



DGITM - Direction générale des infrastructures, des transports et de la mer

CGDD - Commissariat général au développement durable



Big Data et politiques publiques dans les transports

Séminaire conjoint DGITM, CGDD, ENS-Cachan, PSE-Ecole d'économie de Paris

15 octobre 2015, 9h-19h

Inscriptions en ligne : <http://www.predit.prd.fr/predit4/agenda/46774>

9h- 9h15 : Introduction générale (PSE, ENS-Cachan, DGITM, CGDD/DRI)

Première séquence - Introduction au Big Data et à ses enjeux industriels

9h15-10h :

Julien Malaurent (ESSEC), « La révolution des Big Data »

10h-10h50 :

Cédric Mora (DGE), « La solution industrielle française « Économie des données », dispositifs d'incitation et d'aide à l'innovation »

Nicolas Vayatis (ENS-Cachan), « La décision par algorithme - les effets de la transition numérique dans les sciences de l'ingénieur, le marketing et la médecine »

Pause café

11h-11h50 :

Yannig Goude (EDFR&D), « Préviation probabiliste de consommation électrique locale », avec Pierre Gaillard, (EDFR&RD) et Raphaël Nédellec (EDFR&RD)

Argyris Kalogeratos (ENS-Cachan), « Algorithmes efficaces pour contenir des processus de contagion sur des réseaux : applications en épidémiologie et en marketing »

Jean-Michel Tanguy (CGDD), « Les Hackathons »

Pause déjeuner

Deuxième séquence - Big Data et action publique : introduction des enjeux et exemples dans le domaine des transports et de la mobilité

13h – 13h30 :

Mario Barreto (Forum international des transports), « Big Data and Transport »

13h30-16h30 :

Maguelonne Chandesris (SNCF Innovation & Recherche), « Nouveaux modèles de concours : Challenge SNCF Prédiction de la fréquentation des gares SNCF en Ile-de-France, enjeux et exemples concrets d'applications du Big Data »

Latifa Oukhellou (IFSTTAR), « Présentation du projet Mobilletic - Analyse de l'intermodalité par les données de mobilité billettique »

Pierre-Yves Geoffard (PSE), « La concurrence dans les taxis (Uber/plateforme d'intermédiation) »

André de Palma (ENS-Cachan), « Using Big Data in Automated Individual Travel Surveys » avec Fang Zhao, Ajinkya Ghorpade, Francisco Pereira, et Moshe Ben-Akiva (MIT)

Pause-café

Nicolas Coulombel (LVMT), « Staggered work hours at the university, an economic evaluation using smart-card data » avec Mohamed Khalil El Mahrsi (IFFSTAR) et Emmanuel Munch (LVMT-SNCF)

Kamel Gadouche (CASD) et Nathalie Picard, (Université de Cergy-Pontoise et Ecole Polytechnique), « Informatique et traitement des données. Quelle acceptabilité ? Quel respect de la vie privée ? Appariements sécurisés et secret statistique. ».

Troisième séquence Table ronde : enjeux pour l'action publique – premières pistes et sujets à approfondir

Mario Barreto (Forum international des transports)

Ariane Dupont Kieffer (Université Paris 1)

Tristan Guilloux (Cerema)

Olivier Nalin (STIF)

Dominique Place (CGDD/SOeS)

Bernard Schwob (AFIMB)

16h30-18h : Table ronde

18h-18h15 : Conclusions générales : DGITM, CGDD, ENS-Cachan et PSE

18h15 : Cocktail

9h- 9h15 : Introduction générale (PSE, ENS-Cachan, DGITM, CGDD/DRI)

Jean-Marc Tallon est Directeur adjoint de Paris School of Economics en charge de la recherche et Université Paris I Panthéon-Sorbonne, CNRS Centre d'Economie de la Sorbonne. Il est Directeur de Recherche au CNRS. Jean-Marc Tallon effectue des recherches sur la décision dans l'incertain.

<http://www.parisschoolofeconomics.eu/fr/tallon-jean-marc/>

Luc Mathis est ingénieur du corps des ponts, des eaux et des forêts. Il occupe depuis 2012 le poste de chef de bureau de la politique technique à la Direction générale des infrastructures, des transports et de la mer (Ministère de l'écologie, du développement durable et de l'énergie). A ce titre, il pilote le réseau scientifique et technique du ministère dans le champ des transports et coordonne les actions liées à la politique technique, aux études et à la normalisation. Il contribue aux actions ministérielles d'orientation de la recherche et de l'innovation. Il participe à la gouvernance de plusieurs programmes français et européens de financement de la recherche et de l'innovation. Précédemment, il a occupé différentes positions d'expertise et d'encadrement à l'Institut Géographique National (IGN) sur des sujets de systèmes d'information.

Jean-François Sanchez est ingénieur de l'École Centrale Paris. Venu de la RATP où il a fait l'essentiel de sa carrière dans les domaines de l'exploitation et de l'ingénierie, il occupe depuis 2010 le poste de chef de la mission transports à la Direction de la recherche et de l'innovation du Commissariat général au développement durable (Ministère de l'Écologie, du Développement durable et de l'Énergie). A ce titre, il pilote des actions de recherche dans différents domaines du transport (socio-économie, fret et logistique, qualité et sécurité, nouveaux services de mobilité...). Il participe à la gouvernance de plusieurs programmes français et européens de financement de la recherche et représente le ministère auprès Comité conjoint de recherche transport du Forum International des Transports (OCDE).

André de Palma is Professor of Economics at the Ecole Normale Supérieure de Cachan. He holds a PhD in Physics (University of Brussels), and a PhD in Economics (University of Bourgogne). He has coauthored *A Discrete Theory of product Differentiation*, with S. Anderson and J. Thisse (MIT Press) and he is co-editor, with R. Lindsey, E. Quinet and R. Vickerman, of the Handbook in Transport Economics, (Edgar Elgard). He has widely published in Transportation / Urban Economics, in Industrial Organization and in Decision theory. He has developed with Yurii Nesterov and Fabrice Marchal, METROPOLIS, a dynamic traffic meso-simulator.

Première séquence - Introduction au Big Data et à ses enjeux industriels

9h15-10h :

Julien Malaurent (ESSEC), « La révolution des Big Data »

Julien Malaurent est professeur assistant en systèmes d'information à l'ESSEC Business School. Il est titulaire d'un PhD et Doctorat en systèmes d'information obtenu à l'ESSEC. Ses travaux de recherche se focalisent sur l'appropriation des systèmes d'information pour les grandes entreprises, la stratégie des technologies de l'information et plateformes digitales, et la veille technologique. Il a publié de nombreux articles dans des revues et conférences scientifiques internationales. Il est aussi éditeur associé de la revue Information Systems Journal depuis 2011.

Résumé de l'intervention: Le Big Data révolutionne la plupart des secteurs d'activités. Cette présentation proposera des clés pour en comprendre les raisons profondes:

Le Big Data est une révolution et implique un changement de paradigme : ce qui change dans la nature des données, leurs traitements, et les possibilités offertes ;

Comprendre les enjeux du Big Data dans le domaine des transports ;

Illustrations d'applications concrètes et impacts pour les entreprises du secteur.

10h-10h50 :

Cédric Mora (DGE), « La solution industrielle française « Économie des données », dispositifs d'incitation et d'aide à l'innovation »

Cédric Mora est Chargé de mission Cloud Computing et Big Data au sein de la Direction Générale des Entreprises (au Service de l'Economie de Numérique). Intervient notamment sur les appels à projets (rédaction des appels, puis sélection et suivi des projets) et sur la solution Economie des Données de l'Industrie du Futur (phase 2 des 34 plans de la Nouvelle France Industrielle lancés par François Hollande).

Résumé de l'intervention : La Nouvelle France Industrielle entre dans sa deuxième phase avec les annonces réalisées par Emmanuel Macron en mai 2015. Les plans industriels Cloud Computing, Big Data et HPC sont assemblés au sein de la nouvelle solution Economie des données. Cette solution économie des données met l'accent sur trois grands principes qui seront présentés :

Développer l'écosystème de l'Economie de la donnée ;

Favoriser le développement de nouveaux usages de l'économie de la donnée

Soutenir le développement de l'offre technologique française.

Les appels à projets du Programme des Investissements d'Avenir ont permis de lancer des projets de R&D innovants et des projets d'animation à destination de cet écosystème. Les nouveaux appels qui viennent d'être lancés en 2015 seront présentés à l'occasion de cette intervention.

Ces actions sont soutenues par la secrétaire d'Etat en charge du Numérique qui a annoncé récemment vouloir faire de la France un leader mondial du Big Data.

Nicolas Vayatis (ENS-Cachan), « La décision par algorithme - les effets de la transition numérique dans les sciences de l'ingénieur, le marketing et la médecine »

Nicolas Vayatis est directeur du Centre de Mathématiques et de Leurs Applications (CMLA, ENS-Cachan & CNRS) et responsable de la formation de master recherche M2-MVA qui accueille une centaine d'étudiants par an dans le domaine de la modélisation et du traitement numérique des données. Au CMLA, Nicolas anime un groupe de recherche autour de la thématique 'machine learning et traitement de données massives'. Il est l'auteur de plus de 70 publications dans des revues et conférences internationales à comité de lecture, et il participe à de nombreux projets interdisciplinaires aux interfaces des mathématiques dans les domaines de l'énergie, du design de systèmes, du marketing digital, de l'épidémiologie et de la médecine numérique. Nicolas est également membre du comité scientifique de la société 1000mercis et conseiller scientifique au CEA.

Résumé de l'intervention: Dans toutes les branches de l'activité humaine, l'ubiquité des capteurs et l'accumulation de l'information digitale soulèvent de nouveaux défis intellectuels, des rêves technologiques mais aussi des craintes quant au respect de la vie privée, au libre-arbitre, à la diversité culturelle, ... Les avancées récentes dans les domaines de la statistique en grande dimension, la description et la modélisation de processus sur des réseaux, et la deuxième (?) vie de la théorie de l'optimisation ont conduit à des concepts et des algorithmes qui rendent possible l'inférence à partir de données complexes et ouvrent des perspectives quant à l'interaction entre experts et scientifiques dans divers champs d'application. Un point de tension dans cette démarche est l'équilibre à trouver entre la personnalisation des solutions et la reproductibilité des travaux. Pourtant ces deux aspects devront, selon moi, guider les innovations futures dans le domaine du *machine learning*. Dans l'exposé, j'illustrerai ces idées en parcourant quelques réalisations récentes résultant de projets interdisciplinaires dans des domaines comme les sciences de l'ingénieur, le marketing digital, et l'éthomique.

Pause café

11h-11h50 :

Yannig Goude (EDFR&D), « Prédiction probabiliste de consommation électrique locale », avec Pierre Gaillard, (EDFR&RD) et Raphaël Nédellec (EDFR&RD)

Yannig Goude est chercheur à EDF R&D depuis 2008, expert en modélisation statistique et chef de projet "statistiques pour le management d'énergie". Il est également maître de conférence associé à l'université Paris-Sud 11 Orsay depuis 2013. Il a obtenu sa thèse en 2007 à l'université Paris-Sud 11 Orsay, est diplômé de l'ENSAI et de l'université de Rennes 1. Son domaine de recherche est la prédiction pour les marchés de l'énergie, les séries temporelles, les méthodes d'ensembles, les modèles additifs plus généralement les méthodes d'apprentissage statistique.

Raphael Nédellec (EDF R&D) est ingénieur-chercheur à EDF R&D depuis 2011. Diplômé de l'ENSAI, il s'intéresse à la modélisation de séries temporelles, les modèles non-paramétriques et leurs applications à la prédiction de consommation électrique.

Pierre Gaillard (Univ. PSUD-EDF R&D) est en post doctorat à l'université de Copenhague en mathématiques. Diplômé de l'École Normale Supérieure de Paris, il a reçu en 2015 à l'Université Paris Sud un doctorat sous la direction de Gilles Stoltz. Celui-ci s'est déroulé dans le cadre d'un partenariat CIFRE avec le département Osiris d'EDF R&D sous l'encadrement industriel de Yannig Goude. Pierre Gaillard a développé des algorithmes d'agrégation robuste séquentielle d'experts qu'il a appliqués aux marchés de l'énergie. Il s'est également intéressé à des méthodes de prédiction séquentielles non paramétriques et à des méthodes de prédiction quantile. Il a remporté les volets sur le prix et la consommation d'électricité de la compétition GEFCom2014 avec Yannig Goude et Raphaël Nédellec.

Résumé de l'intervention: Pour optimiser la gestion de la production ou du réseau électrique il est nécessaire de disposer de prévision de consommation à différentes mailles géographiques. Par exemple, la gestion du réseau de distribution s'effectue à minima au niveau des 2200 postes sources (voir au niveau des départs HTA) du territoire quand l'optimisation de la production d'EDF s'effectue au niveau national. Nous proposons une nouvelle méthode de prévision des différents quantiles de la consommation électrique à ces différentes mailles passant aisément à l'échelle. Nous la déclinons sur plusieurs jeux de données réelles dont notamment les données de la compétition GEFCom14 à laquelle cette approche est arrivée à la première place.

Argyris Kalogeratos (ENS-Cachan), « Algorithmes efficaces pour contenir des processus de contagion sur des réseaux : applications en épidémiologie et en marketing »

Argyris Kalogeratos est docteur en informatique et chercheur au CMLA (ENS-Cachan & CNRS) dans le domaine du machine learning. Ses domaines d'expertise vont de la segmentation non supervisée de flux de données complexes à l'analyse sémantique de données structurées. Il est actuellement chef de projet dans le cadre de SODATECH (PIA Big Data) mené conjointement avec la chaire Réseaux sociaux de l'institut Mines-Telecom et la société 1000mercis. Argyris poursuit également des travaux de recherche sur les algorithmes d'allocation dynamique de ressources utilisés en épidémiologie computationnelle

Résumé de l'intervention: Dans cet exposé, on s'intéresse à la viralité dans les réseaux et aux stratégies possibles pour contenir une épidémie dès lors qu'on dispose d'un budget limité d'antidotes. L'étude porte notamment sur la dynamique des mécanismes concurrents de contagion-guérison et contribue à l'élaboration de modèles computationnels pour l'allocation dynamique de ressources. On montre que la structure macroscopique et microscopique du réseau joue un rôle clé dans l'explication de ces phénomènes de propagation/endiguement.

Jean-Michel Tanguy (CGDD), « Les Hackathons »

Jean-Michel Tanguy est conseiller du directeur de la recherche et chargé du numérique auprès de la commissaire générale au développement durable. Il met en place un programme d'acculturation au numérique à destination des agents du CGDD, est organisateur d'événements pour la COP21, porte le projet Climate Change Challenge (C3) qui réunit une dizaine d'organismes autour de la diffusion des données publiques et participe dans le cadre du C3 à l'organisation d'un hackathon qui se tiendra du 6 au 8 novembre prochain en simultané à Paris, Toulouse, Lyon et Nantes doublé d'un challenge inter-écoles.

Résumé de l'intervention : Depuis la publication de la directive INSPIRE, la France progresse rapidement dans le domaine de l'ouverture des données publiques : création d'ETALAB, de la mission d'administrateur de données publiques, organisation d'événements (Mardi digital par la SE au Numérique)... Le projet Climate Change Challenge (C3) s'inscrit dans cette dynamique dans le contexte de la COP21. Rassemblant une dizaine d'organismes, son objectif est de susciter la création de solutions innovantes pour lutter contre le changement climatique et/ou s'y adapter. Il met en œuvre une méthode innovante en 3 étapes : (1) faire émerger les besoins et les rêves partagés (2) énoncer les défis auxquels on veut répondre (3) organiser un marathon de l'innovation (hackathon) de 48h pour imaginer et développer des solutions innovantes aux défis posés.

Plus de 150 bases de données ont été ouvertes qui touchent plusieurs thématiques dont l'aménagement du territoire pour lequel plusieurs défis concernant les transports sont identifiés. L'objectif du hackathon qui se déroulera du 6 au 8 novembre sur 4 sites en parallèle : Paris, Lyon, Toulouse et Nantes est de stimuler l'écosystème du numérique de manière à faire émerger des solutions innovantes, impliquer davantage le citoyen pour son action sur le climat et développer de nouvelles technologies.

Pause déjeuner

Deuxième séquence - Big Data et action publique : introduction des enjeux et exemples dans le domaine des transports et de la mobilité

13h – 13h30 :

Mario Barreto (Forum international des transports), « Big Data and Transport »

Mario Barreto est chargé du développement des statistiques de transport au Forum International des Transports (FIT). Le FIT, qui est la division transport de l'OCDE, a développé conjointement avec Eurostat et CEE ONU, un Glossaire de définitions communes à leurs pays membres. L'application de ces définitions assure une meilleure compréhension, comparabilité et harmonisation des statistiques de transports publiées par le FIT.

Résumé de l'intervention : Les données provenant de sources « Big Data » sont précieuses à plusieurs niveaux. Pouvoir y accéder fournit un support d'information important pour diverses utilisations notamment pour l'analyse statistique, les programmes de recherche ou encore l'aide à la décision politique. Cependant l'exploitation des données 'Big Data' se confronte à deux obstacles de taille.

Le premier concerne la confidentialité des informations. Les données géo-localisées utilisées à des fins de politique de transport, comportent une dimension confidentielle. Or l'anonymisation de ces informations de par leur nature spatiale et temporelle, reste relativement difficile et ne permet pas toujours d'assurer la protection des individus.

Le second concerne le mode opératoire, comment ces données sont obtenues et par qui. Souvent ces données sont collectées sur la voie publique, détenues et exploitées par le secteur public et peuvent avoir une valeur commerciale importante. Les Administrations, qui n'ont plus la main sur les informations

cruciales à la gestion des flux de transports, se retrouvent à devoir demander des droits d'accès pour des données clés à la mise en place et au développement de leur politiques. Est-ce un modèle viable ou même souhaitable et si non quels autres modèles pourrait-on instaurer?

13h30-16h30 :

Maguelonne Chandesris (SNCF Innovation & Recherche), « Nouveaux modèles de concours : Challenge SNCF Prédiction de la fréquentation des gares SNCF en Ile-de-France, enjeux et exemples concrets d'applications du Big Data »

Maguelonne Chandesris est responsable de la thématique innovation & recherche "Data, Mobilité et Territoires" pour SNCF. Elle est également en charge de l'équipe "Statistique, Econométrie et Datamining" qui développe des solutions de traitements statistiques avancées et de visualisation d'importants volumes de données au service de différentes activités du groupe SNCF.

Diplômée de l'Ecole Nationale de la Statistique et de l'Analyse de l'Information (ENSAI), elle est également titulaire d'un doctorat de Paris VI en mathématiques.

Résumé de l'intervention: SNCF manipule et récolte quotidiennement d'importants volumes de données, leur valorisation constitue un enjeu majeur pour l'entreprise. En effet, on verra que les nouvelles capacités technologiques modifient profondément le contexte et le jeu des acteurs des transports publics en offrant de nouveaux potentiels. Quelques exemples d'applications concrètes seront exposés.

Latifa Oukhellou (IFSTTAR), « Présentation du projet Mobilletic - Analyse de l'intermodalité par les données de mobilité billettique »

Latifa Oukhellou habilitée à diriger des recherches de l'Université Paris-Est (2010) et docteur de l'Université Paris-Sud Orsay (1997), est directrice de recherche à l'Ifsttar depuis 2011 au sein du laboratoire GRETTIA (Génie des Réseaux de transport terrestre et Informatique avancée). Elle était auparavant maître de conférences à l'Université Paris-Est Créteil entre 1998 et 2011. Ses activités de recherche concernent l'apprentissage statistique et la fusion de données pour des problèmes de diagnostic et pour l'analyse des données de mobilité. Elle anime un groupe de travail au sein de l'Ifsttar sur le traitement de données spatio-temporelles pour la mobilité (anim@tic). Au niveau national, elle est coordinatrice de deux projets : le projet Predit Mobilletic (2013-2016) qui est dédié à l'analyse de l'intermodalité par les données de mobilité billettiques : le cas Rennais et le projet ANR DIADEM (2013-2017) sur le Diagnostic Dynamique et la Maintenance Prévisionnelle de Systèmes Embarqués sur Train.

Résumé de l'intervention: L'analyse des mobilités urbaines peut s'appuyer sur des traces numériques de plus en plus nombreuses générées lors de nos déplacements. Nous disposons en effet aujourd'hui de multiples sources comme les données billettiques, les traces GSM, Wi-Fi ou bluetooth la géolocalisation de nos publications sur les réseaux sociaux, générées tout au long de nos déplacements. Ces dispositifs qui n'ont pas été initialement conçus pour l'analyse des mobilités trouvent ici une utilité évidente. Les traces numériques qu'ils génèrent peuvent servir à la mise en place d'approches de modélisation renouvelées des mobilités urbaines. L'objectif de ce séminaire est de présenter une synthèse des travaux menés dans le cadre du projet Predit Mobilletic. Il s'agit notamment de travaux de fouille de données billettiques collectées sur le réseau de transport urbain de l'agglomération Rennaise (Réseau Star). Les résultats obtenus illustrent le potentiel de ces données à offrir un éclairage sur la dynamique des mobilités urbaines.

Pierre-Yves Geoffard (PSE), « La concurrence dans les taxis (Uber/plateforme d'intermédiation) »

Pierre-Yves Geoffard est directeur de PSE- Ecole d'économie de Paris depuis janvier 2013. Il est directeur de recherche CNRS depuis 2002, directeur d'études à l'EHESS (Ecole des Hautes Etudes en Sciences Sociales) depuis 2008. Il est également membre du Cnis (Conseil national de l'Information statistique) et président de la Commission "Service public, Services aux publics". Pierre-Yves Geoffard est normalien (1985-89) et docteur en mathématiques appliquées aux sciences économiques.

Résumé de l'intervention : Le domaine des transports est l'un des secteurs où l'évolution des technologies de l'information change radicalement l'organisation de l'activité, et soulève d'importantes questions sur les régulations adéquates. L'irruption de plateformes permettant une intermédiation plus directe entre la demande de transport urbain et des conducteurs pouvant répondre à cette demande a bousculé de plein fouet l'activité des taxis, traditionnellement très étroitement encadrée et réglementée par les pouvoirs publics. A tel point que le terme barbare d' "uberisation" est décliné à toutes les sauces pour évoquer les transformations à l'œuvre dans d'autres secteurs très éloignés ("ubérisation de la médecine", "ubérisation des pressings,"...). De facto, la plateforme proposée par Uber ou d'autres permet à une activité directement concurrentielle des taxis d'émerger dans de nombreuses villes à travers le monde. Mais dans ce secteur comme dans d'autres, les plateformes peuvent faire l'objet de rendements croissants, conduisant à une situation de monopole naturel. Comment éviter que cette position dominante ne soit exploitée au détriment des usagers et des conducteurs? Comment introduire, ou réintroduire, la possibilité pour de nouveaux acteurs d'entrer sur ce marché si une entreprise y est déjà solidement installée? L'ouverture de certaines des données des opérateurs de transports est prévue dans le projet de "loi numérique" actuellement en discussion pour ce qui concerne les opérateurs publics, ou les entreprises privées gérant une délégation de service public. Une ouverture plus générale de certaines données de l'ensemble des opérateurs de transport (taxis, VTC), évoquée par le "rapport Jutand" serait-elle à même de réduire les risques d'abus de position dominante?

André de Palma (ENS-Cachan), « Using Big Data in Automated Individual Travel Surveys » avec Fang Zhao, Ajinkya Ghorpade, Francisco Pereira, et Moshe Ben-Akiva (MIT)

André de Palma is Professor of Economics at the Ecole Normale Supérieure de Cachan. He holds a PhD in Physics (University of Brussels), and a PhD in Economics (University of Bourgogne). He has coauthored *A Discrete Theory of product Differentiation*, with S. Anderson and J. Thisse (MIT Press) and he is co-editor, with R. Lindsey, E. Quinet and R. Vickerman, of the Handbook in Transport Economics, (Edgar Elgard). He has widely published in Transportation / Urban Economics, in Industrial Organization and in Decision theory. He has developed with Yurii Nesterov and Fabrice Marchal, METROPOLIS, a dynamic traffic meso-simulator.

Fang Zhao is a Research Scientist at the Singapore-MIT Alliance for Research and Technology (SMART), Future Urban Mobility (FM) group. She has an inter-disciplinary background with expertise in machine learning, communication networks, and travel surveys. She has led a research team in SMART to develop the Future Mobility Sensing technology. She received her Ph.D degree in Electrical Engineering from the Massachusetts Institute of Technology in 2010, and M.Eng and B.Eng degrees from the National University of Singapore in 2001 and 2000, respectively.

Ajinkya Ghorpade is a Research Engineer at the Singapore-MIT Alliance for Research and Technology (SMART), Future Urban Mobility (FM) group. His interests include machine learning, pattern recognition, software engineering and travel surveys. He is currently working in SMART to develop the Future Mobility Sensing technology.

Francisco Pereira has recently joined as professor in the Transport department of the Technical University of Denmark (DTU), where he leads the ITS research group. Previously, he was Senior Research Scientist at MIT/CEE ITS Lab, where he worked on real-time traffic prediction, behaviour modelling, and advanced data collection technologies, both in Boston and Singapore, as part of the Singapore-MIT Alliance for Research and Technology, Future Urban Mobility project (SMART/FM). His main research focus is on applying machine

learning and pattern recognition to the context of transportation systems with the purpose of understanding and predicting mobility behaviour, and modelling and optimizing the transportation system as a whole. Francisco has an MSc (2000) and PhD (2005) in Computer Engineering and Artificial Intelligence from University of Coimbra, Portugal.

Moshe Ben-Akiva is the Edmund K. Turner Professor of Civil and Environmental Engineering at the Massachusetts Institute of Technology (MIT), and Director of the MIT Intelligent Transportation Systems (ITS) Lab. He holds a PhD degree in Transportation Systems from MIT and honorary degrees from the University of the Aegean, the Université Lumière Lyon, the Royal Institute of Technology (KTH), and the University of Antwerp. His awards include the Lifetime Achievement Award of the International Association for Travel Behavior Research, the Jules Dupuit prize from the World Conference on Transport Research Society (WCTRS), and the Institute of Electrical and Electronics Engineers (IEEE) ITS Society Outstanding Application Award for DynaMIT, a mesoscopic simulator with algorithms for dynamic traffic assignment, traffic predictions and travel information and guidance. Ben-Akiva has coauthored two books, including the textbook *Discrete Choice Analysis*, published by MIT Press, and over 200 papers in refereed journals or in refereed journals or conference proceedings.

Résumé de l'intervention: With increasingly pervasive smartphone ownership and advanced sensing technologies, traditional household travel surveys are being supplemented or even replaced by automated travel surveys that can deliver previously unobtainable range of behavioral data and insights. One example of such a system is Future Mobility Sensing (FMS), developed at the Singapore-MIT Alliance for Research and Technology and MIT ITS Lab.

FMS consists of three separate, but inter-connected, components: the smartphone app (or other mobile devices) that collects the sensing data; the server that includes the database as well as the data processing and learning algorithms; and the web interface that users access to view and verify the processed data and answer additional questions to supplement the validated data. The figure below shows the three components and the data flows among them. FMS is a fully integrated and automated travel survey system that infers, in the backend, participants' activity times, locations, types, as well as transportation mode between the activities. User verification through the Activity Diary interface provides valuable input that helps improve the detection accuracy over time.

The context and user data being used in the backend plays an important role in improving the detection accuracy of the machine learning algorithm and reducing participant burden. Current FMS system makes use of Points of Interest (POI), land use, and transit network information. Other Big Data from various sources can potentially be used to further improve FMS. For example, smartcard, traffic, and even weather information can help refine mode detection. Data from social networks provides hints on events and incidents that help us better understand user behavior and activities. In addition, if linked up with smart-meter or credit card transaction information, FMS can reach an even higher level of resolution in activity data including in-home activities, and the types of shops/restaurants visited.

Field test of FMS in Singapore in conjunction with Land Transport Authority's (LTA's) Household Interview Travel Survey (HITS'12) demonstrated that this technology is capable of capturing higher resolution data with better spatial/temporal accuracy. Also, it collects multiday data at zero marginal cost, and reveals large day-to-day variability in the same user's activity pattern. It can be used to provide estimation on participants' emission and carbon footprint information. We will also discuss about the potential impact on economics analysis of this new approach of using Big Data and technology to collect detailed travel data over multiple days.

[Pause-café](#)

Nicolas Coulombel (LVMT), « Staggered work hours at the university, an economic evaluation using smart-card data » avec Mohamed Khalil El Mahrsi (IFFSTAR) et Emmanuel Munch (LVMT-SNCF)

Nicolas Coulombel is an assistant professor at Paris Est University, France. After completing his Ph.D. at the Ecole des Ponts ParisTech on the interactions between transport and land use, he extended his work on this topic by contributing to the SustainCity European project, under the supervision of Pr. de Palma. Since 2011, he has worked as an assistant professor at the LVMT research unit of Paris Est University. His main fields of interest are transportation and housing, be it from an economic or a modeling perspective. Recent research study how to leverage the new sources of data known as "Big Data" to improve our knowledge of housing and transportation markets. This includes the MOBILLETIC project, funded by the PREDIT and carried out in collaboration with GRETTIA and CEREMA, which aims to better understand travel behaviors using smart card data. Last, Nicolas Coulombel also holds a teaching position at the Ecole des Ponts, ParisTech.

Mohamed Khalil El Mahrsi conducted his Ph.D. thesis (that he defended in 2013) at Télécom ParisTech (France) under the advisorship of Prof. Dr. Fabrice Rossi. He is currently working as a postdoctoral researcher at Ifsttar where he is involved in the Mobilletic project. The aim of this project is to study the analysis and visualization of smart card data in the context of public transit.

Emmanuel Munch est diplômé en 2013 du Master Aménagement, Urbanisme et Transports spécialité Transport et Mobilité (IUP- Université Paris- Est Créteil Val-de-Marne / École des Ponts Paris Tech). Il est actuellement doctorant au LVMT depuis le 1er juillet 2014, sous la direction conjointe de Pierre Zembri et Frédéric de Coninck. Dans le cadre d'une Convention Industrielle de Formation par la Recherche (CIFRE), le travail de doctorat est partagé entre le LVMT et SNCF Transilien.

Résumé de l'intervention: Congestion in transportation systems has multiple economic drawbacks, which are now well documented in the literature. This includes travel conditions (time spent in congestion, discomfort, increased accident rate, etc.), but also the repercussions on the activities that follow or precede the trip (delay at destination, loss of productivity, etc.). In many cities, the struggle against congestion, in particular at peak hours, is therefore one of the key priorities in the field of transportation. While building new transport infrastructures used to be the standard response to alleviate congestion, it is costly, requires land, which becomes gradually scarcer in the central parts of metropolitan areas, and may even be counterproductive in the long run as it would actually spur the growth of traffic (Downs, 1962; Duranton and Turner, 2011). Accordingly, attention has shifted toward optimizing the use of existing infrastructures through travel demand management (TDM). Following the seminal works of Vickrey (1969, 1973), congestion pricing has for instance been extensively studied in the academic literature. This policy aims to spread travel demand by encouraging users to shift their departure time through efficient pricing strategies. It is rarely put to practice, however, in particular because of technical complexities and acceptability issues. One alternative to achieve a similar result is to implement staggered work hours schemes. The impact on travel demand is even more direct than in the case of congestion pricing, as travel demand is not changed through the result of a trade-off between the preferred arrival time and the toll level, but directly at the level of user preferences. Such schemes are furthermore easy to implement technically speaking and relatively inexpensive. Several papers have studied the impact of such schemes from a theoretical point of view (starting with Henderson, 1981), but empirical literature on this topic remains rare.

This paper aims to fill this gap and provide a better understanding of staggered work hours schemes by analyzing the case of the University of Rennes 2, France. With more than 15 000 students, this university is an important traffic generator at the local level. Located near a subway station, the significant flows of students lead to heavy congestion at specific times of the day, especially in the morning when classes start. In cooperation with the transit operator Keolis, the University of Rennes 2 has therefore decided to shift the class start time of half its students from one quarter hour - from 8:15 to 8:30 a.m. – starting from September 2013. This paper seeks to evaluate this experiment, which is largely considered a success by all the involved parties. To do so, we first develop a theoretical model to better understand and identify the various impacts of staggered work hours schemes in the case of a public transit line. Next, we combine two

data sources, transit operations data (GTFS) and smart card data, to appraise the effect of the staggered study hours scheme at Rennes 2 on the subway users' behavior and in-vehicle congestion. Finally, we provide a first economic assessment of the measure, in order to appraise the potential economic benefits that one might expect from such a scheme.

Kamel Gadouche (CASD) et Nathalie Picard, (Université de Cergy-Pontoise et Ecole Polytechnique), « Informatique et traitement des données. Quelle acceptabilité ? Quel respect de la vie privée ? Appariements sécurisés et secret statistique. ».

Kamel Gadouche est administrateur INSEE, directeur du CASD. Kamel Gadouche a travaillé pendant près de 15 ans dans la statistique et l'informatique, occupant des fonctions de développeur de chaînes statistiques et de chef de projet au sein de l'Insee. Il a rejoint le Genes en 2006 pour prendre la direction du service organisation et système d'information. Il a parallèlement conduit un projet de mise en place en France du Centre d'Accès Sécurisé aux Données (CASD) dont il est directeur aujourd'hui.

Nathalie Picard est Maître de Conférence Habilitée à Diriger des Recherches, au THEMA, Université de Cergy-Pontoise, où elle enseigne principalement l'économétrie. Elle est agrégée de mathématiques, diplômée de l'ENSAE et docteur en économie. Ses thèmes de recherche vont de l'Economie des transports et économie urbaine à la Théorie de la décision, en passant par l'Economie du développement, la Démographie économique, la Microéconométrie, l'Econométrie des variables qualitatives, l'Analyse du risque, l'Economie expérimentale, la Finance comportementale, les Modèles de décision au sein des couples et les Modèles collectifs. Elle a publié une trentaine d'articles dans des revues internationales à comité de lecture telles que Journal of Economic Theory, Journal of Urban Economics, Transportation Science, Transportation Research, Theory and Decision, Marketing Letters, Mathematical Population Studies, Journal of Population Economics, ainsi qu'une dizaine de chapitres d'ouvrages.

Résumé de l'intervention : Début 2010, Nathalie Picard, Catherine Quantin, Benoît Riandey et Anne Solaz ont publié dans *Courrier de Statistiques* un article faisant le point sur "Les besoins des démo-économistes en matière d'appariements sécurisés". Ils avaient fait le tour des données de panel disponibles en France pour des analyses socio-économiques et/ou démographiques et n'avaient pu que constater leur pauvreté comparée aux panels américain (PSID), britannique (BHPS) ou allemand (GSOEP). Ils avaient alors proposé des solutions pour apparier données d'enquêtes ou de panels et sources administratives dans le respect de la confidentialité. Cinq années plus tard, la situation a beaucoup évolué, avec la création du CASD et de ses fameuses boîtes, qui complètent les avancées dans la diffusion de données administratives, sans oublier les possibilités offertes par le recensement rénové. La situation est amenée à évoluer encore plus dans un avenir proche, en lien avec la loi sur le numérique, en particulier l'article 10, et la consultation associée sur les appariements. Il est donc temps de faire le point sur les données accessibles et les nouvelles perspectives de diffusion.

Troisième séquence Table ronde : enjeux pour l'action publique – premières pistes et sujets à approfondir

16h30-18h : Table ronde

Bernard Jullien est maître de conférences Hors Classe en économie (HDR) en poste à l'Université de Bordeaux. Bernard Jullien a été de septembre 2006 à septembre 2015 le Directeur du Réseau International GERPISA – Groupement d'Intérêt Scientifique soutenu par le MESR et le ME localisé à l'ENS-Cachan. Il a ce titre conduit plusieurs recherches ou études sur l'industrie automobile (projet européen ICaTSEM ; projet « La Transition vers une économie écologique dans l'automobile » 2010-2012 ; mission d'évaluation « Bilan à mi-parcours de l'action du FMEA » 2011-2012 ; expertise auprès du CCFA, de la CSIAM et du CNPA dans le cadre de l'auto-saisine de l'Autorité de la Concurrence sur le marché de l'après-vente automobile en France

2011-2012 ; Livre Blanc de l'ACSIA 2012). Il a co-dirigé l'ANR « Gouvernement Européen des Industries » GEDI avec Andy Smith 2008-2012 et a dirigé cette année le programme « Nouveaux services de mobilité offerts par les usages novateurs de la voiture (véhicule en libre-service, covoiturage, autopartage...) » conduite dans le cadre du Pôle interministériel de prospective et d'anticipation des mutations économiques (PIPAME) et demandée par la DGE et le CGDD.

Mario Barreto est chargé du développement des statistiques de transport au Forum International des Transports (FIT). Le FIT, qui est la division transport de l'OCDE, a développé conjointement avec Eurostat et CEE ONU, un Glossaire de définitions communes à leurs pays membres. L'application de ces définitions assure une meilleure compréhension, comparabilité et harmonisation des statistiques de transports publiées par le FIT.

Ariane Dupont is the dean of the department of economics at Paris PI Panthéon-Sorbone. After ten years of experience of investigating theoretical and methodological issues at IFSTTAR (former INRETS) to contribute to the assessment of road externalities (including, congestion, safety, GHG and local air pollutants), she moved in September 2014 to PHARE at la Sorbonne in order to work on macroeconomics on the one hand and on the economic and philosophical grounds of social justice, and the translation of the theoretical concepts of equity and fairness in policy implementations over the last century and more recently in the transport sector on the other hand. She also works on the history of transport modeling and transport data.

Tristan Guilloux, est architecte urbaniste de l'Etat. Responsable du groupe Analyse et Prospective de la Mobilité (APM) au sein de la Direction technique Territoires et Ville du Cerema. Le groupe APM a en charge la production et la diffusion de bases de données dans le domaine des services de transports collectifs urbains, départementaux et régionaux ainsi que sur le stationnement public. Par ailleurs il contribue à l'élaboration de méthodologies d'enquêtes sur les déplacements pour lesquelles les enjeux de données numériques sont croissants. Il contribue également au pilotage des actions du CEREMA en matière de modélisation et de prospective.

Olivier Nalin est Directeur du développement, des affaires économiques et tarifaires au sein du STIF, l'autorité organisatrice de la mobilité en Ile de France. Le STIF définit l'offre et la tarification des transports, il veille à la cohérence des investissements. Le STIF élabore le plan de déplacement urbain d'Ile de France et anime sa mise en œuvre. Pour assurer ses missions le STIF rassemble de nombreuses données et pilote l'observatoire multimodal de la mobilité en Ile de France (OMNIL). Le système d'information sur les données de validation (SIDV) est une réalisation originale du STIF en matière de Big Data. L'ensemble des données de validation anonymisées sont recueillies et analysées par le STIF. Le SIDV fournit des informations particulièrement riches sur la mobilité et le comportement des voyageurs selon les titres de transport qu'ils utilisent.

Dominique Place est chef de projet « Connaissance de la mobilité des personnes » à la sous-direction des statistiques des transports du SOeS, en charge de l'élaboration du nouveau dispositif national de connaissance de la mobilité des ménages et de la réflexion sur les possibilités d'utilisation des nouvelles technologies et d'exploitation de données massives.

Bernard Schwob est directeur de l'Agence française pour l'information multimodale et la billettique (AFIMB). Opérationnelle depuis début 2011, l'AFIMB est chargée en particulier i) de promouvoir l'interopérabilité dans les domaines de l'information multimodale et de la billettique, et de soutenir la normalisation dans ces domaines ; ii) d'encourager le développement de services d'information à l'intention des usagers, prenant en compte l'ensemble des modes de transport ; iii) de mettre les systèmes d'information multimodale en réseau, en liaison avec les collectivités territoriales volontaires.

[18h-18h15 : Conclusions générales : DGITM, CGDD, ENS-Cachan et PSE](#)

[18h15 : Cocktail](#)