Percorso Autonomo Autorizzato

Title	**Big Data Analytics**
Coordinator	(Dip. Matematica, PoliMi): prof. Anna Maria Paganoni
Supporting coordinators	(DEIB, PoliMi): prof. Letizia Tanca; prof. Stefano Ceri; (Dip. Matematica, PoliMi): prof. Luca Formaggia, prof. Piercesare Secchi, ing. Chiara Montanari; (IBM): ing. Carla Milani, ing. Marco Monti
Scientific collaborations and partnerships	EPFL – Losanna (CH) MRU – Cambridge (UK) École Centrale Paris – Supélec (F)
Description and aims	Modern society is witnessing an increasing availability of data on different platforms - including the web, blogs and social networks, digital libraries - and of different forms such as the temporal profiles of clicks on the content of a web page, the calls by phones on a network of a given region, the satellite images for remote sensing, the biomedical diagnostics images, the medical records and so on. Many of these data are not fully structured, but contains valuable information for the knowledge discovery on the phenomena that govern communities, such as emerging opinions in social networks, buying behaviors of consumers, clinical and epidemiological practice, management and organization of complex systems. The traditional techniques of "knowledge discovery from data" are not always appropriate to the use of big data, which by nature are complex and multidimensional. Therefore it is essential to create a profession with solid skills in computer science, statistics and management, able to design, control, and conduct analysis on complex data. The track aims at training Date Scientists, able to use the analytical tools of research such as statistics, mathematical modeling and data mining, and, at the same time, able to develop capabilities in problem solving . Key features of a Data Scientist can be summarized in : ➢ solid technical skills in: statistics, computer science, data structured and unstructured data processing, data quality management and visualization techniques; ➢ knowledge of technologies and operating systems available for the data management and processing; ➢ knowledge of operational management processes;
Study Plan	The track is part of the PSPA (Major) "Statistics", introducing further aspects of computer science and management process. Students have to choose the following courses, in addition to those required in the Major of " Statistics" : 1. 095976- Advanced programming for scientific computing (sem II. 8 CFU), see Table SC 2. 10 CFU in computer science of which at least 5 CFU between: *089202 – Technologies for information systems* (ING-INF/05, sem. I, 5 CFU) *089183 – Data bases 2* (ING-INF/05, sem. I, 5 CFU) and 5 CFU between: *089165 – Computer security* (ING-INF/05, sem. II, 5 CFU) *089228 – Entreprise digital infrastructures* (ING-INF/05, sem. II, 5 CFU) *090950 – Distributed systems* (ING-INF/05, sem. I, 5 CFU) *090931 – Middleware technologies for distributed systems* (ING-INF/05, sem. I, 5 CFU) *093216 – Distributed software development* (ING-INF/05, sem. I, 5 CFU) *094743- Data management for the Web* (ING-INF/05, sem. I, 5 CFU)

	3. 10 CFU in management between: *096078 – Accounting, finance & control* (ING-IND/35, sem. I, 10 CFU) XXXX - *Multichannel Customer Strategy,* (ING-IND/35, sem. I, 10 CFU)
Past MSc theses	· Rachele Biasi (Ing. Mtm.), *Uso delle misure di profondità per dati funzionali multivariati nella previsione di patologie: un'applicazione ai segnali elettro-cardiografici*, 2013 · Ruben Binda (Ing. Mtm.), *A new approach to evaluate reputational loss: application to spread and oil and gas company*, 2013 · Mara Bernardi (Ing. Mtm.), *Registrazione di curve mediante l'algoritmo k-mean alignment ed estensione all'allineamento locale e di superfici*, 2013 · Claudia Belloni (Ing. Mtm.), *Sviluppo di modelli di stima e analisi del rischio reputazionale nel settore Oil&Gas*, 2013 · Emanuele Giani (Ing. Mtm.), *Metodi grafici ed inferenziali per l'identificazione di outliers: il caso del processo di cura di patologie cardiovascolari*, 2013 · Mirco Patriarca (Ing. Mtm.), *Development of the "fdakma" R package for the joint clustering and alignment of functional data with an application to neuronal spike trains data*, 2013 · Camilla Pezzotti (Ing. Mtm.), *Il processo di acquisto: studio del comportamento multicanale di un campione di consumatori*, 2013 · Nicholas Tarabelloni (Ing. Mtm.), *Metodi numerici e statistici per la simulazione e validazione di ECG*, 2013 · Nicholas Cazzaniga e Stefano D'Ettole (Ing. Mtm.), *Market and Social Media Reactions to Reputational Events in the Electric Sector* · Luca Ferrara (Ing. Mtm.), *Market Reactions to Reputational Events in the Oil&Gas Sector: Comparison between Companies* · Teresa Pietrabissa (Ing. Mtm.), *Ricostruzione di hazard functions per lo studio della progressione di malattia in pazienti affetti da scompenso cardiaco*
Work opportunities	Counselling, engineering companies, software development centers, companies operating in the medical , biomedical and pharmacological context, research and development centres in the area of transports, telecommunication, energy.