



ÉCOLOGIE &
ENVIRONNEMENT

Le cycle de vie des données de Recherche

Amandine Hénon
UAR DoHNEE



Un peu de théorie

Le cycle de vie des données...

Définition

Le cycle de vie des données de recherche décrit le processus d'utilisation des données de leur création à la publication et à leur réutilisation ultérieure (source : INIST)



Le cycle de vie des données...



Et après ?

Le cycle de vie des données...

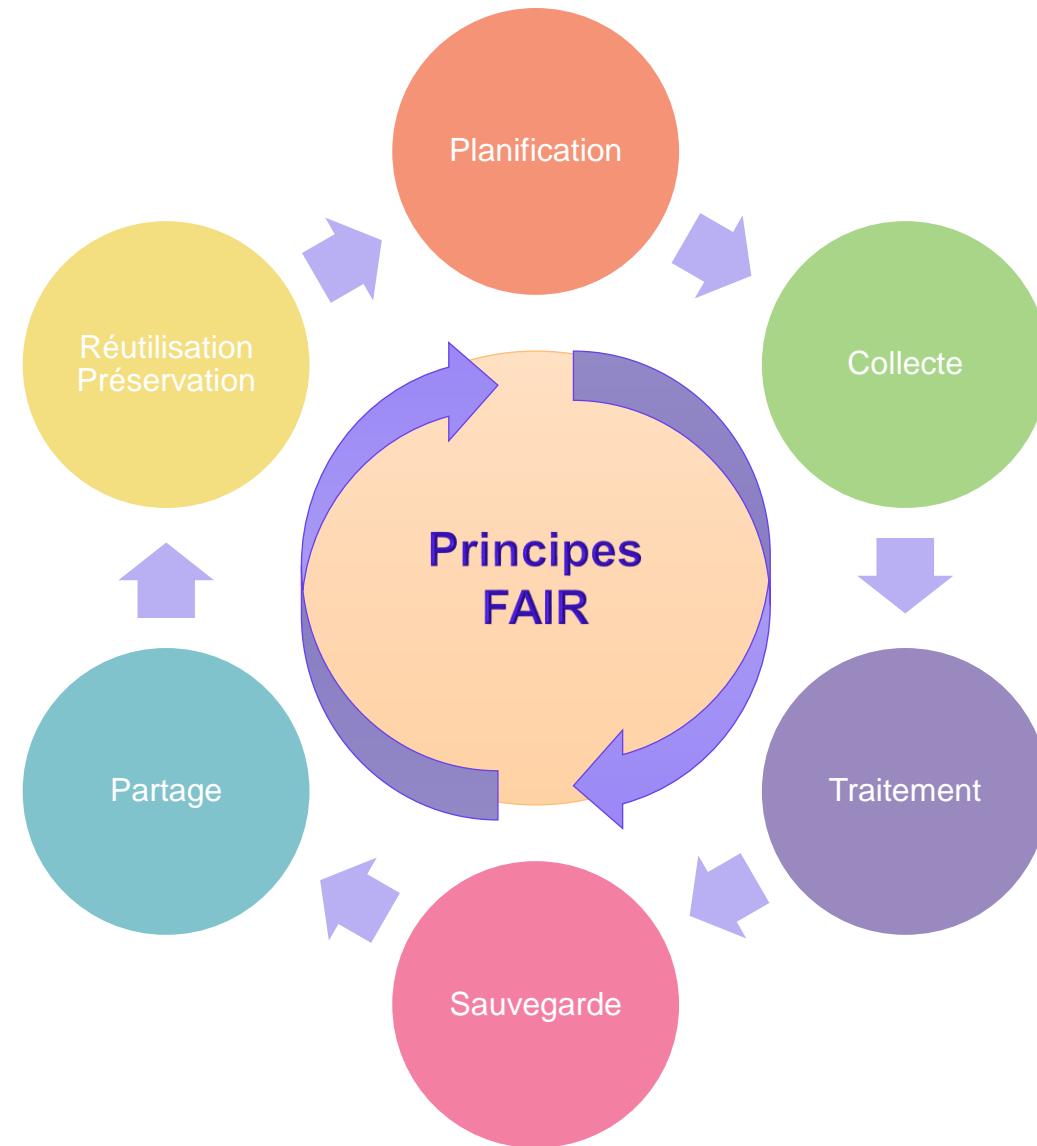
... Un élément structurant pour le Plan de Gestion de Données

- Réfléchir selon le cycle de vie permet d'anticiper les étapes clé et ses problématiques.
- Il permet aussi d'envisager les données sur le moyen et long terme.

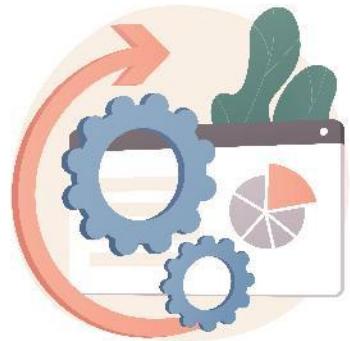
... Il s'agit des premiers pas pour le plan de gestion de données !



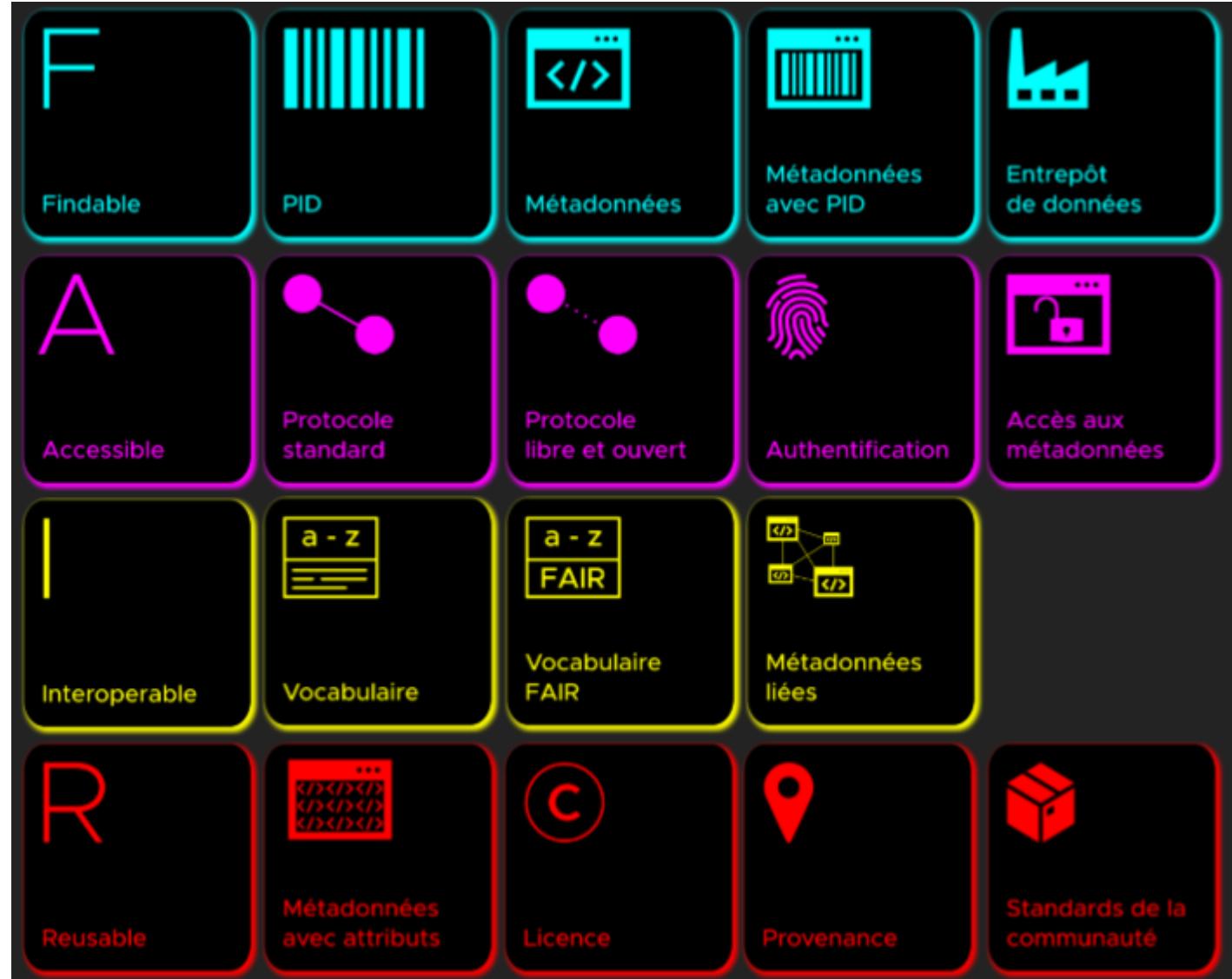
Processus itératif centré sur les bonnes pratiques et les principes FAIR



Processus itératif centré sur les bonnes pratiques et les principes FAIR



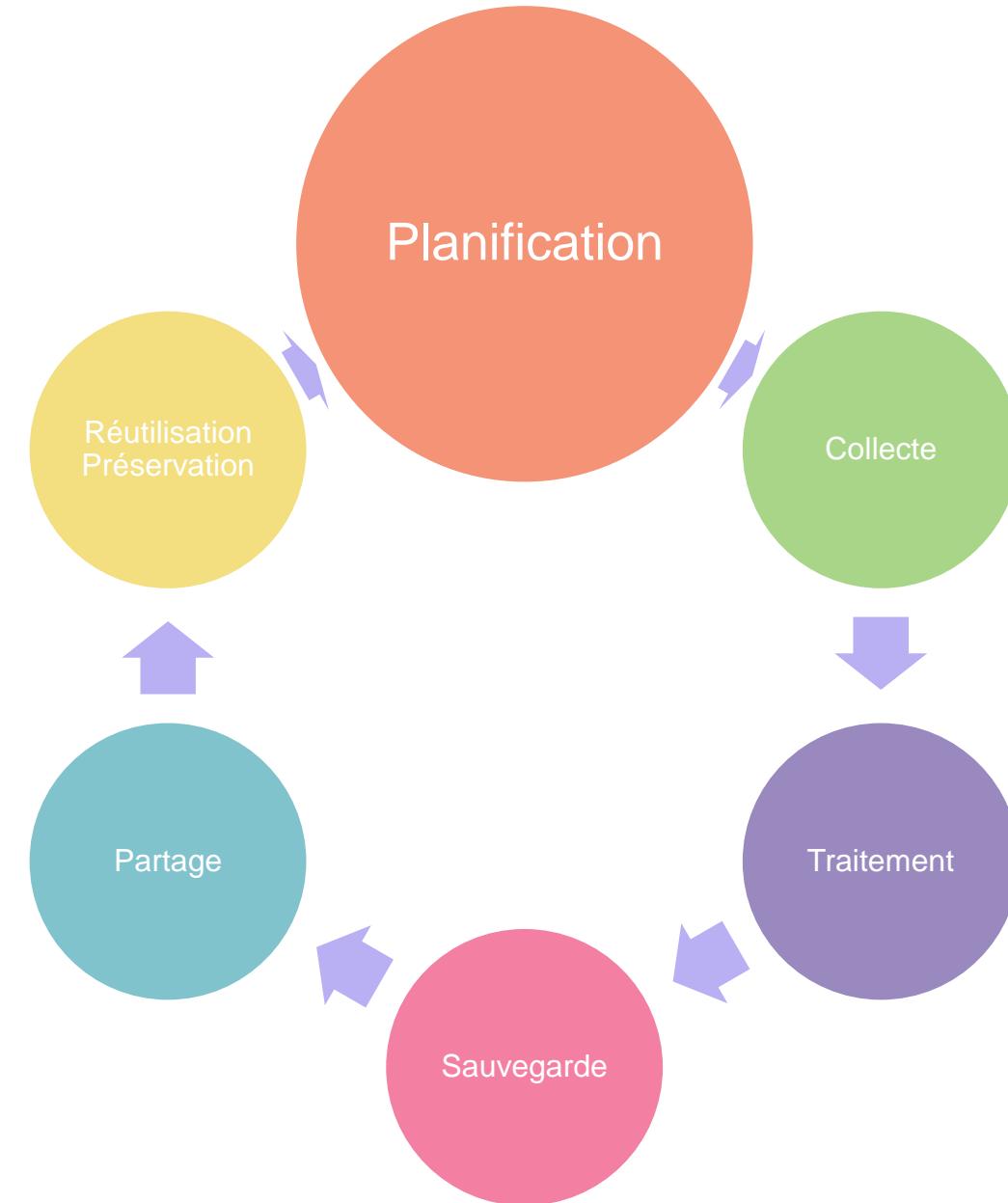
→ Bonnes pratiques de gestion et diffusion des données pour les rendre facilement réutilisables



Source : Doranum, INIST

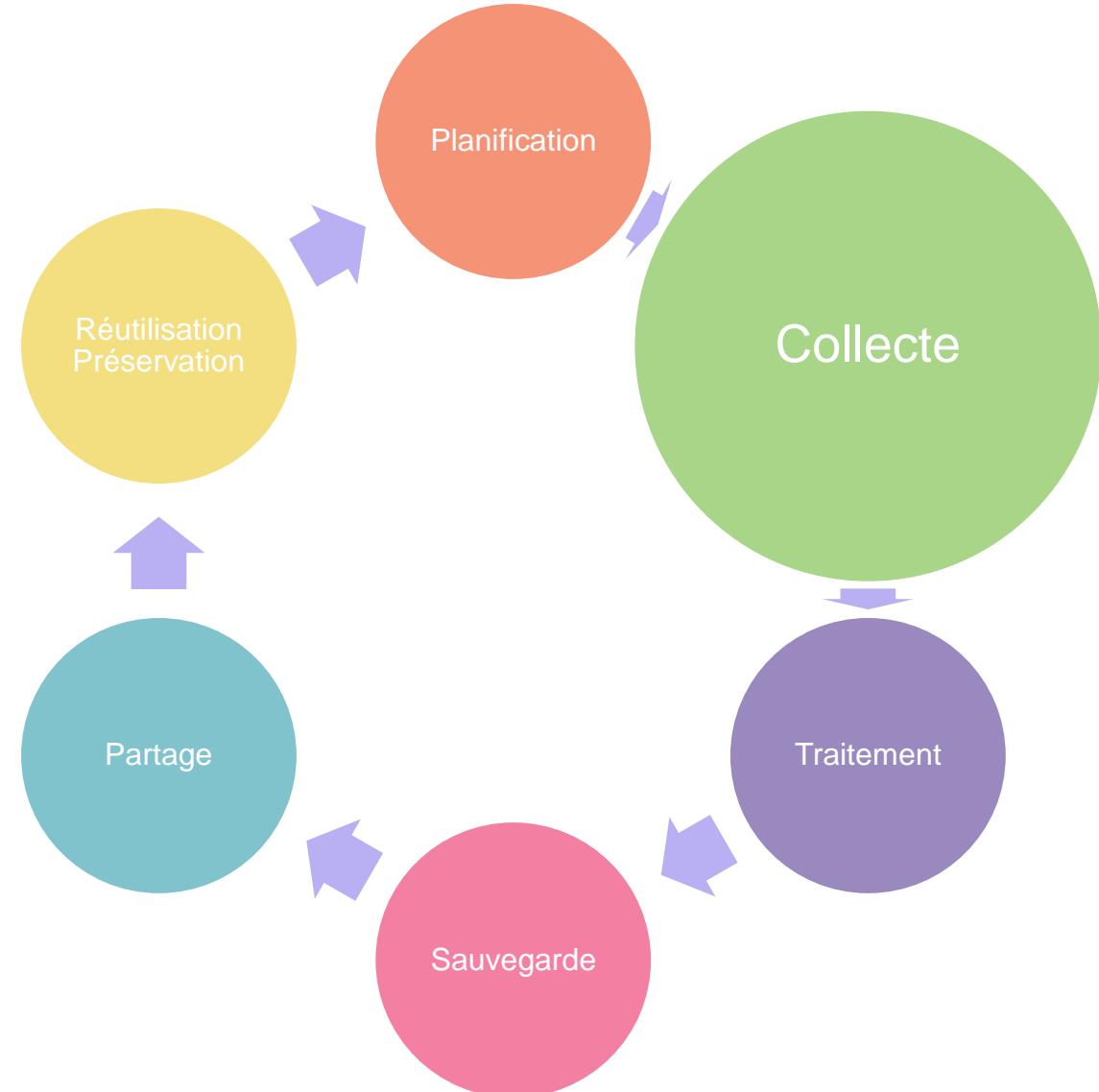
En amont du projet

- Réflexion sur toutes les phases de la vie des données avec les différents acteurs (partenaires, accompagnement...)
- Rédaction du plan de gestion de données.
- Définition des rôles et implications.
→ Phase primordiale pour la suite !



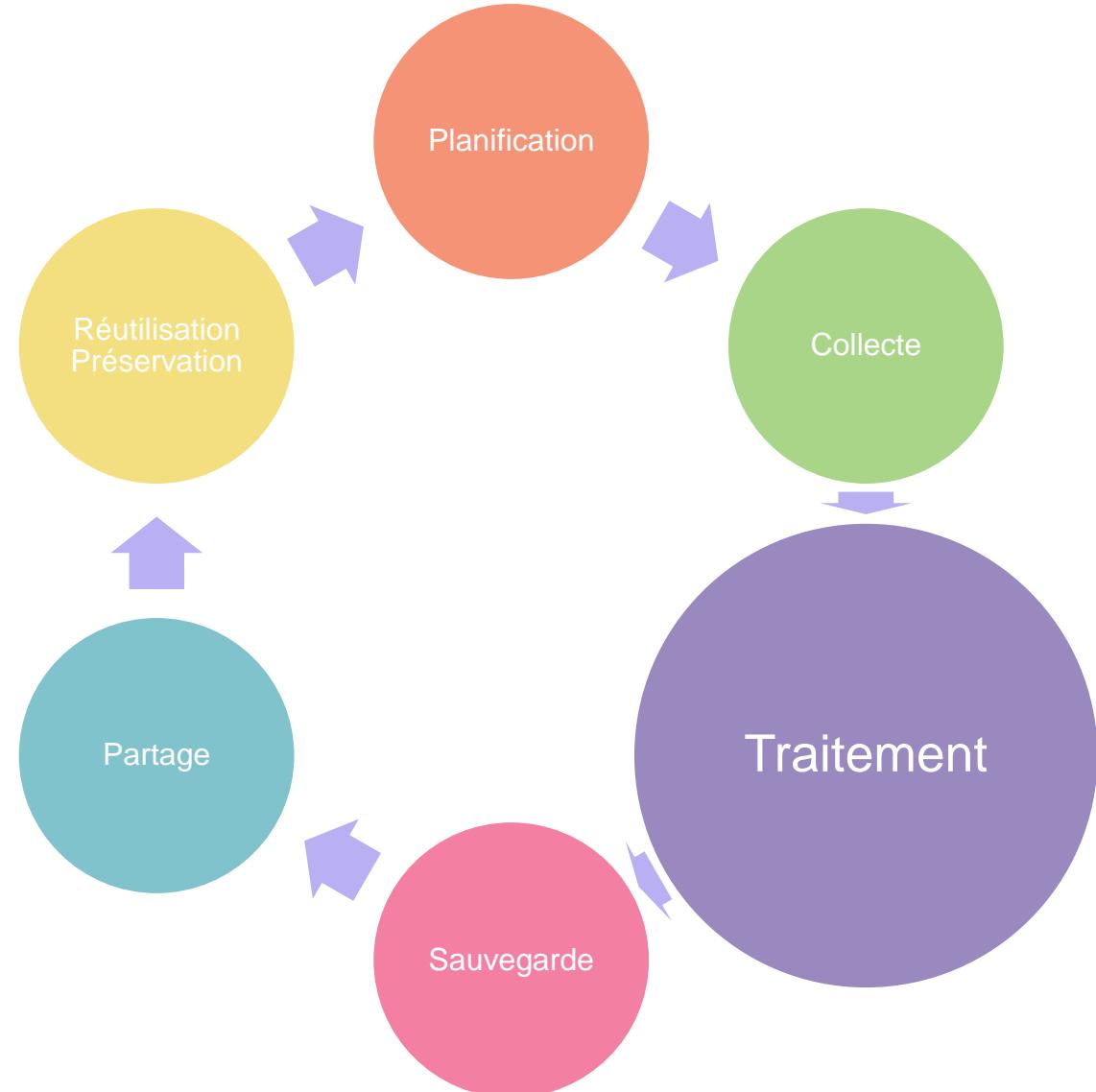
Collecte

- Il peut s'agir de données nouvelles ou de réutilisation de données existantes (attention aux droits et licences !)
- Rédaction des métadonnées.
- Méthodologie et protocole.
- Formats.
- Stockage durant le projet.



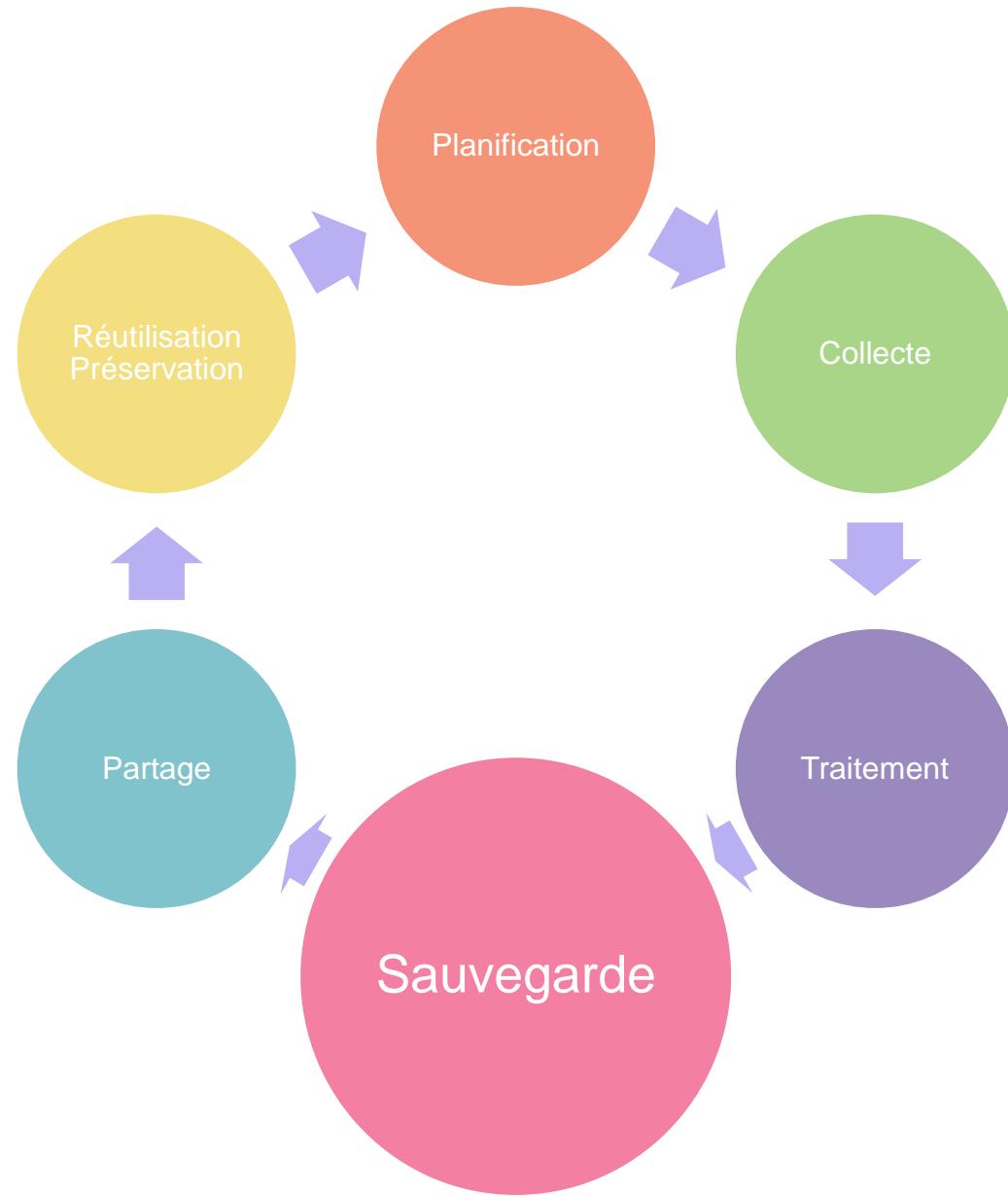
Traitement et analyse

- Nettoyage des données.
- Phase d'analyse (données brutes vs. données élaborées).
- Structuration des données qui seront partagées.



Sauvegarde

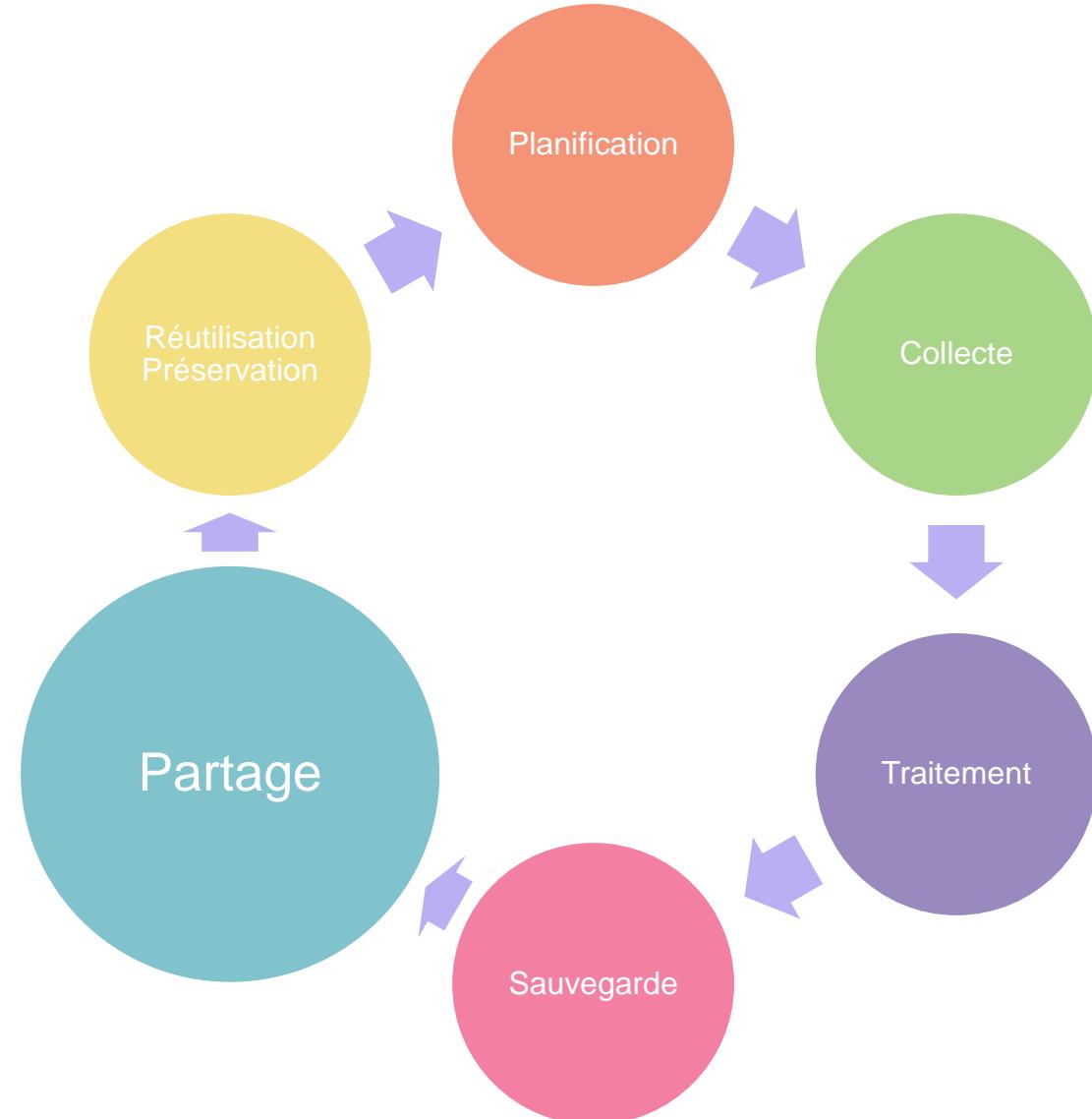
- Sauvegardes sécurisées sur des serveurs.
- Choix des données à pérenniser.
- Documenter les données, codes et logiciels qui seront sauvegardés sans toutefois être diffusés.



Partage et diffusion

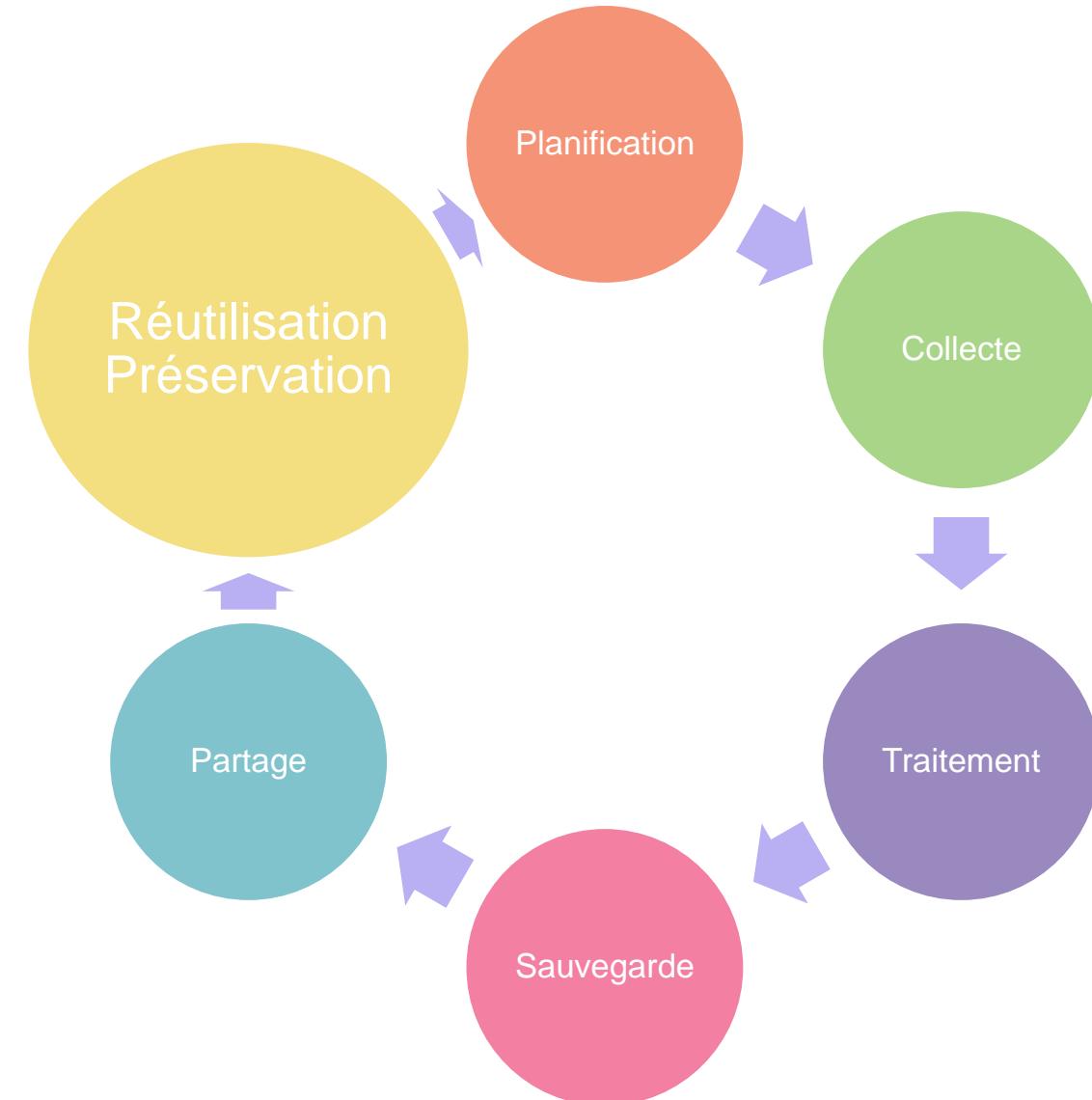
- Diffusion des jeux de données en « supplementary data » des articles scientifiques (attention au choix de l'entrepôt !)
- Publication de jeux de données dans les entrepôts de données.
- Publication de data papers ?

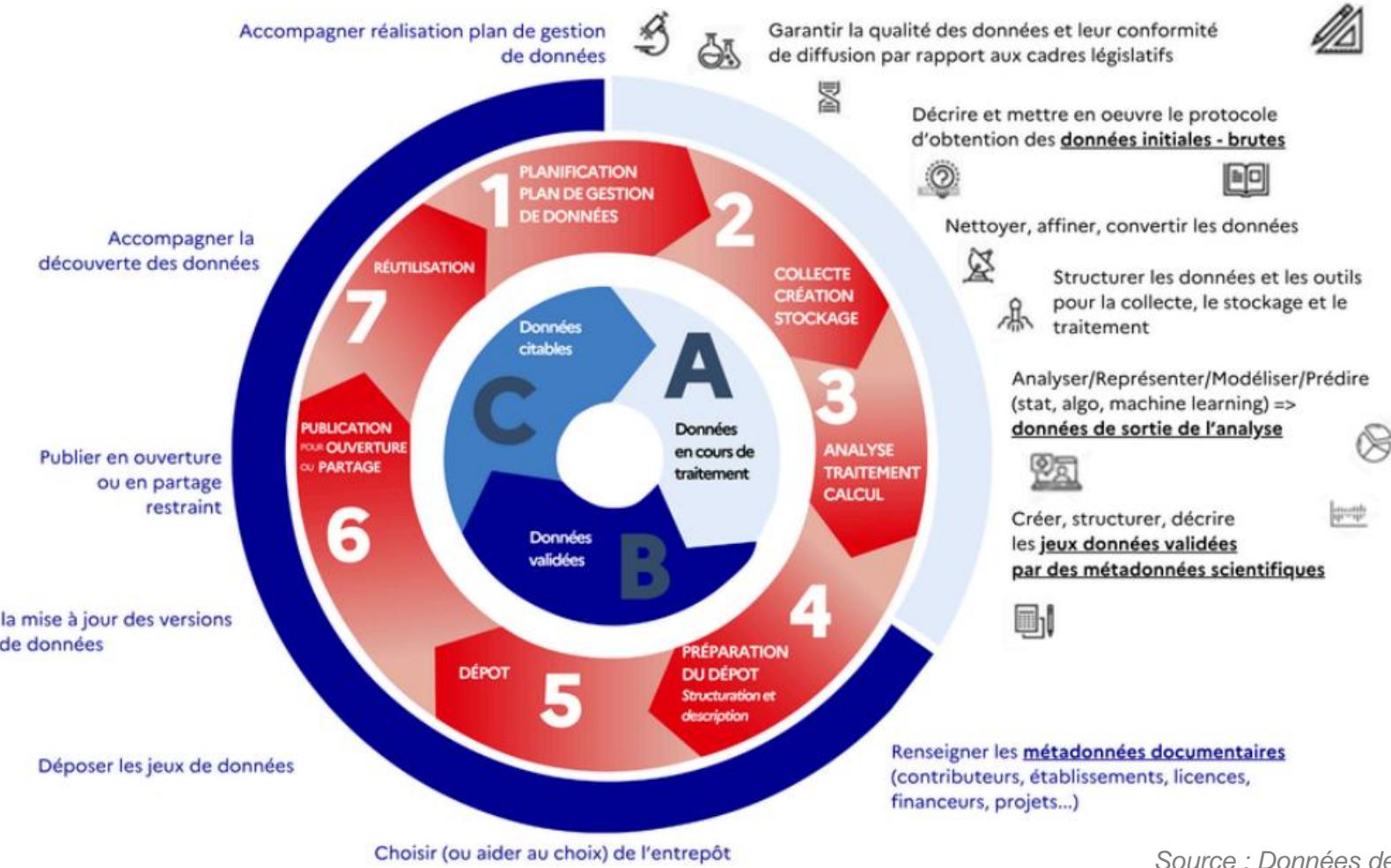
Etape cruciale pour la préservation et la réutilisation (licences, pérennité des systèmes...)



Réutilisation et préservation

- Préservation : des données ont-elles vocation à être archivées ? (le coût de l'archivage est à prévoir dès la planification : 1000€/To./an). Prendre en compte les contraintes de l'archivage (formats et documentation) dès le début.
- Réutilisation : importance des licences apposées. Création de nouvelles données. Réutilisation dans l'enseignement.





Source : *Données de recherche : caractérisation des fonctions clés liées à leur gestion, leur exploitation, leur diffusion et leur utilisation, MESR 2025*

Beaucoup de pratique !

Frise temporelle des données

Outils et acteurs du cycle de vie

But de l'atelier : lister les **outils** utilisés et les **personnes** impliquées à chaque étape du cycle de vie de vos données.

L'idée n'est pas d'être exhaustif mais de refléter vos pratiques.





Personnes
ressources

Planification Conception

Préparation du
projet

Collecte

Nouvelles données
Données existantes

Traitement

Traitement et
analyse des
données

Stockage Sauvegarde

Quelles données
sauvegarder et où ?

Partage Diffusion

Quelles données
partager et où ?

Réutilisation Préservation

Quelles données
préserver ?



Outils



ÉCOLOGIE &
ENVIRONNEMENT

École thématique Data SEE-Life

27 - 30 octobre 2025



Personnes
ressources



Outils

Planification
Conception

Préparation du
projet



Personnes
ressources

Qui peut
m'accompagner ?



Outils

Quels outils
peuvent m'aider ?

Réutilisation
Préservation

Quelles données
préserver ?



Personnes
ressources



Outils



Préparation du
projet

Planification et conception

Etape de préparation du projet

- Réflexion sur les données collectées et/ou réutilisées (types, formats)
- Anticipation des besoins : stockage, volumétrie, analyse, documentation des données...
- Définition des rôles et des acteurs
- Espaces partagés pour la documentation du projet



Personnes
ressources



Outils



ÉCOLOGIE &
ENVIRONNEMENT

Collecte

Nouvelles données
Données existantes

Collecte

Acquisition et réutilisation de données existantes

- Si données réutilisées : licence, droit de réutilisation
- Données collectées : capteurs ? Données personnelles (RGPD) ? Formats des données ?
- Méthodologie et protocole ?
- Bonnes pratiques lors de la collecte (nommage...)
- Espace de stockage pour ces données ?



Personnes
ressources



Outils

Planifi-
Concep-

Prépara-
pro

Traitem-
ent

Traitem-
ent et
analyse
des
données

Traitem- ent

Traitem- ent et analyse des données

- Quel logiciel/équipement est utilisé ? Format de sortie... (opensource ou licence)
- Quel type d'analyses et par qui ?
- Comment documenter cette étape (traçabilité)
- Quels standards ?



Personnes
ressources



Outils

Plan
Con

Prépa

Stockage
Sauvegarde

Quelles données
sauvegarder et où ?

Stockage et sauvegarde

Quelles données sauvegarder pendant le projet et où ?

- Quelles données doivent être partagées entre les membres du projet ? Où et comment elles seront partagées ? Quelle volumétrie ?
- Quelles données devront être sauvegardées ? Où et comment elles le seront ? Quelle volumétrie ?
- Sélection, nettoyage et structuration des données à pérenniser



Personnes
ressources



Outils

Plan
Cor

Prép

Partage
Diffusion

Quelles données
partager et où ?

Partage et diffusion

Quelles données diffuser et où ?

- Quelles données seront diffusées ? Sous quels critères, quelle forme, format ?
- Comment les données à diffuser seront structurées (standards, thésaurus, variables...) ?
- Sur quelles plateformes, quels entrepôts les données seront ouvertes ?
- Comment faire connaître mon projet ?
- Quelle licence sera apposée ?
- Comment seront citées ces données (auteurs...) ?



Personnes
ressources



Outils

Plan
Con

Réutilisation
Préservation

Quelles données
préserver ?

Réutilisation et préservation

Quelles données préserver ?

- Quelles données ont vocation à être préservées sur le long terme ?
- Sous quelles conditions ?
- Comment les citer (PID...) ?
- Combien va coûter cette préservation ?

Atelier pratique

Fin de l'atelier : 17h30
Debrief jusqu'à 18h

- ✓ Formez 3 groupes.
- ✓ Chaque groupe a une frise et son matériel pour la remplir.
- ✓ 17h30 : fin de la réflexion en groupes.
- ✓ 17h30-18h : Debrief pour créer une frise commune regroupant toutes les propositions.



Licence

<https://creativecommons.org/licenses/by/4.0>



Attribution 4.0 International (CC BY 4.0)

Vous êtes autorisé à :

- **Partager** — copier, distribuer et communiquer le matériel par tous moyens et sous tous formats
- **Adapter** — remixer, transformer et créer à partir du matériel pour toute utilisation, y compris commerciale.

Selon les conditions suivantes :

Attribution — Vous devez créditer l'Œuvre, intégrer un lien vers la licence et indiquer si des modifications ont été effectuées à l' Œuvre. Vous devez indiquer ces informations par tous les moyens raisonnables, sans toutefois suggérer que l'Offrant vous soutient ou soutient la façon dont vous avez utilisé son Œuvre.

Pas de restrictions complémentaires — Vous n'êtes pas autorisé à appliquer des conditions légales ou des mesures techniques qui restreindraient légalement autrui à utiliser l' Œuvre dans les conditions décrites par la licence.

Attribution : A. Henon – Ecole thématique Data SEE-Life - 2025