« Police prédictive » : les belles histoires de l'Oncle Predpol

Xavier RAUFER Hiver 2015

1. "Police prédictive" : bobards et réalité

Le scénario est constant. Soudoyé ou naïf, mais toujours extatique, un journaliste annonce la Bonne Nouvelle. À l'automne 2015 encore, c'est "Kevin", du site *Science Post*. Le titre : "Une intelligence artificielle capable de prédire les crimes est en développement" – pas moins. Suit l'inévitable référence à "*Minority Report*" et un déluge de mots techniques visant à fasciner et sidérer le lecteur. On va "arrêter les criminels... avant qu'ils ne commettent de crime". De la science-fiction ? Non, c'est pour demain.

Sauf qu'aujourd'hui, la vraie incertitude n'est pas plus modélisable qu'au temps d'Aristote (sur lequel repose encore notre acception du temps). En clair : une base documentaire contenant tout sur le terrorisme depuis l'antiquité jusqu'au 10 septembre 2001, ne prédira en rien les attentats du 11 septembre. On rougit de devoir rappeler que si le connu d'hier résolvait mécaniquement l'inconnu de demain, tout le monde gagnerait à la loterie...

À les lire, les naïfs propageant ces bobards pour le compte de marchands de *software* semblent ignorer le fonctionnement des logiciels de "police prédictive". Aidons-les : des algorithmes brassent des données sur des crimes passés pour en annoncer de futurs. Or – insistons – cela est absolument *impossible*. Que font vraiment ces logiciels ? Ils assument que, comme hier un crime s'est commis à tel endroit, demain, il s'en commettra peut-être un autre. Ce raisonnement n'est pas de la prédiction, mais du *wishful thinking*.

Et quelle crédulité chez ceux qui diffusent ces contes de fées : en août 2015, "Le Parisien" nous assénait ainsi ainsi qu'à Munich, l'usage d'un logiciel de "police prédictive a fait "baisser les cambriolages de 30%". Vérifions, avec les statistiques officielles du *Bundeskriminalamt* (Office fédéral allemand de police criminelle). De 2003 à 2012, les cambriolages augmentent de 16,9% en Allemagne. Mais en 2014, alors qu'on installe à Munich le fameux logiciel-miracle ? Augmentation encore, de 1,8% – pas même une stagnation. Clairement, l'usage de la baguette magique "prédictive" fait juste jouer l'effet de déplacement : se sentant observés, les cambrioleurs changent de crèmerie et partent piller la ville d'à côté. Mais à l'échelle nationale, effet zéro.

Au mieux, le logiciel-miraculeux permet à la police d'ordonner son travail et d'intervenir plus souvent au bon endroit et au bon moment – mais rien de "prédictif" là-dedans, bien sûr. En prime, cet effet positif est forcément temporaire. Depuis l'âge des cavernes, le réflexe humain est constant : s'il se sent observé, l'homme modifie son comportement – qu'on le regarde à l'œil nu ou électroniquement. Ainsi, vous payez des dizaines de milliers d'euros un logiciel-miracle qui, six mois plus tard, n'annonce plus rien de pertinent. Cela s'appelle une arnaque.

La recherche américaine de pointe confirme ce diagnostic. Discrètement, IARPA lançait en mars 2015 une étude sur un outil informatique visant à "prévoir des événements rares". Rappelons ici qu'IARPA (*Intelligence Advanced Research Projects Activity*) est le laboratoire *high tech* du renseignement des États-Unis. Si IARPA cherche cet outil informatique, c'est bien sûr qu'il ne l'a pas. Or s'il était vrai qu'"une intelligence artificielle capable de prédire les crimes soit en développement", IARPA aurait fait 95% du chemin, car les infréquents crimes graves sont de ces "événements rares" que le renseignement US aimerait prévoir.

Conclusion : les actuels "logiciels de police prédictive" naïvement vantés par certains médias sont considérés comme de la daube par l'élite américaine de la recherche scientifique de renseignement. Ce qu'il fallait démontrer.

2. Plus largement, l'hyper-technologie peut-elle remplacer l'intelligence humaine ?

Il existe en effet des idolâtres de l'intelligence des machines, vantant la capacité des ordinateurs à tout résoudre bien plus vite que la pauvre cervelle humaine. Or face à de tels élans amoureux, le vrai problème ne réside pas dans le progrès technique – Merveilleux ? inquiétant ? Il est là, c'est ainsi. Nul ne veut revenir à la pierre taillée et l'auteur moins encore, enchanté qu'il est du confort de travail offert par cette réelle prothèse du cerveau hu-

main qu'est l'informatique, de bureau ou portable.

Le problème n'est pas non plus dans l'exaltation mystique de Silicon Valley devant ses cyber-créations, l'idée qu'un jour Super-Google règnera sur le monde, empêchant une bête humaine enfin domptée de sévir comme elle le fait depuis toujours. Les Californiens ont l'âme techno-religieuse : laissons-les jouer avec leurs cyber-sectes à la "Singularity".

Le vrai problème, encore, est le traitement que la presse-des-milliardaire (les quotidiens rachetés par de grosses fortunes) réserve à ces affaires. Devenus les serviteurs de milliardaires de la nouvelle économie, d'ex-journalistes désormais voués aux relations publiques réservent un accueil extatique à tous ces progrès techniques, e-commerce, etc. – par une amusante coïncidence, à tout ce qui par ailleurs, enrichit les magnats qui les emploient.

Exemple parmi d'autres : la publicité faite dans ces médias pour les logiciels de "police prédictive" – en fait ni plus ni moins efficaces que ceux des sites de rencontre en ligne, car usant des mêmes algorithmes. Là encore, nulle critique ! L'extase hi-tech, obligée et unique voie du progrès. Qui héberge un doute n'est qu'un ringard.

Pour être édifié, il suffit de lire la presse-des-milliardaires depuis qu'en mars 2014, le vol MH 370 de Malaysian Airlines a disparu corps en bien. Depuis lors, nous avons vécu en direct l'effarant échec d'*Eye in the sky*, Washington en étant réduit à implorer les internautes d'aller scruter les océans. Et bientôt deux ans après le drame on ne sait en fait, toujours rien du sort de cet aéronef. Or là-dessus, la presse dite d'information se tait.

Dans un premier temps, on crie au miracle. Vient le désastre et les mêmes regardent ailleurs. On nous a certes beaucoup parlé de cet avion et du mystère de sa disparition. mais pas un mot sur le fait que les merveilleux systèmes High-tech, soi-disant capables de lire, depuis le fin fond des cieux, la marque d'un paquet de cigarettes jeté au sol, ont manifesté une totale impuissance à retrouver un avion de 300 tonnes et de 70 mètres de longueur et d'envergure. Le fait d'avoir enseveli dans le silence cet échec flagrant d'un high-tech que ces médias encensent chaque jour, là est le vrai scandale.

3. Extase médiatique et "police prédictive"

Dans le domaine de l'immobilier, avez-vous aimé l'arnaque des *Subprimes* ? Alors, vous devriez adorer la suivante, celle de la "police prédictive" en matière de sécurité. Cette arnaque emballe bien sûr les médias. Le *Journal du Dimanche* s'ébahit devant "La machine à déceler le crime", *Le Monde Magazine* vante "Le logiciel qui prédit les délits", etc. A y regarder d'un

peu près, l'enthousiasme semble moins spontané qu'il n'y paraît, puisque tous ces articles sont quasi-identiques et sentent la com' à un kilomètre.

Indice sérieux, ils évoquent pour la plupart le film *Minority Report*, tiré d'une nouvelle de Philip K. Dick. Film dont à l'évidence, nos enthousiastes journalistes ignorent tout, car il n'a rien à voir avec le "*predictive policing*" ! Dans l'œuvre de Dick, des "voyants" ("*Precogs*"), pressentent d'imminents homicides et alertent ensuite la police, ce qui est tout différent.

Venons-en aux cyber-Madoff prétendant "prédire" les crimes, attentats, etc., pour souligner d'emblée cette imparable évidence (développée plus bas) : aujourd'hui encore, et durablement, certains phénomènes irréductiblement complexes sont *imprédictibles* ; exemple, les séismes. D'autres phénomènes à risque seraient-ils, eux, prévisibles ? Au mieux, faiblement – voyons plutôt.

Dans les domaines où l'on pense savoir prédire, et où on le fait de longue date : économie-finance, météorologie, séismes, etc., le bilan de la prévision est plutôt catastrophique, dès lors qu'il s'agit de risques. Prouvons-le : avant la crise des *Subprimes* les agences de notation (Standard & Poor, Moody's,etc.) disposaient toutes d'estimations chiffrées du risque de défaut de remboursement par les acheteurs de maisons à crédit.

Ces risques étaient alors estimés grâce à l'analyse quantitative, discipline issue de la physique des probabilités, supposée maîtriser "scientifiquement" les risques de *trading*. Or quand éclata la crise, le risque réel se révéla... *deux cent* fois pire que les prévisions des agences ! Se croire hors risque grâce aux estimations des agences de notation, ricana alors un expert, revenait à se protéger du feu nucléaire en s'enduisant de crème solaire...

Ceci n'empêche pas les Cyber-Madoff de nous offrir le nouveau miroir aux alouettes qu'aux États-Unis, on nomme "Crime predictor" ou "Predictive Policing" (*Pred Pol* pour faire *cool*). Il s'agit de "logiciels d'analyse prédictive" supposés prévoir les crimes, et même, pourquoi pas, les crises politiques, les révolutions ou les attentats.

Lisons plutôt les articles que la presse leur consacre : "Le logiciel qui prédit les délits... XXX (*la marque du logiciel*) débarque au Royaume-Uni. Un logiciel qui "prédit à quel endroit les criminels peuvent attaquer, et quand", ou qui peut "prédire où des cambriolages, vols, agressions, vont se produire dans l'avenir... Avec des résultats convaincants."

Or les Cyber-Madoffs qui, pour des médias naïfs, prétendent "prévoir les crimes", usent de l'analyse quantitative et des "algorithmes prédictifs" mêmes qui ont provoqué l'effondrement de Wall Street ! Mais au fait, comment nourrit-on ces "logiciels prédisant les crimes" ? Avec un

"algorithme conçu pour prédire où et quand un crime va se produire, grâce à une base de données *recensant les infractions passées*" ; ou par "l'utilisation de *statistiques historiques*" ; ou en usant de "bases de données criminelles *datant des années 1960*". Tous ces logiciels tirent sans exception leurs références du passé. Leur carburant est *strictement* rétroactif.

On ramasse tout ce qui traîne sur Internet en habillant cette récolte, opérée plus ou moins à l'aveuglette, de noms pompeux-high tech : *Big Data*... *Data Mining*... puis l'on passe cette matière brute dans une moulinette à algorithmes. Résultat plausible : sans doute une possible et momentanée optimisation du travail policier – en attendant que les bandits fassent, comme d'usage, jouer l'effet de déplacement en leur faveur – mais capacité de prédiction, certainement pas – et voici pourquoi.

Car posons ici la question cruciale – qu'ignorent (sciemment ou non) nos enthousiastes médias : ce qu'on sait du passé permet-il de prévoir le futur ? Par exemple : le temps qu'il faisait hier certifie-t-il celui de demain ? D'évidence non, du fait de *l'incertitude*, immuable et non-négociable élément de toute prévision – que les Cyber-Madoff et leurs thuriféraires se gardent bien d'évoquer. Et même, les choses ne risquent pas de s'arranger dans l'avenir, par amélioration progressive des systèmes type *Pred Pol* – mais ne peuvent à l'inverse qu'empirer. Pour les experts en effet, les algorithmes et logiciels dits "prédictifs" possèdent une capacité forte à pré-formater la réalité, à influer sur elle – suscitant donc leur propre validité et ainsi, persuadant le client que tout marche au mieux - alors que sur le terrain, dans la vraie vie, là où opèrent les vrais voyous et dealers, rien n'a vraiment changé.

Mirage technologique, ces "logiciels de prédiction des crimes" ne font en fait que prolonger des courbes. Ils partent pour cela d'une idée simple : ce qui était là hier, sera là demain. Comme lundi, des bandits ont opéré dans tel coin, montons mardi une embuscade là où le logiciel l'indique et hop ! On les pince à coup sûr. Or cela ne relève nullement de la prévision, mais de la pétition de principe.

Enfin, les algorithmes type *Pred Pol* servent aussi à des prestations d'un tout autre genre : les rencontres en ligne sur Internet. À leur propos, voici l'avis d'un expert : "les gens paient pour ces services de rencontres quoique, après examen, les algorithmes supposés trouver leur partenaire idéal ne marchent sans doute pas". Après les amoureux numériques, les Cyber-Madoff ciblent maintenant les policiers et les élus. Bienvenue au club des gogos.

4. Plus largement : peut-on tout prédire ?

Parmi les pires excès de cette cyber-propagande, le domaine du prédictif. De gros

ordinateurs, des algorithmes appropriés, le *Big Data* – tout cela réuni en un dispositif approprié, permettrait de prédire tout et le reste... Grâce à Wikipedia par exemple, et à un modèle mathématique approprié, on pourrait désormais annoncer *avec précision, un mois à l'avance,* le succès d'un film au *box-office*. Une recette également utilisable dans l'agroalimentaire, pour les sodas et les sandwiches nouveaux.

Dès 2030 annoncent même certains gourous, nous disposerons de planétaires machines-à-prédire ("*global precognition machines*") pour débarrasser le monde de tout le négatif qui l'encombre. D'ores et déjà, racontent des VRP du logiciel, des spécialistes disposent du pouvoir – l'expression est trop belle en anglais – "*to predict whether you're going to click, buy, lie or die*". Pour ces VRP, prédire le comportement humain permet de prendre de meilleures décisions, dissipe les risques financiers, renforce la santé publique, détruit le *Spam*, stimule les ventes et bien sûr renforce la lutte contre le crime. (Rien à ce jour sur un quelconque effet vermifuge, ou capacité à guérir les écrouelles. Mais espérons.)

Intéressons-nous donc plus précisément à la "prédiction du crime". Depuis 2010 environ, des articles éclosent dans les médias papier ou électroniques du monde anglo-saxon ; tous nous racontent la même histoire. Bientôt, il en ira du crime comme de la météo... Des logiciels spécialisés comme PredPol permettront à la police de pré-dire les crimes. Et ça marche déjà en Californie ! "Quand le système a été testé, la criminalité constatée à baissé de 12% et les cambriolages, de 27%". Et tout le monde se précipite : Silicon Valley bien sûr, les militaires, l'université, l'édition et même Hollywood.

Les livres d'abord : "*Predictive analytics for dummies*" est assez caractéristique du lot. Sur le ton de l'enthousiasme, de l'affirmation sans nuances, les auteurs nous présentent le Saint-Graal. Mais que contient en fait ce livre ? Des banalités-marketing pour école de commerce. On y prolonge des courbes, rien de plus. Surtout, dans une table des matières détaillée de 7 pages, et dans un index qui en fait le double, on ne trouve rien, pas un mot, ni sur le *temps*, ni sur la *temporalité* – concepts d'évidence cruciaux dans tout le domaine du "pré" (prédire, prévoir, présumer, etc.).

Second ouvrage typique : "*Predictive analytics*", œuvre de marchand de gadgets où ce qu'est une *prédiction* n'est jamais défini. Car observer des comportement, ou des habitudes, user de sa jugeote, optimiser, prolonger les courbes, estimer des probabilités : est-ce cela, prévoir ou prédire ?

Exemple, quand un magasin voit une cliente acheter une robe de grossesse, est-il besoin de mobiliser un super-calculateur pour lui proposer ensuite un biberon et des couches ? De même, quand un lecteur commande un roman policier sur un site,

faut-il un cyber-génie pour lui en suggérer d'autres ? N'est-on pas là plutôt dans le classique "*Customers who bought this item also bought...*" des sites marchands ?

Car voilà comment fonctionne ce dispositif plutôt grossier, qui tente de donner du sens à des données diffuses, chaotiques et recueillies en vrac. Un système qui a tout à voir avec de l'optimisation, ou du *marketing* de bon sens, mais rien avec de la prévision. Cas concret : avec la vitesse et la puissance de l'informatique, ici, indéniablement supérieures au cerveau humain, un logiciel trie en quatre catégories les clients du *e-commerce* :

A. Ceux qui achètent un produit en négligeant la publicité : *laissés de côté,*

B. Ceux qui n'achètent un produit qu'en l'absence de publicité : *laissés de côté,*

C. Ceux qui flânent sur Internet sans acheter en ligne : *laissés de côté,*

D. Ceux qui n'achètent rien sans publicité, mais achètent s'ils en voient : ces "réceptifs confirmés" sont placés dans une base de donnée spécifique, puis bombardés de pub'.

La répétition du mot "*predictive*" dans *Predictive analytics* veut exercer un effet hypnotique sur le lecteur. Mais rien n'y permet de vraiment prédire quoi que ce soit. Et puisque les collections portant ce titre sont à la mode, on est plutôt ici dans "le behaviorisme pour les nuls".

Après les livres, la stratégie : on apprend ainsi en novembre 2013 que l'armée américaine expérimente en Afghanistan un "nouveau modèle prédictif", *Global Database Events,* conçu par un professeur de sciences politiques de l'Université d'Etat de Pennsylvanie. Ce logiciel "recueille des nouvelles sur Internet" et "catalogue des événements, des élections locales jusqu'aux génocides". Il en extrait des "prédictions à court et à long terme", utiles pour "gérer les crises" et "prévoir les niveaux de conflits en Afghanistan".

Les séries-télé se joignent à la meute : diffusé sur TF1, "*Person of Interest*" a eu en 2012, 14 millions de téléspectateurs en moyenne sur CBS-États-Unis. Dans cette série "un génie de l'informatique invente une machine sachant déjouer les attentats", qui "prédit aussi les crimes crapuleux". Quelle machine sympa ! Un "ordinateur intelligent et même doué d'émotions... une intelligence artificielle touchante"... Aux prises avec "le gouvernement, la mafia, les policiers corrompus de New York", son héros "prévient bien sûr les crimes avant qu'ils ne soient commis". Une ambiance, dit la critique, "à la *Minority Report*". Notons la référence à ce film, qui est l'immuable marqueur de l'escroquerie au prédictif. Même les universités américaines, se ruent désormais pour proposer aux étudiants "des certificats ou Masters en analytique prédictive".

"Prédire le crime" : ce que racontent étourdiment les médias

Pour tous ces articles, la "police prédictive" "réduit la criminalité par l'analyse de données sur des crimes et le lieu de leur commission"... "S'agissant de personnes à risque, la méthode prédictive peut être efficace"... "Par voie électronique, le Maryland génère des prédictions sur les malfaiteurs placés en liberté provisoire, pour savoir qui tuera et qui sera tué"... "Des chercheurs et des policiers ont conçu des systèmes prédictifs qui, parmi les individus déjà condamnés pour homicide, prédit lesquels tueront encore"... "Grâce à des analyses informatiques sophistiquées, on peut prédire où et quand des crimes seront commis"... "Des programmes de police prédictive fondés sur des algorithmes et des données historiques, estiment la localisation et la nature de crimes futurs"...

Dans ces articles, notons une forte odeur de "communication", pas une critique – au point qu'on dirait plutôt de la publicité rédactionnelle : "Le logiciel de police prédictive est deux fois plus efficace qu'un analyste humain disposant des mêmes données"... "Dans un secteur de Los Angeles, un outil de police prédictive a fait, en 5 mois, baisser d'un tiers les cambriolages"...Dans tous les États-Unis, des dizaines de services de police ont déjà acquis de tels dispositifs"... Pour *The Police Chief*, organe de la puissante *International Association of Chiefs of Police*, "la police prédictive marque le début d'une ère nouvelle".

De mieux en mieux : la "justice prédictive"

Décidément mis à toutes les sauces, le "prédictif" a également contaminé la justice. Car de fait les magistrats s'interrogent fort légitimement : ce détenu en instance de libération peut-il commettre un crime grave dans les prochaines années ? Ce primo-délinquant a-t-il l'étoffe d'un multirécidiviste ? Or d'usage, ces juges n'ont que leur expérience, ou leur flair, pour les aider. Et si des logiciels "prédictifs" pouvaient contribuer à sélectionner les détenus avec pertinence ? orienter leurs choix pour des peines avec sursis, des libérations conditionnelles, sans risque pour la société ?

Bref : peut-on concevoir un système objectif permettant *à la fois* de multiplier les libérations conditionnelles et de réduire le nombre des récidives ? De bien choisir les détenus envoyés dans des programmes de réinsertion, durant la peine ou après celle-ci ? De décider (dans un système de *common law*) qui reste incarcéré, et pour combien de temps ? De détecter les jeunes détenus influençables à ne surtout pas incarcérer avec des criminels confirmés ?

Aux États-Unis, 4/5^{e} des comités d'évaluation de libération sous condition utilisent désormais de tels logiciels "prédictifs". Qu'y introduit-on pour évaluer un détenu ? Son âge, son sexe, ses antécédents ; la date de sa première arrestation, son éducation, les infractions qu'il a com-

mises, son comportement en prison ; aussi, le casier judiciaire de ses plus proches amis et le résultat des tests psychologiques qu'il a subis ; et même, l'éthylisme de sa mère durant sa grossesse. Ensuite, on compare avec des profils analogues.

Que des décisions soient prises sur la base de tels calculs peut inquiéter mais – se doit de préciser ici le criminologue – un peu moins cependant quand on a lu l'effrayante étude "Extraneous factors in judicial decisions" (Princeton, février 2011) qui, ayant comparé des centaines de cas, démontre qu'en fin de matinée, des juges américains affamés tendent à condamner plus sévèrement que quand ils sont repus, deux heures plus tard...

Aléa informatique, aléa stomacal... Aléa toujours. Il demeure que tout ce qui relève de la justice, ou de la police, "prédictives" repose exclusivement sur du rétrospectif : "Les modèles prédictifs analysent des données *historiques*...", "Nous collectons toutes ces données dans des occurrences passées...", etc. Et là, commencent les problèmes conceptuels.

5. Comment fonctionnent les logiciels prédictifs ?

Vu le prix, 73 000 dollars pour acquérir un logiciel de "police prédictive", plus ensuite 45 000 dollars d'abonnement annuel, il importe de voir comment tout cela marche. Là encore, tous les articles publiés fournissent les mêmes explications. On commence toujours par saisir "toutes les données sur des crimes commis, et leur localisation, remontant à l'année XXX" ; ou bien "on analyse les dossiers d'environ 1,5 million de crimes commis de 2003 à 2012"... Pourquoi ? Parce que "les infractions à venir seront souvent commises sur les lieux d'anciens délits ". Là, le criminologue sursaute pour la première fois.

Les algorithmes maintenant. Docilement, les journalistes ou publicitaires ayant rédigé les articles sur la "police prédictive" répètent la même leçon : "l'algorithme comprend *(au sens de "digère")* le schéma criminel et produit une prédiction"... Ici, second sursaut du criminologue.

Preuve, ajoutent ces thuriféraires, que l'affaire est sérieuse "ces algorithmes sont fondés sur les modèles de prédiction des séismes" *(variante)* sur les modèles de prédiction des répliques des séismes". Une "preuve" d'efficacité qui sans cesse revient, article après article. Là, le criminologue effectue un véritable bond car il sait, et le prouvera plus bas, que *les séismes sont, à ce jour, absolument imprédictibles*, ce que nul des rédacteurs des contes de fées précités ne s'est un instant avisé de vérifier.

Mais bien sûr, cela n'empêche pas l'Amérique du Sud de se lancer dans l'aventure : Au Chili, des ingénieurs "mê-

lant la criminologie à la modélisation mathématique", ont entrepris de prédire en quels points de la frontière terrestre (en effet gigantesque, 6 170 kilomètres) du pays, se produiraient des crimes et des migrations illégales. Au Brésil, au printemps 2014, la ville de Sao Paulo acquiert le système *Detecta* de Microsoft, qui permet de combattre le crime en "agrégeant des données" et en "créant entre elles des associations automatiques". Les policiers de la ville sont dotés d'ordinateurs portables, tablettes et *smartphones* pour accéder au système, et orienter ainsi leurs patrouilles dans un esprit préventif.

Mais pourquoi se borner à l'Internet ? Patrouiller sur les réseaux sociaux peut aussi contribuer à "prédire le crime". Comment ? Pour le *Predictive Technology Lab* de l'Université de Virginie, *Twitter* peut en effet "prédire" certains types de crimes. Dans la revue scientifique *Decision Support Systems* de mars 2014, les chercheurs de ce laboratoire affirment que des *Tweets* géo-localisés (qu'on peut donc situer finement), peuvent "prédire de 19 à 25 types d'infractions : harcèlement, vols, agressions...". Comment cela ? "Si assez de *twitteurs* annoncent vouloir se saouler dans le même secteur, *alors* on peut y "prédire" des infractions liées à l'alcoolisme". Il faut en outre comparer ces analyses *Twitte*r avec "les concentrations d'actes criminel historiquement fortes". Cette opération relève-t-elle un seul instant de la prévision ? Non, on le verra plus bas.

Résumons : l'analyse prédictive fonctionne selon le dispositif suivant : *data mining* (recherches de données présentes sur Internet, mais souvent cachées) + statistiques + algorithmes sophistiqués et logiciels spécifiques (*mining tools*) = modélisation, supposée pré-dire.

• *Ces logiciels sont-ils sérieux ?*

En juger nécessite de préalablement fournir au lecteur des éléments contextuels pertinents. Les voici.

– D'abord, le point le plus décisif. Aujourd'hui et pour longtemps encore, l'incertitude... l'entropie.. l'aléatoire.. le désordre (au sens scientifique), relèvent de l'imprévisible et de l'imprédictible. Pour le dire autrement, tout ce qui est à un moment donné possible, n'advient pas nécessairement. Ce que Donald Rumsfeld, ministre américain de la défense lors de la guerre d'Irak, exprima fameusement en ces termes : "Il y a le *connu-connu*, le domaine de ce qu'on sait et sait savoir. Il y a le *connu-inconnu*, ce qu'on sait ne pas savoir. À quoi s'ajoute *l'inconnu-inconnu*, ce qu'on ignore sans le savoir".

Affirmer que l'observation algorithmique permet de modéliser ou de standardiser l'inconnu-inconnu relève clairement de l'escroquerie intellectuelle. Car *le rétroactif n'est jamais prédictif* : sinon, avoir en sa possession la liste de tous les

tirages passés d'une loterie permettrait de gagner au coup suivant. Or si tout ce qu'on introduit dans une moulinette algorithmique provient du passé, il n'en sort qu'une probabilité, souvent maladroite, prolongeant plus ou moins grossièrement des courbes : les criminels étaient là hier à telle heure, ils y seront donc aussi demain.

– Depuis qu'existent les premiers ordinateurs, l'idée de prédire taraude les savants. Au début de la seconde Guerre Mondiale, Norbert Wiener, père de la cybernétique, tenta-t-il de concevoir un modèle prédisant les évolutions des avions de chasse allemands, pour pouvoir les abattre plus aisément. Ce fut un échec, par manque, dit-on alors, de capacités de calcul.

– Depuis trente ans et plus, des organismes de recherche militaire, comme la Darpa aux États-Unis,cherchent, à coup de millions de dollars, comment agréger, combiner, maintes données d'apparence déconnectées, hétéroclites ; pour y découvrir des corrélations permettant de poser des diagnostics ou de faire des prédictions, par exemple sur de futures émeutes, de futurs attentats.

En 2010, la Darpa a lançait ainsi un programme "Data to decisions" doté d'un budget de 92 millions de dollars. Il visait à concevoir un algorithme permettant de connecter, exploiter, donner un sens, anticiper ; à partir des masses d'informations stockées, commercialisées, échangées dans l'Internet. Avec toujours le même rêve : prédire l'agitation sociale... Les attaques terroristes...les événements stratégiquement significatifs. Or à en juger par les présents soubresauts affectant la politique étrangère américaine, de l'Afghanistan à l'Irak, la Darpa ne semble toujours pas avoir trouvé cet équivalent prédictif de la pierre philosophale...

Rien de moins neutre qu'un algorithme...

Les ordinateurs et l'informatique ne sont pas neutres. Et les outils de collecte et d'analyse des données ne le sont pas plus. Les algorithmes ne sont pas le mètre-étalon, mais renferment les biais de leurs créateurs, qui sont des humains faillibles et ont donc pu y instiller du *wishful thinking* en lieu et place de science...

Car ces algorithmes, outils de compilation du monde que les naïfs entourent d'une adoration quasi-théologique, peuvent avoir été truqués par leurs créateurs (en leur faveur à eux) ; ou par des pirates, facétieux ou stipendiés. Entrevoit-on alors ce qu'il advient du dispositif de "police prédictive" ainsi "arrangé" par des *hacktivists* embusqués dans le *dark web*? Les patrouilles de police envoyées là où rien n'arrive ? Et à l'autre bout du quartier, les cambrioleurs fort tranquilles pour agir ? Ici, insister serait cruel.

Plus largement, il est ardu de vérifier si des algorithmes remplissent vraiment leur tâche, du fait de leur capacité à influer sur la réalité, à la pré-formater ; de par leur masse, leur puissance même. Si leur usage est assez massif, ces algorithmes suscitent leur propre validité, exercent un effet "banc de poisson" sur les faits matériels. Tout cela, les médias devraient le savoir, car en matière de truandage d'algorithmes, les exemples récents ne manquent pas :

– Dans la décennie 1970 déjà, le "Black-Scholes Model" "prédit" la valeur future des actions. Mais en 1998, algorithmes géniaux ou pas, le *hedge-fund* Long Term Capital Management s'effondre, conduisant au bord du grouffre le marché mondial du crédit,

– En 2001, un modèle trafiqué – bien sûr fondé sur de fascinants algorithmes – permet à la société Enron d'attribuer une valeur astronomique à des actifs évanescents. Puis Enron s'effondre et ses dirigeants filent durablement en prison.

– Lors de la crise des subprimes, on découvre que des agences de notation "adaptent" leurs logiciels (reposant, on l'aura compris, sur d'ésotériques algorithmes) au résultat souhaité.

– Après ladite crise, le géant bancaire JP Morgan doit "s'excuser" de l'usage d'un logiciel "inadapté" – fondé sur des algorithmes sophistiqués – ayant pourtant conduit la banque à perdre 6 milliards de dollars,

En 2008, trois importants *hedge-funds* subissent des pertes énormes, du fait de "mouvements imprédictibles du marchés". Mouvements que les magiques algorithmes de ces *hedge-funds* étaient censés prévoir...

Au-delà, les algorithmes peuvent carrément permettre de juteuses escroqueries. Dès 2005 un businessman texan prétendument "ex-officier du renseignement militaire" et professeur d'université, se disait l'inventeur d'un algorithme permettant de faire fortune sur le marché des devises étrangères. Il escroque 33 millions de dollars à ses clients naïfs puis écope de 20 ans de prison. Mais venons-en au fond. La "police prédictive" est-elle sérieuse ?

Résolvons d'abord l'affaire de la "prédiction des séismes". Voici ce qu'en dit le géophysicien Bill Ellsworth, chercheur à *l'United States Geological Survey* : "Nul ne sait prédire les tremblements de terre... Si les séismes sont prédictibles, nous ne savons pas le faire. Et ils sont sans doute pour de bon imprédictibles... Nous ignorons même quelles équations régissent la mécanique des séismes..." Commentaire de l'interviewer "Un savant comme Bill Ellsworth, qui *sans limites financières*, (nous soulignons) dispose de toutes les données et de toutes les capacités informatiques possibles, admet qu'il ne voit pas vraiment comment prédire les tremblements de terre".

Or de même, le milieu criminel s'inscrit clairement "dans le monde réel, là où les interactions humaines complexes ne peuvent pas toujours être saisies, même par les modèles les plus sophistiqués", là où les outils d'analyse statistique tendent à produire des résultats dénués de sens. "Nous ne pouvons donc plus nous fier aux expériences en laboratoire pour une analyse de causalité".

L'homme, le crime et la machine

D'abord, une remarque, banale pour tout criminologue, mais qui échappe aux informaticiens et aux thuriféraires de *Big Data*. L'activité humaine n'est pas une ressource naturelle qu'il suffirait d'extraire du *cloud* pour le rentabiliser ; elle n'est pas une sorte de charbon, ou de pétrole, qu'on peut exploiter à son gré. Car – millénaire réalité ! – l'être humain, criminel ou non, *ne se laisse jamais observer passivement*. Dans son cerveau reptilien – dans ses gènes – l'imémoriale trace des millions d'années passées à éviter, pous survivre, les prédateurs équipés de griffes et de crocs mortels. Or l'être humain n'a rien de ces dangereux appendices. Ni de cornes, sabots ou carapace. Pour seule arme, l'homme a son gros cerveau ; il s'adapte, se cache, ruse ou triche : il *réagit*.

Constamment, l'homme joue contre l'observateur/ordinateur. Il est actif. Extraire le *Big Human Data* n'est donc pas une banale activité minière type extraction de charbon ou de pétrole – c'est une partie d'échec, voire un match de boxe.

Dans le domaine de l'illicite, un exemple sur ce point, si parlant. En France, les autorités multiplient les radars au long des routes et autoroutes ? Les conducteurs s'adaptent en trichant. Les usurpations de plaques d'immatriculation des véhicules (ce que "*flashent*" les radars) explosent littéralement : en 2010 : + 98% de ce type d'infractions constatées ; en 2011, + 73%. De 2010 à 2012 on passe de 5079 de ces "usurpations", à 17 479...

L'utilité de Big Data et sa capacité à prévoir

Si donc, la "police prédictive" est clairement illusoire dans sa conception actuelle ; ni plus ni moins efficace que les sites de rencontres en ligne (usant d'analogues algorithmes...). La voie du prédictif n'est pas pour autant bouchée :

- Les outils de *data mining* sont utiles dans le vaste domaine du connu-connu : nul mieux que l'informatique ne sait trier, classer, ordonner – donc *optimiser* ; ce, à toute vitesse,

- Les logiciels spécialisés améliorent les diagnostics médicaux, ou l'effectivité thérapeutique ; *idem* pour le ciblage publicitaire ou l'estimation des primes d'assurances. Là dedans, un travail croisé entre informaticiens et criminologues susciterait à coup sûr

des outils policiers préventifs fort utiles,

– L'exploration du Big Data permet aussi de pré-voir les tendances nouvelles ; de révéler des émergences, d'entrevoir des dynamiques discrètes ; ce qui, en matière de sécurité est d'évidence utile.

Dans l'attente de percées technologiques à ce jour inouïes, aller plus loin serait périlleux.