

LA TECHNOLOGIE PEUT-ELLE PRÉVOIR LE FUTUR ?

Les géants du Web peuvent-ils prévoir notre avenir ? Le 28 avril, *Lyon Capitale* était partenaire de la soirée-débat “À qui profite le big data ?” organisée par la mission Sciences et Société de l’Université de Lyon, dans le cadre du cycle de rencontres “Et si on en parlait” et avec le soutien du CNRS Rhône-Alpes-Auvergne. De nombreux Lyonnais se sont déplacés pour assister à cet événement et découvrir comment toutes nos données numériques peuvent être utilisées au quotidien.



Étaient présents autour de la table (de gauche à droite) :

- ◆ **Stéphane Grumbach**, directeur de recherche à l’Inria, spécialiste du big data
- ◆ **Valérie Peugeot**, prospectiviste à Orange Labs, présidente de l’association Vecam et vice-présidente du Conseil national du numérique
- ◆ **Sébastien Nedjar**, maître de conférences au sein du laboratoire d’informatique fondamentale de Marseille

“**I**l est plus simple de prévoir les comportements humains que la météo.” Stéphane Grumbach nous rappelle la révolution qui se trame aujourd’hui dans les bureaux de Google, Apple, Facebook ou Amazon (les GAFAs), mais aussi dans de nombreuses autres entreprises, pays ou administrations. Tous les jours, nous laissons des traces numériques qu’il est possible de collecter et recouper pour ensuite nous proposer des publicités toujours plus ciblées, des produits que nous voudrions sans le savoir, améliorer les flux routiers dans une ville, voire prédire le futur. L’utilisation de ces données est regroupée sous l’appellation “big data”.

À l’origine, les traces

Les marques et États n’ont pas attendu l’ère numérique pour récolter des données, mais Internet leur a permis d’augmenter la masse des informations reçues, les obligeant à mettre en place de nouveaux outils pour les traiter. Inconsciemment, les citoyens travaillent pour ces géants et fournissent

chaque jour des centaines voire des milliers de données. Il y a les vecteurs de récolte évidents : la navigation sur Internet, l’utilisation des réseaux sociaux, les photos envoyées en ligne où les visages sont automatiquement reconnus, la montre intelligente récupérant nos données de santé comme le rythme cardiaque. Et il y a les plus discrets : le smartphone et son GPS capables de traquer tous nos déplacements. Et puis il y a ceux qu’on ignore, tels les capteurs placés dans les routes, les caméras de vidéosurveillance, la balise dans une boutique qui repère les mobiles qui passent à proximité... Demain, les lunettes connectées type Google Glass “pourront mettre sur le Net les images que l’on voit, c’est la bascule du monde réel dans le monde virtuel”, pour Stéphane Grumbach.

Le mythe de l’anonymat

L’adage moderne veut que, lorsqu’un service est gratuit, c’est souvent le consommateur lui-même qui est le produit. En mettant en place des logiques de “big data”, et les in-

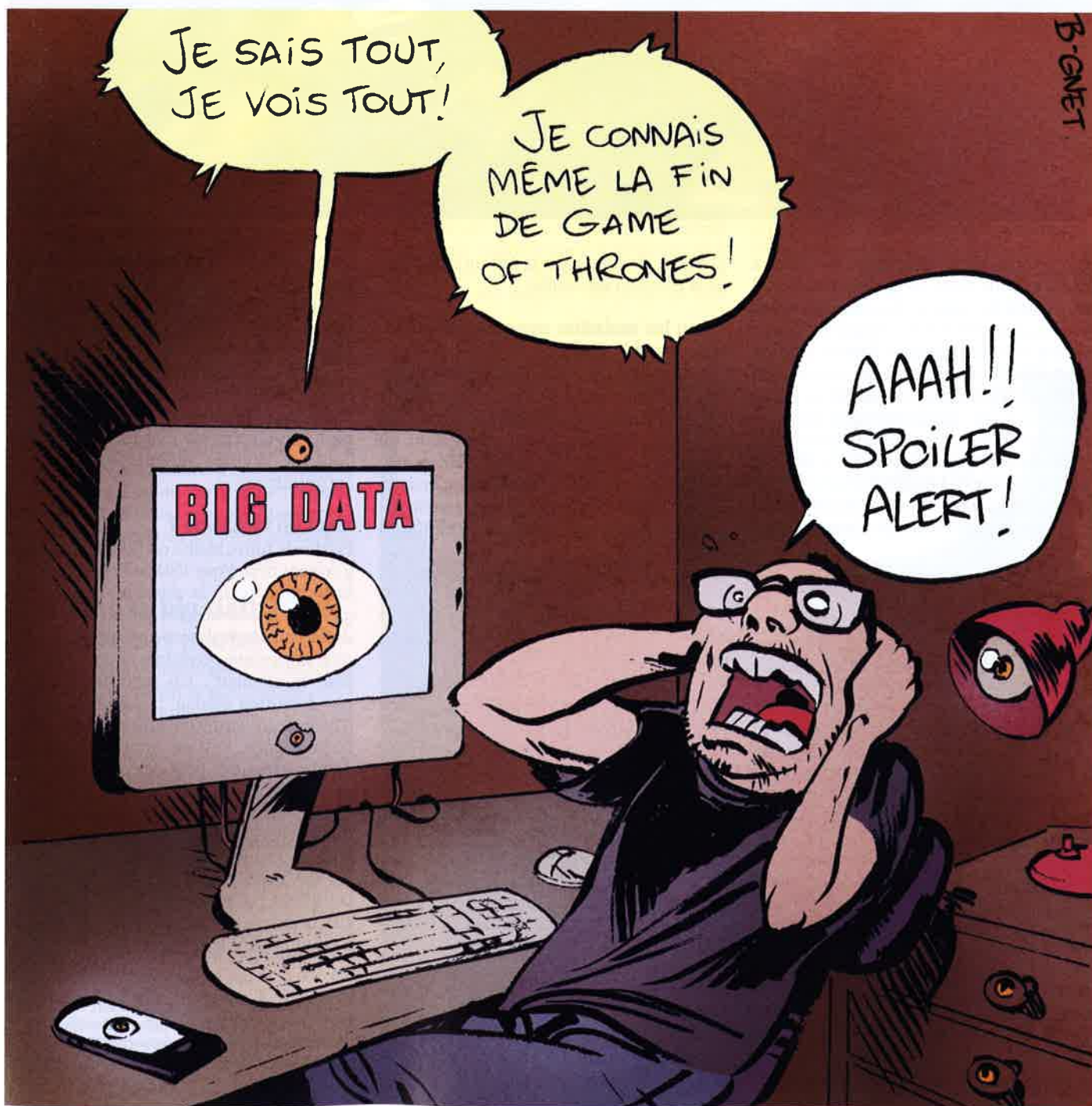
vestissements qui en découlent, les GAFAs veulent pouvoir exploiter de la donnée fiable rattachée à la véritable identité des fournisseurs, qu’ils soient simples citoyens, entreprises ou collectivités. Dans ce contexte, l’anonymat sur le Net devient un mythe difficile à atteindre, même lorsque l’on choisit de ne pas s’inscrire sur les réseaux sociaux. Un rapport commandé par l’équivalent belge de la Cnil, publié début 2015, tend à prouver que Facebook surveillerait tous les internautes, même ceux qui n’ont pas de compte sur le réseau social ou qui se sont déconnectés. Les chercheurs expliquent comment Facebook traquerait tous les internautes à l’aide des boutons “J’aime” qui ont été placés quasiment sur tous les sites Internet. Facebook a réagi à ces accusations en précisant que les données récupérées n’étaient pas utilisées à des fins publicitaires, mais pour la sécurité de son site. Facebook peut dans tous les cas récolter des informations sur autrui quand elles sont partagées par ses connaissances : numéro de téléphone, photos, données personnelles. Pour Sébastien

Nedjar, "à cause des recoupements d'infos, aucune méthode ne garantit à 100 % l'anonymisation des données sur les réseaux sociaux". Même lorsque l'utilisateur tente de se protéger en utilisant les méthodes les plus efficaces (réseaux cryptés, ou déportés), les géants ont déjà la parade : reconnaître un internaute grâce à la manière dont il utilise son clavier, les mouvements de sa souris, la nature de son matériel informatique ou l'enchaînement des sites qu'il visite. Telles des empreintes digitales, nous avons tous nos habitudes qui, recoupées, correspondent à une identité unique.

Prédire le futur...

Forts de bases de données qui grossissent chaque jour, les GAFAs ainsi que certaines start-up imaginent déjà pouvoir prédire le futur. Durant la Coupe du monde de football 2014, Cortana, l'assistant personnel de Microsoft, a été capable de dire quelle serait l'équipe victorieuse pour quinze des derniers matchs sur seize. Mais la capacité de voir le futur peut être plus globale encore. En Israël, la jeune Kira Radinsky, 28 ans, travaille sur un logiciel capable de prédire l'avenir. Kira Radinsky est l'incarnation même de la formule qui veut que l'histoire se répète.

En recoupant des événements qui se sont déroulés dans le passé avec des statistiques et en les comparant avec des milliards de données actuelles, elle affirme pouvoir dire si des événements peuvent se reproduire et quand. Dans une interview donnée à Canal+, Kira Radinsky s'amuse même de pouvoir prédire qui seront ses "concurrents" demain. Enfin, prédire le crime avant qu'il n'arrive n'est plus de l'ordre du fantasme. Le scénario du roman et du film *Minority Report* n'est pas loin. En France, la loi sur le renseignement, censée faciliter le travail des services antiterroristes, permettra l'installation de



DÉBAT

LA TECHNOLOGIE PEUT-ELLE PRÉVOIR LE FUTUR ?

Le nouvel amphithéâtre de l'Université de Lyon a fait le plein pour la conférence sur les Big Data, dans le cadre du cycle "Et si on en parlait".



boîtes noires chez les opérateurs. Grâce à un algorithme dont on ignore la nature, il sera possible de "détecter" les comportements suspects et de "révéler une menace terroriste" avant qu'elle ne se produise. Reste la question de cet algorithme et de ceux qui sont chargés de le rédiger. Difficile de savoir ce qu'est réellement un "comportement suspect". Avec des boîtes noires qui analysent tous les flux, le risque de faux positif* se fait plus fort. En matière de big data, "ce qui a changé ces dernières années, selon Stéphane Grumbach, c'est la capacité de calcul par rapport à

la masse de data. Cela se fait au détriment de la précision du résultat".

... ou les maladies avant qu'elles n'arrivent ?

De son côté, grâce à la somme des informations déjà récoltées, Google imagine déjà prédire le cancer ou les AVC. Le géant américain a publié des outils permettant de suivre des épidémies, souvent en avance sur ceux des gouvernements. Pour y parvenir, il recoupe les requêtes sur son moteur de recherche correspondant aux symptômes

d'une maladie. La santé est l'un des grands chantiers du big data, et les capteurs du type bracelets sportifs ou montres connectées sont chargés de récolter les données nécessaires. Prochaine étape : convaincre les citoyens d'envoyer des prélèvements ADN à l'un des GAFAs, actuel ou futur, pour qu'il réalise un séquençage et indique les prédispositions à certaines maladies ou cancers. Tout est "donnée" aujourd'hui, même le corps humain.

Juste un fantasme ?

Prédire le futur, réalité ou fiction ? Certains y croient déjà. Pour Valérie Peugeot, "notre fascination pour la donnée vient du fait qu'elle a une dimension prédictive pour les déplacements, maladies, conflits. Cependant, cette dimension prédictive ne se vérifie pas pour le moment". Un sentiment partagé par Sébastien Nedjar : "La data est loin d'avoir des capacités prédictives fiables." Dans notre monde, tout ne serait donc pas que data, parfois des petits éléments anodins échappant à la collecte peuvent tout chambouler, telle l'explication imaginée de la théorie du chaos qui veut qu'un papillon battant des ailes à Tokyo entraîne un orage à New York. Les oracles du big data ne comptent pas se laisser décourager. Pour eux, la théorie du chaos n'est qu'une donnée à rajouter dans leur algorithme.

♦ FLORENT DELIGIA

* Lorsqu'une personne est accusée à tort, à cause d'un comportement pouvant prêter à confusion. Exemple : une personne qui rechercherait des informations sur la sécurité d'une centrale nucléaire.

Tous fichés : est-il trop tard pour se protéger ?

Le fichage est une réalité, l'exploitation des données aussi. Impossible de faire machine arrière face aux GAFAs, il est difficile de faire valoir un droit sur l'effacement de ses données. Stéphane Grumbach rappelle la situation : "La protection des données tourne à vide, car les data dépendent des lois américaines – car Google, Facebook et les autres sont américains." Pour Valérie Peugeot, les textes juridiques français se basent sur des lois "anté-numériques" : "La valeur d'usage de ces données est détenue par ces services plateformes et non par l'utilisateur qui les produit." Elle préconise une solution simple : "Remettre le citoyen au cœur des big data" et, si l'on souhaite ne pas tout laisser aux Américains, "il faudra proposer des entreprises alternatives européennes fortes. Pour la loi, un règlement européen est en cours de réalisation et va renforcer la protection des données personnelles". Si cela ne suffit pas, Valérie Peugeot imagine déjà la création d'un "droit collectif, avec la possibilité d'actions de groupe en matière de données". Des propositions qui ne vont pas dans le sens du lobbying des géants, qui militent de leur côté pour le droit de changer d'identité numérique. En 2010, le PDG de Google, Eric Schmidt, imaginait déjà un futur où les jeunes auraient le droit de changer de nom à l'âge adulte, pouvant ainsi repartir sur de nouvelles bases et oublier les erreurs du passé postées sur les réseaux sociaux. Au-delà de cette utopie, surtout marketing, grâce au big data et au recoupement, même changer d'identité ne servira à rien. Sur Internet, les géants savent qui vous étiez hier, qui vous êtes aujourd'hui, mais aussi qui vous serez demain.

Les usages du "big data"

Nous sommes déjà consommateurs d'usages rendus possibles par le big data. Petit tour du monde d'une révolution silencieuse.

◆ Prédire le trafic une heure à l'avance

La ville de Lyon n'est pas en reste avec ce fantasme de prédire le futur. L'application Optymod, développée par Cityway, peut donner le trafic jusqu'à une heure à l'avance et conseiller l'utilisateur pour qu'il choisisse le meilleur mode de transport en fonction de la situation actuelle et à venir. Petite révolution technologique, Optymod échoue sur un point : l'application n'est pas encore assez intuitive et surtout ne peut pas être utilisée dans le métro lyonnais, où il n'y a pas à ce jour de réseau mobile.

◆ Google Now : charmant, mais intrusif

"Bouchon sur la route, pour ne pas être en retard vous devez partir maintenant." Vous n'avez pas dit à l'assistant personnel Google Now où vous habitez, ni où vous travaillez, pourtant, grâce à votre mobile, la firme le sait déjà. Votre GPS intégré a transmis vos déplacements avec la plus grande des précisions. Google peut ainsi recouper vos données de géolocalisation avec votre agenda et vous conseiller de partir plus tôt pour être à l'heure, ou vous rappeler que la route que vous prenez tous les jours est en travaux.

◆ Quand le big data n'est peut-être pas neutre

Quand Amazon vous propose des livres ou objets qui "pourraient vous intéresser", vous êtes en droit de penser qu'il s'agit d'un bel exemple de big data. Grâce à vos précédents achats, Amazon est effectivement capable de dresser un portrait des goûts de chacun pour toujours mieux le "conseiller". Impossible d'évaluer la neutralité de ces informations. Amazon est en droit de mettre en avant des produits susceptibles de plaire à ses clients, mais surtout qui doivent être vendus en priorité. Fin 2014, l'entreprise a déposé un brevet pour "anticiper la livraison des colis" : concrètement, il s'agit d'envoyer des produits avant même qu'ils aient été commandés. En cas d'erreur, il suffira de refuser le colis. Amazon envisage même d'offrir certains produits aux clients mécontents d'avoir reçu quelque chose dont ils ne voulaient pas.

◆ Une accélération grâce aux objets connectés

D'ici à 2020, il devrait y avoir 80 milliards d'objets connectés dans les maisons du monde entier, dont 5 milliards rien qu'en France. Ampoule, thermostat intelligent, table de chevet qui recharge le téléphone sans fil, avec sa lampe capable de s'éclairer en fonction des cycles du sommeil, seront autant d'objets pouvant amasser de nouvelles données. Le big data pourra aller plus loin dans notre intimité et les géants nous vendront des usages qui feront passer la pilule : économie d'énergie, meilleure santé, efficacité dans l'organisation de nos journées, jusqu'au fantasme de science-fiction de l'assistant de maison, majordome 2.0 dirigeant tout grâce à des automatismes et les ordres vocaux de ses maîtres. Tel pourrait être

notre quotidien en 2020. Un nouvel enjeu sera alors au cœur des débats : comment protéger ces données, notamment celles relatives à la santé, et faire en sorte qu'elles ne se retrouvent pas en libre circulation sur la Toile ? Allons-nous voir l'apparition de cyber-rançons, qu'il faudra payer pour récupérer ses données et éviter qu'elles ne soient diffusées, comme c'est déjà le cas aujourd'hui avec des vidéos coquines volées sur des smartphones ? Et comment empêcher les assurances, mutuelles et autres organismes de crédit d'accéder à ces données et d'éventuellement refuser un contrat sous prétexte que votre patrimoine génétique vous prédispose à tel type de cancer ? Les usages seront infinis, tout comme les futures polémiques.

◆ FLORENT DELIGIA

