
Construire et publier collaborativement ses bases de données intersciences avec la plateforme OpenCommonS

Pascal Dayre*^{1,2}

¹Institut de Recherches en Informatique de Toulouse – CNRS, INPT, Université Toulouse 3 Paul Sabatier, Université Toulouse 1 Capitole, Université de Toulouse 2 - Jean Jaurès – France

²Institut des sciences informatiques et de leurs interactions - CNRS Sciences informatiques – CNRS – France

Résumé

Le 2ème Plan National pour la Science Ouverte (PNSO 2021-2024) vise à généraliser les pratiques de Science Ouverte, à partager et à ouvrir les données de la recherche, et à promouvoir les autres produits de la recherche tels que les codes sources et les workflows. Les enjeux de partage et de réutilisation concernent à la fois les scientifiques producteurs et (ré)utilisateurs des données, mais aussi le personnel support, particulièrement les personnes en charge du data stewardship.

Pour contribuer à ces missions, nous proposons la plateforme OpenCommonS pour créer, gérer, rendre exploitable et développer les usages des métadonnées descriptives des données scientifiques par une communauté de données et une communauté de recherche associée (actionnable metadata).

Nous nous intéresserons ici plus particulièrement au scénario collaboratif de construction d'une base d'une communauté de données utilisant toutes les facilités offertes par l'écosystème numérique de la Science Ouverte comme les référentiels des acteurs, des unités et des ressources de la recherche, les vocabulaires et les services de vocabulaires, les entrepôts de données, services de répertoires et tout autre type de service.

La nouvelle base de données pourra être faite à partir de données déjà existante et/ou de données nouvellement créées. Elle pourra intégrer différents types de ressource comme les jeux de données, leur(s) distribution(s) ou encore des services ou d'autres bases de données.

Nous discuterons de comment rendre réutilisable sa base de données et aussi de comment la valoriser sur HAL, RDG, EOSC et/ou autres portails européens de données.

La plateforme générique OpenCommonS est conçu pour rendre des données et des bases de données réutilisables en intégrant les principes FAIR dans sa spécification et sa conception. Elle offre la possibilité de créer des espaces de données dans lesquels des communautés de producteurs et d'utilisateurs de données cohabitent avec les personnels dédiés au support, afin de faciliter l'accès et la production de catalogues de données dans un contexte disciplinaire, multi ou interdisciplinaire ou encore intersciences quelle que soit la nature et la

*Intervenant

localisation des données. Il est possible alors de mettre en place des scénarios collaboratifs comme des processus de modération et remédier à la problématiques des silos de données disciplinaires. La plateforme gère leurs métadonnées descriptives. Les données elles-même sont hébergées par des entrepôts partenaires.

Les utilisateurs peuvent ainsi rechercher, recenser et prpéparer les données avant, pendant et après leurs projets. Ils peuvent référencer en un seul point l'ensemble des jeux de données utilisés ou produites à destination d'une exposition privée, publique ou partagé pour un PGD, des collaborations tout le long des cycles de vie des données et des projets.

URL de la plateforme OpenCommonS : <https://ocs.irit.fr/last>