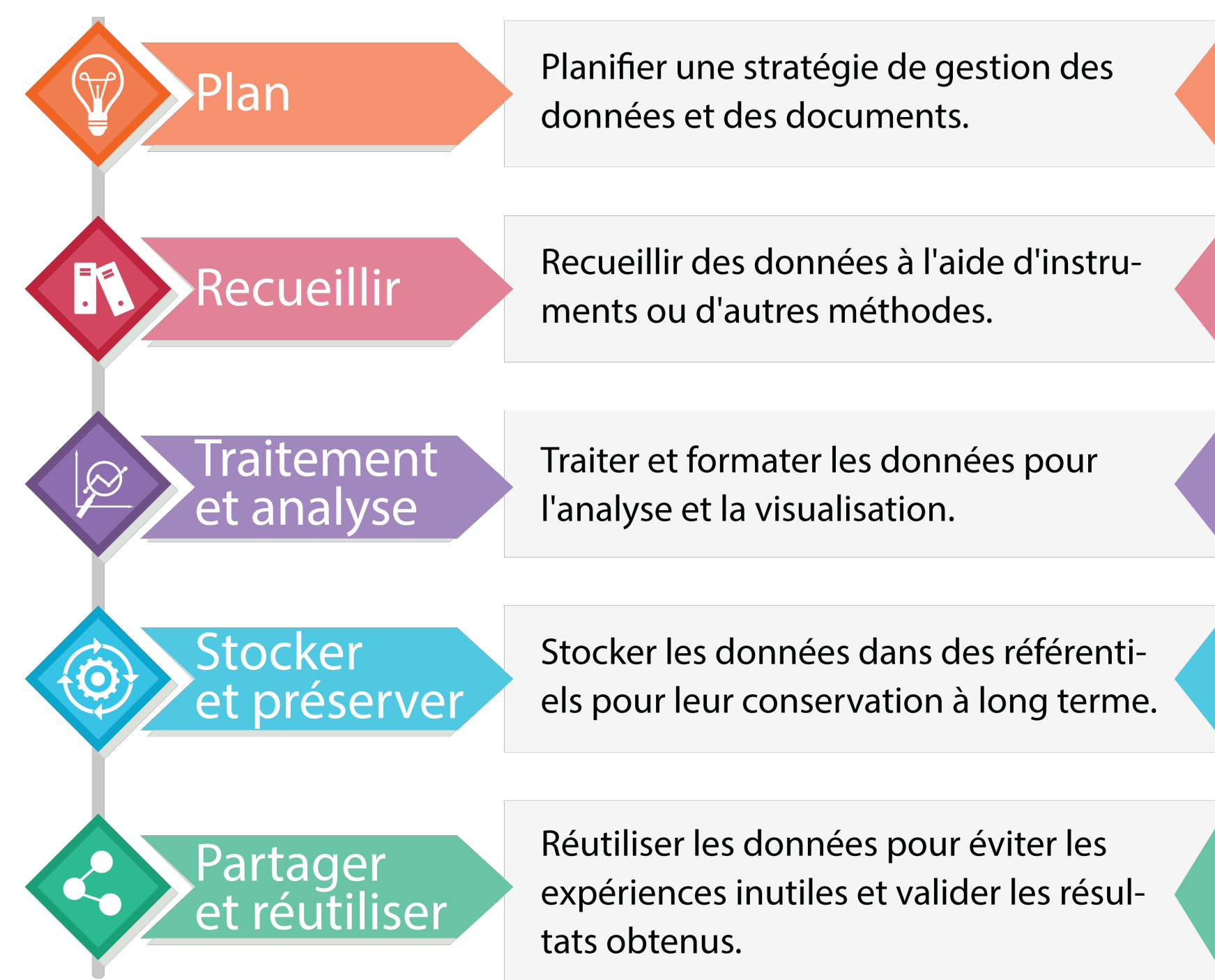


# Partager des données de recherche de haute qualité et des flux de travail reproductibles

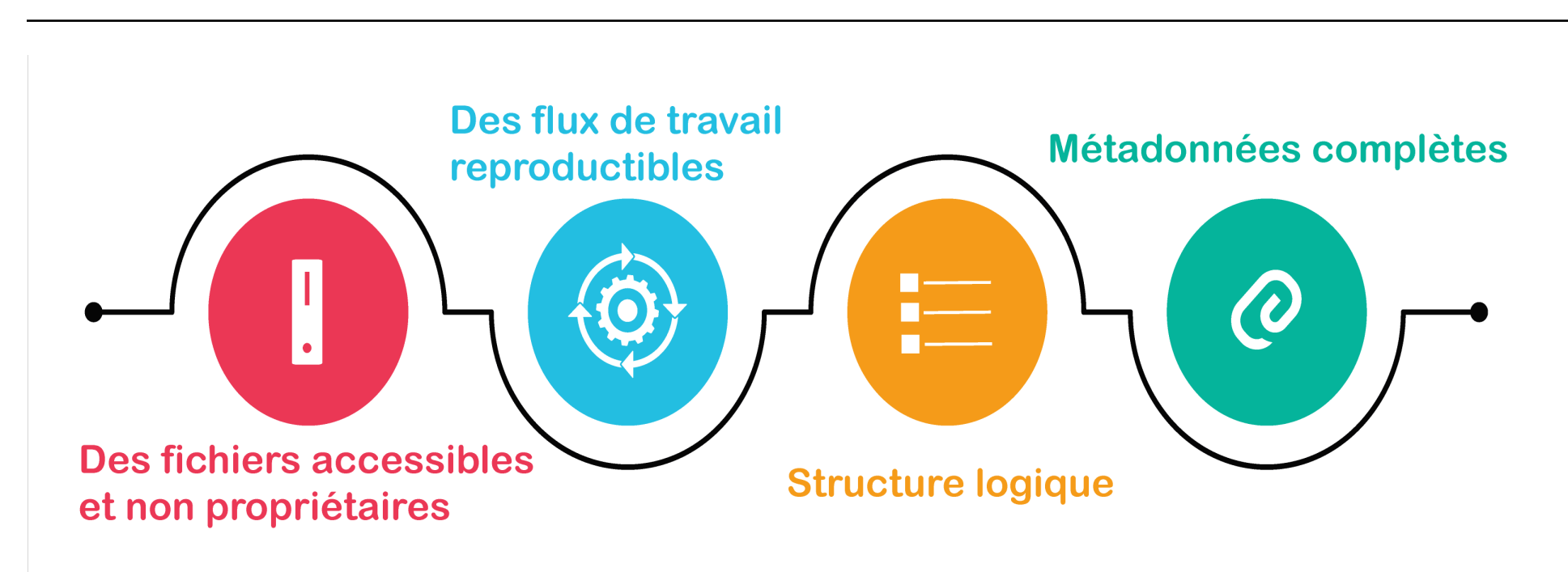
Équipe de curation des données de recherche (Dépôt fédéré de données de recherche -DFDR)

**KEYWORDS** — Gestion des données de recherche, science ouverte, partage des données

## I. CYCLE DE VIE DES DONNÉES DE RECHERCHE



## II. RENDRE LES DONNÉES DE RECHERCHE COMPRÉHENSIBLES



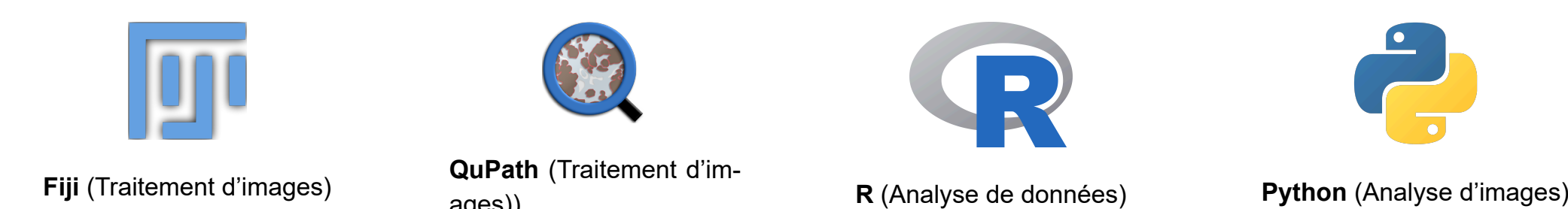
Le partage des données de recherche comprend des procédures systématiques permettant de les découvrir, comprendre et réutiliser au profit de la communauté. Il garantit également des flux de travail reproductibles qui permettent de reproduire et de valider les résultats de recherche.

### i. Privilégier les formats de fichiers ouverts et non propriétaires

Type de données	Formats recommandés	Notes
Données tabulaires	CSV, TSV	Largement soutenu, non propriétaire
Documents textuels	TXT, PDF/A	PDF/A est un format d'archivage à long terme
Images	TIFF, PNG	TIFF pour un résultat sans perte, PNG pour une taille plus petite
Données géospatiales	GeoTIFF, Shapefile	S'assurer que tous les composants du fichier Shapefile sont inclus
Données audio	WAV, FLAC	WAV pour l'audio non compressé
Données vidéo	MP4, MOV	MP4 est largement compatible, mais avec perte
Code	PY, R, QMD	Fournir un code et des scripts bien commentés

### ii. Mettre en oeuvre des flux de travail reproductibles

Utiliser des logiciels libres pour traiter les images et les données. Utiliser des scripts pour le traitement par lots et permettre des flux de travail reproductibles.

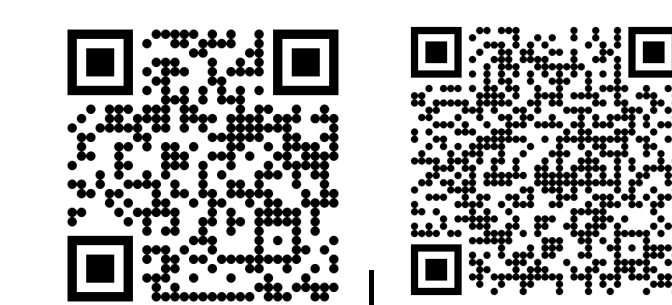


### iii. Organiser vos fichiers de manière logique

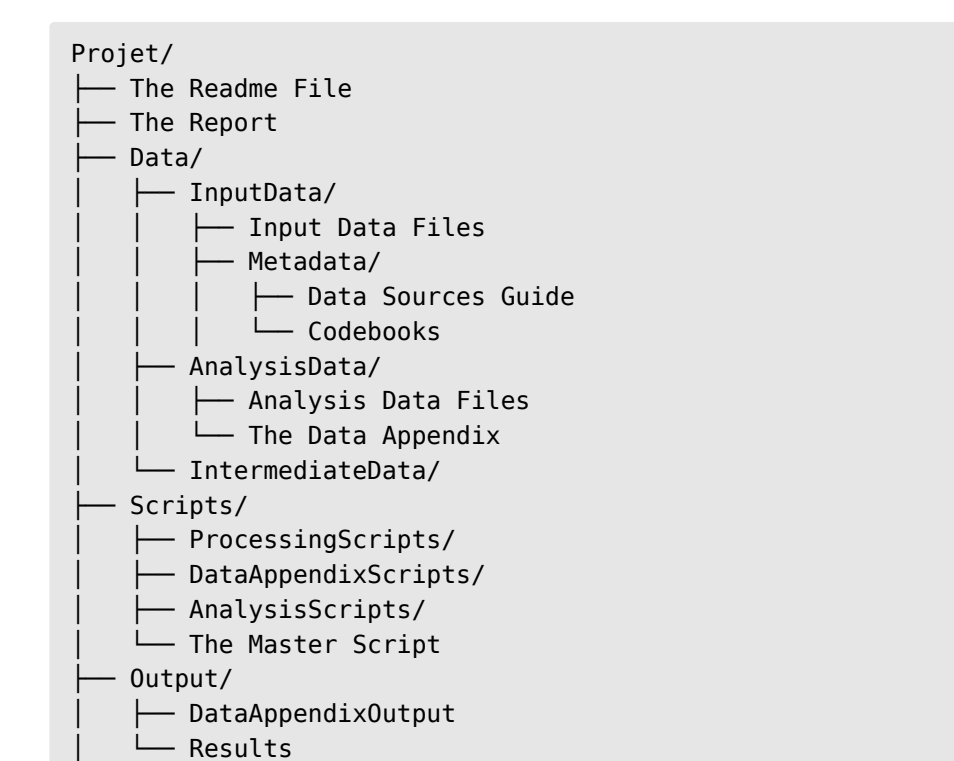
Les fichiers doivent être faciles à retrouver et à identifier pour tout réutilisateur potentiel.

- Mettre au point une **structure de dossiers logique** hébergeant les fichiers bruts et traités, le code d'analyse et les résultats finaux.
- Utiliser des **conventions de dénomination** descriptives pour étiqueter les dossiers et les fichiers.

Exemple de projet / Générateur d'arbres



Protocole TIER 4.0



### iv. Mettre en oeuvre des métadonnées complètes

Décrire et contextualiser ses données à l'aide de **métadonnées**, de fichiers **README** et de **livres de code**. Cela permet une réutilisation correcte des jeux de données. Pensez aux éléments suivants lors de l'élaboration de la documentation:



Télécharger un modèle de README

- **Auteurs et contenu du jeu de données:** Ils permettent à l'utilisateur de comprendre l'ensemble de données indépendamment des articles de recherche associés. Les jeux de données sont des **objets autonomes** utilisables seuls.

Exemple depuis <https://doi.org/10.20383/103.0802>

Le jeu de données contient des données de caractérisation et d'activité pour une série d'échantillons d'électrodes de nitrure de carbone. Le nitrure de carbone a été cultivé sur du FT0, du FT0 nettoyé au plasma et du FT0 aminosilanisé nettoyé au plasma. Du platine a également été ajouté sur le CNx pour certains échantillons...

- **Informations sur l'accès/la licence:** Décrire comment la communauté peut réutiliser les données. Vérifier le consentement de l'Open Brain lorsqu'il s'agit de données sensibles, par exemple.

CC BY 4.0 Attribution 4.0 International

Cette licence standard permet aux utilisateurs de **partager** (-copier et redistribuer) et **d'adapter** (remix, transform) (remixer, transformer) le matériel pour n'importe quel usage, à condition d'en mentionner les auteurs originaux.



Visitez l'Open Brain Consent

- **Informations méthodologiques:** Décrire comment les données ont été générées. Cette information permet de contextualiser et de valider le jeu des données.

Exemple depuis <https://doi.org/10.20383/103.0802>

Les mesures électrochimiques ont été effectuées dans une cellule électrochimique en quartz à l'aide d'un potentiostat PalmSens4. L'électrolyte a été entièrement purgé avec de l'Ar entre les balayages pour créer un environnement électrochimique inerte...

## III. PARTAGEZ VOS DONNÉES DANS DES RÉFÉRENTIELS

Ces bonnes pratiques de gestion des données font que les ensembles de données peuvent être partagés dans des dépôts comme **DFDR** (<https://www.frdr-dfdr.ca/>) et **Borealis** (<https://borealisdata.ca/>). Le DFDR est un dépôt multidisciplinaire, bilingue et curaté, destiné exclusivement aux chercheurs canadiens et spécialisé dans les dépôts de données volumineux. Borealis, l'instance canadienne du Dataverse, est un dépôt de données de recherche bilingue, multidisciplinaire et sécurisé, soutenu par les bibliothèques universitaires et les institutions de recherche à travers le Canada.



Accord tri-agences canadien

Le gouvernement du Canada a établi en 2021 le Cadre des trois organismes pour favoriser une conduite responsable de la recherche, notamment des bonnes pratiques pour mettre en oeuvre des plans de gestion des données (PGD) et partager des données provenant de la science financée par des fonds publics.