

## **Université Paris-Dauphine**

### **Département Mathématiques et Informatique de la Décision et des Organisations**



**Mention :** Informatique des Organisations

**Domaine :** Mathématiques et Informatique de la Décisions et des Organisations

**Responsable :** Jamal Atif, Professeur des Universités, 27<sup>ème</sup> section

e-mail : [jamal.atif@dauphine.fr](mailto:jamal.atif@dauphine.fr)

**Sécrétaire :** Nadine De Matteis

e-mail : [Nadine.deMATTEIS@dauphine.fr](mailto:Nadine.deMATTEIS@dauphine.fr)

#### **Objectifs du parcours**

Cette formation de deuxième année de Master a pour objectif de former des spécialistes de très haut niveau en Informatique Décisionnelle (**Business Intelligence**) en mettant l'accent sur les paradigmes récents des **Sciences des Données** en général et du **Big Data** en particulier, et des **Sciences de la Décision**. A cet effet, les étudiants reçoivent une formation approfondie en outils et paradigmes pour le big data, en datamining/machine learning, en entrepôts de données, en recherche opérationnelle, en aide à la décision multicritères et collective, mais aussi en conduite de projets BI en entreprise et en management et mesure de la performance. Les étudiants peuvent compléter leur formation par des cours optionnels proposés par le parcours ID ou dans les cours proposés par les parcours « Informatique pour la Finance » ou « Systèmes d'Informations et Technologies Nouvelles ».

Il s'agit donc de former des professionnels maîtrisant les outils les plus récents des technologies de Big Data et des Sciences des Données et de la Décision et pouvant accompagner la mutation du domaine de la *Business Intelligence* vers l'intégration de davantage d'éléments décisionnels fondés sur des analyses prédictives à partir de très grandes masses de données.

#### **Organisation**

Le parcours Informatique Décisionnelle ([ID](#)) appartient à une structure pédagogique appelée le pôle Info 3 qui regroupe également les parcours *Informatique pour la Finance* (MIAGE-IF) et *Systèmes d'information et Technologies Nouvelles* ([MIAGE-SITN](#)). Depuis sa création en 2000, ce pôle reçoit approximativement 600 demandes de candidature et diplôme une centaine d'étudiants. Chacune de ces spécialités de Master présente des enseignements techniques et scientifiques orientées vers un secteur professionnel bien identifié, ainsi que des enseignements d'ouverture. Ces secteurs déterminent les enseignements fondamentaux proposés dans chaque spécialité. Il s'agit respectivement :

- de l'informatique décisionnelle, pour la spécialité [ID](#) ([master-id@dauphine.fr](mailto:master-id@dauphine.fr));
- du développement des systèmes d'information et management des technologies nouvelles, pour la spécialité [MIAGE-SITN](#) ([master-MIAGE-SITN@dauphine.fr](mailto:master-MIAGE-SITN@dauphine.fr));
- de la conception, réalisation et maîtrise d'ouvrage d'applications dans le domaine financier pour la spécialité MIAGE-IF ([master-MIAGE-IF@dauphine.fr](mailto:master-MIAGE-IF@dauphine.fr));

La structure de ce pôle permet de mettre en commun les ressources pédagogiques et administratives communes à ces trois spécialités professionnelles. En particulier, du point de vue des étudiants, le pôle info 3 constitue une véritable bibliothèque de cours de haut niveau en informatique de gestion et informatique décisionnelle qui permet de se constituer un parcours à la carte en complément des enseignements obligatoires propres à une spécialité donnée.

Par ailleurs, la spécialité ID et les masters de recherche MODO et ISI ont en commun des thèmes (l'aide à la décision pour MODO et Intelligence Artificielle pour ISI), ce qui explique qu'elles partagent certains cours. Cependant, la spécialité ID est davantage orientée vers la mise en œuvre des outils et les aspects technologiques associés. À ce titre, elle partage l'ensemble des cours optionnels avec les spécialités MIAGE-SITN et MIAGE-IF.

### **Publics de la spécialité**

Cette spécialité s'adresse aussi bien à des informaticiens désireux d'acquérir une spécialisation dans le domaine de l'aide à la décision, qu'à des étudiants n'ayant que des connaissances de base en informatique qu'ils souhaitent développer vers le décisionnel. Par son orientation (double compétence informatique et aide à la décision), cette spécialité s'adresse à des publics variés :

- Etudiants ayant validé les 60 ECTS d'un M1 du Master Informatique des Organisations ou provenant d'autres Master de mention informatique ou Mathématiques,
- étudiants provenant d'un M1 « Mathématiques appliquées aux sciences sociales » ou équivalent,
- étudiants issus d'un M1 de Gestion ayant une « appétence » particulière pour l'informatique notamment au travers des systèmes d'information,
- étudiants sortant d'une école d'ingénieur,
- étudiants sortant d'une école de commerce.

### Formation continue et VAE

Cette spécialité est ouverte à la formation continue pour un nombre limité d'étudiants qui seront intégrés dans le groupe de formation initiale.

### Formation par apprentissage

Cette spécialité est ouverte en apprentissage pour un groupe d'étudiants qui suit un rythme d'alternance : 2-3 semaines à l'université / 4 semaines en entreprise.

**Double diplôme :** Cette spécialité accueillera à partir de l'année 2015-2016 les étudiants inscrits dans le double diplôme avec *Laurea Magistrale in Informatica per l'Economia e per l'Azienda* de l'Université de Pise.

### Organisation de la spécialité

Nous présentons ici la description des Unités d'Enseignements et les modalités de contrôle.

Les Unités d'Enseignements (UE) de la spécialité se décomposent en : 35 ECTS d'UE obligatoires, 9 ECTS de stage et 16 ECTS d'UE optionnelles. Les UE obligatoires se déclinent en deux groupes :

→ les UE relevant de la thématique **Business Intelligent/Systèmes d'Informations** sont :

- ID1 : Entrepôts de données, 33 h – 3 ECTS.
- ID2 : Le décisionnel en entreprise, Cas pratiques + Management et mesure de la performance, 36h – 4 ECTS.
- ID3 : Projet en Business Intelligence, 24h – 3 ECTS.
- ID4 : Data Mining/Machine Learning, 33 heures – 4 ECTS
- ID5 : Systèmes, Langages et Paradigmes pour le Big Data, 24h – 3 ECTS.
- ID6 : Visualisation des données, 24h - 3 ECTS
- L'un des quatre cours de SITN :
  - Management de projet informatique, 24h – 3 ECTS
  - Gouvernance, contrôle et risque des SI, 24h – 3 ECTS
  - Gestion de processus métier, 24h – 3 ECTS
  - ERP, 24h - 3 ECTS

→ Les UE relevant de la thématique **Aide à la Décision/Intelligence Artificielle** sont :

- ID7 : Modélisation et résolution de problèmes en Recherche Opérationnelle, 24 h – 3 ECTS
- ID8 : Systèmes intelligents d'aide à la décision collective et diagnostic, 24 h – 3 ECTS
- ID9 : Aide Multicritère à la décision, 24 h – 3 ECTS

Une UE de 48h d'Anglais (3 ECTS) est également obligatoire pour tous les parcours.

*Les UE optionnelles :* chaque étudiant doit choisir 16 ECTS (modulo un ou deux ECTS en plus en fonction du nombre d'ECTS des options) parmi :

- ID10 : Advanced Big Data Analytics, 24h – 3 ECTS

- ID11 : Gestion et interrogation des données persistantes, 24h- 3 ects
- Tous les UE du master SITN
- Tous les UE du master IF (exceptés les projets)

L'année est structurée en 3 sessions de 8 semaines. Chaque session termine par une semaine d'examen.

*Stage* : un stage de cinq mois minimum, donnant lieu à un rapport de stage et une soutenance, permet la validation de 9 ects

### **Module de « mise à niveau »**

Les cours de mise à niveau décrits ci-dessous sont proposés en début d'année, pendant 2 semaines, en préalable aux enseignements des spécialités. Ils ne donnent pas lieu à une notation et servent principalement à intégrer des étudiants externes à Dauphine qui n'auraient pas suivi le M1 et auraient de ce fait manqué des formations requises pour le M2.

<b>Enseignements du module « Mise à niveau »</b>		
Algorithmique et Java	Optionnel	14 h
Mathématiques pour la Finance	Optionnel	15 h
Bases de Données	Optionnel	10h
Recherche opérationnelle	Optionnel	15 h
UML	Optionnel	9 h
Statistiques	Optionnel	9h

### **Modalités de contrôle des connaissances**

Pour valider l'ensemble des 60 crédits ECTS, l'étudiant devra :

- obtenir une moyenne supérieure ou égale à 10 à l'ensemble des enseignements obligatoires, hors enseignements de parcours (pondérés par les ECTS),
- obtenir une moyenne supérieure ou égale à 10 à l'ensemble des enseignements de parcours et optionnels (pondérés par les ECTS),
- obtenir une note supérieure ou égale à 10 au stage,
- n'avoir aucune note inférieure à 6 sur les enseignements obligatoires ou optionnels.

L'attribution d'une mention Assez Bien, Bien ou Très Bien ne sera possible que si la moyenne de tous les enseignements, hors stage, pondérée par les ECTS, est supérieure

ou égale à 12. Dans ces conditions, si la moyenne générale de tous les enseignements et du stage, pondérée par les ECTS correspondants est :

- supérieure ou égale à 12 et inférieure à 14, l'étudiant pourra obtenir la mention Assez Bien
- supérieure ou égale à 14 et inférieure à 16, l'étudiant pourra obtenir la mention Bien
- supérieure ou égale à 16, l'étudiant pourra obtenir la mention Très Bien

### **Equipe pédagogique**

#### Intervenants universitaires

Jamal Atif – Professeur, Informatique  
Colazzo Dario – Professeur, Informatique  
Gabriel Virginie - Maître de conférences, Informatique  
Galand Lucie - Maître de conférences, Informatique  
Nègre Elsa - Maître de conférences, Informatique  
Pietriga Emmanuel – Chargé de recherche, INRIA, Informatique  
Öztürk Meltem – Maître de conférences, Informatique  
Gabriella Pigozzi - Maître de conférences, Informatique  
Drew Katrina, Rebecca LoxtonTaskin Saadet – Professeurs d'anglais.

#### Intervenants professionnels

Cohen-Solal Cyril – Directeur d'entité, Keyrus  
De Lassence Grégoire - Responsable Pédagogie et Recherche, SAS  
Pitel Guillaume – Président de Exensa  
Madar Sébastien – Directeur pole BI, MCNEXT  
Roman Carlo - Nell'armonia  
Thuleau Stéphane – Nell'armonia

### **Partenaires professionnels**

Keyrus, Sopra Group, CapGemini, SAP-Business Object, Edf, Novedia Decision, A.I.D, SAS, Nell'Armonia, SNCF, Eurodécision, ALTIC, MCNEXT

### **Débouchés**

Les débouchés de la spécialité sont principalement orientés vers :

- les sociétés de service en informatique notamment spécialisées dans le décisionnel;
- les sociétés de conseil et bureaux d'études ;
- les départements fonctionnels et d'études des entreprises et administrations.

De nombreuses propositions d'emplois en CDI provenant de SSII ou de plus grandes entreprises sont régulièrement communiquées au responsable, l'offre étant chaque année

supérieure au nombre d'étudiants diplômés dans cette spécialité. Les entreprises accueillant nos étudiants sont notamment :

- SAP-Business Objects, Business et Décision, Cap Gemini, Eurodecision,...
- Axa, Banque de France, BNP Paribas, Société Générale, Crédit Lyonnais, Barclays, Groupe Hervé, HSBC, CIC
- EDF, Renault, Air France, Orange, Cegetel, L'Oréal, Pierre & Vacances, Amadeus, Printemps, Thales

## **Contenu des enseignements :**

### **ID01-Entrepôts de données**

**Objectif :** Acquérir les notions de base relatives à l'acquisition, l'intégration, la modélisation et au traitement de données multidimensionnelles.

**Plan :** Introduction et définition d'un entrepôt de données, Architecture fonctionnelle, Modélisation conceptuelle, Alimentation, Stockage, gestion et exploitation de l'entrepôt. Utilisation de l'ETL Talend en TP.

TP1 : Introduction talend

TP2 : Business Modeler et construction des premiers jobs

TP3 : Découverte d'autres connecteurs talend

TP4 : Déploiement des jobs talend pour Mise en production

**Volume horaire** 12 h cours / 12 h TP

**Enseignant :** E. Negre, B. Delplace

**Crédits ECTS :** 3

**Référence :** "Data Warehouse Design: Modern Principles and Methodologies" de Mattaeo Golfarelli et Stefano Rizzi, 2009, Ed : Osborne/McGraw-Hill.

### **ID02- Le décisionnel en entreprise, études de cas**

**Partie 1 :** Applications des outils de l'informatique décisionnelle, gestion de projet décisionnel, outils de reporting, exploration de

Données : OLAP, MDX (tp)

#### **Objectifs :**

Donner une visibilité concrète de l'utilisation des applications décisionnelles en entreprise : Qu'apporte un logiciel décisionnel à un PDG, à un Directeur Marketing, à un Directeur commercial ? Quelles formes prennent les outils de l'informatique décisionnelle en entreprise ? Comment transformer un entrepôt de données en tableau de bord fonctionnel permettant d'expliquer le passé et d'anticiper l'avenir ? Ce cours vous donnera les réponses à ces questions à travers la présentation de projets concrets mis en place récemment dans de grands groupes.

**Contenu :**

- Acquérir les concepts de l'analyse décisionnelle.
- Comprendre l'apport d'un logiciel décisionnel aux différents acteurs de l'entreprise.
- Typologie d'applications
- Démarche et présentation de projets décisionnels concrets
- Initiation sur un outil décisionnel
- Mise en forme d'un reporting standard & création d'un cube d'analyse multidimensionnelle
- Réalisation d'un cas pratique

**Volume horaire** 12 h cours / 15 h TP

**Enseignants** : C. Cohen-Solal, S. Madar

**Crédits ECTS** : 3

**Pré-requis** :

**Partie 2 : Management et Mesure de la performance**

**Objectif** : Disposer d'une vue précise de l'Enterprise Performance Management, discipline intégrant, au sein d'un même projet, la mesure de la performance et l'informatique décisionnelle. Les études menées par les organismes indépendants considèrent ce secteur comme hyperporteur. Quels en sont les acteurs ? Quelles sont les briques logicielles concernées ? Quel positionnement dans la cartographie applicative des entreprises ?

**Plan** :

- Les fondamentaux du pilotage : finalités, domaines et départements concernés. Processus EPM. Fonctions et profils concernés.
- Description d'une solution type pour chaque processus
- Acteurs clé du marché de l'EPM : Editeurs et intégrateurs
- Description d'un projet EPM type
- Rôle du consultant EPM sur le projet

**Bibliographie** : pour une introduction, lire « l'informatique décisionnelle » SANDOVAL Librairie Eyrolles.

**Volume horaire** 12 h cours

**Enseignants** : S. Thuleau

**Crédits ECTS** 1

**ID03-Projet en BI**

**Objectifs** : L'objectif est de faire un projet décisionnel dans sa globalité, en partant du système opérationnel pour arriver au Reporting. A l'issue du projet, les étudiants devraient mieux conceptualiser les éléments suivants, car ils les auront pratiqués :

- o Différence entre ETL BI et ETL analytique
- o Différents Data Warehouse normalisé, dé-normalisé, OLAP, Virtuel et/ou en mémoire : avantages et inconvénients de chaque forme logique
- o Les différents cubes OLAP, ROLAP, MOLAP, HOLAP et In-Memory
- o Reporting Ad-hoc, Information Map, Reporting de masse, visualisation,



tableau de bord, procédures stockées

- Intégration des statistiques, du Data Mining, de la recherche opérationnelle et/ou de l'économétrie, dans la BI.

Après une présentation des concepts, l'énoncé du cas et une démonstration sur la plateforme, lors du premier cours ; les étudiants suivront tout d'abord un tutoriel de prise en main avant de travailler à la réalisation du cas complet.

**Volume horaire** 6 h cours / 18 h TP

**Enseignants** : G. de Lassence

**Crédits ECTS** 3

**Pré-requis** :

**Bibliographie :**

- \*\*\* DW 2.0 – The Architecture for the Next Generation of Data Warehousing – Editions Morgan Kaufmann – W.H. Inmon, Derek Strauss, Genia Neushloss – 2008
- \*\*\* Entrepôts de données – guide pratique de modélisation dimensionnelle – 2ième édition – Ralph Kimball et Margy Ross – Editions Vuibert – 2002
- \*\*\* Le datawarehouse – Guide de conduite de projet – Ralph Kimball, Laura Reeves, Margy Ross, Warren Thornthwaite – Editions EYROLLES – 1998
- \*\* Les stratégies absurdes : Comment faire pire en croyant faire mieux, Maya Beauvallet, Editions Seuil
- \*\* SAS. Introduction Pratique : Du Data Management Au Reporting, 2e édition - Sébastien Ringuedé - Edition Pearson
- \* Agile Data Warehouse Design: Collaborative Dimensional Modeling, from Whiteboard to Star Schema - Lawrence Corr - DecisionOne Press
- \* MDM - Enjeux et méthodes de la gestion des données Franck Régner-Pécastaing, Michel Gabassi, Jacques Finet Collection: InfoPro, Dunod/Logica – 2008

**ID04-Data Mining/Machine learning**

**Objectifs** : Il s'agit d'initier les étudiants à la pratique de l'extraction de connaissances à partir des grandes masses de données. Il sera illustré par des cas concrets en marketing et des exemples réalisés en session avec le logiciel R.

**Contenu** :

- Introduction
- Objectifs et panorama du datamining et du Machine learning
- Méthodes non supervisées :
  - Réduction de dimensionnalité



- Clustering :
  - K-means, CAH
  - Approches probabilistes : EM
  - Approches spectrales
  - Application à une segmentation marketing
  - Application au Text Mining
- Règles d'association
- Méthodes supervisées :
  - Régression logistique
  - Arbre de Décision
  - Méthodes à Noyaux
  - Approches neuronales
  - Application au scoring

**Enseignant** : Jamal Atif

**Crédits ECTS** 4

**Pré-requis** : Bonnes connaissances en statistiques et en algèbre linéaire

**Bibliographie :**

**Bishop, C. M. (2006). Pattern recognition and machine learning. springer.**

**Rajaraman, A., & Ullman, J. D. (2012). Mining of massive datasets (Vol. 77). Cambridge: Cambridge University Press.**

**ID05-Systèmes, Langages et Paradigmes pour le Big Data**

**Objectifs :**

Le cours a pour objectif d'apprendre aux étudiants les aspects fondamentaux des technologies Big Data pour la gestion et analyse de données massives. Le cours s'articule en trois parties.

Dans la première, l'accent est sur le paradigme MapReduce et le système Hadoop, avec un focus sur son système de fichiers HDFS. Le cours illustrera les mécanismes de base de Hadoop pour le support de l'exécution parallèle de *dataflow* MapReduce sur des clusters de machines. Une attention particulière sera donnée aux aspects algorithmique et d'optimisation de *dataflow* MapReduce.

La deuxième partie présentera des langages de requête et d'analyse de données caractérisés par des mécanismes de haut niveau et qui sont compilés sur MapReduce. Le focus sera sur les langages Pig Latin et Hive, des langages incluant des mécanismes à la SQL. Les techniques de compilation vers MapReduce seront présentées.

La troisième partie sera consacrée à des évolutions de Hadoop, et en particulier au système Spark et au langage supporté Scala. Le focus sera sur l'architecture de Spark, la notion de RDD, l'évaluation lazy de transformations et actions sur des collections distribuées RDD.

Les notions apprises seront mises en pratique dans un projet où les étudiants devront concevoir un dataflow pour l'analyse de grands volumes de données. L'implémentation sera faite tant en MapReduce que en Spark, et une analyse expérimentale sera effectuée pour comparer les performances des deux implémentations.

**Volume horaire** 24 h cours

**Enseignant** : D. Colazzo

**Crédits ECTS** 3

### **ID06-Visualisation des données**

**Objectifs** : Représenter les données de manière visuelle permet de tirer parti des capacités humaines en termes de perception et de cognition pour identifier des motifs particuliers, des éléments singuliers ou encore des anomalies dans ces données. La visualisation permet aux utilisateurs de systèmes d'information d'explorer leurs données de manière interactive, d'en avoir une vue d'ensemble, d'en extraire de l'information qui pourrait difficilement être obtenue à travers des processus d'analyse automatique (fouille, apprentissage), de formuler de nouvelles hypothèses qui pourront ensuite être vérifiées, par exemple, au moyen de tests statistiques. Au-delà de ces aspects exploratoires, la visualisation de données sert aussi d'aide à la prise de décisions, et de support à la communication entre individus des résultats d'analyses effectuées sur ces données. Le but de ce cours est de donner une vision d'ensemble du domaine, des principes fondamentaux de la perception visuelle humaine aux techniques de visualisation adaptées aux différentes structures de données (données multi-variées, arbres, graphes, séries temporelles, etc.). Les travaux pratiques seront réalisés au moyen de la bibliothèque D3 (Data-Driven Documents).

#### **Contenu :**

- Vue d'ensemble du domaine de la visualisation de données
- Perception visuelle
- Visualisation de données multi-variées
- Visualisation de structures arborescentes
- Visualisation de graphes et réseaux
- Visualisation de données temporelles
- Visualisation multi-échelle
- Systèmes et boîtes à outils pour la visualisation

**Volume horaire** : 12 h cours + 12 h TP

**Enseignant** : E. Pietriga

**Crédits ECTS** 3

Pré-requis : aucun

**Bibliographie :**

- B. Shneiderman, Catherine Plaisant, Designing the User Interface. Pearson Addison Wesley, 2005.
- E. Tufte, The Visual Display of Quantitative Information. Graphics Press, 1983.
- C. Ware, Information Visualization: Perception for Design. Elsevier, 2013.

**ID07-Modélisation et résolution de problèmes en Recherche Opérationnelle**

**Objectifs** : Le cours vise à présenter des modélisations originales de différents problèmes concrets de décision. Il s'agit de développer les aptitudes des étudiants à élaborer et mettre en œuvre des modèles pertinents face à une situation de décision. Puis, un ensemble de méthodes de résolution classiques sont présentées, leur efficacité est discutée.

**Contenu :**

- Concept de modèle en aide à la décision. Description du processus de modélisation et de ses différentes phases.
- Présentation de modélisations non triviales de problèmes de décision.
- Présentation d'outils de modélisation (modeleurs).
- Généralités sur les problèmes d'optimisation combinatoire
  1. Introduction (optimisation continue vs. optimisation combinatoire, exemples,...)
  2. Quelques méthodes heuristiques (algorithmes glouton, relaxation,...)
  3. Utilisation de solveurs : EXCEL et **GLPK**

**Volume horaire** 24 h cours

**Enseignant** : V. Gabrel

**Crédits ECTS** 3

**Pré-requis** : cours de Recherche Opérationnelle d'une licence et M1 informatique

**Bibliographie**

- H.P. Williams. Model building in mathematical programming. J. Wiley, New York, 1999. 4ème édition
- I. Charon, A. Germa, O. Hudry, Méthodes d'optimisation combinatoire, Masson, 1996
- J. Teghem, M. Pirlot, Optimisation approchée en recherche opérationnelle, Hermes, 2002

**ID08-Systèmes intelligents d'aide à la décision collective et diagnostic**

**Objectifs** : Les problèmes de décision impliquant plusieurs intervenants font l'objet de nombreux travaux, aussi bien en informatique (systèmes d'aide à la décision collective, systèmes multi-agents, systèmes répartis, systèmes de diagnostic...) qu'en économie et en

théorie de la décision (choix social, théorie des jeux...). Avec le développement d'applications réparties coopératives (nécessitant des mécanismes de décision collective, de vote ou de consensus) ou d'applications multi-utilisateurs pour le web (commerce électronique, partage de compétences...), le domaine de **la décision collective** est en plein essor. Aussi devient-il nécessaire de concevoir des systèmes efficaces de décision collective.

L'objectif du cours est l'étude de systèmes complexes de résolution de problèmes pour **l'aide à la décision de groupe** et plus précisément, l'étude de systèmes intelligents destinés à automatiser, à diagnostiquer ou à aider la prise de décision de groupe. Le cours vise à présenter différents modèles, outils et méthodologie pour le développement de tels systèmes. Le fil rouge de ce cours s'appuie sur le concept d'agent intelligent, et se positionne assez largement dans le domaine de l'intelligence artificielle distribuée. Les agents intelligents peuvent présenter une certaine autonomie, ont des croyances et des préférences qui leur sont propres, ainsi que des capacités de raisonnement, de communication, et de prise de décision.

**Contenu :** Parmi les thèmes étudiés, citons :

- Diagnostic
- Raisonnement et décision collective dans l'incertain et réseaux bayésiens
- Modèles de décisions collectives
- SMA: modèles d'agents et d'interaction
- Négociation entre agents
- Argumentation
- Enchères électroniques et détermination du gagnant

Les domaines d'applications privilégiés sont : les systèmes d'aide à la décision dans le transport (transport intelligent), les systèmes d'aide à la décision dans la gestion de l'environnement : aide aux pêcheurs, aide aux organisations territoriales (gestion de l'eau, gestion des crues des rivières), commerce électronique et négociation automatique.

**Volume horaire** 24 h cours

**Enseignant :** G. Pigozzi

**Crédits ECTS** 3

**Pré-requis :**

**Bibliographie :**

- **D. Bouyssou, T. Marchant, M. Pirlot, P. Perny, A. Tsoukiàs, et P. Vincke, Evaluation and Decision Models: A Critical Perspective, Kluwer, 2000**
- **A. Darwiche, Modeling and Reasoning with Bayesian Networks, Cambridge University Press, 2009**
- **R. Kishore, H. Zhang, et R. Ramesh, "Enterprise integration using the agent paradigm: foundations of multi-agent-based integrative business information systems", Decision Support Systems, 42: 48-78, 2006**
- **I. Rahwan, G. R. Simari, Argumentation in Artificial Intelligence, Springer, 2009**
- **M. Wooldridge, An Introduction to MultiAgent Systems, Wiley, 2009**

**ID09-Aide multicritère à la décision**

**Objectifs :** Ce cours vise à étudier des concepts fondamentaux en théorie de la décision, ainsi

que leur mise en oeuvre opérationnelle. Il s'articule autour de trois axes majeurs en décision : le choix social, la décision multicritère et la décision dans l'incertain. Plus précisément, ce cours a comme objectif de sensibiliser les étudiants d'une part à la modélisation de problèmes décisionnels dans lesquels différents points de vue sont pris en compte (tels que des critères, l'avis de différents agents ou des scénarii), et d'autre part à l'analyse de ces problèmes sous un angle opérationnel, permettant d'appréhender la difficulté algorithmique inhérente à ces modèles.

### **Contenu :**

Choix social : étude de méthodes de vote, axiomatisation, introduction au choix social computationnel

Aide multicritère à la décision : étude de quelques méthodes d'agrégation (modèles fondés sur un critère unique de synthèse, modèles fondés sur une relation de surclassement),

Introduction à l'optimisation combinatoire multiobjectif

Décision dans l'incertain : étude de quelques modèles en décision dans l'incertain et en décision dans le risque, théorie de l'utilité

**Volume horaire** 24 h cours

**Enseignant** : L. Galand & M. Ozturk

**Crédits ECTS** 3

**Pré-requis** :

**Bibliographie** :

D. Bouyssou, T. Marchant, M. Pirlot, P. Perny, A. Tsoukias, P. Vincke, Evaluation and decision models : a critical perspective. Kluwer Academic, Dordrecht, 2000.

B. Roy, Méthodologie multicritère d'aide à la décision, Economica, 1985.

P. Vincke, L'aide multicritère à la décision, Ellipses Marketing, 1998

## **ID10-Gestion et interrogation des données persistantes**

### **Contenu :**

1. Concepts fondamentaux : modèles de données et d'interrogation de données persistantes
  - Les modèles de données utilisés en gestion de données persistantes
  - Adaptation du modèle de données utilisé par un système au problème traité : données fortement ou faiblement structurées ; stockage document, n-uplet, colonne.
  - L'interrogation : bases de documents ; bases de données fortement structurées notamment OLTP et OLAP
  - Problème : optimisation de requête = minimisation du temps de réponse d'une requête
2. Supports de stockage des données.
  - Hiérarchies de mémoire, goulots d'étranglement
3. Optimisation de requête = optimisation locale
  - Organisation des données, structures d'index mono et multidimensionnels, index bitmaps
  - Vues et redondance pour l'optimisation de requête
  - Algorithmes d'évaluation d'opérations relationnelles unaires ou binaires standard s

- Algorithmes d'évaluation d'opérations spécifiques aux entrepôts de données
- Techniques d'optimisation de requête, incluant vues matérialisées, partitionnement, stockage par colonne

#### 4. Maintien de la cohérence de l'entrepôt

- Transactions.
- Chargement et rafraichissement d'un entrepôt : minimiser le temps de chargement
- Optimisation globale : requêtes et rafraichissement
- Pannes et reprises, incluant répliques.

**Volume horaire** 24 h cours

**Enseignant** : G. Jomier

**Crédits ECTS** 3

**Pré-requis** :

**Bibliographie** :

### ID10- Advanced Big Data Analytics

**Objectifs** : les problèmes liés au traitement des "mégadonnées" sont multifacettes et peuvent être adressés selon plusieurs angles :

- passage à l'échelle du stockage (distribution, redondance, sharding),
- passage à l'échelle du traitement (pensez à Map/Reduce) et des algorithmes,
- passage à l'échelle des algorithmes par approximation.

C'est ce dernier point qui sera étudié dans ce cours, dans lequel nous allons détailler toute une catégorie d'approches dont la caractéristique commune est de jouer sur l'imprécision pour gagner dans les deux dimensions habituellement antagonistes que sont le temps d'exécution et l'espace mémoire.

L'objectif du cours est de donner à l'étudiant accès à une palette d'outils qui viennent compléter les outils algorithmes classiques : structures de listes, ensembles, tas, tableaux associatifs, tris, etc.

Champs applicatifs : Big Data, Large Scale Machine Learning, Internet of Things

**Contenu** : Les techniques étudiées seront les suivantes

- Comptage approximatif (Morris et Flajolet)
- Comptage d'événements uniques (HyperLogLog)
- Appartenance à un ensemble (Bloom Filter)
- Comptage d'événements (Count-Min Sketch, Spectral Bloom Filter)
- Recherche de plus proche voisins (MinHash, Locality Sensitive Hashing)
- Statistiques (Q-digest, Random Stream Sampling)
- Réduction de dimension (Stochastic SVD, Random Projection)
- Stream Clustering

**Pré-requis** :

- Algorithmie avancée

- Programmation
- Bonne compréhension des techniques de hachage

**Volume horaire** 24 h cours

**Enseignant** : G.Pitel

**Crédits ECTS** 3

**Bibliographie** :

- **Mining of Massive Datasets** : <http://infolab.stanford.edu/~ullman/mmds/book.pdf>
- Flajolet, P., Fusy, É., Gandouet, O., & Meunier, F. (2008). **HyperLogLog: the analysis of a near-optimal cardinality estimation algorithm**. DMTCS Proceedings, (1)
- Cormode, G. (2009). **Count-min sketch**. In Encyclopedia of Database Systems (pp. 511-516). Springer US
- Piotr Indyk et Rajeev Motwani, « **Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality**. », Proceedings of 30th Symposium on Theory of Computing
- Shrivastava, N., Buragohain, C., Agrawal, D., & Suri, S. (2004, November). **Medians and beyond: new aggregation techniques for sensor networks**. In Proceedings of the 2nd international conference on Embedded networked sensor systems (pp. 239-249). ACM