



February 11-12, 2021 | Irisa | Web Conf

JOURNÉES DKM

Annual meetings

C O N F E R E N C E P R O G R A M

1 1 F E B R U A R Y 2 0 2 1

13:45 – 14:10 **ACCUEIL ET PRÉSENTATION DU PROJET DU DÉPARTEMENT**
Anne Siegel

14:10 – 14:30 **PRÉSENTATION DE DEUX ARTICLES FIL ROUGE**

Thomas Guyet (Lacodam). "NegPSpan: efficient extraction of negative sequential patterns with embedding constraints", KDD 2018.

[LINK](#)

David Gross-Amblart (Druid). "Overlapping Hierarchical Clustering (OHC)", IDA 2020.

[LINK](#)

Objectif/critère d'évaluation du travail des groupes

14:30 – 14:45 **PAUSE**

14:45 – 16:15 **REPRODUCTIBILITÉ : EXPOSÉS INVITÉS ET POSITIONNEMENT DES ÉQUIPES**

(1) Nicolas Rougier : présentation de Re-science

(2) Arnaud Legrand: intérêt et utilisation d'un cahier de labo en ligne pour la reproductibilité

(3) Olivier Collin : Reproductibilité et gestion des données sur la plate-forme GenOuest

(4) Table-ronde : politique de reproductibilité des équipes, outils à disposition via genouest

16:15 – 16:30 **PAUSE**

16:30 – 17:30 **TRAVAIL EN GROUPES SUR LA REPRODUCTIBILITÉ**

Fil rouge : reproduire un des articles sélectionnés, avec un objectif de reproductibilité (+1 indicateur qualité)

17:30 – 18:00 **RESTITUTION ET REMISE DU PRIX**

CONFERENCE PROGRAM

11 FEBRUARY 2021

(1) NICOLAS ROUGIER

Title : Let's redo Science !

Abstract : ReScience C is a platinum open-access peer-reviewed journal that targets computational research and encourages the explicit replication of already published research, promoting new and open-source implementations in order to ensure that the original research is reproducible. You can read about the ideas behind ReScience C in the article “Sustainable computational science: the ReScience initiative” [1] To achieve this goal, the whole publishing chain is radically different from other traditional scientific journals. ReScience C lives on GitHub where each new implementation of a computational study is made available together with comments, explanations and tests. Each submission takes the form of an issue that is publicly reviewed and tested in order to guarantee that any researcher can re-use it. If you ever replicated computational results (or failed at) from the literature in your research, ReScience C is the perfect place to publish your new implementation. ReScience C is collaborative and open by design. Everything can be forked and modified. Don't hesitate to write a submission, join us and to become a reviewer. [1] Rougier, Hinsen et al, “Sustainable computational science: the ReScience initiative”, PeerJ Computer Science, 2017. DOI: 10.7717/peerj-cs.142

(2) ARNAUD LEGRAND

Title : Using Laboratory Notebooks to Improve the Reproducibility of your Research

Abstract : Although laboratory notebooks are commonly used in experimental science to document hypothesis, experimental conditions and initial analysis, the physical version (a stack of numbered pages) is often inappropriate to modern practices which involve a lot of digital information like data, computations and visualizations. Many so-called computational notebooks (e.g., Jupyter or Rstudio with the Rmd format) have thus emerged and provide literate programming capabilities although they do not necessarily encourage journaling as it is commonly done with classical laboratory notebooks. In this presentation, I will quickly demo some of these tools, present how I use them on a daily basis as an electronic laboratory notebook, and how they allow me to write reproducible articles which comprise all the important research information to the readers.

(3) OLIVIER COLLIN

Title : Reproducibility and data management on the GenOuest facility

Abstract : In the context of FAIR data and Open Science in Bioinformatics, new environments arose that will help scientists improve the reproducibility of their computational workflows. During these last years, GenOuest has built environments to help scientists address reproducibility challenges.

(4) INTERVENANTS DE LA TABLE RONDE "POLITIQUE DE REPRODUCTIBILITÉ DES ÉQUIPES"

Olivier Dameron (Dyliss) / **Charles Deltel** (Genscale) / **Luis Galarraga** (Lacodam) /
François Goasdoue (Shaman) / **David Gross-Amblard** (Druid) / **Fabrice Legeai** (Genscale)

ANNUAL MEETING OF PH-D STUDENTS CONFERENCE PROGRAM 12 FEBRUARY 2021

09:00 – 09:40 **TABLE RONDE**
**"COMMENT UN·E DOCTORANT·E DOIT CHOISIR
SON ÉDITEUR / SA CONFÉRENCE ?"**

Intervenant·es :

Dominique Lavenier (comité national section 6 et CID51)

Elisa Fromont (conseil scientifique INS2I)

Alexandre Termier (responsable d'équipe)

Mireille Ducassé (ex CNU)

Arnaud Martin (ex CNU)

09:40 – 10:00 **PAUSE**

10:00 – 10:30 **RETOUR D'EXPÉRIENCE DES DOCTORANT·ES**

**Comment s'est passé le processus de soumission d'un
article ?**

Camille-Sovanneary Gauthier (Lacodam)

Zuo-Wei Zhang (Druid)

10:30 – 10:45 **PRÉSENTATION DES DOCTORANT·ES DE PREMIÈRE
ANNÉE SUR LEUR SUJET DE THÈSE (VOLET 1)**

1. **Julien Delaunay (Lacodam)** *Automatic Construction of Explanation for AI models*
 2. **Camille Juigné (Dyliss)** *Analyse des données biologiques hétérogènes par exploitation de graphes multicouches pour comprendre et prédire les variations d'efficacité alimentaire chez le porc*
 3. **Garance Gourdel (GenScale)** *Sketch-based approaches to processing massive string data*
 4. **Josie Signe (Lacodam)** *Characterizing the diversity of livestock farming situations with data mining methods used on information from dairy herd sensors*
 5. **Véronne Yepmo (Shaman)** *Détection et explication d'anomalies dans un contexte de commerce international*
 6. **Meziane Aite (Dyliss)** *Intégration et analyse basées sur des connaissances symboliques pour l'identification de nouvelles combinaisons thérapeutiques*
 7. **Olivier Gauriau (Lacodam)** *Fouille de règles numériques pour la prédiction de la dynamique des maladies des plantes*
 8. **Lucas Robidou (GenScale)** *Search engine for genomic sequencing data*
-

ANNUAL MEETING OF PH-D STUDENTS CONFERENCE PROGRAM 12 FEBRUARY 2021

10:45 – 11:15 **PAUSE ICE-BREAKER**

11:15 – 11:30 **PRÉSENTATION DES DOCTORANT·ES DE PREMIÈRE ANNÉE SUR LEUR SUJET DE THÈSE (VOLET 2)**

1. **Hugo Ayats (SemLIS)** *De la prédiction à l'automatisation avec une IA explicable et centrée-utilisateur - application à la construction de graphes de connaissances*
2. **Lénaïg Cornanguer (Lacodam)** *Apprentissage automatique d'automates temporisés à partir de séries temporelles*
3. **Nicolas Buton (Dyliss)** *Nouveaux modèles pour la caractérisation de protéines fonctionnelles par apprentissage profond*
4. **Wafaa EL HUSSEINI (Shaman)** *Gestion efficace de données à l'aide d'ontologies expressives : application au factchecking*
5. **Antonin Voyez (Lacodam)** *Privacy-Preserving Power Consumption Time-Series Publishing, Designing Attacks for Stress-Testing Sanitized Datasets*
6. **Victor Epain (GenScale)** *Développement de méthodes efficaces, précises et conviviales pour corriger, assembler et aligner des lectures issues des technol*
7. **Matthieu Bougueon (Dyliss)** *Modélisation prédictive pour le ciblage thérapeutique du TGF-beta dans les pathologies chroniques hépatiques*
8. **Victor Guyomard (Lacodam)** *Explicabilité des décisions prises par les algorithmes d'apprentissage machine*

11:30 – 12:00 **POSTERS DES DOCTORANT·ES DE DEUXIÈME ANNÉE**

ANNUAL MEETING OF PH-D STUDENTS
CONFERENCE PROGRAM
12 FEBRUARY 2021

TITRES ET NUMÉROS DES POSTERS

1. **Arnaud Belcour** : *Investigating microbiota metabolic complementarity in a biogas plant ecosystem*
2. **Louis Béziaud** : *Ethical and privacy issues of artificial intelligence in law and justice*
3. **Simon Corbillé** : *Multi-branch CNN for online handwriting recognition*
4. **Clara Delahaye** : *Sequencing DNA at low cost with nanopores: troubles and biase*
5. **Samuel Felton** : *Deep learning for visual servoing*
6. **Virgilio Kmetzsch** : *Plasma microRNA signature in individuals with C9orf72-associated FTD and ALS*
7. **Téo Lemane** : *Kmtricks modular k-mer count matrix and Bloom filter construction for large read collections*
8. **Gregory Martin** : *Prediction-Based Car Relocation Assignment for Free Floating Car Sharing Services*
9. **Maria Massri** : *Base de données graphe temporelle pour les objets connectés*
10. **Van Long Nguyen Huu** : *MASCARA (ModulAr Semantic CACHing fRAMework) towards FPGA acceleration for IoT Security monitoring*
11. **François Mentec** : *Toward Robust Recommender System for Recruitment*