



① Prévenir les
② Gaffes et les
③ Déboires...

Mettre en place un
PGD Entité
à partir de zéro

Judith Vallet
Michaël Dumas



- 97 personnes
- 8 équipes, 1 service administratif

Notre entrée dans le Réseau



Nous recherchons un correspondant données par unité de recherche

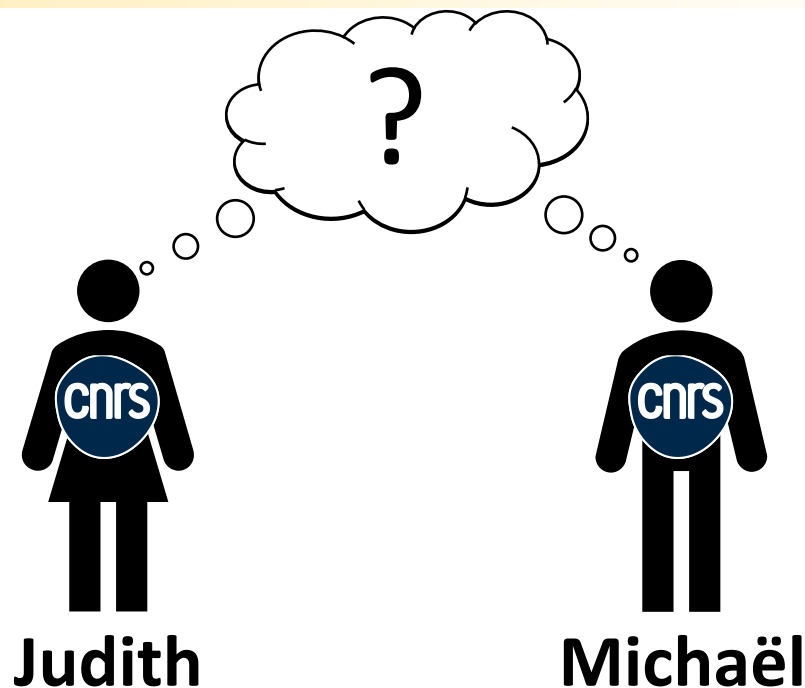
15.12.2023

Est-ce que vous m'autorisez à donner vos noms ?

05.04.2024



Directeur d'unité



Judith

Michaël

- Biologiste
- Bioinformaticien
- Correspondante
- CSSI
- RGPD
- 2024 : pilotes cahier de laboratoire électronique eLabFTW

Notre entrée dans le Réseau

24.04.2024
Discussion avec
Stéphanie, Aline et Inès



29.05.2024
1^{ère} Formation PGD
URFIST



30.05.2024
Volontaires pour
correspondants
données de la
recherche



11.06.2024
Webinaire CNRS
Principes FAIR et cycle
de vie des données



03.12.2024
Séminaire de Virgile
Jarrige (Dnum)
« Introduction à la
science ouverte »



03.10.2024
Invitation de Stéphanie
Cheviron pour un
séminaire « Gestion
des données en
biologie et en chimie »



28.06.2024
Réunion de lancement
du réseau de
correspondants
données



19.06.2024
1^{ère} discussion
technique « stockage
et sauvegarde » :
Infoprox, Dnum



Début 2025
Mise en place de
groupes de travail
dans l'UMR



Actions menées sur 1 an

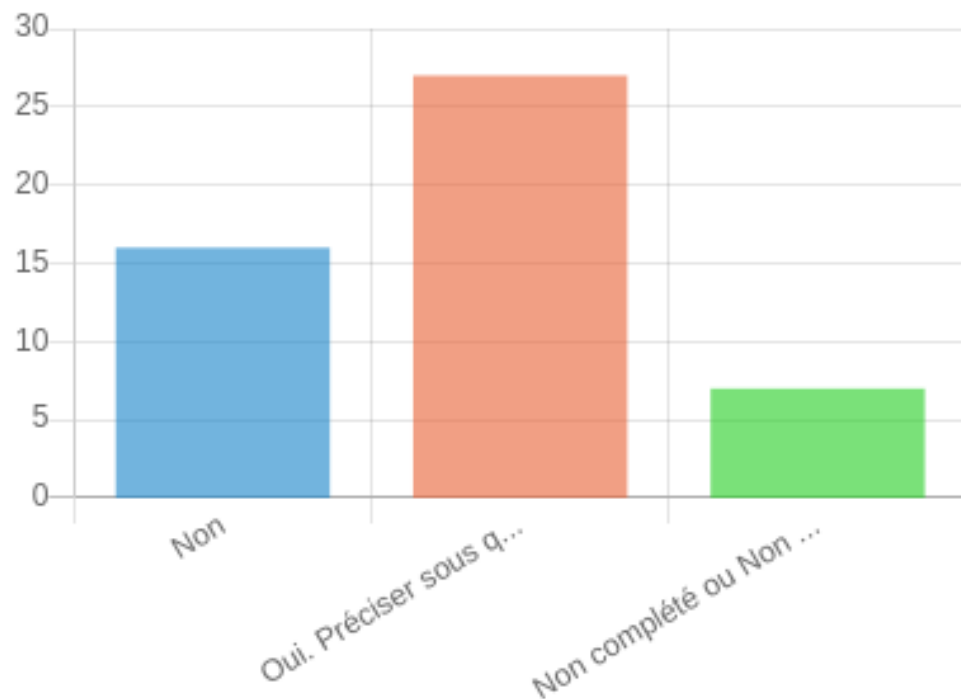
Recensement

- Sondage « état des lieux des données et des usages dans l'UMR » (été 2024)
- Sondage « Volumétrie des données UMR » (28.02.2025)
- Inventaire appareillage et format de données (en cours)



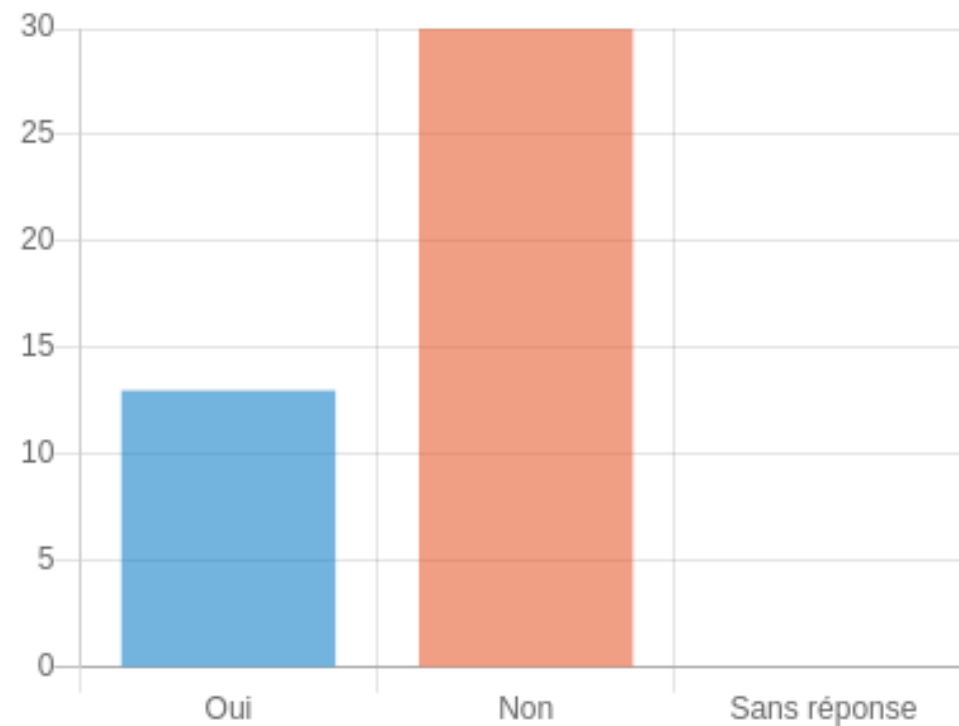
L'état des lieux

Avez-vous été sensibilisé à la gestion des données ?



■ Non
■ Oui. Préciser sous quelle forme :
■

Avez vous déjà entendu parler de principe FAIR pour les données ?



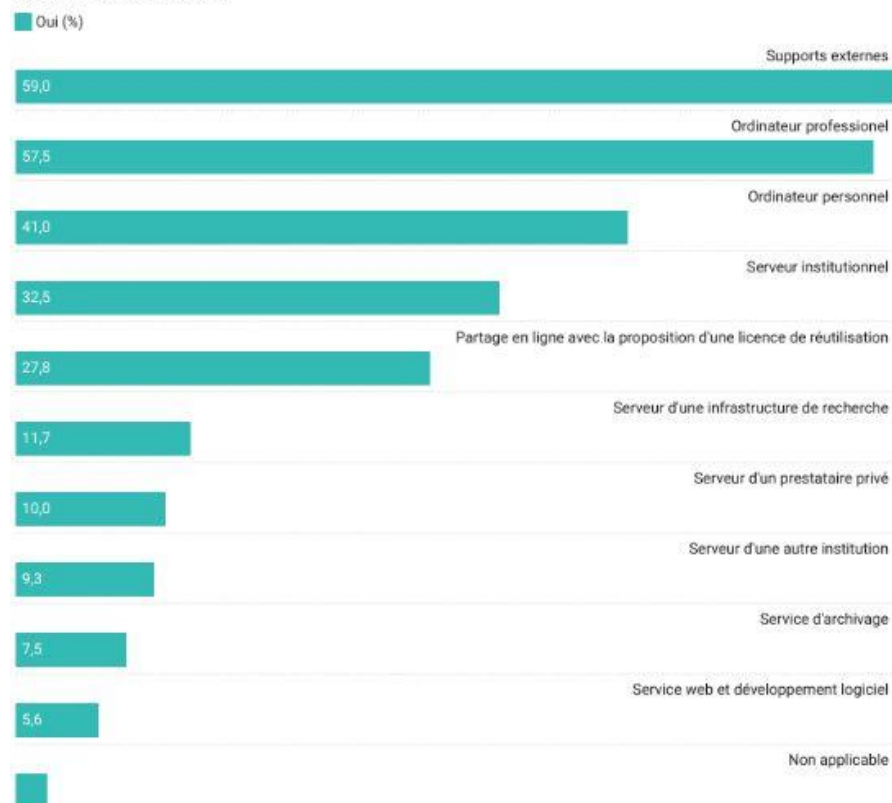
■ Oui
■ Non

Invitation de Stéphanie Cheviron et Aline Grand en séminaire d'unité (03.10.2024)

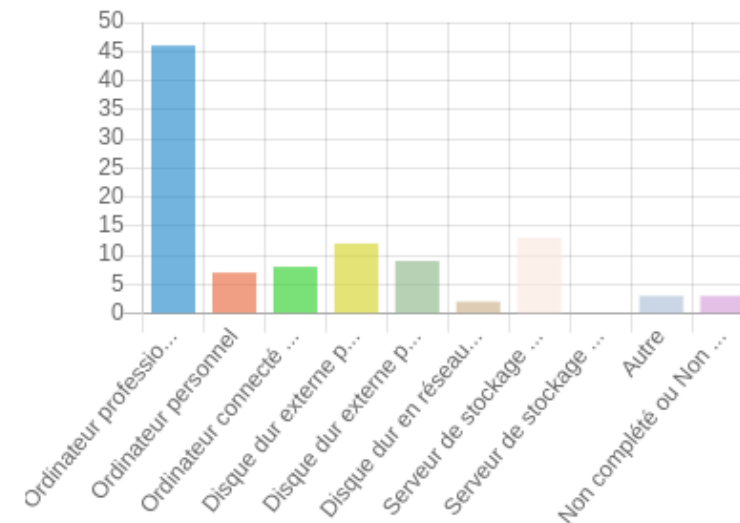
La réalité du terrain : stockage

Actuellement, comment conservez-vous vos données à la fin d'un projet?

Réponse à choix multiples.



Sur quels supports sont STOCKÉES vos données ?
! Attention de bien distinguer stockage et sauvegarde, se référer au glossaire dans le mail



- Ordinateur professionnel
- Ordinateur personnel
- Ordinateur connecté à une machine (qPCR, microscope, WB ...)
- Disque dur externe professionnel
- Disque dur externe personnel
- Disque dur en réseau (NAS, Synology)
- Serveur de stockage professionnel
- Serveur de stockage personnel
- Autre
- Non complété ou Non affiché

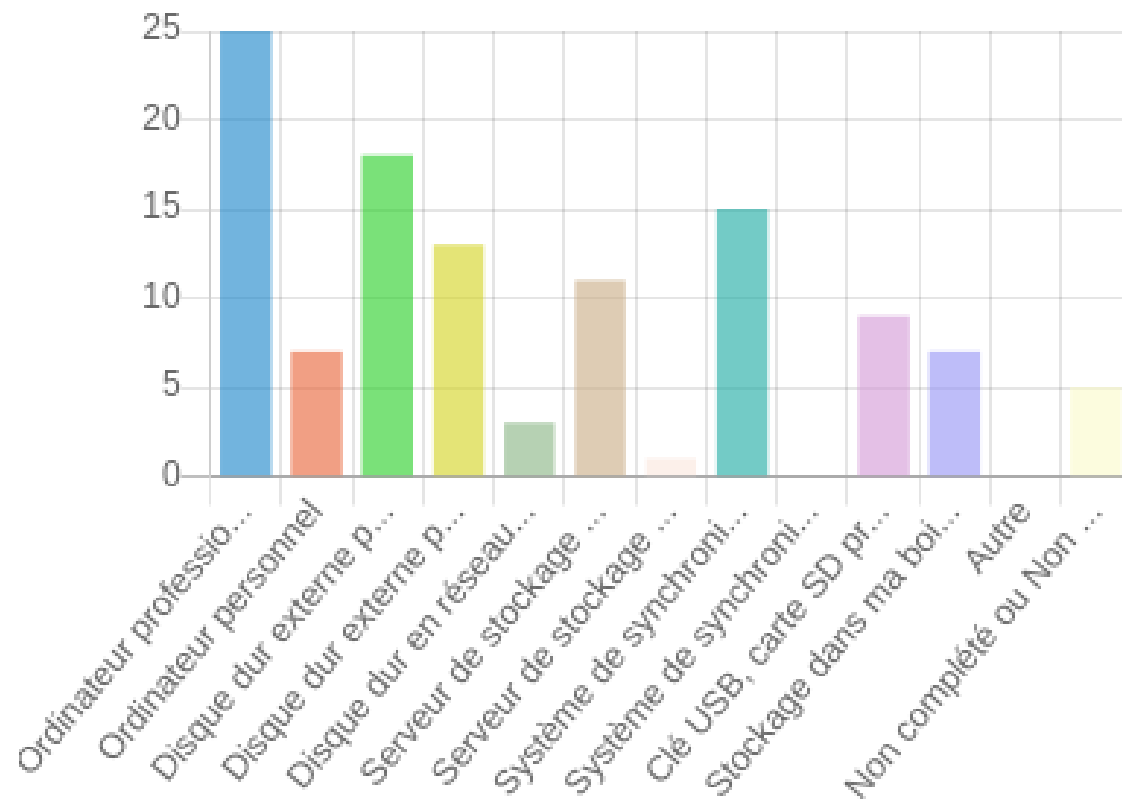
Enquête du Ministère français de l'Enseignement supérieur, de la Recherche et de l'Innovation auprès du CNRS, INSERM, Université...

Source : Décliner-SO – Décliner la science ouverte – Rapport final – 2022 – Licence CC BY 4.0



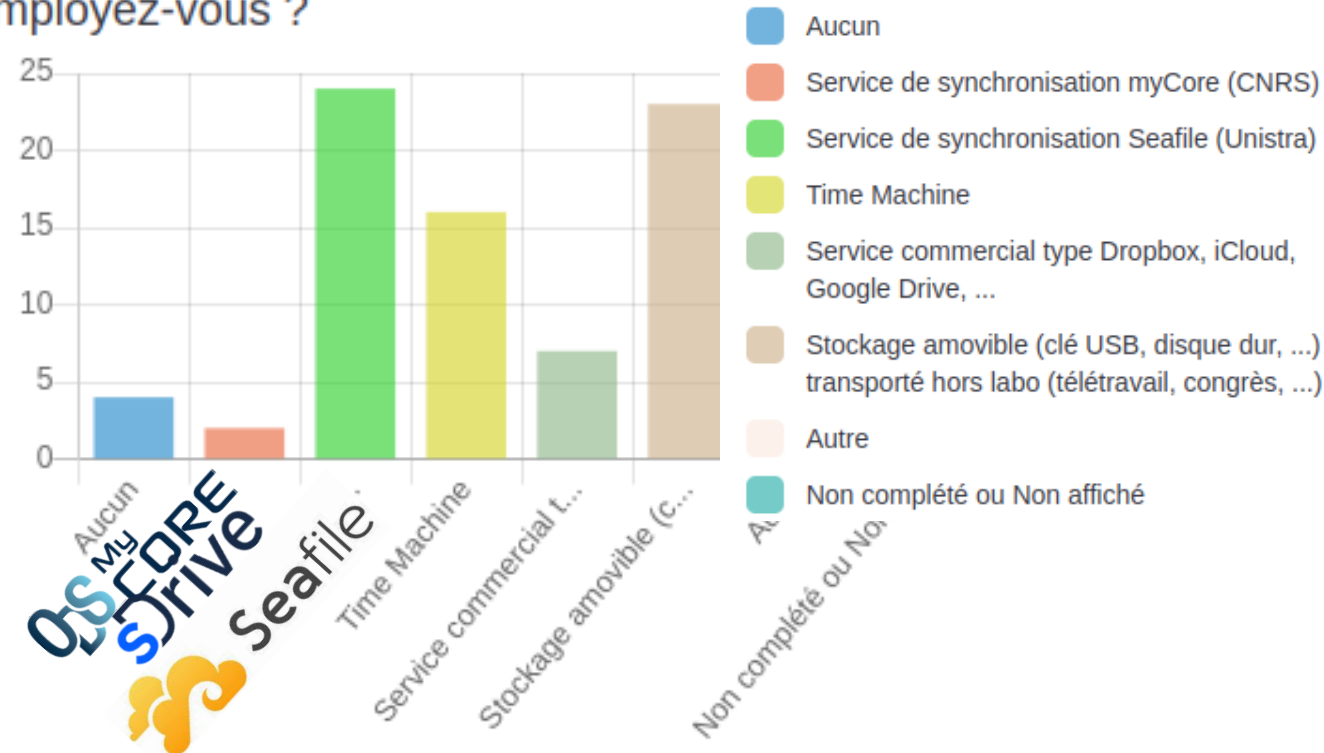
La réalité du terrain : sauvegarde

Si vos données sont SAUVEGARDEES, sur quels support se trouvent elles ? ! Attention de bien distinguer stockage et sauvegarde, se référer au glossaire dans le mail

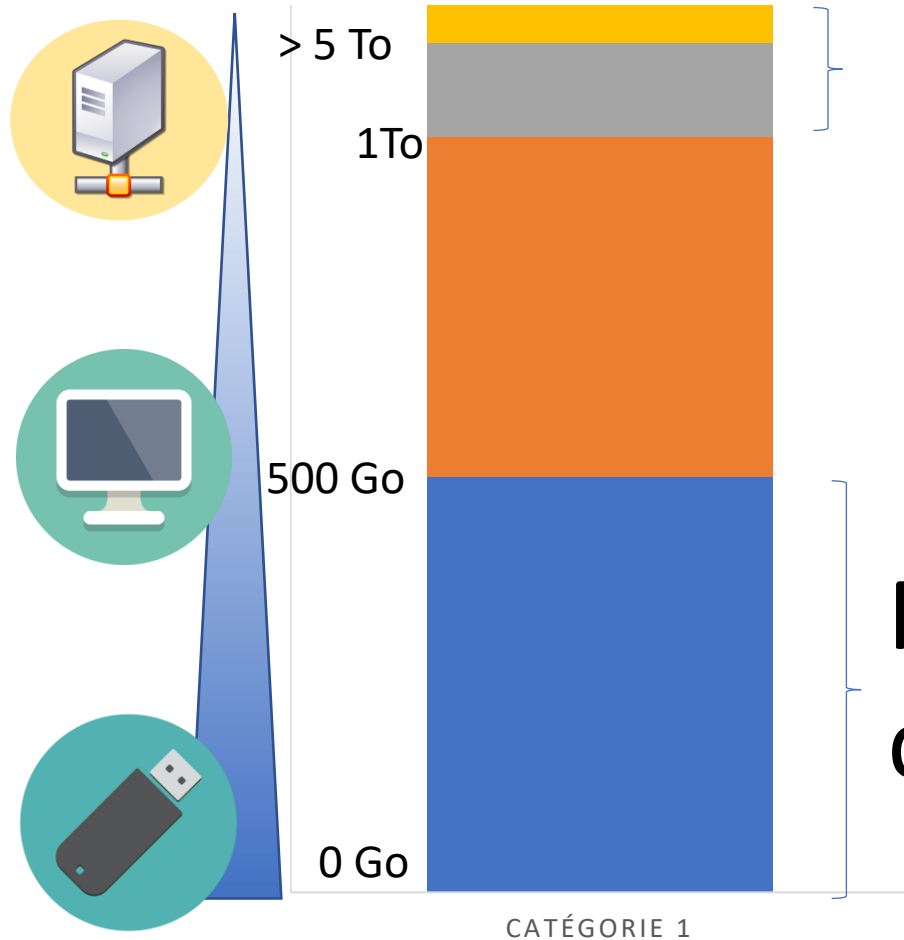


Pas de système de sauvegarde généralisé pour l'unité... mais Seafile se démocratise

Parmi les outils suivants, lequel ou lesquels employez-vous ?



La réalité du terrain : volumétrie



Microscopie
Bioinformatique

Majorité de postes
de travail individuel

Répartition volume
utilisé



Vol utilisé	93 To
Vol libre	99 To
Vol total	192 To



Actions menées sur 1 an

Recensement

- Sondage « état des lieux des données et des usages dans l'UMR » (été 2024)
- Sondage « Volumétrie des données UMR » (28.02.2025)
- Inventaire appareillage et format de données (en cours)

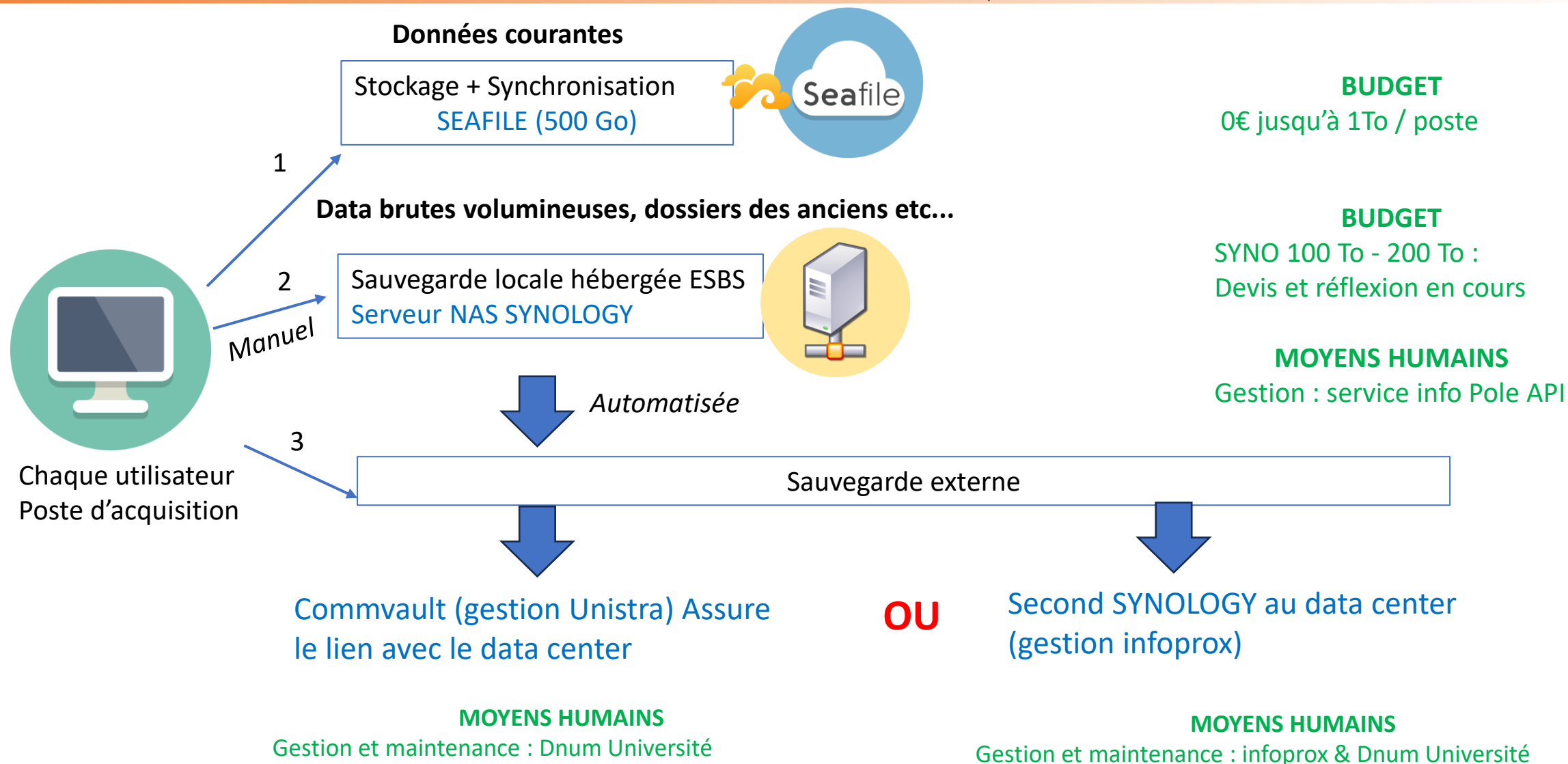
Groupes de travail

- Stockage et la sauvegarde des données dans l'UMR BSC (10.01.2025)
 - Infoprox, Dnum, DU et chefs d'équipes
- Prise en compte des aspects relatifs au RGPD (28.01.2025)
 - Adjointe DU, ADELE, RGPD CNRS

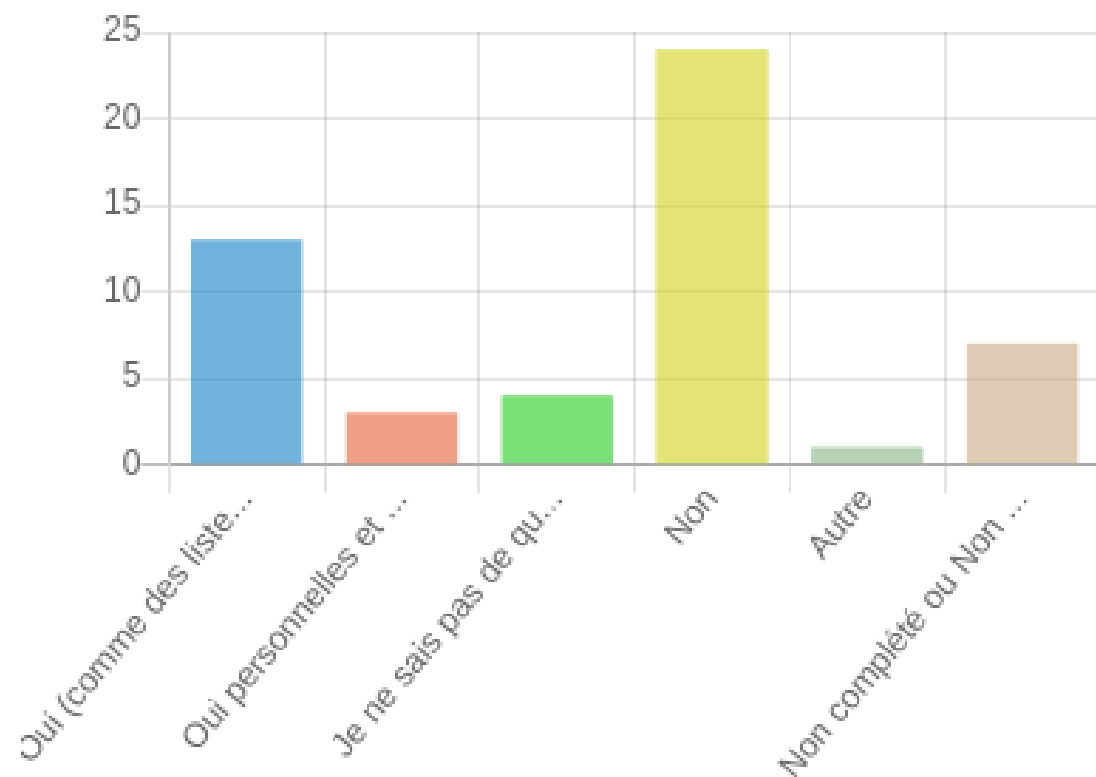


Proposition de fonctionnement

(sans être informaticien !)



Gérez vous des données personnelles ? Se référer au glossaire du mail pour les définitions



Nos interlocuteurs :

- Kodjo KOKOU, Juriste en protection des données au Service général de protection des données du CNRS : dpd.demandes@cns.fr
- Kevin GEIGER, Adjoint au délégué régional CNRS délégation Alsace : kevin.geiger@alsace.cns.fr

+ formations régulières

https://intranet.cns.fr/protection donnees/rgpd/formations/Pages/cafe_RGPD.aspx



Actions menées sur 1 an

Recensement

- Sondage « état des lieux des données et des usages dans l'UMR » (été 2024)
- Sondage « Volumétrie des données UMR » (28.02.2025)
- Inventaire appareillage et format de données (en cours)

Groupes de travail

- Stockage et la sauvegarde des données dans l'UMR BSC (10.01.2025)
 - Infoprox, Dnum, DU et chefs d'équipes
- Prise en compte des aspects relatifs au RGPD (28.01.2025)
 - Adjointe DU, ADELE, RGPD CNRS

PGD entité : en cours de réflexion/rédaction

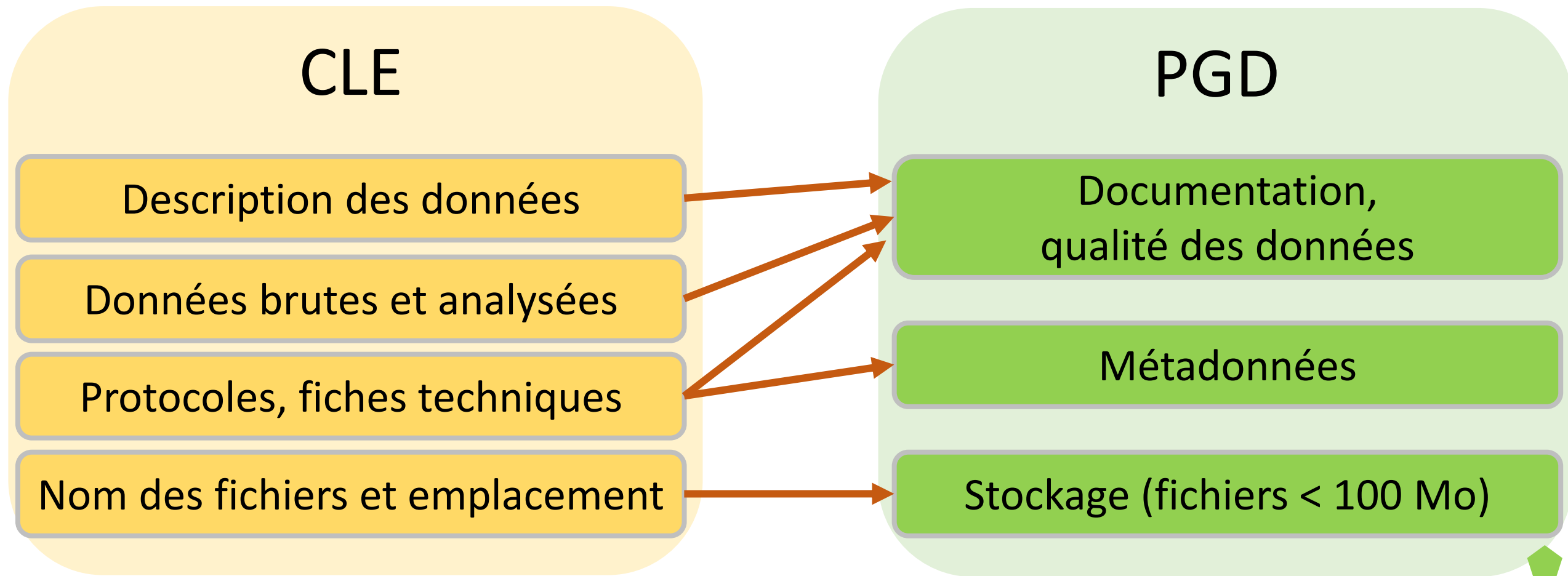
- Description des données
 - Cahier de laboratoire électronique
 - Recensement en cours
- Stockage et sauvegarde
 - Solutions techniques
 - Moyens humains
 - Budget



Cahier de laboratoire électronique (CLE)

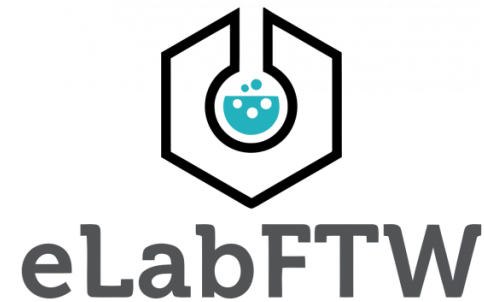
Relation PGD et CLE

La structuration de l'organisation dans le cadre d'un cahier de laboratoire électronique va nourrir le plan de gestion de données



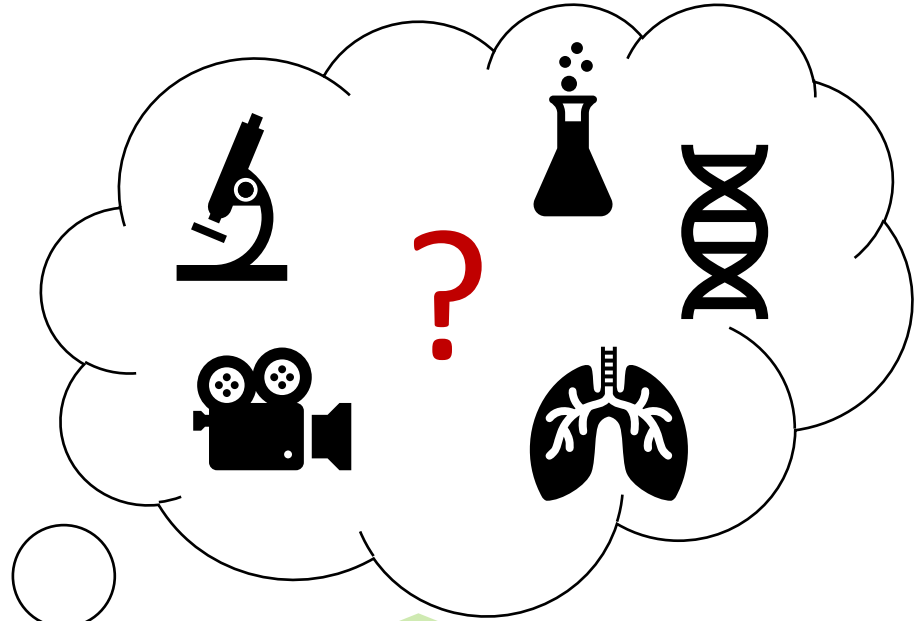
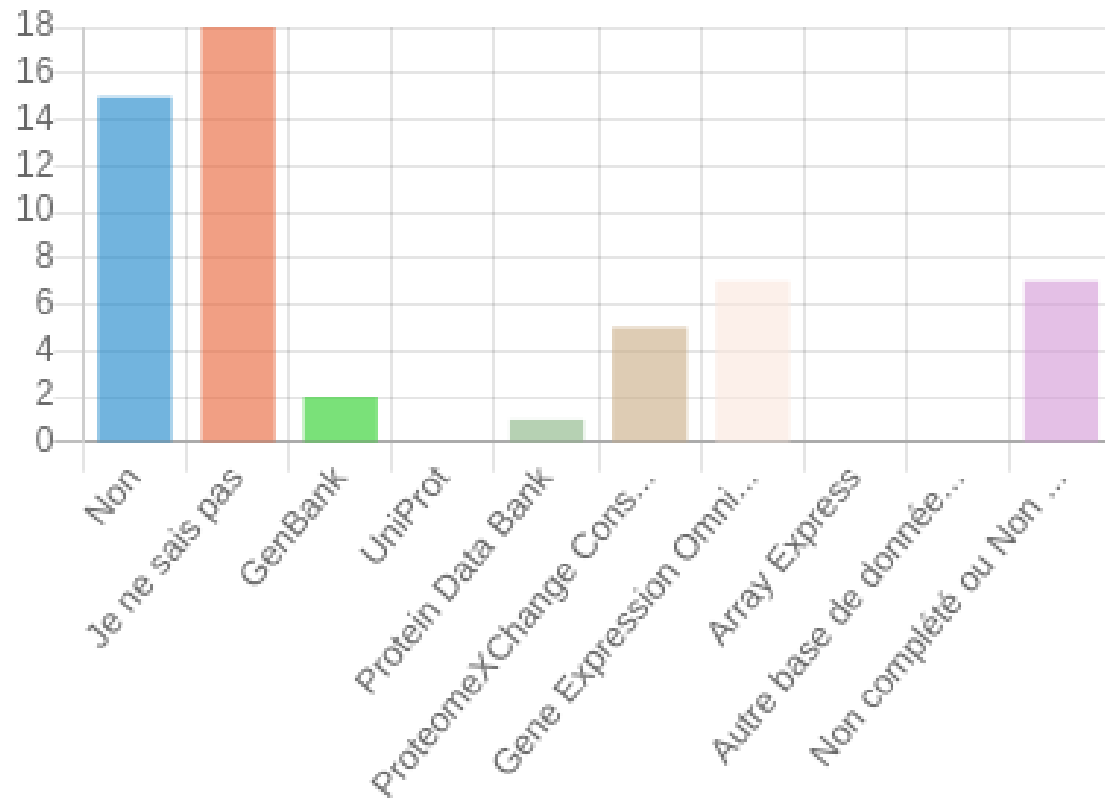
Cahier de laboratoire électronique

- Pilotes Cahier de Laboratoire Electronique (CLE)
- 32 utilisateurs volontaires depuis mars 2024 (1/3 de l'unité)
 - Principalement des permanents
 - 8 administrateurs formés (1 par équipe)
- Utilisateurs globalement satisfaits
- Webinaire enregistrés disponibles en ligne (avec les identifiants Janus) :
https://extra.core-cloud.net/projets/Cahiers_laboratoire_electroniques/SitePages/Description%20de%20l'offre.aspx
 - onglet « Description offre », tout en bas de la page web



Entrepôts de données

Avez vous (ou votre équipe) des données déposées sur des entrepôts publiques ? Pour rappel, Entrepôt de données : service en ligne permettant le dépôt, la description et l'accès à des jeux de données tels que GenBank, GEO, ProteomeXChange Consortium, ...



Nous sommes ouverts
aux suggestions
d'entrepôts



Bilan des actions menées sur 1 an



Recensement

- Sondage « état des lieux des données et des usages dans l'UMR » (été 2024)
- Sondage « Volumétrie des données UMR » (28.02.2025)
- Inventaire appareillage et format de données (en cours)



Groupes de travail

- Stockage et la sauvegarde des données dans l'UMR BSC (10.01.2025)
 - Infoprox, Dnum, DU et chefs d'équipes
- Prise en compte des aspects relatifs au RGPD (28.01.2025)
 - Adjointe DU, ADELE, RGPD CNRS



PGD entité : en cours de réflexion/rédaction

- Description des données
 - Cahier de laboratoire électronique
 - Recensement en cours
- Stockage et sauvegarde
 - Solutions techniques
 - Moyens humains
 - Budget



Pour conclure

Science Europe : modèle structuré	DMP OPIDoR	Institution	Modèle structuré de plan de gestion de données (PGD) basé sur le "Guide pratique pour une harmonisation internationale de la gestion des données de recherche-V2" (janvier 2021) de Science Europe. Organisé en 12 questions, il couvre les six exigences fondamentales pour une bonne gestion et le partage des données dans le respect des principes FAIR*. L'objectif de ce modèle est de rendre le contenu du PGD à la fois lisible par les humains et exploitable par les machines, afin de proposer de nouveaux services ou améliorer certains services existants. Le modèle structuré : • facilite la réutilisation automatique d'informations (projets ANR), • encourage l'utilisation d'identifiants (contributeurs, organisations, partenaires...), • propose des référentiels (standards de métadonnées, entrepôts de données, terminologies...), • permet d'identifier et de lister les types de coûts associés à la gestion des données, • simplifie et automatise les échanges d'informations avec les services impliqués dans la gestion des données (centre de stockage par exemple), • produit un format de PGD conforme aux recommandations RDA. Des recommandations sont proposées pour aider les chercheurs à renseigner ce modèle et à s'assurer que tous les aspects pertinents de la gestion des données sont effectivement couverts. <i>* Acronyme pour Facile à trouver, Accessible, Interoperable, Réutilisable.</i>	17/10/2024	 
Science Europe : modèle structuré pour les entités de recherche	DMP OPIDoR	Institution	Modèle structuré de plan de gestion de données pour les entités de recherche basé sur le "Guide pratique pour une harmonisation internationale de la gestion des données de recherche-V2" (janvier 2021) de Science Europe. Organisé en 12 questions, il couvre les six exigences fondamentales pour une bonne gestion et le partage des données dans le respect des principes FAIR*. L'objectif de ce modèle est de rendre le contenu du PGD à la fois lisible par les humains et exploitable par les machines, afin de proposer de nouveaux services ou améliorer certains services existants.	17/10/2024	 
Science Europe: structured template	DMP OPIDoR	Institution	Science Europe's Data Management Plan (DMP) based on Science Europe's "Practical Guide to International Alignment of Research Data Management - Extended Version" (January 2021). It covers the six fundamental requirements for good data management and sharing in accordance with the FAIR* principles.	17/10/2024	 
Science Europe: structured template for research entities	DMP OPIDoR	Institution	The objective of this model is to make the content of the DMP both human-readable and machine-actionable, in order to provide new services or improve some existing ones.	17/10/2024	 

Questions :

- Description des données
- Exigences légales et éthiques, code de conduite
- Stockage et sauvegarde des données
- Partage



Bilan : Actions... Réactions

On a toujours fait comme ça et ça marche très bien !

On a déjà un très bon serveur et je m'en occupe très bien.

Le collègue gère ça pour toute l'équipe, je ne me sens pas concerné

Je partage déjà mes données, je n'ai pas besoin de PGD pour ça.

J'ai déjà rempli un PGD et « c'est passé »

Qui valide les PGD, qui dit que c'est pas bien ?



Bilan : Actions... Réactions

Combien ça coûte ?
Qui paie ?

Vous passez trop de
temps sur le PGD

J'ai appris des trucs.

Quelles données sont
importantes à
sauvegarder ?

Vous êtes bien courageux
de vous lancer là-dedans !





- Manque de réponses : sondage été 2024 « état des lieux des données et des usages dans l'UMR » : 30% de répondants
- Nombreuses relances nécessaires
- Difficultés à faire changer les pratiques (« on a toujours fait comme ça »)
- Manque de moyens humains :
 - Pas d'informaticien dans l'UMR
 - Equipe Infoprox en manque d'effectifs
- RGPD : au début « jargon » et très théorique
- Pour l'instant : données personnelles « type RH » non gérées
- Réticence du responsable d'équipe à allouer du temps à cette activité





- Formations régulières et pertinentes (URFIST, ADELE et CNRS)
- Mise à disposition de 2 PGD de demandes ANR remplies
- Soutien de la direction
- Budget débloqué
- Infoprox se rend disponible et de bonne volonté
- Equipe ADELE, DNum réactive, à l'écoute ... et d'une grande aide

⇒ Investissement (temps, énergie) indispensable en + des activités de biologistes

⇒ Travail formateur, intéressant et utile à l'UMR

