

UNIVERSITÉ PARIS 1 - PANTHÉON SORBONNE  
ÉCOLE D'HISTOIRE DE LA SORBONNE



## Formations du PIREH (2023-2024)

Master 1 et 2 – Doctorat

**Enseignants** : J. Alerini, J.-C. Balois, P. Brochard, A.-S. Bruno,  
G. Bonnot, L. Dumont, O. Julien, S. Lamassé, H. Noizet

## ORGANISATION GÉNÉRALE

Pour toute information complémentaire : <https://recherche-pireh.pantheonsorbonne.fr/pireh>

Les formations en Histoire et informatique du Master recherche visent à permettre aux étudiants d'aborder leurs sources sous un angle nouveau grâce à une maîtrise suffisamment approfondie de certains outils informatiques. Le détour par l'ordinateur conduit en effet à repenser ses sources et l'application de méthodes sérielles se révèle bien souvent très fructueuse sur le plan scientifique.

Au-delà de leur utilité pour les travaux de master, ces formations permettent d'enrichir sa culture scientifique et technique et offrent des bases solides pour celles et ceux qui souhaitent s'orienter vers d'autres domaines du monde professionnel (édition, patrimoine, gestion, journalisme, économie du web, *data mining*, etc.).

Les enseignements se composent et s'organisent autour :

- de séminaires (facultatifs en fonction du Master d'inscription),
- de TD de formation générale, **obligatoire pour la plupart des étudiants de Master recherche**,
- de diverses formations de spécialité (facultatifs en fonction du Master d'inscription),
- d'un parcours de Master, intitulé « Sciences des données et Histoire » (SDH).

Des stages sont également proposés aux doctorant·e·s.

### Quel parcours adopter ?

Si le parcours de Master « Sciences des données et Histoire » vous intéresse, il est indispensable, dès la licence 3 ou à la rentrée, de contacter les enseignants du PIREH.

Pour les formations dites générales et spécialisées, la démarche la plus logique consiste à suivre la formation générale en M1 (quelle que soit la période étudiée) dans le cadre du bloc « Méthodologie : Outils de la recherche », pour décider en connaissance de cause s'il convient ou non d'avoir recours aux méthodes informatiques, de manière plus poussée, pour le travail accompli dans le cadre du M1 (si la formation spécialisée est au second semestre) et du M2. La validation de la formation générale est un préalable indispensable pour suivre une des formations de spécialité.

Ceux qui n'auraient pas suivi la formation générale en M1 la commenceront donc en année de M2. Par ailleurs, les étudiants ayant bénéficié de la formation « Histoire et numérique » au niveau L3 sont dispensés de la formation générale mais doivent suivre une formation spécialisée en M1 ou M2.

Les étudiants de master se destinant à un master en alternance ayant une dimension informatique, comme le master MIMO, sont invités à valider au moins trois formations spécialisées.

#### Synthèse (pour tous les étudiants de Master, hormis les étudiants du parcours SDH) :

- suivre un TD de formation générale est obligatoire en M1 ou en M2, en fonction des maquettes des différents Master,
- cette formation obligatoire donne accès à des enseignements, facultatifs, de spécialité (ouverts aux doctorants dans la mesure des places disponibles).

## Séminaires

Deux séminaires sont ouverts aux étudiant·e·s intéressé·e·s par l'usage des méthodes informatiques et quantitatives en histoire :

### Séminaire de recherche du parcours Sciences des données et histoire

Le séminaire, dirigé par Stéphane Lamassé et Léo Dumont, a lieu au premier semestre le vendredi, de 10h à 12h, en salle Perroy. Il vise à explorer et penser le lien entre informatique et histoire à travers la présentation de travaux de chercheurs et chercheuses confirmé·e·s ou débutant·e·s (en master ou doctorat) et la discussion collective autour de lectures ou de projets actuels.

Programme sur <https://recherche-pireh.pantheonsorbonne.fr/recherches/seminaire>.

### Construction et catégorisation de données historiques

L'objectif du séminaire, dirigé par Anne-Sophie Bruno, est de montrer les étapes de la construction d'un corpus de données susceptible de se prêter à un traitement statistique ainsi que les problèmes spécifiques posés par la catégorisation de données temporelles complexes. Le séminaire mêlera considérations théoriques, présentations de travaux d'historiens et cas appliqués à partir des données des participants aux séminaires. Le lundi de 10h30 à 12h, les 29 janvier, 5 février, 26 février, 4 mars, 11 mars, 18 mars, 25 mars, 22 avril, et 29 avril. La salle sera annoncée d'ici la rentrée du second semestre.

### La formation générale

Il s'agit d'initier les étudiants à un ensemble de méthodes susceptibles d'être mobilisées dans le cadre de leurs recherches, tout en favorisant une réflexion critique sur l'usage de l'informatique dans la construction du discours historique. Elle prend la forme de TD (**obligatoires**) au 1<sup>er</sup> semestre (pour un volume de 13h – voir ci-dessous les horaires des TD) et est validée par une épreuve écrite et/ou un travail de constitution d'une base de données ou d'un corpus textométrique à partir des sources historiques de l'étudiant·e·s. Pour approfondir ce travail en vue de la réalisation de leur mémoire, les étudiant·e·s peuvent bénéficier d'un suivi individuel assuré par les enseignants du PIREH (sous réserve qu'ils aient bien suivi la formation durant l'année universitaire en cours).

**Cette formation, dite générale, est obligatoire et peut être suivie en M1 ou en M2 (voir ci-dessus « Quel parcours adopter ? », p. 1).** L'inscription à un TD se fait via l'EPI (cf. p. 7).

### Les formations spécialisées

Elles sont destinées à des étudiants ayant déjà validé la formation générale, en L3 ou en M1. Elles permettent d'approfondir des compétences informatiques acquises en formation générale (bases de données, lexicométrie) ou d'en acquérir de nouvelles (XML, programmation, analyse statistique, analyse réseau, cartographie – voir les descriptifs ci-dessous).

L'inscription à une formation spécialisée se fait via l'EPI (cf. p. 7).

### Le parcours « Sciences des données et Histoire »

Le parcours de Master SDH permet de développer des compétences statistiques et numériques venant s'ajouter aux qualités spécifiques qu'offre une formation de sciences humaines et sociales, à savoir une capacité à interroger les sociétés humaines dans le temps et l'espace, en mettant en oeuvre des méthodes de réflexion, d'analyse et de rédaction adaptées. L'ensemble de ces compétences est désormais indispensable pour s'insérer dans la vie professionnelle à des postes d'encadrement intermédiaire ou supérieur.

La familiarisation durant la durée du parcours avec le raisonnement statistique et avec les outils numériques vous aidera à construire votre objet de recherche et votre corpus de sources, à expliciter et à formuler vos hypothèses, puis à proposer une méthode rigoureuse pour traiter l'ensemble des matériaux. Ce parcours doit permettre de développer les capacités de structuration et d'analyse de « données » historiques. Il permettra d'acquérir des savoir-faire dans des domaines variés comme les statistiques appliquées, l'analyse de texte, les systèmes d'information géographiques, l'analyse de réseaux sociaux ainsi que les bases de données.

L'objectif est de former des chercheuses et chercheurs en histoire à des méthodes quantitatives et informatiques. Il se veut ambitieux et s'attache à donner une véritable autonomie technique, historiographique et épistémologique. Les étudiants devront :

- choisir un sujet de recherche dans un autre des parcours de l'École d'histoire de la Sorbonne. La réussite au mémoire conditionne la réussite au parcours. Ils seront donc encadrés par l'enseignant référent de leur choix,
- suivre une charge de cours conséquente (entre 250 et 300 heures sur deux ans).

Pour mener à bien des recherches en histoire, dans le cadre de ce parcours, il est nécessaire :

- d'être attiré par la recherche,
- de ne pas être effrayé par les sciences numériques,
- d'être capable de s'investir dans une formation exigeante,
- d'être titulaire d'une licence.

## **Stages pour doctorants**

Le PIREH et l'École doctorale d'histoire organisent des stages de quelques jours ou d'une semaine destinés aux doctorant·e·s de l'École d'histoire de la Sorbonne. Selon les années, ces stages portent sur la cartographie, la textométrie, les bases de données, l'analyse de réseaux sociaux, etc. S'adresser à l'École doctorale pour plus d'informations.

## DESCRIPTIF DES ENSEIGNEMENTS

### Formation générale

<b>Travaux dirigés de formation générale</b>	Premier semestre 7 séances de 2h
<p>La formation générale est une initiation destinée aux étudiants n'ayant pas suivi la formation « Histoire et numérique » en L3. Pour les étudiants de Master 1, elle permet de structurer leurs données dès le début de leur recherche sans attendre le Master 2. Pour les étudiants de Master 2, elle est obligatoire s'ils ne l'ont pas suivie en M1. Dans le cadre de la formation générale, les étudiants apprendront à constituer et à exploiter, à partir de sources historiques, une base de données ou un corpus lexicométrique (corpus de textes structuré en vue d'analyses statistiques).</p> <p><b>Les TD ne sont pas organisés par période.</b></p> <p>Pour valider cette UE, les étudiants doivent passer une épreuve écrite et/ou rendre un dossier de recherche mobilisant une de ces méthodes pour l'analyse de leurs sources. Si les étudiants n'ont pas accès à leur documentation dans le cadre du M1, l'enseignant pourra leur proposer un petit corpus de sources à partir duquel réaliser un dossier. Les étudiants pourront au second semestre bénéficier d'un suivi individuel pour approfondir leur travail en vue de la réalisation de leur mémoire.</p>	

### Formations spécialisées

Ces formations spécialisées sont destinées aux M2 ayant déjà suivi la formation générale en histoire et informatique, aux étudiants de M1 qui ont suivi l'enseignement « Histoire et numérique » en L3 et, dans la limite des places disponibles, aux doctorants. Les étudiants envisageant de postuler à un master professionnel orienté informatique (MIMO, Pise...) doivent suivre au moins 3 de ces différentes formations spécialisées.

<b>Modélisation et bases de données</b>	Premier et second semestres 2h hebdomadaires
Julien Alerini (S1) et Pierre Brochard (S2)	
<p>Cette formation vise à approfondir les méthodes de formalisation et de modélisation de l'information historique sous forme de bases de données. La numérisation des informations historiques et des sources n'est pas une opération neutre. Elle répond à des choix scientifiques d'analyse ou d'usage des corpus numérisés. À partir de sources distribuées et de celles des étudiants, on étudiera les principes et techniques de modélisation d'une base de données relationnelle (Merise), de son implémentation et de son exploitation (SQL). Nous verrons également comment d'autres approches des bases des données (bases hiérarchiques en XML, bases NoSQL orientées réseaux et documents) permettent des approches différentes de la structuration des données dans la perspective de la recherche historique. Le second semestre sera notamment consacré à la question du web sémantique et à la manière d'utiliser ses technologies avec des données historiques.</p>	

<b>Programmation et exploitation du Web pour l'histoire avec Python</b>	Premier semestre 2h hebdomadaires
Octave Julien	
<p>Savoir programmer peut être un grand atout pour l'historien(ne) : la programmation permet de mettre la puissance de l'ordinateur au service de la recherche et de développer ses propres méthodologies en automatisant le traitement des données, qu'elles soient textuelles, chiffrées, iconographiques, ou encore de faciliter leur récupération sur le Web.</p> <p>Conçu comme une initiation destinée à des débutants, ce cours se concentrera d'abord sur la maîtrise de Python, un langage de programmation moderne, puissant, et facilement utilisable, qui peut être immédiatement utile au travail de Master. On verra ensuite comment associer ce langage aux technologies employées sur le Web pour faciliter la récupération ou la diffusion de données en lignes.</p>	

<b>Analyse de réseaux sociaux</b> Octave Julien	Premier semestre 7 séances de 2h
L'analyse de réseaux en histoire, dans la veine des réseaux sociaux en sociologie, permet d'étudier les structures qui émergent lorsqu'on met en relation des individus ou des choses. Dans le cadre de cette formation, on verra comment constituer, visualiser et analyser un réseau à partir de données historiques. On s'appuiera pour cela sur les fonctionnalités du logiciel Gephi (représentations graphiques, indicateurs statistiques, outils de détection des sous-groupes d'un réseau, etc.) et sur des lectures illustrant les usages des réseaux en histoire.	
<b>Analyse et visualisation de données avec R</b> Gaëtan Bonnot	Premier semestre 2h hebdomadaires
Cette formation vise à exploiter une base de données ou un corpus lexicométrique déjà constitué à travers différentes méthodes, ainsi que des données issues du web. Celle-ci s'attachera à enseigner les méthodes d'analyse de données permettant d'aller plus loin dans l'exploration d'un gisement de données historiques (analyses factorielles, classifications, tests statistiques, régressions, etc.). Après quelques rappels de statistiques descriptives et inférentielles, cette approche sera croisée avec des outils de visualisation des données (analyses réseaux et cartographie). Les étudiants se familiariseront ainsi avec le logiciel d'analyses statistiques <i>open source</i> R.	
<b>Édition de sources (L<sup>A</sup>T<sub>E</sub>Xet XML-TEI)</b> Gaëtan Bonnot	Second semestre 2h hebdomadaires
L'édition de documents historiques peut faire partie intégrante du métier de l'historien·ne. Or, il y a une différence entre la manière dont on peut considérer les ressources textuelles en histoire et la façon dont la plupart des logiciels de traitement de texte, couramment utilisés et avec une maîtrise variable, sont élaborés. Ces systèmes réduisant l'écart entre ce qui est affiché à l'écran et le « résultat final » ( <i>What you see is what you get</i> ) présentent un certain nombre de contraintes et limites : structuration et mise en forme du document, gestion des appareils critique et savant, de la bibliographie, interopérabilité avec des analyses informativantes (base de données relationnelles, lexicométrie, cartographie, etc.), intégration dans une communauté scientifique, etc. Cet enseignement a pour objectif de former les étudiants à des usages de L <sup>A</sup> T <sub>E</sub> X et de la <i>Text Encoding Initiative</i> (TEI), utilisés depuis plus de trente années, en lien étroit avec l'édition critique de sources.	
<b>Cartographie et SIG</b> Hélène Noizet	Second semestre 7 séances de 3h
<p>Cette formation vise à doter les participant·es des connaissances et des compétences nécessaires pour aborder les enjeux spatiaux et cartographiques d'une recherche historique. Dans cette perspective, seront successivement abordées les différentes étapes allant de l'utilisation des données historiques, à leur traitement et leur analyse avec les méthodes géomatiques, pour finalement envisager la création et l'édition concrète de cartes. On s'appuiera pour cela sur les fonctionnalités du logiciel QGIS. Principales notions abordées :</p> <ul style="list-style-type: none"> <li>— Géoréférencement et projection, utilisation et création de couches géographiques, traitements géométriques simples</li> <li>— Création de cartes statistiques</li> <li>— Cartographie statistique avancée : de l'usage de l'AFC et du diagramme triangulaire pour produire des cartes de synthèse à partir de plusieurs variables.</li> <li>— Principes de la sémiologie graphique et de l'édition de cartes.</li> </ul> <p><b>Pour s'inscrire, faire l'inscription en ligne et écrire à <a href="mailto:helene.noizet@univ-paris1.fr">helene.noizet@univ-paris1.fr</a>.</b></p>	

<b>Textométrie</b> Léo Dumont	Second semestre 2h hebdomadaires
<p>Cette formation vise à doter les participants-es des connaissances et des savoir-faire nécessaires à l'analyse d'un corpus de sources textuelles à l'aide des méthodes de l'analyse des données textuelles.</p> <p>Nous aborderons les techniques d'acquisition et de construction de corpus : reconnaissance optique de caractères manuscrits (HTR) et imprimés (OCR), les méthodes de structuration de corpus, de lemmatisation et d'étiquetage morpho-syntaxique du vocabulaire. Seront également enseignées Les différentes méthodes de mesures lexicales et d'interrogation de corpus étiquetés couramment utilisées en textométrie, les analyses multivariées et les méthodes de classification ainsi que les méthodes cooccurentielles permettant l'étude des relations entre les mots. Une attention sera enfin portée aux techniques venues du domaine du traitement automatique des langues et du <i>machine learning</i>.</p> <p>Pour ce faire différents logiciels pourront être mobilisés : Lexico, TXM, Iramuteq, Hyperbase, des paquets R ou bien Python.</p>	

## CALENDRIER ET HORAIRES DES ENSEIGNEMENTS

**IMPORTANT !**

**L'inscription aux formations générale ou de spécialité est obligatoire via l'EPI :**

<https://cours.univ-paris1.fr/course/view.php?id=33675>

Les salles informatiques D627 et D628 se situent à l'entre-sol, au bout de la galerie Dumas côté rue Cujas, à droite. Les salles 02 et 03 (C616, C615) sont au sous-sol (escalier O), voir le plan p. 8.

### **TD de formation générale**

N°	Jour	Horaire	Salle	Début	Fréquence	Enseignant
01	Lundi	10h-12h	D 627	25/09	Tous les 15 jours	J. Alerini
02	Lundi	10h-12h	D 627	02/10	Tous les 15 jours	J. Alerini
03	Lundi	12h30-14h	D 628	02/10		J.-C. Balois
04	Lundi	15h30-17h	02	02/10		J.-C. Balois
05	Mardi	9h30-11h30	D 628	26/09	Tous les 15 jours	L. Dumont
06	Mardi	9h30-11h30	D 628	03/10	Tous les 15 jours	L. Dumont
07	Mardi	10h-12h	02	26/09		G. Bonnot
08	Mardi	13h-15h	D 628	26/09	Tous les 15 jours	L. Dumont
09	Mardi	13h-15h	D 628	03/10	Tous les 15 jours	L. Dumont
10	Mardi	14h-16h	03	26/09		J. Alerini
11	Mercredi	8h-10h	03	27/09	Tous les 15 jours	G. Bonnot
12	Mercredi	8h-10h	03	04/10	Tous les 15 jours	G. Bonnot
13	Jeudi	11h-13h	D 628	12/10		O. Julien
14	Vendredi	13h-15h	D 628	13/10		O. Julien
15	Vendredi	15h-17h	D 628	13/10		O. Julien

### **Formations spécialisées et/ou intégrées au parcours SDH**

#### **1er semestre**

Formation	Jour	Horaire	Salle	Début	Enseignant
Modélisation et bases de données (1)	Lundi	14h30-16h30	02	25/09	J. Alerini
Programmation avec Python	Mardi	10h00-12h00	D 627	26/09	O. Julien
Analyse de réseaux	Mardi	13h30-15h30	D 627	07/11	O. Julien
Analyse et visualisation de données	Jeudi	14h00-17h00	02	28/09	G. Bonnot

#### **2e semestre**

Formation	Jour	Horaire	Salle	Début	Enseignant
Édition de sources (L <sup>A</sup> T <sub>E</sub> X et XML-TEI)	Lundi	15h30-17h30		29/01	G. Bonnot
Cartographie et SIG	Mardi	10h00-13h00		30/01	H. Noizet
Textométrie	Mardi	13h00-15h00		30/01	L. Dumont
Modélisation et bases de données (2)	Jeudi	10h00-12h00		31/01	P. Brochard



FIGURE 1 – Localisation des salles informatiques (centre Sorbonne)

