



Éditorial - La diffusion numérique des données en SHS - Guide de bonnes pratiques éthiques et juridiques

Stéphane Pouyllau

► To cite this version:

Stéphane Pouyllau. Éditorial - La diffusion numérique des données en SHS - Guide de bonnes pratiques éthiques et juridiques. Véronique Ginouvès; Isabelle Gras. La diffusion numérique des données en SHS - Guide de bonnes pratiques éthiques et juridiques, Presses universitaires de Provence, pp.299-302, 2018, Digitales, 9791032001790. hal-01974527

HAL Id: hal-01974527

<https://amu.hal.science/hal-01974527>

Submitted on 8 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

La diffusion numérique des données en SHS

Guide des bonnes pratiques éthiques et juridiques

sous la direction de
Véronique Ginouvès & Isabelle Gras



DIGITALES





DIGITALES

La diffusion numérique des données en SHS

Guide des bonnes pratiques
éthiques et juridiques

sous la direction de

Véronique Ginouvès & Isabelle Gras

2018

PRESSES UNIVERSITAIRES DE PROVENCE

Tous les textes sont placés en licence CC-BY, avec l'accord des auteurs.

© PRESSES UNIVERSITAIRES DE PROVENCE

Aix-Marseille Université

29, avenue Robert-Schuman – F – 13621 Aix-en-Provence CEDEX 1

Tél. 33 (0)4 13 55 31 91

pup@univ-amu.fr – Catalogue complet sur presses-universitaires.univ-amu.fr

DIFFUSION LIBRAIRIES : AFPU DIFFUSION – DISTRIBUTION SODIS

Éditorial

Stéphane Pouyllau

Directeur technique d'Huma-Num, Ingénieur de recherche CNRS

La très grande infrastructure de recherche Huma-Num (TGIR), créée en 2013 de la fusion des infrastructures de recherche Adonis et Corpus, propose à l'ensemble des communautés de recherche en Lettres, sciences humaines et science sociales (LSHS) un ensemble de services numériques interconnectés répondant aux principes d'ouvertures des données selon les principes dit FAIR (*Findable, Accessible, Interoperable, Reusable*). Ces services offrent la possibilité aux équipes universitaires et aux organismes de recherche de stocker, de traiter, de documenter, de diffuser et d'archiver leurs données de recherche. Les données de la recherche sont multiples dans les disciplines des LSHS et suivant que l'on soit historien, sociologue, ethnologue, archéologue, géographe, etc. les objets d'études seront évidemment différents. Ce qui est commun aujourd'hui c'est l'ampleur de la dimension numérique des dispositifs de collecte (de l'enquête sur les terrains, au dépouillement de fonds d'archives en passant par les relevés à l'aide de drones), de la façon d'organiser les corpus, de l'instrumentation d'analyse et enfin des façons d'envisager le partage des données, les chemins pour travailler les données sont multiples. Pour autant, les enjeux de stockage, de conservation, de traitement, d'accès et de partage des données de recherche sont transversaux et la gestion en « silos » disciplinaires est révolue.

L'amélioration des moyens de stockage et de calcul informatique ont permis aux disciplines des LSHS de constituer et d'analyser (ou de réinterroger) des corpus plus accessibles, plus étendus et plus structurés. L'encodage XML et l'analyse de textes, l'indexation de flux audio et vidéo, le traitement de l'information géographique avec les SIG et la visualisation en temps réel d'espaces et objets 3D ont transformé la façon de construire le savoir, de le publier et de le partager. C'est dans ce contexte qu'Huma-Num propose depuis 2013 un écosystème complet formé de communautés de travail et de services numériques dédiés au stockage, au traitement, à l'archivage et à la publication et au signalement des données et documents de recherche organisés ou pas en corpus raisonnés. Les communautés de travail, organisées par Huma-Num sous la forme de consortiums disciplinaires, ancrées dans les maisons des sciences de l'Homme, forment un réseau d'expertise sur des LSHS en matière d'informatisation et traitement des données numériques de la recherche

L'ensemble des services d'Huma-Num offre aux enseignants-chercheurs la possibilité de partager, diffuser et signaler leurs documents et données, dans la logique des principes FAIR et dont l'objectif est de favoriser la découverte, l'accès, l'interopérabilité et la réutilisation des données de la recherche. Qu'il s'agisse de NAKALA, pour déposer et documenter des corpus de données numériques, ou

d'ISIDORE, qui permet de partager et de placer les corpus au cœur d'un moteur de recherche SHS, les principes FAIR sont aujourd'hui au cœur des services de la TGIR. Dans ce cadre, et dans le but d'améliorer en cycles itératifs les services d'Huma-Num, l'apport des réflexions des communautés scientifiques est déterminant. Si les consortiums d'Huma-Num et les MSH constituent le premier creuset naturel dans lequel se développe la « R&D » des futurs services d'Huma-Num, les travaux, prototypes, expérimentations qui sont réalisés dans les domaines de la recherche en informatique, de l'information scientifique et technique et chez les industriels de la connaissance, doivent venir nourrir aussi les programmes de recherche des LSHS afin d'en décloisonner les réalisations. Les services d'Huma-Num sont au service de l'ensemble des communautés LSHS et ne sont pas réservés aux programmes les plus pointus des humanités numériques. Ils permettent, dès les premiers besoins, dès les premiers pas — par exemple le simple partage de fichiers au sein d'un programme de recherche, de partir sur de bonnes pratiques. Elles seront ainsi un investissement pour plus tard. Ainsi, en proposant dans NAKALA une interopérabilité des métadonnées fondée sur le protocole documentaire OAI-PMH, puis en plaçant aussi les métadonnées dans le *Linked Open Data* via la mise à disposition d'une base de données RDF (*triple-store*), NAKALA s'adresse à la fois à ceux qui souhaitent développer des corpus de données numériques partageables et moissonnables par des portails tel que ISIDORE, Gallica (BnF) et autres portails disciplinaires sur le web. La généralisation, d'outils co-développés par les communautés LSHS et la Huma-Num (OpenThésos, WebOAI, Cocoon, ArchéoGIS, NAKALONA, etc.), accroît les processus de partages et la réutilisation des données. Ce mouvement va maintenant plus loin, il est accompagné désormais de travaux interdisciplinaires au sein des LSHS pour développer des outils échangeables, créant ainsi de véritables communautés de développement autour des corpus, des bibliothèques numériques, des plateformes de données et documents numériques. Ces communautés ont un regard critique sur les outils et dispositifs développés, c'est la marque des LSHS. Le rôle d'Huma-Num est aussi d'accompagner ces communautés de recherche qui créent, avec nous, des infrastructures « critiquées » et qui sont les véritables humanités numériques. Ainsi, les principes FAIR sont, au sein d'Huma-Num, à la fois une réalité technologique, mais aussi une réalité humaine dont les communautés LSHS doivent s'emparer afin de co-piloter et de valoriser les réalisations et les infrastructures existantes.

Le présent guide de bonnes pratiques éthiques et juridiques s'adresse précisément à ces communautés qui pourront l'utiliser comme un outil, parmi d'autres, pour construire leur projet de diffusion numérique des données.

Table des matières

Chercheurs, quand je serai mort qui prendra soin de ma page FB, GS, RG, CvHAL, Hypothèses.org ? David Aymonin	5
Éditorial Stéphane Pouyllau	7
Préface Marie Masclet de Barbarin	9
État des lieux sur les bonnes pratiques éthiques et juridiques en matière de diffusion des données en SHS	
Diffuser des données de la recherche dans le respect du droit et de l'éthique Comment faire lorsqu'on n'est pas juriste ? Anne-Laure Stérin	19
Pratiques d'archives Problèmes actuels sur les usages du matériau documentaire Jean-François Bert	31
Preserving Public Domain Collections. Institutional Policies Best Practices Mélanie Dulong de Rosnay	39
La réutilisation des données de la recherche après la loi pour une République numérique Lionel Maurel	49
<i>Big data</i> en sciences sociales et protection des données personnelles Émilie Debaets	61
Dématérialisation et valorisation des matériaux de terrain des ethnologues L'archiviste face aux questions éthiques Marie-Dominique Mouton	73
Comment diffuser les données en SHS ? Réalisations et retours d'expérience Les archives orales, chapitre introduit par Florence Descamps	
Introduction Florence Descamps	91

La parole et le droit Recommandations pour la collecte, le traitement et l'exploitation des témoignages oraux Raphaëlle Branche, Florence Descamps, Frédéric Saffroy, Maurice Vaïsse	103
Two Oral History Projects, Two Countries and the Encountered Issues and Subsequent Solutions to Online Recording Accessibility Issues Leslie McCartney	129
Consent in the digital context The example of oral history interviews in the United Kingdom Myriam Fellous-Sigrist	143
Ouverture de données qualitatives à caractère personnel Approche éthique, juridique et déontologique Marie Huyghe, Laurent Cailly, Nicolas Oppenheim	159
Les archives sonores entre demande sociale et usages scientifiques Quelles modalités pour réutiliser les sources enregistrées ? Francesca Biliotti, Silvia Calamai, Véronique Ginouvès	169
Les données sensibles de la recherche, chapitre introduit par Laurent Dousset	
Données sensibles. Peuvent-elles ne pas l'être ? Laurent Dousset	197
Anonymat et confidentialité des données. L'expérience de beQuali Selma Bendjaballah, Sarah Cadorel, Émilie Fromont, Guillaume Garcia, Émilie Groshens, Emeline Juillard	207
Du remède par les plantes à la sorcellerie Retour sur une expérience de traitement et de diffusion d'archives orales en Bretagne Maëlle Mériaux	223
MEMORIA – la préservation des processus d'étude comme enjeu éthique Iwona Dudek, Jean-Yves Blaise	231
Le traitement des données d'un défunt dans un contexte de recherche Jean-Charles Ize	241
L'évolution du droit en matière de numérique, chapitre introduit par Philippe Mouron	
Droit d'auteur et diffusion numérique des données de la recherche Philippe Mouron	247
Les enjeux éthiques et juridiques du dépôt des travaux scientifiques dans une archive ouverte Isabelle Gras	255



Les robots sont-ils des lecteurs comme les autres ?	267
Émergence et codification d'une exception au droit d'auteur pour le <i>text & data mining</i>	
Pierre-Carl Langlais	
La confiscation des données issues de l'humanisme numérique	283
Un paradoxe résistible	
Marie-Luce Demonet	
Postface	299
Véronique Ginouvès, Isabelle Gras	
Bibliographie	303
Biographie des auteurs	327



La diffusion numérique des données en SHS

Guide des bonnes pratiques éthiques et juridiques

DIGITALES

La collection
« Digitales »
s'intéresse
aux rapports
entre les
sciences humaines
et le monde
numérique, qu'il
fournisse des
outils critiques
ou qu'il soit
un domaine
de création.

Produire, exploiter, éditer, publier ou valoriser des données numériques fait partie du travail quotidien des chercheurs en sciences humaines et sociales (SHS). Ces données sont aujourd'hui disséminées sous de multiples formats dans le monde de la recherche et, au-delà, auprès de citoyens de plus en plus curieux et intéressés par les documents produits par les scientifiques. Dans un contexte de mutation fulgurante des méthodes de travail, ce guide aborde avec simplicité des questions et des enjeux complexes auxquels se confronte quotidiennement la communauté des SHS. De leur collecte à leur réutilisation, les données de la recherche sont manipulées, éditorialisées, interrogées, mises en ligne... par tous les acteurs du monde académique qui ne savent pas toujours répondre aux questions juridiques et éthiques ou même, ne parviennent pas à les poser clairement. C'est à eux que s'adresse cet ouvrage, fondé sur des réflexions et des retours d'expériences qui présentent les bonnes pratiques pour accompagner celles et ceux qui s'inscrivent dans la dynamique de la science ouverte.

conception graphique
et illustration de couverture
J.-B. Cholbi

Véronique Ginouvès est responsable des archives sonores et audiovisuelles à la Maison méditerranéenne des sciences de l'homme (AMU-CNRS) à Aix-en-Provence.

Isabelle Gras est conservatrice des bibliothèques au Service commun de la documentation de l'université d'Aix-Marseille (SCD AMU).

Presses
Universitaires
de Provence



Aix-Marseille
université
Initiative d'excellence

Bibliothèques
universitaires



Maison méditerranéenne
des sciences de l'homme
USR 3125

Huma-Num
la TGR des humanités numériques



20 €