



**HAL**  
open science

## Data Mining et Big Data

Eric Rivals

► **To cite this version:**

Eric Rivals. Data Mining et Big Data. 15ème Rencontre Eveline Markiewicz sur les urgences et complications sévères chez le patient cancéreux, European Lung Cancer Working Party & Institut Jules Bordet, Nov 2015, Bruxelles, Belgique. lirmm-01286788

**HAL Id: lirmm-01286788**

**<https://hal-lirmm.ccsd.cnrs.fr/lirmm-01286788>**

Submitted on 11 Mar 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Data Mining et Big Data

Eric Rivals

14 novembre 2015

Résumé pour intervention aux Rencontres Eveline Markiewicz sur les Urgences et Complications Sévères chez Patient Cancéreux. Samedi 14 novembre 2015 à l'Institut Jules Bordet, Bruxelles.

## Adresse :

Laboratoire d'Informatique de Microélectronique et de Robotique de Montpellier (LIRMM)  
and Institute of Computational Biology (IBC)

CNRS and Université de Montpellier, France – <http://www.lirmm.fr/~rivals/>

De par son comportement, ses déplacements, ses communications, chaque individu laisse des traces digitales qui sont aujourd'hui de plus en plus souvent mémorisées par des systèmes informatiques. Cette masse d'information peu structurées, diverses, réparties géographiquement renferme dans les relations entre ces informations des connaissances sur ces individus. Explorer ces informations par traitement informatique pour en extraire ces connaissances est le rôle de la fouille de données – ou de l'extraction de connaissances. Les techniques utilisées sont pluridisciplinaires : statistique classique, analyse de données, et divers domaines de l'informatique tels que l'apprentissage automatique, les bases de données et l'algorithmique. Aujourd'hui les applications les plus visibles de la fouille de données concerne l'analyse des requêtes internet pour élaborer de la publicité ciblée ou l'analyse des comportements internet pour inférer le profil des visiteurs ou des clients à des fins de gestion ou de marketing. Clairement, ce type de méthodes et d'analyse s'appliquent au domaine médical et aux sciences de la vie. L'hétérogénéité des données et la complexité du domaine rendent leur application plus délicate. En revanche, l'intérêt et les avantages potentiels de ces applications sont énormes : potentiel en médecine personnalisée ou de précision, inférence de règles pour l'aide thérapeutique, économies dues à une meilleure prise en charge. Cette présentation contient une introduction à la fouille de données et sa variante concernant les très grands volumes de données, le « Big Data », liste quelques applications au domaine médical, et donne quelques intuitions sur les méthodes informatiques utilisées.

