Données semi-structurées

ECTS : 4

Enseignant responsable : Dario COLAZZO (https://dauphine.psl.eu/recherche/cvtheque/dario-colazzo)

Langue du cours : Français

Volume horaire : 45

Description du contenu de l'enseignement :

1. XML : Présentation de la norme XML et de ses principaux langages de manipulation (XPath, XSLT, XQuery) et de programmation (DOM, SAX).2. JSON : Définition des données en JSON, validation d'un document JSON, langages de requêtes pour JSON et différences avec XML et passage du XML au JSON.3. RDF : Modèle de données RDF, langage de description des vocabulaires RDF (RDFS) et langage de représentation des ontologies (OWL), interrogation de données RDF (SPARQL) et différences et liens avec XML (syntaxe RDF/XML) et JSON (JSON-LD).

Compétences à acquérir :

Former les étudiants aux normes du World Wide Web Consortium (W3C) pour modéliser et manipuler les données semi-structurées : XML (Extensible Markup Language) et JSON (JavaScript Object Notation) qui permettent de représenter des données textuelles (documents) ayant une structure potentiellement irrégulière, et RDF (Resource Description Framework) qui permet de décrire les ressources du Web sous la forme de graphe en les enrichissant avec de l'information sémantique. L'objectif du cours est d'apprendre aux étudiants comment représenter les données avec ces différentes normes, comment valider la représentation des données, ainsi que les transformer et les interroger. Chaque partie du cours fera l'objet d'un cours magistral et de plusieurs séances de TP.

Document susceptible de mise à jour - 04/11/2026
Université Paris Dauphine - PSL - Place du Maréchal de Lattre de Tassigny - 75775 PARIS Cedex 16