



## Portrait **Le psychologue des voitures autonomes**

Jean-François Bonnefon  
conquiert le monde avec  
son étude des dilemmes  
moraux posés aux  
véhicules sans chauffeur





## RENDEZ-VOUS

LE MONDE • SCIENCE & MÉDECINE  
MERCREDI 31 OCTOBRE 2018

# Jean-François Bonnefon trace sa route en toute autonomie

**PORTRAIT** - Le chercheur français en psychologie a acquis une renommée mondiale en étudiant les dilemmes moraux posés aux véhicules sans chauffeur

Jean-François Bonnefon est proprement désarmant. Posez-lui n'importe quelle question, scientifique ou personnelle, complexe ou ridiculement naïve, et le psychologue vous répondra avec une franchise, une patience, une pédagogie, une douceur déconcertantes. Le tout accompagné d'une modestie rare. Ses amis vantent son savoir littéraire ? Il admet « *une bonne connaissance des romans de science-fiction* ». Ils notent sa culture musicale tous azimuts ? « *Mon père jouait du jazz, ma mère était prof de musique. J'écoutais des opéras à l'école primaire. Mais j'ai arrêté à l'adolescence, et je n'ai jamais été capable de jouer d'un instrument.* » Ses collègues louent son brio scientifique, sa précocité, ce recrutement au CNRS en 2004, six mois après sa soutenance de thèse ? « *Un concours de circonstances*, jure-t-il. *Le mouvement Sauvons la recherche avait abouti à multiplier par deux le nombre de postes cette année-là. J'étais en liste d'attente, je suis rentré.* » Quant au « *niveau impressionnant* » qu'il afficherait en informatique, au dire même des spécialistes du domaine avec lesquels il collabore, il évacue : « *Je suis né en 1973, j'ai grandi avec les premiers ordinateurs domestiques et j'étais un geek.* »

Tout semble simple, transparent. Prenez la voiture autonome, le sujet qui vaut à l'universitaire toulousain une célébrité mondiale, une invitation d'un an au prestigieux Massachusetts Institute of Technology (MIT) et une récente publication dans la revue *Nature*. Dans l'article « *The Moral Machine Experiment* » (« *l'expérience de la machine morale* »), publié le 24 octobre, lui et ses collègues imaginent des dilemmes moraux imposés au véhicule sans chauffeur et demandent aux humains le choix

qu'ils feraient. Ecraser trois chiens ou percuter un mur et tuer un passager ? Privilégier les enfants ou donner la même valeur à toute vie ? Quid du sexe, de l'état de santé, de la condition sociale, du respect des feux par les piétons ?

## « Une sorte d'illumination »

Depuis quatre ans, Jean-François Bonnefon surfe sur le sujet, publiant dans les plus grandes revues avec toujours une vague d'avance. Son secret ? « *Une curiosité hors du commun et un grand talent pour poser des questions originales* », diagnostique Catherine Casamatta, professeure de finance à l'université Toulouse-I Capitole. Lui, oppose une réponse plus simple : « *J'ai eu une sorte d'illumination. J'ai réalisé que les voitures autonomes étaient à la confluence de tout ce qui m'intéressait : la disruption liée à l'arrivée de l'intelligence artificielle dans le quotidien, la dimension morale des choix qui vont s'imposer à la voiture, l'aspect financier – les automobilistes vont-ils acheter ? – et enfin une question politique : les gens vont-ils vouloir que le gouvernement réglemente ?* »

La trajectoire professionnelle de Jean-François Bonnefon semble, il est vrai, un long et brillant prélude vers cette exposition mondiale. On avait lâché le geek à l'adolescence, entre clavier et jeu de rôle. On le retrouve en maths sup, puis maths spé. « *Là, j'ai réalisé que je n'avais jamais rien choisi, juste toujours suivi la pente la plus abrupte, et que ces études menaient à des métiers dont je n'avais pas envie.* » Il arrête et pendant un an, il papillonne. Biologie, langues, lettres classiques et enfin psychologie. « *Comme 80 % des étudiants, je n'y allais pas pour comprendre les autres, mais pour me comprendre moi-même. Et ce luxe de s'autoriser l'in-*



*trospection me plaisait beaucoup. La discipline était très segmentée: psychologie de l'enfant, psychopathologie, psychologie des groupes. Tout m'intéressait mais sans passion. Puis dans un cours de psychologie cognitive, j'ai découvert que des psychologues étudiaient la façon dont les gens faisaient des maths, de la logique. Pour moi qui aimais les maths, c'était une révélation. »*

Le voilà plongé dans la psychologie cognitive. Il s'affaire à voir si l'humain, l'air de rien, ne suivrait pas les modes de raisonnement privilégiés par les chercheurs en intelligence artificielle. « J'étais en thèse, lui en maîtrise, mais il était déjà brillant. Au point que les deux maîtres de l'intelligence artificielle de l'époque, Didier Dubois et Henri Prade, sont venus écouter sa soutenance », se souvient son collègue et ami Eric Raufaste, professeur de psychologie cognitive à l'université Jean-Jaurès. Ce qui, vu par l'intéressé, donne : « Je voulais tester sur les gens la logique possibiliste qu'ils avaient inventée, mais je ne comprenais pas leurs thèses. Ils ont été d'une patience incroyable, sans doute attendris par ma naïveté. »



Depuis, Jean-François Bonnefon a plus impressionné qu'attendri. En 2008, dix ans après cette fameuse maîtrise, il décroche la médaille de bronze du CNRS, le Graal des jeunes chercheurs. Le psychologue n'a pas tout à fait abandonné les machines intelligentes, mais il a surtout creusé le filon humain et la théorie de la décision, la perception du risque, des incertitudes. « *Par exemple, comment les gens comprennent la notion de possibles effets secondaires...* »

Il multiplie les terrains de jeu, accumule les expériences, dirige notamment le laboratoire de psychologie et de linguistique de l'université Jean-Jaurès. Personne ne rêve encore de voiture sans chauffeur, mais lui se demande déjà, avec son collègue Bastien Trémolière, comment la pensée de la mort affecte les choix des gens. Ou encore la part d'intuition et de réflexion dans nos décisions.

#### Analyse des choix

La psychologie morale, le sujet d'une vie pour certains. Sauf que la vie est courte et foisonnante. Jean-François Bonnefon part donc la croquer chez les voisins de l'École d'économie de Toulouse (TSE). « *Ils ont des thèses rafraîchissantes*, explique-t-il. *Les points de départ sont toujours très simples, ensuite on raffine.* » Le voilà qui étudie la rationalité des décisions, incluant cette fois des variables financières ou la capacité à prédire un choix en observant le visage de l'intéressé. Il monte un site Web participatif articulé autour d'un jeu de rôle. Une expérience qu'il mettra encore à profit lorsqu'en 2016 il lancera Moral Machine, ce site viral sur la voiture autonome, qui a rassemblé plus de 2,5 millions d'internautes de plus de 200 pays, impressionnante base de données et support du récent article de *Nature*. Cette dernière aventure ne concentre pas l'intégralité de son temps. Trop

de projets débordent de ses cartons. Dernier en date, lancé cet été, juste avant son départ pour Boston : une demande de financement européen sur les « nouveaux dilemmes », ceux qui se posent dans la gestion des priorités pour les dons d'organe (âge, délai d'attente, condition médicale...), ou dans l'évaluation des risques de récidive avant des remises en liberté conditionnelles de détenus. « *Ce sont toujours des arbitrages entre équité et efficacité, et toujours des algorithmes d'intelligence artificielle qui imposent aux humains de faire des choix* », explique-t-il.

Pas si loin, donc, de la voiture autonome, cœur de son escapade américaine et responsable de son nouveau statut. Tel un jeune footballeur français prêté pour un an au FC Barcelone, le voilà à la conquête du MIT, le saint des saints de la recherche mondiale. Il faut entendre ses nouveaux collègues vanter « *sa rapidité* », « *son originalité* », ou encore, comme le dit l'informaticien Iyad Rahwan, « *sa maîtrise étonnante des outils statistiques, pas si fréquente chez les psychologues* ». D'une pirouette, il évacue une nouvelle fois : « *Je n'avais jamais fait de postdoc [recherche postdoctorale] ni quitté Toulouse. Alors, comme ça, on cessera de me prendre pour le péquenaud du coin...* »

Dans ce paysage d'une déconcertante honnêteté, Jean-François Bonnefon avait toutefois oublié un détail, que l'un de ses collègues a fini par nous confier : il ne conduit pas. Presque un lien d'intérêt pour une vedette des voitures autonomes ? Sa réponse est venue par mail. « *Haha. C'est vrai, j'ai le permis, mais je n'ai pas pris le volant depuis quinze ans. Je suis trop dangereux, et ma famille me l'interdit. Je suis le client parfait pour les voitures autonomes, qui quel que soit leur niveau de sécurité, conduiront toujours mieux que moi.* » ■

NATHANIEL HERZBERG



Jean-François Bonnefon, le 18 octobre.

M. SCOTT BRAUER POUR «LE MONDE»