

## Chargé(e) de recherche de classe normale

### Concours N° CRN15CSS5

**Commission scientifique sectorielle (CSS) :**  
CSS5 « Sciences des données et des modèles »

#### Intitulé

**Développements méthodologiques appliqués aux données massives, multimodales et hétérogènes (Big data)**

#### Mots clés / disciplines

Fouille de données hétérogènes, lacs de données, web sémantique, ontologies, représentation et visualisation des connaissances.

#### Contexte et enjeux au Sud

La multiplication et l'augmentation de la dimension des jeux de données issues de l'observation, de l'expérimentation ou encore de la simulation numérique demandent de définir de nouvelles méthodes permettant l'interopérabilité, l'intégration de données et l'extraction de connaissances au sein de ces masses de données hétérogènes, multi modales et gérées en silo. Cela afin de pouvoir 1) en permettre des représentations spatio-temporelles et thématiques compatibles pour des analyses interdisciplinaires, 2) assurer la reproductibilité des analyses (modélisation, simulation, visualisation) impliquant le croisement de données pluridisciplinaires ou d'algorithmes et enfin 3) en assurer l'exploration visuelle et interactive pour produire de nouvelles interprétations. Cela constitue un domaine de recherche en soi qui est actuellement en plein essor. L'IRD souhaite renforcer les expertises dans ce domaine tout en les positionnant autour des objectifs du développement durable et de la science de la durabilité partout où il s'implique.

#### L'état de l'art et besoins

Le candidat apportera son expertise méthodologique pour l'extraction, la représentation, et le rapprochement de connaissances pluridisciplinaires, ainsi que la visualisation graphique et interactive de sources de données massives multimodales et hétérogènes.

Il contribuera à concevoir des approches innovantes qui utiliseront et intégreront des méthodes et techniques venant de la représentation des connaissances, de la fouille ou de la visualisation de données qui pourront être appliquées aux grandes bases de données structurées et pluridisciplinaires, au web de données ou encore aux lacs de données. Seront plus particulièrement appréciés les profils de chercheurs ayant développé des recherches 1) sur la formalisation et représentation des connaissances 2) les méthodes de fouille de données dans de grands ensembles de données.

Des exemples de thématiques de l'IRD en lien avec ces enjeux autour de l'interopérabilité, l'intégration de données et l'extraction de connaissances sont l'environnement, la santé, et les sciences sociales pour lesquelles l'IRD développe des recherches pour le développement.

#### Compétences / capacités souhaitées

Expérience de recherche dans la représentation formelle des connaissances, la fouille de données au sein d'ensembles hétérogènes et multimodales ou la visualisation de données complexes et de publications dans les conférences et revues de référence du domaine. Un intérêt ou expérience de projets pour une science de la durabilité est vivement souhaité.

Ce profil est transverse aux principales thématiques des Objectifs de développement durable (ODD).