



MI

Mission Interdisciplinarité

Défi interdisciplinaire - Appel à manifestations d'intérêt 2013

MASTODONS

Masses de Données Scientifiques

Au cœur des sciences de l'information et de la communication, les données représentent la matière première consommée ou produite par les applications et les systèmes. De nombreux domaines scientifiques (ex : cartographie du génome, observatoire virtuel en astronomie, simulation en physique et énergie, imagerie médicale), économiques (ex : e-commerce, systèmes décisionnels), ou sociétaux (ex : réseaux sociaux, bibliothèques numériques, distribution de contenus multimédia) produisent et consomment des volumes de données considérables. L'ouverture des données (Open Data) et la corrélation entre plusieurs sources de données (Linked Data) sont devenues des instruments de valorisation des données et posent à ce titre de nombreux problèmes d'hétérogénéité, de sémantique et de droit d'usage. Le CNRS, à travers ses unités de recherche et ses grands instruments, capitalise plusieurs centaines de bases de données et de corpus d'informations dont les volumes croissent de façon exponentielle et dont la valorisation reste un enjeu stratégique. Les modalités d'interopérabilité et de rapprochement entre les données de différents domaines sont encore mal maîtrisées. Les systèmes d'information doivent mettre en œuvre des chaînes de traitement complexes pour obtenir des produits informationnels à forte valeur ajoutée (connaissances, comportements, patterns, événements rares, ...). La conception et la réalisation de tels systèmes ne font pas seulement appel aux seules techniques de bases de données et de calcul intensif mais aussi aux techniques de modélisation, à la recherche d'informations, à la fouille de données et à l'apprentissage, aux connaissances linguistiques, au raisonnement et à la visualisation...

La Mission pour l'Interdisciplinarité (MI) du CNRS a lancé en 2012 un défi sur la gestion, l'analyse et l'exploitation des très grandes masses de données scientifiques (MASTODONS). **Le but est d'identifier et de soutenir des actions de recherche dont les résultats ne pourraient être obtenus sans une fertilisation croisée des disciplines et sans une synergie effective entre chercheurs.** Cet appel, très fructueux a permis de sélectionner une première liste de projets. L'appel 2013 a pour objectif de compléter cette liste pour couvrir au mieux aussi bien les verrous scientifiques posés par l'utilisation à grande échelle des données (sémantique, stockage, recherche, visualisation...) que l'impact sociétal qu'induisent les technologies proposées (protection de la vie privée, propriété intellectuelle, préservation de la connaissance...). L'objectif majeur du défi MASTODONS est de favoriser l'émergence d'une communauté scientifique interdisciplinaire autour de la **science des données**, susceptible de rivaliser avec les grands laboratoires internationaux et de produire des solutions originales sur **le périmètre des données scientifiques**.

Dans cette perspective, la MI du CNRS souhaite soutenir une seconde vague de projets visionnaires, capitalisant des connaissances de plusieurs disciplines, et permettant de franchir un pas significatif dans la valorisation des grandes masses de données (i.e. gestion, analyse et exploitation). Ces recherches peuvent concerner, mais de façon non exclusive, les axes de recherche suivants :

1. Calcul intensif sur des grands volumes de données, parallélisme dirigé par les données, simulation sur des grands volumes de données,
2. Recherche, indexation, exploration et visualisation de grandes masses de données et de connaissances,
3. Qualité des données, confidentialité et sécurité ; problèmes de propriété, de droit et d'usage des données.
4. Sémantique des données et des connaissances résultant de processus de requêtage ou de fouille de données

Les équipes ou consortiums d'équipes intéressées doivent adresser une manifestation d'intérêt sous forme d'un document scientifique de 3 à 5 pages maximum contenant les éléments suivants :

- Vision scientifique de l'équipe/consortium sur les thèmes du défi,
- Les verrous scientifiques et les axes de recherche à moyen terme, avec un focus particulier sur la première année,
- Les acquis scientifiques dans le domaine ou dans un domaine connexe susceptible de contribuer aux problèmes scientifiques ou sociétaux posés (publications significatives, projets passés ou en cours, applications réalisées, logiciels, brevets...),
- Les différentes disciplines impliquées et leurs contributions respectives au projet,
- Une liste de 3 à 5 chercheurs seniors impliqués de façon significative dans la recherche.

Le montant du financement de chaque action retenue sera négocié entre la MI et les instituts en fonction des soutiens déjà apportés par les instituts aux laboratoires et de la stratégie de site du CNRS : Le soutien financier alloué permettra le fonctionnement de projets. Les frais de personnels (CDD, salaires doctorants, post doctorants...) et les indemnités de stage ne pourront pas être financés sur ces crédits. Le financement sera octroyé pour l'année civile 2013.

L'appel à manifestations d'intérêt n'est pas réservé aux personnels du CNRS. Cependant le projet collaboratif doit comprendre au moins une unité de recherche qui relève du CNRS (UPR, UMR...). Un rapport de synthèse des principaux résultats obtenus sera demandé à la fin de chaque année.

=====

Le formulaire de candidature est disponible à l'URL : <http://www.cnrs.fr/mi/spip.php?article297>

Ce formulaire complété (5 pages maximum) doit être obligatoirement déposé sur l'application SIGAP à l'adresse :

<https://sigap.cnrs.fr/sigap/web/connexion.php>

Date limite de dépôt des candidatures : le 15 avril 2013 à minuit.

Pour obtenir des informations : Mokrane Bouzeghoub <mokrane.bouzeghoub@cnrs-dir.fr> (coordinateur scientifique du défi) et la Mission pour l'interdisciplinarité mi.contact@cnrs.fr