

## Introduction à l'analyse de données massives

Groupe 50

Vendredi, de 9h00 à 12h00 PK-4370 (cours)

---

### Responsable(s) du cours

---

**Nom du coordonnateur :** MEURS, Marie-Jean**Nom de l'enseignant :** MEURS, Marie-Jean**Local :** PK-4935**Téléphone :** (514) 987-3000 #6139**Courriel :** meurs.marie-jean@uqam.ca

---

### Description du cours

---

L'objectif est d'offrir aux étudiants des cours à contenus variables et variés sur des sujets très actuels dans le domaine de l'informatique cognitive. Chaque fois d'un cours est offert, le contenu est nouveau et proposé par un professeur à partir de son domaine de recherche. Ce cours peut être offert sur un thème pour une durée de 45 heures ou sur deux ou trois thèmes en fractionnant différemment le temps.

---

### Objectifs du cours

---

Étude des enjeux et des défis liés à la collecte, au stockage et à l'analyse des données massives. Présentation des principaux algorithmes de fouille de données et d'apprentissage machine utiles à l'analyse des données massives. Contextes d'application et perspectives. Le cours mettra l'accent sur l'étude de cas utilisant des données issues de domaines variés (réseaux sociaux, internet des objets, sécurité, génomique, etc.).

Ce cours permettra aux étudiants de :

- se familiariser avec l'analyse de données massives,
- connaître les enjeux et les difficultés particulières au traitement des données massives,
- découvrir et comprendre les différentes approches algorithmiques permettant de traiter ces données,
- connaître les principaux outils logiciels du domaine et leur cadre utilisation,
- approfondir leur connaissance de l'état de l'art sur un thème d'application de leur choix (santé, sécurité, réseaux sociaux, génomique, marketing, etc.)

---

### Contenu du cours

---

- Introduction
- Collecte et stockage des données massives : bases de données non relationnelles, noSQL, Neo4j
- Manipulation des données massives : problèmes spécifiques, framework Apache : Hadoop, MapReduce, Spark, Storm...
- Algorithmes de fouille pour : les systèmes de recommandation, la recherche de motifs, la classification
- Données massives connectées : bases de données graphiques, analyse de liens (PageRank, moteurs de recherche, réseaux sociaux)

## Formules pédagogiques

- Une séance de 3 heures chaque semaine
- Cours magistral
- Présentations des étudiants : articles scientifiques, projets
- Présentations de chercheurs/sécialistes invités

## Modalités d'évaluation

Analyse et présentation d'articles scientifiques	30%
Dossier de recherche	30%
Quiz	10%
Examen	30%

L'analyse et la présentation d'articles scientifiques sont des travaux individuels. Le dossier de recherche peut être réalisé individuellement ou par équipe. Les étudiants devront constituer un dossier sur un thème de leur choix qui pourra être lié à leur sujet de thèse. Le sujet de chaque dossier sera défini en accord avec le professeur. Le dossier devra inclure une proposition de mise en oeuvre détaillée permettant d'adresser la question de recherche étudiée.

### Politique d'absence aux examens

**L'autorisation de reprendre un examen en cas d'absence est de caractère exceptionnel. Pour obtenir un tel privilège, l'étudiant-e doit avoir des motifs sérieux et bien justifiés.**

Il est de la responsabilité de l'étudiant-e de ne pas s'inscrire à des cours qui sont en conflit d'horaire, tant en ce qui concerne les séances de cours ou d'exercices que les examens. **De tels conflits d'horaire ne constituent pas un motif justifiant une demande d'examen de reprise.**

Dans le cas d'une absence pour raison médicale, l'étudiant-e doit joindre un certificat médical original et signé par le médecin décrivant la raison de l'absence à l'examen. Les dates d'invalidité doivent être clairement indiquées sur le certificat. Une vérification de la validité du certificat pourrait être faite. Dans le cas d'une absence pour une raison non médicale, l'étudiant-e doit fournir les documents originaux expliquant et justifiant l'absence à l'examen &ndash; par exemple, lettre de la Cour en cas de participation à un jury, copie du certificat de décès en cas de décès d'un proche, etc. Toute demande incomplète sera refusée. Si la direction du programme d'études de l'étudiant-e constate qu'un étudiant a un comportement récurrent d'absence aux examens, l'étudiant-e peut se voir refuser une reprise d'examen.

L'étudiant-e absent-e lors d'un examen doit, dans les cinq (5) jours ouvrables suivant la date de l'examen, présenter une demande de reprise en utilisant le formulaire prévu, disponible sur le site Web du département à l'adresse suivante : <http://info.uqam.ca/politiques/>

L'étudiant-e doit déposer le formulaire dûment complété au secrétariat de la direction de son programme d'études : PK-3150 pour les programmes de premier cycle, PK-4150 pour les programmes de cycles supérieurs. Pour plus de détails sur la politique d'absence aux examens du Département d'informatique, consultez le site web suivant : <http://info.uqam.ca/politiques>

## Intégrité académique

### PLAGIAT Règlement no 18 sur les infractions de nature académique. (extraits)

**Tout acte de plagiat, fraude, copiage, tricherie ou falsification de document commis par une étudiante, un étudiant, de même que toute participation à ces actes ou tentative de les commettre, à l'occasion d'un examen ou d'un travail faisant l'objet d'une évaluation ou dans toute autre circonstance, constituent une infraction au sens de ce règlement.**

La liste non limitative des infractions est définie comme suit :

- la substitution de personnes;
- l'utilisation totale ou partielle du texte d'autrui en la faisant passer pour sien ou sans indication de référence;
- la transmission d'un travail pour fins d'évaluation alors qu'il constitue essentiellement un travail qui a déjà été transmis pour fins d'évaluation académique à l'Université ou dans une autre institution d'enseignement, sauf avec l'accord préalable de l'enseignante, l'enseignant;
- l'obtention par vol, manoeuvre ou corruption de questions ou de réponses d'examen ou de tout autre document ou matériel non autorisés, ou encore d'une évaluation non méritée;
- la possession ou l'utilisation, avant ou pendant un examen, de tout document non autorisé;
- l'utilisation pendant un examen de la copie d'examen d'une autre personne;
- l'obtention de toute aide non autorisée, qu'elle soit collective ou individuelle;
- la falsification d'un document, notamment d'un document transmis par l'Université ou d'un document de l'Université transmis ou non à une tierce personne, quelles que soient les circonstances;
- la falsification de données de recherche dans un travail, notamment une thèse, un mémoire, un mémoire-crédation, un rapport de stage ou un rapport de recherche;
- Les sanctions liées à ces infractions sont précisées à l'article 3 du Règlement no 18.

Les règlements concernant le plagiat seront strictement appliqués. Pour plus de renseignements, veuillez consulter les sites suivants : <http://www.sciences.uqam.ca/etudiants/integrite-academique.html> et <http://www.bibliotheques.uqam.ca/recherche/plagiat/index.html>