

## Les Conférences Scientifiques de TELECOM ParisTech

Mercredi 6 février 18h30 en Amphi Emeraude

### Compresser pour comprendre : l'analyse des données de 1904 à 2008

Ludovic Lebart (CNRS et TELECOM-ParisTech)

Au début du vingtième siècle, des psychologues, en voulant mesurer l'intelligence, se sont heurtés au multidimensionnel en l'absence d'outils mathématiques adéquats et de moyens de calcul. Leur technique, l'analyse factorielle, a préfiguré néanmoins les outils modernes de l'analyse des données et du /data mining/. Les liens avec la statistique mathématique, avec les méthodes de compression dites structurales et la synergie avec les méthodes de classification automatique vont se tisser très laborieusement au cours de plusieurs décennies. Les cloisons disciplinaires, les différences de langues et de culture vont rendre cette évolution assez chaotique, et mettre en évidence les difficultés de la communication entre scientifiques. Le passage du numérique au qualitatif, puis aux textes a étendu le /data mining/ au /text mining/. Au début du vingt et unième siècle, les données sont pléthoriques mais les méthodes de validation statistiques restent balbutiantes et l'utilisateur, face à une riche panoplie d'outils, ne dispose toujours pas de vraie stratégie d'analyse.



Ludovic Lebart est Directeur de Recherche au CNRS. Ingénieur civil des Mines, Docteur en mathématiques, il est en poste au LTCI, UMR 5141 CNRS-TELECOM ParisTech, au département SES.

Quelques travaux récents de Ludovic Lebart :

- **Lebart L., Piron M. et Morineau A.,** (2005), Statistique Exploratoire Multidimensionnelle, 4ème Edition, Refondue., Paris, Dunod.
- **Lebart L., Piron M. et Steiner J. F.,** (2003), La Sémiométrie, Paris, France, Dunod.
- **Morin A., Bosc P., Hébrail G. et Lebart L.,** (2002), Bases de données et statistiques, Paris, France, Dunod.
- **Lebart L., Morineau A. et Piron M.,** (2002), Statistique Exploratoire Multidimensionnelle, Paris, France, Dunod.
- **Lebart L., Salem A. et Becue M.,** (2001), Analisis Estadístico de Textos, Lleida, Espagne, Milenio Editorial.
- **Lebart L.,** (2002), "Analyse statistique de textes: Quelques pistes de recherche", *Journal de la Société Française de Statistique*, Vol. 142, N. 4, pp. 7-30.
- **Lebart L.,** (2001), "Analyse de Contiguïté et Classification", *Revue MODULAD (INRIA)*, Vol. 27, N. 2, pp. 1-22.
- **Ohsumi N. et Lebart L.,** (2000), "Analyzing open-ended questions: some experimental results in Textual data analysis based on Info-Miner.", *Proc. of the Institute of Statistical Mathematics*, Vol. 48, N. 2, pp. 339-376.

TELECOM-ParisTech, 46 rue Barrault, 75013 Paris  
Métro Corvisard ou Cité Universitaire



