

# **BIG DATA ET ACTION PUBLIQUE**

## **L'EXEMPLE DE LA POLICE PRÉDICTIVE**

Rapport de Boris CHABANEL - NOVA 7  
Mars 2016

# SOMMAIRE

<b>1. Vers une « action publique algorithmique » ?</b> .....	<b>3</b>
Le big data fait apparaître des personnes aux yeux de l'administration.....	3
Les traitements qui donnent du sens aux données collectées deviennent beaucoup plus individualisés et prédictifs .....	3
Des décisions publiques peuvent dorénavant être prises de manière automatique .....	6
<b>2. Zoom sur un exemple d'application du big data à l'action publique : la « police prédictive »</b> .....	<b>7</b>
De quoi s'agit-il ? .....	7
Quelle mise en œuvre ? .....	7
Pour quels résultats ? .....	9
Quelles limites ? .....	9
Quelles recommandations ? .....	12
<b>Ressources documentaires</b> .....	<b>13</b>

## 1. Vers une « action publique algorithmique » ?

L'irruption des technologies du big data ne réinterroge pas seulement les modèles économiques de certains secteurs d'activités ou entreprises, ou les paradigmes épistémologiques. Elle intéresse également de plus en plus la puissance publique, elle-même productrice et détentrice de quantités considérables de données numériques. Les administrations de nombreux pays sont désormais dotées d'un

responsable de haut rang dédié au big data : « Chief data officer » aux États-Unis,

« Chief digital officer » responsable du « Government Digital Service » en

Grande-Bretagne, « Administrateur général des données » en France Comme le

souligne Elisabeth Grosdhomme Lulin dans un rapport publié récemment par l'Institut de l'entreprise et intitulé « Gouverner à l'heure du big data », le big data apparaît aujourd'hui comme un facteur majeur de transformation de l'action publique :

*« Ce qui se préfigure à présent, sous la dénomination encore incertaine et sujette à des interprétations très diverses d'« action publique algorithmique », c'est une transformation profonde, sous l'effet du numérique, des manières d'agir des administrations, et ce faisant, au nom de l'efficacité et de l'efficacit , une mutation substantielle du contrat implicite entre le pouvoir et la soci t . »*

Elisabeth Grosdhomme Lulin explique en effet que nous sommes entr  dans une nouvelle phase du d ploiement du num rique dans l'action publique permettant   cette derni re de devenir plus personnalis e et plus pr dictive.

### **Le big data fait appara tre des personnes aux yeux de l'administration**

D sormais, les donn es collect es par la puissance publique ne sont plus seulement des donn es « statiques », d clar es ponctuellement par le citoyen sous son contr le. Ce sont de plus en plus des donn es comportementales, collect es en continu, qui d crivent tr s pr cis ment la r alit  de notre vie personnelle :   quoi nous d pensons les allocations sociales qui nous sont vers es, combien de temps nous utilisons l'appareil respiratoire qui nous a  t  prescrit,   quelle vitesse nous roulons en traversant telle agglom ration. Dans ces conditions, selon Elisabeth Grosdhomme Lulin l'usager de l'administration n'est plus seulement un sujet de droit, mais bel et bien un  tre de chair et d'os, une personne dont le comportement quotidien devient visible par l'autorit  publique. Par exemple, la gestion des r seaux de transports urbains donne lieu   une collecte massive de donn es en temps r el : validation des titres de transports, localisation des bus, mesure de l' tat du trafic, vid osurveillance, etc.

### **Les traitements qui donnent du sens aux donn es collect es deviennent beaucoup plus individualis s et pr dictifs**

Les donn es collect es ne servent plus seulement   alimenter de simples arborescences destin es   assigner chaque usager   telle cat gorie pr d finie, des

calculs visant à établir le montant de nos créances et dettes vis-à-vis de telle procédure, ou encore des traitements statistiques pour établir des moyennes. Les traitements informatiques qui leur sont appliqués s'approchent peu à peu de l'intelligence artificielle, permettant de caractériser telle situation ou de modéliser tel profil particulier en vue d'en tirer des conclusions pour l'action. Pour Elisabeth Grosdhomme Lulin, ce changement de perspective est fondamental. En effet, en permettant de modéliser la diversité des situations, d'objectiver les différences, et donc de fonder les différences de traitement sur l'hétérogénéité des besoins, la collecte et le traitement massif des données permet d'envisager **un service public universel dans ses finalités, mais personnalisé dans ses modalités en fonction des besoins de chaque usager.**

Parmi les exemples de personnalisation du service permise par le big data évoqués par Elisabeth Grosdhomme Lulin figurent notamment les suivants :

- Le **portail « mes-aides.gouv.fr »** récemment mis en ligne à titre expérimental par le gouvernement permet à chaque usager de connaître quelles sont les aides fiscales ou sociales auxquelles il a droit parmi les multiples aides proposés par l'Etat et ses agences (assurance maladie, Pole Emploi, assurance retraite, allocations familiales, éducation nationale...). Pour ce faire, il suffit d'indiquer les données qui caractérisent les situations personnelles (statut familial, nombre d'enfants, revenus, etc.). Grâce à des algorithmes, le portail gère ensuite la complexité réglementaire et procédurale (conditions, seuils, critères d'éligibilité,...) et informe sur les prestations auquel l'on a droit. Elisabeth Grosdhomme Lulin souligne l'intérêt de ce type de service non seulement pour chaque citoyen mais aussi pour tous les agents des services sociaux (caisses d'allocations familiales, centres communaux d'action sociale, caisses de mutualité sociale agricole...).
- La start up bordelaise Joijoba propose un **moteur intelligent de recherche d'emploi** (<http://www.jobijoba.com/fr/>). Agrégeant les offres d'emploi de 400 sites partenaires (dont Pole Emploi), Jobijoba les classent ensuite de façon intelligente grâce à de puissants algorithmes. Ce qui doit permettre in fine de rapprocher de façon beaucoup plus performante les offres disponibles et les demandes, en identifiant plus précisément les métiers. En effet Jobijoba ne s'appuie pas sur le répertoire des métiers de Pôle Emploi mais construit un référentiel ad hoc en analysant de façon automatisée la façon dont les métiers sont effectivement décrits par les recruteurs. Autre innovation, Jobijoba exploite systématiquement l'ensemble des données des fiches de poste en les confrontant aux données des internautes chercheurs d'emploi (cookies des dernières recherches et CV déposé en ligne). Ainsi le site est-il en capacité de suggérer à un employeur de compléter sa job description, au motif que les autres employeurs ont dans la plupart des cas réclamé une compétence qu'ils ont omise ; de même, il va pouvoir signaler à un demandeur d'emploi non seulement le nombre et la nature des offres disponibles sur sa demande, le nombre de personnes qui s'y intéressent, mais aussi le type de compétences traditionnellement recherchées : il pourra alors lui suggérer, de façon automatisée, d'ajouter une ligne à son CV (exemple : capacité à travailler en équipe, ponctualité) ou de s'engager dans une formation. Troisième innovation majeure, qui résulte des deux précédentes, Jobijoba peut permettre de rapprocher des offres et des demandes qui ne pourraient a priori pas se rencontrer selon les classifications classiques : les algorithmes vont

permettre de dépasser l'approche traditionnelle par métiers et par secteurs pour permettre une approche par compétences et par talents. Ainsi, le site permet de faire une recherche en indiquant comme mot-clé une compétence (rédaction, travail en équipe, gestion, capacité d'adaptation, etc.) et un lieu, et "d'en déduire une liste d'annonces qui vont proposer différents métiers autour de cette compétence (acheteur, administrateur réseau, etc.

De plus, l'« action publique algorithmique » favoriserait les **approches dynamiques, proactive, préventives**. L'analyse approfondie de données beaucoup plus abondantes et détaillées sur les faits passés permet de mieux anticiper la probabilité des événements significatifs de chaque domaine d'action et de s'y préparer, soit pour en prévenir l'occurrence, soit pour en atténuer les effets dommageables.

Elisabeth Grosdhomme Lulin présente plusieurs exemples d'applications prédictives du big data dans le champ de l'action publique.

- Concernant la **régulation des déplacements urbains**, la métropole de Lyon a développé un projet nommé Optimod (« Optimiser la mobilité durable en ville ») qui vise à prévenir la congestion urbaine en facilitant le report des déplacements en véhicules personnels vers des modes de transport alternatifs et en optimisant les opérations de fret urbain. La pierre angulaire de ce projet est un dispositif de collecte et d'analyse en temps réel de toutes les informations pertinentes (flux de circulation, signalement de travaux ou d'accidents, conditions météorologiques) permettant l'élaboration et la diffusion d'une prédiction détaillée de trafic à une heure.
- En matière de **lutte contre la fraude fiscale**, le big data peut permettre de croiser diverses bases de données et de repérer des corrélations suspectes indiquant une forte probabilité de fraude.
- Dans le **domaine social**, London Ventures, l'agence d'innovation créée par la ville de Londres pour inventer les services publics de demain, a entrepris de créer un modèle d'analyse prédictive qui aide les travailleurs sociaux à anticiper le risque de maltraitance des enfants afin d'intervenir à temps. Après tant d'exemples où, au lendemain d'un drame, le constat a été dressé que tel ou tel signal perçu tantôt par l'institutrice, tantôt par le médecin, tantôt par l'assistante sociale n'avait pas été dûment partagé et interprété, l'idée est de rassembler systématiquement les informations dont on dispose et, grâce à un algorithme d'analyse qui se raffinera au cours du temps, de repérer les situations à risque qui émergent de la mise en lien de signaux disparates. Dans le même ordre d'idées, le Medway Youth Trust, l'équivalent du centre d'action sociale du canton de Medway, dans le sud-est de l'Angleterre, a développé l'analyse prédictive pour repérer de manière anticipée les adolescents en risque de décrochage scolaire à partir d'une analyse textuelle des communications électronique entre les conseillers et les jeunes accompagnés et leur conseiller.

Au total, selon Elisabeth Grosdhomme Lulin, les capacités de traitement de données désormais accessibles aux agents publics peuvent les aider à donner sens à des symptômes épars, à confronter leurs observations à une base de connaissances existantes, à choisir entre différentes solutions.

### Des décisions publiques peuvent dorénavant être prises de manière automatique

Sur la base des données ainsi collectées et des traitements qui leur sont appliqués, le numérique est désormais en mesure d'être bien plus qu'une simple aide à l'instruction d'un dossier par un agent public habilité à prendre la décision finale. Il peut devenir un automatisme qui passe directement de l'analyse à l'action. Deux exemples illustratifs de cette forme d'« action publique automatique » peuvent être tirés du rapport de l'Institut de l'Entreprise.

- Aux Etats-Unis, les **allocations d'aide alimentaire** sont versées aux bénéficiaires non pas par virement bancaire ou par chèque, mais exclusivement via des cartes de paiement électroniques prépayées. Il en va de même de 40 % des allocations de chômage (unemployment benefit) et des allocations familiales (child support), ainsi que de 83 % des allocations familiales d'urgence (Temporary Assistance for Needy Families) et de la totalité des allocations de chauffage (Home Energy Assistance Program). Ce mode de versement permet à l'administration de contrôler l'usage qui est fait des allocations ainsi versées. En effet, la carte de paiement peut être paramétrée de manière à interdire les dépenses non autorisées : par exemple, impossible de payer avec cette carte dans certains commerces, typiquement dans les débits de boissons ou les salles de jeu.
- En prospective, certaines expérimentations en matière de **régulation du trafic automobile** laisse augurer d'un pilotage automatique de la vitesse des véhicules en milieu urbain ou sur voie rapide lorsqu'il apparaît nécessaire de fluidifier la circulation.
- En prospective toujours, le big data pourrait conduire à une **automatisation de la réciprocité entre droits et devoirs** : vos déchets ménagers seront dûment collectés si et seulement si vous les avez correctement triés ; votre traitement médical vous sera remboursé si et seulement si vous observez strictement la prescription qui vous a été faite par votre médecin ; vos primes d'assurance maladie seront modulées selon votre hygiène de vie, selon que vous fumez ou non, faites du sport ou non, mangez sainement ou non, etc.

Comme le souligne Elisabeth Grosdhomme Lulin, citant Jay Stanley, senior policy analyst pour l'American Civil Liberties Union : « Nous allons vers un monde dans lequel il devient techniquement possible d'assurer à 100 % la conformité avec la plupart des lois. » Cette possibilité d'automatisation de l'action publique soulève dès lors la question fondamentale de l'arbitrage à établir entre efficacité de l'action publique et liberté de comportement des individus.

## 2. Zoom sur un exemple d'application du big data à l'action publique : la « police prédictive »

Les pratiques de « police prédictive » (« predictive policing ») constituent un bon exemple pour prendre la mesure de l'intérêt, mais aussi des limites, que peut représenter l'application du big data à l'action publique. Elles présentent en effet l'intérêt de porter sur des pratiques sociales intervenant dans un cadre urbain, de concerner une application prédictive du big data et de donner lieu à une application opérationnelle par les services de police de différentes villes américaines mais aussi européennes.

### De quoi s'agit-il ?

Bilel Benbouzid, maître de conférences en sociologie à l'université Paris-Est, définit le « predictive policing » comme « une sécurisation policière assistée par des algorithmes d'anticipation du crime dans l'espace et le temps » (rue 89, 2015). La police prédictive consiste à anticiper consiste à collecter de multiples données sur les circonstances et modalités des crimes et délits passés, à les modéliser, à les confronter en temps réel aux données géolocalisées permettant de caractériser la situation de telle rue, tel quartier ou telle zone d'un territoire, en sorte de prévoir le risque d'occurrence des faits que l'on cherche à combattre et à prépositionner en conséquence des patrouilles de police (Grosdhomme Lulin, 2015).

### Quelle mise en œuvre ?

Elisabeth Grosdhomme Lulin indique que la police prédictive est en voie de généralisation aux États-Unis. En Europe, la police prédictive a également fait son apparition en Allemagne, en Suisse, au Royaume-Uni et arrive même en France (voir encadré ci-dessous).

#### L'arrivée des outils de police prédictive en France

Plusieurs articles de presse signalaient en mai dernier l'arrivée de la « police prédictive » en France. En particulier, un article publié par Médiapart indique que le ministère de l'intérieur vient de dévoiler l'existence d'un algorithme prédictif, à l'œuvre depuis fin 2014, permettant de prédire les grandes tendances en matière de criminalité sur le territoire. Celui-ci intègre des données issues des faits constatés par les forces de l'ordre et des statistiques de l'Insee, pour ensuite fournir des cartes permettant d'analyser la criminalité et de prédire son évolution. « Nous construisons un modèle basé sur les infractions constatées entre 2008 et 2013. S'il est validé et se vérifie sur les faits de 2014, nous le projetons sur l'année 2015 », explique à 20 Minutes Patrick Perrot, chef de la division analyse et investigation criminelle. « Seuls les délits et les infractions en quantité suffisante sont exploitables », poursuit le gendarme. Les homicides ne sont donc pas étudiés. Mais les vols, les agressions sexuelles et les viols, les cambriolages ou encore les trafics de stupéfiants sont ciblés. Sur cette base, une échelle de risques est alors définie et les informations recueillies sont transmises aux gendarmes sur le terrain pour qu'ils augmentent l'intensité des patrouilles, par exemple, dans un secteur défini où les vols par effraction seraient susceptibles d'augmenter au cours des prochaines semaines. Toutefois, à la différence des États-Unis, où les policiers reçoivent directement sur le tableau de bord de leurs voitures de patrouilles les indications de prédictions, ces nouvelles données ne sont transmises qu'aux chefs de service. A eux ensuite d'adapter leurs moyens et d'exploiter au mieux ces renseignements criminels dans leurs zones.

Plusieurs prestataires privés sont présents sur le marché, les plus notables étant l'américain PredPol et l'allemand Precobs. Etant déployé dans une soixantaine de

viles américaines, PredPol apparait comme le leader mondial. Dans un article publié en avril 2015 par Le Monde, le journaliste Yves Eudes explique que la start-up californienne, installée dans la ville universitaire de Santa Cruz et réunissant des anthropologues, mathématiciens et informaticiens spécialistes du big data, a tout d'abord téléchargé sur ses serveurs la totalité des archives informatisées de la police de la ville, en remontant jusqu'à 2004 : rapports, procès-verbaux, comptes rendus, transcriptions d'appels... A partir de cette masse de données hétérogènes, elle a constitué une base structurée et exploitable. Puis des algorithmes ont extrait et classé les délits selon **trois critères : date, lieu et catégorie** (vol à la tire, cambriolage, attaque à main armée, violences physiques, vol de voiture, vandalisme...). Sur cette base, les chercheurs ont produit des équations (tenues secrètes) suggérant que le crime est prévisible à court terme puis les ont testés de façon concluante avec les données de terrain. L'un des fondateurs de la société, Jeffrey Brantingham, professeur d'anthropologie à l'université de Los Angeles, explique au journaliste que son équipe s'est inspirée d'algorithmes inventés par des chercheurs qui essayent de prévoir les tremblements de terre : « *Les sismologues classent les séismes en deux catégories : les secousses primaires, difficilement prévisibles, et les répliques, plus faciles à prévoir car elles sont proches, en temps et en lieu, du séisme initial. C'est la même chose pour les cambriolages. Si un immeuble est cambriolé une fois, les chances qu'il le soit à nouveau augmentent de près de 100 %. Peu importe les causes, les faits sont là. Cette séquence d'événements est modélisable.* » La ville de Santa Cruz ainsi a été ainsi l'une des premières à se doter, en juillet 2011, du programme PredPol (abréviation de predictive policing).

L'enquête du journaliste du Monde auprès des services de police de la ville de Modesto en Californie permet de mieux comprendre l'utilisation de PredPol. Avant de prendre leur service pour patrouiller en ville, les policiers sont briefés par leur supérieur sur les « hot spots », les points chauds à surveiller en priorité. Un écran mural affiche un plan de la ville, sur lequel on distingue dix petits carrés rouges : un carrefour, un square, un centre commercial... Selon l'ordinateur, connecté au système PredPol, les délits vont se concentrer, au cours des douze prochaines heures, dans ces carrés de 150 mètres de côté, représentant une minuscule fraction de la superficie de Modesto. Le sergent zoome sur la carte pour montrer aux différentes patrouilles les quartiers qui leur sont assignés. L'ordinateur affine ses prédictions et affiche cinq « carrés PredPol » par secteur, classés par ordre de priorité. Chaque policier reçoit par e-mail cinq petites cartes, représentant différentes portions de leur secteur. Quand ils seront en patrouille, ils pourront les consulter sur leur téléphone ou sur l'ordinateur de bord de leur voiture.



quelques heures, la scène se répétera pour l'équipe du matin, mais certains points chauds auront changé, car l'ordinateur modifie ses prédictions d'heure en heure. En effet, PredPol reste connecté 24 heures sur 24 au réseau informatique de la police, et télécharge en temps réel tous les nouveaux rapports, à mesure que les policiers

les tapent sur leurs ordinateurs de bord. La base de données est ainsi mise à jour en continu de façon à recalculer en permanence les trois paramètres clés mentionnés plus haut : la nature, le lieu et l'heure du crime.

Source : Le monde, 22/04/2015

### Pour quels résultats ?

Dans sa note sur l'action publique algorithmique, Elisabeth Grosdhomme Lulin indique que les gains d'efficacité constatés après quelques mois ou trimestres de mise en œuvre de ces méthodes de police prédictive s'avèrent spectaculaires. Ainsi comme le met en avant PredPol sur son site internet :

- Le service de police de **Santa Cruz** observe que les agressions ont baissé sur 9%, les cambriolages de 11%, et les vols de 27% au cours de la première année d'utilisation du logiciel (2011-2012). D'une manière générale, les actes criminels ont reculé de 25% en Juin 2013 et de 29% en Juillet 2013 par rapport à ces mêmes mois de l'année précédente.
- A **Modesto**, le service de police fait état des plus bas taux de criminalité observé au cours des 3 dernières années depuis le déploiement de PredPol en Janvier 2014.
- A **Atlanta**, la criminalité globale a diminué de 8% et 9% dans les deux premières zones où a été déployé PredPol en juillet 2013. Sur les quatre zones où PredPol n'a pas été déployée, le taux de criminalité a augmenté de 1 à 8% dans trois d'entre elles et est resté stable dans la dernière. En raison de ces résultats positifs, le département de police d'Atlanta a décidé en Novembre 2013 de mettre en œuvre PredPol dans toute la ville. Le département de la police d'Atlanta a vu la criminalité globale baisse de 19% et attribuent une grande partie de la réduction soutenue pour le déploiement de PredPol.
- Le service de police d'**Alhambra** a enregistré une baisse de 32% des cambriolages et une baisse de 20% des vols de véhicules depuis le déploiement en Janvier 2013. La ville a connu en mai 2014 le mois avec la criminalité la plus faible de son histoire.

Sur un plan plus qualitatif, les témoignages de policiers présentés dans l'article du Monde mettent en évidence le fait que PredPol permet de lutter contre la routine, qui pousse les policiers à couvrir toujours les mêmes endroits. Autrement dit, l'outil permettrait d'identifier des zones à enjeux auxquelles les policiers n'auraient pas pensé spontanément. Si les policiers plébiscitent également PredPol, c'est parce que les carrés ciblés par l'outil permettent de trouver plus fréquemment des personnes recherchées par la justice : « par exemple, des inculpés libérés sous caution qui ne se sont pas présentés à leur procès, ou des personnes en liberté anticipée qui ne se soumettent pas aux contrôles obligatoires – et qui sont enclins à commettre de nouveaux délits. » Enfin, plusieurs responsables de service de Police voient dans PredPol le moyen d'innover dans un contexte marqué par le recul des moyens qui leur sont alloués. Grace à la police prédictive il paraît possible de faire autant, voire mieux, avec moins : « Désormais, nous pouvons être proactifs, concentrer nos moyens limités sur les zones prioritaires. »

### Quelles limites ?

Le développement des pratiques de police prédictive ne va pas sans soulever un certain nombre de critiques touchant tant à leur efficacité supposée qu'à leur légitimité et leur acceptabilité.

#### Une efficacité surjouée ?

Selon Ismaël Benslimane, auteur d'un article intitulé « Étude critique d'un système d'analyse prédictive appliqué à la criminalité : Predpol® », questionner l'efficacité d'un tel outil amène tout d'abord à rappeler les biais inhérents à mesure de la criminalité. D'une part, ne sont comptabilisés que les crimes et délits officiellement constatés. Tous ceux qui ne sont pas déclarés par les victimes ou des témoins restent bien évidemment dans l'ombre. De même, il suffit d'un changement dans les procédures d'enregistrement pour faire changer les chiffres. Or, Predpol est aussi un outil d'enregistrement et de classification des délits. Ce qui pose la question des types de délits enregistrés et de ceux utilisés pour l'évaluation et la prédiction : par exemple, si il y a une réduction des vols de voitures mais une augmentation des vols à l'arraché, que faut-il en conclure sur l'efficacité ? D'autre part, parce que l'utilisation de PredPol se traduit par la concentration des patrouilles de police en certains lieux jugés à risque, la baisse des actes criminels constatée peut résulter d'un simple déplacement de ceux-ci vers des quartiers ou villes alentours. Ce qui pose la question de la zone retenue pour évaluer la baisse de la criminalité.

Une seconde critique proposée par Ismaël Benslimane renvoie au fait qu'il suffit prédire toujours les mêmes lieux « à risque » pour être aussi performant que Predpol. En effet, les crimes ayant lieu majoritairement dans les zones historiquement les plus criminogènes de la ville, le supplément de précision apporté par PredPol n'apparaît pas aussi spectaculaire qu'affiché. Comme le résume le journaliste d'InternetActu Hubert Guillaud dans un article présentant l'article d'Ismaël Benslimane, PredPol renverrait avant tout à une « prédiction des banalités ».

Toutefois, cette critique des prédictions de PredPol pourrait être nuancée. Les témoignages du sociologue Dominique Cardon et de Yves-Alexandre de Montjoye, chercheur au Media Lab du MIT, présentés dans l'article d'Hubert Guillaud laissent à penser que, si la répétition des crimes dans le temps et l'espace constitue certes une banalité de la criminologie et si les policiers n'ont jamais patrouillé au hasard, cela ne veut pas dire pour autant qu'une méthode plus scientifique de ciblage des lieux à surveiller en priorité n'a aucun intérêt.

### La primauté donnée à la coercition sur la prévention ?

Une critique a priori plus incisive des pratiques de police prédictive renvoie à l'utilisation avant tout coercitive qui est faite des prédictions faites par PredPol. Comme en témoignent les policiers interrogés par le journaliste du Monde, le ciblage des patrouilles policières permet d'accroître le nombre d'arrestations, notamment du fait des situations de « flagrant délit », et d'éviter l'occurrence d'actes répréhensibles. En revanche, les anticipations permises par PredPol ne sont pas mises au service d'une prévention des facteurs sociaux des actes de délinquance ou criminels. Pour le dire rapidement, PredPol ne sert pas à amener les travailleurs sociaux auprès de ceux qui en ont besoin mais plutôt la police. C'est aussi ce que soulignent les chercheurs en sociologie et science politique Vincent Dubois, Morgane Paris et Pierre-Edouard Weill à propos de l'utilisation du big data au sein de la Caisse Nationale d'Allocation Familiale : celle-ci vise exclusivement la lutte contre la « fraude sociale », « laissant de côté la prospection des droits et la lutte contre le non-recours ». Au total, selon Ismaël Benslimane, un risque soulevé par la police prédictive est de contribuer à masquer la réalité sociale. PredPol serait surtout une manière d'exprimer de manière politiquement correcte qu'il y a plus de délits dans certaines zones d'une ville, sans rien dire de la précarisation de ces zones.

### Une menace pour la vie privée et le libre arbitre

Dans l'article du Monde évoqué plus haut, le fondateur de fondateur de PredPol, Jeffrey Brantingham, déclarait : « Nous n'effectuons aucun profilage, nous ne nous intéressons pas aux auteurs des infractions. Pour nos prédictions, l'identité des délinquants, ou leurs caractéristiques socioculturelles sont sans valeur. » On peut cependant douter de cette affirmation. Parce qu'elle se nourrit en partie de données à caractère personnel, l'action publique algorithmique soulève le risque de profilage de personnes prises individuellement. Ceci est d'autant plus vrai que l'anonymisation des données n'apparaît pas comme une protection suffisante (Ollion et Boelaert, 2015) : les chercheurs en informatique ont, depuis longtemps, montré qu'on pouvait retrouver l'identité des personnes en croisant des données anonymisées avec des informations publiquement disponibles.

A cet égard, Elisabeth Grosdhomme Lulin pose une question cruciale : à l'heure où tant de données nous concernant sont accessibles aux tiers, que pouvons-nous dire ou faire qui ne puisse, ensuite, nous devenir opposable dans l'exercice de nos droits et libertés de citoyens ? Selon elle, le risque majeur ne réside pas tant dans la perte de confidentialité de nos faits et gestes (déjà largement entamée par les smartphones et les réseaux sociaux, quelle que soit l'action de l'Etat) que dans la protection de notre libre arbitre, de notre capacité à faire des choix et à échapper au conformisme découlant de la structuration de notre « menu d'options » par des algorithmes : les billets qui apparaissent en tête de notre fil d'actualités Facebook, les liens qui ressortent de nos recherches sur Google, les livres qu'Amazon nous propose de lire, etc.

Ce risque d'atteinte aux libertés individuelles prend une acuité toute particulière en matière de prédiction des crimes. En effet, si PredPol relève de la prédiction spatio-temporelle, d'autres outils s'inscrivent résolument dans une logique de profilage individuel. Dans une perspective d'enquête policière, par exemple pour trouver l'auteur d'un meurtre, il peut s'agir de trouver l'identité du coupable en croisant une multitude de données sur le mode de passage à l'acte, des informations ADN, des systèmes de vidéosurveillance, des fichiers, etc. Mais ce ciblage peut également viser à prédire les criminels, c'est-à-dire identifier les individus qui présentent un risque élevé de commettre un crime dans un futur proche. C'est ce que montrent deux articles récents relayés par InternetActu.

- Détecter les enfants risquant de devenir délinquant

Le premier, publié par le Pacific Standard, rend compte d'une expérimentation réalisée par le service de protection de l'enfance du comté de Los Angeles. Celui-ci a mis en place un outil d'analyse prédictive visant à identifier les enfants qui risquent de se retrouver un jour derrière les barreaux, en croisant de multiples données sur les familles, la consommation de drogues, la réussite scolaire, les arrestations... Si ce projet pilote a pris fin en 2014, un rapport du Conseil national sur le crime et la délinquance américain de décembre 2015 suggère non seulement que l'initiative du comté de Los Angeles allait dans le bon sens, mais que d'autres organismes gouvernementaux devraient s'y intéresser. Si ce système était pensé au départ comme un moyen permettant d'aider les travailleurs sociaux à faire face à l'ampleur de la tâche, l'article s'interroge sur le risque que ce type d'analyse finisse par être partagé avec le système judiciaire ou policier, amplifiant alors considérablement le pouvoir coercitif de la puissance publique sur les individus.

- Calculer le score de menace d'un individu

Le second article, publié par le Washington Post, évoque le logiciel « Beware » utilisé par le « Real Time Crime Center » de la police de Fresno. Ce logiciel permet de chercher dans des millions de données (rapports d'arrestation, registres de propriété, bases de données commerciales, recherches sur le Web profond et sur

les réseaux sociaux) pour calculer le potentiel de violence d'un individu, le résultat aboutissant à un code de trois couleurs, dont le rouge pour dangereux.

Au total, comme le suggère Elisabeth Grosdhomme Lulin, l'action publique algorithmique soulève le risque de voir émerger de nouvelles catégorisations des individus pouvant présenter un caractère discriminatoire. Viktor Mayer-Schönberger, professeur de gouvernance et de régulation de l'internet à l'université d'Oxford, évoque quant à lui le risque que le big data soit utilisé « pour assigner des responsabilités à des individus, non pour ce qu'ils ont fait, mais pour ce que l'on prévoit qu'ils pourraient faire ».

### Quelles recommandations ?

Face aux risques soulevés par l'irruption du big data dans le champ de l'action publique, plusieurs pistes de régulation peuvent être évoquées.

#### Equiper notre démocratie d'une capacité d'audit des fameux algorithmes

Selon Elisabeth Grosdhomme Lulin, l'utilisation croissante des algorithmes dans des applications opérationnelles – quels passagers la police va-t-elle interroger et fouiller lors de l'arrivée de tel avion à l'aéroport ? Quelle sera la liberté effective laissée à tel condamné à la liberté surveillée avant que son bracelet électronique ne se mette à sonner ? Quelle affectation universitaire va recevoir tel candidat à l'enseignement supérieur ? – implique en retour de doter les institutions démocratiques d'une capacité d'audit de ces fameux algorithmes. Celle-ci doit permettre non seulement de vérifier si les règles codées par les informaticiens à l'origine sont bien conformes aux orientations politiques qui avaient été annoncées et au respect des droits fondamentaux des personnes, mais également de vérifier si, de son propre chef, l'algorithme n'a pas ajouté d'autres règles ou réinterprété ses premières instructions.

#### Définir un droit des algorithmes prédictifs

L'article de Médiapart cité plus haut rappelle quant à lui que l'article 10 de la loi du 6 juin 1978 relative à l'informatique, aux fichiers et aux libertés interdit bien toute décision « produisant des effets juridiques » prise « sur le seul fondement d'un traitement automatisé de données ». Le journaliste Jérôme Hourdeaux se fait aussi l'écho des préconisations formulées du Conseil d'Etat à l'occasion de son étude annuelle 2014, intitulée « Le numérique et les droits fondamentaux ». Les sages y proposent de « définir un droit des algorithmes prédictifs ». Selon eux, il faut « éviter que des systèmes présentés comme relevant de "l'aide à la décision" soient en réalité presque toujours suivis et commandent la décision, l'intervention humaine n'étant alors qu'apparente ». Le Conseil d'Etat demandait notamment un avis de la Cnil pour « préciser l'interprétation du texte actuel, en prévoyant que l'intervention humaine ne doit pas être que formelle. Cette disposition pourrait indiquer les critères permettant de s'assurer du caractère effectif de cette décision, tels que les compétences et les qualifications de la personne qui prend la décision, la marge de manœuvre dont elle dispose dans le processus défini par son organisation et l'existence d'éléments d'information lui permettant le cas échéant de justifier de la prise d'une autre décision que celle proposée par l'algorithme ».

Concernant l'utilisation des algorithmes dans les procédures judiciaires, l'étude du Conseil d'Etat propose l'introduction du principe du contradictoire dans le traitement des données, en offrant de nouveaux droits aux citoyens : « Lorsqu'une décision produisant des effets juridiques ou une mesure affectant de manière significative les intérêts d'une personne est en partie fondée sur un algorithme, cette personne devrait bénéficier de garanties analogues à celles d'une procédure contradictoire. Elle doit pouvoir en effet être en mesure de faire valoir ses

observations auprès de la personne qui prendra la décision, en produisant des arguments de nature le cas échéant à contrebalancer la proposition de l'algorithme ». Pour cela, la loi devrait « imposer aux auteurs de décisions s'appuyant sur la mise en œuvre d'algorithmes une obligation de transparence sur les données personnelles utilisées par l'algorithme et le raisonnement général suivi par celui-ci. Donner à la personne faisant l'objet de la décision la possibilité de faire valoir ses observations ».

### Ressources documentaires

Ismaël Benslimane, « Predpol : prédire des crimes ou des banalités ? », Cortecs, 10 décembre 2014, <http://cortecs.org/mathematiques/predpol-predire-des-crimes-ou-des-banalites/>

Conseil d'Etat, « Etude annuelle 2014 du Conseil d'Etat - Le numérique et les droits fondamentaux », <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/144000541.pdf>

Vincent Dubois, Morgane Paris et Pierre-Edouard Weill, « Le data mining, un instrument probabiliste de contrôle entre raisons statistique, juridique et gestionnaire », non publié

Yves Eudes, « PredPol, le big data au service de la police », Le Monde, 22/04/2015, [http://abonnes.lemonde.fr/ameriques/article/2015/04/22/predpol-le-big-data-au-service-de-la-police\\_4620855\\_3222.html](http://abonnes.lemonde.fr/ameriques/article/2015/04/22/predpol-le-big-data-au-service-de-la-police_4620855_3222.html)

Elisabeth Grosdhomme Lulin, « Gouverner à l'ère du Big Data. Promesses et périls de l'action publique algorithmique », Institut de l'Entreprise, mai 2015

Hubert Guillaud, « Predpol : la prédiction des banalités », InternetActu, le 23/6/2015, <http://www.internetactu.net/2015/06/23/predpol-la-prediction-des-banalites/>

Jérôme Hourdeaux, « Gendarmes et industriels imaginent un nouveau logiciel pour prédire le crime », Médiapart, 25/05/2015, <https://www.mediapart.fr/journal/france/250515/gendarmes-et-industriels-imaginent-un-nouveau-logiciel-pour-predire-le-crime>

Justin Jouvenal, « The new way police are surveilling you : Calculating your threat 'score' », Washington Post, 10/01/2016

Étienne Ollion, Julien Boelaert, « Au delà des big data. Les sciences sociales et la multiplication des données numériques », Sociologie 2015/3 (Vol. 6)

Camille Polloni, « Police prédictive : la tentation de « dire quel sera le crime de demain » », Rue89, Le grand entretien, le 27/05/2015, <http://rue89.nouvelobs.com/2015/05/27/police-predictive-tentation-dire-quel-sera-crime-demain-259384>

Matt Stroud, « Should Los Angeles County Predict Which Children Will Become Criminals? », Pacific Standard, 21/01/2016, <http://www.psmag.com/politics-and-law/minority-report>