

Master Archives – Technologies numériques appliquées à l’histoire MAQUETTE du M2 (2023-2024)

		Heures	ECTS / modalités d'évaluation	Enseignant
UE 1 - Cycle de vie du document numérique		83	6	
<p>Cette UE porte sur le cycle de vie des documents numériques dans les archives et les bibliothèques : acquisition par numérisation ou collecte, description et signalement, pérennisation de l’information numérique et accès sur le web. Elle aborde sous forme de cours magistraux, de travaux dirigés et de visites en institutions patrimoniales les grandes catégories de documents primaires numérisés ou de documents nativement numériques (textes, images fixes et animées, sons, archives numériques, jeux vidéo, sites web) et leurs principales caractéristiques techniques. Sont abordés les principes, les techniques, les outils et les enjeux majeurs aux différentes étapes du cycle de vie des documents. L’UE contient en outre un module consacré à la description archivistique et un atelier de mise en pratique de l’archivage numérique.</p>				
(ENC 2550) La numérisation de l’image, du texte et de l’audiovisuel	La chaîne de numérisation, de la sélection au contrôle qualité ; aspects juridiques ; formats de numérisation des images et du texte ; référentiels de numérisation des images et de métadonnées ; OCR et HTR	12	Devoir sur table : étude de cas	Emmanuelle Bermès
	(à l’Ina) Présentation des missions de l’Ina ; spécificités et formats de numérisation de l’image animée et du son ; visite en ateliers			Emmanuelle Bermès + intervenants INA
(ENC 2551) Métadonnées, interopérabilité et accès aux documents numérisés	Enjeux du signalement et des métadonnées, historique de l’informatisation des catalogues et inventaires, transition bibliographique ; modèles conceptuels en bibliothèques (FRBR-LRM) et musées (Cidoc-CRM) ; formats de métadonnées et conversions, référentiels et identifiants, introduction aux API (OAI-PMH, SRU, IIIF)	14		Emmanuelle Bermès

(ENC 2552) Introduction à la pérennisation des documents numériques	La production administrative numérique ; les enjeux de la pérennisation de l'information numérique ; concepts, méthodes et normes ; le modèle OAIS ; les réalisations (Vitam et SPAR) ; les standards d'échanges pour l'archivage numérique SEDA et MEDONA ; Présentation d'outils pour la préparation de versements d'archives numériques : enjeux du traitement amont, démonstration des outils Archifiltre, Octave et Resip ; outils pour la création de profils d'archivage (Agape, Sherpa, Pastis)	10		Baptiste Nichele, Violette Lévy
(ENC 2553) La description archivistique	Principes de la description archivistique ; normes et formats pour la production de métadonnées archivistiques (ISAD(G), ISAAR-CPF, XML/EAD et EAC-CPF, RIC) ; mise en pratique : travaux d'encodage. 14h en 2023-24, 10h à partir de 2024-25 (une partie basculant en M1)	15	Travail individuel : production de notices	Jeanine Gaillard, Florence Clavaud
(ENC 1354 et 1355) Nouveaux objets patrimoniaux	Typologie des nouveaux objets patrimoniaux et processus de patrimonialisation ; histoire de l'informatique, des jeux vidéos et du dépôt légal numérique ; introduction au web et aux enjeux de son archivage ; travaux pratiques (BnF DataLab) NB : les étudiants ayant suivi cette option en 2022-23 en M1 ne suivront que 10h	20	Rédaction d'un dossier en binôme	Christophe Gauthier, Emmanuelle Bermès, intervenants BnF
(ENC 2454) Atelier : Les outils de l'archivage numérique	Mise en pratique de la préparation des données pour les versements d'archives numériques ; connaissance et utilisation de Vitam ; cas pratiques de l'accès aux archives numériques (à confirmer)	12	Sur assiduité	Emeline Levasseur, Julien Fennech, Pauline Charbonnier

UE 2 - Édition scientifique numérique	65	6	
<p>Cette UE apporte des savoir-faire dans le domaine de l'édition électronique de sources et de documents, de la modélisation à la mise en ligne. Elle porte sur l'apprentissage, via des travaux dirigés, de trois standards : XML/TEI pour l'édition critique des textes, LaTeX pour la composition de documents scientifiques (utilisable notamment pour le mémoire) et XSLT pour la manipulation et la publication de documents XML. Des outils et logiciels spécifiques à l'utilisation de ces formats et à la publication d'applications ou de produits éditoriaux sont mobilisés afin que les étudiants soient rapidement autonomes s'agissant d'appliquer ces outils dans le contexte d'un projet de recherche. Un atelier sur le langage de requête XQuery et l'utilisation de bases de données orientées document et de moteurs de recherche est proposé en option.</p>			

(ENC 2555) L'édition scientifique à l'ère numérique, la TEI	Concepts, normes, langages et outils logiciels permettant de structurer données et documents ; rappels sur XML et la modélisation XML (DTD, schémas) ; différents modes d'édition : critique, scientifique, paléographique ; encodage de documents en XML/TEI et utilisation de ODD	24	Contrôle de connaissances sur table + réalisation d'un encodage (individuel)	Ségolène Albouy
(ENC 2556) LaTeX et la rédaction de mémoire	Introduction au langage LaTeX ; structure générale des documents, mise en forme typographique, images, tableaux, manipulation et création de commandes et d'environnements... ; utilisation de LaTeX dans le domaine de l'édition scientifique ; méthodologie de rédaction du mémoire	8	réalisation d'un encodage (travail individuel)	Enimie Rouquette
(ENC 2557) Technique et chaîne de publication électronique	Transformation d'un document XML vers un autre document XML ou un format différent (HTML, LaTeX) en utilisant une feuille de style XSL ; navigation dans un arbre XML via XPath et ses principales fonctions ; manipulation des règles XSL basiques, des conditions et des boucles.	20	Contrôle de connaissances sur table + réalisation d'une transformation (individuel)	Jean-Damien Généro
(ENC 2558) Atelier : Interrogation des bases de données document, XQuery	Langage de requête XQuery ; bases de données document orientées XML ; BaseX et eXistDB ; gestion de corpus XML via leur exploitation à base de xQuery et de feuilles XSLT	12	sur assiduité	Rémy Delmotte

UE 3 – Traitement de la donnée	70	6	
---------------------------------------	-----------	----------	--

Cette UE apporte des connaissances et des savoir-faire sur la gestion des données dans le cadre d'un projet ou d'un système d'information. Elle apporte une définition de la notion de donnée et des exemples liés au contexte des sciences historiques et des institutions patrimoniales et aux enjeux actuels avec le développement de l'intelligence artificielle. Un accent particulier est mis sur l'apprentissage de la modélisation, incluant la modélisation conceptuelle et de processus avec UML, la modélisation de graphes dans le cadre du web sémantique et celle des bases de données relationnelles, ainsi que les langages de requête associés. Des outils de traitement de données (nettoyage, enrichissement, alignement) sont utilisés pour déboucher sur la réalisation d'une visualisation de données (datavisualisation). À l'issue de cette UE, les étudiant.e.s sont capables de mettre en œuvre le traitement d'un jeu de données sur l'ensemble de son cycle de vie : acquisition, modélisation, transformation, enrichissement, visualisation, analyse, dissémination.

(ENC 2560) Modélisation et requêtage	Introduction à la notion de données ; principes de la modélisation, initiation à UML, différence entre les grands types de modélisation de données : arbres, graphes, modèles relationnels ; introduction aux modèles conceptuels, logiques et physiques et aux bases de données ; introduction au web sémantique et requêtage SPARQL ; introduction aux bases de données relationnelles et à SQL	18	Travail de groupe : réalisation d'un traitement de données avec datavisualisation	Emmanuelle Bermès
(ENC 2561) Mise au point d'une chaîne de traitement de données, du requêtage à la datavisualisation	Massification des données (big data), fouille de texte et de données (TDM - text and data mining) ; traitement de données en masse et intelligence artificielle ; Nettoyage, enrichissement et alignement de données ; utilisation d'un outil de traitement de données (Dataiku) ; Principes de la visualisation de données ; réflexion autour des enjeux de la représentation graphique de la donnée ; exploration d'un outil de visualisation de données (Tableau) et présentation d'autres outils open source ou du marché ; accompagnement à la réalisation de datavisualisations dynamiques avec Tableau	20		Lauryne Lemosquet, Axel Roche-Dioré
(ENC 2562) Les bases de données relationnelles	Utilisation des bases de données relationnelles ; requêtage SQL (travaux pratiques et évaluation)	10	Contrôle continu sur table	Vincent Jolivet
(ENC 2351) Méthodes quantitatives pour l'historien	Cours commun avec AP1 et HN. Sciences historiques et données : analyse causale, systèmes complexes et modélisation mathématique, intelligence artificielle ; stylistique computationnelle et stylométrie ; lecture distante et modèles de sujet (topic modeling) ; théorie et analyse des réseaux. En 2023-24, cours suivi par les étudiants qui ont déjà suivi le cours NOP en M1	10	Partiel sur table	Florian Cafiero, Marie Puren
(ENC 2553) Atelier : apprentissage automatique appliqué à la culture et au patrimoine	Acquisition de données textuelles (OCR et HTR) ; analyse, traitement et génération d'images ; considérations réglementaires et éthiques	12	sur assiduité	Florian Cafiero

UE4 – Fondamentaux de l'informatique		53	4	
<p>Cette UE vise à permettre aux étudiants d'acquérir des compétences de gestion de projet informatique. Elle comprend des cours magistraux qui apportent des connaissances de base sur l'organisation d'un système d'information, son architecture et les différents rôles et processus en jeu. Un module est plus spécifiquement consacré à la maîtrise d'ouvrage et des exemples de projets sont présentés pour permettre aux étudiants de projeter des mises en situations des compétences acquises au cours de l'année. Des compétences techniques (bonnes pratiques de gestion d'un projet de développement, introduction à Linux et à la gestion de serveur) sont également développées.</p>				
(ENC 2566) Introduction à Linux	Rappel historique sur Unix, GNU, Linux ; utilisation du terminal et des lignes de commande ; arborescence d'un système Unix/Linux, création et manipulation de fichiers ; introduction au shell bash avec GameShell	9	Contrôle continu sur table	Nils Ratusznik
(ENC 2567) Travail collaboratif et bonnes pratiques de développement informatique avec Git	Notions de bases liées à Git ; méthodes de travail autour des interfaces disponibles (Gitlab, Github) ; versionning ; cycles de contrôles et de tests	8	Contrôle continu ou sur assiduité	Vincent Jolivet ou Victor Boby
ENC (2565) Le système d'information et ses enjeux	Définition du SI ; rôles, processus et méthodes ; gestion de la qualité ; architecture ; notions de responsabilité numérique (accessibilité, RGPD, sécurité, impact environnemental) ; Présentation d'exemples de projets par des intervenants externes	10	rédaction d'une note de projet en groupe (en lien avec l'UE3) +	Emmanuelle Bermès
(ENC 2568) Conduite de projet informatique	Conduite de projet ; méthode agile ; rédaction de cahier des charges ; pilotage d'un sous-traitant	14	rédaction d'un retour d'expérience individuel	Jean-Louis Foucard
(ENC 2569) Atelier : introduction à la mise en ligne d'application et à la gestion de serveur	Installation d'une machine virtuelle, d'un serveur web (Apache, Nginx), découverte de SSH, installation de MariaDb, PhPMyAdmin et Wordpress ; conteneurs (Docker) ; gestion des utilisateurs et sécurité (sudo, HTTPS)	12	sur assiduité	Nils Ratusznik

UE5 - Initiation à la programmation informatique et au développement web		72	6	
<p>Cette UE vise à transmettre à la fois un ensemble de compétences techniques et une compréhension des logiques d'un langage de programmation en se basant sur l'exemple de Python : contexte, syntaxe, principes de la programmation orientée objet, utilisation de bibliothèques spécifiques. Les étudiants acquerront des notions d'algorithmique et de langage de scripts basées sur la réalisation d'exercices en Python. A la suite de ce module, les étudiants seront capables d'utiliser Python pour développer une application Web et de comprendre les problématiques de développement liées à la mise à disposition et la consommation de données via des API.</p>				
(ENC 2570) Fondamentaux du développement web	Travaux pratiques d'initiation à HTML et CSS	6	contrôle continu	Chloé Pochon
(ENC 2571) Expressions régulières	Découverte par la pratique des expressions régulières, outil incontournable de l'ingénierie des données textuelles : syntaxe, implémentation ; cas pratiques.	6	sur assiduité	Enimie Rouquette
(ENC 2572) Introduction au langage de programmation Python	Types et opérations basiques ; boucles, conditions et fonctions ; lecture et modification de données et de fichiers (CSV, JSON) ; initiation au scraping, à l'utilisation d'API pour récupérer des données et à l'utilisation de référentiels.	16	contrôle continu sur table	Virgile Reignier
(ENC 2573) Développement d'application web avec Python	Développement d'application web en Python : installation et utilisation d'un module web (Flask) ; utilisation d'API et ORM avec SQLAlchemy	20	contrôle continu + travail de groupe : réalisation d'une application web	Maxime Challon
(ENC 2574) Atelier : développement d'application web avec HTML, CSS et Javascript	HTML, CSS, Javascript, Bootstrap ; principes clés et principales méthodes pour le web design et la conduite de projet web	12	sur assiduité	Chloé Pochon
(ENC 2575) Atelier : Python pour les données de recherche en histoire	Réalisation d'un outil en ligne de commande pour matcher des données tabulaires ; manipuler et visualiser des tables de données avec Pandas ; Introduction à la reconnaissance automatique d'entités nommées (REN) sur des documents historiques avec SpaCy	12	sur assiduité	Bertrand Duménieu

UE6 - Anglais		18	2	
Le but de cette UE est de faire en sorte que les étudiants soient capables de lire et comprendre la littérature en langue anglaise relative aux domaines des sciences de l'information abordés pendant l'année, et de communiquer en anglais par écrit (listes de discussion, articles) et oralement (conférences) sur ces questions. Un niveau moyen de maîtrise de l'anglais lu, écrit et parlé est un pré-requis indispensable.				
(ENC 2580) Anglais langue de l'informatique	Compréhension écrite et orale, expression écrite et orale ; étude d'une conférence TED ; module d'autoformation sur IIIF	18	réalisation d'un travail individuel avec présentation orale	Edward Gray
2^e semestre : stage professionnel			30	
Total heures maquette		362		
Total heures étudiant (après choix d'option)		314		
Total ECTS			60	