Inserm

La science pour la santé
From science to health

Déploiement du cahier de laboratoire électronique à l'INSERM et nouvelles perspectives

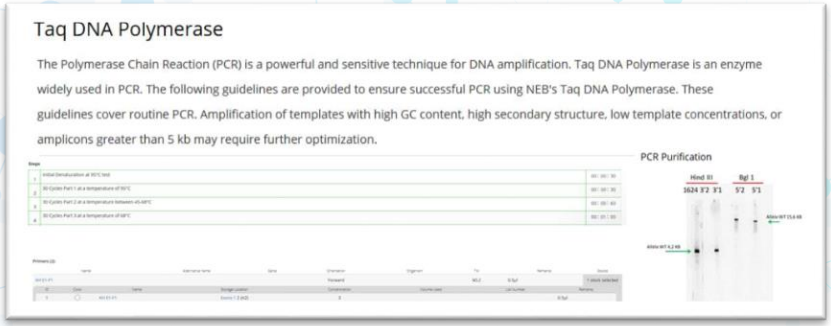
Paul Guy Dupré

Inserm, DSI

Pourquoi le cahier de labo électronique?

Version papier

Version numérique (ELN)



Journal de bord:

- Preuve juridique,
- Traçabilité, qualité,
- Gestion de la connaissance.

- Horodatage, versioning, signature électronique,
- Gestion des inventaires et instruments,
- Indexation, moteur de recherche,
- Intégration des données numériques,

\+ ***Travail collaboratif et accès distant,***

\+ ***Management des équipes et des projets.***

Comment s'est fait le choix entre les différentes solutions?

Pré-étude de 9 solutions

- **Labguru**, Biodata;
- **Hivebench**, Elsevier;
- **LabCollector**, AgileBio;

- Plateforme d'expérimentation

Objectifs

- Expérimenter une solution d'ELN et évaluer son impact.
- Choisir une solution dédiée à la biologie, collaborative, centrée sur le laboratoire et hébergeable en interne.

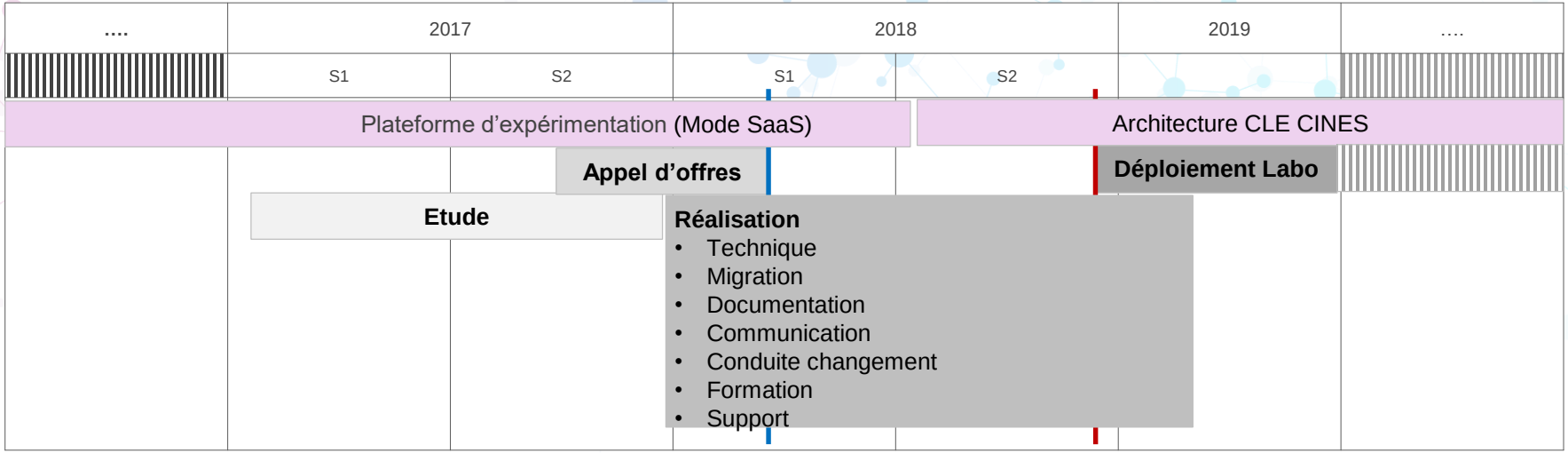
Périmètre

- 250 utilisateurs pendant 2 ans.

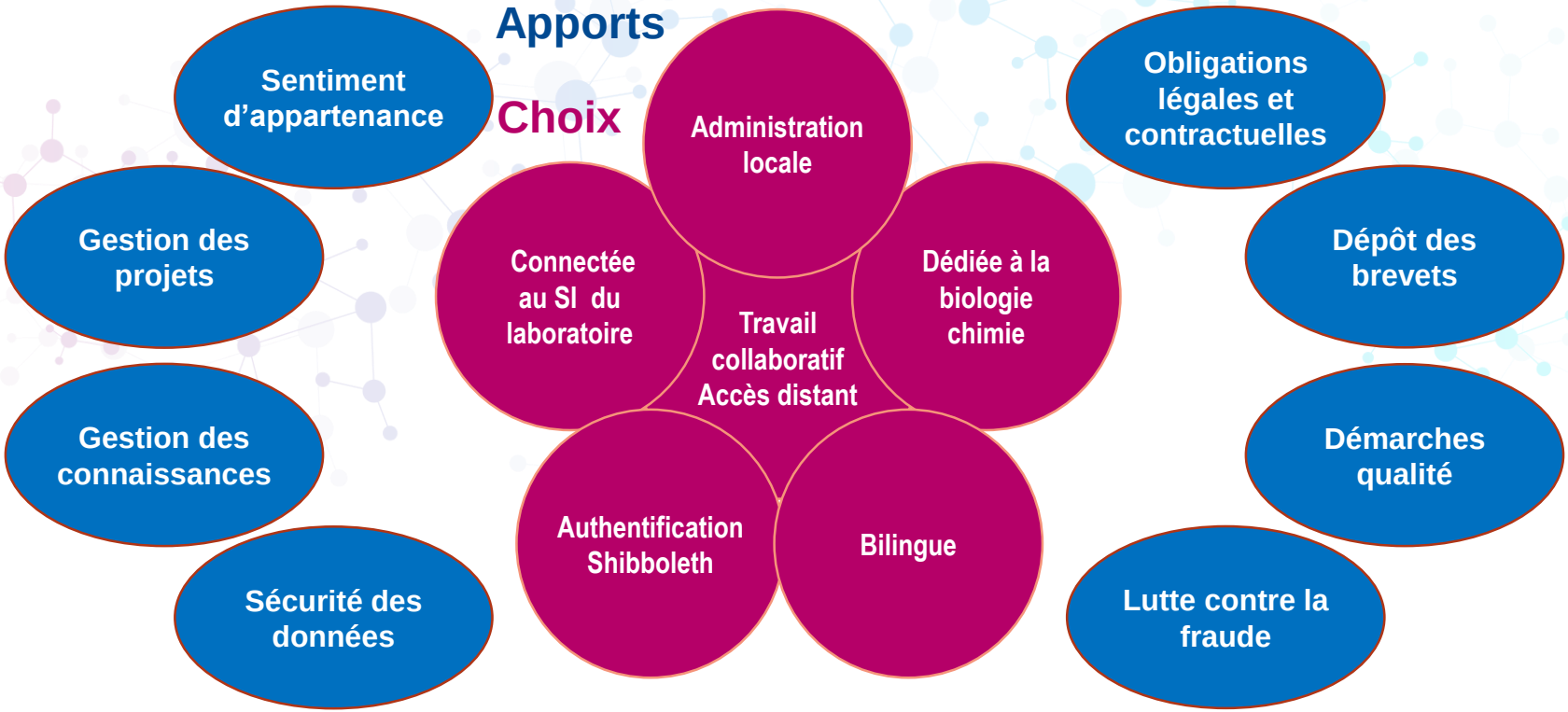
Evaluation

- 2 indicateurs,
- Bilan après 6 mois d'utilisation.

Projet CLÉ (suite à l'expérimentation)



Choix stratégiques et apports de la solution



Ce qui change dans le laboratoire

Le cahier n'est plus personnel, mais partagé

- Travailler en même temps à plusieurs sur une expérience
- Travailler dans plusieurs cahiers
- Centraliser la gestion d'un projet

Le cahier est accessible à distance

CLÉ permet la gestion des projets et de visualiser l'implication de chacun

CLÉ organise la gestion de la connaissance dans le laboratoire

- En cas de départ d'un membre, les données restent dans CLÉ, et ses projets peuvent être réaffectés

CLÉ offre la possibilité de gérer les inventaires et les équipements

Le responsable du laboratoire assure la confidentialité des données

Déployée sous l'autorité des directeurs et organisée autour du responsable du cahier(PI)

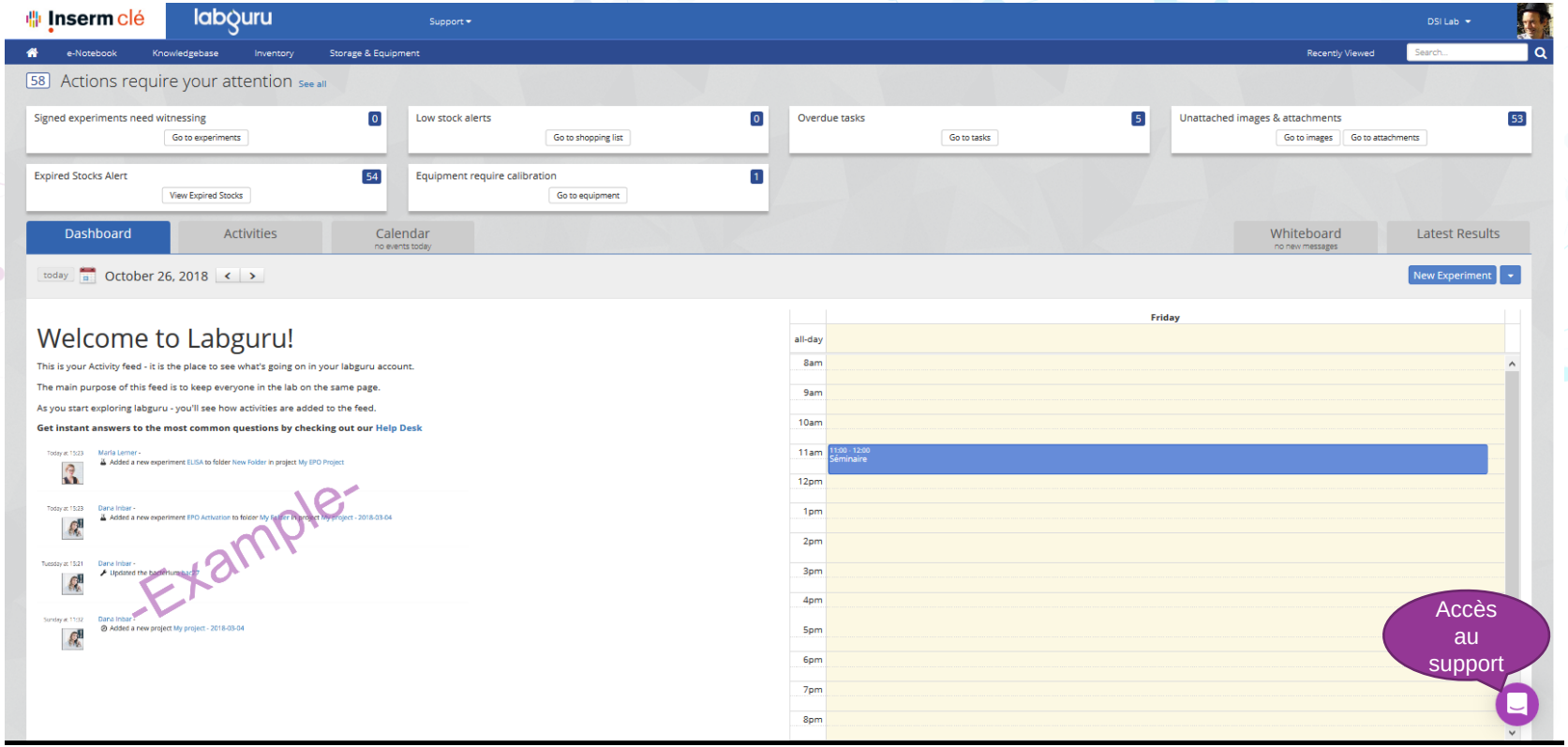
Co-administrée:

- DSI: assure en central l'hébergement, l'exploitation, la maintenance, la sauvegarde et l'archivage
- Responsable du cahier: assure (ou délègue) l'administration des comptes et des droits d'accès

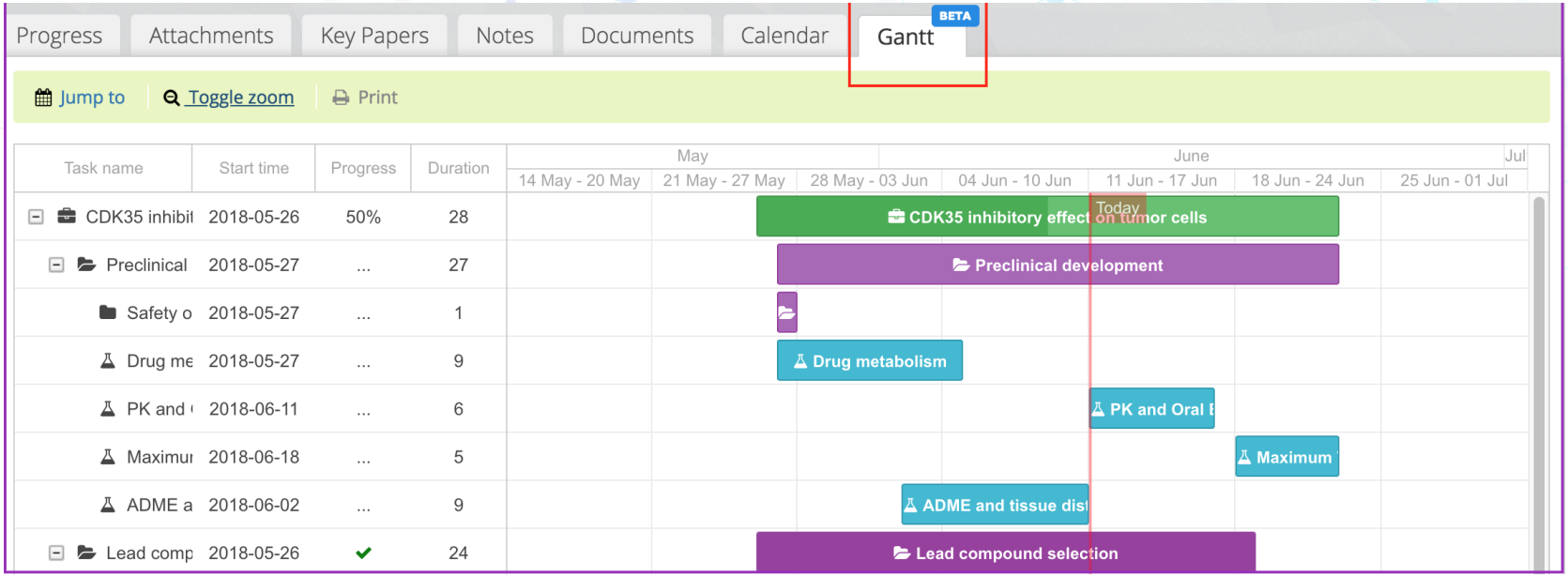
Présentation: Labguru la solution retenue par l'Inserm

- Présentation live
- Evolutions récentes des fonctionnalités

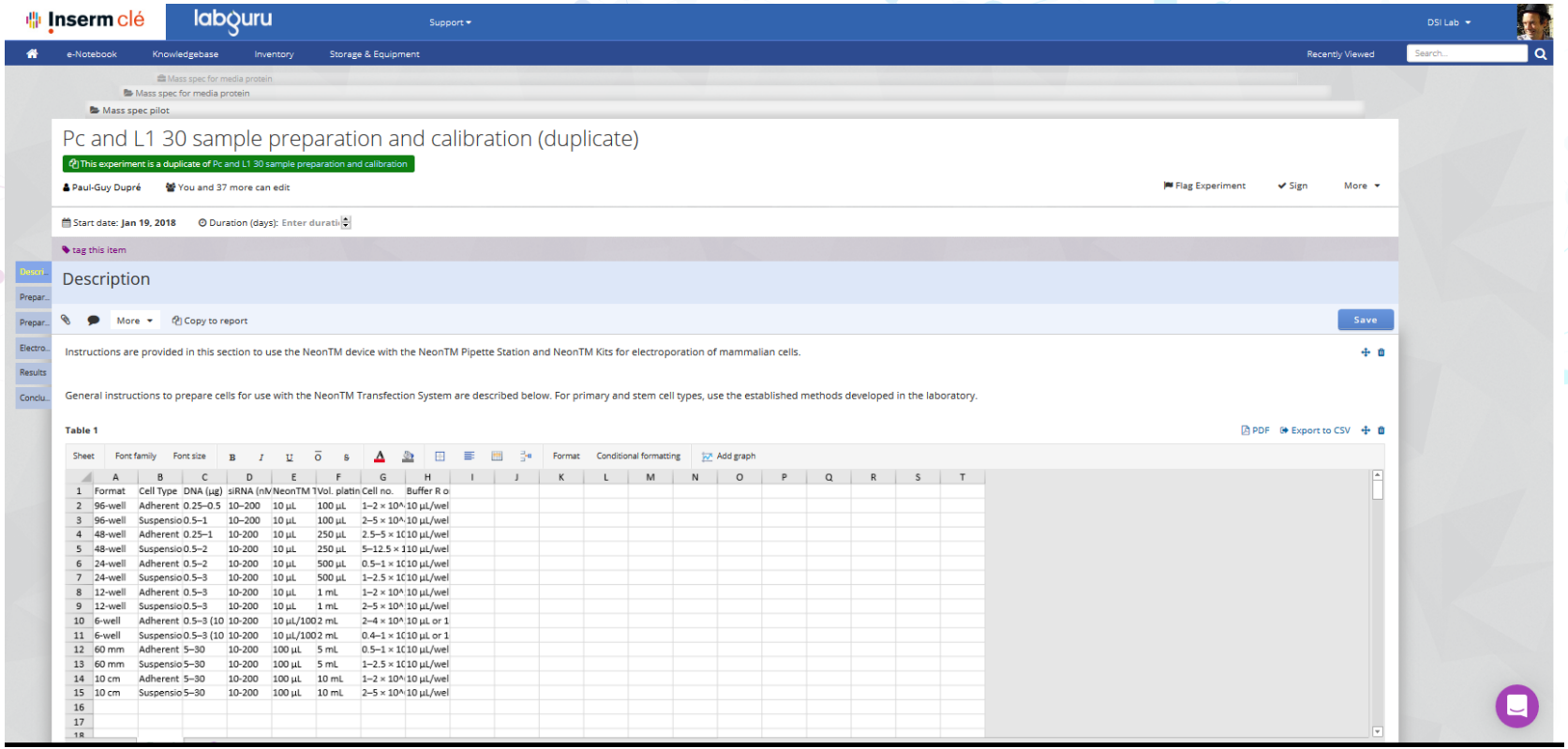
Evolution de la page d'accueil: Tableau de bord collaboratif



Evolution des projets: Piloter la progression et la planification

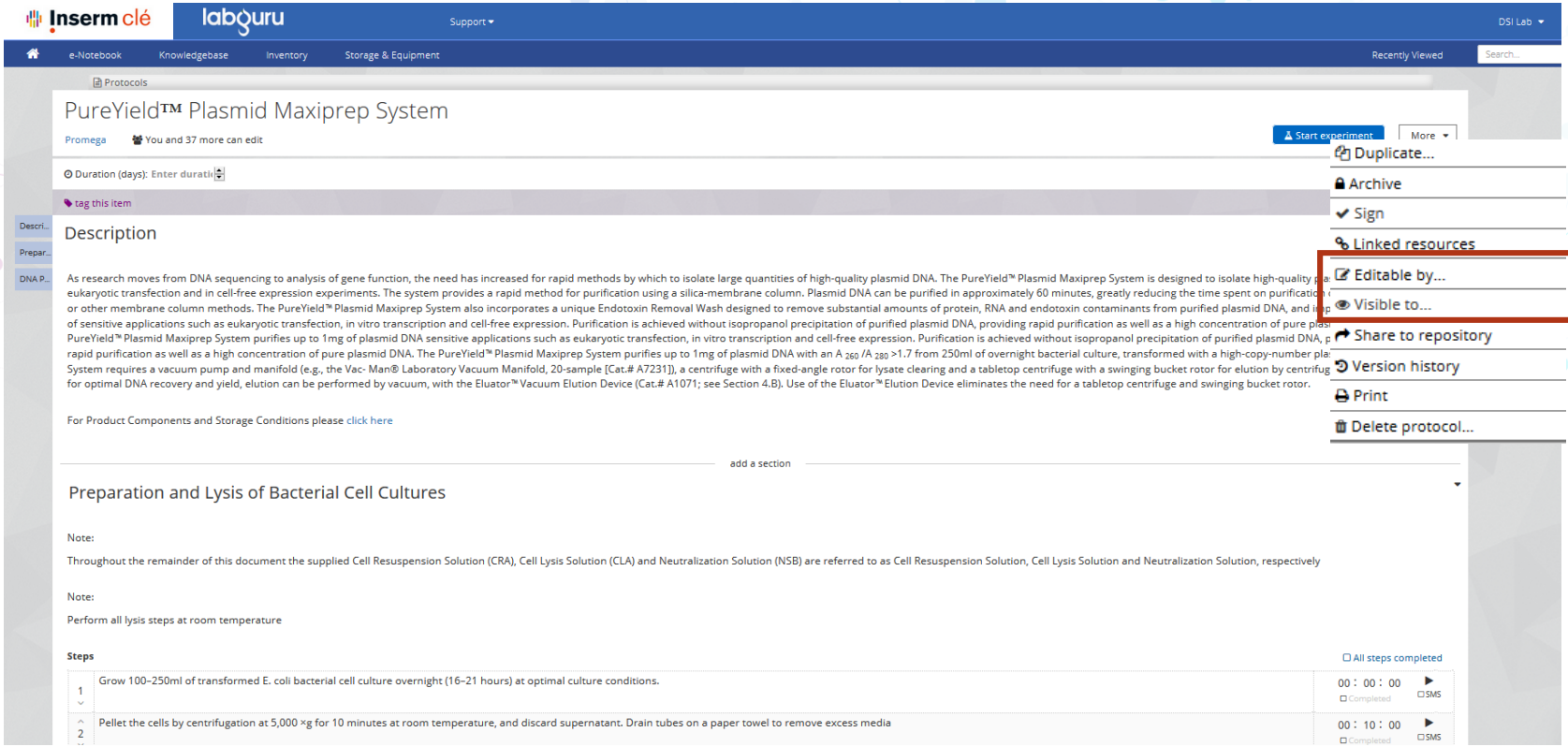


Evolution des expériences: intégration d'Excel dans l'éditeur

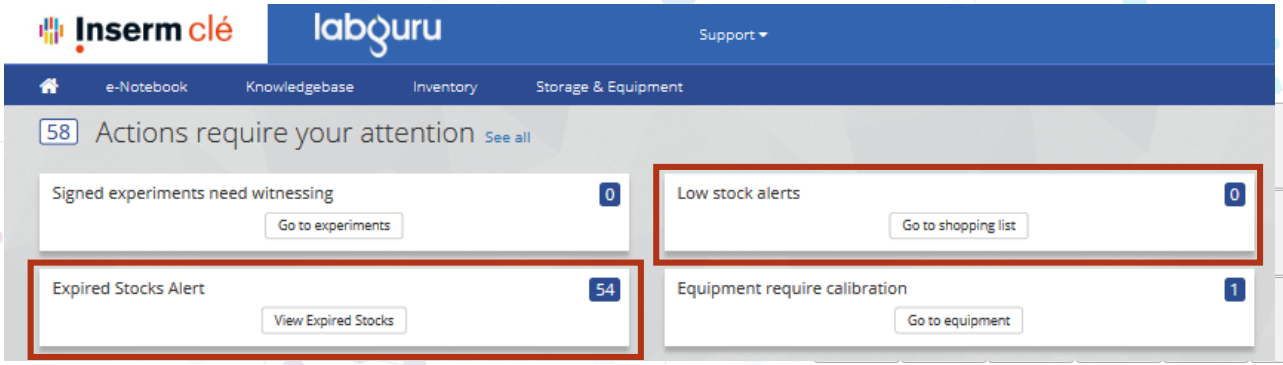


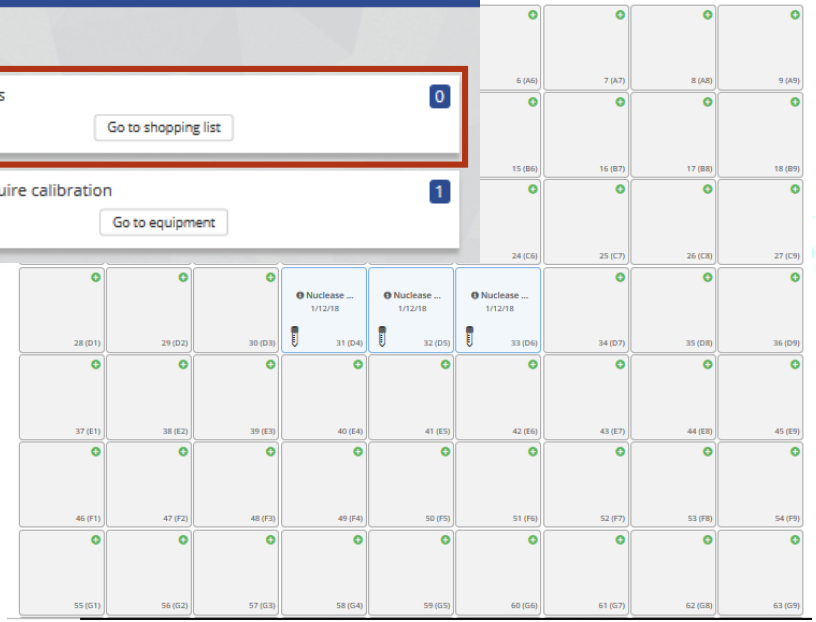
	A	B	C	D	E	F	G	H
1	Format	Cell Type	DNA (µg)	siRNA (nM	NeonTM T	Vol. platin	Cell no.	Buffer R o
2	96-well	Adherent	0.25–0.5	10–200	10 µL	100 µL	1–2 × 10^	10 µL/wel
3	96-well	Suspensio	0.5–1	10–200	10 µL	100 µL	2–5 × 10^	10 µL/wel
4	48-well	Adherent	0.25–1	10-200	10 µL	250 µL	2.5–5 × 10	10 µL/wel
5	48-well	Suspensio	0.5–2	10-200	10 µL	250 µL	5–12.5 × 1	10 µL/wel
6	24-well	Adherent	0.5–2	10-200	10 µL	500 µL	0.5–1 × 10	10 µL/wel
7	24-well	Suspensio	0.5–3	10-200	10 µL	500 µL	1–2.5 × 10	10 µL/wel
8	12-well	Adherent	0.5–3	10-200	10 µL	1 mL	1–2 × 10^	10 µL/wel
9	12-well	Suspensio	0.5–3	10-200	10 µL	1 mL	2–5 × 10^	10 µL/wel
10	6-well	Adherent	0.5–3 (10	10-200	10 µL/100	2 mL	2–4 × 10^	10 µL or 1
11	6-well	Suspensio	0.5–3 (10	10-200	10 µL/100	2 mL	0.4–1 × 10	10 µL or 1
12	60 mm	Adherent	5–30	10-200	100 µL	5 mL	0.5–1 × 10	10 µL/wel
13	60 mm	Suspensio	5–30	10-200	100 µL	5 mL	1–2.5 × 10	10 µL/wel
14	10 cm	Adherent	5–30	10-200	100 µL	10 mL	1–2 × 10^	10 µL/wel
15	10 cm	Suspensio	5–30	10-200	100 µL	10 mL	2–5 × 10^	10 µL/wel
16								
17								
18								

Evolution des protocoles: définir des droits d'accès

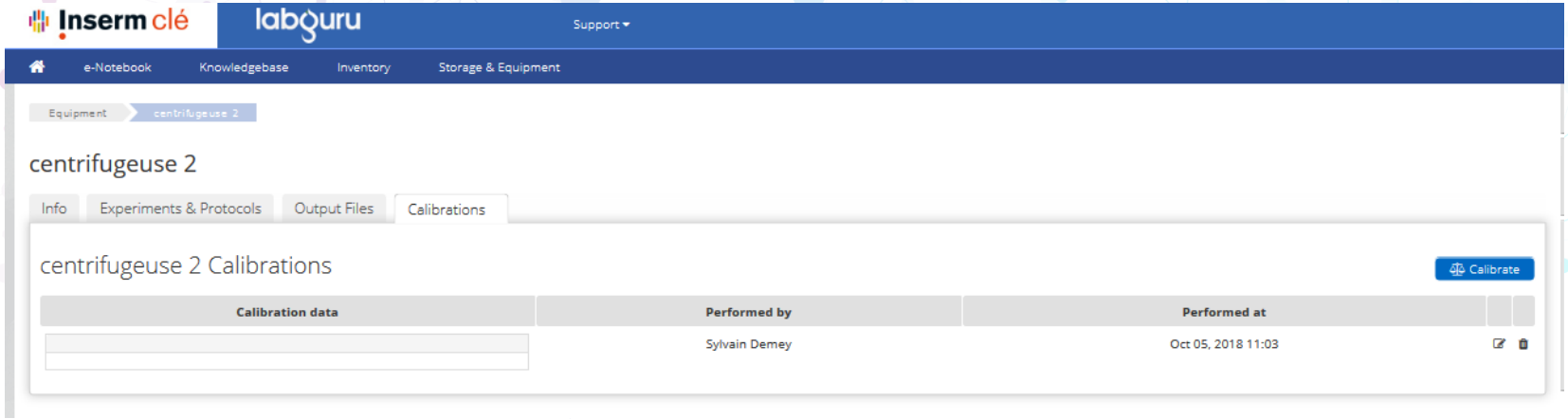


Evolution des inventaires: les alertes



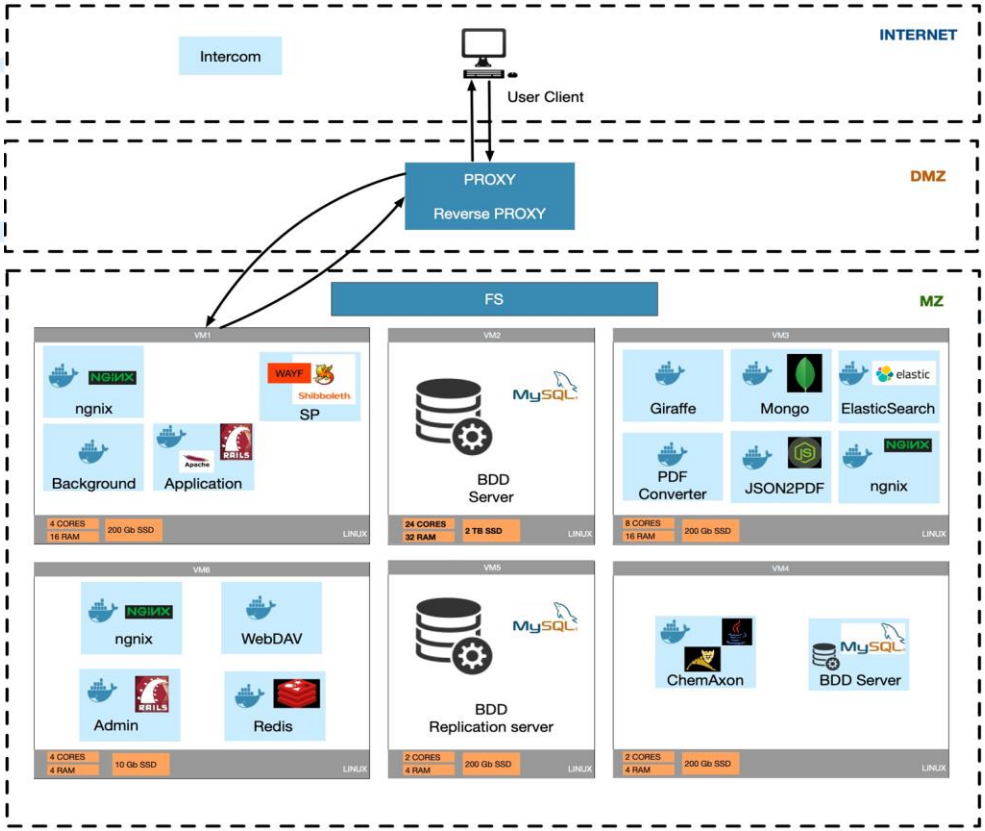


Evolution des instruments: la calibration



Architecture

- **Virtualisation**
  - 6 serveurs VM
  - 14 conteneurs applicatifs Docker
  - ➕ simplification des mises à jour
- **Technologies actuelles**
  - ElasticSearch, MongoDB, NGNIX, Redis, Json, MySQL
- **Authentification shibboleth**
  - ⛔ complexité du support multi-établissements



Déploiement

- **Un cahier pour**
  - un groupe de recherche travaillant sur des projets communs, autour d'un responsable
  - un plateau technique
  - un contrat industriel

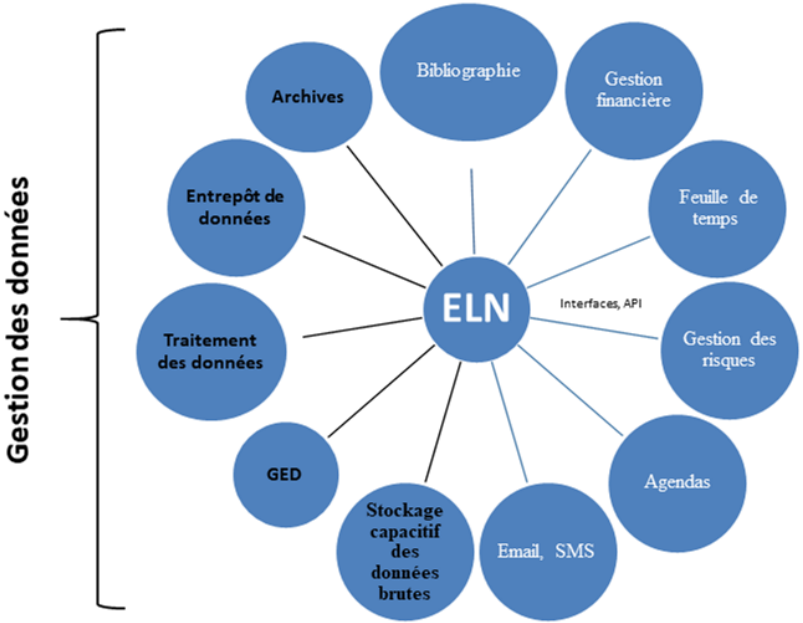
Étapes du déploiement dans une UMR
Inscription – validation
Réunion préparatoire
Demande – ouverture de cahiers
Réunion de lancement dans le laboratoire
Formation – support des utilisateurs

Bilan à 10 mois	
Structures en cours de déploiement	44%
Cahiers créés	550
Utilisateurs enregistrés	2500
Utilisateurs formés (e-learning)	250
Administrateurs formés (Webinar)	170

Rôle fédérateur de l'ELN

Le LIMS du laboratoire (Laboratory Information Management System)

- **Besoins locaux**
  - s'interfacer avec des applications métiers et des bases de données
- **Besoins institutionnels**
  - s'interfacer avec le catalogue de services scientifiques et administratifs



API pour échanger les données
(application programming interface)

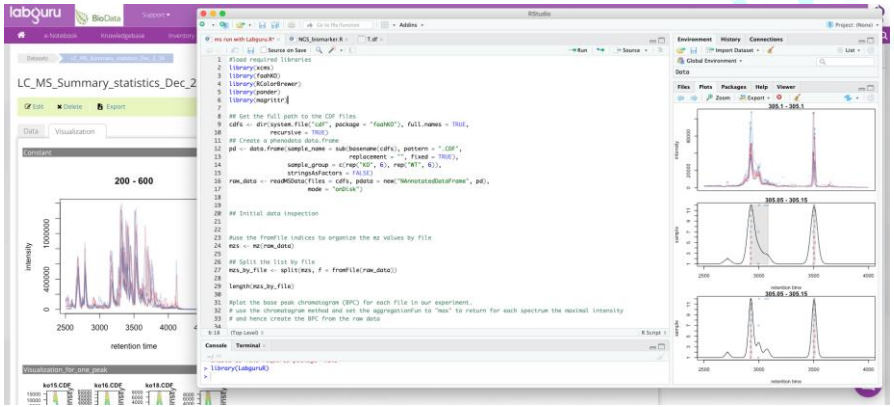
API de Labguru

- basées sur JSON / REST
- publiées sur https://my.labguru.com/api/docs, https://github.com/BioData/labguru-api-examples

- pour consulter, créer, modifier
  - projets de recherche
  - expériences
  - protocoles
  - stocks
  - équipements
  - documents annexes
  - DataSet, SOPs…

LabguruR

- script pour travailler depuis R
  - accéder aux DataSet depuis la Rconsole
  - enregistrer les résultats dans Labguru



Nouveaux enjeux : un ELN comme GED de la recherche

Améliorer la gestion de la donnée

- Retrouver et décrire la manière dont seront obtenues, documentées et analysées les données produites au cours d'un projet de recherche
- Garantir dès le début du projet, dans le cadre du DMP, la production de données de recherche FAIR

Améliorer la reproductibilité des expériences

- Garantir que des expériences ont été faites dans les mêmes conditions et que les démarches de mesures sont strictes
- Garantir la traçabilité et conserver le descriptif et l'ensemble des éléments associés à un projet
  - données brutes (besoins capacitifs)
  - résultats
  - moyens mis en œuvre (protocoles, logiciels et versions utilisées...)

Et le stockage?

Actuellement

- Le plus gros cahier utilisé depuis 2015 est de 60 Go
- Peu de données brutes
  - Contraintes réseaux et navigateurs

- Stockage en mode Fichier (NFS) sur des baies NetApp (FAS8020)

Cible

- Augmenter les quotas pour prendre en charge l'essentiel des données (de l'ordre de 100 To)
- Utiliser des agents de synchronisation automatique

- Nouvelle architecture de stockage capacitif et pérenne (Mode objet)