



Académie Tunisienne des Sciences,
des Lettres et des Arts
Beit al-Hikma

**Le Département des Sciences Naturelles et Mathématiques de
L'Académie Tunisienne des Sciences, des Lettres et des Arts
Organise :**
***Journée scientifique « Big Data & Data science »
Vendredi 28 octobre 2016***

Introduction :

Le champ de production des données numériques (systèmes d'information, sites web, smartphones, médias sociaux, objets connectés) n'a jamais été aussi immense. Aussi, la baisse des coûts de stockage aidant, les relations entre individus et systèmes, professionnelles et privées, n'ont jamais été autant enregistrées.

Le Big Data (la traduction française préfère les termes de « Données massives », ou « Méga-données ») est l'explosion de cette production des données numériques hétérogènes (multistuctures, multimodales, multiversions) en volumes tellement importants qu'il devient impossible de traiter avec les outils traditionnels de gestion de base de données.

Le problème n'est donc plus tant la collecte des données que leur sélection, leur manipulation, leur analyse, leur visualisation et leur valorisation dans tous les secteurs d'activité : industrie, agriculture, commerce, finance, administration, transport, santé, sécurité, éducation, environnement, recherche scientifique, etc... Selon diverses études, le Big Data s'apprête à représenter un business de plus de 50 millions de dollars en 2019, à créer plus de 5 millions d'emplois dans le monde d'ici trois ans et la pénurie de compétences en data scientists et data analysts serait le frein majeur à l'exploitation de cette source importante de valeur ajoutée.

Cette journée "Big Data & Data Science », organisée par le département des Sciences naturelles et mathématiques de Beit al-Hikma, s'adresse aux spécialistes du domaine comme aux néophytes. A travers la vision d'experts reconnus qui apporteront divers éclairages et les questions et les idées qui suivront, nous espérons contribuer utilement à la compréhension et à la réflexion sur les enjeux multiples, sociétaux, techniques, scientifiques de l'ère du "Big Data où la formation, l'interdisciplinarité et les interactions entre chercheurs sont tout autant importantes que les données elle mêmes.

Programme

9h00 – 9h20

Allocution d'ouverture et de bienvenue par Abdelmadjid Charfi, Président de l'Académie tunisienne des Sciences, des Lettres et des Arts (Beit al-Hikma).

Présentation de la journée par Kamel Barkaoui, Président du Comité d'Organisation.

9h20- 10h20

Big Data as a challenge for Data Science, Mustapha Lebbah

L'exposé sera consacré aux opportunités que le Big Data offre aux chercheurs et industriels qui s'intéressent à l'extraction de l'information à partir des données massives. L'exposé abordera aussi l'évolution du domaine.

Mustapha Lebbah est maître de conférences (HDR) et membre permanent de l'équipe "Apprentissage Artificiel et Applications" du laboratoire d'Informatique de l'Université de Paris Nord et du CNRS. Un de ces axes de recherche concerne le passage à large échelle en se basant sur la modélisation statistique et l'utilisation des paradigmes MapReduce ainsi que l'écosystème Big Data. Mustapha Lebbah est aussi secrétaire de la Société Francophone de Classification (SFC), et membre élu du bureau de l'association EGC (Extraction et Gestion des Connaissances) et du bureau du groupe "Data Mining et Apprentissage" de la SFdS (Société Française de Statistique)

10h20-11h20

Big Stat for Big Data, Christophe Biernacki

L'exposé abordera un des challenges actuels en statistique et en machine learning à savoir le développement de modèles de statistique usuels et des modèles complexes (Big Stat) adaptés au traitement de données massives.

Christophe Biernacki est Professeur de statistique au laboratoire de mathématiques Painlevé de L'Université Lille 1 et du CNRS. Il est Responsable scientifique de l'équipe de recherche MODAL chez Inria Lille. Sa recherche est focalisée sur la classification à base de modèles, en particulier dans le cadre des données dites complexes, mélangeant les variables de natures différentes (continues, qualitatives, fonctionnelles...) et aussi manquantes ou partiellement manquantes (intervalles). Il a participé au développement de plusieurs logiciels de ce domaine, par exemple le logiciel MIXMOD.

11h20-12h20

Réunion-débat : Data Science : L'Interdisciplinarité par excellence

Chair : Abdelhamid Hsaïri ; Animateurs : Christophe Bierbacki, Ali Frihida (ENIT), Mohamed Mohsen Gammoudi (ISAM) et Mustapha Lebbah

Internet en général et le web collaboratif en particulier qui offrent une multitude de pratiques d'échange de connaissances et de partage d'informations à faible coût et de façon quasi temps-réel, ont permis à toutes les disciplines scientifiques de disposer d'ensembles de données de plus en plus volumineux. Sans évacuer la question fondamentale de leur accessibilité (l'open access), l'analyse de ces données d'échelle et de nature différentes qui consiste à faire "parler ces données " en tirant du sens et établir des liens, implique plus que jamais la nécessité de faire collaborer les disciplines. D'où la question : Comment les avancées technologiques Big Data et Data Science permettent d'accompagner ces nouvelles démarches scientifiques et favoriser l'interdisciplinarité et la transdisciplinarité ?

12h20-14h00 Pause

14h00-14h45

Big Data & E-Commerce, Fatma Hamdi

L'exposé abordera la démarche scientifique et industrielle adoptée chez Sarenza pour améliorer le e-commerce en utilisant l'écosystème et les modèles statistiques.

Fatma Hamdi est Docteur en informatique. Spécialiste en Datamining et Apprentissage Statistique. Elle commence sa carrière en faisant de l'analyse comparative de moteurs statistiques chez GDF Suez, puis s'est occupée du programme de mise en place de moyens de détection de fraudes à l'octroi de crédits dans différents pays chez BNP Paribas Personal Finance. Actuellement elle est Responsable du Pôle DataScience chez Sarenza.

14h45-15h30

Big Data & Biology and Health, Mourad Elloumi

With the rise of the Next-Generation Sequencing (NGS) technologies, also known as high-throughput sequencing technologies, we are witnessing an exponential growth of both the volume and the complexity of biological databases. In this presentation, we will talk about some challenges related to the collection, storage and analysis of biological big data.

Mourad Elloumi is currently a Full Professor of Computer Science, Faculty of Economic Sciences and Management of Tunis, University of Tunis-El Manar, and Head of the BioInformatics Group (BIG) of the Laboratory of Technologies of Information and Communication and Electrical Engineering (LaTICE), National Higher School of Engineers of Tunis (ENSIT), University of Tunis, Tunisia.

15h30-16h : Pause

16h00-16h45

Big Data & Cybersécurité, Saïd Agrebi

L'ère du Big Data attire de nombreux cybercriminels dont l'activité va des attaques visant à paralyser le système d'information jusqu'au vol de données personnelles pour perpétrer des délits. Dans cette conférence nous allons voir les enjeux sociétaux et économiques du Big Data et de la Cybersécurité.

Saïd Agrebi est Fondateur & CIO de la start up MDM est « ChiefData Scientist officier », expert statisticien diplômé de l'école polytechnique et de l'UPMC. Il fut directeur de projet pendant 13 ans auprès de grandes comptes ALLIANZ, EGF-GDF, SG, L'OREAL, ORANGE, THALES, EADS...) dans des projets BI, statistique et Data Mining.

16h45-17h30

Big-Data : Les défis éthiques et juridiques, Sanaa Aït-Daou

L'exposé abordera les différents défis éthiques et juridiques liés à l'exploitation des données privées ou ouvertes (open data).

Sanaa Aït-Daoud est chercheur en sciences de gestion, membre associé de l'équipe Systèmes d'Information du laboratoire MRM - Université Montpellier. Elle travaille sur les problématiques de Management Responsable et Ethique des Systèmes et Technologies de l'Information.

17h30-18h30

Réunion-Débat : Formation en Big Data & Data Science

Chair : Souad Chouk ; Animateurs : Saïd Agrebi, Rim Faiez (IHEC), Faiez Gargouri (ISIMS) Amel Grissa-Touzi et Meriem Jaidane (ENIT),

Le Big Data entraîne de nouveaux usages, de nouveaux services dans le monde digital et l'écosystème de l'innovation. Cette évolution exige la mise en place de nouvelles formations pour l'acquisition de nouvelles compétences. Ainsi à titre d'exemple, l'étendue des données collectées a nécessité l'évolution de l'enseignement de l'apprentissage automatique et de la reconnaissance des formes en tirant profit du cloud computing et de la programmation parallèle pour réduire le temps d'exécution des algorithmes d'une part et de la taille des bases d'apprentissage pour développer de nouvelles méthodes d'analyse inaccessibles aux approches statistiques classiques telles que le deep learning ou la détection des écarts ou signaux faibles d'autre part. Quant aux métiers, ceux de data scientist et d'algorithmiste apparus récemment, ils sont déjà promis à un bel avenir.

18h30 Clôture de la journée

Pour participer à cette journée inscrivez- vous en cliquant sur le lien suivant :

<https://www.inscription-facile.com/form/07MbINxDvdhPzevPk0P8>