Loi «Renseignement» : Ce que vous avez vu dans les séries TV pourrait bien se passer en vrai | Le Net Expert Informatique

	Loi «Renseignement» : Ce que vous avez vu dans les séries TV pourrait bien se passer en vrai

Quand la réalité rejoint la fiction. Le projet de loi renseignement, qui va être défendu par le gouvernement dans l'hémicycle du Sénat à partir de ce mardi, va « légaliser » certaines pratiques déjà utilisées par les services de renseignement. Les données récupérées avec ces nouveaux outils vont pouvoir être versées au dossier judiciaire des suspects.

Loi «Renseignement»: Les séries TV savent ce… *par 20Minutes*

Si elle fait l'objet d'un large consensus parmi la majorité des parlementaires, cette loi est contestée par les sénateurs communistes qui ont déposé une série d'amendements de suppression et ont dénoncé un risque de « surveillance de masse ». La plupart des techniques sur le point d'être légalisées sont déjà utilisées. Et diffusées dans les séries TV. Florilège…

Poser un mouchard sous une voiture
Dans Breaking Bad (Episode 9, Saison 5), Walt accuse Hank qui travaille pour la DEA, la brigade des stupéfiants américaine, d'avoir posé un tracker GPS sous sa voiture. Le projet de loi prévoit l'emploi de balises « permettant de localiser en temps réel un véhicule ou un objet ».

Mettre un appartement sous vidéosurveillance

Dans la deuxième saison de Scandal, l'appartement de l'avocate Olivia Pope est placé sous vidéo-surveillance par Jake Ballard, le fidèle ami du président. Elle s'en rend compte dans le 18e épisode. Des caméras partout, ainsi que des micros quasiment indétectables sont utilisés. Le projet de loi permettra aux services de renseignement d'appliquer ce type d'écoutes. Les policiers passeront cependant à travers le filtre de la Commission nationale de contrôle des techniques de renseignement (CNCTR). Les plus sceptiques regrettent le pouvoir amoindri de cet organe de contrôle.

Géolocaliser un téléphone portable

Dès le premier épisode de la saison 1 du Bureau des Légendes, Cyclone, un des clandestins du BDL, est arrêté à Alger alors qu'il est ivre au volant d'une voiture. Le Bureau des Légendes va s'inquiéter : Cyclone étant musulman pratiquant, il n'aurait pas dû être saoul. Sisteron décide alors de géolocaliser son téléphone portable. Le signal du mobile indique qu'il se trouve bien au commissariat.

Intercepter les métadonnées d'un téléphone
Dans la série américaine Those who kill, Catherine Jensen, experte en tueurs en série, fait appel à un détective de la brigade des stupéfiants pour mettre sur écoute un suspect. Ce dernier, à l'épisode 9 détaille comment les policiers parviennent à récupérer les données enregistrées sur un téléphone portable, en se faisant passer pour une antenne relais après avoir copié la carte sim. Dans la « vraie vie », les policiers utilisent des Imsi-Catchers qui peuvent intercepter dans un rayon donné toutes les données qui transitent via un téléphone. Cette technologie, rendue possible par la loi renseignement, fait pourtant polémique.

Capter un écran d'ordinateur en direct grâce à un logiciel espion
Les experts de CSI : Cyber (saison 1, épisode 1), mettent au point ce qu'ils appellent un RAT (Remote Administration Tool), un outil d'administration à distance. En clair, un programme permettant la prise de contrôle total, à distance, d'un ordinateur depuis un autre ordinateur. Ils y ont introduit un logiciel espion qui permet, en fonction de mots-clés utilisés dans un mail, d'activer une alarme. Ils peuvent aussi capter en direct le mot-clé qui est tapé. Les défenseurs de la liberté numérique dénoncent à travers la loi Renseignement la surveillance massive des ordinateurs des internautes.

Source : http://www.20minutes.fr/societe/1621371-20150602-video-loi-renseignement-vu-series-tv-ca-pourra-passer-vrai
Par William Molinié

Protection des données personnelles : les entreprises bel et bien contraintes | Le Net Expert Informatique

	Protection des données personnelles : les entreprises bel et bien contraintes

Pensée pour protéger le citoyen, la loi Informatique et libertés est de plus en plus détournée de son objectif premier. Tant par les salariés que par les entreprises elles-mêmes, qui n'hésitent plus à s'en servir comme arme concurrentielle. L'analyse de l'avocat François Coupez.

La protection des données à caractère personnel est née en France avec la loi du 6 janvier 1978 dite « Informatique et libertés ». Le texte a été modifié en 2004 (à la suite de la directive européenne 95/46), et il est destiné à l'être à nouveau par le projet de loi sur le numérique annoncé en grande pompe depuis deux ans maintenant… avant d'être de toute façon complètement remplacé par un projet de règlement européen (http://www.europarl.europa.eu/oeil/popups/ficheprocedure.do%3Freference=2012/0011(COD)%26l=fr) encore en discussion qui unifiera en 2017 ou 2018 le droit de tous les pays de l'Union européenne sur le sujet.

Si ces différents projets visent à accroître de façon très importante les sanctions financières, ils ont également pour but de permettre une application plus efficace des règles (droit à l'oubli numérique/au déréférencement, co-responsabilité des sous-traitants, etc.). Mais en parallèle, on constate depuis quelques années le développement d'une véritable instrumentalisation de cette protection légale, aux règles extrêmement formelles et aux impacts potentiellement dévastateurs[1] sur l'image des entreprises prises en faute.

Salariés et clients, quand le pouvoir change de camp

Historiquement, la CNIL a eu l'occasion d'appliquer les principes de la loi « Informatique et libertés » dans plusieurs domaines, avec la plupart du temps deux points communs : d'une part la protection des clients contre l'utilisation qui serait faite de leurs données en contradiction avec les règles applicables et, d'autre part, la protection des salariés dans des hypothèses de surveillance abusive, de discrimination ou de mode d'évaluation des performances illicites.

Dans les deux cas, l'action de la CNIL conduit souvent l'entreprise fautive à revoir beaucoup plus globalement l'ensemble de ses processus et leur conformité.

Or les difficultés pratiques concernant le respect de cette réglementation pour l'entreprise ne doivent pas être sous-estimées. Elles tiennent tant à son formalisme qu'à ses conditions d'application, étant entendu que les traitements de ce type de données se développent de façon exponentielle avec la transformation numérique. De plus, l'entreprise, confrontée à un lacis réglementaire croissant et dans tous les domaines, alloue parfois ses ressources pour se mettre en conformité en fonction de l'urgence, ou du risque réel de sanction. Les entreprises ne peuvent ainsi pas toujours prétendre réussir un sans-faute en matière de protection des données personnelles… et en sont pleinement conscientes.

En parallèle, un phénomène se développe depuis quelques années, à un point tel qu'il se généralise. Sentant la faille, des clients ou des salariés bien informés n'hésitent plus à l'utiliser, non pour faire valoir leurs droits en la matière, mais pour faire pression dans le cadre d'un contentieux ou d'une revendication autre. La réglementation devient alors un simple prétexte destiné à faire plier son opposant.

Concernant le cas des clients, cela concerne souvent les entreprises disposant de nombreux points de contact avec la clientèle (et disposant de nombreux conseillers clientèles, etc.). Dans les grands réseaux, il est toujours plus difficile de faire respecter à tous les salariés en contact avec la clientèle les règles de base (notamment concernant la zone de « bloc-note » ou de note en champ libre sur les fiches clients, propices à tous les excès), ce qui multiplie les hypothèses de manquements ;

Quant aux salariés, il est de plus en plus fréquent qu'ils fassent jouer leurs droits en la matière. Par exemple, l'une des pratiques les plus fréquentes est de systématiquement avoir recours au droit d'accès aux données personnelles que leur employeur collecte sur eux, lorsque le contrat de travail arrive précocement à son terme, ou que les deux parties se retrouvent aux Prud'hommes. La pratique montre ainsi que, sur l'ensemble des personnes pouvant faire valoir leur droit d'accès dans le cas de traitements réalisés par une entreprise, près de 75% des demandes proviennent de l'interne et donc des salariés. Ainsi, il n'y a qu'à regarder la jurisprudence en droit social ces dernières années pour s'apercevoir qu'il est devenu aussi courant d'alléguer un traitement de données personnelles contraire à la loi, et donc de l'illicéité du moyen de preuve opposé à un salarié, que d'en appeler aux pages Facebook en matière de divorce. Un exemple récent nous vient de l'arrêt de la Cour d'appel de Rouen rendu le 12 mai 2015 qui invalide les preuves concernant d'une part un système de badgeage (pas d'information du comité d'entreprise) et d'autre part un logiciel permettant de contrôler les horaires des salariés (pas de formalités CNIL) : le licenciement est ainsi considéré comme étant sans cause réelle ni sérieuse.

Maintenant les contentieux… entre entreprises ?

Ce qui est plus marquant encore, c'est que ce phénomène est en passe de gagner les relations entre entreprises.

Alors que l'on s'attend à ce que ce soit la victime (client, salarié, etc.) qui fasse valoir les droits qui lui sont reconnus, les tribunaux sont en effet saisis de façon croissante de manquements à cette réglementation allégués par… des sociétés concurrentes.

Pour mettre fin à un partenariat commercial, annuler une vente, tenter de prouver une rupture abusive des relations commerciales ou empêcher un concurrent de commercialiser un service innovant, les hypothèses se multiplient dans lesquelles des tribunaux de tout type sont confrontés à cette situation.

En voici quelques exemples :

le 25 juin 2013, la Cour de cassation a rendu une décision conduisant à l'annulation de la vente d'un fichier de clients informatisé. Dans cette affaire, les associés d'une entreprise avaient vendu pour 46 000 € le seul fichier des clients de l'entreprise, fort de 6 000 clients référencés depuis 1946. Or pour l'acheteur ayant utilisé quelques semaines cette base, celle-ci était une coquille vide de 1 950 clients actifs seulement. Il en demandait donc le remboursement… qu'il obtint : pour la Cour de cassation, l'absence du respect des formalités CNIL rend toute commercialisation du fichier impossible, la vente ayant nécessairement un objet illicite.

À la suite d'une décision de la CNIL du 8 septembre 2011 autorisant pour la première fois une entreprise à traiter pour des raisons commerciales le numéro NIR (aussi appelé « numéro de sécurité sociale »), une entreprise concurrente a formé un recours considérant que l'interprétation était contestable au sens de la loi Informatique et libertés et qu'elle conduisait à un avantage concurrentiel injustifié. C'était le premier recours intenté à l'encontre d'une décision d'autorisation, alors qu'en général – et logiquement – les recours sont formés en cas de refus de la CNIL. Or, le Conseil d'État, s'il a confirmé la décision de la CNIL le 26 mai 2014, a surtout reconnu le droit à agir de la société concurrente dans cette affaire (voir, à ce sujet, l'excellent article de Guillaume Desgens-Pasanau dans Expertises N° 397 de Décembre 2014: « Données personnelles : ouverture de l'usage du NIR au secteur privé »).

Dans une affaire récente de rupture abusive alléguée de relations commerciales, la société se plaignant de la rupture (société B) proposait à l'autre société (A) de numériser pour elle des documents dans lesquels figuraient des données personnelles, et d'effectuer cette prestation depuis le Vietnam. La société A aurait donc dû demander l'autorisation de la CNIL du fait des flux de données vers ce pays, ce qu'elle n'a pas fait. Inaction qui, pour la société B, constitue un élément de preuve que la société A ne croyait en réalité pas au projet et ne comptait pas sérieusement contracter avec elle. La Cour d'appel de Paris toutefois, pour des raisons de défaut de preuve, n'a pas suivi cette analyse et a considéré le 10 avril 2015 qu'il n'y avait pas de rupture abusive.

Le grand classique des contentieux de demain ?

On le voit à travers ces quelques exemples jurisprudentiels récents, le phénomène va croissant. Il est surtout appelé à prendre encore de l'ampleur avec le futur projet de règlement européen, qui conduit à remplacer les formalités préalables par un contrôle constant de conformité et oblige donc à documenter la façon dont les traitements sont opérés à toutes les étapes. Or toutes ces informations forment un vivier de preuves de ce qui a été fait (ou pas), destinées au régulateur… et qui pourraient facilement être utilisées par une société concurrente dans le cadre d'un procès.

Plus globalement, les entreprises doivent prendre conscience de cette évolution et en saisir toutes les opportunités, mais également tous les risques : il semble logique que les études de risque, réalisées préalablement à la mise en œuvre de traitement de données à caractère personnel, aient également à prendre en compte cette nouvelle donne…

A terme en effet, en cas de contentieux et dès que l'on parlera de près ou de loin de données, la vérification de la licéité des traitements de données personnelles de l'entreprise adverse pourrait devenir un préalable aussi convenu que la vérification des pouvoirs du signataire d'un acte.

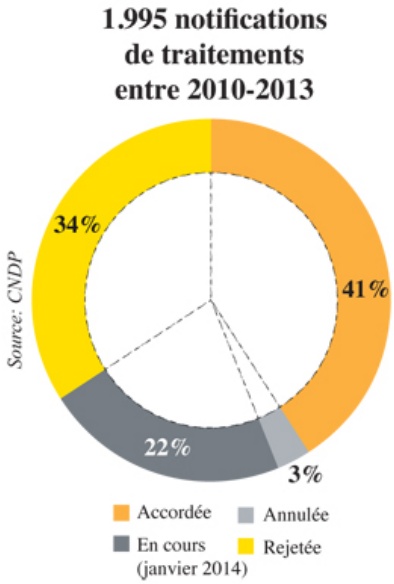
Si cette évolution peut paraître critiquable car compliquant encore les dossiers en justice, elle est malgré tout le signe que la réglementation sur les données personnelles s'ancre profondément dans les habitudes. Un réel progrès, et qui n'était pas chose évidente il y a encore quelques années…

[1] Certes, 17 textes pénaux prévoient une sanction de 5 ans d'emprisonnement et de 1 500 000 € d'amende pour les entreprises qui enfreindraient les règles en la matière, mais les applications jurisprudentielles sont rarissimes. Les sanctions de la CNIL sont quant à elles beaucoup plus fréquentes, avec des montants financiers pour le moment limités à 150 000 € (le double en cas de récidive), seul Google inc. ayant été condamné à une telle peine. Leur efficacité est fortement renforcée par leur publication (fort effet d'image sur les grandes entreprises).

Source : http://www.silicon.fr/protection-donnees-personnelles-loi-instrumentalisee-116895.html
Par François Coupez, Avocat à la Cour, Associé du cabinet ATIPIC Avocat et titulaire du certificat de spécialisation en droit des nouvelles technologies

---

Attention à l'usage abusif de la géolocalisation | Denis JACOPINI



Attention à l’usage abusif de la géolocalisation

Pour les besoins de leurs activités, certains opérateurs de transport et logistique surtout utilisent la géolocalisation. Technologie qui permet de pister un véhicule de service par exemple (voir encadré). Malgré la légitimité de leur prétention, ses utilisateurs sont-ils pour autant en règle avec la loi?

Le traitement de données peut porter atteinte à la vie privée. D'où l'interdiction par exemple de suivre les déplacements d'un salarié hors service. La réglementation en vigueur prévoit des garde-fous: finalité du traitement, nature des données collectées, durée de leur conservation, droits des personnes concernées, consentement des salariés… Une formalité de grande importance à respecter. L'entreprise doit notifier le traitement à la «Commission nationale de contrôle de protection des données à caractère personnel (#CNDP), en France, la Commission Nationale Informatique et Libertés (CNIL). Une demande de déclaration-type est mise à leur disposition. Une demande d'autorisation (et non de déclaration) distincte s'impose à la société qui procède à l'interconnexion ou au recoupement avec d'autres fichiers dont les principales finalités sont différentes. L'employeur doit donc se conformer à toutes ces règles.

Par quoi doit-il commencer?

La finalité du traitement impose que le dispositif de géolocalisation ne soit installé que dans un véhicule à usage professionnel. Une société est en droit de rationaliser la gestion de son parc automobile, d'assurer le secours de son personnel en cas d'incident ou d'accident, facturer une prestation au juste prix (kilométrage, consommation, temps…).
Garantir la sécurité des marchandises et des véhicules est également un motif légitime. L'évaluation du rendement des conducteurs est aussi envisageable. Ce cas-là est verrouillé par l'autorité de contrôle (CNDP ou CNIL): la géolocalisation n'est justifiable que «lorsqu'il n'y pas d'autres moyens» pour jauger la productivité d'un salarié. Cette exception ouvre la porte au débat: un syndicat qui, tout en cautionnant l'installation du système, s'oppose à sa prise en compte dans le rendement des salariés.

Que valent aussi les données emmagasinées et utilisées dans une procédure de licenciement pour faute grave?

La géolocalisation «ne doit être activée que si le véhicule est utilisé à des fins professionnelles», précise la CNDP dans sa Délibération n°17-2014. Voilà qui réduit sensiblement la marge d'exploitation de ce système par un employeur.
Celui-ci doit aussi être particulièrement prudent sur la nature des données collectées. Elles sont réparties en deux rubriques. Nom, prénom, coordonnées professionnelles, sont des informations liées directement au salarié. Il y a ensuite des données qui renseignent plutôt sur le véhicule suivi: numéro de plaque d'immatriculation, position géographique, kilométrage parcouru, horaire et durée d'utilisation du véhicule et de conduite, nombre d'arrêts et la vitesse moyenne de circulation. La durée de conservation est limitée à un an. Au-delà, l'exploitation de ces informations serait illégale. Annulant ainsi leur portée dans une procédure disciplinaire ou judiciaire par respect notamment du principe de la loyauté de la preuve.
Le droit d'information pèse dans la mise en place du système de géolocalisation. Au même titre que l'utilisation des caméras de surveillance, une notice s'impose pour informer les salariés: nom du responsable du traitement, nature et destinataires des données, finalité de la collecte…
A moins de justifier «l'existence d'une dérogation», le consentement «libre et éclairé» des conducteurs est indispensable. Exemple: l'insertion d'une clause «géo-localisation» dans le contrat de travail des futures recrues. Toutefois, l'information préalable des instances représentatives des employés demeure la règle. Une obligation à respecter avant l'installation du dispositif de géolocalisation. Seul le gestionnaire du parc automobile et le service ressources humaines, éventuellement, peuvent accéder aux données. Les responsables de traitement doivent donc être identifiables au sein de l'entreprise et sont, en cas de contrôle, les interlocuteurs des agents assermentés de la CNDP. Ils ont pour charge de veiller à la sécurité et à la confidentialité des données. La divulgation d'une information ou son exploitation abusive engagent la responsabilité civile, voire pénale, du dirigeant et des responsables de traitement.

En droit à double tranchant

La géolocalisation permet d'identifier la position d'un objet ou d'une personne en ayant recours aux bornes wifi, antenne GSM et position de constellation satellitaire. L'une de ses utilisations les plus courantes est le suivi des véhicules de service: optimiser les déplacements, le temps, la consommation de carburant.

«Consciente de son utilité», la commission nationale de contrôle de protection des données à caractères personnel (CNDP) définit des règles aux responsables de traitement afin que le dispositif de géolocalisation soit conforme avec la loi 09-08 protégeant les données privées. Toute infraction à la réglementation expose son auteur à de lourdes amendes, voire dans certains cas à l'emprisonnement.

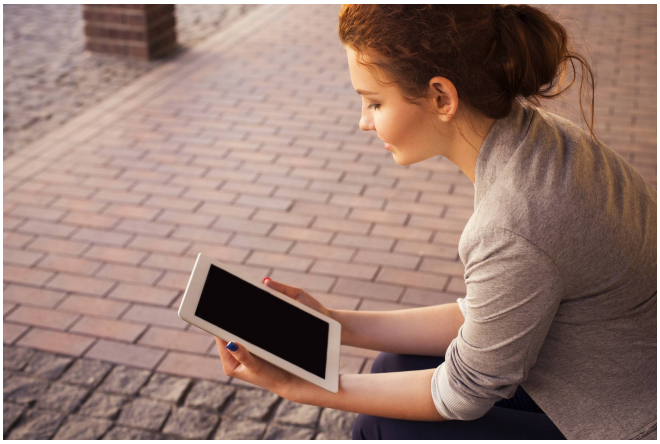
Ces informations concernent le Maroc. Un équivalent pour la France Existe.
Nous organisons régulièrement des **actions de sensibilisation ou de formation** au risque informatique, à l'hygiène informatique, à la cybercriminalité et à la mise en conformité auprès de la CNIL. Nos actions peuvent aussi être personnalisées et organisées dans votre établissement.
Besoin d'informations complémentaires ?
Contactez-nous
Denis JACOPINI
Tel : 06 19 71 79 12
formateur n°93 84 03041 84

Expert Informatique assermenté et formateur spécialisé en sécurité Informatique, en **cybercriminalité** et en **déclarations à la CNIL**, Denis JACOPINI et Le Net Expert sont en mesure de prendre en charge, en tant qu'intervenant de confiance, la sensibilisation ou la **formation de vos salariés** afin de leur enseigner les bonnes pratiques pour assurer une meilleure sécurité des systèmes informatiques et améliorer la protection juridique du chef d'entreprise.
Contactez-nous

Cet article vous plait ? Partagez !
Un avis ? Laissez-nous un commentaire !

Source : http://www.leconomiste.com/article/971816-attention-l-usage-abusif-de-la-geolocalisation
Par Faical FAOUZIHI

Mouchards sur les ebooks : Big brother is reading you ! | Le Net Expert Informatique



Mouchards sur les ebooks : Big brother is reading you !

Grâce aux « mouchards numériques », il est désormais possible de savoir si un livre a été lu jusqu'au bout. Amazon, Apple, Google et Kobo en savent beaucoup plus sur vos habitudes de lecture que vous ne le pensiez…

Eric Zemmour a vendu plus de 400 000 exemplaires de son essai « Le suicide français ». Mais seulement 7,3 % des lecteurs l'ont lu jusqu'à la fin ! L'économiste Thomas Piketty fait un peu mieux : 9,7 % des lecteurs ont terminé son pavé de près de 1 000 pages (Le capital au XXI ème siècle). Encore mieux, le dernier roman de Patrick Modiano, Prix Nobel 2014 (Pour que tu ne te perdes pas dans le quartier) affiche un honorable taux de 44 %. Quant à Valérie Trierweiller (Merci pour ce moment), son score d'achèvement est, de loin, le meilleur : environ 66 % des lecteurs sont allés au terme des mésaventures sentimentales de l'ex compagne de François Hollande.
Comment connaît-on ces taux de lecture avec une telle précision ? Tout simplement grâce aux « mouchards » numériques installés sur nos liseuses et nos tablettes. Ces instruments dédiés à la traçabilité permettent en effet de collecter une série de données sur le comportement des e-lecteurs : nombre de pages lues, vitesse de lecture, temps passé sur une page, heures de lecture, surlignage…
Ces chiffres proviennent des statistiques collectées par Kobo (partenaire de la Fnac) l'un des leaders de la lecture numérique dans le monde. Autant dire qu'ils ne sont pas passés inaperçus, notamment auprès des lecteurs les plus attentifs aux questions de confidentialité des données. Mais il faut reconnaître à Kobo une qualité : il dit ce qu'il fait. L'un de ses responsables, Nathan Maharaj, a publiquement présenté ce dispositif de mesure de « l'engagement du lecteur » lors du salon du Livre de Francfort qui s'est tenu au mois d'octobre dernier. Selon Kobo, ces données ne seraient exploitées qu'à des fins statistiques ; surtout, elles seraient anonymisées et non rattachées à un compte lecteur. En revanche, elles sont revendues aux éditeurs qui peuvent ainsi accéder à des données inédites sur le comportement réel des lecteurs.
Lire la suite….

Source : http://www.archimag.com/bibliotheque-edition/2015/05/22/mouchards-ebooks-big-brother-reading-you

E-marchands : faut-il confier vos données à Google et Facebook ? | Le Net Expert Informatique

Les deux plateformes risquent-elles de réutiliser vos données pour vos concurrents ? Quelles informations leurs fournissez-vous déjà ? Comment vous protéger à l'avenir ?
Google et Facebook ayant des modèles économiques avant tout publicitaires, ils sont amenés à collecter de plus en plus de données auprès de leurs clients annonceurs. Parmi les informations que les marchands leur transmettent déjà, la première est tout simplement ladite publicité, qui elle-même va générer plusieurs données : d'une part qui lui est exposé, d'autre part qui clique ou pas. « Ces données sont collectées par Google et Facebook, et l'annonceur doit négocier pour les obtenir » explique Thibaut Munier, cofondateur et DG de 1000mercis.

 Thibaut Munier, DG de 1000mercis © S. de P. 1000mercis

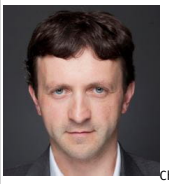
Par ailleurs, l'e-commerçant va communiquer des données de transformation à Google et Facebook, qui placent des tags sur les pages du site, tunnel de conversion compris. Mais si les deux plateformes essaient d'obtenir une meilleure vision de ce qui se passe chez les annonceurs, c'est dans le but de mieux les servir, assure Thibaut Munier.
Qui analyse : « C'est du donnant-donnant et le rapport de force se construit petit à petit, avec les avancées technologiques et les besoins des sites ».

D'autre part, le catalogue produit contient également des données importantes. Que Google les récupère en tant que données publiques sans demander leur avis aux marchands ou que ces derniers les transmettent, par exemple pour personnaliser leurs bannières publicitaires sur Facebook en fonction de leur catalogue, il s'agit encore d'un bloc de données supplémentaire dont les deux plateformes peuvent prendre possession. « Et pour les services de people-based marketing de plus en plus nombreux, comme les 'custom audiences' de Facebook, les marchands sont aussi amenés à charger non plus leurs produits mais leurs clients, afin de cibler soit leurs clients soit leurs non-clients », ajoute Thibaut Munier.

Négocier et ne pas tout donner
A quoi dès lors le marchand doit-il veiller ? D'abord, à bâtir un rapport de force lui permettant de récupérer auprès de Google et Facebook les données que génèrent ses publicités. Et bien sûr à les utiliser, idéalement en les déversant dans sa DMP, qu'il alimentera avec un maximum d'informations.

« Google et Facebook en savent plus que le marchand sur ses prospects »
« Entre les tags et les dispositifs d'identité unifiée comme Facebook Connect et Google+, Google et Facebook ont accès au parcours continu de l'internaute et savent même ce qu'il fait avant et après avoir visité un site marchand, remarque Christophe Camborde, cofondateur et PDG d'Ezakus. Ils en savent donc davantage que le marchand sur ses prospects. » Pour Thibaut Munier, raison de plus pour bien réfléchir à quels tags mettre sur son site. « Le pire est de tout taguer et de ne rien en faire. Si on met des tags, il faut les utiliser, faire des tests et se battre pour récupérer des informations dans l'autre sens », recommande-t-il. Conseil très similaire à propos du catalogue produit (et de la base clients) : se demander si on le charge ou pas et avec quelle granularité. Des questions à considérer aussi à l'aune du contexte concurrentiel plus ou moins sensible du marchand, bien sûr.

La valeur (et la marge) pourrait être transférée avec les données
Le marchand court-il le danger de perdre une partie de sa connaissance client au profit de Google et Facebook ? Christophe Camborde se veut d'abord rassurant : « Jamais ils n'utiliseront les données d'un Cdiscount pour fournir un meilleur service à un Rueducommerce. Garder un secret pareil serait impossible. » En outre, ce qui serait mauvais pour les marchands le serait à terme aussi pour Google et Facebook qui, s'ils « tuaient » leurs clients, n'auraient plus de revenus publicitaires à engranger.

 Christophe Camborde, PDG d'Ezakus © S. de P. Ezakus
« En revanche, une dépendance très forte des marchands va se créer envers Google et Facebook, qui finiront par mieux connaître leurs clients qu'eux, anticipe le PDG d'Ezakus. Lequel prend l'exemple de BigQuery. Cet équivalent de Google Analytics en big data est déjà capable de répondre à une requête du type : montre-moi mes clients qui ont dépensé plus de 200 euros ces quatre derniers mois. « Pour un marchand, pourquoi ne pas utiliser cela plutôt que son CRM interne ? Or avec chaque nouveau service fourni par les deux plateformes, avec chaque morceau de connaissance client et donc de valeur qui se transfère chez elles, c'est une partie de la marge du marchand qui partira aussi chez elles », souligne Christophe Camborde.
Raison pour laquelle il est urgent de monter en expertise sur ces sujets, répond Thibaut Munier. Le marchand est obligé de fournir des données, mais il doit être conscient de ce qu'il donne et de ce qu'il en retire. Pour le DG de 1000mercis, « il faut savoir quelles données ont quelle valeur et comment être pertinent dans leur utilisation. Et éventuellement se doter d'outils pour cela, au premier rang desquels une DMP, meilleure façon pour l'annonceur de protéger ses données. A ces conditions, il est possible d'en retirer des bénéfices. » Le dirigeant établit ainsi un parallèle avec les marketplaces. Certains marchands y commercialisent tout leur catalogue et transmettent leur valeur à Amazon, certains refusent tout en bloc et se privent d'un apport de revenus… et d'autres ne donnent pas tous leurs meilleurs prix, pas tout leur catalogue, et jouent sur plusieurs paramètres afin d'en sortir gagnants.

« On ne peut confier son CRM ou sa DMP à Google ou Facebook »
D'autant que pour Christophe Camborde, pas moyen de faire sans Google et Facebook. « C'est une fatalité, les marchands sont obligés d'y aller. Ceux qui bénéficient d'une clientèle très fidèle, sur une niche, pourront s'en passer. Pas les gros généralistes. »

Se renforcer pour mieux se protéger
Un plan d'action se dégage donc : répartir ses investissements pour ne pas dépendre d'une seule plateforme et travailler la fidélisation et le lien direct avec les consommateurs. « Un fan n'est pas un client », insiste Thibaut Munier, considérant pour sa part qu'on ne peut confier son CRM ou sa DMP à Google ou Facebook. « L'actif du marchand, c'est sa base de clients, sa DMP et son expertise dans ses investissements publicitaires. » Et de marteler : « il existe beaucoup de manières d'être exigeant dans sa relation avec Google et Facebook et beaucoup de manières d'être actif pour tester des choses nouvelles et mesurer ce qu'on en retire. »
Le nombre extrêmement restreint de marchands français disposant d'une DMP montre toutefois que même s'ils sentent qu'il leur faut organiser et protéger leurs données, ils n'investissent encore que très peu dans la data et misent en majeure partie sur le court terme : la publicité. La Redoute a une DMP, selon nos informations Carrefour et Voyages-Sncf.com y travaillent… et Cdiscount et la Fnac y ont réfléchi. La barrière de protection data des e-commerçants français n'est pas encore en place.

Après cette lecture, quel est votre avis ?
Cliquez et laissez-nous un commentaire…

Source : http://www.journaldunet.com/ebusiness/commerce/donnees-e-commerce.shtml?een=4a4b0e45c54d9fed8fc26819a6b6f84f&utm_source=greenarrow&utm_medium=mail&utm_campaign=ml50_e-marchandsetle

Les opérateurs télécoms nous espionnent-ils ? | Le Net Expert Informatique

Les opérateurs télécoms nous espionnent-ils ?

Votre téléphone est-il sur écoute ? Depuis plusieurs semaines, une affaire secoue le milieu des télécoms suite à la révélation du journal arabophone Al Massae quant à l'utilisation d'un logiciel «d'espionnage des données personnelles» par un opérateur télécoms de la place. Il s'agit du LCS, un logiciel en principe prohibé en Europe et aux États-Unis, qui permet de surveiller l'activité des utilisateurs en dehors de tout cadre légal.

La question qui se pose aujourd'hui : les opérateurs télécoms ont-il le droit d'utiliser les outils qu'ils ont pour contrôler et accéder aux informations et aux données des utilisateurs ? «Il s'agit d'un système permettant de retracer toutes les actions d'un utilisateur sur sa ligne téléphonique et d'effectuer les perquisitions numériques, explique Carlo Lando, un expert italien en sécurité des télécoms.

«En principe, les opérateurs téléphoniques doivent avoir des autorisations du tribunal pour utiliser cette technologie de système sur les réseaux, notamment pour les enquêtes», précise l'expert. Interpellé, le DG de l'Agence nationale de régulation des télécoms (ANRT), Azeddine Mountassir Billah, dit tout ignorer de cette affaire et refuse de la commenter.

En revanche, à la Commission nationale de contrôle de la protection des données à caractère personnel (CNDP), on apprend qu'une réunion aura lieu cette semaine, avec à l'ordre du jour, entre autres, cette affaire de «logiciel LCS». Le président du CNDP, Saïd Ihrai, a déclaré que la commission n'a pas encore été saisie sur ce type de «procédé prohibé» permettant de surveiller et de contrôler les données personnelles des utilisateurs.

Source : http://www.leseco.ma/maroc/30623-donnees-personnelles-les-operateurs-telecoms-nous-espionnent-ils.html
Par NAIMA CHERII

L’immatriculation des drones, solution à toutes les craintes ? | Le Net Expert Informatique

L’immatriculation des drones, la la solution à toutes les craintes ?

L’immatriculation des drones de loisir est-elle une solution efficace pour responsabiliser leurs propriétaires ? À vous de juger !

Doter les drones de loisir de plaques d’immatriculation : c’est l’une des pistes de réforme envisagées par le gouvernement pour éviter les survols intempestifs de ces petits robots volants au-dessus de la capitale et aux abords de centrales nucléaires. La soixantaine d’incidents recensés ces derniers mois a en effet révélé les lacunes de la réglementation et des systèmes de détection et d’interception existants.

Deux projets ont été sélectionnés par l’Agence nationale de la recherche (ANR) en vue de relever le défi que ces engins provocateurs lancent aux autorités. Des systèmes de captation de signaux entre le pilote et l’appareil et de brouillage GPS forçant le drone à atterrir sont en cours d’expérimentation. Il est aussi question de doter ces appareils de puces d’identification.

La dissuasion passe également par des sanctions plus lourdes que celles encourues actuellement pour le non-respect des règles de sécurité, à savoir un an d’emprisonnement et 75 000 euros d’amende, outre les peines encourues pour mise en danger de la vie d’autrui. Car les drones présentent des risques, peuvent blesser des gens, s’écraser sur une route ou sur une piste d’aéroport. Une collision avait été évitée de justesse entre un A320 et un drone à l’aéroport de Heathrow en juillet 2014…
Faut-il obliger les propriétaires de drones à les faire immatriculer à leurs frais comme le font les propriétaires d’aéronefs civils ? À vous de juger – et de voter après avoir regardé nos deux expertes, Myriam Quéméner et Christiane Féral-Schuhl, plaider le « pour » et le « contre » en… trois minutes !

Faut-il immatriculer les drones ? *par LePoint*

Cet article vous plait ? Partagez !
Un avis ? Laissez-nous un commentaire !

Source : http://www.lepoint.fr/justice-internet/au-tribunal-de-l-internet-faut-il-immatriculer-les-drones-18-05-2015-1929083_2081.php
Par Laurence NEUER ET Anne-Sophie JAHN

Enjeux et défis du web profond | Le Net Expert Informatique

Le web profond (Deep Web) désigne le sous-ensemble d'internet qui n'est pas indexé ou mal indexé par les grands moteurs de recherche comme Google, Yahoo ou Bing,…On sait que cet ensemble de données reste difficilement mesurable mais qu'il occupe un espace très supérieur à celui de l'ensemble des sites web bien indexés par les moteurs classiques. Certaines études avancent un ratio de 80% de Deep Web contre 20% de web de surface à l'image de la partie immergée d'un iceberg…

Profond comme le web

Le contenu du deep web demeure hétérogène. On y trouve de grandes bases de données, des bibliothèques volumineuses non indexées par les moteurs en raison de leur tailles, des pages éphémères, mal construites, à très faible trafic ou volontairement rendues inaccessibles par leurs créateurs aux moteurs traditionnels.
D'après une étude récente de la Darpa, l'agence américaine en charge des projets de défense, plus de 60 millions de pages à vocation criminelle ont été publiées depuis deux ans dans les profondeurs du web. Les moteurs de recherche classiques, Google en tête, utilisent des algorithmes d'indexation dérivés du puissant Pagerank qui s'appuient sur une mesure de popularité du site ou de la page.
Cette approche qui a fait le succès de Google va de fait exclure les pages à faible trafic, éphémères ou furtives. Ce sont précisément ces pages qui sont utilisées par les acteurs de la cybercriminalité pour diffuser de l'information tout en restant sous les radars des grands moteurs. Lorsque cette information concerne une activité criminelle, c'est dans le Dark Web qu'elle sera dissimulée et rendue accessible aux seuls clients potentiels via des outils d'anonymisation spécialisés comme Tor. Le web profond réunit donc de la donnée légitime, souvent de haute qualité lorsqu'il s'agit de bases de données scientifiques volumineuses peu ou mal indexées par les moteurs.
Il réunit de la donnée sécurisée accessible seulement par mot de passe mais aussi de la donnée clandestine issue de trafics et d'activités criminelles. Cet ensemble informationnel hétérogène intéresse depuis longtemps les grands acteurs du numérique, chacun avec une motivation spécifique. L'accès au web profond constitue un élément stratégique du dispositif global de lutte contre la cybercriminalité qui reste l'une des grandes priorités de l'administration américaine. Les efforts pour obtenir des capacités de lecture du web profond se sont concrétisés avec le développement en 2014 du moteur de recherche Memex tout droit sorti des laboratoires de la Darpa.

Memex, le moteur Darpa

Dans son communiqué officiel publié le 9 février 2014 [1], l'agence Darpa décrit Memex comme « le moteur qui révolutionne la découverte, l'organisation et la présentation des résultats de recherche en ligne. Le programme Memex imagine un nouveau paradigme, où il est possible d'organiser rapidement et intelligemment un sous-ensemble de l'internet adapté à l'intérêt d'une personne ».

Le moteur est construit autour de trois axes fonctionnels:
1. l'indexation de domaines spécifiques,
2. la recherche de domaines spécifiques
3. la mise en relation de deux premiers axes

Après plus d'un an d'utilisation en phase de test par les forces de l'ordre américaines, Memex a permis de démanteler un réseau de trafiquants d'êtres humains. Durant la finale du Super Bowl, Memex a servi pour détecter les pages associées à des offres de prostitution. Ses outils d'analyse et de visualisation captent les données invisibles issues du web profond puis tracent la graphe des relations liant ces données. De telles fonctionnalités s'avèrent très efficaces pour cartographier des réseaux clandestins de prostitution en ligne.

D'après les récents communiqués de la Darpa, Memex ne traite pour l'instant que les pages publiques du web profond et ne doit donc pas être associé aux divers outils de surveillance intrusifs utilisés par la NSA. A terme, Memex devrait offrir des fonctionnalités de crawling du Dark Web intégrant les spécificités cryptographiques du système Tor. On peut raisonnablement imaginer que ces fonctions stratégiques faisaient bien partie du cahier des charges initial du projet Memex dont le budget est estimé entre 15 et 20 millions de dollars… La Darpa n'est évidemment pas seule dans la course pour l'exploration du web profond. Google a parfaitement mesuré l'intérêt informationnel que représentent les pages non indexées par son moteur et développe de nouveaux algorithmes spécifiquement adaptés aux profondeurs du web.

Google et le défi des profondeurs

Le web profond contient des informations provenant de formulaires et de zones numériques que les administrateurs de sites souhaitent maintenir privés, hors diffusion et hors référencement. Ces données, souvent très structurées, intéressent les ingénieurs de Google qui cherchent aujourd'hui à y avoir accès de manière détournée. Pour autant, l'extraction des données du web profond demeure un problème algorithmiquement difficile et les récentes publications scientifiques des équipes de Google confirment bien cette complexité. L'Université de Cornell a diffusé un article remarquable décrivant une infrastructure de lecture et de copie de contenus extraits du web profond [2],[3].

L'extraction des données s'effectue selon plusieurs niveaux de crawling destinés à écarter les contenus redondants ou trop similaires à des résultats déjà renvoyés. Des mesures de similarités de contenus sont utilisées selon les URL ciblées pour filtrer et hiérarchiser les données extraites. Le système présenté dans l'article est capable de traiter un grand nombre de requêtes sur des bases de données non adressées par le moteur de recherche classique de Google [4].

A moyen terme, les efforts de Google permettront sans aucun doute de référencer l'ensemble du web profond publiquement accessible. Le niveau de résolution d'une requête sera fixé par l'utilisateur qui définira lui même la profondeur de sa recherche. Seuls les contenus privés cryptés ou accessibles à partir d'une identification par mot de passe demeureront (en théorie) inaccessibles à ce type de moteurs profonds.

Vers une guerre des moteurs?

Les grandes nations technologiques ont pris en compte depuis longtemps les enjeux stratégiques de l'indexation des contenus numériques. Peu bruyante, une « guerre » des moteurs de recherche a bien lieu aujourd'hui, épousant scrupuleusement les contours des conflits et les concurrences de souverainetés nationales. La Chine avec son moteur Baidu a su construire très tôt son indépendance informationnelle face au géant américain.

Aujourd'hui, plus de 500 millions d'internautes utilisent quotidiennement Baidu à partir d'une centaine de pays. La Russie utilise massivement le moteur de recherche Yandex qui ne laisse que peu de place à Google sur le secteur du référencement intérieur russe puisqu'il détient plus de 60% des parts du marché national. En 2014, Vladimir Poutine a souhaité que son pays développe un second moteur de recherche exclusivement contrôlé par des capitaux russes et sans aucune influence extérieure. Plus récemment, en février 2015, le groupe Yandex a déposé une plainte contre Google en Russie pour abus de position dominante sur les smartphones Android. Yandex reproche en effet à Google de bloquer l'installation de ses applications de moteur de recherche sur les smartphones fonctionnant sous Android. Les constructeurs sont contraints aujourd'hui à pré-installer sur leurs machines les Google Apps et à utiliser Google comme moteur par défaut sous Android…

Le moteur face aux mégadonnées

La course à l'indexation des contenus du web profond apparaît comme l'une des composantes stratégiques de la guerre des moteurs. Si la géopolitique des données impose désormais aux nations de définir des politiques claires de stockage et de préservation des données numériques, elle commande également une vision à long terme de l'adressage des contenus. La production mondiale de données dépassera en 2020 les 40 Zo (un zettaoctet est égal à dix puissance vingt et un octets). L'évolution de cette production est exponentielle: 90% des données actuelles ont été produites durant les deux dernières années. Les objets connectés, la géolocalisation, l'émergence des villes intelligentes connectées et de l'information ubiquitaire contribuent au déluge de données numériques. La collecte et l'exploitation des mégadonnées (le terme officiel français à utiliser pour big data) induiront le développement de moteurs polyvalents capables d'indexer toutes les bases de données publiques quelle que soient leurs tailles et leurs contenus.

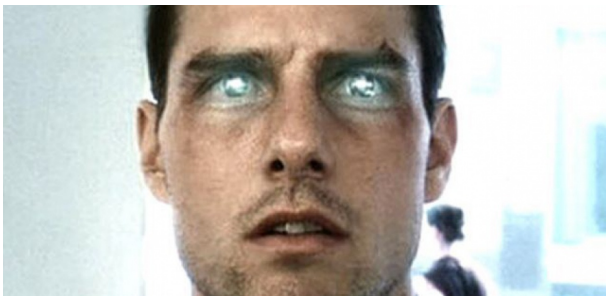
Le moteur de recherche doit être considéré aujourd'hui comme une infrastructure de puissance stratégique au service des nations technologiques. Qu'attend l'Europe pour développer le sien?

_____

[1] La présentation du moteur Memex par l'agence Darpa
http://www.darpa.mil/newsevents/releases/2014/02/09.aspx
[2] « Google's Deep-Web Crawl » – publication de l'Université Cornell
http://www.cs.cornell.edu/~lucja/publications/i03.pdf
[3] « Crawling Deep Web Entity Pages » – publication de recherche, Google
http://pages.cs.wisc.edu/~heyeye/paper/Entity-crawl.pdf
[4] « How Google May index Deep Web Entities »
http://www.seobythesea.com/2015/04/how-google-may-index-deep-web-entities/

Source : http://www.huffingtonpost.fr/thierry-berthier/enjeux-et-defis-deep-web_b_7219384.html
Par Thierry Berthier

Surveillance : votre œil vous trahira bientôt | Le Net Expert Informatique

	Surveillance : votre œil vous trahira bientôt

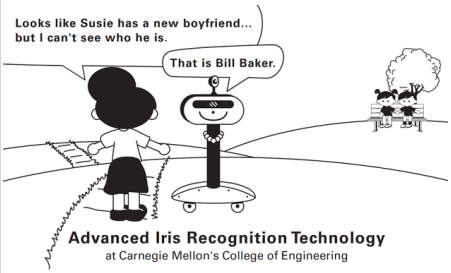
On n'arrête pas le progrès, mais en matière de surveillance, peut-on le qualifier comme tel ? La dernière trouvaille, repérée par « The Atlantic » fait un peu peur : a été mis au point un système permettant d'identifier une personne, à distance, par l'analyse de son œil. Rien de révolutionnaire, direz-vous ? Eh bien si, car les mots importants, dans la phrase précédente, sont : « à distance ».

La reconnaissance d'iris existe certes depuis longtemps, mais jusque là, il fallait que la personne à identifier coopère, qu'elle pose avec précision son œil sur un oculaire. Avec la machine mise au point par Mario Savvides, un professeur de l'université Carnegie Mellon, à Pittsburg, Etats-Unis, l'histoire sera très différente. A plus de dix mètres, assure-t-il, il est désormais possible d'analyser un iris avec la précision d'une empreinte digitale, avant de le confronter à une base de donnée.

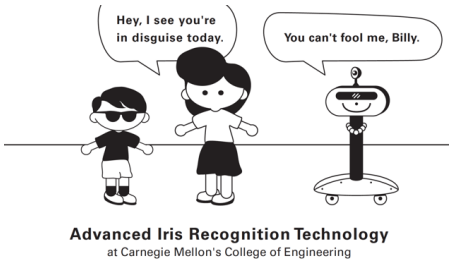
Dans une vidéo mise en ligne, le professeur d'ingénierie donne un exemple d'une utilisation possible d'une telle technologie : un policier repère un automobiliste au comportement suspect et lui demande de garer sa voiture. L'automobiliste jette un coup d'œil dans le rétroviseur et le policier en profite pour vérifier, sans avoir à sortir de son véhicule, s'il n'est pas fiché comme personne dangereuse…

On peut imaginer bien d'autres usages :
avant un match de foot, l'appareil vérifierait l'iris de chaque spectateur pénétrant dans le stade, afin de filtrer les éventuels hooligans fichés ;
un enfant est enlevé, son iris est livré aux autorités des frontières pour éviter qu'il ne soit emmené à l'étranger ;
dans une grande entreprise, une administration, ou un festival, seuls les propriétaires d'iris « VIP » pourraient accéder à certains espaces.
seuls les propriétaires d'ordinateurs ou de voitures pourront démarrer ces derniers, sans mot de passe, sans clé (et sans avoir à poser son oeil sur son volant) ;
à l'aéroport, les voyageurs pourront se passer de montrer leurs papiers.

Mais on peut craindre aussi des usages plus effrayants. Le service de presse de Carnegie Mellon a envoyé deux dessins à « The Atlantic ». Sur le premier, une jeune fille aperçoit un couple enlacé au loin : « On dirait que Susie a un nouveau petit ami… », dit-elle. Et la machine, ce robot-commère, lui dit : « Oui, c'est Bill Baxter ». Qui sait si la Susie en question n'est pas une femme politique et la héroïne du dessin une affreuse paparazzi ?



Autre dessin, la même jeune fille est face à un garçon grimé : « Eh, je vois qu'on s'est déguisé aujourd'hui ! ». La machine, froidement : « Tu ne m'aura pas, Billy ». Mignon ? Pas vraiment : une machine qui rend le déguisement obsolète ne peut guère être considérée comme un grand progrès pour la vie privée.



La police ne manquera pas de tester ce système, mais gageons qu'elle ne sera pas la seule à se pencher dessus. Les usages commerciaux, si cet appareil biométrique fonctionne, ne manqueront pas d'apparaître. On pense à ces scènes du film « Minority Report » où des publicités alpaguent les passants tout en s'adaptant à leurs goûts (« John Anderton ! Vous n'auriez pas envie d'une petite Guinness ? ») et où un hologramme, à l'accueil d'un magasin de vêtements, reconnaît les yeux que s'est greffés le héros, Tom Cruise, pour échapper à la police (« Hello Mr Yakomoto, contente de vous revoir chez Gap »).

Interrogé par « The Atlantic » sur les craintes que soulève cette technologie, Marios Savvides les balaye d'un argument pour le moins fataliste :
Les gens sont traqués, chacun de leurs mouvements, de leurs achats, de leurs habitudes, où ils se trouvent chaque jour, à travers leurs transactions par carte de crédit, leurs cartes de fidélité –si quelqu'un veut vraiment savoir ce que vous faites à n'importe quel moment de la journée, il n'a pas besoin de systèmes de reconnaissance faciale ou de reconnaissance d'iris. Tout ce qu'il faut est déjà en place. »

Autrement dit : bah, la surveillance de masse, un peu plus, un peu moins…
La mise en place d'un tel système de reconnaissance « à distance » sera facilité par la décision prise par plusieurs pays, il y a plusieurs années déjà, de constituer des bases d'iris. Aux Etats-Unis, depuis quatre ans, la police scanne ainsi les yeux des personnes condamnées à des peines de prison. Dans les Emirats arabes unis, l'iris est scanné à l'entrée et à la sortie du territoire. Et l'Inde va plus loin encore : ce sont les iris de l'ensemble de la population qui sont peu à peu associés, dans une base de donnée, à leur numéro unique d'identité.

Cet article vous plait ? Partagez !
Un avis ? Laissez-nous un commentaire !

Source : http://tempsreel.nouvelobs.com/loi-renseignement/20150515.OBS9017/surveillance-votre-il-vous-trahira-bientot.html

Par Pascal Riché

Facebook développe la reconnaissance faciale… de dos ! | Le Net Expert Informatique

Capture d’écran du film « Mon nom est personne » (1973) (Tonino Valerii)	**Facebook développe la reconnaissance faciale… de dos !**

Après la reconnaissance vocale, la biométrie et son cortège (empreinte digitale, ADN, iris…), nous pourrons être reconnus même visage caché. Une technique qui devrait surtout servir à (encore plus de) la surveillance.

Au Far-West, dans les westerns il y avait une règle d'or, un code de bonne conduite, un code d'honneur : on ne tire pas sur les gens de dos. Par exemple dans le chef d'œuvre « Mon nom est personne », Jack Beauregard (aka Henry Fonda) déclare :
« D'accord, Nevada était mon frère, mais c'était aussi un salaud de la plus belle espèce. Pour une poignée de dollars, il tirait dans le dos d'un ami, et je ne vais pas risquer ma peau pour le venger. »
Comme on n'arrête pas le progrès, après avoir développé ceux déjà redoutablement efficaces de la reconnaissance faciale, voici que Facebook annonce avoir mis au point un nouvel algorithme capable de vous reconnaître sur une photo ou dans une vidéo même si vous êtes de profil ou … de dos. Oui oui.

La reconnaissance faciale de dos
L'article dans lequel les chercheurs du labo IA (intelligence artificielle) de Facebook expliquent comment ils cherchent à aller au-delà de la seule reconnaissance faciale au travers de la détection d'une multitude d'autres « indices » est disponible en ligne [PDF]. La méthode utilisée par l'algorithme a même un petit nom sympatoche : elle s'appelle PIPER pour « Pose Invariant PErson Recognition ». En gros : « Les poses identiques pour la reconnaissance de personnes. » [illegible]
Très précisément, il ne s'agit pas de « reconnaître quelqu'un de dos » (mais avouez que ça fait un bon titre), ni même de reconnaître quelqu'un uniquement sur la base de ses mouvements ou attitudes, mais – c'est déjà pas mal – de se servir de la reconnaissance desdites attitudes (baptisées « poselets ») pour améliorer la reconnaissance faciale.
« Nous utilisons la méthode "PIPER", qui agrège les indices recueillis par un système de reconnaissance d'attitudes, entraîné par des réseaux à convolution pour lisser les variations de la pose, combiné avec une reconnaissance de visage et une reconnaissance globale. » [PDF]

Image extraite de l'étude de Facebook (DR)

Tels que présentés dans l'étude, les résultats sont assez bluffants :
« 83,05% de réussite pour les 581 identités du corpus (1 identité étant décrite par plusieurs photos) dans lesquelles on ne dispose pas de photo "de face". De plus, quand une photo de face est disponible, le taux de réussite de DeepFace (l'algo de reconnaissance faciale) passe de 89,3% à 93,4% faisant baisser de 40% le taux d'erreur relatif. » [PDF]
À noter enfin que les datasets utilisés sont directement piochés dans la base Creative Commons de Flickr [PDF].

Image extraite de l'article en question (DR)

Le poids des mots, le choc des photos
Nous sommes en train de vivre, à l'échelle de l'image et de la vidéo, la même révolution scientifique et technologique que celle que nous vécûmes au début des années 80 à l'échelle de l'ingénierie linguistique et de la fouille de corpus de textes avec l'invention du TAL (traitement automatique des langues http://fr.wikipedia.org/wiki/Traitement_automatique_du_langage_naturel) dont Jean Véronis fut un des pionniers.

Le côté positif c'est que grâce à ces progrès nous avons aujourd'hui Google, des livres numériques, nous pouvons travailler à l'échelle de corpus considérables, chercher des mots-clés facilement dans à peu près n'importe quel texte, bénéficier de services de traduction automatique de plus en plus efficaces, de correcteurs d'orthographe et de syntaxe, etc. Le côté obscur c'est que l'ensemble des technologies de surveillance dont on débat aujourd'hui au travers des [illegible]

La même chose se produira, est déjà en train de se produire, à l'échelle des technologies de ce que nous pourrions baptiser le TAIFA (traitement automatique des images fixes et animées). Et de la même manière que nous en retirerons de grands services, elles nous exposeront simultanément à de grandes dérives liées à la surveillance et au contrôle.

Il faut se souvenir qu'il y a à peine 15 ans, au début des années 2000, la plupart des spécialistes de ces questions butaient sur d'immenses difficultés pour simplement parvenir à faire avec les images ce que l'on avait réussi à faire à peu près correctement avec les textes, c'est-à-dire parvenir à les indexer.

[illegible]