

## 2ème année - MSc Data Science - Automne 2017/2018

FCO		Domaine commun aux filières 3A			
<b>UVMCSDSP01-Entreprise et informatique Décisionnell.</b>		Siegfried ROUVRAIS	<b>5 crédits</b>		
F3BP01A	Projet ISA Brest	Nicolas JULLIEN	126h	Fiche n°:1	
<b>UVMCSDS101-Statistique classique et inférence bayé...</b>		Yannis HARALAMBOUS	<b>5 crédits</b>		
F3B101E	Statistiques	Gilles COPPIN	21h	Fiche n°:2	
MSC-CSDS101A	Modèles graphiques probabilistes	Yannis HARALAMBOUS	20h	Fiche n°:3	
<b>UVMCSDS201-Fouilles de données</b>		Philippe LENCA	<b>5 crédits</b>		
F3B213A	Fouille de données	Philippe LENCA	63h	Fiche n°:4	
<b>UVMCSDS301-Ingénierie des connaissances et WEB sé</b>		Serge GARLATTI	<b>5 crédits</b>		
F2B206A	Ingénierie des connaissances	Serge GARLATTI	21h	Fiche n°:5	
F2B206B	Ingénierie des applications Internet	Serge GARLATTI	21h	Fiche n°:6	
F2B206C	Lectures & Interprétations des contenus : le web et les mutations dans le partage et la transmission des connaissances	Ioannis KANELLOS	18h	Fiche n°:7	
<b>UVMCSDS401-Systèmes d'information décisionnels</b>		Laurent BRISSON	<b>5 crédits</b>		
F3B415A	Business Intelligence	Laurent BRISSON	63h	Fiche n°:8	
<b>UVMCSDS501-Aide à la décision</b>		Patrick MEYER	<b>5 crédits</b>		
F3B502A	Aide à la décision	Patrick MEYER	63h	Fiche n°:9	



## UVMCSDSP01-Entreprise et informatique Décisionnelle

Responsable: Siegfried ROUVRAIS

### Modules de l'UV

F3BP01A      Projet ISA Brest

Nicolas JULLIEN

126h    Fiche n°:1

Fiche n°: 1

### Module F3BP01A : Projet ISA Brest

Responsable: Nicolas JULLIEN

Cécile BOTHOREL

Dernière mise à jour le: 06/03/17

#### Présentation

L'UV projet de 126 heures est consacrée à un seul projet. Il est généralement réalisé en groupes de 2 à 4 étudiants, et a pour objectif d'approfondir les connaissances acquises en cours, et en particulier en filière de troisième année.

Il existe trois types de projets, correspondant à chacun des parcours recommandés dans la filière ("système d'information", "systèmes d'information décisionnels" et "consulting").

Les élèves choisissent un des sujets proposés conjointement par un enseignant de Telecom Bretagne et un client, extérieur ou non. Ils élaborent et répondent à un cahier des charges en mettant en pratique diverses connaissances et compétences acquises dans la totalité du cursus, et en particulier dans les enseignements de la filière de 3ème année. Ils doivent notamment relier des connaissances appartenant à des domaines différents. Ils sont amenés, en fonction du projet choisi, à : réaliser des études documentaires ou bibliographiques, des études de terrain, des questionnaires qualitatifs et/ou quantitatifs, utiliser différentes méthodes d'analyse des données recueillies, construire des modèles mathématiques, modéliser des données etc.

**Volume horaire :** 126h

#### Organisation

##### Evaluation

Les étudiants seront évalués selon leur prestation au cours d'une soutenance finale, leur rapport écrit et l'appréciation des résultats fournis au client par l'encadrant du projet.

# UVMCSDS101-Statistique classique et inférence bayésienne

Responsable: Yannis HARALAMBOUS

## Modules de l'UV

F3B101E	Statistiques	Gilles COPPIN	21h	Fiche n°:2
MSC-CSDS101A	Modèles graphiques probabilistes	Yannis HARALAMBOUS	20h	Fiche n°:3

Fiche n°: 2

## Module F3B101E : Statistiques

Responsable: Gilles COPPIN  
Romain BILLOT

Dernière mise à jour le: 06/03/17

### Présentation

Qu'il s'agisse d'études marketing, d'analyse d'usages, de finance ou d'analyse d'image, il est usuel (et même systématique) de devoir traiter des ensembles de données importants. Les premières phases de ces traitements (regroupés classiquement sous le terme analyse exploratoire) permettent de visualiser les données sous différentes formes, de les projeter de façon optimale pour réduire le nombre de dimension, et globalement préparent la suite de la chaîne de traitement (classification, data mining). Grâce à ce module, vous pourrez donc - enfin - découvrir le monde merveilleux des boîtes à moustache et régresser autant que vous le souhaitez... Venez nombreux !

Les étudiants du module maîtriseront les premières étapes d'une analyse de données (analyse exploratoire) et les principes de tests d'hypothèses statistiques. Ces compétences seront instanciées sur la plate-forme R.

### Pré-requis

Les fondamentaux de la plate-forme R (syntaxe, structures de données, etc.) seront à charge des élèves qui disposeront d'un manuel et TP simple avec corrigé leur permettant de préparer les travaux encadrés dispensés dans le cours.

**Volume horaire : 21h**

### Contenu détaillé

- ¿ représentation de données / codage (3h)
- ¿ ¿ différents types de variables, effet sur la représentation
- ¿ ¿ distribution, histogrammes, boîtes à moustache, nuages de point
- ¿ ¿ analyse univariée
- ¿ ¿ coefficients de corrélation, notion de covariance, indices (3h)
- ¿ ¿ échantillonnage et tests d'hypothèse (3h)
- ¿ ¿ bureau d'étude test d'hypothèse (3h00)
- ¿ ¿ régression linéaire (3h)
- ¿ ¿ bureau d'étude régression linéaire (3h00)
- ¿ ¿ réduction dimensionnelle, analyse en composantes principales, analyse factorielle (3h)
- ¿ ¿ bureau d'étude analyse en composantes principales (3h)

## Module MSC-CSDS101A : Modèles graphiques probabilistes

Responsable: Yannis HARALAMBOUS

Dernière mise à jour le: 06/03/17

### Présentation

Les modèles graphiques probabilistes combinent les méthodes probabilistes (en particulier, l'approche bayésienne, les probabilités jointes et conditionnelles) et les méthodes de la théorie de graphe. On modélise une situation en associant à chaque variable aléatoire un sommet de graphe. Les arêtes du graphe peuvent, selon le cas, représenter une dépendance (au sens des probabilités conditionnelles), une causalité, ou tout simplement une influence réciproque. Le calcul de distribution jointe des variables est largement facilité par la structure du graphe. À partir des données obtenues par l'observation de certaines variables aléatoires du réseau, les modèles graphiques probabilistes nous permettent de faire des inférences sur les variables non-observées. D'autre part, en introduisant la notion d'utilité, on obtient des diagrammes d'influence qui permettent de prendre des décisions optimales.

### Pré-requis

Notions de probabilité et de théorie des graphes

**Volume horaire : 20h**

### Contenu détaillé

1. Raisonnement probabiliste

Rappels sur les probabilités, variables aléatoires, distributions, distribution jointe, marginalisation, formule de Bayes, indépendance, indépendance conditionnelle, inférence probabiliste

2. Rappels sur les graphes

Définitions, acyclicité, tri topologique, graphe orienté acyclique, couverture de Markov, clique, arbre couvrant, algorithme de Kruskal

3. Réseaux bayésiens

Réseau de croyances, exemple de l'arroseur, v-structure, transmission de l'information dans un réseau bayésien, d-connexion, d-séparation, équivalence markovienne, immoralité

4. Modèles graphiques probabilistes

Potentiel, réseaux markoviens, distribution de Gibbs, séparation dans un réseau markovien, propriétés de Markov globale et locale, champ aléatoire de Markov, théorème de Hammersley-Clifford, graphes de chaînes, graphes avec facteurs, I-map, D-map, map parfait

5. Inférence dans les arbres

Chaînes de Markov, passage de message, distribution stationnaire

6. Algorithme de l'arbre de jonction

Propagation de Hugin, consistance, propriété d'intersection courante, arbre de jonction, graphes triangulés, ordre d'élimination parfait, algorithme de triangulation de Tarjan, duc de Densmore, propagation de Shafer-Shenoy

7. Décision

Utilité espérée, arbre de décision, diagramme d'influence, potentiels de décision, de probabilité et d'utilité, arbre de jonction fort, processus de décision markovien, équation de Bellman, processus à horizon infini

### Organisation

#### Evaluation

Examen écrit de 2h, documents non autorisés

### Activités programmées

C1 (C)	1h30	Y.Haralambous Cours 1
PC/TP 1 (PC)	1h30	Y.Haralambous PC/TP 1
C2 (C)	1h30	Y.Haralambous Cours 2
PC/TP2 (PC)	1h30	Y.Haralambous PC/TP 2
C3 (C)	1h30	Y.Haralambous Cours 3
PC/TP3 (PC)	1h30	Y.Haralambous PC/TP 3
C4 (C)	1h30	Y.Haralambous Cours 4
PC/TP4 (PC)	1h30	Y.Haralambous PC/TP 4
C5 (C)	1h30	Y.Haralambous Cours 5
PC/TP5 (PC)	1h30	Y.Haralambous PC/TP 5
C6 (C)	1h30	Y.Haralambous Cours 6
PC/TP6 (PC)	1h30	Y.Haralambous PC/TP 6
Eval (CC)	2h	Y.Haralambous Évaluation

### Supports pédagogiques

Polycopié avec exercices

### Lectures recommandées

David Barber, Bayesian Reasoning and Machine Learning, Cambridge University Press, 2012

Daphné Koller and Mir Friedman, Probabilistic Graphical Models: Principles and Techniques, The MIT Press, 2009

## UVMCSDS201-Fouilles de données

Responsable: Philippe LENCA

### Modules de l'UV

F3B213A Fouille de données

Philippe LENCA

63h Fiche n°:4

## Module F3B213A : Fouille de données

Responsable: Philippe LENCA

Dernière mise à jour le: 06/03/17

### Présentation

L'informatique décisionnelle (en anglais : DSS pour Decision Support System ou encore BI pour Business Intelligence) désigne les moyens, les outils et les méthodes qui permettent de collecter, consolider, modéliser, restituer et explorer les données recueillies ponctuellement ou accumulées au fil du temps au sein d'une organisation en vue d'offrir une aide à la décision.

Cette aide à la décision peut avoir différents d'objectifs, par exemple :

- l'identification d'une population particulière (des patients présentant une certaine pathologie)
- la segmentation d'une population (identifier les usages qui sont faits d'un service comme un site web, déterminer des profils d'investisseurs)
- la description de comportement (déterminer des règles du style « les gens achetant de la bière achètent également des chips »)
- la prévision par régression (prévoir la consommation d'électricité, le montant moyen des achats réalisés par un client) ou par classification (prévoir la classe de risque d'un emprunteur)

Des données à la décision, le chemin est long : la mise en place d'un projet d'informatique décisionnelle nécessite de respecter un processus découpé en plusieurs étapes (récupération et préparation des données, constitution d'un entrepôt de données, fouille des données, évaluation des résultats, etc). Cette UV permet de présenter les différentes approches et méthodes qui forment le socle théorique de l'informatique décisionnelle. Après avoir suivi l'UV, l'étudiant disposera d'une variété de méthodes et d'outils qui doivent lui permettre de résoudre la plupart des problèmes relevant de l'analyse de données et de l'informatique décisionnelle.

La discipline connaît depuis plusieurs années une forte expansion. Face aux besoins qualitatifs des entreprises, à la variété des méthodes disponibles, à la diversité des données (voix, image, etc.) et au rythme d'augmentation de leur volume, le domaine draine une large palette d'emplois allant de l'ingénieur d'études ou du chercheur en décisionnel (qui participe à la conception et au développement de systèmes décisionnels), du consultant décisionnel (qui participe aux différentes phases d'un projet : spécifications, conception, modélisation, alimentation, restitution et maintenance) jusqu'au manager (qui contrôle la mise en œuvre des projets). Les emplois sont nombreux et variés tant chez les SSII que chez les éditeurs et les grands comptes. Les compétences scientifiques recherchées sont variées et fortement liées au processus : analyse de données, apprentissage et bases de données, mais aussi les compétences métier du client final.

Les différentes étapes d'un projet décisionnel seront présentées. Un accent particulier sera porté sur l'étape d'exploration des données (la fouille des données) pour des objectifs de classification, de description, de prédiction et de régression. On s'intéressera également à la validité des résultats et aux méthodes de comparaisons des algorithmes.

Les travaux pratiques permettront aux étudiants de s'approprier les différentes méthodes d'analyse et solutions logicielles. En particulier on présentera la méthodologie CRISP-DM, la solution SAS Entreprise Miner et la méthodologie SEMMA associée.

### Objectifs

- Connaître et comprendre les techniques de la fouille de données,
- Savoir appliquer les techniques de la fouille de données sur des problèmes réels à l'aide de logiciels de référence,
- Savoir proposer une solution argumentée,
- Savoir évaluer, comparer et combiner les techniques de la fouille de données

### Pré-requis

Eléments de statistiques, d'analyse de données et de bases de données.

## Liens

Aide à la décision  
Apprentissage  
Bases de données  
Recherche opérationnelle

**Volume horaire : 63h**

## Contenu détaillé

### 1) Cours

Introduction aux problématiques de la fouille de données  
Les étapes d'un projet décisionnel & les aspects méthodologiques  
Méthodologies CRISP-DM et SEMMA  
Méthodes d'exploration des données & qualité des données  
Méthodes de préparation des données  
Méthodes de modélisation : méthodes supervisées et non supervisées

- arbres de décision, règles d'association
- apprentissage statistique et bayésien
- apprentissage connexioniste
- support vector machine
- régression
- etc.

### Évaluation des modèles

- critères de performances et mesures de qualité
- mesures et procédures de validation
- tests statistiques, méthodes graphiques

### Combinaison de modèles

- Méthodes hybrides
- Méthodes ensemblistes & re-échantillonnage

### Fouille interactive de données

Morceaux choisis : text mining, web mining, bioinformatique, etc.

### 2) Bureaux d'études, travaux pratiques

- Maîtrise de la suite SAS, exemples pris dans le cours
- Approfondissement de Sas Entreprise Miner, exemples pris dans le cours
- Étude de cas & projet

## Organisation

### Evaluation

L'évaluation est réalisée sur la base d'un projet en petit groupe permettant sur des données réelles de mettre en œuvre l'intégralité du processus de fouille de données.

### Activités programmées

<b>C1 (C)</b>	3h	P.Lenca Introduction à la fouille de données (et à l'U.V.), données, méthodologie
<b>TP1 (TPE)</b>	3h	R.Billot P.Lenca S.Moga Démarche et méthodologie pour la fouille de données
<b>C2 (C)</b>	3h	R.Billot Apprentissage non supervisé (clustering)
<b>TP2 (TPE)</b>	3h	R.Billot P.Lenca S.Moga Clustering
<b>C3 (C)</b>	3h	P.Lenca Apprentissage non supervisé (motifs fréquents et règles d'association)

TP3 (TPE)	3h	R.Billot P.Lenca S.Moga Règles d'association
TDA1 (TPE)	3h	R.Billot P.Lenca S.Moga Initiation du projet
C4 (C)	3h	P.Lenca Apprentissage supervisé (arbres, comparaison classifieurs, ensembles)
TDA2 (TPE)	3h	R.Billot P.Lenca S.Moga Projet
TP4 (TPE)	3h	R.Billot P.Lenca S.Moga Arbres de décision, comparaison classifieurs, ensembles
C5 (C)	1h30	Y.Haralambous Apprentissage supervisé (support vector machine)
TP5 (TPE)	1h30	Y.Haralambous Support vector machine
C6 (C)	3h	P.Lenca Apprentissage supervisé/non supervisé (méthodes connexionnistes)
C7 (C)	3h	Y.Haralambous Text mining
C8 (C)	3h	Y.Haralambous Text mining
TP6 (TPE)	3h	Y.Haralambous Text mining
TDA3 (TPE)	3h	R.Billot P.Lenca S.Moga Projet
TDA4 (TPE)	3h	R.Billot P.Lenca S.Moga Projet
C9 (C)	3h	E.Gauthier Etude de cas de A à Z
TP10 (TPE)	3h	E.Gauthier Etude de cas de A à Z
TDA5 (TPE)	3h	R.Billot P.Lenca S.Moga Projet
Soutenances (1) (S)	3h	R.Billot P.Lenca S.Moga Soutenances projets (1)
Soutenances (2) (S)	3h	R.Billot P.Lenca S.Moga Soutenances projets (2)
Examen (CC)	3h	R.Billot P.Lenca S.Moga Examen

### Remarques

Pour les TPs prévoir 3 salles proches les unes des autres. Merci.

### Supports pédagogiques

Transparents de cours.

## UVMCSDS301-Ingénierie des connaissances et WEB sémantique

Responsable: Serge GARLATTI

### Modules de l'UV

F2B206A	Ingénierie des connaissances	Serge GARLATTI	21h	Fiche n°:5
F2B206B	Ingénierie des applications Internet	Serge GARLATTI	21h	Fiche n°:6
F2B206C	Lectures & Interprétations des contenus : le web et les mutations dans le partage et la transmission des connaissances	Ioannis KANELLOS	18h	Fiche n°:7

Fiche n°: 5

### Module F2B206A : Ingénierie des connaissances

Responsable: Serge GARLATTI

Dernière mise à jour le: 06/03/17

#### Présentation

L'intelligence artificielle propose un cadre formel et théorique pour le web sémantique. Cette dernière permet de représenter des connaissances et des savoir-faire en machine, mais aussi de raisonner sur celles-ci afin d'obtenir de nouvelles connaissances ou de nouvelles informations. La représentation des connaissances en intelligence Artificielle propose différents paradigmes. La logique et les représentations objets sont ceux qui sont actuellement les plus utilisés dans le cadre du web sémantique. Ces paradigmes amènent à différents langages de représentation de connaissances et mécanismes de raisonnement associés, dont Prolog fondé sur la logique, les logiques de descriptions et la F-Logic fondées sur les représentations orientées objet.

En se fondant sur l'Intelligence Artificielle, le web sémantique a défini des standards qui permettent de proposer de nouveaux services aux utilisateurs : une recherche d'information très précise et donc la réutilisation automatique (par la machine) des informations accessibles sur internet en fonction des besoins utilisateur (cf. <http://www.youtube.com/watch?v=off08As3siM&feature=related>). Ce type de service est impossible à réaliser par des moteurs de recherche de type Google qui délivrent trop de résultats dont la plupart ne sont pas pertinents, notamment car les mots clés utilisés peuvent avoir plusieurs sens (polysémiques). Ces Standards sont fondés sur les deux paradigmes suivants: la logique (SWRL) et les représentations orientés objet (RDF/RDFS, les ontologies et OWL).

Grâce à ces standards, on peut associer aux informations accessibles sur Internet des descriptions sémantiques interprétables par l'homme et par la machine. Ces descriptions sémantiques sont liées à des usages qui donnent du sens aux ressources pour une communauté de pratiques particulière. Les connaissances et savoir-faire explicites issus de ces communautés de pratiques permettent de déterminer le sens des informations. Ce sont donc ces connaissances et savoir-faire qui vont permettre de donner un sens unique aux informations et ainsi proposer de nouveaux services aux utilisateurs tant en entreprise que pour le grand public, notamment web social ou web 2.0.

#### Objectifs

- Être capable de faire des raisonnements sur des connaissances à l'aide des standards actuels du web sémantique (SPARQL, OWL, RDF, RDFS, etc.),
- Être capable de représenter des connaissances avec les paradigmes "logique" et "objets"

**Volume horaire** : 21h

#### Contenu détaillé

1. Introduction (1h30)

La notion de connaissances, d'interprétation et de sens

3. Représentation des connaissances

Les différents paradigmes, et ceux qui sont utilisés actuellement

3.1.1. La logique (3H00)

Le calcul des prédicats, théorie axiomatique, théorie des modèles

Prolog, exemples, démo.

### 3.2. Les représentations Objets (3H00)

Graphes conceptuels, les logiques de descriptions et la F-Logic

#### 3.2.1. Les Logiques de descriptions

#### 3.2.2. La F-Logic

### 4. Le Web 3.0 ou Web sémantique (7H30)

#### 4.1. Introduction/Enjeux

#### 4.2. L'architecture du Web Sémantique

#### 4.3. RDF/RDFS

#### 4.4. Les Ontologies et OWL

#### 4.5 Le langage de règles SWRL

#### 4.6. Les environnements gérants RDF/RDFS, OWL et SWRL

## Travaux personnels encadrés

Le cours est principalement organisé sous forme de résolution de problèmes réalisés en groupe et de synthèses sur les méthodes, standards et technologies abordées.

Ces problèmes ont pour objet d'appréhender les principaux enjeux du cours, de mettre en pratique les méthodes, les standards et les technologies proposées dans des cas concrets, mais aussi en utilisant des environnements du web 2.0.

## Organisation

### Evaluation

L'évaluation du module sera faite à partir des résolutions de problème réalisés en groupe, des rédactions correspondantes publiées sur un blog et de la conception d'une application concrète (réalisée individuellement) des méthodes et technologies abordées en cours. Cette application est laissée au choix de chaque élève.

## Module F2B206B : Ingénierie des applications Internet

Responsable: Serge GARLATTI

Dernière mise à jour le: 06/03/17

### Présentation

La conception de ces applications Internet pose aujourd'hui les principaux problèmes suivants : une augmentation de la complexité liée à la masse d'informations disponibles et à leur hétérogénéité qui pose des problèmes de recherche d'information et de réutilisation de ces dernières. L'accès aux informations se fait dans des contextes très différents, notamment caractérisés par la localisation, l'activité de utilisateurs et/ou des utilisateurs impliqués, le type de terminal (smartphone, Psp, Wii, Ipad, Tablet, ordinateur portable, télévision, etc.), le niveau de connaissances, la langue, etc.. En fonction de ces contextes, les informations pertinentes, leurs organisations et leurs présentations changent. Il devient nécessaire d'adapter les applications Internet à ces contextes d'utilisation. Il n'est en effet pas envisageable de construire une application par contexte et cela d'autant plus que ces contextes sont de plus en plus nombreux et variables.

Il est d'une part nécessaire de retrouver les informations pertinentes et d'autre part d'utiliser des méthodes facilitant la maintenance et la conception de ce type d'applications afin de pouvoir modifier dynamiquement les informations pertinentes, leurs organisations et leur présentation. En d'autres termes, il est nécessaire d'automatiser le plus possible la création de ces applications en les adaptant aux besoins utilisateur. Pour résoudre ces problèmes de recherche d'information et de conception, on utilise une approche orientée modèle qui facilite la conception et la maintenance. Cette dernière est fondée sur les standards du web sémantique et des usages dans une communauté de pratiques particulière qui déterminent le sens des informations et donc la manière de les indexer pour les retrouver et les réutiliser. Parmi ces modèles, on peut notamment citer : le modèle des contextes d'utilisation, le modèle de l'organisation et le modèle des présentations pour ensuite définir le modèle de l'adaptation. Tous ces modèles sont des connaissances acquises à partir des communautés de pratiques impliquées.

### Objectifs

- Etre capable de concevoir, de définir et d'utiliser les modèles sémantiques nécessaires à la conception des nouvelles applications Internet,
- Être capable de concevoir les nouvelles applications Internet

### Pré-requis

Le bloc "le web 3.0 et la représentation formelle de connaissance"

**Volume horaire : 21h**

### Contenu détaillé

- 1) Introduction (1h30)  
Problématique et enjeux des systèmes d'information sémantiques
- 2) Les Documents virtuels personnalisables (3h)  
Comment automatiser la génération de ces systèmes d'information
- 3) La recherche d'information et les métadonnées (1h30)
- 4) La modélisation utilisateur (3h)  
Comment modéliser les utilisateurs pour ensuite s'adapter à leurs besoins
- 5) L'adaptation/personnalisation (3h)  
Principes, techniques et méthodes pour l'adaptation
- 6) Etudes de cas (3h)

## **Travaux personnels encadrés**

Le cours est principalement organisé sous forme de résolution de problèmes réalisés en groupe et de synthèses sur les méthodes, standards et technologies abordées.

Ces problèmes ont pour objet d'appréhender les principaux enjeux du cours, de mettre en pratique les méthodes, les standards et les technologies proposées dans des cas concrets, mais aussi en utilisant des environnements du web 2.0.

## **Organisation**

### **Evaluation**

L'évaluation du module sera faite à partir des résolutions de problème réalisés en groupe, des rédactions correspondantes publiées sur un blog et de la conception d'une application concrète (réalisée individuellement) des méthodes et technologies abordées en cours. Cette application est laissée au choix de chaque élève.

## Module F2B206C : Lectures & Interprétations des contenus : le web et les mutations dans le partage et la transmission des connaissances

Responsable: Ioannis KANELLOS

Dernière mise à jour le: 06/03/17

### Présentation

Le terme « Société de l'Information » n'est pas qu'une invention linguistique pour parler des choses anciennes d'une manière moderne. Il désigne une réalité très concrète, à savoir le passage à une nouvelle façon pour penser le développement de notre monde, de nos sociétés en particulier, où la notion d'information joue un rôle prépondérant, plus que les secteurs traditionnels (matière première, transport, transformation, etc.). Il désigne, aussi, de nouveaux paradigmes économiques. Il embrasse, enfin, les modifications radicales que subissent déjà nos façons de penser et d'agir ; d'être, somme toute, et, plus précisément, d'être avec les autres. Cependant, le terme abrège une réalité bien plus complexe, qu'il conviendrait de mieux désigner par « Société de la Communication, de l'Information et de la Connaissance » : l'information concerne autant la couche physique (les systèmes et les réseaux de communication) que la couche informatique ou celle des contenus. Ce module concerne, spécifiquement, le rapport de l'information comme vecteur de transmission des cultures et des connaissances. La notion de connaissance sera envisagée du point de vue de l'homme en tant qu'utilisateur des machines au moyen desquelles il cherche à communiquer, à produire et à consommer, en quelque sorte, du sens. En effet, les machines sont envisagées comme des dispositifs de médiation dans toute opération impliquant l'information, depuis sa genèse jusqu'à sa valorisation, voire sa destruction.

Dans notre Société de l'Information, le rapport à la connaissance suit les mutations technologiques, en subit même les conséquences : nouveaux supports, nouvelles formes de production de documents, nouveaux documents aussi, en genre et en forme, exigent des stratégies de lecture adaptées ; nos interprétations et, à leur suite, nos compréhensions, en dépendent étroitement. D'autant plus que nos écologies ne sont plus celles d'hier : le virtuel ne vient pas seulement s'asseoir à côté du réel, mais compose avec lui des espaces hybrides qui deviennent de plus en plus des espaces de notre quotidien ordinaire.

Centré sur l'homme, ce module s'attachera à étudier les rémanences, les résistances, les pressions, les modifications mais aussi les opportunités impliqués dans nos efforts de domestiquer la nouvelle donne technologique qui sous-tend le concept d'information et en retracer les enjeux en matière d'appropriation et de transmission des connaissances.

L'objectif du module est de sensibiliser les élèves à la dimension « contenus et connaissances » de la Société de l'Information : à la spécificité des objets qui s'y rangent, aux modes de production et de consommation qu'ils impliquent, aux usages qu'ils convoquent, aux services qu'ils ouvrent. Plus avant, et plus spécifiquement, l'objectif est de montrer la viabilité de certains paradigmes anciens en matière de lecture, d'interprétation et de compréhension ainsi que d'explorer quelques voies que les nouvelles formes sémiotiques du document numérique apportent

### Objectifs

- Être capable d'adapter une lecture d'après le genre du document numérique qu'ils abordent
- Être capable d'envisager des moyens d'adaptation des connaissances véhiculées par un contenu à des profils d'utilisateurs différents
- Être capable de modéliser un contenu suivant plusieurs points de vues

### Pré-requis

Connaissances avérées d'un cursus normal en informatique.

Une sensibilité aux questions qui touchent à la cognition humaine reste souhaitable. Une connaissance de l'histoire de l'IA aiderait par ailleurs grandement à en saisir les enjeux de fond. Des connaissances complémentaires en psychologie cognitive, philosophie, linguistique, sémiologie, épistémologie ne sont pas nécessaires : le cours fournira ponctuellement le cadre conceptuel qui convient pour aborder les questions dans leur environnement interdisciplinaire à chaque fois.

## Liens

De manière naturelle, ce cours se lie aux autres parties qui constituent cette même UV (F2B206 A & B), où l'on trouvera, aussi, les arguments formels dans un espace d'application pérenne, précisément à travers les problématiques de représentation de connaissances que pose aujourd'hui le WEB 3.0.

**Volume horaire :** 18h

## Contenu détaillé

- Position du problème de communication entre les hommes et les machines. Le document numérique (DN) vecteur de nouvelles formes de communication. Forme et sens. Systèmes formels et sémantique.
- Critique du schéma classique de communication (émetteur/récepteur). Ouverture vers des schémas où l'interprétation joue un rôle fondamental dans la communication. Le problème de l'interprétation dans le cadre du DN.
- Lecture, interprétation, compréhension : stratégies, parcours, budget temps, univers d'attentes, plausibilités, régimes de validation. Isotopies, présomptions d'isotopie. Généralisation dans le cas des « rich media ».
- Contexte, objectifs, normes, valeurs. Sociabilité et échange. Le problème du genre du DN. Perception, stratégies de lecture. Interprétants et isotopies sémantiques, présomptions de sens.
- Calcul (machine) et sens (homme) : enjeux et problèmes. Limites. Critique rapide de l'histoire de l'IA et de ses promesses premières en matière de lecture et des connaissances.
- Compréhension et intention. Constructivisme sémantique. Co-construction du sens. La lecture comme action. Énaction.
- Approches centrées sur l'homme : coopération et interaction avec la machine. Départage des rôles. Pratiques et scénarios de coopération.
- Le rôle de la mémoire. Données, informations, connaissances. La nécessité des ontologies.
- Thèmes fondamentaux et paradigmes de traitement dans les échanges avec la machine :

## Travaux personnels encadrés

- Conception d'une brique d'un musée virtuel à profondeur variable et multi-points de vue

## Organisation

### Evaluation

Examen sur 2 heures (3 heures pour les non francophones) sur l'étude d'un cas.

## Supports pédagogiques

Articles et documents distribués ponctuellement.

## Lectures recommandées

- Delorme, A. & Fluckiger, M. (2003). Perception et réalité: une introduction à la psychologie des perceptions. De Boeck.
- Eco, U., Interprétation et surinterprétation. PUF, 1996.
- Imbert, M. (2006). Traité du cerveau. Paris: Odile Jacob.
- Kliekenberg, J-M. : Précis de sémiologie, Seuil, Points.
- Lindsay & Norman : Traitement de l'information et comportement humain. Interéditions, 1980.
- Nadel, J. & Decety, J. (2002). Imiter pour découvrir l'humain. Paris: Presses Universitaires de France.
- Rastier, F. : Sémantique Interprétative, Seuil, 1987 (2000).
- Turing, A. et Girard, J.-Y. : La machine de Turing, Seuil (Sources du Savoir), Paris 1995.
- Varela, F., Thompson, E., & Rosch, E. (1993). L'inscription corporelle de l'esprit. Sciences cognitives et expérience humaine. Paris: Seuil.
10. Les publications de l'initiative « Roger Pédaque », sur le document électronique : <http://rtp-doc.enssib.fr/sommaire.php3>

# UVMCSDS401-Systèmes d'information décisionnels

Responsable: Laurent BRISSON

## Modules de l'UV

F3B415A

Business Intelligence

Laurent BRISSON

63h Fiche n°:8

Fiche n°: 8

## Module F3B415A : Business Intelligence

Responsable: Laurent BRISSON

Dernière mise à jour le: 06/03/17

**Volume horaire** : 63h

### Contenu détaillé

Introduction (6h)

Enchaînement de 3 mini problèmes pour appréhender la notion d'APP :  
Modélisation de la base de données du SI opérationnel d'une entreprise  
Conception de tableaux de bords pour le reporting de son activité  
Proposition d'une architecture décisionnelle

Problème 1 : Conception de tableaux de bords (15h)

Atelier (3h) : Manipulation d'un tableau de bord sous tableau

Problème (9h)

Bilan et évaluation (3h)

Problème 2 : Modélisation de données décisionnelles (15h)

Atelier (3h) : Introduction à la modélisation dimensionnelle

Atelier (3h) : Introduction à SQL Server Analysis Services

Problème (6h)

Bilan et évaluation (3h)

Problème 3 : Extraction, transformation et chargement des données (12h)

Atelier (3h) : Introduction à l'ETL - Découverte de l'outil Talend Open Studio

Problème (6h)

Bilan et évaluation (3h)

Projet : Deuxième itération du projet (15h)

## UVMCSDS501-Aide à la décision

Responsable: Patrick MEYER

### Modules de l'UV

F3B502A Aide à la décision

Patrick MEYER

63h Fiche n°:9

Fiche n°: 9

### Module F3B502A : Aide à la décision

Responsable: Patrick MEYER

Dernière mise à jour le: 06/03/17

#### Présentation

Dans la vie quotidienne et celle de l'entreprise, on est souvent confronté à des situations qui nécessitent des prises de décision complexes. Très souvent, ces décisions portent sur des sujets pour lesquels on n'est pas forcément un expert du domaine.

De telles situations peuvent par exemple concerner:

- Le choix d'un emploi, le choix d'un ou de plusieurs candidats pour un poste dans une entreprise;
- La mise en place d'un système d'évaluation de la qualité de produits;
- Le choix d'un investissement, de la constitution d'un portefeuille de titres;
- La détermination de l'emplacement d'une usine, d'un aéroport, d'un centre de stockage;
- Des décisions dans des situations critiques (aéronautique, nucléaire, médicale, judiciaire, ...);
- ...

Le but de cette UV est de présenter un panorama des modèles mathématiques d'aide à la prise de décision.

Après une introduction aux différents sujets abordés dans l'UV, nous aborderons les théories d'aide à la décision faisant intervenir des objectifs multiples et souvent conflictuels. Ensuite nous nous concentrerons sur la décision dans l'incertain et les approches cognitives de la décision.

#### Objectifs

- Comprendre les méthodes d'aide à la décision en vue de les appliquer correctement dans des situations concrètes,
- Être capable d'identifier la technique d'aide à la décision à utiliser dans une situation pratique

**Volume horaire : 63h**

#### Contenu détaillé

- Introduction à l'Aide à la Décision
- La modélisation des préférences
- Notions fondamentales d'AMCD
- TP: introduction aux logiciels diviz et à XMCD
- Les méthodes à critère de synthèse unique (méthodes directes)
- Les méthodes à critère de synthèse unique (méthodes indirectes, élicitation de préférences)
- Les méthodes de surclassement 1 (Electre, PROMETHEE)
- Décision cognitive 1
- Aide à la décision experte en industrie

- Les méthodes de surclassement 2 (Rubis)
- Cas d'application d'un problème de choix (avec Rubis)
- Décision cognitive 2
- TP: Projet AMCD
- Décision dans le risque et l'incertain 1
- Décision cognitive 3
- Théorie des votes et agrégation des préférences
- Décision dans le risque et l'incertain 2
- Décision dans le risque et l'incertain 3
- Décision dans le risque et l'incertain 4
- Décision cognitive 4

## **Organisation**

### **Evaluation**

- Travail personnel (analyse d'un problème et sa résolution via les techniques d'Aide à la Décision Multicritère vues au cours) et soutenance;
- Examen oral individuel.

## **Supports pédagogiques**

Transparents vus aux cours.

## **Lectures recommandées**

P. Vincke, Multicriteria Decision Aid, John Wiley & Sons Ltd, ISBN-10: 0471931845, 10 march 1992