NTD212 - Les systèmes d'information documentaire. Niveau avancé : linked data et open data

Mis à jour le 16-09-2025

Code : NTD212

Unité d'enseignement de type cours

6 crédits

Volume horaire de référence (+/- 10%) : **50 heures**

Responsabilité nationale :
EPN15 - Stratégies / Gérald KEMBELLEC

Contact national :

EPN15 Information et documentation

2 rue Conté
31.3.12
75003 Paris
01 40 27 20 36
Jérôme Braemer
jerome.braemer@lecnam.net

Présentation

Prérequis

Cette UE, à distance, nécessite d'être inscrit avant son démarrage. Elle s'adresse prioritairement aux spécialistes de l'information (conservateur, documentaliste, bibliothécaire, secrétaire d'édition, secrétaire de rédaction, journaliste, veilleur, etc.) mais peut également concerner les professionnels proches de ces fonctions et désireux d'élargir leur champ d'action. Elle est intégrée au certificat CS22, "Les TIC dans les SID", formation à distance de conception et gestion de systèmes d'information documentaire.
Cette UE peut entrer dans le parcours d'une VAE.

Il est conseillé, outre une acculturation numérique (PIX ou c2i), d'avoir de bonnes notions de HTML et/ou d'avoir suivi l'UE NTD214.

Objectifs pédagogiques

Objectif : comprendre et utiliser la collecte et l'enrichissement des données pour l'environnement documentaire

Ce module introduit les classifications sur le web, la modélisation et la représentation des connaissances, les ontologies et le web sémantique, les modèles de métadonnées en gestion de l'information et les humanités numériques.

Il a pour objectif opérationnel de collecter et préparer les données pour enrichir les collections des humanités numériques et générer des visualisations sur des tableaux de bords ou insérables dans des rapports. Les domaines intéressés sont ceux qui construisent des connaissances ; le secteur culturel y est privilégié.

Programme

Contenu

MÉTHODE D'APPRENTISSAGE

L'étudiant(e) dispose d'un accès à une plate-forme d'enseignement où sont entreposés les cours, les exercices à réaliser en autonomie, le calendrier des visioconférences et des travaux à réaliser, les périodes de tutorat et de restitution des travaux.

PLAN DE TRAVAIL

5 ateliers à distance, 5 réunions de régulation à distance, 5 TP à réaliser

COURS en ligne

XML et les linked data

- Identification et adressage
- Typologie et qualité des identifiants
- Le web de données
- RDF
- De la notice aux métadonnées

Terminologies, web sémantique

- Les ressources sémantiques
- Evolution des travaux en ingénierie des connaissances
- Thesaurus vs Ontologie formelle : l'exemple du MeSH
- Le "cake" du web sémantique et le "cake" du web de données

Humanités numériques

- Exploration, interrogation, exploitation des données
- Processus d'enrichissement
- Création de nouveaux produits documentaires

Collecte et préparation des données ouvertes

- Les données ouvertes (*open data*) : définition, intégrité, consistance, authenticité
- Les outils et services de gestion de données
- Prise en main d'un outil opensource de collecte et de préparation de données

Datavisualisation

- Typologie des structures visuelles (quantitatives, temporelles, spatio-temporelles, spatiales)
- Les outils et services de gestion et de visualisation de données
- Prise en main de services en ligne de génération de visualisations

ATELIERS

- Open data
- Datavisualisation
- *linked data :* Collecte et préparation de données ouvertes
- Web sémantique (nécessite des connaissances HTML de base)
- Humanités numériques

Modalités de validation

- Contrôle continu
- Projet(s)

Description des modalités de validation

Un rendez-vous de travail hebdomadaire : ateliers à distance avec réunions de régulation à distance.
Travaux individuels et en groupe pour le projet