

## Visio du CAN, mercredi 24 février à 11h

*Excusé : Olivier Ali, RDP*

*Présent-e-s*

- Ralf EVERAERS - CBP/ENS de Lyon, Laboratoire de Physique UMR 5672
- Catherine Moali, DR CNRS, LBTI
- Cécile Hilpert, IE CNRS, MMSB/PSF
- Guillaume Launay, MCU bioinformatique structurale, MMSB
- Nicolas Goudemand, physicien, chef d'équipe IGFL + prof. ENS
- Lucie Etienne, CR CIRI, correspondante bioinformatique du CoPil BIBS du CIRI
- Jacques Brocard, IR INSERM, responsable plateau imagerie et microscopie
- Laurent Modolo, IR CNRS au LBMC, responsable du hub "biocomputing"
- Jocelyn Turpin, CR-INRAE, IVPC
- Estelle Devêvre, IE INSERM, responsable du plateau de cytométrie
- Sandrine Hughes, CR CNRS, responsable séquençage NGS à l'IGFL
- Olivier Gandrillon, DR CNRS LBMC/UAR

Olivier : bravo pour le gros succès de sondage autour des formations !

Sandrine : je résume les résultats :

- 82 demandes de formation "beginner" pour R
- 32 demandes de formation avancée en R, mais aussi en imagerie ou en NGS
- format visio OK ; préférence en un bloc ou plusieurs jours selon les répondants
- pas de préférence sur le jour de la semaine (à peine plus pour lundi)
- 2/3 de demandes en français, 1/3 en anglais, sachant que des répondants se sont inscrits dans les deux catégories.

Olivier : toute l'affaire maintenant sera de trier ces demandes et d'y répondre mais de quelle façon, où, avec quels moyens, par qui, etc ! En attendant, Ralf va se présenter.

Ralf : Pr Physique à l'ENS, directeur du mésocentre de calcul CBP/PSMN après Hervé Gilquin qui est maintenant directeur adjoint de l'IFB. La question que je me pose concerne l'utilisation des moyens du mésocentre PSMN vs IFB ? PSMN fait partie de l'EQUIPEX de l'IFB et propose des applications sur GPU au CBP. Est-ce une expérience à étendre ? Même question pour le projet "OpenStack" ou bien la mise en place de serveurs Jupyter pour R ?

Laurent : par défaut, l'utilisation des moyens IFB à l'ENS se dirige vers l'IFB core plutôt que le PSMN, mais on peut le modifier, surtout s'il s'agit d'une recommandation du CAN ? Ce qu'on apprécie à l'IFB c'est l'utilisation de machines virtuelles à jour pour la dernière version de R Studio par exemple.

Ralf : je me demande si l'IFB dispose des informations de connexion de ses utilisateurs, la durée, combien de formation etc ?

Laurent : oui vraisemblablement. On peut spécifier l'utilisation du PSMN mais on ne saura pas s'il est saturé.

Olivier : Pour en revenir aux formations à organiser, combien de sessions et avec combien de formateurs, faudrait-il organiser pour répondre à la demande ?

Laurent : au dernier cycle de formation "débutants" sur R en présentiel au LBMC, j'avais ~ 15 personnes + 1-2 IE et thésards expérimentés en R dans la salle pour m'assister, pendant 1h par semaine sur 12 semaines et ça s'est très bien passé. Avec une formation Unix en distanciel par contre où j'avais 5 participants, j'ai réalisé que dans ces conditions, il n'était pas possible de s'assurer que chacun suive correctement. Pour des formations avancées par contre, les gens sont plus ou moins autonomes et quelques formateurs (4 ?) peuvent suffire pour organiser les questions et y répondre pour un groupe beaucoup plus grand (300 ?).

Sandrine : Nous avons reçu une série de cours en R et stats par le passé à raison de 1h / mois sur un thème donné avec un formateur + 2-3 personnes aidantes et c'était très bien.

Olivier : le format 1 formateur / 5 étudiants ou plutôt une équipe de 3 formateurs / 15 étudiants a l'air de bien fonctionner.

Laurent : oui et sur 12 semaines on est passé d'utilisateurs qui n'avaient jamais codé à des personnes autonomes dans la manipulation de tableau de données complexes, la réalisation de tous les graphes possibles sous ggplot, etc !

Ralf : dans ce cas, est-ce que ces formations ne seraient pas à rapprocher de la formation continue pour les adultes et/ou de la formation des étudiants via le master ?

Laurent : c'est la raison pour laquelle nous voulons nous rapprocher de l'IFB et Elixir.

Lucie : la notion de programmation est abordée en master mais c'est très variable selon les formations/master en Biologie : Certaines formations sont très approfondies alors que d'autres ne donnent qu'une présentation (overview) ou qq heures de TP généralistes.

Ralf : savez-vous que le département de biologie de l'ENS aménage un espace numérique où il serait possible de former des étudiants sur des vrais outils de recherche ?

Olivier : attention, on arrive là aux limites du champ d'action du CAN...

Nicolas : moi je suis très intéressé par l'intervention de Ralf car nous sommes en pleine réflexion des maquettes d'enseignement du département de bio. Pour info, une nouvelle UE d'application de stats avancées sur des données d'écologie collectées sur le terrain pourrait voir le jour. Ainsi qu'une nouvelle UE de M1 pensée comme un tremplin entre la biologie et la physique, notamment pour aider les étudiant.e.s de bio qui s'orientent ensuite en M2 vers le parcours Systèmes complexes. Renforcer le lien entre la recherche et l'enseignement est au cœur de la formation à l'ENS.

Lucie : au CIRI (et des membres LBMC se sont joint aux groupes), nous avons eu une formation R/NGS (analyses RNAseq spécifiquement) à distance en deux groupes de 12 personnes environ pour 2 formateurs (Marie Cariou et Antoine Corbin) en 3 x 3h sur 3 semaines (précédées de 2h de pré-formation pour ceux qui voulaient se familiariser avec l'outil command-line auparavant). Ce qui nous

a marqué c'est la grande diversité de profils ! De plus, nous avons prolongé cette formation par du suivi de projet personnalisé par les IR / CR experts locaux : il n'est pas raisonnable d'espérer que les apprenants puissent être autonomes sur un délai si court !

Nicolas (à Ralf) : serait-il possible d'imaginer en NGS, imagerie, etc, ce qui est déjà proposé au CBP en matière de modélisation ? Pour préciser ma question : je cherchais à obtenir un retour d'expérience de Ralf et du CBP en ce qui concerne les suivis de projet et voir si on peut se servir de cette expérience de suivi en modélisation pour préciser notre offre au niveau du CAN.

Ralf : nous pouvons offrir un socle informatique homogène mais sans le personnel compétent pour ces formations puisqu'il est dans vos labos - et c'est tant mieux ! C'est en cela que la mutualisation peut le mieux fonctionner, par échange entre la mise à disposition de matériel vs la mise à disposition de savoir-faire, de compétences.

Lucie : vu le nombre de demandes, est-il bien nécessaire de proposer des formations localement ? Sauf dans le cas de suivi de projets par des experts locaux ? Cela pourrait prendre la forme d'une "clinique" où l'on se retrouve à un moment fixé à l'avance pour accueillir les questions des étudiants.

Olivier : il y a aussi les formations initiales en R déjà prodiguées par l'IFB ? xxx Quoiqu'il en soit, si on imagine réunir trois groupes de formateurs autour de Lucie, Sandrine et Laurent, on peut espérer former 45 participants, ce qui est déjà bien ! Il faudrait dans ce cas disposer d'un matériel homogène pour que les étudiants puissent naviguer d'un groupe à l'autre ?

Laurent : pas forcément car je crois que le fait de s'adapter au rythme des apprenants plutôt qu'à celui de la formation constitue un gros avantage pédagogique.

Sandrine : et il est toujours intéressant pour un formateur de suivre un même groupe pendant plusieurs sessions.

Laurent : je dispose de quelques supports en ligne et diaporamas mais Guillaume avait parlé d'outils plus interactifs qu'il serait intéressant de mettre en oeuvre en distanciel ?

Guillaume : des outils interactifs utilisés en master ? Pas sûr qu'ils soient adaptés et les refaire représente un gros investissement en temps de la part des formateurs !

Olivier : donc il nous reste la question du tri des demandes pour retenir 45 participants et de la constitution des "équipes" de formateurs.

Laurent : je vous propose de travailler en amont en organisant une réunion avec les personnes pressenties d'une part et en partageant le matériel envisagé d'autre part ?

Lucie, Sandrine : OK ; on va contacter les personnes intéressées et les envoyer vers toi.

Ralf : je reviens un instant sur le modèle économique de l'IFB. Pas de facturation actuellement mais cela pourrait changer prochainement, raison supplémentaire de revenir dans le giron de PSMN :)

Nicolas : Je l'ai simplement mentionné dans le chat mais le concept de classes inversées est à la mode en enseignement. L'idée est de créer ou d'utiliser une série de cours en ligne que les étudiant.e.s vont consulter en amont. Les heures de 'cours' sont consacrées aux questions des étudiant.e.s et à des projets d'application collaboratifs. L'avantage que j'y vois ici, c'est que les cours de base existent déjà

sur le net et qu'on a discuté du fait que le suivi individuel serait notre valeur ajoutée. On pourrait donc utiliser une même base de cours en ligne potentiellement déjà existante pour tous les groupes de formateurs/apprenants et focaliser notre action sur la partie 'questions/projets'. Ca peut aussi permettre de s'adapter à des vitesses d'apprentissage ou de niveau différents au sein du groupe. Possibilité les faire travailler ensemble en mixant a minima les niveaux pour qu'ils et elles s'entraident autour des questions qui les intéressent (travaux de préparation des données pour la NGS ou pour l'imagerie, etc). Les mêmes groupes (ou pas) peuvent enchaîner sur une formation avancée ensuite de façon assez naturelle.

En conclusion :

1. La prochaine réunion du CAN aura lieu le lundi 15 mars à 11h00.
2. En attendant, un groupe de formateurs-rices se réunira autour de Laurent pour commencer à drafte ce qui pourrait constituer le cours R for beginners. Rappel : 3 groupes de formateurs / 3 horaires différents par semaine (chaque groupe a son horaire) (éventuellement deux langues différentes, genre deux groupes feront en français, un en anglais ?); un matériel commun ; chaque groupe peut aller à son allure. On vise à former 45 personnes dans un premier temps.
3. Il reste à discuter le 15 au moins la partie R plus avancée, et la question de la cliniCAN.