

CartoCrime : la délinquance en cartes

Avec plus de dix ans de retard sur les États-Unis, la France vient enfin de se doter d'un outil – pour l'instant encore sommaire – de cartographie de la délinquance sur *Internet*. Piloté par l'Office national de la délinquance, autorité indépendante, *CartoCrime* présentera deux visages : un outil d'information du public, et de travail pour les services de police.

Depuis maintenant le début des années 2000, quel géomaticien n'a-t-il pas eu l'occasion de se rendre sur les pages *Internet* de villes américaines comme New-York ou la Nouvelle-Orléans et de constater que l'on y trouvait des cartes de délinquance extrêmement précises : géolocalisation des faits, classement suivant les catégories, les heures, etc. Jusqu'à présent, la transposition de telles initiatives en Europe, et particulièrement en France, se heurtait à plusieurs obstacles majeurs. D'une part, les incertitudes qui planent sur le recensement des faits (nous y reviendrons) ; d'autre part, la forte réticence des pouvoirs publics face à un phénomène devenu enjeu sociétal et politique, avec, de surcroît, la peur de certains élus locaux de voir leur territoire stigmatisé, puis mis au ban.

Certes, l'interprétation de cartes brutes peut prêter à confusion, et l'on connaît également la propension de certains media à dénicher le sensationnel parmi les informations les plus anodines, ou à biaiser les interprétations. Il n'en reste pas moins que tenir le public dans une sorte d'ignorance quant aux crimes et délits entretient des rumeurs, et parfois des appréhensions, qui ne sont pas nécessairement fondées.

Créé en 2002 suite à un rapport parlementaire mixte (gauche/droite), l'*Office national de la délinquance*, OND, placé sous la tutelle de l'*Institut national des hautes études de sécurité* (INHES), un service du ministère de l'Intérieur, est cependant un organisme indépendant qui possède sa propre autonomie ; il est administré par un conseil éclectique composé

de représentants de la Police, de la Gendarmerie, mais aussi des assurances, des transporteurs, des élus et même des journalistes. Sa mission est triple :

- Analyser les chiffres ;
- Communiquer ces chiffres ;
- Procéder à une grande enquête dite « de victimation ».

Des statistiques biaisées

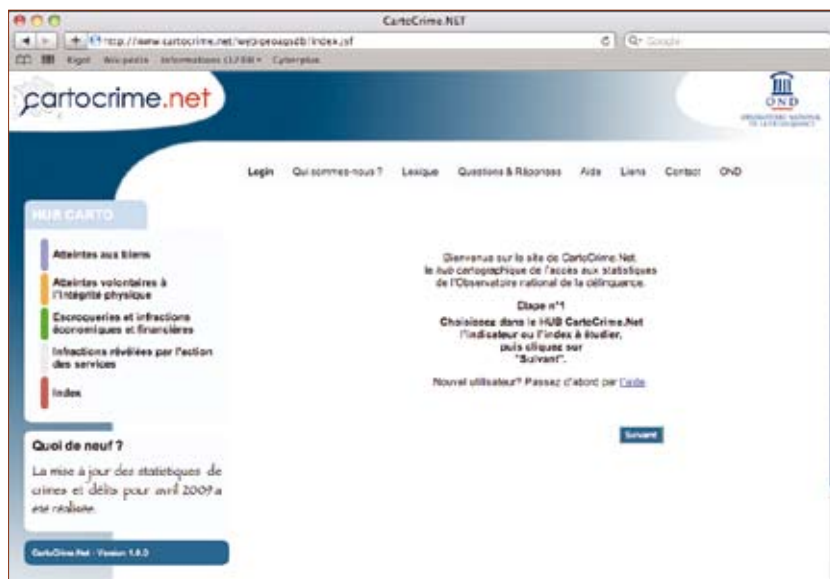
Quand on parle de délinquance, on évoque des phénomènes multiples, qui ne peuvent faire l'objet d'une opération de recensement unique. La plupart des délits violents (crapuleux ¹ ou non) viennent à la connaissance des services de police au travers des dépôts de plainte : cette source reste loin d'être exhaustive. En effet, certains individus ne portent pas plainte, pour des raisons diverses : menaces (dans le cas des violences conjugales, par exemple), crainte de se présenter à la Police (si la victime est en situation irrégulière, ou est elle-même délinquante, cas des toxicomanes), analphabétisme, etc. En outre, les infractions relevées par l'activité des services (trafics divers, fausse monnaie, fraudes...) ne sont pas non plus comptabilisés dans le fichier des plaintes, puisqu'ils passent inaperçus et ne provoquent pas d'atteinte directe à l'intégrité physique.

1. On rappelle qu'un acte est dit *crapuleux* s'il est commis par intérêt. La violence non-crapuleuse correspond à de la « violence gratuite ».

C'est la raison pour laquelle l'OND, en partenariat avec l'Insee, mène l'enquête de *victimation*. Sur une base annuelle, des opérateurs se déplacent sur le terrain à la rencontre des habitants et demandent si les personnes interrogées (sous couvert d'anonymat) ont été ou non victimes d'agression ; les réponses, qui sont plus fidèles à la réalité, permettent d'établir un facteur correctif par rapport aux statistiques officielles, ce qui donne des chiffres plus conformes à la réalité, et précisent l'estimation de la marge d'erreur.

Un autre biais, cette fois géographique, entache également les données du fichier du STIC 2 : les faits sont géolocalisés non à l'endroit où ils ont été commis, mais au commissariat (ou gendarmerie) où ils ont été rapportés. C'est la conséquence de la compétence nationale de la Police française, contrairement à d'autres pays européens comme l'Italie, ou aux États-Unis, où les plaintes ne sont reçues et instruites que par les brigades territorialement compétentes, qui collaborent volontairement avec les services nationaux, comme le FBI. Ce biais perturbe significativement l'interprétation des données en zone urbaine (où l'on peut s'apercevoir d'un vol qu'à l'issue de son trajet pendulaire domicile-travail, par exemple) : ainsi, on estime à 5 – 10 % le nombre de faits enregistrés à Paris mais qui n'y sont pas commis ; en outre, cette pratique rend les statistiques en zone rurale, où la densité de commissariats/gendarmeries est faible, très grossières.

La localisation précise du fait est pourtant enregistré dans le fichier, mais sous la forme d'un attribut qui est filtré lors de sa remontée. En effet, chaque mois l'OND reçoit de la part de la DGPN, la Direction



Après un avertissement liminaire sur le contenu du site, l'internaute arrive sur une page d'accueil où il est invité à choisir parmi quatre grandes catégories de délits.

générale de la Police nationale, organisme qui est chargé de gérer le STIC et de l'archiver, un fichier récapitulatif comptabilisant les faits commis, ventilés selon une nomenclature à 107 postes, baptisé *état 4001*, et classés par départements. Cette matière première extrêmement pauvre serait la seule disponible si l'OND n'avait passé des conventions locales avec quelques services de police, comme la Préfecture de police parisienne (Marseille devrait suivre), aux termes desquelles, sous contrôle de la CNIL, l'Office recueille des données anonymes mais plus précises directement du STIC local.

Ce n'est pas encore la panacée : la localisation issue des fiches du STIC demeure parfois vague. Il suffit pour cela que les faits se passent dans un espace non-clos sans points de repère (route de campagne, espace vert suffisamment vaste, bois...) ou dans des transports en commun (entre telle et telle gare). Seule une localisation par GPS pourrait permettre le recueil d'une information exacte,

mais cela supposerait le déplacement d'une équipe avec la victime, ce qui n'est guère envisageable – en outre, il n'est pas certain qu'une information géolocalisée au mètre près soit véritablement pertinente. Il faut donc se satisfaire de données géographiques parfois imprécises, mais globalement satisfaisantes.

Le fichier de l'état 4001. Comme on le voit, il s'agit de données totalement anonymes et impossibles à interpréter sans le secours d'un ordinateur.

2. Le fichier STIC, *Système de traitement des infractions constatées* recense les personnes et les biens mis en cause (victimes ou agresseurs) dans des crimes ou délits. Wikipédia, citant Le Monde du 20 janvier 2009, écrit à son sujet que : « Dans son rapport publié le 22 janvier 2009, basé sur des contrôles effectués de juin à novembre 2008, la CNIL estime que seulement 17 % des fiches sont exactes. Sur les 5,5 millions de personnes mises en causes et les 28,3 millions de victimes mises en cause dans 36,4 millions d'affaires à la date du 2 décembre 2008, la plupart des erreurs concernent des défauts d'actualisation : en 2007, seulement 21,5 % des classements sans suite ont été enregistrés. »



Étape suivante, l'internaute doit sélectionner un mode de représentation. Le choix par défaut convient à la plupart des situations, les statisticiens pourront expérimenter la réalisation de cartes plus complexes.

Une idée venue d'Outre-Atlantique

Pour satisfaire sa mission de communication, l'OND s'était contenté, dans un premier temps, de publier un bulletin mensuel comportant des statistiques générales sur les faits constatés et leur

évolution. Bien vite, il est venu à l'idée des responsables de l'Office de publier cette information sur l'Internet. Et pourquoi d'en profiter pour réaliser une partie cartographique où le public pourrait accéder à un certain nombre de cartes. « Le premier problème qui s'est posé, explique Jean-Luc Besson, géomaticien à l'OND, a été de définir des indicateurs perti-



L'affichage cartographique se fait sur fond GéoFla. Zoom possible, ainsi que sélection et interrogation des données sur un département particulier.

nents. Il était hors de question de proposer 107 cartes correspondants aux 107 index de l'état 4001³. Nous avons donc choisi de regrouper les infractions en quatre indicateurs catégoriels : les violences, les atteintes aux biens, les infractions économiques et celles révélées par l'action des services.

Bien entendu, nous nous sommes inspirés de ce qui se faisait aux États-Unis, où la cartographie criminologique a débuté dès les années 1980. Il est bon de préciser que, même là-bas, la géolocalisation des faits est soumise à des dispositions légales, dont certaines protègent la vie privée : la localisation des personnes y est également interdite, excepté dans le cas des agresseurs sexuels – où le taux de récidive reste important. La méthodologie est également différente : les Américains travaillent à l'échelle du ponctuel, et rendent publiques leurs statistiques dans l'espoir que tous les citoyens participent à l'effort anti-criminalité, et que les touristes ou visiteurs soient conscients des risques qu'ils courent. »

Outre s'aligner sur les pratiques américaines, il s'agissait également pour l'OND de constituer un guichet unique d'accès aux données, sous une forme aisément compréhensible ; de surcroît, l'application devait permettre de remonter autant que faire se peut dans le temps : « Cependant pas au-delà du début de 1996. Avant, non seulement les statistiques ne sont plus très fiables, mais le Code pénal était différent, donc l'état 4001 ne correspondait plus à la nomenclature que nous avons définie », enchaîne Jean-Luc Besson.

Baptisé Cartocrime 4, ce site cartographique ressemble beaucoup aux autres sites de webmapping. Après une page initiale qui avertit le visiteur sur la limite de validité des statistiques et les éventuelles erreurs d'interprétation qui

3. On pourra trouver la nomenclature de l'« état 4001 » ainsi que des statistiques à cette adresse : http://www.interieur.gouv.fr/sections/a_la_une/toute_actualite/securete-interieure/chiffres-delinquance-2008/downloadFile/attachedFile_1/Etat_4001.pdf?nocache=1232376670.41

4. <http://www.cartocrime.net>

Détails techniques

La partie technique de *Cartocrime* a été confié à la société de service *Galigeo* qui a travaillé sur base de produits *ESRI*, dans la continuité des choix techniques qui avaient déjà été effectués par l'OND. Le projet a donc consisté à réaliser une intégration de technologies *ESRI* dans un environnement assez rigide : le site est hébergé à Strasbourg dans une infrastructure destinée à l'usage des polices européennes (le centre d'hébergement des applications de l'espace Schengen), et il doit, en outre, répondre aux exigences du Ministère de l'Intérieur, ce qui implique l'utilisation de composants qualifiés, certains commerciaux, d'autres libres comme la base de données *PostGres*. Cette dernière a été choisie, plutôt qu'une *Géodatabase*, en raison du caractère essentiellement alphanumérique des données stockées.

En partant des données, la première fonction offerte est la connexion à la base de données au travers d'une interface sécurisée type *https* qui permet à l'administrateur de charger facilement les fichiers de l'état 4001. Derrière ce formulaire se cachent une moulinette de traitement, des opérations de stockage et de calcul écrits directement en *SQL*.

Une fois intégrées à la base, la consultation se fait grâce au trio classique *TomCat/Apache/ArcGIS Server*. Les données géographiques, issues du *GéoFla™* de l'IGN, sont jointes dynamiquement lors de la préparation de la carte, le fond de plan est quant à lui préparé à l'avance sous forme de fichier *Shape* et *mx*. « L'intérêt de cette solution par rapport à une jointure en dur, explique Sébastien Connesson, qui a supervisé le projet chez *Galigeo*, c'est d'offrir de la souplesse pour pouvoir ajouter

de nouveaux indicateurs ou de nouveaux niveaux géographiques. Chaque utilisateur possède sa propre vision dynamique de la base, et peut choisir, grâce à un ETL embarqué, d'exporter les données au format Excel, par exemple. »

Les utilisateurs authentifiés (grâce à un mot de passe devant répondre aux normes sévères du ministère de l'Intérieur) accèdent à des données enrichies de quatorze états supplémentaires correspondant à l'activité des services. Enfin, l'administrateur lui-même a la possibilité d'ouvrir ou de fermer des comptes, de paramétrer des requêtes servies par l'application, ajouter des critères, etc.

« Parmi les évolutions, poursuit Sébastien Connesson, nous allons revoir l'ergonomie – qui peut en dérouter plus d'un – et passer sur une technologie plus moderne à base de Flex, ce qui améliorera le confort d'utilisation, par exemple en permettant la conservation de critères de sélection entre deux requêtes... La présentation graphique pourra se faire sous forme de templates XML pour plus de souplesse. En outre, dans la partie privée et pour l'usage de la *DGPN*, il nous a été demandé de produire des cartes à l'échelle communale.

Je dois dire que nous sommes plutôt satisfaits de la robustesse de l'application. Depuis son lancement le 24 avril, nous n'avons enregistré aucun plantage, malgré le pic de connexion que nous avons enregistré après le lancement (plusieurs dizaines de milliers par jour). Nous avons prévu quatre machines en parallèle, tout s'est bien déroulé. À l'heure actuelle, le dimensionnement a dû être revu à la baisse, mais l'application est très stable. Cela valide les choix techniques que nous avons faits. »

peuvent en résulter, la création d'une carte se déroule en trois étapes : l'internaute est d'abord invité à choisir une des quatre grandes catégories de faits, puis une sous-classe de la catégorie choisie ; enfin, un formulaire assez élaboré lui permet de choisir l'analyse : quantum de temps (mois, année ou période définie), type de classement (intervalles égaux,

quantiles, écarts types...), nombre de classes, calcul du taux pour mille habitants, inclusion éventuelle des statistiques des services centraux et de la PAF 5. « La PAF n'opère que dans les départements frontaliers ou ceux qui possèdent un aéroport international, donc inclure ces statistiques introduit un biais. De même, calculer le taux d'infractions par mille habitants conduit à sur-

représenter les centres touristiques ou économiques : imaginez-vous Paris IV^e arrondissement, quel est le rapport de la population en journée – donc du potentiel d'agression ou de vol – par rapport aux "vrais" résidents ? », s'interroge Jean-Luc Besson. Les comparaisons temporelles se limitent à des indicateurs annuels. Même s'il n'est pas possible de produire des comparaisons

5. Police de l'air et des frontières.



Pour ceux que l'affichage cartographique ne satisfait pas, des statistiques temporelles sont également disponibles sous forme de graphe simple.

automatiquement, la fonction d'exportation au format *Excel* permet de récupérer les valeurs dans un tableur puis d'effectuer les comparaisons.

La version privée, qui nécessite une identification préalable, propose quelques fonctions supplémentaires, comme des statistiques plus poussées sur l'action de services, notamment le nombre de gardes à vue, de mises en examen et d'écrous.

Réactions et évolutions

Du côté des élus, les réactions à l'ouverture du site n'ont pas été nombreuses, même si certaines préfectures se sont déclarées intéressées. La DGPN, en revanche, a immédiatement contacté l'OND : elle avait, semble-t-il, l'intention de monter son propre site cartographique. Comme il n'est pas de

mise, en ces temps de restriction budgétaire, de multiplier les dépenses, les deux organismes vont signer une convention au terme de laquelle le site *Cartocrime* se verra enrichi d'une partie privée dédiée à la DGPN.

Le contenu évoluera prochainement : en effet l'OND va bientôt passer sous la tutelle du Premier ministre, tout en étendant ses compétences au domaine juridique : elle deviendra l'ONDRP où RP signifie *Réponses Pénales*, ce qui signifie qu'elle portera à la connaissance du public des chiffres concernant les condamnations, les incarcérations, les mises sous écrou en général. Jean-Luc Besson réfléchit également sur la façon d'exploiter, en partenariat avec l'*Insee*, les résultats de l'enquête de victimation.

Quant à l'interface, elle devrait dès cette année évoluer vers un peu plus de dynamisme grâce à une API de type *Flex* ; des nouveautés seront progressivement ajoutées, comme la possibilité de séparer les activités de la Police de celle de la Gendarmerie, ou bien la disponibilité de fonctions d'analyse statistiques métier plus pointues. □