

"La mise en œuvre de mécanismes de citation de données : Un socle solide pour bâtir des *data papers*"

Edward J. Gray et Nicolas Larrousse
TGIR Huma-Num CNRS

Méthodologie : De la citation de données aux entrepôts

Quelles recommandations pour la création d'une "bonne" citation ?

- 1) Evaluation des [pratiques](#)¹ de citation en SHS
- 2) Développement de [Recommandations](#)² adaptées aux SHS à partir des principes de Force11
- 3) Analyse des entrepôts de données en SHS à partir des [Recommandations](#)²

Quels critères pour évaluer un entrepôt du point de vue de la citation ?

- **Présence d'un Identifiant Pérenne (PID *Persistent Identifier*)**
- **Présence d'une page d'accueil**
- **Présence des métadonnées structurés actionnables dans le page d'accueil**
- **Présence d'une fonctionnalité "Cite As"**
- Utilisation de vocabulaires standardisés pour la description (e.g. ORCID)
- Présence d'une fonctionnalité d'accès aux différentes versions (versionning)
- Présence de liens vers les publications liées à ces données

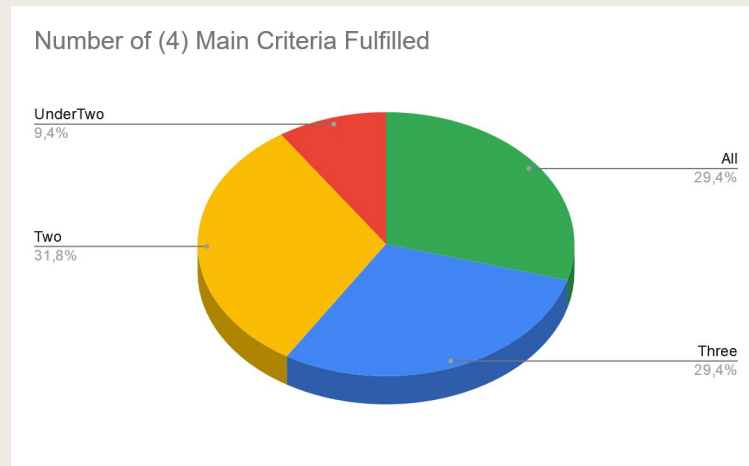
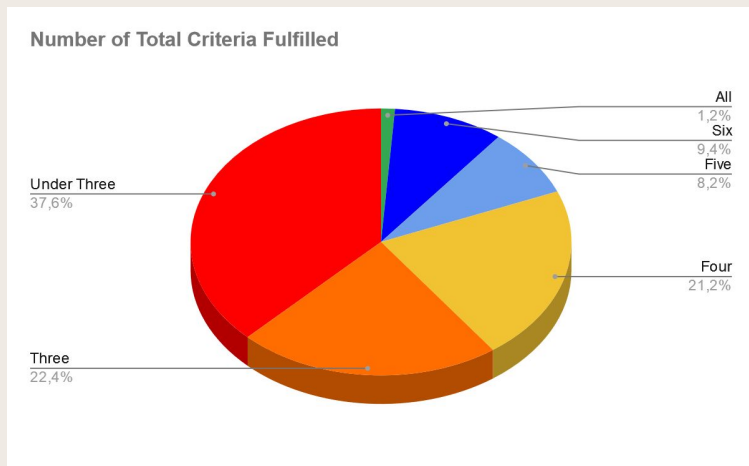
Ces 4 critères représentent les critères principaux

1. Nicolas Larrousse, Daan Broeder, Jan Brase, Cesare Concordia, and Vasso Kalaitzi. "SSHOC D3.2 Inventory of SSH Citation Practices, and Choice for SSHOC Citation Formats and Implementation Planning". Zenodo, December 31, 2019. <https://doi.org/10.5281/zenodo.4436736>

2. Nicolas Larrousse, and Edward J. Gray. "Recommendations for FAIR Data Citation in the Social Sciences and Humanities". Zenodo, September 1, 2021. <https://doi.org/10.5281/zenodo.5361718>

Entrepôts : Résultats de l'évaluation

Analyse de 85 Entrepôts de données en SHS sur 7 critères dont 4 principaux³



Nombreuses pistes d'amélioration de la qualité des entrepôts pour mettre en oeuvre la citation de données, en particulier :

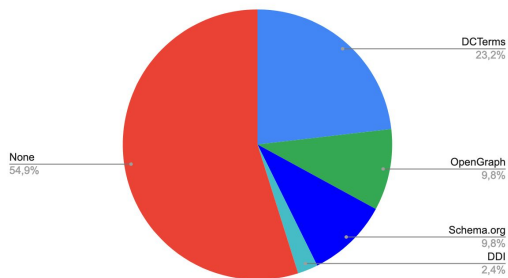
- La fonctionnalité "CiteAs" n'est présente que dans 50% des entrepôts et les chaînes générées ne sont en général pas exploitables par des machines
- Une meilleure structuration des métadonnées en particulier dans les pages d'accueils
- L'utilisation de vocabulaires standardisés

De la citation aux *data papers*

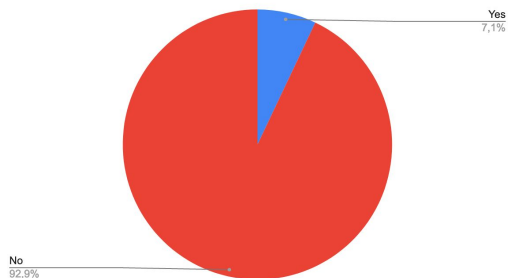
Les **critères** nécessaires à la construction d'une bonne citation vont permettre d'alimenter les informations requises dans un *data paper* (e.g. Titre, Auteurs etc.)

Le fait de rendre les citations "**actionnables**", en particulier par la normalisation des contenus (e.g. structuration de l'information dans un format standard et utilisation de vocabulaires comme ORCID), permet d'envisager **d'automatiser** le transfert de données descriptives pour un *data paper* à partir d'un entrepôt

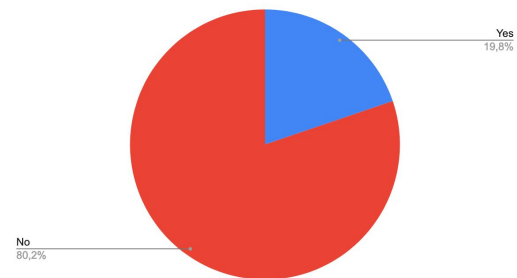
Presence of (embedded) Structured Metadata



Use of Standardized Vocabularies



Link(s) to Related Publications



Les données extraites de l'évaluation des entrepôts montrent que ces aspects sont encore à développer