

# US333U - Interrogation et Manipulation de Données Semi-structurées

## Présentation

### Prérequis

Cette UE est ouverte aux auditeurs Cnam qui suivent le Master STIC mention Informatique spécialité Systèmes d'Information et de Décision.

Elle fait partie du cursus M2.

### Objectifs pédagogiques

Les données échangées sur le Web s'avèrent d'une nature radicalement différente de celles manipulées par les classiques Systèmes de Gestion de Bases de Données : il s'agit, pour la plupart, de données de type documentaire qui ont une certaine régularité mais dont la structure sous-jacente reste très limitée. La nécessité de développer des outils efficaces permettant d'exploiter de telles données et d'en extraire les informations pertinentes pour l'utilisateur a conduit à l'émergence d'un modèle, appelé modèle de Données Semi-Structurées, et à l'adoption du langage XML comme nouveau standard d'échange de données sur le Web.

L'objectif de ce cours est d'une part, de présenter ces nouveaux modèles de données et d'autre part, d'étudier les différentes problématiques posées par leur exploitation : définition de schémas, vérification de schémas, langages de requêtes, contrôle du changement ...

### Compétences

Maîtriser les données semi-structurées.

## Programme

### Contenu

Chaque problématique sera étudiée selon deux perspectives : d'une part, on présentera les nouveaux standards du WEB associés, tels que XQuery, XSLT pour les langages de requêtes ou XML Schéma comme exemple de systèmes de types, et d'autre part, on étudiera les objets théoriques sous-jacents à la définition et au fonctionnement des outils permettant de résoudre les problématiques étudiées (tels que les automates ou les transducteurs d'arbres).

### Description des modalités de validation

Examen

### Bibliographie

Titre					Auteur(s)
Extensible Markup Language (XML) 1.0.	http://www.w3.org/TR/1998/REC-xml-19980210				T. Bray, J. Paoli, and C. M. Sperberg-McQueen
Xml Path Language (xpath) 1999,	http://www.w3.org/TR/1998/REC-xml-19980210				James Clark
Xsl transformation (xslt) specification,	http://www.w3.org/TR/WD-xslt				1999, James Clark
Xml schema part1: Structures,	May	1999.			D. Beech, S. Lawrence, M. Maloney, N. Mendelson, and H. Thompson
http://www.w3.org/TR/xmlschema-1					
Regular Tree and Regular hedge languages over Unranked Alphabets, Theoretical Computer Science Center Research Report, TCSC-2001-05					Anne Brüggemann-Klein, M. Murata, D. Wood

Non valide depuis le 31-08-2019

**Code : US333U**

Unité spécifique de type cours

5 crédits

**Responsabilité nationale :**

EPN05 - Informatique / 1

**Contact national :**

Equipe pédagogique ingénierie des systèmes d'information

2D4P10, 33.1.13, 2 rue Conté  
75003 Paris

Sandra Bosse

[sandra.bosse@cnam.fr](mailto:sandra.bosse@cnam.fr)

---

The Generating Power of Total Deterministic Tree Transducers. Sebastian Maneth  
Information and Computation, 147:111-144, 1998.

---

Typechecking for XML Transformers, In Proceedings of the 19th ACM Symposium on Principles of Database Systems  
(PODS'2000), p. 11-22. ACM Press, 2000

Tova Milo, Dan Suciu,  
Victor Vianu