

Journée Système (JoSy) : La reproductibilité en pratique : méthodes et outils

Rapport sur les contributions

ID de Contribution: **0**

Type: **Non spécifié**

Présentation ARAMIS - DR7 - Univ-Lyon 1 / ICBMS

jeudi 23 mai 2019 09:00 (15 minutes)

Orateur: M. HURTEVENT, Vincent (ARAMIS)

ID de Contribution: 1

Type: **Non spécifié**

Keynote : Les enjeux de “la recherche reproductible”

jeudi 23 mai 2019 09:15 (45 minutes)

Orateur: HINSEN, K. (Centre de Biophysique Moléculaire, CNRS Orléans)

ID de Contribution: 2

Type: **Non spécifié**

SIDUS : c'est un "grand pas" pour la déduplication. Un unique "petit pas" pour la reproductibilité ?

jeudi 23 mai 2019 10:00 (30 minutes)

Au delà de son aspect "pratique" ("on n'installe plus les machines, on les démarre !"), SIDUS apporte une reproductibilité totale du système d'exploitation, dans le temps (sur une même machine à des instants différents) ou dans l'espace (sur des machines déployées au même instant). A travers quelques exemples du Centre Blaise Pascal, nous verrons que, lorsque des variabilités (notamment temporelles) apparaissent sur des installations sous un même SIDUS, les origines sont à rechercher ailleurs : matériel, BIOS, conditions climatiques... Nous concluerons que, si des solutions à base de conteneurs présentent la racine d'un système comme le graal de la reproductibilité, il n'en est rien : SIDUS vous prouvera le contraire !

Orateur: QUEMENER, E. (ENS Lyon)

ID de Contribution: 3

Type: **Non spécifié**

La reproductibilité au service de la Biologie computationnelle

jeudi 23 mai 2019 10:45 (45 minutes)

Les outils de reproductibilité choisis ont tous des équivalents fonctionnels (docker/singularity, snakemake/nextflow, jupyter/Rmarkdown, ...)

Une étude récente publiée dans Nature a montré que près de 70% des expériences en Biologie ne sont pas reproductibles. Il est donc indispensable de mettre en place des bonnes pratiques afin de garantir l'intégrité des données et la reproductibilité des résultats d'analyse. Concernant les données, les principes FAIR-data sont de plus en plus utilisés. Ces mêmes principes peuvent être détournés au service des analyses pour garantir des résultats identiques à partir d'un même jeu de données et au cours du temps. L'objectif de cette présentation est de proposer un panel de fonctionnalités permettant de rendre reproductible une analyse complète de bioinformatique. L'exemple présenté a pour but de sélectionner des gènes qui ne se comportent pas de la même façon entre deux conditions expérimentales. Les fonctionnalités présentées ne sont pas dépendantes de cet exemple. En effet, elles peuvent être appliquées à n'importe quelle autre question biologique. Brièvement, nous récupérons les données depuis les bases de données publiques (ENA/SRA), nous réalisons une analyse reproductible avec un système de workflow (snakemake) dans un environnement virtuel (docker) dont l'ensemble du code, versionné (git), est disponible en open source (Github et docker-hub). La visualisation des résultats est dynamique (shiny app) et un rapport (Rmarkdown) en pdf ou html est disponible. Il regroupe les résultats de l'analyse et détaille l'ensemble des paramètres choisis par l'utilisateur.

Auteur principal: DENECKER, T. (I2BC, Gif-Sur-Yvette)

Co-auteur: TOFFANO-NIOCHE, C. (I2BC, Gif-Sur-Yvette)

Orateur: DENECKER, T. (I2BC, Gif-Sur-Yvette)

ID de Contribution: 4

Type: **Non spécifié**

Panorama des solutions de diffusions et d' installations des codes logiciels dans un contexte HPC largement multi-utilisateurs

jeudi 23 mai 2019 11:30 (30 minutes)

Orateurs: BOUTTIER, P.A. (UGA Grenoble); LOUVET, V. (UGA Grenoble)

ID de Contribution: 5

Type: **Non spécifié**

Lightning talk: Pipelines nextflow

jeudi 23 mai 2019 14:00 (7 minutes)

Orateur: GANOFKY, J. (ENS Lyon)

ID de Contribution: 6

Type: **Non spécifié**

Présentation du MOOC Recherche Reproductible

jeudi 23 mai 2019 14:15 (45 minutes)

Utilisation d'un serveur Gitlab pour réaliser tous les exercices du MOOC, d'un espace personnel Jupyter par apprenant (Jupyterhub) avec un système de gestion de version simplifié.

Orateurs: ROSPARS, B. (INRIA Grenoble); FARHI, L. (INRIA Grenoble)

ID de Contribution: 7

Type: **Non spécifié**

Au-delà des conteneurs: environnements logiciels reproductibles avec GNU Guix

jeudi 23 mai 2019 15:00 (30 minutes)

Orateur: COURTES, L. (INRIA Bordeaux)

ID de Contribution: 8

Type: **Non spécifié**

La plateforme web "Virtual Imaging Platform" pour une science ouverte et reproductible.

jeudi 23 mai 2019 17:00 (30 minutes)

Orateur: POP, S. (Creatis Lyon)

ID de Contribution: 9

Type: **Non spécifié**

Bistro: a library to build large-scale workflows in computational biology (OCaml)

jeudi 23 mai 2019 16:30 (30 minutes)

Orateur: VEBER, P. (UMR 5558, Lyon)

ID de Contribution: **10**

Type: **Non spécifié**

Execo a library to manage unix processes on thousands of remote hosts

jeudi 23 mai 2019 15:45 (45 minutes)

It is well designed for:

- prototyping experiments on distributed systems
- automatize admin tasks
- create reproducible experiments

Orateur: POUILLOUX, L. (Ecole centrale, Lyon)

ID de Contribution: 12

Type: **Non spécifié**

Lightning talk : Portail web pour soumettre et formater des jeux de données / package R

jeudi 23 mai 2019 14:08 (7 minutes)

Orateur: DRAY, S. (LBBE)