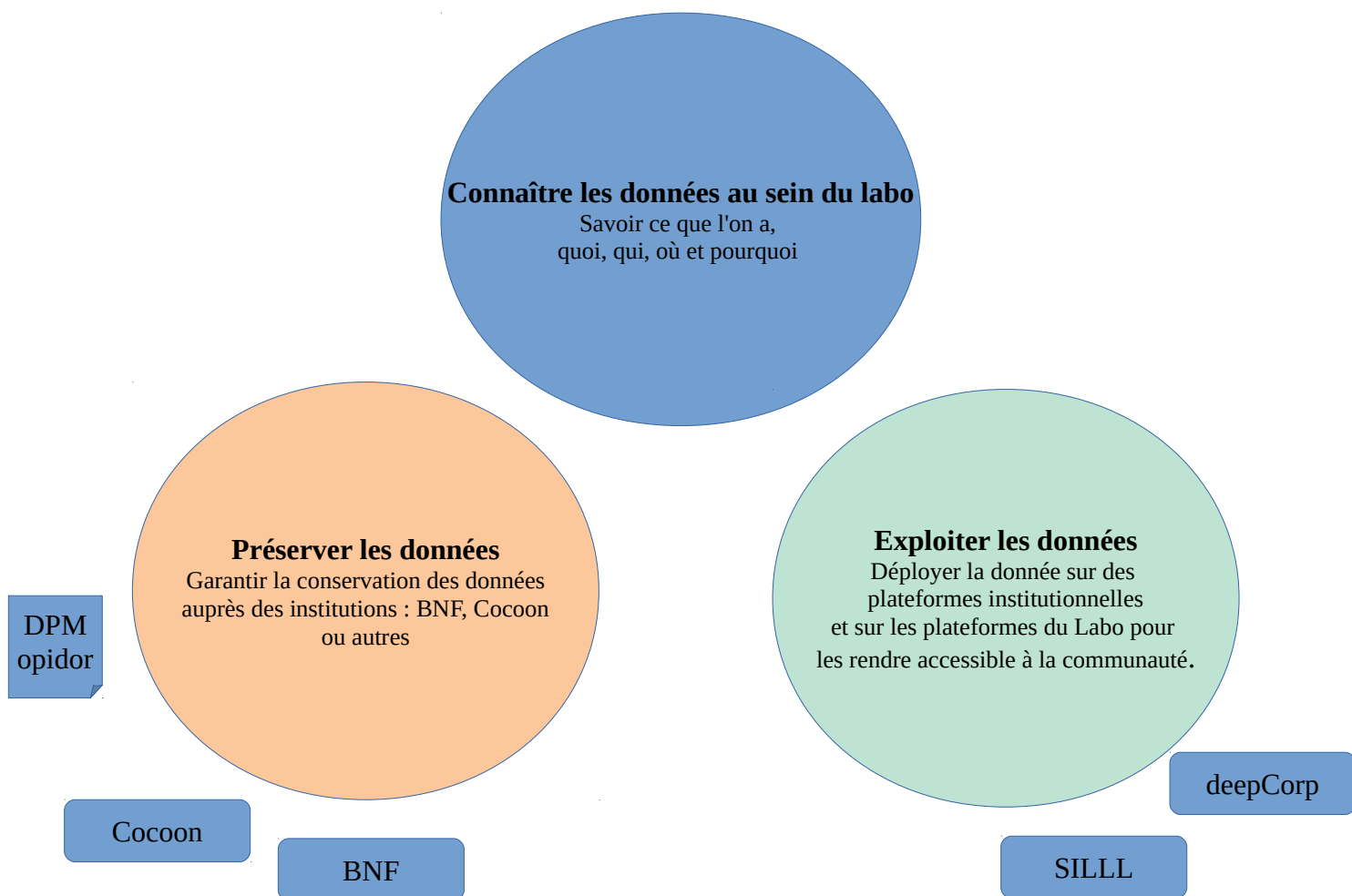
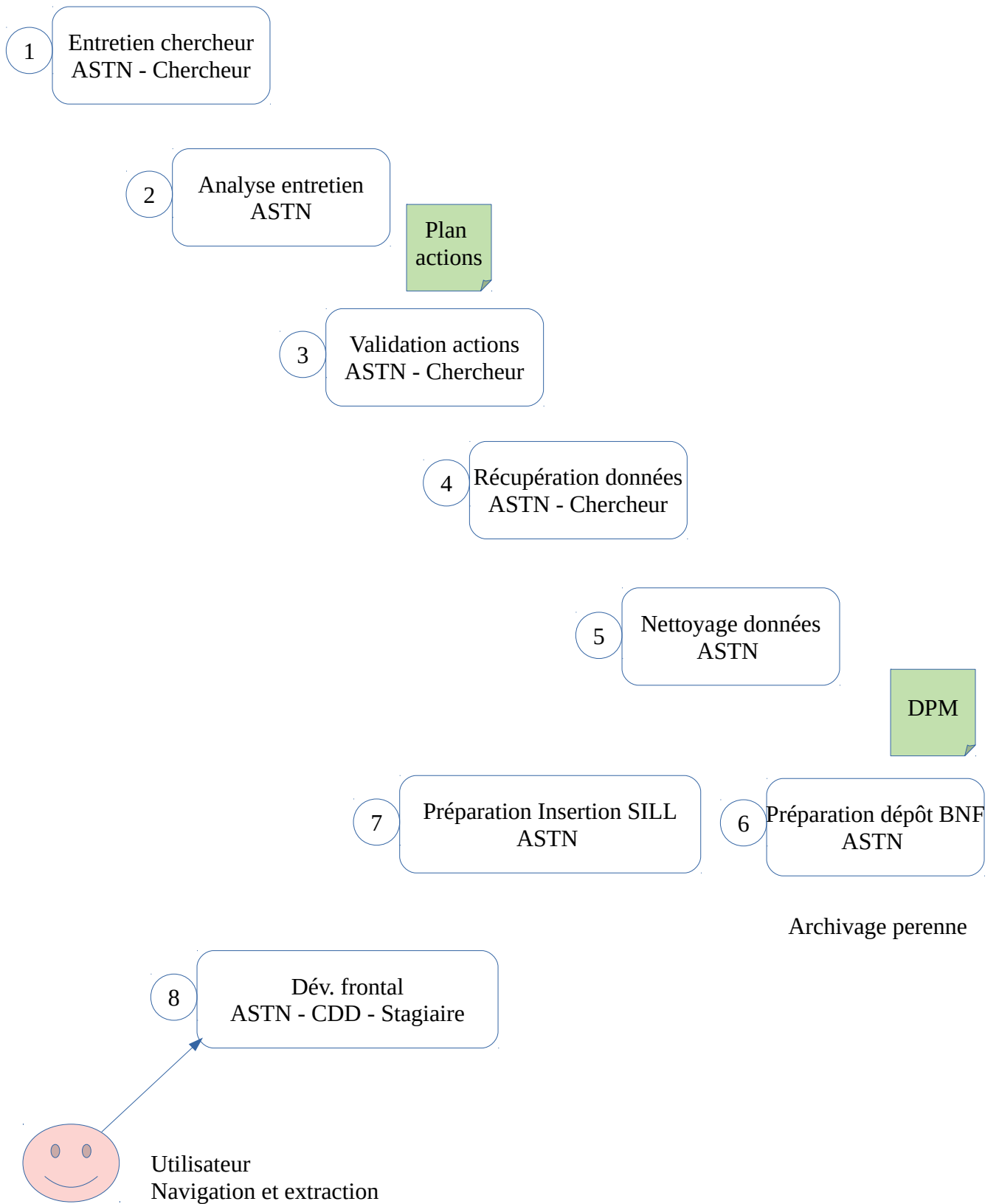


Processus d'accompagnement pour les données
ASTN/LLL 2023
YS

Ce document décrit les tâches à effectuer pour l'intégration des données dans le cadre de l'open Science. Il décrit les acteurs, les actions à effectuer pour mener à bien les opérations avec trois objectifs : connaître les données au sein du LLL, garantir leur conservation et permettre leur exploitation.



Workflow de traitements



Déroulement des tâches à exécuter pour atteindre ces trois objectifs.

1^{ère} tâche :

Acteurs : ASTN --> Chercheur

Intitulé de la tâche : Entretien chercheur

Contenu de la tâche : Identifier les données : collecte des informations sur les données produites par le chercheur. L'auteur de cette collecte, l'objectif et toutes les informations permettant de décrire le contexte de cette collecte ainsi que toutes les productions scientifiques qui en découlent.

Outil utilisé : questionnaire <https://astn.lll-projects.org/data-du-lll/>

Production : questionnaire rempli

2^{ème} tâche :

Acteurs : ASTN

Intitulé de la tâche : analyse du questionnaire issu de l'entretien

Contenu de la tâche : après l'entretien avec le chercheur , analyse du questionnaire et préparation des actions à entreprendre avec indication du degré d'urgence.

Outil utilisé : critère de préservation et de qualité de la donnée. Mesurer le risque de perdre la donnée.

Production : plan d'action dédié à proposer au chercheur. Le déroulement des étapes à suivre pour conserver, compléter et exploiter cette données. Voir les offres de dépôt institutionnel, la préparation des données pour les rendre compatibles avec ces dépôts et de voir comment permettre leur exploitation et dans quelle condition.

3^{ème} tâche :

Acteurs : ASTN --> Chercheur

Document en entrée : plan d'actions

Intitulé de la tâche : discussion et mise au point du plan d'actions avec le chercheur.

Contenu de la tâche : validation du plan d'actions et récupération des données .

Outil utilisé : support de stockage et de protection associé à la confidentialité de cette donnée.

Production : plan d'actions validé

4^{ème} tâche :

Acteurs : ASTN

Document en entrée : plan d'actions + les données récupérées.

Intitulé de la tâche : analyse du contenu des données

Contenu de la tâche : analyse approfondie des données, contrôle de leur contenu et éventuellement opération d'alimentation si besoin. Définir des structurations des données adaptées à une exportation vers les sites de dépôts institutionnels et vers les plateformes ASTN/LLL.

Outil utilisé : éditeur, logiciel d'écoute sonore et de transcription, OCR si les données sont fournies dans un format propriétaire (docx, pdf).

Production : ensemble des données complétée et réactualisée

5^{ème} tâche :

Acteurs : ASTN

Documents en entrée : données et leur description

Intitulé de la tâche : remise en conformité

Contenu de la tâche : reprendre l'ensemble des données pour remise en conformité. Définir des structurations des données adaptées à une exportation vers les sites de dépôts institutionnels et vers les plateformes ASTN/LLL. Voir la rédaction d'un DPM pour un dépôt dans opidor .

Outil utilisé : <https://dpm.opidor.fr>

Production : lot pour les sites d'archivage dépôts et pour l'utilisation au sein des plateformes ASTN/LLL.

6^{ème} tâche :

Acteurs : ASTN --> BNF --> Cocoon

Documents en entrée : lot des données

Intitulé de la tâche : transfert vers le dépôt

Contenu de la tâche : mettre à disposition les données pour un dépôt au sein de la BNF.

Outil utilisé : transfert de fichiers

Production : un accès vers ces dépôts pour un public plus large.

7^{ème} tâche :

Acteurs : ASTN

Documents en entrée : lot des données

Intitulé de la tâche : insertion dans la base de données de SILLL de ASTN/LLL

Contenu de la tâche : transférer les données dans le contexte de SILLL sous la base de données CouchBase en document JSON.

Outil utilisé : script d'insertion en python.

Production : insertion des données dans le réceptacle SILLL et tests de l'exploitation.

8^{ème} tâche :

Acteurs : ASTN --> stagiaires ou CDD

Documents en entrée : contenu de SILLL

Intitulé de la tâche : développement du frontal dédié

Contenu de la tâche : développer en prenant le template sitecorpus, une version d'un frontal permettant de naviguer et d'exploiter les données mises à disposition.

Outil utilisé : Wordpress, mysql

Production : site web avec une url de type https://nom_copus.lll_projects.org

9^{ème} tâche :

Acteurs : ASTN --> stagiaires ou CDD

Documents en entrée : contenu de SILLL

Intitulé de la tâche : intégration dans deepcorp

Contenu de la tâche : vérification de la prise en compte et aménagement de la plateforme multi-corpus deepcorp pour l'exploitation et la prise en compte dans les traitements de ce nouveau corpus.

Outil utilisé : wordpress

Production : site <https://deepcorp.org>