

Lettre d'information GT Atelier Données n° 11 - Mai 2021

présentation d'Isabelle Blanc, administratrice ministérielle des données, des algorithmes et du code de la recherche, de l'enseignement supérieur et de l'innovation
Vers la création d'une plateforme nationale fédérée des données de recherche
<https://www.rd-alliance.org/sites/default/files/2021.05.11%20Recherche%20Data%20Gouv-webinaireRDA.pdf>

Ce message s'adresse à tout le monde, quel que soit votre statut et quelles que soient vos méthodes, même s'il intéressera peut-être particulièrement les quantitativistes. Dans le cadre des discussions actuelles sur la "science ouverte", qui s'étendent de plus en plus à la notion de "données ouvertes" (voir par exemple les "plans de gestion de données" demandés par l'ANR et d'autres financeurs), je participe à un groupe de travail qui vise à mettre en avant les pratiques réelles de recherche propres à chaque discipline, afin d'éviter que des normes "d'ouverture" ignorant ces pratiques, et dès lors inapplicables, soient édictées. Dans ce cadre, une campagne d'entretiens a été menée et doit maintenant être complétée par le recueil de réponses à un questionnaire en ligne. Ce questionnaire vous demande principalement de donner des éléments descriptifs sur vos pratiques de recherche habituelles : travail collectif ou individuel, sur quel type de matériaux, avec quels outils informatiques, etc. Il est anonyme et le remplissage prend environ 25 minutes. Il est ouvert jusqu'au 15 juin 2021 à cette adresse : <https://link.infini.fr/donnees-science-ouverte-coso> Claire Lemerrier

An examination of data reuse practices within highly cited articles of faculty at a research university <https://www.sciencedirect.com/science/article/pii/S0099133321000604>

FAIR or FAIRer? An integrated quantitative FAIRness assessment grid for semantic resources and ontologies <https://hal-lirmm.ccsd.cnrs.fr/lirmm-03208544/document>

Circulaire n°6264/SG du 27 avril 2021 relative à la politique publique de la donnée, des algorithmes et des codes sources <https://www.legifrance.gouv.fr/circulaire/id/45162>
avec la Liste des administrateurs ministériels des données <https://data.gouv.fr/fr/datasets/liste-des-administrateurs-ministeriels-des-donnees/>

Data dictionary cookbook for research data and software interoperability at global scale The objective of this poster is to report on the challenges

met while building data dictionaries in three global projects related to biodiversity and/or disease research: PARSEC, Kakila, ERINHA-Advance <https://hal.archives-ouvertes.fr/hal-03214743/document>

Data Management Plans in Horizon 2020: what beneficiaries think and what we can learn from their experience <https://open-research-europe.ec.europa.eu/articles/1-42/v1>

GO FAIR US Data Stewardship Interest <https://gofair.us/news/2021-05-8-data-stewardship-ig-mtg.html>

Générer des données aléatoires avec faker <https://www.docstring.fr/blog/generer-des-donnees-aleatoires-avec-faker/>

Pandas Profiling: Exploratory Data Analysis <https://sercangl.medium.com/pandas-profiling-exploratory-data-analysis-9a2e3ef5471c> avec Pandas Profiling <https://github.com/pandas-profiling/pandas-profiling>

Affiche « Cycle de vie des données : un outil pour améliorer la gestion, la mise en qualité et l'ouverture des données » <https://hal.inrae.fr/hal-03204351/document>

video

csv-detective: solving some of the mysteries of open data <https://www.youtube.com/watch?v=iCjmiya6JMU>
Rdataretriever: A platform for downloading, cleaning, and installing publicly available datasets <https://www.youtube.com/watch?v=lJr18H-vW7g>

et d'autres choses <https://csvconf.com/speakers/>

goldMEDAL : une nouvelle contribution à la modélisation générique des métadonnées des lacs de données <https://hal.archives-ouvertes.fr/hal-03231527>
<https://arxiv.org/pdf/2103.13155.pdf>

Cartographier les laboratoires d'analyses médicales, notre projet du mois de mai 2021 https://projetdumois.fr/projects/2021-05_laboratory

Entrepôt de données santé IQVIA : la CNIL rappelle les conditions et le cadre légal ayant permis son autorisation en 2018 <https://www.cnil.fr/fr/entrepot-de-donnees-sante-iqvia-cnil-rappelle-conditions-et-cadre-legal-ayant-permis-son-autorisation>

La CNIL publie son rapport d'activité 2020 <https://www.cnil.fr/fr/la-cnil-publie-son-rapport-dactivite-2020>

15 Juin 2021 : Webinaire PIN-Formats : Digital Preservation Handbook Le webinaire sera l'occasion de présenter ce Manuel de préservation numérique destiné aux professionnels qui souhaitent déployer dans leur institution une stratégie de préservation à long terme des objets numériques. <https://www.association-aristote.fr/webinaire-pin-formats-digital-preservation-handbook/>

04 June 2021 : Common Minimum Metadata for Semantic Artefact <https://www.fairsfair.eu/events/common-minimum-metadata-semantic-artefact>

23 juin 2021 : Workshop sur la recherche reproductible <https://mtp-rr2021.sciencesconf.org/>

26 May 2021 : Webinar: Discovering the FAIR Cookbook <https://fairplus-project.eu/about/news/cookbook-webinar> FAIR Cookbook que l'on peut retrouver à cette adresse <https://fairplus.github.io/the-fair-cookbook/content/home.html>

7-10 juin 2021 10 ans avec CAHIER: des corpus d'auteurs pour les humanités à leur exploitation numérique <https://cahier10.sciencesconf.org/program/details> avec [FAIR_Line] Permanence toute la journée - La permanence [FAIR_Line] est organisée toute la journée. Elle est destinée à vous aider à FAIRiser vos données avec l'entrepôt de données Nakala

15, 16, and 17 June 2021 : Dataverse Community Meeting 2021 <https://projects.iq.harvard.edu/dcm2021/home>

27 May 2021 : Versioning Data Is About More than Revisions: A Conceptual Framework and Proposed Principles <https://www.rd-alliance.org/data-versioning-wg-webinar-2021>

(l'événement est passé mais j'espère qu'un enregistrement sera disponible) 25
Mai 2021 Open Data et épidémiologie citoyenne <https://www.rencontresantepubliquefrance.fr/sessions/open-data-et-epidemiologie-citoyenne/>