

Document structuré et écriture numérique

Infos pratiques

- > ECTS : 3.0
- > Nombre d'heures : 24.0
- > Période de l'année : Enseignement neuvième semestre
- > Méthodes d'enseignement : En présence
- > Forme d'enseignement : Cours magistral et Travaux dirigés
- > Ouvert aux étudiants en échange : Oui
- > Composante : Philo, Info-Comm, Langues, Littératures & Arts du spectacle

Présentation

Le cours portera sur l'écriture numérique dans le contexte de l'open data et de la science ouverte (open science) car de nos jours les publications (articles scientifiques, thèses, mémoires, rapports, littérature grise) embarquent des données issues de bases de connaissances, de bases de données, d'API, du Web sémantique. Dans ce contexte, les publications embarquent de nos jours non seulement du texte et des références bibliographiques, mais aussi des données (sérielles, documents, etc.) et des programmes informatiques qui traitent ces dernières. Quels sont les enjeux de ces nouvelles formes de publication ? Comment « écrire » des programmes dans un document ? Quels rôles jouent les vocabulaires documentaires, mais aussi les API et les SPARQL endpoint ? Quels sont les standards qui s'en dégagent ? Est-ce une nouvelle forme de publication ? Comment pérenniser ces documents ? Dans ce contexte l'utilisation des langages du Web, de XML à JSON, pour décrire des documents semi-structurés et des métadonnées nécessite généralement d'effectuer des transformations sur ces documents afin de les utiliser dans des systèmes d'information documentaires. L'utilisation des langages XSLT et XPATH, mais de plus en plus Python est utilisé afin de montrer le type de transformation qu'il est possible

de réaliser. Les langages RDF, RDFS et OWL (utilisés dans le web sémantique) sont présentés et mis en perspective du monde XML afin d'en montrer les principales finalités dans une perspective de construction de SI documentaire. Le cours est centré sur la réalisation de "mashups" web utilisant flux XML issus de requêtes en SPARQL du moteur de recherche ISIDORE (<https://isidore.science>). Il permet de voir les principaux aspects de la gestion d'un projet de développement de SI documentaire.

Objectifs

Le cours permettra d'acquérir :

- Les enjeux des données dans la publication scientifique et technique
- La conception de publications de données
- Le traitement de données
- La compréhension des vocabulaires documentaires structurés
- Les enjeux de l'interopérabilité des métadonnées et données

Évaluation

M3C en 2 sessions

- Régime standard session 1 – avec évaluation terminale (1 seule note) : **Un projet par groupe**
- Régime dérogatoire session 1 : **Un projet par groupe**
- Session 2 dite de rattrapage : **Un projet par groupe**

Compétences visées

Comprendre les principaux aspects de la gestion d'un projet de développement de SI documentaire.

Bibliographie

Bibliographie et ressources disponibles sur <http://blog.stephanepouyllau.org/1009>