



École
nationale
des
chartes

Master Humanités Numériques et Computationnelles

Université PSL

Année universitaire 2024 - 2025

Table des matières

Maquette de la première année	2
Maquette de la seconde année	4
Séminaires extérieurs	5
Liste des établissements partenaires et offres de formation	5
Parcours thématiques	5
Parcours "Littérature"	5
Parcours "Histoire contemporaine"	6
Parcours "Histoire moderne"	7
Parcours "Histoire médiévale"	8
Parcours "Histoire ancienne"	9
Parcours "SHS computationnelles"	10
Parcours "Philologie langue et littérature médiévales"	11
Parcours "Paléographie et Image"	12
Semestre 1 - Tronc commun, liste des cours	13
Semaine intensive de rentrée	13
Initiation aux humanités numériques (EPHE 100)	13
Introduction aux mathématiques (ENC 2313)	13
Fondamentaux de l'informatique (ENC 2435)	14
Méthodes quantitatives en SHS (ENC 2310)	14
Algorithmique et programmation pour les humanités : introduction à Python (ENC 2436)	15
Méthodologie et mémoire (ENC 2314)	15
Séminaire de recherche du master (ENC 2316)	16
Semestre 2 - Tronc commun, liste des cours	16
Structuration XML et XML/TEI (ENC 2437-1)	17
Algorithmique et programmation pour les humanités : Python avancé (ENC 2320)	18
Apprentissage machine (ENC 2321)	18
Acquisition de données et annotation en DataScience	18
Séminaire de recherche du master (ENC 2316)	18
Mémoire	19
Semestre 3 - Tronc commun, liste des cours	19
Fondamentaux de mathématiques et Python pour DataScience (ENC 2325)	19
Traitement automatique de l'image	19
Approches quantitatives et modélisation mathématique en SHS (ENC 2322)	19
Langue vivante (au choix)	20
Séminaire de recherche du master (ENC 2316)	20
Cours de spécialisation : Philologie computationnelle (ENC 2323)	20
Cours de spécialisation : Humanités numériques spatialisées	22
Cours de spécialisation : Traitement automatique de la langue et analyse sémantique (ENS N104)	22
Mémoire – bilan d'étape	22
Semestre 4 - Tronc commun, liste des cours	22
Algorithmique et programmation pour les humanités : Python avancé (ENC 2320)	23
Cours de spécialisation : Traitement automatique de l'image (EPHE P102)	23
Cours de spécialisation : Analyse de réseaux (ENC 2324)	23
Cours de spécialisation : Paléographie computationnelle	23
Séminaire de recherche du master (ENC 2316)	23
Stage de découverte – 3 ECTS	24
Mémoire et soutenance	24

Maquette de la première année

La première année de Master se concentre sur l'acquisition des méthodologies de base en SHS computationnelles ainsi que dans les champs disciplinaires et thématiques de spécialisation de l'étudiant. Il couvre notamment un large spectre de cours en programmation (Python, XML, bash, Git), méthodes quantitatives et apprentissage machine. Le programme se subdivise entre:

- un tronc commun de méthodes computationnelles en SHS (40 ECTS, répartis sur les deux semestres);
- un programme pédagogique individualisé, composé essentiellement d'enseignements méthodologiques et thématiques propres à la discipline de l'étudiant : histoire, philologie, lettres, etc. (20 ECTS). Des parcours pré-définis sont proposés aux étudiants, mais ces derniers peuvent composer un parcours sur-mesure, en concertation avec leur directeur de mémoire et le responsable pédagogique de la formation.

Les étudiants réalisent en outre un mini-mémoire, qui compte pour 1/6e de la moyenne de l'année, d'une trentaine de pages. Les cours du tronc commun humanités numériques se concentrent sur les jeudis et vendredis durant les deux semestres.

Intitulé de l'enseignement	Heures	ECTS
S1 - Tronc commun	127	20
Semaine intensive de présentation du master	20	
Initiation aux humanités numériques	12	3
Introduction aux mathématiques	20	3
Fondamentaux de l'informatique	22	3
<i>Systèmes d'exploitation, algorithmique, Linux</i>	8	
<i>Introduction au bash</i>	4	
<i>Expressions régulières</i>	4	
<i>Bonnes pratiques de développement, Git</i>	6	
Méthodes quantitatives en SHS	10	2
Algorithmique et programmation pour les humanités : introduction à Python	20	4
Méthodologie et mémoire	18	4
<i>Méthodologie de la recherche en SHS computationnelles (ateliers)</i>	6	
<i>Outils pour la rédaction du mémoire</i>	12	
Séminaire de recherche du master	5	1
S2 - Tronc commun	71	20
Structuration XML et XML/TEI	20	4
Algorithmique et programmation pour les humanités : Python avancé	20	4
Apprentissage machine	14	3
Acquisition de données et annotation en DataScience	12	2
Séminaire de recherche du master	5	1
Mémoire		6

Intitulé de l'enseignement	Heures	ECTS
S1 et S2 - Enseignements thématiques et d'ouverture	env. 100	20
Bloc enseignements thématiques (3 au choix) <i>Cours de méthodologie de votre discipline</i> <i>2 × enseignements thématiques au choix</i>		12
Bloc enseignements d'ouverture (1 au choix) <i>1 × PSL week + 2 × langues</i> <i>1 × PSL week + 1 × langue + 1 × sport/théâtre</i> <i>2 × PSL week</i> <i>1 × cours thématique</i>		4
Option technique (1 au choix) <i>Option technique de M2</i> <i>Langage SQL (voir master TNAH)</i> <i>Langage R (voir offre de formation EPHE / EHESS)</i>		4
TOTAL	298	60

Maquette de la seconde année

La deuxième année de master se construit sur un tronc commun, permettant d'approfondir l'initiation reçue en première année, tandis que des cours et séminaires optionnels permettent à chaque étudiant d'approfondir sa spécialisation dans les méthodes propres à son champ disciplinaire. Les étudiants suivent également plusieurs séminaires de recherche. Un séminaire du master donne l'occasion aux étudiants de deuxième année de présenter leurs travaux aux étudiants de première année. La deuxième année est validée pour 1/3 par l'écriture et la soutenance d'un mémoire de recherche, encadré par un directeur en SHS computationnelle (enseignant du master) et un directeur thématique. Le mémoire devra mettre les méthodes computationnelles (données structurées, algorithmes, méthodes quantitatives, ...) au service de la recherche, via une démarche de test d'hypothèse appuyée sur des données, et une bonne connaissance de l'état de l'art (tant en humanités numériques qu'en SHS traditionnelles).

Pour les étudiants visant une poursuite en doctorat, le choix des séminaires de recherche optionnels est crucial. De même, pour les étudiants envisageant une professionnalisation dès la sortie du master 2, le choix des cours et séminaires doit refléter cette ambition.

Intitulé de l'enseignement	Heures	ECTS
S3 - Tronc commun	79	20
Fondamentaux de mathématiques et Python pour DataScience	10	2
Traitement automatique de l'image	10	2
Approches quantitatives et modélisation mathématique en SHS	20	4
Langue vivante (au choix)	10	1
Spécialisation de SHS computationnelles (1 au choix)	24	5
<i>Philologie computationnelle</i>		
<i>Humanités numériques spatialisées</i>		
<i>Traitement automatique de la langue et analyse sémantique</i>		
Séminaire de recherche du master	5	1
Stage de découverte en laboratoire (4 à 8 semaines)		3
Mémoire (bilan d'étape)		2
S4 - Tronc commun	49	10
Algorithmique et programmation pour les humanités : Python avancé	20	4
Spécialisation de SHS computationnelles (1 au choix)	24	5
<i>Traitement automatique de l'image</i>		
<i>Analyse de réseaux</i>		
<i>Paléographie computationnelle</i>		
Séminaire de recherche du master	5	1
S3 et S4 - Enseignements thématiques et d'ouverture	62	12
Bloc enseignements thématiques (2 au choix)	52	10
<i>Enseignement thématique au choix</i>		5
<i>Enseignement thématique au choix</i>		5
Bloc enseignements d'ouverture (1 au choix)		2
<i>Enseignement d'ouverture (séminaire, cours, PSL week, etc.)</i>	10	
S4 - Mémoire et soutenance		18
TOTAL	190	60

Séminaires extérieurs

Les enseignements et séminaires extérieurs sont à sélectionner dans les **établissements partenaires exclusivement**, sous réserve de l'accord de l'enseignant concerné et de l'approbation par le jury d'admission. Voir la section pour les cours recommandés.

Liste des établissements partenaires et offres de formation

- **École normale supérieure - PSL** : voir les séminaires proposés dans chaque départements de l'ENC
- **École pratique des hautes études - PSL** : Voir la liste des conférences des sections de sciences humaines ouverts sous réserve de l'accord de l'enseignant concerné.
- **École nationale des chartes - PSL** : Voir les séminaires assurés par les professeurs de l'École des chartes
- **École des hautes études en sciences sociales** : Voir la liste des enseignements
- **PSL week** : Voir les PSL week proposées

Parcours thématiques

Plusieurs parcours ont été définis par les enseignants du Master, afin d'identifier les enseignements les plus adaptés à votre champ disciplinaire. Les cours listés ci-dessous sont indicatifs mais **hautement recommandés et doivent être choisis en priorité**. Il vous est conseillé de prendre contact avec l'enseignant référent ou le responsable pédagogique de la formation en cas de doute dans le choix des cours.

Parcours "Littérature" / Enseignant référent : Jean Barré

- **École normale supérieure - PSL**
 - Approches du récit
 - Recherches contemporaines en narratologie. 2023-24 : L'acceptabilité des récits
 - Linguistique et stylistique : formation accélérée
- **EHESS**
 - UE450 - Histoire et littérature, 1800-1950
 - Écritures du passé. Littérature et histoire : méthodes, théories, terrains

Parcours "Histoire contemporaine" / Enseignant référent : Marie Puren

● **École normale supérieure - PSL**

- Méthodologie :
 - * L'histoire en chantier (cours de méthodologie trans-période)
 - * Cours d'initiation à l'histoire contemporaine
- En fonction du sujet de mémoire :
 - * Histoire contemporaine
 - * Séminaires associés
 - * Histoire transnationale et globale
 - * Séminaires d'élèves
 - * Cours associés au département

● **EPHE - PSL**

- En fonction du sujet de mémoire :
 - * Histoire de la guerre (XIXe-XXe) (séminaire de Martin Motte)
 - * Histoire des spectacles à l'époque contemporaine (séminaire de Jean-Claude Yon)
 - * Histoire de la diplomatie et des relations internationales (séminaire de Yves Bruley)
 - * Archives orales et histoire des organisations (séminaire de Florence Descamps)

● **École nationale des Chartes-PSL**

- Méthodologie :
 - * Historiographie
 - * Diplomatie - période contemporaine (Edouard Vasseur)
- En fonction du sujet de mémoire :
 - * Histoire du droit - période contemporaine (Katia Weidenfeld)
 - * Histoire des médias contemporains
 - * Nouveaux objets patrimoniaux
 - * Histoire des institutions - période contemporaine (Edouard Vasseur)
 - * Histoire du livre
 - * Géographie historique - période contemporaine (Edouard Vasseur)

Si votre emploi du temps ne permet pas d'assister au cours d'initiation à la recherche en histoire proposé par l'ENS, l'EHESS propose 2 cours d'initiation :

- Le passé au présent. Séminaire commun du master Histoire
- Méthodologie de la recherche en histoire (le lien renvoie au groupe 1 mais il y a 4 groupes en tout ; à voir avec l'enseignant et en fonction de votre emploi du temps)

- **École nationale des Chartes-PSL**

- Méthodologie :
 - * Historiographie
 - * Diplomatique - période moderne (Olivier Poncet)
- Paléographie :
 - * Initiation à la paléographie française 1500-1700
 - * Paléographie française : XVe-XVIIIe siècle (actes et documents)
 - * Paléographie française : XVe-XVIIIe siècle (actes, documents, typographie)
 - * Paléographie française : XVe-XVIIIe siècle, approfondissement
- En fonction du sujet de mémoire :
 - * Histoire des institutions - période moderne (Olivier Poncet)
 - * Histoire du livre - période moderne
 - * Géographie historique - période moderne (Olivier Poncet)

- **École normale supérieure - PSL**

- Méthodologie :
 - * L'histoire en chantier (cours de méthodologie trans-période)
 - * Cours d'initiation à l'histoire moderne
- En fonction du sujet de mémoire :
 - * Cours associés au département
 - * Séminaires d'élèves
 - * Histoire transnationale et globale
 - * Régulations du travail et de la société
 - * Les jeudis de l'histoire et de la philosophie des sciences

- **EHESS**

- Méthodologie (si impossibilité de suivre les cours de l'ENS) :
 - * Le passé au présent. Séminaire commun du master Histoire
 - * Méthodologie de la recherche en histoire
 - * Les outils de l'historien.ne
 - * Pratiquer les archives 1
 - * Pratiquer les archives 2
- En fonction du sujet de mémoire :
 - * Séminaires en histoire
 - * Mémoire en histoire du genre

- **EPHE - PSL**

- En fonction du sujet de mémoire :
 - * Séminaire d'Hubert Bost
 - * Séminaire de Jean-Pierre Brach
 - * Séminaire d'Emmanuelle Chapron
 - * Séminaire d'Olivier Christin
 - * Séminaire de Silvio de Franceschi
 - * Séminaire de Sabine Frommel
 - * Séminaire de Michel Hochmann
 - * Séminaire de Vassa Kontouma
 - * Séminaire de Jean-Louis Quantin

- **École nationale des Chartes-PSL**

- Méthodologie :
 - * Historiographie
 - * Matériaux d'archives médiévales : panorama
 - * La tradition de l'acte médiéval
 - * Sigillographie. Genèse des actes. Faux.
 - * Édition des textes documentaires
 - * Le latin des documents de la pratique
- En fonction du sujet de mémoire :
 - * L'épopée de Virgile à l'Arioste (trans-période)
 - * Histoire du droit - droit civil et droit canonique (Patrick Arabeyre)
 - * Histoire des institutions - période médiévale (Olivier Guyotjeannin)
 - * Géographie historique - période médiévale (Olivier Guyotjeannin)
 - * Cours de paléographie latine ou française

- **École normale supérieure - PSL**

- Méthodologie :
 - * L'histoire en chantier (cours de méthodologie trans-période)
 - * Cours d'initiation à l'histoire médiévale
 - * A la source : diversité et matérialité du métier d'historien
- En fonction du sujet de mémoire :
 - * Séminaires d'histoire médiévale

- **EHESS**

- Méthodologie :
 - * Le passé au présent. Séminaire commun du master Histoire
 - * Méthodologie de la recherche en histoire
 - * Les outils de l'historien.ne
- En fonction du sujet de mémoire :
 - * Archéologie du haut Moyen-Âge
 - * Liste des séminaires du département d'histoire

- **EPHE - PSL**

- En fonction du sujet de mémoire :
 - * Séminaires du master Civilisations, Cultures et Sociétés, spécialité Études médiévales
 - * Séminaire d'Astrid Castres
 - * Séminaire d'Émilie Dorgeix (trans-période)
 - * Séminaire de Corinne Leveaux-Teixeira (trans-période)
 - * Séminaire de Philippe Lorentz

- **École nationale des Chartes-PSL**

- Méthodologie :
 - * L'épopée de Virgile à l'Arioste (trans-période)

- **École normale supérieure - PSL**

- Méthodologie :
 - * L'histoire en chantier (cours de méthodologie trans-période)
 - * Cours d'initiation à l'histoire ancienne
 - * A la source : diversité et matérialité du métier d'historien
 - * Aux sources de l'histoire antique (département Antiquité)
 - * Introduction à l'histoire de l'Antiquité (département Antiquité)
- En fonction du sujet de mémoire :
 - * Séminaires en histoire ancienne
 - * Décor et architecture antiques d'Orient et d'Occident
 - * Archéologie et histoire de l'Orient hellénistique et romain I. L'Orient méditerranéen (Égypte, Syrie, Palestine, Anatolie, Chypre)

- **EHESS**

- Méthodologie :
 - * Le passé au présent. Séminaire commun du master Histoire
 - * Méthodologie de la recherche en histoire
 - * Les outils de l'historien.ne
- En fonction du sujet de mémoire :
 - * Une première sélection (AnHiMA)
 - * Séminaires du département d'histoire

- **EPHE - PSL**

- En fonction du sujet de mémoire :
 - * Histoire de l'art et archéologie du monde byzantin et de l'Orient chrétien - cours de master
 - * Séminaire de Pascal Butterlin
 - * Séminaire de Geoffrey Herman
 - * Séminaire de Frédérique Duyrat
 - * Séminaire d'Ivan Guermeur
 - * Séminaire d'Antony Hostein
 - * Séminaire de François Queyrel
 - * Séminaire de Catherine Saliou
 - * Séminaire de William Van Andringa
 - * Séminaire de Philippe Bernard (trans-période)

Parcours "Philologie langue et littérature médiévales" / Enseignants référents : Lucence Ing et Jean-Baptiste Camps

- **EPHE - PSL**

- Linguistique romane (séminaire de Martin Glessgen)
- Philologie romane (séminaire de Fabio Zinelli)
- Langue et littérature latines du Moyen Âge (Anne-Marie Turcan-Verkerk)

- **École nationale des Chartes-PSL (philologie romane)**

- Initiation à la philologie romane : l'ancien français
- Initiation à la philologie romane : le moyen français
- Philologie romane approfondie : occitan, gascon et franco provençal
- Philologie romane approfondie : dialectologie médiévale
- Introduction générale au français médiéval et à la philologie
- Édition des textes, Généralités
- Édition de textes, L'édition des textes littéraires latins et gallo-romans
- Le livre manuscrit, Cours théorique de codicologie
- Paléographie, Paléographie latine : XIIe-XVe siècle (initiation)
- Paléographie, Paléographie française : XIIe-XVe siècle (initiation)

- **École nationale des Chartes-PSL (philologie (medio)latine)**

- Le latin des textes littéraires : approche linguistique
- Latin des textes littéraires : approche stylistique
- Histoire des textes : l'épopée de Virgile à l'Arioste
- Histoire des textes : les courants littéraires, du IIIe au XIIe s.
- Édition des textes, Généralités
- Édition de textes, L'édition des textes littéraires latins et gallo-romans
- Le livre manuscrit, Cours théorique de codicologie
- Paléographie, Paléographie latine : XIIe-XVe siècle (initiation)
- Paléographie, Paléographie française : XIIe-XVe siècle (initiation)

- **École Normale Supérieure-PSL**

- Cours de Nathalie Khoble
- Cours de formation accélérée à la langue et à la littérature médiévales en français
- Littérature médiévale et théorie littéraire
- Médiévie. Séminaire de recherche en littérature médiévale (Nathalie Khoble)

Parcours "Paléographie et Image" / Enseignant référent : Chahan Vidal-Gorène

- **EPHE - PSL**

- Humanités numériques et computationnelles appliquées à l'étude de l'écrit ancien
- Méthodologie de la recherche en histoire de l'art et archéologie
- Méthodes et applications de l'imagerie 3D en SVT et SHS T et SHS (enseignement à distance)
- Histoire de l'art et archéologie du monde byzantin et de l'Orient chrétien - cours de master

- **EHESS**

- Atelier méthodologique en histoire de l'art et études visuelles

- **École Normale Supérieure - PSL**

- Initiation à l'épigraphie grecque
- Initiation à l'épigraphie latine et à la civilisation romaine

- **École nationale des Chartes-PSL**

- Ensemble des cours de paléographie et d'épigraphie dispensés à l'école des Chartes
- Le processus archéologique : objets et méthodes

Semestre 1 - Tronc commun, liste des cours

Semaine intensive de rentrée

La première semaine de l'année est occupée par des présentations des enseignants du master, visant à présenter leur domaine, leurs recherches actuelles, les problèmes qu'elles adressent, et les possibles sujets de recherches futures (notamment les sujets de mémoire qui pourraient être traités par les étudiants du master). L'objectif est de plonger d'emblée les étudiants et étudiantes dans la recherche la plus actuelle en SHS computationnelles, et de leur faire entrevoir ce qu'ils ou elles pourraient y contribuer.

Initiation aux humanités numériques (EPHE 100)

Enseignant : Peter Stokes (EPHE)

Mode d'évaluation : dossier à rendre en janvier

Volume horaire : 12h

ECTS : 3

Programme :

1. 3h par séance, sauf 2h la dernière
2. Qu'est-ce que ce sont les HN. Principes de l'ordinateur; langages de programmation; séance pratique en programmation basique.
3. Le livre numérique ; HTML, PDF et ePub ; numérisation de textes et de documents éditions et représentations; séance pratique en HTML et TEI (Oxygen).
4. Humanités visuelles : images ; séance pratique 2D et élaboration d'images.
5. Données structurées; bases des données, ontologies et RDF ; séance pratique en BDD
6. TAL; authorship attribution; stylometry; distant reading; séance pratique en concordances (Word-smith et TMX) et expressions régulières.
7. Intelligence Artificielle et Apprentissage de la Machine
8. Modélisation, prototypage et maquettes

Éléments bibliographiques :

1. MOUNIER, Pierre, dir. (2012). Read/Write Book 2 : Une introduction aux humanités numériques. Nouvelle édition [en ligne]. Marseille : OpenEdition Press.
2. CLIVAZ, Claire (2019). Ecritures Digitales : Digital Writing, Digital Scriptures, esp. Sauf le nom : nommer les Digital Humanities (p. 44–81). Leiden : Brill.
3. MCCARTY, Willard (2004). Modeling : A Study in Words and Meanings. In Siemens, R. & Schreibman, S. A Companion to Digital Humanities. Oxford : Blackwell, p. 254–70.
4. FLANDERS, Julia, et JANNIDIS, Fotis (2015). Data Modeling. In Schreibman, S., Siemens, R. and Unsworth, J. (dir.), A New Companion to Digital Humanities. Chichester, UK : John Wiley & Sons, pp. 229–37 doi :10.1002/9781118680605.ch16

Introduction aux mathématiques (ENC 2313)

Enseignant : Michel Capot (SPIDEO)

Mode d'évaluation : participation

Volume horaire : 20h

ECTS : 3

Programme : Introduction générale aux mathématiques : le langage symbolique et ses notations, éléments de logique, fonctions et dérivées, algèbre linéaire et éléments de probabilité et de statistique.

Éléments bibliographiques :

1. Patrick Juola et Stephen Ramsay, Six Septembers : Mathematics for the Humanist, 2017.

Fondamentaux de l'informatique (ENC 2435)

Enseignants : Marc Bui (EPHE), Vincent Jolivet (ENC), Jean-Victor Boby (ENC).

Mode d'évaluation : devoir sur table

Volume horaire : 22

ECTS : 3

Programme : Ce module permet de découvrir les fondamentaux du fonctionnement du système d'exploitation Linux, du mode texte, des réseaux et des serveurs Web d'une part, de la programmation d'autre part, ce dernier point grâce à une introduction à l'algorithmique ponctuée d'exercices. Il propose également une première approche à la fouille de texte via l'utilisation d'expressions régulières, ainsi qu'aux bonnes pratiques du développement collaboratif, par une initiation à Git.

1. Contenu 1 : Séminaire intensif / install Party
2. Contenu 2 : Linux et Shell
3. Contenu 3 : Expressions régulières
4. Contenu 4 : Git et bonnes pratiques de développement

Éléments bibliographiques :

1. Mathieu Nebra, "Reprenez le contrôle à l'aide de Linux", OpenClassrooms, en ligne
2. Documentation Ubuntu Francophone, notamment "Initiez-vous à GNU/Linux avec Ubuntu", en ligne.
3. Eric Sanchis, Introduction à la programmation en Bash, en ligne.
4. Mendel Cooper, Guide avancé d'écriture des scripts Bash, en ligne.
5. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest et Clifford Stein, Algorithmique, Paris, 2010.
6. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein, Introduction to Algorithms, 3e édition, MIT, 2009, en ligne
7. Ebel Franck et Rohaut Sébastien, Algorithmique. Techniques fondamentales de programmation. Exemples en Python, Edition ENI, 2014.

Méthodes quantitatives en SHS (ENC 2310)

Enseignants : Florian Cafiero (CNRS), Marie Puren (CNRS)

Mode d'évaluation : exposé portant sur la lecture d'un article de recherche et devoir sur table

Volume horaire : 10h

ECTS : 2

Programme : Ce cours est conçu comme une première approche à l'utilisation de méthodes quantitatives et statistiques dans les sciences humaines. Son objectif principal est de permettre la lecture et la compréhension autonomes de travaux de recherche se situant dans le champ des méthodes quantitatives en sciences humaines et des humanités numériques, en donnant les clefs et concepts nécessaires pour accéder à la partie mathématique du raisonnement. Il doit notamment permettre de pouvoir apprécier en connaissance de cause les partis pris méthodologiques ou les résultats présentés et leur significativité. On présentera également la démarche nécessaire à l'application de ces méthodes à une recherche académique, en mettant en lumière leur potentiel comme leurs limites. Dans un premier temps, ce cours se concentrera sur quelques notions et concepts fondamentaux, et sur les clefs de lecture les plus indispensables. Dans un second temps, il sera consacré à la lecture critique d'articles et l'étude de cas concrets d'application.

1. Introduction et clefs de lecture ; SHS et données

2. Analyse de réseaux
3. Analyse de données textuelles
4. Lecture distante
5. Modèle de sujet
6. Stylométrie
7. Analyse causale
8. Systèmes complexes et modélisation mathématique
9. intelligence artificielle

Algorithmique et programmation pour les humanités : introduction à Python (ENC 2436)

Enseignant : Jean Barré (ENC).

Mode d'évaluation : contrôle continu (1/2) et devoir maison final (1/2)

Volume horaire : 20h

ECTS : 4

Programme : Ce module a pour objet de donner les connaissances de base en programmation informatique. Il revient sur les fondamentaux de l'algorithmique, mis en œuvre autour d'un langage très usité dans la recherche en SHS, Python.

Détails sur le mode d'évaluation : Note finale = $(A + B) / 2$, où A (note sur 20) = Somme des 4 meilleures notes des devoirs à rendre au cours du semestre (6 à rendre). Chaque devoir produit une note sur 5, chaque devoir non rendu équivaut à une note de 0.

Méthodologie et mémoire (ENC 2314)

Enseignants : Enseignants du master.

Mode d'évaluation : mini-mémoire à rendre le 1ER JUIN (PREMIER JUIN, [CALENDIS JUNII]).

Volume horaire : 20h

ECTS : 4

Programme :

- Cours 1 : Methodologie de la recherche en SHS computationnelles (ateliers - 6h)

Ce module de tutorat permet à chaque étudiant de se lancer dans l'application de méthodes et outils des humanités numériques à ses problématiques de recherche et à son mémoire. À l'issue de la première année de master, les étudiants devront rendre un mini-mémoire de dimensions raisonnables (une trentaine de pages), accompagnées d'annexes numériques (données, scripts, etc.) permettant la vérification et la reproduction des analyses présentées. Ce mini-mémoire doit être conçu comme un premier pas vers une démarche de recherche approfondie en M2, pour permettre d'en vérifier la faisabilité. Partant d'une question de recherche et d'un état de l'art du sujet, il doit se concentrer sur un sous-ensemble du corpus envisagé, et proposer un premier protocole d'analyse, ainsi que quelques pistes d'interprétation des résultats. Il est impératif d'identifier des directeurs, de définir le sujet et de commencer l'état de l'art et les recherches aussi tôt que possible dans l'année. N'hésitez pas à solliciter les conseils de l'équipe pédagogique : la démarche de préparation du mémoire vous est propre, doit être faite en autonomie, et il n'appartient pas aux directeurs de recherche de venir vous chercher, mais à vous des les solliciter quand vous en éprouvez le besoin, ainsi que pour faire des points réguliers. Réaliser un mémoire peut (doit) être une entreprise passionnante, et c'est le travail où vous mettez le plus de vous-mêmes. Son succès dépend avant tout de votre investissement et de votre intérêt pour la problématique. Il dépend toutefois aussi de la faisabilité et de la pertinence de l'approche, points sur lesquels vous aurez besoin des conseils de l'équipe pédagogique, ainsi que du test grandeur nature fourni par le mini-mémoire de M1. Dans ce contexte, le premier semestre devra être le moment de débiter ce travail en :

- déterminant le sujet, avec l'accord de directeurs de recherche (directeur de spécialité et directeur "numérique") et élaborant un questionnement et une problématique;

- faisant l'état de l'art;
- collectant les données.
- Ateliers du premier semestre
 1. Qu'est-ce qu'un mémoire? Construire une démarche de recherche.
 2. Les données et le protocole d'analyse.
 3. Résultats et interprétation.
- Jean-Baptiste Camps, biblatex-enc, BibLaTeX style files for the bibliographic standards of the École nationale des chartes, en ligne : (<https://ctan.org/pkg/biblatex-enc>), <https://github.com/Jean-Baptiste-Camps/biblatex-enc>.
- Imprimerie nationale, Lexique des règles typographiques en usage à l'Imprimerie nationale, Paris, 2002
- Mémoires des années précédentes, prochainement en ligne sur Zenodo.
- A companion to digital humanities, dir. Susan Schreibman, Raymond Georges Siemens, et John Unsworth, Oxford, 2004. Une version assez largement révisée de cet ouvrage est parue en 2016 : A new companion to digital humanities, dir. Susan Schreibman, Raymond Georges Siemens et John Unsworth, Malden, etc., 2016.
- Jim Gray, " Jim Gray on eScience : A transformed scientific method". In The fourth paradigm : Data-intensive scientific discovery, édité par Tony Hey, Stewart Tansley, et Kristin Tolle, xviixxxi. Washington, 2009.
- Digital Scholarship in the Humanities (olim Litterary and Linguistic Computing).
- Digital Humanities Quarterly.
- Actes des conférences annuelles Digital Humanities.
- Cours 2 : Outils pour la rédaction du mémoire (ENC 2314-2) (12h)
Présentation de quelques outils pour la constitution de bibliographies (Zotero), pour la rédaction du mémoire (L^AT_EX) et de bonnes pratiques (sauvegardes fréquentes, organisations des fichiers, etc.).

Séminaire de recherche du master (ENC 2316)

Enseignants : Enseignants du Master

Mode d'évaluation : participation et discussion (M1); présentation de 15 minutes (M2)

Volume horaire : 5h

ECTS : 1

Programme : Ce séminaire, qui s'étend sur les quatre semestres de la formation, permet aux étudiants de première année de découvrir les recherches les plus actuelles dans le domaine des humanités numériques, notamment celles menées par des étudiants plus avancés. Aux étudiants de deuxième année, elle fournit également l'occasion de confronter leurs expériences et leurs idées, et de s'initier à la présentation de leurs travaux.

Détails sur le mode d'évaluation : Pour les étudiants de M2, ils doivent une fois dans l'année réaliser une présentation de leurs recherches de 15 minutes; Pour les étudiants de M1, lightning talk sur leur mini-mémoire. **Éléments bibliographiques** :

- Les compte-rendus des années précédentes sont disponibles sur le carnet de recherche du séminaire : Master HN, Données et sciences humaines : Carnet de recherche du master Humanités numériques de l'Université Paris, Sciences & Lettres, dir. Jean-Baptiste Camps & Marie Puren & Chahan Vidal-Gorène, <https://humanodon.hypotheses.org/>, 2019-

Semestre 2 - Tronc commun, liste des cours

Structuration XML et XML/TEI (ENC 2437-1)

Enseignants : Marguerite Vernet, Lucence Ing (ENC)

Mode d'évaluation : dossier d'édition électronique

Volume horaire : 20h

ECTS : 4

Mode d'évaluation : Évaluation finale : proposer un encodage XML-TEI pertinent et argumenté pour un corpus donné. Le fichier XML devra être régi par un schéma ODD et accompagné d'une documentation.

1. Séance 1 Introduction à XML
2. Séance 2 Présentation de TEI
3. Séance 3 Structuration générale d'un fichier TEI
4. Séance 4 Les TEI guidelines
5. Séance 5 Initiation à l'édition scientifique et description des sources manuscrites
6. Séance 6 Encoder une édition scientifique
7. Séance 7 Les schémas XML (1) : Les différents formats
8. Séance 8 Les schémas XML (2) : Personnalisations avancées de ODD
9. Séance 9 Les schémas XML (3) : Mise en pratique
10. Séance 10 Initiation à Xpath
11. Séance 11 Transformer et interroger son XML
12. Séance 12 Retour d'expérience

- Ouvrages à consulter

- BUREAU, Bruno, "Quelques réflexions sur la notion de littérarité à partir de l'édition numérique de commentateurs anciens", *Interférences*. *Ars scribendi*, janvier 2012, En ligne : <http://journals.openedition.org/interferences/186>.
- BURNARD, Lou et BURGHART, Marjorie, *Qu'est-ce que la Text Encoding Initiative ?*, 2015, en ligne : <http://books.openedition.org/oep/1237>.
- DUVAL, Frédéric, "Pour des éditions numériques critiques. L'exemple des textes français", in *Le texte à l'épreuve du numérique*, Saint-Denis, France, Presses universitaires de Vincennes, 2017, p. 1330.
- HAROLD, Elliotte Rusty, MEANS, W. Scott, ENSARGUET, Philippe [et al.], *XML en concentré*, Paris, O'Reilly, 2005.
- KAY, Michael R., *XPath 2.0 programmer's reference*, Indianapolis, IN, Wrox Press, 2004.
- PALLUAULT, Florent, "Informatiser des descriptions complexes : l'utilisation de l'EAD et de la TEI pour les manuscrits et les livres anciens en France", *IFLA 2012 (Helsinki)*, en ligne.
- PIERAZZO, Elena, *Digital scholarly editing : theories, models and methods*, Farnham Burlington (Vt.), Ashgate, 2015.
- RAHTZ, Sebastian et BURNARD, Lou, "Reviewing the TEI ODD System", *ACM*, 2013, p. 193 196.
- ROBINSON, Peter, "Where We Are with Electronic Scholarly Editions, and Where We Want to Be", *Jahrbuch für Computerphilologie*, vol. 5 / 5, 2003, p. 126146.
- Ron Van den Branden, Melissa Terras, et Edward Vanhoutte, *TEI by Example*.

- Ressources numériques

- BURGHART, Marjorie, PIERAZZO, Elena et al., Online course : Digital Scholarly Editions : Manuscripts, Texts, and TEI Encoding – Digital Editing of Medieval Manuscripts, en ligne : <https://www.digitalmanuscripts.eu/digital-editing-of-medieval-texts-a-textbook/>.
- TEIguidelines, en ligne : <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>.
- XML Tutorial, en ligne : <https://www.w3schools.com/xml/default.asp>.
- XPath Tutorial, en ligne : https://www.w3schools.com/xml/xpath_intro.asp.
- XSLT Tutorial, en ligne : https://www.w3schools.com/xml/xsl_intro.asp.
- Éditions à consulter
 - Geste : un corpus de chansons de geste, <http://dev.chartes.psl.eu/elec/geste/>.
 - L'année 1437 dans la pratique de Pierre Christofle, notaire du Châtelet d'Orléans.
 - Perseus Digital Library, <http://www.perseus.tufts.edu/hopper/>.

Algorithmique et programmation pour les humanités : Python avancé (ENC 2320)

Enseignant : Chahan Vidal-Gorène (ENC)

Mode d'évaluation : contrôle continu (1/4) et projet de groupe (3/4)

Volume horaire : 20h

ECTS : 4

Programme : Ce cours est consacré au renforcement des compétences en Python et en mathématiques, notamment en mathématiques appliquées à l'apprentissage machine (algèbre linéaire, mesures, métriques, clustering, ...). Ce cours est commun avec les M2. Le programme alterne une année sur deux.

Apprentissage machine (ENC 2321)

Enseignants : Daniel Stockholm (EPHE)

Mode d'évaluation : devoir sur table

Volume horaire : 14h

ECTS : 3

Programme : Ce module propose une introduction aux concepts et aux outils d'apprentissage machine.

- Ian Goodfellow, Yoshua Bengio et Aaron Courville, Deep Learning, MIT, 2016, en ligne : <http://www.deeplearningbook.org/>.

Acquisition de données et annotation en DataScience

Enseignants : Chahan Vidal-Gorène, Florian Cafiero (ENC)

Mode d'évaluation : à définir

Volume horaire : 12

ECTS : 2

Programme : Ce module propose une prise en main des bibliothèques Python de webscrapping, de découverte des API pour l'acquisition d'images et de sons, et leur annotation avec la plateforme LabelStudio. Le cours propose une ouverture à l'utilisation de LLM pour l'annotation de données.

Séminaire de recherche du master (ENC 2316)

Enseignants : Enseignants du Master

Mode d'évaluation : participation et discussion (M1); présentation de 15 minutes (M2)

Volume horaire : 5h

ECTS : 1

Programme : Ce séminaire, qui s'étend sur les quatre semestres de la formation, permet aux étudiants de première année de découvrir les recherches les plus actuelles dans le domaine des humanités numériques, notamment celles menées par des étudiants plus avancés. Aux étudiants de deuxième année, elle fournit également l'occasion de confronter leurs expériences et leurs idées, et de s'initier à la présentation de

leurs travaux.

Détails sur le mode d'évaluation : Pour les étudiants de M2, ils doivent une fois dans l'année réaliser une présentation de leurs recherches de 15 minutes; Pour les étudiants de M1, lightning talk sur leur mini-mémoire. **Éléments bibliographiques** :

- Les compte-rendus des années précédentes sont disponibles sur le carnet de recherche du séminaire : Master HN, Données et sciences humaines : Carnet de recherche du master Humanités numériques de l'Université Paris, Sciences & Lettres, dir. Jean-Baptiste Camps & Marie Puren & Chahan Vidal-Gorène, <https://humanodon.hypotheses.org/>, 2019-

Mémoire

Mode d'évaluation : mini-mémoire à rendre le 1er juin

ECTS : 6

Semestre 3 - Tronc commun, liste des cours

Fondamentaux de mathématiques et Python pour DataScience (ENC 2325)

Enseignant : Chahan Vidal-Gorène (ENC)

Mode d'évaluation : Évalué dans le cadre du cours de programmation python avancé

Volume horaire : 10h

ECTS : 2

Programme : Le cours vise à revoir les bibliothèques essentielles pour la manipulation et la visualisation des données textuelles et non textuelles (image). Il comporte pour moitié des rappels en mathématiques pour l'analyse numérique, le traitement de l'image, ainsi que de l'analyse (étude de fonctions).

Traitement automatique de l'image

Enseignants : Daniel Stockholm (EPHE)

Mode d'évaluation : à définir

Volume horaire : 10h

ECTS : 2

Programme : Module d'initiation aux méthodes et approches en traitement automatique de l'image et computer vision.

Approches quantitatives et modélisation mathématique en SHS (ENC 2322)

Enseignants : Jean-Pierre Nadal (EHESS), Julien Randon-Furling (Univ. Paris 1), Annick Vignes (ENPC).

Mode d'évaluation : Exposé à présenter sur la lecture d'un article de recherche.

Volume horaire : 20h

ECTS : 4

Programme : Le cours, pluri- et trans-disciplinaire, porte sur une palette d'approches quantitatives pour l'analyse et la modélisation mathématique de données : analyse multi-variée; méthodes économétriques; modélisation multi-agent; modèles dynamiques (dynamiques de population...), théorie des jeux... Le cours est structuré de manière à présenter les concepts, méthodes, modèles, dans le cadre d'une démarche globale d'analyse et de modélisation. Les notions et techniques sont systématiquement illustrées sur des cas concrets d'intérêt en SHS. Le cours demande une forte implication des étudiants : mise en œuvre des théories ou méthodes sur des données réelles, lecture critique d'articles. La validation se fait sur un travail personnel s'appuyant sur un article scientifique.

- Cours et ressources d'apprentissage en ligne Statistiques, <http://wikistat.fr/>.
- Cours Statistique et Big Data Mining, Philippe Besse, INSA Toulouse <http://www.math.univ-toulouse.fr/besse/enseignement.html>.

- Charpentier A., Flachaire E. et Ly A., Économétrie et Machine learning, 2017, en ligne.
- Juola, P. & Ramsay S., Six Septembers : Mathematics for the Humanist, Zea Books, 2017.
- Se former avec le numérique, site du Ministère de l'enseignement supérieur, de la recherche et de l'innovation (cours, exercices, moocs, etc, dans divers domaines, tous niveaux).
- Cours Systèmes dynamiques. Chaos et applications de Frédéric Faure, UJF Grenoble, en ligne: https://www-fourier.ujf-grenoble.fr/%7Efaure/enseignement/systemes_dynamiques/index.html.
- T. C. Schelling, Micromotives and Macrobehavior, Norton & Co, 1978. Traduction française : La tyrannie des petites décisions, PUF, 1980. Voir sur ce livre le compte-rendu de lecture de E. Reynaud, paru dans la Revue française de sociologie en 1982.
- Thomas C. Schelling, The Strategy of Conflict, Boston, Harvard University Press, 1960.
- Traduction française Stratégie du conflit, Paris, Presses universitaires de France, 1986.
- Sur ce thème, voir l'interview de Schelling par Erhard Friedberg et Martha Zuber, 27 oct. 2000 (traduite en français par J. Bouyssou).
- J.-P. Nadal et M. B. Gordon, "Choix sous influence sociale : heurs et malheurs de la coordination", dans Qu'appelle-t-on aujourd'hui les sciences de la complexité?, sous la direction de G. Weisbuch & H. Zwirn, Vuibert, 2010, en ligne : <http://freakonometrics.hypotheses.org/files/2017/07/econometrics-ML-final-1.pdf>.
- Voir aussi les articles donnés comme supports de cours.

Langue vivante (au choix)

Volume horaire : 10h

ECTS : 1

Séminaire de recherche du master (ENC 2316)

Enseignants : Enseignants du Master

Mode d'évaluation : participation et discussion (M1); présentation de 15 minutes (M2)

Volume horaire : 5h

ECTS : 1

Programme : Ce séminaire, qui s'étend sur les quatre semestres de la formation, permet aux étudiants de première année de découvrir les recherches les plus actuelles dans le domaine des humanités numériques, notamment celles menées par des étudiants plus avancés. Aux étudiants de deuxième année, elle fournit également l'occasion de confronter leurs expériences et leurs idées, et de s'initier à la présentation de leurs travaux.

Détails sur le mode d'évaluation : Pour les étudiants de M2, ils doivent une fois dans l'année réaliser une présentation de leurs recherches de 15 minutes; Pour les étudiants de M1, lightning talk sur leur mini-mémoire. **Éléments bibliographiques** :

- Les compte-rendus des années précédentes sont disponibles sur le carnet de recherche du séminaire : Master HN, Données et sciences humaines : Carnet de recherche du master Humanités numériques de l'Université Paris, Sciences & Lettres, dir. Jean-Baptiste Camps & Marie Puren & Chahan Vidal-Gorène, <https://humanodon.hypotheses.org/>, 2019-

Cours de spécialisation : Philologie computationnelle (ENC 2323)

Enseignants : J.B. Camps (ENC), F. Cafiero (Univ. Sorbonne Paris Cité).

Mode d'évaluation : dossier d'analyse de données textuelles à rendre en janvier.

Volume horaire : 24h

ECTS : 5

Programme : Ce cours aborde à la fois les méthodes de production et d'analyse de corpus textuels (stylométrie notamment). Ce cours s'adresse à tous ceux qui souhaitent acquérir des notions fondamentales de philologie computationnelle, et veulent découvrir les tenants et aboutissants pratiques et théoriques de la constitution et de l'analyse de corpus. Une très large place est laissée aux travaux pratiques, effectués sur plusieurs corpora. Ce cours envisagera ainsi les méthodes computationnelles en philologie, tant dans la constitution de corpus ou l'édition de textes que dans leur étude et l'établissement de connaissances les concernant. Il suivra la distinction traditionnelle entre haute et basse critique, c'est-à-dire entre problématiques liées à l'attribution des textes, leur datation ou leur localisation, et démarche ecdotique, de la recension des manuscrits à l'établissement de la généalogie des témoins. Le cours envisagera à chaque fois les questions de production ou collecte et formalisation des données, ainsi que leur analyse. Sur un plan technique, les langages utilisés seront :

- R : le langage d'analyse statistique, utilisé via l'environnement de développement RStudio;
- Python : qui servira notamment pour ses bibliothèques d'apprentissage machine;
- XML : langage descriptif incontournable pour la structuration des données textuelles
- Des outils d'administration de corpus et de textométrie seront également présentés, en particulier :
- TXM : plateforme dédiée à la création et l'interrogation de corpus.

Programme :

1. Introduction : fondamentaux de philologie computationnelle;
 2. Ecdotique 1 : manuscrits et données matérielles;
 3. Ecdotique 2 : transcription, reconnaissance des écritures et système graphique;
 4. Ecdotique 3 : alignement des textes et collation;
 5. Stylométrie 1 : première approche à la stylométrie : méthodes exploratoires;
 6. Stylométrie 2 : méthodes de partitionnement;
 7. Stylométrie 3 : méthodes supervisées (KNN, SVM, NN);
 8. Séance d'approfondissements et de pratique selon les problématiques des étudiants.
- Tara Andrews, "The third way : philology and critical edition for a digital age", *Variants : the Journal of the European Society for Textual Scholarship*, 10 (2012), <http://boris.unibe.ch/43071/>.
 - Jean-Baptiste Camps, "Où va la philologie numérique?", *LHT Fabula*, 20 (2018)
 - E. Stamatatos, "A Survey of modern authorship attribution methods", *Journal of the American Society for information Science and Technology*, 60-3 (2009), p. 538-556.
 - Mike Kestemont, "Function words in authorship attribution : From black magic to theory?", *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*, 2014, p. 59-66, en ligne.
 - Moshe Koppel, Jonathan Schler et Shlomo Argamon, "Computational methods in authorship attribution", *Journal of the American Society for Information Science & Technology*, 60-1 (2009), p. 9-26.
 - *Digital Humanities for Literary Studies : Theories, Methods, and Practices*, et notamment l'article de M. Eder, J. Rybicki, M. Kestemont, "Measuring Style".

- Stefan Evert, Thomas Proisl, Fotis Jannidis, Isabella Reger, Steffen Pielström, Christof Schöch et Thorsten Vitt, "Understanding and explaining Delta measures for authorship attribution", *Digital Scholarship in the Humanities*, 32 (suppl_2[2017]), p. ii4–ii16.

Cours de spécialisation : Humanités numériques spatialisées

Enseignants : Carmen Brando (EHESS, CRH)

Mode d'évaluation : dossier à rendre

Volume horaire : 24h

ECTS : 5

Programme : Pour les inscriptions et plus d'informations, voir le site de l' atelier : <https://psigehess.hypotheses.org/1879>.

Les Systèmes d'Information Géographique (SIG) et leur développement en ligne (webmapping) apparaissent comme un moyen de faire travailler concrètement les chercheurs ensemble. Le lieu géographique fait le lien pour des chercheurs de différentes disciplines (géographes, archéologues, historiens, sociologues, anthropologues, informaticiens, etc.). Les SIG induisent la mise en place de nouvelles logiques de recherche. Pour être pleinement opérationnels, ils ne peuvent être dissociés des acteurs, qui dialoguent, négocient et adaptent des cultures et solutions techniques qui leur sont propres. L'examen des modalités de mise en œuvre de référentiels géo-historiques permet de mettre au jour des lignes directrices collectives et de préciser les éléments qui sont à la base de ces nouvelles pratiques de recherche et qui sont des expressions concrètes de ces nouvelles formes d'interdisciplinarités. Ce séminaire se propose d'être un lieu d'échange sur différentes expériences menées en géomatique à l'EHESS et dans d'autres institutions. Il sera associé à une formation initiale sur les SIG permettant d'acquérir les bases du maniement des outils géomatiques et de l'analyse des données cartographiques et de l'imagerie aérienne.

- Taylor Arnold, Lauren Tilton, *Humanities Data in R Exploring Networks, Geospatial Data, Images, and Text*, 2015 (Quantitative Methods in the Humanities and Social Sciences).
- Patricia Bordin, *SIG : Concepts, outils et données*, Paris, 2002.
- Philippe Rigaux, Michel Scholl, Agnès Voisard, *Spatial Databases : With Application to GIS*, San Francisco, etc., 2002.

Cours de spécialisation : Traitement automatique de la langue et analyse sémantique (ENS N104)

Enseignants : Thierry Poibeau, Jean Barré

Mode d'évaluation : devoir à rendre en rapport avec le contenu du cours (compte-rendu critique d'un article, petit programme centré sur une tâche de TAL comme la reconnaissance des entités nommées, etc.). La nature exacte du travail à rendre devra faire l'objet de l'accord préalable du responsable du cours.

Volume horaire : 24h

ECTS : 5

Programme : Le but de ce cours est de montrer en quoi le traitement automatique des langues peut être utile dans le cadre des humanités numériques. Les séances seront essentiellement pratiques : une large part sera consacrée à la présentation d'outils et à des manipulations simples par les participants, même si les bases théoriques du domaine ne seront pas oubliées. Le cours portera en premier lieu sur la manipulation de corpus (analyse, annotation, extraction d'information) à des fins d'analyse linguistique, littéraire ou de stylistique. Les données et les corpus considérés seront déterminés en fonction des objectifs de la séance et de l'intérêt des participants.

- Dan Jurafsky et James H. Martin. *Speech and Language Processing*. 3e édition en cours, disponible en téléchargement en ligne : <https://web.stanford.edu/~jurafsky/slp3/>. (C'est sans conteste la meilleure introduction au domaine, constamment mis à jour et très pédagogique, en accès libre de surcroît).

Mémoire – bilan d'étape

ECTS : 2

Évaluation de l'avancement du mémoire à l'issue du premier semestre 2, par les directeurs du mémoire. Validation du plan, des recherches bibliographiques et des approches expérimentales. Note sur 20.

Semestre 4 - Tronc commun, liste des cours

Algorithmique et programmation pour les humanités : Python avancé (ENC 2320)

Enseignant : Chahan Vidal-Gorène (ENC)

Mode d'évaluation : contrôle continu (1/4) et projet de groupe (3/4)

Volume horaire : 20h

ECTS : 4

Programme : Ce cours est consacré au renforcement des compétences en Python et en mathématiques, notamment en mathématiques appliquées à l'apprentissage machine (algèbre linéaire, mesures, métriques, clustering, ...). Ce cours est commun avec les M2. Le programme alterne une année sur deux.

Cours de spécialisation : Traitement automatique de l'image (EPHE P102)

Enseignants : Daniel Stockholm (EPHE)

Mode d'évaluation : à définir

Volume horaire : 24h

ECTS : 5

Programme : Ce module présentera des outils et notions utiles au traitement automatique de l'image, depuis des opérations élémentaires (échantillonnage et quantification, transformation en niveau de gris, binarisation), jusqu'à la détection de contours ou régions ou les opérateurs morphologiques. Chaque séance donnera lieu à un travail pratique : un exercice de programmation appliqué à l'image.

- Diane Lingrand, Introduction au traitement d'images, 2e édition, Paris, 2008.

Cours de spécialisation : Analyse de réseaux (ENC 2324)

Enseignants : Katarzyna Kapitan (ENC)

Mode d'évaluation : à définir

Volume horaire : 24h

ECTS : 5

Programme : Ce module présente les concepts et outils nécessaires à l'analyse de données relationnelles et de réseaux sociaux, dans des perspectives historiques, sociologiques, ethnologiques, etc.

Cours de spécialisation : Paléographie computationnelle

Enseignants : Chahan Vidal-Gorène, Peter Stokes, Dominique Stutzmann, intervenants invités

Mode d'évaluation : à définir

Volume horaire : 24h

ECTS : 5

Programme : Ce module présente les approches et méthodes en analyse des ressources historiques, notamment en computer vision pour l'analyse des écritures, la datation des documents, l'extraction de motifs, etc. Aucun pré-requis en paléographie n'est demandé, seulement les méthodes d'annotation, des rudiments en traitement de l'image et de solides connaissances en programmation python. **Lectures recommandées** : Publications publiées dans le cadre de l'International Conference on Document Analysis and Recognition (ICDAR) ces dernières années

Séminaire de recherche du master (ENC 2316)

Enseignants : Enseignants du Master

Mode d'évaluation : participation et discussion (M1); présentation de 15 minutes (M2)

Volume horaire : 5h

ECTS : 1

Programme : Ce séminaire, qui s'étend sur les quatre semestres de la formation, permet aux étudiants de première année de découvrir les recherches les plus actuelles dans le domaine des humanités numériques, notamment celles menées par des étudiants plus avancés. Aux étudiants de deuxième année, elle fournit également l'occasion de confronter leurs expériences et leurs idées, et de s'initier à la présentation de

leurs travaux.

Détails sur le mode d'évaluation : Pour les étudiants de M2, ils doivent une fois dans l'année réaliser une présentation de leurs recherches de 15 minutes; Pour les étudiants de M1, lightning talk sur leur mini-mémoire. **Éléments bibliographiques :**

- Les compte-rendus des années précédentes sont disponibles sur le carnet de recherche du séminaire : Master HN, Données et sciences humaines : Carnet de recherche du master Humanités numériques de l'Université Paris, Sciences & Lettres, dir. Jean-Baptiste Camps & Marie Puren & Chahan Vidal-Gorène, <https://humanodon.hypotheses.org/>, 2019-

Stage de découverte – 3 ECTS

Volume horaire : 4 à 8 semaines

ECTS : 3

Chaque étudiant devra, au cours de la deuxième année, réaliser un stage de découverte d'un laboratoire, d'une durée maximale de deux mois. Ce stage pourra avoir lieu soit en alternance durant l'année universitaire, soit en continu hors période des cours, et peut prendre place durant l'été précédant le début des cours ou suivant le rendu du mémoire. Ce stage peut se voir substituer une mobilité internationale ou un séjour de recherche à l'étranger.

Mémoire et soutenance

ECTS : 20 (18 + 2)

Mémoire de recherche :

- donnant lieu à un bilan d'étape à la fin du S3 semestre (2 ECTS) ;
- présenté devant la communauté des M1, M2 et enseignants dans le séminaire du master ;
- rendu avant le 1er juin (première session) ou le 1er septembre (deuxième session). Tout retard injustifié aura des conséquences.
- défendu dans une soutenance publique (rendu et soutenance, 18 ECTS).

Pour obtenir son diplôme, l'étudiant devra réaliser et soutenir un mémoire de recherche, mettant en œuvre ses compétences disciplinaires et numériques, et qui sera suivi par un directeur de recherche spécialiste du sujet et un spécialiste des humanités numériques. Le mémoire devra mettre les technologies numériques (données structurées, algorithmes, méthodes quantitatives, ...) au service de son sujet de recherche, et accompagner son mémoire de fichiers techniques.

Comme en M1, il ne faut pas hésiter à solliciter les conseils des encadrants : ici encore, la démarche de préparation du mémoire vous est propre, doit être faite en autonomie, et il n'appartient pas aux directeurs de recherche de venir vous chercher, mais à vous de les solliciter quand vous en éprouvez le besoin, ainsi que pour faire des points réguliers. Réaliser un mémoire peut (doit) être une entreprise passionnante, et c'est le travail où vous mettez le plus de vous-mêmes. Son succès dépend avant tout de votre investissement et de votre intérêt pour la problématique.

Entreprise de dimensions plus importantes que le mini-mémoire, ce mémoire de M2 doit capitaliser sur les expériences menées au cours du M1. Le travail qui lui est consacré ne doit pas se concentrer sur la fin de l'année, mais avoir démarré dès la rentrée, voire avant (pause estivale).

Pour les M2, le mémoire est l'aboutissement de vos études de master – voire le premier pas vers une thèse de doctorat – et il est également le travail dans lequel vous avez le plus d'autonomie et de liberté. Il importe néanmoins de faire un point régulier sur l'état d'avancement de vos travaux avec vos directeurs.

Les mémoires ne sont pas jugés sur la quantité de pages qu'ils contiennent, mais sur la qualité et rigueur de leur démarche scientifique et numérique, tant du point de vue des sujets abordés que de la méthode (données, outils d'analyse, etc.). En M2, 80 pages peuvent suffire, si elles sont claires, bien construites et fondées sur une démarche pertinente. Il s'agit là toutefois d'indications et non pas de règles absolues.

Le mémoire doit être rédigé dans une langue claire et rigoureuse, évitant les anglicismes ou le jargon. Il s'adresse à un public relativement large et a vocation à pouvoir être diffusé en ligne. À ce titre, d'une part, la qualité de son contenu vous engage autant que PSL et, d'autre part, le texte du mémoire doit

être clair et compréhensible, tant que faire se peut, par des chercheurs même non spécialistes du sujet traité. Les concepts, normes, outils... mentionnés devront être définis et présentés. Tous les tableaux et illustrations devront être légendés (et repris dans la table des illustrations).

Les mémoires de master 2 donnent lieu à soutenance publique. Le jury est composé de trois membres, dont le ou les directeurs de recherche et un président de jury. Durant cette soutenance, l'étudiant dispose de 15 minutes pour présenter sa démarche et ses recherches, les résultats et limites de celle-ci et les perspectives qu'elle ouvre, par un exposé oral aidé ou non d'un diaporama, et qui peut être étayé par une démonstration technique. Son exposé laisse ensuite la place à une discussion, incluant les remarques de chaque membre du jury et des questions. La durée maximale de la soutenance est d'1h30.