



CNRM, UMR 3589

## SEMINAIRE CNRM

*mardi 30 mars 2021 à 14h*

### **PRÉSENTATION DU PROJET ANR SEMANTICS4FAIR : PASSER DE DONNÉES OPEN À DONNÉES FAIR PAR L'INTÉGRATION DES USAGES**

**Nathalie Aussenac-Gilles et Amina Annane – IRIT,  
Pascal Gaillard – MSHS-T  
Christophe Baehr – DESR/CNRM**

**en visioconférence**

URL de la réunion : [https://bluejeans.com/233805338?src=join\\_info](https://bluejeans.com/233805338?src=join_info)

Par tél : composer l'un des numéros suivants: <https://www.bluejeans.com/numbers>)

Entrer l'ID et le code ID réunion suivi d'un #

Connexion depuis sys de salle? Appeler: bjn.vc ou 199.48.152.152 et entrer ID réunion et code ID

#### Résumé :

A Météo-France et au CNRM en particulier, nous produisons une masse importante de données, tant comme des produits directs de nos recherches que comme des produits dérivés. Nous avons eu jusqu'alors une culture du cloisonnement de nos données que la volonté de la société de généraliser la science ouverte remet en question. Ouvrir les données par une publication sur un entrepôt de données n'assure pas que nos données soient facilement accessibles par les acteurs de la société civile. Si nos données paraissent absentes, le risque est alors que ces acteurs se tournent vers des sites de données d'autres fournisseurs, étatiques ou privés. Pourtant, nos données produites dans nos expériences de recherche peuvent potentiellement être exploitées dans de nombreuses actions de recherche aussi bien en interne qu'en externe.

Rendre les données ouvertes et accessibles nécessite cependant des efforts considérables afin de garantir la qualité des données et leur conformité aux principes FAIR (données **F**aciles à trouver, toujours **A**ccessibles, facilement **I**nteropérables et **R**éutilisables dans le temps et par d'autres).

Le but du projet ANR SEMANTICS4FAIR est de faciliter la tâche de recherche et d'accès aux données scientifiques résultant à la fois de la recherche et de la production, afin de soutenir le développement de nouveaux usages par d'autres communautés que celle de production. L'originalité de l'approche proposée est double : (i) une méthode basée sur l'ergonomie vise à saisir les besoins et le vocabulaire des utilisateurs ; et (ii) une approche sémantique relève le défi de la découverte de données.

Comme cas d'application nous considérons les données publiques de Météo-France utilisées dans le cadre d'une action de recherche sur la propagation de l'Ambroisie menée par l'OMP.

Nous verrons sur cet exemple quels étaient les verrous qui ont entravé la recherche de données météorologique du chercheur de l'OMP et nous présenterons des pistes d'amélioration, notamment par le recueil des usages, l'ajout de métadonnées sémantiques et une réflexion sur l'affichage des métadonnées dans des référentiels accessibles aux potentiels utilisateurs.

**Pour tout renseignement, contacter Y. Poirier (05 61 07 96 55)**

Centre National de Recherches Météorologiques  
42, Avenue G. Coriolis - 31057 Toulouse Cedex