

Évaluer un **data paper** retour d'expérience de la revue



european journal of geography
revue européenne de géographie

Clémentine Cottineau, CNRS
Christine Kosmopoulos, CNRS
Denise Pumain, Paris 1

Webinaire du groupe de travail inter-reseaux
Atelier Données de la MITI (CNRS)
Jeudi 5 novembre 2020

D'une revue de géographie en ligne
À la publication des modèles et des données

Évaluer un data paper pour Cybergeo
Principes et processus

Retour d'**expérience** et **difficultés** rencontrées

D'une revue de géographie en ligne À la publication des modèles et des données

Cybergeo créée en 1996 par Denise Pumain et Ky Nguyen

Ancrage dans l'UMR Géographie-cités (comités, relecteurs)

Commande européenne ECTQG (correspondants, lecteurs)

Revue généraliste, à comité de lecture, bilingue, sans numéro

Revue 100% en ligne, libre d'accès (jusqu'à aujourd'hui), pas d'APC

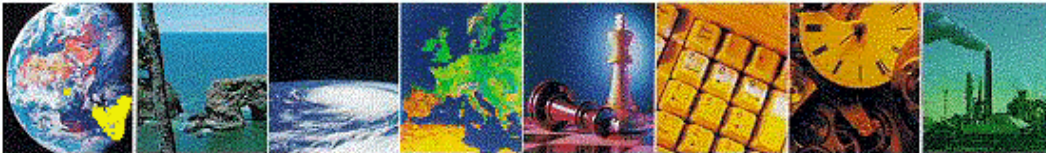
Évolution avec le web et le paysage des revues

2002

1996

Bienvenue

CyberGEO



Revue Européenne de Géographie
ISSN 1278-3366

European Journal of Geography
<http://www.cybergeog.presse.fr>

In English
Click here

2020

<https://journals.openedition.org/cybergeog/28929>

25 années de science ouverte

- 1997.** Arrivée de Christine Kosmopoulos, secrétaire de rédaction
- 2007.** Cybergeo apporte à *Revue.org* son format bilingue
- 2010.** Plateforme de gestion éditoriale *Manuscripts*
- 2012.** Cybergeo adhère au modèle *Freemium* (libre accès équitable)
- 2014.** Création d'une rubrique dédiée à la publication de modèles géographiques (code, docker, métadonnées) > *GeOpenMod*
- 2017.** Création d'une rubrique dédiée à la publication de données géographiques (données, métadonnées) > *DataPapers*

2017 - Une nouvelle rubrique: les data papers

Evolution du contenu de la revue avec l'intégration de nouveaux formats d'articles : Les « *Data Papers* »

- Décrivent, documentent et mettent en partage librement et gratuitement les jeux de données produites par les auteurs afin de permettre leur réutilisation et une science reproductible.
- Contrairement à un article de recherche classique, le *data paper* décrit uniquement des données scientifiques et les circonstances et méthodes de leur collecte.
- Donne une visibilité aux données et permet leur citation.

Première revue en géographie à publier des *data papers*

Une nouvelle rubrique: les data papers

Les « *data papers* » de Cybergegeo sont plus spécifiques que leurs équivalents célèbres (Nature's Scientific Data, Open Health Data, etc.) car ils ne retiennent que des bases de données géographiques. C'est pourquoi les recommandations aux auteurs demandent de préciser les échelles concernées, les composantes spatiales, la géométrie et la compatibilité avec les systèmes d'information géographique des éléments inclus dans la base de données, en plus des informations requises habituellement telles que les sources originales, la méthode de construction des données et les procédures de validation.

<https://journals.openedition.org/cybergegeo/28929>

Une nouvelle rubrique: les data papers

La rubrique ambitionne de devenir un espace rassemblant des bases de données géographiques élaborées par des auteurs garantissant que :

1. la production des données est transparente et citable (il ne s'agit pas de données non publiques)
2. le travail réalisé par les auteurs pour construire la base de données géographique est documenté afin d'être réutilisable et transférable vers de nouveaux objectifs
3. les bases de données sont évaluées par les pairs, ce qui offre un plus par rapport aux simples entrepôts de données
4. les ensembles de données sont accessibles et explorables avec des moteurs de recherche sémantiques et géographiques

<https://journals.openedition.org/cybergeogeo/28929>

Une nouvelle rubrique: les data papers

2020 | 2019 | 2018 | 2017

Céline Rozenblat

Extending the concept of city for delineating large urban regions (LUR) for the cities of the world [Full text]

Extension du concept de ville pour délimiter de grandes régions urbaines pour toutes les villes du monde

Ampliando el concepto de ciudad para delimitar las grandes regiones urbanas en las ciudades del mundo

Article 954, 29 September 2020

Mikhail Rogov and Céline Rozenblat

Delineating Russian cities in the perspective of corporate globalization: towards Large Urban Regions [Full text]

Délimiter les villes russes dans la perspective de la mondialisation des entreprises : vers les grandes régions urbaines

Delimitando las ciudades rusas en la perspectiva de la globalización de las empresas: hacia grandes regiones urbanas

Article 949, 23 July 2020

Antoine Peris, Willem Jan Faber, Evert Meijers and Maarten van Ham

One century of information diffusion in the Netherlands derived from a massive digital archive of historical newspapers: the DIGGER dataset [Full text]

Un siècle de diffusion de l'information aux Pays-Bas extrait d'une archive de journaux historiques numérisés : la base de données DIGGER

Un siglo de difusión de información en Países Bajos producido a partir de un archivo digital masivo de periódicos históricos: base de datos DIGGER

Article 928, 14 January 2020

- 7 articles publiés
- En anglais
- Nombre croissant (1 en 2017; 3 en 2020)
- 26 citations

Évaluer un data paper pour Cybergéo

Principes et processus

Evaluation par les pairs. Un nouveau comité de lecture

Des exigences liées aux données. Dépôt pérenne et ouvert

Des données géographiques. Métadonnées inédites

Consignes aux auteurs. Ajouts spécifiques

Chaîne de traitement. Dépôt > Évaluation > Publication

Principes et consignes de construction

“Les Data Papers devant être le plus souvent possible rédigés en anglais pour une plus large diffusion internationale, les recommandations aux auteurs concernant cette rubrique sont rédigées dans cette langue.”

<https://journals.openedition.org/cybergeogeo/23412#tocto1n7>

- **Visibilité et réutilisation maximale**
- **Comité de lecteurs internationaux**

Les normes de présentation de votre article

Les métadonnées

Le titre

Le résumé

Les mots clés

Le texte

Les images et les fichiers supplémentaires

Les images

Rappels (carto)graphiques

Le titre de vos illustrations

La source de vos illustrations

Le dépôt des fichiers supplémentaires

Le format des images

Quelques recommandations essentielles

Dans quels cas utiliser ces formats d'image

Pour obtenir une qualité optimale

La bibliographie

La Feuille de style (Zotero)

Les articles

Les extraits d'ouvrages

Les livres

Les articles électroniques

Guide pour les Data Papers et modèle des métadonnées

Cybergeogeo Data Papers Guidelines and Metadata template

Specific guidelines for GeoData Papers

Metadata template

Format for metadata

Consignes spécifiques: dépôt

“Make sure to deposit your data in an appropriate and sustainable repository and provide the appropriate authorization certificates if the data includes elements under proprietary license.”

- **Dépôt extérieur** (figshare, zenodo, nakala etc.) de préférence institutionnel
- Recommandé : Attribution d'un DOI par l'entrepôt institutionnel
- **Attention aux licences d'utilisation**
- **Responsabilité des auteurs**

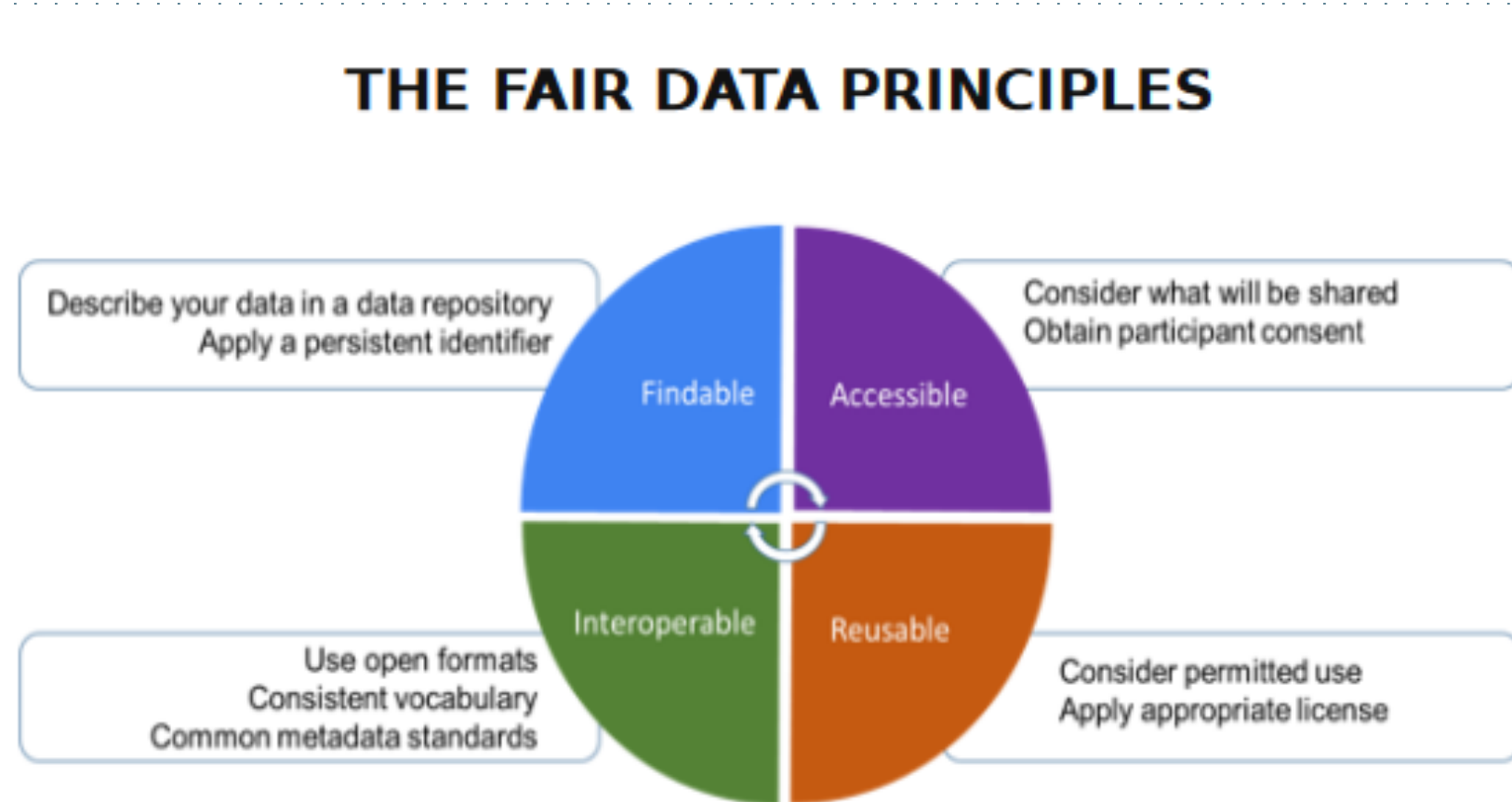
Consignes spécifiques: objectifs

“In your data paper, you should **document, explain and justify all the procedures used to transform the original records into the final geodatabase, demonstrate its need and use, and guide the reader on how to reuse it.** The reader, given this information and the appropriate authorisation certificates, should be in a position to replicate your work entirely.”

- **Data paper ≠ dépôt de données**
- Exigence de **reproductibilité**

Consignes spécifiques: respect des principes FAIR

“In line with FAIR principles (<https://www.go-fair.org/fair-principles/>) ...



Consignes spécifiques: respect des principes FAIR

- “...it is very important that your data paper provides **detail** on:
- all the **sources of the original records** which were used to construct the database (organisation, timestamp, format, variables selected, link to the repository, etc.).
 - all the **processes which have led from the original records to the final database** (data cleaning, aggregation, estimation, interpolation, modelling, etc.)”

➤ Description **exhaustive** et **détaillée**

Consignes spécifiques: respect des principes FAIR

- **“the spatial and scalar features of the data** (points, lines, polygons, local, regional, etc.). What spatial units are they composed of? Are they compatible with any available GIS geometry? How do the output objects relate to the input objects (aggregation, transformation, etc.)?
- **the validation process of the geodatabase which guarantees its quality** (technical validation, sensitivity and robustness analyses, comparisons with other data benchmarks, etc.). Make sure to mention the level of adequacy between the results of this validation process and the initial objectives for the construction of the geodatabase. State very clearly the strengths and limitations of the published dataset compared to other existing materials if any.
- **the value added of this database and any specific usage instructions to potential future users** if necessary.”

Consignes spécifiques: accès aux données

“If necessary, *Cybergeo* reviewers may ask you to **give temporary access to any intermediate data** (that you do not intend to publish) so as to validate the process of data production and its added value.”

- **Accès anonyme**
- **Vérification** d'étapes intermédiaires

Consignes spécifiques: metadata

Format for metadata

> Normes ISO

Title			text	250 characters
Abstract			text	100 words
Temporal reference				
	Time lapse			
		time_begin	date	dd-mm-year
		time_end	date	dd-mm-year
	Publication date		date	dd-mm-year
	Latest update		date	dd-mm-year
	Creation date		date	dd-mm-year
Responsible organization			text	250 characters
Responsible role			text	250 characters
Keywords			text	250 characters
Use			text	50 words

Type of spatial representation			controlled list	vector - grid - table - tin - estereoscopic model - photo - video
Spatial resolution (scale or minimum cartographic unit)			text	50 characters
Language			text	50 characters
Themes			text	100 characters
Geographic extension			text	
	min y		number - float	
	min x		number - float	
	max y		number - float	
	max x		number - float	
Reference system			text	50 words
Source			text	100 words

Un comité spécifique

Le comité de lecture des DataPapers

BACCAINI Brigitte, Commissariat Général du Développement Durable et INSEE, France

CAUVIN Colette, CNRS/UMR 7362 Image, Ville, Environnement, France

COTTINEAU Clémentine, CNRS/UMR 8097 Centre Maurice Halbwachs, France

ESCOBAR, Francisco, Geography and Environmental Sciences, University of Alcala, Espagne

LAURINI Robert, INSA-Lyon, France

MATHIAN Hélène, CNRS/UMR 5600 EVS-Lyon, France

OLIVEAU Sébastien, Aix Marseille Université, France

PFAENDER Fabien, UTSEUS-Shanghai City, Chine

RAIMBAULT Juste, University College London, Grande-Bretagne

RUAS Anne, IFSTTAR, France

Statisticien.nes
Géographes
Ingénieur.es

+ relecteurs extérieurs

- Par thème
- Internationaux

Processus d'évaluation

Première évaluation par les éditrices. Format et adéquation à la rubrique

Annonce auprès du comité Datapaper. Évaluation Datapaper

Demande de relecteur extérieur au Comité Datapaper. Évaluation thématique

Synthèse éditrice(s). Et itération si corrections exigées...

Échelle temporelle d'un cycle d'évaluation: ~ 2 mois

Retour d'expérience et difficultés rencontrées

Globalement: expérience positive, attente des auteurs

Difficultés rencontrées liées :

- Au dédoublement des articles
- Au dépôt anonyme des données
- À l'objet de l'évaluation

Dédoubllement des articles

“*Cybergeog* Data Papers are stand-alone scientific articles that are devoted to the construction of geographical databases. **Their publication can be followed by another article exploiting the analysis of such geographical data, but the two would be distinct in their scope and writing.** The publication of a data paper does not ensure the acceptance of any other article, and vice versa.”

- Pas toujours respecté
- Data Paper confondu avec matrice des métadonnées

Dépôt de données et anonymat

Dépôt de données antérieur à la soumission du data paper

Anonymat du dépôt > envoi d'un jeu anonyme + stockage intermédiaire par la revue (pas possible pour le moment)

Validation des articles vs. validation des données?

Vérification du processus de reproductibilité des données

Analyse du jeu de données correspondant

Pas de validation des données elles-mêmes > auteurs

Attention portée aux licences... mais responsabilité des auteurs

Conclusion

Data Papers dans la continuité de l'ouverture de la revue. Anticipe le plan national de la science ouverte lancé en 2018.

Spécificité des données géographiques. (échelle, couverture, etc)

Spécificité de l'évaluation. Nouveau comité, consignes spécifiques

Publication de Data Papers et non des données elles-mêmes

Difficultés nouvelles. Anonymat et dédoublement des articles

À suivre. Potentiel d'espace commun pour les données géographiques, accessibles et explorables (index géographique)

Merci pour votre attention

<https://journals.openedition.org/cybergegeo/28545>

<https://cybergegeo.hypotheses.org>

[@RevueCybergegeo](https://twitter.com/RevueCybergegeo)

Data Papers Cybergegeo

Blog Cybergegeo Conversations

Compte Twitter de la revue



Clémentine Cottineau, CNRS
Christine Kosmopoulos, CNRS
Denise Pumain, Paris 1

Webinaire du groupe de travail inter-reseaux
Atelier Données de la MITI (CNRS)
Jeudi 5 novembre 2020