

Data Papers

Midi de l'info scientifique, 30 avril 2017

1. Data Paper : les données comme sujet de publication

Le Data Paper est un type de publication scientifique qui décrit un ou des jeux de données accessibles en ligne.

Ces articles sont nommés différemment selon l'éditeur et les journaux dans lesquels ils sont publiés.

Data article
Data descriptor Data in brief
Data note
Data original article Data paper
Database article Database paper
Dataset paper

2. Modalité d'accès aux données

Il existe plusieurs modes de publication des données. Celles-ci peuvent être intégrées directement à l'article, être jointes comme matériel supplémentaire à l'article sur la page de l'éditeur ou être déposées dans un entrepôt dédié à accueillir et préserver des jeux de données. Dans ce cas un lien réciproque entre l'article et les données par le biais d'un identifiant numérique - le DOI - sera créé.

3. Avantages et inconvénients de ce type de publication

Le data paper est une publication citable et, selon les revues, référencée dans les bases de données (ex. Web of Science, PubMed). Elle permet ainsi de tracer l'usage des données et leur réutilisation. Comme toute publication, le data paper fait le plus souvent l'objet d'un peer review, qui concerne non seulement la publication, mais également les données.

Il faut cependant noter que comme toute autre publication, un data paper demande du temps et que des frais de publication (APCs) peuvent s'appliquer. Obtenir le consentement des différentes personnes impliquées ou concernées par les données prend du temps, tout comme la préparation des données en vue de leur anonymisation ou standardisation. Enfin, il convient de s'assurer que l'on dispose bien des droits sur les données et de définir les conditions sous lesquelles on les met à disposition d'autrui.

Ces étapes ne sont toutefois pas propres au data paper mais à tout dépôt des données dans un entrepôt, ce qui est de plus en plus une exigence des bailleurs de fonds (*open science*).

Dans tous les cas, un data paper donne de la visibilité aux données de recherche auxquels il se réfère, accélère leur diffusion, et favorise leur réutilisation par d'autres chercheurs. De plus, le fait de décrire les données et leur contexte de production assure qu'elles seront réutilisées de manière appropriée et que les auteurs seront bien crédités. Enfin, cela peut offrir des opportunités de collaboration et favorise la recherche transdisciplinaire.

4. Publication : choix du journal

Le Data Paper peut être proposé à un *journal classique* qui publie ce type de publication en plus des articles originaux (tel que PLoS ONE, F1000Research) ou à un Data Journal, *journal spécifique* à ce type d'article (ex. Data in Brief).

La majorité de ces journaux publie selon le principe de l'Open Access favorisant ainsi l'accès public aux données de la recherche

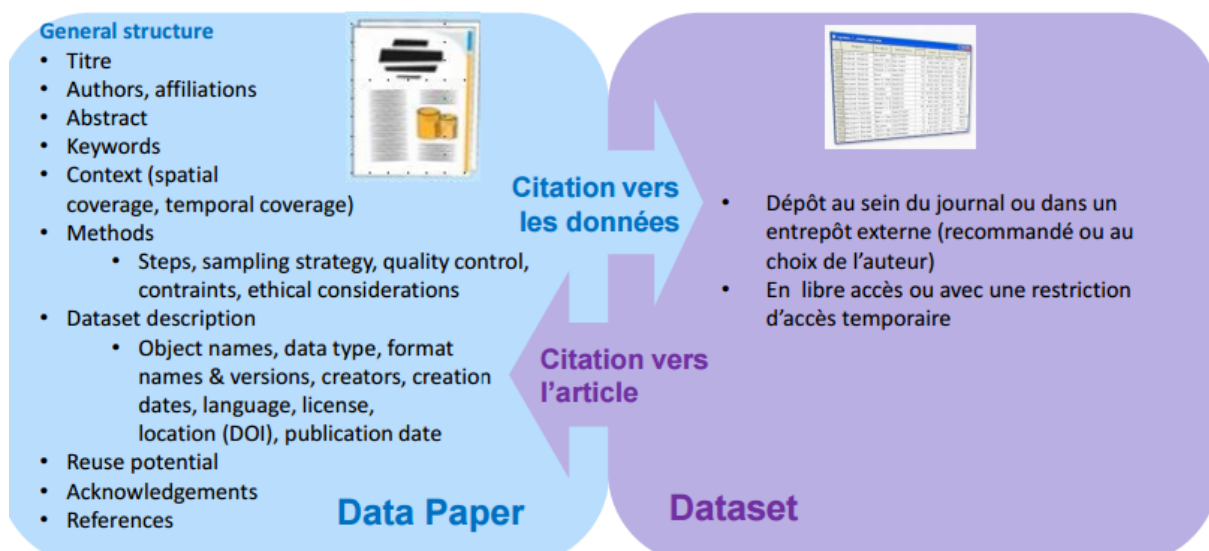
Les titres du domaine biomédical et des sciences de la santé sont en grande partie listés dans le Directory for Open Access Journal (DOAJ). Comme pour les journaux en Open Access, les auteurs paient alors des frais de soumission d'article - Article Processing Charge (APC).

Avant de choisir sa revue, il peut valoir le temps de vérifier l'éditeur scientifique derrière la revue, ses conditions d'accès (open access ou non), le montant des frais de publication (APCs), les critères de peer review ainsi que les bases de données dans lesquelles la revue est indexée. Les instructions aux auteurs recommandent parfois des entrepôts de données spécifiques et/ou proposent des modèles (« template ») de data papers, auxquels il convient de se conformer (ex. Data in Brief ; Scientific in Data).



5. Structure

Comme les Data Papers décrivent des données, ils contiennent des éléments descriptifs - métadonnées -, des informations techniques comme les méthodes d'obtention des données, et des informations concernant le potentiel de réutilisation des données :



Open Science : gestion et partage des données de la recherche / Esther Dzale & Dominique l'Hostis. CC-BY-NC-SA

Au contraire des articles traditionnels, ils ne contiennent par contre pas d'hypothèse de recherche, d'expérience réalisée en vue de confirmer/d'infirmer l'hypothèse et de conclusions suite à l'analyse de données.