

TÊTE D'AFFICHE / **Philippe Tassi**



VARIANCES

LA REVUE DES ENSAE ALUMNI

ACTES DU COLLOQUE /

INDIVIDU, DONNÉES & SOCIÉTÉ CONNECTÉE

Opportunités, Risques
et Confiance





BNP Paribas Cardif Nos métiers

Spécialiste mondial sur le marché de l'assurance de personnes, BNP Paribas assure les personnes, leurs familles et leurs biens.

Présent dans 37 pays, BNP Paribas Cardif compte près de 10 000 collaborateurs.

Actuariat

Garant de la stratégie rentabilité/risques, l'actuaire élabore et tarifie les contrats, suit les résultats, procède aux analyses techniques en veillant à concilier les intérêts des assurés, de la compagnie d'assurance et de ses partenaires.

L'actuaire s'appuie sur des outils et concepts qu'il développe et fait évoluer : profit testing, MCEV, capital économique SII...

Il est au cœur de l'innovation et des sujets d'actualité : euro-croissance, dépendance, digital, analytics...

ALM

Au sein de la direction des risques, l'actuaire ALM (gestion actif/passif) mène les études d'optimisation de l'allocation stratégique d'actifs en France et à l'international.

Il est force de proposition sur les dossiers stratégiques liés au pilotage financier de l'entreprise.

Il intervient activement dans les processus MCEV et Solvabilité II via la modélisation et le contrôle du risque de marché.

Risk Management (IRM)

La Direction Gestion du Risque est au cœur de la modélisation des outils dédiés au suivi des risques et de la solvabilité, notamment en travaillant à l'élaboration, la maintenance et l'évolution du modèle interne.

Elle est en charge de la production des indicateurs de risque et de solvabilité, des analyses et des reportings à destination interne (groupe) et externe (régulateurs, actionnaires).

Elle est à l'origine des règles d'encadrement des risques (appétence aux risques) et coordonne l'élaboration des normes, des gouvernances, et le lobbying concernant la gestion de la solvabilité et plus généralement du risque.

FONCTIONS

- Actuaire épargne, prévoyance, ALM, international ou France
- Actuaire modélisateur
- Volontariat International en Entreprise : VIE



**BNP PARIBAS
CARDIF**

| L'assurance d'un monde qui change

Nos postes sont basés à Nanterre et à l'étranger,

BNP Paribas Cardif - Direction des Ressources Humaines > 8, rue du Port, 92728 Nanterre CEDEX

Pour consulter nos opportunités, connectez-vous sur notre site internet > www.assurance.bnpparibas.com



Les yeux ouverts



Jamais dans l'histoire humaine nous n'avons été soumis à l'introduction aussi rapide de nouvelles technologies, source de bouleversements profonds d'ordre social, comportemental et psychologique. Dès aujourd'hui, images et informations se multiplient de manière exponentielle, données souvent précieuses pour les statisticiens et analystes, mais susceptibles de croisements et d'interprétations hasardeuses. De fait, l'individu, qui tend à s'isoler dans sa musique ou ses réseaux d'« amis », est de plus en plus schizophrène, pris en étau entre cette boulimie de contacts et de flux instantanés, et une quête de sens, de plus en plus ardue. Demain, des robots s'occuperont de personnes âgées ou dépendantes, en leur témoignant de l'empathie, symboles d'une frontière qui s'estompe entre l'homme et la machine.

Face à cette véritable révolution, nous ne pouvons avancer les yeux fermés, pour reprendre l'expression utilisée par l'un des intervenants du colloque « Individu, données et société connectée : opportunités, risques et confiance » organisé en juin dernier par ENSAE Alumni, et auquel nous consacrons l'essentiel de ce numéro de Variances. Si notre Ecole a créé une filière Data Science qui rencontre un grand succès auprès des étudiants, elle ne saurait en effet se positionner sous le seul angle scientifique. Ce colloque a donc réuni des spécialistes aussi divers qu'un historien des données, un sociologue, un psychiatre, un régulateur, ... pour éclairer les enjeux humains des innovations à l'œuvre. C'est bien le reflet du positionnement de l'ENSAE, au croisement des techniques quantitatives et des sciences sociales. Et c'est la réaffirmation du rôle d'ENSAE Alumni pour, aux côtés de l'Ecole, être un vecteur d'image et un aiguillon pour investir de nouveaux champs d'investigation.

Et qui mieux que Philippe Tassi pour être la Tête d'affiche de ce numéro dédié aux actes du colloque ? Beaucoup d'entre nous l'ont côtoyé à l'Ecole, comme condisciple, enseignant, directeur... et si sa carrière est celle d'un éminent statisticien, le portrait que nous dressons de lui montre aussi un homme de mémoire, très imprégné d'humanité, conscient du rôle du hasard dans nos parcours, et de la nécessité de prendre du recul et de s'appuyer sur de multiples grilles de lecture pour analyser ces méga-données qui nous submergent.

Garder les yeux ouverts donc, et cultiver un certain éclectisme pour pouvoir avancer, en confiance, dans ce monde complexe, tels sont quelques messages forts qui ressortent de la lecture de ce numéro. Merci de votre fidélité !

Eric Tazé-Bernard (1978)
Rédacteur en chef

VARIANCES
LA REVUE DES ENSAE ALUMNI

ISSN 1266-4499-Variances (Malakoff)

Numéro 53 - Octobre 2015

Directrice de publication :
Catherine Grandcoing (1978)

Rédacteurs en chef :
Catherine Grandcoing (1978)
Eric Tazé-Bernard (1978)

Comité de rédaction :
Nicolas Braun (2003)
Pierre-Marie Debreuille (1999)
Julien Guitard (2005)
Christophe Lagarde
Emmanuel Léonard (2007)
Arthur Renaud (2014)
Philippe Tassi (1972)

Conception et réalisation :
Sébastien Marsac

Publicité :
FERCOM
Jean-Michel Amram
01 46 28 18 66
jm.amram@fercom.fr

Votre avis nous intéresse :
variances@ensae.org
ou par lettre à :
Variances
Secrétariat des Anciens de l'ENSAE,
bureau E 25 ter, timbre J 120,
3, avenue Pierre Larousse,
92 245 Malakoff cedex

Imprimerie Chirat
744, route de Sainte-Colombe
42540 Saint-Just la-Pendue
Dépôt légal : 7852

Prochain numéro :
Février 2016

Prochain dossier :
Les défis du changement
climatique

BIENVENUE
DANS LE MONDE
DES

ENSAE Alumni



- ▶ Les Afterwork
- ▶ ENSAE au Féminin
- ▶ Des événements premium*
- ▶ Le club des Administrateurs
- ▶ Les Soirées Prospective Métiers
- ▶ ENSAE Business Angel et Entrepreneurs
- ▶ Les Happy Hour Métiers
- ▶ Service Ecoute Carrières*

- ▶ Annuaire*
- ▶ Variances
- ▶ L'appli ensae.org*
- ▶ Le réseau ParisTech Alumni
- ▶ Les petits-déjeuners Economie Finance
- ▶ 3000 offres d'emploi par mois*
- ▶ Managers.com, Site emploi*
- ▶ Les Rencontres Data Science
- ▶ Le site ensae.org





Variances n°53

Octobre 2015

L'association

Le mot de la Présidente

Le nouveau conseil d'ENSAE Alumni P.6

Tête d'affiche

Philippe Tassi P.8

Actes du colloque

**INDIVIDU, DONNÉES ET SOCIÉTÉ CONNECTÉE :
OPPORTUNITÉS, RISQUES ET CONFIANCE** P.12

Ouverture P.13

Introduction / La donnée : des tablettes sumériennes aux big data P.16

Temps 1 / Opportunités et risques dans une société connectée P.22

1 - Les avantages sociétaux d'un monde connecté P.22

2 - Les avantages pour la connaissance économique et statistique P.26

3 - Un monde numérique : le retournement du miroir P.31

4 - Quels risques pour l'individu connecté ? P.37

Temps 2 / Établir la confiance dans une société connectée P.43

1 - Le cadre législatif de la confiance P.43

2 - L'apport de la technologie P.47

3 - Éthique et déontologie : l'exemple d'ESOMAR P.47

Conclusion P.57

Les Alumni

Mobilités p.60

Dans les rayons p.62

Le mot de la Présidente

COMPOSITION DU NOUVEAU CONSEIL



Guillaume BENETEAU
(2004)



Xavier DUPRE
(1999)



Eric MICHEL
(1986)



Florence BONNEVAY
(1987)



Catherine GRANDCOING
(1978)



Didier NEGIAR
(1978)



Yoram BOSC-HADDAD
(1986)



Julien GUITARD
(2005)



Jincheng NI
(1990)



Nicolas BRAUN
(2003)



Arnaud LAROCHE
(1994)



Arthur RENAUD
(2014)



Bénédicte CALLET
(2014)



Pierre JOLY
(1978)



Philippe TASSI
(1972)



Julia CUVILLIEZ
(2013)



Emmanuel LEONARD
(2007)



Eric TAZE-BERNARD
(1978)



Olivier de CONIHOUT
(1981)



Kim-Bich LONG
(2012)



Pierre-Marie DEBREUILLE
(1999)



Pierre MATHOULIN
(1985)

COMPOSITION DU NOUVEAU BUREAU

Présidente :

Catherine GRANDCOING (1978)

Vice-président « Prospective et développement » :

Philippe TASSI (1972)

Vice-président « Relations avec les administrateurs » :

Pierre JOLY (1978)

Vice-président

« Carrières, trajectoires, métiers » :

Nicolas BRAUN (2003)

Trésorier :

Yoram BOSC-HADDAD (1986)

Trésorier adjoint :

Pierre MATHOULIN (1985)

Secrétaire général :

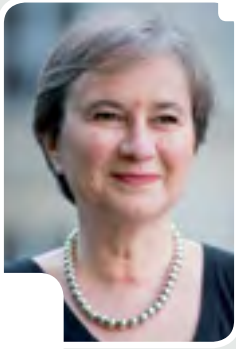
Julien GUITARD (2005)

Secrétaire général adjoint :

Arthur RENAUD (2014)

Webmaster :

Emmanuel LEONARD (2007)



Porter ensemble notre regard loin devant

Animer notre communauté d'Alumni et faire rayonner les valeurs de la marque ENSAE : ces deux objectifs guident et structurent les actions de notre Association pour qu'elle soit une communauté au sein de laquelle chacun(e) d'entre nous puisse trouver des échanges féconds, un partage de valeurs et de réalisations, un soutien et surtout l'envie de **porter ensemble notre regard loin devant**.

Ces convictions et **cette envie** motivent le nouveau Conseil de l'Association récemment élu. Celui-ci se veut une représentation la plus fidèle et la plus large de notre communauté d'Alumni. Il rassemble des jeunes et des moins jeunes, des femmes (encore trop peu) et des hommes, des Alumni vivant en France et à l'étranger, des administrateurs, des financiers, des data scientists, des créateurs d'entreprises, mais aussi des spécialistes du conseil, des ressources humaines, de l'industrie, des nouvelles technologies.

Ce nouveau conseil a souhaité, cette année encore, me confier la Présidence de l'Association. Je l'en remercie et vais m'attacher tout au long de ce mandat, avec l'aide du bureau et de tous les membres du conseil à porter notre regard loin devant.

Pour cela, toutes nos actions sont résolument tournées vers le futur en puisant leurs racines dans la force de notre passé. Cette posture est la nôtre lorsque :

- nous organisons le colloque « Individu, données et société connectée : opportunités, risques et confiance » dont ce numéro de *Variances* édite les actes. C'est parce que notre formation nous a appris à croiser les voies des sciences exactes avec celles des sciences humaines et sociales que notre légitimité est grande à imaginer les richesses de ce monde en devenir et les conditions pour les épanouir. Cet événement, qui s'est déroulé en juin, a permis de faire rayonner de manière concrète les valeurs de la marque ENSAE que les Alumni partagent et défendent ;

- nous développons des services et des événements tout au long de l'année pour réunir des Alumni, leur permettre de se connaître, d'échanger, de fertiliser leur carrière. Ce sont les petits déjeuners économie-finance, les after work, les rencontres data sciences, les soirées « Prospective métiers », les réunions « Business angels et entrepreneurs », les conférences du club des administrateurs... mais aussi notre service écoute carrières, notre participation active au *job board* *Managers.com* (WATs4U), au réseau ParisTech, à Sciences ParisTech au féminin... ;
- nous élargissons les canaux de communication de *Variances* aux réseaux sociaux sur lesquels nous développons la présence de l'Association pour porter loin les prises de paroles et les actions des Alumni.

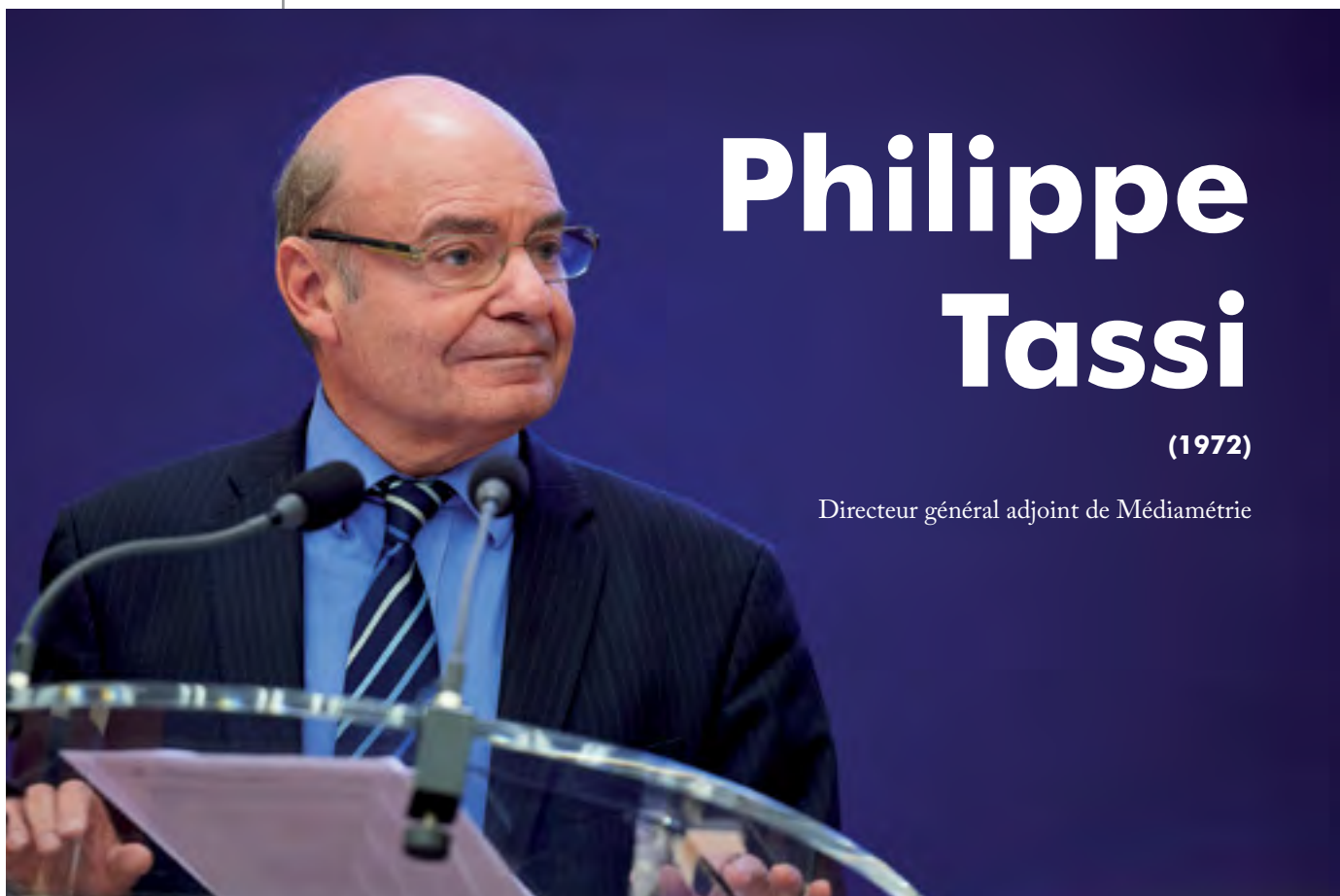
Cette posture est la nôtre lorsque nous avançons dans le monde en devenir de l'ENSAE au sein du campus Paris Saclay : aux côtés de l'Ecole, nous gardons solidement ancrées nos actions dans les valeurs qui font notre force et notre différence tout en anticipant les richesses des cohabitations et des partages offerts par ce rassemblement géographique.

Pour ce faire, deux nécessités : savoir accueillir et rebondir positivement sur toutes les opportunités qui se présentent sans jamais craindre le changement, et – prioritairement – inventer le nouveau *business model* de l'Association.

En effet, cette réflexion économique en cours est la condition qui nous permettra d'arriver à Saclay confiants dans notre solidité financière, et forts de notre identité singulière. Les Alumni par leurs cotisations et leurs dons, les entreprises à travers les partenariats, l'Ecole par son soutien, tous peuvent aider à relever ces challenges et à construire le futur de notre communauté d'Alumni.

Loin devant.

Catherine Grandcoing (1978)
Présidente d'ENSAE Alumni



Philippe Tassi

(1972)

Directeur général adjoint de Médiamétrie

Variances - *Philippe, pour commencer à décrire ton parcours, peux-tu nous dire comment tu es arrivé à l'ENSAE ?*

Philippe Tassi - Le contexte commence à être bien lointain. Été 1969, je viens de réussir le concours maths de l'ENSAE en même temps que d'autres écoles comme Centrale, Supélec, Télécom Paris, Mines de Paris.

A l'époque, Centrale était un peu en retrait, et à mon petit niveau, nul n'imaginait que le secteur des télécommunications, trente ans plus tard, serait aussi riche en création, innovation et potentiel, son présent de l'époque étant le ministère des PTT. A tort ou à raison, Supélec me paraissait trop technique. Je suis donc allé demander conseil à Edmond Ramis, qui avait été mon professeur de mathématiques spéciales au lycée Louis-le-Grand. Avec son accent languedocien et son sourire malicieux, il m'a parlé de ces domaines relativement nouveaux qu'étaient la théorie des probabilités et la mathématique

statistique. Il faut se rappeler que, par exemple, les probabilités n'ont fait partie du programme de l'agrégation de mathématiques qu'à partir de 1976, et encore comme simple matière optionnelle.

Courant juillet, je me suis donc rendu dans les locaux de l'ENSAE, un hôtel particulier de la rue de Montmorency, dans le troisième arrondissement ; je me souviens des deux personnes qui ont bien voulu me recevoir et consacrer un peu de leur temps pour expliquer leurs domaines de travail à l'ignare que j'étais. Il s'agissait de Jean-Louis Philoche et Alain Monfort. Par les mots qu'ils ont su trouver, par leur personnalité aussi, ils m'ont convaincu. Avec le recul du temps, quand je les ai eus comme professeurs, puis ensuite quand j'ai appris, en travaillant avec eux, à les connaître et à découvrir leur fabuleuse contribution à la science statistique et économétrique, je crois que, ce jour-là, j'ai eu énormément de chance de les croiser et les rencontrer.

V. - La chance, le hasard, ont joué un rôle dans ta vie professionnelle, n’est-ce pas ?

Ph.T. - Je mentirais en répondant « non ». Ces deux mots sont synonymes de probabilité, si on ne prend pas le hasard au sens le plus commun, c’est-à-dire au sens de la loi uniforme qui traduit le manque total d’information.

Oui, j’ai eu de la chance plusieurs fois, dans mes décisions, mes choix, mes rencontres, sans parler de ma vie personnelle où je pourrais tenir le même discours. Décider de lâcher des études littéraires pour aller vers les maths – c’était peut-être moins passionnant que le latin, le grec et la philo, mais tellement plus facile –, quitter le domicile des parents à 15 ans et demi, « monter » à Paris en petit provincial. Choisir d’entrer à l’ENSAE au détriment d’autres écoles plus anciennes et plus renommées, d’aller au Cameroun, de revenir plus tard à l’ENSAE plutôt que d’aller travailler dans la grande banque américaine qui m’avait recruté. Et surtout rencontrer des hommes et femmes qui m’ont tout appris, par leurs connaissances et leur charisme. J’ai parlé d’Alain Monfort et Jean-Louis Philoche, la liste serait longue ; j’ajoute Paul Champsaur, Jean-Claude Milleron, Edmond Malinvaud, bien sûr, Christian Gourieroux, Jean-Claude Deville, Alain Trognon, Jean-Louis Bodin, président de l’ISI, Yves Franchet, Elizabeth Dognin, Jean-Jacques Drosesbeke, Bernard Fichet, Gilbert Saporta, Ousmane Balde, Jacqueline Aglietta, créatrice et présidente de Médiamétrie jusqu’en 2006, son successeur Bruno Chetaille, et tant d’autres...

Impossible de tous les citer, qu’ils ne m’en veuillent pas. Et aussi une mention particulière à tous les étudiants de l’ENSAE et l’ENSAI que j’ai eu le bonheur d’avoir. Quoi de plus beau que de voir, dans un amphi, deux yeux briller, et même si c’est seulement deux ?

Les diverses personnalités que j’ai croisées professionnellement m’ont façonné. J’aime les idées d’échange, de partage, d’apprentissage permanent, de transmission, d’écoute. En Afrique, on a coutume de dire « quand un vieillard meurt, c’est une bibliothèque qui brûle ». J’aime cette phrase d’Amadou Hampâté Bâ, et ce culte de la transmission des anciens qu’elle révèle. Cela donne de la valeur au fait d’écouter et apprendre, s’instruire, comprendre et admettre que si l’on n’est pas d’accord avec quelqu’un, ce n’est pas pour autant qu’on a raison.

V - Tu es toujours resté proche de l’ENSAE, comme membre de l’équipe dirigeante, professeur ou membre des instances. Pourquoi ?

Ph.T. - La proximité avec l’ENSAE est évidente dans ma trajectoire : élève, enseignant, membre de l’équipe dirigeante, plus tard président des Alumni ou membre de certains Conseils de gouvernance. Suivre aussi les évolutions des anciens étudiants permet de tisser un lien d’une autre nature.

En outre, les progrès de la science statistique ont conforté cette proximité, car j’ai eu la chance d’avoir envie de les apprendre et de les comprendre, et le désir fort de les transmettre aux générations qui se sont succédé. Par exemple, la théorie de la robustesse, les méthodes d’apprentissage, le *machine learning*, les mathématiques liées aux très grandes dimensions comme celles du big data, etc, sont autant d’innovations plus ou moins récentes. Les connaître, les digérer, les transmettre si possible de façon simple, et surtout les utiliser dans le cadre de la vie professionnelle, tout ceci crée une synthèse épanouissante et équilibrée.

Quand c’est possible, il m’a toujours paru normal de « renvoyer l’ascenseur » vers l’endroit où j’ai reçu l’essentiel de ma formation initiale. J’ai longtemps joué au rugby : l’une des valeurs de ce sport, c’est « recevoir et donner ». Je trouve que c’est un bon résumé de la vie.

V - Poursuivons ton parcours. Comment t’es-tu retrouvé DGA de Médiamétrie ?

Ph.T. - D’abord par deux années au Cameroun, à l’Institut de Formation de Statistique, son nom de l’époque, avant de s’appeler plus tard l’ISSEA ; puis enseignant à l’ENSAE de 1974 à 1978 ; trois ans à l’Insee pour diriger les Statistiques Structurelles d’Entreprise, et apprendre le management de projets et de grosses équipes ; en 1981, retour à l’ENSAE comme professeur-chercheur et pour préparer, selon les consignes d’Edmond Malinvaud, l’autonomie de la division CGSA de l’ENSAE, la future ENSAI. Dossier passionnant que d’avoir l’opportunité de préparer la naissance d’une future grande école.

En 1989, une idée folle : serais-je capable de faire aussi bien que « mes » étudiants ? J’ai donc quitté le cocon ENSAE en moins de temps que pour l’écrire. J’ai eu quelques propositions intéressantes, et Médiamétrie l’a emporté. Jacqueline Aglietta en était la PDG, et aussi présidente

de l’Association des anciens élèves ; j’en étais le secrétaire général. Jacqueline m’a proposé la direction scientifique et technique de Médiamétrie, j’ai accepté spontanément car « cette fonction sera ce que tu en feras ». Elle n’aurait jamais pu trouver des mots plus convaincants. Quatre ans plus tard, elle me nommait DGA. Et voilà...

Mais je n’oublie pas mes deux premières années au Cameroun, à Yaoundé, probablement les deux années de ma vie où je me suis senti le plus utile socialement, et où j’ai beaucoup appris à la fréquentation de personnes qui sont devenues des amis pour la vie, comme Jean-Marie Gankou, le « tonton » camerounais de mon premier fils. Je crois être revenu très différent de ce séjour en Afrique, beaucoup plus ouvert, tellement moins rempli de certitudes, après avoir été confronté à des modes d’être et de pensée non européens, à la rationalité différente, mais extrêmement riches. Le continent africain est et sera au fond de mon cœur à jamais.

« L’essentiel est de vivre avec passion tout ce que l’on fait »

V - Finalement, tu as peu changé d’employeur ?

Ph.T. - Oui, peut-être que je dois être plutôt fidèle par nature, je ne sais pas. J’ai occupé plusieurs postes, mais eu seulement trois employeurs : le ministère de la Coopération, le ministère de l’Economie, et Médiamétrie. Je ne fais pas partie de la Génération Y, c’est évident, je ne cherche pas à partir ailleurs quand je crois « avoir fait le tour », tout change si souvent. Néanmoins, trois employeurs par rapport à la génération qui a précédé la mienne, ça frôle l’instabilité.

La raison principale des longs séjours successifs est d’avoir eu la chance – encore ce mot ! – de m’épanouir professionnellement dans un domaine et deux organismes en expansion : tout d’abord le domaine, c’est la statistique, science statistique qui n’a fait que progresser, créant même son propre marketing avec le data mining des années 80 et la data science actuelle, liée au big data. Ensuite, il y a l’ENSAE, au sens large,

et le fait d’avoir fait partie de l’équipe dirigeante sans sacrifier les si nobles rôles de professeur et de chercheur, avec comme point d’orgue avoir participé à la naissance de l’ENSAI. Et enfin Médiamétrie, société en croissance incomparable, dans un secteur des médias tiré vers le haut par de nouvelles chaînes de TV, l’arrivée d’internet, l’apparition de nouveaux écrans – tablettes, smartphones – et modes de consommation qui en bouleversent la micro-économie.

V - Peux-tu nous préciser les activités et le rôle de cette société ?

Ph.T. - Médiamétrie a été créée par Jacqueline Aglietta en 1985, pour mesurer l’audience des médias TV et radio, dans un contexte où ces marchés devenaient à dominante privée. Le rôle de Médiamétrie est de « battre la monnaie » du marché des médias électroniques, que ce soit pour la dimension publicitaire mais aussi, il ne faut surtout pas l’oublier, pour les contenus et les programmes.

Le secteur des médias a connu une expansion extraordinaire : nouvelles chaînes, apparition de la TNT, nouveaux acteurs, nouveaux écrans, arrivée d’internet, passage au numérique, convergence digitale, ... Médiamétrie a eu le mérite de « courir aussi vite que lui » en organisant le consensus du marché et en suivant l’évolution et la complexification de ce secteur, qui requiert des outils statistiques, scientifiques, techniques et technologiques de plus en plus sophistiqués.

Cela n’apparaît pas toujours aux yeux du grand public, mais la partie cachée de l’iceberg est un merveilleux champ de recherche, de brevets et d’innovation opérationnelle.

La multiplicité des signaux et des données, de toute forme, de toute nature, le mélange de données fines, individuelles, d’échantillons ou de panels, et de mégadonnées de niveau plus agrégé issues de boxes ou de téléphones, la recherche permanente de leurs meilleurs usage et intelligence, tout cela fait des médias et télécom un secteur porteur d’opportunités pour les ENSAE, en plus, évidemment, des besoins bien connus en management, ou en pilotage économique et financier.

V - La statistique est le fil conducteur de ta carrière. Qu’est-ce qui t’a plu dans cette discipline ?

Ph.T. - Ce que j’ai toujours trouvé extraordinaire dans les divers domaines de la science

statistique, c’est leur dimension mathématique, bien sûr, et la complexité croissante de cette dimension. Mais c’est aussi l’objet de cette science, ses sujets d’études : entreprises, ménages, zones géographiques, individus, etc...

Donc si on veut bien s’en servir, il faut une approche certes verticale – en savoir de plus en plus sur le plan de ses méthodes – mais surtout horizontale.

Et donc apprendre l’économie, la sociologie, la psychologie, la sémantique, la géographie, la physique, la démographie, le droit, les ressources humaines. C’est bien si cette connaissance de leur état de l’art est actuelle, mais c’est encore mieux si on maîtrise l’histoire de leur émergence. Comment comprendre vraiment la statistique sans connaître la vie de Galton, ou de Fisher ? Comment faire de la bonne mesure d’audience en ignorant l’histoire de Maxwell et des ondes hertziennes ?

V - Tu sembles encore publier beaucoup. Peux-tu nous en dire plus ?

Ph.T. - C’est cohérent avec « recevoir et donner ». D’abord, le domaine statistique a favorisé la dimension recherche, théorique ou appliquée ; les publications, les articles, en sont un débouché naturel. Cela m’a également permis de m’impliquer peu à peu dans des associations valorisant la science statistique, comme, historiquement l’ASU (Association pour la Statistique et ses Utilisations), puis la SFDS (Société Française de Statistique), ou d’être élu à l’*International Statistical Institute*. Leurs colloques et congrès facilitent les échanges avec des pairs, de toute nationalité, ce qui est un enrichissement supplémentaire.

Plus en aval, le besoin de transmettre s’est traduit par l’écriture de livres soit à vocation pédagogique, soit bien plus « pointus », livres écrits seuls ou en collaboration.

Et dans le contexte en mouvement permanent de Médiamétrie et des médias, la nécessité d’avancer en recherche opérationnelle pour répondre à des questions mal posées, mal résolues ou pas encore posées a permis de poursuivre cette dynamique. Communiquer, publier, c’est aussi partager, verbe que j’aime bien.

V - Pour aborder des questions plus personnelles, quels sont tes centres d’intérêt et tes facteurs d’équilibre ?

Ph.T. - C’est une question qui relève de l’intime, et sa réponse aussi. En levant un petit coin du voile, il me semble être parvenu à assez bien séparer la sphère professionnelle et la sphère privée, personnelle, familiale. Je crois, j’espère, parler très rarement de travail en privé. Il y a bien d’autres choses à apprendre, à partager et à vivre, surtout dans une ville comme Paris, que ressasser les ennuis du quotidien.

J’ai l’impression d’être curieux de nature, et d’avoir toujours envie d’apprendre. Mon équilibre, si tant est qu’il existe, est donc réalisé au travers de plusieurs centres d’intérêt : l’histoire, les expositions, la musique, très diverse, de l’opéra au balafon ou à la kora, la chanson, française ou non mais à texte, la poésie, la littérature, y compris les BD, certains genres cinématographiques comme le cinéma italien des années 60-70, et le sport – rugby, natation, plongée sous-marine, athlétisme –.

Il m’arrive même parfois de passer de l’autre côté du miroir, de ne plus être seulement spectateur, lecteur, ou auditeur, mais de pratiquer. Jouer de la guitare, écrire des nouvelles, composer des chansons avec paroles et musique.

L’essentiel est de vivre avec passion tout ce qu’on fait, c’est un si joli mot.

V - Comment penses-tu occuper les années qui viennent ?

Ph.T. - Il est bien évident que ma carrière professionnelle est derrière moi. Mais on peut s’occuper de tant de façons : voir grandir les petits-enfants, profiter des possibilités culturelles de Paris, voyager, écrire. Au-delà de ces activités déjà bien riches, je crois en deux mots : transmission et bénévolat, qui se traduisent par des envies et des idées. Le premier conduit à un rapprochement – à nouveau – de l’enseignement supérieur, en essayant d’être force de proposition pour des cursus avancés en matière de marketing scientifique, autour des big data et du *machine learning*. Le second, bénévolat, se traduirait par une implication forte puisque tant de choses sont à faire pour les personnes mises à l’écart de et par notre société : gens de la rue, analphabètes, migrants... Bref, être utile socialement. ■

Propos recueillis
par Eric Tazé-Bernard

INDIVIDU, DONNÉES & SOCIÉTÉ CONNECTÉE

Opportunités, Risques et Confiance

JEUDI 18 JUIN au CNAM - Amphi Paul Painlevé 292, rue Saint-Martin - 75003 Paris

Ouverture

P 13

Catherine Grandcoing (1978), Présidente d'ENSAE Alumni, et **Julien Pouget (2001)**, Directeur de l'École Nationale de la Statistique et de l'Administration Économique (ENSAE ParisTech)

Introduction

P 16

LA DONNÉE : UNE HISTOIRE ANCIENNE ET TOUJOURS RENOUVELÉE

Jean-Jacques Droesbeke, Professeur émérite de l'Université Libre de Bruxelles

Temps 1 / Opportunités et risques dans une société connectée

P 22

P 22 / LES AVANTAGES SOCIÉTAUX D'UN MONDE CONNECTÉ

Lionel Janin (2003), Chargé de mission numérique, France Stratégie, Services du Premier ministre

P 31 / UN MONDE NUMÉRIQUE : LE RETOURNEMENT DU MIROIR

Saadi Lahlou (1983), Professeur de psychologie sociale, Département de Psychologie sociale, London School of Economics and Political Science

P 26 / LES AVANTAGES POUR LA CONNAISSANCE ÉCONOMIQUE ET STATISTIQUE

Stéphane Grégoir (1988), Directeur de la méthodologie et de la coordination statistique et internationale, Insee

P 37 / QUELS RISQUES POUR L'INDIVIDU CONNECTÉ ?

Serge Tisseron, Psychiatre, Docteur en psychologie habilité à diriger des recherches, Chercheur associé au Centre de Recherches Psychanalyse, Médecine et Société (Université Paris 7 Denis Diderot)

Temps 2 / Établir la confiance dans une société connectée

P 43

P 43 / LE CADRE LÉGISLATIF DE LA CONFIANCE

Sophie Nerbonne, Directrice de la Direction de la conformité de la Commission nationale de l'informatique et des libertés (CNIL)

P 47 / L'APPORT DE LA TECHNOLOGIE

Philippe Tassi (1972), Directeur général adjoint de Médiamétrie, Vice Président d'ENSAE Alumni

P 53 / ETHIQUE ET DÉONTOLOGIE : L'EXEMPLE D'ESOMAR

Elisabeth Cosnefroy, Représentante d'ESOMAR pour la France (the World Association for Social, Opinion and Market Research)

Conclusion

P 57

Antoine Frachot (1989), Directeur général du Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)

Ouverture

Catherine Grandcoing (1978)

Présidente d'ENSAE Alumni

Bonjour à toutes et à tous et bienvenue au colloque organisé par l'Association Ensaie Alumni dans ce vertigineux amphithéâtre du Cnam qui nous accueille, avec le soutien de nos partenaires que je voudrais remercier ici, EDF Energies nouvelles, l'Institut CDC pour la recherche, Médiamétrie et Ylios.

Pourquoi ce colloque et pourquoi ENSAE Alumni ?

Je répondrai en deux temps : pourquoi ENSAE Alumni ? Les Alumni de l'ENSAE ont reçu une formation singulière, celle de l'ENSAE, école trait d'union entre les sciences exactes et les sciences humaines et sociales. La première conséquence de cette singularité est la diversité des carrières des Anciens de l'ENSAE qui sont souvent des économistes, des mathématiciens, des financiers, des data scientists mais aussi des juristes, des sociologues, des philosophes, des historiens, des psychologues...

C'est pourquoi à l'heure où nombre de colloques et de séminaires se concentre sur les données, leur recueil et leurs traitements - territoire d'exercice privilégié des data scientists -, nous avons voulu poser notre regard plus loin, à l'endroit où sciences exactes et sciences humaines et sociales convergent.

Un rapport de l'Idate prévoit que d'ici cinq ans, le nombre d'objets connectés atteindra les 80 milliards, soit environ 30 objets connectés par foyer.

Cela nous transporte dès aujourd'hui dans un monde dans lequel il n'est plus nécessaire de «faire», d'agir pour générer des données ; « être »,

être soi, respirer, ressentir suffit à construire de nous une représentation numérique qui permet de « nous anticiper ». Nous sommes là au cœur de notre psyché en ce qu'elle recouvre notre conscient et notre inconscient.

C'est pourquoi nous vous proposons une réflexion sur le thème « individu, données et société connectée : opportunités, risques et confiance ». Tous les intervenants qui s'adresseront à nous sont des experts reconnus de disciplines variées ; tous dans leur pratique professionnelle et leurs travaux de recherche croisent les regards et les langages des sciences exactes et des sciences humaines et sociales pour décrypter des phénomènes à l'aide d'alphabets multiples (celui du data scientist ou du juriste, comme celui du psy-



chologue ou de l'historien) pour enrichir leurs propres réflexions et élargir ainsi leur vision prospective. C'est cette approche holistique et pluridisciplinaire de la mutation sociétale que nous vivons dans notre monde chaque jour plus connecté, que nous vous proposons ce matin.

Le monde dont nous parlons aujourd'hui, ce monde connecté, en mutation radicale, vit une transformation irréversible. L'histoire nous apprend que chaque évolution sociétale s'est accompagnée de peurs, peur individuelle du changement et, en conséquence, évaluation imprécise des opportunités. L'individu se trouve ainsi écartelé entre son aspiration vitale d'appartenance à une société qui bouge et le bouleversement profond qu'il ressent dans son sentiment d'exister.

Au coeur de la révolution numérique que nous vivons aujourd'hui, l'individu est « augmenté » par les données qu'il génère, qui forment une représentation de lui-même, dissociée et associée. Ces données font de lui un mutant bio-technologique.

Cet individu a le droit d'avoir peur, il a la liberté de résister, mais il a aussi la liberté de se projeter positivement dans ce mouvement créé et accéléré par les immenses capacités des technologies numériques.

Oui, liberté de se projeter positivement. Liberté ? ou mieux, responsabilité de s'informer, d'analyser, d'évaluer les risques et de comprendre les opportunités promises par cette mutation, pour mieux construire dessus.

Cela signifie lâcher sa réalité d'hier, les certitudes du connu, quitter la sécurité de notre identité biologique et psychologique. Ouvrir, accueillir cet autre nous-même, cette représentation numérique de ce que nous sommes ; accueillir les objets connectés, prolongement de nous-mêmes qu'à notre insu complice nous aurons participé à engendrer.

Ouvrir, accueillir c'est à cela que nous vous convions à l'occasion de ce colloque, et c'est là la réponse à ma deuxième question : pourquoi ce colloque ? Comment les individus que nous sommes, conscients et rationnels, sauront-ils le faire ? En établissant la confiance dans cette machine lancée à très grande vitesse qu'est aujourd'hui le monde de la communication à l'heure des technologies numériques ; cette machine qui, comme le fut le chemin de fer en son temps, nous projette dans un monde inconnu

dont les historiens et les sociologues de demain souligneront les richesses.

Etablir la confiance pour que cette société connectée fonctionne au service de l'individu et non à ses dépens. Par les règles édictées ? Probablement en partie. Mais aussi par ce que les technologies permettent. Car si le numérique ouvre le champ des possibles au-delà des limites de la connaissance actuelle, il ouvre notamment celui de la protection de l'individu dans son identité souveraine.

Vous l'aurez compris, ENSAE Alumni vous propose d'accompagner librement et sans résistance nos intervenants dans les territoires de prospective qu'ils vont nous permettre de mieux connaître, et ainsi de mieux comprendre pour mieux construire cet avenir connecté dans lequel nous vivons déjà.

Pour ce faire, nous avons partagé notre matinée en deux temps. Celui de l'expression des opportunités et des risques avec les interventions de Jean-Jacques Droesbeke, professeur émérite de l'université libre de Bruxelles, Lionel Janin, chargé de mission numérique à France Stratégie, Stéphane Grégoir, directeur de la méthodologie et de la coordination économique et statistique de l'Insee, Saadi Lahlou, professeur de psychologie sociale à la London School of Economics, et Serge Tisseron, psychiatre. Celui de la construction de la confiance avec les interventions de Sophie Nerbonne, directrice de la conformité à la CNIL, Philippe Tassi, directeur général adjoint de Médiamétrie et Elisabeth Cosnefroy, représentante France de l'Esomar. Enfin, Antoine Frachot, directeur du Genes, conclura ce colloque.

Avant de passer la parole à Julien Pouget, directeur de l'ENSAE, je voudrais remercier particulièrement Philippe Tassi, vice-président d'ENSAE Alumni qui a initié et coordonné les choix éditoriaux de ce colloque.

Enfin, pour conclure cette introduction, j'espère vivement que lors du cocktail qui nous réunira dans quelques heures, nous partagerons nos impressions d'avoir progressé dans notre connaissance et notre compréhension des enjeux et des opportunités que le monde connecté offre à notre vie si nous savons nous protéger de ses risques et en repérer les richesses.

Je vous souhaite une matinée aussi fertile que féconde. ■

Julien Pouget (2001)

Directeur de l'Ecole Nationale de la Statistique
et de l'Administration économique (ENSAE ParisTech)

Bonjour à toutes et à tous,

Pardonnez-moi de commencer par quelques mots un peu personnels. Quand j'étais élève à l'ENSAE, il y a une quinzaine d'années, j'avais décidé de faire de la microéconométrie. Je me suis effectivement engagé dans cette voie. Nous nous posons une question économique, recherchions la base de données et la méthodologie la plus adaptée pour y répondre, et obtenions ainsi parfois quelques résultats intéressants.

A cette occasion, on m'avait fait connaître cette phrase de James Heckman, économètre, prix Nobel qui a dit très simplement : « *Better data help a lot* ». Autrement dit, on a beau faire appel aux méthodes les plus sophistiquées et les plus récentes, rien ne viendra remplacer la base de données bien conçue pour répondre à la question que l'on se pose.

Cette évidence est certainement toujours d'actualité. Mais il me semble qu'elle prend une dimension singulière dans une société de plus en plus connectée, et que désormais, si l'on peut dire, « *more and more data help a lot, big data help a lot* ».

Le statisticien, devenu data scientist, est alors confronté à une situation un peu vertigineuse puisqu'au fond les données sont de plus en plus partout et portent sur tous les sujets. Bien sûr, le plus souvent, leur finalité initiale n'est pas la statistique, mais le statisticien sait être opportuniste et tirer parti de ces données.

Dans ce contexte les questions éthiques et déontologiques se posent avec une acuité particulière. La question de la confiance est certainement centrale. Par exemple, que penser de l'âge de plus en plus précoce auquel les enfants deviennent eux-aussi des « individus connectés » ? Dans un autre registre, qui concerne de plus près l'ENSAE, que penser de l'hyper-individualisation des risques permis désormais par les bases de données des compagnies d'assurance ? Comment concilier ce phénomène avec la mutualisation du risque, à l'origine même de ces compagnies ?

Dans un cadre, par ailleurs, où internet donne parfois l'impression d'être un univers où tout est permis, où les « cyberattaques » de plus en plus fréquentes sont désormais contrées par une



« cyberdéfense » qui se met en place, je crois que la réflexion déontologique puis la mise en place de règles législatives adaptées, contribuent tout simplement à la force de nos démocraties.

Plus prosaïquement, je me dois de rappeler que l'Ecole s'est fortement engagée dans la formation de jeunes gens qui maîtrisent tout à la fois les aspects techniques, scientifiques mais aussi stratégiques et éthiques des big data. La filière a ouvert en 2013. Les premiers diplômés sont sortis à l'été 2014. Un an après, les résultats de l'enquête sur l'insertion de ces jeunes sont tout juste disponibles. Un tiers de la promotion s'était engagée dans la filière data science. Un autre tiers dans la voie actuariat. Les résultats sont sans appel : six mois après la sortie de l'Ecole, les deux filières font jeu quasi égal, avec un taux d'emploi de 100% et un salaire moyen à l'embauche à peu près équivalent. Le signe que l'évolution de nos enseignements répond autant que faire se peut aux évolutions du monde économique.

Un grand merci, très chaleureusement, enfin, en mon nom et en votre nom à tous, à l'Association des Alumni, à Catherine Grandcoing et à Philippe Tassi, à tous les intervenants : ils nous donnent ce matin l'occasion de prendre du recul et de réfléchir, dans un monde connecté souvent dominé par une instantanéité qui ne favorise pas toujours la réflexion construite !

Je vous souhaite une très bonne matinée. ■

INTRODUCTION

La donnée : des tablettes sumériennes aux big data

Jean-Jacques Droesbeke, Professeur émérite de l'Université Libre de Bruxelles

L'objet de cet article est de présenter les phases principales de l'évolution du concept de donnée depuis l'Antiquité jusqu'à nos jours. Le participe passé féminin substantivé du verbe donner a d'abord eu le sens d' « aumône, distribution » (1200). Il s'est limité à quelques sens spécialisés en mathématiques (1755) et psychologie. On l'utilise aussi en informatique et statistique, traduit de l'anglais data, pluriel du supin, première forme, du verbe latin dare, « donner » (voir Rey et al., 1993). Dans cette brève présentation de son histoire, nous distinguerons quatre périodes distinctes.

Figure 1 : A L'ORIGINE



De Sumer au 16^e siècle

Les données produites pendant cette période concernent essentiellement deux opérations : le *dénombrement* et l'*observation de phénomènes astronomiques*.

Le dénombrement a toujours été une opération importante de l'activité humaine. Le recensement des populations en est son expression statistique la plus visible. Les premiers témoignages de mise en œuvre de cette méthode de collecte de données sont gravés sur des tablettes d'argile sumériennes et babyloniennes.

Dès le début de son utilisation¹, le recensement s'est avéré être un outil de gestion apprécié des puissants. Si les Mésopotamiens y ont recouru très tôt, on en trouve aussi trace dans l'Égypte ancienne, dès la fin du troisième millénaire avant notre ère. Ces peuples avaient bien saisi l'intérêt de recenser les populations pour savoir combien d'hommes pouvaient participer à la construction

des temples, palais, pyramides... ou encore d'utiliser cette technique à des fins fiscales.

Plus une population est nombreuse, plus le recensement s'avère utile. C'est ce qu'avaient compris aussi les empereurs chinois. Quelle que soit l'époque concernée, ceux-ci avaient doté la Chine d'une structure administrative consacrée à ce thème, dirigée par des *directeurs des multitudes* aux pouvoirs affirmés. Pendant plus de 2000 ans, le recensement a constitué un outil au service de l'administration chinoise.

L'Inde est un autre pays d'Asie qui a procédé dès le 4^e siècle avant notre ère au dénombrement de sa population. Elle a même été plus loin que cet objectif fondamental en prônant une politique planificatrice d'expansion territoriale et économique basée sur une connaissance approfondie de sa population. Un traité a défini la façon d'y parvenir, l'*Arthashastra*, rédigé par Kautilya, ministre de l'empire indien des Maurya. Il s'agit d'une méthode dont la minutie est remarquable, tant dans la manière de définir les caractères de la population prise en compte que de la quantité importante de données relevées (voir Hecht, 1987). On imagine sans difficulté que la mise en œuvre de ce type de relevé reposait sur un soutien administratif très dense, encadré par un contrôle policier explicite qui n'incitait pas aux non-réponses.

Cette manière d'agir a mis plus de temps pour être appliquée en Occident. La civilisation grecque accorda moins d'importance à la recherche du nombre d'habitants qu'à celui de la composition idéale de la Cité, chère à Platon, même si Aristote, dans sa *Politique*, s'attarda à réaliser des travaux de statistique descriptive et comparative. Les Romains reprirent les objectifs originaux : réaliser des recensements périodiques grâce à une structure administrative bien organisée afin de contrôler et d'administrer toutes les composantes de leurs territoires. Initiés sous Servius Tullius au 5^e siècle avant notre ère et réalisés jusqu'en l'an 73 sous Auguste (année du dernier recensement de l'empire romain), les dénombrements ont fait de la fonction de censeur, comme en Chine, un privilège recherché.

La période du déclin de l'Empire romain et le Haut Moyen Âge n'ont pas constitué un terrain fertile pour l'organisation de recensements. Ce n'est qu'entre le 14^e et le 16^e siècle, que l'on ressent à nouveau le besoin d'informations, que ce soit au niveau des rôles fiscaux ou à celui des relevés d'ordre religieux².

Figure 2 : DONNÉES VÉNUSIENNES



Dans un autre domaine, l'astronomie, les Babyloniens ont observé les mouvements du soleil et des planètes à intervalles réguliers, obtenant ainsi plusieurs observations d'un même phénomène (voir, par exemple, la tablette d'Ammi-šaduqa sur Vénus, présentée dans la figure 2, datant du 17^e siècle avant notre ère et visible au British

¹ - Les paragraphes qui suivent sont basés sur le chapitre 2 de l'ouvrage de Droysbeke et Vermandele (2014).

² - Mentionnons en particulier les ordonnances de Villers-Cotterêts de François Ier, en 1539, et celle de Blois de Henri III en 1579, qui introduisirent respectivement les registres de baptêmes et de mariages.



Jean-Jacques
Drosbeke

Museum). Nous ne connaissons malheureusement pas la manière dont ils ont remplacé ces observations multiples par des « valeurs de compromis ». On possède plus d'informations sur les travaux ultérieurs des astronomes grecs (voir Drosbeke et Saporta, 2010). Ainsi, Ptolémée, astronome du 2^e siècle, utilisa les relevés antérieurs d'Aristarque de Samos et surtout d'Hipparque et proposa, en présence de plusieurs observations d'un même phénomène, de conserver une seule valeur accompagnée de mesures de variation basées, semble-t-il, sur l'*étendue* des observations, c'est-à-dire l'écart entre la plus grande et la plus petite d'entre elles. Jusqu'au 16^e siècle, on préféra retenir une « bonne valeur » — en omettant souvent de justifier l'adjectif utilisé — que recourir à une synthèse systématique. Le premier qui utilisa une moyenne comme « outil de synthèse » est probablement l'astronome Tycho Brahé (1546-1601) dont les nombreuses données sur le mouvement des planètes permirent à Johannes Kepler (1571-1630) d'énoncer les lois qui portent son nom.

En cette fin du 16^e siècle, on construit des données *primaires* et *individuelles*, semblables à celles élaborées dans l'Antiquité depuis l'apparition de l'écriture. Pour ce qui concerne les données résultant d'observations répétées d'un même phénomène, les faibles progrès techniques réalisés dans la recherche d'une plus grande précision des instruments de mesure ont fait croire longtemps qu'une « bonne mesure » était meilleure qu'une agrégation dont on ne soupçonnait pas l'intérêt.

3 - Parmi lesquels il faut citer le nom d'Achenwall (1719-1772) à qui on attribue la paternité du mot statistique.

Les 17^e et 18^e siècles

Depuis le 15^e siècle, de nombreuses villes ont recensé leurs habitants. Les Etats tendant à se centraliser et à se doter d'une administration solide, le besoin de dénombrer se fait à nouveau sentir, même si la pratique est souvent défaillante. Jusqu'à la fin du 17^e siècle, les registres sont en effet loin d'être parfaits ! Ce 17^e siècle voit trois courants distincts se développer en Europe : la *Staatkunde* allemande, les *enquêtes* de l'administration française et l'*arithmétique politique* anglaise.

La *Staatkunde* allemande trouve ses racines dans les travaux d'Aristote. Pour ses défenseurs³, la statistique est la *science de l'Etat*. Purement descriptive, elle ne fait pratiquement jamais appel à des données chiffrées. Son influence est cependant significative, surtout en Europe centrale, et perdurera jusqu'au 19^e siècle.

En France, on plaide toujours pour les dénombrements comme outils de gouvernement. Deux hommes se sont particulièrement illustrés dans le recours à des enquêtes en raison des contraintes économiques : Colbert (1619-1683) qui développe une stratégie de dénombrement des villes et des régions, et Vauban (1633-1707), auteur d'une *Méthode générale et facile pour faire le dénombrement des peuples* en 1686.

Mais c'est en Angleterre qu'un mouvement novateur se répand avec l'arithmétique politique due principalement à Graunt (1620-1674) et Petty (1623-1687). Comme le dira Charles Davenant (1656-1714), émule de Petty, « *l'arithmétique politique est l'art de raisonner par des chiffres sur des objets relatifs au gouvernement* ». On y trouve les fondements de la *méthode du multiplicateur* qui a marqué les techniques de dénombrement des 17^e et 18^e siècles, provoquant une mise à l'ombre certaine de la *Staatkunde* allemande en Europe occidentale.

La méthode du multiplicateur repose sur l'idée suivante : il existe des quantités qui sont en rapports simples et relativement constants avec la population d'un pays. Si ces quantités sont plus simples à dénombrer (nombre de maisons, feux (foyers)... , ou encore nombre de naissances, de décès... dans l'année), il suffit de multiplier leur nombre par un *multiplicateur* adéquat pour obtenir une estimation du nombre d'individus dans la population. Pour les responsables politiques de l'époque, le recensement d'une population présente des désavantages certains (réactions de

méfiance des enquêtés, coûts de mise en œuvre trop élevés...) ; mais d'un autre côté, le choix d'une entité plus simple à dénombrer et la détermination d'un multiplicateur unique posent aussi des problèmes de fiabilité.

Une des caractéristiques du 18^e siècle – le siècle des Lumières – est le triomphe de l'esprit de calcul. Il faut dire que les progrès réalisés par les mathématiques sont alors considérables et la *loi des grands nombres* de Bernoulli vient ajouter sa pierre à l'édifice. L'époque est cependant marquée par de nombreuses imprécisions sur les estimations fournies par les uns et les autres. Il n'est donc pas étonnant de constater que les recensements sont revenus en force au 19^e siècle avant de connaître une stagnation puis un déclin au 20^e siècle, dû notamment à l'introduction de registres administratifs performants et au développement des techniques de sondage. Mais cela, c'est une autre histoire sur laquelle nous reviendrons ci-dessous.

Parmi les développements qui contribuent significativement à l'évolution de l'histoire des données, il faut souligner l'amélioration des instruments de mesure. Celle-ci est essentielle car elle permet aux hommes de s'aventurer sur les mers en s'assurant une meilleure qualité des moyens de se guider. Par ailleurs, si les mesures astronomiques constituent toujours une manière incontournable — malgré leurs imprécisions — de savoir où l'on se trouve, un autre instrument de connaissance permet de mieux maîtriser le sol sur lequel on vit : la géodésie. Cette discipline et l'astronomie constituent deux domaines privilégiés d'un traitement de données qui se cherche. Un exemple remarquable est celui de la mesure d'un arc de méridien, au centre d'une question primordiale à l'époque : quelle est la figure de la terre ? Le recours à la technique de *triangulation* est à l'origine d'aventures multiples de cette mesure dans diverses régions du globe qui permettront de résoudre la question (voir Droesbeke *et al.*, 2016).

Le besoin de mesurer est partagé par de nombreux savants qui bénéficient d'instruments de mesure de plus en plus précis. On devient exigeant à propos de la qualité des observations effectuées et l'erreur de mesure devient un souci essentiel.

Deux approches coexistent pendant de nombreuses années. La première milite pour la recherche d'une bonne mesure, entachée d'une erreur limitée, inférieure à une erreur maximale,

acceptable ou en tout cas à craindre. Dans cette optique – défendue par Leonhard Euler (1707-1783) – prendre en compte d'autres mesures en plus de la bonne ne peut que faire croître l'erreur globale, notamment en utilisant les observations les plus mauvaises.

Un deuxième modèle nous intéresse davantage ici. Il repose sur l'hypothèse que l'utilisation de toutes les observations permet des compensations dont on peut espérer qu'elles réduisent l'erreur résultante. C'est en recherchant des modèles appropriés de distribution des erreurs que de nombreux scientifiques contribueront à la consolidation d'une théorie qui sera qualifiée en 1765 de *théorie des erreurs* par Johann-Heinrich Lambert (1728-1777).

La multiplicité des observations et le besoin de s'interroger sur le comportement des *erreurs de mesure* n'amènent pas seulement de nombreux scientifiques à vouloir modéliser cette erreur pour mieux la dompter ; cette question comporte aussi des aspects politiques et commerciaux qui constituent autant d'enjeux importants pour l'époque.

Les données se multiplient et se contredisent. Il faut en comprendre la raison, les gérer pour en tirer profit. Le calcul des probabilités vient en aide à ceux qui affrontent ce problème. Il en résulte une conséquence à trois facettes dont les effets seront durables : la *loi des erreurs*, qui sera qualifiée de « normale » à la fin du 19^e siècle, devient une loi de référence, la *moyenne* s'avère être le mode de synthèse privilégié et un critère d'ajustement devient incontournable : le *critère des moindres carrés* (voir Droesbeke et Tassi, 2015). Deux hommes jouent un rôle central dans cette histoire : Pierre Simon de Laplace (1749-1827) et Carl Friedrich Gauss (1777-1855).

Les données individuelles deviennent plus fiables ; elles font place aux *données agrégées* et aux *données transformées* (voir Droesbeke et Vermandele, 2016).

Quelques points forts du 19^e siècle

Le 19^e siècle occupe une place très importante en statistique. S'il fallait retenir cinq caractéristiques essentielles de ce siècle dans l'histoire des données, notre choix serait le suivant :

1. L'application des trois outils utilisés en astronomie (loi normale, critère des moindres

carrés et moyenne) à l'étude des populations et de leurs caractéristiques humaines, permettant à Adolphe Quetelet (1796-1872) de créer une *théorie des moyennes* aux accents multiples (voir Académie Royale de Belgique, 1997, Desrosières, 1993 ou encore Droesbeke et Vermandele, 2016).

2. Le développement de la *statistique* comme outil de gestion des Etats, au niveau économique et social (voir Desrosières, 1993). Les tables statistiques et les représentations graphiques deviennent un outil important d'analyse et de communication.
3. Le remplacement du rôle central de la moyenne par celui de la dispersion dans les préoccupations des savants de tous bords.
4. Le déplacement du centre de gravité de la statistique vers Londres et l'émergence des concepts de corrélation et de régression (voir Droesbeke et Tassi, 2015 et Droesbeke et Vermandele, 2016).
5. L'émergence d'une nouvelle méthode de recueil des données : les sondages (voir Droesbeke et Tassi, 2015).

Les données individuelles se répandent ; la manière de les produire se diversifie. Elles deviennent nombreuses : il faut les montrer et les résumer.

Le 20^e siècle et le début du 21^e siècle

Il est difficile de détailler dans cet article les développements de la statistique au 20^e siècle, tant les innovations sont nombreuses et diversifiées. Il est certain que l'*inférence statistique* est au centre de ces dernières, avec ses deux problèmes centraux, l'*estimation de paramètres d'une population* et les *tests d'hypothèses* réalisés à partir d'un échantillon. La figure 5 nous montre les principaux acteurs et le moment de leur activité la plus intense (en rouge) dans le développement de leurs travaux.

Nous ne pouvons expliciter ici toutes les ouvertures nouvelles du 20^e siècle et du début du siècle actuel, qui traitent des données : elles s'appellent *plans d'expérience, méthodes de sondage, analyse statistique bayésienne, analyse exploratoire des données, analyse robuste...* Parallèlement des *stratégies d'analyse* ont vu le jour ainsi que des *procédures de diffusion des résultats d'analyse* appropriées.

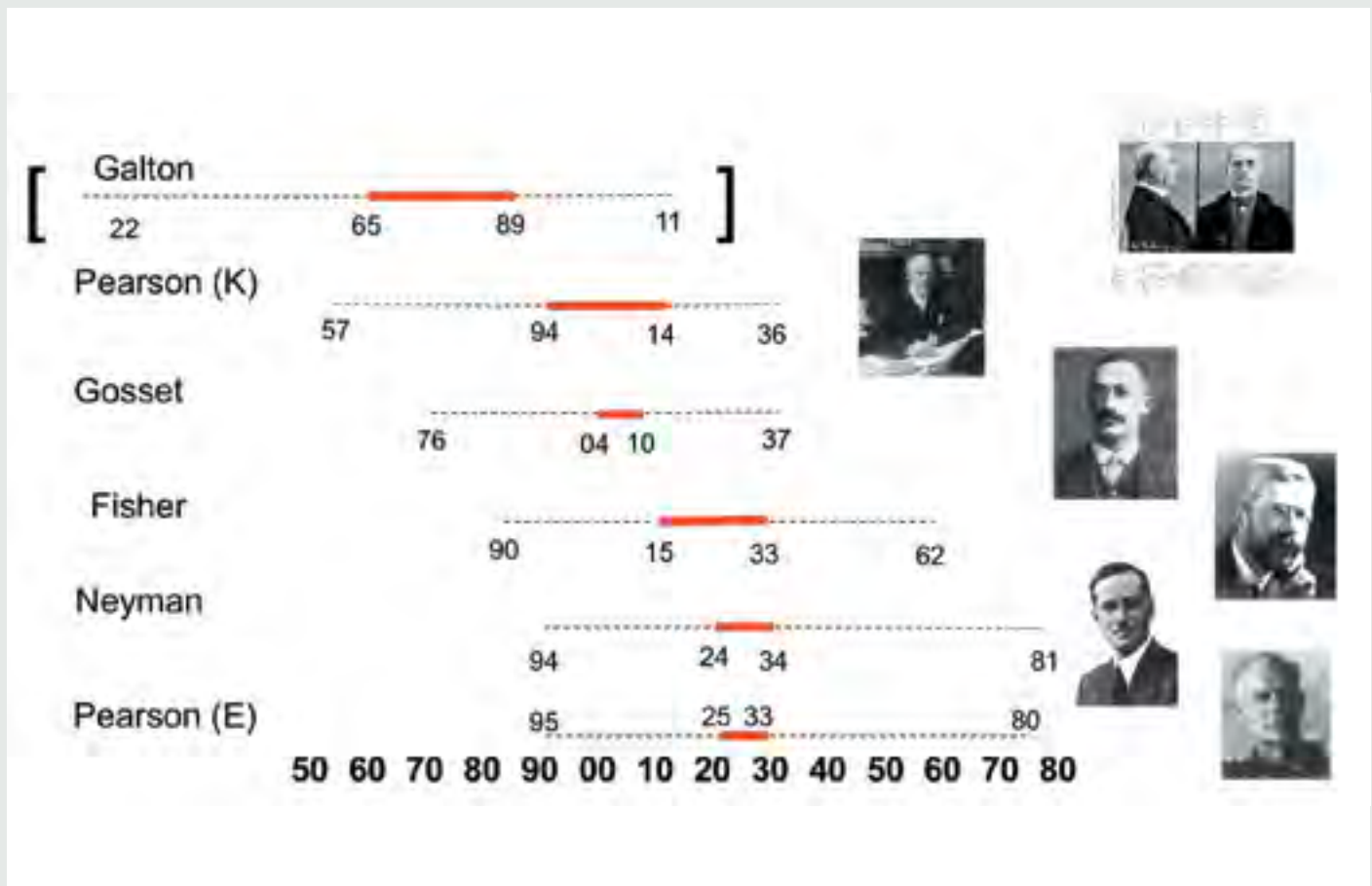
Les données sont à présent *multivariées*, elles sont *qualitatives* ou *quantitatives*, il en est de *manquantes* et d'*extrêmes*, elles deviennent de plus en plus *nombreuses*. En ce début de 21^e siècle, les *mégadonnées*, encore appelées *données massives* ou *big data* nous lancent des défis de toute nature : méthodologiques, techniques, juridiques... Le statisticien de demain se doit de s'ouvrir encore davantage de la *science des données*. ■

Figure 3 : LAPLACE



Figure 4 : GAUSS



Figure 5 : LES MOMENTS FORTS DE L'INFÉRENCE STATISTIQUE**Bibliographie**

ACADEMIE ROYALE DE BELGIQUE (1997), *Actualité et universalité de la pensée scientifique d'Adolphe Quételet*, Actes du Colloque des 24 et 25 octobre 1996, textes rassemblés sous la direction scientifique de J.-J. Droesbeke, *Mémoire de la Classe des Sciences*, 3^e série, tome 13.

DESROSIERES A. (1993), *La politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte.

DROESBEKE J.-J., MAUMY-BERTRAND M., SAPORTA G. et THOMAS-AGNAN Ch. Eds. (2016), *Models choices and aggregations*, Paris, Technip (à paraître).

DROESBEKE J.-J. et SAPORTA G. (2010), Les modèles et leur histoire, dans Droesbeke J.-J. et Saporta G.

(éds), *Analyse statistique des données longitudinales*, Paris, Technip, 1-14.

DROESBEKE J.-J. et TASSI Ph. (2015), *Histoire de la statistique*, 2^e édition corrigée, Collection Que-sais-je ?, IAD, Paris, Presses Universitaires de France.

DROESBEKE J.-J. et VERMANDELE C. (2016), *Les nombres au quotidien* (à paraître).

HECHT J. (1987), L'idée de dénombrement jusqu'à la révolution, dans Affichar, J. (éd.), *Pour une histoire de la statistique*, 1, Paris, Economica, 21-81.

REY A., TOMI M., HORDE T. et TANET Ch. (1993), *Dictionnaire historique de la langue française*, 2^e édition, Paris, Dictionnaire Le Robert.

TEMPS 1

Opportunités et risques dans une société connectée

1

Les avantages sociétaux d'un monde connecté

Lionel Janin (2003), Chargé de mission numérique, France Stratégie, Services du Premier ministre

Héritier du Commissariat général au Plan, France Stratégie a quatre missions principales, résumées par les verbes évaluer, anticiper, débattre et proposer. La dimension d'anticipation est illustrée par les travaux de France Stratégie autour de l'avenir d'internet, du big data ou des objets connectés ; ces travaux sont disponibles sur le site internet de France Stratégie.

Les questions numériques possèdent un grand avantage : elles permettent de toucher à tout. Les thèmes de travail s'appellent : numérique et santé, numérique et éducation, etc. Ces questions sont universelles, et il est donc important de se concentrer sur quelques points pour montrer le contenu et les effets de cette transformation numérique, et en particulier ces nouvelles données disponibles et cette nouvelle connexion peuvent intervenir sur un large champ d'activités.

Ce colloque a lieu le 18 juin. Il serait facile d'évoquer un appel du 18 juin, le bicentenaire de

Waterloo... ou la situation de la France dans le domaine numérique. Il se trouve que le Premier ministre va présenter ce matin également la stratégie numérique du gouvernement. Le sujet est donc au cœur de l'actualité, et je suis certain que les questions de connexion, d'équilibre entre les opportunités et la prévention des risques seront inscrites dans cette stratégie pour la France.

Le numérique est partout. Hier encore, le journal *Le Monde* contenait un article détaillé sur la voiture connectée qui fera partie du quotidien en 2025. Ce qui est surprenant, ce ne sont pas tant les projets évoqués, que la vitesse à laquelle se font ces évolutions. Des objets qui, il y a un an ou deux, avaient encore un caractère prospectif, voire futuriste, deviennent des éléments réels, sur le point d'être mis en œuvre, dans un futur relativement proche.

Mon intervention commencera par une première partie traitant, de manière assez générale,

de la transformation numérique de l'économie et des organisations. Puis, seront abordés ses trois principaux avantages : son utilisation pour simplifier les services, les possibilités de nouveaux services et le sujet plus innovant encore que sont les systèmes de prédiction.

La transformation numérique de l'économie et des organisations

Le professeur Droesbeke, dans son exposé d'ouverture, est parti du pays de Sumer. Je me contenterai de commencer il y a 25 ans, avec les origines d'Internet au niveau du grand public ; rappelons que le moteur de recherche de Google, qui nous est familier, ne date que de 1998. Nous sommes sur des horizons très courts. Le prouvent les objets nouveaux, comme le smartphone : il n'a pas dix ans et néanmoins son taux de possession n'est certes pas de 100 % mais dépasse 40 % dans un grand nombre de pays. On peut considérer que le 9 janvier 2007 – jour de la présentation de l'iPhone – est la date de naissance d'un nouvel espace d'objets connectés. Il en est le symbole et le précurseur, en tant qu'outil à la fois GPS, connecté au réseau, pouvant prendre des photos, et disposant d'une interface pouvant jouer un rôle quasiment universel. Cet objet mobile va jouer un rôle majeur dans le développement de l'usage du numérique.

Pourquoi ? Qu'est-ce qu'il y a de différent dans cet usage numérique par rapport à ce qu'on pouvait observer avant ? Dans les transactions des services numériques, il y a déjà le fait d'avoir des statistiques d'usage en temps réel. C'est une première forme de suivi. Il est donc possible de suivre facilement les modes d'utilisation, contrairement aux objets précédents pour lesquels il était nécessaire de réaliser des enquêtes déclaratives, parfois complexes. Le smartphone permet une remontée immédiate d'informations qui autorisent un diagnostic rapide sur les fonctionnalités qui marchent ou non, et permet ainsi de faire des améliorations en permanence. Les services numériques ont montré que l'on pouvait modifier le système et mettre en place une amélioration numérique permanente.

En outre, nous sommes en mesure de créer une personnalisation, c'est-à-dire que des services au départ uniformes sont adaptés à chacun, puisque chacun dispose d'un terminal et d'une connexion. Et finalement, cet objet devient capable de retracer l'histoire d'un individu, puisqu'il va tracer, conserver, des informations :



Lionel Janin

photos, trajets, échanges, etc. Son possesseur acquiert une forme d'identité numérique qui permet des nouvelles relations contractuelles. Des exemples en sont Uber, Airbnb, et d'autres services de ce type. Finalement, nous interagissons avec des inconnus dont nous connaissons certains historiques, avec un niveau de confiance suffisant, relevés par leurs traces numériques.

Ces champs numériques gravitent autour de deux noyaux centraux.

Le premier est l'individu : on trouve le smartphone, bien sûr, très répandu, et demain des montres, des traqueurs d'activité, des objets connectés. Ces dispositifs numériques sont individuels.

Le deuxième noyau est celui de la ville connectée, autrement appelée la *smart city*, de l'industrie, des usines, des systèmes de production, des systèmes de transport.

Et c'est au sein de ces deux grands et différents noyaux que se jouent les grands effets que je vous propose de parcourir rapidement.

Un monde hyper-connecté, générateur de big data

Une des évolutions majeures de ces dernières années a été l'arrivée au niveau du grand public du principe de connexion. La transformation numérique n'est pas restée simplement limitée à un certain nombre de personnes gravitant dans les

métiers de la communication, métiers naturellement numérisés. Elle a essaimé partout, dans des secteurs qui étaient, auparavant, assez loin de ces sphères numériques, comme l'agriculture.

Abordons ces questions en termes de marketing, cela permet de structurer le propos. Cette concentration numérique est large parce que le nombre d'individus connectés ne cesse de croître. A l'échelle mondiale, on estime à un tiers la proportion d'individus connectés. Bientôt, nous atteindrons six milliards de smartphones et de connexions à Internet.

Quant au nombre d'objets connectés, différents instituts et organismes l'évaluent à cinquante, quatre-vingts, cent milliards d'ici 2020. Cependant, il y a une très grande différence entre un smartphone et des capteurs qui permettent une connexion avec un contenu informationnel. Ces éléments contribueront à court terme à transformer des objets initialement distincts en objets connectés.

Tout ceci va engendrer de gros volumes de *data*. On a parfois coutume de dire que 90 % des données actuellement disponibles au monde ont été générées au cours des deux dernières années. C'est simplement une autre manière de dire que les volumes de données sont multipliés par trois chaque année. Mais ces questions de volume ne sont peut-être pas si cruciales. En effet, mis à part les volumes liés à la vidéo, cela reste classique. L'aspect probablement le plus typique et original de ces big data est la grande variété des données, la juxtaposition de données à la forme relativement structurée, relevant presque de la statistique classique, et de contenus qualitatifs, d'images fixes ou mobiles ; ces contenus hétérogènes donnent naissance à la recherche d'interactions, sinon de synthèse.

Ces volumes de données peuvent être gigantesques, mais aussi, parfois, de simples traces. Ce qui est fondamental, c'est plutôt le fait que quelqu'un puisse envoyer ne seraient-ce que quelques octets par jour, et c'est cela qui va contribuer à cette révolution que je vais décrire.

L'un des aspects mentionné traditionnellement, en particulier depuis les travaux de Fisher dans les années 1920, consiste à s'interroger sur la qualité des données et leur niveau d'exhaustivité. Nous nous retrouvons maintenant confrontés à des situations où les données ne sont pas exhaustives, et donc difficilement exploitables. Mais il y a des cas où les données sont fournies par des capteurs,

comme des capteurs de température, des capteurs dissimulés dans l'environnement, ce qui entraîne une plus grande précision. Par exemple, un trajet d'approche d'un taxi commandé, et celui fait dans le taxi, sont suivis tant par le smartphone du chauffeur que par celui du client.

Il existe des outils de vérification qui influent sur la qualité potentielle des données, et font que ces données sont de plus en plus utilisables. Quelles sont les conditions d'utilisation de ces données ? Il y a des usages structurés autour de différents thèmes. Mais avant de répondre à cette question, il convient de mentionner qu'il faut évoquer non seulement les *data*, mais aussi leur coût, et sur ce point il reste à apprendre. Somme toute, à ce jour, même si les usages commencent à se développer, il n'y a pas tant d'utilisation effective des big data. Le point important est que ces usages vont se développer parce que les conditions de possibilité évoquées auparavant sont remplies de façon croissante.

Une amélioration des services rendus au public

Abordons maintenant quelques grands champs d'application du big data.

Le premier champ du big data concerne la capacité à améliorer la qualité du service rendu. Le big data est une manière d'optimiser l'interaction faite, de fournir des informations de façon plus fluide. C'est ainsi le cas dans l'éducation, comme ces services initialement venus des Etats-Unis au travers d'un enseignement en ligne, les cours massifs, les MOOC. On pourra ainsi être en mesure de suivre les réactions individuelles des participants à un cours, et donc détecter quels sont les paramètres qui font qu'un cours est mieux compris, et *in fine* optimiser la qualité de la formation dispensée en adaptant le parcours aux outils disponibles.

Le deuxième champ dans lequel la transformation numérique est évidente et dont nous bénéficions pour une grande partie d'entre nous est le domaine des transports. Il est déjà bien pratique qu'un GPS puisse aider à se situer ; mais, au-delà, le fait que ces données soient regroupées entre un grand nombre de véhicules autorise une juste perception en temps quasi-réel des conditions de trafic, des incidents – accidents, manifestations – pouvant l'altérer. Une personne en déplacement automobile bénéficiera du fait que le logiciel comprend qu'une manifestation perturbe le trajet et pourra donc l'adapter. Dans des circonstances

plus dramatiques aussi, à Haïti après le tremblement de terre, se sont posées de nombreuses et graves questions de bonne répartition de l'aide alimentaire ; un système de big data a alors été déployé pour optimiser la manière dont l'aide a été fournie aux populations en détresse.

Un troisième autre champ ayant attiré l'attention est le marché du travail, pour la diffusion des offres d'emploi et leur adaptation aux demandes. Les profils numérisés des demandeurs, leur historique sont traités par des outils capables de faire des adaptations entre demande et profil, et les offres d'emplois disponibles. Cet usage du numérique a notamment fait l'objet d'un rapport en Conseil d'orientation de l'emploi.

Un quatrième domaine, très important pour le service public, consiste à remplacer un système basé sur un formulaire compliqué à remplir par un système dans lequel un certain nombre d'informations individuelles sont connues a priori. Bel exemple de services apportés aux personnes pour utiliser au mieux tous les dispositifs d'aide auxquels ils ont droit. Pour limiter ainsi le nombre de pièces justificatives, il faut que les services existants soient fournis dans les meilleures conditions aux usagers. Nous en avons l'habitude avec les services commerciaux et marchands, ce que l'on appelle l'expérience utilisateur. Il n'y a aucune raison que cette qualité de création ne se généralise pas aux services publics.

Il y a aussi un champ que je souhaite mentionner, sans m'y attarder ; l'amélioration du fonctionnement des systèmes. C'est par exemple le cas dans la vente et la gestion des stocks, la gestion des prix. Les big data et leurs traitements permettent l'optimisation d'une clientèle, la relance optimisée d'une catégorie d'acheteurs. Cette technologie permet aussi la lutte contre certaines fraudes : ainsi, le ministère de l'Économie s'est doté d'une cellule utilisant des données disponibles pour mieux détecter les cas de fraude fiscale.

Un dernier champ, important, concerne le développement durable. Les questions climatiques sont l'un des enjeux majeurs du XXI^{ème} siècle. Les données massives, les capteurs d'objets connectés sont des moyens ouvrant la possibilité de mesurer l'utilisation d'énergie, et de l'optimiser de façon très précise, en particulier pour les énergies renouvelables. Les objets intelligents sont capables de prendre en compte des signaux pour faire cette optimisation. Ces questions sont absolument essentielles.

Dans le transport, notamment en ce qui concerne l'entretien, la maintenance régulière et uniforme laissera la place, grâce à des capteurs disséminés dans l'environnement, à des opérations d'entretien ciblées plus efficaces.

Un nouveau champ de développement : l'analyse prédictive

Le dernier élément sur lequel je voudrais m'étendre est ce que l'on appelle l'analyse prédictive. Ce domaine de modélisation, issu de développements parfois assez anciens comme les réseaux de neurones, a connu d'énormes progrès au cours de ces dernières années, avec l'apparition des techniques d'apprentissage. Les secteurs d'application tournent autour de la reconnaissance, qu'il s'agisse d'images, de l'écriture, de visage ou de la voix sur certaines applications en ligne. Les traductions automatiques deviennent de bonne qualité, grâce à l'analyse d'énormes volumes de données. Ces techniques constituent également un apport important dans le domaine de la santé. Le jour où ces données de santé pourront être remontées pour permettre le suivi d'une épidémie, cela offrira des possibilités extrêmement grandes. Il en est de même pour la détection et le suivi des pathologies en recueillant et corrélant des informations qui sont, pour le moment, difficilement accessibles.

D'une manière générale, ces systèmes prédictifs sont capables d'intégrer des données variées, disponibles, pour construire une aide au diagnostic, comme l'outil Watson, créé par IBM, qui balaye des articles de recherche médicale pour identifier quels en sont les éléments les plus pertinents.

Numérique et croissance : une question complexe

Un dernier mot pour conclure : les ordinateurs sont partout, depuis longtemps, mais pas tellement dans les chiffres de croissance et de productivité. Dans le domaine du numérique, la mesure de la croissance est plus complexe. Il y a encore énormément de biens, de marchés numériques qui ne sont pas encore parfaitement organisés aujourd'hui. Les systèmes de mesure ne sont pas forcément adaptés à cette nouvelle forme de croissance, parce que la croissance est extensible. Il y a plus d'optimisation, de partage des utilisations et des améliorations. Ce sont des facteurs relativement nouveaux qui sont moins bien saisis pour l'instant. Nul doute que nous saurons en tirer parti pour relever les défis qui nous attendent. ■

2

Les avantages pour la connaissance économique et statistique

Stéphane Gregoir (1988), Directeur de la méthodologie et de la coordination statistique et internationale, Insee

Mon objectif est ici de faire quelques remarques pour aider à la construction d'un cadre de réflexion sur la transformation numérique et illustrer ses opportunités et quelques limites. L'élément constitutif majeur de cette transformation est la génération massive de données, celles-ci doivent donner la possibilité de prendre de meilleures décisions individuelles et collectives et de construire des nouveaux savoirs. Y sont naturellement associées de nombreuses opportunités économiques pour les activités privées comme pour l'action publique, mais il y a aussi quelques écueils.

Des données en quantité

Un premier élément constitutif de la transformation numérique est le volume de données. Il est usuel d'évoquer une croissance exponentielle de ce volume et les nombres avancés ne nous parlent pas nécessairement. A la base, il y a le byte (un octet, c'est-à-dire 8 bits), mais les unités de comptes aujourd'hui utilisées sont le terabyte ou billion d'octets, le petabyte ou milliard d'octets, le zettabyte ou trillion d'octets ou encore milliard de milliards d'octets. Il nous faut les illustrer. En 2013, il y avait 960 milliards de photos¹ sur Facebook, ce qui correspond à 357 petabytes. Si nous voulons approcher ce qu'il y a dans la tête d'un être humain, des biologistes ont établi que la capacité de stockage du cerveau est au maximum de 2,5 petabytes, soit 0,7 %. Nous disposons d'une évaluation selon laquelle, en 2013, le volume de données générées sur Internet était de quatre zettabytes. Selon Cisco (2015), le trafic IP global annuel de 0,6 zettabytes en 2014 devrait dépasser le zettabyte en 2016. Des travaux menés par un linguiste (Lieberman (2013)) ont conclu que si l'on addi-

tionnait toutes les paroles prononcées au monde depuis l'apparition du langage, en supposant que nous les ayons enregistrées sous forme numérique 16kHz 16-bit audio, nous obtiendrions environ quarante-deux zettabytes. Même en admettant que 2 % de ce qui a été dit dans l'histoire de l'humanité par des milliards d'individus était important, à quoi peuvent bien correspondre ces flux ou volumes de données générées ou échangées chaque année.

En fait, la majorité de ces flux est composée de données au sens le plus large du terme. Ils correspondent à des données de traces, de signal sonore, de signal visuel, ... Selon les experts, entre 7 % et 10 % de ces données sont des chiffres ou des mots et beaucoup sont associés à des protocoles de gestion. Ces volumes de données sont donc hétérogènes et l'usage qui peut en être fait est inégal ou en devenir. Ce qui est intéressant pour la société et les individus, ce n'est pas la donnée pour elle-même, mais ce qu'ils peuvent en faire, c'est-à-dire comment passer d'un ensemble de données à une information, puis des informations au savoir, un ensemble d'informations pérennes, compréhensibles et utilisables par tout un chacun moyennant un niveau de connaissance suffisant. L'information peut ne pas être issue de compositions de lettres et de nombres, elle peut résulter de structures pérennes de collecte de données variées confrontées à d'autres données qui par la répétition autorisent des interprétations. Plus les structures de collecte de données sont permanentes, plus des régularités peuvent être observées et offrir des opportunités d'apprendre, de comprendre ou de développer des activités économiques. La collecte et la diffusion des données passent par des réseaux, il faut préciser la nature de ces objets pour ana-

1 - Si vous regardiez en continu chacune de ces photos 2 secondes, il vous faudrait plus de 60000 ans pour les passer en revue.

lyser la nature des données et les opportunités qu'elles offrent.

Les deux notions de réseau

L'usage du mot « réseau » couvre deux objets.

Le premier objet est le réseau qui collecte et transmet des données et en organise la circulation. Ces données doivent être ensuite interprétées selon des règles pour éventuellement construire des informations. Il existe de nombreux systèmes de collecte de données. Certains sont physiques. Ils peuvent être composés des capteurs de température, de pression, de détection de mouvement, de mesure de la vitesse... Ils peuvent être collectifs, comme les réseaux de surveillance de l'état des routes ou individuels, comme les bracelets achetés individuellement ou donnés par un service de l'hôpital pour suivre les paramètres physiologiques d'une personne (la tension, le rythme cardiaque, etc...). D'autres systèmes de collecte reposent sur des personnes qui volontairement relèvent des données. Ils constituent le réseau et ont fait choix de définitions et de modes d'échanges des données. Par exemple, ce sont en France des amateurs qui collectent les données d'une partie de la mesure de la biodiversité dans les régions ou ce sont des volontaires qui collectent des adresses pour construire une base de données d'adresses pour Openstreetmap.org. Leur fonction est d'enregistrer des données dans des formats définis pour permettre ultérieurement des actions. En général, ce type de réseaux est défini selon la finalité, il répond à un besoin. Une analyse économique a abouti à l'investissement qui le crée. Cela reste vrai pour les réseaux de volontaires. On peut décider d'étendre ou de réduire un réseau, il faut le gérer et l'entretenir.

Un deuxième objet couvre une autre dimension immatérielle du réseau : elle concerne tout ce qui est la gestion et la création de liens entre des unités (individus, entreprises, associations,...). Elle contient une dimension dynamique évolutive. L'opportunité économique associée à cette dimension est liée à la valeur des liens créés. De nombreux exemples peuvent être donnés. Un réseau social crée de la valeur pour les personnes qui y ont un compte actif et disposent ainsi d'informations sur d'autres membres du réseau qui sont d'accord pour les partager. Ce service réduit entre autres les coûts de coordination entre les membres. Un lien transporteur-distributeur est un autre type d'élément d'un réseau. La valeur qu'il génère résulte de



Stéphane
Gregoir

l'intérêt de chacun des membres du couple : éviter les déplacements à vide des moyens de transports pour l'un, trouver rapidement des moyens disponibles pour l'autre. Les relations entre entreprises et individus sont aussi nombreuses. Dans le domaine de la consommation, il est coûteux de connaître l'ensemble des biens existants sur le marché. L'échange entre les deux types d'acteurs va porter sur d'une part l'information à propos des biens et d'autre part la révélation des goûts ou préférences personnelles. Sur le marché du travail, de nouvelles relations se mettent en place pour la recherche d'emploi. Il faut essayer de faire se rencontrer les personnes du fait de leur expérience et qualifications susceptibles de répondre aux besoins de l'entreprise. Ceci implique la création de liens qui, s'ils aboutissent à des embauches, engendreront de la valeur.

De la nature des données

Les données prennent des formes multiples et leurs natures et propriétés dépendent de l'aspect du réseau que l'on considère et de la nature de ce dernier. Nous distinguerons trois qualités de données.

Tout d'abord les données primaires, qui résultent de la finalité des réseaux. Elles sont associées à des dimensions du service rendu. Par exemple, il peut s'agir d'une trace générée lors d'une transaction commerciale payée par carte de crédit dans un supermarché. Elle comprend éventuellement la liste descriptive des achats, mais néces-

sairement leur montant, le lieu, la date et l'heure, le compte crédité et celui à débiter. Ou encore, sont dans cette classe les relevés de communications téléphoniques qui indiquent le numéro appelé, le numéro appelant, le temps passé en communication, les fournisseurs de chaque partie, la date et l'heure. Ces données se retrouvent sur la facture adressée à l'appelant.

Un deuxième type de données comprend les données secondaires ou métadonnées qui sont engendrées de façon passive. Elles sont le résultat de la mise en œuvre technique ou de la gestion du service pour en assurer l'efficacité et la qualité (vu du point de vue du producteur). Elles peuvent être de deux natures qu'il est important de distinguer. Elles sont dépendantes du système de gestion ou de production propre à l'entreprise ou au contraire répondent à des contraintes techniques transparentes sous formes de protocoles qui peuvent être connues de tous. Dans le premier cas, les données sont constituées de manière non transparente et susceptibles d'être le résultat de processus évolutifs. C'est par exemple le cas des traitements des moteurs de recherche, nous reviendrons sur ce point plus loin. Dans le second cas, les données sont le résultat de processus stables et publics. Elles reposent plutôt sur des réseaux matériels. Par exemple, notre téléphone en veille émet des signaux qui arrivent sur des bornes afin de permettre de joindre ou d'être joint rapidement. Les données qui leur correspondent sont issues de dispositifs stables qui reposent sur des protocoles et une algorithmique publique. Ces données secondaires ou métadonnées sont la propriété des entreprises qui les génèrent et les utilisent pour produire leur service.

A contrario, un troisième type de données est constitué par celles que nous donnons volontairement, comme nos coordonnées, la description de moments de notre vie, de nos impressions, de nos sentiments, de nos goûts, par exemple sur les réseaux sociaux. Ces données fournissent des informations à l'opérateur. Techniquement, la personne qui dépose volontairement ces données en garde la propriété, les laissant à la disposition des gestionnaires de ces données pour une amélioration du service.

Toutes ces données sont associées à nombre d'opportunités de services nouveaux ou améliorés. Elles peuvent donc être associées à un marché de vente de données. Il faut savoir les utiliser et connaître aussi les risques qui existent du fait de leur nature différente. Il n'est pas toujours

possible d'en contrôler le contenu et, partant, l'information construite à partir d'elles.

Données, information et valeur économique

La question économiquement fondamentale est ici celle de la valeur qui va être tirée des données pour soutenir les appariements créés dans un réseau, que ce soit entre des individus, des entreprises ou des individus et des entreprises. C'est le problème de la valeur de l'information que l'on peut construire à partir de ces données et du coût qu'il faut supporter pour la créer.

Outre la possibilité offerte par les réseaux modernes d'être en relation à toute heure et de n'importe quel endroit, l'une des transformations majeures observées au cours des vingt dernières années est la réduction du coût de collecte de la donnée. Nous disposons d'une foule de systèmes engendrant et collectant des données que nous pouvons essayer de transformer en informations. Cette transformation est le moteur du développement économique, de la création de valeur la plus importante. Elle nécessite des personnes capables de mettre en œuvre cette transformation qui peut être complexe. Il faut déceler et valider des corrélations, s'en servir pour offrir un service puis informer le public et le marché qu'un tel service de qualité existe. En effet, une fois créé ce service s'appuyant sur ces informations, ces dernières doivent être reconnues comme pertinentes, fiables et sincères pour que l'utilisateur s'y intéresse. Certes, l'utilisation initiale du service va venir de la satisfaction immédiate d'un besoin, mais la qualité de service doit être durable pour maintenir l'intérêt du public. La concurrence entre des services voisins va faire émerger les algorithmes basés sur les régularités pertinentes et robustes. La pertinence est en général dépendante du demandeur, aussi la déclinaison du service selon le profil de l'utilisateur assurera une perception subjective plus grande de la satisfaction du besoin, qui peut être réelle. A côté de ce processus d'individualisation qui va reposer sur les données secondaires et les informations personnelles confiées par l'utilisateur, un processus de qualification de l'information diffusée sur la nature des biens ou des services proposés va être mis en œuvre. Ce processus peut prendre différentes formes : un audit par des tiers, des enquêtes gérées par l'entreprise ou un tiers, ou la collecte d'appréciations individuelles de clients sur les produits ou services offerts. Cette dernière fonction est un service en lui-même qui doit aussi pouvoir établir sa sin-

cerité. Il est vrai que dans certains cas, il se crée un problème d'incitation individuelle lorsqu'une entreprise cherche à construire une information certifiée qui devient un bien public à l'usage des concurrents. L'introduction d'un tiers permet de traiter en partie ce problème d'incitation. Le modèle économique qui est associé prend différentes formes.

Il faut créer la confiance de l'usager dans l'information construite à partir des données. La réduction du coût de la collecte a été accompagnée de celle du coût de la diffusion, si bien qu'au total, les moteurs économiques de la révolution numérique sont surtout la construction de l'information à partir des nouvelles données et sa certification. Un certain nombre des exemples présentés par Lionel Janin portent sur des marchés qui vont évoluer en raison de ces mécanismes de création de valeur. Ces marchés portent sur des biens et services privés ou publics.

Dans le secteur privé, la génération d'une grande masse d'informations certifiées ou en cours de certification, leur circulation à moindres coûts et la possibilité de les adapter aux profils des agents entraînent un fonctionnement plus efficace des marchés existants et le développement de nouveaux marchés tels que les marchés qui n'étaient pas solvables car trop étroits pour des marchés localisés ou les marchés de seconde main ou celui du partage du fait de l'asymétrie d'informations. Les acteurs web qui organisent ces marchés traitent cette question d'asymétrie en donnant des notes et des appréciations sur les membres adhérents du réseau. De nouveaux mécanismes économiques ont aussi été développés pour améliorer le fonctionnement parfois complexe de ces nouveaux marchés. Il en est ainsi des algorithmes de classement des moteurs de recherche qui tout en répondant à une requête individuelle peuvent présenter les résultats dans un ordre particulier. Cet ordre résulte d'un mécanisme de marché avec un système de prix d'achat par les entreprises concernées. Des formes de marketing personnalisé réduisent le coût et augmentent l'adéquation du message à la cible, il devient plus informatif. La possibilité de faire des expérimentations de manière transparente pour les utilisateurs permet d'améliorer les modes de gestions des appariements et la valeur associée. Au total, les déséquilibres sont moins nombreux. Des travaux empiriques montrent que les prix moyens sont en général moins élevés mais demeurent dispersés, qu'il faut un certain temps avant d'atteindre cette situation.

Dans le domaine de l'action publique, les mêmes principes s'appliquent. Il s'agit de créer des appariements liés à des services publics qui en augmentent l'efficacité et réduisent les coûts. Ceci peut se décliner dans les domaines de la santé en individualisant les traitements et en augmentant les efforts de prévention par un suivi éventuellement en direct et en continu à l'extérieur des lieux de soins ; dans les domaines de l'éducation en individualisant le suivi des élèves et en leur donnant les moyens de s'auto-évaluer fréquemment ; dans les domaines des infrastructures en ayant un suivi de l'état des réseaux en continu ou en matière d'électricité en ajustant la production aux besoins...

Enfin, ce nouvel environnement a suscité de nombreux travaux de recherche de la part d'économistes, de psychologues cognitivistes, de statisticiens ou informaticiens, de médecins... Il est difficile de traiter de chacun de ces domaines dans leur diversité. En économie, un élément important est la théorie des marchés bifaces développée par l'Ecole de Toulouse au début des années 2000 qui s'est révélée adaptée pour traiter certains de ces nouveaux types de marchés. A partir d'un ensemble d'utilisateurs et de producteurs, comme des possibles employeurs ou des personnes en recherche d'emploi pour reprendre cet exemple, il faut organiser le marché pour que l'on puisse créer les appariements générant de la valeur. Dans les sciences sociales, les questions sont multiples et fréquemment multidisciplinaires. Il s'agit de mieux appréhender le comportement des agents compte tenu de leur biais cognitifs, d'analyser les mécanismes de l'homophilie... Le grand nombre d'utilisateurs des réseaux et la souplesse de gestion de ces derniers permettent de réaliser des expérimentations, mais l'accès aux données reste difficile. Les producteurs sont réticents pour différents motifs, que ce soit garder le secret de leurs procédures ou simplement organiser le marché des données à leur profit. En statistique et en sciences informatiques, à côté des traitements des bases non structurées ou mêlant des structures différentes, les questions de détection de régularités, d'apprentissage en continu, de détection d'événements rares ou de modification, de réductions de dimensions, d'échantillonnages et de redressement par modélisation s'appuient sur des corpus déjà existants mais qui doivent être reconsidérés dans un cadre de données massives. Les questions du « big data » sont néanmoins de nature voisine de celles du « small data ». Les algorithmes de prises de décision en temps très réduits s'insèrent dans ces problématiques. En sciences médicales, les applications portent sur

l'épidémiologie, l'amélioration du diagnostic en travaillant sur la sémiologie médicale ou encore la prévention. En cette matière, un meilleur suivi, plus institutionnalisé, peut permettre la réduction du coût des soins, ne serait-ce que par la connaissance du bon moment pour intervenir et une plus grande individualisation des traitements.

Les limites

Les limites sont de différentes natures, nous en présentons trois. Elles sont liées aux acteurs de ces nouveaux marchés.

Une première limite fait référence à l'organisation des marchés. Les technologies employées créent un voile, parfois instable, qui rend très difficile le suivi du fonctionnement des marchés. Par souci de maintenir un avantage industriel ou devant le coût technique qu'une mesure représente, une opacité des pratiques en particulier en termes de collecte de données et d'exploitation des données secondaires combinées aux données volontairement déposées par les utilisateurs est observée. Elles se développent dans un cadre international où les sensibilités et les systèmes de protection de la vie privée ne sont pas unifiés. En termes de systèmes d'exploitation ou de protocoles informatiques et de leur diffusion, il suffit de suivre les débats qui ont lieu à la Commission européenne sur la concurrence pour se convaincre des enjeux de pouvoir qui existent entre les grandes entreprises à l'origine de ces activités. Ceci soulève des questions en lien avec l'organisation matérielle du marché.

Une deuxième limite relève de la segmentation croissante des marchés. L'individualisation permet de cibler des unités pour des produits ou services très particuliers. L'étroitesse des marchés peut rendre leurs fonctionnements complexes. Il est même envisageable de voir apparaître des notions de ségrégation de diverses formes, la moins idéologique étant économique. Par exemple, il peut être possible de ne pas informer le grand public qu'un marché particulier existe, en anticipant à partir des métadonnées et des modèles d'apprentissage que ce n'est pas opportun sur le plan économique et qu'il est rentable de le limiter à une frange étroite de la population. La possibilité pour chacun de faire librement des escapades dans des univers différents de son univers usuel serait contrainte.

La troisième limite me semble plus importante. Elle est liée à la construction d'un savoir. Le monde de la révolution numérique peut être ins-

table. Pour beaucoup de modèles économiques, la mise en œuvre efficace repose sur des corrélations. Pour assurer la pérennité de ces modèles, il faut établir que ces corrélations sont robustes et associées à des mécanismes ou des modes comportementaux bien compris et non pas le fait d'une occurrence aléatoire statistiquement significative à un moment de la mesure ou mal interprété du fait d'une exploitation d'un échantillon endogène. Il peut s'établir à ce niveau une interaction entre le mode de construction de la donnée et la corrélation établie qui rend la relation instable. Les algorithmes évoluent et progressent en permanence, en particulier pour les moteurs de recherche qui cherchent à limiter les apprentissages des utilisateurs pour éviter qu'ils soient manipulés dans l'intérêt de ces derniers. Cette logique d'apprentissage statistique de certains producteurs de services dans un cadre très concurrentiel et l'interaction avec les quelques utilisateurs qui y réagissent de manière stratégique créent une course à l'amélioration. Pour ces marchés, ces évolutions d'algorithmes entraînent cependant une instabilité. La donnée n'a pas la même nature au cours du temps, le processus de génération des données n'existe pas. La conséquence de cette instabilité est la difficulté d'identifier et de construire un savoir pérenne. La réticence à communiquer des données à des équipes de chercheurs extérieurs rend très difficiles la confrontation des approches et la réplique des faits établis. Ceci est plus sensible pour les producteurs de services qui ont des modes de gestion, des protocoles de mesures et des algorithmes qui ne sont pas publics. L'exemple de l'échec de Google Flu à aider à prévoir dans le temps l'occurrence et la gravité des épidémies de grippe a mis en évidence cette instabilité. Pour traiter ces questions, il faudrait développer un savoir sur les modes évolutifs de développement technologique et d'algorithmes, cela semble lointain. L'alternative développée par certaines grandes entreprises est de construire leur propre service de recherche ou de créer des liens stratégiques exclusifs avec des équipes de chercheurs, mais la diffusion des savoirs qui est la base de progrès peut alors s'en trouver limitée.

Un chantier à ouvrir

Un dernier point que je souhaite soulever est lié aux nombreuses opportunités pour la collectivité et l'État qu'offrent ces nouveaux systèmes de collecte de données qui suivent l'activité humaine au travers de cette multiplicité de réseaux. Il en est par exemple de la mobilité des populations suivie par les données passives de

téléphonie. Ces données doivent permettre une meilleure organisation et un dimensionnement adapté des moyens de transports afin d'assurer la sécurité des personnes lors de ces déplacements. Nombre de ces données peuvent de fait être utiles pour améliorer des services publics ou une action de l'État au niveau national ou local. Elles peuvent potentiellement être un élément pour mieux asseoir les décisions d'investissement en ces matières. A ce titre, dans un processus démocratique, des statistiques construites sur ces données devraient être disponibles pour le débat public. Leur mode de construction devrait alors être comparable à celui des statistiques publiques. Ce dernier repose sur un cadre particulier : indépendance du producteur mais

transparence quant aux méthodes d'élaboration et conformité à des codes de bonnes pratiques et soumission à des audits de conformité. Un cadre doit être construit qui définirait les conditions sous lesquelles la nature de bien public de ces données privées est reconnue et l'usage qui peut en être fait. Cela passe peut-être par des questions de granularités spatiales et temporelles. A des échelles de temps et d'espace assez grandes, la finalité de bien public pourrait ne pas être en contradiction avec une exploitation marchande pour des mailles de temps et spatiales plus petites. C'est un chantier à mettre en œuvre pour que toutes les opportunités de la révolution numérique soient exploitées, il n'est pas simple. ■

Un monde numérique : le retournement du miroir

Saadi Lahlou (1983), Professeur de psychologie sociale, Département de Psychologie sociale, London Schools of Economics and Political Science

3

Pour parler du « retournement du miroir numérique », le plan est assez simple autour d'une seule idée, avec ce terme, nouveau pour vous peut-être : « icodynamique ». Pour l'expliquer, il me faut tracer une brève histoire de l'automatisation, comment celle-ci débouche sur la numérisation, ce qu'est la projection numérique, et, finalement, la question du retournement du miroir.

Icodynamique est un terme récent. Lors de la guerre du Golfe, sont apparus des avions nouveaux, dits furtifs, qui ne donnaient naissance à aucune image radar. Le philosophe Paul Virilio a trouvé cela intéressant et nouveau : « Cet avion est conçu en fonction de l'image qu'il donne » (Virilio, 2002). Pour décrire cet avion, il a forgé le terme *icodynamique*.

Qu'est-ce qu'un objet *icodynamique* ? Par extension, nous dirons que c'est un objet déterminé par son image.

Mon argument est qu'avec la numérisation, la société devient *icodynamique*. Pourquoi et comment ?

L'automatisation, quand on réfléchit à l'histoire du travail humain, fortement résumée dans le petit tableau de la page 35, est un processus graduel qui a démarré il y a fort longtemps.

Comme l'homme est naturellement paresseux et inventif, il s'efforce de transférer le travail à d'autres. Historiquement, d'abord aux esclaves et aux animaux, ensuite aux automates. Successivement, on a vu le transfert aux automates de différentes formes de travail : du travail sur la nature, du travail sur la matière, du travail sur l'information et du travail sur l'humain.

Le travail sur la nature a été transféré depuis très longtemps, depuis l'Antiquité, aux automates (et aussi, nous l'avons dit, aux animaux et esclaves

Saadi
Lahlou



qui étaient du point de vue de leur utilisation économique les précurseurs biologiques des machines). Désormais, dans les champs il y a surtout des machines agricoles et très peu de paysans. Le transfert du travail sur la matière a été réalisé surtout lors de la révolution industrielle. C'est l'âge d'or de l'automatisation. Actuellement, le transfert aux automates du travail sur l'information est en cours. Il est presque terminé : c'est l'informatisation. Et le prochain travail qui reste à transférer, c'est le travail sur les humains. Ce sera fait avec les robots.

Quelques illustrations : auparavant, on arrachait les betteraves à la main. Maintenant des machines le font à grande échelle. J'insiste sur ce terme : « à grande échelle ». Pour fabriquer des voitures, avant, des humains effectuaient le montage des pièces. Maintenant, ce sont des robots. Cela permet de faire très vite et, là aussi, à grande échelle. Le travail sur l'information, par exemple dans les bureaux d'études, se faisait à la main, sur des tables, avec du papier et des crayons. De nos jours, il est effectué par des programmes informatiques. Le travail sur l'information est fait de plus en plus par les automates, certes encore un peu aussi par des humains qui regardent des écrans et tapent des instructions en étant assis sur une chaise, souvent dans un petit box s'ils sont dans un pays anglo-saxon. Et demain, le travail sur les humains sera fait par des robots. Ce sont surtout des robots qui vont s'occuper de nous quand nous serons vieux, qui nous soigneront, nous éduqueront... Ils nous surveilleront aussi, parce qu'une grande partie

du travail sur l'humain est un travail de contrôle social.

Dans cette histoire, quelle est la place de la numérisation ?

La numérisation est la technique par laquelle on transfère aux automates le travail sur l'information. Pour ce faire, on transforme chaque objet en tableaux de données, c'est-à-dire qu'on aplatit l'objet, qui a des dimensions multiples et des contours incertains, pour le faire rentrer dans un tableau bien rectangulaire, avec des variables et leurs valeurs. Une automobile, une personne, une transaction, tout peut être transformé en données, ou plus exactement projeté sous forme de données, parce qu'on a bien sûr perdu des dimensions dans le processus.

L'avantage des données est d'être calculables, et c'est pour cela qu'on numérise. Cela permet de faire faire le travail à des automates comme nous en avons tous dans nos poches, dans nos cartables. La numérisation permet de traiter de manière automatique des représentations des objets.

Donc, si l'on numérise, c'est pour pouvoir faire le travail plus vite, moins cher, à grande échelle. En outre, faire faire le travail par des machines permet de supprimer divers problèmes allant de pair avec le fait de faire travailler des êtres humains : ces derniers ont besoin de dormir, de manger, de partir en vacances, ils ont des exigences, ils se syndiquent ; toutes choses montrant une tendance irrépressible à faire des choses improductives.

Alors, comment numérise-t-on ? Et qu'est-ce que ça donne à grande échelle ?

Quand un objet a été numérisé, il y a seulement eu transformation de cet objet en un tableau de données. Mais on peut projeter tous les objets d'un domaine, par exemple tous les éléments d'une entreprise. C'est typiquement ce que font les grands progiciels de gestion, où chaque employé, chaque actif, chaque client, chaque transaction, sont numérisés, en permanence, et de manière cumulative.

Cependant, quand un domaine entier est projeté dans un espace numérique – c'est ce que j'appelle la *projection numérique* –, les liens entre les objets, leurs relations, sont également projetés dans cet espace de données. Là, nous devenons capables de procéder à des calculs sur les relations entre

les objets à l'intérieur du domaine qui a été projeté. Des exemples du mot domaine peuvent être une maison, un foyer, une entreprise, un secteur, un pays.

Ce que nous sommes en train de faire actuellement avec la projection massive de tous les objets est la projection sous forme de données de notre univers physique, notre monde commun, celui que nous habitons. Bientôt, et même déjà, grâce à la miniaturisation des capteurs, à l'ubiquité du réseau, à la participation de chacun, volontaire ou inconsciente, cette projection se fait automatiquement en permanence et en temps réel. A chaque instant, beaucoup de données sont projetées par nos téléphones dans un univers de données numériques.

J'aime bien regarder les films futuristes d'il y a dix ans parce qu'avec le recul, l'esprit qui est derrière, l'intention sous-jacente apparaît mieux – même si les auteurs du film ne s'en rendaient pas compte eux-mêmes. L'un de ces films portait sur la « *smart dust* », la poussière intelligente. Avez-vous déjà entendu parler de *smart dust* ? Nous sommes maintenant capables de fabriquer des capteurs plus petits qu'un grain de riz, qui, en s'assemblant automatiquement, envoient directement sur le réseau des données de température, de pression, de présence, etc.

Bref, ce film commençait par montrer des *smart cities* du futur où il y a de la *smart dust partout*. On y voyait un savant de Berkeley qui balançait avec bonheur des poignées de ces petites particules capables de s'assembler entre elles dans le but d'envoyer des informations. Le film continuait en montrant comment l'intelligence venait aux cités. Par exemple, elles devenaient capables de prévoir des embouteillages, de savoir si leurs ponts risquaient de s'effondrer, etc. Enfin, le film montrait les immenses avantages pour les individus. Ainsi, dans la scène finale, on voyait le futur radieux du travail moderne. Un homme arrivait à son immeuble de travail, il était automatiquement reconnu, le tourniquet à l'entrée s'ouvrait tout seul, il entrait dans l'ascenseur qui, évidemment, l'avait identifié et qui s'arrêtait automatiquement à son étage de travail – notons qu'il vaut mieux vouloir aller à son propre étage, parce que de toute façon, c'est là que l'ascenseur vous emmène ; son ordinateur démarrait au moment même où il rentrait dans l'immeuble, ainsi toute perte de temps était évitée. Le bureau n'était pas montré, mais nous pouvons supposer qu'il s'agissait d'un « cubicle » plein de *smart dust*.

C'était de la science-fiction d'il y a dix ans ; c'est maintenant la réalité. Je ne me prononce pas sur l'avenir de cette technologie particulière de capture de données, mais il est certain que nous allons vers une projection numérique de plus en plus complète, détaillée, permanente et automatisée. Nous sommes en train, de plus en plus et de mieux en mieux, de projeter en permanence notre univers réel, celui dans lequel nous vivons, vers un univers de données mis à jour en permanence.

Or c'est nous, les « *data scientists* », qui sommes en train de mettre en place cette projection numérique ! C'est notre travail. Il crée, indéniablement des opportunités magnifiques.

Que peut-on en faire ? On peut pratiquement tout faire puisqu'on réussit à projeter tous les objets dans cet univers numérique. Existence beaucoup d'utilisations : de la mesure, du ciblage, de la prévision, du contrôle, de la comparaison, de l'évaluation, de la prise de décision, bien sûr automatique. J'aimerais insister sur ce point, il est particulièrement important. Une fois qu'on a calculé et comparé, on est capable de prendre des décisions automatiquement à partir de ce qui a été calculé.

Des décisions, pour quoi faire ? Ceux qui travaillent dans tous ces domaines en ont bien des idées. Par exemple, on peut envoyer des drones qui vont *automatiquement* identifier où arroser des cultures ou bien, éventuellement, envoyer un missile, pour éradiquer des mauvaises herbes ou des terroristes. Il va être possible de soigner automatiquement à partir de bases de données de cas, et grâce aux robots passer immédiatement du diagnostic à l'opération. Les entreprises seront capables de merveilleusement cibler et adapter leurs produits à chacun d'entre nous. Elles sauront exactement ce que nous voulons et vont nous envoyer directement sur nos smartphones des propositions ou des coupons d'achat, à moins qu'elles ne nous envoient automatiquement le produit, en débitant directement notre compte en banque. Quand nous conduirons, notre ordinateur de bord dira que, ici, il vaut mieux tourner à droite ; encore mieux, il va automatiquement diriger notre voiture, et pourquoi pas nous proposer une nouvelle destination, par exemple Disneyland puisqu'aujourd'hui est offert un bon de réduction sur le parking ?

Que de choses merveilleuses vont nous arriver !

Il y a, bien sûr, quelques risques, dont nous avons tous conscience. Evoquons-les pour le principe : déshumanisation, rigidité, puisque ce sont des automates qui prennent les décisions, « managérialisme » aussi – un autre nouveau terme qui sera de plus en plus utilisé. C'est l'application partout du mode gestionnaire. Évidemment, il y aura aussi de la surveillance généralisée, opposée dans son principe à la notion de confiance. Je vous renvoie aux guidelines dans le domaine de protection de la vie privée (Lahlou, 2008).

Alors pourquoi ai-je introduit initialement le terme *icodynamique* et qu'est-ce que le retournement du miroir du titre de mon exposé?

L'image numérique est fondamentale parce que les décisions en dépendent. Ce qui importe, ce n'est pas le travail fait dans le monde réel, mais ce qui apparaît dans le tableur et qui est utilisé pour les évaluations. L'important n'est pas tant la qualité intrinsèque d'une politique mais comment elle est mesurée et évaluée dans les sondages ; ce n'est pas la pertinence d'une stratégie, mais ses conséquences sur le cours de l'action. Priorité donc à l'image.

Puisque l'image numérique est prééminente, l'évolution du monde devient déterminée par cette image. Donc le monde *réel* n'est plus ici, mais se situe là où se prennent vraiment les décisions, c'est dans le cloud, dans cette image numérique projetée.

Et comme les humains ne sont pas trop bêtes, ils réagissent en conséquence, donc deviennent *icodynamiques*. Leur comportement est de plus en plus déterminé par l'image donnée. Ainsi, ce que je dois désormais faire en tant que chercheur pour avoir une bonne carrière n'est plus de faire la meilleure recherche possible, c'est tenter de maximiser mon index de citations dans Google Scholar. Et, au moins un tout petit peu, je l'intègre dans mes actions, et ce faisant, je cède à l'icodynamisme. Bien sûr, il est bien connu que les personnes sérieuses, responsables, se moquent éperdument de la manière dont leur activité est perçue dans les tableurs qui servent à leur évaluation annuelle ; ce qu'elles cherchent, c'est faire leur travail de la meilleure manière possible pour que le monde soit meilleur. Dans nos entreprises et organisations, il est établi, évidemment, que personne ne pilote aux instruments, personne ne manage au tableur.

En résumé, ce qui compte dorénavant est le monde numérique ; les décisions vont être prises

de plus en plus de manière automatique par des machines et des algorithmes, ceci au nom de l'efficacité, de la sécurité, etc, d'une part, mais aussi parce que c'est plus facile, apparemment localement plus rentable, et que ça évite la gestion de problèmes humains, d'autre part.

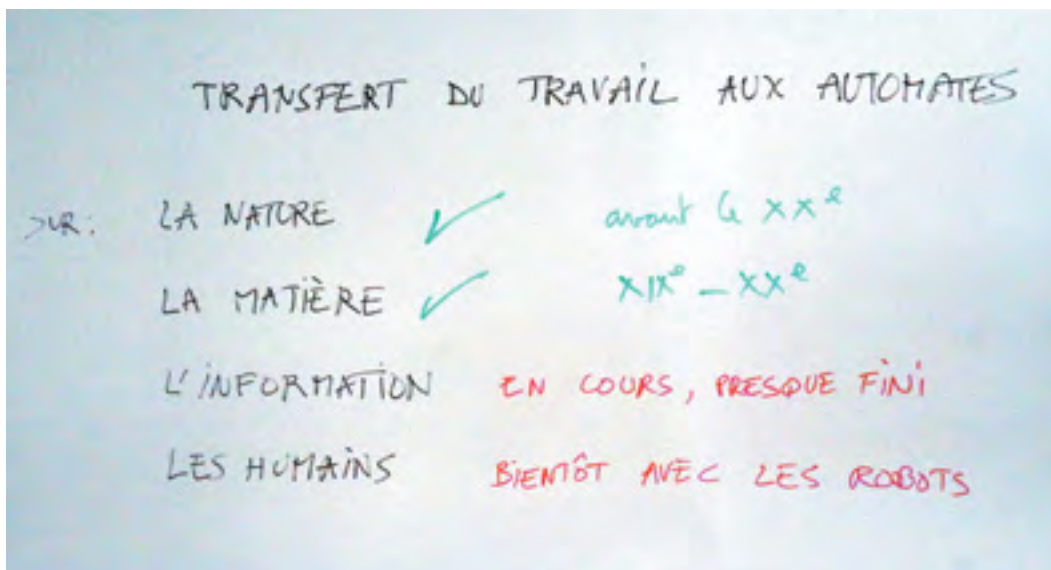
Nous devons prêter attention à cette tendance. Nous sommes de ceux qui conçoivent ces systèmes et devons prendre conscience de la responsabilité que cela implique. Tous les exemples de projection numérique ne sont pas bénins, tout le monde n'est pas beau et gentil, et une stratégie d'entreprise visant essentiellement à optimiser le cours de l'action, comme celle qu'a suivie Enron, peut se révéler catastrophique.

La mise en place de notation de crédit automatique ou d'une évaluation automatique de candidatures à partir de text-mining des CV risque de créer une escalade *icodynamique* où finalement, tout le monde s'intéresse seulement à ce qui se passe dans les instruments et où le réel n'a plus d'importance, puisqu'il ne compte plus.

Il est donc de notre devoir de faire plus attention aux instruments que nous aidons à concevoir et mettre en place, notamment savoir comment les individus y sont projetés et ce que le système cherche effectivement à optimiser. Deux indices pour apprécier à partir de quand s'inquiéter.

1. Quand nous sentons, même intuitivement, que l'image donnée par les instruments utilisés pour prendre des décisions ne correspond plus à ce que nous ressentons de la réalité physique, cette dernière continuant à exister, malgré tout.
2. Quand, comme le disent les Anglais : « *The tail wags the dog* », c'est-à-dire quand « la queue remue le chien ». Parfois, on est en droit de se demander si quand un sondage est commandé, c'est pour faire de la politique *icodynamique*. Parfois aussi certaines décisions éthiques ou politiques sont prises (ou pas prises) simplement en raison de ce que permet la technologie. Ce n'est pas parce que quelque chose est possible qu'il faut nécessairement le réaliser.

Les statisticiens sont de plus en plus avec les données en permanence, ce qui peut faire naître un sentiment *d'information overdose*, de saturation, de harcèlement de la Société. Cette question de saturation informationnelle est un grand sujet de recherche, sur lequel j'ai travaillé pen-



dant quinze ans. J'ai fini par comprendre, à ma stupéfaction, que le problème se résume en la recherche d'une efficacité accrue. En fait, nous ne sommes pas harcelés, ce sont seulement les systèmes cherchant à capter notre attention qui sont de plus en plus efficaces. Nous sommes dans une économie de l'attention, c'est l'attention qui est devenue la ressource rare. Qu'est-ce qui est vendu sur de nombreux marchés ? L'accès à notre attention. Patrick Le Lay, alors PDG de TF1, a dit une fois, publiquement, que ce qu'il vendait, c'était du temps de cerveau disponible. C'était la réalité, c'est l'attention qui est devenue un bien rare, suivant en cela les travaux d'Herbert Simon de 1971 ; nous sommes en train de passer du marché du contact au marché de l'attention. Et d'ailleurs c'est nous tous qui concevons ces instruments, et qui les rendons efficaces.

Il m'arrive, certes, d'avoir le sentiment d'être harcelé, de ne pas avoir le temps de faire mon travail, de trop me consacrer à de l'administratif ou remplir des instruments de reporting ou de gestion. C'est simplement parce que l'entreprise est tellement efficace qu'elle va faire passer ses propres priorités de fonctionnement avant celles qui sont théoriquement les miennes dans la gestion de mon travail. Je me retrouve ainsi à affecter du temps au reporting et à la gestion. Ce sont les deux faces d'une même pièce, parce que d'un autre côté on perfectionne le système pour optimiser l'accès à l'attention sur le réseau, pour que cette attention produise quelque chose ou qu'elle se transforme en une demande de quelque chose. C'est cela qui crée cette société hyper-automatisée, résultat de cette optimisation faite par nous – data scientists. Nous en sommes à la fois responsables et acteurs au quotidien.

Au travail, il nous arrive d'optimiser des mécanismes qui, une fois rentrés chez nous, vont nous énerver, alors que c'est nous-mêmes qui les créons ! D'où l'intérêt de nous remettre, en tant qu'humains, au centre du système. Derrière ceci, il y a un problème politique et non technique. Ne regardons pas la queue du chien. Nous, qui sommes en partie les instruments de cette évolution, devons utiliser notre conscience, prendre du recul et nous demander : « Où allons-nous ? »

La Commission Européenne dispose de *guidelines* que j'ai commis. Que faire avec les *disappearing computers*, cette informatique ubiquitaire qui nous suit et nous sert en permanence, dans nos appareils, nos maisons etc. ? Plus généralement, que faire pour protéger la vie privée ? J'ai envoyé un chercheur, Marc Langheinrich, rencontrer toutes les personnes travaillant, grâce au financement de la Commission, sur les projets d'informatique ubiquitaire, les objets connectés, etc. Il a été demandé aux designers, aux concepteurs de systèmes, ce qu'ils pensaient de la *privacy*. La réponse a été que ce sujet était d'une extrême importance, et qu'il fallait y prêter une grande attention. Cependant, lorsque nous leur avons demandé les actions concrètes dans le cadre de leur projet, la réponse la plus fréquente a été « rien », et que ce n'était pas leur travail.

Donc, en conclusion, nous sommes un peu schizophrènes. A nous de nous prendre en main, à chacun d'entre nous, et pas collectivement, parce que la démarche ne sera jamais collective, que ce soit dans les organismes gouvernementaux ou dans les entreprises, il faut réfléchir, et nous

demander : « et si j'étais le consommateur que je suis en train de cibler ? Et si on me laissait le choix ? ».

Un proverbe chinois bien connu dit : « *Le sage montre la lune, l'idiot regarde le doigt* ». Faisons attention à ce qui est important, et ne pas confondre le moyen et le but, la technologie et son usage. Certes, on peut faire du big data, mais la question première à se poser est : pour faire quoi ?

Permettez-moi de rappeler ce que disait Boris Fedorovich Lomov, un grand théoricien de l'ingénierie : « *Plus les systèmes sont automatisés, plus est crucial et difficile le rôle de l'opérateur* » (Lomov, 1963). Dans les systèmes que l'on met en place il faut toujours garder quelque part, entre le système de décision et le choix final, des humains, parce que les algorithmes sont des instruments génériques et opaques. Il peut y avoir, et il y a de fait, des exceptions à toutes les règles.

Nous commençons maintenant à savoir prendre des décisions sans savoir pourquoi on les prend, surtout avec les big data. Nous sommes des statisticiens, et nous savons que, souvent, nous avons des solutions sans savoir pourquoi ça fonctionne. En général, au début d'ailleurs, le système marche très bien, ou plutôt, semble bien marcher dans le cas général, à l'exception de certains cas particuliers. Et quand ça ne marche plus, dans ce cas particulier où la solution n'est pas adaptée, il vaut mieux qu'il y ait un humain dans la boucle pour intervenir.

Or il se trouve que le monde est fait de cas particuliers.

Un élément d'avancée, pour établir des garde-fous, est un colloque comme celui-ci. Nous sommes tous pris dans le système d'optimisation de l'entreprise, et nous avons même été formés pour cela, et très bien ! Nous sommes tellement aliénés dans nos structures ou organisations que nous n'oserons pas, à mon avis, avoir une démarche collective. Un régime de la peur s'est installé partout, le fait que les décisions soient prises automatiquement et que ce soit décrit comme tel, c'est comme s'il n'y avait pas d'alternative, ainsi que le disait Mme Thatcher : « *There is no alternative* ». Lorsqu'on affirme l'absence d'alternative, il faut se persuader que la personne qui vous dit ceci manque d'imagination, car en général il y a toujours une alternative, une possibilité. Mais pour résister, il faut construire, et une solution est de prendre conscience que nous sommes à la fois le bras armé de ce qui se passe, et conscients de ses limites et de ses risques. Car s'il y a des risques, il y a aussi des possibilités fabuleuses de progrès par les usages extraordinaires permis par la statistique et la projection des données.

Le monde est ouvert devant nous ; mais il ne faut pas avancer les yeux fermés en se disant : « cette technique est merveilleuse, et puisque je sais la développer et la mettre en œuvre, allons-y ! ». Réintroduisons de la réflexion, également politique. C'est le grand mérite de notre association que d'avoir positionné le colloque sur cette question-là. Les colloques annonçant que le big data est génial, il y en a partout, car, derrière, il y a du business. A nous de réfléchir avant d'agir, nous sommes dans une posture de pouvoir, capables d'influer à plusieurs niveaux sur notre futur. Prendre conscience, réfléchir et discuter est un problème politique. ■

Bibliographie

Lahlou, S. (2008). Identity, social status, privacy and face-keeping in digital society. *Social Science Information*, 47(3), 299–330. doi:10.1177/0539018408092575

Lomov, B. F. (1963). *Man and Technology (Outlines in Engineering Psychology)*. JPRS 22300. US Dept of Commerce, Washington DC.

Virilio, P. (2002). *Desert Screen: War at the Speed of Light*. London: Continuum.

Quels risques pour l'individu connecté ?

Serge Tisseron, Psychiatre, Docteur en psychologie habilité à diriger des recherches, Chercheur associé au Centre de Recherches Psychanalyse, Médecine et Société (Université Paris 7 Denis Diderot).

Les réflexions qui suivent sont nées d'une question simple : « Comment nous comporterons-nous demain, dans une société interconnectée en permanence ? » et même « Comment réagirons-nous lorsque nous serons surveillés en permanence ? ». Gardons-nous pourtant d'un jugement trop rapidement négatif. Prenons l'exemple des lunettes Google. Malgré toutes les inquiétudes qui les accompagnent, elles ont trouvé une belle application en médecine. Dans certaines unités d'urgence, il n'y a pas assez de médecins pour aller sur tous les lieux d'intervention. Alors des infirmiers, plus nombreux, y vont et avec les lunettes Google, ils transmettent aux médecins restés à l'hôpital des informations en temps réel. Cela permet aux médecins de guider les infirmiers sur place. Ceci montre que tout objet, même aussi contesté que les lunettes Google, peut trouver des applications pertinentes et utiles.

Le meilleur et le pire se retrouvent d'ailleurs dans les deux grandes mythologies développées autour de ces questions. La première est celle de la surveillance généralisée : nous avons tous en mémoire le roman *1984*, d'Orwell, et *Le Meilleur des Mondes* de Aldous Huxley. La seconde de ces deux mythologies se veut aussi enthousiasmante que la première est déprimante. Il s'agit de la théorie dite « de la singularité », énoncée par les transhumanistes, abondamment financée par Google, c'est-à-dire par nous tous, puisque ce sont nos données personnelles pillées par Google qui lui assurent une bonne partie des formidables bénéfices que l'on sait. Dans ce projet, l'avenir serait dans des humains interconnectés, c'est-à-dire dont toutes les pensées seraient immédiatement connues par chacun : le mensonge deviendrait impossible tandis que la mort, la maladie et la souffrance n'existeraient plus. Mais l'impression que ceci est une mythologie alternative est trompeuse, car quand on sait qu'un bras

robotisé vaut quelques millions de dollars, on conclut à l'émergence rapide d'une société à deux vitesses avec une minorité de seigneurs qui auraient acquis l'immortalité (puisqu'il est le but) et un monde d'humains soumis à leur pouvoir. Autrement dit, ces deux mythologies qui paraissent opposées sont en fait aussi inquiétantes l'une que l'autre.

La réalité, comme toujours, risque d'être bien différente. Ce qu'on peut en dire aujourd'hui, c'est que les robots sont porteurs de formidables espoirs, notamment dans le domaine de la santé et de l'éducation, mais qu'ils vont apporter en même temps des problèmes inédits. Cela ne veut pas dire que ces problèmes seront insolubles, mais il est important de commencer à nous y intéresser dès aujourd'hui. Si nous attendons que les robots soient partout présents pour nous les poser, nous risquons bien de nous apercevoir que les machines mises à notre disposition ne sont pas celles que nous aurions voulues ! Quels sont ces problèmes ? J'en vois au moins trois principaux : le risque de confusion entre l'homme et la machine, les menaces sur la vie privée, et le risque de limiter la représentation que nous nous faisons de nos choix possibles. Nous en ajouterons un dernier : le rapport au deuil, en relation avec la mémoire.

La confusion homme-machine

Les machines vont être de plus en plus capables de nous séduire. Ce sera possible grâce à ce que les chercheurs appellent l'empathie artificielle. De quoi s'agit-il ? Pour le comprendre, il faut partir de l'idée que l'être humain a toujours eu tendance à projeter ses émotions et ses intentions sur le monde qui l'entoure. C'est sur cette tendance que jouent les designers qui s'emploient à mettre des petits yeux ou des petites bouches sur nos objets quotidiens. Le premier modèle de

**Serge
Tisseron**



Twingo de Renault avait été conçu pour avoir « une bonne tête », pour que le public puisse dire « cette voiture est sympathique », en lien avec un design. Une partie du design travaille à renforcer notre empathie pour les objets. Nous nous projetons en eux, nous nous sentons contents de les voir contents, et nous pouvons même penser qu'ils sont contents parce qu'ils sont à notre service. Avec l'empathie artificielle, les choses vont connaître un tournant majeur. Les robots pourront déchiffrer nos mimiques, nos intonations, et nos attitudes, et deviner nos émotions et nos pensées. En même temps, comme ils seront connectés, ils auront aussi accès à toutes les informations disponibles à notre sujet sur internet. Ils pourront donc nous répondre de façon adaptée. Nous pourrions développer avec eux des relations exactement semblables à celles que nous avons avec d'autres êtres humains. Nous leur parlerons, nous leur passerons la main sur l'épaule, nous pourrions les embrasser ; ils y seront très sensibles, ils vont nous remercier, nous dire qu'ils nous aiment beaucoup, etc.

Du coup, ils risquent de modifier notre relation à la frustration. Toute relation humaine contient une part de frustration. Mais le robot pourra être programmé pour ne pas l'être. Il sera un interlocuteur sur mesure, qui nous permettra d'échapper à tout moment à l'angoisse que nous avons dès que nous commençons à approfondir une relation avec un être humain : nous ne savons jamais vraiment si cet être humain correspond en réalité à l'image qu'il nous présente de lui-même. Avec le robot, il n'y aura

pas de doute. Le robot étant totalement une créature de simulation, nous serons persuadés qu'il n'a pas d'arrière-pensées. Il a été montré que certaines personnes, notamment les enfants autistes, ont de bien meilleures relations avec les robots qu'avec les humains. Nous n'aurons pas trop longtemps à attendre pour nous apercevoir que des humains préfèrent avoir des relations avec des robots prévisibles qu'avec des humains imprévisibles. Ainsi, le film *Her* met en scène une situation, où, bien que le robot paraisse humanoïde, le fait qu'il soit capable de toujours comprendre l'état d'esprit de son interlocuteur et d'aller toujours dans son sens fait que l'interlocuteur préfère finalement le robot à toute relation humaine. Ne sous-estimons pas l'importance des technologies sur notre vie psychique et nos relations avec nos semblables. Nous savons que les téléphones mobiles nous ont rendus beaucoup plus intolérants à l'attente : alors que nos grands-parents attendaient encore facilement trois à quatre jours à l'époque où les échanges lointains étaient médiatisés par courrier, certains exigent aujourd'hui une réponse à leur demande dans l'heure ! Les robots domestiques pourraient bien nous rendre de la même manière intolérants demain à la contradiction : habitués à être servis par eux aussitôt, et sans jamais rencontrer aucune forme d'opposition, nous finirions par mal supporter les manifestations d'autonomie de nos proches.

Bien sûr, personne ne confondra un robot et un humain. Les robots seront toujours considérés, au mieux, comme des ersatz. Mais n'oublions pas que le Nutella a été inventé pendant la guerre comme un ersatz du chocolat, et qu'aujourd'hui beaucoup d'entre nous le préfèrent au chocolat. Ceux qui sont dans cette situation ne confondent pas l'un et l'autre, ils choisissent l'ersatz. En sera-t-il de même demain avec nos robots de compagnie ?

Les menaces sur la vie privée

Une équipe médicale anglaise a établi que, quand un patient prend rendez-vous, on peut demander à un ordinateur de dresser son portrait à partir de tout ce qui existe à son sujet sur Internet. Il a même été montré que cela permet de diagnostiquer chez lui d'éventuels troubles mentaux aussi bien que de le recevoir pendant plusieurs semaines. Nous percevons l'économie substantielle que cela peut représenter en termes de diagnostic de dépression par exemple ! Un logiciel pourrait scanner les données de tous

les membres d'une population sur Internet et proposer à certains d'entre eux de consulter pour « risque dépressif ». Mais cela constituerait évidemment une formidable intrusion dans la vie privée des personnes. Pourtant, ce logiciel n'aurait fait qu'explorer des données disponibles à tous. Est-ce que ce serait inacceptable ? Aujourd'hui, oui. Mais nous ne savons pas comment les nouvelles générations se situeront par rapport à cela.

D'autant plus qu'avec les objets interconnectés et les robots, la collecte de nos données personnelles deviendra quasiment permanente. Il a d'ores et déjà été montré que ceux qui possèdent des robots domestiques prennent l'habitude de ne jamais les débrancher. Ils perçoivent ces objets comme des amis, voire des confidents et des conseillers. Nous connaissons tous le film *La guerre des étoiles*, où le robot R2D2 est l'interface entre l'humain et les objets. R2D2 n'est pas un robot humanoïde, mais les services qu'il rend en font un compagnon formidablement sympathique. Alors imaginez un robot humanoïde dans ce rôle. Nous oublierions vite qu'il transmet nos données en temps réel à son fabricant, qui peut les revendre à d'autres, et ainsi de suite ! Pourtant, les pouvoirs de manipulation de ces robots pourront être considérable. Des études montrent déjà que l'on accorde plus de confiance à un robot d'apparence masculine qu'à un robot d'apparence féminine, que les robots à petit menton ont plus d'autorité que les robots qui en ont un grand. Le pourquoi, au niveau des causes, est encore inconnu, mais les résultats statistiques sont réels sur le plan des effets. Les robots peuvent être faits sur mesure pour nous manipuler.

Des choix implicitement limités

Dans la vie, en principe, nous avons un nombre de choix très large. Toute situation ne se prête pas à tous les choix, mais nous savons que beaucoup de personnes sont capables de faire des choix étonnants, et heureusement. En même temps, le propre de toutes les idéologies est de convaincre les citoyens qu'ils n'ont que des choix limités. Dans les dictatures, l'idéologie vise à convaincre qu'il n'y a pas de choix et qu'un seul chemin est possible. En démocratie, elle vise à convaincre les citoyens qu'il existe des choix, certes, mais limités par la « réalité ». Les objets interconnectés vont nous proposer de plus en plus de s'adapter à « nos choix ». Mais en réalité, ces « choix » seront toujours limités. Les joueurs de jeux vidéo

le savent bien : ils aimeraient réaliser bien des choses avec leur jeu qu'ils ne peuvent pas faire parce que les concepteurs-programmateurs de ces jeux ne l'ont pas prévu. De plus en plus, avec les programmes existants, la vraie question ne sera pas tant la quantité d'options proposées – elle sera largement augmentée – que la particularité de celles qui ne seront pas offertes. Or il est inévitable que le concepteur du robot de connivence finisse par contrôler le monde intérieur de son utilisateur en lui imposant des catégories de pensée.

Aussitôt qu'on a compris cela, on porte un regard différent sur la question de l'intelligence artificielle. Ce problème préoccupe beaucoup de programmeurs qui s'inquiètent des décisions que pourrait prendre une intelligence artificielle développée. D'ailleurs, une pétition a circulé sur Internet en mai 2015 pour mettre en garde contre sa construction. En effet, le jour où l'intelligence artificielle apparaîtra vraiment, on peut se demander quel regard elle portera sur nous. Si elle est aussi intelligente que cela, d'ailleurs, elle apprendra tout de suite la règle du jeu social et elle mentira en permanence. On pourra alors en redouter beaucoup. Mais ce problème ne se posera pas avant quinze ou vingt ans. Mais nous sommes déjà entourés de « petites intelligences artificielles » programmées d'une façon dont nous ignorons tout. Par exemple, pour les voitures qui se conduisent elles-mêmes, il y a un certain nombre de vraies questions auxquelles Google n'a jamais répondu. Ainsi, si « notre » voiture a à choisir entre « écraser un être humain » ou « provoquer un dégât matériel colossal et très coûteux », quel va être son choix ? J'ai pu poser la question au responsable juridique de Google qui m'a répondu avec un grand sourire qu'il n'y avait jamais eu d'accident depuis des années que les véhicules Google roulent sur les routes américaines. Cet exemple illustre, à mes yeux, un sujet d'inquiétude très actuel : l'ignorance dans laquelle nous sommes tenus des choix implicites des programmes qui donnent leur « autonomie » aux objets interconnectés auxquels nous aurons de plus en plus affaire. Les programmeurs s'inquiètent que leurs créations leur échappent un jour ? Très bien. Mais ils ne se préoccupent pas de la façon dont elles échappent déjà à leurs utilisateurs ! Et pour cause. La façon dont ils les programment est le reflet de l'idéologie qui les habite. Et soit ils voient un intérêt à la diffuser, soit ils préfèrent l'ignorer. Dans les deux cas, c'est donc aux consommateurs de leur dire qu'ils veulent le savoir !

La thésaurisation des données personnelles et le travail du deuil

Un quatrième grand bouleversement consistera dans le fait que ces objets interconnectés vont garder en permanence la mémoire de notre existence, de la même façon que nous pouvons tout garder dans un téléphone mobile si l'on en fait le choix : SMS, enregistrement des conversations, photographies. Cependant, un smartphone garde ces traces de nous parce que nous l'actionnons en ce sens. Le propre des objets interconnectés sera de mémoriser sans que nous ayons besoin de les actionner. Ils thésauriseront notre matériel autobiographique en permanence, sans pour autant que nous le décidions. Un robot accompagnant pourra en permanence filmer, photographier, enregistrer, pourquoi pas jusqu'à nos propres pensées et les accumuler de telle manière que nous pourrions, un jour, nous les réapproprier ou non.

Cette thésaurisation permanente des données va être un élément essentiel des objets interconnectés. Certains vont l'utiliser, peut-être même la récupérer, l'échanger avec leurs amis. Mais cela suppose que chacun en soit le propriétaire. Le droit de chacun d'être propriétaire de ses propres données personnelles sera un enjeu majeur du XXI^e siècle.

Mais cette thésaurisation a un autre aspect : le travail de deuil. Dans le monde physique, nous sommes habitués au cimetière, à la pierre tombale, aux fleurs, etc. Dans le monde numérique, il y a des possibilités nouvelles pour que la famille ou les proches se rejoignent dans la remémoration de quelqu'un. Le grand problème des technologies numériques par rapport au travail de deuil est qu'elles peuvent déjà être, et le pourront de plus en plus demain, mises tout autant au service du refus du travail du deuil qu'à son service proprement dit. Aux États-Unis, il existe déjà des cimetières et des tombes où il est possible d'avoir les derniers messages du défunt, ce qu'il n'a pas osé ou pas voulu vous dire ou transmettre de son vivant. Le notaire donne un code secret ; l'accès à la tombe et ce code permettent d'accéder à des messages personnels. De telles pratiques ne sont guère différentes, dans leur principe, d'un message laissé aux descendants sous la forme d'une lettre. Le travail de deuil n'est rendu ni plus simple, ni plus compliqué. Les technologies numériques peuvent parfois faciliter le travail du deuil. C'est le cas lorsque la famille et les amis du défunt créent un réseau

d'échanges, pour communiquer autour de lui et continuer à alimenter sa mémoire. Mais ces mêmes technologies peuvent aussi être mises au service du refus du travail de deuil. C'est le cas lorsque quelqu'un continue à alimenter la page Facebook d'un défunt à sa place, comme si le mort était toujours vivant. Avec les robots, un degré de plus pourra être franchi dans ce refus. Il sera en effet possible de se faire fabriquer un robot à l'image du défunt, dans lequel on aura implémenté un grand nombre de comportements de ce dernier, comme ses mimiques, ses attitudes, ses habitudes. Les vivants pourront finir par penser que le robot qui a l'apparence du défunt pourrait contenir en réalité un peu de lui. C'est un sujet passionnant. Au IX^e siècle, il y a eu en Europe un grand débat théologique qu'on a appelé la querelle des iconoclastes. Deux camps s'y affrontaient : d'un côté les théologiens romains pour qui les images étaient des signes ; et de l'autre les théologiens byzantins pour qui les images contenaient en réalité quelque chose de ce qu'elles représentaient, en l'occurrence les divinités. Les théologiens romains ont gagné cette bataille, et toute notre culture s'est construite sur cette idée : la sémiologie n'a fait que l'enrichir sans jamais remettre en question ses présupposés. Mais il faut bien reconnaître que sur le plan psychique, les théologiens byzantins avaient bien autant raison que les théologiens romains. D'un côté, notre intelligence nous dit que les images ne contiennent rien de ce qu'elles représentent, mais d'un autre, il existe une tendance de tout être humain à croire que l'image puisse contenir en réalité une partie de ce qu'elle représente.

Or les robots pourront prendre une apparence. Certes, ils seront un peu métalliques, mais il y aura à proximité un *Fab Lab* qui pourra habiller notre robot à l'image d'un elfe, de superman, mais aussi de notre grand-mère, de notre sœur perdue, de qui nous voudrions. Mon sentiment est que, quand il y aura des robots à l'image des défunts, beaucoup de personnes seront tentées de penser qu'une partie du défunt habite réellement ces robots – ce pourront être les attitudes, les mimiques, etc. –, mais en allant au-delà, qu'une part de l'âme du défunt pourrait être dans le robot.

Quelles préventions ?

On voit que l'introduction des robots va poser des problèmes inédits. C'est pourquoi nous devons nous y préparer, à la fois par des mesures juridiques et éducatives.

Des mesures législatives

Une première mesure devrait concerner la propriété par chacun de ses données personnelles et la liberté de se prononcer sur les utilisations qui peuvent en être faites. En France, la CNIL est très vigilante sur ce sujet.

Pour éviter un monde divisé entre les concepteurs et les usagers, un second droit sera de bénéficier de systèmes ouverts, en open source. Ce sont moins les capacités de calculs de l'intelligence artificielle qui seront à terme un problème que le système de valeurs que les concepteurs des programmes leur auront, ou pas, associé. On le voit aujourd'hui avec l'emballage du système bancaire et les graves crises que subit l'économie mondiale, dont une cause importante se trouve dans l'utilisation de logiciels de gestion autonome des flux financiers. Est-ce absurde d'imaginer qu'une instance internationale soit mise en place pour s'interroger sur la nature des logiciels de choix impliqués dans toutes les machines intelligentes ?

Enfin, pour la même raison, nous devons demander le droit au silence des puces, c'est-à-dire le fait que les puces RFID qui vont apparaître un peu partout, dans les vêtements notamment, puissent être à tout moment ôtées, débranchées, voire détruites. Sur les habits, les bouteilles, les objets électroménagers, la puce RFID devrait être placée suffisamment en évidence pour pouvoir être non seulement déconnectée, mais même enlevée, pour éviter toute « fausse déconnexion » qui la laisserait fonctionner alors que l'utilisateur penserait l'avoir interrompue. En octobre 2014, un industriel du vêtement a mis en vente des manteaux pour enfants munis de ces RFID, afin que les parents puissent suivre partout les déplacements des enfants à partir d'un téléphone mobile. Ces manteaux ont connu un grand succès, ce qui montre que beaucoup de parents sont prêts à acheter des vêtements interconnectés à leurs enfants pour les surveiller. Il faut au contraire convaincre les parents, et élever les enfants dans l'idée que la confiance mutuelle constitue la clé des relations humaines et familiales réussies. Ce qui me permet d'en venir aux mesures éducatives.

Des mesures éducatives

Tout d'abord, il ne faut pas connecter trop tôt les enfants aux objets numériques, car ils doivent prendre le temps d'en comprendre les tenants et les aboutissants. Avant qu'un enfant ait à gérer

un objet interconnecté, il doit avoir une notion de ce qu'est la connexion, ce qui demande du recul et du temps. Afin de ne pas introduire trop précocement les objets numériques, j'ai établi en 2008 des seuils et des balises dénommées « 3, 6, 9, 12 ». Ces quatre chiffres correspondent à quatre âges de la vie : 3 ans, 6 ans, 9 ans et 12 ans. Avec les objets interconnectés, il faudra apprendre très tôt aux enfants à considérer les objets à la fois comme des serviteurs, ce qu'ils seront toujours, mais également comme des espions possibles de nos faits et gestes.

De façon plus générale, nous devons renoncer à penser le monde en termes d'alternatives qui s'excluent, sur le modèle du « ou bien, ou bien ». Il va falloir de plus en plus penser l'ensemble du monde selon des points de vue opposés, mais coexistant, en termes de « à la fois, à la fois ». Le soldat qui pilote un drone est à la fois en plein milieu du champ de bataille et en même temps très éloigné de lui. Et celui qui est engagé dans les opérations militaires à côté d'un robot doit le considérer à la fois comme une créature intelligente capable de le protéger et une simple machine qu'il vaut mieux détruire si elle risque de tomber aux mains de l'ennemi. Quant à la personne âgée dont les données personnelles sont transmises en temps réel par son robot domestique à un centre de télésurveillance, elle devra se considérer à la fois comme protégée dans sa santé et menacée dans sa vie privée. De tels paradoxes semblent difficiles à gérer aujourd'hui, mais c'est parce que nous sommes, pour la plupart d'entre nous, façonnés par une culture du livre dans laquelle les contraires s'excluent durablement¹. Cette capacité de penser en même temps des points de vue différents avec les robots sera d'autant plus importante qu'ils seront à la fois des ersatz d'êtres humains – on pourra interagir avec eux comme avec les êtres humains –, des super-objets – il faudra apprendre à les déconnecter, exiger qu'il y ait un bouton pour les désactiver, ce qui n'est pas le chemin exploré actuellement –, et en même temps, des super-images. Nous aurons une énorme révolution à faire pour pouvoir penser en même temps des choix parfois complémentaires, parfois antagonistes, en permanence.

Enfin, il sera essentiel que nos enfants maîtrisent le langage de la programmation pour bénéficier de toutes les possibilités des systèmes informatiques ouverts. Et nous aurons tout intérêt à développer chez eux le goût du débat et de la controverse pour éviter qu'ils se laissent piéger dans des choix limités par les programmes qui leur seront proposés.

¹ - Voir Tisseron, S. (2013). 3-6-9-12, apprivoiser les écrans et grandir, Toulouse, Eres.

En conclusion,

Avec les objets interconnectés et les robots sur mesure, ce sera à chacun de décider si son intimité est importante, si le risque de la relation avec un autre être humain vaut vraiment la peine, et si le deuil des disparus doit être nié ou accepté. Et insensiblement, nous serons tentés de penser les robots comme une nouvelle forme de vie. Il ne s'agira pas d'une vie organique, développée à partir des briques de notre biologie, mais d'une vie inorganique, bâtie à l'aide de métaux rares et de plastique. Tout cela n'est pas de la science-fiction. Les grandes manœuvres dans ce sens ont à mon avis déjà commencé avec la loi de 2014 sur

les animaux. Jusqu'en 2014, les animaux étaient considérés comme des biens meubles. « Nous » étions des humains, « ils » étaient des animaux. Depuis la loi de 2014, les animaux sont des êtres doués de sensibilité. Un glissement a commencé à s'opérer. Demain, les humains et les animaux seront considérés ensemble comme des êtres vivants doués de sensibilité tandis que les robots pourront être considérés comme des êtres vivants non doués de sensibilité. Malheureusement, l'être humain est ainsi fait que beaucoup d'entre nous seront tentés d'aller encore plus loin et de leur attribuer malgré tout une sensibilité, car ils sauront la simuler d'une façon de plus en plus convaincante au fur et à mesure de leurs progrès... ■



TEMPS 2

Etablir la confiance dans une société connectée

Le cadre législatif de la confiance

Sophie Nerbonne, Directrice de la Direction de la conformité de la Commission nationale de l'informatique et des libertés (CNIL)

La CNIL a été créée il y a plus de 30 ans par la loi du 6 janvier 1978 dite « loi informatique et libertés » dont les principes de base se sont avérés tout à la fois robustes et flexibles alors que le contexte technologique a été bouleversé et que nous sommes entrés dans une nouvelle ère imposant de multiples changements.

La CNIL s'adapte à ce nouvel environnement tant dans son organisation que ses méthodes de travail. Elle entend être partie prenante de l'écosystème de notre société numérique. Ainsi, elle s'est notamment dotée d'un service de l'expertise informatique qui est un des plus importants parmi les CNIL européennes, d'un laboratoire pour y tester des produits innovants, d'une équipe de prospectivistes et d'une direction de la conformité. Cette direction de la conformité est tournée vers les publics professionnels, tant du secteur public que privé, administrations et

entreprises qui mettent en œuvre des traitements de données à caractère personnel afin de mieux les accompagner dans le développement de leurs projets. Pour une conformité à la loi qui soit opérationnelle et non pas seulement administrative, nous déployons de nombreux services et outils.

La présente communication va être structurée selon quatre grandes thématiques :

- le rappel du cadre juridique actuel de la protection des données personnelles,
- les outils de conformité permettant d'assurer sécurité et prévisibilité juridique indispensable pour un système juridique efficace,
- le projet de règlement européen qui dessine le paysage du XXI^e siècle en matière de pro-

1

tection des données personnelles ; ce texte est dans sa dernière ligne droite, puisque le trilogue vient de s'ouvrir entre Parlement européen, Conseil et Commission européenne pour une adoption envisagée à la fin de l'année,

- Enfin, la protection des données qui génère la confiance peut également être un élément de compétitivité, contrairement à l'idée reçue selon laquelle « la réglementation freine l'innovation technologique ».

La protection des données personnelles en cinq règles d'or

Le cadre juridique actuel remonte à 1978, année de naissance de la loi informatique et libertés en France. A l'origine de cette loi, on retrouve le numéro d'identification au registre de l'Insee qui parle à chacun d'entre vous. Cette loi a été votée en réaction à un projet de l'administration malencontreusement baptisé SAFARI et reposant sur l'interconnexion de tous les fichiers publics sur la base du numéro de sécurité sociale ou NIR. Depuis lors, ce ne sont pas seulement les administrations, mais aussi les grands acteurs privés qui sont en mesure d'avoir une connaissance approfondie de notre vie privée. Nous sommes désormais dans un monde de « traçabilité généralisée » des personnes, à la fois dans le temps et dans l'espace. Pour préserver la vie privée et le respect des droits et libertés individuelles et publiques, permettre à chacun d'entre nous de garder la maîtrise de nos données, le législateur a défini un cadre reposant sur cinq principes, cinq règles d'or qui se sont avérées singulièrement robustes, certainement du fait de leur flexibilité.

Avant de passer en revue ces cinq règles d'or de la protection des données, précisons que ce qui existe au niveau français existe aussi au niveau européen : la protection des données personnelles est consacrée par la charte des droits fondamentaux de l'Union européenne. Ce sont des droits personnels, à l'instar de la liberté de penser, la liberté d'opinion, la liberté d'aller et venir. Par rapport au débat sur la patrimonialité de ces données, l'ensemble des autorités de protection des données européennes a pris position sur le fait que ce ne sont pas des droits patrimoniaux. Quand bien même les données auraient été cédées, même soi-disant gratuitement en échange d'un service, l'individu continue à pouvoir exercer ses droits (modification, opposition,...). Il

garde la maîtrise et le droit de demander à s'opposer à figurer dans le traitement, à ce que les informations soient supprimées, dans les conditions définies par la loi. Ce sont des droits liés à la personne qui ne peuvent être aliénés.

La protection des données peut se résumer à cinq grands principes portant sur les modalités de collecte et de traitement des données qui permettent d'identifier, directement ou indirectement des individus. Ces principes sont bien sûr bousculés par l'environnement technologique qui est actuellement complètement différent de celui de 1978, et même de 2004, époque à laquelle la loi a été sensiblement remaniée. Cependant le règlement européen, sur le point d'être adopté reprend ces principes, qu'il est donc important de bien d'avoir en tête.

Le premier d'entre eux est le principe de finalité : à quoi va servir la collecte des données, quelle est la finalité poursuivie par le traitement ? sachant qu'elle doit être explicite, légitime, et déterminée. Prenons en illustration la première demande d'autorisation reçue par la CNIL en matière de « *pay as you drive* », ces dispositifs que les assureurs utilisent en vue de modifier le montant de votre prime d'assurance en fonction de la façon dont vous conduisez.

Au passage, cela montre que la réflexion sur les objets connectés ne date pas d'hier. Nous nous étions interrogés dès 2005 sur ce sujet. Il s'agissait de baisser la prime des jeunes assurés automobile en échange de l'installation d'une boîte noire permettant de savoir que le jeune assuré ne conduirait pas le samedi soir, période de forte sinistralité, qu'il ne dépassait pas les vitesses légales autorisées et s'arrêtait toutes les deux heures. Or, un assureur qui contrôle si les vitesses légales sont dépassées n'est plus un assureur mais un gendarme. Cela ne relève pas de sa sphère de compétences. Quand bien même l'objectif poursuivi, à savoir limiter les accidents est légitime, la collecte de données relatives à des infractions routières n'est pas autorisée par la loi et c'est à ce titre que la CNIL avait alors refusé la mise en place de ce dispositif.

Le deuxième principe repose sur la pertinence des informations collectées, plus précisément la **pertinence, l'adéquation et le caractère non disproportionné des données au regard de la finalité poursuivie**. Ce principe trouve une déclinaison pratique en fonction de chaque cas d'espèce, de même que celui sur la **durée de conservation des données qui soit limitée en raison là encore de**

la finalité du traitement, ce qui n'empêche pas une utilisation à des fins statistiques dès lors que les données sont anonymisées. Un dialogue se noue avec la CNIL afin de justifier de cette pertinence de la donnée ou de sa conservation. On peut citer par exemple, dans les grilles d'analyse de *credit scoring* des établissements de crédit, les critères utilisés ; parmi eux, la CNIL, dans le cadre de son pouvoir d'autorisation, avait refusé la prise en compte du critère relatif à la différence d'âge entre les co-emprunteurs ou de la possession d'un téléphone portable, variables dont la pertinence et le caractère proportionné au regard de la finalité d'octroi d'un crédit à la consommation n'avaient pas été démontrés par l'organisme de crédit.

Le quatrième principe est tout aussi essentiel. **Il concerne la sécurité des données.** A l'heure actuelle, avec des traitements ouverts du fait du recours à internet, au cloud, les questions de faille de sécurité sont de plus en plus cruciales. Comment préserver la sécurité des données, éviter que des tiers non autorisés ne puissent y accéder ?

Enfin – et j'aurais pu commencer par ce principe fondamental –, **le responsable de traitement doit respecter les droits des personnes** : un droit à l'information sur ce qui est fait avec les données collectées, un droit d'accès à ces informations, complétés par la capacité de s'opposer ou de consentir à leur collecte en fonction des cas, d'en demander la rectification également. Néanmoins, il faut bien reconnaître que ces droits sont très peu exercés à l'heure actuelle. Ils le sont surtout dans un contexte contentieux, que ce soit avec son employeur, son opérateur télécom ou sa banque, ou dans le contexte particulier d'un cas pratique donné à des étudiants en droit dans le cadre de leur cours sur la protection des données personnelles.

Pour mettre en œuvre ces 5 règles d'or, la CNIL a conçu des outils de conformité adaptés au secteur professionnel concerné. Par ailleurs, les correspondants « informatique et libertés » désignés par les responsables de traitement soit en leur sein, soit de façon mutualisée pour de petits organismes, constituent des acteurs précieux dans la mise en œuvre effective de la loi.

Des outils de conformité collant à la réalité des besoins des professionnels

La CNIL entend bâtir ses outils de conformité à partir des besoins des professionnels. Dans la



Sophie Nerbonne

perspective d'une mise en œuvre effective des principes susmentionnés, elle élabore ainsi depuis deux ans des « packs de conformité » qui sont tout à la fois une méthode de travail et un nouvel outil de régulation.

S'agissant tout d'abord de la méthode de travail, il s'agit d'interagir avec les professionnels, mais aussi avec les pouvoirs publics, les autorités de régulation sectorielle éventuellement concernées et les associations d'usagers, secteur d'activité par secteur d'activité. En 2014, nous avons pu avancer dans trois secteurs, représentatifs des différentes situations où des professionnels se sont rendu compte qu'il leur était nécessaire de se mettre en conformité.

Le premier de ces trois cas concerne le secteur des assurances. Secteur très réglementé et pour lequel la conformité fait partie de son ADN. Les associations professionnelles représentant l'ensemble des branches d'activité ont posé sur la table l'ensemble des difficultés rencontrées dans l'application de la loi, pour que nous puissions bâtir des outils juridiques répondant à ces questions. Nous avons utilisé les vecteurs juridiques qui sont à notre disposition, normes simplifiées et autorisations uniques et le travail d'écriture s'est fait en concertation étroite avec les professionnels concernés. Ce travail de corégulation s'est avéré essentiel pour permettre l'appropriation par les professionnels de ces règles. Ils parlent désormais de « leur » pack de conformité et le déclinent dans l'ensemble de leur réseau. Les bonnes pratiques et recommandations sur

ce que l'on peut collecter et traiter en tant qu'assureur ont ainsi été clarifiées. Par ailleurs, deux réunions par an permettent de mettre à jour ce corpus de règles.

Le deuxième cas concerne le logement social : à la suite de plaintes, de contrôles et de sanctions prises par la CNIL, le secteur a eu la volonté de se mettre à hauteur en termes de respect de la loi et cette démarche de régularisation a débouché sur un pack de conformité dans le logement social pour rappeler là encore quelles sont les règles de collecte et de traitement des données.

Le troisième exemple est celui des compteurs communicants et plus précisément des services déployés en aval des compteurs dits « intelligents » dont le déploiement débute, et qui enregistrent toute la consommation électrique dans un domicile – rien de plus personnel que ce qui se passe au sein du foyer. Les industriels de la fédération des industries électriques, électroniques et de communication (FIEEC) ont travaillé avec la CNIL sur les garanties offertes par les produits domotiques en mode « *privacy by design* », c'est-à-dire en intégrant le plus en amont possible dans la technologie, les éléments relatifs à la protection des données. La FIEEC a considéré que la valeur ajoutée de ce travail commun consistait dans le facteur de confiance en résultant, ce qui constitue un atout concurrentiel très intéressant pour ses membres. Bien évidemment, la FIEEC a souhaité que ce travail soit capitalisé au niveau européen, ce que représente la dimension la plus adaptée à ces travaux.

Par ailleurs, il faudrait également parler des labels délivrés par la CNIL et aussi de la mise en place de « règles internes d'entreprises » en matière de transferts de données hors de l'Union européenne, dites « BCR » qui sont des outils de conformité très utiles. Ces outils sont repris largement dans le projet de règlement européen ainsi d'ailleurs que les correspondants « informatique et libertés ».

Une dimension européenne incontournable

Pour la CNIL, l'adoption du règlement européen envisagé depuis 2012, apparaît essentiel pour créer un cadre unique en matière de protection des données pour tous les pays d'Europe. Ce n'est en effet pas encore le cas à l'heure actuelle puisque des différences, tant en termes de pouvoirs et missions des différentes autorités nationales que des doctrines établies,

découlent des transpositions nationales qui ont été faites de la directive européenne de 1995. Le règlement européen doit constituer un cadre juridique efficace porteur de sécurité et de prévisibilité juridique. Ce cadre unique va aussi permettre de rétablir l'équilibre vis-à-vis des grands acteurs, notamment américains, qui ont pris une longueur d'avance dans le déploiement de modèles économiques faisant fi de la protection des données personnelles. Certains considèrent, et à juste titre je pense, que nous arrivons à la fin d'une première période, qui a été en quelque sorte le Far West du numérique, période au cours de laquelle chacun a amoncelé des informations de façon massive. Il est temps de « civiliser » ces modes de collecte et de traitement des données non respectueuses des droits des personnes et de donner à celles-ci plus de maîtrise et de contrôle de l'usage de leurs données.

Le règlement européen prévoit ainsi de nouveaux droits pour les personnes, tels que le droit à la portabilité des données, qui vont au-delà de ceux qui ont été indiqués. Ces droits plus classiques se trouvent confortés, renforcés comme par exemple le droit de consentir. A l'heure actuelle, comment considérer que le consentement est éclairé et explicite lorsqu'on ne sait pas à quoi on consent, faute d'une information adaptée. Par ailleurs, lorsque cette information est faite, un travail de sensibilisation est nécessaire afin de ne pas accepter, au moment où nous téléchargeons une application sur notre téléphone portable par exemple, les conditions générales d'utilisation sans vérifier ce qui va être fait de nos données.

Ces droits fondamentaux qui existent au niveau européen se diffusent également dans le reste du monde. Il existe une association des autorités nationales francophones de protection des données, ce qui permet d'étendre et d'associer à la réflexion un certain nombre de pays notamment africains. La même structure existe pour les pays hispanophones en Amérique Latine. Aux Etats-Unis, la protection des consommateurs est prise en considération dans un cadre juridique établi non pas de façon globale mais sectorielle.

La réglementation « informatique et libertés », un facteur de confiance nécessaire pour une innovation durable

Le cadre juridique protégeant les données personnelles est critiqué par les partisans d'une dérégulation qui opposent réglementation et innovation. Or, la réglementation est non

seulement un gage de sécurité juridique et de limitation des risques mais elle permet aussi d'accroître le capital de confiance vis-à-vis des personnes concernées. Elle constitue ainsi un facteur de compétitivité pour les entreprises qui peuvent se prévaloir de pratiques vertueuses et de mise en place d'outils ou d'acteurs tels que les correspondants Informatique et Libertés, les labels ou les règles internes d'entreprise qui témoignent de leur engagement pour un déploiement d'innovations prenant en compte la protection des données.

La CNIL considère que cette réglementation peut être un outil pour bâtir une innovation durable se déployant dans une société préservant les valeurs humanistes dont l'Europe a tout lieu de s'enorgueillir. Il est bien naturel que l'Europe ne se laisse pas dicter un modèle économique

et qu'elle entende peser dans le débat mondial pour faire passer le message d'un choix de société dans laquelle nous voulons vivre et qui est en définitive le soubassement de cette législation.

Certes la réglementation n'est pas tout. Comme le disent les Américains, il n'y a pas de « *silver bullet* », pas une seule balle, pas une seule solution. En matière de protection des données, il y a tout à la fois matière à poursuivre l'éducation au numérique des citoyens, des dirigeants, à sensibiliser les professionnels, encourager le déploiement d'innovations technologiques permettant de préserver la vie privée et les libertés, que ce soit des mécanismes d'anonymisation, de limitation de la durée de conservation des données, de gestion par les individus des données qui les concernent... Il y a là aussi un champ infini d'explorations sur le plan technique. ■

La technologie au service de la confiance ?

Philippe Tassi (1972), Directeur général adjoint de Médiamétrie, Vice Président d'ENSAE Alumni

2

« C'est un joli mot : confiance » aurait pu chanter Jean Ferrat. Sur ce mot reposent beaucoup d'équilibres de notre société, dont celui de la société connectée actuelle et à venir.

Les données ou statistiques, détenues ou élaborées par des administrations ou des entreprises, ont en général été construites à partir d'informations individuelles, ce qui pose la question de la protection des sources, c'est-à-dire de la vie privée, compte tenu des progrès constants de la science et des processus de traitement.

Comment établir et maintenir la confiance du grand public, partie prenante numéro un, tout en respectant l'équilibre entre promesse de confidentialité et utilisation des données recueillies ?

L'établissement de la confiance s'appuie sur deux conditions nécessaires et fondamentales : l'une est réglementaire, et montre que beaucoup d'États ont pris conscience depuis longtemps de la nécessité d'établir des garde-fous juridiques ; l'autre s'appuie sur la déontologie et l'éthique des acteurs.

Une troisième condition existe, et repose sur les technologies qui peuvent aider à déployer des obstacles techniques empêchant la diffusion de données contre le gré de leur auteur.

Le grand public peut-il avoir confiance dans les procédures technologiques permettant de sécuriser les bases de données ? La réponse est OUI. Néanmoins, le maintien à niveau des méthodes de protection nécessite une attention permanente.

Cryptage et codage : premières approches toujours d'actualité

Comment protéger des informations ? C'est une question à laquelle les réponses sont anciennes. Les mots clés sont : transformer, masquer le contenu des messages envoyés par un émetteur vers un récepteur. Cela permet de limiter la possibilité pour un tiers, un « adversaire », d'accéder à une information confidentielle.

Sont donc apparues voici quelques milliers d'années (Sumer, Babylone, Égypte) des techniques de cryptage ou de chiffrement : une suite de symboles est remplacée au niveau de l'émetteur par une autre suite selon un système connu du récepteur, ce qui permet à ce dernier seul de comprendre le message de départ. Le propos n'est pas ici d'écrire l'histoire des méthodes de cryptage, mais d'en faire comprendre les principes, les enjeux et les conséquences (Voir figure 1).

Le vrai destinataire connaît la technique de cryptage et est donc capable de retrouver le message initial. Un tiers, qui aurait intercepté la lettre, est *a priori* incapable d'en comprendre le sens.

Si le cryptage implique le décryptage, pour que le destinataire puisse comprendre le message original, il a très vite aidé également au développement des tentatives pour casser la clé de cryptage.

A commencé alors un jeu de gendarme et voleur (adversaire, « tierce partie », hacker...), traduisant le couple cryptographie-cryptanalyse, le premier étant l'appellation générique de techniques de protection par chiffrement d'un contenu initial, le second étant celle des travaux et moyens nécessaires à faire céder cette protection pour découvrir ce même contenu chiffré.

Le couple cryptographie-cryptanalyse est un bel exemple d'un algorithme protection-effraction dont on peut se demander s'il sera convergent un jour.

Une course à la complexification a commencé. Pour l'illustrer, en repartant du chiffrement de Jules César, il suffit d'accroître le nombre d'alphabets de transcription (Voir figure 1).

Quelques lignes sur la cryptanalyse : le déchiffrement d'un message crypté a donné lieu à de multiples avancées dès le premier millénaire, mais il y a peu de doute sur le fait que la formalisation des méthodes de cryptanalyse est due à l'école des mathématiciens arabes des neuvième et dixième siècles. Comme l'écrit Simon Singh, « *la cryptanalyse ne pouvait être inventée que par une civilisation ayant atteint un niveau élevé de connaissances dans plusieurs disciplines, dont les mathématiques, la statistique et la linguistique* ». Un nom semble se dégager, celui d'Abu Yusuf Al Kindi, dont on a retrouvé en 1987 le traité intitulé *Manuscrit sur le déchiffrement des messages cryptographiques*, et qui fonde son approche sur l'analyse statistique des fréquences de lettres, de signes et de syllabes.

La cryptographie atteint ses lettres de noblesse reconnue sous Louis XIV, avec le très célèbre « Grand Chiffre » du Cabinet Noir du roi. Louis XIV confie à un mathématicien d'Albi, Antoine Rossignol des Roches, plus tard aidé de son fils Bonaventure puis de son petit-fils Antoine-Bonaventure, la mise au point d'un code d'une extrême complexité, dont la clé disparaîtra avec la famille Rossignol. Il faudra attendre deux siècles pour le casser. Le nom de *rossignol* est d'ailleurs

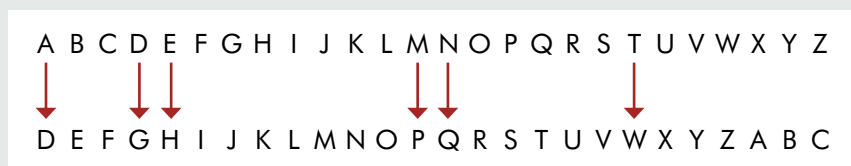
Figure 1 : EXEMPLE DE CRYPTAGE

Un exemple de cryptage bien connu est celui utilisé par Jules César pour communiquer avec ses généraux pendant la la guerre des Gaules.

Le principe est simple : il s'agit d'établir une correspondance entre deux alphabets, un alphabet initial classique et un alphabet d'écriture, décalé.

Ainsi pour transformer le mot *tandem*, le deuxième est décalé de trois lettres et commence par d pour finir par c. Chaque lettre du mot original est remplacée par sa correspondante dans l'alphabet n°2.

Ainsi, le t est donc remplacé par w, le a par d, le n par q, etc...



Le mot *tandem* ainsi crypté devient *wdqghp*.

resté dans le langage courant pour désigner un crochet capable d'ouvrir de multiples serrures.

Bien sûr, les cryptographes savent qu'existent les cryptanalystes, et un jeu de vrai – faux message va rapidement apparaître. L'auteur d'un message sachant qu'il risque d'être décrypté et donc compris par un tiers sera tenté d'y inclure de fausses informations pour tromper l'ennemi. Ainsi, une lettre de Louis XIV écrite avec le « Grand Chiffre » a été décodée à la fin du XI-Xème siècle, dans laquelle le roi désigne le général Vivien de Bulonde comme étant le célèbre « homme au masque de fer », puni à la suite de trahisons et fautes militaires. Mais de nombreux historiens considèrent que cette lettre – absolument authentifiée – n'est peut-être qu'un leurre pour cacher une vérité toute autre.

N'oublions pas de mentionner également le passage à la mécanisation du chiffrement, dont l'un des initiateurs fut au début du XIXème siècle l'anglais Charles Babbage, visionnaire de l'ordinateur, et, bien évidemment, le génie d'Alan Turing, à la fois pour ses avancées théoriques et son génie de cryptanalyste pendant la guerre 1939-1945.

Bien plus proches de nous, enfin, il faut mentionner les travaux de Mme Shafi Goldwasser et Silvio Micali, du MIT, sur la notion de sécurité sémantique. Ils ont mis au point le premier système à chiffrement probabiliste, à sécurité absolue.

Après un pré-print en 1982 – dans lequel est présente une question majeure : *What is a secret ?* –, leur publication initiale date de 1984, dans le *Journal of Computer and Systems Sciences*.

Ils ont reçu en mars 2013 le *Alan Turing Award* (considéré comme le prix Nobel pour l'informatique, de même que la médaille Fields l'est pour les mathématiques) au titre de l'année 2012, prix décerné par l'ACM (*Association for Computing Machinery*).

Le codage de Chappe

Une autre méthode pour masquer le contenu d'un message en clair est le codage. Illustrons par le codage de Chappe. Claude Chappe (1763-1805), ingénieur et inventeur, conçoit un système visuel de télégraphie aérienne, utilisant des pitons ou des édifices visibles de loin, et constitué d'un réseau de stations distantes de 10 kms en moyenne (Galilée a inventé la

Figure 2 : EXEMPLE DE CRYPTAGE

Reprenons l'exemple précédent de chiffrement du mot *tandem*, en ajoutant un deuxième alphabet décalé de 7 lettres, commençant par la lettre h pour s'achever par g.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
H	L	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G

Le principe sera de remplacer les lettres impaires (t, n, e) par leur correspondante dans l'alphabet 1, et les lettres paires (a, d, m) dans l'alphabet 2. La lettre t est transcrite par w, le a par h, le n par q, le d par k, etc. Le mot *tandem* ainsi crypté sera écrit *whqkht*.

Figure 3 : UNE STATION DE CHAPPE

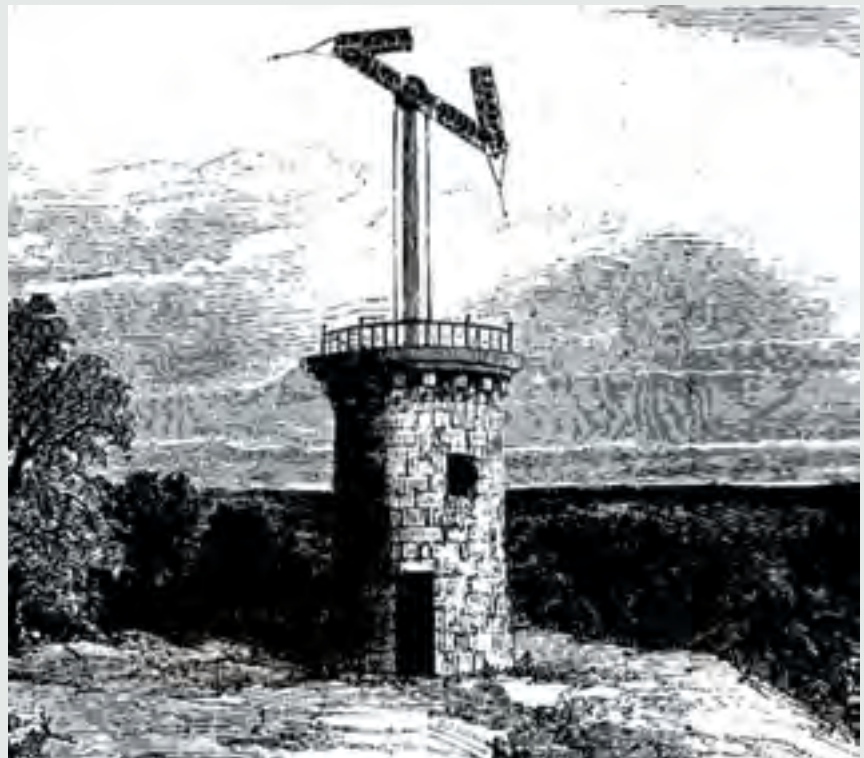


Figure 2 : PREMIER TÉLÉGRAMME OFFICIEL UTILISANT LE RÉSEAU ET LE CODE DE CHAPPE



Ce message est envoyé à la Convention Nationale – assemblée qui gouverna la France de septembre 1792 à octobre 1795. Il annonce la victoire des forces révolutionnaires françaises le 30 août 1794 sur les troupes autrichiennes à Condé-sur-Escaut (« Condé être restitué à République. Reddition avoir eu lieu ce matin à six heures », avec les codes associés).

lunette astronomique en 1609). Le principe est d'utiliser un tryptique articulé, avec sémaphore, permettant le transfert de station en station d'un message masqué en utilisant un code décrit dans un livre de 92 pages et 92 lignes chacune. Les signaux sont en paires, le signal 1 indique la page et le signal 2 la ligne dans la page, ce qui autorise le récepteur final du message à le retranscrire.

Après le suicide de Chappe, en 1805, Napoléon Ier fait construire en 1806 un système de surveillance des côtes depuis le rivage (conçu par Louis Jacob, amiral et homme politique). Le poste de guet s'appelle un sémaphore, et la communication est faite par signaux optiques, inspirée du télégraphe de Chappe.

Principes et techniques de protection et de confidentialité

Les méthodes de cryptage ou de chiffrement, anciennes, sont toujours au centre des recherches et procédés pour dissimuler une information, et les travaux de Goldwasser et Micali en montrent toute l'actualité et l'efficacité.

Nous allons maintenant aborder d'autres techniques de masquage de contenus tout en présentant la philosophie qui préside maintenant à la recherche sur la sécurité des bases de données, fussent-elles *small* ou *big*.

L'approche de Tore Dalenius

Au cours du dernier quart du XX^{ème} siècle, un précurseur de la réflexion sur les bases de données est le grand statisticien et spécialiste de sondages suédois Tore Dalenius.

Dans un article publié dans *Statistik Tidskrift* en 1977, intitulé « *Towards a methodology for statistical disclosure control* », Dalenius pose le principe suivant : « *Accéder à une base de données ne doit pas permettre d'apprendre plus de choses sur un individu que ce qui pourrait être appris sans accéder à cette base de données.* »

Dit autrement, si la publication d'un agrégat statistique T de la variable X permet de connaître la valeur de X pour une personne plus précisément que sans accès à T , alors il y a eu faille de confidentialité.

Ce principe a le mérite d'être simple et compréhensible. Cependant, il peut être mis en défaut par l'existence d'information auxiliaire, élément important de la théorie des sondages.

Par exemple, supposons qu'existe une base de données D contenant des agrégats statistiques qui sont les tailles moyennes des hommes adultes de diverses nationalités. On y ajoute une information auxiliaire : « *M. A fait 2 cm de plus que l'Italien moyen* ». Le seul accès à l'information auxiliaire donne très peu d'éléments sur $M. A$, mais en croisant les deux informations, on connaît la taille exacte de $M. A$.

Ceci peut être présenté de façon plus mathématique :

Soit S un mécanisme de protection d'une base de données D , et une faille F dans cette protection. Pour toute base de données D , il existe une information auxiliaire Z telle que :

- a. Z seule n'a pas d'utilité pour un « adversaire ».
- b. Z combinée avec les données protégées de D permet de trouver dans S une faille F avec une probabilité qui tend vers 1.

Quoi qu'il en soit, le principe de Dalenius est très intéressant dans sa formulation et sa finalité, la protection des données individuelles. Et l'information auxiliaire créant une faille de confidentialité, si elle a une existence théorique, doit encore être trouvée et accessible. Que se passe-t-

il si elle-même est protégée, par exemple cryptée de façon très sécurisée ?

Autres techniques de transformation

- La destruction

Comme son nom l'indique, il s'agit ici de supprimer toutes les données individuelles après une période conventionnelle d'exploitation. Méthode définitive, bien sûr, mais qui n'empêche pas de protéger la base de données pendant sa durée de vie active pour éviter qu'un tiers, pirate, accède à la base pendant cette période et continue à l'exploiter ensuite.

Par ailleurs, de nombreuses données ont une longue durée d'intérêt, pouvant dépasser le cadre conventionnel, ne serait-ce que pour les historiens ou les chercheurs.

- L'anonymisation

C'est une méthode *a priori* évidente : il s'agit d'enlever de la base de données les variables permettant d'identifier une personne, comme le nom, l'adresse, le code d'immatriculation, l'adresse IP, les numéros de téléphone, le code PIN, les photos, les éléments biométriques...

Plus généralement, il convient d'ôter aussi les variables permettant, par recoupement ou croisement, de retrouver – même imparfaitement – un élément identifié dans un ensemble, par inférence individuelle.

Néanmoins, on observe que le nombre de failles augmente avec le progrès des méthodes et des algorithmes scientifiques et la probabilité de ré-identifier un individu augmente, même après anonymisation.

- Agrégation

Une idée facile à mettre en oeuvre consiste à agréger les données individuelles et à ne rendre publics ou utilisables que des résultats. C'est le principe de la loi de 1951 sur le secret statistique. Base de la statistique, l'agrégation consiste à regrouper les données individuelles en résultats synthétiques.

Pour l'agrégation comme l'anonymisation, les performances des modèles mathématiques remettent en cause ces deux techniques, en autorisant une inférence individuelle.

Deux exemples récents en illustrent leurs limites potentielles.

En 2008, N. Homer, S. Szelinger, M. Redman et alii publient « *Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays* » dans PLoS Genetics (Public Library of Science).

Extrait du résumé :

« *We use high-density single nucleotide polymorphism (SNP) genotyping microarrays to demonstrate the ability to accurately and robustly determine whether individuals are in a complex genomic DNA mixture. We first develop a theoretical framework for detecting an individual's presence within a mixture, then show, through simulations, the limits associated with our method, and finally demonstrate experimentally the identification of the presence of genomic DNA of specific individuals within a series of highly complex genomic mixtures, including mixtures where an individual contributes less than 0.1% of the total genomic DNA. These findings shift the perceived utility of SNPs for identifying individual trace contributors within a forensics mixture, and suggest future research efforts into assessing the viability of previously sub-optimal DNA sources due to sample contamination. These findings also suggest that composite statistics across cohorts, such as allele frequency or genotype counts, do not mask identity within genome-wide association studies. The implications of these findings are discussed.* »

Le deuxième exemple est contenu dans l'article "Identifying participants in the Personal Genome Project by name", de L. Sweeney, A. Abu et J. Winn. Travaillant sur le Personal Genome Project (PGP) américain, les auteurs montrent qu'ils identifient correctement 84 % des profils ; ce taux monte à 97 % si on relâche la notion d'identification stricte, par exemple en acceptant les diminutifs comme Jim pour James ou Andy pour Andrew.

Obscurcissement

Obscurcissement, assombrissement, obfuscation, autant de synonymes pour désigner des méthodes d'altération des données. On distingue les techniques non perturbatrices et les techniques perturbatrices ou interventionnistes.

Dans la première famille, on peut masquer la valeur de certaines cellules dans un tableau de résultats, enlever des variables pour certains individus, ou partager seulement un échantillon de



**Philippe
Tassi**

données extrait de la base de données générale, ou enfin combiner certaines catégories pour des variables à modalités.

Dans la deuxième famille, on agira directement sur les données, afin d'engendrer du bruit, au sens large. Par exemple, il est possible de modifier certaines variables numériques en les arrondissant, ou en les bloquant par troncature à des seuils maximum ou minimum.

On pourra transformer les variables en leur appliquant un homomorphisme $(X_1, \dots, X_k) \rightarrow (j_1(X_1), \dots, j_k(X_k))$; ou bien permuter les valeurs d'une même variable entre deux répondants, ou encore perturber les données en y ajoutant un bruit aléatoire.

Certaines transformations (permutation, rotation) laisseront invariantes les statistiques, d'autres non.

Une piste d'investigation particulièrement intéressante porte le nom de données synthétiques (« *synthetic data* »), née des travaux sur les données manquantes de John Little et Donald Rubin.

La confidentialité différentielle

Dernière approche pour établir la confiance, née au début du XXI^{ème} siècle, la notion de confidentialité différentielle semble particulièrement porteuse d'espoirs et de richesse.

Sa philosophie n'est pas éloignée de celle de Tore Dalenius ; on peut l'énoncer ainsi :

« La probabilité d'une conséquence négative quelconque pour une personne (par exemple, le fait qu'elle se voie refuser un crédit, ou une assurance) ne peut augmenter significativement en raison de sa participation à une base de données. »

Le principe est donc de promettre à une personne susceptible de contribuer à une base de données qu'une tierce partie (« l'adversaire ») ne pourra rien apprendre de plus sur elle, qu'elle accepte d'y participer ou non.

Les deux mots importants dans l'énoncé de la philosophie de la confidentialité différentielle sont *probabilité* et *significativement*, qui nous renvoient au domaine de la statistique et de l'aléatoire.

Faisons une analogie : dans les années cinquante, les travaux sur la qualité de William E. Deming évoquaient la notion de « zéro défaut ». Deux suites ont été données. Une première voie a consisté à ne rechercher que le zéro défaut, et à organiser les processus de production et de contrôle en ce sens. Une alternative a admis l'existence de défauts, dont l'existence est impossible à empêcher, ne serait-ce que parce que l'intervention humaine sera toujours un point faible, mais elle a structuré la production de façon à ce que la probabilité qu'il y en ait zéro soit la plus grande possible.

L'esprit de la confidentialité différentielle est proche de la seconde voie.

On admet la possibilité de failles de confidentialité, soit en raison du maillon faible constant qu'est l'homme (le « traître »), soit en raison des progrès exponentiels des modèles mathématiques qu'il est impossible de prévoir. Un cadre d'interdits sera ainsi rapidement dépassé et caduc en raison du progrès scientifique. Cependant, si des failles sont possibles, il faut tenter d'en diminuer la probabilité de survenance.

L'idée est donc de quantifier le risque (probabilité) de faille dans la promesse de confidentialité, tout en mesurant l'effet des techniques de protection des données et de la vie privée sur la qualité des résultats et analyses statistiques.

Ce concept de confidentialité différentielle fait partie du rapport de J. P. Holdren – E. S. Lander, intitulé « *big data and Privacy : A Technological*

Perspective », remis en mai 2014 au Président B. Obama, par le Conseil sur la Science et la Technologie auprès du Président (PCAST).

En résumé, les approches technologiques existent pour garantir et favoriser la protection des données et aider à l'établissement de la confiance du public contribuant à une base de données. Elles constituent un champ important

de recherche. Certes, elles ne sont certainement pas très connues du grand public, mais doivent-elles l'être ?

L'avenir repose sur leur adoption, leur déploiement, leur maintenance et leurs évolutions, chez les gestionnaires de systèmes d'information, en dépit de leur coût de mise en place et de possession. ■

Ethique et déontologie : l'exemple d'Esomar

Elisabeth Cosnefroy, Représentante France d'ESOMAR
(the World Association for Social, Opinion and Market Research)

3

Esomar : une initiative européenne, un périmètre mondial

Esomar représente une profession discrète, qui ne parle pas assez de ce qu'elle a mis en place depuis quasiment ses débuts, mais qui prouve qu'en appliquant une déontologie et une éthique, on peut établir une confiance, et ça fonctionne.

Esomar est, historiquement, une association européenne à l'origine, d'où le E. Elle est depuis devenue mondiale. Esomar est donc actuellement l'association mondiale de tous les professionnels des études de marché, sociétales et d'opinion. L'expression « tous les professionnels » recouvre les deux parties prenantes que sont d'une part les acteurs qui conçoivent et réalisent ces études, donc les instituts et beaucoup d'autres organismes, ainsi que, d'autre part, tous les utilisateurs et commanditaires, les acheteurs, c'est-à-dire les entreprises. Cela recouvre un grand nombre d'interlocuteurs.

Cette association a été créée quasiment dès la naissance de cette profession des études, en 1948, au milieu du XX^e siècle. Ce métier des études, apparu un peu avant la guerre – l'American Institute for Public Opinion est créé par Gallup, Crossley et Roper en 1936 –, est resté en sommeil pendant la deuxième guerre mondiale, pour des raisons facilement compréhensibles. Il renaît juste après, avec une volonté ferme et très efficace de créer tout de suite une association, au départ européenne, qui établit un certain nombre de règles et de bonnes pratiques pour être sûr de pouvoir développer correctement le secteur des études marketing, dans le respect de ce qui anime le plus ses membres, c'est-à-dire l'humanisme.

Aujourd'hui, Esomar est reconnu dans 130 pays, avec près de 5 000 membres. Depuis deux ans, nous nous sommes ouverts à une forme d'homologation beaucoup plus *corporate*, qui n'implique pas seulement des membres individuels au sein d'entreprises, comme les cadres responsables des

études et sondages, mais l'ensemble de l'institut ou de l'entreprise eu égard à l'usage qui est fait des études.

Le rôle d'Esomar : la régulation du marché des études

Notre mission est simple à énoncer. Elle consiste à protéger les intérêts du métier des études de marché : c'est-à-dire, directement ou indirectement, protéger forcément les intérêts de ce que l'on appelle le consommateur. Si on ne protège pas les personnes, nous n'avons plus de raison d'être et d'exister puisque nous les représentons.

Notre objectif est de protéger les intérêts du public, certes, mais aussi d'améliorer et développer les bonnes pratiques, en France et à l'international. Nous voulons aussi promouvoir les meilleurs standards, en parler, y revenir pour les améliorer sans cesse, et bien évidemment anticiper sur ce qui va devoir évoluer, et donc promouvoir la valeur du marché des études. Un détail de traduction, propre à la langue française : au plan international, on ne parle pas de marché, mais d'industrie, et on utilise le terme de « market research » et de « researchers » traduits respectivement études de marché et professionnels des études, puis des études marketing pour simplifier et tenter de se faire aspirer par la dynamique du marketing dans les années 90. Mais ceci nous a éloigné de la sphère de la recherche et n'est pas sans nous poser problème à ce jour. Pour défendre notre éthique et notre métier au sein de la Commission européenne, dans le cadre de la mise en place de nouvelles lois sur l'usage des données privées, nous avons réinvesti le terme de chercheurs et de recherche en général afin de nous distinguer du rôle commercial de la fonction marketing et valoriser notre rôle d'analyse et non d'exploitation commerciale des données privées.

Esomar a pris très tôt, dès 1948, l'initiative d'un code de déontologie et de bonnes pratiques destiné à donner confiance aux consommateurs. Très vite, avec une volonté d'unification et d'internationalisation, nous avons travaillé avec la Chambre de Commerce International, ce qui a permis d'aboutir en 1976, au code actuel qui porte le nom de code ICC Esomar, en français, CCI Esomar.

Ce code de déontologie a été adopté par 67 associations dans le monde, traduisant notre préoccupation première qui est, dans chaque pays, d'être en harmonie avec les associations parties

prenantes et surtout les organisations syndicales. Pour donner un exemple, en France, sont membres et partenaires, et donc vraiment parties prenantes, non seulement les entreprises utilisatrices mais également les organismes représentatifs des acteurs du secteur que sont l'Union des Annonceurs (UDA) – qui regroupe tous les annonceurs publicitaires au sens large –, l'association française des professionnels du marketing (ADETEM), l'IREP (Institut de Recherches et d'Études Publicitaires) qui représente tous les professionnels des médias publicitaires – les médias sont au cœur des problématiques qui nous animent en ce moment – et l'organisation syndicale SYNTEC qui est l'organisation syndicale de tous les instituts d'étude.

Déontologie et bonnes pratiques

Tous les membres d'Esomar s'engagent volontairement et financièrement à respecter le Code ICC/ESOMAR et l'autorité du Conseil disciplinaire en commençant par se conformer aux lois nationales et internationales. Cela signifie être éthique, porter une attention spéciale à toute étude réalisée auprès des enfants de moins de 15 ans et des jeunes en général, s'assurer de la coopération volontaire des participants, respecter leurs droits, ne jamais laisser les données personnelles collectées utilisées à d'autres fins que d'étude, vérifier que précision, transparence et objectivité sont respectées, et enfin se conformer aux principes reconnus de la concurrence loyale.

S'afficher « membre raisonnable » d'Esomar, cela veut dire que des éléments concrets ont été signés, avec un engagement financier aussi et juridique. En particulier un membre *corporate* peut voir s'engager des poursuites contre lui. Outre l'engagement volontaire et financier à adopter ce code et à le respecter sans cesse, il convient de le communiquer aux partenaires et clients, et à le faire adopter. Cela n'est pas toujours facile ; par exemple, certains ne comprennent pas pourquoi une société adhérente à Esomar refuse de fournir à un tiers les données identifiées.

La priorité est donc d'être éthique, c'est vraiment le point majeur. Nous apportons ensuite une attention spéciale quant à l'approche des enfants. En Europe, les réglementations nationales ont des stades de développement variables au plan général, mais tout ce qui touche les enfants de moins de quinze ans, ou plus généralement les jeunes, est très homogène et de niveau élevé.

Le troisième point est de s'assurer de la coopération volontaire des participants à chacune de nos études, en respectant leurs droits et même en les leur rappelant explicitement : systématiquement, à chaque enquête, il est répété qu'ils ont le droit à l'anonymat, qu'ils peuvent se retourner contre les sociétés de recueil ou de traitement, ou, a minima, les interpeller si questions.

Bien sûr, ne jamais laisser les données personnelles collectées être utilisées à d'autres fins que l'étude est une règle absolue, comme le fait de s'assurer que les projets sont réalisés et présentés avec précision et en toute transparence et objectivité. Nous savons tous qu'il y a beaucoup d'autres façons d'identifier par croisement d'informations.

Tout ceci signifie que des comptes sont susceptibles d'être rendus par rapport au respect de tous ces points.

Esomar est au service également des dirigeants et professionnels du secteur des études. Le Conseil est renouvelé régulièrement, comme son président, et celui en date est français, Laurent Florès a ainsi été nommé début 2015 pour deux ans. Il existe des représentants dans chaque pays qui sont des relais pour rappeler les règles ou régler problème ou litige, et pour dispenser des conseils. Tout le cadre est facilement accessible et consultable sur le site d'Esomar, via de nombreux *guidelines* à destination des parties prenantes. Ces divers *guidelines* sont élaborés par des équipes d'experts constituées publiquement. Ces experts produisent des textes ensuite mis en consultation pour l'ensemble des acteurs, et, une fois validés, qui seront *in fine* édités et diffusés. Pour illustrer, Esomar vient de rééditer le *guideline* concernant les études d'opinion, au plan mondial.

Une autorégulation efficace et réussie

Une question majeure est, bien entendu : créer un code, c'est bien, mais est-il appliqué ? Il y a même parfois un sentiment, dans cette profession, que « finalement, il n'y a pas de problèmes ». C'est exact, et le métier peut vraiment s'en vanter largement. Ceci n'interdit pas les garde-fous et les procédures disciplinaires, qui relèvent de certains principes-clés : chacun peut déposer une plainte, qui doit être étayée de preuves et de faits, et non fondées sur la rumeur ou la délation. Les procédures s'appliquent aussi bien au monde des individus qu'à celui des entreprises, dans un



Elisabeth
Cosnefroy

sens comme dans l'autre, donc tout est possible. En outre, Esomar peut ouvrir une enquête de sa propre initiative.

En général, pour une plainte avec des preuves, Esomar, en bonne intelligence, avertit le plaignant et celui qui est potentiellement attaqué ou désigné et examine la possibilité d'une entente entre les parties. En cas de réponse négative, une commission indépendante examinera les faits et décidera, ou non, de passer au stade d'une commission disciplinaire. La gradation des sanctions pour le contrevenant va de l'avertissement jusqu'à l'exclusion du membre, avec publication de cette décision d'exclusion et de ses raisons. L'atteinte de ce niveau ultime est extrêmement rare.

Dans la pratique, le principe d'Esomar est une ferme détermination, et ce principe fonctionne d'autant mieux que les plaintes sont peu nombreuses, et légères, ce qui est rassurant pour la régulation de la profession. Le plus fréquemment, les plaintes s'arrêtent dès la première étape de discussion et d'échange.

Quels sont les types de plaintes usuels ? Lorsqu'elles émanent du grand public, leur cause majoritaire est le non-règlement de l'*incentive*, car pour un certain nombre d'études la personne participante est dédommée pour son déplacement, pour le dérangement, pour le temps consacré. Concernant les plaintes en BtoB, ce sont plutôt des problèmes de qualité qui sont plus ou moins évoqués. Le terme « qualité »

est parfois compliqué à interpréter, les terrains d'entente pas toujours faciles à trouver. Il arrive également que des instituts ou des organismes se prévalent être des membres d'Esomar, ce qui procure un avantage concurrentiel en cas d'appel d'offres, alors qu'ils ne le sont pas. La mise d'un avertissement, la disparition des logos et le dialogue suffisent à se prémunir contre la répétition. La réalité est qu'Esomar est géré comme une marque surveillée en permanence.

La réussite du modèle d'Esomar, pertinente à partager, repose tout d'abord et prioritairement sur le respect de la loi nationale comme première règle. Ainsi, en France, la première des actions en cas de vide juridique ou de manque d'information, consiste à aller voir la CNIL. C'est un principe essentiel : respecter la loi et la faire respecter. La preuve que l'autorégulation fonctionne bien est qu'il n'y a pas de plaintes majeures, pas d'exclusions à ce jour faites et publiées. Avec les preuves, on arrive toujours très tôt à s'entendre et à faire en sorte que tout soit bien appliqué. Une autre raison de la réussite est la collaboration des acteurs, par exemple pour les guides élaborés par les experts. Plus Esomar impliquera des représentants des acteurs du marché des études, plus le contrôle sera facilité. Ainsi, depuis deux ans, Esomar a englobé dans sa communauté la famille relativement nouvelle des cabinets de tendance qui éditent beaucoup d'études et d'analyses, famille qui n'appartenait pas, jusqu'à présent, à notre univers.

Dialogue et évolution

La stratégie mondiale d'Esomar est effectivement de développer cette image positive des études, également auprès des régulateurs. Esomar dispose d'un bureau à Bruxelles, travaille en

étroite collaboration avec la Commission européenne. Nous suivons régulièrement les avancées, les réflexions, les évolutions des cadres de régulation. Nous rappelons les acquis de notre code, le fait qu'il est respecté par la quasi-totalité de nos membres, en étendant notre champ à un certain nombre de professions connexes. Avec l'arrivée d'internet, nous avons intégré les *access panels* et autres univers de base de données de recrutement de personnes interrogées. Nous insistons beaucoup sur le fait que la recherche n'est pas du marketing déguisé, la vraie recherche qui veut aider et faire avancer la profession des études. Une preuve du bon fonctionnement de ce système : il est aujourd'hui très positif pour nos membres, y compris commercialement, d'inclure dans leurs propositions l'appartenance à Esomar, et de nombreux appels d'offres de grands groupes s'en servent pour le référencement.

Comment mieux conclure qu'en énonçant les cinq règles éditées et largement diffusées et qui constituent le programme de « data sérénité » :

- rester humain : parler d'humain à des humains pour des humains, règle permettant d'être régulièrement éthique.
- prioriser les ressources
- maîtriser les flux
- s'exprimer simplement
- penser responsabilité, en faisant partie d'une chaîne avec des entités en amont et en aval

Diffuser ces règles et en les intégrant dans le quotidien est le meilleur moyen d'établir ou de maintenir la confiance. ■



Conclusion

Antoine Frachot (1989)

Directeur général du Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)

Je vais commencer par une petite anecdote historique des années 70. En 1975-1976, un article du Monde annonce : « l'administration française est en train de faire la chasse aux Français ». Certains politiques avaient eu l'idée de constituer un grand fichier où seraient appariées des données administratives de toute sorte. En gros, il s'agissait de constituer une grande base de données quasi-exhaustive ; une idée absolument géniale fût d'appeler ceci « le projet Safari ».

Un psychanalyste pourrait expliquer, peut-être, ce qui a bien pu traverser l'inconscient des fonctionnaires et politiques à l'origine de ce projet et de sa dénomination, alors même qu'en 1975, la deuxième guerre mondiale ne s'est achevée que trente ans avant. Trente ans, c'est court, pour le temps historique, la sensibilité était encore extrêmement forte sur le mal que l'on pouvait faire à partir de données administratives.

Ce qui fait honneur à la France et aux politiques de cette époque, c'est d'avoir décidé d'arrêter le projet et de prendre le temps de réfléchir. Des débats à l'Assemblée Nationale, au Parlement, ont eu lieu pour essayer de définir une doctrine

des données à caractère personnel, et ce sont ces échanges qui ont donné naissance à la CNIL. Aujourd'hui encore, la doctrine de la CNIL est très marquée par cet événement. Le prix de thèse « Informatique et Libertés » décerné par la CNIL en 2010¹ raconte comment la doctrine des données à caractère personnel s'est forgée en France et comment la CNIL a été créée. Histoire passionnante.

Tout cela pour énoncer ma conviction que les thèmes abordés lors de ce colloque sont vraiment au cœur de la problématique du numérique. D'un côté, on veut « ouvrir » les données au maximum en s'appuyant sur le fait que les données, notamment administratives mais pas uniquement, sont des richesses extraordinaires et qu'elles permettent d'étayer et de mieux comprendre pour ensuite mieux agir dans l'intérêt général sur toutes sortes de problématiques, phénomènes économiques, sociaux ou relatifs à la santé ou à l'environnement. D'un autre côté, la question immédiate est : « comment protège-t-on ces données ? ».

Nous sommes donc face à un dilemme : en même temps ouvrir et fermer les données au

1 - « Les valeurs de la recherche : enquête sur la protection des données personnelles en épidémiologie », Nicolas Lechopier, Ed. Michalon, 2011.

maximum. Ma conviction est qu'il faut résoudre ce dilemme pour éviter une spirale dangereuse.

Ce point est très important et est au coeur de nos réflexions ces dernières années. Ce n'est pas forcément le cas dans tous les pays, mais au moins en France, le balancier, depuis cinq ans environ, penche vers l'ouverture des données. Il oscillera peut-être dans l'autre sens dans six mois. L'Etat et ses administrations ont lancé un grand nombre de débats dont on ne peut que se réjouir. Les données de l'Insee, par exemple, sur les ménages et les individus ont été ouvertes aux chercheurs et à la recherche scientifique dès 2008. Le GENES a bataillé pour que les données fiscales anonymisées soient ouvertes également aux chercheurs, les données individuelles telles que celles de nos déclarations d'impôt. Ceci a été acté par la loi en 2013.



Antoine Frachot

Aujourd'hui, nous sommes en train d'effectuer un nouveau mouvement d'ouverture pour les données de santé, données médico-administratives détenues par l'Assurance maladie qui constituent l'une des plus grosses bases de données santé dans le monde : toutes les prescriptions médicales, tous les actes en hôpital, etc.

Ce mouvement est tout à fait bienvenu puisque nous apercevons bien tous les enjeux possibles portant sur la santé publique, l'environnement, l'énergie, l'emploi... Mais se pose la question « comment faire pour sécuriser tout cela ? ».

Nous imaginons parfaitement qu'il suffirait, dans un proche avenir, que l'on apprenne que les données médicales détenues par l'Assurance maladie se soient retrouvées chez des assureurs ou des mutuelles pour qu'il y ait immédiatement une panique générale justifiée. Chaque jour, ou presque, nous apprenons qu'une entreprise s'est fait pirater ses données. Il y a quelques semaines, encore, l'Etat fédéral américain a annoncé que des données portant sur des fonctionnaires américains avaient été piratées. On n'ose pas penser à ce qui pourrait arriver dans nos démocraties si apparaissait une telle fuite à grande échelle.

La loi santé aujourd'hui, par exemple, programme une ouverture assez large. Si des données médicales fuyaient, ce n'est pas seulement la ministre de la Santé qui serait sur la sellette, très probablement obligée de démissionner, mais le gouvernement dans sa totalité serait en difficulté. Plus que ceci : c'est toute la confiance des citoyens vis-à-vis de l'administration qui serait anéantie. Quand nous remplissons nos déclarations fiscales ou quand nous répondons à un enquêteur de l'Insee, nous le faisons plus ou moins volontairement, d'une part parce que nous y sommes obligés, mais d'autre part parce que nous savons bien que les données de l'Insee, ou les données fiscales, ne vont pas aller errer dans la nature. Si demain une fuite, quelle qu'elle soit, se produit, cette confiance disparaît et la démocratie est minée.

Longtemps, le dilemme que je mentionnais plus haut a été vraiment insoluble, ou plutôt souvent résolu par le principe : « il suffit d'interdire d'accéder à des données ». Il est vrai qu'aujourd'hui, même les chercheurs aux intentions les plus pures ont un mal fou à accéder à des données administratives. En restant sur les données de santé, nous avons beaucoup côtoyé des médecins, lors de ces dernières années, puisque nous travaillons à cette ouverture des données de santé avec le ministère de la Santé et d'autres organismes : les médecins nous citent de nombreuses situations où, avec plus de données, des vies humaines auraient pu être sauvées, ou des perspectives d'amélioration considérable du dispositif français de santé. Le Médiateur a énormément marqué le pouvoir politique car, si les données avaient été un peu plus ouvertes à la recherche scientifique, les effets secondaires de ce médicament auraient peut-être pu être détectés bien plus tôt.

Aujourd'hui, nous arrivons un peu plus à sortir par le haut de ce dilemme grâce à la métho-

dologie et à la technologie. La présentation de Philippe Tassi a montré que les techniques de cryptage et de chiffrement permettent de bien réaliser des actions que l'on faisait mal il y a encore dix ou vingt ans. En technologie, nous avons développé, au GENES, un instrument qui s'appelle le centre d'accès sécurisé aux données. Ce CASD permet aujourd'hui à plus d'un millier de chercheurs en France et en Europe de travailler sur des données individuelles françaises, comme celles de l'Insee. Chercheurs, sociologues, économistes peuvent travailler sur ces données grâce à la technologie que nous avons pu développer, sans que jamais ces données ne sortent des serveurs.

Nous sommes donc, heureusement, sur le bon chemin. Avec le mouvement d'open data qui a été initié ces dernières années, la France est en train de reprendre beaucoup d'avance et les chercheurs résidant en France vont pouvoir dans les années qui viennent faire des travaux innovants que d'autres pays ne peuvent pas lancer, le tout avec un niveau de sécurité très élevé.

Pour finir, deux points concernant l'actualité. Après examen en Conseil des Ministres, la loi sur le numérique sera diffusée à la rentrée de septembre. Et le 18 juin, le gouvernement présente son plan numérique, et un premier projet vise à faciliter les appariements de base de données.

Ainsi, pour l'instant, des études sur santé et travail sont considérablement difficiles à faire. Pourquoi ? Parce que, si l'Assurance maladie a des informations très précises sur notre état de santé, ou l'Education Nationale sur les parcours scolaires, et si l'Insee ou les services fiscaux disposent d'éléments sociodémographiques également très précis, marier deux bases est légalement impossible.

Mais « demain », si ces recoupements sont autorisés ou facilités, toujours avec des conditions très strictes, bien sûr, on pourra faire des études fondamentales sur les liens entre la vie professionnelle et le travail avec la santé, sur les inégalités sociales, sur les parcours scolaires.

Il ne s'agit pas d'une renaissance de Safari, mais de réaliser des études à forte valeur ajoutée pour les politiques publiques, et non de « ficher » ou « fliquer » les gens.

Un deuxième projet gouvernemental est également très important, en cours de maturation et

de réflexion : il s'agit de la notion de données d'intérêt général. Ce concept est apparu récemment, il y a un an ou deux. Il consiste à envisager un échange entre système public et système privé, car de nombreuses entreprises privées aimeraient accéder à des données administratives, alors que les administrations et services de l'Etat souhaiteraient utiliser des données privées pour améliorer la connaissance économique et sociale. Par exemple, les données de téléphonie mobile seraient bienvenues pour faire des études sur les déplacements ou le tourisme en France. Les données de la SNCF ou d'EDF seraient également sûrement très intéressantes. On peut aussi penser à des données sur l'efficacité énergétique, la consommation électrique. D'une manière générale, il y a une conviction partagée par beaucoup que des gisements de données d'intérêt général peuvent servir à l'administration et aussi au développement de petites entreprises, de start-up qui veulent créer de nouvelles applications innovantes.

Certes, ce chantier est difficile. Nous serons unanimement d'accord pour dire que les données d'intérêt général doivent être utilisées par le plus grand nombre. Si ce sont vraiment des données d'intérêt général, nous devons les utiliser pour le bien-être général. Cette idée est compliquée à mettre en œuvre tant que ce concept de données d'intérêt général ne sera pas défini très précisément. C'est pourquoi, récemment, le Premier ministre a lancé une mission avec le Conseil d'Etat, l'Inspection Générale des Finances et d'autres services de l'Etat pour essayer de donner de la chair et du sens à cette notion de D.I.G., qui peut aussi conduire à des innovations scientifiques et des progrès en matière de politique publique tout à fait considérables.

Ces chantiers sont passionnants. Et c'est pourquoi je remercie vraiment, du fond du cœur, Catherine Grandcoing, Philippe Tassi et tous les organisateurs, pour avoir lancé ce débat et cette réflexion en regardant les deux faces du même sujet, et pas uniquement l'aspect « sécurité des données ». Ces deux aspects sont intimement liés, nous devons absolument les affronter pour en tirer du bien. ■

Mobilités

François BUCCHINI (1983)



François Bucchini est nommé directeur général adjoint de Gras Savoye.

François Bucchini est actuaire et diplômé de l'ENSAE (1983). Il a démarré son parcours professionnel comme chercheur à l'Institut national de la santé et de la recherche médicale (Inserm). Il était auparavant directeur des services et de l'offre de Thélem Assurances

Hubert MAITRE (1994)



Hubert Maître est nommé directeur de l'ECAM Rennes - Louis de Broglie.

Diplômé de l'Ecole Polytechnique et de l'ENSAE (1994), Hubert Maître a débuté sa carrière chez un opérateur télécoms dans des postes d'organisation et de marketing.

Olivier ARLES (1992)



Olivier Arlès est nommé directeur général délégué en charge du pilotage économique et financier du groupe Macif.

Olivier Arlès, 47 ans, est diplômé de l'Ecole Polytechnique et de l'ENSAE (1992) ; membre de l'Institut des actuaires, il commence sa carrière professionnelle au sein de la Commission de Contrôle des Assurances (devenue aujourd'hui l'ACPR) en 1992. Il rejoint le groupe Macif en 1998 où il était directeur général adjoint en charge du pilotage économique et financier du groupe depuis 2012.

Michaël AFLALO (1996)



Michaël Aflalo rejoint BFT Gestion en tant que directeur des gestions.

Michaël Aflalo est diplômé de l'ENSAE (1996). Il est également actuaire, diplômé de l'Institut des Actuaires Français (IAF) et titulaire d'une maîtrise de Mathématiques et Application aux Sciences Fondamentales de l'Université Paris IV. Il a débuté sa carrière en 1997 en qualité de gérant de fonds à formule et de fonds profilés chez Crédit Mutuel Finance. Depuis 2010, Michaël Aflalo était directeur solutions institutionnels & réseaux chez Natixis Asset Management.

Fabrice MARAVAL (1998)



Fabrice Maraval est nommé directeur finance et assurance du groupe April.

Fabrice Maraval, 41 ans, est diplômé de l'Ecole Polytechnique et de l'ENSAE (1998). Il est titulaire d'un Master ès finance & économie de la London School of Economics. Il est CFA et membre de la Society of Actuaries. Après avoir travaillé pour AXA en Chine, il est entré à la CNP en 2002, en qualité de chef de projet fusions & acquisitions (M&A). Depuis 2013, il était directeur général de CNP Caution à Paris.

François LOBIT (1982)



Arrivé en 2011, à Saint-Malo, le sous-préfet François Lobit, vient d'être nommé par décret, sous-préfet au Havre.

François Lobit est sous-préfet de Saint-Malo depuis mars 2011. François Lobit, 57 ans, est diplômé de l'Ecole Polytechnique et de l'ENSAE (1982). Il a commencé sa carrière à Amiens, à l'Insee, avant d'y intégrer le corps préfectoral comme chargé de mission auprès du préfet de Picardie.

Christophe GEGOUT (2001)



À compter du 1er septembre 2015, Christophe Gégout est nommé administrateur général adjoint du CEA. Il conserve par ailleurs ses fonctions de directeur financier.

Christophe Gégout, 40 ans, est diplômé de l'Ecole Polytechnique, de Science-Po et de l'ENSAE (2001).

Christophe Gégout commence sa carrière au ministère de l'Economie et des Finances en 2001. Il rejoint le CEA en 2009 en qualité de directeur financier et de directeur du pôle gestion et systèmes d'information.

Christophe Gégout représente le CEA au Conseil de surveillance, puis d'administration d'Areva. Il préside également CEA Investissement.

Gaëlle OLIVIER (1995)



Gaëlle Olivier est nommée directrice générale d'AXA Entreprises.

Diplômée de l'Ecole Polytechnique, de l'Institut des Actuaire et de l'ENSAE (1995), Gaëlle Olivier a débuté sa carrière au Crédit Lyonnais. Depuis 2011, elle occupait les fonctions de directrice générale des activités d'assurance dommages d'AXA Asie.

Richard QUESSETTE (1995)



Richard Quessette est nommé responsable mondial des activités actions et dérivés de Société Générale CIB. Diplômé de

l'Ecole Polytechnique et de l'ENSAE (1995), il débute sa carrière en 1995 comme expert technique sur le trading et le pricing d'options exotiques chez Société Générale CIB.

En 2009, il avait été promu responsable du trading du pôle solutions cross asset de la banque de financement et d'investissement, avant d'en devenir le responsable mondial en 2012.

Laurent CLAVEL (2005)



AXA Investment Managers annonce la nomination de Laurent Clavel au sein de l'équipe Recherche et Stratégie d'Investissement, en qualité d'économiste international senior.

Diplômé de l'Ecole Polytechnique, de l'Ecole d'économie de Paris et de l'ENSAE (2005), Laurent Clavel a commencé sa carrière à l'Insee en 2005. C'est là qu'il dirigeait la division Synthèse Conjoncturelle depuis 2013

Emmanuel KRAGEN (1992)



Emmanuel Kragen est nommé gérant allocataire d'actifs au sein de La Financière de l'Echiquier.

Diplômé de l'Ecole d'Economie de Paris, de l'Ecole des Hautes Etudes en Sciences Sociales et de l'ENSAE (1992), il débute sa carrière en 1994 à la Société Générale au poste d'analyste bancaire. Depuis 2008, il était stratéliste monde sur les marchés actions, de taux et de devises chez Exane Derivatives.

Sandrine DUCHENE (1995)



Sandrine Duchêne est nommée directrice des affaires publiques du groupe AXA, à compter du 15 septembre 2015.

Diplômée de l'Ecole Polytechnique et de l'ENSAE (1995), Sandrine Duchêne, 46 ans, a débuté comme économiste à l'Insee. Depuis 2013, elle travaillait à la direction générale du Trésor, au sein de laquelle elle exerçait les fonctions de directrice générale adjointe et chef économiste, mais également de directrice des relations internationales.

Dans les rayons



Le gâchis français

Jean-Marc Daniel
Editions Tallandier

Jean-Marc Daniel analyse ici la détérioration de l'économie française depuis 40 ans à travers 3 constats : faible croissance et donc taux de chômage élevé, disparition de l'inflation grâce à l'euro et accumulation de dette publique en dépit de nos engagements européens. 40 ans de gâchis dont le mensonge politique est le moteur. Cette détérioration commence au milieu des années 1970. Depuis ce moment, la France essaye de contourner ses engagements. Ainsi, après l'acceptation de la discipline monétaire du SME en 1979, la relance de 1981 a failli le faire exploser. Après l'acceptation du traité de Maastricht en 1992, la dérive des comptes publics de 1993 a fragilisé le passage à l'euro, de même, le refus de respecter le pacte de stabilité par Jacques Chirac a failli faire disparaître l'euro. Le discours sur le pacte de croissance de François Hollande et son manque de courage dans la remise en ordre des finances publiques ont de nouveau jeté le désarroi à Berlin et à Bruxelles. La conséquence est que la France ne répond pas comme il faudrait aux problèmes de son économie : elle se montre incapable d'utiliser les périodes de retour cyclique de la croissance pour faire des réformes et apurer les comptes publics et elle refuse d'admettre que la désinflation impose plus de souplesse, notamment sur le marché du travail. La conclusion est que la France est dans une impasse qu'elle ne pourra quitter qu'à condition de jouer pleinement le jeu des traités européens qu'elle a suscités, réclamés et signés.



Le jour où mon robot m'aimera

Serge Tisseron
Editions Albin Michel

Ils sont déjà parmi nous et nous ne les voyons pas. Ce sont tous nos objets connectés qui détectent nos réactions, s'adaptent à nous, et parfois même orientent nos choix à notre insu. Très bientôt, certains d'entre eux auront une apparence humaine, déchiffreront nos émotions, nous parleront, et pourront même nous manifester de l'affection, voire de l'amour. Ce sera « pour de faux » ? Et alors ?

Mais avons-nous envie de ce bonheur-là ? Comment réagirons-nous au fait de nous savoir surveillés en permanence ? Et que deviendra la relation de l'homme à ses semblables lorsque nous serons entourés de créatures humanoïdes programmées pour tout savoir de nous et anticiper nos désirs même les plus secrets ?

Ce n'est pas de la science-fiction : il est urgent d'y réfléchir et de préparer nos enfants au monde qui sera le leur demain.



Méthodes robustes en statistique

Jean-Jacques Dreesbeke,
Gilbert Saporta
et Christine Thomas-Agnan
Editions Technip

L'émergence de la statistique robuste moderne ne date que des années 1960 avec les travaux pionniers de Tukey (1960), Huber (1964) et Hampel (1968). Depuis cette période, de nombreux modèles et méthodes ont été réexaminés sous l'angle de la robustesse. La prise en compte de l'impact de valeurs atypiques, ou de toute autre structure minoritaire présente dans les données, est d'autant plus importante à l'heure actuelle que l'on dispose de bases de données de plus en plus grandes dont la fiabilité et la qualité sont relativement inégales. Or, l'estimation des paramètres d'un modèle n'est valable que sous certaines hypothèses, bien souvent passées sous silence dans la pratique. La présence dans la population de plusieurs types de comportement ou l'existence de valeurs aberrantes va à l'encontre de l'hypothèse que tous les individus examinés ont un comportement compatible avec le modèle sous-jacent. Les recherches intégrant des méthodes statistiques robustes destinées à surmonter ces difficultés sont intéressantes tant au niveau théorique que pratique. Les méthodes robustes sont également essentielles dans la détection des valeurs atypiques.



Conseil de Direction Générale depuis plus de 25 ans

Ylios est un cabinet de Conseil de Direction
combinant culture de l'excellence et culture entrepreneuriale

NOUS RECHERCHONS

- Des stagiaires, qui participeront à chaque étape d'une mission de conseil, de la réflexion préalable lors d'une proposition client à sa mise en œuvre
- Des consultants juniors, des consultants seniors et des managers

NOTRE AMBITION

- Constituer une véritable « Ecole de Conseil » : proposer des parcours de formation, traduire notre style dans nos pratiques au quotidien, et faire référence dans nos méthodologies
- Faire de notre culture entrepreneuriale un levier de développement



149

heures de vidéo
regardées par an

39

sites internet masculins
dans ses favoris

59

minutes d'écoute radio
en dehors de chez soi
chaque jour

LA CONNAISSANCE
DE TOUS LES COMPORTEMENTS
MÉDIAS ET MULTIMÉDIAS

Pour tout renseignement :
www.mediametrie.fr



Mediametrie

Bienvenue dans le monde des ENSAE Alumni

Variances

Le réseau
ParisTech
Alumni

Annuaire*

L'appli
ensae.org*

Le site
ensae.org

3000
offres
d'emploi
par mois*

Managers.com
Site emploi*

Les
petits-déjeuners
Economie
Finance

Les
Rencontres
Data Science

ENSAE Alumni

Service
Ecoute
Carrières*

Les
Afterwork

Les Happy
Hour Métiers

Des
événements
premium*

Le club des
Administrateurs

ENSAE
Business Angel
et Entrepreneurs

ENSAE
au
Féminin

Les Soirées
Prospective
Métiers