

Fiche récapitulative

RCP216 | Ingénierie de la fouille et de la visualisation de données massives



51

Total d'heures d'enseignement



6

Crédits ECTS



15/09/2025

Début des cours prévu

Programme

1. Introduction : applications, typologie des données, typologie des problèmes
2. Approches : réduction de la complexité, distribution
3. Passage à l'échelle de quelques problèmes fréquents
 - a. Recherche par similarité, systèmes de recommandation
 - b. Classification automatique
 - c. Fouille de données textuelles
 - d. Fouille de flux de données
 - e. Apprentissage supervisé à large échelle
 - f. Fouille de graphes et réseaux sociaux
4. Visualisation d'information : historique, applications, outils
5. Enjeux perceptifs de la visualisation d'information : couleurs, formes, immersion, lecture
6. Techniques de représentations : graphes, hiérarchies, lignes de temps
7. Techniques d'interaction : association focus/contexte, distorsion, filtrage

Le cours est complété par des travaux pratiques (TP) permettant de mettre en pratique des techniques présentées. Pour la partie fouille de données, les TP seront réalisés à l'aide de Apache Spark. Pour le travail sur le projet, l'auditeur devra installer le logiciel Spark (gratuit) sur un ordinateur personnel de capacité suffisante, suivant les instructions disponibles en ligne.

Les supports de cours et de TP, ainsi que d'autres explications concernant le déroulement de l'UE sont accessibles à partir de

Objectifs : aptitudes et compétences

Objectifs :

Cet enseignement s'intéresse à l'impact des caractéristiques des données massives (volume, variété, vitesse) sur les méthodes de fouille de données. Sont examinées les approches actuelles qui permettent de faire passer à l'échelle les méthodes de fouille, en insistant sur les spécificités des opérations de fouille en environnement distribué.

Les caractéristiques mentionnées sont ensuite considérées de façon plus spécifique pour certains problèmes fréquents dans le traitement des données massives. Sont ainsi abordés les systèmes de recommandation et la recherche efficace par similarité, la classification automatique et l'apprentissage supervisé sur une plate-forme distribuée, les opérations spécifiques au traitement des données textuelles souvent hétérogènes, les implications de la vitesse sur la fouille de flux de données, l'analyse de grands graphes et de réseaux sociaux.

L'UE s'intéresse également au rôle de la visualisation et de l'interaction, non seulement dans la présentation des résultats mais aussi dans les opérations de fouille de données.

Compétences :

Capacité à mettre en oeuvre des techniques de fouille de données, de modélisation décisionnelle et de visualisation sur des données massives. Maîtrise de techniques adaptées à quelques problèmes fréquents rencontrés dans la fouille de données massives.

Prérequis

Bonnes connaissances mathématiques et statistiques générales, maîtrise de méthodes statistiques pour la fouille de données, connaissance de techniques de gestion de données massives faiblement structurées, connaissance de techniques de passage à l'échelle par distribution. Connaissance d'au moins un langage de programmation.

Vous êtes encouragés à évaluer votre capacité à suivre cette UE en répondant au questionnaire en ligne accessible sur Vous pouvez répondre sans vous identifier, le résultat vous est donné immédiatement et n'est pas enregistré.

Délais d'accès

Le délai d'accès à la formation correspond à la durée entre votre inscription et la date du premier cours de votre formation.

- UE du 1er semestre et UE annuelle : inscription entre mai et octobre
- UE du 2e semestre : inscription de mai jusqu'à mi-mars

Exemple : Je m'inscris le 21 juin à FPG003 (Projet personnel et professionnel : auto-orientation pédagogique). Le premier cours a lieu le 21 octobre. Le délai d'accès est donc de 4 mois.

Planning

Légende:

-  Cours en présentiel
-  Cours 100% à distance
-  Mixte: cours en présentiel et à distance

Modalités	Lieux	Disponibilités	Prochaines sessions *	Tarif indicatif
	En ligne	Semestre 1	15/09/2025	De 0 à 1.020 €
	En ligne	Semestre 1	Prévue en 2026-2027	De 0 à 1.020 €
	En ligne	Semestre 1	Prévue en 2027-2028	De 0 à 1.020 €

*Selon les UEs, il est possible de s'inscrire après le début des cours. Votre demande sera étudiée pour finaliser votre inscription.

Modalités

Modalités pédagogiques :

Pédagogie qui combine apports académiques, études de cas basées sur des pratiques professionnelles et expérience des élèves. Équipe pédagogique constituée pour partie de professionnels. Un espace numérique de formation (ENF) est utilisé tout au long du cursus.

Modalités de validation :

Note finale = ((note de projet + note d'examen) / 2).

Tarif

Mon employeur finance	1.020 €
Pôle Emploi finance	510 €
Je finance avec le co-financement Région	Salarié : 156 €
Je finance avec le co-financement Région	Demandeur d'emploi : 124,80 €

Plusieurs dispositifs de financement sont possibles en fonction de votre statut et peuvent financer jusqu'à 100% de votre formation.

- Salarié : Faites financer votre formation par votre employeur
- Demandeur d'emploi : Faites financer votre formation par Pôle emploi
- Votre formation est éligible au CPF ? Financez-la avec votre CPF

Si aucun dispositif de financement ne peut être mobilisé, nous proposons à l'élève une prise en charge partielle de la Région Nouvelle-Aquitaine avec un reste à charge. Ce reste à charge correspond au tarif réduit et est à destination des salariés ou demandeurs

d'emploi.

Pour plus de renseignements, consultez la page [Financer mon projet formationopen_in_new](#) ou contactez nos conseillers pour vous accompagner pas à pas dans vos démarches.

Cette unité d'enseignement n'est valorisable que dans cette certification.

Avis des auditeurs

Les dernières réponses à l'enquête d'appréciation de cet enseignement :

↓ [Fiche synthétique au format PDF](#)

Taux de réussite

Les dernières informations concernant le taux de réussite des unités d'enseignement composant les diplômes

↓ [Taux de réussite](#)