

***D'un  
monde  
de la  
donnée  
à l'autre***

***entre appropriation  
et dépossession***

# Sommaire

*Sommaire*

Introduction	04
Informations et pédagogie <i>Le travail des algorithmes</i>	06
Métiers à venir <i>Nouveaux profils</i> <i>Postes publics</i>	08
Tiers de confiance <i>Portabilité de la réputation</i> <i>Nouvelles certifications</i>	12
Frictions socio-techniques <i>Obsolescence des choix techniques</i> <i>Le marché gris du profilage</i>	16
Nouvelle forme de protestations <i>Dévoisement des données</i> <i>Luddisme de la donnée</i> <i>La valeur économique des secrets</i> <i>Modélisation et données</i>	20
Données et collaborations <i>Collaboration à bénéfices mutuels</i> <i>Complétion des données publiques</i>	28
Enjeux et implications	32
Lexique	34

# Introduction

## *Gouvernance de la donnée: scénarios prospectifs*

La mise en place des plateformes *Open Data* de multiples collectivités vise à répondre aux enjeux de la ville intelligente en donnant un large accès aux données publiques afin de favoriser la création de services innovants et la participation citoyenne. Or, après une période prometteuse de démarrage de ces initiatives, il convient d'interroger tant les évolutions possibles de la plateforme en fonction de paramètres contextuels (évolution des usages, environnement juridique, nouvelles formes d'expertise), mais aussi de nouvelles opportunités. En d'autres termes, il apparaît nécessaire pour les collectivités de réaliser une prospective de la place de la donnée publique afin de faire évoluer la gouvernance et la plateforme *Open Data* de demain, sur les manières d'aborder les données à moyen/long terme, et sur leurs rôles dans la contribution aux politiques publiques avec des données.

Ce livret propose une série de scénarios prospectifs concernant les enjeux et les opportunités de la gouvernance des données de la Métropole. Ces voies possibles soulignent le renouvellement des défis actuel et à venir. Leur rôle consiste davantage à interpeller, en particulier en accentuant les risques et dérives, pour favoriser une prise de conscience et l'anticipation de réactions qui ne doivent pas être exclusivement techniques. Plus spécifiquement, ils ont pour objectif de :

- Décrire l'incarnation des politiques des données, en particulier en montrant les conséquences – positives ou négatives – afin de susciter un débat sur les limites et les opportunités.
- Aborder le rôle d'une collectivité, et sur les façons de mener une politique qui peut créer du rejet ou de la défiance chez les citoyens, voire une crainte de l'accaparement par les entreprises.
- Croiser le thème de la réutilisation externe des données publiques, avec celle de l'utilisation interne (services de la collectivité), en faisant un lien avec les données personnelles des citoyens.
- Marquer ce que ces changements représentent pour un acteur public comme la métropole de Lyon, avec ses compétences.

## Méthodologie

La constitution de ces scénarios a reposé sur la réalisation de deux études conjointes (une exploration des imaginaires de la donnée publique par *Pop Up Urbain*, et un benchmark international mené par Antoine Courmont). Ces deux travaux ont servi de points d'entrée à la réalisation d'un atelier mené par *The Near Future Laboratory/Pop-Up Urbain* en février 2017 qui a fait ressortir une cinquantaine de pistes de scénarios probables. Lesquels ont été filtrés et présentés lors d'un atelier mené au Grand Lyon fin juin 2017, qui a permis également d'enrichir et de nuancer 14 scénarios principaux.

Chacun de ces scénarios, qui reflètent différents axes d'évolution envisageables, est décrit dans ce livret sous une forme textuelle accompagnée d'illustrations matérialisant les changements.<sup>1</sup>

## Usage des scénarios

Notons que si ces scénarios sont des pistes possibles et probables, leur présence dans ce livret ne signifie nullement que celles-ci sont souhaitables pour les collectivités publiques, leurs habitants et leurs entreprises. Outre la description de ces scénarios et de leurs implications présentées dans ce livret, leur objectif consiste à soutenir un débat plus large concernant le positionnement à adopter sur le thème de la gouvernance des données. Notre parti pris, dans la rédaction de ces scénarios, a consisté à décrire des changements insidieux légers et en surface mais problématique dans leurs implications générales; et cela, plutôt que de formuler des descriptions-chocs qui font réagir, mais qui ne témoignent pas des nuances et de la complexité de la situation. Il s'agit de modèles-types pointant des problématiques et des points de tension. Ces scénarios ont pour vocation d'être utilisés dans le cadre d'ateliers avec différentes parties prenantes du partage et de la réutilisation des données publiques, et de faire réagir et d'échanger sur ces sujets. Enfin, si l'ensemble développé ici peut parfois être perçu comme négatif, c'est aussi pour susciter des réactions et prendre la mesure des défis auxquels nous sommes confrontés.

1. Nous adoptons ici une démarche de design fiction, une approche prospective qui propose des scénarios sur le futur d'une technologie/situation donnée en utilisant les manières de faire du design. En particulier en créant des objets standards qui expriment ces futurs possibles (catalogue d'objets fictionnels, magazine fictif, etc.) et qui permettent d'en discuter ou d'intervenir pour en changer la trajectoire.

# Informations et pédagogie

## *Le travail des algorithmes*

Si le thème du *Big Data*, et plus largement des données, a pris une place prépondérante dans les médias, un autre « objet du numérique », a pris une importance tout aussi grande ces dernières années dans le débat public : l'algorithme. Comme l'a montré l'étude menée par *Pop-Up Urbain*,<sup>2</sup> cette notion est ainsi devenue le nouveau référentiel de la mise en données du monde. Ce qui n'est d'ailleurs pas sans conséquence puisqu'il s'agit d'un objet difficile à cerner, et qui renvoie actuellement à de multiples craintes sur leur fonctionnement, à un manque de transparence, et *in fine* à des imaginaires négatifs du contrôle. En outre, l'absence de discours public sur ces questions a deux conséquences. D'une part les représentations sont principalement construites par les industriels, et par les descriptions qu'en font les médias avec plus ou moins de précisions. D'autre part, cette absence de positionnement public clair renforce le sentiment d'une dépossession, puisque seules les sociétés technologiques semblent avoir prise sur ces questions.

*Le travail des algorithmes*

C'est la raison pour laquelle il est souhaitable que les structures publiques communiquent sur ce sujet afin de démystifier le thème, et dans le même temps, de montrer le rôle croissant des algorithmes dans différentes couches de la société (place jouée dans la production de services urbains, importance pour une collectivité publique, lien avec les données personnelles). Un tel projet pourrait par exemple prendre la forme d'un manuel pédagogique sur ce thème, comme représenté ci-contre. Lequel s'adresserait tout autant aux enfants qu'aux enseignants avec un double objectif de sensibilisation à ces questions, mais aussi à donner des rudiments de compréhension; et ce, en complément, mais pas forcément des initiatives actuelles d'apprentissage à la programmation.

2. *Mise en données du monde: imaginaires en équation*, 2017: <http://www.millenaire3.com/ressources/mise-en-donnees-du-monde-imaginaires-en-equation>



### Implications pour l'acteur public

Quelle que soit la forme que peut prendre ce type de communication pédagogique, il y a une nécessité pour la Métropole de contextualiser et d'expliquer la place des enjeux de données et d'algorithme à très court terme. Si de multiples actions positives sont menées, par exemple « des ateliers du TUBA (Tube à expérimentation à Lyon) » de même que ses activités d'accompagnement à la compréhension de l'exploitation des données, le grand public serait aussi un acteur pertinent à toucher sur ces sujets.<sup>3</sup> De plus, il apparaît également pertinent pour la collectivité de définir elle-même – et d'informer à ce propos – sa vision sur des questions d'ordre a priori techniques, mais qui ont une influence certaine sur la société. Si elle ne fait le pas, ce seront les sociétés technologiques qui le feront elles-mêmes, à la manière de Microsoft qui produit des livres pour enfants sur le matériel informatique.<sup>4</sup>

3. Notamment du fait de la Loi pour une République numérique qui instaure une « Obligation de communication de l'algorithme ayant participé à la prise d'une décision individuelle par une administration » et ouvre ainsi la voie à une communication « obligatoire » et donc vertueuse sur les algorithmes.

4. Comme avec le livre pour enfant 'Mommy, why is there a server in the house?' (Microsoft, 2008).

# Métiers à venir

## *Nouveaux profils*

Comme à chaque période de changement technologique, on constate un double mouvement d'apparition de nouveaux métiers en parallèle des craintes récentes et avérées de disparition de nombreuses places de travail. Si ces nouveaux postes ne seront pas nécessairement occupés par les personnes ayant perdu leur emploi du fait de ces changements, des profils nouveaux émergent sur le marché du travail. Force est de constater que cette évolution est déjà en cours. C'est ainsi que des métiers liés aux enjeux de gestion et maintenance des systèmes d'information (archiviste, *Data janitor*/ nettoyeur de bases de données), mais aussi d'algorithmes.

Un exemple actuel<sup>5</sup> concerne le besoin nouveau de pouvoir « entraîner » les programmes informatiques basés sur les technologies d'intelligence artificielle (*machine learning*). Cette notion d'entraînement correspond à la nécessité d'une forme de supervision leur permettant d'« apprendre » les résultats attendus par leurs programmeurs et leurs usagers. C'est par exemple le cas des personnes qui doivent alimenter des programmes en contenu afin que ceux-ci améliorent leur capacité de reconnaissances de formes, les recalibreurs d'objets connectés. Il s'agit aussi aujourd'hui d'inculquer à des objets domotiques, tels que le thermostat connecté nest, comment réguler efficacement la température domestique, ou de faire comprendre à des assistants vocaux (SIRI, Cortana) quels sujets aborder (ou éviter). Par extension, on peut donc s'attendre à ce qu'il faille faire de même pour les objets connectés ou les véhicules autonomes. Même si l'on peut imaginer des usagers avancés pour faire cela, il est fort probable que cette expertise sera prise en charge par des professionnels d'un nouveau genre. Dans le cas d'un système domotique, il est vraisemblable que le fait d'atteindre un confort thermique qui convienne aux habitants d'un même domicile, qui tienne compte des enjeux écologiques de thermorégulation, et qui évolue avec les saisons ou les activités de membres d'une même famille, va nécessiter une aide ciblée tant sur le plan technique sur les modes opératoires, avec cette logique d'entraîner le système à mieux s'adapter à un foyer. L'exemple est valable également avec le véhicule autonome qui devrait lui aussi être « entraîné » à se comporter de la manière souhaitée par son usager.

À côté de cette dimension infrastructurelle, il n'est pas interdit de penser que le secteur de la prestation de services (conseil, formation, sous-traitance) va intégrer de façon croissante des savoir-faire et des techniques liés à ce sujet. L'émergence de cabinets de conseil tournés vers le conseil sur les données – hébergement, maintenance, sécurité ou conformité (*compliance*) des données et contenus numériques – sont un autre exemple de compétences nouvelles qui va être amené à se développer plus largement. De même, les données publiques mises à disposition pourront être analysées par de tels experts dans des logiques d'optimisation en fonction de toutes sortes de risques calculables à partir de celles-ci (différentiel de taxes locales, risques environnementaux ou de pollution, etc.).

5. Voir par exemple cet article du *New York Times* du 28 avril 2017: <https://www.nytimes.com/2017/04/28/technology/meet-the-people-who-train-the-robots-to-do-their-own-jobs.html>



### Implications pour l'acteur public

Outre les problématiques de formation professionnelle et de réorientation posées par l'arrivée de ces nouveaux métiers (sur le territoire ou ailleurs), une telle situation appelle aussi une réflexion quant au rôle des collectivités dans la mise à disposition de données publiques pour faciliter, ou empêcher, les tâches effectuées par ceux-ci. S'il apparaît légitime de faciliter l'émergence de ces métiers en libérant des données qui les permettent, il s'agit aussi de réguler, voire d'interdire des usages potentiellement déviants et problématiques pour la collectivité. Afin de garantir un cadre de vie, un espace et des services publics bénéficiant à tous, il convient par exemple de surveiller l'éclosion de métiers qui, sur la base des données publiques, pourraient permettre de tricher, voire même de contourner des règles établies.

# Métiers à venir

## *Postes publics*

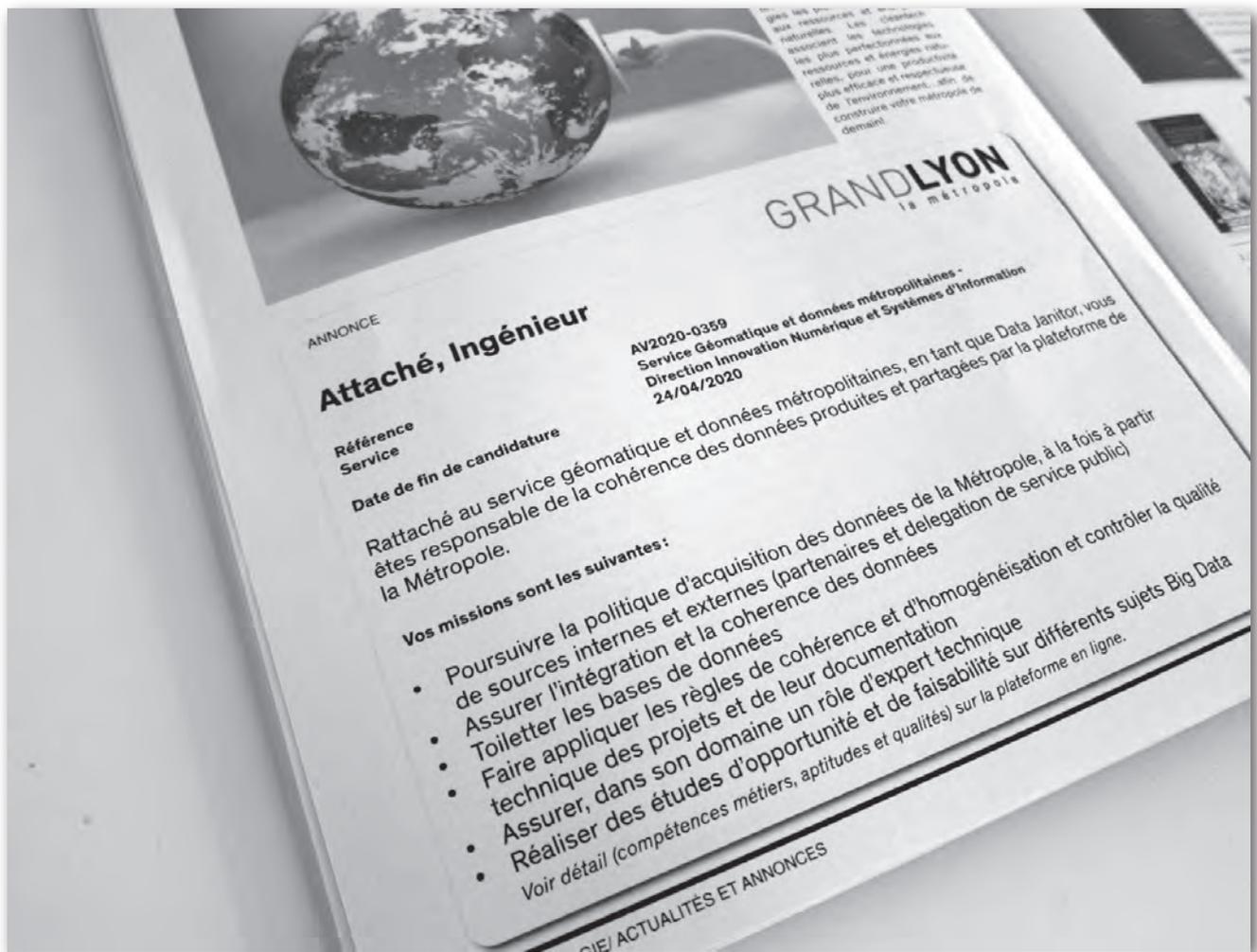
Si la plupart des nouveaux métiers liés à l'expertise concernant les données et les algorithmes risquent de concerner l'économie privée, l'emploi public risque aussi de connaître une évolution liée aux enjeux posés par l'usage croissant des méga-données, ouvertes ou non. Concernant la mise en place de systèmes d'information, il est vraisemblable que les missions de service public vont nécessiter des expertises qui ne seront pas forcément couvertes par des organisations privées. Ce qui peut mener à l'émergence de profils de poste nouveaux dans les collectivités publiques.

Citons en premier exemple la nécessité sans cesse renouvelée pour la puissance publique de jouer pleinement son rôle dans la gouvernance des données. Si l'on comprend aisément que celle-ci se joue en particulier sur un plan technique (mise à disposition de données, administration de celles-ci), il ne faut pas oublier la place fondamentale des différents échelons publics sur les questions de régulation, de certification (par exemple sur la base des données produites par les usagers), ou d'arbitrage (conflits avec les grandes plateformes, avec les prestataires).

Nous pouvons aussi nous attendre à voir évoluer les métiers de la statistique publique, ou de la gestion d'archives (gestion du cadastre, histoire du patrimoine) sur la base des technologies de collecte, d'enregistrement et d'archivage des données. De la même manière, les métiers gravitant autour de la sciences des données<sup>6</sup> sont déjà en train d'être intégrés aux compétences de la Métropole.

En outre, le nettoyage des données implique une quantité de travail colossale, et souvent invisible, qui serait assurée par des *Data janitor* en charge de la fiabilité des données. À cet égard, comme certaines municipalités commencent à l'envisager, il serait par exemple pertinent, à croiser les expertises de l'archivage, de la numérisation et de l'*Open Data* pour constituer des archives mises à jour de manière dynamique en temps-réel sur la base de différentes sources informationnelles (réseaux de capteurs, photographies et caméras, etc.). Le dialogue entre ces nouveaux profils et les services métiers « traditionnels » est également un défi, tant il pourrait reconfigurer les habitudes de travail des agents opérationnels.

6. La science des données (Data Science) est une discipline nouvelle qui s'appuie sur des outils mathématiques, de statistiques, et d'informatique pour établir des méthodes d'analyse de données de masse et de sources plus ou moins complexes, afin d'en extraire des informations pertinentes.



## Implications pour l'acteur public

Outre la question du financement de ces postes, ces constats interrogent sur la capacité des collectivités publiques à intégrer de façon proactive des profils pluridisciplinaires (expertise juridique et technique, voire géographique ou historique), à participer à la formation de ceux-ci, ainsi qu'à mettre en place une politique de recrutement qui va dans cette direction. Or, on constate depuis plusieurs années un phénomène problématique : il y a d'un côté dans la stratégie de la collectivité l'intérêt, et le soin, de constituer une expertise autour des métiers du numérique et des systèmes d'information. Par exemple en ayant une politique de recrutement à la mesure des ambitions sur ces questions. Mais d'un autre côté, l'explosion des besoins dans ces métiers, particulièrement autour des expertises concernant le *Big Data* (programmation, design de services, design d'interface), est une concurrence rude pour les acteurs publics ; menant à une difficulté à trouver, ou à retenir, des profils ayant ces compétences. Le décalage des moyens à disposition entre les grands groupes industriels et la collectivité, ou entre l'enthousiasme vécu dans l'aventure entrepreneuriale de certaines *start-ups* et une organisation publique, rend la situation difficile. Ce qui donne parfois l'impression d'une véritable fuite des cerveaux pour tout profil compétent dans ces domaines. Le problème est donc double puisqu'il se situe autant du côté de l'expertise concernant la mise en place, la diffusion et la maintenance des plateformes d'*Open Data* et des systèmes d'information en général, que dans le besoin pour la collectivité de garder une compétence interne d'analyse, de régulation et de négociation comme compétence interne.

# Tiers de confiance

## *Portabilité de la réputation*

Les données compilées et analysées du fait de l'économie des plateformes permettent aujourd'hui d'établir un ou plusieurs indicateurs de réputation pour les usagers. C'est ce dont on peut se rendre compte avec des services tels que Uber ou Airbnb; chacun gardant pour lui et la maîtrise, et l'usage des évaluations des utilisateurs du service ou des prestataires. Or, à ce stade les collectivités publiques en France n'ont aucune maîtrise de ce processus; ni dans leur régulation, ni dans leurs usages. Et cela, alors qu'il implique des considérations éthiques et stratégiques fondamentales, et que ces données privées pourraient aussi être combinées à celles générées par les services publics ou para-publics à propos des citoyens.

S'il apparaîtrait clair que ce type de possibilité ne peut naître que d'une forme de régulation particulièrement affirmée auprès des grands acteurs de l'économie des plateformes, le cas chinois montre que c'est une possibilité, et qu'elle interroge sur sa souhaitabilité. Le Sesame Credit Score (芝麻信用分) est éloquent pour questionner ce type de service. Géré par les sociétés privées Alibaba et Tencent, pour le compte du gouvernement, il s'agit d'un système de notations évaluant chaque citoyen en fonction de ses activités quotidiennes. Celui-ci repose sur la collecte des données des internautes chinois, des échanges sur les réseaux sociaux au respect du Code de la route. Tout élément pouvant décrire le comportement d'un citoyen est ainsi comptabilisé et analysé afin de leur donner une note évaluant à quel degré ceux-ci sont dignes de confiance; par exemple pour l'attribution d'un logement ou de crédits. La logique de centralisation ici à l'oeuvre interroge sur les conséquences sociales d'une telle situation.

Sans forcément singer le cas chinois, il est pertinent de s'interroger sur les manières dont la collectivité pourrait aborder ces questions. Sur la base d'expériences actuelles en France, en particulier le projet *Mes Infos* mené par la FING<sup>7</sup>, il est par exemple vraisemblable que des logiques de rapatriement, d'hébergement, voire de croisement avec les données publiques ouvertes soient mises en place. Mais ce type de démarche pourrait aller plus loin, avec par exemple une opportunité de portabilité des réputations qui permettrait de certifier pour un usager ses données et leur utilisation ou leur transfert. Un acteur public pourrait par conséquent devenir garant de ses citoyens pour avoir accès à tel ou tel service. De cette manière, la collectivité pourrait certifier les habitants dans leurs démarches auprès d'organisations publiques (déplacement dans un autre Etat), para-publics (services de santé), voire privées (compagnie d'assurances). Un tel cas pourrait en effet trouver sa pertinence dans le cadre de déplacements internationaux pour lesquels les procédures actuelles sont rigides, et pour lesquelles des formes de certification sont mises en place de manière croissante.<sup>8</sup>

7. <http://mesinfos.fing.org/lepilote/> (une initiative prometteuse dont la Métropole de Lyon est partenaire).

8. Pensons par exemple aux autorisations de voyage aux Etats-Unis (ESTA) ou au Canada (AVE) et qui intègrent progressivement des données concernant la présence sur les réseaux sociaux.



## Minimisez le coût de votre déménagement à Montréal

### Transfert total des données

**Adulte: 100 CAD**

**Famille: 175 CAD**

- ✓ e-Résidence franco-canadienne de 6 mois
- ✓ Migration complète des données
- ✓ Valable dans toutes les provinces

Rendez-vous sur [ambassadenumérique.ca/France](http://ambassadenumérique.ca/France)



AFFAIRES ÉTRANGÈRES, COMMERCE  
ET DÉVELOPPEMENT CANADA  
Service des délégués commerciaux

FOREIGN AFFAIRS, TRADE  
AND DEVELOPMENT CANADA  
Trade Commissioner Service

## Implications pour l'acteur public

Ce principe de réputation portable interroge quant aux enjeux éthiques d'un tel croisement de données ainsi que sur les difficultés techniques d'une compilation de données provenant de services et de supports divers. Il permet néanmoins d'aborder la place de la collectivité publique dans la régulation ou l'orchestration des relations/partenariats entre acteurs de l'économie des plateformes. Le fait de se positionner pour certifier des citoyens sur la base des données partagées apparaîtrait judicieux pour plusieurs raisons: (a) afin de ne pas laisser des acteurs privés le faire de manière hasardeuse, (b) dans le but d'améliorer la qualité des indicateurs (par exemple en croisant avec des données publiques libérées), (c) afin d'exercer une influence sur la dimension normative de ces indicateurs.

# Tiers de confiance

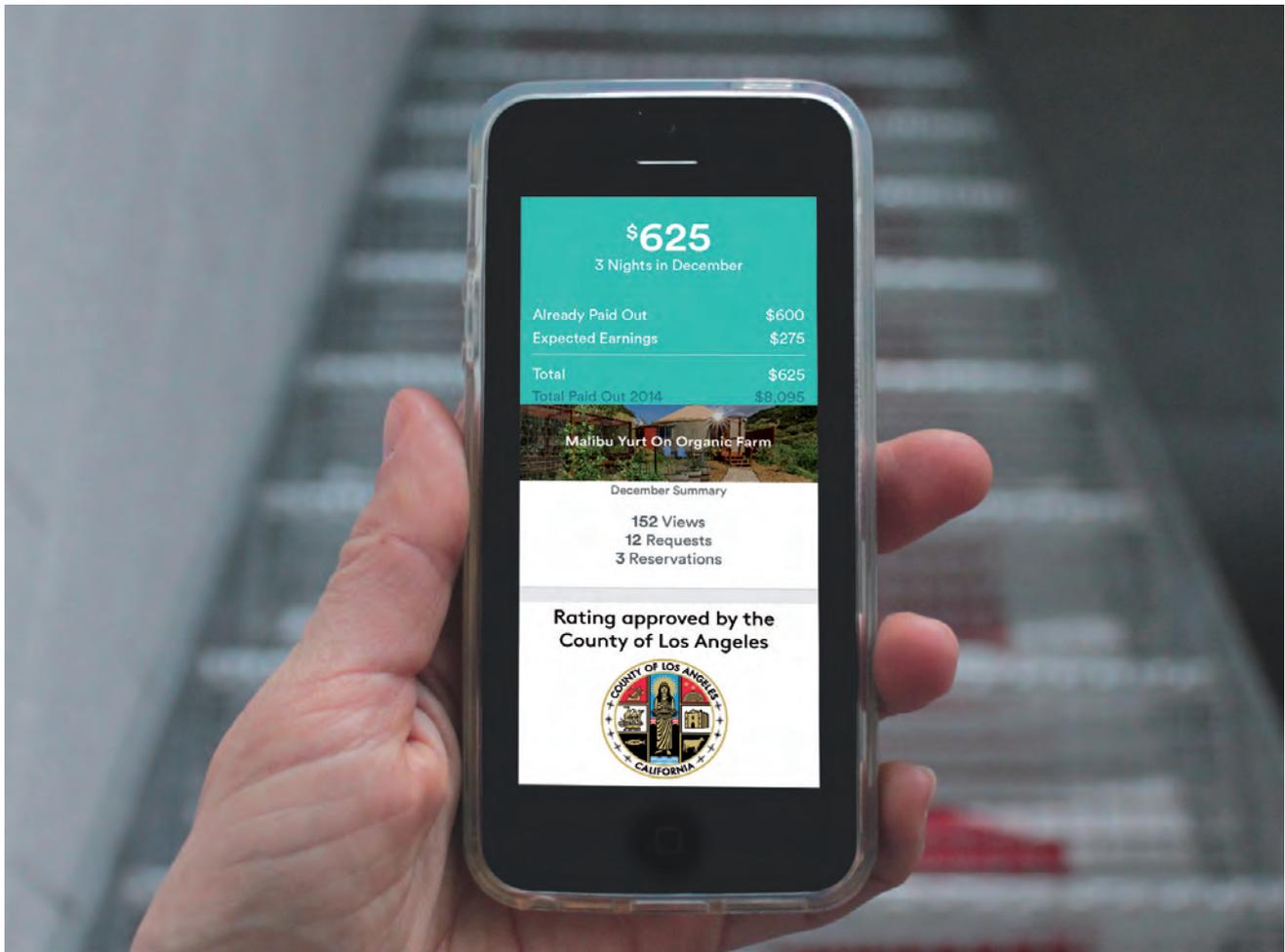
## *Nouvelles certifications*

Une autre posture de tiers de confiance, variante moins problématique que celle décrite dans le scénario précédent, peut consister à certifier la qualité des données et leur mise en circulation entre différentes plateformes privées ou para-publiques. L'objectif étant alors demain de s'assurer de l'éthique et de la responsabilité des données garant d'une politique publique d'un secteur donné. Ce positionnement semble cohérent avec le suivi des actions mises en place par des acteurs privés délégataires de services publics (gestion des déchets, chauffage, eaux, etc.) pour déterminer si ceux-ci répondent ou non à des enjeux de politique publique. La mise en place d'une telle labellisation pourrait participer à la définition d'une responsabilité contractuelle dans les usages des données : quel lien entre les objectifs fixés par la collectivité et les indicateurs à disposition ? Quelles initiatives de respect de la sphère privée sont-elles inscrites dans la plateforme considérée ? Quelles formes de réappropriation des données sont-elles proposées aux usagers ? Dans ces différents cas, la certification pourrait concerner les multiples dimensions du cycle de vie des données : hébergement, sécurisation, types de croisement de données autorisées ou non, pérennité dans le temps (avec les enjeux éventuels de droit à l'oubli), etc.

En outre, il est vraisemblable qu'à l'avenir une telle certification soit aussi applicable à des acteurs tiers impliqués dans d'autres types de services. Il s'agirait alors de garantir des services dans lesquels la réputation des citoyens, ou des prestataires seraient engagée; comme par exemple dans le cas de logeurs ou d'usagers d'AirBnB. La certification serait alors moins sur les données elles-mêmes que sur les indicateurs utilisés, et leur construction. De la même manière que toutes les balances de pesage de la ville de Los Angeles doivent être certifiées par un organisme de ce Comté (Los Angeles County), il serait pertinent que les indicateurs de réputation d'une plateforme telle qu'AirBnB, ou les propositions d'itinéraires d'un Waze, le soient aussi, établi sur des standards fixés par la collectivité.

Dans le même ordre d'idées, la mise en place d'un tel système pourrait amener les municipalités à interdire, ou punir un opérateur commercial si celui-ci ne répond pas à des critères de qualité imposés par la collectivité publique; comme le montre le cas récent du non-renouvellement de la licence d'Uber dans la ville de Londres, cette société n'ayant pas respecté les standards de responsabilité sociétale des entreprises.

Par comparaison avec le scénario précédent, ce genre d'initiative relève d'une autre posture de la collectivité publique, qui se positionne moins en tant qu'acteurs que dans une logique de labellisation des initiatives dans lesquelles des données sont échangées ou partagées.



### Implications pour l'acteur public

Malgré le caractère logique d'une posture de collectivité qui labellise et certifie, un tel scénario interroge quant aux leviers d'action à disposition pour jouer un rôle de régulation. S'il est envisageable de jouer sur les conditions d'octroi de concessions publiques ou de régies municipales de manière contractuelle, la situation est plus complexe dans le cas des plateformes externes (AirBnB, Uber, Waze). Cependant, les exemples d'Austin et Londres dans le cas d'Uber, et de Barcelone pour AirBnB montrent que des municipalités ont pu intervenir. En outre, ce scénario implique la mise en place d'un rapport de force avec des acteurs tiers (délégataires de services publics, plateformes privées) afin de positionner les acteurs publics dans une situation d'arbitre. Or, celle-ci n'est pas forcément évidente de prime abord, et implique une prise de risque avec une dimension politique. En effet, comme dans le cas d'Uber à Londres, les interventions publiques de ce type pouvant être mal perçues par les entreprises innovantes. Cet état de fait suppose cependant que les collectivités reprennent confiance dans son pouvoir de négociation, sans uniquement travailler sur ces questions pour se placer dans le jeu concurrentiel entre grandes villes.

# Frictions socio-techniques

## *Obsolescence des choix techniques*

Le décalage entre le temps des décisions démocratiques (qui requiert une analyse posée de la situation ainsi qu'un débat entre les parties prenantes) et le temps des techniques qui évoluent rapidement est un thème classique des récriminations contre les collectivités quant à leur efficacité, ou à leur éventuel manque de modernité. Cette même célérité des changements techniques va d'ailleurs de pair avec une difficulté pour les grandes organisations à construire une expertise quant aux choix et aux usages des dispositifs technologiques à faire.

C'est la raison pour laquelle les médias, ou des groupes de « citoyens veilleurs » attentifs aux dépenses de deniers publics, critiquent de temps à autre les développements de plateformes ou les achats qui sont effectués par les collectivités. La mise en place de nouveaux systèmes d'information dans les services publics est un cas typique à cet égard; ceux-ci étant mis en avant comme des échecs (alors même que le nombre d'expériences coûteuses et problématiques n'est pas forcément majoritaire, et que celles-ci permettent souvent la constitution d'une expertise), cristallisant une forme de rancœur auprès des observateurs.

De multiples choix à la fois techniques et juridiques (choix des licences, fonctionnement de celles-ci) vont être discutés et critiqués dans le débat public, sous des formes plus ou moins simplificatrices, et plus ou moins agressives.



Accueil > Recherche > Recherche Question écrite N° 28546 de Mme [redacted]



## 14ème LEGISLATURE

Questions N°  
28546

de Mme [redacted] (Ecologiste - Calvados)

Question écrite

Ministère interrogé > Budget	Ministère attributaire > Budget et comptes publics	
Rubrique > fonctionnaires et agents publics	Titre > statut	Analyse > logements de fonction, réglementation.

Question publiée au JO le: **30/08/2017** page: [7595](#)  
Réponse publiée au JO le: **11/04/2017** page: [2914](#)  
Date de changement d'attribution: **07/12/2017**  
Date de renouvellement: **03/01/2018**

### Texte de la question

Mme [redacted] attire l'attention de Mme la ministre déléguée auprès de la ministre de la réforme de l'État, de la décentralisation et de la fonction publique, chargée de la décentralisation, sur l'application de la circulaire n° 5608 du 19 septembre 2012 du Premier ministre, définissant les orientations pour l'usage des logiciels libres dans l'administration. La circulaire incitait les ministres à l'utilisation des logiciels libres dans leurs services. Elle souhaite savoir quelles suites ont été données à cette circulaire, notamment les études d'opportunités de migration de logiciels, l'intégration de ce critère dans les appels d'offres, les projets de migration de logiciels propriétaires vers des logiciels libres ou encore la mise à disposition des sources de logiciels développés en interne ou par un prestataire, au sein du ministère et de l'intégralité des administrations qui en dépendent. Elle souhaite de plus connaître le montant des dépenses en logiciel, en distinguant les logiciels propriétaires des libres, au sein du ministère et des administrations qui en dépendent, pour chaque année de 2008 à 2012.

## Implications pour l'acteur public

La mise en place de tout système d'information, quel qu'il soit, et même dans une logique vertueuse telle que celle de l'*Open Data*, va amener de manière croissante des observateurs externes à évaluer et émettre un jugement quant au bien-fondé et au bon fonctionnement de ses plateformes. D'où le besoin d'une justification publique de ces décisions qui va être demandée plus fréquemment, même sur des dimensions a priori très techniques. Et ce, d'autant plus que l'*Open Data*, et la problématique des données, sort de son caractère de niche pour devenir un enjeu sociétal majeur; ce qui peut amener des discours simplificateurs à son propos. C'est la raison pour laquelle le fait de privilégier et de promouvoir des services open source peut favoriser l'appropriation de technologies nouvelles, et leur adaptation aux évolutions à venir, au-delà des formats propriétaires dont la pérennité peut être plus limitée, comme dans le scénario ci-contre.

# Frictions socio-techniques

## *Le marché gris du profilage*

Avec la généralisation des objets numériques produisant et collectant des données, risquent d'apparaître de multiples métiers officiels comme nous l'avons vu plus haut. Mais il faut également compter sur des profils plus officieux pour accompagner, aider, voire détourner les services actuels et à venir. Lesquels pourraient impliquer ou non des données venant de plateformes *Open Data*, mais plus probablement celles des délégations de services publics (eau, électricité) ou de prestataires de santé.

Si ce phénomène paraît exagéré à première vue, il existe hors de France différents cas récents de piratage de comptes/profils d'assurance automobile, de remplacement de données podométrique pour tromper les assurances santé, de reprogrammation de boîte noire de véhicule (fraude à l'assurance basée sur les données GPS). Ou dans un secteur plus urbain, des affaires de fraude au comptage de piétons et de cyclistes (dans des visées politiques) pourraient aussi exister.

Malgré les dires des fabricants et des fournisseurs de services qui insistent toujours sur le caractère sécurisé et non-attaquables de leurs produits, ces nouveaux objets numériques que sont les appareils domotiques ou les véhicules munis de capteurs vont faire l'objet de services frauduleux; de la même manière que les décodeurs Canal+ en leurs temps, et que les déblocages de smartphone actuellement.

Il semble par conséquent probable qu'un marché gris et frauduleux du calibrage et du profilage va émerger; par exemple avec les compteurs électriques du genre Linky, les objets connectés dans le secteur de la santé, les services domotiques de thermorégulation et de sécurisation du domicile. Si une telle offre pourrait concerner des travailleurs potentiellement ancrés sur le territoire (comme dans le cas de l'annonce ci-contre), il est aussi envisageable que des acteurs distants offrent toute une gamme de coups de main à distance. Par exemple en substituant des données, en utilisant et en reprogrammant des faux-profils, etc.



### Implications pour l'acteur public

L'existence d'un marché gris détournant ou piratant les technologies de l'époque est quasi-inévitable. Si la dimension de recours légal contre ces services frauduleux n'est pas à négliger, il apparaît néanmoins pertinent de mieux connaître ce type d'offre. Sans vouloir faire d'angélisme sur le comportement à adopter, il semble pertinent de ne pas négliger ce marché gris de l'intervention sur les données. Une meilleure compréhension de celui-ci permettrait d'abord de mieux saisir des besoins et enjeux éventuels des populations considérées, mais aussi de savoir comment mettre en place une réponse pertinente (démarche de labellisation, accompagnement et valorisation des services certifiés, etc.). En effet, des services frauduleux à première vue peuvent aussi être vus en tant que moyen de défense ou de réappropriation des citoyens contre la dimension perçue comme prédatrice de certains services technologiques. L'existence de tels services serait du coup un indicateur intéressant dans une logique de régulation publique ou de certification.

# Nouvelles formes de protestations

## *Dévolement des données*

Si l'*Open Data* repose sur une logique de compilation et de partage de connaissances avec un objectif vertueux d'optimisation du fonctionnement de la ville et de construction de connaissances sur les multiples phénomènes sociaux qui s'y déroulent, rien n'empêche des acteurs moins scrupuleux de faire un usage orienté des données mises à disposition. Dans le contexte de la crise épistémologique actuelle – qui se traduit tant par les polémiques autour des « fausses nouvelles » (*fake news*) que par les conflits autour des statistiques publiques de chômage ou d'immigration – on voit apparaître de tels usages dévoyés des données partagées. C'est la manière avec laquelle le climatocceptique Bjorn Lomborg semble procéder dès 2001 puisqu'il agrège des jeux de données différents et sans précaution pour en tirer des conclusions politiquement orientées.<sup>8</sup>

Il apparaît donc vraisemblable que de telles pratiques puissent être amenées à prendre de l'ampleur. En particulier du fait de la mise à disposition croissante de jeux de données sur toutes sortes de sujets, et à cause des tensions grandissantes sur diverses polémiques de société. Le thème des statistiques ethniques qui est particulièrement sensible en France pourrait par exemple se voir saisi par des activistes. Lesquelles pourraient croiser de façon plus ou moins solide des données partagées ou bientôt mises à disposition, avec d'autres sources pas forcément valides afin de montrer des corrélations et communiquer sur une version simpliste de leur interprétation.

9. Voir en particulier <http://grist.org/article/for/>



Un screenshot of the website **MEDIAPART**, dated **MAR. 29 AOÛT 2017 - ÉDITION DE LA MI-JOURNÉE**. The page features a navigation bar with categories like **LE JOURNAL**, **LE STUDIO**, **LE CLUB**, **DEPUIS 48 HEURES**, **LES BLOGS**, **LES ÉDITIONS**, **LAGENDA**, **LA CHARTE**, and **PARTICIPEZ**. The main headline is **Un dévoiement du Big Data pour légitimer une idéologie nauséabonde**, dated **18 OCT. 2019** by **PAR STEPHANE DESHAYES (CSDP)**. The article text begins with: "Une bonne statistique publique est indispensable au bon fonctionnement de la démocratie : cela semble évident. Pourtant, depuis quelques mois, la statistique française est minée par une nouvelle forme de menace. En effet, depuis les dernières élections européennes, les multiples factions de la droite conservatrice et de l'extrême-droite ont repris de manière quasi systématique des statistiques ethniques qui n'ont pas leur place dans le débat français. Or, et c'est bien là le cœur du problème, celles-ci ne proviennent pas d'organismes publics – puisque leur collecte et leur analyse est rigoureusement illégal au regard du droit français – mais de think tanks basés en Suisse et à Malte. Lesquels expliquent (voir article de nos confrères du journal Le Temps) produite ces statistiques eux-mêmes sur la base de croisement de jeux de données compilés sur les différentes plateformes de partage étatique et municipales françaises. En d'autres termes, c'est la logique de l'Open Data elle-même qui est dévoyée pour des buts peu avouables. Et naturellement, comme l'ont souligné plusieurs chercheurs dans la presse récemment, la combinaison de ces données n'est que peu fiable; ce qui semble attester d'un objectif politique de propagande et non de construction de savoirs valides." The author's bio for **CDSP** is also visible, listing **1 ÉDITION**, **21 ARTICLES D'ÉDITIONS**, **3 LIENS**, **1 FAVORI**, and **1 CONTACT**.

## Implications pour l'acteur public

Ce scénario de détournement des données ouvertes montre bien que tout jeu de donnée n'est pas neutre et relève d'une construction.<sup>10</sup> Est-ce qu'une telle situation est évitable ? Il est possible d'anticiper les critiques en travaillant sur deux dimensions. D'une part, l'acteur public doit continuer, comme il le fait déjà, à s'engager pour améliorer l'approche de la donnée, de sa qualité, de sa description, de son interopérabilité, et de sa compréhension par tous. Et d'autre part, s'il est difficile d'empêcher ce type d'utilisation orientée des données, il apparaît pertinent de construire dès aujourd'hui un contre-argumentaire, par exemple une typologie, des mésusages qui pourraient être fait des données partagées. Il s'agirait par exemple de fournir des explications plus claires sur les limites de celles-ci, sur les intentions attachées à ces données, et comment éviter des sur-interprétations à partir de données qui n'ont pas été construites pour tel ou tel objectif.

10. Voir en particulier <http://grist.org/article/for/>

# Nouvelles formes de protestations

## *Luddisme de la donnée*

Malgré le caractère vertueux de l'*Open Data* en général, il faut néanmoins garder à l'esprit que la dimension quantitative et statistique du partage de données n'est pas forcément perçue de manière positive par tous les citoyens. D'une part, la critique de l'*Open Data* peut être faite sur son principe même, décrié par certains du fait d'une logique simplificatrice de la mise en chiffre du monde, ou d'un reproche général fait aux objets techniques proches du luddisme.<sup>11</sup> D'autre part, certains collectifs ont développé une compréhension des limites et des écueils des données numériques ainsi que de leurs usages. Ces mêmes usagers-experts pourraient ainsi construire un point de vue critique très élaboré sur différents phénomènes de société et du coup trouver discutables les interprétations effectuées par les collectivités sur la base de données publiques.

Cela pourrait par exemple être le cas de toutes données visant à dévaloriser certains quartiers du fait de statistiques montrant un décalage par rapport aux moyennes sur des thèmes que la criminalité, le chômage, l'éducation ou des enjeux de santé. L'analyse de la dimension spatiale des données publiques induit un effet de compétition et de classement qui peut provoquer un sentiment de stigmatisation inévitable à ce genre d'exercice. Notons cependant qu'un tel phénomène fonctionne à double-sens, les activistes développant une expertise technique ou statistique peuvent être amenés à déconstruire les jugements produits à partir des données disponibles et proposer de nouveaux indicateurs.

De la même manière, il serait possible, mais non-souhaitable pour les acteurs publics qu'un partenaire de l'*Open Data* (DSP, communes, associations) se déclare *data free* avec la volonté de ne plus partager des données relevant de ses prérogatives (ou en empêchant tout projet statistique ou de pose de nouveaux capteurs par exemple). Un tel phénomène interroge sur la mise en place d'une politique de données ouvertes lorsqu'un partenaire ne veut pas jouer le jeu ; il est lié au risque de constituer des jeux de données incomplets, difficiles à valoriser auprès d'organisations tierces qui ne pourraient pas les utiliser pour proposer des services pertinences aux habitants ou aux visiteurs.

11. Voir par exemple <http://www.piecesetmaindoeuvre.com>



### Implications pour l'acteur public

S'il apparaît difficile d'empêcher tout phénomène de contestation de la production de données<sup>12</sup>, il n'en demeure pas moins pertinent de faire participer des associations et des usagers-experts de ce type, même lorsqu'il s'agit d'activistes de divers bords politiques, dans les débats concernant la constitution d'indicateurs composites à partir de données publiques, voire dans la mise en place d'une stratégie pour la plateforme *Open Data*.

12. En particulier car les positions de certains activistes sont résolument contre la discussion sur ce type de sujet, ou même à la participation à des réunions de concertation.

# Nouvelles formes de protestations

## *La valeur économique des secrets*

Depuis longtemps déjà l'avantage concurrentiel d'acteurs économiques s'est construit sur l'asymétrie informationnelle. Le fait de posséder des informations plus ou moins précises sur un secteur donné a favorisé l'émergence d'entreprises, en particulier dans l'immobilier ou dans le secteur du bâtiment. Du fait d'une littérature grise ou de réseaux sociaux bien constitués, la connaissance de futures opérations de reconversion, de l'établissement de périmètres potentiellement constructibles, ou d'infrastructures permettant une valorisation accrue d'un territoire sont des cas classiques à cet égard.

Il est donc vraisemblable que la mise à disposition de données nouvelles via l'*Open Data* – des informations de nature géospatiale, mais pas uniquement – pourrait venir bousculer ces comportements. Autrement dit, l'ouverture des données va détruire des rentes. Non pas en les supprimant, mais en reconfigurant les réseaux d'informations, avec l'avènement d'acteurs potentiellement nouveau et plus à même de tirer parti du croisement d'informations numériques. Verra-t-on des groupes immobiliers modéliser les phénomènes de gentrification à partir des jeux de données *Open Data* croisées avec leurs propres statistiques? Est-ce qu'une telle automatiser des analyses sur la base de ces données ne risquerait pas de privilégier de gros acteurs distants? Ou est-ce qu'au contraire la mise à disposition de ces données fera disparaître cette valeur du secret pour permettre une plus grande égalité des chances entre les acteurs économiques?

La valeur économique des secrets

### Implications pour l'acteur public

Comme il faut s'attendre d'une part à des protestations potentielles d'acteurs qui bénéficiaient sans trop le communiquer d'informations et de statistiques qui n'était pas largement partagées, il apparaît important de construire un argumentaire concernant les raisons et le bien-fondé d'un tel partage dans une logique *Open Data*. Par exemple en soulignant le fait que des rentes acquises il y a longtemps doivent être contrées dans une logique d'égalité, sans forcément pointer les organisations ou les personnes qui en ont bénéficié. Ou encore, en rappelant que même sans l'ouverture de telles données, des entreprises trouveront ou créeront elles-mêmes ces données ou des données équivalentes. De plus, il conviendrait aussi de superviser, et donc de réguler, l'apparition de nouvelles positions dominantes du fait de la libération des données; surtout si celle-ci provient d'entreprises non-situées sur le territoire. Et cela, car la collectivité se doit de garantir l'égal accès au marché via les données. Par contre, il semblerait malvenu de modifier l'accès aux jeux de données au gré de cette supervision. Et ce, car la stabilité de l'information est une condition importante pour la mise en place d'une offre de services continue et pérenne.

# Le Point

Hebdomadaire d'information du jeudi 20 décembre 2018 n° 3000



## Les nouveaux génies de la data

Flairer les bons coups  
avec les algorithmes

Quels risques pour nous?  
Quels risques pour nos industries?

# Nouvelles formes de protestations

## *Modélisation et données*

La mise à disposition toujours croissante de données, publiques ou non, est utilisée pour à la fois pour bâtir de nouveaux services aux usagers, mais également afin de nourrir les travaux de modélisation prévisionnelle de multiples phénomènes. Et cela, particulièrement par les organisations intéressées par l'évaluation des risques.

C'est notamment le cas de l'activisme citoyen qui s'empare parfois de données partagées pour construire ses propres modèles; un phénomène dont on peut se rendre compte également avec les initiatives de data-journalisme. Il est par exemple possible que des associations de mobilisation contre des nouvelles portions d'autoroutes utilisent des données libérées de mobilité pour chercher à contester l'utilité de tels projets.

Dans un autre registre, des compagnies d'assurances, et de ré-assurances, voient dans l'analyse de mégadonnées un moyen de qualifier plus finement le profil de risque de ses clients, ou d'évaluer plus précisément l'avènement de phénomènes qui engageraient un risque prudentiel. C'est par exemple le cas de l'analyse des données météorologiques réalisée sur la base de données extraites de plateformes *Open Data* locales et internationales, et qui sert à estimer l'occurrence de catastrophes telles que des tempêtes ou des sécheresses.

Les données partagées sont donc une opportunité nouvelle et un risque pour les collectivités. La mise à disposition des données peut certes permettre la création de services censés améliorer la qualité de vie des usagers de la ville, mais, d'autre part, il est vraisemblable que la construction des modèles prédictifs risque de susciter des frictions nouvelles, engageant la responsabilité de la collectivité, ou de ses liens avec des délégations de services publics: comment réagir si les prévisions réalisées sur la base de données alimentent une logique conflictuelle avec la collectivité? Que se passe-t-il si une simulation prédictive contredit des décisions prises dans le passé?

# Comment réagir si la simulation prédictive a sous-estimé les risques financiers ?



## Implications pour l'acteur public

Comme souvent avec les technologies de l'information et de la communication, on constate que celles-ci mettent en lumière des phénomènes qui restaient discrets auparavant. Dans une logique de transparence, on peut donc faire l'hypothèse que les avantages de tels changements excèdent ses limites. Néanmoins, pour la collectivité, il apparaît nécessaire de cartographier les nouveaux écueils émergents et leurs conséquences, afin d'anticiper des modalités de réaction en temps voulu. Cela implique la constitution d'une expertise interne de modélisation, sur la base des données de la plateforme *Open Data*, des statistiques municipales qui ne sont pas forcément partageables, et de celles fournies par des acteurs tiers (délégations de service-public). Cette opportunité d'ouverture des données est en dernière analyse liée à la notion d'intérêt général, et en particulier à qui la défend. Elle part du principe que la position de la collectivité – censée assurer l'intérêt général – soit connue, mais elle implique simultanément d'accepter un débat démocratique, quelles que soient les conséquences et les dérives potentielles.

# Données et collaboration

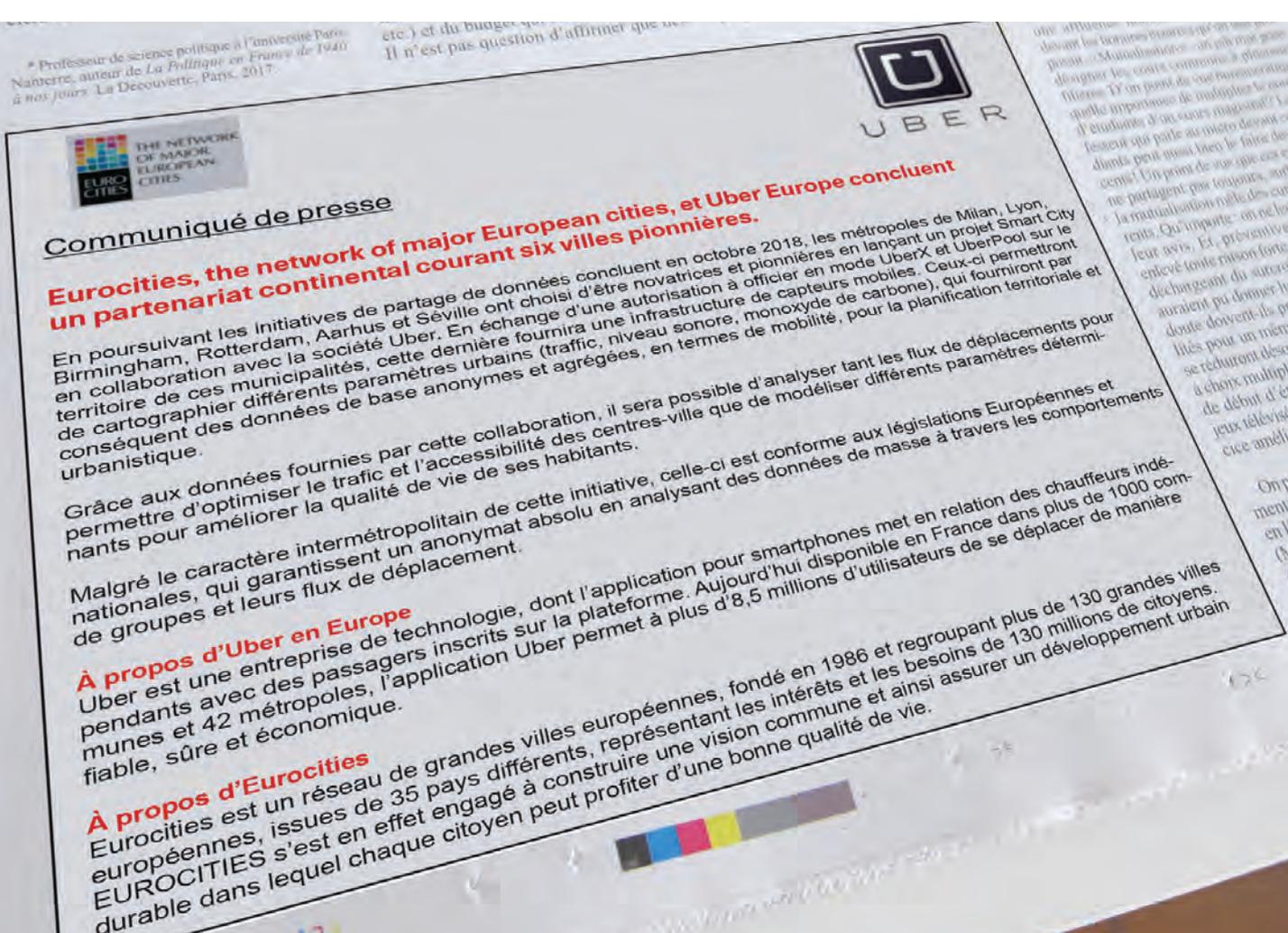
## *Collaboration à bénéfices mutuels*

La majorité des données mises à disposition sur les plateformes *Open Data* des collectivités publiques proviennent de leurs propres services, mais l'on voit émerger de plus en plus de collaborations avec d'autres acteurs tiers (DSP, Vélos en libre service).

Cette opportunité interroge sur le fait que la collectivité puisse pousser un opérateur à partager ses données à la fois sur la plateforme de diffusion des données territoriales, mais également sous les conditions des licences spécifiques. Si les délégations de services publics vont dans cette direction de manière légale (Loi pour une République Numérique) voire pour certaines données issues du secteur privé, contractuelles, et nécessitant des négociations certainement ardues, il peut aussi être pertinent de s'interroger sur les moyens de contraindre des opérateurs tels qu'Uber, Blablacar ou AirBnB à le faire de manière poussée. Certes, ces services partagent des données avec certaines des municipalités avec lesquelles elles ont conclu des accords, mais des doutes subsistent sur la qualité de celles-ci, et sur les limites possibles de leurs usages.<sup>13</sup>

On peut toutefois imaginer des modes de collaboration plus doux et qui passent par des échanges de services minimisant les risques. Par exemple, comme dans l'exemple ci-contre, il ne serait pas absurde de voir un acteur tel qu'Uber fournir des données de pollution urbaine (monoxyde de carbone, bruit) enregistrées sur la base sur la base de capteurs dans les véhicules installés en collaboration avec la collectivité. Un tel échange interroge toutefois sur les modalités de négociation et les moyens de contrainte ou de régulation des municipalités qui ne sont pas nécessairement l'échelon administratif adapté pour cela.

13. Voir par exemple les cas de partage de données de mobilité entre Uber ou Waze et de la ville de Boston.



## Implications pour l'acteur public

Dans une perspective de construction de partenariats ciblés et pertinents pour la collectivité, et en complément avec les initiatives actuelles qui vont dans ce sens, il serait intéressant d'adopter une démarche proactive en définissant des organisations hors-service public avec lesquelles privilégier un partenariat dans l'échange des données. Et, en second lieu, de rencontrer des plateformes numériques afin d'évaluer les opportunités et limites de négociation sur ces aspects.

Dans ce contexte, le fait de garantir un usage des données des citoyens conformes aux règles françaises, européennes (RGPD) va devenir un axe fondamental pour les collectivités publiques. En particulier afin de favoriser une négociation à bénéfices mutuels tout autant pour les individus que pour les acteurs du secteur privés du territoire. Et ce, puisqu'en définitive l'encouragement du partage des données privées contribue à un meilleur pilotage des politiques publiques, et donc à l'intérêt général.

# Données et collaboration

## *Complétion des données publiques*

Bien souvent les jeux de données publiques présents sur les plateformes d'*Open Data* n'existent que grâce à l'historique d'initiatives des collectivités publiques. Suivant les territoires, au gré des expérimentations et des partenaires disponibles pour mener celles-ci, des thématiques sont plus abordées que d'autres. C'est la raison pour laquelle certaines municipalités ont construit une expertise et une communication autour de thèmes spécifiques tels que la mobilité ou la culture. Une des conséquences d'une telle situation est qu'il y a parfois des absences ou des manques cruciaux dans les jeux de données; une situation qui peut se manifester lorsque des nouveaux acteurs entrants sollicitent les collectivités pour lancer de nouvelles initiatives, ou que des réglementations pourraient imposer l'usage de certaines données.

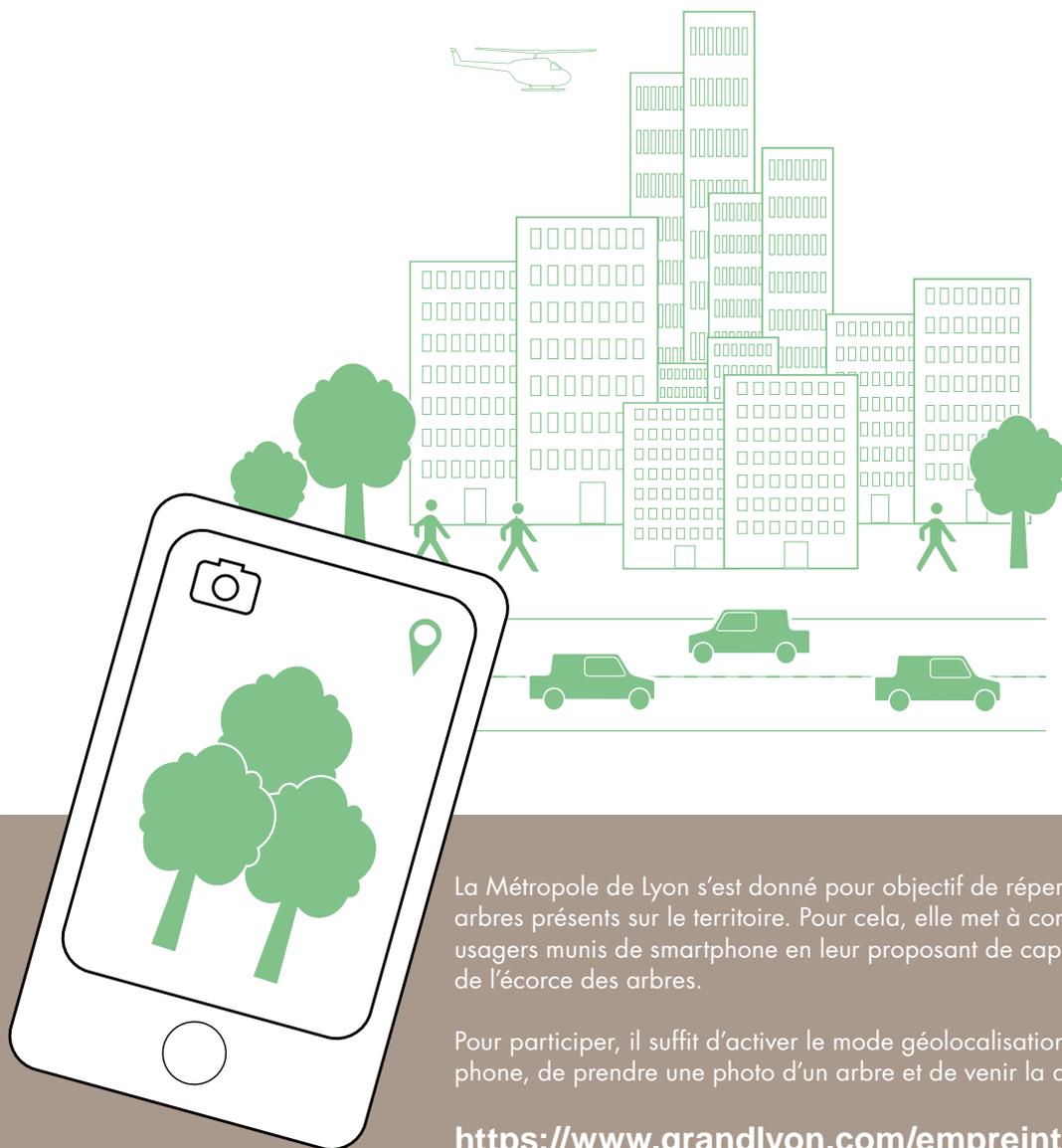
Lorsque de telles situations se présentent apparaît alors la nécessité de constituer des jeux de données nouveaux, ce qui est évidemment coûteux et parfois difficile à mettre en place rapidement. On peut donc imaginer la mise en place d'initiatives nouvelles, comme par exemple la création d'événements dédiés, afin de remplir cet objectif et de poursuivre d'autres buts. Si le cas de Respiralyon vient à l'esprit<sup>14</sup> comme exemple d'une telle démarche volontaire, on peut imaginer la création d'initiatives comme des jeux urbains afin d'allier collecte de données et dimension ludique pour les citoyens.

### Implications pour l'acteur public

Tous les « manques » dans les jeux de données publiques ne nécessitent pas le développement d'un jeu grand public, mais il semble pertinent de trouver des formes de complétion et de constitution de ceux-ci. Et ce, par exemple avec des partenariats à bénéfices mutuels ou par des actions dédiées. Malgré son intérêt, une initiative de ce genre aura pour limite le fait que des citoyens pourront demander des comptes à la collectivité; par exemple en insistant sur les contreparties pour les citoyens d'un tel travail (les initiatives de crowdsourcing étant parfois perçues comme un moyen de faire travailler des individus sans rémunération). Que faire si une partie de la population souhaite une contrepartie? Ou limiter l'usage qui sera fait plus tard des données générées? La mise en place de mécanismes contributifs implique donc une éventuelle évolution du cadre de l'*Open Data* ou l'élaboration d'un argumentaire autour de la ville contributive qui valorise son apport pour la communauté qui dépasserait l'intérêt marchand basé sur la rémunération individuelle.

14. Un dispositif de surveillance et de lutte contre les nuisances olfactives sur l'agglomération lyonnaise qui fonctionne pas signalement volontaire : [www.respiralyon.org](http://www.respiralyon.org)

# Empreintes d'écorces



La Métropole de Lyon s'est donné pour objectif de répertorier tous les arbres présents sur le territoire. Pour cela, elle met à contribution les usagers munis de smartphone en leur proposant de capturer des images de l'écorce des arbres.

Pour participer, il suffit d'activer le mode géolocalisation de votre téléphone, de prendre une photo d'un arbre et de venir la déposer sur

**<https://www.grandlyon.com/empreintedecorces>  
entre décembre 2018 et juin 2019.**

Toutes les photographies réunies seront montrées sous la forme d'une installation interactive et d'une exposition en ligne à l'automne 2019, réalisée avec la société transmedia faun.tv.

Mais ces images permettront surtout de constituer un répertoire des arbres et des espèces pour le service géomatique et données métropolitaines de la Direction Innovation Numérique et Systèmes d'Information de la Métropole. L'analyse et le traitement des images sera réalisé en partenariat avec le laboratoire d'informatique LIRIS (Université de Lyon).

**GRANDLYON**  
la métropole

# Enjeux et implications

## Entre appropriation et dépossession

D'un point de vue général, les scénarios présentés ici reflètent la situation complexe à laquelle les collectivités publiques font face. Tout d'abord, même si l'objet de ce document concernait les données ouvertes, force est de constater la difficulté à ne pas traiter des données provenant de l'économie des plateformes (AirBnB, Uber, etc.) tant l'existence de celles-ci est liée aux enjeux discutés ici. Une prospective sur la donnée publique partagée implique donc d'en tenir compte, quitte à montrer une situation difficilement lisible, mais cohérente avec la réalité du terrain. Ce caractère confus n'est cependant pas inéluctable comme on l'a vu à plusieurs reprises, en particulier à propos des mentions au pouvoir d'intervention des collectivités: régulation, labellisation, accompagnement des acteurs qui certifient, etc.

Ensuite, comme l'indique le sous-titre de ce document, nous nous trouvons dans une période singulière qui marie appropriation et sentiment de dépossession. D'où un ensemble de scénarios qui jouent sur cette dualité. Si les questions d'*Open Data* sont maintenant installées dans les discours et les pratiques publiques, la stabilité de ceux-ci est toute relative. Les mutations dans l'environnement économique (les nouvelles entreprises qui en tirent parti ou viennent empiéter sur ce terrain) ou politique (changement du cadre légal national et européen) en attestent, et donnent l'impression d'un constant besoin de réinventer les choses et donc d'une phase d'appropriation sans fin. Laquelle se traduit en particulier par une nécessité de repenser la plateforme de partage de données, de la repositionner au gré des évolutions de cet écosystème, ou encore d'un besoin de communiquer sur ces mutations. En parallèle de ce sentiment d'appropriation perpétuel, la collectivité est confrontée à une logique de dépossession de ses prérogatives. Si d'un côté les collectivités mettent à disposition des données publiques qui bénéficient ensuite à des acteurs aussi gigantesques que lointains, et que les moyens de gouvernance et de régulation de l'activité de ceux-ci sont limités, comment ne pas y percevoir une perte de souveraineté? Est-ce que la mise à disposition des données ouvertes et libérées ne risque pas de faciliter l'hypercentralisation des grands acteurs industriels du secteur? Ou, peut-on encore croire qu'elle va offrir la possibilité à de nouveaux entrants de rebattre les cartes?

C'est donc le passage d'un monde de la donnée à un autre qui est présent en filigrane dans ces scénarios. Cette transition insidieuse fait percevoir le risque d'une collectivité débordée d'un côté par le haut, par de grands acteurs distants qui tirent parti des données partagées; lesquelles viennent encore renforcer les positions acquises. Et ce débordement peut d'un autre côté apparaître par le bas, avec toutes sortes de dévoiement des données, de l'apparition d'une économie grise (Cf. « Le marché gris du profilage »), ou des moyens de contourner l'action publique (Cf. « Nouvelle forme de protestations »).

## Quels leviers d'action ?

Malgré la marge de manoeuvre a priori limitée des acteurs publics sur ces questions actuellement, tant due à l'échelle de prérogatives publiques qu'elle peut mobiliser que liée à la présence d'entreprises technologiques très puissantes, nous soulignons avec certains scénarios le besoin de réaffirmer la position de la collectivité comme un acteur au coeur de ces sujets. La nouvelle réglementation européenne à ce sujet (Règlement européen sur la Protection des Données) étant en outre l'occasion de poursuivre les initiatives menées jusqu'à ce jour.

En continuité avec l'action publique actuelle qui va déjà dans ce sens, ce positionnement correspond à quatre leviers à mener conjointement, décrits dans les scénarios correspondants :

— Des actions de communication et d'information (Cf. « Information et pédagogie ») par exemple envers de jeunes publics, mais pas exclusivement; ce qui nécessite aussi de construire une position cohérente d'une part, mais également la constitution d'une expertise interne sur ces métiers (« Métiers à venir. Postes publics ») d'autre part.

— Des actions de participation et de co-élaboration des jeux de données de même qu'une explication sur les jeux de données qu'elle produit (Cf. « Complétion des données publiques »).

— De mener des expérimentations grandeur nature à un niveau métropolitain qui permettent ensuite de changer d'échelle et potentiellement de faire évoluer les législations nationales ou européennes.

— D'être pleinement un acteur de la régulation (et la confiance qu'elle doit avoir en celle-ci), qui peut être déclinée en plusieurs actions concrètes: tiers de confiance (Cf. « Portabilité de la réputation », « Nouvelles certifications »), collaboration dans l'échange de données (Cf. « Collaboration à bénéfices mutuels »).

# Lexique

## *Algorithme*

suite d'opérations ou d'instructions, mise en œuvre sous forme de calculs, de programmes, permettant d'effectuer des traitements complexes, de résoudre un problème, d'obtenir un résultat.

## *Big Data (ou mégadonnées)*

ensembles de données produites par l'utilisation des technologies numériques à des fins personnelles ou professionnelles, et caractérisées par des critères de volume (les masses de données à traiter sont sans cesse croissantes), vitesse (la collecte, l'analyse et l'exploitation de ces données étant réalisées en temps réel) et variété (hétérogénéité des formes et formats).

## *Datafication*

néologisme désignant le processus de mise en données du monde, c'est à dire de traduction des comportements complexes humains en jeu de données.

## *Data-janitor*

nouvelle profession apparue avec l'avènement du *Big Data*, et qui consiste à reformater, condenser et interpréter des méga-données en vue de leur usage.

## *Données (data)*

description élémentaire, une observation, une mesure.

### *Intelligence artificielle*

ensemble de techniques, de théories et programmes mis en oeuvre pour simuler des fonctions cognitives (reconnaissance de formes, apprentissage de tâches complexes, résolution de problèmes, raisonnement).

### *Licence*

accord légal qui définit le transfert de propriété intellectuelle du cédant de licence au licencié et qui en spécifie les conditions.

### *Machine-learning*

un champ d'étude de l'intelligence artificielle qui concerne la conception et l'implémentation de méthodes permettant à un programme informatique d'apprendre et d'évoluer par un processus systématique, et ainsi de remplir des tâches difficiles ou problématiques par des moyens algorithmiques plus classiques.

### *Open Data (ou donnée ouverte)*

données numériques dont l'accès est public, dont le format est ouvert, et leur usage libre : exploitation et réutilisation. Ces données pouvant être d'origine publique ou privée, produites notamment par une collectivité, un service public (éventuellement délégué) ou une entreprise.

### *Statistique inférentielle*

ensemble de méthodes permettant de tirer des conclusions fiables à partir de données d'échantillons statistiques.

# Impressum

Créé et produit par:  
The Near Future Laboratory  
Nicolas Nova  
Fabien Girardin

Participants à l'atelier prospectif  
(juin 2017): Fabrice Bardet, Martin  
Cahen, Alessandro Cerioni, Gilles  
Gesquière, Stéphane Grumbach,  
Tamam-Rose Hannouche-Yonis,  
Pierre Houssais, Angela Lanteri,  
Nicolas Leprêtre, Christine Malé,  
Jean-Loup Molin, Nicolas Nova,  
Nathalie Vernus-Prost.

En collaboration avec:  
Philippe Gargov (Pop-Up Urbain)  
Design graphique  
Constance Delamadeleine  
Lyon — Genève  
Février 2018



La plateforme Data du Grand Lyon vise à répondre aux enjeux de la métropole intelligente en donnant un large accès aux données des secteurs publics et privés du territoire afin de favoriser la création de services innovants et la participation citoyenne. Or, après une période prometteuse de démarrage, il convient d'interroger tant les évolutions possibles de la plateforme en fonction de paramètres contextuels (évolution des usages, environnement juridique, nouvelles formes d'expertise), mais aussi de nouvelles opportunités.

L'objectif de ce livret est de soutenir un débat plus large sur la gouvernance des données. Il propose une série de scénarios prospectifs concernant ses enjeux, ses opportunités, ainsi que le positionnement de la collectivité.



---

WWW.  
MILLENAIRE3.  
COM

RETROUVEZ  
TOUTES LES ÉTUDES SUR

MÉTROPOLE DE LYON  
20, RUE DU LAC  
CS 33569  
69505 LYON CEDEX 03