



Données de la Recherche : enjeux, perspectives, politique(s)

Journée-atelier du
29 mai 2012

CNRS
Campus Gérard Mégie
3, rue Michel-Ange
Paris 16^e

Les pratiques de recherche ont considérablement évolué ces dernières décennies : essor des coopérations internationales, de la spécialisation à l'interdisciplinarité, financement par projets, recours massifs aux grands équipements et aux moyens de calculs partagés... Ces évolutions ne sont pas limitées aux sciences de l'observation, mais touchent également les sciences humaines et sociales. Il y a une dizaine d'années a été forgé le concept d'escience (enhanced science, John Taylor, 1999), faisant référence à cette nouvelle science collaborative et distribuée, s'appuyant sur l'accès à des grands réservoirs de données, des réseaux à haut débit et des moyens de calcul et de visualisation à haute performance.

Mais une des plus marquantes caractéristiques accompagnant cette transformation de la pratique scientifique est la formidable quantité de données numériques produites par les capteurs, senseurs, instruments divers d'observation et de mesure, des télescopes ou sondes embarquées, mais aussi générées par des processus de simulation.

Si l'efficacité de la science moderne passe par sa capacité à traiter et exploiter ces gigantesques masses de données dans des environnements partagés, son efficience pour les années à venir se mesurera à sa capacité à GÉRER, PARTAGER et autoriser la RÉUTILISATION de ces mêmes données.

C'est dans cet esprit que sont nées les réflexions à l'échelle européenne en faveur d'une infrastructure de données capable d'offrir dans 2 ou 3 décennies, de façon transparente, l'accès à des interfaces d'utilisation et réutilisation de données validées et dignes de confiance.

Atteindre cet objectif nécessite de fédérer des moyens et compétences variées sur le long terme, en se nourrissant des expériences de chaque communauté scientifique, en définissant des cadres de collaborations internationales, mais aussi à l'échelle du laboratoire, de l'équipe, en sensibilisant et formant les chercheurs aux enjeux du partage et de la réutilisation des données scientifiques.

Certaines disciplines scientifiques sont déjà largement sensibilisées à ce problème des données de recherche et leurs communautés ont largement anticipé ces réflexions en se dotant d'infrastructures de gestion des données : les astronomes, les biologistes, les physiciens...

Certaines nations européennes, certaines institutions ont déjà initié des programmes visant à accompagner les scientifiques dans ces approches.

L'objectif de cette journée est modeste : il sera dans un premier temps de présenter quelques réalisations et réflexions sur le problème des données de recherche à l'échelle internationale et européenne.

L'après-midi sera consacrée à une revue de quelques initiatives françaises disciplinaires du CNRS, de l'Inserm, de l'Inra... non pas tant pour mettre en valeur des « success stories » que pour faire toucher du doigt les difficultés et les écueils de cette activité de soutien à la recherche qu'est la gestion des données.

En rassemblant des chercheurs de disciplines variées des professionnels de l'information scientifique et technique ainsi que des informaticiens, intéressés et/ou impliqués dans des processus de gestion de données, nous ambitionnons de confronter les positions et points de vue sur des thèmes aussi variés que :

- la propriété des données de recherche et la responsabilité des chercheurs
- la formation des spécialistes du data management : data librarians, data scientists...
- les processus de la curation de données numériques
- le financement des infrastructures
- la prise en compte de ces activités dans la carrière des chercheurs
- les liens entre publications et données
- les positions institutionnelles *versus* les pratiques des communautés de recherche.

Une table-ronde rassemblant décideurs de la recherche, experts, juristes... clôturera la journée. Elle éclairera les rôles des différentes parties prenantes du débat : organismes de recherche, professionnels de l'information, agences de financement, chercheurs, comités d'éthique...

Sans prétendre résoudre les problèmes posés par la gestion des données de recherche, cette journée doit être un levier de sensibilisation, un temps de partage d'expériences, une occasion de prendre conscience de l'intérêt d'une approche pluridisciplinaire du sujet.

Paris, le 29 mai 2012

Le comité d'organisation

