

# Services de gestion et de partage des données de recherche : ce qu'en pensent les chercheurs

Violaine Rebouillat & Ghislaine Chartron

EA 7339 DICEN-IdF (Cnam, Paris)

ISKO-France 9-11 octobre 2019



# CADRE CONTEXTUEL

# Les données de la recherche

- Une notion qui recouvre des réalités différentes
- Il n'existe pas de données en elles-mêmes (**Leonelli, 2015**)

« La représentation d'une observation, d'un objet ou de toute autre entité devient une donnée de recherche dès lors qu'elle est utilisée comme preuve d'un phénomène, à des fins scientifiques. »  
[trad. libre] (**Borgman, 2015**)

# Pratiques autour des données

Une diversité de pratiques, qui varient en fonction des individus et des communautés de recherche

- **Tenopir et al., 2015**
- **Serres et al., 2017**
- **Schöpfel, 2018**
- **Stuart et al., 2018**

# Politiques d'ouverture des données

- Déclaration de Berlin (2003)
- Déclaration de l'OCDE (2004)
- Communication de Commission européenne (2012)
- Plan national pour la science ouverte (2018)

# Premiers acteurs du mouvement d'ouverture des données en France

- Les professionnels de l'information scientifique et technique
  - Une profession en renouvellement
  - Le développement de services pour la gestion et la diffusion des données de recherche



beQuali



DataPartage



# Le catalogue Cat OPIDoR

- Objectif : répertorier les services français dédiés aux données scientifiques
- Commandité par : le groupe *Données de la recherche* de l'ancienne Bibliothèque Scientifique Numérique
- Conçu et modéré par : l'Inist-CNRS
  
- Des services nés du mouvement d'ouverture, mais aussi des services « disciplinaires »

RESIF Seismic Data Portal  
Entrepôt du CDDP  
Webservice-Energy Catalog  
BASS 2000  
EELS Data Base  
AVISO+  
VRP-REP  
GAPHYOR  
CDS Portal  
InfoTerre

Ortolang  
Nakala  
ArkeoGIS  
BeQuali  
CocoON  
Kinsources.net  
ArchiPolis  
Réseau Quetelet

**SCIENCES & TECHNOLOGIES**

**SCIENCES HUMAINES & SOCIALES**

**SERVICES MULTIDISCIPLINAIRES**

PerSciDo  
MédiHAL  
MISTRALS Database  
Annuaire du Cirad  
IO Data Science  
DMP OPIDoR  
Agence DataCite  
Equipe Valorisation  
Accompagnement USPC  
Accompagnement Unistra  
DoRANum  
Site Données de la Recherche  
CoopIST Cirad  
Archivage du CINES

Portail Data INRA  
Data BRGM  
Coriolis  
ILL Data Portal  
IrsteaData  
Datapartage

Annuaire BBEES

**VIE & SANTE**

SEANOE  
ESTHER DB  
Carbohydrate-Active enZYmes Database  
Portail Epidémiologie France  
ECOSCOPE





# PROBLÉMATIQUE ET MÉTHODOLOGIE

# Utilisation des services de données par les chercheurs

- Quels services de données utilisent les chercheurs ?
- Connaissent-ils Cat OPIDoR et les services qui y sont répertoriés ?
- Ce catalogue et ces services ont-ils une utilité pour eux ?

# Une enquête auprès de chercheurs

- À l'Université de Strasbourg
- Sous forme d'entretiens semi-directifs

Domaine scientifique	Laboratoire	Discipline du laboratoire	Nombre de chercheurs contactés	Nombre d'entretiens menés	Taux de participation
Vie & Santé	Laboratoire 1	Biologie	1	1	100 %
	Laboratoire 2	Chimie et biologie	30	3	10 %
	Laboratoire 3	Chimie et biologie	22	5	23 %
	Laboratoire 4	Écologie	19	3	16 %
	Laboratoire 5	Neurosciences	20	4	20 %
	Laboratoire 6	Neurosciences	10	1	10 %
Sciences & Technologies	Laboratoire 7	Astronomie	16	7	44 %
	Laboratoire 8	Sciences de l'ingénieur	33	5	15 %
	Laboratoire 9	Sciences de la Terre	23	3	13 %
Sciences humaines et sociales	Laboratoire 10	Géographie	22	4	18 %
	Laboratoire 11	Sciences sociales	20	8	40 %
	Laboratoire 12	Théologie	17	2	12 %
<b>Total</b>			<b>233</b>	<b>46</b>	<b>27 %</b>



# RÉSULTATS

# Les chercheurs et Cat OPIDoR

- Inconnu des chercheurs interrogés
- Jugé potentiellement utile
- Inconvénient : son périmètre national
  
- Services répertoriés peu connus des chercheurs

[Synchrotron SOLEIL](#)

[PROGEDO](#)

[Recolnat](#)

[CDS Portal](#)

[Pôle HPC Unistra](#)

# Les chercheurs et les services de données (1/2)

- Recours peu fréquent à des services de données
  - Méconnaissance des services existants
  - Approche non outillée des données (en droit, histoire et théologie)
  - Nécessaire proximité du chercheur avec les données (en sociologie et anthropologie)

# Les chercheurs et les services de données (2/2)

- Recours plus fréquent à des services de collecte
  - Télescopes en astronomie
  - Panels de population en démographie
  - Séquenceurs ADN en biologie, chimie et écologie
- Et plus généralement à des services « disciplinaires »
  - Développés au sein des communautés
  - A l'échelle internationale



# Les chercheurs et l'ouverture des données (1/2)

- Un sujet parlant mais éloigné de leurs préoccupations
  - « Economie de la réputation » (**Fecher et al., 2015**)
  - Système d'évaluation fondé sur les publications
  - Données reléguées au second plan

# Les chercheurs et l'ouverture des données (2/2)

Un rapport privilégié à la publication

- Enquête Springer : 63% des répondants partagent les données sous-jacentes à une publication (**Stuart et al., 2018**)
- Dans notre enquête : 70%
- Résultats à nuancer
  - Des pratiques ponctuelles
  - Des pratiques guidées par l'importance de publier

Domaine scientifique	Laboratoire	Discipline du laboratoire	Nombre de chercheurs rencontrés	Nombre de chercheurs ayant déjà partagé des données sous-jacentes à une publication	Proportion
Vie & Santé	Laboratoire 1	Biologie	1	1	100 %
	Laboratoire 2	Chimie et biologie	3	3	100 %
	Laboratoire 3	Chimie et biologie	5	5	100 %
	Laboratoire 4	Écologie	3	3	100 %
	Laboratoire 5	Neurosciences	4	4	100 %
	Laboratoire 6	Neurosciences	1	1	100 %
Sciences & Technologies	Laboratoire 7	Astronomie	7	2	29 %
	Laboratoire 8	Sciences de l'ingénieur	5	2	40 %
	Laboratoire 9	Sciences de la Terre	3	2	67 %
Sciences humaines et sociales	Laboratoire 10	Géographie	4	3	75 %
	Laboratoire 11	Sciences sociales	8	2	25 %
	Laboratoire 12	Théologie	2	0	0 %
		<b>Total</b>	<b>46</b>	<b>28</b>	<b>70 %</b>

# Les chercheurs et l'ouverture des données (2/2)

Un rapport privilégié à la publication

- Enquête Springer : 63% des répondants partagent les données sous-jacentes à une publication (**Stuart et al., 2018**)
- Dans notre enquête : 70%
- Résultats à nuancer
  - Des pratiques ponctuelles
  - Des pratiques guidées par l'importance de publier



# CONCLUSION

# Conclusions de l'enquête

Les services développés et mis en avant par les professionnels de l'IST semblent correspondre pour une faible partie à ceux qu'utilisent les chercheurs.

- Un échantillon peu représentatif des usages de la communauté scientifique ?
- Une offre de service qui arrive trop tôt ?
- Un pouvoir d'influence limité ?

# Perspectives

- Quel avenir pour les services développés par les professionnels de l'IST ?
- Vers une coordination entre acteurs publics de la recherche et éditeurs scientifiques ?
- Comment tenir compte du cadre institutionnel, épistémique et social dans le développement de services de données ?

# Merci de votre attention

## le cnam

### **Violaine Rebouillat**

Doctorante en sciences de l'information et de la communication

EA 7339 DICEN-IdF

[violaine.rebouillat.auditeur@lecnam.net](mailto:violaine.rebouillat.auditeur@lecnam.net)

### **Ghislaine Chartron**

Professeur en sciences de l'information et de la communication

EA 7339 DICEN-IdF

[ghislaine.chartron@lecnam.net](mailto:ghislaine.chartron@lecnam.net)



# Bibliographie

- BORGMAN, C. L. (2015). *Big data, little data, no data: scholarship in the networked world*. Cambridge: The MIT Press.
- FECHER, B. ET AL. (2015). 'A Reputation Economy: Results from an Empirical Survey on Academic Data Sharing'. *DIW Berlin Discussion Paper No. 1454*. <http://dx.doi.org/10.2139/ssrn.2568693>
- LEONELLI, S. (2015). 'What Counts as Scientific Data? A Relational Framework'. *Philosophy of Science*. 82 (5): 810–21. <https://www.jstor.org/stable/10.1086/684083>
- SCHÖPFEL, J. (2018). *Vers une culture de la donnée en SHS. Une étude à l'Université de Lille*. Université de Lille. <https://hal.archives-ouvertes.fr/hal-01846849>
- SERRES, A. ET AL. (2017). *Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l'Université Rennes 2*. Rapport de recherche. Université Rennes 2. <https://hal.archives-ouvertes.fr/hal-01635186v2>
- STUART, D. ET AL. (2018). *Whitepaper: Practical challenges for researchers in data sharing*. [https://figshare.com/articles/Whitepaper\\_Practical\\_challenges\\_for\\_researchers\\_in\\_data\\_sharing/5975011](https://figshare.com/articles/Whitepaper_Practical_challenges_for_researchers_in_data_sharing/5975011)
- TENOPIR, C., ET AL. (2015). 'Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide'. *PLOS ONE*. 10 (8). <http://dx.doi.org/10.1371/journal.pone.0134826>

# Sources

*Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* (2003).

<https://openaccess.mpg.de/Berlin-Declaration>

**COMMISSION EUROPÉENNE (2012).** *Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au comité des régions. Pour un meilleur accès aux informations scientifiques: dynamiser les avantages des investissements publics dans le domaine de la recherche.*

<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:52012DC0401>

**MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION (2018).** *Plan national pour la science ouverte.*

[http://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/DI AN\\_NATIONAL\\_SCIENCE\\_OUVERTE\\_978672.pdf](http://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/DI AN_NATIONAL_SCIENCE_OUVERTE_978672.pdf)

**ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES (2004).** *Déclaration sur l'accès aux données de la recherche financée par des fonds publics.* <https://legalinstruments.oecd.org/fr/instruments/157>