



Appel à communications

Colloque interdisciplinaire APD-DAP

L'action publique des données et les données de l'action publique

Données ouvertes, action publique et #dataESR

Ce colloque interdisciplinaire se tiendra du 17 au 19 octobre 2022 à la MISHA de Strasbourg. Il sera consacré à l'exploration des liens entre les données ouvertes et l'action publique, avec une attention particulière, mais non exclusive, portée au champ institutionnel de l'enseignement supérieur et de la recherche (ESR).

Ce colloque rassemblera des acteurs du monde académique (chercheur·se·s et chercheur.se.s, jeunes et confirmé·e·s, ingénieur·e·s, bibliothécaires...) ainsi que du monde socio-économique (journalistes data, représentant·e·s d'organismes...) afin de croiser les points de vue dans une approche pluridisciplinaire (science des données, sciences politiques, sociologie, science de gestion, droit, philosophie, ...). Il a notamment vocation à permettre aux participants d'explorer ensemble de nouvelles perspectives méthodologiques dans la production et l'analyse des données ouvertes par et pour l'action publique.

Nous proposons de structurer ce colloque en trois temps :

- **l'action publique des données** : présentations et des discussions des travaux et projets sur la donnée, sa production et son utilisation dans l'action publique ;
- **les données de l'action publique** : présentations et des discussions des travaux et qui utilisent les données pour étudier des objets publics ;
- **appel à problèmes #DataESR** : ateliers participatifs pour co-construire des connaissances sur des problèmes rencontrés par les participant.e.s lors d'études dans le domaine de l'éducation et l'ESR, en utilisant les données ouvertes.

Organismes partenaires : MISHA, Laboratoire SAGE (UMR 7362), Université de Strasbourg, CNRS et CPESR.

Pour participer à ce colloque, veuillez nous faire parvenir **avant le 15 juin 2022** à l'adresse suivante :

<https://sondagesv3.unistra.fr/index.php/313383?lang=fr>

- **un titre, le nom des auteurs et l'email du contact**
- **un résumé de 5 lignes maximum**
- **un document pdf présentant vos travaux ou un problème #DataESR**, deux pages maximum hors graphiques et bibliographie.

Une attention sera portée à une démarche de science ouverte.

Comité scientifique : Paola Bertelli, Antoine Blanchard, Pauline Boyer, Séverine Chauvel, Florent Figon, Julien Gossa, Hugo Harari-Kermadec, Cédric Huguée, Ivana Milic, Romuald Normand, Mathieu Rossignol-Brunet et Mélanie Sargeac.

Contacts : julien.gossa@unistra.fr et pauline.boyer@cnrs.fr

Description de l'action publique des données

Outre la volonté de valoriser tous les travaux exploitant des données pour décrire, comprendre et agir dans l'action publique, incluant les travaux sur les données comme sur la conception d'indicateurs, de représentations, ou de documents, rapports ou tableaux de bords automatisés, ce colloque a vocation à encourager la réflexion collective sur les liens entre données ouvertes et action publique, autour de trois axes principaux :

- **La donnée ouverte comme objet de l'action publique** : Quelle(s) évolution(s) ont connues les politiques d'ouverture des données, pour quels objectifs ? Comment répandre la culture de l'ouverture des données au sein de l'administration, et de la société ? Quelle place occupent et devraient occuper les acteurs privés ? Peut-on parler d'une « privatisation numérique » (Jeannot and Cottin-Marx, 2022) ? Quelles sont les effets de l'ouverture des données sur les services publics ?
- **La donnée ouverte comme outil d'évaluation de l'action publique** : La donnée ouverte peut-elle être considérée comme un outil d'émancipation « statactiviste » (Bruno *et al.*, 2014) ? Peut-on évaluer l'action d'un organisme public à partir de données qu'il rend disponible ? Quelle information est créée et perdue lors de l'ouverture des données ? Comment se réapproprier la donnée et comment dépasser son indexicalité ? Qui sont les réutilisateurs des données ouvertes ? Quels sont les obstacles à la lisibilité et réutilisation des données ouvertes, particulièrement dans le champ de l'ESR, et comment les dépasser ? Quels indicateurs et quels outils de data visualisation prioriser ?
- **La donnée ouverte comme outil d'amélioration de l'efficacité de l'action publique** : Qui sont les « petites mains » (Denis and Pontille, 2012) de l'ouverture des données au sein des administrations ? Quels sont les processus et pratiques d'ouverture des données ? Comment simplifier les démarches administratives avec les données ouvertes ? Qu'apportent les données ouvertes dans l'échange d'informations entre administrations ? Les citoyens ont-ils un rôle à jouer dans l'amélioration de l'efficacité de l'action publique, et quel est-il ?

L'ouverture des données désigne la démarche qui vise à rendre gratuitement accessibles à tous et réutilisables les données numériques produites tant dans le secteur public que privé (Dulong de Rosnay, 2017). Des acteurs de la société civile ont joué un rôle clef dans la formulation des revendications de ce mouvement techno-politique, avec par exemple la rédaction de l'Open Definition (2005) et la création de l'Open Data Index (2012) par l'Open Knowledge Foundation, la conférence de Sébastopol en Californie (2007), ou encore la publication du modèle en 5 étoiles (2010) par Tim Berners-Lee (Goeta, 2016). Ces initiatives ont permis l'appropriation de ce sujet par un nombre croissant de gouvernements, avec l'adoption d'une charte par les chefs d'Etat du G8 (2013) et la progressive mise en œuvre de l'Open Government Partnership (2011) qui rassemble aujourd'hui 78 gouvernements et plusieurs milliers d'organisations de la société civile au sein d'une même structure.

La France s'est emparée de la question assez tôt, en créant dès 2011 la mission Etalab auprès du Premier Ministre, rapidement suivie par l'ouverture de la plateforme nationale de la donnée ouverte, data.gouv.fr. En 2016, la loi pour une République numérique (art. 6) fixe le principe d'ouverture par défaut pour toutes les données administratives, à quelques exceptions près, dans le respect du RGPD. On assiste donc à une augmentation significative de l'ouverture des données : début mai 2022, data.gouv.fr abrite 41 819 jeux de données.

L'enseignement supérieur et de la recherche (ESR) n'est pas en reste dans l'ouverture de ses données : en 2014, le MENESR a lancé la première plateforme ministérielle de données ouvertes, #dataESR, développée par une start-up en collaboration avec la mission Etalab, et qui contient aujourd'hui 146 jeux de données. Par ailleurs, ce portail ne contient pas toutes les données ouvertes disponibles sur l'ESR, puisque certains opérateurs de l'Etat ont mis en ligne des données sans passer par #dataESR (ABES, ANR, ...).

Cependant, donnée ouverte ne veut pas nécessairement dire donnée réutilisable, et quantité de données ouvertes ne satisfont pas aux critères des différents référentiels existants pour évaluer la qualité des données ouvertes, tels que le modèle en 5 étoiles, mais également les 8 principes de l'ouverture des données publiques adoptés lors de la conférence de Sébastopol ou les recommandations du World Wide Web Consortium (W3C). A titre d'exemple, un fichier pdf ou une photo librement disponible en ligne peuvent, selon certaines perspectives, être considérés comme des données ouvertes, alors que des données libres mais peu structurées et non documentées peuvent être inexploitable.

Cela nous invite à nous interroger sur le travail qu'exige l'ouverture des données, et à battre en brèche deux mythes répandus à propos des données et *a fortiori* des données ouvertes : d'une part, le présupposé d'une disponibilité des données qui pré-existeraient à leur 'libération' et réutilisation, et d'autre part le postulat que les données peuvent aisément circuler d'une structure à une autre, et donc être réutilisables, sans friction. En effet, le concept-même de donnée brute est oxymorique (Gitelman, 2013) et, pour reprendre la phrase célèbre de Bruno Latour, « on ne devrait jamais parler de *données* mais d'*obtenues* » (Latour, 2007: 188). L'ouverture des données représente une tâche chronophage, avec des choix stratégiques à opérer (notamment l'identification des données à ouvrir) et de gestes techniques à accomplir pour les *brutifier* (reformatage, nettoyage, désindexicalisation ; Denis and Goëta, 2013). Une fois ouvertes les données, demeure le travail de maintenance (mise à jour des données périmées, enrichissement des données et métadonnées, ... ; Attard *et al.*, 2015). La circulation des jeux de données suscite également des frictions (Edwards *et al.*, 2011) et implique donc une charge de travail supplémentaire pour minimiser ces frictions en documentant les données (les métadonnées).

L'exploitation des données ouvertes se révèle particulièrement complexe lorsque les jeux de données portent sur l'ESR. En effet, le périmètre de l'ESR est en constante évolution, en raison notamment des différentes politiques de regroupement successivement expérimentées (Aust and Crespy, 2009; Musselin, 2017; Guiselin, 2019; Eyraud, 2021). On comptait ainsi en 2005 81 universités, quand elles ne sont plus que 71 en 2020. A ces modifications du paysage s'ajoutent la difficulté de la coexistence de plusieurs périmètres (établissement, COMUE, EPE, association, coordination territoriale), et celle des multi-appartenances. Cette complexité institutionnelle se reflète dans la structuration des jeux de données, dont la manipulation appelle donc à une fine connaissance du paysage institutionnel en plus des compétences techniques.

Ainsi, tout au long du cycle de vie des données ouvertes, on crée de l'information, et on en perd. Cette information, créée ou invisibilisée, n'est jamais neutre, ce qui nous invite à nous questionner sur les conditions de réutilisation des données ouvertes : peut-on évaluer l'action publique à partir de données mises à notre disposition par l'institution qu'il s'agirait d'évaluer ?

Au-delà de l'enthousiasme suscité par l'ouverture des données, il s'agit de conserver une attitude critique face à un discours que partis politiques et acteurs du monde économique ont très largement et très rapidement repris à leur compte : si le discours institutionnel vante l'amélioration de l'efficacité des services publics, Bates (2012) soutient que ce mouvement faciliterait leur néolibéralisation et marchandisation, tandis que Jeannot et Cottin-Marx (2022) parlent d'une « privatisation numérique » (voir aussi : Gurstein, 2011; Longo, 2011). Quels sont donc les effets de l'ouverture des données sur les services publics ?

Bibliographie indicative

- Attard, J. et al. (2015) 'A systematic review of open government data initiatives', *Government Information Quarterly*, 32(4), pp. 399–418. doi:10.1016/j.giq.2015.07.006.
- Aust, J. and Crespy, C. (2009) 'Napoléon renversé ? : Institutionnalisation des Pôles de recherche et d'enseignement supérieur et réforme du système académique français', *Revue française de science politique*, 59(5), p. 915. doi:10.3917/rfsp.595.0915.
- Bates, J. (2012) "'This is what modern deregulation looks like" : co-optation and contestation in the shaping of the UK's Open Government Data Initiative', *The Journal of Community Informatics*, 8(2). doi:10.15353/joci.v8i2.3038.
- Bogdanović-Dinić, S., Veljković, N. and Stoimenov, L. (2014) 'How Open Are Public Government Data? An Assessment of Seven Open Data Portals', in Rodríguez-Bolívar, M.P. (ed.) *Measuring E-government Efficiency: The Opinions of Public Administrators and Other Stakeholders*. New York, NY: Springer (Public Administration and Information Technology), pp. 25–44. doi:10.1007/978-1-4614-9982-4_3.
- ter Bogt, H.J. and Scapens, R.W. (2012) 'Performance Management in Universities: Effects of the Transition to More Quantitative Measurement Systems', *European Accounting Review*, 21(3), pp. 451–497. doi:10.1080/09638180.2012.668323.
- Boltanski, L. (2014) 'Quelle statistique pour quelle critique ?', in Bruno, I., Didier, E., and Prévieux, J. (eds) *Statactivisme*. Paris: La Découverte (Zones), pp. 33–50.
- Bruno, I. et al. (2014) 'Introduction. Pour un statactivisme !', in Bruno, I., Didier, E., and Prévieux, J. (eds) *Statactivisme*. Paris: La Découverte (Zones), pp. 5–30.
- Denis, J. and Goeta, S. (2013) 'La fabrique des données brutes. Le travail en coulisses de l'open data', in *Penser l'écosystème des données. Les enjeux scientifiques et politiques des données numériques*. Available at: <https://halshs.archives-ouvertes.fr/halshs-00990771> (Accessed: 5 October 2021).
- Denis, J. and Pontille, D. (2012) 'Travailleurs de l'écrit, matières de l'information', *Revue d'anthropologie des connaissances*, 61(1), pp. 1–20.
- Desrosières, A. (2008) *Pour une sociologie historique de la quantification : L'Argument statistique I, Pour une sociologie historique de la quantification : L'Argument statistique I*. Paris: Presses des Mines (Sciences sociales).
- Desrosières, A. and Kott, S. (2005) 'Quantifier', *Geneses*, 58(1), pp. 2–3.
- Dulong de Rosnay, M. (2017) 'Données ouvertes (Open Data)', in *Dictionnaire des biens communs*. PUF (Quadrige), pp. 403–407. Available at: <https://halshs.archives-ouvertes.fr/halshs-01575643> (Accessed: 9 March 2022).
- Edwards, P.N. et al. (2011) 'Science friction: data, metadata, and collaboration.', *Social studies of science*, 41(5), pp. 667–690. doi:10.1177/0306312711413314.
- Eyraud, C. (2021) 'L'Université française : après la LRU, les fusions, les PIA..., voici venu le temps des EPE', *Revue Française de Socio-Economie*, 27(2), pp. 5–9.
- Eyraud, C., Miri, M.E. and Perez, P. (2020) 'Les enjeux de quantification dans la LOLF. Le cas de l'enseignement supérieur', *Revue Française de Socio-Economie*, (en lutte), pp. 219–240.
- Gao, Y., Janssen, M. and Zhang, C. (2021) 'Understanding the evolution of open government data research: towards open data sustainability and smartness', *International Review of Administrative Sciences*, p. 00208523211009955. doi:10.1177/00208523211009955.
- Gitelman, L. (ed.) (2013) *'Raw Data' Is an Oxymoron*. Cambridge, MA, USA: MIT Press (Infrastructures).
- Goeta, S. (2016) *Instaurer des données, instaurer des publics : une enquête sociologique dans les coulisses de l'open data*. phdthesis. Télécom ParisTech. Available at: <https://pastel.archives-ouvertes.fr/tel-01458098> (Accessed: 12 March 2022).
- Goeta, S. (2018) 'Une petite histoire d'Etalab : comment l'open data s'est institutionnalise en France?',

Statistique et Société [Preprint]. Available at: <https://hal.archives-ouvertes.fr/hal-01724275> (Accessed: 8 March 2022).

Guiselin, E.-P. (2019) 'Les regroupements d'établissements dans l'enseignement supérieur et la recherche : enjeux politiques et cadrage juridique', *Revue française d'administration publique*, N° 169(1), pp. 37–50.

Gurstein, M.B. (2011) 'Open data: Empowering the empowered or effective data use for everyone?', *First Monday* [Preprint]. doi:10.5210/fm.v16i2.3316.

Harari-Kermadec, H. (2019) *Le classement de Shanghai. L'université marchandisée*. Le Bord de l'eau (L'économie encastrée).

Jeannot, G. and Cottin-Marx, S. (2022) *La privatisation numérique*. Raisons d'agir. Paris.

Kitchin, R. (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. SAGE.

Lascombes, P. and Galès, P.L. (2005) *Introduction : L'action publique saisie par ses instruments*. Presses de Sciences Po. Available at: <https://www.cairn.info/gouverner-par-les-instruments--9782724609492-page-11.htm> (Accessed: 6 April 2022).

Latour, B. (2007) 'Le « pédofil » de Boa Vista – montage photo-philosophique', *Poche / Sciences humaines et sociales*, pp. 171–225.

Longo, J. (2011) '# Opendata: Digital-Era Governance Thoroughbred or New Public Management Trojan Horse?', *Public Policy & Governance Review*, Vol. 2, No. 2, p. 38 ... [Preprint]. Available at: https://www.academia.edu/1045781/_Opendata_Digital_Era_Governance_Thoroughbred_or_New_Public_Management_Trojan_Horse (Accessed: 18 March 2022).

Lourenço, R. (2013) 'Open Government Portals Assessment: A Transparency for Accountability Perspective', in: *12th International Conference on Electronic Government (EGOV)*, Springer, p. 62. doi:10.1007/978-3-642-40358-3_6.

Mennicken, A. and Salais, R. (2022) 'The New Politics of Numbers: An Introduction', in Mennicken, A. and Salais, R. (eds) *The New Politics of Numbers*. Cham: Springer International Publishing, pp. 1–42. doi:10.1007/978-3-030-78201-6_1.

Mocquet, B. (2019) *Gouvernance, numérique et enseignement supérieur*. Paris: Presse des Mines (Design numérique).

Musselin, C. (2017) *La grande course des universités*. Paris: Presses de Sciences Po.

Ollion, E. and Boelaert, J. (2015) 'Au-delà des big data. Les sciences sociales et la multiplication des données numériques', *Sociologie*, 6(3).

Ruijter, E.H.J.M. and Martinius, E. (2017) 'Researching the democratic impact of open government data: A systematic literature review', *Information Polity*, 22(4), pp. 233–250. doi:10.3233/IP-170413.

Salais, R. (2022) "'La donnée n'est pas un donné": Statistics, Quantification and Democratic Choice', in Mennicken, A. and Salais, R. (eds) *The New Politics of Numbers*. Cham: Springer International Publishing, pp. 379–415. doi:10.1007/978-3-030-78201-6_12.

Strathern, M. (2000) 'The Tyranny of Transparency', *British Educational Research Journal*, 26(3), pp. 309–321. doi:10.1080/713651562.

Thévenot, L. (2019) 'Measure for Measure: Politics of Quantifying Individuals to Govern Them', *Historical Social Research / Historische Sozialforschung*, 44(2 (168)), pp. 44–76.

Yang, T.-M., Lo, J. and Shiang, J. (2015) 'To open or not to open? Determinants of open government data', *Journal of Information Science*, 41(5), pp. 596–612. doi:10.1177/0165551515586715.

Zuiderwijk, A. and Janssen, M. (2014) 'The negative effects of open government data - investigating the dark side of open data', in *Proceedings of the 15th Annual International Conference on Digital Government Research*. New York, NY, USA: Association for Computing Machinery (dg.o '14), pp. 147–152. doi:10.1145/2612733.2612761.