

8

JUILLET

2019

# NOUVELLES MODALITÉS DE RÉGULATION

## **La régulation par la donnée**

Autorité de la concurrence, AMF, Arafer, Arcep, CNIL, CRE, CSA

## Nouvelles modalités de régulation – la régulation par la donnée

L'environnement numérique se caractérise par une évolution rapide des technologies, qui engendre une innovation permanente tant sur les usages que sur les modèles d'affaires. La réglementation *a priori* classique peut s'en trouver dépassée par la difficulté à appréhender un environnement en évolution continue et à l'horizon inconnu.

L'action de l'Etat (et *de facto* des régulateurs) peut en complément s'inscrire dans le cadre évolutif que permet la régulation par la donnée. Celle-ci combine responsabilisation des acteurs, capacité renforcée d'analyse du régulateur, et information des utilisateurs<sup>1</sup> et de la société civile.

La régulation par la donnée vient compléter les outils traditionnels du régulateur. Au lieu de prescrire aux acteurs économiques un certain comportement, il s'agit de créer un réseau d'informations et d'incitations pour réduire les asymétries d'information et démultiplier l'impact de l'action du régulateur en mobilisant les utilisateurs et leurs relais. Cette approche appelle une nouvelle culture et de nouvelles compétences au sein de l'Etat.

Les travaux des AAI/API ont permis de mettre en évidence deux grands objectifs associés à la régulation par la donnée :

- amplifier la capacité d'action du régulateur, notamment dans une logique de supervision ;
- éclairer les choix des utilisateurs et mieux orienter le marché.

En pratique, cela passe non seulement par la collecte d'informations auprès des acteurs régulés mais aussi par un élargissement des données, par des outils de *crowdsourcing*, par des démarches de simulation, par l'animation d'un écosystème d'acteurs de la mesure, de comparateurs...

Le développement de la régulation par la donnée engendre, pour les régulateurs, de nouveaux besoins en termes de compétence, d'outils ou encore d'appropriation des nouvelles technologies.

---

<sup>1</sup> Entendu comme l'ensemble des utilisateurs finals des services sur le marché de détail (consommateurs, usagers, professionnels, etc.)

La démarche de régulation par la donnée consiste à utiliser la puissance de l'information pour comprendre le marché et éclairer son fonctionnement de manière factuelle pour mieux l'orienter et mieux assurer la protection d'ordre public du consommateur sur ces différents marchés.

Un premier type d'action procède d'une logique de supervision. Il s'agit de collecter et analyser des sommes d'information pour détecter des signaux faibles ou des problèmes systémiques, accélérer et rendre plus efficace la régulation ou veiller au respect d'obligations par les acteurs régulés.

Un deuxième type d'action consiste à fournir aux utilisateurs et aux acteurs relais (acteurs publics, associations, société civile, outils de mesure ou comparateurs, etc.) un contre-pouvoir par une information précise, individualisée voire personnalisée. A travers leurs choix, les utilisateurs peuvent devenir des acteurs de la régulation à part entière en pesant sur les décisions des acteurs économiques dont ils sont les clients. Dans ce cadre, le régulateur a un rôle important : il peut éclairer les choix des utilisateurs pour mieux orienter le marché. Cette capacité régulatrice des utilisateurs ne se substitue pas à celle des régulateurs. Elle peut contribuer à atteindre les objectifs de politiques publiques propres à chaque secteur et définis par les autorités publiques. Les acteurs économiques peuvent aussi au regard des informations publiées anticiper les réactions négatives des utilisateurs et acteurs relais, voire modifier leur comportement en amont, sachant leurs actions observées.

Il est essentiel de souligner que la régulation par la donnée complète les modalités d'intervention de l'Etat (prise d'engagements, injonctions, etc.) et ne peut exister sans une implication des pouvoirs publics. A cet égard, elle ne doit pas être confondue avec la simple transparence, notamment dans le cadre de l'ouverture progressive des données détenues par l'administration : d'abord parce que le régulateur intervient pour définir des priorités et mettre en lumière spécifiquement certains paramètres, ensuite parce que le régulateur vient normer certaines notions et le cas échéant centraliser l'information. Enfin, car la régulation par la donnée s'appuie sur un cadre juridique : s'il ne s'agit pas d'imposer des prescriptions quant au comportement sur le marché des acteurs régulés, c'est bien de manière uniformisée et organisée, et selon des paramètres définis par le régulateur que les acteurs doivent produire l'information.

Pour permettre ces évolutions, il est nécessaire d'accroître le volume et la qualité des données accessibles pour les régulateurs, ainsi que de multiplier leurs sources. L'un des exemples phare en est la mobilisation des utilisateurs pour signaler les problèmes rencontrés ou les informations observées (qualité de service ou communication faite aux utilisateurs par exemple) en passant d'une logique de plainte consommateur à un acte citoyen. Les municipalités ont développé ainsi des applications offrant l'opportunité à leurs administrés de faire remonter des dysfonctionnements divers (trottoir cassé, etc.). Dans le domaine de la régulation, ce type d'outil peut être très utile pour détecter des signaux faibles ou disposer d'informations directes permettant de *réduire les asymétries d'information avec* les acteurs régulés. Dans le domaine des marchés publics, cet outil pourrait être aussi pertinent pour permettre la détection d'irrégularités.

La régulation par la donnée, en définitive, complète les outils traditionnels du régulateur. Au lieu de prescrire aux acteurs économiques un certain comportement, il s'agit d'armer le régulateur et le consommateur/utilisateur pour donner des incitations au marché. Cette approche appelle certainement un changement de culture au sein de l'Etat. L'Etat ne garantit pas seul l'intérêt général : chaque acteur, chaque utilisateur peut prendre sa part à la défense de l'intérêt général et chaque citoyen est mis en capacité de devenir un acteur de la régulation. De plus, l'Etat n'est pas toujours le mieux placé pour apporter l'information pertinente aux consommateurs et il peut être plus efficace d'agir dans une logique d'*empowerment* et d'animation d'un écosystème « RegTech », c'est-à-dire de startups ou d'acteurs de la société civile susceptibles de développer des outils d'information.

Plusieurs régulateurs français mettent en œuvre cette approche au sein de leurs secteurs respectifs depuis plusieurs années suivant des objectifs qui leurs sont propres. Ces objectifs relèvent, le plus souvent, des catégories suivantes.

## A. Grands objectifs de la régulation par la donnée

Plusieurs objectifs peuvent être identifiés au développement d'une démarche de régulation par la donnée. Listés ci-dessous, ces objectifs sont parfois menés concomitamment par les régulateurs. Parfois, a contrario, l'un d'eux est développé en premier avant que les autres soient abordés.

- *Amplifier la capacité d'action du régulateur sur son cœur de métier, notamment par une meilleure supervision des acteurs et une collecte de données élargie*

La régulation par la donnée permet avant tout de renforcer la capacité d'action du régulateur sur son cœur de métier en augmentant sa capacité à exercer un suivi fin des acteurs et de leurs pratiques, en lui permettant de détecter plus aisément des signaux faibles, de suivre de manière plus fluide les évolutions du marché ou en accélérant sa capacité de réaction. Par ailleurs, en analysant finement les données unitaires des opérateurs, il est possible de contrôler les effets et la cohérence des choix des acteurs régulés.

Ainsi, par exemple, au-delà de la surveillance de marché, les données collectées par l'AMF servent aussi à la supervision des acteurs régulés par l'AMF. Depuis mai 2018, l'AMF a créé une équipe dédiée à l'exploitation de ces données à des fins de supervision des entreprises d'investissement (*data-driven supervision*). Cette équipe, outre le suivi de la bonne qualité des déclarations qui alimentent l'ensemble des directions de l'AMF, va développer des outils *ad-hoc* permettant de détecter des anomalies, et d'aider les superviseurs dans l'analyse de l'activité des entités dans leur portefeuille. Au-delà des données quotidiennes reçues par le régulateur, ces outils de contrôle ont vocation à intégrer également des données issues de *reportings* périodiques (mensuels ou annuels) fournis par les entités, ainsi que des données externes pertinentes. Cette supervision, l'AMF l'exerce en collectant et mettant en cohérence les données issues d'acteurs divers : les régulés eux-mêmes (chaque contrepartie aux transactions), mais également les plateformes.

Au-delà, le régulateur peut parfois déporter, décentraliser une partie de sa régulation au travers notamment de mécanismes de *compliance*. Le secteur financier par exemple connaît une régulation par la *compliance* incitant les acteurs régulés à internaliser une partie de l'analyse et du contrôle du risque ainsi que du *reporting*. Par cette approche, le régulé est obligé de construire ses processus en intégrant les enjeux de la régulation et se voit responsabilisé dans sa démarche vis-à-vis du régulateur.

D'autres outils, comme le contrôle de conformité au RGPD ou le traitement des notifications de violations de données personnelles réalisés par la CNIL ou l'attribution dynamique de fréquences expérimentée par l'Arcep peuvent également être cités comme des exemples d'outils exploitant une information plus précise à des fins d'une régulation plus efficace.

- *Eclairer les choix des utilisateurs et mieux orienter le marché*

Il s'agit pour le régulateur de mettre en place une démarche d'information enrichie pour permettre à tous les utilisateurs (opérateurs économiques, consommateurs, entreprises, décideurs, collectivités etc.) de faire des choix plus éclairés. Par ces outils, le régulateur cherche à créer un « choc de transparence » et orienter en conséquence, les choix des acteurs de marché.

Pour ce qui concerne les réseaux de communications électroniques mobiles, l'Arcep, au travers de monreseauomobile.fr, apporte deux types d'information sur la performance des réseaux :

- les cartes de couverture des opérateurs, conçues à partir de simulations numériques ;
- des indicateurs de qualité de service, issus de plus d'un million de mesures réalisées annuellement par l'Arcep sur le terrain, en conditions réelles.

Pour ce qui concerne les réseaux fixes, l'Arcep prépare un observatoire complet de la couverture du territoire par les réseaux fixes toutes technologies confondues, à l'échelle de l'adresse. En attendant,

elle continue d'enrichir [cartefibre.arcep.fr](http://cartefibre.arcep.fr) qui présente déjà commune par commune<sup>2</sup> et immeuble par immeuble l'avancement des déploiements en fibre optique jusqu'à l'abonné.

Au-delà des utilisateurs eux-mêmes, la politique de régulation par la donnée doit accompagner les initiatives de réutilisation de données par des relais issus du secteur public, du monde associatif ou de la société civile. Une politique *open data* et ouverte vis-à-vis de ces acteurs est ainsi essentielle. Depuis la loi *pour une République numérique* du 7 octobre 2016, les autorités de régulation doivent, comme toutes les administrations publiques, « par défaut » ouvrir les données dont elles disposent et qui ne revêtent pas un caractère confidentiel. La plupart d'entre elles ont, pour cela, mis en place des plates-formes *open data*<sup>3</sup> et répertorié leurs données sur [data.gouv.fr](http://data.gouv.fr).

Pour l'Arcep, la publication en open data des données correspondant à ces publications sur le site [data.gouv.fr](http://data.gouv.fr), la plateforme ouverte des données publiques françaises, permet de faciliter la réutilisation des données et d'aider start-ups, collectivités ou analystes à se les réapproprier.

L'AMF a également mis en place une politique de mise à disposition des données, notamment à destination des chercheurs, sur la base de conventions établies de manière ponctuelles. Une première mise à disposition de données par l'AMF aux chercheurs est en place depuis 2013, dans le cadre d'un partenariat noué avec le CNRS (Eurofidai) et l'Institut Louis Bachelier (ILB). Il permet d'alimenter en données d'ordres de la Bourse de Paris une base de données financières européennes mise à la disposition des chercheurs. D'autres pistes sont à l'étude pour élargir l'accès aux données collectées par l'AMF. Par son Observatoire de l'épargne et son Baromètre, elle éclaire également les relais d'opinion sur les problématiques de l'épargne.

L'Arjel publie, quant à elle, des analyses trimestrielles du marché des jeux d'argent et de hasard en ligne. Ces rapports présentent l'activité trimestrielle des trois segments du marché – paris sportifs, paris hippiques et poker en ligne – à travers l'agrégation, l'analyse et la comparaison des données de supervision communiquées par l'ensemble des opérateurs. Toutefois si ces documents sont utiles aux opérateurs, leur impact sur l'orientation des choix des utilisateurs est limité. C'est dans l'optique de développer cet axe de travail, notamment, que l'ARJEL a commencé à publier, sur son site internet, des communications, pouvant être utiles aux joueurs, qui donnent une analyse juridique de certaines pratiques contestées mises en œuvre par les opérateurs.

Depuis 2005, la CRE publie tous les trimestres des observatoires sur les marchés de gros et de détail de l'électricité et du gaz naturel. Ceux-ci donnent à un lectorat principalement constitué des professionnels du secteur un panorama complet des faits marquants et des tendances constatées sur ces marchés et constituent une référence pour le secteur. Les données brutes ayant servi à la constitution des graphiques et des indicateurs sont désormais également proposées sur la page *open data* du site internet de la CRE.

L'Arafer et la CRE ont par ailleurs organisé en mars 2019 un « datathon » dans l'objectif, entre autres, de mettre en place des outils « BtoC » basés sur une réutilisation adéquate des données dont elles disposent, afin de favoriser une interactivité avec les utilisateurs finaux qui pourraient d'une part obtenir de l'information qui n'est aujourd'hui pas transmise par les opérateurs (ou pas suffisamment précise) et d'autre part, remonter des informations au régulateur (sur des dysfonctionnements, par

---

<sup>2</sup> Lorsque cela est possible, les informations permettent à présent de visualiser les zones arrière de point de mutualisation

<sup>3</sup> À titre non exhaustif : AMF, Arafer, Arcep, Autorité de la concurrence, CNIL, CRE, Hadopi.

exemple). Ce premier datathon commun « Énergie et Mobilité » a ainsi permis aux deux Autorités de :

- bénéficier d'un œil nouveau avec des capacités techniques et créatives très importantes (de fortes compétences en programmation, interface et expérience utilisateurs) pour se réappropriar la donnée<sup>4</sup> dans des outils innovants (applications web/mobiles) au service des actions des régulateurs ou au bénéfice directement des utilisateurs finaux, et ce en l'espace d'un seul week-end ;
- disposer d'un support de communication externe important, pour manifester aux parties prenantes des marchés régulés la volonté du régulateur de concourir activement à leurs bons fonctionnements ;
- repenser l'innovation par la donnée au sein des Autorités, avec notamment la volonté de poursuivre l'accompagnement des projets développés sur le week-end du datathon et potentiellement de reconduire l'évènement.

Le CSA s'inscrit également dans cette démarche de publication de données et d'études, qu'elle soit récurrente (observatoires d'usages, bilans annuels des chaînes, mises en demeure, etc.) ou ponctuelle (études thématiques ou générales). Ces publications viennent éclairer le fonctionnement du secteur mais également apporter des mesures de référence notamment quantitatives.

La CNIL mise elle aussi sur plusieurs leviers. Sa politique d'ouverture de ses données<sup>5</sup> comporte des indicateurs quant aux points de vigilance de son activité répressive et donc de protection de la vie privée. Par ailleurs, les constats qualitatifs (bonnes et mauvaises pratiques) relevés dans le cadre de son activité nourrissent l'élaboration régulière de recommandations pratiques génériques<sup>6</sup> à l'attention des consommateurs, responsables de traitements, éditeurs de solutions, etc. La CNIL entend par ailleurs renforcer l'exploitation qualitative de certains flux de données, comme les notifications de violations de données, rendues obligatoires par le RGPD dans certaines hypothèses, qui pourront permettre à terme de dresser un panorama des vulnérabilités et points de vigilance significatifs.

## *B. Élargir les informations collectées*

Le développement d'outils numériques plus performants, visant à identifier des signaux faibles et des risques systémiques, comme la transmission d'informations aux utilisateurs ou à leurs relais doit s'accompagner d'un élargissement constant des sources de données accessibles pour offrir au régulateur une approche complète. Plusieurs sources ont été identifiées.

### *1. S'ouvrir aux signalements, donner aux utilisateurs la possibilité de jouer un rôle dans la régulation*

La première remontée possible des utilisateurs est « l'enquête utilisateur ». Cette solution permet souvent de mieux identifier les enjeux que rencontrent les utilisateurs sur le secteur.

Ainsi, l'Arafer a par exemple mis en place des « enquêtes utilisateurs », afin de recueillir des informations plus qualitatives sur le profil des voyageurs et les déterminants des choix de déplacement. Deux enquêtes ont ainsi été menées en 2017 et 2018 auprès des clients des nouveaux

---

<sup>4</sup> Donnée en open data ou acquise par le régulateur et désensibilisée si nécessaire (bruitage de données soumises au secret des affaires) à l'occasion du datathon.

<sup>5</sup> Par exemple, sur les sanctions prononcées et la typologie des manquements relevés, sur les organismes contrôlés

<sup>6</sup> Par exemple, récemment, sur les assistants vocaux ou les pratiques de développement « privacy by design »

services interurbains librement organisés de transport par autocar mis en place conformément à la « loi pour la croissance, l'activité et l'égalité des chances économiques » dite loi « Macron » d'août 2015, et une enquête auprès des utilisateurs de TGV/OUIGO est menée actuellement.

L'Arcep a également eu recours à des études utilisateurs. En 2016, afin d'orienter sa démarche de régulation par les données et de prioriser les chantiers à venir, l'Arcep a souhaité évaluer les attentes des utilisateurs en matière d'information sur le secteur des télécoms et les critères de choix d'offre sur le marché, révélant alors un besoin fort d'information sur les sujets de couverture et de qualité de service. L'Arcep s'est également engagée dans une démarche d'enquête utilisateur (d'abord consommateurs, bientôt entreprises) régulière sur la satisfaction de ceux-ci sur le secteur (offres et services des opérateurs et qualité de leurs services clients) afin de remettre en perspective sa compréhension des enjeux des utilisateurs.

Les signalements, notamment portés par des utilisateurs, sont un moyen efficace de développer une vision plus précise des difficultés sur un secteur et d'alimenter un faisceau d'indices de régulations en offrant la possibilité à des acteurs utilisateurs de faire un geste citoyen.

La plateforme en ligne *J'alerte l'Arcep* permet ainsi à chaque utilisateur, qu'il soit un particulier, une entreprise ou une collectivité, d'alerter l'Arcep de dysfonctionnements rencontrés dans leurs relations avec les opérateurs fixes, mobiles, internet et postaux. Par leur signalement, les utilisateurs ont la possibilité de contribuer à leur échelle à la régulation du marché, pour inciter les opérateurs à améliorer leurs services et à développer leurs réseaux. Des conseils adaptés à leur situation leur sont également proposés.

Pour l'Arcep, les alertes recueillies permettent de se rapprocher du terrain, de suivre en temps réel les difficultés rencontrées par les utilisateurs, d'identifier les dysfonctionnements récurrents ou les pics d'alertes. L'objectif étant ainsi de gagner en efficacité en ciblant son action et en pensant des réponses systémiques améliorant le fonctionnement du secteur.

Bien que l'Arcep ne soit pas en charge du traitement des litiges individuels, *J'alerte l'Arcep* a répondu à un réel besoin des utilisateurs d'obtenir un éclairage direct sur les problèmes qu'ils rencontrent (34 000 signalements reçus en un an). Ces signalements ont permis d'alimenter le faisceau d'indices de la régulation, contribuant par exemple à la décision d'ouverture de procédures formelles.

L'Arjel a, quant à elle, mis en place une adresse de messagerie permettant aux joueurs et parieurs de remonter les dysfonctionnements qu'ils identifient. Les informations recueillies ont permis à l'ARJEL d'améliorer ses actions de contrôle de l'activité des opérateurs agréés et de lutte contre les sites illégaux et l'addiction. En premier lieu, l'ARJEL peut orienter son activité de contrôle et d'enquête en fonction des réclamations que font remonter les joueurs, surtout lorsque elles portent sur des difficultés récurrentes. Ainsi, une réclamation portant sur des faits similaires de la part de plusieurs joueurs peut conduire l'ARJEL à effectuer des recherches dans le « *frontal* » de l'opérateur concerné pour apprécier leur réalité. Dans le cas où ces contestations s'avèrent fondées, l'ARJEL est alors amenée à ouvrir une enquête administrative.

Les canaux de saisines de la CNIL sont, en complément notamment des constatations effectuées par ses contrôleurs, autant de capteurs permettant au régulateur de disposer d'informations sur les pratiques constatées et les enjeux de vie privée, afin de guider l'ensemble de ses activités de contrôle, d'accompagnement, etc. Le dispositif de remontée de signalements peut être, en tant que de besoin, concentré sur certaines thématiques ou périodes : par exemple, un formulaire dédié est mis à disposition du public en période électorale. Cette activité peut, enfin, faire l'objet de partenariats. En matière de lutte contre les messages publicitaires non sollicités, la CNIL a par exemple renforcé en 2017 son dispositif conjoint avec l'association « Signal Spam », qui centralise des signalements, permet à la CNIL de disposer d'une vision globale du phénomène, et donc de conduire des investigations et une politique répressive plus efficaces

Début 2019, afin de donner à l'ensemble des acteurs la possibilité de plus aisément saisir son comité de règlement des différends et des sanctions (CoRDiS), la CRE a décidé de modifier son règlement intérieur pour autoriser la saisine par voie électronique. À cette occasion, elle a mis en place une plateforme sur son site internet, qui, tout en veillant à réorienter l'utilisateur le cas échéant (notamment vers le Médiateur national de l'énergie), aide celui-ci à fournir les renseignements et pièces utiles, selon les différentes situations prévues par le code de l'énergie. Il s'agit d'une modernisation et d'une fluidification évidentes de la procédure de saisine, en particulier saluée par les avocats spécialisés du secteur.

Les plaintes des téléspectateurs qu'enregistre et instruit le CSA constituent des indicateurs forts, tant pour le régulateur que pour les publics, des orientations et tendances éditoriales du secteur audiovisuel. Ils sont autant de rappels aux limites fixées en matière de communication audiovisuelle et de protection des publics.

L'AMF a mis en place dès 2010 un centre d'appels pour le public qui permet de capter les tendances de commercialisation, les dysfonctionnements du monde régulé et d'identifier les arnaques financières. A partir de ces données, l'AMF agit auprès de ses régulés, exerce son pouvoir d'alerte auprès du grand public et mène des actions de prévention et enfin répond aux sollicitations de la représentation parlementaire et des medias (statistiques). Elle travaille aujourd'hui sur une application Websapp qui permettra à l'épargnant depuis son smartphone de contacter directement l'AMF pour effectuer un signalement (sortie prévue en septembre 2019)

## *2. Enrichir les décisions de collecte pour les rendre plus pertinentes*

Pour démultiplier le potentiel de la donnée, il est nécessaire de pouvoir graduellement améliorer la qualité des données dont dispose le régulateur et la fréquence de leur mise à disposition.

L'Arcep procède ainsi à un « dégroupage de la donnée » en imposant aux opérateurs de publier des données plus complètes. Ces données ne servent plus uniquement à contrôler les obligations réglementaires mais permettent notamment d'apporter aux utilisateurs une information plus proche de leur expérience d'usage. Ces données peuvent être complétées de données produites par l'Arcep et sont ensuite mises à disposition du public.

De même, dans le cadre de mécanismes d'observation de signaux faibles, le volume de données disponibles et la régularité de leur mise à disposition peuvent être fortement accentués sur le modèle, par exemple, des collectes de l'AMF qui visent à recueillir quotidiennement le détail de chaque transaction réalisée par les acteurs de la finance.

A l'Arafer, le renforcement des pouvoirs de collecte de données avec la loi « Macron » d'août 2015 a permis de mettre en place des collectes régulières de données dans les 3 secteurs (ferroviaire, autocars, autoroutes). Compte tenu des enjeux liés à l'asymétrie d'information entre les parties prenantes, la première utilisation d'intérêt général des données collectées par l'Arafer s'est matérialisée par la production de publications de référence dans les secteurs régulés pour assurer de la transparence et éclairer ainsi la décision publique, notamment dans le secteur ferroviaire. L'Arafer a également engagé dès 2017 une démarche « open-data » (référéncée sur la plateforme data.gouv.fr), en publiant les jeux de données correspondant aux publications de l'Observatoire.

## *3. Bâtir une « alliance avec la multitude<sup>7</sup> »*

---

<sup>7</sup> La « multitude » renvoie à l'idée d'« une communauté créative, reliée et mobile » formée par des individus disposant « d'une puissance de communication et de coordination sans précédent. » in N. Colin et H. Verdier, *L'âge de la multitude*, Armand Colin, 2012, p. 12. Référence y est faite à A. Negri et M. Hardt, auteurs des ouvrages *Empire* et *Multitude*.



Afin d'optimiser son action et compléter les données dont il assure lui-même la production, le régulateur peut s'engager dans une démarche de *crowdsourcing* avec une série d'acteurs tiers. L'Arcep s'est engagé dans une telle démarche avec des acteurs tels que les éditeurs d'applications, les acteurs de la protection des utilisateurs, ou encore les acteurs des secteurs des transports, de l'immobilier, du tourisme.

Pour suivre la qualité de service et la couverture des réseaux télécoms, fixes ou mobiles par exemple, l'Arcep a notamment initié une approche partenariale en vue de favoriser l'ouverture des données et des mesures toujours plus fiables et représentatives des usages des utilisateurs. Elles se concentrent aujourd'hui sur deux axes.

D'une part, l'Arcep a travaillé à une feuille de route pour Monreseaumobile dont l'objectif est d'enrichir les données publiées, et de s'ouvrir pleinement aux mesures réalisées par les territoires et au *crowdsourcing*. Pour cela l'Arcep a d'une part publié « le kit du régulateur » pour permettre aux collectivités et aux acteurs qui conduisent eux-mêmes des campagnes de mesures de qualifier, eux-mêmes, la qualité de l'expérience mobile. De plus, elle a renouvelé son appel aux acteurs de la mesure et notamment du *crowdsourcing* à s'associer à ces travaux sur la base de méthodes de mesures transparentes et pertinentes.

D'autre part, l'Arcep conduit une démarche innovante de co-construction sur les sujets de qualité de service internet. Elle a ainsi fédéré des outils de mesure, des fournisseurs d'accès à internet et des acteurs académiques afin de permettre aux outils de répondre au mieux aux besoins des consommateurs en termes d'information sur la qualité d'internet. Les premiers travaux dans le cadre de ce chantier ont permis de définir un Code de conduite des acteurs de la mesure, dont la publication est imminente et qui sera enrichi avec le temps. Ils ont par ailleurs mis sur pieds une API « carte d'identité de l'accès », qui sera à terme présente dans les boîtes des opérateurs et accessibles aux outils de mesure respectant le Code de conduite afin de leur permettre de caractériser l'environnement de l'utilisateur qui réalise le test (technologie d'accès, offre souscrite, utilisations parallèles de la connexion, etc.).

L'Arafer souligne de son côté que la connaissance des flux de déplacement pourrait être grandement améliorée par l'analyse des traces de téléphone portable ou de l'historique de géolocalisation des terminaux mobiles (dans le respect du cadre de protection des données, notamment personnelles, en vigueur). Compte tenu des coûts très conséquents d'accès à ces données détenues entre autres par les opérateurs mobiles, des solutions partenariales pourraient être envisagées, regroupant plusieurs administrations intéressées par ces mêmes informations, et pourraient permettre d'améliorer le système national d'enquête existant<sup>8</sup>.

Le CSA s'appuie également sur ce principe de participation collective à la régulation en faisant appel au volontariat de ses agents. Le Conseil a ainsi pu mener des études empiriques sur des plateformes consistant notamment à observer les mécanismes de recommandation algorithmique sur des échantillons larges d'utilisateurs et en s'appuyant sur des procédés de collecte automatisée de données en volumes massifs.

Un autre acteur particulier de la multitude est le monde académique. L'Arafer, par exemple, s'est ainsi rapprochée de la sphère académique dans la perspective de démultiplier les formes de réutilisation des données collectées, et a signé deux contrats de recherche avec des universitaires et des laboratoires reconnus<sup>9</sup>, ainsi qu'un partenariat avec un programme de master<sup>10</sup>, permettant selon des modalités différentes d'aborder des sujets variés comme l'analyse des effets sur le marché

---

<sup>8</sup> Il existe une enquête nationale transports et déplacements, réalisée tous les 10/15 ans par l'INSEE et le Ministère chargé des transports, le dernier point date de 2008, or le paysage de la mobilité a beaucoup évolué depuis 10 ans avec le développement du smartphone et des nouvelles mobilités

<sup>9</sup> Le premier avec deux chercheurs de l'université de Montpellier et de Nice, le deuxième avec le CREST, centre de recherche de l'ENSAE et de Polytechnique en économie et statistique.

<sup>10</sup> Le Msc (« Master of science ») EDCBA (« Economic decision & cost benefit analysis ») de PSE et de l'Ecole des Ponts

de la fusion d'opérateurs, l'analyse de l'effet de développement du marché des cars issus de la loi « Macron » sur les modalités de déplacement des jeunes ou encore l'analyse coût/bénéfice des petites lignes ferroviaires, en comparaison avec un basculement vers des lignes routières. Cette démarche est amenée à se poursuivre, afin de promouvoir la recherche en économie des transports au bénéfice d'une meilleure connaissance des secteurs régulés.

De la même manière, le CSA a établi un partenariat avec le monde de la recherche en innovations sociétales pour mieux cerner les usages et les modes de consommation propres à l'environnement numérique. Ce type de partenariat prend tout son sens dans la perspective de mener des travaux faisant converger analyses empiriques et modélisations théoriques.

L'appel à la multitude peut également être l'occasion pour un acteur public de perfectionner ses outils de régulation et de contribuer à la création de référentiels communs, l'Arcep a par exemple publié en open data une base de population sur les territoires outremer, construite à partir de données disponibles en open data. Ses données, utilisées par l'Arcep pour contrôler la couverture mobile des opérateurs sont en réalité utiles à nombre d'autres usages et une construction dans un cadre partenarial permettrait d'aboutir à des données plus précises et partagées entre acteurs.

Le régulateur peut également susciter la création et la maturation de « communautés » d'acteurs autour de valeurs sous-tendant sa propre régulation. Ainsi, dans son cahier « Innovation et prospective » n°6, *La Forme des choix*, la CNIL a proposé une exploration des enjeux du design dans la conception des services numériques, au prisme de la protection des données. Pour promouvoir l'émergence d'un design des interfaces plus responsable et respectueux de la vie privée, la Cnil a souhaité accompagner les professionnels de la conception de service pour construire une approche non-concurrente et open source des bonnes pratiques de design, en créant une plateforme dédiée *Données & Design*. Celle-ci doit permettre l'acculturation des grands principes du RGPD, diffuser des études de cas, et fournir un espace d'échanges entre professionnels.

L'AMF a conclu un partenariat avec l'association Signal spam pour bénéficier d'une remontée des spams signalés par les internautes dans son secteur d'activité (2000 par jour). Afin de les exploiter, elle construit un système d'intelligence artificielle. De plus en plus traversé par des sujets sociétaux, l'AMF a choisi de travailler de manière interdisciplinaire sur ces sujets : partenariat avec le laboratoire de psychologie sociale d'Aix Marseille sur les arnaques financières, partenariat avec l'université Paris Dauphine pour prévenir la mauvaise commercialisation des personnes âgées. Elle intègre par ailleurs les apports de la finance comportementale en déployant des tests consommateurs sur ses messages et documents réglementaires (ex : lettre au porteur pour la gestion d'actifs).

### *C. Défis et perspectives*

Le développement de la régulation par la donnée engendre, pour les régulateurs de nouveaux besoins en termes de compétence, d'échange de données ou encore d'appropriation des nouvelles technologies.

#### *1. Développer les compétences techniques des régulateurs*

L'accompagnement du développement de la régulation par la donnée engendre, pour les régulateurs, de nouveaux besoins en compétence technique, notamment en matière d'analyse de données et d'algorithmes mais également de stockage et gestion de gros volumes de données. D'une manière générale, il est indispensable que les actions du régulateur puissent s'appuyer sur une expertise technique précise et étayée.

Or, les volumes de données à analyser sont croissants. Il est donc nécessaire que les régulateurs puissent donc développer des compétences en matière :

- d'hébergement/stockage/gestion/échange/mise à disposition des données,

- d'analyse de données, notamment par le recrutement de quelques profils pointus tels que des *data engineer* et des *data scientists*.

L'Arcep conduit à titre d'exemple actuellement des travaux de modernisation de ses outils cartographiques et la disponibilité de telles compétences et de compétences en matière de développement est essentielle. Par ailleurs, les régulateurs doivent disposer des compétences nécessaires en matière d'analyse et de compréhension de fonctionnement des algorithmes qui sont au cœur de nombreuses technologies et modèles économiques. Au-delà, c'est vers le développement d'équipes d'un nouveau genre que les besoins vont, mi-développeurs informatiques / mi-statisticiens pour analyser des jeux de données mais également analyser les algorithmes des fournisseurs de service utilisant les données.

S'il est bien entendu possible aux régulateurs de recourir à une sous-traitance de certains travaux, le maintien d'un niveau de compétence élevé en propre est indispensable pour piloter de tels travaux. Il est à noter que ces compétences sont rares et particulièrement recherchées, et il est donc nécessaire de doter les régulateurs des moyens (y compris financiers) de réaliser de tels recrutements.

Partageant ces constats, le CSA s'est progressivement doté au cours des deux dernières années de compétences en matière d'analyses quantitatives. Il les mobilise sur des questions propres au secteur traditionnel des médias mais également sur des enjeux tenant à sa transformation numérique et à la place occupée par les acteurs d'internet. Ces sujets réclament un recours croissant à la data science pour être abordés efficacement.

Cette tendance à l'accroissement des besoins en compétence technique sur l'analyse des données et les algorithmes ne devrait pas s'inverser à l'avenir, de nouveaux besoins étant susceptibles d'apparaître avec la diversification des modes de régulation (développement, algorithmique etc.). Par ailleurs, ces besoins en compétences s'accompagnent de besoins en termes de moyens logistiques (serveurs etc.) mais également de compétences dans les équipes support (administration de la donnée, maintenance applicative et systèmes d'informations).

## *2. Favoriser la disponibilité d'outils de traitement de la donnée performants pour les régulateurs*

L'accès à des outils performants est indispensable pour permettre aux régulateurs d'exploiter pleinement le potentiel des données. Le partage d'outils ou de briques technologiques communes, voire le co-développement d'outils communs pourrait être encouragé.

## *3. Favoriser l'exploitation des nouvelles technologies de pointe, notamment le big data et l'intelligence artificielle*

Les nouveaux outils numériques associés à des technologies telles que l'intelligence artificielle ou encore telles que le big data offrent des perspectives pour rendre plus efficaces les vérifications, les traitements de données et plus généralement le contrôle du régulateur (gains de temps, traitement plus exhaustif des données, gains en précision, etc.).

Ainsi, par exemple, un centre d'excellence Data-Science / Intelligence Artificielle (Datalab) a été créé en 2018 au sein de l'AMF. Son objectif est de répondre à des problématiques métiers identifiées et de développer des solutions d'améliorations à partir des données. Les premiers projets concernent ainsi le suivi du non régulé : l'objectif est de combiner des méthodes d'apprentissages automatiques (machine learning) à des techniques propres au traitement du langage (Natural Language Processing, NLP) afin de proposer une classification binaire des informations reçues (campagnes de mailing) et de détecter plus facilement des tendances, et de possibles arnaques de grande ampleur. Les thèmes émergents sont identifiés en croisant les différents flux d'informations disponibles, puis en appliquant des méthodes d'intelligence artificielle pour repérer des signaux faibles. D'autres applications sont en cours en matière d'enquêtes sur des manquements d'initiés afin d'améliorer le

ciblage des suspects, entre autres. Le Datalab a par ailleurs noué un partenariat avec un laboratoire de recherche et d'autres partenariats seront explorés dans le futur (DataLab de la Banque de France notamment).

La capacité à traiter ces sources de données multiples pourrait ouvrir de nouveaux champs pour les régulateurs en particulier en matière de détection des infractions.