

Modélisation et traitement des données massives

CODE : 7T4/6-MTD

DOMAINE

- Bibliothèques
- Archivistique
- Veille et gestion de l'information
- Technologies de l'information
- Professions de l'information
- Compétences méthodologiques

NIVEAU

- Module obligatoire
- Module à choix

RESPONSABLE DU MODULE : Arnaud Gaudinat

Charge de travail	Crédits	Semestre
H	5 ECTS	4/6

Unité de cours et code	Nb heures de cours	Nb heures de travail personnel	Enseignant·e·s
Data skills en Python 7T4/6-MTD-1	44H	44H	Julien Gobeill
Modélisation des systèmes d'information 7T4/6-MTD-2	16H	16H	Patrick Ruch Emilie Pasche

RÈGLE DE VALIDATION DU MODULE

- Module validé si la moyenne des unités de cours est égale ou supérieure à 4
- Autre : « acquis » ou « non acquis »

EVALUATION DU MODULE

Évaluation commune

Évaluation par unité de cours

- Évaluation exprimée par une échelle de notes chiffrée de 1 à 6
- Évaluation exprimée par une appréciation « acquis » ou « non acquis »

PONDERATION DES UNITES DE COURS AU SEIN DU MODULE

7T4/6-MTD-1 : 70%

7T4/6-MTD-2 : 30%

COMPETENCES DEVELOPPEES

1. Décrire
2. Analyser – Synthétiser - Visualiser
3. Evaluer

OPTIONS

Module obligatoire pour obtenir un diplôme avec Option en Technologies de l'information.

Module à choix pour obtenir un diplôme avec Option en Archivistique ou Bibliothèques ou Veille et gestion de l'information.

ORGANISATION - DESCRIPTIFS PAR UNITE DE COURS

7T4/6-MTD-1 Data skills en Python

Ce cours vous plongera dans l'univers passionnant des données, et vous fournira des bases solides pour vous orienter vers des compétences de data engineer, librarian ou scientist.

Il s'adresse aux étudiants ayant suivi le cours Introduction à la programmation (en Python), et souhaitant approfondir leur formation. Vous acquerrez des compétences essentielles en matière d'importation, nettoyage et modelage de données, de gestion de bases de données avec des outils tels que MongoDB et Elasticsearch, ainsi que de maîtrise des expressions régulières pour extraire des informations pertinentes. De plus, vous serez initié au Machine Learning, vous permettant d'exploiter tout le potentiel des données pour des prédictions précises.

Objectifs d'apprentissage

- Importer, nettoyer, remodeler des données ;
- manipuler le texte avec les expressions régulières ;
- acquérir des données via du Web scraping ou des APIs ;
- gérer facilement des données massives dans MongoDB, déployer un moteur de recherche ElasticSearch ;
- découvrir les fondamentaux du Machine Learning, méthodologie et classification.

Contenu

- Types listes et dictionnaires ;
- manipulation de fichiers ;
- expressions régulières ;
- importation, nettoyage de remodelage de données ;
- acquisition de données en ligne (Web scraping et interrogation d'APIs) ;
- gestion des documents avec MongoDB et Elasticsearch ;
- initiation aux méthodes de Machine Learning et à la classification.

Enseignant : Julien Gobeill.

Méthodes d'enseignement

Chaque cours comporte un peu de théorie développée sur diapos, puis un TP qui comprend un récapitulatif théorique, puis des manipulations, et enfin des exercices à faire pour explorer le concept. Chaque TP contient une correction à la fin.

Format du cours

- Présentiel
- En ligne
- Hybride (cours en ligne transmis depuis une classe de la HEG)
- Flex (classe en ligne et/ou en présentiel selon un calendrier précis)

Salle

- Sèche
- Informatique
- Laboratoire

Groupe

- 1
- 2

Langue d'enseignement

- Français
 Anglais

Organisation temporelle

- Cours hebdomadaire de 2 périodes
 Cours bimensuel de 2 périodes
 Cours de 4 périodes
 Cours en bloc
 Autre :

Modalité d'évaluation :

- Contrôle(s) continu(s) : 100 % de l'unité de cours

- | | | |
|--|--|--|
| <input type="checkbox"/> Evaluation écrite sur table | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |
| <input type="checkbox"/> Evaluation écrite sur PC | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |
| <input checked="" type="checkbox"/> QCM sur PC | <input checked="" type="checkbox"/> Individuel (50%) | |
| <input type="checkbox"/> Travail à rendre | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |
| <input checked="" type="checkbox"/> Travail pratique | <input checked="" type="checkbox"/> Individuel (25%) | <input checked="" type="checkbox"/> Groupe (25%) |
| <input type="checkbox"/> Oral | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |
| <input type="checkbox"/> Autre : | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |

- Examen en session : 0 % de l'unité de cours

- | | | |
|---------------------------------------|-------------------------------------|---------------------------------|
| <input type="checkbox"/> Oral | | |
| <input type="checkbox"/> Ecrit | <input type="checkbox"/> Sur papier | <input type="checkbox"/> Sur PC |
| <input type="checkbox"/> QCM | <input type="checkbox"/> Sur papier | <input type="checkbox"/> Sur PC |
| <input type="checkbox"/> Ecrit et QCM | <input type="checkbox"/> Sur papier | <input type="checkbox"/> Sur PC |

Modalités de remédiation

- Remédiation possible si le résultat du module est compris entre 3.5 et 3.9/6
 Examen complémentaire - Travail additionnel
 Pas de remédiation (Unité de formation pratique)

7T4/6-MTD-2 Modélisation des systèmes d'information

Objectifs d'apprentissage

Ce cours fournira les bases de la représentation de systèmes d'information grâce aux modèles UML.

A la fin du cours, l'étudiant-e saura décrire un processus métier avec des diagrammes UML, utiliser un outil de modélisation, décrire ou lire des modèles dans le cadre d'un projet de développement IT.

Contenu

- Génie logiciel ;
- modèles UML
- diagramme UML (e.g. use case, activité, classes).

Enseignants : Patrick Ruch, Emilie Pasche.

Méthodes d'enseignement

Cours théorique accompagné de travaux pratiques.

Format du cours

- Présentiel
- En ligne
- Hybride (cours en ligne transmis depuis une classe de la HEG)
- Flex (classe en ligne et/ou en présentiel selon un calendrier précis)

Salle

- Sèche
- Informatique
- Laboratoire

Groupe

- 1
- 2

Langue d'enseignement

- Français
- Anglais

Organisation temporelle

- Cours hebdomadaire de 2 périodes
- Cours bimensuel de 2 périodes
- Cours de 4 périodes
- Cours en bloc
- Autre :

Modalité d'évaluation :

- Contrôle(s) continu(s) : 100% de l'unité de cours

- | | | |
|--|-------------------------------------|---------------------------------|
| <input type="checkbox"/> Evaluation écrite sur table | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |
| <input type="checkbox"/> Evaluation écrite sur PC | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |
| <input type="checkbox"/> QCM sur PC | <input type="checkbox"/> Individuel | |
| <input type="checkbox"/> Travail à rendre | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |
| <input type="checkbox"/> Travail pratique | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe |

Oral

Autre :

Individuel

Individuel

Groupe

Groupe

- **Examen en session : % de l'unité de cours**

Oral

Ecrit

QCM

Ecrit et QCM

Sur papier

Sur papier

Sur papier

Sur PC

Sur PC

Sur PC

Modalités de remédiation

Remédiation possible si le résultat du module est compris entre 3.5 et 3.9/6

Examen complémentaire - Travail additionnel

Pas de remédiation (Unité de formation pratique)

Descriptif validé par le responsable du module, le 27.11.2023

Descriptif validé par le responsable de la filière en Information Science, le 01.12.2023