# DataLead 2014 à San Francisco

## Un projet, une réussite, une grande fierté pour notre Ecole

Guillaume Gaudron (2005), directeur des masters et du développement, ENSAE ParisTech


L'**Université de Berkeley** vue par Wikipédia : premier campus de l'université de Californie. Il est situé à Berkeley, en Californie, sur la rive est de la baie de San Francisco, donnant sur le Golden Gate. Sur le campus travaillent 33 000 étudiants et plus de 1 800 enseignants. Il s'agit du seul campus autorisé à se prévaloir indifféremment des noms : University of California, Cal ou Berkeley. Parmi les anciens étudiants, les professeurs ou les chercheurs associés à l'université, on recense 65 prix Nobel, 19 Oscars et 11 prix Pulitzer. Elle fait souvent partie des dix meilleures universités mondiales, et est considérée comme une des universités les plus prestigieuses.

Du 30 septembre au 2 octobre dernier, l'ENSAE a fait venir un peu de la *french touch* sur le campus de Berkeley en organisant conjointement avec la Haas Business School de « Cal » une conférence franco-américaine sur les données massives.

Près de soixante intervenants se sont ainsi succédé pendant trois jours pour expliquer en quoi les big data révolutionnent leurs activités, bousculent leurs stratégies, ouvrent des horizons nouveaux. Données majoritairement par des professionnels, les interventions ont été de grande qualité et couvert des champs thématiques comme la finance ou le marketing, exploré certaines innovations portées par des start up, abordé des enjeux de société comme la *privacy* des données ou la mutualisation en assurance face au risque d'ultra-segmentation. Parmi eux, la douzaine d'intervenants français a été particulièrement appréciée, apportant une touche de variété et d'ouverture qui allait bien au-delà du « *french accent* » qui a résonné dans les couloirs :

en France on maîtrise les data pour le jeu vidéo, l'assurance en ligne, la gestion moderne de la relation client dans la banque, le marketing et la publicité en temps réel, le *job market* via les réseaux sociaux, la rédaction automatique de notes stratégiques conjoncturelles, etc. Et sans rougir !!

La conférence s'est terminée par deux mini-cours très appréciés, donnés par les professeurs de l'ENSAE, Arnak Dalalyan et Manasa Patnam, respectivement sur l'apprentissage statistique et sur l'économie du numérique.

Pourquoi Berkeley, et pourquoi l'ENSAE ? Pour notre Ecole, il est évident qu'une telle collaboration a du sens pour la notoriété qu'elle apporte et pour la cohérence avec le développement à l'ENSAE des formations autour de la data science. Comme le rappelait récemment le directeur de l'Ecole, Julien Pouget, « cette conférence s'inscrit dans le cadre de la création du Mastère spécialisé et de la filière Data Science au sein de notre cycle d'ingénieurs ; elle vient concrétiser par ailleurs notre volonté de nous rapprocher davantage, de manière ciblée, de quelques établissements d'excellence internationaux ». Pari au départ, succès et fierté finalement, ce projet a réussi grâce à la conjonction de trois facteurs :

1. La réussite exceptionnelle d'une ancienne élève de l'école pendant son année passée à Berkeley, Aude Barthélémy, qui a mis en lumière la valeur des élèves et des enseignements de notre « petite » grande école et permis à Linda Kreitzman, responsable du Master of Financial Engineering de Haas, et Rodolphe Pauvert, responsable des relations internationales à l'ENSAE, de nouer une relation de confiance. Où qu'ils soient, nos élèves portent bien sûr leurs propres projets individuels, mais ils sont aussi les ambassadeurs de notre formation et des acteurs de notre avenir.

2. L'engagement rapide de l'ENSAE sur le terrain de la formation dans les Big data, à l'heure où des formations se mettent partout en place et n'existent souvent que partiellement.

3. L'émulation interne au sein même de l'UC Berkeley, entre des *schools* qui peuvent se sentir en concurrence sur le sujet des big data. Dans ce contexte l'ENSAE a été considérée comme une partenaire susceptible de renforcer la Business School.

Et la suite ? Pour les deux partenaires la collaboration doit se poursuivre, sous deux formes : tout d'abord, il s'agit de faire fructifier la marque DataLead en espérant qu'elle s'installera. Pour cela le directeur de l'ENSAE a annoncé qu'après San Francisco, Paris accueillera l'événement en 2015, probablement à l'automne. Ensuite, Berkeley et l'ENSAE poursuivent leurs discussions pour proposer une formation sous un sceau commun, dans laquelle les professeurs de l'ENSAE seront impliqués à part égale avec leurs homologues américains. Au-delà de ces annonces, nous espérons tous que d'autres initiatives naîtront des contacts noués.

La conférence aurait difficilement pu exister sans le soutien de nos sponsors français : AXA en premier lieu (sponsor « Platinum »), ainsi que la Société Générale, le CASD (filiale du GENES), le Labex Ecodec, et le Paris Saclay Center for Data Science. Merci à eux.

Elle n'aurait pas pu avoir lieu sans l'engagement de l'équipe transnationale qui a porté le projet : un grand merci à Linda, Charles, Aaron pour Berkeley, Rodolphe et Arnak pour l'ENSAE ! Cheers !!

Le programme de la conférence et les résumés des interventions sont consultables en ligne à l'adresse http://www.datalead2014.com/ ■

Rodolphe Pauvert
Linda Kreitzman
Julien Pouget

## Discussion with Stéphane Guinet

### CEO AXA Global Direct, member of AXA Group Executive Committee

**Variances -** *Why is AXA the main sponsor of the Conference DataLead? Could you outline your key issues of this conference?*

**Stéphane Guinet -** Leading the space in Big Data is an absolute key priority for AXA and several initiatives have been launched across the Group in the recent years. One of them is around increasing our proximity with the Academic world and partner with world-class Universities; Berkeley, as ENSAE ParisTech, is certainly one of them as they are at the forefront of the latest research in Big Data applied to the Financial Industry. They now also have a dedicated master for Big Data, like many other universities at once in the US, Europe and Asia.

It is also a way to connect with leading Professors but also students who may become future AXA employees. I was there with colleagues from the Data Innovation Lab, from the AXA Lab in San Francisco, from ABS (AXA Business Services in India) and from AXA Rosenberg. And I enjoyed it.

What I could really see is that Big Data has now come from theory to real tangible Use cases and implementations all across the entire value chain: Marketing, Product, Pricing, Claims, and Operations. Even the HR space is profoundly impacted by players like LinkedIn which is just and simply a specialist Big Data company, applied to Human Capital.

I also realized that the Capital markets/Asset Managers have been investing in that field for now quite a long time (High Frequency Trading…).

**V. -** *Which were your main messages as keynote speaker?*

**S.G. -** Big Data originally described the practice in the consumer Internet Industry of applying algorithms to increasingly large amounts of data to solve business problems that had suboptimal solutions with smaller data sets.

Many features, signals, patterns can indeed only be observed by collecting and processing massive amounts of data – this was difficult, or even impossible a couple of years ago but the advent of technologies like MapReduce or Hadoop now make it easier and hence possible.

The Insurance Industry has always been a data-driven Industry. We all use "small" data (i.e. a rather limited number of data points) to assess risks, to understand customer needs, to improve customer satisfaction and retention. We now have to do it much more massively and systematically. More than ever we need to become a Data-Driven Company.

In that respect, I believe that the upcoming revolution is a huge challenge but also a once-in-a-life opportunity if we do it properly.

Challenges for AXA are obviously on building capabilities and expertise but probably even more in the cultural shift – this is not an evolution but a true revolution in the way we approach business problems – innovation spirit, "exploration", test and learn mode, acceptance of failure - and I believe that all leaders at AXA should put Big Data at the very top of their agenda and align their priorities and organization accordingly.

Finally, I insisted on the Privacy and Trust aspect. Those are fundamental questions that need to be addressed even if I felt that this seemed to be slightly less of an issue in the West Coast….

**V. -** *Is the recruitment of data scientists strategic for the Group, and can AXA become the stuff of dreams for young data scientists?*

Recruiting data scientists is obviously key because they are talents who have just been trained with the latest technics and can bring them to our AXA entities.

I believe that AXA is a unique and challenging environment for young data scientists: first of all the activity of protecting people and properties is one of the noblest which exists. I

have been deeply involved in this Industry for now more than 20 years, and I never regretted this choice.

Secondly, the data are at heart of insurance and risks, probably more than in any other Industries. And for AXA it is clearly identified as a major priority in our strategy.

Finally, we offer the opportunity to be confronted with real and concrete problems, in more than 50 countries, in a working environment mixing entrepreneurial agility (Big Data projects are usually organized in teams of 4-6 persons) and the power of the institution (access to deep and rich data, to the best experts,…).

If I was 25 years old again, I am sure that I would more than ever follow the same path that brought me to AXA and the Insurance Industry at the beginning of the 90's.

## Discussion with Manasa Patnam

### Assistant Professor at ENSAE ParisTech

**Variances -** *What is your personal interest in Big Data? Is it a growing field for econometricians too?*

**Manasa Patnam -** Big Data is emerging as an important topic of analysis in economics. It has changed the way we approach economic analysis in two ways.

Firstly, economists now have access to highly disaggregated micro data on various topics, which have allowed them to answer questions that were previously not possible with aggregated datasets. For example, it is now possible to exactly estimate the elasticity of taxation i.e. the extent to which consumers decrease their demand based on a one unit increase in the tax for a product. This is possible by analyzing the purchase behavior of thousands of consumers on eBay, who interact globally with sellers located in regions facing varying tax rates. Previously, only data on aggregated sales tax and total consumption for each region were made available, that did not allow for the precise estimation of the tax elasticity.

Secondly, day-to-day business activity and transactions have transformed in a tremendous way. A large part of business activity is now increasingly being carried out digitally and on the internet. Digital products and markets have features that differ from those that have been traditionally assumed by economic models. This necessitates an understanding and analysis, both theoretical and empirical, on issues surrounding the digital economy. As a consequence there is a growing need to understand how firms and consumers behave online, compared to in traditional brick-and-mortar stores.

My own interest lied in the intersection of the two themes that I have just discussed above. I am specifically interested in how online markets help alleviate certain frictions, commonly encountered in physical markets, subsequently affecting consumer behavior. One example of a widely studied friction is limited information. Traditionally, when consumers shop, for example, in supermarkets, they are faced with an information problem in a way that they are not fully able to access all the information needed to make the right purchase decision. For instance, the physical limitation of the size of the store may prevent them from exploring other products in the same category because they are not stocked in that store. On the other hand, this problem is easily alleviated in internet markets because, potentially, all the

information on each products is just a "click" away. However this gives rise to a different, more complex, problem called limited attention. This is a situation where even though consumers have access to all the information, they have a scarcity of attention and may not be able to devote all their time considering the full catalogue of products on offers. Many internet platforms recognize this issue and attempt to alleviate and refine user search by providing "recommendations" for products that they "may like". Popular examples are websites such as Amazon and Netflix that target products to consumers based on past purchases, visual similarity of currently browsed products etc. Designing a recommendation system can be prohibitively costly, both financially and computationally. Given the high costs it is imperative to understand whether and by how much recommendation systems affect product sales. This is exactly the question I address in my research. I examine the effect of product recommendations on product sales. Specifically we analyse whether the effect can operate through an increase in the visibility (or salience) of the product. This is an important mechanism to understand because substantial salience effects indicate lack of consumer attention, and this has further implications on how we model and estimate consumer choice behavior econometrically.

**V. -** *Could you outline your key issues of this conference?*

**M.P. -** I think there were three key issues in this conference, consistent across the various topics being discussed.

The first was how to think about the "big data" influx. How can firms or startups use the massive data that is available to them to build better products and how can academics help build techniques and methods to help the analysis of available data.

The second was how "big data" is changing the market structure. To me this was quite interesting, because it shed light on new ways of consumer-firm interaction as well as firm-to-firm interaction. Many startups have developed ways to interact with consumers, which allow them to build a large base of users but also use information from this base to better target their product in the future. The use of novel means of communication, through social networks and media, seemed to dominate this form of interaction. On firm-to-firm interaction, I found that there was a drastic transformation taking place compared to traditional forms of retail market interactions. Mostly, entrepreneurs or startups operating on the internet complemented each other's products through the exchange of media and information. Technology is largely shared and there seemed to be muted competition effects. However these are still early days to draw inference on firm- to-firm interaction.

Finally, one key issue that was recurrent throughout the conference was the implication of big data for education. The emphasis was on understanding how academic institutions can build programs and curriculum to build "big data" scientists. Given the growing demand for the use and analysis of big data techniques and skills, there was a need felt to encourage educational institutions to address this specific requirement in their curriculum. With regard to curriculum there was a mixed consensus on which specific skill (management vs quantitative) to emphasize in the curriculum but there was a general agreement that a mix of both, for example through teamwork, would be ideal.

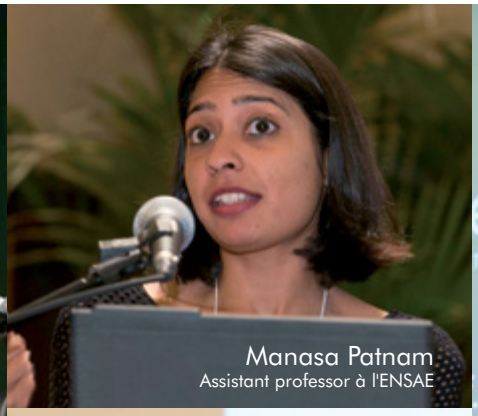
**V. -** *What do you think of this partnership between ENSAE and Berkeley?*

**M.P. -** I think it's fantastic! On the academic front, it enabled an interaction between academics working on big data topics and leading edge entrepreneurs who harvest and apply statistical techniques to big data to build their product. This is very useful for academics, because it helps us understand the current dynamics of the online marketplace. For example, I learnt a lot about what are the current trends in designing online platforms, what is the state of art technology that platforms are employing to attract consumers etc. It enabled me to frame and think more deeply about future research questions.

On the education aspect, both ENSAE and Berkeley would benefit from the exchange of talent and bright students. The exchange would be particularly beneficial because students can trade their skills (management or quantitative) to ensure a full development of their portfolio for their career in data science.

Overall, I envisage that this exchange will lead to many creative collaborations, both amongst academics and students. This will eventually serve to enhance the value of services being provided by the "big data" industry.


Joseph-Emmanuel Trojman
(1997)


Manasa Patnam
Assistant professor à l'ENSAE


Alexis Deudon
(2007)


Loïc Anton


Arnak Dalalyan
Professeur à l'ENSAE


Xavier Spinat

## UC Berkeley MFE and ENSAE ParisTech to Team up to prepare worldwide executives in Big Data Management Decisions

**By Linda Kreitzman,**
Executive Director, Master of Financial Engineering, UC Berkeley

The Master of Financial Engineering Program at the University of California at Berkeley (MFE) is a top ranked program in the world, and that has been keen on seeking strategic alliances with top tier schools in other countries and one of them is ENSAE. We are extremely pleased with the very strong relationship and programs that have been already established with ENSAE.

The financial community has considered the University of California Berkeley MFE program as a leader in training young financial engineers to become successful traders, quants, structurers and sales per-

sons on the sell side, and asset managers on the buy side over the last 15 years. On the technical front, the emphasis in our curriculum has been on pricing models, their calibration and the caution in using these models in illiquid markets, what is known as model risk. Statistics and econometrics have always been used in investment banking, but it seems now that these skills are valued more so than ever before. Also, in the past, banks were struggling with the storage of data and basically ignored historical data, especially high frequency data, in the calibration of their pricing models. Calibration was done through optimization techniques which ensured that the pricing of exotic products was consistent with the market prices of vanilla products. Only recently algorithmic traders have developed the expertise in exploiting high frequency data.

The environment has changed dramatically. Large amount of data are produced every day, every minute, every second. We now have the technology to store this data. With this explosion of data, comes the issue of not only how to analyze this data, but also to think through how this new capability of dealing with huge amount of data, what is known today as Big Data, will change the business model of banks and other financial institutions. For example, trading is becoming automated and traders are progressively replaced by robots. The focus has now shifted to developing algorithms to adjust automatically in real time the price of securities proposed to clients, to execute automatically the hedge of trading books taking into account the order book of many liquidity providers. This technology requires the most advanced statistical expertise that ENSAE has already started to teach its students. ENSAE has a long history in teaching advanced statistical methods and has been a pioneer in the development of the new statistical and econometric methods associated with Big Data.

The partnership between ENSAE and the Berkeley MFE, by combining each exclusive expertise, will allow both institutions to be at the forefront of the new education model in finance, banking, asset management and insurance. All these industries are currently adapting their business model and practices to Big Data, and few institutions are prepared to train the new generations of financial engineers to this new world. Our collaboration has already led to the launch of a very successful conference in Big Data, the topics of which can be found on www.datalead2014.com.

Our cooperation initially stemmed from our interest in the ENSAE students. These students are technically very well prepared to join the finance industry; in particular, they have a very strong theoretical background, not only in mathematics but also in statistics and econometrics. They have, however, little experience in applying the theory they have learned at ENSAE, which is precisely what we offer them at Berkeley throughout the program itself and, of course, the (paid) internship (as an associate) which is an integral part of the one year program at Berkeley. The combination ENSAE/UCB MFE makes the students from ENSAE unique; our experience with ENSAE's students has been very positive among the financial firms. We know there is a demand for ENSAE's students with the background they get and the experience and professional degree that we provide. When we bring on board an ENSAE student at Berkeley, we know we get a person with strong mathematics and statistics background, which enables us to «do the rest» which is to provide the applications to finance and direct exposure to the industry. Therefore, this dual degree we have is tremendously valued on Wall Street and other firms. This year particularly, we had three students from ENSAE and all three were placed in an internship with two different investment banks five weeks into their program. The courses they take at ENSAE such as econometrics, time series, stochastic calculus, C++, and so on…, are very useful. The financial industry values the technical expertise and skills ENSAE provides. Interestingly, you must have these skills in order to pass the interviews in the first place, so ENSAE's students tend to be «snatched fast» for internships and subsequently for full time positions.

Our greatest collaboration has yet to come but is under way. We are currently developing a certificate in Data Analysis and Decision for executives and plan a launch for early 2015. With ENSAE being the leaders in Big Data in France and the Haas School being known for its entrepreneurship spirit and extensive connection to Silicon Valley, we look forward to working together. The conference we put together reached out to experts in the field both in the US and France and they told us what we already knew: there is a need for preparing executives in Big Data in the months to come and that need can only grow. This is an exciting time and two countries will pour its intellectual resources to provide what no other school has ever done.

# Un dîner new yorkais

Zakaria Benjazia (1999), VP, Metlife

Près d'une trentaine d'anciens élèves de l'ENSAE s'est réunie au mois de septembre dans un restaurant coréen à Manhattan en présence de Julien Pouget, directeur de l'ENSAE ParisTech, Guillaume Gaudron, directeur des masters et du développement, et Rodolphe Pauvert, responsable des relations internationales.

La majorité des Alumni présents exerce dans le secteur financier privé ou dans la recherche universitaire où les compétences acquises au sein de l'ENSAE sont très demandées. Il était intéressant de témoigner que les Anciens de l'ENSAE ont atterri, non seulement au sein des filiales de sociétés françaises à New York, mais également au sein de nombreuses sociétés américaines. Ainsi, la discussion a pu porter sur les moyens d'améliorer la visibilité de l'école aux Etats-Unis et, notamment, au sein des groupes américains. En ce qui concerne la recherche universitaire, et plus particulièrement dans le domaine de l'économie et de la finance, les ENSAE Alumni continuent leur parcours dans les meilleures universités américaines, signe fort de la valeur et de la reconnaissance du diplôme de notre école.

La préparation des étudiants de l'ENSAE à exercer des postes de management, et comment trouver le bon équilibre entre préserver la formation technique de haut niveau et s'inspirer du modèle américain en ce qui concerne la communication, le management et l'entrepreneuriat, ont été des sujets largement abordés par les participants avec les représentants de l'école présents ce soir-là.

Enfin, le manque de moyens alloués à l'ENSAE par l'Etat a aussi été questionné et des suggestions ont été formulées pour créer une fondation ENSAE.

Après le dîner une partie du groupe est allée joyeusement finir la soirée sur un *roof-top* new yorkais :-). ■