

CR de la réunion du CAN du 03/10/2022

Date de la prochaine réunion : 1er décembre 14h

Présent·e·s

Guillaume Launay

Jacques Brocard

Sandrine Hughes

Estelle Devevre

Olivier Gandrillon

Carine Rey

Jocelyn Turpin

Lucie Etienne

Laurent Gilquin

Excusée:

Cécile Hilpert

1. Stockage de données

Généralités

La problématique du stockage de données est récurrente, globalement mal comprise/questionnée et complexe à aborder. Il existe différentes solutions :

LBMC, CIRI, RDP : le serveur partagé BIODATA (20 To) administré par la DSI de l'ENS. Le serveur est accessible seulement depuis le réseau ENS (pas par VPN). Le serveur peut être utilisé pour du stockage pérenne.

IVPC - serveur interne de 12 To, utilisable pour stockage pérenne.

IGFL : un serveur FTP pour stockage temporaire (non pérenne) de donnée de séquençage, administré par Chérif Kabir. Accès externe possible dans le cadre de collaborations avec autres labos, entreprises, etc.

tous : le CCIN2P3 pour archivage, les archives publiques (ebi, ncbi).

La majorité des services disponibles a été recensé par Laurent Modolo : http://www.ens-lyon.fr/LBMC/intranet/services-communs/pole-bioinformatique/ressources/data_management.

Les questions/besoins émergents des différents labos interrogent sur la définition d'une stratégie globale de stockage. Il paraît à minima nécessaire d'informer les éventuels utilisateurs des moyens de stockage à disposition.

La question des stockages et des gestions de données est aussi abordée lors de la mise en place d'un DMP Data Management Plan (financements nationaux ou européens...). Solliciter les équipes maintenant via un sondage, pour leur proposer par la suite un cadre pour ces DMPs, semble particulièrement cohérent.

β mise en place d'un sondage pour mieux cibler les besoins/utilisations

IMPORTANT : La problématique du stockage des données ne doit pas être à la charge des plateaux d'analyse.

Sondage

Les questions du futur sondage sont disponibles à cette adresse :
https://mensuel.framapad.org/p/data_management_ums-9wtl?lang=fr
La mise en forme du sondage est à la charge de Sandrine Hughes.

Points importants :

le sondage sera diffusé via l'UMS à destination des équipes de recherche de chaque labo,
le sondage doit permettre de remplir deux champs : nom du laboratoire/institut, nom de l'équipe de recherche,
identification de deux types de publics distincts : permanents et non-permanents. Le sondage doit contenir des questions adaptées à chaque public (quid de l'intégration de cet aspect dans l'élaboration du sondage ?)
le sondage est à envoyer aux plateaux d'analyse pour réponse en tant qu'équipe de recherche

2. Formations

Perspectives : Recycler les formations/guidelines faites à l'INRAE pour/par les réseaux référents données opérationnels et stratégiques?

En cours

Formation "R pour les débutant(e)s" du 03/10 au 19/12

organisatrice : Carine Rey

9 séances programmées + séances supplémentaires possibles

4 groupes dont 1 anglophone,

3 formateurs par groupe

40 inscrits (même nombre que l'année dernière) dont quelques permanents

ß Essayer de programmer plus en avance la réservation des salles de formation pour les futures formations R.

Formation "UNIX ligne de commandes" du 03/10 au 19/12

organisateur : Laurent Modolo

9 séances programmées

1 groupe de 18 personnes

Formation ImageJ

organisateur : PLATIM

6 séances du 08/11 au 15/12 pour deux groupes / semaine, un francophone, un anglophone

À venir

Formation "Cytoscape" pour la visualisation et l'analyse de graphe d'interactions protéines-protéines :

concept : les participants viennent avec leur jeu de données à analyser/visualiser

sondage préliminaire pour recenser les diverses ressources bioinformatiques utilisées

potentielles discussions avec le LBBE pour mutualiser la formation

nécessité de bien cibler les objectifs de la formation au vu des multiples usages du logiciel

Questions diverses

Aucune