

Enregistré à la présidence de l'Assemblée nationale
le 9 juillet 2015

Enregistré à la présidence du Sénat
le 9 juillet 2015

RAPPORT

au nom de

**L'OFFICE PARLEMENTAIRE D'ÉVALUATION
DES CHOIX SCIENTIFIQUES ET TECHNOLOGIQUES**

sur

**LA PLACE DU TRAITEMENT MASSIF DES DONNÉES (*BIG DATA*)
DANS L'AGRICULTURE : SITUATION ET PERSPECTIVES**

*Compte rendu de l'audition publique du 2 juillet 2015
et de la présentation des conclusions du 8 juillet 2015*

PAR

M. Jean-Yves LE DÉAUT et Mme Anne-Yvonne LE DAIN, députés, et M. Bruno SIDO, sénateur

Déposé sur le Bureau de l'Assemblée nationale
par M. Jean-Yves LE DÉAUT,
Président de l'Office

Déposé sur le Bureau du Sénat
par M. Bruno SIDO,
Premier vice-président de l'Office

Composition de l'Office parlementaire d'évaluation des choix scientifiques et technologiques

Président

M. Jean-Yves LE DÉAUT, député

Premier vice-président

M. Bruno SIDO, sénateur

Vice-présidents

M. Christian BATAILLE, député
Mme Anne-Yvonne LE DAIN, députée
M. Jean-Sébastien VIALATTE, député

M. Roland COURTEAU, sénateur
M. Christian NAMY, sénateur
Mme Catherine PROCACCIA, sénatrice

DÉPUTÉS

M. Gérard BAPT
M. Christian BATAILLE
M. Denis BAUPIN
M. Alain CLAEYS
M. Claude de GANAY
Mme Françoise GUÉGOT
M. Patrick HETZEL
M. Laurent KALINOWSKI
M. Jacques LAMBLIN
Mme Anne-Yvonne LE DAIN
M. Jean-Yves LE DÉAUT
M. Alain MARTY
M. Philippe NAUCHE
Mme Maud OLIVIER
Mme Dominique ORLIAC
M. Bertrand PANCHER
M. Jean-Louis TOURAINE
M. Jean-Sébastien VIALATTE

SÉNATEURS

M. Patrick ABATE
M. Gilbert BARBIER
Mme Delphine BATAILLE
M. Michel BERSON
Mme Marie-Christine BLANDIN
M. François COMMEINHES
M. Roland COURTEAU
Mme Dominique GILLOT
M. Alain HOUPERT
Mme Fabienne KELLER
M. Jean-Pierre LELEUX
M. Gérard LONGUET
M. Jean-Pierre MASSERET
M. Pierre MÉDEVIELLE
M. Christian NAMY
Mme Catherine PROCACCIA
M. Daniel RAOUL
M. Bruno SIDO

SOMMAIRE

	Pages
INTRODUCTION	7
M. Jean-Yves Le Déaut, député, Président de l'OPECST.....	7
M. Marcel Deneux, ancien sénateur, ancien Président de la Caisse nationale du Crédit agricole	9
PREMIÈRE TABLE RONDE : QU'EST-CE QUE LE TRAITEMENT MASSIF DES DONNÉES (BIG DATA) ?	13
<i>Présidence : M. Jean-Yves Le Déaut, député, président de l'OPECST</i>	13
M. Serge Abiteboul, professeur à l'ENS Cachan, directeur de recherche à l'INRIA, membre de l'Académie des sciences	13
M. Thomas Hauet, maître de conférences à l'Université de Lorraine	16
M. Claude Kirchner, directeur de recherche, conseiller du président de l'INRIA.....	18
M. Alexandre Termier, professeur à l'université de Rennes 1, responsable d'une équipe au laboratoire INRIA/IRISA	20
M. Xavier Vincelot, <i>Datascientist Senior</i> , cabinet Quantmetry	23
M. Laurent Gouzènes, membre du Conseil scientifique de l'OPECST et expert scientifique du groupe Pacte novation	25
DEUXIÈME TABLE RONDE : COMMENT LE <i>BIG DATA</i> A-T-IL PRIS PIED DANS L'AGRICULTURE ?	27
<i>Présidence de M. Jean-Yves Le Déaut, député, président de l'OPECST</i>	27
M. Hervé Pillaud, secrétaire général de la chambre d'agriculture de Vendée et membre de la commission « Élevage » de l'APCA	27
M. Jean-Louis Peyraud, chargé de mission, direction scientifique agriculture, INRA.....	29
M. Stéphane Grumbach, directeur de l'IXXI, Institut rhônalpin des systèmes complexes, chef de l'équipe INRIA DICE, INRIA Rhône-Alpes.....	32
Mme Marie-Laure Neuburger, entrepreneur.....	34
M. Vivien Mallet, chargé de recherche, INRIA	36
M. Jean-Pierre Chanet, chercheur et directeur de l'unité de recherche technologies et systèmes d'information pour les agrosystèmes d'IRSTEA.....	39
M. Sébastien Lafage, directeur marketing et communication, ISAGRI	41
M. Laurent Gouzènes, membre du conseil scientifique de l'OPECST, expert scientifique du groupe Pacte novation.....	43
DÉBAT	44

TROISIÈME TABLE RONDE : COMMENT FAVORISER LE DÉVELOPPEMENT D'UNE AGRICULTURE DE PRÉCISION EUROPÉENNE ?	51
<i>Présidence : Mme Anne-Yvonne Le Dain, députée, vice-présidente de l'OPECST</i>	51
M. Hervé Pillaud, représentant de la Chambre d'agriculture de Vendée.....	52
M. Frédéric Garcia, chef du département Mathématique et Informatique appliquée à l'INRA, délégué à la Transition numérique.....	54
M. Stéphane Grumbach, directeur de l'IXXI, Institut Rhône-Alpin des Systèmes complexes, chef de l'équipe INRIA DICE, INRIA Rhône-Alpes.....	56
Mme Gaëlle Kotbi, enseignante-chercheuse dans le département « Stratégie et entrepreneuriat » de l'Institut polytechnique LaSalle Beauvais.....	58
M. Stéphane Marcel, directeur général de SMAG.....	61
M. Pierrick Givone, directeur général délégué à la recherche à l'innovation de l'IRSTEA. Merci Madame la Présidente.....	63
M. Pierre-Philippe Mathieu, responsable du programme <i>EO4Food</i> à l'ESA.....	65
DÉBAT	67
CONCLUSION	81
Mme Anne-Yvonne Le Dain, député, vice-présidente de l'OPECST.....	81
EXTRAIT DE LA RÉUNION DE L'OPECST DU 8 JUILLET 2015 PRÉSENTANT LES CONCLUSIONS DE L'AUDITION PUBLIQUE	83
ANNEXES	89
ANNEXE 1 : INTERVENTION DE M. THOMAS HAUET, MAÎTRE DE CONFÉRENCE À L'UNIVERSITÉ DE LORRAINE	91
ANNEXE 2 : INTERVENTION DE M. JEAN-LOUIS PEYRAUD, CHARGÉ DE MISSION, DIRECTION SCIENTIFIQUE « AGRICULTURE », INRA	95
ANNEXE 3 : INTERVENTION DE MME MARIE-LAURE NEUBURGER, ENTREPRENEUR	97
ANNEXE 4 : INTERVENTION DE M. PIERRE-PHILIPPE MATHIEU, RESPONSABLE DU PROGRAMME E04FOOD, ESA	99

INTRODUCTION

M. Jean-Yves Le Déaut, député, Président de l'OPECST. Nous allons commencer. Notre collègue Anne-Yvonne Le Dain arrivera plus tard, car elle a une réunion ce matin.

Je vous remercie d'abord d'être venus pour cette audition publique sur *Big Data* et agriculture. Cette audition est la claire retombée de la visite que Bruno Sido, Premier vice-président de l'OPECST, et moi avons rendue au Salon de l'agriculture en février de cette année. Lors de nos échanges sur place, nous avons été alertés sur le fait que des constructeurs d'équipements placeraient des capteurs et des systèmes de transmission de données sur les machines qu'ils vendent pour recueillir des informations et alimenter les traitements massifs de type *Big Data*.

Comme nos interlocuteurs comptaient parmi les meilleurs experts du secteur, et je parle ici de François Houllier, Président de l'INRA, de Jean-Marc Bournigal, Président de l'IRSTEA à qui l'on a confié la mission Innovation 2025, la sollicitation nous a paru très sérieuse, d'autant que les organisations agricoles que nous avons vues par la suite, deux organisations syndicales, nous ont dit la même chose.

D'un côté, les capteurs implantés dans le matériel agricole permettent d'obtenir des informations toujours plus nombreuses et précises sur l'état des sols, des semences, sur les techniques culturales. De l'autre, des technologies nouvelles de traitement massif des données (ou *Big Data*) permettent d'extraire des directives techniques précises et efficaces de cet afflux de données.

Aujourd'hui, les agriculteurs, comme beaucoup d'autres professionnels, ont le téléphone portable vissé à l'oreille et n'ont aucune réticence à intégrer les nouvelles technologies. En l'occurrence, l'agriculture dite « de précision » que permet le *Big Data* est pleine de promesses. Aux États-Unis, où le mouvement a déjà pris, la hausse des rendements à l'hectare va de pair avec la diminution des quantités d'intrants divers, donc avec une meilleure maîtrise des coûts. Mais il ne faut pas perdre de vue les risques associés à ces nouvelles technologies, en lien avec la gestion des données. C'est d'ailleurs une question qui ne se pose pas seulement dans le domaine de l'agriculture, c'est une question beaucoup plus générale sur laquelle nous devons réfléchir.

D'une part, la souveraineté de la France pourrait être menacée si sa filière agricole devenait tributaire de services conçus par de grandes sociétés étrangères. D'autre part, les conditions de transmission et de réutilisation de données sensibles collectées ne sont pas toujours claires. Ainsi, nous pouvons craindre que certains tracteurs, construits dans d'autres pays, n'envoient les données prises par leurs capteurs à l'insu de l'utilisateur. En tout cas, c'est ce que nous allons essayer de préciser avec vous ce matin.

Nous sommes donc revenus du Salon de l'agriculture avec l'idée de creuser ce sujet dans le cadre d'une audition publique, suggestion que le bureau de l'OPECST a appuyée. Ce sujet fait en effet typiquement appel aux compétences de l'OPECST telles qu'elles ont été définies par la loi du 8 juillet 1983. L'OPECST, l'Office parlementaire d'évaluation des choix scientifiques et technologiques, a pour mission d'explorer les questions relevant de l'évolution des sciences et des technologies, pour détecter à l'avance les enjeux législatifs qui leur sont associés.

L'OPECST a donc été conçu pour travailler en amont de la fabrication de la loi dans ces domaines. Au besoin, nous pouvons mener une étude complète pour identifier les enjeux, en consultant largement la communauté scientifique et professionnelle concernée et en prenant appui sur les situations étrangères analysées sur place pour disposer d'une information de première main.

Mais l'expérience de l'OPECST, depuis trente ans qu'il existe, a démontré qu'il pouvait être aussi utile d'aborder une question dans le cadre d'une audition publique, collective, contradictoire, sur une demi-journée comme ce matin. Quelquefois, cette forme d'investigation suffit pour en tirer des recommandations. Ainsi récemment, en novembre 2014, (je pense que le rapport correspondant est disponible sur les tables à l'entrée de la salle), nous avons investi la question des survols illicites de drones pour réclamer des moyens de recherche de détection d'objets non métalliques par des techniques audio et vidéo, et ces moyens ont été rapidement dégagés.

Dans d'autres cas, l'audition publique est préalable à une étude plus complète. Ainsi deux de nos collègues, le député Patrick Hetzel et la sénatrice Delphine Bataille, approfondissent aujourd'hui la question des terres rares qui avait déjà été abordée lors d'une audition publique en 2010.

Autre exemple : le sénateur Bruno Sido et la députée Anne-Yvonne Le Dain ont publié en février 2015 un rapport sur la sécurité numérique, né des observations (notamment concernant la globalité du problème, au-delà des problématiques apparemment disjointes dans les sphères civile et militaire) qui ont été tirées d'une audition publique sur le risque numérique de février 2013. Messieurs Claude Kirchner, Serge Abiteboul et Stéphane Grumbach ici présents doivent se souvenir de cette audition.

Au passage, vous observerez que nous construisons autant que possible nos binômes de rapporteurs sur une triple mixité. Théoriquement, il n'y a qu'une mixité mais notre triple mixité renvoie aux trois appartenances de chambre (Assemblée nationale/Sénat), de sexe (hommes/femmes) et de courant politique (majorité/opposition). Nous essayons de travailler à l'abri des positions partisans. L'OPECST a la chance de pouvoir s'exprimer en toute liberté, de défendre son point de vue au nom de l'intérêt général.

S'agissant aujourd'hui – je vais terminer par là – de la question du *Big Data* dans le domaine de l'agriculture, elle soulevait pour nous beaucoup d'interrogations, notamment parce que c'est la première fois que l'OPECST aborde cette problématique du *Big Data* en tant que telle.

Vous allez devoir nous renseigner sur diverses questions qui s'imposent d'emblée : comment s'opère la collecte des données ? Est-elle transparente ou clandestine ? Quelle est la finalité économique du traitement en mode *Big Data* ? S'agit-il d'une démarche utilisée dans de nombreux secteurs ? La souveraineté nationale est-elle en jeu ? L'avenir de l'agriculture est-il biaisé par ces pratiques ? Nous avons organisé des débats autour de ces questions en trois tables rondes.

La première table ronde vise à présenter le principe du *Big Data*, en essayant de mettre en lumière les ruptures techniques qui ont permis l'accumulation d'un nombre considérable de données, la mise au point d'algorithmes adaptés au traitement de gros volumes d'informations. Avant de regarder plus particulièrement ce qui se passe dans l'agriculture, il s'agit ainsi d'avoir une idée en général sur les types de services qui sont rendus possibles par le *Big Data*.

La deuxième table ronde s'attachera à essayer de comprendre comment le *Big Data* peut avoir sa place dans le secteur de l'agriculture, où se trouvent et comment fonctionnent les objets connectés qui sont utilisés pour recueillir les données, quels services sont produits à partir du *Big Data*.

La troisième table ronde abordera la dimension stratégique du développement des *Big Data* dans l'agriculture, en examinant d'une part quelle maîtrise les agriculteurs et les pouvoirs publics peuvent et doivent avoir sur l'installation de capteurs et d'objets connectés, et d'autre part quelles conséquences a ou aura le développement des services fondés sur le *Big Data* pour la souveraineté nationale et européenne dans le domaine de l'agriculture.

Avant cela, je donne d'abord la parole, en qualité de grand témoin, à Marcel Deneux, ancien sénateur, ancien collègue et même vice-président de l'OPECST, ancien rapporteur de plusieurs rapports à l'OPECST, pour qu'il nous apporte son témoignage sur la manière dont l'agriculture française a su faire face aux défis liés à l'émergence de nouvelles technologies au cours des dernières décennies.

En effet, Marcel a été un des hauts responsables du Crédit agricole au niveau national et a mené une longue carrière d'agriculteur. Je lui laisse la parole pour resituer cette question du *Big Data* dans le cadre de l'évolution des technologies en agriculture. Après nous aborderons les tables rondes.

M. Marcel Deneux, ancien sénateur, ancien Président de la Caisse nationale du Crédit agricole. Merci, Monsieur le Président. Merci de m'avoir invité. Comme tu viens de le dire, Président, je suis ancien sénateur, ancien vice-président de l'OPECST ; le fait de m'inviter permet ma formation continue, mon

recyclage en matière de travaux parlementaires. Déjà, au bout de huit mois, je m'aperçois du décrochage, alors continuez à m'inviter comme cela pendant une vingtaine d'années pour que je ne devienne pas complètement obsolète ! Merci.

Le privilège des vétérans est de voir l'histoire se répéter et relancer des séquences qui se déroulent autour d'objets nouveaux, inconnus auparavant, mais selon un schéma déjà bien identifié. J'ai la chance d'avoir, comme tu l'as dit Président, une vision assez longue de l'agriculture. Je ne vais pas vous raconter toute ma vie, mais pour faire court, après avoir été responsable des jeunes agriculteurs dans un département, j'ai élaboré les lois d'orientation agricole avec Edgard Pisani et le Général de Gaulle. Cela situe l'époque.

Après, j'ai eu un cheminement multiple où, notamment, à partir de la Caisse nationale du Crédit agricole, j'ai abordé les possibilités de développement technologique dans l'agriculture qui n'en était qu'à ses débuts. J'ai été invité en 1982 (vous voyez l'époque) par le constructeur japonais NEC, à Tokyo, pour visiter un site sur lequel on gérait plus de 10 000 comptes. Nous étions, nous, Caisse nationale, en recherche d'une solution pour gérer de multiples comptes. Nous ne savions pas comment nous y prendre. Tout cela, c'est mon passé.

J'ai vécu les progrès de la mécanisation et de la motorisation, depuis l'usage du semoir, de toutes les machines qui sont arrivées au fur et à mesure, aujourd'hui du GPS pour ce qui est des cultures et des drones. Je viens d'un département où (j'y ai un peu contribué) la chambre d'agriculture a acheté deux drones l'année dernière. Cette agriculture picarde, en principe, n'est pas en retard.

J'ai vécu les progrès de l'intensification et de la standardisation avec les progrès du remembrement, de la spécialisation, l'organisation des filières, jusqu'à la distribution par les grandes surfaces.

J'ai vécu les progrès de rendement procurés par la sélection des espèces et la stimulation par de multiples ingrédients, dont des engrais toujours plus performants.

J'ai vu l'agriculture découvrir puis intégrer successivement les diverses avancées de la technologie : d'abord les machines, les méthodes du taylorisme, les forces de la chimie, les possibilités de la génétique, puis j'ai vu monter les revendications pour sortir de la révolution verte et s'orienter vers une agriculture biotechnologique. J'avais créé un centre de développement à Amiens il y a trente-et-un ans pour développer des techniques renouant avec des pratiques ancestrales comme les améliorations des semences, les rotations des cultures et d'autres techniques que nous avons un peu laissé tomber.

Ainsi, lorsque j'entends parler du *Big Data*, j'avoue que je ne vois pas exactement de quoi il s'agit (mais vous allez nous le dire), sinon que c'est une manifestation de cette révolution numérique qui a atteint toute la sphère économique et qui envahit aujourd'hui la vie sociale. Mais je ne doute pas que l'agriculture française saura se mettre en ordre de marche pour l'intégrer avec le

même appétit de connaissance et la même capacité d'apprentissage qu'elle a toujours su montrer depuis des générations, et plus spécialement depuis soixante-dix ans que j'y participe.

Il y aura des soubresauts, des réactions, donc des ajustements à faire, comme il en a fallu pour concilier la recherche de productivité et la protection du capital naturel, et éviter ainsi l'appauvrissement irréversible des sols. En l'occurrence, il faudra s'affranchir du risque de mainmise des fournisseurs américains, et aussi gérer les risques pour les libertés publiques qui sont liés à l'accumulation des données.

Mais l'agriculture française est toujours parvenue à surmonter les difficultés d'ordre technique pour ne garder que la meilleure part des technologies, donc je n'ai pas de doute que cela se produira de nouveau. J'observe d'ailleurs que l'utilisation à grande échelle de l'informatique est peut-être une nouveauté dans le domaine de l'agriculture en général (et encore), mais qu'elle a déjà une certaine antériorité dans le domaine de l'élevage.

Je voudrais à ce propos vous rappeler une péripétie que j'ai connue, pour bien montrer où nous en sommes – le grand public n'a pas toujours compris tout cela. Nous avons voté en 1966, je crois, à un an près, avec Edgar Faure, une loi sur l'organisation de l'élevage. Il s'agissait de rendre performants l'élevage français et notre système de sélection génétique, le développement de l'insémination artificielle. Il fallait croiser les fichiers tenus à la main des livres de race (que l'on appelait improprement *herd-books*, dans un français anglicisé) et les performances qui étaient enregistrées par des syndicats de contrôle de performance, tout cela à la main, pour converger vers la mise à disposition des centres d'insémination.

Il a fallu une loi pour obliger à cela. Avec Jacques Poly, Directeur adjoint de l'INRA qui était au cabinet, nous savions l'objectif mais nous nous demandions comment faire. C'était un gros truc. Nous avons découvert à l'époque qu'il n'y avait que l'INRA qui avait la capacité de calcul. Nous avons passé une convention avec toutes les organisations d'élevage (*via* l'obligation que nous avons introduite par la loi et quelques arrêtés) qui nous ont aidés à tout regrouper.

Cela a donné ce que nous appelons dans notre jargon d'éleveurs « le numéro à dix chiffres ». La France était en avance lorsque nous avons été confrontés à la maladie de la vache folle, il y a onze ou douze ans, je ne sais plus, par rapport à l'Angleterre qui n'identifie même pas les individus. Nous avons déjà une capacité de suivi de nos animaux, et encore aujourd'hui, je le dis par rapport à des péripéties dont nous entendons parler, nous sommes l'un des rares pays où en simultané, au jour le jour, nous connaissons le nombre d'animaux dans le cheptel bovin français, puisque tous les éleveurs sont tenus de déclarer sous 24 heures les entrées et les sorties. Nous avons mis cela en place voici une vingtaine d'années et cela fonctionne. Par rapport au recensement de personnes, du côté des bovins, nous ne sommes donc plutôt pas en retard.

Vous ayant dit tout cela, je me réjouis que l'Office aborde ce sujet de la place du *Big Data*. Et comme sur les domaines qu'il explore il est par nature très pédagogique, je ne doute pas d'en savoir bientôt un peu plus sur ce sujet intrigant. Je connais bien les méthodes de travail de l'Office puisque j'ai eu l'occasion, je vous l'ai dit, de les utiliser.

À cet égard, je voudrais me réjouir, en terminant, de constater, au vu du programme que j'ai reçu, que le Président Le Déaut a invité des entreprises qui sont en prise directe avec le sujet comme SMAG et ISAGRI. Il donnera la parole à une jeune créatrice d'entreprise fraîchement diplômée, titulaire d'un doctorat. Celle-ci a fait appel à une enseignante de l'Institut polytechnique LaSalle Beauvais (encore une publicité pour la région Picardie) où se forment les acteurs du monde agricole de demain, et sans doute dès l'année prochaine ceux de l'agromachinisme, domaine dans lequel on ouvre une formation à Beauvais qui est assez unique en France.

Permettez-moi, Président, avant de m'arrêter, de conclure que je suivrai avec beaucoup d'intérêt vos travaux et que cela contribuera à me recycler, ce que j'espère. Merci.

M. Jean-Yves Le Déaut. Merci à Marcel Deneux de cette introduction. Nous allons attaquer la première table ronde. Je voudrais d'abord lui dire qu'il se refuse décidément à vieillir parce que la crise de la vache folle, ce n'était pas il y a dix ans, c'était en 1996, donc voici dix-neuf ans. Le virus du prion a été découvert en 1982 par Prusiner. Cela fait trente-trois ans. C'est un peu plus ancien que ce que tu disais.

Deuxièmement, notre rôle est d'informer nos collègues sur des sujets complexes. Les problèmes politiques, très souvent aujourd'hui, ont des connotations technologiques et s'attachent à des sujets complexes que nous, parlementaires, n'arrivons pas toujours, il faut je crois le dire, à appréhender. Vous, vous êtes des spécialistes et votre rôle est de nous aider à décortiquer les sujets pour essayer de voir s'il y a une nécessité de régulation.

**PREMIÈRE TABLE RONDE :
QU'EST-CE QUE LE TRAITEMENT MASSIF DES DONNÉES (BIG DATA) ?**

Présidence : M. Jean-Yves Le Déaut, député, président de l'OPECST

M. Jean-Yves Le Déaut. Cette première table ronde s'attache à décrire ce qu'est le *Big Data*. Dans son intitulé, vous l'avez vu, j'ai choisi d'utiliser le terme de « traitement massif de données » car un traitement massif suppose implicitement un grand nombre de données, et la valeur ajoutée du *Big Data* me semble située du côté du traitement plutôt que du côté de l'accumulation massive de données. C'est ce que nous allons essayer de mieux comprendre.

C'est donc une possibilité technique nouvelle fondée sur l'alliance entre des stocks gigantesques de données et des capacités de traitement démultipliées. Nos intervenants vont avoir pour mission de nous éclairer là-dessus, même si nous avons déjà essayé de trouver des définitions.

Notre premier intervenant est Serge Abiteboul, membre de l'Académie des sciences. Il est déjà venu nous voir à l'Office une première fois en 2013 lors de notre audition publique sur le risque numérique, et avait dessiné des pistes pour mieux maîtriser la situation. J'espère qu'il a pu constater que l'on a tenu compte de ses conseils dans le rapport qui a été rendu depuis. Je vais lui donner la parole pour qu'il nous indique comment selon lui le *Big Data* renvoie à d'autres évolutions déjà constatées dans le domaine de la technologie.

M. Serge Abiteboul, professeur à l'ENS Cachan, directeur de recherche à l'INRIA, membre de l'Académie des sciences. Merci, Monsieur le Président. Je vais commencer par dresser rapidement un panorama du domaine scientifique et technique.

Le point de départ, ce sont des avancées dans les matériels, des mémoires beaucoup moins chères, des processeurs beaucoup plus rapides qui ont conduit à la possibilité de construire des puissances de calcul considérables, les fameuses « fermes de machines ». Maintenant, on sait mettre ensemble des milliers de machines avec des mémoires de plusieurs téraoctets sur chaque machine, ce qui fait que nous avons une puissance d'analyse de quantités d'informations considérable. C'est de cela que nous parlons. Cela n'aurait pas été possible il y a vingt ans.

Cela se combine à un autre aspect de l'évolution technologique : le développement d'objets communicants très bon marché. Il faut voir là-dedans aussi des améliorations techniques, des questions de prix. Des objets communicants, nous en avons depuis très longtemps. Maintenant, on sait construire des objets communicants qui ne valent rien du tout et qui peuvent vous permettre de couvrir tout un domaine agricole si vous voulez, d'en mettre des tas dans des machines agricoles, *etc.*

Nous avons des données et des objets communicants et déjà là, nous rencontrons des problèmes de la technologie : Qu'est-ce que l'on fait comme analyses ? Qu'est-ce que l'on vous garantit sur la qualité des résultats que vous obtenez de votre analyse ? Ces objets communicants, que récupèrent-ils comme données ? Qu'envoient-ils ? Il y a le problème de transparence des données, des problèmes de confidentialité, plein de problèmes très intéressants et importants.

Quand on confronte une industrie à ces nouvelles possibilités, typiquement elle les regarde d'abord de façon un peu frileuse en disant : « *Cela va changer complètement notre façon de travailler.* » Le problème est que nous n'avons pas le choix. Si ce n'est pas l'industrie directement concernée qui s'en empare, ce seront d'autres, peut-être les grandes plates-formes Internet (pas besoin de donner de nom), qui vont s'installer là-dedans et utiliser leurs compétences en traitement de données, en *Data Center*, en analyse de données, en récupération de données.

Nous avons commencé à avoir cela dans un premier temps dans des domaines comme le Web avec la recommandation, le commerce comparatif. Il est important de voir que c'est là un beau laboratoire pour ce qui est arrivé ailleurs. Nous voyons des entreprises venues d'ailleurs qui pénètrent un domaine (par exemple, l'hôtellerie), qui récupèrent des tas de données et qui, d'une certaine façon, vont se mettre à concurrencer les professionnels du domaine, ceux qui savent vraiment faire fonctionner des hôtels mais qui ne savent pas gérer de l'information, gérer des données, et qui se retrouvent, d'une certaine façon, coupés de la médiation avec leurs clients, avec tous les risques que cela implique.

Les domaines d'application, nous les voyons arriver les uns après les autres. Il y a tout ce qui est transport avec les voitures hyper-connectées. Maintenant, quand vous construisez une voiture, à peu près 50 % du développement c'est du logiciel et de l'informatique. Et puis, la santé, l'assurance, et (pour un enseignant-chercheur comme moi, c'est une évolution importante) l'éducation. Dans l'éducation, il y a aussi cette course à la récupération de données avec des grandes plates-formes qui s'interposent devant les étudiants et qui pourraient à terme remplacer les centres classiques d'enseignement.

Dans tous les cas, le phénomène est à peu près le même. La concentration de données, l'existence de déluges de données ouvrent des possibilités. Il faut bien comprendre que ce sont des possibilités, extrêmement positives, de faire des économies d'échelle, de développement, et d'offrir ainsi des fonctionnalités complètement nouvelles.

Pour ce qui est des risques, au Conseil national du numérique, nous avons écrit un rapport sur la neutralité des plates-formes, où l'on étudie dans différents domaines ce genre de problèmes, en mettant en évidence de grandes similarités. On constate au départ l'irruption de nouvelles sociétés, typiquement des grands du Web, qui viennent s'installer dans un secteur industriel et qui utilisent leur avantage en matière de technologie numérique pour, d'une certaine façon, pousser sur le côté les industriels du domaine, voire les transformer en *commodities*.

Comment lutter contre cela ? Évidemment, cette concentration, ce n'est pas quelque chose que nous voulons laisser arriver sans rien faire, surtout d'un point de vue national (quand toutes ces sociétés sont américaines).

Il y a évidemment le rôle des législateurs et de l'État, en particulier le rôle de la communauté européenne à mon avis dans ces domaines-là parce que, très souvent, dans le numérique, les problèmes transcendent les frontières. Je crois aussi beaucoup à ce que peuvent faire les utilisateurs ; chaque utilisateur peut décider de fonctionner autrement, de ne pas se laisser imposer des choix. Il y a tout un ensemble de travaux en cours sur les systèmes d'information personnels ou ce genre de choses. Enfin, je crois surtout aux réponses des professionnels du domaine.

Là, il ne faut pas se fermer les yeux. Ce n'est pas simple. Il existe évidemment une solution de facilité qui consisterait à dire : « *Moi, je continue à faire mon boulot super bien, j'ai mes clients, tout se passe bien ; en fait, je suis en train de tomber du centième étage (c'est une image qu'utilise François Bourdoncle), je suis au quarantième et tout se passe toujours très bien.* »

En fait, il faut prendre le problème en amont. Je crois que dans ces cas-là, la solution, pour contrecarrer ces grandes sociétés qui sont extrêmement puissantes, est de développer des écosystèmes ; on ne peut pas se lancer isolément à l'attaque de Google ou de Facebook. C'est la construction d'écosystèmes en s'appuyant sur la recherche, dans le domaine concerné, en l'occurrence l'agriculture (vous avez beaucoup cité l'INRA qui est extrêmement avancé), en impliquant l'INRIA dont je fais partie (qui est aussi très avancé dans tout ce qui est numérique), le tissu des *start-up*, les industriels du domaine.

Je veux vraiment insister là-dessus. L'écosystème doit être le plus large possible. Par exemple, il ne faut pas regarder que l'agriculture. Je pense que dans ce cadre-là, c'est aussi toute l'industrie agroalimentaire qui est concernée, c'est aussi l'industrie de la distribution. C'est un problème qui est va bien au delà de ce qui se passe dans les champs.

Pour conclure, je vais recourir à des analogies. Je ne connais pas grand-chose à l'agriculture, pour ne pas dire que je n'y connais rien. J'ai essayé de trouver des analogies avec les domaines que nous avons analysés au sein du Conseil national du numérique. Quand vous voyez un tracteur bourré d'objets communicants, vous êtes très près de ce qui se passe dans le transport routier. Via des capteurs, on peut apprendre énormément de choses sur le comportement des agriculteurs, comme on peut apprendre énormément de choses sur le comportement du conducteur. Il y a un autre domaine qui, à mon avis, présente beaucoup de similarités, c'est la santé. Dans le domaine de la santé, ce que l'on a essentiellement développé avec la médecine moderne, ce sont des méthodes générales pour soigner les êtres humains, car, grâce au numérique, on peut s'approcher de méthodes beaucoup, beaucoup plus personnalisées. On peut connaître chaque personne, chaque individu, voire modéliser chaque personne. Si

ce sujet vous intéresse, la leçon inaugurale de Nicolas Ayache au Collège de France a brillamment porté là-dessus, sur la santé personnalisée. Je pense que dans l'agriculture (encore une fois, excusez mon incompetence sur le sujet), c'est un peu la même chose. On peut personnaliser beaucoup plus l'agriculture. Personnaliser, dans ce cadre-là, cela veut dire la localiser beaucoup plus. Il y a très longtemps, le paysan connaissait parfaitement bien chaque recoin de ses champs, et les inflexions de la météo locale. Avec l'agriculture intensive, on a perdu un peu de cette proximité. Avec le numérique, il y a la possibilité peut-être de revenir à un traitement beaucoup plus proche du terrain, et évidemment de mettre en œuvre des agricultures plus écologiques, plus intelligentes.

M. Jean-Yves Le Déaut. Vous avez essuyé le feu le premier, donc je ne vous ai rien dit, mais vous avez vu qu'il y a un bouton rouge signalant qu'on a dépassé son temps. À l'Assemblée nationale, on est habitués à des temps contraints. Le Président vous coupe la parole. Ce que nous ne ferons pas, mais je vais demander aux suivants d'essayer d'être concis, en vous remerciant de ce que vous venez de nous indiquer.

Je vais donner la parole à Thomas Hauet, maître de conférences à l'université de Lorraine. Ce spécialiste des matériaux nanomagnétiques va évoquer les changements formidables qui se sont opérés au cours des dernières décennies dans les capacités de stockage des données qui ont rendu possible le *Big Data*.

Thomas Hauet, je le connais depuis longtemps parce que j'ai donné un cours à Sciences-Po Paris pendant une dizaine d'années sur les grands enjeux technologiques du XXI^e siècle. Il faisait le même cours sur le campus européen franco-allemand de Sciences-Po Paris à Nancy. Il m'a invité un jour dans son cours sur la vulgarisation des évolutions technologiques. À vous la parole.

M. Thomas Hauet, maître de conférences à l'Université de Lorraine. Merci. J'ai préparé quelques diapositives rapides pour vous décrire comment nous sommes passés de la bibliothèque papier dans les années 1980 aux serveurs Web aujourd'hui, en remplaçant un livre par un disque dur ⁽¹⁾.

Aujourd'hui nous pouvons distinguer les méthodes de stockage d'information à partir de leurs grandes caractéristiques majeures : la contenance totale de chaque objet, la densité d'information, c'est-à-dire la quantité d'information que l'on peut mettre au mètre carré (par la suite, je vais parler d'*inch* carré, 2,5 centimètres par 2,5 centimètres, c'est la valeur généralement utilisée), la vitesse d'accès à l'information, le coût, la pérennité ou la volatilité de l'information et la mobilité de l'objet.

Une chose dont je ne vais pas parler beaucoup par la suite, mais qui a quand même un gros impact sur le *Big Data* : les mémoires semi-conductrices avec les DRAM et les SRAM, mémoires volatiles et les flash, non-volatiles. Les

(1) Voir Annexe 1.

mémoires flashes sont de moins en moins coûteuses, très mobiles, avec une densité d'information de plus en plus importante et une vitesse d'accès très grande. Ce sont vraiment des objets mobiles très intéressants pour le *Big Data* qui pourraient être amenées à remplacer les disques durs pour certaines applications serveurs Web.

Le papier, qui a été supplanté de plus en plus, existe toujours comme support d'information.

Les disques gravés fonctionnent avec un système optique. On vient graver simplement un disque, on regarde là où il y a un trou et là où il n'y en a pas. Cela fait un 1 et un 0 ; c'est le principe du langage binaire qui est utilisé dans les ordinateurs. Le problème de ces disques gravés est que la taille du *spot laser* limite la densité d'information que l'on peut mettre sur ce genre de média.

Les bandes magnétiques fonctionnent à peu près comme les disques durs magnétiques : on met des petits grains magnétiques sur une bande plastique, l'aimantation d'une centaine de grains va former un 1 ou un 0. Les bandes magnétiques sont encore très utilisées, contrairement à ce que l'on pourrait penser, notamment dans le domaine de la banque. Le grand avantage des bandes magnétiques est son très bas coût, et la densité d'information stockable qui est énorme. Le problème est que pour accéder à l'information, il faut rembobiner ou débobiner la bande entière. C'est très intéressant pour conserver une information très longtemps de façon pérenne. En revanche, il ne faut pas avoir à accéder très souvent à l'information. C'est très souvent le domaine bancaire qui va être intéressé, où certains serveurs doivent conserver l'information longtemps.

Le disque dur va être au cœur de mon propos d'aujourd'hui. Nous allons voir dans la diapositive suivante que l'ensemble de ses caractéristiques a évolué énormément, depuis le premier disque dur dans les années 50 avec le RAMAC d'IBM. On avait alors une grosse boîte ici avec une cinquantaine de disques qui faisaient jusqu'à 60 centimètres de large. On n'avait que 5 mégabytes d'information avec un coût énorme du gigabyte (10^9 informations) de l'ordre de 10 millions de dollars.

En 1981, l'objet disque dur fait 30 à 40 centimètres de large avec des coûts moindres (100.000 dollars du gigabyte), des vitesses un peu plus grandes d'accès à l'information (20 Mbit/s).

En 1994, nous avons la structure du disque telle que nous le connaissons aujourd'hui, nous sommes à 100 dollars le gigabyte. Quand nous avançons encore, de 2000 jusqu'à aujourd'hui, en 2015, nous arrivons à des disques durs permettant de stocker 4 000 gigabytes d'information avec des densités nettement plus grandes par *inch* carré (500 Gbit/inch^2) et des coûts de 0,04 dollar le gigabyte. Nous sommes passés de 10 millions à 0,04 dollar le gigabyte en 50 ans. C'est vraiment un changement drastique.

Les trois grands points que nous pouvons noter quant à l'évolution technologique du disque dur concernant finalement : le moteur (il va tourner de plus en plus vite, donc permettre l'accès à l'information de plus en plus rapidement), la taille des petits grains magnétiques qui va diminuer, ainsi que le nombre de grains par bit (par 1 et par 0). Ici on peut voir des données enregistrées avec chaque grain portant une aimantation (une flèche magnétique) un peu comme la terre qui fait un aimant (soit nord-sud, soit sud-nord), et cela vous marque 1 et un 0. Enfin la tête de lecture et la tête d'écriture ont dû être adaptées, leurs dimensions diminuées pour écrire de plus en plus d'informations, avec le moins d'erreurs possible. Un point très important : je vous représente ici la densité d'information, la quantité d'information par mètre carré (encore une fois, ici, c'est en *inch* carré) et en comparaison, le coût de l'information. Ici, nous voyons que du RAMAC en 1956 jusqu'à aujourd'hui, nous avons une augmentation de 10^{-6} à 1 000 de la densité d'information, pour vous donner les échelles de grandeur, et en parallèle, le coût diminue drastiquement de 10 million à 0.01. Le prix et la densité du papier sont indiqués en vert.

Je veux noter que chaque fois qu'il y a eu un changement de pente dans cette augmentation de stockage, cela a été lié à des inventions et des découvertes scientifiques. Nous pourrions noter par exemple la GMR (*Giant Magneto-Resistance*) ici, qui a donné lieu à un prix Nobel français, Albert Fert en 2007. La recherche fondamentale et le développement dans les instituts publics comme privés ont vraiment eu un gros impact pour améliorer les disques durs.

Avec cela, je vous remercie pour votre attention.

M. Jean-Yves Le Déaut. Merci beaucoup. Vous pourrez bien sûr tous intervenir à nouveau dans le débat s'il y a des aspects sur lesquels vous souhaitez revenir.

Maintenant, je vais donner la parole à Claude Kirchner, conseiller du président de l'INRIA, mais aussi un peu le contact entre l'INRIA et l'OPECST. Nous avons, dans les organismes, des contacts privilégiés. Il m'a beaucoup aidé à organiser cette audition publique, et je l'en remercie.

Il fait partie d'un groupe de travail mis en place par l'Académie des technologies pour essayer de cerner les enjeux du *Big Data*. Une des idées clés de ce travail collectif est de faire ressortir que le *Big Data* n'est pas développé pour le plaisir d'offrir des services nouveaux, mais pour conquérir des marchés. Il vise à créer un avantage commercial disruptif fondé sur la baisse des coûts et la satisfaction plus fine des clients. Pouvez-vous nous indiquer quelles sont ces caractéristiques du *Big Data*, lui permettant de créer ces avantages potentiels ?

M. Claude Kirchner, directeur de recherche, conseiller du président de l'INRIA. Merci Monsieur le Président. Pour prendre la suite des deux orateurs précédents, je voudrais revenir sur le type de données qu'aujourd'hui nous avons l'occasion, l'opportunité de gérer.

Bien sûr, il y a les données dont nous avons parlé tout à l'heure, les données issues des capteurs, les données de géolocalisation, les données que nous allons stocker ensuite sur différents médias – dont les disques que nous venons d'évoquer. Quand on parle d'agriculture par exemple, il s'agit, entre autres, de toutes les informations qui peuvent être relatives aux données météorologiques, aux calculs de surfaces, *etc.* On eut y ajouter toutes les données produites par le fonctionnement des marchés.

Un deuxième type de données qui me semble fondamental, et dont Serge Abiteboul a précédemment parlé, concerne les données générées à partir de ces données. Typiquement, quand on est dans un réseau social, on va échanger des données qui vont conférer de la valeur ajoutée aux informations dont on est en train de discuter. Ces données, que l'on peut appeler des « données dérivées », sont issues des informations humain-machine, ou bien humain-humain, ou bien machine-machine. Ce type de données est produit également lors des recherches d'informations : quand on demande une information en utilisant un outil comme Google, on va donner de l'information suivant la manière dont on va poser la question.

Bien sûr, un autre type fondamental de données concerne celles issues de la recherche scientifique.

Mais ces différents types de données ne servent à rien si l'on n'a pas d'intelligence pour les traiter. Aujourd'hui, on peut dire que les connaissances, c'est d'une part des données diverses et variées, et d'autre part de l'intelligence. Cette intelligence provient de différents médias : par exemple la visualisation, qui permet aux hommes de mieux comprendre, mais aussi tout ce qui est algorithmes, tout ce qui est dans certaines catégories d'algorithmes – typiquement le *machine learning*, l'apprentissage, *etc.*

Les progrès actuels sont donc également corrélés aux progrès du côté du matériel, pour parvenir à exploiter les informations, les données diverses et variées que nous avons.

Aujourd'hui, nous passons (pour reprendre une image astronomique) du télescope optique, où l'on pointait la lunette à la main et où l'on traitait l'ensemble des choses de façon totalement humaine, à un télescope numérique où l'on a besoin de la machine, du traitement massif d'informations avec les différents algorithmes pour obtenir une vision en termes de connaissance de l'ensemble des données que nous allons traiter.

Ce télescope numérique offre des opportunités absolument formidables comme cela a déjà été souligné, mais aussi une dépendance très forte à la qualité et à la probité des données. Qu'est-ce qui m'assure que les données que je suis en train d'analyser sont correctes ? Si l'entité qui me les donne ou qui me les vend me vend les données qui sont effectivement correctes, quelle est leur niveau de

qualité ? Quand on les a récoltées, quelle est la précision de ces données, que ce soit leur date, leur précision géographique, *etc.* ?

Il s'opère dès lors un changement profond dans la mise au point des stratégies, et aussi dans la définition des modalités opérationnelles, parce que l'on n'utilise plus du tout un télescope optique – la connaissance de l'individu – mais un télescope numérique qui permet de bien mieux comprendre l'ensemble des données agrégées.

Cela pose la question de la souveraineté numérique, au niveau de l'État, de l'individu, aussi du point de vue scientifique. S'agissant de l'agriculture, elle va conduire à soulever quelques questions essentielles : quelles sont les informations que je peux utiliser de façon à comprendre quel sera le rendement, quelle sera la façon de mieux gérer l'exploitation ? Comment utiliser les données recueillies de façon à mener la recherche scientifique le mieux possible, de façon à effectuer les avancées dont on sait qu'elles sont fondamentales pour la suite de la science, et bien sûr la maîtrise du domaine ?

Dans ce contexte et pour terminer, la maîtrise des données constitue un enjeu fondamental. Cette maîtrise des données (je n'ai pas dit « *la possession des données* », j'ai dit « *la maîtrise des données* ») et le fait que l'on soit capable de déterminer leur qualité et leur probité est un point fondamental qu'il faut mettre en lien avec toutes les initiatives autour de ce que l'on appelle les *commons* (la notion de bien commun).

On peut penser aux initiatives de Pierre Bellanger sur le sujet, mais j'attire également votre attention sur le fait que l'ensemble de la numérisation, en particulier de l'agriculture, pose nettement un problème de la cybersécurité puisque l'on va dépendre des données et de leur traitement. Typiquement, si l'on est capable de modifier par exemple un épandage d'engrais, ou de contrôler différemment des robots, des drones, *etc.*, on peut avoir un impact sur le milieu qui peut être considérable. Merci beaucoup.

M. Jean-Yves Le Déaut. Je donne maintenant la parole à Alexandre Termier, professeur à l'université de Rennes 1, responsable du laboratoire INRIA/IRISA, et qui a accepté de venir nous parler du *Data Mining*, le traitement massif que l'on utilise dans le *Big Data* pour extraire des informations pertinentes. Je crois que vous nous donnerez au passage votre avis sur la place qu'il reste pour les algorithmes linéaires dans ces traitements, c'est-à-dire les algorithmes simples reposant sur l'hypothèse que le résultat varie comme les variables de base.

M. Alexandre Termier, professeur à l'université de Rennes 1, responsable d'une équipe au laboratoire INRIA/IRISA. Merci Monsieur le Président. Vous m'attribuez un peu trop de grade : je suis juste responsable d'une équipe au sein du laboratoire INRIA/IRISA.

Je vais entrer dans le détail de ce télescope numérique dont parlait Monsieur Claude Kirchner à l'instant, car le *Data Mining* poursuit le but de trouver automatiquement des connaissances dans les données, des connaissances dont on espère qu'elles vont être nouvelles, utiles et compréhensibles. Ce type de connaissances (je vais vous donner quelques exemples), cela peut être des modèles de classement. Vous en utilisez un tout le temps, peut-être même actuellement : ce sont les filtres de spam ou de non-spam dans les mails.

Il y a d'autres choses que l'on peut faire. Partant des données brutes, on va essayer d'opérer des répartitions en plusieurs groupes qui font du sens. C'est comme cela, par exemple, que Facebook peut regrouper ses utilisateurs selon leur profil et faire de la publicité ciblée.

Une autre chose que le *Data Mining* peut faire, c'est extraire des motifs, trouver des régularités dans les données. Un exemple sympathique est fourni par les supermarchés, qui font cela depuis assez longtemps : en regardant le comportement des clients, en étudiant le contenu de leur caddie, on trouve, par exemple, que 10 % d'entre eux achètent à la fois du pain, du beurre et du chocolat. C'est un motif. Cela va aider à gérer les produits concernés : on va essayer de les situer loin les uns des autres dans le magasin, comme cela les clients vont allonger leur parcours et dépenser davantage, ou au contraire on va faire des promotions conjointes ciblées.

Le problème inverse consiste à détecter des anomalies ; par exemple pour essayer d'identifier, dans du trafic réseau, des attaques potentielles à travers des anomalies du trafic.

Quel est le lien avec le *Big Data* ? De fait, le *Data Mining* est plus ancien que le *Big Data*, c'est quelque chose qui existe depuis un moment. L'intérêt de toute la publicité qui a été faite au *Big Data*, en particulier avec Google et consorts, c'est que beaucoup de gens se sont rendu compte de toute la valeur ajoutée contenue dans les données. Cela a mis encore plus en lumière l'intérêt du *Data Mining*.

Au niveau technique, comment est-ce que cela se passe entre le *Big Data* et le *Data Mining* ? Suffit-il d'appuyer sur un bouton pour que toutes les connaissances soient extraites ? Bien sûr, cela ne se passe pas toujours aussi bien. La première génération d'outils de traitement massif qui sont sortis pour le *Big Data* étaient très bien pour les algorithmes simples – les fameux algorithmes linéaires. Tout se passait bien. Il suffisait de compter des « choses » dans les données. Il est particulièrement important pour Google et Facebook de bien compter des « choses » dans les données (nombre de visites, de *likes*, etc.).

C'était bien, mais malheureusement, le *Data Mining* pour compter ne suffit pas. Il utilise souvent des algorithmes compliqués qui ont besoin de faire plusieurs passes sur les données. Avec de très grosses données, il arrive cependant

qu'après une première passe, vous génériez des données aussi grosses ; si vous devez repasser encore une fois, cela ne se passe donc pas très bien.

En fait, c'était un tel problème, et il y avait un tel besoin d'utiliser ces méthodes de *Data Mining* que cela a motivé la création d'une deuxième génération d'environnement de traitement massif. Je pense notamment à des travaux effectués à Berkeley, où l'on a conçu des environnements qui sont en train de s'imposer partout dans l'analyse de données *Big Data*. Ces approches sont adaptées aux algorithmes qui font plusieurs passes sur les données comme le supposent les algorithmes de *Data Mining*, et donnent de bons résultats.

Maintenant, grâce à ces techniques, nous disposons d'algorithmes de *Data Mining* que nous maîtrisons bien, qui font les tâches dont je vous ai parlé auparavant, et qui sont robustes au changement d'échelle, c'est à dire qu'ils peuvent traiter des téraoctets, même des pétaoctets de données. Dans certains cas, c'est suffisant. Si je veux faire des regroupements, je vais pouvoir les faire sur un grand nombre de données. Cependant, quand on a vraiment beaucoup de données, certaines tâches vont mal se passer – notamment si j'essaye de trouver des motifs.

Je vais vous donner une image. Si j'analyse les données de la supérette du coin – pas une supérette à Paris, nous allons nous situer dans un tout petit village –, les gens ont à peu près la même culture, il y a très peu de produits disponibles dans la supérette. Si je cherche des motifs dans les comportements, il ne va pas y en avoir beaucoup et je vais être capable de comprendre le nombre relativement faible de comportements différents.

Si je passe à un hypermarché parisien, il y a beaucoup plus de produits, beaucoup plus de choix, beaucoup plus de gens qui ont tous des cultures très différentes. Les comportements d'achat vont se démultiplier en un nombre énorme. Là, déjà, nous sommes ennuyés. Mais si je passe au commerce mondial, si je passe à un hypermarché au niveau mondial, cela devient Amazon. Là, vous vous doutez bien qu'il y a un nombre encore plus gigantesque de comportements différents possibles : l'algorithme marche, mais l'humain, derrière, doit analyser beaucoup de résultats : il peut y avoir plus de résultats que de données en entrée ! C'est un problème qui n'est pas si simple que cela à régler. Il y a beaucoup de travaux à faire dans ce domaine, qui sont au cœur des développements actuels dans le *Data Mining*.

Notre but, chercheurs en *Data Mining*, va être de pouvoir offrir le bouton qui vous permet d'extraire les données que vous voulez – c'est un but un peu utopique. Nous avons besoin de travailler avec les utilisateurs pour voir avec eux ce qu'ils veulent trouver dans les données (ils ne le savent pas toujours, au début).

D'une façon générale, pour conclure, et cela s'applique donc aussi dans le cas de l'agriculture, le *Data Mining* s'utilise désormais dans un contexte où l'on s'efforce de trouver les moyens de co-explorer les données, d'offrir des outils très

puissants d'exploration de données mais en donnant à l'utilisateur son mot à dire pour savoir où il veut aller, sachant ce qu'il peut trouver dans ces données.

M. Jean-Yves Le Déaut. Merci beaucoup. Pour terminer cette première partie pédagogique (nous allons aborder l'agriculture par la suite), Monsieur Xavier Vincelot, qui est *Datascientist Senior* au cabinet Quantmetry, spécialisé dans le traitement massif des données. Je compte sur vous, Monsieur, pour nous donner à travers divers exemples que vous connaissez bien une illustration des types de produits et services qui reposent sur le *Big Data*. À vous la parole.

M. Xavier Vincelot, *Datascientist Senior*, cabinet Quantmetry. Merci Monsieur le Président. Je vais essayer de compléter, mais je crois que la plupart des choses ont été dites. Je vais peut-être commencer par resituer les applications dans leur contexte.

Il faut avoir conscience que le *Big Data* est une technologie qui a émergé dans les années 2000 grâce aux géants du Web qui à l'époque étaient confrontés à un problème : avec la croissance d'Internet et cette explosion de nouvelles technologies et d'afflux de données, ils avaient besoin de préserver leur *business model*. Du coup, ils ont trouvé des technologies qui leur permettaient de traiter toujours davantage de données sans que le « coût marge nette » de la donnée n'augmente. Notre définition : le *Big Data* est une technologie inventée par les géants du Web pour les géants du Web.

Ces technologies qu'ils ont introduites ont plusieurs propriétés. Notamment, elles sont *scalables*. Si demain le volume de données est multiplié par dix, on n'aura pas besoin de jeter sa machine et d'en prendre une nouvelle : on pourra juste augmenter l'architecture, donc le coût de traitement ne va pas croître exponentiellement. Il va juste être multiplié par le volume de données.

Il y a d'autres aspects technologiques importants, notamment la tolérance aux pannes qui est essentielle. Il ne faut pas que les serveurs « plantent », mais quand on a des milliers et des milliers de serveurs, le « plantage » fait partie du système. Toutes les minutes, des disques « crashent », et il faut l'avoir intégré.

Du coup, il s'est développé beaucoup d'outils à très bas coût. Ces outils ont bénéficié du support des communautés *Open Source* (notamment le plus connu : la fondation APACHE), puis ont été diffusés dans l'industrie. Comme nous l'avons dit, les objets connectés fournissant de plus en plus de données, nous nous sommes retrouvés avec une technologie qui permettait de traiter beaucoup plus de données, à des coûts beaucoup plus faibles. C'était une opportunité superbe et qui touchait surtout beaucoup de secteurs.

C'est important de prendre en compte que le *Big Data* touche à la fois beaucoup de secteurs (les assurances, la finance, les transports, l'urbanisme, l'industrie), mais aussi beaucoup de fonctions (le marketing, la logistique, la finance, et même les ressources humaines).

Ce sont souvent les thématiques marketing qui servent d'illustration aux possibilités ouvertes par ces évolutions : grâce au *Big Data*, on va pouvoir travailler, approfondir toutes les notions d'acquisition de clients, de rétention, de fidélisation. Je prends le cas d'une assurance : elle ne va pas traiter les dossiers uniquement grâce aux informations que vous donnez. Elle va pouvoir établir une vision beaucoup plus approfondie en essayant de capter le plus possible de signaux issus de sources qu'on n'avait pas l'idée auparavant d'exploiter.

Par exemple, quand vous êtes sur Internet, toute votre navigation, toutes vos actions sur une page Internet vont être sauvegardées : la rapidité avec laquelle vous allez remplir le formulaire, les documents scannés que vous allez envoyer vont pouvoir être exploités facilement ; les données de l'*Open Data*, les renseignements que vous communiquez sur vous sur les réseaux sociaux vont pouvoir être aussi exploités. On va pouvoir mieux vous cerner, et ainsi vous adresser des produits plus ciblés, ou essayer d'anticiper quand vous allez quitter le contrat d'assurance par exemple, et effectuer les actions idoines au bon moment avec le maximum de précision.

Au niveau de l'interaction avec le client, énormément de choses évoluent. On va pouvoir intégrer beaucoup plus de sources de données qu'auparavant, traiter beaucoup plus d'historiques, avoir une connaissance du client bien plus approfondie.

Heureusement, il y a d'autres thématiques que les thématiques marketing. Par exemple, nous travaillons aussi dans les transports. Aujourd'hui, les wagons de trains par exemple ont beaucoup de capteurs, récoltent beaucoup de signaux. En traitant tous ces signaux qui arrivent, on va essayer de trouver, d'anticiper les pannes. On va chercher les motifs statistiques qui se répètent, pour identifier avec une grande précision le matériel défectueux, et prédire la panne. Ainsi, on va pouvoir effectuer de la maintenance prédictive. Il se conduit actuellement beaucoup de travaux autour de la détection d'anomalies.

Pour conclure, un sujet peut-être plus d'actualité ou avec un plus grand enjeu, ce sont les thématiques autour des *smart cities*, les villes intelligentes. Aujourd'hui, la donnée est en train d'être récoltée. On a déjà bien avancé sur ces sujets. On a les moyens technologiques pour traiter cette donnée, pour essayer de détecter des signaux faibles, les signaux cachés qu'auparavant on n'était pas capables de trouver. Maintenant, il y a réellement besoin d'une volonté pour mettre en place des actions à partir de cela, pour essayer d'avoir une vision davantage « pilotée par la donnée », une vision plus quantitative qui s'appuie vraiment sur toutes ces données que nous sommes en mesure de récolter aujourd'hui.

M. Jean-Yves Le Déaut. Merci beaucoup. Nous allons passer tout de suite à la partie agriculture (puisque nous avons fini la partie pédagogique sur le *Big Data*), avec la question : comment le *Big Data* a-t-il pris pied dans l'agriculture ?

Néanmoins, sur les interventions de cette première table ronde, y aurait-il, sans évidemment lancer un débat, un point que vous voudriez compléter, un élément qui aurait été oublié ? Seul Laurent Gouzènes réagit ? Je vais le laisser se présenter, car c'est un membre de notre conseil scientifique. L'Office parlementaire a un conseil scientifique de 24 membres qui représentent toutes les disciplines. Commencez par vous présenter un peu plus, Laurent.

M. Laurent Gouzènes, membre du Conseil scientifique de l'OPECST et expert scientifique du groupe Pacte novation. Merci. Je suis informaticien et j'ai travaillé aussi vingt ans chez *STMicroelectronics*. J'ai beaucoup travaillé sur les sujets d'objets microscopiques.

Je voudrais revenir sur les enjeux technologiques du développement du *Big Data*. En premier lieu, je constate que les sources de données se multiplient. Nous allons trouver les capteurs *in situ* qui délivrent des informations à l'instant T, mais toujours au même point, en continu. Nous avons d'autre part des capteurs de nature agrégée, c'est-à-dire typiquement des vues d'avion ou des vues du spatial. C'est d'ailleurs un sujet intéressant dans le contexte de cette audition.

M. Jean-Yves Le Déaut. Nous en parlerons à la fin de la troisième table ronde.

M. Laurent Gouzènes. Nous avons une troisième source qui est l'ensemble de machines, en l'occurrence les machines agricoles, qui fournissent des informations de suivi de leur *process*.

Or nous avons besoin d'infrastructures de captage de ces informations (pour tous ces objets connectés, il faut bien qu'on rapatrie les informations qu'ils accumulent), et des infrastructures de stockage et d'analyse. Il faudra traiter tous ces volets-là.

Les progrès techniques sur les objets connectés devront intégrer la problématique de l'énergie : si pour faire fonctionner un équipement pendant dix, quinze, vingt ans, le plus longtemps possible, il faut mettre beaucoup de capteurs et les renouveler en permanence, cela peut finir par coûter très cher. Dans cette logique, il sera utile de développer l'ensemble des protocoles basse consommation qui permettront de récupérer les données collectées, sans consommer d'énergie.

Un exemple de mise en œuvre très récent, c'est le pilotage automatique des drones. Je ne sais pas si vous avez fait du pilotage de petit avion quand vous étiez gamins ou même aujourd'hui, mais c'est très difficile de piloter un avion. Aujourd'hui, nous avons des outils qui permettent de faire voler un avion tout seul et de le faire revenir à sa base. C'est un très grand progrès technologique, en plus de lui donner une vraie mission et de le faire revenir sans qu'il y ait de casse, à peu près en sécurité.

Tout cela repose sur des algorithmes, nous en avons parlé. Je voulais confirmer que nous pouvons désormais, grâce à eux, déduire des éléments qui ne sont pas, de façon évidente, dans les données. Un exemple connu est celui d'EDF qui, en analysant votre courbe de consommation, sait quand le frigo fonctionne, quand la télévision se met en route, quand on se met sur Internet, *etc.* C'est complètement traçable. On peut récupérer des informations très détaillées sur le comportement d'une maison. On le fait aussi dans les ateliers de fabrication : quand on suit des machines, on peut déterminer exactement leur productivité. C'est très intéressant. C'était le petit point technique que je voulais faire.

M. Jean-Yves Le Déaut. Merci Laurent.

**DEUXIÈME TABLE RONDE :
COMMENT LE *BIG DATA* A-T-IL PRIS PIED DANS L'AGRICULTURE ?**

Présidence de M. Jean-Yves Le Déaut, député, président de l'OPECST

M. Jean-Yves Le Déaut. Nous passons tout de suite à la partie agriculture qui va nous permettre de comprendre comment les *Big Data* peuvent fonctionner dans le domaine de l'agriculture, notamment en identifiant les données connectées et les services fournis. Ce qui nous intéresse là, c'est de dresser un état de la situation, pour mieux nous donner les bases d'une analyse stratégique en troisième partie.

Je vais donner la parole d'abord à M. Hervé Pillaud, secrétaire général de la Chambre d'agriculture de Vendée, et membre de la commission « Élevage » de l'Assemblée permanente des chambres d'agriculture (APCA), pour qu'il nous dresse un panorama des services rendus par le *Big Data* dans l'agriculture tels qu'ils sont d'ores et déjà déployés en France, et qu'il nous pose la problématique.

M. Hervé Pillaud, secrétaire général de la chambre d'agriculture de Vendée et membre de la commission « Élevage » de l'APCA. Merci Monsieur le Président. Je pense qu'ils ne sont pas encore véritablement déployés. Ils vont se déployer. À ce titre, si vous voulez bien, je voudrais revenir une seconde sur la table ronde à laquelle nous venons d'assister, pour aborder cela d'une autre façon. Je pense que le numérique est à la fois une révolution, et une rupture qui nous amène vers une renaissance.

Pourquoi une révolution ? Parce qu'il impacte trois choses : la multitude d'innovations qu'il nous amène chaque jour opèrent des changements fondamentaux en matière de communication, et également en matière de captation et d'agrégation de données. C'est l'objet de nos travaux de ce matin.

Une rupture, pourquoi ? Parce qu'il remet en cause les notions d'espace. Il n'y a plus de frontières, il se crée une espèce de flou à ce niveau-là ; il remet en cause la notion de temps, également fondamentalement revu ; et enfin, à travers tout cela, la notion de gouvernance, je pense, va être remise en cause.

Pourquoi une renaissance ? Non pas uniquement parce qu'il faut finir par une note optimiste, mais parce que les repères qui organisent la vie de la cité sont déstabilisés (nous voyons déjà apparaître de nouveaux paradigmes ici et là). Nous ne sommes pas forcément si mal, à la fois chez nous en France et également dans l'agriculture.

Pour l'agriculture, je pense que c'est une véritable opportunité. L'agriculture est un peu à la croisée des chemins. Marcel Deneux l'a dit tout à l'heure : nous avons vécu une première révolution, peut-être une deuxième. Je

pense que nous entrons dans une troisième. Nous allons passer d'une agriculture utilisatrice d'intrants à une agriculture utilisatrice de connaissances.

Pourquoi ? Pour différentes raisons. Il y a les règlements mais surtout, un certain nombre d'intrants sont en quantité de plus en plus rare et il va falloir trouver d'autres éléments. Si ces intrants sont de plus en plus rares, en revanche, s'agissant des connaissances dont nous pouvons disposer, nous sommes à l'aube de ce que nous pouvons découvrir. Là, le numérique est une véritable opportunité. Il nous permettra d'en agréger bien au-delà de ce que les cerveaux humains sont capables de faire.

Vous l'avez dit tout à l'heure : les agriculteurs ont toujours connu leur ferme, leur petit univers. Maintenant, nous avons l'opportunité de la partager avec des agriculteurs du monde, avec ce que nous apportent les machines, ce que nous apportent les capteurs, pour, à partir des algorithmes dont on nous a parlé précédemment, créer de nouvelles possibilités.

Je suis personnellement convaincu qu'avec beaucoup moins d'intrants, nous allons pouvoir augmenter considérablement les rendements. C'est pour cela que je vous dis que nous ne sommes qu'à l'aube de ce qui va pouvoir se faire.

D'ores et déjà, beaucoup de choses sont automatisées. L'alimentation des animaux, la traite des animaux sont déjà automatisées. Ce n'est pas généralisé, mais c'est largement vulgarisé. C'est émetteur d'une quantité énorme de données. Il en va de même pour ce qui est de la pulvérisation, de l'irrigation, qui sont pilotées par des capteurs, par une quantité énorme de possibilités.

On parle souvent des drones, des satellites, des capteurs embarqués, comme si ces choses-là étaient indépendantes. Elles ne sont pas indépendantes et justement le *Big Data* va permettre, par les algorithmes, d'agréger les données qu'elles fournissent pour nous ouvrir beaucoup plus de possibilités.

En revanche, il y a un problème peut-être en agriculture par rapport à d'autres métiers : chaque exploitation est une TPE, une microentreprise. Chaque agriculteur n'aura pas la capacité d'agréger lui-même une quantité suffisante de données pour pouvoir les exploiter. Nous sommes donc condamnés (le terme n'est pas bon !) à travailler en groupe là-dessus, mais nous avons toujours également su le faire. Les réseaux sont culturels en agriculture, particulièrement en France, depuis des années. M. Marcel Deneux l'a dit dans son exposé préalable.

Pour moi, il y a six domaines dans lesquels le *Big Data* va impacter l'agriculture.

Le premier, c'est la gestion des risques. Nous exerçons un métier où l'on passe son temps à gérer les risques pour dompter la nature, pour en tirer la quintessence. Le numérique va nous permettre d'aller plus loin, à la fois dans la prévision mais également dans la gestion. À ce titre-là, *Monsanto* a racheté une très belle *start-up* de la *Silicon Valley* qui s'appelle *Climate Corp.*, 1 milliard de

dollars, pour en faire sûrement un géant du domaine de l'assurance. Cette *start-up* est capable de porter assistance aux agriculteurs sinistrés beaucoup plus vite qu'aucune autre compagnie d'assurances.

Deuxièmement c'est le financement. Le financement de l'agriculture va évoluer aussi et là, le *Big Data* donne des indications sur ce qu'il y a dans une ferme. Il y a un point auquel il faut faire attention : c'est l'utilisation des données. Je ne parlerai ni de protection, ni de propriété. Nos confrères américains du *Farm Bureau* ont déjà travaillé là-dessus. Au niveau de la FNSEA, un travail se fait avec eux pour regarder ce que nous pouvons en tirer.

Troisième élément : la recherche et le développement vont également être fondamentalement revus. Certains en reparleront, je pense, et sont beaucoup plus connaisseurs que moi du domaine du *smart agri*, des objets connectés au service de l'agriculture. Là, il faut bien veiller à ce que l'agriculteur reste maître de ses décisions, qu'émerge une agriculture connectée au service de décideurs connectés, que nous ne perdions pas cette dimension-là.

Le conseil et la formation seront fondamentalement revus, à partir des outils d'aide à la décision (OAD) qui vont s'appuyer sur l'exploitation de toutes les données disponibles.

Et enfin l'approche du marché va se modifier, car le marché est en train de devenir à la fois très local et très global. Là aussi, l'exploitation des données nous permettra d'appréhender différemment ce nouveau contexte. L'enjeu est très important pour les agriculteurs, qui devront apprendre à le maîtriser, mais je reviendrai tout à l'heure, en remplacement de Michel Masson, pour vous donner quelques pistes sur la façon dont nous voyons les choses.

M. Jean-Yves Le Déaut. Merci beaucoup Monsieur Pillaud. Je donne maintenant la parole à Jean-Louis Peyraud, qui est chargé de mission à l'INRA et qui travaille au sein de la direction scientifique et agriculture, en lui demandant de nous donner des exemples de capteurs utilisés dans l'agriculture pour alimenter en données les services s'appuyant sur le *Big Data*.

M. Jean-Louis Peyraud, chargé de mission, direction scientifique agriculture, INRA. Merci, Monsieur le Président. Mon exposé est basé sur quelques diapositives ⁽¹⁾.

La première permet de repositionner un peu notre sujet sur la place des capteurs, du *Big Data*, de la robotique au sein de l'exploitation agricole : elle rappelle qu'historiquement, l'agriculteur observait ses animaux, observait ses champs, et en fonction de ce qu'il observait, prenait des décisions pour agir.

(1) Voir Annexe 2.

Aujourd'hui, avec les capteurs, nous avons au niveau de l'agriculteur une masse de données qui arrivent. Les capteurs mesurent ces données. Des logiciels intégrés (ou pas) aux capteurs permettent de transformer ces données observées en données prédites, de fournir des informations, qui sont, en retour, utilisées par l'agriculteur pour prendre ses décisions, éventuellement à travers des outils d'aide à la décision pour le conseiller par exemple sur les niveaux de fertilisation compte tenu du passé cultural de sa parcelle.

S'ajoute à cela aujourd'hui un deuxième niveau de sophistication qui est la robotique. Les capteurs sont aussi capables de piloter directement des robots. Vous voyez ici un robot qui travaille sous serre, un tracteur connecté, et évidemment l'emblématique robot de traite ; tous permettent d'accomplir des actions cette fois-ci complètement automatisées ce qui, du coup, modifie la forme de la prise de décision : l'agriculteur agit en spécifiant des seuils d'action à ses robots, mais n'intervient plus directement.

Et évidemment, à travers le *Big Data*, aujourd'hui il est possible de mettre en interrelation les différentes exploitations agricoles, de faire du *benchmarking* entre exploitations agricoles, de les comparer, et de prendre des décisions compte tenu de ces informations. Je ne reviens pas sur tout ce qui s'est dit autour du *Big Data* et du *Data Mining* qui a toute sa place ici.

L'ensemble de ces informations peut aussi servir à émettre des alertes (nous parlions tout à l'heure de résilience des systèmes) : alertes sanitaires – nous voyons arriver peut-être plus précocement que par le passé des aspects sanitaires – ou bien sûr des alertes météo, nous en avons parlé en évoquant *Monsanto* juste à l'instant.

Voilà le nouveau cadre dans lequel peut fonctionner une exploitation agricole : avec les capteurs, la robotique, le *Big Data* et les outils d'aide à la décision, c'est un centre stratégique de prise de décision.

Les capteurs se retrouvent maintenant partout en agriculture. J'en ai sélectionné ici quelques exemples. Nous pourrions avoir des dizaines de diapositives comme celle-ci, sur le suivi des rendements et de la qualité, avec éventuellement des observations par drone ou par satellite qui permettent de faire, en cours de végétation, des états des parcelles, de voir des hétérogénéités intra-parcellaires. Je vous ai symbolisé ici les machines : à travers une récolte de blé ou l'ensilage de maïs pour les fourrages, nous avons en temps réel le rendement de la parcelle, les irrégularités de rendements dans la parcelle, la composition (éventuellement teneur en matière sèche, en protéines) du produit avec des capteurs infrarouges. Au niveau de la salle de traite, nous avons en temps réel des données sur les volumes de lait, la composition du lait, *etc.*

Un deuxième groupe très important, concerne les systèmes de localisation de traitement au champ. Ce sont des systèmes qui sont aujourd'hui vendus clés en main par exemple par *John Deere*. À travers une demande de photographie

satellitaire ou de drone, on peut avoir une information sur l'hétérogénéité des parcelles. Cette information est traitée, elle est ensuite retransmise à la cabine du tracteur qui finalement module les apports d'engrais, de traitements phytosanitaires, avec des doses variables selon les zones dans les parcelles.

Un exemple très emblématique, je trouve, concerne tout ce qui est la régulation d'ambiance dans les bâtiments d'élevage, notamment pour la volaille, mais aussi sous serre, à travers des capteurs mesurant l'hygrométrie, les températures, la qualité du sol. On arrive à ajuster l'éclairage, la température, la ventilation dans les bâtiments.

Autre exemple également très emblématique : aujourd'hui, on sait mesurer beaucoup d'indicateurs sur les animaux, en particulier sur les vaches laitières. Il y a une profusion d'appareils qui existent dans la littérature : on mesure les déplacements des animaux, les positions (debout, couché), les températures, les rythmes cardiaques. Quelqu'un a dit que l'on se rapprochait de la médecine : oui, complètement. Au niveau du monde animal, du bien-être des animaux, nous disposons aussi désormais d'indicateurs. Avec les robots de traite, il y a plusieurs dosages qui peuvent être effectués en temps réel sur le lait, et qui permettent d'anticiper, par exemple, des mammites.

Cela permet aussi du progrès génétique. Ici, je vous ai représenté la « Phénomobile », véhicule qui parcourt de manière autonome les parcelles et recueille des informations sur l'état des plantes grâce à des capteurs embarqués. Des travaux se font à l'INRA, chez ARVALIS (Institut du végétal) aussi, pour observer en plein champ les phénotypes des cultures, leur forme par exemple. De même pour les animaux, où les données des robots de traite repartent directement chez les fabricants, à l'insu de l'éleveur. Les fabricants de robots sont ceux qui disposent du plus de données pour faire du phénotypage à haut débit sur les vaches laitières. Ce ne sont plus les centres de sélection qui ont la main ; on change d'acteur. Cela soulève des questions de propriété de données, de leur utilisation.

Et pour la recherche, il y a bien sûr, au niveau du *Big Data*, des enjeux énormes d'acquisition de connaissance entre les données qui viennent des sciences en *omics* (génomique, métabolomique ⁽¹⁾...) et celles qui viennent de ces capteurs.

Dernier exemple : les réseaux d'épidémiosurveillance qui permettraient, à travers des réseaux de capteurs disséminés sur le territoire, de voir apparaître de façon précoce des pathologies ou des risques variés.

En guise de conclusion, je voudrais dire que tout cela soulève des questions. Nous avons déjà parlé d'interconnectivité des données, de cybersécurité, de propriété de données. J'ajouterai à cela que nous allons vers une évolution très forte du conseil en agriculture, avec l'apparition de nouveaux acteurs. Aujourd'hui, *John Deere* est un conseiller en agriculture très présent pour

(1) Science qui étudie l'ensemble des métabolites (sucres, acides aminés, acides gras, etc.) présents dans une cellule.

celui qui s'est équipé chez *John Deere*. Des sociétés comme IBM rachètent aussi des *start-up* pour faire du conseil en agriculture. Nous voyons que le monde évolue très vite, ce qui nous amène à nous interroger sur la façon dont nous aurons à travailler demain.

Un autre enjeu devient aussi très important : c'est celui de la couverture en haut débit du territoire. Tout ce transfert d'information nécessite du haut débit. Là, on risque de se heurter à des questions d'hétérogénéité au niveau de la France. Ce sera peut-être plus simple d'obtenir une bonne couverture quand on sera céréalier près de Chartres que quand on sera éleveur en zone de montagne.

M. Jean-Yves Le Déaut. Merci beaucoup de cette question dont nous allons discuter dans la conclusion tout à l'heure.

Je donne tout de suite la parole à Stéphane Grumbach, Directeur de l'Institut rhônalpin des systèmes complexes, et chef (j'espère que le titre est juste, cette fois), au sein de l'INRIA, de l'équipe qui se consacre aux données de l'Internet au cœur de l'économie, la DICE.

Comme Serge Abiteboul, vous aviez participé en février 2013 à l'audition publique sur les risques numériques, notamment pour dresser un panorama stratégique de la maîtrise des données que vous pourrez transposer au cas du *Big Data* au cours de notre troisième table ronde. Pour le moment, vous allez illustrer le fonctionnement du *Big Data* dans l'agriculture à travers le cas de la société *Climate Corporation*.

M. Stéphane Grumbach, directeur de l'IXXI, Institut rhônalpin des systèmes complexes, chef de l'équipe INRIA DICE, INRIA Rhône-Alpes. Merci Monsieur le Président. Nous avons parlé de l'évolution des technologies. Nous avons parlé également des données. C'est le sujet de cette audition. Nous en parlons sous le vocable « *Big Data* », qui désigne la capacité à gérer des quantités de données astronomiques, et d'en extraire des connaissances extrêmement riches.

Je voudrais mettre l'accent maintenant sur le « *Big User* », c'est-à-dire la capacité à développer des systèmes qui touchent les utilisateurs en nombre astronomique. Certains de ces systèmes touchent un nombre d'utilisateurs qui est à peu près l'ensemble des habitants de la planète. Les grands systèmes, comme vous le savez, dépassent le milliard d'utilisateurs.

Je crois que cette focalisation sur l'utilisateur est importante. Elle fait fortement défaut dans le débat en Europe. Les sociétés engagées dans le *Big Data* qui connaissent le plus de succès, d'une part au niveau économique, en étant capables de lever de grandes quantités de capital-risque, en connaissant des ascensions fulgurantes de leur capitalisation, d'autre part au niveau sociétal, en réalisant des percées disruptives qui les amènent à se trouver en conflit sur les territoires avec les États, la justice ou d'autres sociétés, sont toutes des sociétés qui cherchent à acquérir de très nombreux utilisateurs. Au niveau des données, elles combinent une analyse d'une part, sur les données accessibles en général (je pense

par exemple aux données du Web ou aux données de la météorologie qui vont nous intéresser pour l'agriculture), et d'autre part, sur les données de leurs utilisateurs.

La plus grande et la première entreprise de ce secteur est un moteur de recherche, et je voudrais rappeler comment elle a construit sa position, car cette construction est fondamentale, et *Climate Corporation* a été exactement construite sur ce modèle. Ce moteur de recherche a eu à sa disposition les pages Web librement accessibles. Il a traité ces données pour offrir le service d'une recherche efficace et pertinente, parmi un océan de pages Web. À partir de ce service (qui s'il est bien conçu, va être adopté par de très nombreux utilisateurs), la société qui gère le moteur récupère les traces d'utilisation. À partir des traces d'utilisation (les données des utilisateurs), on peut générer de nouveaux services, et dans le cas du moteur de recherche en question, deux types de services : le service personnalisé sur lequel la société a basé son modèle économique de publicité, et les services d'extraction de connaissances, dont Google a démontré l'efficacité.

Uber s'est construit sur ce modèle, et *Climate Corporation* fonctionne exactement de la même manière. C'est une société qui a été créée à San Francisco en 2006 par deux ingénieurs de chez Google, dont l'objectif était de fournir aux secteurs industriels qui dépendent de l'aléa météorologique une information extrêmement précise, extraite des informations météorologiques produites par les grandes agences.

Elle s'est spécialisée dans le domaine agricole où l'enjeu particulier de l'aléa climatique devient de plus en plus important. Elle a mis en place à la fois un flux descendant d'informations vers les agriculteurs, permettant de les conseiller en particulier sur l'arrosage, mais aussi, grâce à l'installation de capteurs dans les exploitations agricoles, un flux montant d'informations.

C'est cette combinaison entre cette information détaillée du terrain et l'information globale qui fait sa force. Elle a annoncé récemment sur son site ses derniers résultats. Le nombre d'hectares de terres cultivées en lien avec les services de *Climate Corporation* a augmenté de 50 % entre 2014 et 2015, passant de 20 millions d'hectares à 30 millions d'hectares. Sa part de marché atteint aujourd'hui, dans le domaine céréalier, pour ce qui est du maïs et du soja, 45 % des surfaces cultivées.

Comme l'a indiqué Monsieur Pillaud, et c'est un aspect absolument fondamental de ce genre de plate-formes, les services de base permettent la définition de nouveaux services. *Climate Corporation* est devenue société d'assurance. Elle assure les agriculteurs dans le domaine céréalier. Elle n'a pas besoin de faire appel à une expertise pour évaluer les dommages puisqu'elle dispose de suffisamment de données pour cela. Elle fait l'économie de l'expertise, elle peut agir et indemniser l'agriculteur de manière directe.

On comprend dès lors le potentiel et le pouvoir d'une société comme *Climate Corporation*, rachetée par Monsanto, comme Monsieur Pillaud l'a rappelé, pour un milliard de dollars, conférant ainsi à Monsanto un pouvoir absolument impressionnant. Je reviendrai tout à l'heure sur ce qu'il convient de faire, mais il est clair que, même si *Climate Corporation* s'est développée sur le territoire américain, elle a bien sûr une visée internationale, comme l'ensemble des plates-formes américaines. Je vous remercie.

M. Jean-Yves Le Déaut. Merci beaucoup. Vous soulevez des questions importantes. 30 millions d'hectares, ce que vous indiquez là, c'est deux fois les grandes surfaces françaises, les grandes cultures. Cela signifie que cela va se développer de manière très forte dans les prochaines années.

Je donne maintenant la parole à Laure-Marie Neuburger, ingénieure et titulaire d'un doctorat d'AgroParisTech qui l'a conduite à travailler sur la mise au point d'un biocapteur. Elle s'est lancée récemment, cela a été dit en introduction, dans la création d'une entreprise pour développer une offre de services utilisant les technologies au service de l'agroécologie et de l'environnement. Nous comptons sur vous pour nous donner quelques exemples d'innovations récentes en agriculture, nous indiquer s'ils se créent des *start-up* en France, et si nous sommes au niveau de ce qui se passe dans d'autres pays. Au niveau international, vous avez aussi un éclairage sur ce qui se passe, je crois, en Israël, et vous allez peut-être nous en dire un mot. Merci.

Mme Marie-Laure Neuburger, entrepreneur. Merci beaucoup. Je vais vous donner rapidement ma vision, étant tout juste arrivée dans l'écosystème de l'entrepreneuriat, puisqu'il m'a été proposé de faire un tour d'horizon rapide du monde entrepreneurial des *start-up* ⁽¹⁾.

Pour commencer ce tour du monde, sans originalité, je vais vous proposer d'aller directement aux États-Unis. Là-bas, *Big Data* et agriculture, comme nous l'avons vu, sont en confrontation directe.

Google Venture, fonds d'investissement de Google, a investi récemment dans la *start-up Farmers Business Network* qui a levé au total 28 millions de dollars. C'est un réseau social pour agriculteurs qui axe sa communication sur la sécurité des données. Globalement, c'est aux États-Unis et plus particulièrement en Californie, que sont investis la majorité des fonds de capital-risque, tous secteurs confondus. Il en va de même dans les *AgTech*, c'est-à-dire les technologies de l'agriculture et de la nourriture. Un rapport récent (*AgFunder*) fait état de 2 milliards de dollars investis pour l'année 2014 en capital-risque dans les *AgTech* aux États-Unis, dont la moitié uniquement dans l'État de Californie.

Ainsi, avec les États-Unis, ce tour du monde semble presque déjà terminé ; on peut avoir l'impression que tout se passe dans la vallée de Salinas, où se tient d'ailleurs la semaine prochaine le *AgTech Summit* « Réinventer l'agriculture ».

(1) Voir Annexe 3.

Mais c'est sans compter les écosystèmes bouillonnants de *start-up* que l'on peut trouver un peu partout dans le monde. Peut-être connaissez-vous par exemple, la *Ladybird*, qui est un robot de culture autonome fonctionnant à partir de capteurs solaires, développé en Australie ?

L'investissement dans les *AgTech* a en effet une dimension véritablement internationale, et connaît effectivement une forte croissance. Plus de vingt nouveaux fonds se sont créés depuis le début de l'année pour financer ce domaine de l'*AgTech*, notamment en Asie, en Chine particulièrement, première plate-forme de l'e-commerce au monde, comme l'illustre la création de ce fonds sino-russe annoncé récemment qui va mobiliser 2 milliards de dollars pour l'agriculture.

Les secteurs les plus ciblés sont la distribution de produits et la vente en ligne, la robotique, les capteurs, les objets connectés, la culture hors-sol, l'hydroponie, l'aquaponie, et l'agriculture de précision qui a concentré 12 % des investissements selon le rapport *AgFunder* en 2014, soit 284 millions de dollars.

Personnellement, je pense que cette tendance concerne aussi d'autres activités émergentes comme le mouvement *Maker*⁽¹⁾, ou l'innovation ouverte en règle générale, ainsi que l'agro-écologie. Nous voyons aussi le développement de l'agriculture personnalisée, comme évoqué précédemment, au même titre que la médecine personnalisée. Par exemple, poussés par le même besoin que la Californie pour vaincre l'aridité, Israël est, comme vous l'avez dit, en pointe dans l'agriculture de précision et notamment l'irrigation, aussi bien avec des acteurs historiques (*Netafim*), qu'avec des *start-up* (*Crop*). Israël est *leader* dans le système de la laiterie, avec une production moyenne de 11 900 litres par an et par vache, un rendement 10 % supérieur à celui en moyenne des États-Unis, grâce notamment à des robots de traite connectés (*Afimilk*). C'est un pays très en pointe dans l'aquaponie⁽²⁾, l'hydroponie⁽³⁾, secteur tiré notamment par la culture du cannabis à des fins médicinales.

Dans les autres écosystèmes nationaux, l'ouverture et la communication sont les moteurs du développement. Parmi les projets liés à l'*Open Data*, nous pouvons citer le projet *Groenmonitor* aux Pays-Bas, ou le CIAT en Colombie qui utilisent des données ouvertes, publiques et privées concernant des études sur le riz pour constituer une offre d'information en matière d'itinéraires techniques, de rendements, de données météorologiques, pour développer un outil d'aide à la gestion agricole, non seulement utile pour les producteurs de riz, mais aussi disponible pour tous. Autre exemple, en Afrique : les téléphones mobiles sont moteurs de l'innovation. Des plates-formes de paiement, mais aussi des services aux agriculteurs *via* les téléphones mobiles, sont de plus en plus répandus. Bref,

(1) Il s'agit d'un mouvement organisé depuis 2006 autour de manifestations diverses (ateliers, présentations, conférences) mettant en valeur les thèmes de la créativité et de la fabrication « Do it yourself » (Faites-le vous-même). Ses adeptes y viennent pour présenter leurs réalisations et partager leurs connaissances.

(2) Culture de végétaux dans un aquarium, les déjections des poissons servant d'engrais.

(3) Méthode de culture sans sol, utilisant des solutions nutritives.

les innovations sont partout : au Canada, en Chine, en Europe, et comme nous l'avons vu, en Australie.

Tout à l'heure, je crois avoir entendu le mot « révolution ». Oui, l'idée est bien d'introduire une innovation disruptive, perturbante, voire une innovation de rupture.

Mais qu'est-ce qui motive ces innovations ? Une *start-up* dans l'*AgTech* se projette dans un univers où ces milliards d'animaux sont connectés, où les pieds de vigne, voire les épis de maïs, twittent leur manque d'eau, où les tracteurs autonomes sont suivis à la trace. Pour 7,325 milliards d'êtres humains, on compte 22,9 milliards de volailles, 2 138 millions d'ovins et de caprins, 1,5 milliard de bovins, 980 millions de porcins, 25 millions de tracteurs (données FAO).

Les *start-up* les plus financées voient la donnée comme un outil, un objet, mais pour autant celle-ci n'est pas au centre de leurs propositions de valeurs.

Les modèles économiques d'exploitation des données sont tournés souvent autour de trois grands modèles : la revente de données, la publicité, la redirection de trafic. Mais ce qui caractérise bien une *start-up*, c'est la capacité à inventer et réinventer son modèle économique. Les données sont une matière brute qui nécessite généralement beaucoup de travail – comme nous l'avons vu – de nettoyage, d'analyse, afin de construire et d'inventer des modèles pertinents, dans le respect de règles éthiques. Aujourd'hui, trouver des moyens de capter la donnée pour l'exploiter utilement reste encore à mon sens un enjeu.

Mais à quoi peut conduire une stratégie privilégiant la donnée dans le domaine de l'agriculture ? Nous l'avons vu, dans le domaine marketing, cela a créé les nouveaux géants de la publicité numérique.

En ce qui concerne l'agriculture, on peut imaginer deux voies possibles d'évolution qui devraient coexister : la première, celle d'une agriculture recourant à une automatisation et une robotisation poussées à l'extrême, une agriculture moderne, productive, mais peut-être déshumanisée ; la deuxième, celle d'une agriculture aux impacts maîtrisés, peut-être par le biais de l'*impact investing*, une agriculture anticipant les aléas climatiques, l'appauvrissement des terres, alliant écologie et technologie. Merci.

M. Jean-Yves Le Déaut. Merci beaucoup. Vivien Mallet, qui est chargé de recherche à l'INRIA, va essayer de nous expliquer quel est l'apport des méthodes d'assimilation de données au contexte agricole. Peut-être pourriez-vous commencer par nous expliquer ce que l'on entend par l'assimilation de données.

M. Vivien Mallet, chargé de recherche, INRIA. Je vous remercie, Monsieur le Président. L'assimilation de données vise à exploiter de manière conjointe l'observation, c'est-à-dire les masses de données dont on parle aujourd'hui, et la simulation numérique, c'est-à-dire des modèles de culture, des

programmes informatiques qui vont simuler au cours du temps l'évolution des cultures et en particulier de leur biomasse, à différentes échelles.

Cela peut aller d'une simulation à l'échelle d'une exploitation, même d'une parcelle, jusqu'à la simulation à l'échelle d'un pays ou d'un continent pour les plus grands modèles. L'objectif de l'assimilation de données est d'essayer d'estimer, de la manière la plus précise possible, l'état réel des cultures, notamment la biomasse, son évolution au cours du temps et *in fine* le rendement des cultures.

On peut le faire sur le passé. On va ainsi générer des analyses. Et puis on peut utiliser ces techniques pour améliorer les prévisions. Ces techniques d'assimilation de données ont été beaucoup appliquées en météorologie – cela a commencé dans les années 1980 – et avec beaucoup de succès. On peut dire que, dans les vingt dernières années, les prévisions météorologiques se sont principalement améliorées du fait de l'assimilation des données, notamment des données satellitaires.

En agriculture, l'assimilation de données a fait une entrée plus récente, qui relève encore beaucoup du domaine de la recherche, avec l'objectif de corriger les calculs des modèles dont je parlais, qui vont notamment décrire les sols, modéliser la croissance des plantes, tenir compte des interactions avec l'atmosphère, et éventuellement intégrer les pratiques agricoles d'irrigation, de fertilisation.

Ces modèles, pour faire leurs calculs simulant la croissance des cultures, vont avoir besoin de paramètres sur les plantes (leur variété, les paramètres biologiques de ces plantes), de paramètres sur les sols comme la profondeur des sols, de paramètres météorologiques, ce qui concerne, bien entendu, les précipitations, le rayonnement, mais aussi les vents, la température, l'humidité, *etc.*

Les modèles de culture vont être ensuite capables de calculer l'évolution d'une culture, son développement foliaire et, à la fin, son rendement. L'objectif de l'assimilation de données va être de corriger ou d'améliorer les calculs faits par ces modèles-là en utilisant des observations, souvent satellitaires, qui vont indirectement mesurer le développement foliaire ou l'état de la biomasse.

Nous allons ainsi combiner les observations et ce qui a été calculé par le modèle pour avoir une meilleure représentation de l'état de la culture à un certain instant. À partir de cette description améliorée de la culture, le modèle va être capable de faire des prévisions de meilleure qualité à échéance de plusieurs semaines.

Une autre manière de faire de l'assimilation de données n'est pas de corriger directement ce que le modèle a calculé mais d'améliorer des paramètres que le modèle utilise, qui sont ici incertains. Un exemple : nous pouvons essayer de retrouver la date de semis lorsqu'elle est inconnue, de façon que lorsque le modèle fait des simulations avec cette date de semis améliorée, il reproduit mieux

ce que l'on a observé. Ainsi, on peut avoir une meilleure vision de ce qu'a été l'évolution des cultures sur le passé, et aussi faire des prévisions améliorées.

Vous avez compris : l'assimilation de données est cette combinaison, que l'on essaie de rendre optimale, entre d'une part, la simulation numérique qui va apporter une information assez complète, avec beaucoup de variables sur l'état des cultures, mais qui va être assez incertaine, et d'autre part, les observations, plus précises, mais incomplètes parce que l'on n'observe pas toutes les variables qui décrivent les cultures.

Par exemple, si vous faites une observation satellitaire, vous aurez des problèmes de couvertures à cause des nuages. Vous avez donc besoin des modèles, et il vous faudra essayer de tirer parti du meilleur des deux mondes que sont la simulation et l'observation.

L'assimilation de données permet de fournir non seulement des résultats, mais aussi une incertitude sur ces résultats. Cela peut être important lorsque l'on pense éventuellement au calcul des risques, notamment sur les rendements. Je voulais souligner que l'assimilation de données est capable d'utiliser des observations de différentes natures. Là, je souscris tout à fait à ce que disait Monsieur Pillaud : il faut être capable d'utiliser à la fois les données satellitaires et les données *in situ*. L'assimilation de données est capable d'utiliser toutes ces données hétérogènes, toutes les données dont Monsieur Peyraud a parlé, en prenant en compte la qualité de ces données-là, en donnant des poids différents selon l'incertitude associée à chacune de ces données.

Je voulais conclure en soulignant le succès du lancement du satellite Sentinel 2, la semaine dernière, qui va nous donner des images à une résolution très fine (de l'ordre de 10 mètres tous les dix jours). Il sera suivi d'un autre satellite qui permettra de réduire la période à cinq jours. Ces nouvelles données à haute résolution seront très intéressantes pour l'assimilation de données.

M. Jean-Yves Le Déaut. Merci beaucoup. J'en profite pour dire que notre prochaine audition, mardi 7 juillet, portera sur la politique spatiale. Nous y aurons l'occasion d'aborder un certain nombre des problématiques que vous venez d'aborder.

Jean-Pierre Chanet est directeur de l'unité de recherche technologies et systèmes d'information pour les agrosystèmes de l'IRSTEA. Vous allez nous dresser un panorama sur la réalité de l'utilisation des systèmes de *Big Data* dans l'agriculture française d'aujourd'hui : beaucoup de collectes de données, mais assez peu de traitement de ces données pour la fourniture de services. En tout cas, c'est un diagnostic qui a été fait ; je ne sais pas s'il sera partagé dans la discussion tout à l'heure. C'est une évolution technologique en devenir. Il faut veiller à ce que l'accès aux données reste ouvert pour ne pas freiner l'éclosion de l'offre. Les accès aux données sont-ils ouverts aujourd'hui ? J'ai entendu parler tout à l'heure

de *John Deere* qui devenait un conseiller très important. Cela ne pose-t-il pas le problème d'accès ouvert aux données ?

M. Jean-Pierre Chanet, chercheur et directeur de l'unité de recherche technologies et systèmes d'information pour les agrosystèmes d'IRSTEA. Merci Monsieur le Président. Tout d'abord, comme cela a été dit, l'agriculture a une longue tradition de gestion des données – nous parlions notamment de l'élevage en introduction. Je pense que le *Big Data* change surtout le rapport à la donnée pour les agriculteurs. Jusqu'à présent, ils avaient l'habitude de fournir volontairement leurs données dans un cadre réglementaire, à des fins d'aide à la décision, mais ils étaient toujours acteurs dans la fourniture de données.

Avec le *Big Data*, ce qui change, c'est que la donnée, ils la fournissent non seulement volontairement, mais aussi à leur insu, via les différents équipements de leur exploitation. Il y a vraiment un changement de paradigme par rapport à la donnée, pour l'agriculteur. Mais c'est une situation qui lui est familière, car il est très technophile en général. Quand les technologies arrivent dans le grand public, peu de temps après elles sont embarquées dans les machines d'exploitation (cf. les smartphones, les tablettes) ; tous les objets connectés arrivent très vite dans le secteur agricole. Ce n'est pas un secteur en retard par rapport aux technologies de l'information et de la communication. En revanche, le rapport à cette donnée est en train de changer complètement.

Je pense qu'il faut que les acteurs s'interrogent sur leur place dans l'écosystème et leur rapport à la donnée. Pour l'instant, ce sont les acteurs privés qui sont moteurs, majoritairement : les constructeurs, les fournisseurs de solutions logicielles ; ceux-ci ont pris cette vague du *Big Data* à temps, par rapport à d'autres secteurs. Ils sont vraiment là, ils se sont appropriés cette nouvelle dimension. Mais les autres acteurs agricoles, comme les instituts techniques, les agriculteurs, les conseillers, doivent à leur tour s'approprier réellement la question et analyser en quoi ce changement de paradigme de leur métier impacte leur rapport au conseil, à la décision.

Donc oui, il faut que les données restent relativement ouvertes – ouvertes ne veut pas dire gratuites – car l'agriculture n'est pas un *Big Data* comme on peut en voir dans le domaine du Web, par exemple. Nous sommes plutôt dans le cas des « petits *Big Data* » parce qu'il y a un cloisonnement. Chaque constructeur (qui n'a pas l'envergure des grands groupes du Web), chaque coopérative a ses propres données.

Si nous prenons le cas de l'agriculture, c'est un contexte très particulier. La donnée n'est pas générique, elle est très contextualisée. Il y a le contexte pédoclimatique ⁽¹⁾, le rapport au temps est également complètement différent par rapport au Web. Si nous considérons les rotations culturales, nous n'avons des retours d'expérience que tous les quatre ans. Nous avons une massification des

(1) Caractérisant les conditions extérieures au niveau du sol, du pied des plantes.

données moindre, par rapport au Web. C'est une dimension à prendre en compte aussi dans les traitements de *Big Data*.

Tous ces aspects doivent également être pris en considération, à mon avis, par la recherche. Cette dimension contextuelle est importante pour la prise en compte des données du *Big Data* dans la décision, dans les modèles. Nous l'avons vu précédemment : les questions d'assimilation de données vont prendre une place importante dans l'aide à la décision.

Jusqu'à présent, nous avons des modèles assez génériques qui étaient censés marcher dans des conditions communes aux différentes exploitations. Il va falloir désormais les contextualiser fortement, ce qui est souhaitable d'ailleurs, ce qui est pertinent, pour les rendre plus performants et plus efficaces.

Le savoir-faire des agriculteurs n'est pas que dans les données. Il comprend aussi une connaissance implicite. Nous allons entrer dans l'ère de la connaissance pour l'agriculture. Le *Big Data* est-il capable de faire émerger par la fouille de données, par ce type de choses, la connaissance implicite que les agriculteurs mettent en œuvre pour produire ? Cette dimension-là est importante car, de plus en plus, les agriculteurs ne sont pas issus de familles agricoles. Ce sont des nouveaux venus dans la profession. Il faut que ce savoir-faire puisse émerger et être transmis aux nouveaux agriculteurs.

Enfin, à mon avis le *Big Data* peut apporter des choses importantes au niveau de l'action. Nous l'avons vu, il va y avoir plein de robots, plein de choses automatisées. La boucle complète sera assurée depuis l'acquisition, la prise de décision, jusqu'à l'action. La dimension robotique en permettant de mieux travailler avec des intrants en moindre quantité, avec des applications beaucoup plus localisées, va permettre aussi à l'agriculture des marges de progrès importantes.

Mais je pense, pour conclure, qu'à l'heure où les technologies sont là (ou elles vont être là bientôt), il serait important que les acteurs créent un écosystème à l'échelon national, tous, acteurs privés, acteurs publics, compagnies, agriculteurs. Il faut que tout le monde s'interroge et voie comment on peut travailler ensemble sur ces données qui, au-delà de l'agriculture et de la dimension économique, ont des dimensions nationales, qui concernent tous les citoyens, notamment en termes de sécurité sanitaire, d'évolution du territoire, de compréhension du changement climatique par exemple. Ces données ont des potentiels au-delà des données purement agricoles. Elles ont un enjeu au delà du secteur agricole.

M. Jean-Yves Le Déaut. Pour terminer cette table ronde, je donne la parole à Sébastien Lafage, Directeur marketing et communication de la société ISAGRI qui fournit des solutions logicielles aux agriculteurs, pour la comptabilité comme pour l'assistance à la production.

Vous êtes bien placé pour répondre à la question que je posais juste avant concernant l'offre de marchés dans le domaine du *Big Data*, et nous indiquer si l'utilisation de ces techniques se diffuse dans votre clientèle.

M. Sébastien Lafage, directeur marketing et communication, ISAGRI.
Merci, Monsieur le Président. Juste deux mots pour présenter l'entreprise pour laquelle je travaille. Je représente aujourd'hui le groupe ISAGRI. C'est une entreprise familiale française. Nous sommes treizième éditeur de logiciels au niveau national et nous sommes leader européen sur le marché des solutions informatiques pour l'agriculture.

Les 1 400 salariés du groupe travaillent pour 115 000 utilisateurs de nos produits, principalement en France mais également dans les onze pays dans lesquels nous avons des filiales, vers lesquels nous exportons nos produits logiciels. Ces filiales sont à ce jour principalement situées dans la Communauté européenne.

Notre métier est de concevoir, développer, commercialiser et maintenir des applications de gestion à destination des agriculteurs et de leurs partenaires. Les partenaires des agriculteurs sont les associations de gestion et de comptabilité, ou les cabinets d'expertise comptable, les coopératives, les négoce et les agro-fournisseurs.

Une petite parenthèse : à la faveur du développement du *Cloud* (le stockage et le traitement sur réseau), nous avons investi récemment dans la construction d'un *Data Center* situé à Reims afin de garantir l'intégrité et la sécurité des données de nos clients. Je précise cela parce que nous n'en avons pas encore parlé ce matin, mais la localisation des données est une des problématiques potentielles du *Big Data*.

Je vais être un peu provocateur. Jusqu'à présent, j'ai entendu beaucoup de choses derrière les mots « *Big Data* ». Il y a de la robotique, de l'électronique, du Web. Je dirais que nous, cela fait quinze ans que nous faisons du *Big Data*. Plutôt d'ailleurs du *Small Data*, car cela dépend de ce que l'on appelle « *Big* ». Je dirais par exemple : nos outils depuis quinze ans permettent à une association de gestion et de comptabilité de compiler les données comptables de l'ensemble de ses adhérents, et de fournir des analyses de groupes pour que chaque acteur puisse se positionner et améliorer sa performance technico-économique.

Pour les coopératives, nous en avons parlé tout à l'heure, c'est pareil. Elles regroupent aujourd'hui dans des bases de données uniques, à l'échelle de chaque coopérative, les données techniques de leurs adhérents ou de leurs clients. Aujourd'hui, par exemple, nous pouvons estimer que 40 % de la SAU française est gérée par des outils produits par notre groupe.

Nous développons également depuis de nombreuses années des outils d'aide à la décision, en collaboration avec des agro-fournisseurs. En gros, l'idée est d'optimiser les choix techniques en s'appuyant sur le jeu de données associées, notamment des données météo ou des données pédoclimatiques.

Évidemment, j'ai aussi entendu ce matin – et nous faisons nous aussi ce constat – qu'avec le développement du *Big Data*, ce qui changeait par rapport aux dernières années, c'est une promesse nouvelle, à savoir fournir des services à haute valeur ajoutée grâce à deux choses : un, le développement du *Cloud*, avec des capacités de stockage et de traitement qui permettent une centralisation immédiate et du temps réel ; et deux, l'émergence de capteurs qui fournissent de la donnée de manière automatique, et sans effort de la part de l'utilisateur.

Ce qui est vraiment nouveau, c'est que cette valeur ajoutée peut s'obtenir de manière automatique, via les algorithmes, de façon transparente pour l'agriculteur. Je vais dans le sens de ce qui a été dit précédemment : tout cela peut remettre en cause la position de certains intermédiaires nationaux, certains conseillers qui jouaient un rôle de régulateur, et qui pourraient s'en voir dépossédés au profit des outilleurs – qui sont dans un grand nombre de cas nord-américains, rarement européens.

Nous pouvons donc imaginer que des multinationales du machinisme agricole arrivent à intégrer l'ensemble de la chaîne de valeur en enfermant les agriculteurs dans un système propriétaire – ce qui est déjà le cas dans un certain nombre de situations – et en excluant, par là même, les structures de conseil. Nous pouvons évoquer à cet égard le domaine des productions végétales – nous avons parlé de *John Deere* tout à l'heure –, nous savons que c'est déjà une réalité dans le domaine de l'élevage, notamment dans le cas des automates de traite.

La position de notre groupe aujourd'hui sur le *Big Data* est la suivante : nous nous interrogeons sur la meilleure manière d'y aller. Il y a deux voies possibles. Soit chaque acteur économique le fait tout seul et forcément avec une partie des données, ou alors en essayant de tisser de gré à gré des partenariats avec, par exemple, des entreprises qui opèrent dans le domaine des objets connectés, en essayant de développer dans son coin des modèles et en raisonnant évidemment sur une taille de marché *a minima* européenne. Il n'est pas raisonnable de penser que le marché intérieur français suffise à absorber les investissements nécessaires pour développer des outils suffisamment innovants et performants.

Soit – c'est la seconde voie –, si l'on considère le domaine des productions végétales, il s'agit d'arriver à mettre autour de la table les trois grands acteurs qui travaillent sur ces systèmes d'information. Ce n'est évidemment pas la solution naturelle parce que ce sont trois acteurs concurrents, qui ont des natures très différentes : ISAGRI est une structure privé ; la seconde structure est plutôt intégrée dans le système coopératif français, et la troisième est consulaire, et la concurrence avec elle n'est pas forcément facile.

La première solution est probablement la plus simple et la plus rapide. La seconde porte en elle la promesse d'une offre plus ouverte, avec plus de potentialités, car elle s'appuie sur un jeu de données plus important.

M. Jean-Yves Le Déaut. Merci beaucoup. Nous avons eu les avis de tous, d'abord sur ce qu'étaient les *Big Data*, puis sur la manière dont on gérait ces déluges de données ou ces données astronomiques (c'est le terme que vous avez employé), pour les extraire, puis les rendre utiles à l'agriculture.

Laurent Gouzènes, notre témoin du conseil scientifique, une réaction ?

M. Laurent Gouzènes, membre du conseil scientifique de l'OPECST, expert scientifique du groupe Pacte novation. Très rapidement, je voudrais aborder trois points.

Le premier, c'est qu'il existe un danger de déqualification de l'agriculteur qui va devenir de plus en plus informaticien et de moins en moins agriculteur. C'est une évolution qui mérite d'être analysée.

Deuxième sujet : l'assurance. Nous avons parlé d'assurance de précision pour l'agriculture de précision, qui peut cibler un type de culture, voire une commune. Il va devenir possible d'aller très loin dans le détail du ciblage, en raison de l'abondance des données : nous verrons sans doute de nouveaux systèmes d'assurance se mettre en place, pour assurer les cerises, dans telle région, tel département. Cela va sans doute arriver. Au niveau public aussi, dans le domaine de la gestion des catastrophes naturelles : la couverture d'assurance va jouer de façon plus fine qu'au niveau du département. Il y aura des évolutions dans ce sens.

Le troisième point qui est important selon moi concerne le logiciel embarqué dans les capteurs ou les machines. Est-ce que ceux qui achètent les machines et les capteurs ont le droit de le modifier ? Souvent, ils ont des licences d'usage et ne sont pas propriétaires. Si le logiciel qui est embarqué ne vous plaît pas, s'il ne collecte pas les données que vous voulez, si vous n'avez pas envie qu'il rapproche entre elles certaines données, vous n'avez *a priori* pas le droit de le modifier, de le corriger, *etc.* Cela pose des problèmes à la fois d'usage et même de propriété. Vous achetez une machine, vous n'êtes pas propriétaire du logiciel qu'il y a dedans. C'est une situation assez bizarre : vous possédez les éléments mécaniques et matériels mais pas les contenus immatériels. On ne sait pas très bien ce que l'on a acheté, parce que la machine ne peut pas fonctionner sans son logiciel. C'est un point qui mérite examen aussi dans le cadre de cette audition.

DÉBAT

M. Jean-Yves Le Déaut. Merci, nous allons revenir sur ce point. Marcel Deneux voulait réagir brièvement. Après, que ceux qui souhaitent faire une remarque lèvent la main ; d'abord parmi les intervenants, ensuite parmi les autres personnes dans la salle.

M. Marcel Deneux, ancien sénateur, ancien président de la Caisse nationale du Crédit agricole. Je voulais porter témoignage de la qualité et des compétences informatiques de la jeune génération d'agriculteurs. Je suis impressionné : il y a une proportion très forte de TPE du monde agricole qui interagissent tous les jours par informatique, y compris sur les marchés.

De ce côté-là, voici une expérience qui m'a marqué, et la preuve, c'est que je vous la raconte : j'étais rapporteur de la résolution du Sénat sur la volatilité des prix agricoles pour le G20, il y a je ne sais plus combien d'années. J'avais auditionné à cette occasion l'un des grands traders français de matières agricoles. Il m'avait expliqué comment il disposait d'un programme informatique qui combinait les mercuriales (listes des prix de marché) de Chicago, les données météo, la vitesse des cargos pour influencer les cours. Tout cela a pris de l'ampleur avec le développement de l'informatique, et influe sur le comportement des agents économiques que sont les agriculteurs, mêmes lorsqu'ils sont isolés. Ils prennent des décisions individuelles sur une base qui n'existait pas autrefois.

M. Jean-Yves Le Déaut. Merci. Qui souhaite intervenir en réaction à ce qui s'est dit, ou pour apporter un complément ?

M. Hervé Pillaud. Par rapport à ce que le membre de votre conseil scientifique (j'ai oublié son nom) vient de dire, je ne pense pas que le risque soit important que les agriculteurs soient transformés en informaticiens. Ce n'est pas dans notre culture. J'y reviendrai tout à l'heure. En revanche, il faut veiller à ce que nous gardions la maîtrise, que nous ne soyons pas simplement utilisateurs que mais que nous puissions être également acteurs, ce qui n'est pas tout à fait la même chose.

Après, pour ce qui concerne les systèmes assurantiels, ce n'est qu'un aspect des évolutions à prévoir. En termes de financement de l'agriculture, de R&D, de gestion des marchés, du conseil, de l'information, il y aura les mêmes avancées que pour les systèmes de gestion des risques. Ce n'est pas uniquement sur ce domaine-là qu'il va y avoir des avancées.

Sur la capacité à rentrer dans le logiciel, il est évident qu'il n'y a pas d'autorisation. En revanche, la capacité à « hacker » le matériel est la même dans le monde agricole qu'ailleurs, que ce soit directement par les agriculteurs (il y en a peu qui le peuvent, mais il y en a), ou par d'autres personnes qui pourraient le faire pour eux. Je ne sais pas qui a évoqué tout à l'heure le fait que la génération d'agriculteurs qui arrive n'est pas forcément d'origine agricole. Nous avons dans

ces nouveaux venus-là des gens qui viennent des secteurs tertiaires, notamment des secteurs informatiques, qui ont cette capacité-là. Ils viennent pour bien d'autres raisons, surtout pas pour refaire de l'informatique, mais ils sont capables d'améliorer le matériel.

Nous nous rendons compte qu'il y a deux façons de gérer cela. Soit l'on met en place des procédures, soit les entreprises (et je connais un exemple) « dealent » avec la personne concernée pour améliorer leur propre matériel.

M. Jean-Yves Le Déaut. Je vais rebondir en vous posant la question : ce qui m'apparaît comme un danger, comme une rupture en tout cas, est le fait que les sociétés informatiques (parce qu'elles ont la maîtrise des données) en viennent à investir d'autres secteurs. Nous allons en parler pour le domaine spatial la semaine prochaine, dans la mesure où le lancement de constellations de satellites par les GAFA (Google, Amazon et d'autres sociétés) fait que les sociétés de télécommunication peuvent finir par être prises en étau entre ceux qui auront la maîtrise de ces constellations et les utilisateurs.

Ce qui a été décrit quant aux évolutions aux Etats-Unis peut arriver chez nous : une dépendance accrue vis à vis d'un petit nombre d'acteurs, qui finalement modifiera profondément l'économie. Je souhaiterais que vous réagissiez à cette analyse parce que c'est un peu la question que nous avons à anticiper si nous voulons y réfléchir au niveau politique, au préalable.

Monsieur Marcel Deneux, vous deviez prendre la parole lors de la table ronde suivante, mais vous avez raison d'intervenir tout de suite. Vous êtes à la tête d'une société informatique qui vient de signer une convention avec une coopérative du Languedoc-Roussillon – c'est notamment pour cette raison que j'ai souhaité que vous soyez là aujourd'hui.

M. Stéphane Marcel, directeur général de SMAG. Effectivement, à l'échelle mondiale, nous voyons émerger de grosses entreprises puissantes avec des surfaces financières importantes qui se sont constituées grâce à leur réussite dans les secteurs du numérique. Je crois qu'il ne faut pas avoir peur de cela. C'est un parcours semblable à celui d'autres acteurs économiques qui existent sous d'autres formes, des multinationales venant d'autres secteurs qui ont également diversifié leurs actifs dans de nouveaux domaines. Monsanto, et même le groupe LVMH, par exemple, possèdent bon nombre d'actifs dans différents domaines qui leur permettent d'avoir une diversité dans la gestion du risque financier et une capacité à développer et à créer de la valeur.

Il faut prendre le temps d'analyser les choses, mais je crois que c'est le produit de la révolution numérique que le monde entier connaît aujourd'hui. Auparavant, ce processus d'émergence de multinationales s'est produit dans le domaine de l'agroalimentaire par exemple ou d'autres secteurs ; aujourd'hui c'est dans le numérique et elles réinvestissent l'argent gagné dans d'autres secteurs

d'activité. Maintenant, il peut se produire la même chose en Europe si nous nous organisons bien. Nous en reparlerons tout à l'heure.

M. Jean-Yves Le Déaut. Monsieur Pillaud ?

M. Hervé Pillaud. Je pense que ce n'est pas uniquement dans le secteur informatique qu'il faut chercher les dangers potentiels. C'est partout. Je vais vous donner deux exemples.

Le premier est que les GAFAs ont d'abord bousculé le secteur de l'informatique lui-même. Quand Apple a sorti ses premiers outils, il a mis à mal IBM qui dominait l'informatique de l'époque.

Autre exemple, dans le secteur de la traite. La robotisation de la traite des vaches a été réalisée d'abord par Lely, constructeur de matériels de travail du sol et de fenaison qui n'avait rien à voir au départ avec la traite.

Les évolutions viendront des acteurs qui auront compris (je l'ai dit en introduction) le nouveau fonctionnement des marchés de demain. C'est vrai qu'il peut en résulter un risque potentiel.

Je pense qu'il faut intégrer ces éléments-là. Ces acteurs-là iront chercher les informaticiens pour faire le travail. Mais en aucun cas ce sont les informaticiens eux-mêmes qui vont *a priori* y aller. C'est un nouveau système de fonctionnement qu'il faut intégrer pour pouvoir s'adapter à ces choses-là. Si nous ne l'intégrons pas, nous nous mettrons en situation de défensive, alors qu'il faut être dans une situation plutôt proactive que défensive.

M. Jean-Yves Le Déaut. Merci. Monsieur Abiteboul ?

M. Serge Abiteboul, professeur à l'ENS Cachan, directeur de recherche à l'INRIA, membre de l'Académie des sciences. Je voudrais intervenir sur le cœur du débat, qui est, à mon avis, la propriété et le contrôle des données. Après, on peut envisager des tas de *business model* autour de cela, mais partons du fait que les données sont produites par l'agriculteur quand il conduit son tracteur, par exemple. La question est : à quel niveau intervient la propriété ou plutôt le contrôle puisque la propriété est un terme confus ? Qu'est-ce qu'on fait avec ces données ? Qui contrôle ces données ? À mon avis, nous ne devons avoir une attitude angélique. Si on laisse faire n'importe quoi, il se passera n'importe quoi. Dans le cadre du régime des données personnelles, on commence à voir apparaître des associations d'utilisateurs. La Communauté européenne ou le Gouvernement français commencent à se mobiliser pour essayer d'instaurer un peu de réglementation, mais aussi des préconisations pratiques. Quand vous allez sur *Facebook* et que vous signez un document comme quoi vous donnez toutes vos informations à *Facebook*, c'est vous qui l'avez accepté. Les agriculteurs doivent aussi – ainsi que leurs associations et leurs organismes – comprendre que laisser toutes leurs données à la disposition des constructeurs, c'est quelque chose qui ne devrait pas être acceptable. Ils peuvent être aidés par les réglementations,

par les pouvoirs publics, mais il leur revient d'abord de comprendre que la donnée et le contrôle de cette donnée ont une valeur considérable.

M. Jean-Yves Le Déaut. Merci beaucoup. Monsieur Peyraud ?

M. Jean-Louis Peyraud, chargé de mission, direction scientifique « Agriculture », INRA. L'agriculteur produit aujourd'hui des céréales, de la viande et du lait, mais il produit aussi de la donnée. Cette donnée a de la valeur. Il faut donc trouver des modèles économiques pour valoriser cette donnée. Actuellement, un groupe assez important de *farmers* américains est en train de se retourner contre *John Deere* pour obtenir des indemnités suite à l'utilisation de leurs données par *John Deere*. Les *farmers* ont estimé que ces données étaient utilisées indûment car elles ont été transmises sans qu'ils ne le sachent depuis leurs tracteurs. *John Deere* possède toutes les informations sur le rythme moteur du tracteur, sur les parcelles. *John Deere* est capable à partir de là de faire beaucoup de choses, mais les *farmers* se retournent aujourd'hui contre lui. Un nouveau modèle économique doit donc être mis en place.

Le deuxième point sur lequel je voudrais insister : un agriculteur a une exploitation à gérer. Pour l'instant, les données fournies ne permettent d'en gérer qu'une partie. Un polyculteur éleveur peut avoir du matériel *John Deere* pour ses cultures et du matériel *Lely* pour alimenter ou traire ses vaches. Comment fait-il pour gérer son exploitation si les deux systèmes *Lely* et *John Deere* ne communiquent pas ? Il y a un vrai problème d'inter-connectivité derrière. Je l'avais mentionné dans ma conclusion mais je voulais revenir sur ce point, car c'est très important.

M. Sébastien Lafage. C'est en train d'être géré aujourd'hui avec une association de niveau international qui s'appelle l'AEF.

M. Jean-Louis Peyraud. Oui, je sais.

M. Sébastien Lafage. Elle vise justement à uniformiser l'image de l'ISOBUS avec un protocole d'échange ISO XML qui permet aux agro-équipementiers de s'accorder sur le fait d'échanger leurs données. Ce sont ensuite des acteurs, comme celui que je représente aujourd'hui, qui viennent capter cette donnée et permettre à l'agriculteur d'avoir une approche globale de son exploitation agricole, de pouvoir gérer et créer de la valeur sur sa donnée. Mais attention à ne pas trop fantasmer sur la donnée ! La donnée brute n'a de valeur qu'à partir du moment où on sait la traiter. La donnée de télémétrie sur la consommation en temps réel du tracteur, sur son carburant, doit pouvoir être croisée avec d'autres informations structurées ou non. Quand je prends ma voiture, je n'ai pas l'impression d'être générateur de données. Je ne pense pas que l'agriculteur, lorsqu'il prend son tracteur, ait l'impression d'être en train de générer de la valeur pour autrui.

M. Jean-Yves Le Déaut. L'agriculteur non ; mais tout à l'heure vous parliez de deux modèles possibles pour collecter à un niveau plus important la totalité de ces données. On peut donc les exploiter. La question qui se pose est « *Est-ce que l'accès à ces données est libre et gratuit ? Est-ce qu'il est possible de les transmettre à une coopérative ou à une organisation nationale ou européenne ? Ou est-ce que l'accès appartient seulement à celui qui a installé le logiciel dans le matériel agricole ?* ». Quant à l'harmonisation des données, vous avez raison ; mais regardez ce qui se passe entre Apple et ses concurrents pour les chargeurs par exemple. Rien que sur cette question, vous voyez que l'harmonisation des données est quelque chose de compliqué. Quand vous avez un avantage compétitif à ne pas harmoniser, vous n'harmonisez pas.

M. Sébastien Lafage. Dans le monde agricole, on a quand même évolué. L'ISOBUS est une vraie avancée aujourd'hui. À une époque, un tracteur ne savait tirer qu'un engin de la même marque que la sienne. Aujourd'hui, on sait associer plusieurs engins, et je pense que l'on peut arriver à transférer les données.

M. Jean-Yves Le Déaut. Il faut y arriver mais ce n'est pas un long fleuve tranquille.

M. Sébastien Lafage. C'est sûr.

M. Jean-Yves Le Déaut. Monsieur Pillaud ?

M. Hervé Pillaud. Je pense que la problématique est extrêmement complexe. On essaye là encore de regarder cela avec les outils dont on dispose actuellement. Je ne pense pas que l'on soit en capacité actuellement d'avoir des règlements qui nous permettent de cadrer ces éléments. Je vais donner un exemple. La Révolution française a engendré le Code civil, qui sert encore de cadre pour un certain nombre de choses. La Révolution industrielle a engendré le Code du travail. Je pense que la Révolution numérique devra engendrer son propre code. Là où il y a un souci, comme je l'ai dit tout à l'heure dans mon intervention, c'est que le numérique remet en cause les notions de frontière et les notions de temps. Il est donc extrêmement complexe d'établir ce nouveau code. Je n'en suis pas capable évidemment. Je pense que beaucoup de gens s'interrogent mais que peu de gens sont capables de pouvoir cadrer ces éléments. C'est pourtant bien à partir de là que l'on pourra progresser. L'utilisation des données par l'agriculteur, c'est assez facile à gérer. Ce qui est beaucoup plus compliqué à gérer, c'est la capacité à réguler l'utilisation extérieure des données par leur propre détenteur. Je pense que l'interdire serait une erreur. Laisser le champ libre est aussi une erreur. Où est la vérité dedans ? Ce n'est pas forcément simple.

M. Jean-Yves Le Déaut. Claude Kirchner, puis Monsieur. Après on va peut-être terminer cette première partie pour laisser Anne-Yvonne Le Dain, députée de L'Hérault, prendre la suite parce que je vais devoir aller à une autre réunion. Anne-Yvonne Le Dain suit ces questions de près, car elle a été chercheur au CIRAD dans une vie professionnelle antérieure. Elle conduira la fin de

l'audition et la troisième table ronde. Claude Kirchner, puis Monsieur dans la salle. Vous vous présenterez, s'il vous plaît ?

M. Claude Kirchner, directeur de recherche, conseiller du président de l'INRIA. Juste un mot à partir de ce qui vient d'être dit, sur les aspects standards et normes. En prenant l'exemple de ce qui peut se passer ailleurs dans le numérique, on s'aperçoit qu'être novateur et capable de construire des normes ou des standards de faits, est essentiel dans la maîtrise des données, des protocoles, des échanges. Je crois qu'il faut que l'on soit porteur par rapport à cela, en particulier en France et en Europe. Aujourd'hui, nos amis américains sont très bons pour arriver à faire du *lobbying* dans cette direction, et pour mettre en place des standards de faits qui ensuite risquent de nous bloquer.

M. Pascal Poidevin. Bonjour, je suis Pascal Poidevin, consultant sur les questions du numérique. J'ai travaillé quinze ans chez Apple, sept ans sur les systèmes d'information de santé en France et cinq ans sur les systèmes d'information d'élevage. La question de la maîtrise des données, de la propriété ou de l'accès aux données, est effectivement un sujet assez compliqué. Il me semble qu'aujourd'hui, si j'achète un robot ou un tracteur, je vais forcément accepter en tant que client les termes et conditions juridiques préparés par *John Deere* ou le fabricant de robots. Ces clauses indiquent un certain nombre de choses, notamment en rapport avec les données. À titre individuel, je ne lis pas ce texte la plupart du temps. *De facto*, j'accepte donc ce qui est écrit mais je ne sais même pas ce qu'il contient. Je pense que cela rejoint une question de gouvernance soulignée notamment par Hervé Pillaud. Aujourd'hui, il y a une responsabilité collective à faire face à ces termes et conditions, comme l'ont fait les exploitants américains. Il faut faire en sorte que quand il achète un équipement de ce type qui génère de la donnée, le client puisse être en situation de vraiment accepter ou pas, en conscience, ce qu'il signe. Un client ne peut pas à l'échelle individuelle s'adosser à un juriste pour faire cela. Cette démarche de régulation doit se situer forcément à un niveau collectif. Il me semble qu'il y a un chantier urgent à ouvrir comme l'ont certainement fait les exploitants américains, pour faire face aux termes et conditions. Les sociétés franco-françaises, voire les organisations professionnelles agricoles, les coopératives, se présentent également vis-à-vis de leurs propres adhérents avec des termes et conditions auxquels il faut s'engager en préalable au traitement des dossiers.

M. Jean-Yves Le Déaut. On passe à la troisième table ronde et c'est Anne-Yvonne Le Dain qui va la présider. Il y aura de nouveau une discussion longue à la fin de cette table ronde. Anne-Yvonne ?

**TROISIÈME TABLE RONDE :
COMMENT FAVORISER LE DÉVELOPPEMENT D'UNE AGRICULTURE DE
PRÉCISION EUROPÉENNE ?**

Présidence : Mme Anne-Yvonne Le Dain, députée, vice-présidente de l'OPECST

Mme Anne-Yvonne Le Dain. Merci Monsieur le Président, chers collègues. Bonjour. Je suis députée de l'Hérault, à Montpellier, et par ailleurs effectivement agronome de formation. J'ai longtemps travaillé au CIRAD. Je suis aussi géologue et spécialiste en sciences du sol. Nous avons donc déjà tenu ce matin deux tables rondes essentielles. Je voudrais remercier mon collègue qui est parti, Jean-Yves Le Déaut, pour avoir organisé cette demi-journée sur la question de la robotique dans le milieu agricole. C'est quelque chose dont personne ne parle, alors que c'est une révolution qui est plus qu'en marche. Elle est là. La question des données, de qui les traite, et comment on les traite, est complètement essentielle. Notre troisième table ronde aborde le cœur de cette préoccupation : l'enjeu stratégique qui existe derrière le développement du *Big Data* dans l'agriculture elle-même.

D'emblée, cette question doit être abordée en termes de souveraineté nationale, à la fois transversale et très française.

Cette question est transversale parce qu'elle rejoint la préoccupation d'une forme de dépendance vis-à-vis des grands acteurs industriels américains du numérique. Les *GAFAs*, comme cela a déjà été évoqué, donnent à nos partenaires, à nos concurrents, à nos alliés, que sont nos amis d'outre-Atlantique, des moyens d'emprise sur notre économie et notre société. Personne n'en parlait ou bien peu il y a deux ans. Tout le monde en parle depuis deux ans. Comme quoi nos sociétés évoluent de manière saccadée, brutale, par sauts et non d'une manière constante, continue, calme et paisible, comme le voudrait le cheminement de la raison ! Elles évoluent selon l'itinéraire de la passion, des enjeux et de l'intérêt économique.

Cette question est aussi très française car ce souci de préserver cette indépendance n'est pas une priorité pour nos voisins, y compris européens. La semaine dernière, une table ronde organisée sur la question de l'agriculture de précision à l'Ambassade de Grande-Bretagne n'a pas une seule fois fait mention de cette préoccupation d'indépendance, pourtant de l'ordre du réflexe culturel en France. Nous avons donc une spécificité particulière, pas toujours facile à vivre ni pour les Français, ni pour les entreprises, ni pour le monde politique ou le monde administratif. Défendre systématiquement un modèle qui nous paraît une évidence alors que les autres n'en comprennent pas les prémices, ce n'est pas forcément simple. Nous considérons souvent que leurs prémices à eux ne sont par nature, pas les bonnes. On a donc un petit problème à progresser dans ce domaine. C'était aussi une des raisons pour laquelle nous avons tenu à organiser cette table ronde.

Cette interrogation sur la manière d'avoir un meilleur contrôle de cette situation – la question est aussi celle-là – et de s'en servir aussi pour fabriquer de la valeur et de l'économie – c'est aussi l'enjeu essentiel – peut être déclinée à deux niveaux. Elle peut tout d'abord être déclinée au niveau du recueil des données. Il s'agit de savoir si ce recueil s'effectue toujours dans la pleine conscience des acteurs concernés. Cela a été évoqué tout à l'heure dans le débat. Elle peut également être déclinée au niveau du risque d'une offre de services liés. Il faut éviter que l'accès aux informations issues du *Big Data* soit conditionné par des prestataires à l'achat d'équipements - cela a été évoqué – ou par des fournisseurs de données. L'objet de cette table ronde sera d'indiquer si ces risques sont des fantasmes ou s'ils correspondent à une forme de réalité. Il s'agira de tracer les voies d'une nouvelle fantasmagorie qui serait pertinente économiquement.

Cette table ronde doit aussi nous permettre d'évoquer les pistes qu'il faut suivre pour s'affranchir de ces risques en construisant une offre de services *Big Data* autonome à l'échelle nationale ou européenne. Je dirais presque « et européenne », et ce n'est pas rien de le faire ! L'Europe est un peu une *terra incognita* actuellement.

J'interrogerai d'abord sur ces questions Monsieur Hervé Pillaud, puisque Monsieur Masson, président de la Chambre d'agriculture du Loiret, a été retenu par un incident de dernière minute. Je remercie Monsieur Pillaud de le remplacer au débotté. Les Chambres d'agriculture sont bien placées pour voir émerger cette problématique d'une dépendance éventuelle des exploitants agricoles vis-à-vis des fournisseurs étrangers de services de *Big Data*. Qu'en est-il selon vous, bien que vous l'ayez déjà un tout petit peu évoqué ? Mais on aborde maintenant les choses sous un autre angle. Merci.

M. Hervé Pillaud, représentant de la Chambre d'agriculture de Vendée. Merci Madame la Présidente. Je crois que vous avez bien posé le problème. Je ne pense pas qu'il s'agisse de lutter contre quoi que ce soit. Je pense qu'il y a de vraies chances à saisir. Il y a un champ suffisamment large pour se mettre en position de leader dans le domaine au niveau du secteur agricole. L'enjeu est que les agriculteurs passent d'une situation d'utilisateurs à une situation d'acteurs. Je pense qu'on ne pourra le faire que d'une seule façon : en réunissant nos forces. Certes, il peut y avoir un cadrage par un règlement, mais ces règlements ne feront que limiter les risques ou le temps avant que ces risques n'arrivent. En aucun cas, ils ne nous mettront dans une situation durable de pouvoir peser sur les choses.

Comment pouvons-nous y parvenir ? Je sais qu'il y a un travail déjà enclenché autour de ce qui a été confié à Jean-Marie Bournigal et d'autres par les ministères et par le Conseil de l'agriculture française (CAF). Celui-ci regroupe l'ensemble des organisations professionnelles pour conduire une réflexion sur la mise en place d'une plate-forme ou de plate-formes agro-numériques interopérables. Vous me permettez de ne pas rentrer dans les détails techniques, je n'en ai pas la capacité. D'autres le feront mieux que moi, même si je ne pense

pas que cette question relève du domaine technique, mais bien du domaine de la volonté politique. Il s'agit de pouvoir y arriver. Déjà, on peut faire un état des lieux de ce que l'on a chez nous en France, voire en Europe, je suis d'accord avec vous. Je pense toutefois que nous devons essayer de le faire en France pour ensuite l'extrapoler au niveau européen. On peut faire un état des lieux de toutes les données publiques de la PAC, de la recherche fondamentale, de l'INRA, de l'IRSTEA, des Chambres de l'agriculture, des Instituts, des COP, des centres de gestion. Les *start-up* françaises qui arrivent et qui sont des belles entreprises, ont aussi un certain nombre de données que bien souvent elles ne savent pas où stocker.

Cette plate-forme, je l'appelle de mes vœux. Je ne suis pas le seul. Je pense qu'il faut que nous la mettions en place. Ce n'est pas facile. Il faut une vraie volonté politique pour mettre les moyens financiers afin de pouvoir la créer. Après, je pense que cette plate-forme devra être ouverte vers l'extérieur de façon à pouvoir y agréger un maximum de données, comme le font les GAFAs. Pour attirer la multitude des acteurs concernés à ce qui peut se faire, notre plate-forme doit grossir et devenir incontournable en termes de données. Je pense que ce deuxième élément d'ouverture est aussi important que le premier.

Troisième élément : nous devons initier un certain nombre de choses pour y arriver. Nous avons lancé des actions au niveau des Chambres de l'agriculture, et j'en ai été un peu à l'initiative, au moment du Salon de l'agriculture. Nous l'avons fait avant en Vendée. Nous avons mis en place l'*Agreen'Startup* qui est un concours de *start-up* comme les *start-up* week-end ou les Hackathons. Il y a également eu un Hackathon dans l'Eure-et-Loire, voici quinze jours. Ces actions nous permettent d'attirer le plus grand nombre. Le challenge est simple : tout le monde a une idée de l'agriculture donc « *Banco, venez le faire avec nous. Ne nous dites pas comment faire, mais faites-le et apportez vos compétences. Nous en avons besoin et nous avons besoin de vous pour créer cette immense plate-forme qui, demain, aura suffisamment de données pour que nous puissions peser* ». J'en ai discuté avec Madame Axèle Lemaire, Secrétaire d'État au numérique. Je pense qu'il faut que nous ayons une vraie volonté d'inscrire ces éléments dans la *French Tech*. L'*Agro French Tech* doit être une réalité. Nous devons offrir dans les pôles ou les métropoles *French Tech* une dimension regroupant cet aspect territorial rural et agricole afin de permettre aux *start-up* d'émerger et de venir participer à cette plate-forme. Nous avons une réelle opportunité de nous positionner en *leader*.

Le champ est encore libre, mais il ne va pas le rester longtemps parce qu'effectivement il intéresse de grandes firmes. Les géants de la *Silicon Valley* commencent à investir en direction des *Food* vus sous l'angle du numérique. Il faut que nous relevions ce challenge. Si vous me permettez un peu de publicité : j'ai essayé de compiler tout cela autour d'un essai qui s'appelle « *Agro Numericus : internet est dans le pré* » qui sortira à la mi-septembre. Merci.

Mme Anne-Yvonne Le Dain. Au moins, vous lancez le débat d'une manière claire et offensive, c'est bien ! Je voudrais maintenant que l'on entende Monsieur Frédéric Garcia. Le *Big Data* fait l'objet aujourd'hui de réflexions au sein de l'INRA et vous allez nous indiquer comment l'INRA compte se positionner par rapport à ce sujet. Vous nous parlerez de ses analyses d'orientation stratégique et de leurs déclinaisons pour nos laboratoires ainsi que des relations que nos laboratoires entretiennent avec le monde économique.

M. Frédéric Garcia, chef du département Mathématique et Informatique appliquée à l'INRA, délégué à la Transition numérique. Merci Madame la Présidente. Je voudrais déjà dire que cette Révolution numérique que l'on vit dans le monde agricole, a des implications au niveau de nos orientations stratégiques de recherche, à l'Institut. Cela se traduit aujourd'hui par le fait que ces questions autour du *Big Data*, de *l'Open Science* ou de l'agriculture numérique, deviennent des choses explicitement écrites et sur lesquelles il faut que l'on ait un affichage clair. Ce n'était pas encore le cas voici quelques années. Le document d'orientation, actuellement en cours de rénovation, porte sur la période 2010-2020, et ces éléments n'y étaient pas fortement inscrits à l'origine. Les choses vont vite. Nous nous en rendons compte.

Je vais vous indiquer nos principales pistes concernant l'évolution de nos recherches et de nos partenariats. En quoi le numérique impacte-t-il ces deux dimensions ? Notre première réflexion concerne la manière dont le numérique va impacter l'inflexion que l'on est en train de vivre au niveau de l'Institut depuis plusieurs années et qui s'intensifie. L'institut de recherche INRA va être encore plus orienté vers l'innovation, la transformation et l'impact. On avait un logo au niveau de l'institut qui était « *Sciences et impact* ». On s'oriente vraiment vers un rôle plus marqué de l'institut dans les domaines du transfert et de l'innovation. La question des techniques autour du numérique traverse tout cela. Elle va faciliter mais aussi contraindre. Il est essentiel pour nous de ne pas nous transformer en institut du numérique, c'est évident, mais le numérique est transversal à toutes nos problématiques. Il impacte toutes ces questions.

Au niveau de la recherche, nous avons un certain nombre de départements directement impliqués par cette révolution du *Big Data* et de l'agriculture numérique : ceux qui se consacrent aux recherches tournant autour de l'agriculture, de l'environnement et de l'élevage. Cela touche également sans doute les départements de l'économie, de l'économie du numérique dans le monde agricole. Plus concrètement, l'INRA avait développé beaucoup de recherche scientifique au cours de ces vingt à trente dernières années sur l'étude des processus. La question de la modélisation est devenue très centrale dans l'Institut. Aujourd'hui, on a quasiment opéré un virage inverse avec un retour vers les *Big Data*, la question des données et de leur rôle. Cela se voit en ce moment au niveau des départements de recherche que j'ai évoqués : on est passé des efforts de recherche sur la modélisation vers des efforts de recherche sur l'utilisation intelligente des données.

On a entendu tout à l'heure un exposé sur la simulation des données et le lien avec la modélisation. On voit aussi ce lien se renforcer au niveau du département de recherche que je dirige à l'INRA, le département de « Math Info ». C'était historiquement un département de statisticiens, puis de biométriciens ; la donnée était la clé. On a observé une évolution forte consistant à doter le département de plus d'informaticiens et de modélisateurs. On a accueilli aussi des chercheurs en intelligence artificielle. Le département a changé de nom. Il était plus lié à la modélisation. En ce moment, on vit un retour vers le *Big Data*, vers la donnée. Il s'agit d'un mix entre informatique, mathématique, statistique, statistique algorithmique et ingénierie des connaissances, *etc.* Je ne dis pas que c'est un retour en arrière, mais ce virage clair mérite réflexion.

Au-delà de ces évolutions dans le domaine de la recherche, l'autre question importante est : « *En quoi le numérique impacte-t-il nos partenariats ?* ». On a depuis dix ans un partenariat de recherche assez fort avec l'INRIA. Il me semble qu'il ne sera que renforcé par l'évolution en cours. On devra aussi développer nos partenariats avec l'IRSTEA ou le CEA. La volonté de l'INRA de prendre plus en compte la question des données depuis leur production – les capteurs, *etc.* – jusqu'à leur utilisation, va sans doute nous amener à nous rapprocher, en France, de certains instituts nationaux avec lesquels on était peut-être un peu distant ces dernières années. Mais s'il nous faudra renforcer nos partenariats avec d'autres instituts de recherche, nous devons aussi développer les partenariats socio-professionnels avec l'industrie. Les données seront, comme cela a été dit ce matin, de plus en plus produites par le secteur privé, chez les constructeurs. Je ne suis pas favorable à la passivité face à cette tendance, et je pense que cette position est partagée au niveau de l'Institut. On pense que laisser de côté ces données n'est pas une hypothèse admissible. On va donc être dans l'obligation de travailler avec les groupes privés concernés. C'est un fait bien réel, les données sont là. Nous devons penser et développer des solutions nationales sur la collecte de données publiques ou privées. Nous devons également chercher à travailler avec ceux qui, aujourd'hui, possèdent la plupart des données du monde agricole. On considère donc comme un axe fort de développement les partenariats avec les grands groupes, mais également avec toutes les coopératives qui produisent et utilisent les données.

Je finirais juste par un message. Je dépasse un peu mon temps de parole, pardon. Je voudrais insister sur l'importance de conserver une approche *Open Data*, c'est à dire d'ouverture de l'accès aux données. Cette nécessité ressort notamment de nos réflexions sur les infrastructures de recherche. C'est un message clé actuel de l'Institut que cette nécessité de l'ouverture de l'accès aux données, mais pas à n'importe quelle condition : tout ne doit pas être gratuit et disponible instantanément. Nous faisons passer aux équipes de recherche le message qu'il faut éviter les solutions de repli au niveau des instituts. C'est en allant vers des solutions d'ouverture, de partenariat, que l'on s'en sortira.

Mme Anne-Yvonne Le Dain. Merci. Votre exposé montre à quel point c'est compliqué quand même, y compris conceptuellement. Dans une logique d'allers-retours où il y a beaucoup d'allers et peu de retours, ou beaucoup de retours et très peu d'allers, on n'a pas beaucoup de *feedback*. On n'est pas dans un système bijectif. Il faut donc inventer autre chose. C'est ce que disait quelqu'un tout à l'heure : il va falloir que le monde numérique invente ses codes. C'est une approche qui m'apparaît pertinente – même si ce n'est pas mon rôle à ce stade de prendre position – et que je trouve, en tous cas, intéressante à travailler. Cela prouve que la vie ne s'arrête jamais et que l'humanité doit toujours progresser. Merci Monsieur Garcia, d'avoir souligné que la transition à opérer n'est pas simple du tout. La grande maison INRA, se trouve, face à l'importance prise par le numérique, dans le même cas que beaucoup d'autres maisons. Est-ce que cette emprise croissante rend service aux autres disciplines ? Est-ce qu'elle doit être l'objet d'une discipline à part entière ? Est-ce que c'est l'invention de quelque chose de nouveau ? Cette incertitude sur sa signification pèse au quotidien sur les choix budgétaires, les choix d'allocation de moyens et les choix de solutions de recherche, y compris dans la formulation des demandes de crédit sur appels d'offres.

M. Frédéric Garcia. C'est clair. Juste une précision. On peut faire l'analogie avec les grands Groupes que l'on évoquait tout à l'heure. On disait que finalement, les industries qui ne prenaient pas le pli du *Big Data* et ne se mettaient pas en capacité d'offrir des services dans ce domaine, allaient perdre leur capacité à rester des acteurs majeurs du marché. Ils laisseront leur place à des nouveaux entrants qui n'y connaissent rien quelques années plus tôt, mais qui auront su développer une capacité dans ce domaine. Pour l'INRA, c'est la même chose : nous ne pouvons pas passer à côté du *Big Data*. Nous n'allons pas devenir l'INRIA mais nous ne pouvons pas ignorer cette dimension devenue centrale.

Mme Anne-Yvonne Le Dain. Il est central et traversant, ce qui n'est pas tout à fait la même chose que transversal. Je suis un peu coquine, je le sais ! Je voudrais passer la parole à Monsieur Stéphane Grumbach, de l'INRIA qui est déjà intervenu tout à l'heure. Je crois que vous souhaitez souligner que Le *Big Data* constitue une telle révolution qu'une offre autonome de services dans ce domaine ne pourra véritablement se constituer qu'à partir de l'émergence de sociétés nouvelles, en contournant les acteurs publics ou privés existant. On est au cœur du sujet. Il est donc vital d'encourager les plate-formes. Monsieur Grumbach, exposez-nous cela. Tentez de respecter les cinq minutes, merci.

M. Stéphane Grumbach, directeur de l'IXXI, Institut Rhône-Alpin des Systèmes complexes, chef de l'équipe INRIA DICE, INRIA Rhône-Alpes. Merci Madame la Présidente. S'intéresser aux données revient à s'intéresser aux moyens et non à la finalité. Les données sont un outil. Les techniques autour de la donnée constituent un outil pour mettre en œuvre des services qui ont des finalités. Je crois que l'on a beaucoup tendance sur le continent européen à s'intéresser trop aux données, même si on ne s'y intéresse que depuis récemment, et même à se focaliser sur les données - et cette matinée en donne l'exemple – alors que nous

devrions nous intéresser avant tout aux objectifs, et à la manière dont on veut impacter le monde. Nous devrions voir les données comme un simple moyen. Je dois dire très honnêtement que même dans les instituts de recherche, nous sommes beaucoup plus tournés vers les technologies de la donnée que vers l'impact sur la société.

Les grandes plateformes du numérique, elles, maîtrisent complètement les technologies de la donnée, mais elles s'intéressent également à la finalité. Celle-ci est en général assez novatrice pour les grandes plateformes parce que les services qu'elles font émerger sont non seulement d'une grande diversité, mais également très éloignés des services premiers qu'elles offrent. Je vais revenir à ce que je disais tout à l'heure. En général, les services premiers sont offerts sur des données accessibles. Dans le cas de l'agriculture, les données sont offertes par les grandes agences de météorologie. Par le suivi des traces d'activité des exploitants, on arrive à des services dérivés qui sont basés à la fois sur les données au sol, les données personnelles des exploitants, et sur les données des grandes agences. Parmi ces services dérivés, on va trouver l'assurance de l'agriculteur. C'est une évolution déjà observée : le moteur de recherche a un modèle économique basé sur les annonceurs ; d'une manière générale, le modèle économique de l'ensemble de ces plateformes n'est pas associé aux premiers services qu'elles rendent, mais aux services dérivés, qui s'appuient sur les traces des utilisateurs.

La libération de la donnée, le fait que les données deviennent accessibles, soit parce qu'elles sont créées (cela a été le cas des pages web ou des relevés faits chez les agriculteurs, des données qui n'existaient pas auparavant), soit parce qu'elles sont libérées (comme les données de *l'Open Data* du Gouvernement), cette libération change complètement l'équilibre du pouvoir. Avoir des données, c'est avoir du pouvoir. Ouvrir l'accès aux données, c'est re-répartir le pouvoir. Ce changement d'équilibre du pouvoir s'opère déjà dans nos sociétés. On le maîtrise très mal. Il conduit à déplacer le pouvoir vers le bas, vers l'utilisateur. J'imagine que ce type de rééquilibrage s'effectuera dans l'agriculture. On le voit globalement s'établir à l'avantage des utilisateurs *lambda* dans tous les secteurs dominés par les grandes plateformes. Il déplace aussi le pouvoir vers le haut, vers la plate-forme qui concentre beaucoup de pouvoir. Ce déplacement du pouvoir vers le bas et vers le haut laisse en grande difficulté les acteurs intermédiaires dont le modèle économique peut devenir complètement caduc. D'autres équilibres, comme celui sur lequel s'organise le rapport employeur/employé, s'effondrent.

Il y a une chose qu'il faut comprendre et qui est absolument caractéristique du numérique : la concentration. Plus on a de données et d'utilisateurs, plus on a de pouvoir, et la logique est celle d'une croissance rapide. Les premières capitalisations boursières aujourd'hui, qui correspondent aux premiers acteurs de l'intermédiation, le prouvent. Dans ce domaine, les grandes plateformes numériques se retrouvent en compétition avec les pétroliers. *Apple* aujourd'hui pèse le double d'*Exxon*. On voit aussi *Google* et *Microsoft* dans les dix premières capitalisations mondiales. On constate aussi des croissances très rapides de certaines sociétés comme *Uber* qui pèse aujourd'hui le quart d'*IBM*.

Pour ces nouvelles grandes entreprises, les rapports entre le nombre de leurs employés et le nombre de leurs clients ou utilisateurs diffèrent complètement de ce qu'on connaissait dans les secteurs traditionnels. Cette concentration n'est pas une chose à laquelle on échappera à mon sens.

Pour conclure, la difficulté pour l'Europe face à cette évolution est de ne disposer d'aucune plate-forme, de ne pas avoir su déployer ses compétences pour développer ses propres plateformes. Je pense que le domaine de l'agriculture, particulièrement pour la France, va être impacté – j'en ai parlé avant – par le développement de *The Climate Corporation*. C'est un domaine dans lequel nous disposons pourtant d'une très grande compétence mais également d'une spécificité culturelle qui justifieraient un effort majeur avec l'ambition très forte d'être aussi puissant que *The Climate Corporation*, ainsi qu'une ambition internationale. Je vous remercie.

Mme Anne-Yvonne Le Dain. Merci. Vous soulevez de fait deux interrogations. D'abord, pourquoi n'avons-nous pas démarré il y a vingt ans pour être un peu forts il y a dix ans et puissants maintenant ? Ensuite, quelles sont les réticences, quelles sont les résistances ? D'où sont-elles venues, du système lui-même ou des usagers, ou des forces en présence ? Ce n'est pas qu'une question financière, vous le savez. Les réticences psychorigides et sociétales ne sont pas anodines. On est là pour en parler. Comment faire pour changer ? Ce n'est pas simple mais on est également là pour en discuter.

Je voudrais maintenant passer la parole à Madame Gaëlle Kotbi, qui est Docteur en économie, enseignante-chercheuse. Je le mets au féminin délibérément car ce n'est pas écrit comme cela dans le programme ; c'est écrit autrement, j'en rajoute. Vous êtes enseignante-chercheuse dans le département « Stratégie et entrepreneuriat » de l'Institut polytechnique LaSalle Beauvais. Cet institut est donc situé à Beauvais ?

Mme Gaëlle Kotbi, enseignante-chercheuse dans le département « Stratégie et entrepreneuriat » de l'Institut polytechnique LaSalle Beauvais. Absolument.

Mme Anne-Yvonne Le Dain. Je vous remercie. Personne n'est mieux placé que vous pour nous indiquer comment vous préparez vos étudiants à la création d'entreprises innovantes dans l'agriculture actuelle, en particulier pour l'offre de services à partir du *Big Data*. Je vous en prie. Allez-y.

Mme Gaëlle Kotbi. Merci Madame la Présidente. Face à la nécessité de manager davantage les risques climatiques, énergétiques, économiques, compte tenu de la volatilité des cours, mais aussi des besoins croissants en investissements dans un contexte français de renouvellement démographique important- ce sont chaque année douze à quinze mille nouveaux agriculteurs qui s'installent et qui vont continuer à s'installer d'ici 2025-2030 -, on assiste à deux phénomènes importants pour notre débat d'aujourd'hui. D'une part, la montée en compétence

fait face à une complexité accrue en agriculture et à la nécessité d'une double performance, à la fois environnementale mais aussi économique. On le voit à travers l'élévation des diplômes. Maintenant, un tiers des agriculteurs a au moins un bac+2. D'autre part, un nouveau profil d'agriculteurs entrepreneurs innovants, émerge. Il est en rupture avec un modèle agricole d'après-Guerre à deux UTH [Unités de travail humain], comme le souligne un certain nombre de rapports ces dernières années.

Compte tenu des attentes collectives et sociétales, des défis colossaux pour nourrir la planète avec des produits durables, l'agriculture a besoin que des ingénieurs agriculteurs entrepreneurs soient formés. Je vais faire référence à LaSalle Beauvais pour prendre un exemple, mais de nombreuses autres écoles et formations vont dans ce sens. L'idée est de former des gens qui vont œuvrer pour la valorisation des filières, qui, par leur implication sur le territoire mais également par leur ouverture d'esprit sur l'extérieur – un pied dans la ferme, un pied à l'extérieur – vont permettre à l'agriculture de continuer sa mutation, en particulier sur la voie du numérique et du *Big Data*. Très concrètement, l'école à laquelle je contribue, a mis en place un certain nombre de programmes dont un qui s'appelle « Innover, entreprendre et servir ». Il vise à sensibiliser l'ensemble des étudiants, quel que soit leur parcours, dans ces domaines. On est également en réseau avec l'ensemble des établissements supérieurs picards et bientôt du Nord-Pas-de-Calais pour former, sensibiliser et accompagner les étudiants à l'entrepreneuriat. Je vais prendre une action concrète pour vous montrer la façon dont cela peut se décliner. Un parcours de formation de niveau Master créé depuis septembre 2013 sur deux années est proposé en formation initiale, mais également ouvert à la voie par apprentissage ou en formation continue, en particulier pour les porteurs de projets d'entreprise dans le monde agricole. Il permet à ces étudiants de devenir les ingénieurs agriculteurs entrepreneurs de demain.

La première année du programme est essentiellement centrée sur une approche disciplinaire. On parcourt l'économie, le management, la finance, le droit, mais également les sciences et techniques de l'information et de la communication, les statistiques, les mathématiques, les techniques quantitatives, et bien entendu l'agronomie et les sciences agro-industrielles. Pour faire le lien avec notre sujet, je peux mentionner, par exemple, comme faisant l'objets de cours et de travaux des étudiants, les systèmes d'information opérationnels de décision, les nouvelles technologies de l'information ou les enjeux de l'agriculture mondiale et les politiques européennes. Durant cette année, un quart du temps passé par les étudiants et un quart des crédits d'enseignement qui nous sont confiés, sont destinés à la participation à un concours de niveau international de création d'une *start-up*. Je pourrai, si vous le souhaitez, revenir pendant les débats sur les exemples concrets menés par les groupes d'étudiants cette année. Durant l'été, entre la quatrième et la cinquième année, les étudiants sont par ailleurs invités à s'ouvrir à d'autres horizons en menant des missions de solidarité à l'international. Cet été, j'ai par exemple des étudiants en mission au Bénin et au Pérou.

La deuxième année, elle, vise davantage la transversalité des compétences en bousculant les idées reçues. Comment bouscule-t-on les idées reçues, parce que c'est vite un vœu pieux ? On participe à des week-ends *start-up* ou des hackathons, à un maximum de visites et de rencontres avec les acteurs des réseaux, les collectifs d'agriculteurs, les entrepreneurs du monde agricole et les coopératives. On réalise des travaux de recherche fondamentale et appliquée, des séminaires avec les équipes internes mais également externes à l'Institut. Récemment des travaux ont été menés dans le cadre de deux chaires créées à l'école, l'une concerne l'agro-machinisme et l'autre le management des risques en agriculture. Vous voyez tout de suite les liens qu'il peut y avoir avec le *Big Data*.

Durant cette deuxième année, les étudiants passent un quart de leur temps sur un projet : la création d'un modèle économique innovant et la construction d'un plan d'affaires dédié à ce projet réel de développement de l'agriculture. Ils ne travaillent pas sur des projets fictifs. Ce projet peut être tout à fait personnel. J'ai par exemple cette année deux étudiants qui ont acquis le statut d'étudiants-entrepreneurs du Ministère de l'Enseignement supérieur pour mener à bien et développer leur propre projet. Ce projet peut être commandité par un porteur de projets ou une entreprise existante. La seconde période de cette deuxième année est consacrée à un stage et à un mémoire de fin d'étude. Nos étudiants sont présents chez la plupart des intervenants autour de cette table.

Pour conclure : qu'apporte une formation d'ingénieur en agriculture pour le *Big Data* ? Pourquoi est-elle importante ? Les compétences numériques, informatiques et mathématiques sont essentielles dans le *Big Data*. Elles rejoignent des compétences de l'ingénieur même s'il faut les spécifier. Ce sont aussi des connaissances économiques et agronomiques qui, pour nous, permettront que demain, de nouvelles pratiques agricoles se développent, soient respectueuses de l'environnement, et soient économiquement durables. Si on veut limiter la dépendance d'une agriculture française vis-à-vis de l'extérieur et de fournisseurs qui se seraient rendus incontournables à l'insu des agriculteurs, si l'on veut éviter une agriculture à deux vitesses – de ce point de vue, les agriculteurs entrepreneurs disposent des cartes dans leur jeu, mais tous les agriculteurs ne mesurent pas encore ces défis –, il faudra l'implication de chacun. Les structures aussi bien publiques que privées, au centre desquelles les agriculteurs se placent, devront s'investir dans la co-construction d'une offre de services en *Big Data*. Je vous remercie.

Mme Anne-Yvonne Le Dain. Merci. C'est ambitieux. Il faut qu'ils créent leur marché ou qu'ils le trouvent. L'adéquation entre créer son marché et le trouver est difficile dans le monde d'aujourd'hui. En tout cas, on y reviendra dans la discussion.

Je voudrais maintenant passer la parole à Stéphane Marcel qui est montpelliérain ou presque. Il est directeur général de SMAG, une entreprise dédiée à la solution informatique pour l'agriculture, notamment pour l'optimisation de la productivité, qui revendique quatre cents clients et

35 000 utilisateurs. La préparation de cette audition a montré que vous étiez au sein de la communauté professionnelle, une figure reconnue de l'innovation en matière d'agriculture de précision. Vous n'hésitez pas à vous engager dans des actions de portée générale puisque vous venez de signer début juin une convention de coopération avec la Fédération des coopératives de la région Languedoc-Roussillon. Votre analyse d'entrepreneur sur les perspectives de développement d'une offre nationale et européenne de services en matière d'agriculture de précision nous intéresse donc tout particulièrement puisqu'on est dans l'opérationnel.

M. Stéphane Marcel, directeur général de SMAG. Merci Madame la Présidente. En tant que Languedocien, je suis ravi de prendre la parole autour de cette table. En quelques mots, je suis le CEO de SMAG et un des fondateurs de ce Groupe qui est un *pur player* du web agricole. Il y a dix ans ou un peu plus, on a misé sur le fait qu'on vivrait aujourd'hui une vraie Révolution numérique dans le secteur agricole, et que le secteur traditionnel du logiciel agricole porté sur la comptabilité, changerait. On est donc au cœur de cette révolution aujourd'hui. Nous proposons pour le secteur agricole, les agriculteurs et les coopératives, des solutions intelligentes pour produire plus et mieux. Nous sommes une filiale d'une union de coopératives qui s'appelle InVivo. Il s'agit d'un important Groupe coopératif à l'échelle mondiale, pesant six milliards d'euros de chiffre d'affaires. Je précise cela parce qu'on constate aujourd'hui qu'autour de la problématique du traitement des Data et de la création de valeur autour de la Data, la profession agricole s'organise à tous les niveaux.

Hervé, tu as parlé du CAF tout à l'heure. Il y a bien une prise de conscience. Des entreprises coopératives font des acquisitions, investissent dans des entreprises comme SMAG, aux quatre coins de la planète. Je vais en citer deux ou trois. Land O'Lakes est une grosse coopérative aux États-Unis. Elle pèse quinze milliards de dollars de chiffre d'affaires. Elle vient de racheter une entreprise toulousaine, Geosys. Baywa qui est le numéro un allemand, une grosse coopérative qui vient de racheter une entreprise allemande, EuroSoft. Agravis, qui est le numéro deux allemand, vient de faire une joint-venture avec Geo-Koncept, une entreprise également allemande, dans le monde de la Data et de la télédétection. Cela illustre tout simplement le fait que, au même titre que la Révolution numérique que l'on vit aujourd'hui dans beaucoup de secteurs d'activité, l'agriculture vit cette révolution. On est en plein changements de paradigmes dans l'agriculture. On passe d'une agriculture qui se disait raisonnée à une vraie agriculture mesurée et pilotée grâce à l'explosion d'un certain nombre de technologies, notamment les technologies d'observation et de capture des données automatiques.

On pense, s'agissant de la problématique de la souveraineté française autour de ce dossier, que si la profession agricole, et notamment l'agriculteur et ses partenaires, s'organise aujourd'hui pour sanctuariser la donnée, nous pourrions garder notre souveraineté. Le fameux scénario que tout le monde redoute aujourd'hui, se résumant au schéma « agriculteur-*John Deere-Monsanto* »,

pourrait ne pas se dérouler. On montre bien souvent du doigt des *success stories* comme *Apple* ou autres, mais *Apple* tire son succès du fait que c'est *Apple* jusqu'au bout des ongles, dans *iTunes* ou votre *iPhone* ou votre tablette. Vous êtes dans ce monde. *John Deere* est un peu le « *Apple* » de l'agriculture avec l'ensemble de ses matériels connectés, avec son « *myjohndeere.com* » qui est le *iTunes* de *John Deere*. Je pense donc qu'en France et en Europe, nous pouvons nous organiser. Il n'est pas trop tard. Nous pouvons constater que la *smart agriculture* et le *Big Data* agricole sont générateurs de valeurs sociétales, économiques, environnementales, agronomiques.

L'agriculteur doit également pouvoir compter sur sa coopérative ou son négoce pour comprendre les enjeux, comprendre que son activité génère une masse de données qui peut partir non pas dans une entreprise mais dans sa propre coopérative. C'est bien l'objet d'une coopérative d'organiser le partage et la mise en commun de moyens. Cette donnée collectée va servir au bien collectif et aux biens individuels en créant de la valeur additive. Nous pensons qu'aujourd'hui la numérisation de l'exploitation agricole doit être favorisée. Cela peut passer par des propositions incitatives et financières par exemple, afin d'accompagner l'équipement et la digitalisation de l'exploitation agricole. Nous pensons que le projet porté par l'IRSTEA concernant la mise au point d'un sas en *Open Data* pour permettre de l'*Open Innovation*, est une bonne initiative. Nous contribuons aujourd'hui à la mise au point de ce dossier.

Nous pensons aussi que les pouvoirs publics doivent faire en sorte de rapprocher le monde du privé et le monde du public, qui est lui aussi générateur d'un grand nombre de données. Je pense à FranceAgriMer avec, par exemple, toute la déclaration PAC. Il faut rapprocher les acteurs du privé et du numérique avec ces acteurs-là. Il faut probablement injecter plus de capitaux. Je sais qu'il y a beaucoup d'aides actuelles autour de la création d'entreprises innovantes dans les *High-Tech*, mais je pense qu'on peut faire encore mieux. Je crois qu'Hervé a parlé des métropoles labellisées *French Tech*. Montpellier en est une. On peut peut-être capitaliser autour de ces labels qui vont nous permettre demain de rayonner au niveau européen et international.

Je conclurai en disant qu'il faut arrêter de fantasmer sur le *Big Data*. Le *Big Data* n'est ni plus ni moins qu'une technologie mathématicienne qui nous permet de faire de l'exploitation en masse de la donnée, éventuellement du *machine learning* pour simplement passer d'un monde finalement post-traumatique à un monde plus prédictif. C'est comme ça. Il s'agit d'évolution. Ne fantasmons pas là-dessus. Organisons-nous pour éviter qu'un *Climate Corporation*, par exemple, ne vienne capter la valeur à notre place. Merci.

Mme Anne-Yvonne Le Dain. Il est trop tard ?

M. Stéphane Marcel. Non, il n'est sûrement pas trop tard.

Mme Anne-Yvonne Le Dain. Voilà !

M. Stéphane Marcel. Il n'est vraiment pas trop tard. Je prends un exemple dans l'hôtellerie : *Airbnb*, en sept ans, se retrouve mieux capitalisé aujourd'hui en bourse que le Groupe Accor qui existe depuis cinquante ans. *Uber*, en six ans, a une cotation boursière égale au quart d'un géant mondial de l'informatique, IBM, qui s'est créé en plusieurs dizaines d'années. Donc, il n'est pas trop tard mais dans un sens comme dans l'autre : on peut aussi se faire très vite doubler, « ubériser » pour reprendre des termes à la mode.

Mme Anne-Yvonne Le Dain. C'est aussi une question de protection. Le cas d'*Uber* n'est pas hyper simple à traiter. La question du modèle économique global, et pas seulement de l'entreprise, est quand même sur la table. « Global » inclut la question des services publics et de l'activité des autres. Tout cela doit être géré et c'est compliqué. En Europe, nous avons un consortium de pays qui ont tous une histoire. On a un peu de mal à sortir notre drapeau aussi de temps en temps, ce que les Américains n'hésitent pas à faire sans état d'âme. On est dans un changement de culture, vous l'avez très bien dit. On passe d'un monde post-traumatique à un mode prédictif. Je l'ai noté, j'ai trouvé la formule très jolie !

M. Stéphane Marcel. Il faut que l'on soit fier de nos entreprises. De notre côté, on gère dix millions d'hectares. Sur la SAU française, cela peut paraître beaucoup mais à l'échelle internationale, on est des nains. Je pense que les Américains savent capitaliser sur leurs *success stories*. On doit rentrer dans ce modèle et porter des entreprises, encore plus si elles sont capitalisées par la profession agricole. Du coup, d'une pierre deux coups, nous parviendrons à rivaliser avec les meilleurs à l'échelle de la planète.

Mme Anne-Yvonne Le Dain. La grande question étant : qui perdons-nous en chemin et que gagnons-nous en chemin et pour qui ? C'est une grande question qui est aussi d'ordre politique. C'est aussi pour cela qu'elle est sur cette table. Stéphane, je vous remercie. Je voudrais maintenant passer la parole à Monsieur Pierrick Givone.

M. Pierrick Givone, directeur général délégué à la recherche à l'innovation de l'IRSTEA. Merci Madame la Présidente. Je voudrais en introduction rappeler une chose qui a été peu dite à ce stade : l'agriculture et l'agriculteur sont aussi des acteurs du développement du territoire. Lorsque l'on parle de données agricoles, on ne parle pas simplement de données de production agricole végétale ou animale, on parle aussi du développement territorial. Les enjeux du *Big Data* en agriculture sont aussi ceux-là et non simplement les enjeux de la production agricole au sens stricte.

L'agriculture, cela a été dit abondamment, génère, mobilise toutes sortes de données, y compris symboliques. Les savoir-faire de l'agriculteur doivent être intégrés aux données en question. Cela rend aussi l'agriculteur acteur de ces *Big Data*, en particulier par ces données symboliques de savoir-faire spatialisées. Le traitement de ces données, sa maîtrise et son contrôle qui sont au cœur du débat, constitue fondamentalement un élément de compétitivité. Vous l'avez rappelé

Madame la Présidente, avec un peu de malice, il constitue également un élément de souveraineté. Je ne vais pas en rajouter sur la souveraineté mais je voudrais simplement dire qu'il ne faut pas être naïf en la matière.

Je rappellerai un fait et un élément d'histoire. Le fait est que 50 % des tracteurs français sont équipés de GPS. Je vous rappelle que l'armée américaine, pour aller vite, a le pouvoir de couper le signal et/ou de le dégrader fortement. *Galiléo*, le système européen, va se mettre en place un jour ou l'autre. On finira bien par le mettre en place mais pour l'instant, il n'est pas tout à fait au même niveau d'opérationnalité que le système GPS. Pour l'histoire, je vous rappelle que les satellites Landsat ont été lancés par les Américains, à l'époque de la Guerre froide, dans le but de prédire les évolutions agricoles de l'URSS. Ceci permettait de prédire quelle attitude allait adopter l'URSS dans certaines négociations géopolitiques : si la récolte était bonne, ils auraient moins besoin d'importer du blé. Les éléments de compétitivité sont donc majeurs mais les éléments de souveraineté ne sont jamais très loin. En tout cas, il ne faut pas être naïf en la matière et il faut veiller à ne pas abandonner ces éléments de compétitivité à quelques majors comme on les a citées, *Monsanto* ou *John Deere* ou autres.

IRSTEA est porteur du projet - avec d'autres et pour d'autres - de construire un véritable écosystème ouvert de l'innovation mobilisant les données les plus larges au profit de l'agriculteur, de l'agriculture et du développement territorial. Nous pensons qu'il faut veiller à constituer une plate-forme la plus ouverte possible. La dimension d'*Open Data* est fondamentale. Cette plate-forme doit être réalisée en veillant à ce que les agriculteurs et la profession les représentant - la CAF, cela a été également dit - soient au plus près de la gouvernance du système.

Unique ou multiple - c'est un point à traiter - elle doit permettre dans tous les cas - c'est mon deuxième point - d'offrir à un certain nombre d'entreprises et de *start-up* la possibilité de construire la véritable finalité du système : offrir les services aux agriculteurs et des outils d'aide à la décision. Certains de ces services sont d'ailleurs inconnus aujourd'hui. Lorsqu'il a commencé à envoyer des images sur le web, Google ignorait absolument quels seraient ses services actuels. Nous voulons parier qu'une plate-forme ouverte en capacité et maîtrisée par les agriculteurs sera capable d'initier un certain nombre de services. Les outils d'aide à la décision sont bien connus. Ils peuvent être améliorés mais il y aura d'autres services. Cela implique une grande vigilance de la part des agriculteurs, et je terminerai là-dessus. Ces agriculteurs génèrent parfois des données sans le savoir et n'en maîtrisent souvent pas le devenir. Cela nécessite la vigilance des représentants de la profession mais aussi sans doute, celle du législateur européen et pas uniquement français. Je vous remercie.

Mme Anne-Yvonne Le Dain. Merci, Monsieur Givone. Je voudrais maintenant passer la parole à Monsieur Pierre-Philippe Mathieu, qui est le responsable du programme *EO4Food* (*Earth Observation for Food*) à l'Agence spatiale européenne. Nous avons été mis en contact avec lui grâce à l'astronaute

Claudie Haigneré qui est membre de notre Conseil scientifique et qui vient de prendre de hautes responsabilités au sein de l'ESA. L'ESA s'engage d'emblée sur une démarche de fourniture de données en mode *Open Source*. Comment l'ESA envisage-t-elle le développement de ses services de *Big Data* autour de la fourniture de ces données satellitaires ? À travers le programme *EO4Food*, y a-t-il déjà des exemples d'agriculture de précision mis en œuvre en Europe à partir de ces données ?

M. Pierre-Philippe Mathieu, responsable du programme *EO4Food* à l'ESA. Bonjour, Madame la Présidente et merci de m'avoir invité. J'ai préparé quelques planches ⁽¹⁾. Je ne sais pas si on peut les projeter. J'ai la grande chance de pouvoir vous annoncer quelques bonnes nouvelles aujourd'hui. Le but de l'ESA est de créer, lancer et exploiter des satellites. Ce n'est pas purement pour la technologie mais pour créer de l'information qui peut vous aider à créer du *business* par la suite. Nous avons une vue européenne et même globale. L'idée est de construire un système européen de *monitoring* de la planète, que l'on appelle Copernicus, nous permettant de gérer les ressources naturelles et les risques. Vous voyez ici un exemple du type de risques que l'on est en train d'essayer de gérer avec nos missions : la pression sur l'environnement due à la population qui va vers neuf milliards de personnes, et évidemment les terres arables qui diminuent et qui subissent également d'autres types de pression, sociale, en matière de compétition avec les forêts ou économiques. On essaie de mettre en place un système d'information qui permette de prendre des décisions de manière objective. Pour cela, on lance une série de missions opérationnelles avec une *Open Data Policy* qui s'appelle les « Sentinelles », dans le cadre de Copernicus. Vous voyez ici Sentinel-2, une mission optique. J'en parle avec beaucoup de fierté parce qu'on a lancé la mission il y a une semaine. Vous voyez la première image après quatre jours. On est vraiment très content. Tout s'est bien passé. On est dans l'orbite nominale. Il s'agit d'une image optique à résolution de dix mètres. L'image varie entre dix et trente mètres sur treize bandes de fréquence, dont certaines sont optimisées pour l'agriculture, pour comprendre la santé de la végétation. Ce genre d'image sera délivré globalement de manière systématique tous les dix jours cette année, et tous les cinq jours d'ici quelques années, quand on aura lancé une deuxième mission. C'est une révolution. Il n'y a pas d'autres missions équivalentes qui fassent cela de manière opérationnelle en *Open Data*. Il y avait des missions commerciales il y a quelques années. Cela veut dire, selon moi, que l'on va vers la démocratie environnementale. On aura un outil qui permettra de regarder ce qui se passe sur la planète. Les gens ne pourront plus se cacher. Cet outil pourra également nous aider à mieux gérer les ressources. Un des problèmes de cette mission est qu'elle est optique : s'il y a des nuages au-dessus, ce qui se passe souvent en Europe, on ne verra rien. C'est pour cela qu'il faut un *revisit time* assez rapide. C'est aussi pour cela qu'on envoie d'autres types de missions comme Sentinel-1, lancée il y a un an. Il s'agit d'un radar : cette mission voit à travers les nuages. Elle n'a pas été faite pour l'agriculture. Elle a été pensée pour regarder les

(1) Voir Annexe 4.

mouvements du sol à quelques centimètres, en cas de tremblements de terre, en cas de pompage d'eau, etc : on arrive à voir la subsidence (affaissement de la surface terrestre) des citées ; on arrive également à voir la glace, ses mouvements, le changement climatique. Elle a évidemment un impact sur l'agriculture parce qu'on voit par exemple, ici, la production de riz en Asie. Vous voyez que le riz a une réponse sur le radar et on arrive à comprendre comment il grandit. Ce genre d'informations est très utile. Il a été fait dans le cadre d'un programme qui s'appelle GEOGLAM (*Global Agricultural Monitoring*), un système de coordination des satellites pour le monitoring de l'agriculture, lui-même poussé par un autre programme qui s'appelle AMIS (*Agricultural Market Information System*). Celui-ci essaie de créer de la transparence dans le marché. Les fluctuations et spéculations mettent un certain nombre de gens dans la pauvreté dès qu'il y a un événement climatique ou un spéculateur. Avoir une vision globale de l'état des récoltes aide à augmenter la transparence et à diminuer les fluctuations. C'était l'objet de *Landsat* : comprendre ce qui se passe dans d'autres pays.

Notre problème est de remplacer le *Big Data* – parce qu'on aura un pétaoctet de données –, par du *Small Data*. On voudrait arriver à envoyer un SMS à l'agriculteur en lui disant « *Vous pouvez planter maintenant* ». Pour traverser toute la chaîne de valeurs concernée, on a besoin de mettre en place un système d'entreprises. On subventionne l'industrie pour créer des produits à valeur ajoutée qui renseignent directement l'utilisateur, et transforment nos données en informations. C'est donc l'un de nos objectifs de créer ce genre de nouveaux produits. On entretient aussi un dialogue avec des *citizen scientists*, dont certains pourraient tout à fait figurer ici dans l'assistance. Ce sont des personnes qui, avec leur téléphone portable, font des photos pour nous, afin de valider nos produits. Ils classent aussi nos produits en les regardant simplement parce que l'œil humain est capable d'une grande performance de classement. Cela permet de les impliquer pleinement. On travaille en partenariat avec la Banque mondiale et d'autres acteurs du développement pour essayer d'étendre le champ d'utilisation des données. On organise également des hackathons et des week-ends d'entrepreneurs pour recueillir les meilleures idées concernant l'usage de nos données. On a donc créé une infrastructure avec des APIs qui permet à ceux qui ne comprennent rien à l'espace d'utiliser nos données et de créer des produits. Voici là une bouteille créée en une journée, dotée de capteurs d'humidité, dont les résultats sont comparés à ceux fournis par nos données. D'autres capteurs écoutent également les criquets pèlerins, et quand ils repèrent un criquet, ils envoient un bruit pour que le criquet s'en aille. C'est un système d'alerte.

Je conclus sur ce slide pour présenter notre MOOC, c'est-à-dire un *Massive Open On line Course*. Si vous voulez en apprendre plus sur ce que l'espace peut faire pour le climat et l'agriculture, vous pouvez vous enregistrer sur ce site ⁽¹⁾. L'enseignement passe par des vidéos et le cours débouche sur un diplôme. On a huit mille personnes enregistrées. Elles sont très contentes. Merci.

(1) À l'adresse suivante : <https://www.futurelearn.com/courses/climate-from-space>.

Mme Anne-Yvonne Le Dain. Vous êtes dans un timing parfait ! Métronomique. Je vais peut-être lancer la discussion entre vous. Je crois que l'on a eu une primeur non, il me semble ?

M. Pierre-Philippe Mathieu. Oui, c'est un domaine où l'Europe est en situation de leadership. On parlait de ce qui se passait aux États-Unis. De fait, les Américains nous envient ce système qui n'a pas d'équivalent opérationnel dans le domaine de la continuité des données sur les trente prochaines années. On a un budget opérationnel de la Commission au titre du programme Copernicus. Cela permet aux gens qui utilisent les données de faire le saut. Si la mission s'arrêtait après trois ans, comme nos recherches, jamais ils n'auraient investi. Ici, ils ont vraiment une garantie. Je précise tout de suite qu'avec une *Open Data Policy*, les Américains pourront utiliser nos données comme le feront les Chinois, *etc.* Il nous faut donc continuer à bouger plus vite !

DÉBAT

Mme Anne-Yvonne Le Dain. C'est clair, mais depuis le début il faut bouger plus vite ! Depuis le début, on s'autocontrôle, donc maintenant il faut se lâcher si j'ai bien compris, pour utiliser une terminologie un peu rapide. Je voudrais passer la parole à mon collègue Patrick Hetzel, député UMP d'Alsace, qui est arrivé. Patrick, si tu veux nous donner rapidement tes impressions ?

M. Patrick Hetzel, député. Ce sont Les Républicains maintenant Anne-Yvonne mais bon !

Mme Anne-Yvonne Le Dain. Tu m'excuseras mais je ne peux pas. Ce sont de vieux réflexes horribles !

M. Patrick Hetzel. Cela viendra ! Merci beaucoup. Je voudrais tout d'abord saluer cette initiative de l'OPECST. C'est vrai qu'à vous entendre ce matin, on comprend bien qu'il y a un sujet et qu'il y a vraiment quelque chose à poursuivre avec le *Big Data* et l'agriculture. Je voudrais m'adresser plus particulièrement aux chercheurs et responsables d'organisme présents ce matin même s'il ne s'agit pas d'avoir un propos exclusif. Je vous ai entendus les uns et les autres, les représentants de l'IRSTEA, de l'INRIA et de l'INRA – d'ailleurs je me demande si le CNRS n'a pas un certain nombre de choses à dire sur le sujet – mais ma question est assez simple. Vos organismes ont été créés en leur temps avec des objectifs bien précis et on le comprend. Vous avez été plusieurs à indiquer – et Monsieur Garcia, vous y faisiez référence tout à l'heure – que vous alliez sans doute travailler davantage les uns avec les autres, besoin faisant nécessité et nécessité devenant loi. Ce qui est très frappant aussi c'est de constater qu'il y a là un vrai sujet par excellence, qui est interdisciplinaire. À l'OPECST, notre démarche consiste, à un moment donné, de partir d'un diagnostic et de faire en sorte qu'il débouche sur un certain nombre de recommandations. On est bien

d'accord, nous avons une philosophie : il faut que les chercheurs puissent chercher. Il ne s'agit pas pour l'OPECST de se substituer à eux et d'aller dans une approche micro vis-à-vis des chercheurs. Il faut plutôt leur laisser un cadre qui leur permet d'effectuer un travail de recherche. Néanmoins, la question qui se pose à nous aujourd'hui sous un angle de politique publique est : est-ce que vous avez besoin d'une évolution du cadre ? Deuxièmement, faut-il pousser à davantage de coopérations ? Quand je dis cela, on voit bien qu'une politique incitative peut passer par des appels à projet ANR, ou éventuellement une plus grande prise en compte de ces aspects dans la programmation des appels à projets. Est-ce qu'au contraire, vous considérez qu'aujourd'hui la boîte à outils fonctionne ? Qu'à travers les appels blancs, vous avez de quoi faire ? Cela nous intéresse d'avoir votre retour là-dessus parce que cela peut déboucher aussi sur des recommandations précises et concrètes dans le droit fil de cette question d'une coordination des travaux de recherche entre les organismes. D'ailleurs, de manière incidente, puisque vous indiquez aussi qu'il y a une nécessité de travailler avec les données disponibles dans le secteur privé, est-ce qu'il y a des choses à repenser dans l'articulation public/privé ? Pardon d'avoir été un peu long, mais on voit bien que pour nous, il y a aussi un certain nombre d'enjeux qui se situent plus dans la partie « évolution de la politique publique en matière de recherche ».

Mme Anne-Yvonne Le Dain. Je vois que l'ancien Directeur général de la recherche que tu as été ne peut pas s'empêcher de le rester un peu ! Merci en tout cas. Nous avons une table ronde passionnante comme les précédentes. J'ai assisté à la fin de la précédente et c'était passionnant. Est-ce que vous avez des interrogations ou des réflexions les uns envers les autres ? Avez-vous des réactions par rapport à ce qu'ont dit les unes et les autres ? Pierrick Givone ?

M. Pierrick Givone. Merci Madame la Présidente. Tous les sujets complexes concernant l'agriculture, l'environnement et le territoire nécessitent qu'on y réponde via de grandes collaborations interdisciplinaires. De ce point de vue, le sujet ne fait pas exception à la règle. Est-ce que tous les mécanismes sont en place pour que ces collaborations soient les plus fécondes possibles ? D'abord, est-ce qu'il y a des mécanismes qui les initient, les développent ou les renforcent ? C'est selon ce que vous considérez : le verre à moitié plein ou à moitié vide. Je vais vous répondre que cela dépend. Je sais que vous détestez quand les chercheurs vous répondent cela. Il y a un certain nombre de choses claires sur des éléments dont on a parlé aujourd'hui, de type robotique agricole, des collaborations sont instituées entre l'INRIA, le CNRS, l'INRA, l'IRSTEA, *etc.* Mon collègue Jean-Pierre Chanet peut en témoigner. Il y a des pans entiers du débat qui nous occupent aujourd'hui pour lesquels des collaborations sont en place ou se développent. Merci de les faire financer le plus possible ! Je m'arrêterai là mais le nerf de la guerre est quand même celui-là.

Ce que l'on vise, ce n'est pas seulement l'ANR. C'est une remarque que je vous demande de porter. On vise également le PIA puisque celui-ci a très peu financé l'agriculture d'une manière générale. Nous avons fait un certain nombre de démarches. Mon Président a rencontré Monsieur Gallois la semaine passée.

C'est une piste. On pense à un labex dédié à la robotique agricole. On pense à un certain nombre de choses de ce style. Il y a des initiatives fortes en matière de *Big Data*. Je pense à un autre élément de notre zoo de la recherche, qui commence à nous intéresser : les *living labs*. Il serait intéressant de regarder ce qu'un *living lab*, qui concerne un territoire entier avec une vision transverse, pourrait apporter comme élément de mise en place.

Mme Anne-Yvonne Le Dain. Monsieur Givone, ne passez pas commande s'il vous plaît. Monsieur Hetzel n'est plus Directeur général de la recherche, il est député, donc il n'a pas un sou. Il est comme moi. On essaie de faire avancer le système. Je voudrais aussi rappeler à ceux qui l'auraient un peu oublié, – j'en ai pour ma part gardé le souvenir –, que, dans les années soixante-dix, – je n'étais pas très vieille – il y avait de grands plans : le plan génie biologique médicale (GBM), le plan calcul. Ils ont tué toutes les entreprises françaises du secteur concerné. Ce n'est pas drôle ce que je viens de dire. On avait des entreprises du GBM, des entreprises de radiologie qui ont disparu. Elles ont été rachetées par Siemens ou d'autres. Il faut faire attention. Bien sûr qu'il faut une coordination, mais il faut aussi de l'impulsion, de l'investissement, des jeunes qui s'engagent dans la création d'entreprise, des entreprises qui les accueillent, non pas pour les éteindre mais pour les allumer. Tout cela n'est pas simple. C'est un changement culturel profond.

L'idée même de *l'Open Data* est un changement culturel profond. Tout le monde l'a en tête. Le système fait en sorte que cela ne marche pas. Je n'ai pas identifié de grands méchants qui manipulent le système. Je n'ai pas identifié non plus de grandes clés qui l'ouvrent. Le problème est d'essayer de tracer un chemin pour avancer. Je crois que la loi sur le numérique qui arrive pourra donner la force d'avancer, sans ignorer les inquiétudes inévitables. Mais si on est tellement inquiet qu'on interdit tout, les autres pays se développeront dans ce domaine sans nous. On utilisera ce qu'ils ont créé mais on ne créera pas la même quantité de valeur ajoutée. On n'aura donc pas la même puissance d'investissement. Je viens de dire un certain nombre d'horreurs ; j'en ai conscience ; je ne suis ainsi pas une républicaine très orthodoxe, même si j'appartiens au groupe socialiste, républicain et citoyen. Je te remercie de m'avoir donné l'occasion de préciser ma pensée à cet égard, Patrick !

M. Hervé Pillaud. Je partage totalement ce que vous venez de dire. J'avais noté deux ou trois choses par rapport à cela. Je pense que se protéger est une erreur. On a toujours cet instinct de commencer d'abord par se protéger. C'est ce qui nous tue, je pense. C'est un des deux ou trois éléments qui font que l'on n'avance pas, comme vous venez de l'expliquer. La protection est une erreur. Elle nous empêche d'avancer. Il y a des mots qui comptent. Quand nous disons que pour faire quelque chose nous allons « prendre des risques », les Anglo-saxons disent « saisir des chances ». C'est fondamentalement différent comme façon d'aborder les choses. Il faut que l'on arrive à aller « saisir nos chances ». Pour cela, franchement, je ne crois plus beaucoup aux procédures qui sont censées déclencher de la créativité. Cela ne marche pas. C'est exactement ce que vous

avez dit. Cela n'a jamais marché. Il faut que l'on passe à un mode *design*. Il faut que l'on passe dans le mode créatif. Bien sûr, il y a le fantasme du garage où se crée *Apple*, mais ce n'est pas qu'un fantasme. Vous avez évoqué tout à l'heure des hackathons et d'autres initiatives. Il faut que l'on cherche à fédérer toutes ces initiatives. Je suis d'accord avec vous : ce n'est pas une question d'argent. Bien sûr, il faudra de l'argent. Et il y en a. Il suffit de le mettre là où il faut, et de ne pas le perdre en procédures pour arriver à la fin à ne plus en avoir pour créer. Des erreurs seront faites, mais au moins des choses auront été tentées. Au final, c'est ainsi qu'apparaîtront les vraies pépites. Le vrai secret de la *Silicon Valley* est là, dans la créativité de tous ses acteurs.

Navi Radjou est Franco-indien, établi dans la *Silicon Valley* comme d'autres. Lorsqu'il évoque l'innovation frugale, il décrit très bien cette créativité. Les *fab labs*, les *living labs* ou les *farm labs* demain – il s'en ouvre un en Eure-et-Loire, à Chateaudun – me semblent très intéressants. Vous avez parlé tout à l'heure, Madame, d'initiatives à l'Institut LaSalle Beauvais qui m'ont bien plu. Il faut que nous fédérions avec les autres, celles que nous mettons en place. Pour l'instant, ce ne sont que des points qu'il faut maintenant réunir par l'action. C'est par la créativité que l'on y arrivera, et non pas par des *process*.

Mme Anne-Yvonne Le Dain. Je partage assez votre point de vue. Il n'empêche qu'une entreprise, quand elle est *start-up*, devient ensuite une petite entreprise. Il faut quand même qu'elle décolle. À un moment donné, il faut quand même du capital. Il faut donc de l'impulsion, puis du capital.

M. Hervé Pillaud. Il faut du capital et des procédures, mais celles-ci doivent venir *a posteriori* et non pas *a priori*.

Mme Anne-Yvonne Le Dain. Oui, je partage cela. J'ai été présidente d'une société de capital-risque pendant cinq ans. Je voudrais donner la parole au Monsieur au fond là-bas. Vous pouvez vous présenter s'il vous plaît ?

M. Antoine Henri. Merci Madame la Présidente. Je vais dire quelque chose plus en lien avec ma vie personnelle. Je vous rejoins entièrement, ainsi que vous, Monsieur de la Chambre d'agriculture, dans le sens où j'entends beaucoup de choses sur les institutions scientifiques. Moi-même actuellement, je suis en doctorat, donc je ne peux pas m'opposer à ce secteur.

Il y a des acteurs de la société civile qui se développent, qui font des choses très intéressantes. Par exemple, à Paris, je fais partie de La Paillasse que certains d'entre vous connaissent. Il s'agit d'un *Bio Hacker Space* extrêmement dynamique. Il conduit beaucoup de projets aussi bien avec la société civile qu'avec des entreprises. Il a signé un partenariat avec *Roche* qui est quand même un grand laboratoire. Le week-end dernier, autour de Paris, au château de Millemont, il y avait l'*Open Château* qui regroupait des personnes de la société civile qui font des projets collaboratifs. Ils avancent et font des choses, mais ils ont besoin de financements. Ils sont là, ils n'attendent qu'une chose, c'est qu'on aille

les voir. Eux, ils avancent sur des projets. Par exemple, j'ai échangé avec des personnes de la région lyonnaise, travaillant sur des projets de méthanisation à domicile. Il y a des choses qui se font. Ce sont des acteurs à ne pas manquer, alors qu'on a peu parlé de la société civile, des citoyens. Ceux qui s'impliquent sur des projets touchant à l'agriculture sont concernés aussi. C'est ce sur quoi je voulais revenir. Merci.

Mme Anne-Yvonne Le Dain. Je partage tout à fait vos analyses. Ces gens qui innovent et inventent restent d'ailleurs pour certains dans une logique semi-privée, tandis que d'autres essayent d'avancer sur le chemin de l'entreprise économique. À ce moment-là, ils peuvent aussi trouver localement ou régionalement des partenaires qui investissent. Les boutiques de région et d'autres exemples existent en France même si on en parle peu. Elles sont très efficaces. Certains ont envie d'exister, de travailler et de vivre correctement sans vouloir grossir. D'autres ont envie de grossir. Il faut qu'une société offre les capitaux à tout le monde pour le faire. Je vous rassure, on le sait. Je vous remercie de l'avoir mis ici sur la table. Je vous passe la parole Monsieur, merci.

M. Claude Kirchner. Je voulais revenir sur les questions de formation et de recherche avec quelques remarques. Mon premier point sera pour dire que je trouve extraordinaire que l'ESA fasse des MOOC pour communiquer, et s'occupe ainsi d'enseigner les avancées technologiques et la façon de les utiliser. Mais je voudrais attirer votre attention sur le fait que vous gérez ce MOOC, d'après ce que je l'ai lu, sur une plate-forme qui s'appelle *FutureLearn*. Mais *FutureLearn* donne ces données à qui ? On a *FUN (France Université Numérique)* en France qui va nous permettre de gérer les données dans un contexte éthique et déontologique largement mieux contrôlé. Faisons attention, lorsqu'on génère de nouvelles données - c'est ma première observation - à ce que celles-ci soient maîtrisées, en particulier en France. C'est très bien quand même que vous fassiez un MOOC sur le sujet.

Mon deuxième point concerne la recherche. Je pense aujourd'hui que le numérique s'affiche clairement comme une priorité au niveau national, mais également européen et international. Malheureusement, cette priorité n'est pas nécessairement traduite en termes d'action, ni concrétisée notamment en termes de financement. Cette priorité doit donc devenir effective à un moment donné. Cela nous permettrait, au niveau de l'ensemble des établissements de recherche et des universités, de faire en sorte qu'on puisse dynamiser la relation de collaboration que l'on a à l'heure actuelle. Avec ma casquette INRIA, je participe au développement de cette relation avec l'INRA, le CIRAD, l'IRSTEA ou le CNRS, bien sûr. Pour aller dans votre sens, Monsieur Hetzel, je confirme que le numérique est effectivement identifié comme une priorité. Rendons-la simplement effective.

Dernier point : je pense que l'on a un vrai sujet de préoccupation quant à la souveraineté scientifique. J'ai parlé au tout début de la souveraineté numérique. La souveraineté scientifique, cela renvoie à la question : « *Avons-nous*

aujourd'hui, ou aurons-nous demain, en France, les capacités à développer de la recherche au niveau international, c'est-à-dire en ayant la capacité de la financer, mais également en ayant accès aux données, qu'elles soient issues de l'agriculture, des publications, ou de la recherche ? ». C'est un point sur lequel il est important d'attirer votre attention.

Mme Anne-Yvonne Le Dain. Je vous remercie. Les questions sont lourdes quand même. C'est aussi pour cela que l'on fait ces auditions : pour que les questions soient sur la table et qu'on puisse continuer à travailler en tant que législateurs, et à faire notre travail auprès du Gouvernement.

J'avais une question. Vous répondrez tout à l'heure. Je vous en prie.

M. Frédéric Garcia. Sur la question relative à la recherche. J'ai quand même l'impression qu'au niveau national, les forces de recherche sur le numérique, qu'elles soient portées par l'INRIA, le CNRS ou les autres instituts, sont d'un très bon, voire d'un excellent niveau. Je n'ai pas de doutes là-dessus. Je pense qu'il faut continuer à les soutenir. La stratégie nationale de recherche soutient bien le numérique. Je n'ai pas trop de doute sur la volonté nationale de continuer. Effectivement, il faut que cela se traduise par des soutiens financiers, des postes, *etc.* Je trouve qu'aujourd'hui, la plus grande difficulté se trouve dans l'après-recherche. Comment traduire ces forces nationales de recherche autour du numérique en *start-up* ou « petites boîtes » qui deviendront de plus en plus grosses ? C'est plus là que j'ai l'impression que le bât blesse en France, et peut-être en Europe, en particulier dans l'agriculture. On a quelques exemples de création de « petites boîtes », mais on n'en a pas cinquante non plus. La difficulté est là. Est-ce que ce sont les instituts de recherche qui doivent faire cela ? Je ne crois pas. Ils doivent aider, ils sont un élément clé de l'écosystème de l'innovation mais on ne peut pas tout demander à ces instituts.

Mme Anne-Yvonne Le Dain. Puis-je me permettre d'apporter une réponse qui n'est qu'une mini-fraction de la réponse ? Dans ma région, on finance, à la demande des institutions scientifiques et universitaires – écoles et organismes – ce que l'on appelle de la recherche collaborative. Le mot est très beau. J'ai toujours été extrêmement réservée par rapport à la recherche collaborative. Je continue de l'être. J'avais imposé qu'il y ait au moins trois acteurs, deux laboratoires et une entreprise ou deux entreprises et un laboratoire, mais pas de bilatéral. Cette contrainte a pas mal fonctionné pour éviter un certain nombre de biais. Le programme est en place depuis cinq ans. Chaque soutien est prévu pour deux ans ; le principe est de s'engager au stade du pré-développement ; la recherche collaborative doit aboutir à quelque chose qui doit pouvoir aller sur le marché. C'est l'idée : on collabore pour finir une opération, et partir. Or, à chaque fois, systématiquement, on nous demande une troisième année !

On a donc un vrai problème de construction de la nature du projet : une idée de recherche est trop souvent poursuivie sans avoir en perspective une idée d'entreprise ou de création de valeur ajoutée. Il faudrait vraiment opérer la

reconnexion en ces deux aspects dans nos établissements publics. Certes, on ne peut pas adresser des injonctions de recherche, d'innovation, de création d'entreprise aux chercheurs ; c'est donc une évolution compliquée à construire. D'un autre côté, les jeunes formés dans nos établissements n'iront pas tous dans la recherche publique ou l'enseignement supérieur. Il faut donc que le secteur privé existant, grand et petit, les accueille, et pas seulement lorsqu'ils contribuent directement à le créer. Cependant, le secteur privé en France a tendance à penser que ce n'est pas à lui de faire de la recherche. Notre structure mentale est bien plus compliquée qu'aux États-Unis ou en Angleterre, ou en Allemagne. Je crois quand même qu'on n'est pas idiot et qu'on va y arriver ; tout le monde partage cette ambition, et en parle dans des termes différents, mais toujours avec beaucoup d'intelligence et de finesse. Oui, Monsieur Grumbach. Tu auras la parole après, Patrick, ne t'inquiète pas !

M. Stéphane Grumbach. Vous avez dit essentiellement des choses très proches de ce que j'allais dire.

Mme Anne-Yvonne Le Dain. Vous renoncez à parler ? Je m'en excuse !

M. Patrick Hetzel. Je voulais juste revenir sur ce que vous dites. C'est très juste. Chacun doit rester dans son rôle. Les organismes de recherche ne sont évidemment pas là pour se substituer aux entreprises. Par contre – et je pense qu'on en est tous d'accord – la question essentielle qui se pose, quand on fait un peu de *benchmark* avec l'étranger, est celle de la création d'un continuum, comme on le voit fonctionner dans la *Silicon Valley*. Je voudrais là rendre hommage à ce que fait l'Institut LaSalle Beauvais. Cet exemple est intéressant, car on voit bien qu'un continuum est mis en place entre la recherche et les dispositifs de formation. Il faut qu'on arrive à mettre en place de plus en plus de circuits courts. Fondamentalement, que ce soit l'université ou les écoles, l'objectif est de faire en sorte que ce qui est développé à un moment donné en matière de recherche, soit très vite traduit en formation, et également en apport au monde économique.

J'avais une autre question en liaison avec ce que vous dites. Sur le papier, quand vous regardez la boîte à outils pour résoudre la problématique à laquelle vous faisiez référence, on a mis en place les SATT de manière systématique sur l'ensemble du territoire national, dans le cadre du PIA. Je suis le premier à dire qu'il y a là un potentiel énorme aussi bien en matière de recherche qu'au niveau économique. Mais si on ne maîtrise pas l'ensemble du processus, la création de richesse risque de nous échapper, ce qui serait redoutable. On se retrouverait dans des situations comme celle que tu mentionnais tout à l'heure, c'est-à-dire celle des années soixante-dix : on était très fort sur le volet recherche sans pouvoir le traduire en matière économique. Question très simple : est-ce qu'aujourd'hui, dans ce domaine du numérique en général, et du numérique dans le secteur agricole, les SATT jouent leur rôle ou pas ? C'est une question un peu brutale. Elle mérite en tout cas d'être posée. Avec Alain Claeys, on a fait un rapport sur une évaluation du PIA Phase 1 en matière d'enseignement supérieur et de recherche. Globalement, on ne peut que se féliciter de la dynamique lancée. Les SATT ne

fonctionnent pas toutes de la même manière. On a la chance en Alsace d'avoir Conectus, sans doute l'une des meilleures, sinon la meilleure des SATT. Est-ce que les SATT font le job dans ce domaine ou pas ?

Mme Anne-Yvonne Le Dain. En Languedoc-Roussillon, la nôtre n'est pas mal non plus ! C'est un dispositif pertinent de mon point de vue. Globalement, il marche. Sur le terrain, il donne aux gens l'impression d'une proximité. Cette impression de proximité joue un rôle important. Si on construit des métropoles dans des régions puissantes, c'est aussi pour créer de l'activité partout, et pour ne pas vivre dans un pays splendide, où la région-Capitale aurait l'immense gentillesse de nous redistribuer sa richesse. C'est un enjeu de performance, d'économie mais aussi de société. Ce n'est pas simple.

Une question a été posée tout à l'heure à l'ESA à propos du site sur lequel on peut utiliser les données. C'est une question passionnante.

M. Pierre-Philippe Mathieu. C'est une question que l'on ne s'est d'ailleurs pas posée au moment où l'on a mis en place le MOOC, simplement parce que c'est le premier MOOC qu'on crée, et que l'idée était vraiment d'abord d'optimiser la visibilité et le contenu. On a quand même choisi une plate-forme européenne parce que les MOOC sont dominés par les États-Unis. On a donc donné un avantage compétitif à l'Europe, à l'Angleterre en l'occurrence. La prochaine étape sera de faire notre propre plate-forme ESA pour tous les MOOC spatiaux. Le *feedback* du MOOC est tellement positif que cette question de la maîtrise des données est restée en retrait pour l'instant ; on a touché une grande variété de personnes, des diplomates, des scientifiques, des gens de la société civile. Certains nous ont dit : « *Je ne croyais pas aux changements climatiques, vous m'avez démontré par des preuves objectives que c'est là. J'ai donc changé d'avis* ». D'autres en sont venus à vouloir créer de l'innovation dans le spatial. On a réussi à inspirer des gens qui, je l'espère, avec les mécanismes que vous évoquez ici, pourront transformer leurs projets en produits commerciaux.

Nous sommes en prise avec le problème de l'innovation en général, et pas seulement dans l'agriculture, à travers l'utilisation de nos données. Notre approche consiste à essayer de rassembler des communautés totalement différentes dans le même endroit physique de manière à ce qu'elles échangent des idées, de leur mettre à disposition une infrastructure de recherche qu'elles ne paient pas. On a porté celle-ci sur le *cloud* avec une possibilité de la lier directement à des services sur le *cloud*, du *processing on-demand*, toujours pour permettre de gérer des applications sans les payer. Nous avons ainsi suscité beaucoup d'idées et beaucoup d'applications.

Ce qui nous manque, c'est le fameux continuum. D'une part, il y a un besoin d'éducation des gens parce que le *Data Scientist* est un nouveau type de scientifique qui a besoin de nouvelles compétences. D'autre part, les innovateurs qui ont obtenu le *funding* pour la *start-up*, mais n'ont pas de capital pour traverser la fameuse « vallée de la mort » dans les années qui suivent, disparaissent tout de

suite. Malheureusement, l'ESA n'a pas les ressources pour offrir ce genre de capital. Pouvoir disposer de mécanismes par lesquels différents *Funders* et autres *Private Ventures* arriveraient à un certain moment serait certainement utile pour créer de la *sustainability* dans le système.

Mme Anne-Yvonne Le Dain. Merci. Je voudrais passer la parole à Monsieur. Ensuite, ce sera le tour de Stéphane Marcel.

M. Thomas Hauet. Je voulais revenir sur la *Silicon Valley*, dont on parle beaucoup, en faisant part de mon expérience. J'ai passé quelques années dans une entreprise qui faisait des disques durs en Californie. En cohérence avec l'intervention qui vient d'être faite, je pense que le rôle du financement privé est très important. Il est assez peu présent en France, notamment dans les nouvelles technologies. On parle souvent d'investissements de l'Etat français dans la recherche, qui est certes important par rapport à ce qu'on constate dans certains pays concurrents. En revanche, il y a un gros manque d'investissement privé. Je le vois dans notre domaine de l'enregistrement magnétique avec une entreprise grenobloise qui s'appelle *Crocus Technology*. Elle a ouvert un bureau dans la *Silicon Valley* pour pouvoir trouver des investisseurs. Sans cela, elle ne peut pas récupérer suffisamment d'argent pour se développer, et devenir pérenne.

Comme le montre le cas de la *Silicon Valley*, un écosystème doit se constituer. J'ai peu d'expérience du domaine de l'agriculture, mais l'écosystème m'y paraît déjà assez fort pour pouvoir se développer. C'est un point très positif. Si on prend l'histoire de la *Silicon Valley*, on voit qu'elle a pu se développer pour plusieurs raisons. Son histoire lui a permis d'aller toujours de l'avant dans une psychologie de développement, de « ruée vers l'or ». Elle dispose de deux universités qui font finalement assez peu de transfert direct, et créent assez peu de *start-up* par rapport au nombre d'étudiants de haut niveau formés par elles, lesquels se font embaucher par des entreprises éloignées, qui se développent ensuite grâce à leurs connaissances. Une part du développement des *start-up* s'effectue en lien avec les universités, mais ce n'est pas la part essentielle.

Par ailleurs, quand une entreprise est créée, le financement public est faible alors que le financement privé est très fort. Il y a énormément de déchets. Toutefois les investisseurs peuvent se permettre de se tromper dix fois, car si cela marche une fois, cela rapporte énormément d'argent, suffisamment pour réinvestir. Par rapport à cela, il ne faut pas se leurrer : n'importe quelle innovation coûte de l'argent. On aura besoin de financement privé également.

Mme Anne-Yvonne Le Dain. Je partage complètement l'analyse sur le besoin de financement privé. Pour autant, en France, on arrive à en créer, surtout depuis quelques temps avec la redynamisation du système d'investissements. La création de Bpifrance a rebattu les cartes en la matière. Le travail fait par les régions a également joué un rôle à travers le soutien aux sociétés de capital-risque, ou les modalités d'accompagnement des entreprises à l'innovation. Je crois que le paysage et les mentalités ont changé. L'approche institutionnelle et administrative

est également en train de changer. Ce n'est pas gagné, et c'est aussi pour cela que l'on fait ce genre de réunion.

J'avais dit Stéphane Marcel donc ensuite, ce sera Monsieur Grumbach. Après, nous ferons la conclusion, si vous voulez bien. À moins que quelqu'un d'autre ne souhaite parler ? Pardon, Gaëlle, excusez-moi.

M. Stéphane Marcel. Je vais être très rapide, parce que plein de choses ont été dites. Il s'avère que je suis aussi président d'un fonds d'investissement et d'amorçage en Languedoc-Roussillon. Je voudrais juste attirer l'attention des personnes présentes sur le fait que je rejoins beaucoup des propos indiquant que, finalement, l'enjeu de l'amorçage se situe à ce niveau. On commence à avoir des dispositifs qui permettent à un créateur d'entreprise ou quelqu'un qui a envie d'entreprendre dans le monde de l'innovation, de trouver quelques financements assez rapidement pour renforcer ses fonds propres. Là où le bât blesse, c'est au niveau du temps nécessaire pour capter ces financements : on s'évertue à les mobiliser alors que l'appétence au risque des investisseurs n'est pas suffisamment élevée aujourd'hui. L'entrepreneur ou l'équipe perdent une énergie folle à convaincre. Pendant ce temps, à côté, le même délai pour mobiliser les fonds nécessaires est divisé par deux ou trois. Il en va ainsi de la compétitivité de nos entreprises sur le marché, surtout sur des innovations où le *time to market* est court, lorsqu'il faut se positionner et investir rapidement dans du marketing ou dans du commerce. C'est dommage, car bien souvent, nos entreprises y arrivent mais elles peinent à atteindre leur premier million d'euros de chiffre d'affaires, alors que leur concurrente étrangère en est déjà à dix millions. On risque alors de se retrouver avec une problématique de rachat de ces entreprises par des majors notamment américaines, et là, nous perdons nos pépites. De tels rachats n'ont pas que des aspects négatifs, parce que cela permet parfois d'attirer ces majors sur nos territoires – en Languedoc-Roussillon, on en a quelques-unes – mais on a ce phénomène de perte des pépites.

Pour rebondir sur ce que vous disiez, Monsieur Garcia, je crois que la France est connue dans le monde pour sa recherche agricole et agronomique mais aussi pour sa recherche informatique et numérique. On a les meilleurs *Data Scientists* du monde. C'est une fierté franco-française, sauf que nos *Data Scientists* partent aujourd'hui parce que les entreprises du numérique ne sont pas suffisamment attractives. On rouvre là le débat de l'attractivité et de la compétitivité des entreprises.

D'un autre côté, on a un certain nombre de dispositifs dont je tiens à souligner l'intérêt : le crédit d'impôt recherche (CIR), par exemple. Il nous permet d'être attractifs pour des entreprises comme *Facebook* qui est, par exemple, en train de mettre en place, je crois en région parisienne, son laboratoire de recherche dans le domaine du *Big Data* appliqué à l'économie, et des technologies neuronales au niveau informatique. On a quand même des atouts. Mais cela va plus vite qu'avant ; le paradigme change et il faut en prendre conscience. Il faut

que l'on se fasse du mal au niveau de nos dispositifs de réflexion ; car l'appétence au risque doit devenir plus importante en France.

Mme Anne-Yvonne Le Dain. Merci de l'avoir dit. Je voudrais à titre personnel, avant de conclure et de passer la parole à Gaëlle et Stéphane, dire qu'en France, on a un peu de mal à formaliser les choses. La recherche scientifique produit des connaissances, qui sont diffusées de deux manières : d'une part, elles sont transmises aux étudiants, qui vont tout simplement créer leur entreprise, les utiliser dans les entreprises qui les recrutent ; d'autre part, elles se diffusent à travers la recherche collaborative avec des entreprises. Mais la recherche ne peut pas créer toute seule la valeur ajoutée future de la Nation ; elle ne peut pas en supporter toute la charge. Les entreprises doivent prendre leur part, c'est une question de confiance, laquelle n'exclut pas le contrôle.

Il faut quand même changer le postulat : les docteurs formés dans nos universités et nos écoles ont, pour la plupart, vocation à être recrutés par les entreprises, et non pas à faire douze postes entre Israël, les États-Unis, l'Australie, et maintenant la Tanzanie ou la Russie, avant de se demander s'ils ne vont pas créer une supérette une fois arrivés à quarante ans. Ce que je viens de dire est un peu violent ! C'est une question de mentalité. Je crois qu'on est en train de changer, mais il faut aussi que les entreprises jouent le jeu. C'est pour cela qu'il y a le CIR, vous l'avez dit. Il est quand même performant, et c'est une manière gentille d'en parler. Mais il y a aussi le CICE que les entreprises utilisent, qui marche. Il permet notamment de recruter des gens à des niveaux de rémunération bas, via le remboursement des charges. Le travail fait par la puissance publique en direction du monde de l'entreprise pour lui dire de prendre des risques, de recruter, de travailler avec des scientifiques en innovant dans le monde entier, existe. Les entreprises doivent saisir la main tendue. Gaëlle ?

Mme Gaëlle Kotbi. Merci. Effectivement, le changement est en cours pour la mise en réseau, et la création d'une « atmosphère » permettant de développer l'écosystème. Je voudrais insister sur un point : les synergies entre l'enseignement et la recherche sont importantes et se nourrissent de voyages d'étude et d'un mixage des publics. Ce sont des moments où les esprits vont avancer et échanger. J'apporte mon soutien à la discussion là-dessus ; je rejoins les positions des précédents intervenants, et suis prête à envisager des collaborations en ce sens.

Je voulais ensuite ajouter des chiffres sur l'entrepreneuriat. Et en agriculture, je pense que la question de la temporalité est particulièrement forte. Tous secteurs confondus, les statistiques nous donnent un ordre d'idée – même si ce ne sont que des moyennes qui cachent les extrêmes : on considère globalement qu'il faut deux à trois années pour faire émerger l'idée d'un projet ; c'est le temps nécessaire pour que le porteur d'un projet – pas forcément un étudiant – parte d'une envie, d'une idée, puis la porte jusqu'au point de créer une entreprise en déposant des statuts. Ensuite, il faudra entre cinq et sept années pour que l'on commence à savoir si la structure montée a une chance de vivre et de se

développer. On constate qu'on est sur des temporalités très éloignées des deux ou trois années qui ont été évoquées pour les mesures de soutien. Ce problème est essentiel.

Mme Anne-Yvonne Le Dain. J'en ai bien conscience. Le projet collaboratif ne sert pas qu'à créer de l'entreprise. La fiction est qu'au bout, on va trouver quelque chose de valorisant. Je pense qu'il faut quand même trouver un petit marché avant de trouver le grand. En Languedoc-Roussillon, c'est ce que l'on fait. Une entreprise qui vivote pendant trop longtemps n'est pas dans le bon mode, on le sait. Pour marcher, une entreprise doit comprendre un technicien, un commercial et un patron. Ce n'est pas facile. Des fois, il y a des échecs. Des fois cela va lentement. On est tous conscients de cela, mais c'est le chemin qu'il faut prendre.

Je voudrais souligner qu'en France, depuis près de quinze ans, le secteur financier ne va pas sur le risque. On le dit peu. L'argument présenté est qu'il gère l'argent des Français et qu'il doit donc être prudent. On a pu constater pourtant certaines imprudences et leurs conséquences assez récemment. Ce discours est donc assez peu audible. Malheureusement il est tenu. Sur ces questions, je peux vous dire qu'une entreprise doit aujourd'hui s'appuyer sur la puissance publique pour arriver à un million d'euros de chiffre d'affaires. Elle est aidée par les collectivités territoriales, à travers leurs sociétés de capital investissement, et le système parallèle. Des chefs d'entreprise ayant bien gagné leur vie ont en effet créé des fonds d'investissements parallèles pour donner des coups de main. Il y a aussi OSEO, et maintenant la BPI. On est dans un système un peu compliqué. Tant qu'une entreprise n'a pas créé un voire deux millions d'euros de chiffre d'affaires, selon le secteur, elle ne trouve pas de gros investisseurs.

Le deuxième écueil majeur, c'est que lorsque les premiers investisseurs arrivent, ils entrent dans le capital et l'écrasent. Les fondateurs n'ont ensuite quasiment plus rien, sauf leur intelligence et leurs bras s'ils continuent à perdurer. On a donc une vraie difficulté de mentalité dans le rapport à la création d'entreprise et à l'argent nécessaire, au niveau des capitaux à mettre. Cela change cependant car le monde change. Cela a été dit par certains d'entre vous. Il y a chaque année entre dix et douze mille nouveaux agriculteurs qui arrivent avec une autre philosophie et d'autres compétences. Ces jeunes vont changer la donne. Il ne faut pas croire que tout est résolu, car au final, nous sommes toujours en retard. Soyons dans le système et dans le monde tel qu'il va. Je devais passer la parole à Stéphane Grumbach.

M. Stéphane Grumbach. Juste une petite remarque : je crois qu'il est vraiment fondamental pour l'Europe de regarder comment l'Asie fait face, de son côté, à la domination des systèmes numériques américains, mais avec une résilience plus forte. Dans un paysage politique divers – si on compare la Chine et le Japon, ou la Corée – elle a une capacité à répondre. Il y a là des modèles dont on doit s'inspirer et qu'on ignore généralement. Je vois très peu de connaisseurs

de l'Asie, ou de gens qui regardent ce qui se passe en Asie, dans les discussions auxquelles je participe. Je voulais juste faire cette remarque. Merci.

Mme Anne-Yvonne Le Dain. Ce n'est plus tellement une *terra incognita*, un espace que l'on doit parcourir avec doigté et talent. Si vous le permettez, je vais vous dire maintenant quelques mots de conclusion. Je voudrais remercier les services de l'OPECST qui font un travail remarquable et accompagnent les députés que nous sommes avec beaucoup de doigté. Après, j'interprète, rassurez-vous !

CONCLUSION

Mme Anne-Yvonne Le Dain, député, vice-présidente de l'OPECST.

Cette audition publique a permis de bien prendre conscience de ce qu'est le *Big Data* et plus encore de ce qu'il n'est pas, ce qui n'était pas évident au départ. Nous avons pris conscience de la manière dont il peut être mis en œuvre ou optimisé dans l'agriculture, y compris au niveau international, et des raisons pour lesquelles c'est une étape incontournable de l'évolution de l'agriculture. Nous n'en sommes plus au stade de la charrue et des bœufs ; suivre les évolutions est indispensable. Le *Big Data* présente un intérêt économique pour les utilisateurs à travers des services offerts, et un intérêt collectif en matière d'effort pour la préservation de l'environnement. Il est un atout pour la productivité, l'efficacité, la performance, les ratios mais aussi pour la prise en compte de la problématique environnementale extrêmement importante. L'agriculture est souvent accusée. L'urbanisme est souvent accusé. La question de l'environnement est première parce qu'on est sur un territoire qui vit au-dessus et en dessous. Nous en sommes tous tributaires alors que la population mondiale ne fait qu'augmenter, comme cela a été souligné. Les territoires agricoles ne font que diminuer en surface ou en usage, ce qui n'est pas tout à fait pareil. La grande question est la préservation de l'environnement.

L'autre grande question porte sur la prise en compte du changement climatique et la lutte contre le changement climatique. Je sépare bien les deux concepts même si on sent bien qu'ils sont liés, mais je pense qu'on ne doit pas tout globaliser. Les approches trop globales ne laissent pas d'espace pour saisir les perspectives et les opportunités de travail, voire d'affaires. La prise en compte du changement climatique est une chose. La lutte contre le changement climatique va de pair, mais il faut simplement voir ce qui se passe et en tenir compte et proposer les services afférents. On peut aussi en proposer pour diminuer les impacts du changement climatique, voire éviter qu'il n'aille trop loin. Le *Big Data* est une piste pour la gestion de ces gros enjeux. On en a peu parlé.

Ensuite, je constate que les pistes qui se dessinent concernant la reprise en main de l'offre de service s'appuyant sur le *Big Data*, renvoient à des problématiques très transversales en matière de numérique. La première dimension est le développement de l'innovation. On en a longuement parlé dans le cadre de cette dernière table ronde. C'est la seule manière saine de développer une alternative à l'offre américaine qui a pris de l'avance, et aux offres qui vont venir d'autres continents – indien, européen et latino-américain, comme cela a été souligné. On parle peu de l'Amérique du Sud mais elle est extrêmement active dans ce domaine.

L'OPECST s'est beaucoup intéressé aux conditions de l'innovation, aux mécanismes de soutien financier, à l'adaptation « aux » ou « des » barrières réglementaires. Les nouveaux services du *Big Data* doivent évoluer au profit de nos entreprises, de nos territoires, et sur notre territoire, pour l'emploi en France. Le terme « territoire » est un mot-valise assez compliqué. Je le décline de différente manière. Ce mot devient d'ailleurs dangereux, car on ne sait plus très bien de quoi on parle.

La deuxième dimension transversale est la gestion des données dans les conditions de prudence qui s'imposent face aux risques d'atteinte à la vie privée bien sûr, et aux intérêts économiques. En même temps, il y a une obligation d'ouvrir l'accès à ce gisement précieux qui doit être une ressource au service de l'humanité. C'est la grande question. Les portails des autres sont grand ouverts. L'usage qu'ils en font crée de la valeur. Nous n'avons ni grand portail ni grande valeur alors qu'on a beaucoup de talent. Nous devons construire maintenant. On fait déjà plein de choses. Stéphane Marcel a cité des initiatives, y compris dans le monde agricole. Il s'est créé une dynamique d'ouverture et de puissance dans laquelle il faut s'inscrire. La démarche doit reprendre l'approche *Open Source* de l'Agence spatiale européenne qui a notre total soutien ; il sera nécessaire de prohiber toute tentative d'exploitation exclusive des fonds et des flux de données, ainsi que des flots de données.

Voilà ce que je voulais vous dire. J'en profite pour vous signaler que vous trouverez dehors ce rapport que nous avons établi, mon collègue sénateur Bruno Sido, Premier vice-président de l'OPECST, et moi-même. On l'a sorti à la fin de l'année dernière sur le thème de la sécurité numérique et des risques. Au lieu de s'interroger en termes de peur, on s'est interrogé en termes d'enjeu et de chance pour les entreprises et pour nos territoires. Dans ce document, qui comprend deux parties, il y a une petite synthèse qui montre que nous sommes bien naïfs de croire que tout va bien, et que nous sommes bien frileux de croire qu'il faut tout fermer. Le petit dessin sur la couverture tente de résumer cela : une cage à oiseaux avec une France dont les données s'échappent de partout par l'ouverture de la cage. On peut l'interpréter comme une peur ou comme une opportunité. Mon analyse personnelle est de dire que c'est une chance.

Je vous remercie d'avoir participé à cette audition. En ce qui concerne les MOOC, je voudrais souligner qu'on a commencé à parler en France globalement de la question des MOOC quand les Américains ont sorti les leurs. Tous ceux qui ont tenté de faire des MOOC dans les universités, écoles d'ingénieurs et établissements français se sont heurtés à un mur de désintérêt total des institutions. C'est une illustration du message d'ouverture des portes, des cages aux oiseaux. Il faut que les oiseaux s'envolent.

Bonne journée et merci.

EXTRAIT DE LA RÉUNION DE L'OPECST DU 8 JUILLET 2015 PRÉSENTANT LES CONCLUSIONS DE L'AUDITION PUBLIQUE

M. Jean-Yves Le Déaut. L'audition publique sur « La place du traitement massif des données (*Big Data*) dans l'agriculture : situation et perspectives » s'est tenue le 2 juillet à la demande conjointe du premier vice-président Bruno Sido et de moi-même.

Elle a été organisée pour répondre à une préoccupation manifestée par des responsables de l'INRA et de l'IRSTEA et par des représentants du monde agricole lors de la visite que nous avons effectuée ensemble au dernier salon de l'agriculture, en mars dernier.

Nous avons été alertés sur le fait que des constructeurs d'équipements (tracteurs et autres) placeraient des capteurs couplés à des systèmes de transmission de données (c'est-à-dire ce qu'on appelle des « objets connectés ») sur les machines qu'ils vendent, pour recueillir des informations et alimenter des traitements de type *Big Data*.

La sollicitation nous a paru très sérieuse, et nous sommes revenus du salon avec l'idée de creuser ce sujet dans le cadre d'une audition publique ouverte à la presse, suggestion que le bureau de l'OPECST a appuyée le 10 juin.

Ce sujet fait en effet typiquement appel aux compétences de l'OPECST telles qu'elles lui sont dévolues par la loi du 8 juillet 1983, c'est-à-dire explorer les questions relevant de l'évolution des sciences et des technologies pour détecter à l'avance les enjeux législatifs qui peuvent leur sont associés.

Cette audition a fourni l'occasion d'aborder, pour la première fois à l'OPECST, la problématique du *Big Data* et des objets connectés en tant que tel ; c'était l'objet de *la première table ronde* de l'audition, qui avait une vocation pédagogique pure.

Cette table ronde a retenu comme définition du *Big Data* la capacité à effectuer des traitements massifs de données et à offrir des services associés. Elle a mis en lumière les ruptures techniques à l'origine de la capacité d'accumulation d'un nombre considérable de données, multipliée par un milliard de 1950 à nos jours, le prix de stockage à l'unité diminuant lui-même parallèlement d'un facteur de l'ordre du milliard. Entre temps, les progrès de la programmation ont conduit à la mise au point d'algorithmes adaptés au traitement de volumes considérables d'informations, qui s'attachent à découvrir, *via* ce qu'on appelle le *Data Mining*, des motifs récurrents dans les flux de données.

Cela a rendu possible des services d'un type nouveau permettant de mieux cibler les offres commerciales à partir d'une analyse plus fine des comportements de consommation, mais aussi de repérer des « signaux faibles » au-delà des phénomènes directement accessibles, ce qui est utile, par exemple, pour identifier la survenue imminente de pannes, et organiser en conséquence une maintenance prédictive.

S'agissant des conditions pratiques de la mise en œuvre du *Big Data* en agriculture, qui faisait l'objet de *la deuxième table ronde*, les données proviennent d'objets connectés qui analysent, par exemple, l'hétérogénéité de l'état des parcelles, les paramètres d'ambiance au sein des bâtiments et des serres, ou renseignent sur la santé des animaux. Le traitement des données fournit des outils d'aide à la décision, par exemple concernant le moment optimal d'épandage d'un fertilisant, mais aussi pilote des systèmes entièrement automatisés comme les régulateurs de température ou d'hygrométrie, les robots de traite ou les doseurs d'engrais embarqués s'ajustant sur la productivité différenciée des parcelles. Ce traitement des données alimente aussi des systèmes de surveillance pouvant déclencher des alertes sanitaires. La précision de ces dispositifs est encore accrue par l'élargissement de la base du traitement à l'ensemble des exploitations, permettant de caler les analyses sur des références multiples.

À partir du constat que la diffusion des techniques du *Big Data* devient clairement un enjeu de compétitivité pour l'agriculture, *la troisième table ronde* avait pour objet d'envisager les démarches stratégiques possibles, sachant que la constitution d'une capacité d'offre performante suppose un minimum de souveraineté.

Il s'agissait notamment d'examiner le degré de maîtrise que les agriculteurs et les pouvoirs publics peuvent et doivent avoir sur l'installation de capteurs ou objets connectés, compte tenu des conséquences que le développement des services fondés sur le *Big Data* ont, ou auront, pour la souveraineté nationale ou européenne dans le domaine de l'agriculture.

Aujourd'hui, une partie des données produites par les objets connectés est récupérée directement par les fabricants d'équipements comme *John Deere*. Et comme l'a signalé M. Stéphane Grumbach de l'INRIA, le rapide succès, depuis 2006, de *Climate Corporation* illustre le risque pour l'Europe de tomber bientôt sous une nouvelle forme d'emprise américaine à l'image de celle déjà à l'œuvre dans le numérique : à partir d'un retraitement de données météorologiques permettant de piloter l'arrosage en lien avec de capteurs installés dans les champs, l'activité de *Climate Corporation* s'est étendue à la fourniture d'assurances pour aujourd'hui près de la moitié des surfaces céréalières aux États-Unis ; en 2013, elle est passée sous le contrôle de *Monsanto*.

L'offre américaine occupe ainsi une position de force sur les services associés au *Big Data*, et l'emprise qui en résulte pourrait se traduire par une atteinte à l'indépendance de la production agricole française et européenne.

M. Pierrick Givone, directeur général délégué à la recherche et à l'innovation de l'Irstea, a rappelé que l'information sur l'état des récoltes avait une valeur stratégique au temps de la guerre froide, au point que les Américains avaient lancé les satellites Landsat pour surveiller l'agriculture soviétique, et que, d'ores et déjà, en l'absence d'alternative européenne, tant que Galiléo ne sera pas pleinement opérationnel, la moitié des tracteurs français équipés de GPS verront leur performance dépendre entièrement du bon vouloir des autorités américaines.

Vis-à-vis du risque de mainmise nord-américaine, il a été constaté que l'ensemble des participants rejetaient tout à la fois l'angélisme et le protectionnisme, et prônaient une reprise en main collective de la gestion des données par l'ensemble des acteurs du secteur ; cette démarche collective étant unanimement perçue comme plus efficace qu'une réaction en ordre dispersé. Cela passe par un soutien politique fort à la mise en place d'une plateforme commune de gestion des données, d'abord à l'échelle de la France ; mais cette démarche est conçue d'emblée comme ouverte, en mode *Open Source*, et ensuite comme annonciatrice d'une solution élargie au niveau européen.

À cet égard, M. Pierre-Philippe Mathieu de l'ESA a présenté l'apport potentiel des satellites Sentinel 1 et 2, qui prennent en charge la composante spatiale du programme européen Copernicus d'observation de la Terre et qui fournissent des images précises, à dix mètres près, des zones agricoles.

Ce dispositif n'a pas d'équivalent dans le monde. La transparence qu'il induit sur l'évolution des récoltes permettra de mieux préserver les habitants de la planète contre les tentatives de manipulation des cours.

Les données sont accessibles en mode *Open Source*. En lien avec la Banque mondiale notamment, des aides sont prévues pour les entrepreneurs qui participent à l'exploitation des données, afin de les transformer en informations directement utiles, par exemple pour envoyer un SMS à l'agriculteur lui signalant que c'est le bon moment pour effectuer la récolte.

Les pistes qui se dessinent ainsi s'agissant de la reprise en main de l'offre de services s'appuyant sur le *Big Data* renvoient finalement à des problématiques transversales dans le domaine du développement du numérique.

La première dimension transversale concerne l'encouragement à l'innovation pour développer une alternative à l'offre américaine qui a pris de l'avance : tous les mécanismes de soutien financier, et toutes les simplifications réglementaires prévues pour encourager l'innovation doivent évidemment jouer au profit des nouveaux services du *Big Data*.

La deuxième dimension transversale concerne la gestion des données, avec les conditions de prudence qui s'imposent pour éviter les risques d'atteinte à la vie privée ou aux intérêts économiques et, en même temps, les obligations d'ouvrir l'accès à ce gisement précieux qui doit jouer comme une ressource au service de l'humanité.

De là, notre soutien total à la démarche *Open Source* de l'Agence spatiale européenne, et la nécessité de prohiber toute tentative d'exploiter de façon exclusive des fonds de données.

En conclusion, cette audition publique a permis de bien comprendre ce qu'est le *Big Data*, la manière dont il est mis en œuvre dans l'agriculture, les raisons pour lesquelles il s'agit d'une étape incontournable de l'évolution de l'agriculture. En effet, il présente à la fois un évident intérêt économique pour les utilisateurs, et un intérêt collectif dans le cadre de l'effort pour la préservation de l'environnement et la lutte contre le changement climatique, puisqu'il permet globalement la diminution des intrants sans pénaliser les rendements.

Cette audition a permis de vérifier que la communauté agricole et scientifique française n'avait pas tort de s'inquiéter des conditions dans lesquelles s'effectuait cette nouvelle révolution en cours de l'agriculture, car il y a bien, en arrière-plan, un enjeu de souveraineté face au dynamisme de certains pays et notamment des États-Unis.

Un maître mot résume les débats : la nécessité de maîtriser les données dans un monde ouvert, où les frontières n'existent plus. Les nouvelles technologies de la communication remettent en cause l'espace, le temps et la gouvernance. Les problèmes d'usage, de propriété et de contrôle des données doivent être précisés comme les questions liées à la revente de ces données, au risque d'atteinte à la vie privée ou à des intérêts économiques ou encore à la sauvegarde de renseignements sensibles.

Cette communauté agricole et scientifique est prête à se mobiliser pour défendre son propre modèle du *Big Data* à partir de la constitution d'une plateforme commune de gestion des données, à l'échelle de la France, en *Open Source*, conçue comme ouverte. Ce gisement précieux doit être exploité comme une ressource au service de l'humanité. Une convention internationale doit donc prohiber l'utilisation exclusive de ces données. C'est la principale proposition qui ressort de cette audition publique.

Une autre proposition consiste à entériner le besoin d'un soutien fort du Gouvernement, non pas pour mettre en place une protection mais pour disposer d'un cadre juridique clair pour la gestion de ces données et pour favoriser la création d'une offre de services innovants. Dans ce domaine comme dans d'autres, le soutien à l'innovation est essentiel et doit être une des priorités gouvernementales. C'est la seule manière de développer une alternative à l'offre nord-américaine qui a pris de l'avance. À cette fin, l'État doit mobiliser les mécanismes spécifiques de soutien financier et d'adaptation aux barrières réglementaires.

Il serait d'autant plus justifié que ce dossier fasse l'objet d'une implication politique forte – et l'OPECST va agir en ce sens – que la préservation de la souveraineté en matière d'agriculture a pour enjeu, en arrière-plan, la maîtrise encore plus cruciale du développement du territoire.

Un futur rapport de l'OPECST pourrait permettre de clarifier les liens entre les différents types d'objets connectés, le traitement massif des données et les services associés.

M. Bruno Sido. Cela fait longtemps que, sur les moissonneuses-batteuses, des dispositifs assurent un calcul du rendement en temps réel et qu'une exportation de données sur une clef USB permet d'obtenir la carte des rendements. On peut ainsi éditer cette carte à partir d'un ordinateur ou insérer ensuite cette clef USB dans un pulvérisateur pour moduler automatiquement les doses, car c'est important d'ajuster les apports d'engrais aux besoins différenciés des différentes zones du sol. Au passage, on réalise des économies sur les quantités d'intrants. Mais tout cela s'est mis en place en l'absence de ce qu'on appelle les objets connectés, et je ne vois pas en quoi il faudrait forcément lier la récupération de données ou la constitution de nouveaux gisements de données, par satellite par exemple, au développement des objets connectés.

M. Jean-Yves Le Déaut. Un objet connecté est un capteur, qui a la capacité en plus de transférer les informations qu'il récupère.

M. Bruno Sido. Mais quel est l'intérêt ? Qu'est-ce que cela apporte de nouveau ?

M. Jean-Yves Le Déaut. Ce qu'il se passe, c'est que, à l'image de certaines stratégies liées au dépôt automatique de *cookies* sur les ordinateurs, l'accord implicite donné à une surveillance, par le fournisseur, de l'usage qu'on fait du service, conditionne la possibilité d'accéder à ce service. Cet accord est considéré comme donné implicitement du simple fait qu'on achète l'équipement.

M. Bruno Sido. Si c'est de cela qu'il s'agit, je trouve légitime de s'insurger, car ce n'est pas normal que l'agriculteur livre ainsi ses secrets, sauf s'il le fait volontairement tout en ayant réellement la possibilité de ne pas le faire.

M. Jean-Yves Le Déaut. Tu rejoins là nos conclusions.

Mme Anne-Yvonne Le Dain. J'aurai deux observations. D'abord, il faut se rendre compte que nombre de produits intégrant des dispositifs de gestion numérique des données ne sont pas nécessairement payés à leur coût réel, car leur prix inclut une décote correspondant à la rémunération des services que le fournisseur va construire par ailleurs sur la base des informations sur les profils de consommation qu'il recueille implicitement auprès des utilisateurs. En ce cas, les utilisateurs, pour ne pas faire partie de l'échantillon de recueil des données, doivent accepter de payer plus cher.

Ma seconde observation concerne l'approche américaine des données, qui sont considérées outre-Atlantique comme un matériau brut, ne prenant une véritable valeur économique que lorsqu'elles ont fait l'objet d'un traitement en vue de la fourniture d'un service. Cette approche conduit à promouvoir un régime d'accès ouvert aux données, pour favoriser l'apparition, à l'initiative du marché, des stratégies de retraitement (*Data Mining*) qui vont permettre d'en extraire des informations à valeur commerciale. C'est dans cette conversion des données en services que les Américains ont un temps d'avance.

M. Bruno Sido. D'ores et déjà, s'agissant des moissonneuses-batteuses, les concessionnaires disposent d'un moyen, sur l'équipement lui-même, et à l'insu de son propriétaire, pour éditer à tout moment un état complet de son utilisation, y compris les moments où le moteur a tourné sans que les fonctions de moissonnage ou de battage aient été utilisées. Bien entendu, ces informations concernent aussi les performances en service, comme le nombre de quintaux réalisés à l'hectare. Je me suis rendu compte de cette possibilité en voyant faire le technicien lors du contrôle d'un matériel avant une vente d'occasion.

Mme Anne-Yvonne Le Dain. Il ne faut avoir aucun doute sur le fait que ces données sont remontées jusqu'au fabricant. À partir de là, pour les gestionnaires publics, il s'agit de savoir s'il faut lutter contre ces pratiques, en ce cas au niveau mondial, ou s'il faut, au contraire, s'y adapter pour conquérir des positions *via* une offre alternative.

M. Jean-Yves Le Déaut. C'est le sens même des conclusions que je propose, qui soulignent la nécessité d'un meilleur contrôle des données et de leur usage. La récupération des données n'est pas qu'un enjeu commercial ; elle peut avoir aussi une dimension stratégique en repérant les années de mauvaises récoltes et, ainsi, les périodes de fragilité d'une société.

M. Bruno Sido. C'était déjà l'objet d'un programme de surveillance par satellite des Américains pendant la guerre froide.

Mme Anne-Yvonne Le Dain. L'enjeu commercial sera de savoir si les Européens sont capables de battre les Américains au jeu de la récupération et du traitement des données.

M. Bruno Sido. Le Président Obama a dit ouvertement que l'Internet était un instrument de domination américaine.

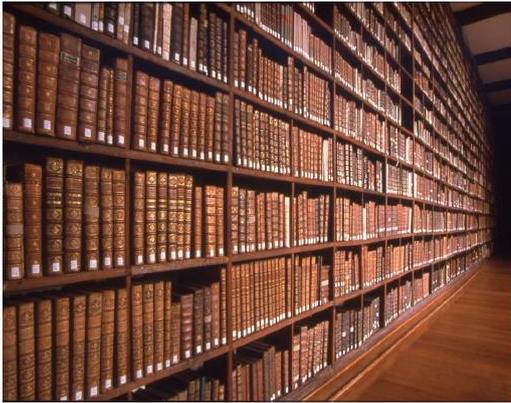
Mme Anne-Yvonne Le Dain. L'Internet a été inventé par les Européens mais n'a pas trouvé tout de suite de relais d'intérêt dans les sociétés européennes, car celles-ci n'ont pas su s'approprier les potentialités de cette innovation.

L'OPECST a approuvé à l'unanimité les conclusions et a autorisé la publication du rapport.

ANNEXES

**ANNEXE 1 :
INTERVENTION DE M. THOMAS HAUET, MAÎTRE DE CONFÉRENCE
À L'UNIVERSITÉ DE LORRAINE**

Disque dur magnétique: la base du Big Data



Bibliothèque papier

1980



Serveur WEB

2015

Thomas Hauet (Université de Lorraine)



Méthodes de stockage de l'information

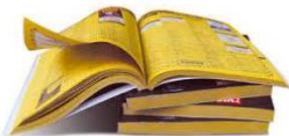
Mémoire semi-conducteur (DRAM, SRAM, Flash)



Caractéristiques majeures :

- Contenance totale
- Densité d'info
- Vitesse d'accès
- Coût
- Pérennité/ Volatilité
- Mobilité

Papier



Bande magnétique



Disque dur magnétique



Disques gravés



Evolution du disque dur: 1956 à aujourd'hui



RAMAC

- 5 MegaBytes
- 2 kbits/in²
- 150 kbit/s
- \$10,000,000/Gbyte

1956



3380 system

- 1.26 GigaBytes (GB)
- 12.2 Mbits/in²
- 20 Mbit/s
- \$100,000 /Gbyte

1981



Microdrive

- 1 GB
- 15.2 Gb/in²
- 38.8 Mbits/sec
- \$1 /Gbyte

2000



TravelStar LP

- 0.72 GB
- 644 Mbits/in²
- 39.5 Mbits/sec
- \$100 /Gbyte

1994

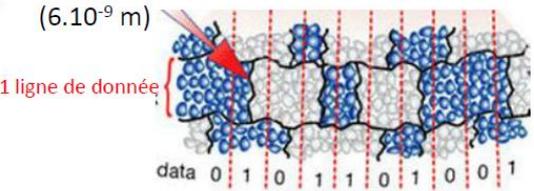
Evolution du disque dur: 1956 à aujourd'hui

**4 Tbytes
HDD**

Moteur (rotation)

Média magnétique
Grains magnétiques
(6.10⁻⁹ m)

1 ligne de donnée



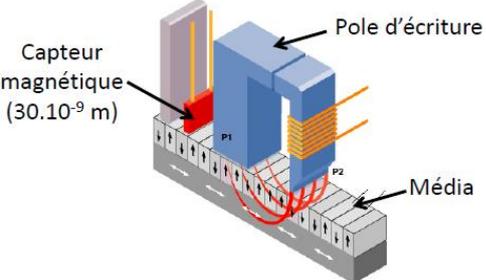
data 0 1 0 1 1 0 1 0 0 1

Microdrive

- 4000 GB
- 450 Gb/in²
- 6 Gbits/sec
- \$0.04 /Gbyte

2014

Tête de lecture / Tête d'écriture

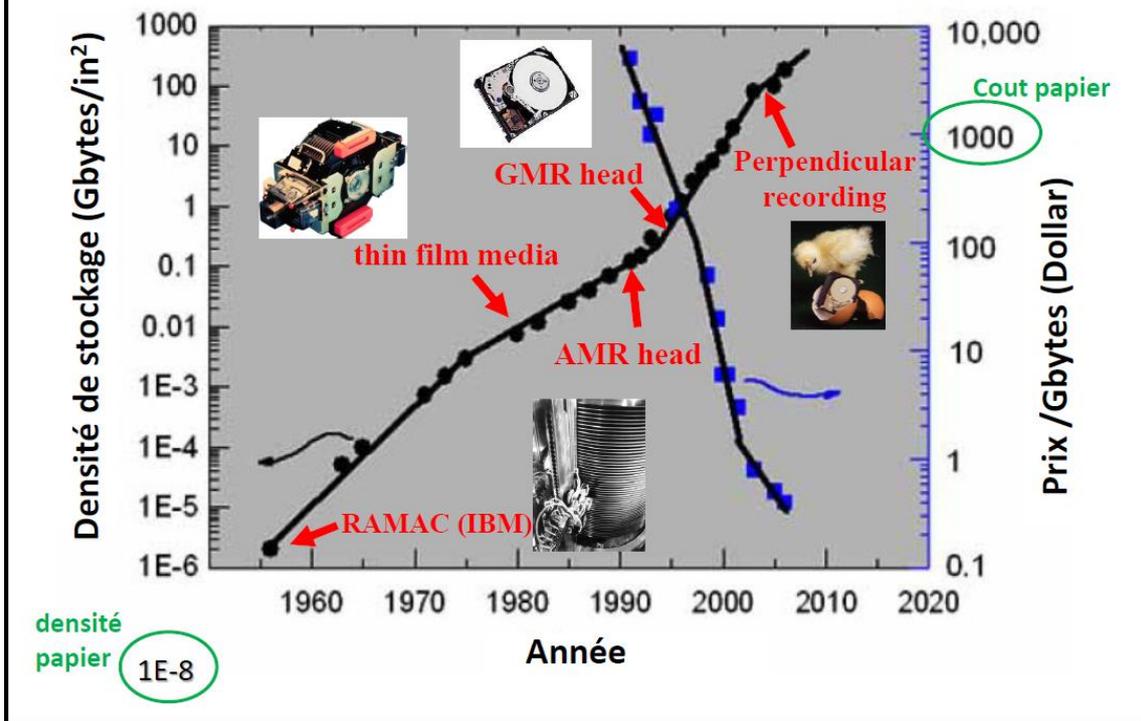


Captur magnétique (30.10⁻⁹ m)

Pole d'écriture

Média

Sciences fondamentales et sauts technologiques



Disque dur magnétique: la base du Big Data



Bibliothèque papier 1980



Serveur WEB 2015

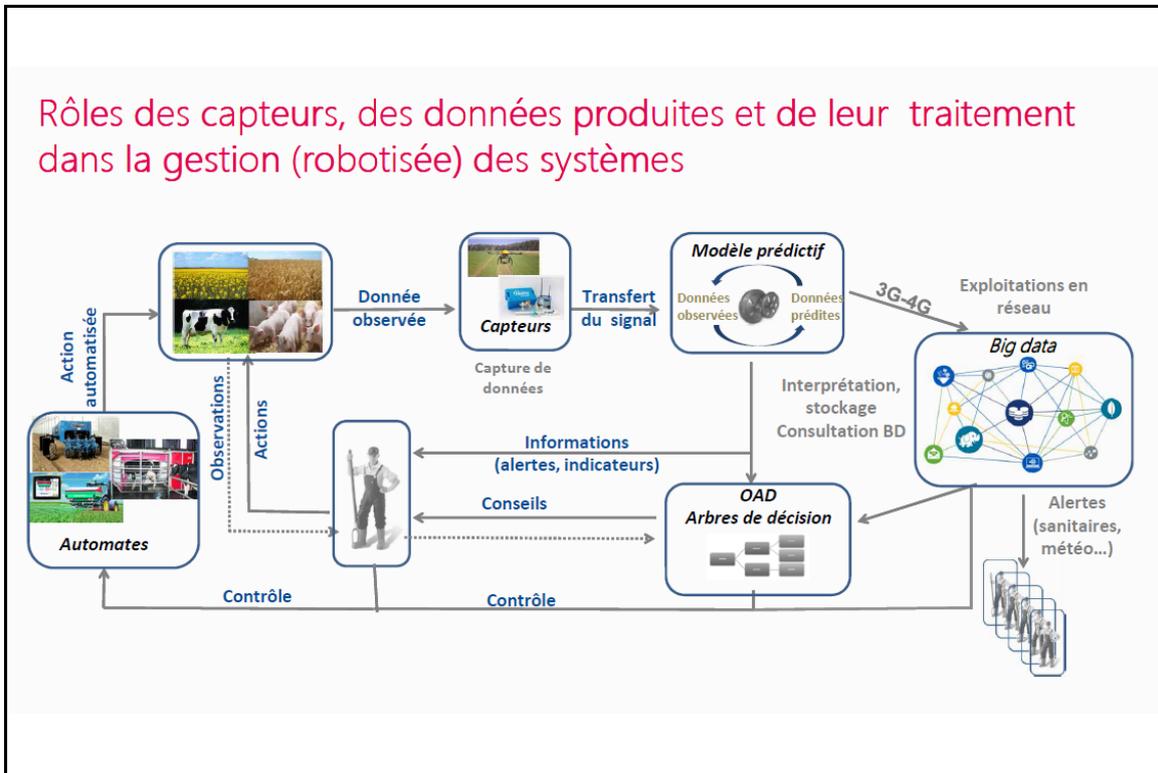
Conséquence : $3 \cdot 10^{12}$ Gbytes de stockage dans le monde
60 millions de fois la quantité d'information papier

Références : <http://www.crn.com/slide-shows/storage/240142353/the-history-of-the-hard-drive-and-its-future.htm/pgno/0/3>
http://www.refletsdelaphysique.fr/images/stories/news/Dossier_spintronique-Reflets_de_la%20Physique.pdf
https://www.canal-u.tv/video/universite_de_tous_les_savoirs/l_enregistrement_magnetique_une_nanotechnologie_grand_public_claude_chappert.4315



ANNEXE 2 : INTERVENTION DE M. JEAN-LOUIS PEYRAUD, CHARGÉ DE MISSION, DIRECTION SCIENTIFIQUE « AGRICULTURE », INRA

Rôles des capteurs, des données produites et de leur traitement dans la gestion (robotisée) des systèmes



Exemples de capteurs et d'utilisation

Suivi des rendements et de la qualité



Localisation des traitements au champ



Progrès génétique



Conduite des troupeaux

La vache connectée

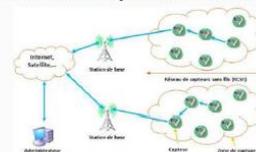


À la récolte

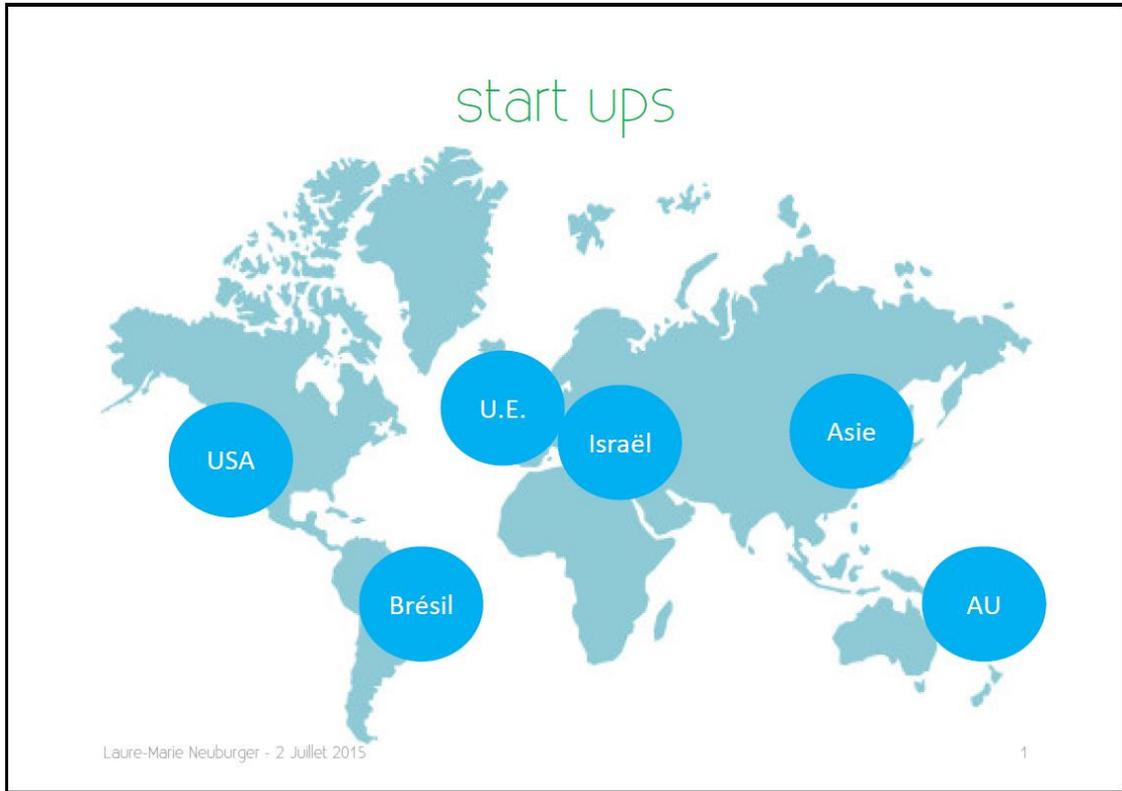
Régulation d'ambiance dans les bâtiments et les serres



Réseaux d'épidémiologie



ANNEXE 3 : INTERVENTION DE MME MARIE-LAURE NEUBURGER, ENTREPRENEUR



Tendances émergentes

Vente directe / circuits courts

Distribution en ligne

Cultures hors sol - hydroponie

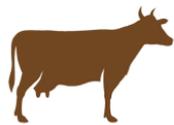
Robotique + capteurs + internet des objets

Mouvement « Maker » + « DIY »

AB / Agroécologie

Laure-Marie Neuburger - 2 Juillet 2015

3



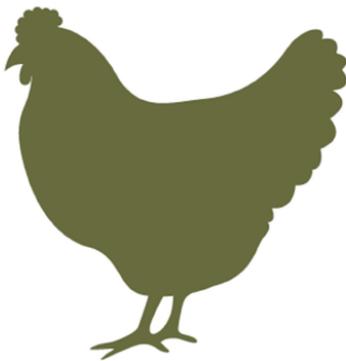
1 500 000 000
BOVINS



7 325 138 103
êtres humains



25 000 000
tracteurs



22 900 000 000
volailles



2 138 000 000
ovins et caprins



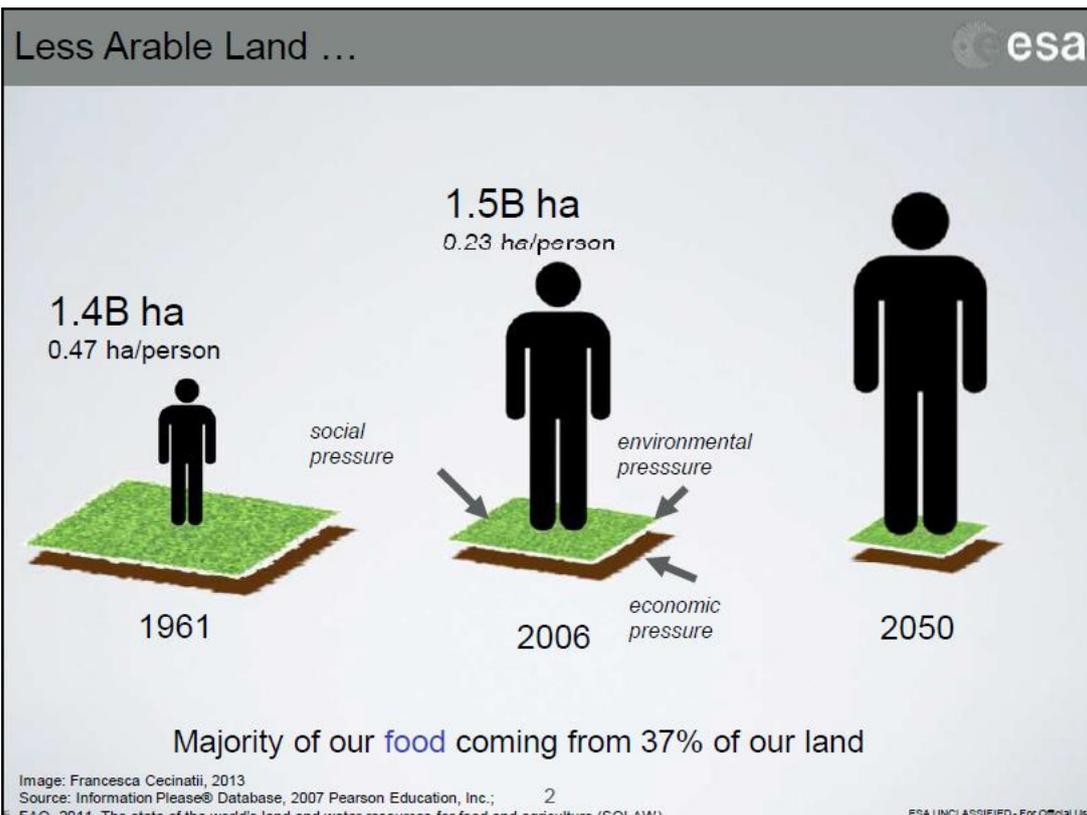
5 000 000 000 ha
surface agricole
utile



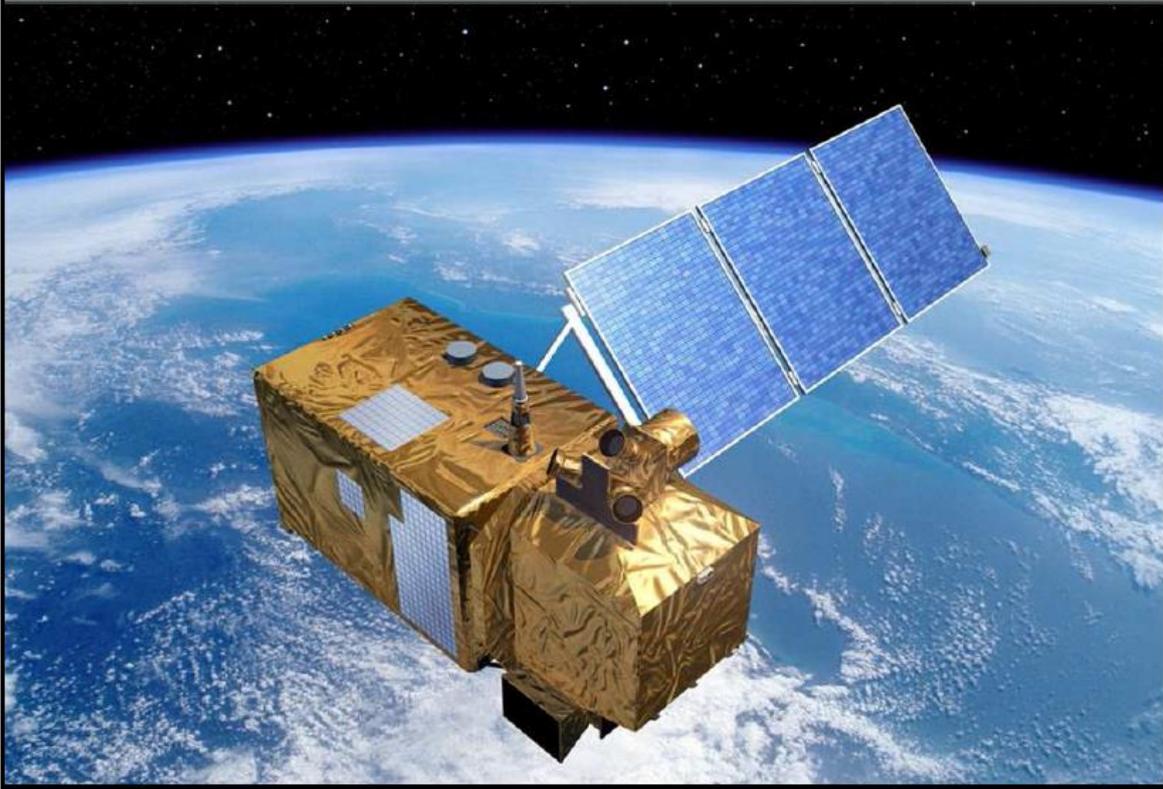
980 000 000
porcins⁴

Laure-Marie Neuburger - 2 Juillet 2015

**ANNEXE 4 :
INTERVENTION DE M. PIERRE-PHILIPPE MATHIEU,
RESPONSABLE DU PROGRAMME E04FOOD, ESA**



Sentinel-2 European Optical Observatory



S-2A first images – Pavia / Po Valley, Italy



S2-A launched 23 June 2015,

- 13 Bands (VIR, NIR, SWIR, red edge),
- 290km swath,
- 10-20-60m resolution,
- 10d revisit.

Image of Pavia acquired on 27 June 2015 at 10:25 UTC,

Copyright: Copernicus data (2015) / ESA

Sentinel-1 European Radar Observatory Mission



Monitoring Rice with S-1



Sentinel-1 time series (Oct.2014-Jan.2015)
GEOGLAM Asia-RICE Site: An Giang (Mekong River Delta, Vietnam)



Towards small data, but big insight

esa

Petabytes **SMS**

Geophysical parameters: **Indicators :** **Information Service :**

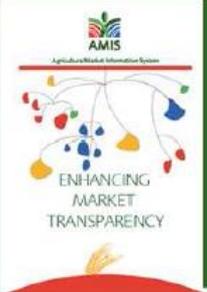
From Data ... **... to Information ...** **... to Decision-support**



G20 Final Declaration

44. We commit to improve market information and transparency in order to make international markets for agricultural commodities more effective. To that end, we launched:

- The "Agricultural Market Information System" (AMIS) in Rome on September 15, 2011, to improve information on markets ...
- The "Global Agricultural Geo-monitoring Initiative" (GEO-GLAM) in Geneva on September 22-23, 2011. This initiative will coordinate satellite monitoring observations systems in different regions of the world in order to enhance crop production projections and weather forecasting data.



ENHANCING MARKET TRANSPARENCY

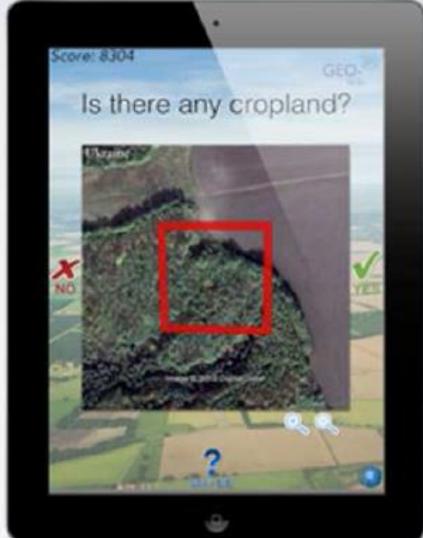
7

ESA UNCLASSIFIED - For Official Use

Driving R&D - New validation techniques

esa

crowdsourcing ground truth and classification (EO Science 2.0)



Score: 8304

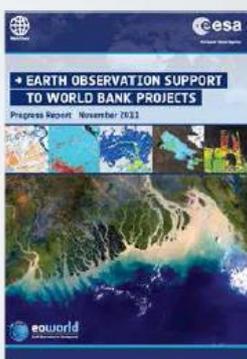
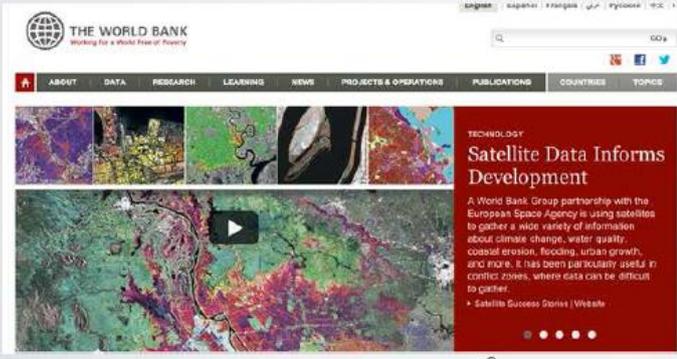
Is there any cropland?

NO YES

ESA UNCLASSIFIED - For Official Use

Partnership with the Public Sector

esa

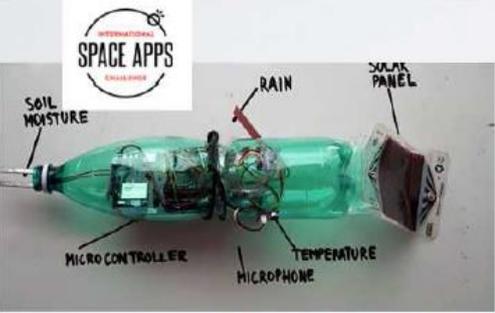


9

ESA UNCLASSIFIED - For Official Use

Driving Innovation & Web Entrepreneurs

esa



Opportunity to develop a **Innovation Pulse Lab @ ESRIN** bringing together academia, private sector and regional actors to address Big Data issues for society and develop new applications.

10

ESA UNCLASSIFIED - For Official Use

MOOC Climate from Space

<https://www.futurelearn.com/courses/climate-from-space>



Monitoring Climate Change from Space

Explore our planet from Space and learn how we can monitor climate change through Earth observation techniques.

Go to course

- FREE online course
- Duration: 5 weeks
- 3 hours per week

ABOUT THE COURSE

We are now at a time on planet Earth where significant and rapid changes to the climate are being witnessed. It is becoming increasingly essential for us to study the climate and observe changes all across the planet at the highest level of detail possible, but how can we achieve such a comprehensive

EDUCATORS

 Neil Kaye

Q2

What is passive remote sensing?

- Passive sensing is used during the day when the sun is illuminating the earth.
- Passive sensing detects energy reflected or emitted radiation from the earth.
- Passive sensing detects infrared radiation.
- Passive sensing detects energy which is scattered by the atmosphere.

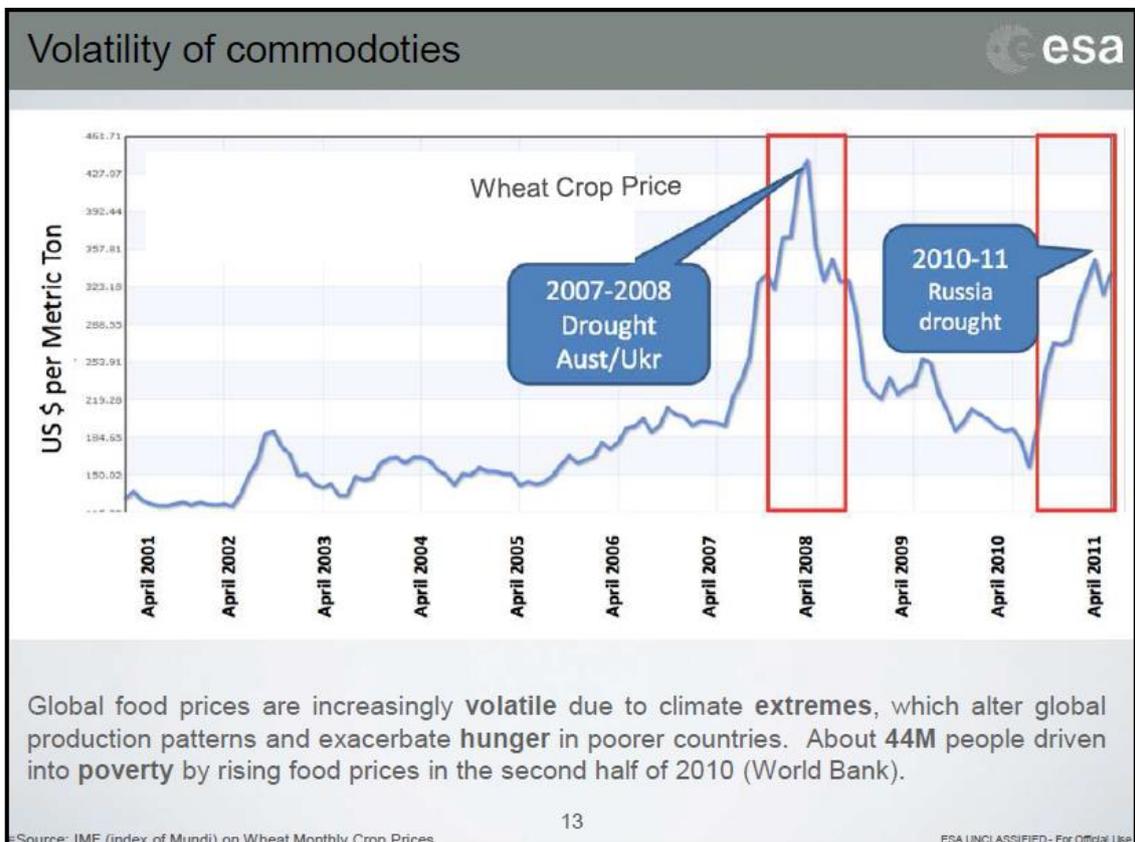
Future Learn Courses About Partners [Sign in](#) [Register](#)

Connecting people and ideas

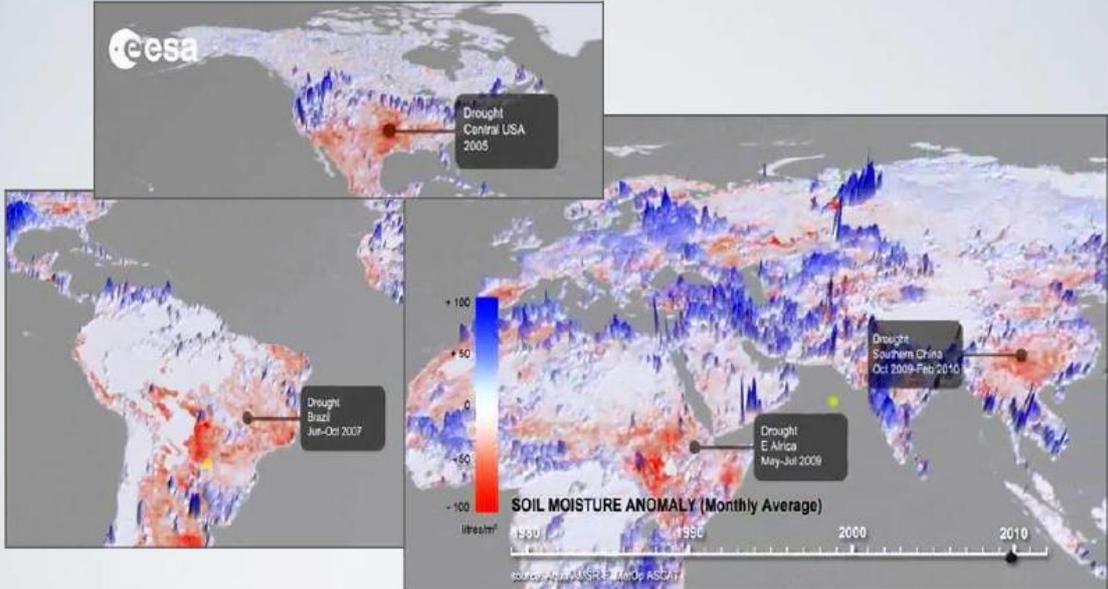
The best free online courses in the world.

[Browse courses](#)

ESA UNCLASSIFIED - For Official Use



Soil Moisture & water-induced food insecurity



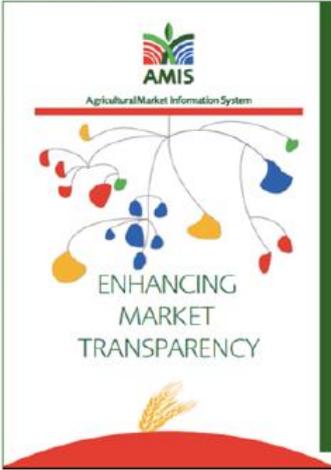
Improved Soil Moisture monitoring from space enables identification of **dry** conditions to support appropriate actions to avoid food insecurity.

Source: Maxime Trompier (ESA)

14

ESA UNCLASSIFIED - For Official Use

Global Framework for Monitoring Agriculture



G20 Final Declaration

44. We commit to improve market information and transparency in order to make international markets for agricultural commodities more effective. To that end, we launched:

- The "Agricultural Market Information System" (AMIS) in Rome on September 15, 2011, to improve information on markets ...;
- The "Global Agricultural Geo-monitoring Initiative" (GEO-GLAM) in Geneva on September 22-23, 2011. This initiative will coordinate satellite monitoring observation systems in different regions of the world in order to enhance crop production projections and weather forecasting data.

Global Framework for monitoring of agriculture and limit volatility. EO from space has a key role to play. ESA is developing EO products to quantify food insecurity risks.

Source: G-20, GEO-GLAM

15

ESA UNCLASSIFIED - For Official Use