

Atelier SoDUCo-BnF - 4<sup>e</sup> Séance

RES(T)ITUER  
les  
ANNUAIRES COMMERCIAUX  
et les  
ÉVOLUTIONS  
de  
L'ESPACE PARISIEN  
du  
XIX<sup>e</sup> siècle

Partager l'expérience et les ressources du  
PROGRAMME ANR SoDUCo

6 et 7 Novembre 2023

Bibliothèque nationale de France, Site François Mitterrand, Paris.



## 4<sup>e</sup> Séance de l'atelier SoDUCo-BnF

# Res(t)ituer les annuaires commerciaux et les évolutions de l'espace Parisien du XIX<sup>e</sup> siècle

Partager l'expérience et les ressources du programme ANR SoDUCo

Séminaire de clôture du programme ANR SoDUCo

(<https://soduco.github.io/>)

Lundi 6 (09h30-17h30) & Mardi 7 (09h30-12h30) Novembre 2023

Bibliothèque nationale de France, Site François Mitterrand, Paris.

La 4<sup>ème</sup> séance de l'Atelier SoDUCo-BnF propose de poursuivre les discussions engagées entre chercheurs et spécialistes de la conservation patrimoniale des documents. Pour rappel, les 3 précédentes séances concernaient : les procédures de conservation, de signalement et de diffusion des documents cartographiques (*Le circuit de traitements des documents cartographiques des XVIII<sup>e</sup> et XIX<sup>e</sup> siècles à la BnF*, 23 mai 2023), les moyens de reconstituer l'histoire des atlas et plans de la ville et d'en proposer une édition critique numérique (*Paris, capitale géodésique. Atlas et plans à grande échelle: 1780 – 1950*, 10 avril 2022), la définition et l'exploitation du corpus des listes d'adresses des *Annuaire commerciaux du XIX<sup>e</sup> siècle* (*Res(t)ituer les adresses des almanachs et annuaires commerciaux parisiens du XIX<sup>e</sup> siècle. Un corpus de localisations urbaines à grande échelle*, 10 novembre 2022), et enfin un aperçu des analyses qui peuvent être mises en œuvre sur les données qui en sont extraites (*Analyses spatiales formalisées de données géo-historiques*, 23 mai 2023).

Cette dernière séance de l'Atelier est organisée sur deux jours, les 6 et 7 novembre 2023. Elle fait office de séminaire de clôture du programme de recherche SoDUCo (*Dynamiques Sociales en contexte urbain: outils, modèles et données libres – Paris et ses banlieues, 1789-1950*). Elle s'attache à présenter un bref bilan de l'ensemble des réalisations, ainsi que les outils et données mis à la disposition du public.

Lors de son montage, nous avons souhaité inscrire le projet SoDUCo dans une démarche de *Science ouverte*. Nous souhaitons discuter des implications de cette décision sur la nature, l'organisation et les choix techniques du projet. Une telle démarche est encore inégalement partagée et pratiquée dans la communauté scientifique et les travaux et ressources qui en relèvent prennent des formes très diverses, notamment dans le domaine des SHS. Appuyée sur le développement des technologies numériques et les collaborations interdisciplinaires, elle est désormais fortement encouragée et encadrée par les institutions et tutelles de la recherche notamment via la mise en place d'un écosystème sociotechnique dédié (infrastructures et outils, principes et méthodologies, spécialisations professionnelles, ...).

Il s'agit dans un premier temps de situer les réalisations du programme de recherche SoDUCo dans ce contexte. Le parti pris assumé de la reproductibilité, de l'ouverture des données et des outils a orienté la manière d'aborder l'étude de la *co-évolution sociale et spatiale de Paris au XIX<sup>e</sup>*. Il a aussi présidé à de nombreux choix tant techniques que méthodologiques. L'ensemble de ces principes et contraintes a sans aucun doute contribué à façonner procédures, résultats et productions. Il nous semble utile de faire le point sur ces aspects et de confronter l'expérience et les questionnements du consortium SoDUCo avec celles des acteurs et experts de différentes communautés et institutions intéressées par ces questions (*conservation patrimoniale, ingénierie des connaissances, analyse des documents, humanités numériques, recherche en sciences sociales, science participative*).

Peut-on mesurer les effets d'une telle démarche sur nos objets et problématiques de recherche ? Quel est le statut des données numériques produites de manière automatique ? Comment et à quelles conditions est-il possible de les utiliser ? Comment les évaluer et améliorer les procédures de traitement ? Comment ré-utiliser les outils dans d'autres contextes et avec d'autres sources ? Peut-on parvenir à cumuler résultats et travaux issus de l'exploitation des corpus du projet ? Comment mettre à jour, enrichir ou encore pérenniser de tels objets et outils numériques ? Comment interagir avec l'écosystème actuel de la conservation patrimoniale et de la recherche et anticiper ses évolutions ?

# 4<sup>e</sup> Séance de l'atelier SoDUCo-BnF

## Res(t)ituer les annuaires commerciaux et les évolutions de l'espace Parisien du XIX<sup>e</sup> siècle.

Programme prévisionnel

Lundi 6 novembre 2023, 9h30-12h15 en salle 70

\*\*\* Matinée, accueil à partir de 9h00 \*\*\*

09h30. Ouverture de la journée : **SoDUCo team**, *Le programme de recherche SoDUCo. Bilan*

### SESSION METADONNEES, CATALOGAGE & DIFFUSION DE DONNEES

Cette demi-journée revient sur les questions du catalogage, de la diffusion des données et des métadonnées pour les projets de recherche en humanités numériques impliqués dans une démarche de *science ouverte*. Nous présenterons le géo-catalogue développé dans le cadre de SoDUCo pour élaborer à partir de sources éparées des corpus de documents adaptés aux questions de recherche, puis conserver la trace des divers traitements qui leur sont appliqués ainsi que les états intermédiaires et finaux des données numériques qui en sont extraites. Nous discuterons notamment les notions de collections et de corpus, d'interopérabilité, de pérennité, de même que nous poursuivrons les réflexions engagées lors des précédentes journées sur rapports entre les catalogages patrimoniaux inter-institutionnels et le catalogage orienté vers l'usage de la recherche. Sur quelles bases chercheurs et spécialistes de la documentation peuvent-il collaborer ? Comment pouvons-nous nous organiser pour utiliser les acquis de SoDUCo et de l'atelier SoDUCo- BnF et créer les conditions d'une cumulativité à la fois du point de vue de la recherche et du point de vue des institutions patrimoniales. Quels principes, quels outils et quel modèle de travail pouvons-nous adopter pour aller plus loin ensemble sur d'autres sujets/objets et avec d'autres acteurs/publics.

Présidence :

10h00. **SoDUCo team**, *Catalogage et diffusion des données dans le programme*.

10h30. Questions

10h45. **J. Schoonman, B. Spaan & M. Claeys Bouuaert** ([Allmaps](#)), *Présentation du consortium Allmaps*

11h05. **M. Denel et al.** (Wikimedia France), *Science participative, science ouverte. L'expérience de Wikimedia France*

11h25. **Adeline Joffres** (Huma-num), *Sous réserve*

11h45. Questions

12h00. Discussion – Table ronde

12h15. Pause-déjeuner.

# 4<sup>e</sup> Séance de l'atelier SoDUCo-BnF

## Res(t)ituer les annuaires commerciaux et les évolutions de l'espace Parisien du XIX<sup>e</sup> siècle.

Programme prévisionnel (suite)

Lundi 6 novembre 2023, 14h00-17h30 en salle 70

\*\*\* Après-midi, accueil à partir de 13h45 \*\*\*

### SESSION EXTRACTION & STRUCTURATION DE DONNEES.

Cette demi-journée revient sur les questions d'analyse automatique de documents anciens semi-structurés, de la détection de leur mise en page à la structuration spatio-temporelle des informations qu'ils contiennent. Nous y présenterons les travaux réalisés dans le projet SoDUCo sur les annuaires du commerce parisien du XIX<sup>e</sup> siècle pour permettre le suivi spatio-temporel de chaque commerce recensé dans ces annuaires. Sur la base de ces expériences, il nous semble utile d'élargir les discussions sur deux aspects ouvrant encore de nombreuses perspectives de recherche : l'analyse automatique de documents anciens et la représentation de connaissances géohistoriques.

Présidence :

14h00. **SoDUCo team**, *Extraction et structuration de données dans le programme SoDUCo*.

14h30. **C. Duvette et P. Kervegan** (*Département des études et de la recherche - Institut national d'histoire de l'art, INHA*), *Chapeliers, architectes, nouveautés : géo-analyse des activités visibles et invisibles d'un quartier*.

14h50. Questions.

15h05. Pause - café

15h20. **N. Hernandez, W. Charles** (*Institut de Recherche en Informatique de Toulouse – IRIT, Université de Toulouse Jean Jaurès*) et **S. Poublanc** (*France, Amérique, Espagne, Sociétés, Pouvoirs, Acteurs – Framespa, Université de Toulouse Jean Jaurès*), *Représenter les territoires de la France moderne dans des graphes de connaissances : le projet OBARDI*

15h40. **C. Kermorvant** ([TEKLIA](#)), *Nouvelles approches pour l'extraction d'informations dans des documents numérisés*.

16h00. Questions.

16h15. Discussion – Table ronde

16h30. Pause - café

### KEYNOTE.

16h45. **M. Barthelemy** (*Institut de Physique Théorique, Commissariat à l'Energie Atomique et aux Energies Alternatives*) et **J. Gravier** (*Centre de recherches historiques - CRH et Centre d'analyse et de mathématique sociales - CAMS, Ecole des hautes études en sciences sociales*)

17h30. Clôture de la journée.

# 4<sup>e</sup> Séance de l'atelier SoDUCo-BnF

## Res(t)ituer les annuaires commerciaux et les évolutions de l'espace Parisien du XIX<sup>e</sup> siècle.

Programme prévisionnel (suite)

Mardi 7 novembre 2023, 09h30-12h15 en salle 70

\*\*\* Matinée, accueil à partir de 9h30 \*\*\*

### DONNEES HISTORIQUES NUMERIQUES EVALUATIONS CRITIQUES ET PERSPECTIVES D'UTILISATIONS

La matinée « Données historiques numériques » vise à confronter l'expérience de chercheurs qui construisent et analysent régulièrement des corpus de données numériques élaborés à partir de documents d'archives (texte et cartes). Dans un moment où l'on assiste à l'explosion de la production de données numériques ouvertes issues de sources anciennes, les questions classiques des conditions d'appropriation et de réutilisation de ces nouvelles « sources » par des tiers se posent de manière accrue.

Cette matinée invite plus particulièrement à réfléchir à ces questions dans le cas où les corpus de gros volumes de données numériques sont produits à partir de processus entièrement automatisés d'extraction de documents et d'enrichissement. Partant des travaux réalisés dans le cadre du programme SoDUCo, notamment du corpus des *Annuaires historiques parisiens* et de la *géolocalisation historique* des entrées d'annuaires sur les plans anciens, nous souhaitons aborder les questions suivantes au cours de la matinée : Comment et à quelles conditions peut-on s'emparer efficacement de telles « sources numériques » ? Comment les comprendre, évaluer leur « qualité » et leur rapport avec les sources initiales ? Comment comparer ou rendre cumulable les travaux très divers qui peuvent être issus de l'exploitation de ces jeux de données ? Enfin, nous souhaitons discuter collectivement les perspectives d'analyse et d'exploitation du corpus des *Annuaires historiques parisiens*.

Présidence :

- 10h00. **SoDUCo team**, *Les données géohistoriques produites au cours du programme SoDUCo. Evaluation d'une « source numérique »*.
- 10h30. **L. Kesztenbaum, G. Postel-Vinay** (*Institut national d'études démographiques, INED*), *Construction de bases de données historiques: retours d'expérience*.
- 10h50. Questions
- 11h05. **C. Lemerrier** (CNRS, *Centre de sociologie des organisations - Sciences Po-Paris*), *Histoire quantitative, données ouvertes en SHS*
- 11h25. [intervention 3 ??]
- 11h45. Questions
- 12h00. Discussion – Table ronde
- 12h15. Déjeuner.



# Dynamiques sociales en contexte urbain : outils, modèles et données libres Paris et ses banlieues, 1789-1950

Le programme de recherche *Social dynamics in urban context* (SoDUCo, <https://soduco.github.io/>) est financé par l'agence Nationale de la Recherche (ANR). Il réunit des chercheurs et ingénieurs du groupe *Geohistoricaldata* membres de l'Ecole des hautes études en sciences sociales (EHESS, LaDéHis-CRH-CAMS), du Laboratoire des Sciences et Technologies de l'Information Géographique (LaSTIG) de l'Institut national de l'information géographique et forestière (IGN), du laboratoire de recherche et développement (LRDE) de l'École Pour l'Informatique et les Techniques Avancées (EPITA), et le Centre de topographie historique de Paris (CTHP) des Archives Nationales.

Voir aussi : progr. [SoDUCo](#), Groupe [Géohistoricaldata](#) (Site et Wiki) et les dépôts associés sur [Github](#), [Harvard Dataverse](#), [Didomena](#), [Nakala](#) et [Zenodo](#)



{BnF

Séminaire BnF/SoDUCo



En 2022-2023, l'équipe SoDUCo et la BnF organisent un séminaire commun, l'Atelier SoDUCo-BnF. Il a pour perspective de diffuser et mettre en question le travail réalisé dans le cadre de SoDUCo et de faire dialoguer chercheurs et spécialistes de la conservation des documents sur des objets et projets communs allant de l'exploitation des fonds (cartes et plans, annuaires du commerce) à la production et diffusion de données libres (géocodage historique, analyse spatiale, science ouverte). Informations et programmes détaillés : [https://soduco.github.io/soduco\\_bnf\\_seminars/](https://soduco.github.io/soduco_bnf_seminars/)

Equipe organisatrice : A. Aniesa, P. Chevallier, J. Garel-Grislin, C. Ion, O. Loiseaux, J.-D. Mellot, J.-P. Moreux, E. Netchine (BnF), N. Abadie, C. Bernard, M. Hersent, J. Perret, J. Raimbault, S. Tual (IGN), E. Carlinet, J. Chazalon, M. Puren (EPITA), S. Baciocchi, M. Barthelemy, P. Cristofoli, B. Dumenieu, J. Gravier, M. Fernandez (EHESS), V. Rebolledo-Dhuin (UPEC) et P.A Le Ny (LeNy Conseil).

- 1<sup>ère</sup> Journée : *Paris, capitale géodésique. Atlas et plans à grande échelle: 1780 - 1950*. Lundi 11 avril 2022, 1<sup>ère</sup> Journée de l'Atelier SoDUCo-BnF, Paris, Bibliothèque nationale de France, Belvédère du site François-Mitterrand.
- 2<sup>ème</sup> Journée : *Res(t)ituer les adresses des almanachs et annuaires commerciaux parisiens du XIX<sup>e</sup> siècle. Un corpus de localisations urbaines à grande échelle*, Jeudi 10 novembre 2022, 2<sup>ème</sup> Journée de l'Atelier SoDUCo-BnF, Paris, Bibliothèque nationale de France, Site François-Mitterrand, en salle 70.
- 3<sup>ème</sup> Journée : *Le circuit de traitements des documents cartographiques des XVIII<sup>e</sup> et XIX<sup>e</sup> siècles à la BnF - Analyses spatiales formalisées de données géo-historiques*, 3<sup>ème</sup> Journée de l'Atelier SoDUCo-BnF, Mardi 23 Mai 2023, 9h30-17h30, BnF Site Richelieu, 5 Rue Vivienne, 75002 Paris.
- 4<sup>ème</sup> Séance : *Res(t)ituer les annuaires commerciaux et les évolutions de l'espace Parisien du XIX<sup>e</sup> siècle. Partager l'expérience et les ressources du programme ANR SoDUCo*, 4<sup>ème</sup> séance de l'Atelier SoDUCo-BnF, Lundi 6 et Mardi 7 novembre 2023, Paris, Bibliothèque nationale de France, Site François-Mitterrand, en salle 70

La 4<sup>e</sup> séance de l'atelier SoDUCo-BnF est ouverte au public.

Elle se déroule à la BnF, Site François-Mitterrand,

Lundi 6 novembre (09h30-12h30 et 14h00-17h30)

Et Mardi 7 novembre 2023 (10h00-13h30)

en **salle 70** (à côté du **petit auditorium**),  
accès par le **hall Est**.

**Inscription obligatoire aux trois demi-journées**  
**via le lien suivant :**

<https://evento.renater.fr/survey/inscriptions-a-la-4eme-seance-soduco-bnf-6-et-7-novembre-2023-0j6blqtl>

Les sessions seront aussi diffusées en visio-conférence avec accès sur inscription.

Organisation et informations :

P. Cristofoli ([pascal.cristofoli@ehess.fr](mailto:pascal.cristofoli@ehess.fr)) & P. Chevallier ([philippe.chevallier@bnf.fr](mailto:philippe.chevallier@bnf.fr))

