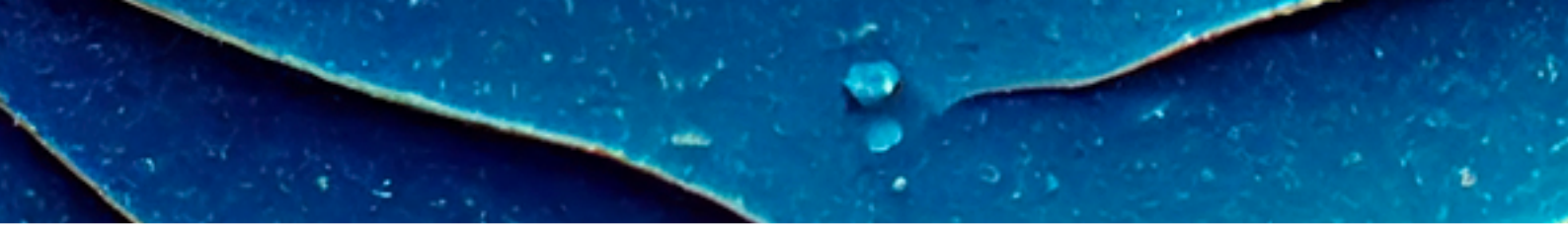
BONNES PRATIQUES DE MISE EN PLACE D'UN CAHIER DE LABORATOIRE ÉLECTRONIQUE

Exemple d'eLabFTW



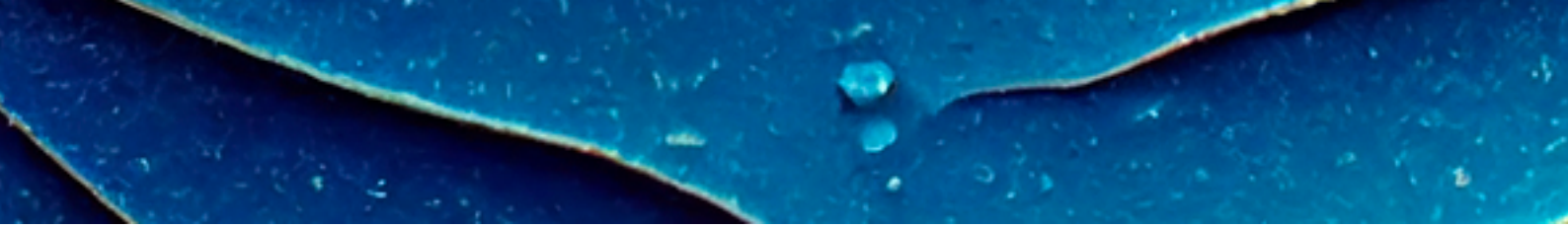
Bonnes pratiques de mise en place d'un cahier de laboratoire électronique

Exemple d'eLabFTW

Yaël Hersant, Nathalie Léon, Alain Rivet et Henri Valeins

Mission pour les Initiatives Transverses et Interdisciplinaires

2022



AUTEURS

Yaël Hersant
Responsable qualité et développement durable
Institut de Chimie de Clermont-Ferrand (ICCF) - UMR6296
Campus Universitaire des Cézeaux / 24 avenue Blaise Pascal
CS 60026
63178 AUBIERE Cedex
yael.hersant@uca.fr

Nathalie Léon
Chargée de projet Cahiers de laboratoire électroniques & Démarche d'amélioration Continue
MPR – Mission aide au pilotage et relations avec les délégations régionales et les instituts
3 rue Michel Ange
75794 PARIS Cedex 16
nathalie.leon@cnrs.fr

Alain Rivet
Responsable qualité et système d'information
Centre de Recherches sur les Macromolécules Végétales
CS 40700
38041 Grenoble cedex 9
alain.rivet@cermav.cnrs.fr

Henri Valeins
Responsable système d'information et qualité
Centre de de Résonance Magnétique des Systèmes Biologiques
146 rue Léo Saignat
33076 Bordeaux cedex
henri.valeins@rmsb.u-bordeaux.fr

REMERCIEMENTS

Les auteurs remercient le CNRS à travers sa Mission pour les Initiatives Transverses et Interdisciplinaires (MITI) qui soutient le réseau Qualité en Recherche. Pour les échanges fructueux dans la relecture de ce document, ils sont particulièrement reconnaissants à :

- Jean-Marie Burgio de la MPR/DGDR du CNRS
- Nicolas Carpi de l'Institut Curie
- Michel Chabanne RSSI du CNRS
- Marie-Laure Bachèlerie DAJ du CNRS
- Bérénice Baranger
- François Baranger pour la sagacité de ses illustrations

Ce guide a été réalisé en collaboration avec le groupe de travail du CNRS sur les « cahiers de laboratoire électroniques » co-porté par le DGDS et la MPR du DGDR. Il s'intègre parfaitement dans la démarche d'accompagnement au changement engagé par l'établissement.

A travers ce guide de bonnes pratiques, le réseau Qualité en Recherche souhaite sensibiliser les personnels des unités de recherche à la mise en place et à l'utilisation d'un cahier de laboratoire électronique au sein d'une unité de recherche en apportant une vision « terrain » à cette nouvelle organisation des activités de recherche.

TABLE DES MATIÈRES

1.	**Préambule**	**7**
2.	**Introduction**	**8**
2.1	- Le contexte national	9
2.2	- Le cahier de laboratoire, un outil d'enregistrement et de traçabilité	10
2.3	- Numérique versus papier	11
3.	**Les initiatives nationales**	**13**
3.1	- INSERM	13
3.2	- DATACC	13
3.3	- CNRS	14
3.4	- Comité pour la science ouverte (CoSO)	16
4.	**Caractéristiques fonctionnelles et techniques**	**17**
5.	**Mise en place d'un CLE au sein d'une unité de recherche**	**20**
5.1	- Conduite du changement	20
5.2	- Pilotage du CLE	22
5.3	- Organisation humaine	23
5.4	- Création des zones de travail collaboratif	24
6.	**Gestion des données**	**25**
6.1	- La gestion des données externes	25
6.2	- Définir des règles de nommage	26
6.3	- Choisir un plan de classement	27
6.4	- Sauvegarde et archivage	28
6.5	- Le plan de gestion de données	31
7.	**eLabFTW, un exemple de CLE**	**33**
7.1	- Les fonctionnalités	33
7.2	- Clôture du CLE	38
8.	**Conclusion**	**39**
	Bibliographie	41
9.	**Annexes**	**42**
	Annexe 1 : Exemple de trame de plan de gestion de données	42
	Annexe 2 : Etapes pour la mise en place d'une "zone cahier"	44

1. PRÉAMBULE

Bonnes pratiques de mise en place d'un cahier de laboratoire électronique
Exemple d'eLabFTW

Le réseau Qualité en Recherche (QeR) est un des réseaux métiers et technologiques regroupés au sein de la Mission pour les Initiatives Transverses et Interdisciplinaires du CNRS (MITI). Le réseau QeR est le réseau référent de la démarche qualité au sein des unités de recherche. Il en assure la promotion et la diffusion au sein des différentes communautés du CNRS et des autres partenaires de la recherche.

La mise en place d'une démarche qualité répond au besoin d'assurer la maîtrise des activités de recherche telles que l'optimisation des ressources humaines et matérielles, la formalisation et la transmission du savoir-faire, la gestion des connaissances ou encore l'amélioration de la sécurité. A ce titre, le cahier de laboratoire est un outil essentiel de suivi des activités de recherche. Il évolue dans une structure de plus en plus marquée par la numérisation des activités de recherche et dans un contexte de science ouverte.

De ce fait, le but de ce guide est d'apporter une vision pratique sur la mise en place et l'utilisation des cahiers de laboratoire électroniques en abordant :

- Les notions d'organisation au préalable à l'utilisation de tels supports
- Les recommandations pratiques dans leur utilisation
- Les mesures à adopter dans la gestion des données brutes attenantes

Dans ce guide, nous proposons des bonnes pratiques d'organisation, de gestion rigoureuse et cohérente des données de la recherche dans le cadre de la mise en œuvre de cahiers de laboratoire électroniques dans une entité (laboratoire, équipe de recherche, etc.). Nous abordons également les contraintes liées à leur utilisation ainsi que des recommandations dans le cadre de leur déploiement afin de lever certains freins.

Pour ce document, nous avons choisi d'utiliser le terme générique "cahier de laboratoire électronique" et l'acronyme "CLE". Vous pouvez aussi retrouver, dans certaines citations, son équivalent anglais "ELN" pour "Electronic Laboratory Notebook.

2. INTRODUCTION

La gestion rigoureuse et cohérente des données de la recherche constitue aujourd'hui un enjeu de taille pour la production de nouvelles connaissances scientifiques. Guidés par le « Plan National pour la Science Ouverte », les différents organismes de recherche et les Instituts du CNRS se sont emparés de ces questions primordiales pour participer à la réflexion et à la mise à disposition des outils, méthodes et infrastructures répondant aux besoins des communautés scientifiques en matière de gestion et de partage des données scientifiques.

Dans ce cadre, le réseau Qualité en Recherche du CNRS souhaite apporter une vision pratique sur la mise en place de ces cahiers de laboratoire électroniques en abordant les notions d'organisation préalable à l'utilisation de tels supports, les recommandations pratiques dans leur utilisation et les mesures à adopter dans la gestion des données brutes attenantes.

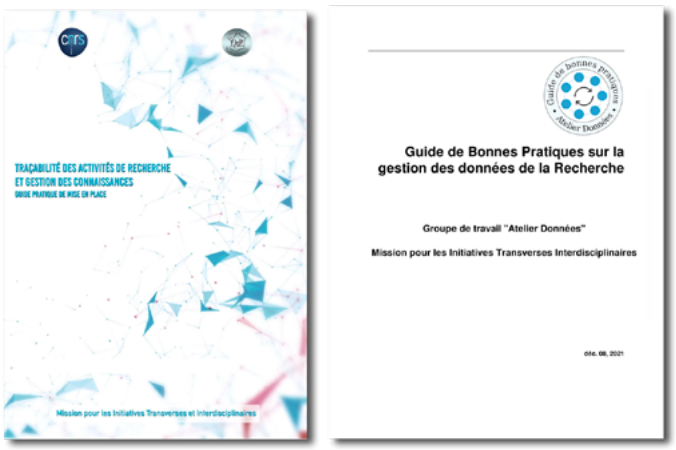


Figure 1 - Guides pratiques

A l'instar des guides « Traçabilité des activités de recherche et gestion des connaissances » (2018) et « Guide de bonnes pratiques de gestion des données de la recherche » (2021) respectivement publiés par le réseau Qualité en Recherche et l'Atelier Données de la Mission pour les Initiatives Transverses et Interdisciplinaires du CNRS, l'originalité de ce guide réside dans son application aux données de la recherche sous l'angle des pratiques métiers de la recherche en fournissant une approche « terrain » en complément d'une politique institutionnelle.

2.1 - Le contexte national

Dans la ligne du “Plan National pour la Science Ouverte” annoncé par Frédérique Vidal le 4 juillet 2018, destiné à rendre obligatoire l’accès ouvert pour les publications et pour les données issues de recherches financées sur projets, le “Plan Données de la recherche du CNRS » (2020) a été mis en place pour inciter les scientifiques à rendre leurs données accessibles et réutilisables.

La science ouverte introduit la notion de FAIR data (Figure 3) qui consiste à rendre les données Faciles à trouver, Accessibles, Interopérables et Réutilisables. Ces principes permettent de guider les stratégies de gestion des données et d’aider tous les acteurs qui œuvrent à les produire, à en contrôler la qualité, à les traiter, les analyser, à assurer leur publication et leur diffusion sur des plateformes de partage ou d’archivage.

Figure 2 - Plan données de la recherche

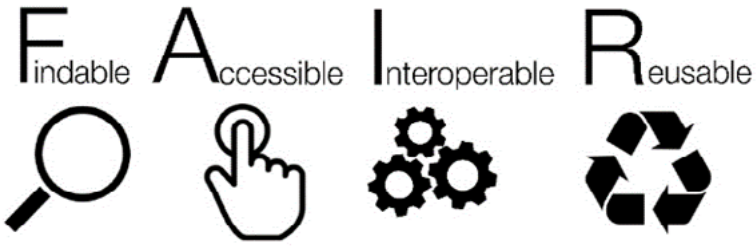


Figure 3 - Les 4 principes du Fair data (Sangya Pundir [CC BY-SA 4.0])

Dans ce contexte, rappelons également l’existence de la charte nationale de déontologie des métiers de la recherche (figure 4) qui s’adresse à l’ensemble des femmes et des hommes d’un établissement ou d’un organisme, permanents ou non, contribuant à l’activité de recherche avec l’objectif « d’expliciter les critères d’une démarche scientifique rigoureuse et intègre, applicable notamment dans le cadre de tous les partenariats nationaux et internationaux » (Charte de déontologie, 2015).

Figure 4 - Charte de déontologie

Relevons deux items de cette charte qui concernent plus particulièrement notre problématique :

« La description détaillée du protocole de recherche dans le cadre des cahiers de laboratoire, ou de tout autre support, doit permettre la traçabilité des travaux expérimentaux »

« Tous les résultats bruts (qui appartiennent à l’institution) ainsi que l’analyse des résultats doivent être conservés de façon à permettre leur vérification. »

2.2 - Le cahier de laboratoire, un outil d'enregistrement et de traçabilité

L'ensemble des données produites par la recherche doit ainsi être répertorié et enregistré dans l'objectif d'une réutilisation potentielle. Nous disposons pour ce faire de supports tels que les cahiers de laboratoire.

Le cahier de laboratoire est un outil non obligatoire mais fortement recommandé pour toute structure générant des données donnant lieu à des connaissances diffusables et valorisables notamment dans les domaines des sciences de la vie, de la chimie…. Il constitue un véritable outil scientifique, et ce dès le commencement d'un projet. Les cahiers de laboratoire répondent également aux obligations légales et contractuelles : en apportant la preuve de l'invention, de ses inventeurs et ils concourent à l'intégrité scientifique.

En 2007 un cahier de laboratoire national (figure 5) a été mis en place en lien avec les préconisations du Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche (MENESR) sur la propriété intellectuelle de façon, d'une part, à renforcer l'esprit d'appartenance à la recherche publique française, et d'autre part, à donner une image d'excellence et de confiance de la recherche française.

Figure 5 - Cahier de laboratoire national

Les plaquettes du réseau CURIE « Le cahier de laboratoire national : Pourquoi l'utiliser ? » et « Le cahier de laboratoire national : Comment l'utiliser ? » présentent des recommandations sur sa bonne gestion.

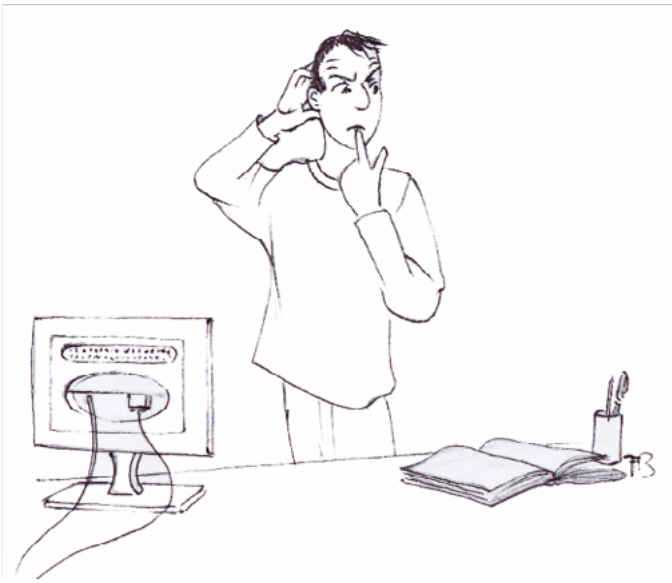
De nos jours, l'outil papier apparaît de moins en moins adapté aux pratiques de la recherche compte tenu de l'évolution de la nature et du volume des données produites (figure 6). En effet, la recherche connaît actuellement une croissance extrêmement forte de ses données numériques dans tous les secteurs d'activité : dématérialisation de nombreuses activités administratives, production massive de données par les équipements scientifiques, etc. Les unités de recherche sont confrontées à une déferlante d'octets qui nous fait entrer dans l'ère du « Big Data ». Dès lors, il apparait nécessaire que les outils de traçabilité accompagnent cette évolution via la numérisation des cahiers de laboratoire électroniques.

Figure 6 - Utilisation du cahier de laboratoire papier

2.3 - Numérique versus papier

Les principales limites de l' « outil papier » sont :

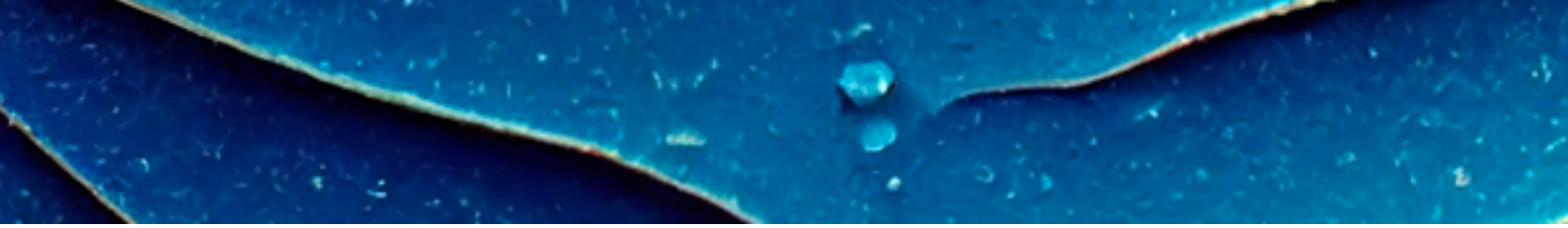
- Sa structure libre (pas de modèle de saisie…)
- Sa relecture (écriture manuscrite, langue…)
- La difficulté pour rechercher, partager et diffuser l'information
- La complexité pour insérer des formats numériques
- Sa conservation (gestion et suivi des cahiers)



Les plaquettes du réseau CURIE « Le cahier de laboratoire national : Pourquoi l'utiliser ? » et « Le cahier de laboratoire national : Comment l'utiliser ? » présentent des recommandations sur sa bonne gestion.

Les apports du numérique sont en effet multiples et améliorent la traçabilité des recherches, l'intégrité scientifique et la gestion des données à travers par exemple, le partage de l'information avec un rattachement des données brutes, une recherche d'informations facilitée et une datation des expériences par l'horodatage. Cette transition a été longtemps ralentie du fait que le cahier de laboratoire électronique ne répondait pas juridiquement en matière de preuve de l'invention et de ses inventeurs.

Dès lors, voici une liste des avantages et des inconvénients d'un cahier de laboratoire électronique.



Avantages :

- Assure une sauvegarde numérique (qui scanne son cahier papier ?)
- Centralise les données (tableurs, schémas, images, spectres, CR, publications…)
- Compile des données simplifiées (pdf, zip, export…)
- Permet un accès sécurisé à distance possible via VPN (collaboration, missions…)
- Offre un puissant moteur de recherche (titre, date, étiquette, corps de texte, code…)
- Permet une datation des expériences par l'horodatage
- Facilite l'encadrement (suivi à distance des manips, verrouillage…)
- Facilite le travail collaboratif (base de données commune, expériences partagées, regroupement par projet…)

Inconvénients :

- Nécessite un accès facile à un support numérique et à une source d'énergie (cahier de « terrain »)
- Perte de « l'objet livre » et de l'agilité manuscrite
- N'offre pas de garantie sur la pérennité des formats et supports intégrés
- Nécessite de prévoir un hébergement des données

Figure 7 - Le cycle de vie de la donnée (Atelier-Données, 2021)

La Science Ouverte et les principes de FAIR data se traduisent par le cycle de vie de la donnée (figure 7) qui se décompose en différentes étapes. Le cahier de laboratoire électronique intervient pendant toute la phase de réalisation interne du projet de recherche, de la phase 2 (Concevoir – Planifier) à l'étape 5 (Réaliser – Analyser). Il est, de ce fait, une réponse pertinente aux enjeux de gestion des données et de reproductibilité des expériences

.

Le cahier de laboratoire électronique va permettre de fournir les métadonnées nécessaires pour apporter toutes les informations utiles à la description des données brutes elles-mêmes (libellés des paramètres, unités de mesure, localisation, étiquettes, propriétaires, etc.), ainsi que sur les dispositifs d'acquisition (équipements, capteurs de mesures, modèles numériques…). Les métadonnées, étymologiquement « les données sur les données », représentent toute l'information sur le document, information qu'il va falloir conserver avec le document pour en assurer l'intelligibilité dans le futur

3. LES INITIATIVES NATIONALES

L'objectif n'est pas d'être exhaustif dans la présentation des études et expérimentations sur la mise en place de cahiers de laboratoire électroniques dans le domaine de la recherche mais d'aborder quelques initiatives significatives. Nous nous limiterons donc :

- À l'expérience de mise en place du logiciel Labguru au sein de l'INSERM
- Au projet dattacc.org porté depuis 2019 par les bibliothèques universitaires de Lyon et Grenoble
- Aux travaux portés par le groupe de travail du CNRS
- A la collaboration du CNRS avec le Comité pour la Science Ouverte (CoSO)

3.1 - INSERM

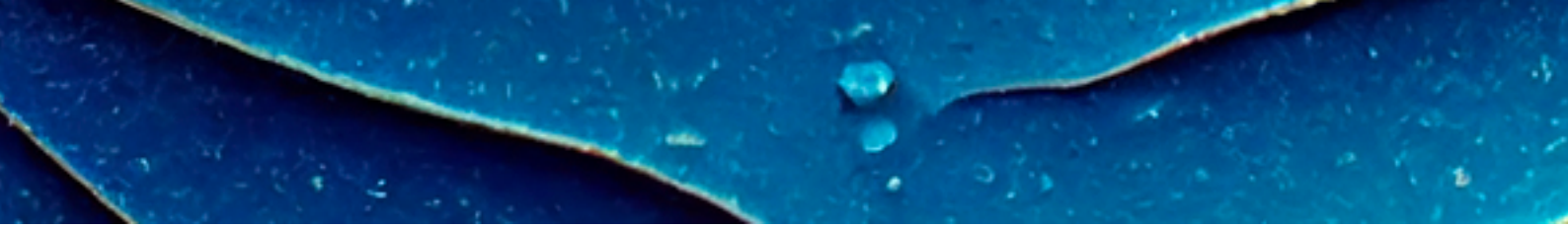
L'INSERM s'est intéressé dès 2013 à la version numérique des cahiers de laboratoire, comme une réplique du cahier papier (Dupré, Gallina-Muller, 2019). Après une phase d'expérimentation d'un CLE menée entre 2015 et 2018 sur une trentaine de laboratoires destinée à recueillir les besoins, avis, perspectives d'évolutions ainsi que les difficultés de mise en œuvre, la solution Labguru de l'éditeur Biodata, dédiée à la biologie a été retenue dans le cadre d'un appel d'offres. L'outil est proposé depuis décembre 2018 à l'ensemble des laboratoires INSERM.

L'INSERM conclut que si la version électronique reste une solution d'enregistrement au quotidien des expériences scientifiques, c'est désormais devenu un outil différent, fortement axé sur la qualité, la gestion de la connaissance, la gestion de projets et le travail collaboratif tout en prenant également en charge la gestion des inventaires et des équipements.

3.2 - DATACC

Le projet DATACC est un dispositif d'accompagnement sur les données de recherche en physique et chimie porté par les bibliothèques de Lyon 1 et de Grenoble Alpes Université. En tant que dispositif d'ACCompagnement sur les DATA, ce projet s'inscrit dans la dynamique des projets soutenus par le Groupement d'intérêt scientifique CollEx-Persée.

Ce projet vise à fournir une réponse opérationnelle aux exigences croissantes en matière de science ouverte avec, en autres, le déploiement d'un projet-pilote sur les cahiers de laboratoire électroniques.
Le site fournit ainsi des contenus nourris sur les cahiers de laboratoire électroniques, issus d'une expérimentation menée avec des chimistes de Lyon 1 et de Grenoble, assortis de bonnes pratiques sur leur utilisation (DATACC).



3.3 - CNRS

Suite à une réflexion menée par un groupe de travail initié en 2020, le CNRS a choisi de proposer une offre de service basée sur eLabFTW. Une note de préconisation du CNRS, adressée en juin 2022 à l'ensemble des unités, présente à la fois cette démarche, le choix réalisé et des recommandations générales à respecter.

Dans un premier temps, une réflexion sur la mise en place de cahiers de laboratoire électroniques suite aux besoins remontés par les agents en laboratoire, en alternative au cahier de laboratoire national (format papier) a été lancée. L'objectif de cette réflexion était d'identifier des solutions électroniques adaptées aux pratiques de la recherche en prenant en compte la diversité des activités du CNRS tout en assurant la mémoire, la traçabilité, la sécurité, la confidentialité et la pérennité des résultats de recherche et tout en respectant les exigences liées à la Science Ouverte et à la protection du patrimoine scientifique et technique de l'établissement.

A la demande conjointe du DGD-S et du DGD-R, un groupe de travail a été mandaté pour se pencher sur cette réflexion réunissant les expertises suivantes :

- Scientifique avec les différents Instituts du CNRS
- Juridique avec la Direction aux Affaires Juridiques
- Qualité avec le réseau Qualité en Recherche
- Systèmes d'informations avec la Direction des Systèmes d'Information
- Science ouverte avec la Direction des Données Ouvertes de la Recherche, les correspondants données et information scientifique et technique des instituts
- Sécurité avec la Direction de la sûreté, le responsable de la Sécurité des Systèmes d'Informations et la Déléguée à la Protection des Données

Ce projet a été mis en place avec une double approche :

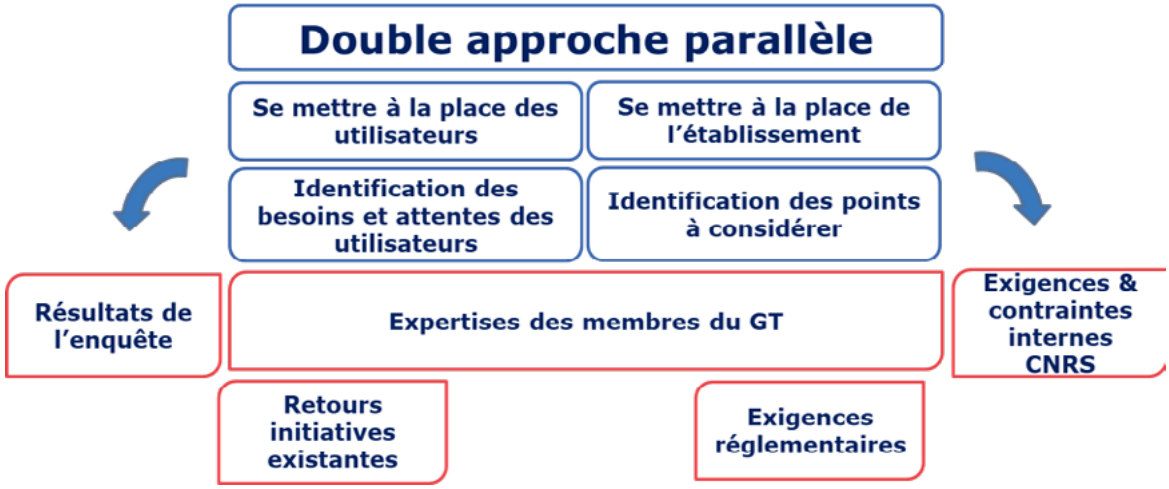


Figure 8 - L'approche CNRS sur les CLE

La première action de ce groupe a été de déployer une enquête auprès de l'ensemble des unités de recherche impliquant le CNRS soit 1250 unités, destinée à répondre aux objectifs suivants :

- Informer la communauté de la réflexion en cours
- Inventorier le nombre de personnes potentiellement intéressées par les versions électroniques des cahiers de laboratoire
- Identifier les solutions déjà utilisées et leur nombre d'utilisateurs, les fonctionnalités souhaitées, la variété et la volumétrie de données et des résultats de recherche à stocker, les besoins d'encadrement, de formation et de support ainsi que les points éventuels d'inquiétude ou de vigilance quant à l'utilisation d'un tel outil

Cette enquête de 19 questions a été diffusée aux directeurs d'unités en juin 2020. Les répondants pouvaient le faire au titre de leur unité, de leur équipe ou de leur propre utilisation personnelle. L'enquête, clôturée le 4 septembre 2020 comptabilise 1072 réponses complètes. Les résultats de cette enquête sont accessibles sur l'intranet du CNRS (Léon, Libri, 2021). Nous présentons simplement brièvement les avantages attendus et les craintes qui ressortent de cette enquête (figure 9).

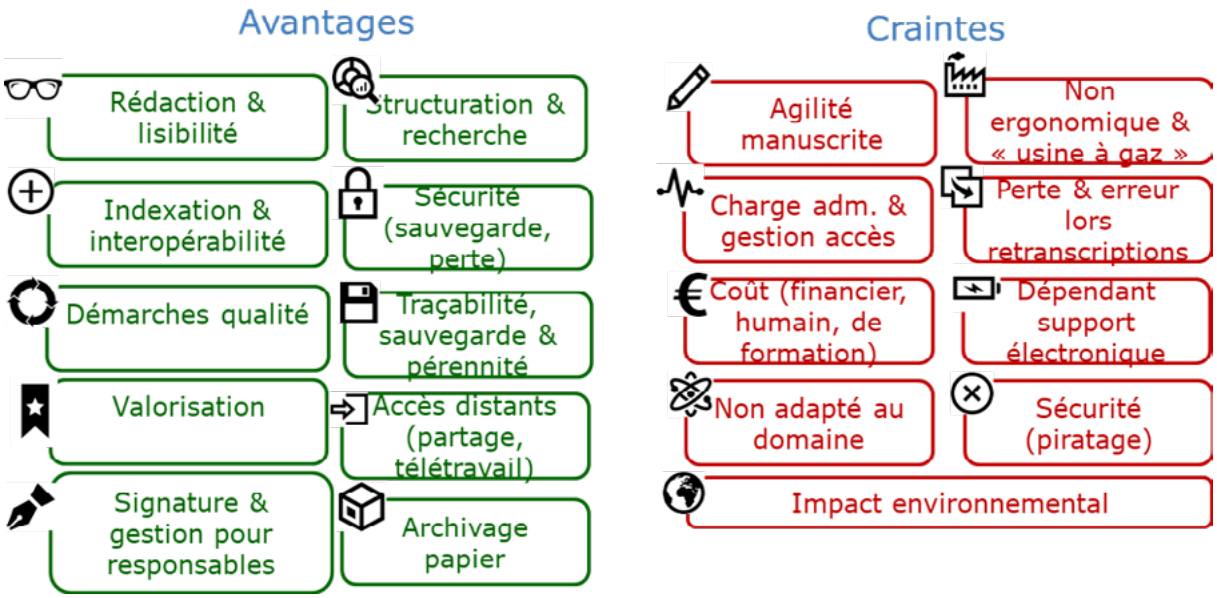
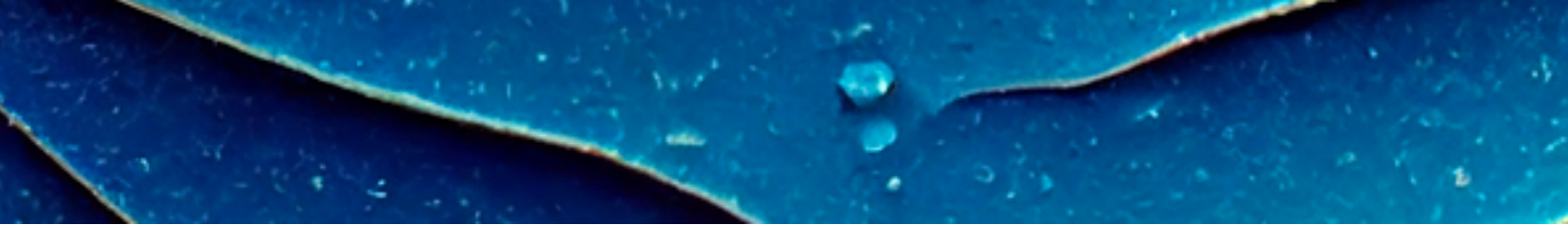


Figure 9 - Principaux avantages et craintes issues de l'enquête du CNRS

Ce projet est un des huit projets USER FIRST, lauréats du Fonds pour la Transformation de l'Action Publique (FTAP) pour un budget 1,532 M€ à jalonner jusqu'en 2023 et permettant de financer les coûts d'investissement et de mise en œuvre.



3.4 - Comité pour la science ouverte (CoSO)

Le Comité pour la science ouverte (CoSO) rassemble des collèges et des groupes d'expertise permanents et conduit des projets. Ils concourent à la mise en œuvre des mesures du Plan national pour la science ouverte et à leur déclinaison par thématique. Le CNRS a ainsi intégré le groupe de travail du CoSO sur les Cahiers de Laboratoire Électroniques.

« Le rapport du groupe de travail « Cahier de laboratoire électronique » (ELN) présente une vision partagée sur la définition, le cadrage, les usages et le périmètre fonctionnel de l'ELN, qui doit pouvoir s'intégrer dans les environnements informatiques et institutionnels existants. Il émet un ensemble de recommandations sur les critères de choix d'un outil et intègre une liste comparative d'outils existants. » (Ouvrir la science, 2021).

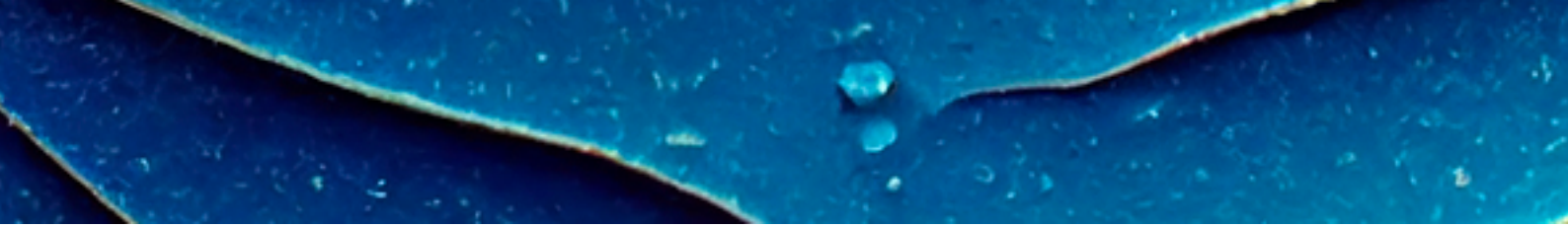
Ce groupe de travail définit les cahiers de laboratoire comme des outils scientifiques et juridiques qui permettent de garantir la traçabilité des expériences de laboratoire, de maintenir la continuité du cheminement intellectuel aboutissant aux résultats scientifiques, d'identifier les contributions de chacun et d'établir les dates d'acquisition des résultats. Ils contribuent à la reproductibilité et au caractère cumulatif des connaissances scientifiques et s'inscrivent donc pleinement dans une démarche de science ouverte.

4. CARACTÉRISTIQUES FONCTIONNELLES ET TECHNIQUES

Voici les caractéristiques fonctionnelles et techniques auxquelles devraient répondre les cahiers de laboratoire électroniques. Celles-ci, validées par le CoSO, sont présentées dans le tableau suivant :

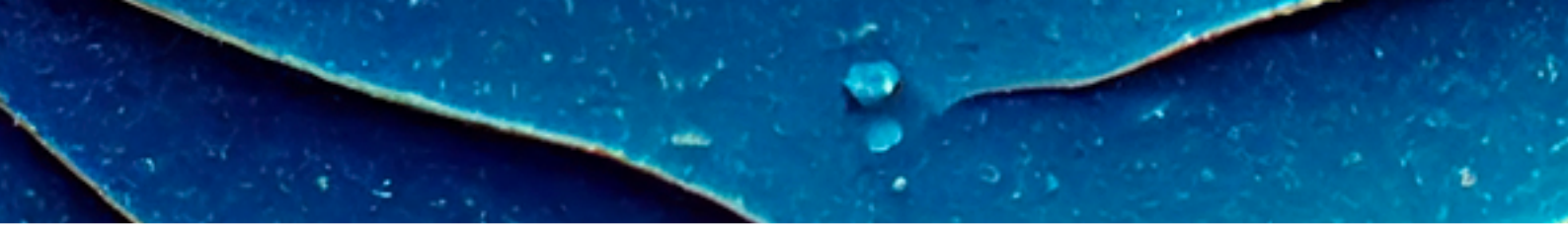
Fonctionnalités
Description des expériences et résultats
Possibilité de dupliquer des expériences (sans les résultats de recherche), comme un modèle/template pour une nouvelle expérience Description des protocoles
Possibilité de dupliquer des protocoles, comme un modèle/template pour un nouveau protocole
Gestion de pièces jointes : bureautiques (PDF, audios, vidéos, etc.) ou scientifiques et techniques (git, chromatogrammes, spectres, images, etc.)
Possibilité d'exporter des éléments de l'ELN à un format imprimable
Accès à des données externes (c'est-à-dire non stockées dans l'ELN), brutes ou non, pour lecture et référencement
Structuration avec moteur et filtre de recherche interne ; intégration de mots-clés (tags)
Contenu accessible (pour fouille/mining)
Possibilité d'associer des métadonnées aux jeux de données manipulés, qu'il s'agisse de métadonnées génériques (dates, auteurs, etc.) ou spécifiques (ontologies, taxonomies, etc.)
Utilisation et maintenance des équipements (y compris remontées d'alertes)
Gestion des stocks (réactifs, lignées biologiques, etc.)
Outils pour des calculs simples (fonctions mathématiques de base, conversion d'unité, etc.)
Outils pour des calculs complexes (statistiques ou analyses nécessitant un langage de programmation)
Outil de construction/dessin de molécules
Notes sur les expériences ou le protocole (pendant l'expérimentation), comptes rendus, etc.
Possibilité de citer des contenus issus de bibliographies (type labtex.bib Text)
Accès aux outils bureautiques du poste de travail
Gestion des outils logiciels et de codes informatiques (intégration continue, déploiement continu, gestion des versions, etc.)
Gestion du partage par les utilisatrices et utilisateurs
Agenda partagé
Edition partagée de documents (en ligne et en simultané par plusieurs utilisatrices ou utilisateurs)
Annotation de contenu et ajout de commentaires

Tableau 1 - Liste des fonctionnalités retenues pour les cahiers de laboratoire électroniques (Ouvrir la science 2021)



Le rapport du CoSO présente également la liste des caractéristiques nécessaires :

Catégorie	Caractéristique
Protection des données	Respecter le RGPD pour les données personnelles des utilisatrices et utilisateurs
	Respecter le RGPD pour les données de recherche : données personnelles des participant-e-s à la recherche si elles sont contenues dans les CLE
	Avoir un stockage maîtrisé, redondant et sauvegardé (sites distants)
Intégrité	Respecter des contraintes réglementaires ou législatives concernant la localisation de l'hébergement des données de recherche
	Permettre la signature des documents (hachage)
	Permettre l'horodatage des actions et des documents
	Identifier de manière unique les entrées (protocole, expérience, note, etc.) – UUID
	Assurer un versionnement/suivi des modifications
	Permettre d'implémenter des workflows prédéfinis (par exemple, suite d'actions obligatoires dans la description d'une expérience ou d'un résultat)
Authentification, sécurité	Permettre une signature électronique des actrices et acteurs
	Utiliser un système d'authentification sécurisé
	Permettre de définir les règles de gestion des utilisatrices et utilisateurs : niveau d'autorisation d'accès selon rôles et projets et cycle de vie (départ/arrivée agent, ouverture/fermeture projet, etc.)
Sauvegarde des données	Chiffrer le transport des données (TLS, algorithmes conformes RGS)
	Chiffrer les données
	Intégrer des sauvegardes régulières et des restaurations
	Définir les durées de conservation et modalités d'archivage ou destruction des données
Confidentialité	Assurer la confidentialité des données exportées pour sauvegarde ou archivage.
	Préserver la confidentialité des informations pouvant constituer un savoir-faire valorisable ou relever du secret des affaires (préserver la nouveauté en cas d'invention brevetable)
	Préserver la confidentialité des informations échangées avec un-e partenaire dans le cadre d'une collaboration



Compatibilité, interopérabilité	Être compatible avec les différents systèmes informatiques (Windows, Linux, etc.) pour la partie serveur
	Permettre la gestion (Import/Export) des données sous un format structuré (par exemple en cas de changement de logiciel ou d'évolution administrative)
	Permettre l'intégration avec un système d'information local (bases de données, annuaires, etc.)
	Permettre les liens avec des entrepôts de données
	Utiliser, quand ils existent, les standards ouverts des données en entrée ou en sortie pour permettre la ré-exploitation des informations de façon pérenne
	Permettre les échanges de données en proposant des services d'intégration et d'exposition, si possible via des Web-Service ou API REST
Usage, utilisation	Pouvoir installer facilement les mises à jour et évolutions, en garantissant l'accessibilité, dans le cas de l'utilisation à partir d'un navigateur Web
	Être ergonomique et intuitif
	Permettre la modularité/personnalisation/flexibilité
	Disposer des références documentaires (guide, tutoriels, consignes d'utilisation, etc.)
	Avoir un code source accessible pour contrôler, si nécessaire, l'évolution de l'outil, sa maintenance et surtout les données qu'il contient
	S'appuyer sur une communauté d'utilisatrices et d'utilisateurs
	Proposer l'aspect multilingue (a minima l'anglais)
	Permettre une utilisation hors-ligne avec synchronisation a posteriori

Tableau 2 - Liste des caractéristiques retenues (Ouvrir la science 2021)

Le groupe de travail du CoSO a analysé plusieurs outils dont :

- eLabFTW, solution open source
- Labguru, solution commerciale, choisie par l'INSERM
- Chemotion, solution open source
- LabCollector, solution commerciale adoptée par des laboratoires du CEA

5. MISE EN PLACE D'UN CLE AU SEIN D'UNE UNITÉ DE RECHERCHES

Pour la suite de ce guide, nous avons fait le choix de l'outil eLabFTW, sur lequel se base le CNRS pour construire son offre de service, pour illustrer l'utilisation d'un cahier de laboratoire électronique.

Le cahier de laboratoire électronique n'est pas une stricte réplique informatique du cahier papier. Il va favoriser une meilleure gestion des données, une meilleure description des expériences, une valorisation des résultats facilitée, un accès distant aux données, le travail collaboratif mais en contrepartie, va nécessiter le management des équipes adapté et des projets.

Dès lors, il convient d'anticiper la mise en place d'un cahier de laboratoire électronique au sein d'une structure. Cela va nécessiter, d'une part, de réfléchir à la conduite du changement, à l'organisation humaine et à l'organisation des expériences à mettre en place et d'autre part, à la gestion des données externes. La mise en place d'un CLE doit constituer un véritable projet de l'unité et, de ce fait, être suivi en tant que tel.

5.1 - Conduite du changement

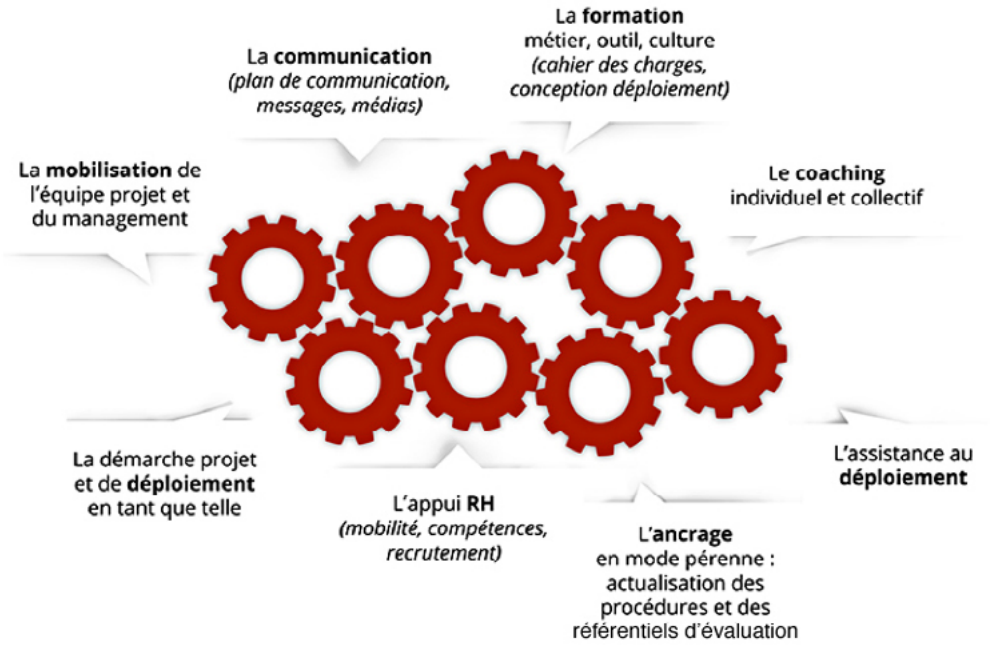


Figure 10 - Les facteurs de succès d'un déploiement de solution logicielle (repris du livre «Pro en Conduite du changement», Julie Ricou et Valérie Moissonnier,2018, Ed. Vuibert)

La question qui se pose dans ce type de projet est : comment accompagner le changement avec l'arrivée d'un nouvel outil amenant une nouvelle organisation et finalement une nouvelle culture ?

Car effectivement, la mise en œuvre de cahiers de laboratoire électroniques amène à se demander comment accompagner le changement lié à un nouvel outil numérique mais aussi à une nouvelle organisation due au format numérique de l'outil, amenant à une exploitation et une gestion différente d'ouverture de cahiers, d'autorisation d'accès, de gestion des données, etc

.

Cette transformation vers le numérique implique de nombreux changements d'ordre organisationnel, culturel et surtout opérationnel (figure 10).

Les impacts relevés nécessitent un accompagnement en termes :

- De communication pour comprendre le sens global de la démarche, les bénéfices apportés aux utilisateurs et au CNRS plus globalement
- De formation pour que chaque collaborateur puisse facilement prendre en main le nouveau process, le nouvel outil modifiant leurs usages du cahier jusqu'ici au format papier
- D'accompagnement afin que les principaux utilisateurs de la solution puissent s'approprier et promouvoir toutes les bonnes pratiques et que les responsables adoptent également la solution comme outil de traçabilité et de collaboration

Les actions associées doivent être un juste équilibre entre :

- De la communication ciblée à plusieurs niveaux afin de sensibiliser au sens global du projet, convaincre des bénéfices du projet et promouvoir les bonnes pratiques
- Des formations aux utilisateurs sous forme de tutoriels vidéos, ateliers, sessions de retours d'expérience, etc.
- De l'accompagnement technique : installation, utilisation et documentation (manuel utilisateur)
- De la création de supports : guide opérationnel, forum, newsletter, espace CORE

5.2 - Pilotage du CLE

Sur le plan opérationnel, cette conduite du changement peut se décliner selon différentes étapes.

Au préalable, il serait intéressant que la mise en place d'un CLE soit annoncée en conseil de laboratoire/ scientifique comme un projet de l'unité et validé par celui-ci. Un plan de déploiement pourra ainsi être présenté et validé.

Il est de la responsabilité du directeur d'unité de s'assurer, avec l'appui des personnels compétents et en particulier, du chargé de SSI (sécurité de Systèmes d'Information) et du chargé de PPST (protection du potentiel scientifique et technique) de son unité ou de sa délégation, du respect des préconisations d'implémentation de l'établissement.

Un pilote de projet pourra être nommé par le directeur avec les rôles suivants :

- Présenter les éléments essentiels d'organisation liés au CLE
- Être le point d'entrée des informations du CLE
- Assurer l'ouverture et le suivi des CLE de l'unité (disposer d'indicateurs sur l'utilisation effective du CLE)

Le pilote de projet pourra présenter l'outil à l'ensemble des personnels de l'unité lors d'un séminaire avec une présentation du guide et des règles d'utilisation du CLE. Il pourra aussi fournir les documents d'aide pour l'organisation et l'utilisation.

Il pourra aussi aider à déployer le CLE dans les équipes (voir les détails dans l'annexe 2).

5.3 - Organisation humaine

Le CLE étant un outil collaboratif, il est indispensable de s'organiser pour contrôler les interactions en créant par exemple une zone de travail. Pour cette zone du cahier de laboratoire électronique, il est conseillé de bien définir son périmètre. En général, les « zones de travail » ou « zones cahier » étant imperméables entre elles, il est indispensable de se poser quelques questions pour que toutes les personnes impliquées puissent collaborer comme elles le souhaitent.

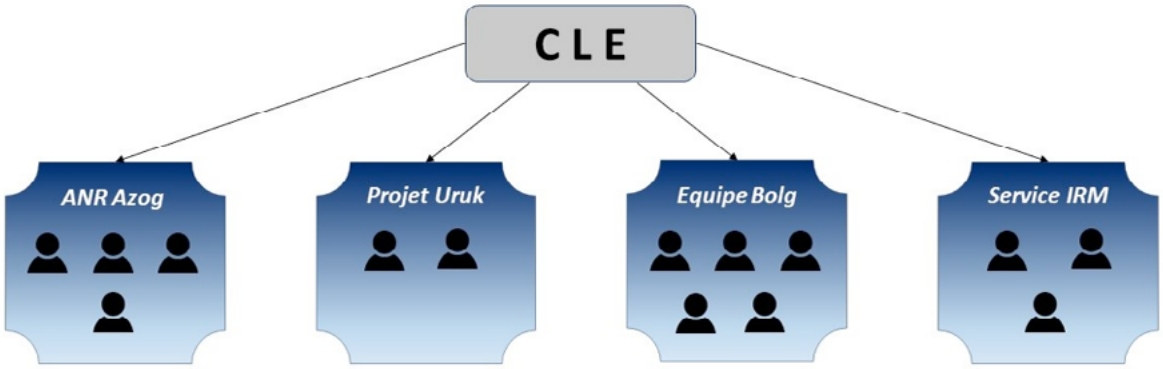


Figure 11 - Exemple de « zones cahier » d'un CLE

Est-ce une zone réservée à un projet ou une zone inter utilisateurs ?

Il est possible de choisir de créer une « zone cahier » avec toute une équipe (intérêt pour créer du partage d'informations générales) mais en général il est préférable de limiter l'espace à une thématique ou un projet.

Par exemple, regrouper toutes les personnes participant à un projet ANR permet de garder la confidentialité des données, de centraliser au même endroit toutes les informations du projet : expériences, comptes-rendus de réunion, rapports, etc., en les rendant disponibles à tous facilement.

Une « zone cahier », centrée sur une thématique, est également intéressante pour regrouper doctorants, post-doc et stagiaires travaillant sur un même sujet. L'encadrement dans ce cas est facilité et le partage d'information entre étudiants devient fluide.

Dans le cas d'un projet sensible ou confidentiel, la possibilité de créer une zone réservée à une seule personne garantit la sûreté des informations (aucun partage possible).

Lors de sa connexion au CLE, il convient d'entrer dans une zone définie et organisée qui permet d'avoir une séparation des données en fonction des projets et ainsi faciliter l'encadrement et l'organisation.

Qui doit avoir accès à la « zone cahier » ?

Toutes les personnes souhaitant collaborer entre elles. Le principe d'une « zone cahier » est de faciliter le partage d'informations et de permettre l'encadrement. Il faut donc au départ déterminer quelles sont les personnes qui travailleront ensemble et quelles sont celles qu'il faut exclure.

Quels droits d'accès pour qui ?

De manière générale, il est choisi un « administrateur » unique par zone qui se charge des aspects techniques (configuration, accès, droits).

Les étudiants et stagiaires ont des droits limités ; globalement il est conseillé de leur donner des droits ne concernant que leurs propres données.

Les encadrants ont des droits supplémentaires permettant d'intervenir sur les données de leurs étudiants.

Cas particulier : quand une zone de cahiers de laboratoire est créée pour un service par exemple, il peut être envisagé de donner à tous les membres du service les mêmes droits pour faciliter le travail collaboratif : expériences communes, gestion des bases de données et des métadonnées communes.

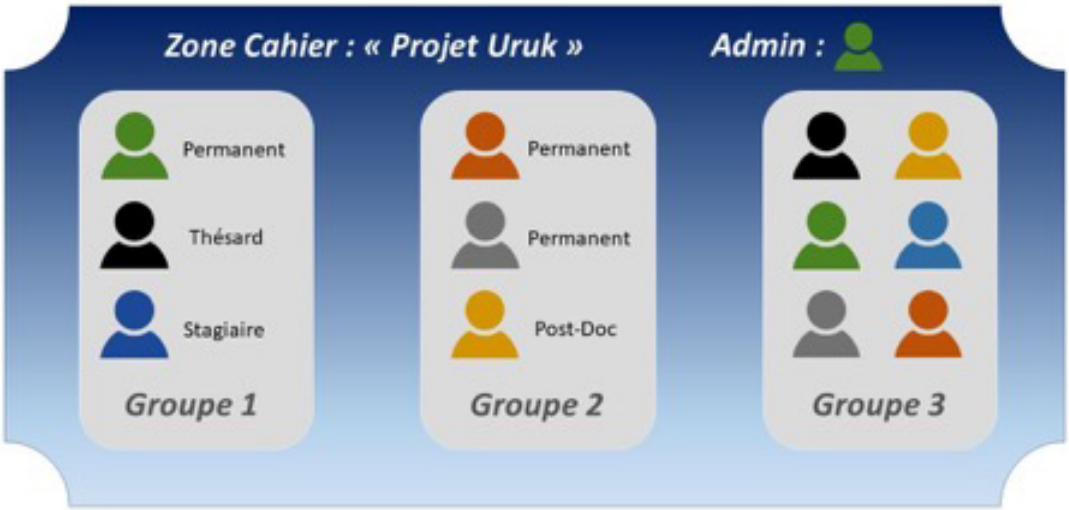


Figure 12 - Exemple de composition d'une "zone cahier"

Au sein d'une même « zone cahier », différents groupes peuvent être créés pour faciliter l'encadrement, le travail collaboratif et bien gérer les expériences (permissions, partage, etc.).

5.4 - Création des zones de travail collaboratif

Une fois l'organisation humaine définie, les différentes « zones cahier » nécessaires sont à élaborer.

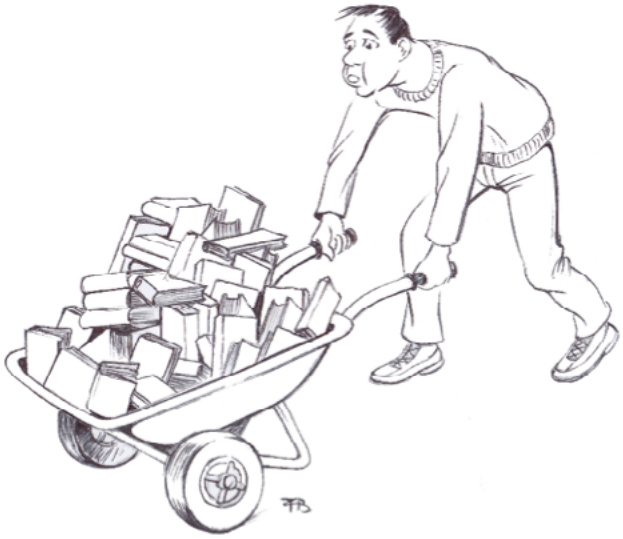
Il est conseillé de planifier, avec rigueur, la formation des personnels à l'utilisation de base du CLE pour bien démarrer. Dans un premier temps il peut également être décidé que la « zone cahier » créée est un « bac à sable » pour apprivoiser l'outil et faire des tests pendant quelque temps (encadrant vs étudiants, partage de documents, horodatage, etc.).

Une phase « d'apprentissage » peut permettre de rassurer les utilisateurs. Certains encadrants peuvent préférer utiliser le CLE au départ pour encadrer uniquement un stagiaire de manière à assurer leur maitrise du logiciel avant de créer une zone avec plus d'intervenants

6. GESTION DES DONNÉES

6.1 - La gestion des données externes

Par données externes, nous faisons référence aux données associées aux différentes expériences mais qui ne sont pas stockées dans le CLE : données brutes, volume important de données, etc.

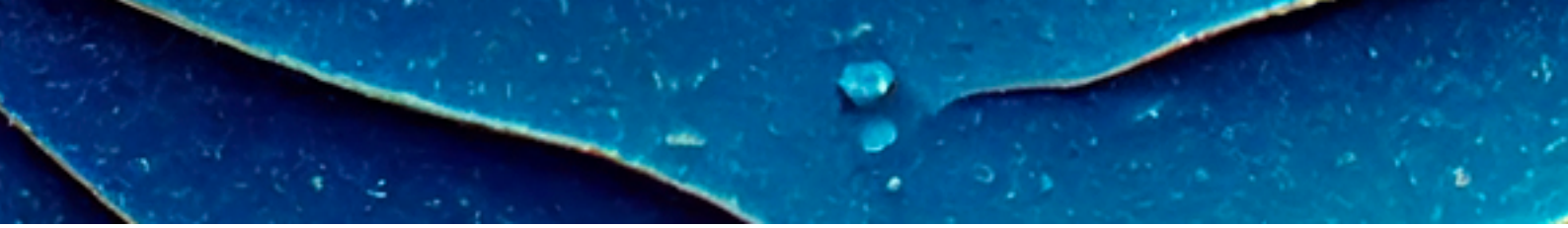


Comme nous l'avons mentionné, le cahier de laboratoire électronique, s'il permet d'intégrer divers documents (textes, images…), ne permet pas, à l'heure actuelle, le stockage en son sein de grands volumes de données que constituent les résultats des expériences. Dans ce cas, il conviendra, dans le CLE, d'intégrer un lien avec les données stockées séparément (espace de stockage, entrepôt de données…).

Comme le rappellent Dupré et Gallina-Muller (2019), « Il est essentiel que les ELN (Electronic Laboratory Notebook ou CLE en français) intègrent rapidement le besoin de stockage capacitif. Ce besoin s'exprime en pétaoctets et se heurte techniquement pour l'instant à des problématiques réseau : comment, en effet, peut-on récupérer et centraliser les données brutes, le plus souvent générées sur les sites lors des phases d'acquisition et de traitement ? ».

La conservation des données brutes constitue un enjeu majeur de traçabilité et de reproductibilité des expériences. De ce fait, il convient de définir et mettre en place une organisation des données brutes référencées dans le cahier de laboratoire électronique.

A cet effet, identifier efficacement les fichiers numériques en fixant des règles de nommage et de définir un plan de classement des dossiers sur les espaces de stockage référencés est essentiel.



6.2 - Définir des règles de nommage

Il est capital d'élaborer et de respecter des règles communes de nommage pour faciliter et pérenniser l'accès à l'information et optimiser le partage et le tri des documents. Chaque unité/laboratoire/service doit définir sa méthodologie, la décrire et la communiquer à ses agents.

Les intitulés des fichiers :

- Doivent être succincts et précis
- Doivent être uniques
- Peuvent être caractérisés a minima par une date, un sujet et un type de document
- Ne doivent pas excéder 31 caractères (extension comprise)

Il convient d'éviter :

- Les accents (é, è ë, ê), cédille (ç), caractères spéciaux (, ; . : ! ? * »%...@ &)
- Les mots vides : le, la, les, un, une, des, et, ou…
- Les dénominations vagues, par exemple « divers », « autres », « à classer »
- Les espaces pouvant être remplacés par des « _ » (tiret 8)

Elément	Exemples de règle	Exemples de nommage
Date	Date de création/modification du document, de l'événement. Format à l'américaine : AAAAMMJJ	20150108 2015_01_03
Type de document	Qualifie la nature du document. Utilisation préconisée d'abréviations clairement définies pour être accessibles à tous les agents. Toute abréviation sera en lettres majuscules.	PO (protocole) CR (compte-rendu) RA (rapport d'activités)
Titre ou sujet	Il s'agit du sujet principal traité au sein du document. Utiliser des noms communs, écrits en lettres minuscules.	commission analyse01 colloque
Version du document	Distingue les différentes versions d'un document. Peut être signalée par un « V » en majuscule suivi de 2 chiffres, VP version provisoire, VF version finale.	_VP1 _VF Peuvent aussi être utilisés les initiales de l'auteur puis des « relecteurs » jusqu'à la version finale.
Exemple	Classement chronologique par sujet : date, type, titre, signataire, _version	20150108_CR_commission_DT_VP1.doc

Tableau 3 - Eléments d'identification des fichiers

6.3 - Choisir un plan de classement

Il est important de disposer d'un plan de classement des dossiers numériques pour améliorer la recherche ultérieure de fichiers et ainsi :

- Gagner du temps et aider chacun à retrouver plus facilement les données
- Rationaliser le contenu du serveur en organisant les dossiers de manière ordonnée
- Faciliter les échanges entre utilisateurs
- Pérenniser les informations et faciliter leur transmission

Le plan de classement n'est pas figé et peut évoluer selon les données à enregistrer. Il doit respecter les principes de structuration et d'organisation suivants :

- Utiliser un serveur disponible, sécurisé et sauvegardé comme lieu de stockage et de partage des informations d'une unité/service
- Exclure les disques durs aussi bien du poste informatique que d'un disque dur externe (non sauvegardé) et des sessions personnelles (non partagées)
- Disposer d'une hiérarchie de répertoires et de dossiers avec des intitulés intelligibles par tous allant du général au particulier
- Choisir une organisation thématique aux niveaux supérieurs puis alphabétique et/ou chronologique pour les dossiers des niveaux inférieurs et les fichiers
- Limiter à cinq ou six niveaux hiérarchiques les répertoires et dossiers à partir de la racine du serveur

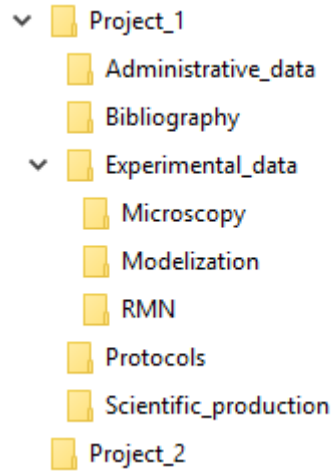
	– Un répertoire « administrative_data » regroupe les informations du projet (conventions, planning, budget, comptes rendus, rapports, etc.) – Un répertoire « bibliography » rassemble les références bibliographiques et fichiers PDF des publications étudiées – Un répertoire « experimental_data » présente les données (brutes et traitées) classées par méthode, technique d'analyse… – Un répertoire « protocols » recense les documents techniques, modes opératoires élaborés – Un répertoire « scientific_production » regroupe les présentations, communications, posters et publications réalisés

Tableau 4 - Exemple d'arborescence dans le cadre d'un projet

6.4 - Sauvegarde et archivage

Stocker, sauvegarder, archiver les données sont des phases essentielles d'une gestion rigoureuse des données, mais il n'est pas toujours aisé de faire la distinction entre ces notions.

Le **stockage** sert des usages de partage de fichiers et de documents. C'est généralement l'étape première qui consiste à déposer les données sur un support numérique pour les rendre accessibles. A ce stade, la donnée n'est ni sauvegardée, ni sécurisée.

La **sauvegarde** consiste à dupliquer des données à l'identique pour pouvoir les restaurer en cas de dommage ou de perte. Une sauvegarde est régulière et les données stockées sont régulièrement modifiées. Cette étape de sauvegarde doit s'accompagner d'une réelle politique qui détermine, en fonction de la criticité et de la sensibilité des données, combien de copies de sauvegarde sont établies par jour, par semaine, par mois. Les sauvegardes se font le plus souvent avec des logiciels spécialisés qui permettent de définir ce qui est sauvegardé et leur fréquence. Le logiciel permet également de restaurer, c'est-à-dire de rétablir les données d'une certaine sauvegarde choisie.

L'**archivage** a pour finalité de préserver vos données anciennes sur une longue durée notamment à des fins de référence. L'archivage consiste à rendre accessible en lecture des données immuables (données de mesures expérimentales, résultats de simulations couteuses à produire, etc.) bien que leur classification, leur format, puissent évoluer dans le temps.

Pour l'archivage, il convient de différencier l'archivage intermédiaire et pérenne. L'archivage intermédiaire est une obligation portant sur l'établissement qui doit mettre à disposition des unités un service d'archivage électronique (SAE) validé et conforme. Le CINES a été souvent retenu pour le CNRS (mission nationale confiée par la tutelle).

La problématique de l'archivage pérenne c'est-à-dire sur le long terme (supérieur à 10 ans) repose sur les formats de fichiers et leur capacité à être interprétés dans le futur ainsi que sur les supports utilisés. Pour cela il convient de privilégier des formats ouverts, non-propriétaires, d'un usage très répandu au sein d'une communauté de recherche. Quelques exemples de formats déconseillés et à privilégier sont publiés dans le tableau 5. Le Centre Informatique National de l'Enseignement Supérieur (CINES) qui est un acteur français dans le domaine de l'archivage pérenne des documents électroniques, dispose d'un service de validation de formats destiné à vérifier l'éligibilité de documents sur sa plate-forme d'archivage (https://facile.cines.fr/).

Format de fichier déconseillé sur le moyen et le long terme	Format de fichier à privilégier sur le moyen et le long terme
Excel (.xls, .xlsx) Word (.doc, .docx) Powerpoint (.ppt, .pptx) Photoshop (.psd) Quicktime (.mov)	Comma Separated Values (.csv) Texte ascii (.txt) ou PDF/A si formatage PDF/A (.pdf) TIFF (.tif, .tiff) MPEG-4 (.mp4)

Tableau 5 : Exemples de formats de fichiers

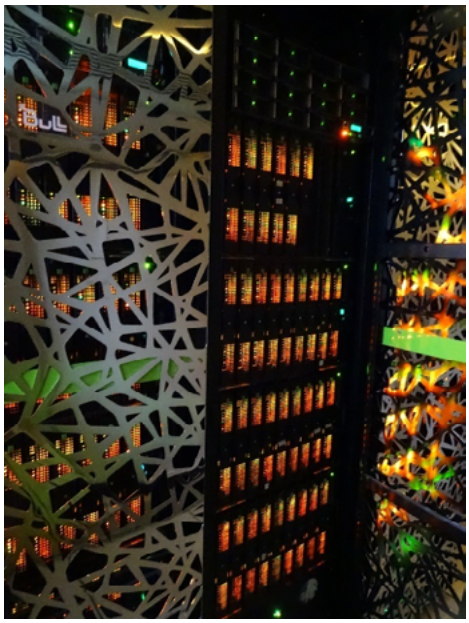
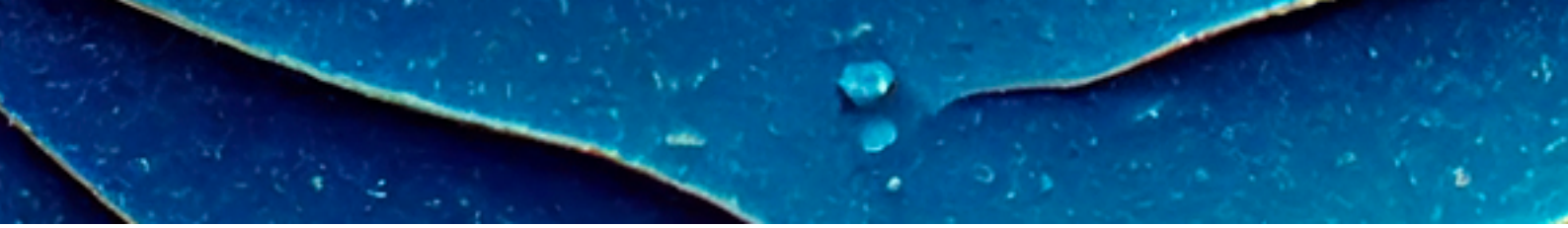


Figure 13 - Serveurs informatiques

Une étude portant sur 25000 ordinateurs a montré que 20% des disques durs cessent de fonctionner après 4 ans d'utilisation. La duplication des données par stockage redondant sur des supports différents de ceux de l'équipement utilisé (poste de travail fixe, mobile, serveur, etc.) est un des principes de base d'une bonne conservation. En général, il convient de privilégier un stockage centralisé, la règle du 3-2-1 étant généralement recommandée, c'est-à-dire 3 copies sur 2 supports différents dont 1 sur un lieu déporté. Chaque copie est protégée en confidentialité et intégrité au même niveau que le support original.



Concernant la sauvegarde, il est important d'établir une véritable politique, déterminant :

- La fréquence de sauvegarde
- La mise en lieu sûr des supports
- La sauvegarde des clés de chiffrement
- Des vérifications de restauration régulières des données

Concernant l'archivage, des recommandations sont disponibles pour assurer la pérennisation des données numériques à travers les normes NF-Z-42-013 et ISO 14721.

La question à se poser concerne les liens que pourraient avoir les entrepôts de données en tant qu'outil de gestion des données externes. Un entrepôt de données est un dépôt central informatique contenant des données décrites par un ensemble minimum de métadonnées (titre, licence, créateur, etc.) permettant leur identification (identifiant ou moyen d'accès), leur diffusion et leur réutilisation. Il convient toutefois de garder à l'esprit que vous n'êtes pas maître de la durée de préservation des données dans un entrepôt, cette information de durée dépend de la politique de l'entrepôt. Un entrepôt ne peut donc pas être considéré comme une véritable archive.

Il existe beaucoup d'entrepôts de données, certains sont des entrepôts institutionnels (DataSuds, Didomena, …), d'autres thématiques (Pangaea pour les données environnementales, Nakala pour les SHS, …) ou généralistes (Dryad, Zenodo, …). Des entrepôts spécifiques peuvent être suggérés (ou imposés) par la revue dans laquelle vous publiez, mais aussi par le financeur, le consortium du projet ou votre institution. Il est habituellement recommandé de déposer les données dans un entrepôt de données de confiance (certification CoreTrustSeal).

Une nouvelle plateforme nationale fédérée des données de la recherche, Recherche Data Gouv (Ouvrir la Science - Recherche Data Gouv : plateforme nationale fédérée des données de la recherche) a été mise en place au printemps 2022 sous l'égide du Ministère de l'enseignement supérieur, de la recherche et de l'Innovation (MESRI). Cette plateforme permet de favoriser le partage et l'ouverture des données produites par la recherche en fournissant aux chercheurs un entrepôt pluridisciplinaire pour le dépôt des données qui ne trouveraient pas place au sein d'un entrepôt thématique de confiance

6.5 - Le plan de gestion de données

Comme nous l'avons vu précédemment, la première étape pour la mise en place d'un cahier de laboratoire électronique implique de réfléchir à l'organisation de sa recherche et plus particulièrement des données utilisées et produites. Cette organisation des données peut être facilitée par l'utilisation d'un plan de gestion des données.

Le DMP (Data Management Plan) ou Plan de Gestion de Données (PGD) est un « document formalisé explicitant la manière dont seront obtenues, documentées, analysées, disséminées et utilisées les données produites au cours et à l'issue d'un processus ou d'un projet de recherche » (Cartier et al. 2018). De nombreuses documentations sur le sujet sont disponibles auprès de l'Institut de l'information scientifique et technique (INIST) qui a déployé DMP OPIDoR, une application web permettant de réaliser un plan de gestion de données en ligne à partir de divers modèles (ANR, etc.).

A titre d'exemple, la trame de plan de gestion de données (Deboin, 2018) présentée en annexe 1 est une adaptation du formulaire en ligne développé par le Digital Curation Centre. Cette trame a pour objectif de recenser toutes les informations nécessaires à la création, la fourniture, la maintenance, la conservation et la protection des données à travers le renseignement des points suivants :

- Informations administratives
- Collection de données
- Documentation et métadonnées
- Ethique, cadre légal
- Stockage, sauvegarde, sécurité
- Sélection et conservation
- Partage des données
- Responsabilités et moyens

En prenant exemple sur le plan de gestion des données en annexe et l'exemple d'organisation proposée dans ce document, certaines catégories sont communes. De manière inverse, la structuration de l'organisation dans le cadre d'un cahier de laboratoire électronique va nourrir le plan de gestion de données. Cela montre bien les interactions fortes entre ces deux outils. Voici ci-dessous un exemple de ces relations :

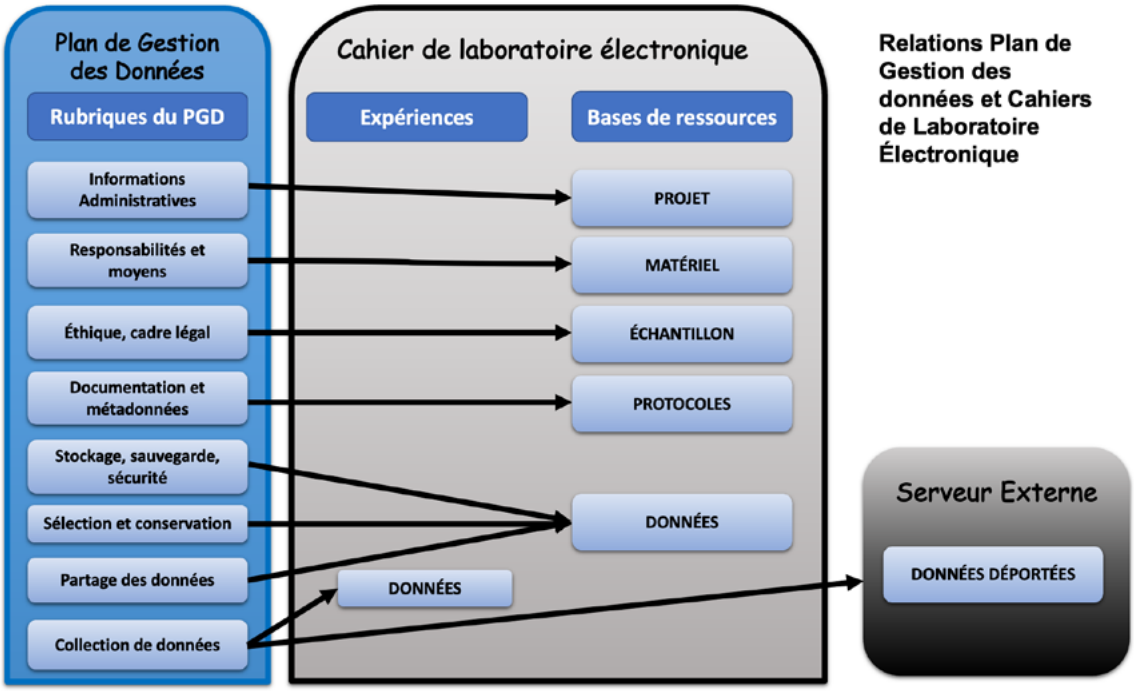


Figure 14 - Les connexions entre PGD et CLE

Bien sûr, ces liens sont très dépendants du type de recherche et du type de données mais ce qui est indispensable, dans tous les cas, est d'avoir réfléchi à l'organisation technique mais aussi humaine pour répondre au mieux aux critères de la science ouverte.

L'organisation à mettre en place peut se découper par exemple selon différents axes :

- Droits et accès au cahier de laboratoire qui correspond à la partie sécurité d'un DMP
- Projet correspondant aux « informations administratives »
- Gestion des données
- Etc.

Le point capital est de bien garder à l'esprit que l'organisation qui est mise en place sert de socle pour la structuration du cahier de laboratoire et du plan de gestion des données. L'un dépendra de l'autre et inversement. Par exemple, dans le plan de gestion de données sont définies les règles de nommage des fichiers qui sont utilisées dans le cahier de laboratoire électronique. De même, la description des expériences dans le CLE est la première source des métadonnées à associer aux données gérées par le PGD.

7. eLabFTW, UN EXEMPLE DE CLE

eLabFTW est un cahier de laboratoire électronique développé par Nicolas Carpi de l'Institut Curie (Paris). C'est un outil sous licence libre, le code source est donc accessible et modifiable. Il est, dès lors, possible d'utiliser eLabFTW pour ses propres besoins en le modifiant pour l'adapter à son usage. Il s'agit d'un outil généraliste, très personnalisable et qui bénéficie d'une communauté très active (mises à jour et améliorations régulières). C'est un outil Web, il est donc possible d'en installer sur des serveurs différents dits instances.

eLabFTW est utilisé par de très nombreux laboratoires en France, en Europe et dans le monde.
eLabFTW est un outil Web compatible avec de nombreux navigateurs, utilisable à partir d'un poste Windows, Mac OS, Linux, qu'il s'agisse d'un poste fixe, un téléphone ou une tablette, grâce à son interface responsive.

Le langage par défaut d'eLabFTW est l'anglais, mais il est possible de changer entièrement la langue de l'interface utilisateur en français, ainsi qu'une quinzaine d'autres langues.

7.1 - Les fonctionnalités

L'implémentation locale d'un tel outil doit obligatoirement être conforme aux instructions du CNRS, et en particulier respecter :

- **Les préconisations d'architecture technique et applicatives**
- **Les consignes de chiffrement du stockage**
- **Les consignes d'authentification renforcée reposant sur au moins deux facteurs**
- **De manière générale, les dispositions de la PSSI du CNRS.**

Gestion des expériences

La fonctionnalité principale d'eLabFTW est bien sûr le stockage et l'organisation des expériences. Chacune dispose d'un titre, d'une date et d'un champ texte, éditable grâce à un éditeur riche (l'édition en Markdown est également disponible).

À cela viennent s'ajouter les « tags », qui permettent une excellente organisation des entrées. Les « tags » correspondent à la notion de mots clés dans eLabFTW.

Il existe la possibilité d'ajouter des fichiers joints, des liens vers les bases de données, de définir des étapes, etc. À chaque expérience est également associé un statut et chaque équipe est libre d'éditer la liste des statuts disponibles.

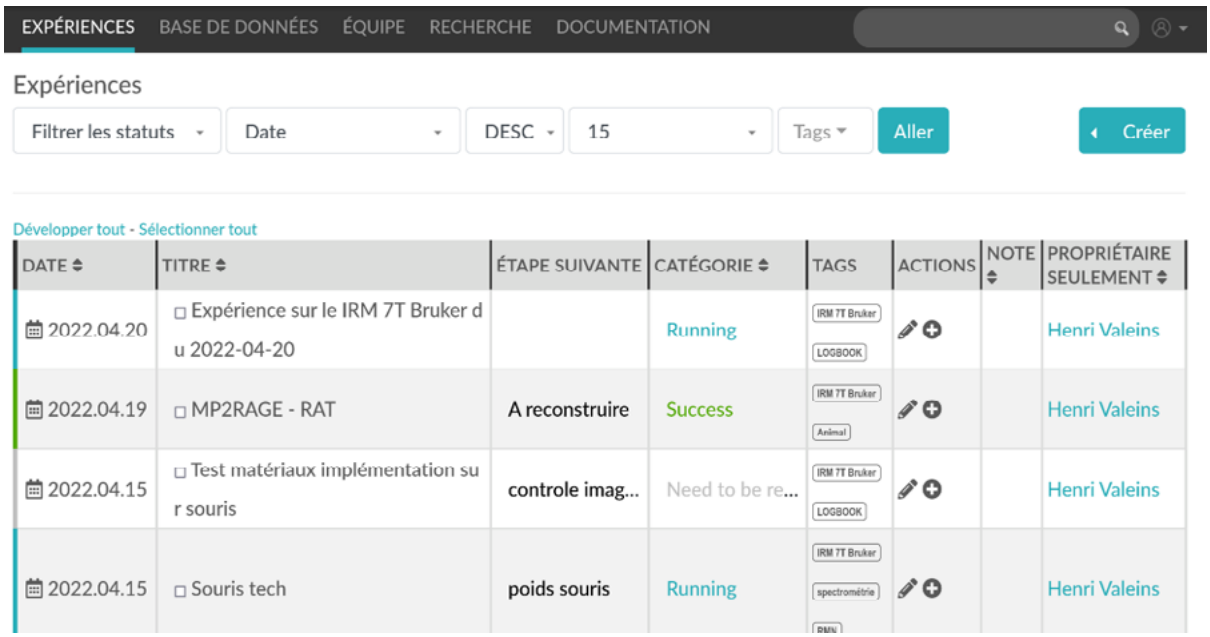


Figure 15 - La page d'accueil d'eLabFTW

Traçabilité des changements et horodatage

Les différentes versions successives d'une expérience sont enregistrées et visualisables. Il n'est pas possible de les supprimer. Une fois une expérience terminée, il est possible pour un administrateur de l'équipe (par exemple le chef d'équipe), de verrouiller ou « figer » l'expérience afin d'empêcher toute modification future. Cela correspond à la signature en bas de page du cahier de laboratoire papier.

Il est aussi possible d'horodater une expérience avec une Time Stamping Authority (TSA) qui supporte l'horodatage de confiance (Trusted Timestamping) défini par le standard RFC 3161.

Modèles d'expériences (templates)

Afin de faciliter la saisie des informations liées à une expérience, il est possible d'utiliser des modèles d'expériences, qui sont comme des squelettes d'expérience et permettent à l'utilisateur d'avoir des protocoles reproductibles et partageables au sein de l'équipe (ou même en dehors). Il est également possible d'intégrer différents modèles au sein d'une même expérience, ce qui permet de composer son expérience à partir de pièces existantes.

Gestion d'une base de ressources

Accolé à la gestion des expériences, il existe un item "Base de données" (figure 16) qui permet d'organiser et de référencer différents éléments. Il s'agit en réalité d'une liste de ressources qui peuvent être très variées car les catégories sont éditables par chaque équipe. Une équipe de biologistes gère ainsi probablement les catégories Anticorps, Plasmides et Lignées cellulaires, une équipe de chimistes nécessite peut-être Molécules, Solvants, Protocoles. Ces ressources peuvent également correspondre à divers équipements de mesure ou d'analyse, de protocoles expérimentaux auxquels peuvent se référer les différentes expériences. C'est un des points forts d'eLabFTW : sa grande flexibilité et modularité, qui permet à chaque équipe d'adapter l'outil à ses besoins.

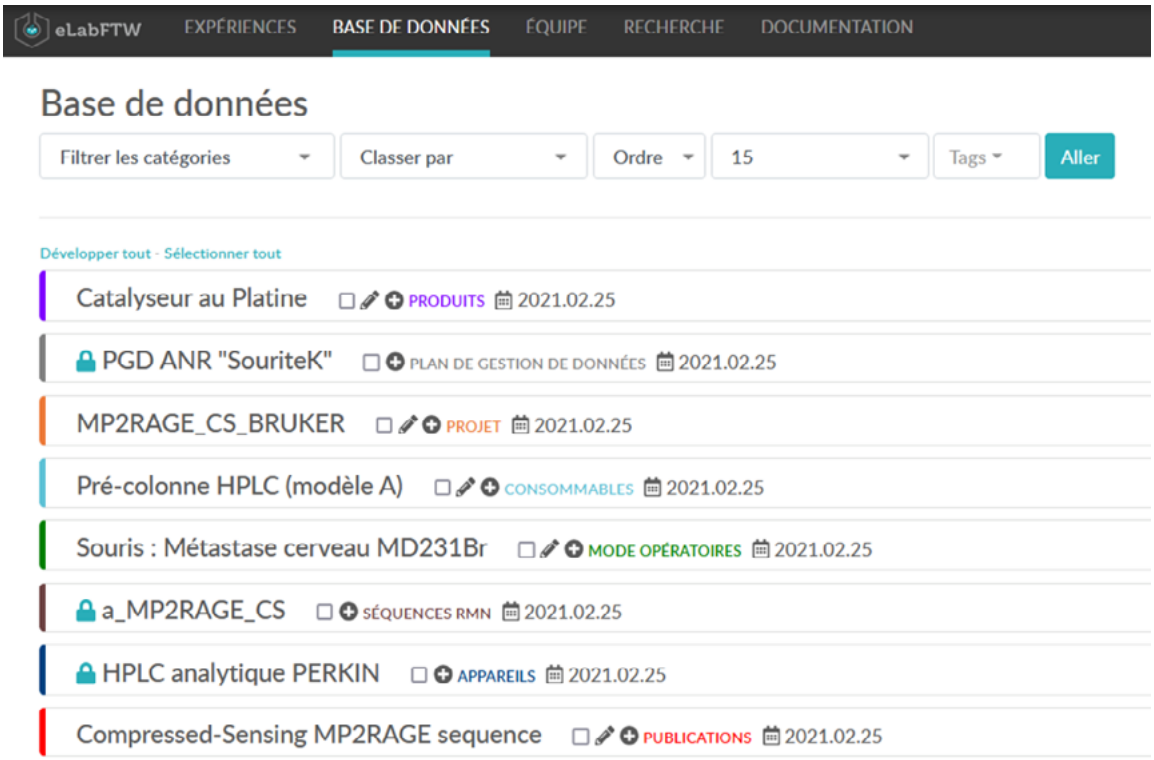


Figure 16 - Les bases de données dans eLabFTW

LeLabFTW peut convenir à tout type de laboratoire (chimie, biologie, physique, astronomie, etc.), en plus des fonctionnalités décrites ci-dessus, les fonctionnalités suivantes sont aussi disponibles :

- Description des expériences avec insertion (images, tableaux, liens)
- Affichage du statut d'une expérience (personnalisable)
- Filtres et moteur de recherche interne multiple (intégration de mots clés)
- Association de métadonnées aux expériences
- Duplication d'une expérience comme un modèle
- Intégration de pièces jointes à chaque expérience (images, PDF, vidéos, tableurs, rapports, publications, spectres, chromatogrammes, etc.)
- Travail collaboratif facilité par le partage de protocoles
- Les expériences ou entrées de la base de données sont exportables au format PDF, ZIP, CSV ou JSON et peuvent même être réimportées d'une instance d'eLabFTW à une autre
- Accès à des données externes pour lecture ou référencement
- Possibilité de commenter les expériences

Les fonctionnalités citées dans le tableau du CoSO (tableau 1) sont exhaustives mais, en fonction du contexte et du domaine, n'ont pas la même importance.

Dans eLabFTW, il y a deux types de concepts :

- Les expériences, qui sont uniques (correspondant à la notion « classique » de cahier de laboratoire) auxquelles sont associées différentes données (spectres, calculs, TAG, images, liens, publications, etc.).
- La base de ressources, communes (classées en items) qui peut être référencées dans les expériences dans lesquelles se trouvent : plan de gestion de données, projets, réactifs, protocoles, appareillages, comptes-rendus, documentations, etc.

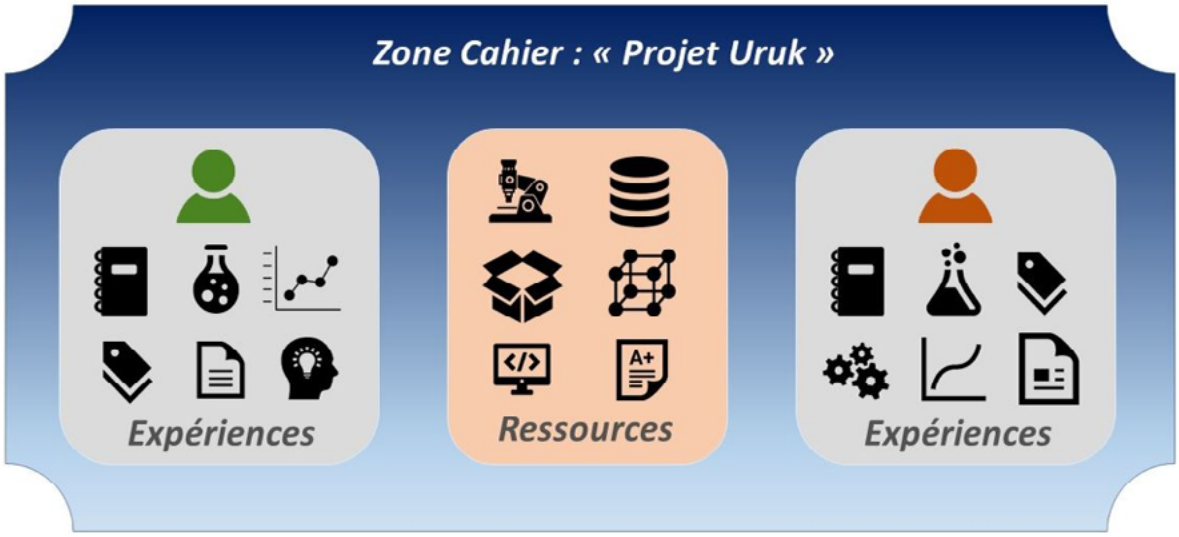


Figure 17 - Exemples de bases de ressources

Il existe généralement deux niveaux d'information, les informations uniques que sont les différentes expériences et les informations réutilisables (réactifs, protocoles…) qui peuvent être référencées dans plusieurs expériences (figure 18). Le schéma suivant introduit les deux notions expériences, base de données/ressources ainsi que la gestion les données externes.

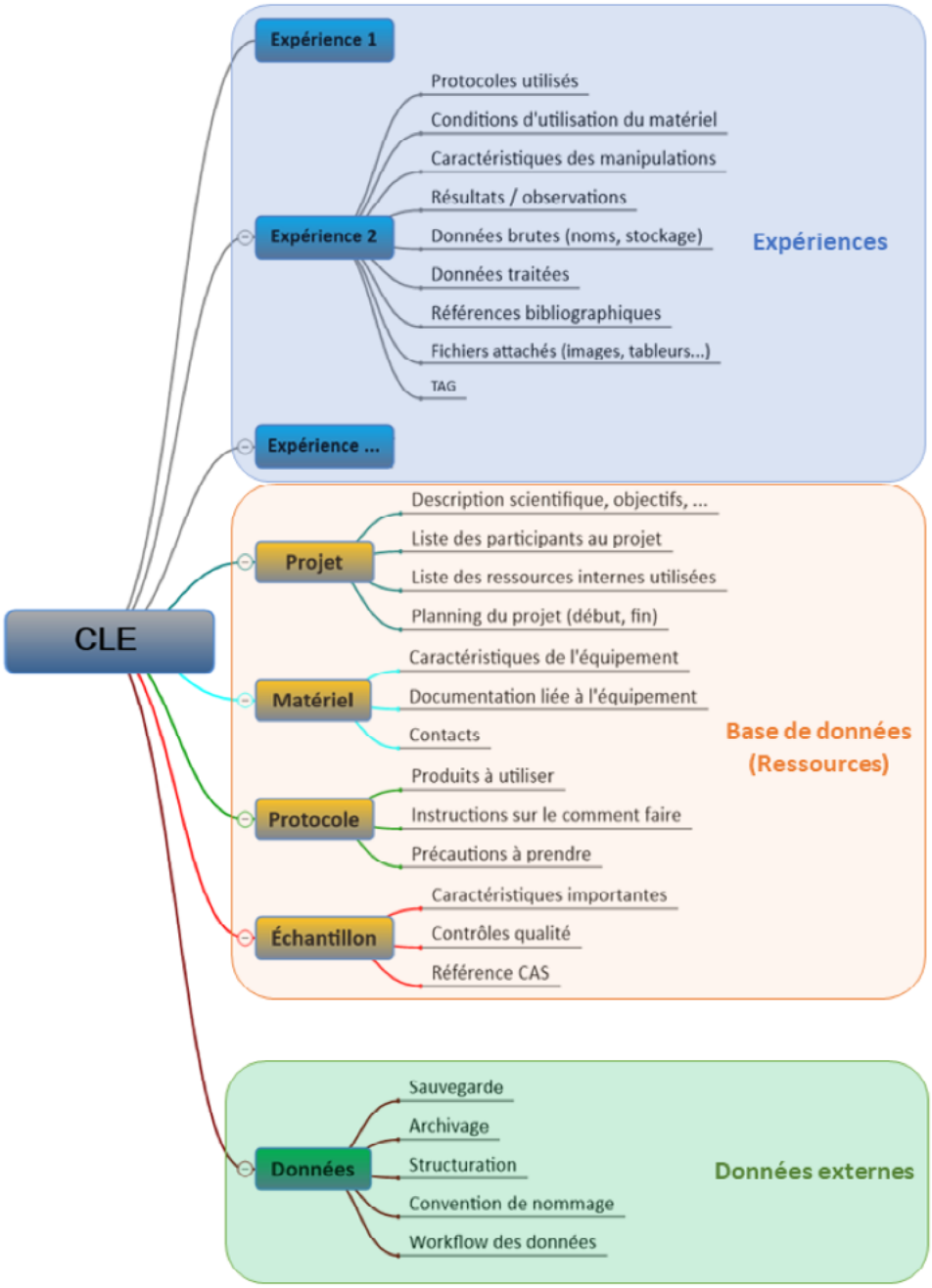


Figure 18 - Exemple d'architecture d'eLabFTW

Il convient de réfléchir aux différentes bases de données qui sont associées aux expériences. Il peut s'agir d'une base de données des équipements qui peut contenir les informations nécessaires lors de la rédaction d'une publication, d'une base de données de protocoles ou modes opératoires ou d'une base de données de consommables. De la même manière, pour faciliter les recherches, il peut être intéressant de définir un référentiel des mots clés (thématiques de recherche, molécules, technologie, …).

Automatisation

Une fois que le cahier de laboratoire électronique est une pratique maîtrisée, il est possible d'envisager certaines automatisations de tâches au travers de l'API d'ELabFTW et ce sans oublier de prendre en considération les aspects de sécurité liés à toute automatisation.

Il est, par exemple, possible de créer un formulaire renseignant diverses données d'une expérimentation. L'API peut permettre d'intégrer automatiquement ces données dans le CLE. L'utilisateur peut ensuite modifier et compléter les informations ainsi récupérées.

7.2 - Clôture du CLE

L'organisation décidée lors de la mise en place du CLE doit prendre en compte la clôture de celui-ci. Elle peut être envisagé à différents niveaux :

Niveau utilisateur

Au départ d'un utilisateur il est conseillé de verrouiller l'ensemble de ses expériences de manière à ce qu'elles ne soient plus modifiables mais toujours accessibles (par exemple avec l'encadrant d'une thèse). Une copie des expériences de l'utilisateur est possible sous forme d'un fichier archive (pdf,zip…).

L'administrateur de l'équipe peut « archiver » un utilisateur ce qui à pour effet de verrouiller toutes les expériences de ce dernier et de lui bloquer son accès login :

Il est toujours possible de revenir en arrière plus tard si l'utilisateur revenait au sein de la structure :

Attention, une suppression d'un utilisateur efface irrémédiablement son compte, ses expériences et les données liées à celles-ci.

De manière générale il ne faut pas oublier de gérer les départs des personnes et vérifier les modalités de transferts des droits et d'accès aux expériences des personnes sortantes.

Niveau projet

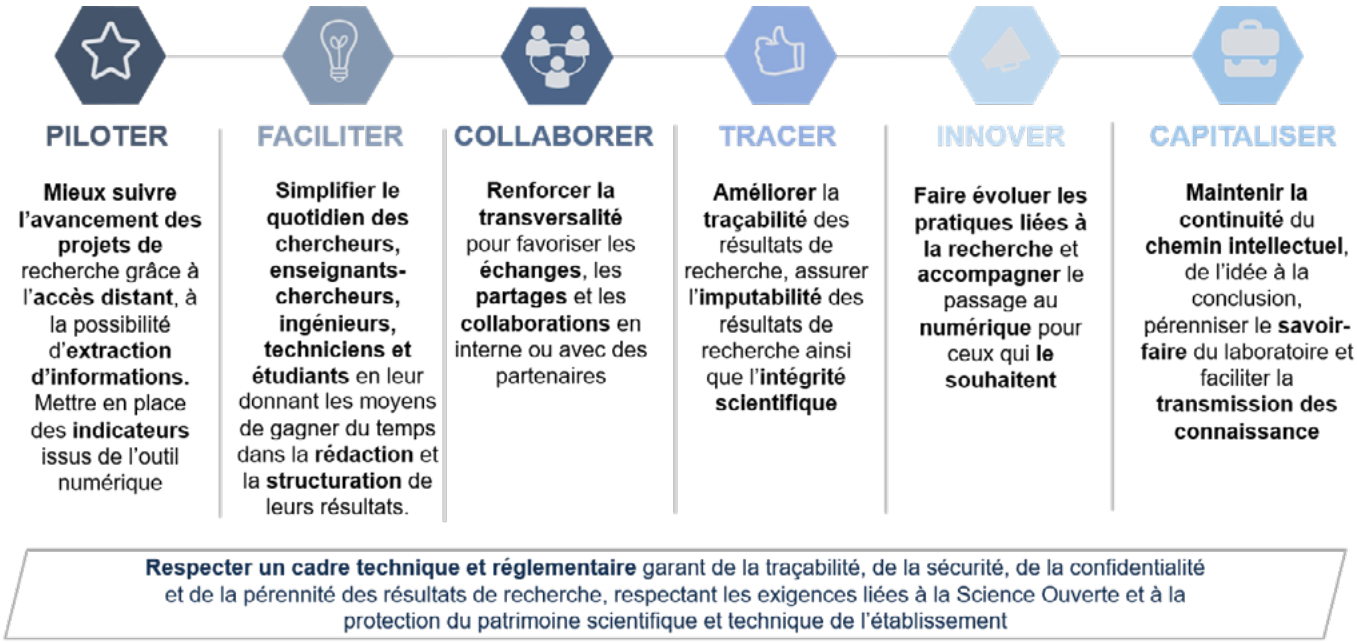
Si une zone a été créé dans le cadre d'un projet, une fois ce dernier terminé il faudra verrouiller l'ensemble des expériences et bases de données, créer une sauvegarde contenant l'ensemble des informations du projet (CLE et données externes) dans le but de les archiver.

8. CONCLUSION

Au sein d'une structure de recherche, assurer la traçabilité des différentes activités est essentiel pour assurer la reproductibilité des expériences et donc concourir à l'intégrité scientifique de la recherche mais également pour permettre le partage et la réutilisation des connaissances. Dès lors, la gestion de l'ensemble des données produites et les métadonnées associées doit être une préoccupation majeure de toute organisation performante avec l'objectif de disposer de données FAIR soit Findable, Accessible, Interoperable and Reusable (figure 3). Dans ce cadre, le cahier de laboratoire constitue un véritable outil scientifique présent pendant les nombreuses étapes du cycle de vie des données.

L'évolution numérique du cahier de laboratoire (CLE), tout autant que celle des modes de travail actuels rendent désormais incontournable ce qui est bien plus qu'un changement de support : l'entrée dans des pratiques compatibles avec les objectifs de la science ouverte, l'amélioration continue vers plus de performance tout au long du processus de recherche, de la naissance des projets, leur réalisation, leur valorisation et à leur archivage.

Au-delà des aspects techniques et pratiques du CLE abordés dans ce guide, il est également important de prendre en compte les enjeux de leur mise en place. En effet, l'utilisation d'un CLE dans un laboratoire nécessite de mener une vraie réflexion sur l'organisation des recherches entreprises. D'une approche individuelle avec le cahier de laboratoire papier, l'approche est désormais plus collaborative avec le CLE. Celui-ci n'est pas une simple transposition numérique du cahier papier :



Il apporte une vraie plus-value permettant d'améliorer :

La traçabilité

J'ai une meilleure **traçabilité** de mes résultats de recherche. Je peux, par exemple, lier mes données sauvegardées sur des lieux de stockage (entrepôt de données, etc.) à mon CLE.
Je bénéficie de l'**horodatage** permettant de définir la temporalité de la production de mes résultats et de prouver l'antériorité d'une découverte.
L'**authentification** permettant d'accéder à mon CLE est **renforcée** afin de garantir la propriété intellectuelle et l'authenticité des résultats.

La gestion des informations

Je peux joindre et centraliser des images, tableurs, schémas, rapports, spectres, équipements, etc. et **personnaliser mon CLE** de manière à disposer des informations de base utiles à l'élaboration de mes recherches (liste de produits, appareils, protocoles, etc.).
J'améliore la **structuration** de mes résultats, l'utilisation d'un puissant **moteur de recherche** avec filtres et étiquettes facilite mes recherches d'informations.

La collaboration & l'encadrement

La **collaboration** au sein de mon équipe/projet est facilitée grâce au partage des expériences (gestion des droits de lecture et d'écriture, regroupement par projet)
L'encadrement et la gestion des projets de recherche est amélioré grâce à l'accès distant permettant un meilleur suivi des manipulations (possibilité de figer les expériences)

La sécurité des résultats de recherche

Je bénéficie d'une **sauvegarde numérique de mes résultats de recherche** dans des serveurs en **conformité avec la Politique de sécurité du système d'information** en vigueur **de l'établissement**.
L'accès aux données est **sécurisé** par l'authentification renforcée et le chiffrement du stockage.
La réglementation autour de la **protection de mes données** est ainsi respectée.

La gestion des expériences - l'automatisation

J'ai la possibilité de **dupliquer** mes expériences et de créer des **modèles de protocoles** .
Si je le souhaite, je peux **intégrer** la **gestion de mes appareils et équipements** (réservations, notifications, rappels).
Par la suite, je peux envisager certaines automatisations de tâches…

Ainsi, ce projet fait évoluer les pratiques liées à la recherche et nécessite d'accompagner le changement d'organisation et le passage au numérique.

Au-delà de la perte de l'objet papier et des habitudes de travail, il convient de bien anticiper les obstacles liés au domaine et à la technique, d'où les nombreux aspects pratiques de ce guide. Plus encore, il apparaît important de souligner que les CLE ne sont pas en soi une finalité ni une garantie de bon usage, mais plutôt des outils à adapter selon les objectifs des équipes et des projets. La question du « comment les utiliser » doit donc s'articuler étroitement à celle du « pourquoi » et faire l'objet d'une réflexion spécifique sur le terrain.

A travers ce guide de bonnes pratiques, le réseau Qualité en Recherche a souhaité sensibiliser les personnels des unités de recherche à la mise en place et à l'utilisation d'un cahier de laboratoire électronique en apportant une vision « terrain » à cette nouvelle organisation des activités de recherche.

BIBLIOGRAPHIE

AFNOR. Norme NF 3 Z42-013. Recommandations relatives à la conception et à l'exploitation de systèmes informatiques en vue d'assurer la conservation et l'intégrité des documents stockés dans ces systèmes.

Hadrossek C., Janik J., Libes M., Louvet V., Quidoz M.C., Rivet A., Romier G.. Guide de bonnes pratiques sur la gestion des données de la Recherche. 2021. hal-03152732. https://mi-gt-donnees.pages.math.unistra.fr/guide/

Rivet A., Bachèlerie M.L., Denis-Meyere A., Tisserand D..Traçabilité des activités de recherché et gestion des connaissances. Mission pour les Initiatives Transverses et Interdisciplinaires. CNRS. 2018. https://qualite-en-recherche.cnrs.fr/wp-content/uploads/2021/08/guide_tracabilite_activites_recherche_gestion_connaissances.pdf

AFNOR. Norme ISO 14721. Organisation et fonctionnement d'un centre d'archivage pour la pérennisation des données numériques.

Cartier A., Delemontez R.,Moysan M., Reymonet N.. Réaliser un plan de gestion de données « FAIR » : guide de rédaction. Université Paris-Diderot, Paris-Descartes. 2018. http://www.bu.parisdescartes.fr/doc/recherche/Realiser_un_DMP_V2_2018_Def_HRS4R.pdf

Charte française de déontologie des métiers de la recherche. Janvier 2015 (ratifications au 22 janvier 2019). https://comite-ethique.cnrs.fr/charte/

DATACC, Accompagnement à la gestion des données de recherche en physique et en chimie, bibliothèques de Lyon 1 et de Grenoble Alpes Université. https://www.datacc.org/bonnes-pratiques/utiliser-un-cahier-de-laboratoire-numerique/

Deboin M.C.. Se familiariser avec les plans de gestion de données de la recherche, en 6 points. Montpellier (FRA). CIRAD, 6p. 2018. https://doi.org/10.18167/coopist/0056.

Dupré P.G., Gallina-Muller C.. Déploiement du cahier de laboratoire électronique à l'INSERM et nouvelles Perspectives. 2019. JRES. Dijon.

Léon N., Libri D.. Analyse de l'enquête sur les cahiers de laboratoire électroniques au CNRS.2021. https://intranet.cnrs.fr/directions/mpr/Documents/Analyse_enqu%C3%AAte_CLE_2020.pdf

Ouvrir la science. Rapport du Groupe de Travail sur les Cahiers de Laboratoire électroniques du CoSO. 2021. https://www.ouvrirlascience.fr/rapport-du-groupe-de-travail-sur-les-cahiers-de-laboratoire-electroniques/

9. ANNEXES

Annexe 1 : Exemple de trame de plan de gestion de données

Informations administratives

- Nom et identifiant du projet
- Description du projet
- Agence(s) de financement
- Nom et identifiant éventuel du responsable principal de projet
- Contact pour les données de projet
- Date de la 1re version
- Date de la dernière mise à jour
- Politiques associées au projet, incluant les instructions ou recommandations de l'agence de financement et de l'institution

Collection de données

- Description des données, incluant le type de données, le format et le volume
- Jeux de données existants qui seront utilisés
- Méthodes de collecte et de création des données
- Système d'organisation, de nommage et de gestion des répertoires et des fichiers
- Processus d'assurance qualité mis en œuvre

Documentation et métadonnées

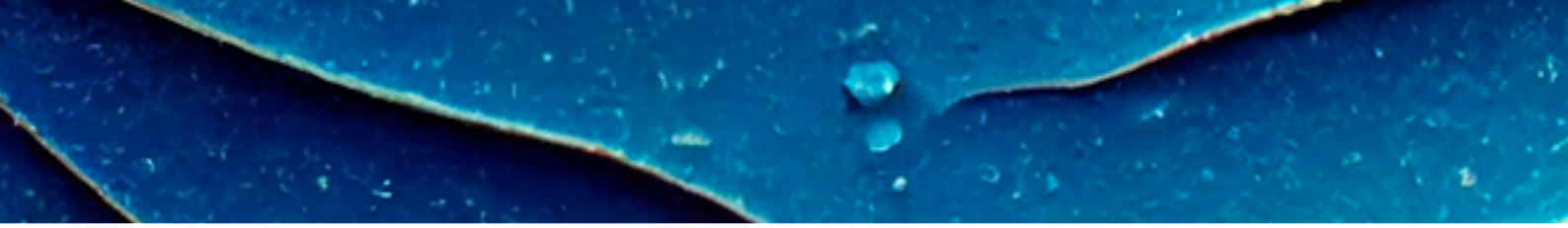
- Informations nécessaires pour lire et interpréter les données
- Organisation de la collecte et de la documentation
- Standards de métadonnées adoptés

Ethique, cadre légal
Ethique

- Détails de l'accord de conservation et de partage des données
- Étapes pour la protection de l'identité des participants
- Étapes pour la sécurité du stockage et du transfert de données sensibles

Cadre légal : droits de propriété intellectuelle et copyright

- Nom de(s) propriétaire(s) des données
- Licence(s) pour la réutilisation des données (par exemple, une des licences Creative Commons ou Open Data Commons)
- Restrictions d'utilisation par les tierces parties
- Délai requis pour le partage de données (embargo lié à la publication dans une revue ou délai d'application d'un brevet



Stockage, sauvegarde, sécurité
Stockage et sauvegarde

- Lieu de stockage des données
- Plan de sauvegarde
- Personne ou équipe responsable de la sauvegarde
- Procédures de récupération

Sécurité

- Risques et leur gestion
- Dispositif d'accès
- Dispositif éventuel pour le transfert sûr et intègre des données collectées sur le terrain

Sélection et conservation

- Informations détaillées sur les données qui seront retenues, partagées et/ou conservées, et référence aux obligations contractuelles, légales ou réglementaires
- Utilisations de recherche prévues des données
- Durée de conservation des données au-delà du projet
- Entrepôt ou archive de conservation des données et responsabilités associées
- Temps et effort nécessaire à la préparation des données pour leur conservation et leur partage

Partage des données

- Étapes à mener pour faciliter la prise de connaissance (discovery) des données par les autres
- Conditions de restriction du partage des données et détails de leur application dans l'accord de partage de donnés
- Mécanisme de partage de données (via un entrepôt, sur demande expresse ou tout autre processus)
- Délai de publication
- Procédure éventuelle d'obtention d'un identifiant persistant pour les données

Responsabilités et moyens

- Nom de la personne responsable de la mise en œuvre du plan de gestion de données
- Nom de la personne responsable de chaque activité de gestion des données
- Équipements et logiciels requis (en addition à ceux existants fournis par l'institution)
- Besoins additionnels d'expertise ou de formation
- Charges imposées par les entrepôts de données

Annexe 2

ETAPES POUR LA MISE EN PLACE D'UNE « ZONE CAHIER »

Cette annexe décrit les étapes pour la mise en place du CLE au sein d'une équipe après présentation de celui-ci à l'unité.

Etape 1 : Séminaire de présentation de l'outil

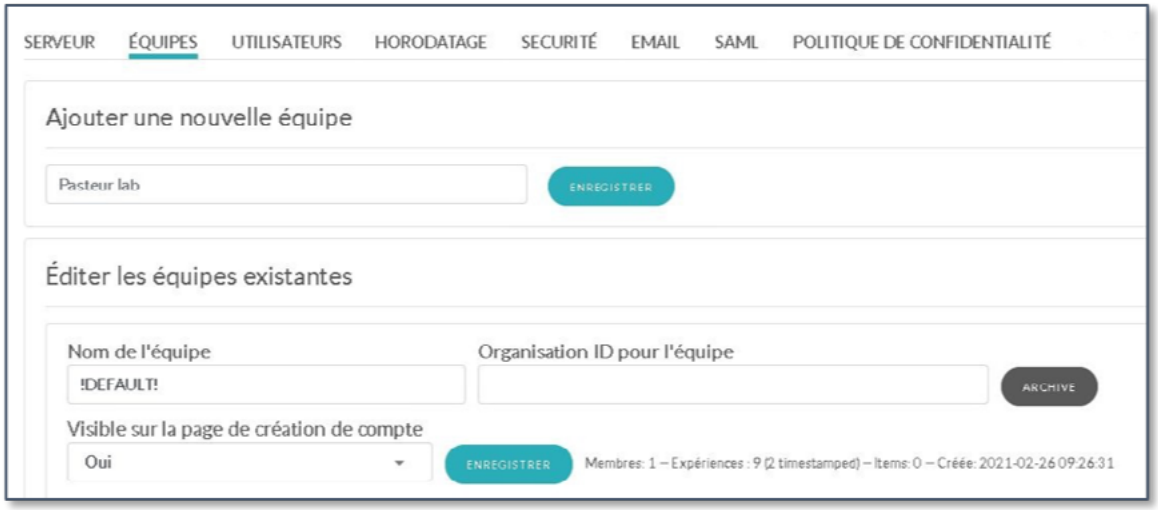
Avant d'activer un cahier de laboratoire électronique pour une équipe, il est fortement conseillé que le pilote du projet CLE (cf. paragraphe Pilotage du Projet) de la structure organise une présentation (visioconférence ou présentation interactive). L'objectif de cette réunion rapide est de présenter les généralités sur les CLE (sécurité, différences avec la version « papier », importance des métadonnées, plan de gestion des données, etc.) puis de montrer les différentes fonctionnalités de l'outil.

Etape 2 : Création d'une « zone cahiers »

Le pilote CLE crée un nouveau cahier de laboratoire (attention au nom de l'équipe, il peut y en avoir beaucoup rapidement, définir un mode de nommage, ex : « EQUIPE_Thématique »). Quelques conseils sur l'organisation humaine sont disponibles au chapitre 5 de ce présent guide.

Ce cahier est un projet (ex : ANR), une équipe, un service (ex : Microbiologie), une thématique… Travailler en mode « projet » est parfaitement compatible avec l'utilisation d'un CLE (travail collaboratif, PGD…) Dans cette « zone » plusieurs personnes peuvent utiliser eLabFTW et avoir la même base de ressources

Panneau du Sysadmin (pilote CLE) :



Etape 3 : Formation de l'administrateur « équipe »

Une personne doit administrer ce nouveau cahier de laboratoire. L'équipe est de ce fait autonome et peut choisir son organisation interne. L'administrateur ayant des droits étendus, il doit porter une attention particulière à ses actions en mode « admin ». Avant la réunion préliminaire de l'équipe, le chef de projet CLE peut approfondir avec l'administrateur les aspects organisationnels par rapport aux besoins de l'équipe et l'aider à définir un mode de fonctionnement (quel genre de base de données, combien de groupes, gestion des métadonnées, etc.).

Seul « l'admin » a accès au « Panneau d'administration » :

Etape 4 : Réunion préliminaire de l'équipe

Il est conseillé à l'ensemble des personnes qui utilisent le nouveau cahier de laboratoire électronique de se réunir pour clarifier son utilisation. Chaque équipe, en fonction de ses habitudes, peut ainsi réfléchir à son mode de fonctionnement (groupe d'encadrement, utilisation de modèles spécifiques, travail collaboratif, base de données adaptée au groupe, règles à respecter, plan de classement, etc.).

Etape 5 : Définitions des règles de fonctionnement

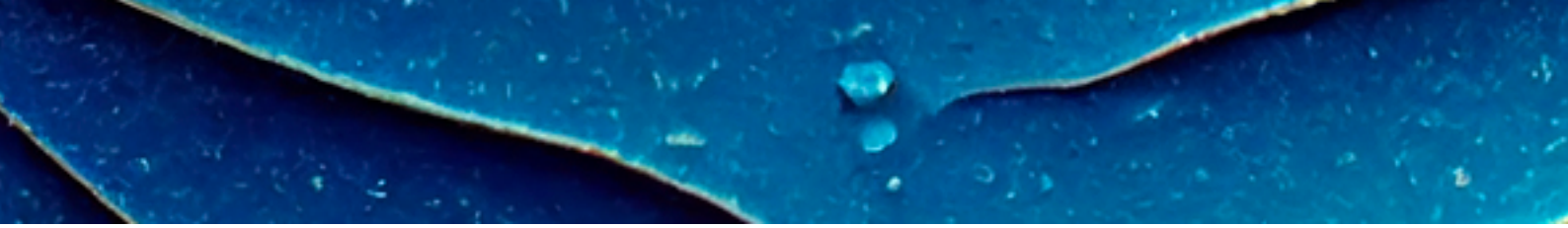
Pour maintenir une organisation optimale pérenne (il y a régulièrement de nouveaux utilisateurs dans l'équipe), il est avantageux de rédiger quelques documents à mettre à disposition de tout le monde dans la base de données.

Exemples :

- Plan de nommage des expériences et des fichiers (cf. « Règles de nommage » § 6.2)
- Plan de classement des fichiers numériques associés (cf. « Choisir un plan de classement » § 6.3)
- Gestion des étiquettes (TAG)
- Plan de gestion des données du projet (cf. « Plan de gestion de données » § 6.5)
- Règles de sauvegarde et d'archivage (cf. « Sauvegarde et archivage » § 6.4)

Etape 6 : Inscriptions des utilisateurs

Chaque nouveau membre du projet doit se créer un compte sur la page d'accueil d'eLabFTW en choisissant « la zone cahier » nouvellement créée. L'administrateur doit valider chaque demande. Un même compte peut faire partie de plusieurs « zones cahier », le choix de la zone de travail se fait à chaque nouvelle connexion.



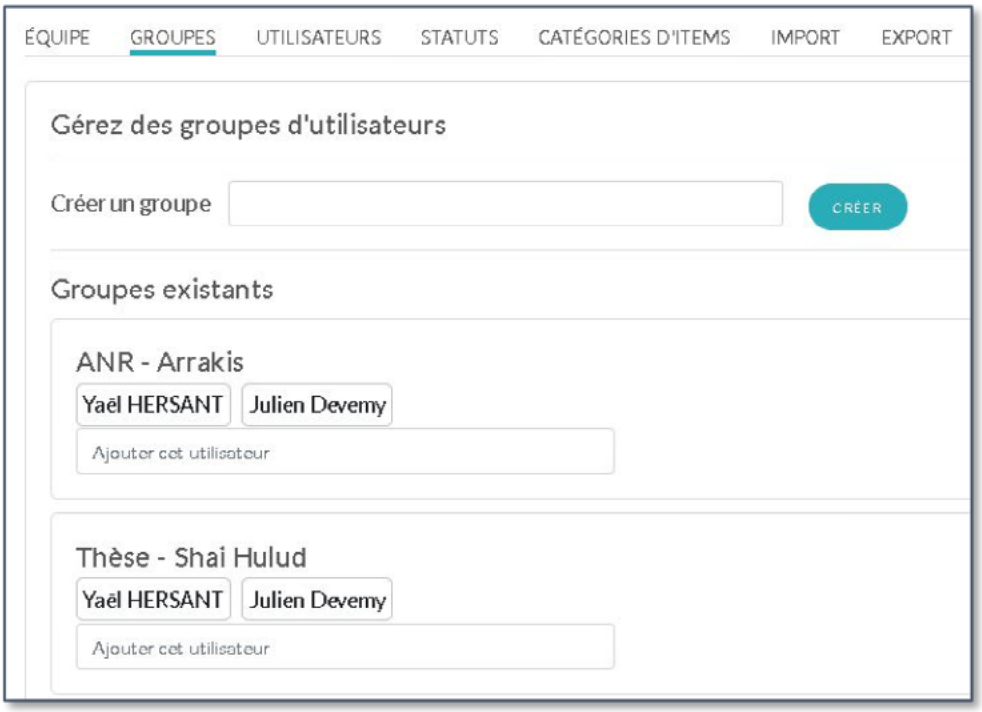
Etape 7 : Configuration du cahier

Onglet « Equipe »

L'administrateur peut configurer sa « zone cahier » pour avoir un fonctionnement qui correspond aux attentes de son équipe. Il peut ainsi définir les droits des « utilisateurs » sur les expériences, la base de données/ressources, les tags, etc.

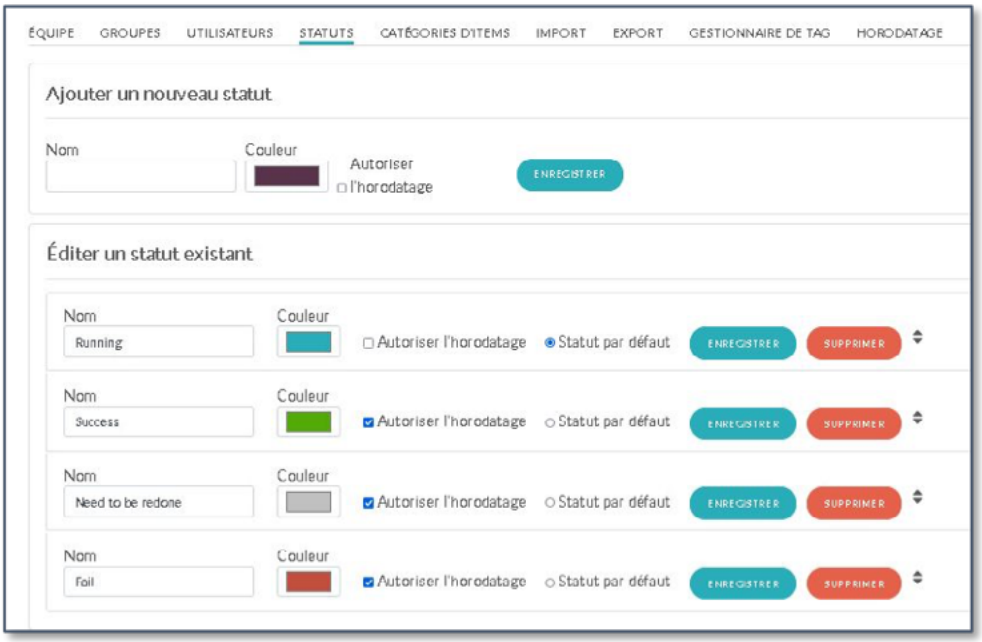
Onglet « Groupes » :

L'administrateur peut définir des groupes de travail qui permettent d'ajouter un niveau dans la zone et ainsi avoir une gestion des expériences facilitée. Il est possible, par exemple, de définir un groupe composé d'une encadrante, d'un thésard et de stagiaire(s) travaillant sur le même sujet. Ces groupes permettent ensuite de contrôler qui peux voir/éditer ou pas une série d'expériences



Onglet « Statuts » :

L'administrateur peut éditer et personnaliser les statuts des expériences

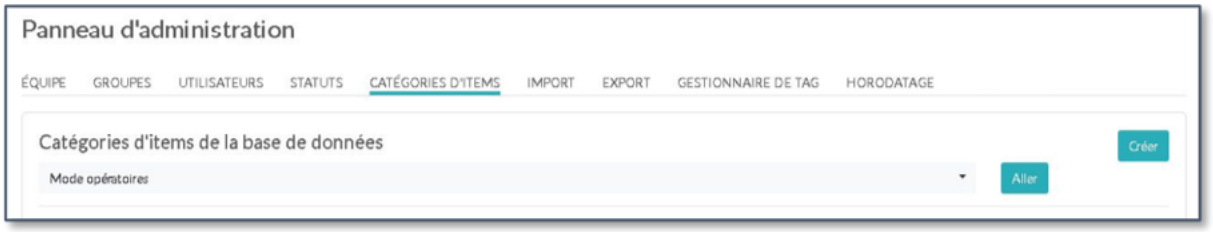


Onglet « Base de données » :

L'administrateur peut créer et personnaliser différents items pour la base de données/ressources en fonction des besoins définis au cours de la réunion préparatoire. Il est ainsi possible de lier un mode opératoire, un appareil ou une documentation présents dans la base aux expériences des utilisateurs.

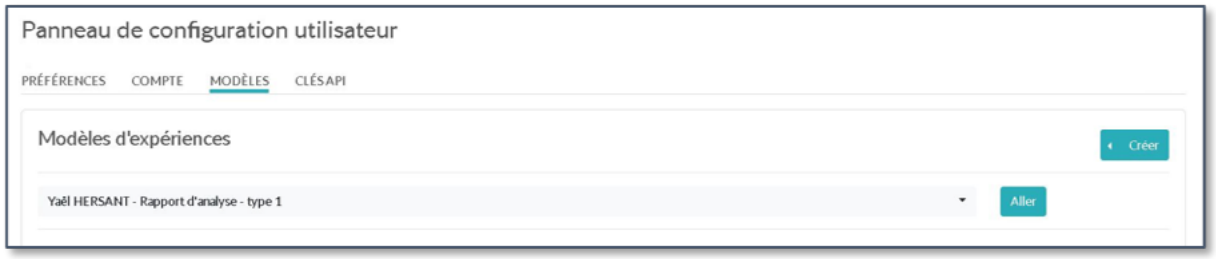
Exemples :

- Modes opératoires
- Documentations (règles de fonctionnement, etc.)
- Publications
- Appareils (une option de réservation est activable)
- Etc.



Onglet « Modèles » :

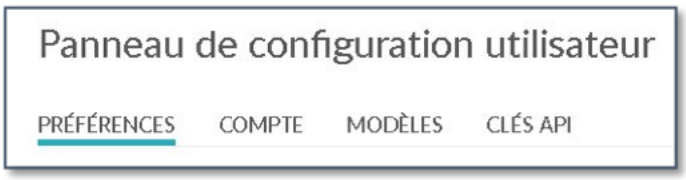
Afin d'obtenir une uniformité sur certaines expériences, il est conseillé de créer des modèles qui permettront : de gagner du temps, de bien gérer les autorisations, d'avoir une meilleure gestion des données, etc.



Etape 8 : Phase de test

Configuration des profils :

Chaque utilisateur peut personnaliser son interface (langue, options d'affichage, modèles, etc.) en ouvrant le « Panneau d'utilisateur ».



Gestion des permissions :

Pour chaque expérience, le choix peut être réalisé entre :

- Qui pourra la voir :

- Qui pourra l'éditer :

Il est conseillé de définir ces permissions directement dans les modèles mais il est également possible de le faire pour chaque expérience. Pour ce faire, il suffit d'éditer une page, de cliquer sur l'icône « trois petits points » en haut à droite et de régler les droits.

Essais :

Pour s'approprier l'outil et vérifier que la configuration du cahier est adéquate, il est conseillé de faire quelques tests :

- Créer des expériences avec les différents modèles
- Vérifier que les groupes remplissent leur rôles (qui voit les expériences, qui peut les éditer, etc.)
- Créer des documents dans la base de données
- Tester l'horodatage d'une expérience
- Tester les aspects sécurité et de confidentialité (en demandant à un tiers s'il voit vos expériences)

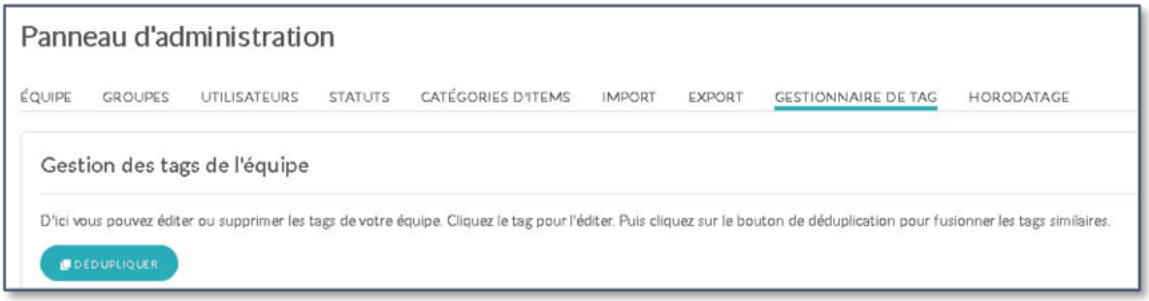
Etape 9 : Actions correctives

Après quelques mois d'utilisation, il peut être intéressant de faire un point sur l'utilisation du cahier électronique. Le mode de fonctionnement de la « zone cahier » est-il optimal ? La gestion des fichiers annexes (métadonnées, classement, sauvegarde…) est-elle bien réalisée ?

Il est conseillé de corriger dès que possible son organisation si besoin car compte tenu de la croissance rapide du nombre d'expériences et de données dans le CLE.

Gestion des tags :

L'administrateur de l'équipe peut gérer l'ensemble des tags pour éviter les doublons et uniformiser leur utilisation :



Ce guide a été réalisé par un groupe de travail du réseau “Qualité en Recherche” soutenu par la plateforme des réseaux de la Mission pour les Initiatives Transverses et Interdisciplinaires du CNRS.

A travers ce guide de bonnes pratiques, le réseau Qualité en Recherche souhaite sensibiliser les personnels des unités de recherche à la mise en place et à l’utilisation d’un cahier de laboratoire électronique au sein d’une unité de recherche en apportant une vision « terrain » à cette nouvelle organisation des activités de recherche.

CNRS
3, rue Michel-Ange 75016 Paris
www.cnrs.fr

Photos, © CNRS
Mise en page et impression : CNRS/IFSeM/Secteur de l’imprimé
Novembre 2022