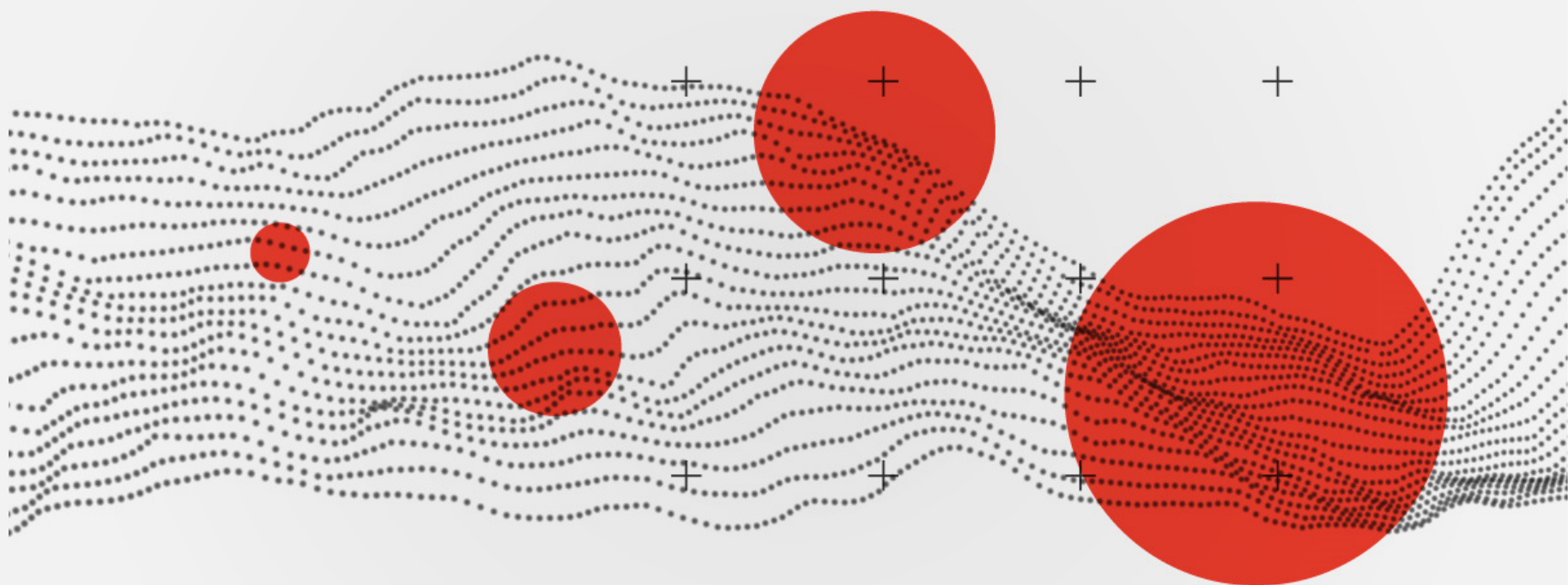


RES  
PA  
DON

## Webinaire - Atelier 3



Groupe de réflexion

[À distance](#)

Date : 23/06/22

© Login/Shutterstock, mise en page © Sandrine Lancereau/Sciences Po



Université  
de Lille

{BnF | Bibliothèque  
nationale de France

SciencesPo

CAMPUS  
CONDORCET  
PARIS - AUBERVILLIERS

gériico

CollEx-Persée

# Cycle d'ateliers – 2021-2023

PUBLIÉ 27/09/2021 · MIS À JOUR 28/09/2021



The Data Lifecycle ©Mushonz via Wikimedia Commons [https://commons.wikimedia.org/wiki/File:The\\_Data\\_Lifecycle.jpg](https://commons.wikimedia.org/wiki/File:The_Data_Lifecycle.jpg)

Le Réseau de Partenaires pour l'exploration et l'analyse de données numériques (ResPaDon) propose un cycle d'ateliers trimestriels.

## Objectifs

Il se fixe pour objectif d'échanger et de consolider les savoir-faire et l'expertise des professionnels de l'information scientifique et technique (IST) autour de l'exploitation des archives du dépôt légal du web.

Pendant deux années, les rendez-vous trimestriels du groupe de réflexion ouvrent un espace de dialogue et offrent aux participants l'opportunité de mettre en évidence des pratiques et des savoir-faire professionnels afin de consolider leur expertise et de mutualiser leurs expériences pour favoriser une montée en compétence des professionnels et in fine, encourager et faciliter l'usage des archives web et des données par les chercheurs.

A terme, les travaux du groupe de réflexion donneront lieu à la formalisation de préconisations qui recouvriront les différents enjeux du [projet ResPadon](#).

# Introduction

Thème (3)

*Enjeux juridiques liés à l'accès et l'exploitation des archives du web.*

Cycle d'ateliers programmé par :

- Emmanuelle Bermès  
BnF, Direction des services et des réseaux, Adjointe chargée des questions scientifiques et techniques
- Madeleine Géroutet  
Université de Lille, SCD, Responsable du Département Services à la recherche et aux chercheurs

# Enjeux juridiques liés à l'accès et à l'exploitation des archives du web

Session coordonnée par Arnaud Laborderie (BnF) et Madeleine Géroudet (Université de Lille), avec la contribution de Marie Cros (Université de Lille) pour l'animation

- Cadre juridique actuel en France et en Europe
- Nature juridique de l'archive : document composite, œuvre de l'esprit, donnée de la recherche ?
- Diversité des points de vue sur ces enjeux juridiques : chercheur, bibliothécaire, juriste...
- Réutilisation des archives du web et diffusion des résultats de recherche

The screenshot shows the 'data.gouv.fr' website interface. At the top, there is a navigation bar with the BnF logo, a search bar, and a 'Recherche initiale' button. Below the navigation bar, the main content area displays the title 'LEGI: Codes, lois et règlements consolidés'. The page includes a description of the dataset, a 'Producteur' section with the French Republic logo, and an 'Informations' section with details about the license and update frequency. A 'Ressources' section at the bottom lists available data formats like XML and DTD.

# Programme de la séance

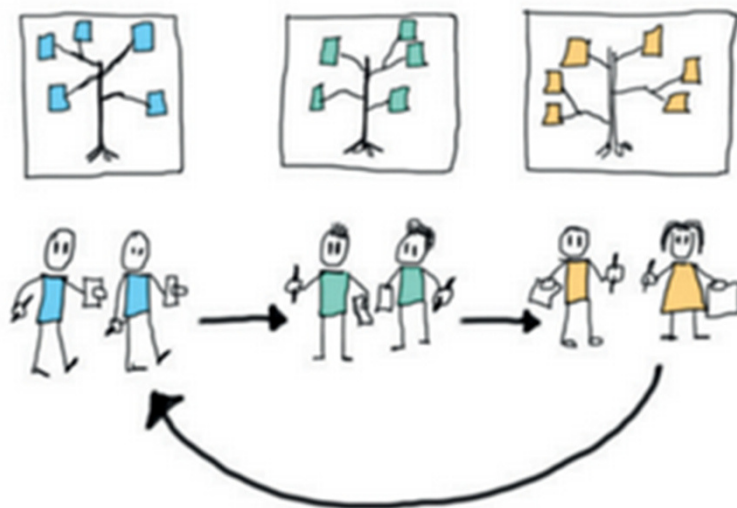
- **Harold Codant** (Adjoint au directeur du département des affaires juridiques et de la commande publique / Chef du service juridique, BnF) et **Marie-Amélie Conte-Bourges** (Service juridique, BnF)  
*Le cadre juridique des archives du web : un panorama en 2022*
- **Emmanuel Cartier** (Maître de conférences en linguistique informatique, LIPN, Université Paris 13)  
*Point de vue d'un enseignant chercheur sur les conditions juridiques pour l'accès et la réutilisation des archives du web*
- **Barbara Mutz** (Responsable du Département des affaires juridiques et réglementaires, INA)  
*Point de vue juridique sur l'accès et l'exploitation des archives du web préservées à l'INA*
- **Yves Maurer** (Responsable technique, Bibliothèque Nationale du Luxembourg)  
*Les conditions juridiques d'accès aux archives du web à la Bibliothèque nationale du Luxembourg*

# Cadre juridique des archives du web

- **Loi Dadvsi du 1 août 2006** relative au droit d'auteur et aux droits voisins dans la société de l'information
  - *« Sont également soumis au dépôt légal les signes, signaux, écrits, images, sons ou messages de toute nature faisant l'objet d'une communication au public par voie électronique. »*
- **Loi Darcos dite du 30 décembre 2021** visant à conforter l'économie du livre et à renforcer l'équité et la confiance entre ses acteurs (décret d'application en cours)
- **Les exceptions mises en œuvre dans le cadre de la recherche :**
  - **Exceptions au titre du dépôt légal** : exception au droit d'auteur, aux droits voisins et au droit des producteurs de bases de données
    - *« L'auteur ne peut interdire aux organismes dépositaires, pour l'application du présent titre :  
1° La consultation de l'œuvre sur place par des chercheurs dûment accrédités par chaque organisme dépositaire sur des postes individuels de consultation dont l'usage est exclusivement réservé à ces chercheurs »*
  - **Exceptions au titre du TDM** : exception au profit des organismes de recherche et des institutions du patrimoine culturel, exception au profit des fouilles réalisées par toute personne

# Atelier

Objectif : réfléchir collectivement aux défis juridiques que pose l'utilisation des archives du web et aux meilleures façons d'y répondre



Jardin des idées

1) Quelles recommandations peut-on formuler dans le contexte légal actuel ?

2) Quel serait le cadre juridique idéal pour les archives du web ?

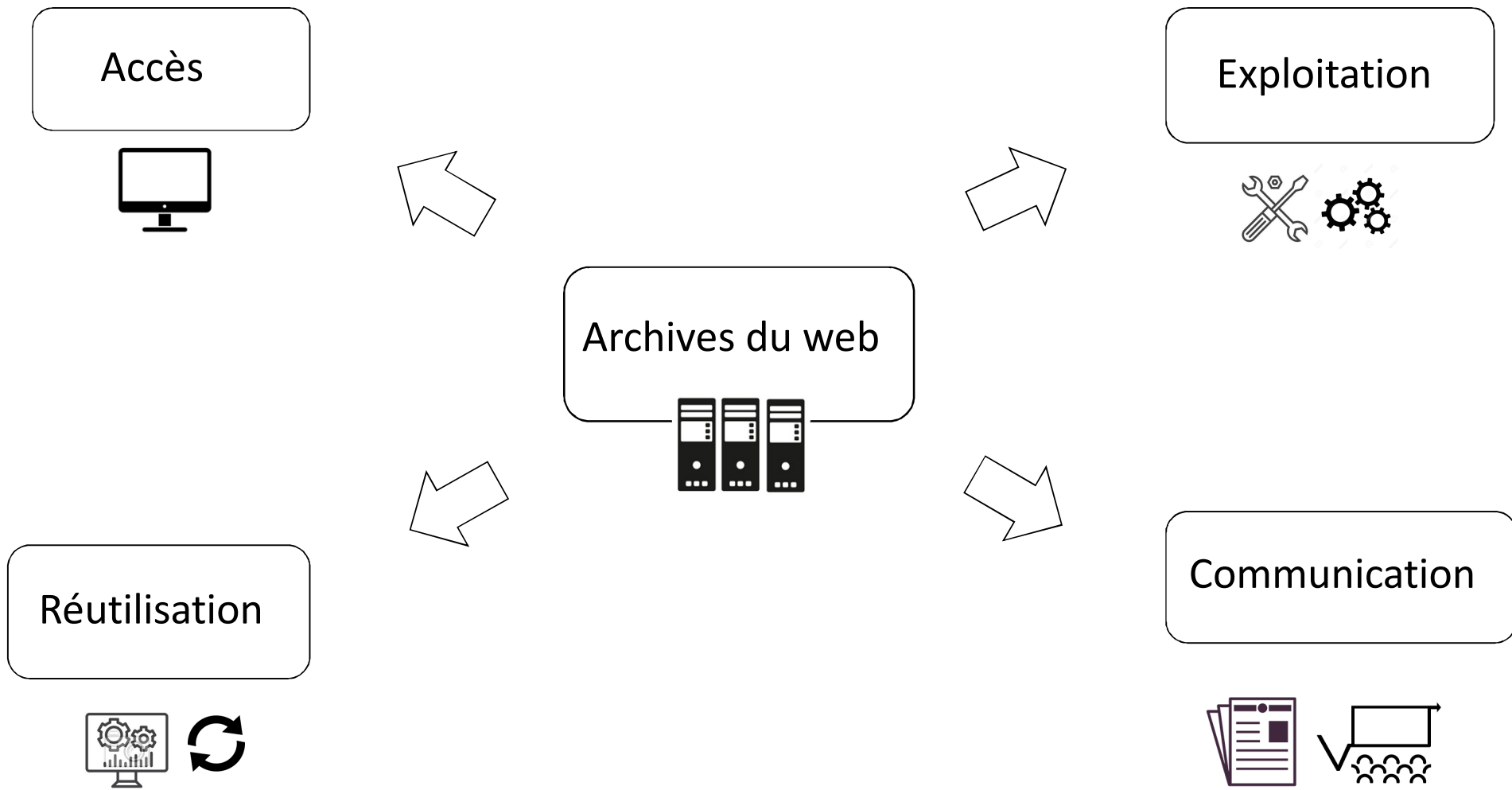
# Webinaire - Atelier 3

## Enjeux juridiques liés à l'accès et à l'exploitation des archives du web

Atelier collaboratif et participatif :

- 1/ Quelles recommandations faire dans le contexte légal actuel ?
- 2/ Quel serait le cadre juridique idéal pour les archives du web ?

© Login/Shutterstock, mise en page © Sandrine Lancereau/Sciences Po



Exception Dépôt légal  
(CP art. L.131-132)

Accès



- Accès libre (néanmoins sous droit)
- Des collectes dans le domaine fr
- Point d'entrée pour les chercheurs

▶ **Qui ?**

- **Accréditation** – tous chercheurs accrédités par les institutions depositaires

➤ Accès simplifié : – référentiel ?  
– conventionnement ?

▶ **Où ?**

- **Sur place** – à la **BnF** (tous sites)  
– à l'**INA** (tous sites)
- **Accès distant** – **INA** : consultation experte (7 lieux)  
/ consult. autonome (50 PCM chez part.)  
– **BDLI** : 21 sites en région (BM, etc.)  
– exp. capsule **Respadon** (BU Lille)

Capsules sécurisées

➤ Extension de la notion de « sur place » à des emprises sécurisées ?

➤ Refonder le statut de partenaires :  
– élargir la liste / penser par catégories ?

➤ Accès différenciés :  
– par chercheurs, projets, institutions ?  
– libre (gratuit) / intramuros (payant) ?

▶ **Quoi ?**

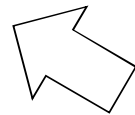
- **Sources** – accès complet sous conditions
- **Métadonnées** – accès libre (licence EtaLab)
- **Résultats** – données de la recherche  
➤ quelle licence ?

➤ Accès partiel libre : agrégation de données, prévisualisation d'extraits, embedding... ?

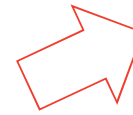
➤ Catalogue, documentation de la donnée collectée, données agrégées, API, etc.

➤ Quel accès aux résultats de la recherche ?

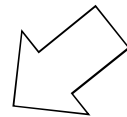
Accès



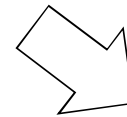
Archives du web



Exploitation



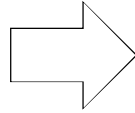
Réutilisation



Communication



## Exploitation



### Trois niveaux :

- données brutes (extraction des archives)
- données transformées (dérivées, agrégées, etc.)



Corpus intermédiaire

- quel statut ?
- conservation possible ?

- résultats de la recherche
  - sous quelle licence ?
  - intégration d'extraits, fragments, captures d'écran ?

## Exception TDM 1

Recherche (CPI art. L.122-5-3-II)



- pas d'autorisation et pas d'opposition possible de titulaires des droits
- fouille réalisée sur place par les chercheurs / par les institutions
- pas d'utilisation commerciale
- production libre de données dérivées
- respect du RGPD
- conservation possible des copies
- stockage sécurisé des données



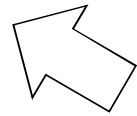
Capsules sécurisées

RES  
PA  
DON

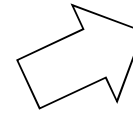
- convention
- accès à distance
- expérimentation d'outils
- corpus restreint

- traitement en masse des archives
- libéraliser les produits dérivés d'analyse des collectes ?
- service de copie à la demande ?
- autoriser la fouille à distance via un système une sorte de « boîte noire » sans accès direct aux sources ?
- fouille possible dans toute institution conservant les copies ?

Accès



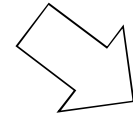
Archives du web



Exploitation



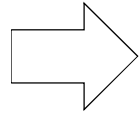
Réutilisation



Communication



## Réutilisation



## Citabilité



### Conservation des copies

- corpus des fouilles : données brutes + dérivées
  - (1) pérenniser l'extraction des archives ?
  - (2) conserver le corpus intermédiaire ?
- quid corpus sans fouille ?
  - usage restreint à copie privée
  - niveau minimum TDM ?
- entrepôt des données – à la fois sécurisé et accessible
  - BnF « tiers de confiance » / Huma-Num ?

- des URL par identifiant pérenne (ark)
- montrer les archives n'est pas abordé dans l'exception TDM

- recourir à d'autres exceptions :
  - droit de courte citation
  - exception pédagogique
- obtenir l'autorisation des ayants-droit

### Activité de recherche :

- transparence des sources
- reproductibilité des traitements



### données de la recherche

- description du corpus (méthodo, etc.)
- publication des permaliens

- (!) pas de cumul des exceptions
- expliciter la citation des archives

- s'ils sont inconnus : mention D.R. avec opt-out ?

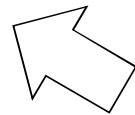
- comment vérifier les hypothèses de recherche, voire les réfuter ?

- accès au corpus pour le retravailler ?

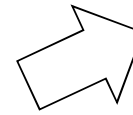
- accès aux résultats de la recherche ?

- ouverture des données ?

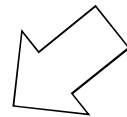
Accès



Archives du web



Exploitation



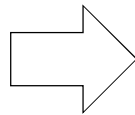
Réutilisation



Communication



## Communication



## Diffusion



## Médiation

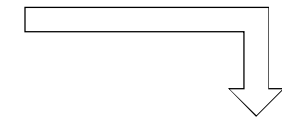
- Comment montrer des captures et extraits des archives du web pour communiquer sur les archives du web ?
- Formation des bibliothécaires et des chercheurs
  - un démonstrateur des archives du web
  - valorisation de la recherche : que montrer ?

- pas de dispositions spécifiques pour les archives du web
  - obtenir l'autorisation des titulaires de droits (voire redevance)
  - **opt-out si ayants-droit inconnus**
- pas de restrictions de re-diffusion pour les contenus en creative commons
- respect du RGPD et de la légalité des contenus

Pour **limiter les risques juridiques**, respecter

trois critères :

- se limiter à certains cas spéciaux,
- ne pas porter atteinte à l'exploitation normale de l'œuvre ou de l'objet protégé,
- ne pas causer de préjudice injustifié aux intérêts légitimes des titulaires de droits.



## Bonnes pratiques

- **guide pour les chercheurs**
  - conditions de consultation, d'exploitation et de diffusion dans un cadre de recherche
  - conditions de dépôt des données ouvertes



## Cycle d'ateliers – 2021-2023

PUBLIÉ 27/09/2021 - MIS À JOUR 28/09/2021



The Data Lifecycle ©Mushonz via Wikimedia Commons [https://commons.wikimedia.org/wiki/File:The\\_Data\\_Lifecycle.jpg](https://commons.wikimedia.org/wiki/File:The_Data_Lifecycle.jpg)

Le Réseau de Partenaires pour l'exploration et l'analyse de données numériques (ResPaDon) propose un cycle d'ateliers trimestriels.

### Objectifs

Il se fixe pour objectif d'échanger et de consolider les savoir-faire et l'expertise des professionnels de l'information scientifique et technique (IST) autour de l'exploitation des archives du dépôt légal du web.

Pendant deux années, les rendez-vous trimestriels du groupe de réflexion ouvrent un espace de dialogue et offrent aux participants l'opportunité de mettre en évidence des pratiques et des savoir-faire professionnels afin de consolider leur expertise et de mutualiser leurs expériences pour favoriser une montée en compétence des professionnels et in fine, encourager et faciliter l'usage des archives web et des données par les chercheurs.

A terme, les travaux du groupe de réflexion donneront lieu à la formalisation de préconisations qui recouvriront les différents enjeux du [projet ResPadon](#).

## Prochain RDV – Atelier (5)

5<sup>ème</sup> rencontre trimestrielle du Groupe de réflexion

Coordonnée par Laurence Favier et Emmanuelle Bermès

### Thème

« Des usages aux services. A partir de l'enquête réalisée sur les usages, quelle offre de services définir ? »

Date : 12 Juillet 2022

Webinaire : 12 septembre, 17:00

Inscriptions : [https://univ-lille-fr.zoom.us/webinar/register/WN\\_UIhAEjOxT6qGMy0HNuE4FA](https://univ-lille-fr.zoom.us/webinar/register/WN_UIhAEjOxT6qGMy0HNuE4FA)

naires pour l'analyse et l'exploration de données numériques – Cycle d'Ateliers

RES  
PA  
DON

## Merci de votre attention

Madeleine Géroutet

Université de Lille, SCD, Responsable du Département Services à la recherche et aux chercheurs

[marie-madeleine.geroutet@univ-lille.fr](mailto:marie-madeleine.geroutet@univ-lille.fr)

Arnaud Laborderie

BnF, département de la Coopération, service de Coopération numérique et de Gallica  
Chef de projet Exploitation des données pour la recherche

[arnaud.laborderie@bnf.fr](mailto:arnaud.laborderie@bnf.fr)

© Login/Shutterstock, mise en page © Sandrine Lancereau/Sciences Po