se référer à une période historique pour l'archéologie.
- Getty's Art and Architecture Thesaurus : employé par la *Chronique des fouilles en ligne* pour ses mots-clefs, le vocabulaire d'autorité proposé par la fondation Getty rend accessible des listes de concepts définis et hiérarchisés pour ce qui relève de l'histoire de l'art et de l'architecture.

Référentiels internes à l'EFA

Développés en interne à l’EFA, ayant principalement pour sujet les actions et acteurs de l’École, servent également de référentiels.

- Liste d’autorité des auteurs : disponibles dans les principales bases de l’EFA (Archimage, publications), ces autorités recensent les auteurs ayant travaillé en lien avec l’EFA et permet d’identifier leurs productions dans ce cadre. Elles sont alignées notamment avec IdRef, qui fournit sa liste de référence au catalogue de la bibliothèque.
- Liste d’autorité des secteurs : disponibles dans Archimage et formant la base des webSIG, les secteurs sont arrêtés par les chercheurs en charge du site de fouilles et identifiés pour pouvoir s’y référer.
- Liste d'autorité des programmes : en cours de réalisation, elle devra permettre d'identifier un programme au sein de la base Missions, tout en étant agrégée aux autres bases de données de l'EFA comme Archimage.

On peut également se référer aux bases de données existantes de l'EFA, qui exploitent ces listes d'autorité : voir les outils.

Quels usages pour ces référentiels ?

Les référentiels catalogués ici peuvent être employés par les chercheurs de l’EFA dans le cadre de programmes de recherche pour correctement identifier et référencer les personnes, les lieux, les sujets et concepts sur lesquels ils effectuent leurs travaux, à l’aide de leur identifiant pérenne. Les problèmes d’homonymie ou de translittération sont par là-même écartés.

Dans le cas d’une publication par exemple, c’est l'intitulé et l’identifiant qui peuvent être donnés en note. Dans le cas plus avancé d’une base de données, des liens peuvent être établis pour mettre en relation les différents éléments entre eux (web sémantique). Cette interopérabilité permet l’accès, la réutilisation des données (pour reproduction des travaux par exemple) ainsi que leur citation. L'EFA recommande fortement ces pratiques en faveur de la science ouverte pour ses chercheurs et associés.

Que deviennent mes données quand je les verse à l'EFA ?

Du versement au signalement

Gestion de la donnée

- À partir des métadonnées transmises, chaque document versé donne lieu à la **création d'une notice** dans la base de gestion des archives, Archimage.
- Chaque notice sera accompagnée du **document numérique** en relation.
- Dans le cas de données composites (photogrammétrie, 3D, etc.), le **dossier des fichiers ayant généré le rendu** sera lui aussi lié à la notice du rendu final.

Mise à disposition

- Toutes les **notices descriptives** sont affichées en mode public.
- Les document numériques sont affichés ou non en fonction de la législation en vigueur.
- **Les dossiers de fichiers** ne sont pas disponibles en ligne.

Les données versées au service des archives de l'EFA entrent dans le cadre des **archives publiques** ; ces fichiers constituent des documents administratifs dont certains sont soumis à droit d'auteur. Pour plus d'informations sur leurs conditions d'accès et d'Exploitation, voir les fiches pratiques Législation sur les archives, *Consulter*, *Reproduire* et *(Ré)utiliser des archives*.

Signalement

Les **notices descriptives** seront ultérieurement moissonnées et figureront dans le portail national France Archives.

> Si vous êtes à l'origine de données versées à l'EFA et que vous souhaitez assurer leur visibilité, voir la fiche pratique *Comment ouvrir mes données ?*

Quelle licence pour les données ?

Les données produites dans le cadre de missions de recherche de l'EFA et ensuite versées au service des archives sont déposées dans la plateforme Archimage. Toujours soumises à droit d'auteur, elles sont mises en ligne en collaboration avec producteurs dans le respect de la législation en vigueur et selon une procédure établie par l'établissement (cf. Fiche Comment ouvrir mes données ?).

Pour la plateforme Archimage, la licence qui a été choisie est la suivante : Creative Commons Attribution-NonCommercial-ShareAlike 4.0. Les termes de cette licence spécifient que toute personne a le droit de réutiliser la donnée à condition de :

- nommer l'auteur original,
- ne pas en faire une utilisation commerciale,
- la partager sous les mêmes conditions.
- L'adaptation et la modification restent possibles.

Comment valoriser mes données produites à l'EFA ?

Les données versées dans Archimage

À l'EFA, le versement des documents par le chercheur au service des archives, avec un bordereau, implique un transfert de responsabilités : c'est l'établissement qui s'engage à les conserver et documenter selon les normes en vigueur. Concrètement, après avoir traversé les étapes de description, les données sont versées avec leur visuel et métadonnées associées dans l'entrepôt institutionnel (et donc développé et maintenu en interne) Archimage pour leur sauvegarde, en ce qu'elles représentent les travaux fournis par l'École. Le chercheur garde les droits moraux et patrimoniaux sur ses documents jusqu'à une éventuelle cession non-exclusive de droits pour partage avec la communauté scientifique. En effet, la mise à disposition se fait par le biais du service des archives, en accord avec les règles de diffusion et de communication ainsi que les principes de la science ouverte.

La plateforme Archimage n'est pas uniquement un lieu de gestion des données à visée archivistique, mais peut aussi être un espace d'exposition des documents versés à la photothèque, à la planothèque et aux archives manuscrites. Peuvent être recueillis tous types de données, allant du texte simple à la base de données plus complexe en passant par des analyses photogrammétriques (voir à ce sujet les fiches pratiques par typologie de document). En ce sens, elle permet une valorisation du patrimoine de la recherche à l'EFA qui ne peut être maintenue que par l'implication des chercheurs et membres associés dans la pratique de l'*open data*, soit l'accès et le partage libre pour tout le monde des données de la recherche, sous une forme normalisée et lisibles par des machines. Cette pratique se fait sous une certaine licence, de type Creative Commons, afin d'être en mesure de les transformer, combiner et partager, toujours en citant ses sources.

Les données déposées dans un entrepôt externe

À l'inverse du versement au service des archives par un bordereau, c'est le cas de figure du dépôt par le chercheur dans un entrepôt externe qui est présenté ici. Il en existe de plusieurs types et de différentes envergures, en voici quelques exemples :

- Recherche Data Gouv est un entrepôt interdisciplinaire et gouvernemental, mis à disposition par le ministère de l'enseignement supérieur et de la recherche pour ses communautés. Il propose une documentation fournie et notamment un outil d'aide à la rédaction de *data paper* pour les données qu'il accueille.
- Nakala est un entrepôt de données spécialisé dans les sciences humaines et sociales, mis à disposition par la TGIR (pour *Très grande infrastructure de recherche*) HumaNum. Il propose différents outils d'exposition notamment par le biais de son plug-in *Nakala-Press* qui permet de créer un site comme une vitrine pour les jeux de données accueillis.
- Zenodo, enfin, est une infrastructure à l'échelle européenne et interdisciplinaire qui propose un stockage sécurisé et pérenne pour des jeux de données académiques.

Ces entrepôts se caractérisent par le fait que le dépôts est effectué par le chercheur lui-même sur la plateforme, sans accompagnement de la part du service des archives. Il reste donc entièrement responsable de leur mise à disposition, par le choix ou non de l'application d'une barrière de consultation ou d'un embargo. Les entrepôts externes, de leur côté, se limitent à offrir une solution d'hébergement pérenne et sécurisé aux données, avec l'application d'un identifiant stable de type DOI, Handle ou ARK.

Les identifiants pérennes et référencements/liens établis entre les différents corpus sert de base au travail d'analyse des données, pour une exposition par le biais d'une publication ou d'une actualité sur un site internet, d'une base de données publiée ou d'un *data paper*. Une fois les résultats publiés, il est recommandé d'établir un lien entre la publication et les jeux de données entreposés grâce à l'identifiant pérenne, afin d'assurer une meilleure reproductibilité et capacité de citation.

> Attention cependant, il faut signaler que le dépôt dans un entrepôt externe n'affranchit pas le chercheur de verser ses archives à l'EFA

Comment valoriser les données à l'EFA

À l’EFA, un certain nombre d’outils sont à disposition afin de permettre aux chercheurs de réutiliser leurs données dans des travaux de recherche. Certains sont proposés en interne, d’autres sont développés en lien avec des partenaires. Ces outils sont là pour améliorer la connaissance que la communauté scientifique et le public ont des travaux de l'EFA mais aussi afin que les chercheurs puissent se réapproprier les données versées ou déposées en les appelant depuis leur entrepôt de stockage.

- Pour documenter la vie d’un projet de recherche, OpenEdition propose l’ouverture de carnets Hypotheses pouvant recevoir des billets d’actualité, des réflexions en cours, etc. Il s’agit d’un blog que les chercheurs peuvent alimenter en nouvelles et ressources en lien avec leur activité.
- Par le biais d’outils numériques, il est possible de concevoir un site plus complet pouvant accueillir par exemple une bibliothèque numérique (grâce aux CMS WordPress, Drupal ou Omeka) : ces sites-vitrines sont l'occasion de diffuser et d'exposer les données créées lors des travaux de recherche et de partager les premiers résultats ou analyses préliminaires.
- Les publications (Voir le catalogue des éditions de l’EFA) qui accueillent des travaux de recherche dans ses collections et revues. Ces publications se font la plupart du temps en *open access* sur les plateformes d’OpenEdition et de Persée. On peut y retrouver des données également versées dans Archimage et mises en interopérabilité par le biais d'un lien.
- La rédaction de *data papers* est encouragée par le ministère de l’enseignement supérieur et de la recherche : voir à ce sujet la fiche pratique *Qu’est-ce qu’un data paper ?*
- Enfin, le signalement des données stockées dans Archimage se fait en interopérabilité avec de grandes plateformes nationales et européennes de référencement pour pouvoir trouver et exploiter les documents mis à disposition : on peut citer parmi elles OpenAIRE, Ariadne plus, Calames ou France Archives. L'idée est de permettre une plus grande visibilité aux collections de l'EFA et de permettre, par les moyens numériques, une interopérabilité et une ré-exploitation optimisées.

On peut se référer à ce sujet au schéma explicitant le cycle de vie de la donnée à l’EFA, mettant en lumière les étapes entre le dépôt d’une donnée et la mise en valeur des résultats acquis.

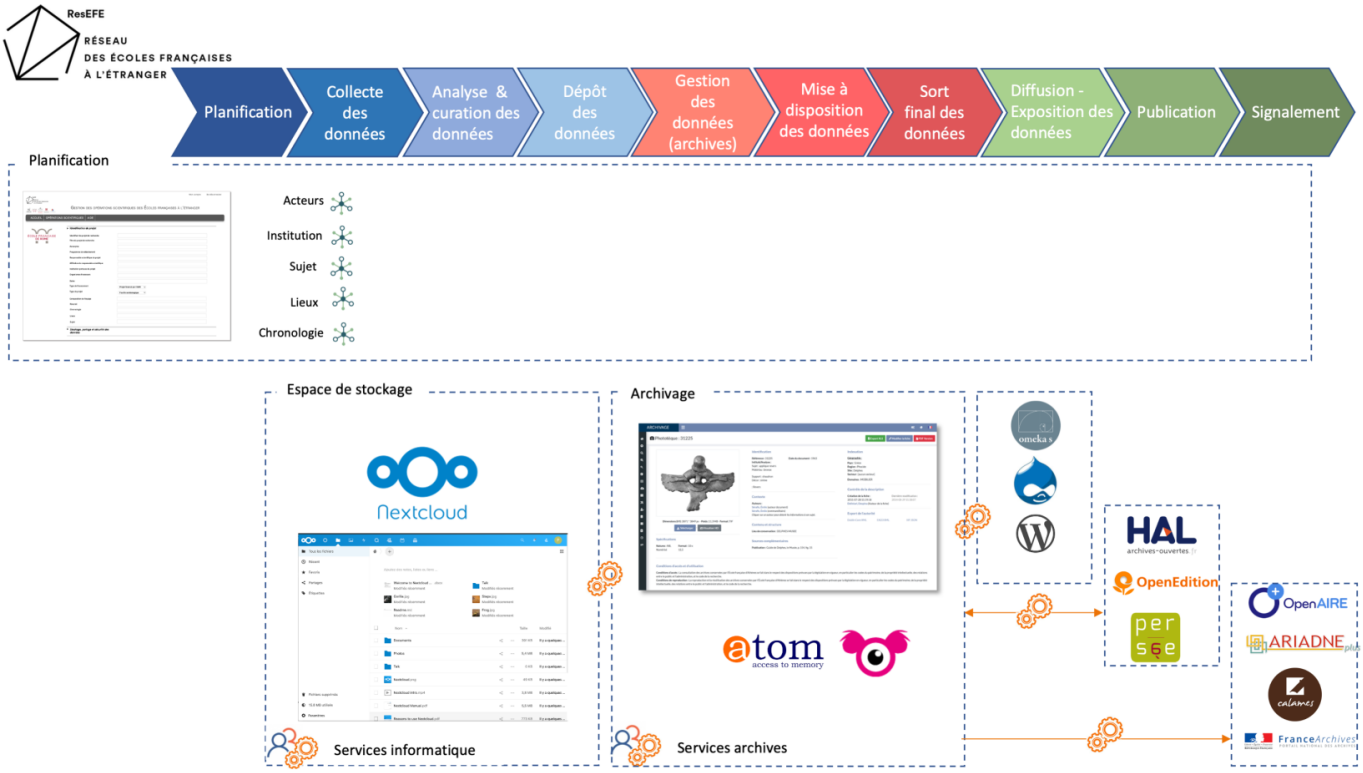


Figure : le cycle de vie de la donnée à l'EFA (crédit Bruno Morandière, ResEFE)

Ici sont exposés tous les outils et les acteurs permettant de mettre en avant la recherche effectuée ou en cours à l'EFA. On y voit bien que le versement ou dépôt sur un service d'entrepôt de données n'est pas une fin en soi, mais une étape vers la valorisation du travail des chercheurs en lien avec la communauté scientifique et les infrastructures de recherche partenaires. L'exposition des données est ce qui permet de préparer leur publication, et le signalement l'amorce d'une réutilisation potentielle dans le milieu académique ou de l'enseignement.

Ces possibilités de valorisation impliquent que les données aient été versées sous une licence libre, permettant leur partage dans les termes de l'*open data*. Il s'agit également d'être vigilant aux métadonnées associées à tout document, car cette documentation de la donnée est ce qui garantit sa qualité et son intégrité. Il convient donc d'y prêter attention dès le moment de la collecte des données, au début du projet (Voir à ce sujet la fiche pratique *Qu'est-ce qu'une donnée de qualité ?*).

Comment ouvrir mes données ?

Les prérequis

Cette fiche a pour objectif de montrer les méthodes et procédures en vue d'une ouverture de données à l'EFA, par le producteur ou par l'établissement directement. Avant toute chose, il y a un certain nombre de prérequis que les données doivent remplir pour pouvoir être ouvertes :

- **Les données doivent avoir été versées au service des archives de l'EFA**, et traitées de manière à déjà être présentes dans la base Archimage. Sont ainsi exclues les données réalisées en-dehors des missions portées par l'établissement et les données stockées dans un entrepôt externe.
- **Les données doivent répondre aux principes FAIR et aux normes de la science ouverte**. Cela implique notamment des données liées et métadonnées harmonisées et correctement lisibles par l'humain et/ou la machine. Pour en savoir plus sur ce point, consulter la fiche pratique *Qu'est-ce qu'une donnée de qualité ?*
- **Le producteur doit fournir la garantie qu'il possède bien les droits d'exploitation de ces documents** : cela signifie qu'il en est bien l'auteur ou le co-auteur, et que dans le cas de co-auctorialité les autres producteurs sont d'accord avec les principes de l'ouverture. Pour en savoir plus sur ces questions de droits, voir la page suivante.

Ouverture à l'initiative du producteur

Le producteur d'une donnée ou d'un ensemble de données peut prendre l'initiative de les ouvrir pour en faire bénéficier la communauté scientifique avec l'accompagnement de l'EFA.

Un point de vigilance à prendre en compte est que seul le titulaire des droits sur un jeu de données est habilité à l'ouvrir. Cela pose plusieurs questions du point de vue de la législation quant à savoir qui détient les droits d'exploitation : en effet, cela varie en fonction du statut du producteur.

- Un **chercheur ou enseignant-chercheur** conserve sur sa production scientifique l'ensemble des droits patrimoniaux et moraux sur les documents soumis à droits d'auteur. Les droits patrimoniaux sont transmissibles aux ayants droit et ne s'éteignent que 70 ans après la mort du producteur : ce sont eux qui peuvent être cédés par un contrat à un établissement public tel que l'EFA pour l'ouverture sur ses plateformes. Les droits moraux sont quant à eux incessibles et inaliénables dans le temps et assurent l'intégrité de l'œuvre. Dans ce cas, le chercheur devra formaliser son souhait par un contrat avec l'EFA.
- Un **agent public non enseignant-chercheur** cède automatiquement certains droits patrimoniaux ou d'exploitation de sa production à l'État. Dans ce cas, l'ouverture des données n'a pas besoin d'être contractualisée.
- Un **prestataire sous contrat** qui effectuerait par exemple une campagne de relevés voit les termes de la cession de droit décrits dans le contrat qui le lie à l'établissement.
- Un **personnel de droit local** voit ses droits sur sa production protégés au même titre que dans une entreprise privée.

Selon les cas de figure, le passage par un contrat de cession non-exclusive des droits patrimoniaux s'imposera ou non.

Pour procéder à une ouverture de données, le producteur prend contact avec le service des archives de l'EFA qui l'accompagne dans sa démarche jusqu'à l'exposition des données sur Archimage.

Ouverture à l'initiative de l'EFA

Seconde possibilité, l'EFA peut souhaiter procéder à une ouverture de lots de données conservées au service des archives pour répondre à une demande ou pour aller de pair avec une manifestation en cours. Deux cas de figure sont envisageables :

- **L'ouverture peut être effectuée par lots** pour un site de fouilles en particulier. Des lots cohérents sont alors déterminés en concertation avec la direction et les responsables des programmes de recherche, par secteur ou par type de données par exemple, en fonction de l'état des publications relatives au site.
- **L'ouverture peut être complète**, soit concerner l'ensemble des données d'un producteur. Cela se fait lorsque l'ensemble des données d'un producteur est jugé cohérent et pertinent pour l'ouverture, et que le producteur (ou ses ayants droit) accorde son autorisation.

La procédure dans ces deux cas reste la même : après identification du périmètre de l'ouverture des données, les chercheurs reçoivent une demande de cession de droits accompagnées d'un contrat et d'une annexe, constituée de la liste des métadonnées associées aux documents concernés (date, légende, etc.) en conservant la possibilité de signaler une donnée ou un ensemble de données à garder fermées sous embargo pour des questions d'études notamment. Si le producteur est décédé, c'est vers ses ayants droit qu'il convient de se tourner. Une fois le contrat signé et retourné, les données sont mises à disposition de la communauté scientifique sur les plateformes de l'EFA, notamment Archimage pour le service des archives.

Accéder au formulaire

L'ouverture est formalisée par un contrat de **cession non-exclusive de droits d'exploitation** entre l'auteur des documents et l'EFA. L'auteur garde ainsi l'**entière liberté d'exploiter ces documents dans un autre cadre**.

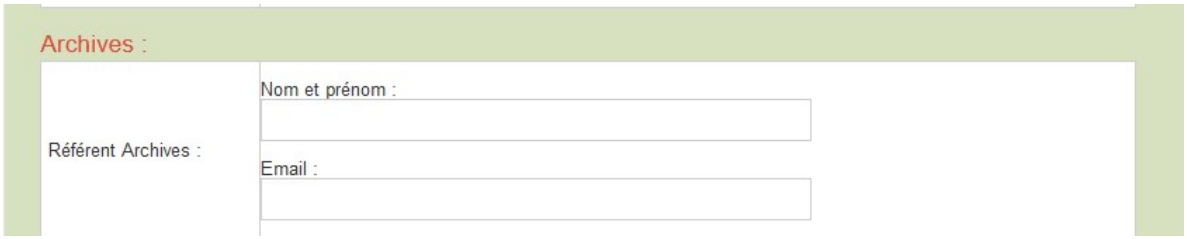
Vous pouvez télécharger un modèle du contrat de cession par un simple clic.

Vous pouvez consulter le tableur modèle listant les notices, qui accompagne en annexe le contrat.

Quel est le rôle d'un référent archives ?

Le référent archives est l'interlocuteur de l'EFA pour la gestion des documents et données du programme de recherche concerné, de leur planification à leur versement en fin de chaque année.

Il est désigné par le responsable du programme et son nom et adresse électronique sont renseignés dans la fiche mission de la plateforme Missions, au stade de la demande.



Le référent archives prend connaissance des fiches pratiques des étagères Gestion des données de la recherche et Service des archives.





Avant la mission, il effectue les démarches afin d'obtenir les cotes photothèque, si nécessaire.

Il guide et conseille les membres de l'équipe sur la bonne gestion des données et documents produits tout au long de la mission, en s'aidant notamment des Fiches par typologie de données et de la Boîte à outils

Il contrôle la conformité et la sauvegarde des documents papier et numérique et **des informations** relatives à leur identification (auteur, date de création, intitulé, etc.).

Il coordonne la préparation du versement des archives produites par l'équipe et **verse les documents et données** au service des archives.

> Pour connaître les rôles de chaque intervenant, consulter Le parcours de la donnée à l'EFA.