

---

# PARConnector : sans détour du Langage R au Cloud Big Data - Application à l'analyse Metagenomique

Iyad Alshabani , Jean-Michel Batto<sup>\*1</sup>, Magali Berland<sup>2</sup>, Denis Caromel<sup>†</sup> , Ndeye Aram Gaye<sup>‡§</sup> , Emmanuelle Le Chatelier , Laurent Pellegrino<sup>‡¶</sup> , Nicolas Pons , Edi Prifti , and Fabien Viale

<sup>1</sup>INRA US MetaGenoPolis 1367 – Institut national de la recherche agronomique (INRA) : US1367 – INRA DOMAINE DE VILVERT Unité MGP - Bâtiment 325 78352 Jouy-en-Josas Cedex, France

<sup>2</sup>INRA (Metagenopolis) – Institut national de la recherche agronomique (INRA) : US1367 – Domaine de Vilvert, 78 352 Jouy-en-Josas Cedex, France, France

## Résumé

L'analyse quantitative du microbiote humain a été développé conjointement par l'équipe INRA-MetaGenoPolis ([www.mgps.eu](http://www.mgps.eu)) dans le cadre du projet Européen MetaHIT ([www.metahit.eu](http://www.metahit.eu)) qui impliquait 13 partenaires académiques et industriels. Cette approche explore l'information génomique des bactéries vivant dans le tractus digestif de manière globale et quantifiée. L'analyse repose sur 'R' qui présente l'avantage d'avoir une gratuité et un vaste choix en modules et traitements. Dans un environnement de traitement HPC, l'orchestration permet de présenter à l'utilisateur une puissance de traitement sans en exposer la complexité sous-jacente. Nous présentons ici notre approche et les outils utilisés pour effectuer du traitement Big Data en lien avec la metagénomique quantitative à travers l'utilisation de l'orchestrateur open source ProActive.

---

\*Auteur correspondant: [batto@jouy.inra.fr](mailto:batto@jouy.inra.fr)

†Auteur correspondant: [denis.caromel@activeeon.com](mailto:denis.caromel@activeeon.com)

‡Intervenant

§Auteur correspondant: [nagaye@jouy.inra.fr](mailto:nagaye@jouy.inra.fr)

¶Auteur correspondant: [laurent.pellegrino@activeeon.com](mailto:laurent.pellegrino@activeeon.com)