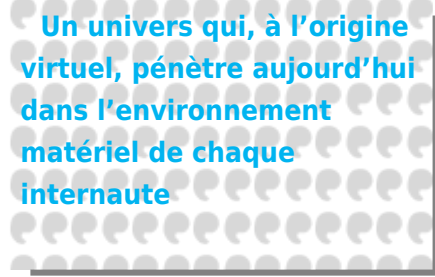


Du *medium* au supercalculateur
Une histoire du *big data* : l'origine
Le *big data* à l'heure sociale
Le *big data* immersif

Entré très récemment dans le discours des décideurs du numérique, le *big data* naît à la fin des années 1990, au moment même où le web, en se déployant, allait dépasser les capacités humaines de traitement de l'information. Depuis cette date, le traitement de l'information ajoute couche de données sur couche de données, et multiplie les relations entre objets connectés. Il reconstitue ainsi par des programmes algorithmiques un univers qui, à l'origine virtuel, pénètre aujourd'hui dans l'environnement matériel de chaque internaute.

Du *medium* au supercalculateur

S'il est difficile, voire impossible, de donner une définition précise des médias, au moins le terme indique-t-il par son étymologie (le pluriel de *medium* en latin, l'intermédiaire, ce qui se trouve au milieu) que la fonction du média est d'abord d'organiser la circulation d'un message entre un émetteur et un récepteur. Et cet émetteur, par le prisme du média, façonne inmanquablement le message, participant *de facto* à l'intermédiation entre, d'un côté, la nouvelle ou le contenu de divertissement qui est transmis et, de l'autre, celui qui va s'y exposer.



Un univers qui, à l'origine virtuel, pénètre aujourd'hui dans l'environnement matériel de chaque internaute

Cette intermédiation, opérée par le média, quel qu'il soit, a d'ailleurs conduit les chercheurs et les essayistes des années 1950 et 1960 à s'interroger, voire à s'inquiéter, sur le pouvoir des médias à l'heure de leur massification, ou plus précisément sur le pouvoir de ceux qui contrôlent les médias et élaborent les messages qu'ils relayent. Avec les *mass media*, s'imposer comme média, prendre à sa charge l'intermédiation, pouvait sembler conduire rapidement à un scénario totalitaire, parce qu'il n'y a qu'un pas entre la libre circulation des idées et la propagande, que certains n'hésitent pas à franchir, renonçant à leurs responsabilités, celles qu'ont inmanquablement les médias et ceux qui les font (les journalistes par exemple ont des chartes de déontologie).

En s'intéressant aux médias comme prolongements technologiques de l'homme, Marshall McLuhan va, dans les années 1970, redonner au terme sa pertinence, reléguant au second plan l'expression *mass media* jusqu'ici dominante (voir Francis Balle, *Médias et sociétés*, 16^e édition, LGDJ, Paris, 2013, p. 10). En tant que techniques, les médias constituent une intermédiation entre l'homme et son environnement, faisant des médias une institution essentielle dans le fonctionnement de la société. Cette institution a été historiquement incarnée par tous ceux qui se sont retrouvés dans une situation d'intermédiation, au sens large du mot « média », tel que l'entend McLuhan (l'accès à la parole, l'accès à l'imprimé, l'accès aux moyens modernes de diffusion...), qu'il s'agisse donc de transmettre des idées ou des convictions (politique, essayiste), une œuvre (artiste, écrivain), des savoirs (enseignement), du divertissement (animateur) ou de l'information (journalisme). Mais en même temps, tous ces intermédiaires sont tenus par le support de communication qu'ils utilisent, obligés en fait de se plier à ses règles (on ne fait pas un journal télévisé comme on fait un quotidien), ces règles qu'impose la technique ou le *medium*, le terme latin étant utilisé tel quel par Marshall McLuhan, et non celles qu'une corporation ou une communauté peut se donner par surcroît.

De ce point de vue, les techniques liées au développement du numérique, et surtout quand ces techniques s'insèrent dans un univers connecté, redonnent à l'interrogation sur le *medium* une plus grande urgence, car force est de constater que les médias historiques cèdent du terrain face à des *media* (*medium* au pluriel) d'un nouveau genre. De manière très schématique, qu'est-ce que le numérique ? Un langage universel pour coder tous les messages sous formes de *bits*. Qu'est-ce qu'un univers connecté ? Un système de délocalisation des données permettant d'identifier et de gérer à distance des demandes, dès lors qu'elles passent principalement par l'internet, lequel permet leur acheminement dans un langage numérique. Comment sont produites ces demandes ? Par des prolongements technologiques de l'homme, en fait des ordinateurs plus ou moins miniaturisés et reliés à l'internet : le PC du foyer, le *notebook*, et surtout le smartphone, la tablette, et déjà certains de nos objets connectés, « notre » montre Galaxy Gear, « notre » voiture, « notre » pèse-personne, l'intérêt des objets connectés étant d'être reliés à l'internet et surtout à un utilisateur identifié.

La question se pose donc de la nature de l'intermédiation quand cette dernière est progressivement prise en charge par une forme d'intelligence artificielle, à l'origine toujours programmée, il est vrai, par un cerveau humain. Si le *medium* impose ses règles, quelles sont donc celles de l'ordinateur, lieu de cette intelligence artificielle qui essaime dans notre environnement, un smartphone, un pèse-personne connecté n'étant que des ordinateurs « cachés » dans des objets destinés à des usages différents de ceux du PC ? La plus évidente de ces règles, sur laquelle nous nous attarderons ici, est sans aucun doute la capacité de

Les nouveaux media sont donc les éditeurs de logiciels dans les univers connectés

calcul de l'ordinateur, qui dépasse de très loin toutes les possibilités du cerveau humain. D'ailleurs, dans la course à la puissance des ordinateurs, le terme supercalculateur l'emporte désormais sur celui de superordinateur, montrant bien ce qui fait la spécificité de tout ordinateur (en 2013, le plus puissant des supercalculateurs est chinois : baptisé Tianhe en chinois et Milky Way-2 pour le reste du monde, il permet d'effectuer 33,86 mille milliards d'opérations par seconde, soit 33,86 PetaFlops/s). Cette course à la puissance de calcul rend possible l'exploitation du **big data**, c'est-à-dire l'exploitation de ces gigantesques masses de données produites chaque jour par les individus et les objets connectés, donc par les environnements ou écosystèmes sociotechniques.

S'il faut encore ici identifier une intermédiation, elle semble se loger moins dans l'information collectée automatiquement que dans la manière de la collecter ou de la traiter, laquelle déterminera ensuite les possibilités de son exploitation, le potentiel d'une base de données étant lié en grande partie à l'architecture de la base, à la manière dont l'information a été classée et recensée. De ce point de vue, le *medium* pour un ordinateur connecté, c'est le *software*, le logiciel, et non le *hardware*, à savoir le processeur. Les nouveaux media sont donc les éditeurs de logiciels dans les univers connectés, ceux qui éditent les systèmes d'exploitation, les services en ligne et les applications et, surtout, ceux qui collectent et organisent les informations qui transitent par l'intermédiaire de ces logiciels.

Une histoire du **big data** : l'origine

L'histoire du **big data** est l'histoire de la collecte, – dans un univers numérique en ligne (à distance) et par l'intermédiaire de logiciels (par des services) – des données de ceux qui se connectent, internautes ou objets. Cette histoire se constitue donc par étapes, à mesure que se déploie la connectivité. Mais l'histoire du **big data** ne saurait commencer *stricto sensu* avec la récolte des données de connexion. Elle commence véritablement quand cette récolte se double d'un traitement de l'information permettant d'adapter la réponse à une demande en ligne émise par un internaute ou l'un de ses objets. De ce point de vue, le premier « clic », celui qui va historiquement connecter les tout premiers internautes, celui qu'il a fallu faire au milieu des années 1990 sur l'icône d'un navigateur (Mosaïc, Netscape), n'a fait qu'amorcer un processus qui allait ensuite se poursuivre en s'amplifiant.

La logique du **big data** naît véritablement avec Google Search. Certes, il existait des moyens de récolter et de traiter l'information avant que ne soit lancé, en 1998, le moteur de recherche. Les navigateurs permettent en effet d'enregistrer tous les sites visités par les internautes et le temps qu'ils y passent. Mais avant l'émergence de la recherche, ces informations ne donnaient pas lieu à une proposition adressée à l'utilisateur, donc à un traitement de l'information débouchant sur une recommandation.

Il y avait également des activités automatisées de recherche en ligne avant le lancement de Google Search, les premiers moteurs de recherche textuels (analyse des mots dans les pages web) s'appelant Lycos, Excite (1995), ou encore AltaVista (1996). Ces moteurs sont dans une logique de recommandation puisqu'ils apportent une réponse à une requête de l'internaute, mais il leur manque un élément essentiel, la clé qui va permettre le déploiement du [big data](#) : au-delà de l'indexation du web, être en mesure d'identifier la popularité d'un contenu par un traitement automatisé de l'information. C'est Google qui opère cette révolution (voir sur ce sujet la synthèse de Philippe Torres et Mathieu Soule dans [Big Data, Big Culture, étude de l'Atelier BNP-Paris pour le Forum d'Avignon, 2013](#)).

Identifier la popularité d'un contenu par un traitement automatisé de l'information

La particularité de Google Search est de proposer une organisation des données disponibles sur les sites web qui s'inscrit d'emblée dans une logique propre au [big data](#). Alors qu'en 1998, Yahoo! domine le marché de la recherche en proposant un annuaire qui référence uniquement les adresses de sites web, Google inverse l'approche. Il ne s'agit plus de référencer les sites web, en présupposant la qualité de leurs contenus, mais bien de référencer le contenu de chaque page web, ce qui démultiplie subitement le nombre de références à recenser. En conséquence, le référencement selon Google s'éloigne alors de la logique de marque qui est celle des sites web, notamment des portails, lesquels s'appuient sur une identité forte et proposent, comme un titre de presse, un « contrat de lecture » à l'internaute. Ce sont dès lors les contenus qui doivent faire l'objet d'une valorisation, d'un traitement algorithmique capable d'en apprécier l'intérêt pour l'internaute. Une fois ces contenus indexés, il faut donc les noter : c'est à cet endroit que Google inventera le [big data](#), l'algorithme du moteur de recherche estimant la valeur d'un contenu en fonction des liens hypertextes qui pointent vers ce dernier et du taux de satisfaction des internautes, exprimé dans le pourcentage de clics générés sur les liens affichés pour un mot-clé donné. Ce calcul dynamique de la pertinence d'un contenu suppose en effet d'avoir collecté l'ensemble des contenus du web, au moins leur grande majorité, et de calculer en permanence les rapports qu'ils entretiennent entre eux ainsi que leurs scores respectifs, en termes de clics, dans les réponses faites aux requêtes des internautes.

En identifiant ensuite l'internaute en fonction de ses requêtes, une couche supplémentaire d'[intelligence artificielle](#) peut être ajoutée qui, en plus de la popularité d'un contenu sur le web et de sa popularité statistique auprès des internautes en général, peut affiner la réponse apportée en fonction du profil de l'utilisateur. Sont donc superposés un index des contenus, un réseau de liens des contenus entre eux, un score sur la pertinence de chaque contenu eu égard à une requête et, enfin, des profils d'internautes constitués par l'analyse de leurs requêtes successives. C'est en croisant ces informations, grâce à des algorithmes, que le moteur de recherche propose ensuite des résultats, faisant ainsi office de véritable *medium*, d'intermédiaire, en listant des recommandations à l'utilisateur. Il s'agit ici du référencement

naturel, un système d'enchères en temps réel étant ajouté pour le positionnement des liens sponsorisés.

Le référencement du web par Google Search ne correspond pourtant qu'à la première grande vague de collectes de données qui a conduit à l'émergence du **big data**. En effet, l'index de Google Search est d'abord un index textuel, qui référence tous les mots, potentiellement, de toutes les pages du web. L'index s'est ensuite étendu à d'autres contenus (vidéos avec YouTube, cartographie avec Maps, livres avec Google Books, etc.). Mais la valeur de cet index, on l'a vu, repose essentiellement sur l'identification des liens entre les contenus, ainsi qu'entre les contenus et les utilisateurs. Or les liens entre contenus et utilisateurs se multiplient très rapidement à mesure que le temps de connexion augmente (plus je suis sur internet, plus je laisse de traces). Deux grandes tendances caractérisent donc l'évolution du traitement des données à des fins de recommandation, qui correspondent à un deuxième et troisième temps de l'histoire du **big data** : d'abord la collecte des données dites sociales ; ensuite la collecte des données communiquées par les objets connectés.

les liens entre contenus et utilisateurs se multiplient très rapidement à mesure que le temps de connexion augmente

Le **big data** à l'heure sociale

Après Google, l'histoire du **big data** a connu une deuxième révolution grâce à Facebook, le service étant lancé sur le web en 2004 et accessible à tout internaute, et non plus aux seuls étudiants ; dès 2006 la seule condition étant que les internautes s'identifient au préalable grâce à une adresse électronique valide. Cette exigence – le fait de devoir s'inscrire sur Facebook – va révolutionner l'approche du web en la personnalisant, le fondateur de Facebook, Mark Zuckerberg, préférant parler de « socialisation » du web. La logique du réseau social suppose en effet que ses membres soient identifiés et qu'ils communiquent des informations personnelles pour qu'une conversation s'engage en fonction d'affinités partagées. De ce point de vue, Facebook opère ici une première grande rupture par rapport au modèle du web selon Google. On doit en effet à Google d'avoir lancé la première grande vague de récolte de données personnelles afin d'optimiser son service de recherche : pour que le moteur vous réponde, vous devez renseigner une requête, donc indiquer à Google votre centre d'intérêt sur le moment, à court terme.

Avec Facebook, une deuxième grande vague de collecte de données personnelles commence : pour profiter d'un réseau social, il faut que celui-ci connaisse vos centres d'intérêt, vos goûts sur le long terme et dispose d'un moyen de vous contacter. C'est donc sous la forme de

Ce graphe associé à chaque objet du web est une véritable révolution

profils que vont être engrangées les données personnelles, l'enjeu étant de disposer de profils véritables et non de fausses identités. Une fois les profils engrangés, il faut les fidéliser par une augmentation de la durée passée à utiliser le service. La logique de collecte des données personnelles de Facebook est en effet complètement différente de Google. Google s'appuie sur des requêtes, donc des données contextuelles, avant de rediriger l'internaute vers des sites tiers. Facebook s'appuie sur la « profondeur » d'un individu : la gestion du temps court, contextuel, lui échappe en partie. Pour connaître cette profondeur, les traits intimes de la personnalité d'un individu, Facebook doit s'assurer que l'utilisateur passe du temps sur son service, c'est-à-dire entre dans « une logique d'engagement ».

dans la manière de concevoir le réseau

C'est au nom de cette logique d'engagement que Facebook va être amené à ajouter une couche supplémentaire d'information au cœur de son réseau social. En plus de l'information relative à ses utilisateurs et de l'actualité des amis de chaque utilisateur, il lui faudra disposer également de l'information concernant les « objets » dont parlent ses utilisateurs. Ces objets sont, dans les faits, les contenus du web, notamment l'actualité et les contenus de divertissement, ceux qui meublent les sujets de conversation et permettent d'échanger facilement autour de centres d'intérêt communs. A l'inverse des actions de chaque utilisateur intéressant un cercle d'amis très restreint, ces objets bénéficient d'un renouvellement permanent qui garantit une attention de l'utilisateur, donc son engagement : l'actualité peut en effet générer une forme d'addiction, et les industries culturelles tentent de créer cette addiction en multipliant les propositions de divertissement. De ce point de vue, Facebook est moins un service de communications privées, interpersonnelles, qu'un service de discussions publiques sur les marques, l'actualité et les contenus des médias. En identifiant nos goûts personnels, les objets dont nous parlons, ce qui peut être qualifié de « données personnelles culturelles » (voir sur ce point : forum-avignon.org), Facebook apprend par ailleurs à mieux nous connaître. Mais il a dû pour cela intégrer dans sa logique le moyen de référencer les objets constituant des sujets de conversations.

Cette intégration s'est faite en deux temps, le premier étant celui qui a permis à Facebook d'accueillir au sein du réseau social les contenus par ailleurs disponibles sur le web. Pour y parvenir, Facebook a proposé aux éditeurs, dès 2007, de créer leur propre application au sein du réseau. Mais tant que l'accès aux contenus passait par des applications développées pour Facebook, le réseau social est resté une alternative aux autres plates-formes intégrées, qu'il s'agisse à l'époque d'Apple avec iTunes ou de MySpace avec ses pages de musiciens... Il fallait donc que Facebook arrive à exister en dehors des seules activités du réseau social, à essaimer par conséquent à travers le web, sans avoir à rapatrier tous les contenus au cœur du réseau social. Facebook s'est donc retrouvé à cet instant face à Google, qu'il sait être son premier concurrent. En effet, Google sait essaimer à travers le web sans avoir à l'intégrer dans ses services : avec son moteur, il référence les sites et devient leur « fournisseur d'audience », sans pour

autant demander quoi que ce soit aux sites, et notamment sans leur demander la moindre autorisation. De ce point de vue, Facebook s'est inspiré du référencement de Google pour proposer, lui aussi, une nouvelle manière de référencer le web. Qui dit référencer le web dit passer en revue ce qui existe sur la Toile, donc maîtriser l'offre de contenus et services, même quand celle-ci n'est pas intégrée au cœur du réseau social. Ce pas a été franchi une première fois avec le lancement, en 2008, de Facebook Connect. Cette fonction permet d'identifier un utilisateur de Facebook, y compris quand il se rend sur un site tiers, dès lors que celui-ci est partenaire du réseau social. Un échange d'informations entre le site et Facebook est à ce moment mis en place, le site partenaire récupérant des informations sociales sur ses utilisateurs, Facebook publiant sur le profil de ses utilisateurs les informations consultées ou commentées sur le site tiers.

Dès lors, l'ensemble des liens sociaux traités par Facebook, qualifié de *Social Graph* (le réseau d'amis) se double potentiellement d'un graphe d'objets (les contenus visités et commentés par l'internaute et ses amis). Cette évolution majeure de Facebook est ce qui va conduire de Facebook Connect à l'*Open Graph* en 2010. Concrètement, *Open Graph* permet d'associer un bouton *Like* à chaque objet du web, multipliant *de facto* les interactions entre utilisateurs de Facebook qui *likent* des objets et les milliards d'objets répartis sur les milliards de pages du web. Ainsi, chaque objet *liké* va se retrouver dans un réseau d'utilisateurs l'ayant recommandé et dans un réseau d'objets en lien avec ses utilisateurs et leurs préférences. A partir de 2011, Facebook multipliera les boutons (lire, regarder, écouter), afin de mieux cibler l'information attachée par les utilisateurs à chaque objet du web.

Identifier l'individu
derrière une multitude de
terminaux utilisés

Ce graphe associé à chaque objet du web est une véritable révolution dans la manière de concevoir le réseau : il remplace le réseau des liens hypertextes, celui exploité par Google, par un réseau de liens sociaux, des « objets » pouvant être socialement reliés entre eux par les recommandations sans l'être par des liens hypertextes. Avec le bouton *Like*, Facebook se met donc en situation de proposer un référencement social du web, une alternative à Google en matière de référencement des contenus disponibles en ligne. La recommandation adressée à chaque utilisateur implique ainsi le traitement simultané des graphes sociaux de l'utilisateur et de ses amis, par extension des amis de ses amis, ainsi que le traitement des graphes d'objets concernés par les discussions de l'utilisateur ou de ses amis, et enfin le croisement de ces graphes entre eux afin de dénicher, parmi la masse des objets (contenus ou utilisateurs) identifiés par Facebook, celui pour lequel le degré d'affinité risque d'être le plus élevé.

Le *big data* immersif

En indexant les contenus du web et en établissant la cartographie dynamique de leurs interactions, Google a lancé la première grande vague de collecte de données personnelles, dès lors que cette cartographie dynamique est mise au service d'un processus de recommandation en réponse à la requête d'un internaute. Mais l'internaute ne révélait directement que ses centres d'intérêt contextuels (ce que je cherche en ce moment) dont il fallait déduire ses centres d'intérêt en général (ce que j'aime en général). Avec Facebook, la « socialisation » du web entraîne une plongée dans les goûts de l'individu, et notamment dans ses goûts culturels (les boutons « regarder », « lire », « écouter »). L'individu sera ainsi d'autant mieux connu qu'il est possible de croiser les sources d'information, par exemple croiser la cartographie dynamique du web selon Google, et la cartographie sociale du web selon Facebook. Il s'agit bien sûr ici d'une hypothèse, ces deux concurrents ne risquant pas de s'échanger leurs précieuses données. Mais l'on comprend avec cet exemple que la multiplication des terminaux connectés, smartphones, tablettes, téléviseurs, outre les ordinateurs, menace paradoxalement l'établissement de bases de données cohérentes et exploitables dans une logique [big data](#).

Si le même individu apparaît de manières différentes à deux endroits de la base de données, par exemple avec l'adresse IP du foyer et celle de son smartphone, et s'il est impossible de superposer ces deux adresses en géolocalisant les lieux de connexion récurrents, l'individu sera mal connu et ses interactions en ligne mal identifiées. L'enjeu est donc, à mesure que se développe le multi-équipement, de parvenir à rapatrier vers un individu identifié une succession de connexions réparties depuis différents terminaux. Facebook n'a pas ce problème dès lors que l'utilisateur accepte de se connecter en étant identifié avec son compte. Il perd en revanche son utilisateur quand celui-ci navigue sur des sites ou applications sans s'être au préalable identifié. A l'inverse, Google a eu ce problème : une recherche depuis deux adresse IP, et c'est la compréhension fine des désirs de l'internaute qui est tronquée. Pour y remédier, Google est entré dans une logique d'individualisation de la connexion, s'inspirant ici de Facebook, mais à une bien plus grande échelle. En favorisant les identifiants Gmail grâce à Android (environ 80 % des smartphones vendus dans le monde), en imposant un identifiant unique pour l'ensemble de ses services ([voir REM n°22-23 p.55](#)), en contrôlant aussi le logiciel qui fournit le plus grand nombre de données de navigation, à savoir le navigateur avec Chrome (Chrome est devenu le premier navigateur au monde en 2013, tous terminaux connectés confondus), Google est aujourd'hui capable de suivre un internaute lors de ses différentes connexions depuis ses différents terminaux, quand bien même les adresses IP changeraient.

C'est d'ailleurs cette capacité à identifier l'individu derrière une multitude de terminaux utilisés qui sera, demain, l'enjeu des régies publicitaires en ligne. Ainsi, la société française Criteo, introduite au Nasdaq le 30 octobre 2013, a connu une croissance très forte grâce à son offre de *retargeting*, laquelle ne doit sa pertinence qu'à la connaissance fine des internautes dans un univers fixe, grâce notamment aux cookies envoyés

C'est aussi l'environnement physique de l'internaute que les géants de la donnée et du [big data](#) commencent à cartographier

sur les ordinateurs. Mais avec l'évolution des usages vers plus de mobilité, Criteo devra, pour résister en Bourse, étendre impérativement à la mobilité sa capacité à identifier les internautes, et croiser les informations reçues des connexions fixes et mobiles.

En identifiant les terminaux par lesquels se connecte l'internaute, c'est aussi l'environnement physique de l'internaute que les géants de la donnée et du [big data](#) commencent à cartographier. Avec ces terminaux connectés, l'individu voit son profil enrichi de données de géolocalisation, d'itinéraires, d'un agenda, de la reconnaissance des mouvements et de la voix. La connectivité pénètre alors progressivement l'environnement physique. Orange est ainsi parvenu à mesurer les flux touristiques dans les Bouches-du-Rhône durant l'année 2013 grâce à un système baptisé Flux Vision, en suivant près de deux millions de clients mobiles à l'intérieur du département. L'entreprise a pu établir la première cartographie précise des flux touristiques sur tout le territoire, un résultat que seul le traitement de données à grande échelle rend possible. De la même manière, Orange a pu déterminer assez précisément le nombre de personnes présentes sur le Vieux-Port pour l'inauguration de Marseille Provence 2013, relativisant d'ailleurs les chiffres des organisateurs. Il lui a fallu pour cela faire plus que comptabiliser le nombre de mobiles Orange présents sur la zone. Cité par *Les Echos*, Jean-Luc Chazarain, représentant d'Orange explique la manière d'aboutir à ce résultat : « *Une batterie d'algorithmes recoupe ces données [de connexion au Vieux-Port] qui sont toutes confidentielles, analyse leur itinérance, élimine les redondances et extrapole le résultat en fonction de nos parts de marché dans les zones correspondantes* ». De la même manière, le traitement à grande échelle des données s'invite progressivement dans la veille sanitaire. En 2008, Google a ainsi lancé Google Flu Trends, un service qui permet de suivre l'évolution d'une épidémie de grippe en suivant l'apparition géolocalisée des requêtes liées à la maladie. En France, Celtipharm produit une carte dynamique des épidémies en croisant les achats réalisés dans un réseau de 4 600 pharmacies avec les prescriptions médicales les plus courantes, les achats de médicaments traduisant en creux le nom de la maladie qu'ils sont censés combattre !

En multipliant également les objets connectés (voiture connectée, montre, abribus, etc.), c'est l'environnement extérieur dans son intégralité qui, progressivement, va entrer dans l'univers de l'internaute et qui pourra s'y adapter, le [big data](#) permettant ici une inversion complète des perspectives. Alors que l'individu devait s'adapter aux contraintes physiques de son environnement, on pourra demander à ce dernier de se modifier à souhait, d'entrer donc dans une logique immersive, pour se présenter au mieux en fonction des attentes de chaque individu connecté qui y évolue. C'est donc notre vie quotidienne, dans sa matérialité, qui pourrait faire l'objet de recommandations permanentes, d'une gestion par le [big data](#), dès lors que l'environnement aura été connecté.