

Valorisation des données et des codes dans un contexte de science ouverte

- **Valorisation par le partage**

Des plateformes comme GitHub et GitLab permettent de mettre à disposition des données et codes accessibles à tout le monde.

Ces plateformes sont couramment utilisées pour transmettre les ressources supplémentaires lors de la publication d'un article.

Il a d'ailleurs été souligné que le partage des données et des codes qui ont servi pour les analyses est obligatoire lors de la publication d'un article (loi Lemaire).

Si les plateformes et l'accessibilité des données est simple et rapide, les utilisateurs/trices rencontrent parfois un manque d'accompagnement par des experts dans l'utilisation des données (visualisation, interprétation). A la manière d'une « clinique des données », on pourrait mettre en place des « cliniques d'accès à la donnée » ou des tutoriels directement sur les GitHub pour utiliser les données à disposition.

- **Valorisation au sein du monde scientifique**

Plus que la diffusion, la valorisation des données peut passer par différents processus comme les analyses, les publications et la représentation graphique (les sites web de type « data-lyses » ont été mentionnés).

Se pose alors la question de la frontière entre des données et des résultats. Comment valoriser une donnée sans montrer un résultat ? L'exemple des cartes a été développé, car une carte peut être une donnée si l'on ne représente qu'une série de valeurs dans l'espace. Par contre, une carte peut être un résultat dès lors qu'on croise plusieurs données entre elles qui permettent de tirer des conclusions et interprétations.

Pour la valorisation des données au sein du monde scientifique, il a été mentionné la publication d'un « data paper » associé à une publication académique. Ces publications plus techniques de valoriser en détail les données disponibles, mais aussi les producteurs/trices de ces données (techniciens, bénévoles, stagiaires, etc.).

- **Valorisation pour le grand public**

Dans les sciences et en particulier dans le domaine des sciences humaines et sociales, il a été mentionné l'importance des témoignages pour la valorisation des travaux de recherche (documentaires, podcasts, interviews, etc.).

Concernant la valorisation des données au sens strict, il a été souligné l'importance de mettre en forme les données brutes afin qu'elles soient compréhensibles et exploitables par le plus grand nombre, y compris des utilisateurs en dehors du monde scientifique. Cependant, il faut être « vigilant/e » à l'information véhiculée car certains contextes permettent l'application des travaux de recherche par des entreprises pour la productivité. Une erreur serait d'utiliser une simplification de la représentation des données pour le grand public, ce qui peut mener à une mauvaise interprétation de la part des utilisateurs/trices des données. Des formats de valorisation de données pour le grand public permettent d'éviter ce risque, notamment les conférences et la valorisation scientifique.

Enfin, la revue « The conversation » a été mentionnée comme un moyen de publier des travaux de recherche en ciblant spécifiquement le grand public. Publier dans une telle revue peut aussi constituer un bon exercice de synthèse et de vulgarisation scientifique, en particulier pour des étudiant/e/s en thèse qui se forment à la recherche.