

# "Défis de la réutilisation des données ouvertes et stratégies alternatives".

Samuel Goëta Metsem du 21/10/20

## *Résumé de la séance :*

Depuis octobre 2018, la loi pour une République numérique prévoit que toutes les administrations de plus de 50 agents (et pour les collectivités locales de plus de 3500 habitants) doivent diffuser les données qu'elles produisent sous une licence et un format ouverts facilitant leur réutilisation. Les administrations publient plusieurs dizaines de milliers de jeux de données, regroupés essentiellement sur [data.gouv.fr](https://data.gouv.fr) en France, avec des niveaux de valeur et d'utilisabilité très variables. Dans cette séance du MetSem, Samuel Goëta, co-fondateur de la coopérative Dataactivist et maître de conférences associé à Sciences Po Aix, proposera de débattre des défis posés par la réutilisation des données ouvertes. Nous nous concentrons sur 4 défis essentiels de l'open data : la découvrabilité, la documentation, la qualité et les échanges entre producteurs et réutilisateurs. Le travail des données, son ampleur et sa répartition entre producteurs et réutilisateurs seront le fil rouge de ce séminaire qui décrira le contexte administratif, politique et juridique dans lequel se déploie l'ouverture des données publiques. Que faire si les données publiques sont inexistantes sur un sujet ou inexploitable ? Le séminaire se conclura par quatre stratégies alternatives d'obtention de données : le recours au droit d'accès aux documents administratifs, l'extraction (*scraping*) des données du web, la constitution de bases de données à partir d'informations disponibles et l'appel à une foule de contributeurs.