Gouvernance et régulation des données

Synthèse de conférence

Conseil d'État, 23 octobre 2020

Dauphine | PSL
CHAIRE GOUVERNANCE ET RÉGULATION

Colloque organisé en partenariat
avec le Conseil d'Etat

Synthèse n°57
23 octobre 2020
Imprimé en France
Université Paris Dauphine-PSL
Décembre 2020

Table des matières

Bruno Lasserre
Vice-président du Conseil d'État

Alain Fuchs
Président de l'Université PSL

Sylvie Hubac
Présidente de la section de l'intérieur du Conseil d'État

Anton'Maria Battesti
Responsable des affaires publiques de Facebook France

Marie-Laure Denis
Présidente de la Commission nationale de l'informatique et des libertés

Roch-Olivier Maistre
Président du Conseil supérieur de l'audiovisuel

Thierry Vedel
Chercheur au CEVIPOF, Sciences Po & CNRS

Éric Brousseau
Professeur à l'université Paris Dauphine PSL, directeur scientifique de la Chaire Gouvernance et Régulation

Jamal Atif
Professeur à l'université Paris Dauphine PSL, directeur scientifique adjoint de l'institut PRAIRIE

Joris Plingers
Délégué à la protection des données, Cour de justice de l'Union européenne

Alexandre Lallet
Rapporteur public à la section du contentieux du Conseil d'État

Winston Maxwell
Directeur d'études, droit et numérique, Télécom Paris

Werner Stengg
Expert au cabinet de Margrethe Vestager, vice-présidente exécutive de la Commission européenne, pour une Europe adaptée à l'ère du numérique

Joëlle Toledano
Professeur émérite d'économie, associée à la Chaire Gouvernance et Régulation

Dominique Jamme
Directeur général des services de la Commission de régulation de l'énergie

Jean-Yves Ollier
Conseiller d'État

François Lhemery
Vice-président des affaires réglementaires de Criteo

Laurent Lafaye
Co-fondateur et directeur général de Dawex

Jean-Denis Combrexelle
Président de la section du contentieux du Conseil d'État

Miguel Amaral
Économiste principal, OCDE

Henri Isaac
Maître de conférences à l'université Paris Dauphine-PSL, membre du Conseil national du numérique

Yves-Alexandre de Montjoye
Professeur associé, Imperial College (Londres)

Fabienne Siredey-Garnier
Vice-présidente de l'Autorité de la concurrence

Cédric O
Secrétaire d'État chargé de la transition numérique et des communications électroniques

Gouvernance et régulation des données

Les données revêtent une importance de plus en plus capitale dans de nombreux domaines. Leur collecte et leur exploitation sont la clé de voûte de la création de valeur dans l'économie du numérique. Méthode de production et de conservation des données, règles de partage et d'accès, élaboration de principes guidant leur traitement, création de régimes d'appropriation et de partage des fruits de leur exploitation : des efforts de clarification, de concertation et de normalisation semblent plus que jamais nécessaires. La question de l'éventuelle régulation est également posée. Les regards croisés d'universitaires, de praticiens du droit, d'acteurs économiques et de régulateurs apporteront un éclairage utile sur la problématique globale de la gouvernance des données.

Introduction

Bruno Lasserre
Vice-président du Conseil d'État

Les données bouleversent l'économie et, plus largement, les démocraties à l'heure où l'information et la fabrique de l'opinion s'émancipent de leurs canaux traditionnels pour se retrouver entre les mains d'acteurs, essentiellement privés, qui contrôlent leur collecte et leur exploitation. La gouvernance et la régulation des données représentent, à cet égard, un enjeu démocratique majeur, et les États n'ont pas d'autre choix que trouver les moyens d'encadrer la production et l'utilisation de ces données, sous peine de devenir – au mieux – captifs de ceux qui les auront devancés.

Les données bouleversent également les politiques publiques, dont l'efficacité dépend de plus en plus de la quantité et de la qualité des données dont disposent leurs auteurs. Aussi les pouvoirs publics ont-ils commencé à définir une gouvernance afin de mieux valoriser les données qu'ils produisent et détiennent, au service de leurs missions d'intérêt général. Cette gouvernance en construction s'est déjà traduite par la création d'un véritable service public de la donnée.

Les risques que comporte l'avènement de la donnée concernent aussi les juges. Il s'agit notamment de concilier les intérêts indéniables de l'open data des décisions de justice avec la nécessité de construire et maintenir une jurisprudence évolutive et, surtout, hiérarchisée. Cela passe avant tout par la maîtrise des algorithmes, lesquels bousculent dans leurs habitudes la juridiction administrative comme les administrations. Aussi le Conseil d'État réfléchit-il à la manière d'adapter son contrôle de légalité pour appréhender l'usage des algorithmes dans la décision publique. Très prochainement, le juge devra être en mesure d'identifier les biais que ces algorithmes véhiculent et de les confronter au droit – notamment les principes d'égalité et de non-discrimination. Un usage renouvelé du contradictoire s'avérera sans doute nécessaire. Enfin, les juridictions devront peut-être se doter d'une expertise scientifique et technologique propre pour préserver leur indépendance et leur impartialité. Les enjeux sont immenses !

Le Département de justice américain vient d'annoncer le lancement d'une procédure antitrust contre Google, faisant écho aux poursuites diligentées par la Commission européenne et plusieurs autorités nationales depuis 2010. Il reproche à l'entreprise d'avoir abusé de sa position dominante sur le marché des moteurs de recherche. Ces poursuites, inédites outre-Atlantique à l'encontre d'une entreprise qui incarne à elle seule les promesses de la donnée, mais aussi ses risques, pourraient remodeler de manière très significative l'usage d'internet et être suivies d'une cascade de procédures visant les autres géants du secteur. Elles pourraient aussi conduire à un ajustement décisif du droit de la concurrence à l'économie numérique. Quid de la notion de marché dans des champs où les frontières – poreuses – évoluent à grande vitesse ? Comment appliquer et réévaluer le prisme du bien-être du consommateur, qui est la boussole du droit de la concurrence, dans des marchés bifaces où les produits sont offerts gratuitement et où la protection des données personnelles n'est pas toujours ressentie, par le consommateur, comme un élément intrinsèque de la qualité du service rendu ? Comment définir un droit de la concurrence qui limite les barrières à l'entrée tout en favorisant l'accès, la portabilité ou l'interopérabilité des données collectées par les opérateurs dominants, sans freiner l'investissement et l'innovation dans la production et la valorisation des données ? Faut-il un droit des concentrations et des règles de preuves spécifiques ?

La Commission européenne s'est activement engagée dans le chantier de l'adaptation du droit de la concurrence et de son cadre réglementaire aux évolutions du numérique. De nombreux avis ont été recueillis dans le cadre de la confection d'un livre blanc sur l'intelligence artificielle et une large consultation est encore en cours sur sa stratégie pour les données. Il s'agit de créer un environnement juridique propice à l'innovation et à la croissance, car l'Europe doit devenir plus attractive sur les marchés des données et rattraper le retard accumulé vis-à-vis des États-Unis et de la Chine.

Pour cela, elle a besoin d'investissements, de formations et surtout de règles – étant précisé qu'en la matière, les règles générales ne sauraient suffire. Qui recueille les données ? Quel type de données recueille-t-on ? Quelles sont les garanties nécessaires ? Etant entendu que les spécificités des marchés concernés appellent souvent des réponses particulières et, le cas échéant, une régulation sectorielle propre. Par exemple, la sensibilité de certaines données, notamment en matière de santé, impose un équilibre plus protecteur qu'en matière de commerce électronique. Dans les secteurs du transport, ce sont les enjeux d'interopérabilité qui sont prépondérants. Cette variété des réponses complique la donne et implique que nous nous interrogions sur l'articulation entre principes généraux et règles spécifiques, mais aussi sur l'architecture institutionnelle la mieux à même de la mettre en œuvre efficacement.

Quoi qu'il en soit, le marché n'est pas tout. D'où l'importance d'aborder la question des données aussi sous l'angle des droits et libertés fondamentaux. De fait, l'économie numérique a besoin de la confiance des citoyens et des consommateurs pour fonctionner. Dans sa jurisprudence récente Schrems-2, la Cour de justice de l'Union européenne a validé le dispositif de *privacy shield* encadrant le transfert de données personnelles vers les États-Unis. Celui-ci pourrait d'ailleurs servir de modèle en cas de Brexit sans accord, lorsque le Royaume-Uni, État européen sans doute le plus actif et le plus puissant dans l'économie des données, se retrouverait dans la situation de pays tiers. L'arrêt La Quadrature du net rendu le 6 octobre dernier est tout aussi important en ce qu'il porte sur la question du stockage des métadonnées. Sauf exceptions en matière de sécurité nationale et de lutte contre le terrorisme, le cadre juridique qui en découle apparaît restrictif. L'accès en flux est-il tout ce qui reste pour les poursuites administratives et judiciaires ? Quid de l'effectivité des enquêtes et de l'intérêt général qu'il y a à réprimer les infractions de toutes sortes ?

Sur tous ces sujets, le Conseil d'État entend jouer un rôle de premier plan. Il vient ainsi d'examiner en commission permanente le projet de loi en prorogation de l'état d'urgence sanitaire, dont certaines dispositions concernent la collecte et le stockage des données liées à la crise. Au contentieux, la jurisprudence du Conseil d'État est déjà riche et a vocation à s'étoffer toujours plus en matière de numérique. La série d'arrêts du 6 décembre 2019, par exemple, précise le mode d'emploi du droit au déréférencement. Par sa décision du 19 juin 2020, le juge a validé la sanction de 50 millions d'euros prononcée par la Cnil à l'encontre de Google, à qui il était reproché de ne pas permettre aux utilisateurs de son système d'exploitation de donner un consentement libre et éclairé au traitement de leurs données personnelles aux fins de personnalisation des annonces publicitaires. Peut également être mentionnée l'ordonnance du juge des référés du 26 juin dernier, imposant à la commune de Lisses de cesser l'usage des caméras thermiques déployées dans une école pour lutter contre la prorogation du coronavirus au motif qu'un fonctionnement continu et sans information revient de facto à collecter des données de santé sans le consentement des personnes concernées.

Quant à l'ordonnance du juge des référés rendue à propos de la plateforme *Health Data Hub*, il s'agissait de savoir si l'architecture et les contrats passés, notamment avec Microsoft, permettaient de parer efficacement aux risques de transferts des données vers les États-Unis. Le juge du référé liberté a considéré qu'en l'absence de tels transferts, il n'y avait pas d'atteinte grave et manifestement illégale à la liberté fondamentale qu'est le droit à la protection des données. Il a cependant, en lien avec la Cnil, adressé des injonctions pour préciser et sécuriser le contrat et prévenir efficacement ce risque.

Le Conseil d'État s'intéresse enfin aussi aux enjeux du numérique dans le cadre de sa fonction d'études. Il est ainsi à l'origine de trois rapports pionniers – en 1998 sur internet et les réseaux numériques, en 2014 sur le numérique et les droits fondamentaux et en 2017, sur les plateformes numériques et l'uberisation. Ces rapports étaient nourris de l'expérience accumulée par le Conseil d'Etat dans le cadre de ses fonctions consultatives et contentieuses et les ont nourries en retour. Et cette réflexion reste en cours même lorsqu'un rapport n'est pas en préparation.

Alain Fuchs
Président de l'université PSL

Les données sont souvent présentées comme la clé de la création de valeurs et porteuses de nouvelles dynamiques de croissance. Elles permettent de mobiliser les ressources économiques disponibles, de renforcer la capacité à produire des services sur-mesure plus efficaces et d'explorer de manière systématique les opportunités d'innovation. Néanmoins, la réalisation concrète de ces promesses impose la mise en œuvre de cadres institutionnels appropriés. En effet, la donnée n'existe pas à l'état naturel. Elle n'est engendrée par les utilisateurs ou les opérateurs de l'économie qu'en fonction des usages et des appropriations de valeurs dont ils peuvent bénéficier : pourquoi accepter de révéler le détail de ses usages à travers un terminal intelligent si aucune amélioration de service ou baisse des coûts n'est proposée en contrepartie ?

Par ailleurs, la donnée n'a de valeur que dans la mesure où elle peut être qualifiée, authentifiée et interopérable avec des données complémentaires. Son exploitation effective requiert des compétences rares et des investissements considérables, pour engendrer des algorithmes mais aussi pour adapter les infrastructures, les processus et les organisations. Elle suppose des efforts de clarification, de concertation, de normalisation, de régulation des méthodes de production et de conservation, des règles de partage et d'accès, mais aussi l'élaboration des principes guidant le traitement de création des régimes d'appropriation et le partage des fruits de l'exploitation.

Les données personnelles constituent une dimension essentielle de la problématique. D'une part, elles sont une source de création de valeur potentiellement très importante. D'autre part, elles permettent aux opérateurs économiques de s'approprier une large partie de la valeur créée, en particulier parce que les services personnalisés deviennent difficilement comparables en matière de prix ou de qualité. Ni les utilisateurs ni les autorités en charge de la concurrence ne disposent de références pour s'assurer que les conditions de service sont loyales. Cependant, il convient de se pencher également sur les données non personnelles, second moteur essentiel de la transformation numérique en ce qu'elles permettent une analyse systématique du fonctionnement des organisations, des dynamiques collectives et des mécanismes de transmission dans les systèmes complexes. Elles ouvrent la voie des décisions fondées sur une évaluation précise des performances et de leur explication –*evidence-based strategies* ou *policies*.

Les régimes d'accès et d'utilisation des données sont susceptibles d'avoir des effets très importants sur la concurrence, du fait des avantages compétitifs irréversibles bénéficiant au primo-accédant. En effet, les données sont le fondement de l'apprentissage algorithmique, à l'origine de nombreuses dimensions de la disruption numérique. Elles peuvent aussi avoir d'importants effets distributifs, puisque leur manipulation permet de jouer sur la répartition des gains de productivité et de la transition numérique. Dans la lignée des travaux menés sur le numérique, tant au Conseil d'État que dans le cadre de la chaire Gouvernance et régulation, il importe de discuter des initiatives qu'il serait souhaitable de prendre au plan national et européen, en tenant compte du jeu des acteurs privés et du caractère *a-territorial* de nombreuses dimensions de la gestion des données et des algorithmes.

Cette journée propose d'opérer un effort de réflexion sur ces pratiques émergentes, pour les faire dialoguer avec les grands principes juridiques dont le Conseil d'État est, entre autres, le gardien. Il s'agit aussi de permettre aux chercheurs impliqués dans le développement de l'intelligence artificielle de porter un effort réflexif sur les enjeux de leurs travaux en matière de gouvernance collective.

À la croisée des décisions publiques et des développements techniques, l'université PSL – dont Dauphine est l'une des composantes – dispose également, à travers ses chaires partenariales et ses instituts thématiques, des plateformes permettant de réaliser ces interactions entre des mondes qui ne se parlent pas toujours. Cela souligne combien l'université, et singulièrement les nouvelles universités de recherche comme PSL peuvent et doivent être des outils au service de l'intérêt général et du pilotage des collectivités face aux grands défis que sont le numérique, le changement climatique ou encore la cohésion économique et sociale.

Table ronde n°1
La protection contre les manipulations individuelles

Modératrice : Sylvie Hubac
Présidente de la section de l'intérieur du Conseil d'État

Cette table ronde ouvre le colloque sur un sujet complexe, multiforme et grave : la protection contre la manipulation concerne en effet autant la vie privée et l'intime que les fondements de la société démocratique.

Sur le sujet s'est opérée, à la suite d'évènements récents et récurrents sur lesquels nous aurons l'occasion de revenir, une prise de conscience des opérateurs et du grand public sur certaines des conséquences inquiétantes que porte en elle-même l'économie des données personnelles.

L'extraction et l'exploitation de données sur les comportements individuels des internautes et leur revente à des tiers constituent la clé de voûte de l'économie du numérique. Ce que nous lisons en ligne, ce que nous achetons, là où nous aimons aller chercher des nouvelles ou de la musique, comment nous naviguons sur la Toile, constituent un ensemble d'informations collectées à tout instant.

Ces informations sur qui nous sommes, que nous fournissons de manière active et apparemment de notre plein gré, sont brassées, travaillées par des algorithmes et croisées avec d'autres données comportementales observées. Ce travail de l'ombre, le plus souvent invisible, permet de dessiner nos profils d'utilisateurs de manière de plus en plus fine et d'attribuer à la donnée sa grande valeur. On sait que cette sophistication du profilage est liée au basculement vers la publicité personnalisée il y a déjà plus d'une décennie.

L'exploitation du profilage a des aspects évidemment positifs ; il améliore le service rendu à l'utilisateur, facilite ses choix, ou lui permet de se rapprocher de groupes avec lesquels il a des affinités. Mais elle sert aussi à peser sur les comportements et conduire la personne ciblée vers des prises de décisions qu'elle n'a pas nécessairement souhaitées affectant ainsi le libre arbitre et l'autonomie de la volonté, même s'il est vrai qu'entre influencer et manipuler la frontière n'est pas toujours évidente à tracer.

Il existe en revanche de véritables et intentionnelles manipulations inspirées par le but de nuire ou de tromper délibérément en ayant recours à diverses techniques, parmi lesquelles la diffusion d'informations fausses ou tronquées ou le vol de données ensuite réutilisées à des fins malveillantes ou illicites. Les fonctionnalités affinitaires et la viralité des réseaux sociaux peuvent également être exploitées pour provoquer des comportements favorables au manipulateur. On pense ici notamment au scandale Cambridge Analytica ou aux faux comptes russes pendant la campagne présidentielle américaine de 2016.

Pour combattre ces dérives, des instruments de protection ont été mis en place ces dernières années, comme la directive sur la protection des données de 2005, le RGPD de 2016, les lois de décembre 2018 sur la lutte contre la manipulation de l'information, la loi de juin 2020 sur la lutte contre les contenus haineux – ou ce qu'il en reste – ou encore les recommandations et la régulation de la Cnil et du CSA.

La protection des données personnelles repose sur les principes fondamentaux de licéité, loyauté, transparence, minimisation, sécurité et consentement de la personne à l'utilisation de ses données. Elle appelle toutefois plusieurs interrogations. Les droits qui s'y attachent sont-ils bien connus et réellement exercés ? Le consentement peut-il vraiment être exprimé ? N'est-il pas souvent davantage abandonné que donné ? Les interfaces sont-elles toujours loyales dans l'organisation du recueil de ce consentement ? S'agissant de la transparence des méthodes et du fonctionnement des marchés des données, des exigences nouvelles ne doivent-elles pas être formulées, à quel niveau et avec quel type d'encadrement ? Si les plateformes se sont engagées à mettre en place des mécanismes de contrôle pour détecter, prévenir et neutraliser les manipulations, quel bilan peut-on tirer de cette autorégulation ? Les plateformes sont-elles légitimes à assurer seules la conciliation entre liberté d'expression et protection des publics ? Peut-on envisager une coopération loyale de tous les acteurs, et comment ? C'est à ces questions que la table ronde s'efforcera de répondre.

Facebook et la lutte contre les manipulations

Anton'Maria Battesti
Responsable des affaires publiques de Facebook France

Un des reproches le plus souvent adressés à Facebook est celui du temps passé. En l'occurrence, la plateforme vise surtout à développer du temps utile. Des modifications algorithmiques sont donc régulièrement apportées pour proposer un produit plus personnalisé et en phase avec les attentes des utilisateurs.

Autre reproche, « si c'est gratuit, c'est vous le produit ». Cette phrase choc a le mérite de poser le débat, mais il importe de savoir que Facebook ne vend pas de données personnelles aux annonceurs – un business model autour de la publicité et de l'expérience publicitaire ne saurait se construire ainsi –, mais la capacité à accéder aux audiences très affinées que nous construisons. La différence est fondamentale, même si elle n'est pas toujours bien comprise.

La question des algorithmes est importante également. Le fil d'actualité Facebook peut être comparé à une route, avec différentes étapes. L'algorithme n'a rien de malicieux ou de gentil : il est conçu pour créer des arbitrages entre ce qui y est présenté ou pas. Dans le cadre d'une expérience personnalisée, ce que vous

regardez ou commentez le plus a tendance à se retrouver plus régulièrement dans votre fil d'actualité. De manière très imparfaite, l'expérience numérique reflète l'expérience sociale : depuis quand n'a-t-on pas pris directement des nouvelles de cet ami qui n'apparaît plus dans notre fil d'actualité Facebook ? La question des manipulations se situe aux confluents de tous ces éléments.

L'outil publicitaire est un levier marketing très puissant, tant pour les grandes marques que pour les entreprises qui ont besoin de se développer. Les plateformes permettent d'y avoir accès à moindre coût. Cet accès est également ouvert à d'organisations activistes, qui utilisent ces campagnes pour promouvoir leurs idées. Malheureusement, certains acteurs malfaisants ont compris qu'ils pouvaient communiquer de façon plus ciblée qu'en distribuant des tracts, par exemple. Pour éviter les dérives, deux fondamentaux doivent être mis en œuvre : la transparence et la responsabilité. S'agissant de la transparence, le consentement est un outil légal très puissant à condition d'être bien compris. Aussi Facebook a-t-il créé des « chemins de consentement », ou *consent flow*. Nous proposons aussi des outils de contrôle a posteriori de l'expérience publicitaire, afin que chaque utilisateur sache pourquoi il a été ciblé, et même qui permettent de refuser d'être suivi en dehors de Facebook. En matière démocratique et électorale, qui plus est, la « bibliothèque publicitaire » est un gage de transparence. Elle recense toutes les publicités chargées de promouvoir des idées ou des causes d'intérêt général, en précisant les auteurs, les sujets, les budgets engagés, les audiences ciblées. Ces informations sont accessibles au régulateur (le CSA), aux journalistes, aux experts académiques et aux utilisateurs – qui peuvent alors demander des comptes.

De fait, la notion de responsabilité de Facebook, ou *accountability*, est au cœur des mesures prises avec le régulateur américain, outre le règlement d'une amende de 5 milliards de dollars. Entre 2014 et 2016, l'abus des faiblesses ou le manque de précision des pratiques de ciblage ont permis des dérives comme l'affaire Cambridge Analytica. Aujourd'hui, ces phénomènes sont beaucoup plus difficiles à mettre en œuvre, grâce à un meilleur encadrement, une plus grande transparence et l'accroissement des responsabilités. Par exemple, Facebook interdira la publicité politique sur sa plateforme à une semaine des élections américaines, pour éviter les manipulations de dernière minute ou la diffusion de fausses informations.

Enfin, dans quelques mois, le Digital Services Act (DSA) devrait renforcer la régulation européenne, utile et nécessaire, des réseaux sociaux et des plateformes d'expression.

Echanges

Sylvie Hubac

Il y a encore beaucoup à faire ! Êtes-vous en capacité de mieux détecter aujourd'hui qu'hier les manipulations ? Quelles sont selon vous les prochaines étapes à franchir pour parvenir ?

Anton'Maria Battesti

Tous les mois, les services Facebook/Instagram regroupent en moyenne 3 milliards d'individus. La technologie ne peut pas tout régler, et le facteur humain ne doit pas être sous-estimé. Malheureusement, la situation est celle du gendarme et du voleur : une fois qu'un front est maîtrisé (par exemple, la fermeture des applications à l'origine des dérives de Cambridge Analytica et l'engagement de poursuites), un autre s'ouvre (comme la propagation virale de *fake news* via les messageries privées). Il est indispensable de procéder de manière à la fois empirique et humble. C'est un travail que nous ne pouvons pas mener seuls. Nous devons également nous ouvrir davantage à la recherche.

Roch-Olivier Maistre

Les opinions publiques mondiales ont évolué très rapidement. Le discours sur l'espace de liberté absolue que devrait être Internet est maintenant derrière nous. En Europe, beaucoup de voix se font entendre pour appeler à davantage de régulation. Aux États-Unis, le débat est vif également, tant sous l'angle de l'antitrust que sous celui de la régulation – la section 230 ayant été remise en cause par le président. Le sujet est également soulevé sur le continent africain et en Océanie, a fortiori depuis l'épisode de Christchurch. Ce qu'on ne supporterait pas dans les médias traditionnels, précisément parce qu'ils sont régulés, on le supporte de moins en moins sur Internet.

Conscientes que leur modèle économique pouvait être remis en cause, les plateformes commencent à être dans l'initiative. Cela étant, elles sont dans des situations très diverses. Facebook a une capacité économique très puissante et dispose de nombreuses ressources pour développer nombre d'initiatives de modération humaine ou par intelligence artificielle. D'autres ne sont pas dans la même configuration. Du fait de ces différences de situation, l'autorégulation ne peut pas être la seule réponse. L'Europe a certes joué cette carte, avec les codes de bonne conduite, mais elle en a constaté les limites. Il faut donc inventer un autre modèle de régulation que celui qui s'applique aux médias traditionnels.

L'apport du droit à la protection des données

Marie-Laure Denis
Présidente de la Commission nationale de l'informatique et des libertés

Une condition nécessaire pour combattre les manipulations consiste tout d'abord à en comprendre les ressorts. En l'occurrence, trois logiques sont à l'œuvre qui, combinées, rendent possible pour les acteurs du numérique d'influencer les comportements individuels : la dissémination de données dans l'espace numérique par les personnes elles-mêmes, de façon plus ou moins consciente ; le chaînage et le décloisonnement des différents comportements numériques,

autrefois compartimentés ; la capacité technologique inédite d'exploitation, de combinaison et d'enrichissement des données.

Face à ces menaces, le RGPD a pour objectif de fournir aux individus des outils de protection et de contrôle de leurs données, d'une part en leur conférant de nouveaux droits, comme celui à la portabilité, d'autre part en renforçant deux exigences : le consentement au traitement de certaines données et la transparence. La Cnil a publié le 1er octobre dernier des lignes directrices et des recommandations en matière de ciblage publicitaire et d'acceptation des cookies qui visent à concrétiser ces exigences dans ce domaine de la vie quotidienne, en garantissant la liberté de choix des internautes.

Les individus se saisissent de ces outils de protection offerts par le RGPD. Ils se tournent davantage vers la Cnil, ce qui démontre la sensibilité du grand public à la protection de la vie privée : le nombre des plaintes à la Cnil est ainsi passé, environ, de 8 000 à 14 000 en trois ans. Autre exemple : la Cnil a prononcé une sanction contre Google à hauteur de 50 millions d'euros, en janvier 2019, à la suite d'une plainte collective, nouveau dispositif mis à disposition des citoyens par le RGPD. Ce texte permet en effet aux autorités de régulation de l'Union européenne de prononcer des sanctions plus élevées qu'auparavant, jusqu'à 4 % du chiffre d'affaires mondial d'une entreprise. Mais plus encore que le montant des amendes, le risque de réputation lié à la publicité des sanctions et la capacité d'injonction sont des outils efficaces.

L'encadrement des traitements sensibles – ceux qui portent sur des données de santé ou biométriques notamment, mais également les traitements massifs, invasifs ou reposant sur le profilage des individus – répond lui aussi à l'exigence de renforcement de la protection des données et de lutte contre les manipulations individuelles. Il impose par exemple des obligations fortes de sécurité et des analyses d'impact : l'ICO, l'homologue anglais de la Cnil, vient de prononcer, dans le cadre de la coopération européenne, une sanction de vingt millions de livres à l'encontre de British Airways pour avoir rendu accessibles les données bancaires de 200 000 personnes, dont le traitement peut affecter directement la situation des personnes.

Le droit à la protection des données présente toutefois des limites. Trois d'entre elles peuvent être mentionnées, qui sont aussi parfois des forces. La première tient à la plasticité des règles établies par le RGPD : ses principes généraux doivent être appliqués selon une approche casuistique, ce qui garantit une certaine souplesse mais s'oppose à des interdictions générales et absolues. La deuxième vient du fait que si le bon niveau de gouvernance est nécessairement européen, la réactivité des autorités de protection des données peut s'en trouver affectée, d'autant plus que cela nécessite de combiner des approches juridiques nationales différentes. Enfin, la troisième limite intrinsèque tient au champ de la législation en matière de protection des données : la lutte contre les manipulations individuelles ou collectives ne peut passer par ce seul prisme et nécessite une inter-régulation effective des acteurs du numérique. La clé est la convergence des régulations plutôt que le suivi de sillons régulatoires qui s'ignoreraient.

Echanges

Sylvie Hubac

Sur quoi portent les plaintes à la Cnil ? Certaines questions de protection sont-elles davantage mises en avant que d'autres ?

Marie-Laure Denis

Les deux premiers sujets concernent le marketing et le démarchage commercial abusif d'une part, la surveillance au travail d'autre part. Mais la Cnil intervient également dans de nombreux autres domaines de la vie quotidienne : suppression de contenus sur internet et déréférencement des moteurs de recherche, fichage bancaire, surveillance en ligne des examens, etc. Globalement, les personnes sont sensibles à leurs droits : les difficultés liées à l'exercice de leur droit d'accès aux données qui les concernent ou la possibilité de modifier ces données lorsqu'elles sont inexactes constituent les principaux motifs des plaintes adressées à la Cnil.

Le point de vue du régulateur sectoriel

Roch-Olivier Maistre
Président du Conseil supérieur de l'audiovisuel

Le régulateur des médias audiovisuels est en pleine transformation, puisque ses compétences s'élargissent progressivement à de nouveaux acteurs, en particulier les plateformes de contenus et les réseaux sociaux.

S'agissant des données, la position du CSA est ambivalente. Si l'une de ses missions historiques est la protection des publics, notamment les plus jeunes, force est de constater que l'utilisation des données représente une opportunité croissante pour les médias audiovisuels, dont le modèle économique est appelé à se transformer sous l'effet de la transition numérique. La publicité segmentée à la télévision vient d'ailleurs d'être facilitée par un récent décret du 5 août 2020. Les ressources qui en découlent pourraient permettre aux chaînes de repenser leur modèle publicitaire, mais aussi de lever des asymétries concurrentielles très fortes : les régies publicitaires de ces médias pourront offrir aux annonceurs ce que seule la publicité digitale permettait jusqu'ici, le gain potentiel étant estimé autour de 200 millions d'euros. La difficulté consistait à trouver un équilibre entre deux impératifs contradictoires : répondre à la forte demande des médias audiovisuels d'avoir accès à cette opportunité économique nouvelle, et assurer une protection des données. C'est pour cette raison que le décret fixe notamment une limite horaire, une obligation d'identification du message et une interdiction autour des programmes pour enfants.

D'autres dispositifs de régulation sectorielle peuvent être imaginés pour lutter contre les diverses manipulations des données. La France a légiféré contre la manipulation de l'information avec la loi du 22 décembre 2018, qui met à la charge

des opérateurs un devoir de coopération avec le régulateur, lequel a pour mission de superviser les obligations de moyens imposées aux plateformes et d'en rendre compte par un rapport public. Cette approche diffère de la régulation des médias audiovisuels, car elle vise à responsabiliser ces nouveaux opérateurs, en leur imposant des obligations de moyens et de transparence. Le premier bilan, en ligne sur le site du CSA, montre que les plateformes se sont mises en mouvement pour lutter contre la manipulation de l'information. Cela étant, d'importants efforts doivent encore être fournis, notamment en matière de transparence du fonctionnement des algorithmes.

S'agissant de la lutte contre les contenus haineux, la proposition de loi de la députée Laetitia Avia visait à imposer aux plateformes une obligation de retrait dans un délai de 24 heures sous peine de sanctions pénales, mais aussi des obligations de moyens supervisées par le CSA et la création d'un Observatoire de la haine en ligne. Son texte a été invalidé en quasi-totalité par le Conseil constitutionnel, à l'exception des dispositions relatives à la création de l'observatoire, lequel a été mis en place cet été par le CSA.

Pour la suite, la réponse est nécessairement européenne. C'est tout l'objet du *Digital Services Act* (DSA), en cours d'élaboration. Inventer un nouveau modèle de régulation propre à notre continent est assurément la bonne orientation. L'enjeu est de dépasser le contrôle des contenus eux-mêmes, en privilégiant une logique de supervision, par un régulateur, d'obligations de moyens et de transparence. Cette mutation suppose une évolution des compétences du régulateur, un très grand renfort de l'inter-régulation, ainsi qu'un dialogue accru avec le monde académique.

Echanges

Sylvie Hubac

Quelles ont été les initiatives prises par Facebook pour lutter contre les manipulations d'informations de santé durant la crise sanitaire ?

Anton'Maria Battesti

Outre le retrait de certains contenus, nous avons apporté notre soutien aux autorités publiques pour diffuser des informations sérieuses et crédibles. De fait, des règles très strictes nous permettent de retirer les informations sur de faux traitements ou des traitements non-validés par les autorités compétentes. Nous retirons aussi tout type de contenu visant à minimiser l'importance de la maladie ou à créer de la panique. Nous avons également travaillé avec le Service d'information du Gouvernement, pour une diffusion aussi massive que possible des messages de prévention. Ce travail se poursuit. C'est un bon modèle de coopération avec l'État.

Sylvie Hubac

Comment repérez-vous les tentatives de manipulation ?

Anton'Maria Battesti

Les informations nous parviennent par des filtres utilisant des mots-clés, par des signalements, mais aussi grâce au travail des modérateurs. Comme je le disais, nous suivons une démarche empirique et itérative.

Sylvie Hubac

Après examen de la proposition de loi Avia, le Conseil d'État est parvenu à la conclusion que le seul bon niveau pour lutter contre la haine en ligne est européen, ne serait-ce que parce que la directive e-Commerce ne reconnait pas la responsabilité des fournisseurs de services à raison des contenus transportés et ne leur impose aucune mission générale de surveillance. Elle ne permet par ailleurs pas de contraindre les plates formes opérant en France quel que soit leur lieu d'émission.

Roch-Olivier Maistre

Dans les années 2000, quand a été adoptée la directive e-Commerce, nombre d'acteurs n'existaient pas encore. La problématique était principalement de nature économique et concurrentielle. Le texte du DSA, dont un volet sera relatif à la régulation des contenus, devrait être connu le 2 décembre. Il importe que son champ soit aussi large que possible, pour couvrir les zones grises et pas seulement les contenus illicites. Ce texte devrait conserver une approche se fondant sur le pays d'implantation, plutôt que sur celui de destination, même si la charge de la mise en œuvre devrait incomber à des régulateurs nationaux coordonnés. En tout état de cause, imaginer un régulateur paneuropéen me semble une vue de l'esprit.

Les manipulations de données à des fins politiques

Thierry Vedel
Chercheur au CEVIPOF, Sciences Po & CNRS

La manipulation implique une action délibérée par des moyens détournés pour conduire une personne à faire quelque chose de précis – ce qui est particulièrement difficile dans le domaine électoral et politique.

Dans la plupart des pays démocratiques, une campagne électorale sert d'abord à mobiliser les soutiens d'un candidat. Il ne s'agit pas véritablement de convaincre, car la persuasion requiert un temps plus long. Les études montrent d'ailleurs que les transferts entre camps adverses sont très rares et que les changements d'intention se font entre l'abstention et le vote effectif, ou entre des candidats voisins. La stratégie vise donc surtout à identifier les électeurs susceptibles de voter pour son candidat et à faire en sorte qu'ils aillent aux urnes. Ces pratiques dépendent largement du contexte réglementaire. En l'occurrence, il convient de distinguer

les pays anglo-saxons et les pays plus protecteurs des données personnelles en matière électorale. En Grande-Bretagne, la pratique du porte-à-porte est très ancienne. En France, en revanche, il est interdit de dresser des listes nominatives de préférences politiques. Une autre façon de procéder consiste à faire du profilage, pour définir des catégories d'électeurs susceptibles de voter pour tel candidat ou parti. La sociologie électorale française est assez célèbre pour son approche de géographie électorale, à un niveau très fin. Les enquêtes de sondage permettent aussi de faire le lien entre les intentions de vote, les caractéristiques socio-démocratiques, les croyances, les valeurs, etc.

La grande révolution du numérique réside dans la capacité à engranger plus de données et à fusionner la sociologie et la géographie électorales, pour construire des modèles prédictifs de plus en plus précis, intégrant des variables auxquelles on n'aurait pas nécessairement pensé. Avec le profilage, les plateformes permettent de toucher des catégories de public identifiées. Certes, le commerce de l'accès à des catégories de publics est légitime. Mais il pose de nombreuses questions. Dans quelle mesure une plateforme en situation de position dominante voire de monopole peut s'arroger le droit de définir ses propres règles – interdire la publicité politique une semaine avant les élections américaines, par exemple ? Le sujet touche à la nature ambivalente des plateformes – « réseaux sociaux » en français et « *social media* » en anglais : si l'on a tendance à réguler les médias, le principe de neutralité des opérateurs de télécommunications prévaut.

Par ailleurs, les études montrent que le fait d'exposer des catégories d'électeurs à un message ciblé ne signifie pas nécessairement que ces derniers seront influencés dans leur vote. C'est même plutôt l'inverse qui se produit, car nous ne voyons que ce que nous croyons et préférons déjà.

Echanges

Sylvie Hubac

Force est de constater que les plateformes se comportent comme des États, en définissant leurs propres règles et en organisant leur propre police. Cette autorégulation est un phénomène nouveau, qui doit évoluer vers davantage de régulation par les autorités publiques. Comme le dit Thierry Breton « dans bien des cas l'univers numérique est une zone de non-droit ».

De la salle

Les manipulations publicitaires et individuelles présentent aussi des risques pour l'état physique et mental des personnes. Qui plus est, le cloisonnement nuit à l'esprit critique. La crise sanitaire ne risque-t-elle pas d'accélérer ce phénomène ? Au-delà de la régulation, comment éduquer à la data ?

Sylvie Hubac

L'enfermement par le système est un phénomène préoccupant et inattendu. Il est à l'opposé de la plus grande ouverture de chaque individu sur la diversité et la complexité du monde qui était la première promesse de l'Internet.

Marie-Laure Denis

L'éducation au numérique et le développement d'outils qui permettent l'empowerment des individus sont également des axes de travail indispensables. La Cnil porte ainsi de nombreuses actions éducatives, notamment au travers du collectif EducNum. Il est également indispensable de favoriser le dialogue entre les générations, qui ne doit pas être uniquement descendant, bien au contraire. Par ailleurs, l'outil Cookieviz est un bon exemple d'outil clé en main, mis par la Cnil à la disposition des personnes, qui permet de savoir si un site sur lequel on navigue a déposé des traceurs – utilisés aux fins de constitution de profils publicitaires qui sont ensuite revendus.

Roch-Olivier Maistre

Le risque d'enfermement est réel : on ne s'informe qu'au sein de sa propre communauté et la seule vérité est celle de sa communauté.

Thierry Vedel

L'éducation est la clé, car l'autorégulation sera très difficile à mettre en œuvre. Encore faut-il qu'elle soit bien faite et ne conduise pas au relativisme culturel.

Par ailleurs, la pratique du ciblage est, sans doute acceptable, dès lors qu'on est consentant pour recevoir tel ou tel type d'informations en relation avec ses intérêts. L'idéal serait qu'il n'y ait que des pratiques de opt-in, et pas de opt-out.

De la salle

Confier la modération des contenus à une plateforme, en particulier en période électorale, est assez choquant. Le modèle du contrôle de deuxième niveau ne présente-t-il pas une limite ?

Roch-Olivier Maistre

L'autorégulation présente des effets pervers. Il ne va pas de soi, dans une démocratie, qu'une plateforme décide *proprio motu* d'éliminer un tweet du président des États-Unis, par exemple. La proposition de loi de la députée Laetitia Avia avait aussi soulevé la question d'une forme de « privatisation » de la justice, en faisant encourir une sanction très lourde aux plateformes qui n'élimineraient pas les contenus haineux dans un délai de 24 heures après leur signalement. Le risque était que les plateformes sur-éliminent des contenus pour ne pas engager leur responsabilité, ce qui porterait atteinte à la liberté d'expression, qui est une liberté fondamentale. Il est donc indispensable que les États se saisissent de la question, que la démocratie s'exprime et qu'un cadre juridique soit posé.

De la salle

Le Parlement européen a voté en grande majorité un rapport qui demande à la Commission de se prononcer sur l'interdiction, au nom de la liberté d'expression, les outils de sélection automatique a priori des contenus – à l'exception des contenus « manifestement illicites ». Qu'en pensez-vous ?

Sylvie Hubac

Nous aborderons la question de la transparence des algorithmes dans la deuxième table ronde.

Un grand merci pour ces échanges dont je retiens que l'opinion publique est de plus en plus demandeuse de règles, que l'autorégulation présente des limites qui appelle l'élaboration d'un autre modèle, que le niveau européen s'impose comme la bonne échelle de régulation, que l'inter-régulation entre autorités administratives est indispensable en raison de la convergence des supports, et enfin que l'éducation à l'ouverture et à l'esprit critique est la clé pour éviter le repli dans la « bulle de filtres » et donner aux citoyens les moyens de naviguer en sécurité , en connaissant et en exerçant leurs droits.

Table ronde n°2
La régulation et la transparence des algorithmes

Modérateur : Éric Brousseau
Professeur à l'université Paris Dauphine-PSL, directeur scientifique de la Chaire Gouvernance et Régulation

Les données n'ont de sens que parce qu'elles sont utilisées par des algorithmes, lesquels sont eux-mêmes issus des données ; que ces dernières soient utilisées pour modéliser les problèmes de décision afin de programmer des outils d'aide ou de prise décision, ou que ces données nourrissent des processus d'apprentissage automatisé qui aboutissent, eux aussi, à des décisions en partie automatisées. Ces algorithmes facilitent l'utilisation de la profusion d'information et l'identification des informations les plus pertinentes. Selon les cas, ils assistent, suggèrent voire prennent des décisions, allant jusqu'à se substituer au libre arbitre de l'utilisateur de l'algorithme ou de celui qui lui fournit une prestation. Souvent, en effet, ces suggestions automatisées ont de tels avantages pratiques (instantanéité, réduction des efforts cognitifs, etc.) qu'on risque de ne plus interroger leur bien-fondé.

Pourtant, les risques d'erreurs et de biais sont inhérents à la décision algorithmique qui n'est fondée que sur des situations et individus statistiques, en ignorant leurs spécificités. D'où les questions relatives aux droits individuels et aux manipulations potentielles. Ces questions se posent également pour les personnes morales. Les conséquences potentielles concernent tout aussi bien les sphères sociales, politiques, économiques.

Au-delà des conséquences de la décision algorithmisée pour les individus ou les organisations, se pose le problème des interactions entre algorithmes. Ces dernières peuvent produire des effets indésirables. Plusieurs études ont montré, par exemple, que sur certains marchés les algorithmes peuvent être conduit à adopter des comportements collusifs.

Ces questions très générales doivent bien entendu être abordées en tenant compte des entités qui utilisent des algorithmes pour fonder leurs décisions. Les « plateformes », très diverses au demeurant, ne sont pas les seules à les mettre en œuvre. D'autres opérateurs économiques et les pouvoirs publics sont aussi des utilisateurs d'algorithmes et les conséquences de leurs décisions respectives peuvent avoir une portée différente sur les droits individuels et les mécanismes de régulation collective. Cette variété requiert certainement différents modèles de régulation.

Les failles potentielles des algorithmes ne doivent pas faire oublier que les processus de décision fondée sur l'humain ont aussi leur faille et que les organisations, publiques ou privées, de même que nos cadres institutionnels intègrent des mécanismes pour corriger les erreurs de diagnostic ou de jugement, l'arbitraire

des décisions, les défaillances des décideurs. L'esprit de ces mécanismes doit certainement guider la mise au point de dispositifs de vérification a posteriori des décisions prises par les algorithmes. Deux spécificités du numérique doivent, cependant, être prises en compte : l'extrême rapidité de la prise de décision et le risque du non-respect a priori de certains droits par des algorithmes trop orientés par la « performance ».

La boîte noire des algorithmes

Jamal Atif
Professeur à l'université Paris Dauphine-PSL, Directeur scientifique adjoint de l'institut PRAIRIE

Avant de parler de gouvernance des algorithmes, il convient de préciser de quels algorithmes il est question ici. Plusieurs classes d'algorithmes existent. Certains sont déterministes, d'autres stochastiques, et enfin, et c'est ceux-là qui font l'objet d'attention ces derniers temps, il me semble, il existe une classe d'algorithmes qui émergent des données ou de l'expérience. On parle d'apprentissage automatique (*Machine Learning*), ou d'Intelligence Artificielle, IA (appellation impropre puisque l'IA est un domaine plus large avec des contours mal définis). L'idée générale de l'apprentissage machine est de faire émerger à partir d'un ensemble d'observations une règle de décision qui se généralise sur des situations, justement non observées. Quand les observations sont constituées de données étiquetées, c'est à dire une donnée et son étiquette (par exemple une image et une étiquette encodant son contenu), on parle d'apprentissage supervisé, en l'absence d'étiquette, on parle d'apprentissage non supervisé, et enfin quand il s'agit d'acquérir des observations par expérience ou en agissant sur un environnement, on parle d'apprentissage par renforcement. D'autres paradigmes existent comme l'apprentissage semi-supervisé, autosupervisé, etc.

Réguler un algorithme d'apprentissage automatique suppose de vérifier sa conformité avec des règles prédéfinies par le législateur, ou d'encoder ces mêmes règles au sein de la fonction objectif de ce dernier. Dans les deux cas, il s'agira in fine d'auditer ces algorithmes pour vérifier leur conformité à la règle. Cela a pour préalable l'interprétabilité et la transparence du modèle de décision. Or les algorithmes d'apprentissage automatique les plus performants aujourd'hui, les réseaux de neurones profonds, sont encore mal compris et fonctionnent comme des boites noires. Par ailleurs, on constate que plus le modèle est complexe, en ce sens où il contient beaucoup de paramètres, plus il est performant. A contrario, plus le modèle contient de paramètres, moins il est interprétable. L'exemple le plus patent est le dernier modèle de génération de langage, GPT3, développé conjointement par OpenAI et Microsoft. Ce modèle ne comporte pas moins de 175 milliards de paramètres et atteint des performances de rupture dans des applications de traitement automatique des langues. Son entraînement a nécessité de déployer des ressources en calcul inédites jusqu'alors. Cela pose aussi des questions écologiques dont il faudra se saisir.

Des travaux émergent cependant sur le développement d'algorithmes dits responsables ou de confiance. On parle de Trustworthy AI ou Trustworthy Machine Learning. Cela concerne la protection, by design, des données personnelles, l'absence de biais, l'équité des décisions apprises. L'interprétabilité post-hoc attire de plus en plus l'attention des chercheurs. Ces résultats sont encore préliminaires, sauf peut-être pour la protection des données personnelles où le paradigme de confidentialité différentielle représente une vraie avancée. Toutefois, force est de constater, pour l'instant du moins, que réguler se fera au détriment des performances intrinsèques des algorithmes. Se pose alors la question de savoir quel algorithme utiliser et dans quel cadre. Peut-on déployer un algorithme d'apprentissage profond, par ailleurs non robuste aux attaques adverses, dans des domaines critiques ? Quel compromis entre performances et auditabilité faut-il fixer pour tel ou tel champ d'applications ? Ce sont des questions qui nécessitent une coopération entre spécialistes de l'apprentissage automatique, les sciences humaines et sociales, et les autorités de régulations.

Remarque

Éric Brousseau

L'apprentissage automatisé pose des problèmes particulièrement difficiles à résoudre pour les décideurs, qu'il s'agisse du juge ou du politique.

La jurisprudence de la Cour de justice de l'Union européenne

Joris Plingers
Délégué à la protection des données, Cour de justice de l'Union européenne

La question de la protection des données à caractère personnel est posée depuis les années 1970, avec le développement des bases de données et des algorithmes traditionnels. Outre les premières résolutions relatives à la protection de la vie privée des personnes physiques vis-à-vis des banques de données électroniques, plusieurs actes ont été posés pour garantir à la fois la libre circulation des données à caractère personnel et la protection des droits fondamentaux.

Le droit à la protection est également inscrit à l'article 8 de la charte de droits fondamentaux de l'Union européenne et à l'article 16 du traité sur le fonctionnement de l'Union, et repris dans le RGPD qui consacre notamment, en son article 22, le droit de ne pas faire l'objet d'une décision individuelle automatisée – lié à l'obligation de transparence imposée au responsable du traitement des données quand ce dernier implique une prise de décision automatisée ou un profilage. Si la Cour de justice n'a pas encore eu l'occasion de se prononcer directement sur la mise en œuvre de ce droit, la protection des données à caractère personnel a déjà fait l'objet de plusieurs arrêts.

Ainsi, dans son avis 1/15 dit PNR, Passenger Name Record, la Cour a fixé pour les algorithmes d'analyse automatisée de données à caractère personnel les exigences suivantes : des critères préétablis spécifiques et fiables, permettant d'aboutir à des résultats ciblant les individus à l'égard desquels pourrait peser un soupçon raisonnable de participation à des infractions terroristes ou de criminalité transnationale grave, et non-discriminatoires. En outre, la fiabilité et l'actualité des modèles, des critères préétablis et des bases de données utilisées doivent faire l'objet d'un examen régulier tenant compte de données statistiques et des résultats de la recherche internationale. De la même façon, dans l'affaire de La Quadrature du net, concernant la mise en œuvre de traitements automatisés destinés à détecter sur les réseaux de télécommunications et sur instruction des autorités nationales des connexions susceptibles de révéler une menace terroriste, la Cour a considéré que des règles claires et précises devaient offrir des garanties suffisantes contre le risque d'abus, a fortiori en cas de risque d'accès illicite aux données. Elle a également précisé que le fait qu'une analyse automatisée ne permette pas l'identification d'une personne ne faisait pas obstacle à la qualification de « données à caractère personnel ». D'autant que la procédure en question vise à identifier la personne concernée si ses agissements paraissent suspects et semblent caractériser l'existence d'une menace terroriste.

S'agissant de la transparence, la Cour a toujours souligné l'importance de fournir une information au moins de nature générale, et individuelle dès lors qu'il est procédé à une identification. Par ailleurs, compte tenu du taux d'erreur inévitable lié aux traitements automatisés, un examen par des moyens non-automatisés peut être demandé avant l'adoption d'une mesure individuelle produisant des effets préjudiciables.

La jurisprudence de la Cour contient également des éléments intéressants pour la régulation des algorithmes, notamment en termes de responsabilité. À la suite de l'arrêt Google Spain, par exemple, le droit au déréférencement et à l'effacement a été codifié dans le RGPD. La Cour a également précisé que l'activité d'un moteur de recherche doit être considérée comme un traitement de données à caractère personnel lorsque les informations affichées contiennent ce type de données. En outre, l'exploitant de ce moteur doit être considéré comme responsable dudit traitement. À cet égard, un traitement initialement licite peut devenir problématique lorsque les données ne sont plus pertinentes ou qu'elles paraissent inadéquates ou excessives au regard de la finalité pour lesquelles elles ont été collectées. Aussi les résultats d'un algorithme doivent-ils être adaptés et actualisés afin d'assurer le respect des principes de la protection des données à caractère personnel. S'agissant des moteurs de recherche, cette mise à jour est faite à la suite de demande de déréférencement des personnes concernées, étant précisé que ce dernier ne concerne que les résultats et non le contenu des pages indexées.

Par ailleurs, dans l'affaire Wirtschaftsakademie Schleswig-Holstein, la Cour de justice a décidé que les opérateurs de pages dites « fan » sur Facebook et le réseau social étaient responsables du traitement des données à caractère personnel des visiteurs. Cette responsabilité partagée contribue à une protection plus complète

des droits des utilisateurs. Elle n'est toutefois pas sans limite. Ainsi, dans l'affaire Fashion ID, la Cour a considéré que le gestionnaire d'un site internet pouvait être considéré comme responsable du traitement découlant de l'utilisation d'un « plugiciel » dans une page web, en l'occurrence le bouton « j'aime » de Facebook. De fait, cette insertion implique le transfert de données à caractère personnel vers Facebook, qui peut alors les utiliser pour améliorer ses algorithmes. Mais la responsabilité de ce gestionnaire est limitée à l'opération dont il détermine effectivement les finalités et les moyens. Il ne peut donc être tenu responsable des opérations ultérieures effectuées par Facebook.

Au total, une interprétation large comme celle que fait la Cour semble la plus appropriée pour la protection des données à caractère personnel.

La régulation juridique des algorithmes de l'administration

Alexandre Lallet
Rapporteur public à la section du contentieux du Conseil d'État

Les données d'apprentissage du juge sont les affaires contentieuses qu'on lui soumet, lesquelles sont très rares s'agissant des algorithmes. En l'occurrence, deux décisions de principe sont fichées au recueil Lebon, sur un total de 172 406 décisions figurant dans la base de jurisprudence de la juridiction administrative. À ce stade, le contrôle des algorithmes est encore un sujet marginal dans la jurisprudence administrative et le juge n'a qu'une connaissance floue et « médiatique » de l'intelligence artificielle. La plus grande modestie est donc de mise.

Le juge administratif a davantage conscience des risques attachés à l'utilisation abusive des algorithmes. Au point, peut-être, d'envoyer un signal négatif, occultant les opportunités ouvertes par l'intelligence artificielle pour l'activité administrative et juridictionnelle. Personne ne s'interroge en France, à ma connaissance, sur un droit opposable à un traitement algorithmique, que le Conseil d'État italien a esquissé en avril 2019 en s'appuyant sur le principe constitutionnel de bonne gestion administrative. L'objectif de sauvegarde des données publiques, le principe de mutabilité du service public et celui d'égalité devant la loi pourraient converger pour reconnaître une forme de droit *à* l'algorithme. Toutefois, je parlerai ici, plus classiquement, du droit *des* algorithmes, c'est-à-dire des règles qui encadrent leur déploiement dans la sphère administrative.

La diversité des algorithmes et de leur contribution à la prise de décision administrative est grande, et appelle des règles juridiques elles-mêmes variées.

S'agissant du contrôle de la « transparence », l'une des conditions nécessaires – mais non suffisantes – pour auditer un algorithme administratif est d'avoir accès à

la documentation et au code source. À cet égard, nous sommes assez bien pourvus, avec le RGPD et le code des relations entre le public et l'administration (ex-loi CADA). Le droit à communication trouve toutefois plusieurs limites. Notamment, si un tiers à l'administration détient des droits de propriété intellectuelle, le code source n'est pas communicable. Une autre limite tient aux secrets protégés par la loi, comme l'a illustré la décision du Conseil constitutionnel sur Parcoursup. Quoi qu'il en soit, la transparence formelle est assez largement assurée dans le droit positif. Reste la question de la transparence réelle, au sens de l'intelligibilité du traitement et de ses conséquences, que le droit néglige à ce stade.

Quant au contrôle de la légalité, il se heurte à une première difficulté : comment saisir un acte administratif derrière un programme informatique ? Quand il existe un décret créant un traitement de données à caractère personnel, la question ne se pose pas. Alicem, par exemple, qui est une application sécurisée pour accéder à des téléprocédures administratives reposant, lors de la création du compte, sur la reconnaissance faciale, est créée par un décret en bonne et due forme, qui peut être contrôlé. Mais dans certains cas, et tout particulièrement depuis le RGPD qui a allégé les formalités préalables, il n'existe pas d'acte formel créant le traitement de données. Le juge peut alors soit décider que l'existence d'un acte créateur est révélée par celle du traitement de données, soit s'appuyer sur la branche de la jurisprudence qui ouvre le prétoire du juge administratif aux documents et actes de droit souple ayant des « effets notables ». De fait, il est difficile de contester qu'un algorithme qui fonde en tout ou partie des décisions administratives n'aurait pas d'effet notable.

Enfin, concernant le contrôle de fond, le cadre juridique est clair pour ce qui est de la prise de décision automatisée : l'article 47 de la loi Cnil éclairé par la décision du Conseil constitutionnel de 2018 pose une obligation de transparence, un droit au recours administratif examiné par un humain (lequel devrait être obligatoire, à notre avis, pour que le juge administratif contrôle non pas la décision de la machine, mais celle de l'agent public), la non-utilisation de données sensibles, la maîtrise du fonctionnement de l'algorithme (ce qui paraît exclure tout le *machine learning*) ou encore la légalité des critères implémentés. Faut-il aller au-delà ? Pas nécessairement. Ce qui compte avant tout, c'est le résultat, donc la décision elle-même, et pas tout le détail du raisonnement dont elle est issue ou la question de savoir quel est le neurone, artificiel ou humain, qui a causé l'erreur, le cas échéant.

Explicabilité, non-discrimination et responsabilité

Winston Maxwell
Directeur d'études, droit et numérique, Télécom Paris

En matière de transparence, les explications ne sauraient être les mêmes selon leur destinataire. En effet, elles ne seront pas les mêmes selon qu'on s'adresse à un mathématicien, un régulateur ou un citoyen. Quoi qu'il en soit, la loi tente de définir « logique sous-jacente » portée par le RGPD ou celle d'explication « utile pour le commerçant » qui figure dans le règlement *Platform to business*. En tout état de cause, pour les juristes et les régulateurs notamment, la transparence au sens de l'accès au code source n'est pas toujours très utile.

Pour les algorithmes privés, aucun texte ne fixe une obligation globale de transparence et d'explicabilité, même s'il peut y en avoir par secteur d'activité, bancaire par exemple. Dans le public, en revanche, cette obligation repose sur un fondement constitutionnel – tant aux États-Unis (*due process*) qu'en Europe (droit de recours effectif, accès aux documents de l'administration, etc.). Trois cas de jurisprudence méritent d'être mentionnés en la matière. Dans l'arrêt Loomis vs Winsconsin, relatif à un algorithme de prédiction du risque de récidive d'un détenu, le tribunal a considéré qu'il n'y avait pas de violation du *due process* dans la mesure où le score utilisé était un facteur mineur dans la décision rendue. En revanche, il a considéré qu'une notice d'utilisation détaillée devait notamment mettre en avant les limites et les faiblesses de l'algorithme. Au Texas, en revanche, dans l'affaire Houston Teachers le tribunal a donné raison à un syndicat d'enseignants qui contestait un algorithme de gestion de carrière des enseignants au motif qu'il était opaque et que les enseignants n'étaient pas en mesure de le tester et le critiquer – il y avait donc bien un déni de *due process*. Une décision similaire a été rendue aux Pays-Bas à propos d'un algorithme de détection de la fraude à la sécurité sociale, même s'il n'était qu'un élément parmi d'autres d'aide à la décision, laquelle relevait d'une intervention humaine.

Par ailleurs, la non-discrimination algorithmique ne saurait être totale en tout instant. Certaines lois mathématiques, par exemple rendent incompatibles la non-discrimination individuelle et la non-discrimination de groupe. Dans le domaine informatique, qui plus est, pousser trop loin cette notion se fait souvent au détriment de la performance. Par ailleurs, les contraintes sont moins fortes aux États-Unis qu'en Europe en ce qui concerne l'utilisation de données sensibles (couleur de peau, origine ethnique) pour tester les algorithmes et réduire les biais. L'université de Berkeley a poussé la question de non-discrimination plus loin, constatant que l'algorithme de sélection à l'entrée le plus efficace pour réduire les discriminations était celui qui tenait compte de la couleur de peau ou du quartier dans lequel ont grandi les candidats, et ce afin de donner des poids différents à certaines données d'entrée. De fait, cet algorithme de quasi-discrimination

positive aboutit à un résultat non-discriminatoire plus satisfaisant en termes d'objectifs sociétaux. En France, cela se heurterait au principe constitutionnel d'égalité. Il est dommage que l'Europe soit freiné dans la recherche sur les biais et discriminations algorithmiques en raison des contraintes liées à l'utilisation des données sensibles. Ces contraintes existent pour de très bonnes raisons, mais idéalement il faudrait trouver une solution pour pouvoir avancer plus vite sur ce sujet important.

Enfin, s'agissant du cadre réglementaire de la responsabilité, l'exemple le plus abouti est la loi votée par l'État de Washington en mars 2020 pour réguler des algorithmes de reconnaissance faciale. Elle impose au fournisseur de diffuser une notice d'information très détaillée. Pour sa part, l'exploitant doit élaborer un cahier des charges pour l'utilisation. Une étude d'impact sur les droits fondamentaux doit également être effectuée et des actions correctives doivent être proposées, le cas échéant. L'algorithme doit être ouvert aux tests indépendants. Enfin, un cadre institutionnel de contrôle est prévu. Ce dispositif est perfectible, mais c'est une bonne source d'inspiration pour la régulation des applications IA créant un risque pour les individus.

La construction d'une régulation algorithmique européenne

Werner Stengg
Expert au cabinet de Margrethe Vestager, vice-présidente exécutive de la Commission européenne, pour une Europe adaptée à l'ère du numérique

Ces trois initiatives européennes fournissent un bon exemple de la façon dont la législation peut essayer de faire face aux défis technologiques : le règlement Platform to business (P2B) de juillet 2020, qui promeut l'équité et la transparence pour les entreprises utilisatrices de services d'intermédiation en ligne ; le Digital Services Act (DSA) à venir, qui propose une révision de la direction sur le commerce électronique ; le Livre vert sur l'intelligence artificielle, qui donnera lieu à un règlement sur les risques liés à cette technologie.

Le règlement P2B est le premier acte législatif qui traite de la question des algorithmes des fournisseurs de services d'intermédiation – Amazon, Booking.com, app stores et autres places de marché. Un de ses articles concerne le classement proposé par les plateformes, c'est-à-dire leur façon de prioriser les biens et services offerts au terme de la recherche effectué par l'utilisateur. De fait, ce classement a une incidence importante sur le choix des consommateurs, donc sur la réussite commerciale des entreprises utilisatrices. C'est la raison pour laquelle cet article s'applique aussi aux moteurs de recherche en ligne. En l'occurrence, pour établir leur classement, les plateformes utilisent des

algorithmes dont le fonctionnement est complètement opaque pour les utilisateurs – ce qui s'avère particulièrement problématique lorsqu'ils sont en concurrence avec les services directement offerts par les plateformes. Afin que le classement ne soit pas arbitraire, le règlement impose aux plateformes de décrire au préalable les principaux paramètres de classement – il s'agit d'en identifier un nombre limité parmi les plus importants. Pour autant, les plateformes ne doivent pas fournir le fonctionnement détaillé de leur algorithme de classement. L'objectif visé n'est pas seulement la transparence, mais aussi l'utilité.

Le projet de DSA suit la même logique d'accroissement de la transparence. Les très grandes plateformes, dont la définition reste à finaliser, devront faire l'objet d'une surveillance approfondie par le régulateur compétent voire d'un audit externe s'agissant de leurs critères de transparence ou de leur pratique de modération des contenus.

Enfin, dans la mesure où l'intelligence artificielle porte de nombreux risques du point de vue du citoyen – discrimination, biais informationnel, etc. – un cadre réglementaire devrait être proposé au printemps 2021 pour fixer des exigences de transparence et lutter contre l'opacité.

Echanges

De la salle

D'un côté, les données augmentent de façon exponentielle. De l'autre, elles sont de plus en plus réglementées. Cela ne risque-t-il pas d'être source de complexité pour les entreprises ?

Par ailleurs, les scientifiques ont réussi à reproduire le système de stockage des neurones et des synapses. Il existe des outils de data management, notamment de *data lineage*, pour tracer la donnée. Mais l'environnement est souvent très complexe. Récemment, par exemple, deux chatbots de Facebook qui communiquaient entre eux avec un langage compris des scientifiques ont subitement changé de langage. Si le chercheur n'est plus apte à comprendre la machine, ce sera plus difficile encore pour le juge et les institutions.

Enfin, quand la machine deviendra une entité à part entière, ne faudra-t-il pas la traiter comme un être humain, puisque nous sommes nous-mêmes des entités complexes ?

Éric Brousseau

Le sujet de l'éthique de l'intelligence artificielle est de plus en plus étudié.

Jamal Atif

Il importe de distinguer le fantasme de la prospective à très long terme de la réalité. Pour reprendre cet exemple, faire communiquer les chatbots avec un langage que

nous ne comprendrions pas ne présenterait pas de grand intérêt. La plus grande prudence est également de mise face aux annonces de la création de neurones artificiels.

En tout état de cause, nous formons tous nos étudiants à l'éthique de la production scientifique.

Éric Brousseau

Nous avons clos la première table en soulevant la question de l'éducation des utilisateurs. La question de l'éducation de ceux qui fabriquent les machines se pose également.

Table ronde n°3
Le commerce et les échanges de données

Modératrice : Joëlle Toledano
Professeur émérite d'économie, associée à la Chaire Gouvernance et Régulation de l'université Paris Dauphine-PSL

Les tables rondes de cette journée illustrent bien l'ambivalence du numérique. Les débats autour de ce sujet renvoient souvent à des questions autour des données et de leurs usages. L'économie, notamment, pointe à la fois de formidables opportunités, des blocages qu'il faut lever et des risques dont il convient de tenir compte.

Les données ont une multiplicité d'usages et permettent de nombreux développements. Toutes les activités devraient en profiter, à condition d'y avoir accès et de savoir comment les utiliser. En principe, le marché sert à cela. Or force est de constater que tel n'est pas vraiment le cas. Les échanges marchands ne concernent qu'une partie limitée des données utilisables, essentiellement personnelles.

La donnée n'est pas un bien comme un autre et le marché est souvent défaillant pour en optimiser l'usage. On ne peut pas compter sur lui ni pour assurer une collecte suffisante et une diffusion efficace associée à un partage équitable de la valeur tout en évitant une collecte excessive qui ne respecterait pas la vie privée, ni pour prévenir les problèmes de verrouillage d'accès ou de cybersécurité. De fait, ce sont les défaillances du marché des données qui justifient l'intervention publique et la mise en place d'une nouvelle organisation et d'une gouvernance.

Par ailleurs, la donnée est un bien non rival : elle peut être utilisée plusieurs fois dans différents contextes sans perdre de son utilité ou de sa valeur. C'est la raison fondamentale des politiques publiques dites d'open data pour disséminer et valoriser les données publiques. Ne pas y avoir accès crée donc une barrière, qu'il convient de lever. Quand les données ne sont pas publiques, des incitations à investir dans les différentes dimensions (collecte, mise en forme, traitement, etc.) sont nécessaires. Sans compter que certaines données sont aussi monopolisées par des acteurs, qui en tirent des gains substantiels. Ils ne souhaitent pas les partager et, par certaines pratiques, verrouillent et opacifient les marchés. D'où la nécessité de trouver un ou des cadres qui favorisent la transparence tout en protégeant à la fois la concurrence et la vie privée.

Les courtiers en données sont les seuls acteurs du marché des données. Or les informations et les chiffres disponibles les concernant sont très limités. Se pose aussi la question des données stockées dans des systèmes interconnectés. Des mesures politiques complémentaires au RGPD sont nécessaires pour garantir une protection adéquate des données sensibles.

L'exemple du secteur de l'énergie

Dominique Jamme
Directeur général des services de la Commission de régulation de l'énergie

Les réseaux de transport et de distribution d'électricité et de gaz sont des monopoles naturels, régulés. En revanche, les secteurs de l'amont (production) et de l'aval (commercialisation) sont concurrentiels – même si en France, EDF est le seul producteur de nucléaire, le renouvelable est subventionné par l'État et il existe encore des tarifs réglementés de vente.

Les données sont omniprésentes et leur quantité ne cesse de croître avec la multiplication des capteurs connectés. Elles ont une grande valeur économique, car le prix de l'électricité dépend fortement de l'heure et du jour de l'année.

Plusieurs cas pratiques se posent au régulateur. Le premier concerne le statut des fichiers clients issus des anciens monopoles des tarifs réglementés de vente dont ont hérité EDF et Engie à l'ouverture du marché. Ces opérateurs ont l'interdiction d'utiliser ces fichiers pour proposer des offres de marché en électricité ou en gaz, et ils ont l'obligation de communiquer des données clients aux fournisseurs concurrents lors des étapes de disparition des tarifs réglementés de vente. Ainsi, en 2023, Engie devra communiquer à ses concurrents qui en feront la demande le fichier des 3,5 millions de clients toujours concernés par ces tarifs pour le gaz (sauf opposition expresse du client). Conformément aux recommandations de la CRE et des pouvoirs publics, ce fichier ne devra pas mentionner le numéro de téléphone des consommateurs : en effet, l'intérêt de ces derniers est de recevoir des offres, pas de se faire harceler.

Un autre cas concerne les comparateurs d'offres de fourniture. Ces outils, en fort développement depuis deux ou trois ans, animent le marché, accroissent l'information des consommateurs et aident les clients à choisir la meilleure offre – sous réserve de renseigner différentes données (adresse, facture actuelle, habitudes de consommation, type de logement, etc.). Ces opérateurs sont rémunérés par le fournisseur, soit au clic soit au contrat signé. Le régulateur peut s'assurer de l'existence d'une transparence minimale. En revanche, il ignore ce que ces plateformes font des données collectées et si elles les valorisent.

Un troisième cas concerne les données fines de consommation et les courbes de charge. Ces données sont protégées, car ce sont des informations commercialement sensibles pour l'entreprise et personnelles pour les consommateurs résidentiels. Ainsi, dans un logement, connaître la consommation de l'occupant précédent relève d'une procédure très encadrée.

Du côté du système électrique, les fournisseurs ne connaissent pas aujourd'hui la consommation exacte de leurs consommateurs pour chaque demi-heure, mais uniquement des moyennes par profil. Ils ne sont donc pas incités à agir sur la consommation de leurs clients : même si un client réduit sa consommation

en cas de pic tarifaire, le fournisseur est facturé sur la base de la moyenne de tous les consommateurs du même profil. Un système de règlement des écarts sur la base des courbes de charge par demi-heure serait plus efficace, puisque les fournisseurs seraient facturés sur la base de la véritable courbe de charge de consommation des clients en portefeuille – ce que les systèmes d'information permettent de faire. Aussi la CRE a-t-elle décidé que pour les 500 000 entreprises au-dessus de 36 kVA, le règlement des écarts se fera progressivement (en deux étapes, fin 2020 et fin 2022) sur la base des courbes de charge.

De son côté, le consommateur peut accéder à sa courbe de charge et à son historique en se rendant sur le site d'Enedis. Pour l'instant, 4,5 millions de foyers l'ont fait sur un total de 28 millions équipés de compteur connecté. Par ailleurs, conformément à un avis Cnil de 2015, Enedis ne stocke pas les données par défaut : même si Enedis est opérateur régulé en monopole et s'il ne peut pas vendre les données, leur stockage requiert un accord explicite du consommateur. Or pour demander une offre pertinente, il importe de disposer d'un historique de 12 mois. La situation n'est donc pas parfaite, et il convient de réfléchir à la façon d'optimiser le système.

En somme, les enjeux de la courbe de charge sont d'un côté la transition énergétique et l'intérêt même du consommateur, de l'autre la protection de la vie privée et des données individuelles.

La régulation des données : passer des principes à la pratique

Jean-Yves Ollier
Conseiller d'État

Pourquoi le cadre juridique et de gouvernance de l'espace européen des données est-il toujours en construction, alors que les principes de protection et de libre circulation des données ont été posés il y a plus de 20 ans ? L'échange de données ne se laisse pas aisément appréhender par la notion de marché, dans les différentes acceptions du terme : l'échange n'est pas toujours marchand, du moins pas explicitement, et son objet ne se laisse pas facilement saisir ni localiser. Les échanges de données ne caractérisent pas un secteur. Enfin les contrats sont souvent implicites et opaques. On peut examiner quelques-uns des enjeux de cette construction à travers les différentes briques de ce qui constitue classiquement les règles communes d'un marché.

1. Quel est l'objet de l'échange et quel est le champ des transactions possibles ?

Les données elles-mêmes ne sont pas des biens susceptibles d'appropriation exclusive, non seulement parce qu'elles sont non rivales, mais surtout parce que des droits fondamentaux s'y attachent : protection des données personnelle,

liberté de communication, accès aux données publiques. En revanche, les droits de producteurs des bases de données sont protégés au titre d'un régime spécifique de propriété intellectuelle, lorsqu'ils ont investi substantiellement dans le développement de ces bases. Le réexamen de ce cadre est au programme de la Commission, pour encourager à la fois l'investissement et le partage des données.

En outre, certaines données sont placées tout ou partie hors du marché par la loi les données personnelles collectées par un opérateur ne sont pas accessibles à des tiers, ni susceptibles d'échange, sans le consentement de l'utilisateur, qui doit en principe conserver un contrôle sur leur utilisation. Les termes du contrat entre les plateformes et les utilisateurs posent plus largement des questions de la protection du consommateur (notamment pour les services fournis sans contrepartie financière) et de son autonomie. Celle-ci passe par la portabilité des données personnelles et non personnelles au sein et vers l'extérieur des écosystèmes dans lesquelles elles sont accumulées, pour prévenir les pratiques de verrouillage. Le RGPD pose un principe de portabilité externe des données personnelles, ce qui ne résout qu'une partie du problème. En effet, la portabilité est en partie liée à l'interopérabilité des systèmes, sujet difficile à appréhender par le droit. Globalement, les utilisateurs pâtissent encore de l'absence d'outils et de normes techniques pour rendre l'exercice de leurs droits simple et effectif à l'autre bout du spectre, les données des administrations publiques et des entreprises investies de mission de services public sont un bien public. C'est le droit national qui décide si elles sont communicables. Le cas échéant, le droit communautaire indique qu'elles doivent être librement réutilisables.

L'ensemble des données à caractère non personnel qui ne relèvent pas du secteur public est plus difficile à appréhender, mais cet espace est appelé à fortement se développer avec les objets connectés.

2. Quelles sont les modalités de l'échange ou de la prestation de service liés aux données ?

Des règles doivent venir encadrer les marchés de données internes aux écosystèmes intégrés, dans lesquels les transactions s'effectuent le plus souvent selon un mode hiérarchique. C'est l'objet du règlement platform to business, dit P2B, de juin 2019.

L'un des enjeux de la stratégie européenne pour les données est la clarification du cadre réglementaire et de gouvernance applicable à certains maillons essentiels de la chaîne de valeur, comme l'informatique en nuage.

3. S'agissant de la délimitation des espaces géographiques des échanges, les règlements posent un principe de libre circulation dans l'Union, ou plus exactement de libre flux des données personnelles dans limites permises par la protection des données personnelles. Or les flux physiques ne correspondent pas toujours à ceux des transactions. Les flux vers les pays tiers doivent se faire dans le respect d'un niveau de protection adéquat. Par l'effet du Brexit, le RGPD devrait cesser à la fin de l'année 2020 de s'appliquer directement au Royaume-Uni, qui représentait près

de 20 % de la valeur de l'économie des données dans l'Union à 28. La circulation des données personnelles avec l'UE après la fin de cette période transitoire sera soumise à une décision d'adéquation ou à la mise en place des autres outils prévus par le RGPD. , deviendra un pays tiers.

4. Les règles et le cadre de gouvernance peuvent aussi définir des espaces sectoriels d'échanges de données. Ainsi, la France a décidé de l'ouverture de certaines données de transport : la loi d'orientation des mobilités a confié à cet égard une mission à l'autorité de régulation des transports.

La directive PSI III de 2019 a ouvert le chantier des ensembles de données de forte valeur, c'est-à-dire celles des opérateurs publics dont la réutilisation est associée à d'importantes externalités positives pour la société, l'environnement ou l'économie – comme la cartographie des observations satellites.

En somme, si les principes généraux sont l'une des forces de l'Union européenne, la construction d'espaces de données porteurs de croissance passe par la résolution des questions que pose leur mise en œuvre effective (conditions du consentement, modalités de la portabilité, interopérabilité, effectivité des recours…), dans l'environnement des échanges de données, en tenant compte des caractéristiques et des comportements des principales plateformes et de spécificités sectorielles.

Les données de la publicité ciblée en ligne

François Lhemery
Vice-président des affaires réglementaires de Criteo

Criteo est un acteur mondial de la publicité en ligne. Il y a deux ans, cette entreprise française a également ouvert un laboratoire de recherche en IA – le deuxième laboratoire privé en Europe, qui héberge actuellement 80 ingénieurs et chercheurs.

Tous les mois, Criteo touche près de 2 milliards d'utilisateurs connectés dans une centaine de pays. Et ce, en utilisant trois types de données : celles sur les utilisateurs (via un cookie ou un identifiant publicitaire), celles sur les produits (via des partenariats avec 20 000 annonceurs) et celles sur les publicités (via les clics sur telle ou telle publicité). Le modèle de fonctionnement est celui du consentement. Par ailleurs, Criteo a inventé le modèle des bannières publicitaires à la performance, qui est devenu la norme sur le marché. Il convient également de préciser que nous ne collectons pas de données d'identification directe des utilisateurs (nom, adresse, mail, activité sur réseaux sociaux), mais uniquement des données qui permettent de comprendre la navigation, d'en inférer des intentions d'achat et de proposer le bon produit au bon endroit. De fait, ce n'est pas la quantité mais la qualité des données qui importe, et la capacité à en tirer de l'intelligence.

Ainsi, chaque acteur collecte les données dont il a besoin, et la valeur qu'il crée tient du traitement qu'il en fait et de sa capacité à en inférer des intentions d'achat. À cet égard, même si nos clients annonceurs partagent leur catalogue avec nous, pour permettre l'indexation, il n'existe pas vraiment de marché sur lequel se ferait de l'achat/revente de données.

Le problème n'est donc pas le marché, mais la concentration des données dans les mains d'une poignée d'acteurs dominants dont le modèle très intrusif crée des distorsions de concurrence. Les deux premiers acteurs, Google et Facebook, qui ne sont pas des *pure players* de la publicité, représentent environ 70 % du marché mondial de la publicité en ligne, et les trois suivants sont en train de devenir très conséquents. Cette concentration pose deux grands enjeux sociétaux : le maintien d'un internet ouvert, peu cher et accessible à tous ; le financement de la pluralité des médias. Une étude Google conduite l'an dernier montre que la suppression des cookies, donc de la capacité à faire de la publicité à performance, entraînerait une baisse de revenus de 52 à 75 % pour les éditeurs de presse.

Qui plus est, la crise du Covid a dégradé la situation chez tous les acteurs de la publicité en ligne. L'intervention des pouvoirs publics est donc nécessaire, car la concentration est telle sur les marchés qu'ils ne sont plus en mesure de se corriger eux-mêmes.

Dans ce contexte, quelques pistes de réflexion méritent d'être ouvertes.

Comment mieux anticiper les effets sur la concurrence des règles sur la protection des données ? L'interprétation trop stricte du RGPD par certains régulateurs a eu pour effet néfaste de renforcer la domination de quelques acteurs dominants au détriment de l'internet ouvert et des plus petits acteurs, comme le montrent les articles du professeur Gérardin. Autre exemple, le projet de règlement e-Privacy présente des risques d'incohérence avec d'autres textes comme le DSA, s'agissant notamment des cookie walls.

Par ailleurs, comment permettre au DSA de tenir ses promesses ? Faut-il interdire le *self-preferencing* ? Faut-il ouvrir l'accès aux jardins fermés des grandes plateformes ?

La technologie n'est pas neutre. La réglementation peut et doit se structurer dans le monde virtuel comme dans le réel, et la réglementation en matière de vie privée doit à tout prix prendre en compte la dimension de concurrence.

La valeur des données

Laurent Lafaye
Co-fondateur et directeur général de Dawex

L'émergence du big data – la valorisation des données collectées par les entreprises pour leur propre usage – a été suivie de l'open data, qui a surtout concerné les entreprises qui travaillaient avec des acteurs publics. Puis l'arrivée du RGPD s'est traduite par d'importants investissements des entreprises. Enfin, le sujet des cyber-menaces incite les dirigeants à appréhender la data comme un enjeu face auquel il faut investir.

Le sujet est donc avant tout traité sous l'angle du coût. Mais quelle peut être la valeur économique de cet actif qu'est la donnée ? La grande distribution, par exemple, a longtemps diffusé gratuitement à ses fournisseurs des informations de stock et de points de vente pour faire des panels. Aujourd'hui, elle entend les commercer, en définissant très précisément leurs conditions d'utilisation. Dans un tout autre domaine, un fabricant de batteries électriques a tout intérêt à comprendre les usages de ces équipements, donc à autoriser l'accès à ses données. Même en l'absence de valeur marchande directe de la transaction de données, il existe une valeur indirecte au travers de l'amélioration de la connaissance des usages donc des produits. D'où l'intérêt de réfléchir à un modèle d'accès aux données non personnelles – le règlement européen en la matière étant très peu connu.

De nombreuses opportunités peuvent être trouvées. Certaines filières commencent ainsi à s'organiser entre elles pour gouverner l'accès et le partage de leurs données, sans nécessairement y associer un modèle économique direct. Des représentants du monde agricole français ont ainsi mis en commun des moyens pour créer une infrastructure d'échange de données industrielles. Le modèle économique n'est pas la valeur des données échangées, mais les conditions dans lesquelles elles sont échangées, qui permettent de se protéger contre les GAFA qui pourraient être tentés de les revendre aux brokers de matières premières sur les marchés financiers. L'exemple des villes intelligentes est également intéressant. Avec l'arrivée massive des objets connectés, les pouvoirs publics ont un rôle à jouer dans l'accès et la circulation des données entre tous les acteurs du territoire, pour favoriser l'activité économique mais aussi renforcer l'attractivité.

En tout état de cause, un véritable effort d'information et de pédagogie est nécessaire.

Par ailleurs, la question de l'adéquation entre les différentes régulations est de taille. Certains principes du RGPD, notamment celui de responsabilité, pourraient s'appliquer à des données non personnelles par exemple.

Enfin, des infrastructures de circulation de la donnée sont nécessaires, ce qui passe aussi par l'identification de tiers de confiance. Plus on créera des cadres de confiance, plus on sera proche de ce que font les marchés financiers : définir des règles du jeu pour tous les opérateurs.

Echanges

De la salle

Nous vivons une guerre réglementaire, géopolitique. La régulation est positive. Dans certains pays comme la Russie ou la Chine, elle se fonde sur le stockage dans le pays. En Europe, le projet Gaia-X vise à permettre de gérer l'interopérabilité et la portabilité des données. Mais pour certaines entreprises, la digitalisation peut s'avérer compliquée. Comment rendre la réglementation plus claire et plus fluide pour faciliter l'accélération digitale tout en la sécurisant ?

Jean-Yves Ollier

Il n'existe pas d'obligation de localisation. Au contraire, le règlement sur les échanges de données non personnelles interdit aux États membres d'imposer ce type de contrainte au sein de l'UE. En revanche, la question de l'adéquation du niveau de protection peut faire obstacle aux échanges avec certains pays tiers, comme l'illustrent les récentes décisions sur le *privacy shield*.

Par ailleurs, la Commission travaille à définir un cadre d'usage permettant de structurer une industrie européenne du *cloud* et de réduire les asymétries, notamment s'agissant des clauses abusives qui limiteraient la réversibilité. Les outils sont essentiellement contractuels.

De la salle

De quel accès aux données de Linky disposent les fournisseurs d'énergie ? Qu'en font-ils concrètement ?

Dominique Jamme

Les fournisseurs ont accès d'office à toutes les données nécessaires à la facturation, c'est-à-dire essentiellement les consommations mensuelles… Pour le reste, le consentement exprès du client est requis – ce qui génère de la complexité, puisque ce sont les opérateurs de réseau qui doivent vérifier que le fournisseur a bien reçu ce consentement.

Par ailleurs, les fournisseurs utilisent ces données pour proposer des offres innovantes, comme des offres spécialisées « week-end » ou « véhicule électrique ». Celles-ci restent encore marginales, au regard des offres vertes et à prix fixe, mais nous espérons qu'elles se développeront davantage. Aujourd'hui, le verre n'est qu'à moitié plein : on n'exploite pas pleinement le potentiel des compteurs évolués.

De la salle

Est-il nécessaire de légiférer pour organiser un accès forcé aux données d'intérêt général détenues par les entreprises ?

Jean-Yves Ollier

Nous aborderons en partie ce sujet dans la table ronde suivante. En tout état de cause, les documents des entreprises chargées de mission de service public sont accessibles depuis la loi d'origine du 2 juillet 1978, actualisée par la loi pour une République numérique de 2016. L'enjeu de la directive PSI-III, qui doit être transposée en 2021, vise précisément à s'assurer de l'applicabilité effective des principes de mise à disposition et d'interopérabilité. L'extension de ce type de contraintes à d'autres catégories de données est une autre affaire !

De la salle

Dans les deux années à venir, les navigateurs devraient limiter la collecte des cookies tiers. Par ailleurs, la publicité exclusivement contextuelle, sans utilisation de données personnelles, est plutôt utopique. Dans le cadre de l'initiative Privacy Sandbox, Google tente de proposer un modèle alternatif. Criteo investigue-t-il les possibilités de collecter différemment ou de faire du ciblage prédictif sans données personnelles ?

François Lhemery

Bien sûr ! Non seulement nous étudions avec attention la proposition de Google de faire de la publicité de ciblage par cohorte, suivant la méthode dite Turtledove, mais nous avons aussi soumis une proposition complémentaire, Sparrow. Nos trois axes de réflexion sont les suivants : comment faire de la publicité ciblée avec ou en dehors du cookie ; comment faire de la publicité ciblée par cohorte ou groupe d'intérêt ; comment faire de la publicité contextuelle.

La publicité contextuelle considère que vous pouvez être intéressé par un billet d'avion ou une chambre d'hôtel lorsque vous regardez un site sur Londres. La publicité ciblée, elle, se focalise sur les intentions d'achat que vous n'avez pas concrétisées. L'ambition est d'allier pertinence et efficacité. C'est là que se trouve la limite du contextuel par rapport au ciblage.

Joëlle Toledano

Dawex a-t-il des concurrents, ou doit-il élargir son marché ?

Laurent Lafaye

Les deux sujets sont très liés. Le marché croît rapidement et nous devons faire de même, en ayant face à nous deux géants américains – alors qu'au démarrage, en 2015, il n'y avait que des startups.

Les acteurs publics, notamment les collectivités territoriales, peuvent participer à la définition du cadre des nouvelles mobilités face à l'arrivée de nouveaux acteurs et de nouvelles formes de collecte de données. En tant que tiers de confiance, ils ont un rôle à jouer tant vis-à-vis des acteurs privés déjà installés que des nouveaux entrants.

Éric Brousseau

Il n'y a pas encore grand-chose à réguler, puisque les marchés n'existent pas. Pourtant, des enjeux majeurs sont d'ores et déjà posés. Les acteurs privés ne devraient-ils pas s'organiser par filière, à l'instar de l'exemple cité dans le secteur agricole ? Ne faudrait-il pas avant tout promouvoir la gouvernance d'écosystèmes de données, plutôt qu'une régulation ?

Laurent Lafaye

Pour créer les conditions de la confiance, un cadre est nécessaire. Des initiatives émergent, notamment dans l'agriculture, dans l'agroalimentaire ou dans le tourisme – dont les données étaient captées jusqu'ici par des sociétés ni françaises ni européennes, qui créaient des modèles économiques qui les dépassaient. Il importe d'en faire la promotion, car la France et l'Europe peuvent avoir un leadership sur ces sujets.

Joëlle Toledano

Nous sommes à l'aube d'un système incroyablement ouvert !

Il est stupéfiant de constater que le seul acteur économique capable de chiffrer ce que coûterait à la presse la disparation est cookies est Google. Il est urgent que toutes les entreprises et organisations prennent conscience qu'elles doivent impérativement maîtriser les données de leur propre vie économique, pour des questions d'indépendance.

Table ronde n°4
La régulation de l'accès aux données

Modérateur : Jean-Denis Combrexelle
Président de la section du contentieux du Conseil d'État

Dans les années 90, la question de la régulation des données portait essentiellement sur les données publiques : quelles étaient les conditions de leur diffusion, le principe de gratuité s'appliquait-il ? Dans les années 2000 on s'est aperçu que les opérateurs privés disposaient de données tout aussi, voire plus intéressantes. La régulation des données a pris alors un ton nouveau autour des interrogations sur les distorsions de concurrence...

Données et fabrique de la loi

Miguel Amaral
Économiste principal, OCDE

Les quatre principaux enjeux, liés aux technologies émergentes, identifiés par l'OCDE en matière de politique règlementaire sont les suivants : enjeu lié à la célérité des nouvelles technologies ; enjeu pour les cadres de réglementation existants ; enjeu en matière de mise en œuvre des règles ; et enjeu lié à l'organisation institutionnelle de la régulation et questions transfrontalières. Face à ces enjeux, il conviendrait que les gouvernements développent des approches plus agiles. Cela implique, en particulier, de renforcer la coordination, car l'échange de données entraîne souvent des chevauchements entre les mandats des différents régulateurs. Il convient aussi d'aller vers plus d'expérimentation (au travers, par exemple, des bacs à sable réglementaires) et de renforcer les capacités d'anticipation au moyen, par exemple, d'analyse prospectives sur les opportunités et les risques associés à l'émergence d'une nouvelle technologie.

Par ailleurs, les mérites des approches de régulations dites souples (co-régulation, auto-régulation, standards) ne doivent pas être sous-estimés, en complément des approches traditionnelles de règlementation.

L'OCDE suggère aussi de repenser la fabrique de la loi au regard des enjeux soulevés par les technologies émergentes. Forte de plusieurs années d'études de cas, elle est notamment convaincue qu'il est urgent de ré-examiner la façon dont l'étude d'impact ex ante ou la consultation des parties prenantes sont mises en œuvre. Un regard plus rapproché et une discussion plus étroite avec les entreprises privées semble, en particulier, indispensables. Quant à l'évaluation ex post, il s'agit souvent d'une lacune de la politique réglementaire, alors qu'il est crucial de pouvoir apprécier les effets de la réglementation pour, le cas échéant,

l'adapter. Il convient également de renforcer la coopération internationale et d'en tirer toutes les opportunités pour répondre aux problématiques transnationales que soulèvent les technologies digitales dans leur grande majorité.

S'agissant des données, deux questions en particulier soulèvent des difficultés : la première question est peut-être *qui régule* (sous réserve que la régulation soit l'option la plus adaptée) ? Comment articuler l'organisation de la régulation avec des technologies qui créent des chevauchements entre les mandats actuels de plusieurs agences gouvernementales et de régulateurs ? Comment articuler l'approche concurrentielle avec le besoin qui semble aujourd'hui partagé de renforcer la régulation ex ante de l'information et des données ? A défaut de pouvoir mettre en œuvre des solutions institutionnelles à court terme, il parait difficile de faire l'économie d'une coopération renforcée entre les agences gouvernementales (y compris au-delà du territoire national), dans une approche qui doit inclure les régulateurs économiques. Une deuxième question est de déterminer l'objet de la régulation. Au-delà des complexités liées la diversité des données, cette question touche certainement aux problématiques aujourd'hui non résolues de définition des droits de propriété sur les données. Des travaux restent à conduire sur ces questions et l'OCDE entend y contribuer.

Remarque

Jean-Denis Combrexelle

Le sujet de l'expérimentation a donné lieu à un rapport de la section du rapport et des études du Conseil d'État. Les questions relatives à la fabrique de la loi occupent également largement nos réflexions.

Quelles modalités d'élargissement de l'accès aux données ?

Henri Isaac
Maître de conférences à l'université Paris Dauphine-PSL, membre du Conseil national du numérique

S'agissant de l'ouverture de l'accès aux données, la question du « comment » est souvent laissée de côté, ce qui complique le débat. De la même façon, la grande variété des données est peu prise en considération dans la volonté d'équilibrer les marchés en ouvrant les données des plateformes structurantes.

À qui ouvrir les données ? La dimension concurrentielle, par exemple, mérite d'être élargie : l'ouverture des données ne doit pas être uniquement à destination des concurrents. Elle peut se faire au profit des utilisateurs également. Dès lors qu'il s'agit de partager des données personnelles, des questions juridiques se poseront nécessairement. Par ailleurs, l'accès des chercheurs est rarement discuté. Ce sujet

devient pourtant déterminant au fil de la numérisation des usages et des échanges, marchands ou non. De fait, comment faire de la recherche sans avoir accès aux données ? Le silence auquel donne lieu cette question est assourdissant.

Alors que les collectivités locales doivent ouvrir leurs données publiques dans le cadre de l'open data, seules 10,96 % d'entre elles l'ont fait quatre ans après l'entrée en application de la loi pour une République numérique. Cela témoigne du chemin qu'il reste à parcourir, mais aussi des difficultés qui le jalonnent.

Comment organiser l'accès aux données ? Les portails d'ouverture de données posent la question de la licence, c'est-à-dire des conditions de réutilisation des données. Il existe différents types de licences, dont Etalab et Open Database Licence (ODBL). Cette dernière, utilisée par les collectivités françaises, impose des restrictions en matière de réutilisation des données, ce qui revient à limiter leur valeur future. Au niveau européen, la logique du *data altruism* revient à spécifier les bons usages futurs de la réutilisation des données. Cela pose de nombreuses questions, rarement évoquées. De fait, bloquer la valeur future des données pour des usages qu'on ne connaît pas est quelque peu utopique. C'est aussi méconnaître que la valeur dépend fondamentalement des usages.

Il importe également de définir les interfaces d'échange des données en temps réel, notamment dans le secteur du transport ou sur les marchés publicitaires ? Les API, que l'on utilise généralement, ont des coûts d'usage. Qui va les financer ?

Une autre façon d'imaginer l'accès aux données est l'interopérabilité, qui implique un accès à la circulation des données entre les plateformes. Or quel est le cadre juridique ? Le droit de la concurrence pourrait obliger à cette circulation. Mais dans le cas des réseaux sociaux, il se heurterait à plusieurs difficultés – ne serait-ce que parce qu'on est en présence de données personnelles, encadrées par le RGPD. Si des données personnelles collectées par un service A sont transférées à un service B, le nouveau responsable du traitement n'a pas collecté le consentement pour ce traitement. Il existe alors un risque de perte de contrôle de l'utilisateur quant au traitement de ses données, et une dispersion des responsabilités de traitement.

En somme, la simple question du « comment » en soulève d'autres, techniques et juridiques en l'occurrence. Il s'agit notamment d'articuler la liberté de choix et la protection de la vie privée. Le cadre adéquat n'existe pas encore complètement.

La question des formats de portabilité se posera aussi au fur et à mesure de l'élargissement de l'accès aux données et de leur interopérabilité. Qui définira les règles, les normes et les standards des données et des métadonnées pour faciliter les échanges ? Les plateformes dominantes imposeront-elles leur format ? Une instance devra-t-elle s'en charger ? Laquelle – le W3C ou un régulateur ? Ces sujets techniques sont souvent considérés comme secondaires. Ce sont pourtant eux qui rendront effectif ou non l'accès aux données.

Dans l'open data, on a ouvert des données. Mais celles qui ne sont pas normalisées ne peuvent pas être réutilisées. Les seules pour lesquelles il existe des standards,

aujourd'hui, sont les données de mobilité. Il faut aller très loin dans le détail et la nature des données pour que la régulation permette effectivement d'élargir l'accès aux données.

Finalement, la régulation n'arrive-t-elle pas trop tard ? Un récent article montre qu'il est désormais possible d'entraîner des algorithmes de *deep learning* avec dix fois moins de données. Compte tenu de la vitesse d'évolution et de progrès technologique, le sujet de l'accumulation de données par les big techs n'est-elle pas déjà un faux problème ?

Remarque

Jean-Denis Combrexelle

Nous avons soulevé toutes ces questions dans le cadre de l'élaboration de plusieurs lois, notamment relatives aux données de santé ou aux mobilités.

Le rapport « Politique de concurrence à l'ère du numérique »

Yves-Alexandre de Montjoye
Professeur associé, Imperial College (Londres)

Le rapport Crémer, Schweitzer, de Montjoye, sur la politique de concurrence à l'ère du numérique, montre combien il est essentiel d'aller dans le détail pour « designer » le fonctionnement et l'utilisation des données, en particulier l'interopérabilité et les types d'accès.

Il s'agit avant tout d'entrer dans le détail du type de données, tant elles sont variées. Données volontaires (contribuées par les utilisateurs), données observées (la grande majorité des données utilisées pour entraîner les algorithmes d'intelligence artificielle), données inférées (données brutes, observées ou contribuées, travaillées pour générer des profils d'utilisateurs ou des algorithmes de recommandation) : de quoi parle-t-on, et pour quoi faire ?

Il faut aussi détailler le type d'usages. Le rapport en distingue quatre : l'usage non-anonyme de données (pour offrir un service personnalisé à un individu), l'usage anonyme de données individuelles (les algorithmes d'intelligence artificielle, par exemple, n'ont pas besoin de données identifiées), l'usage de données agrégées de manière irréversible (statistiques nationales, par exemple), l'usage de données contextuelles (qui ne dérivent pas de données collectées au niveau individuel). La question de l'usage se pose d'autant plus que la très grande majorité des données est collectée au niveau des individus. Il s'agit donc de données identifiables et difficilement utilisables de manière anonyme.

Le rapport propose plusieurs solutions. La première concerne l'accès à des données personnelles et leur interopérabilité sous le contrôle de l'individu, indépendamment du lieu où elles sont stockées. Deux manières dont l'interopérabilité pourrait être imposée ont été étudiées : soit une régulation spécifique à certains secteurs, soit le régime de l'article 102 pour les entreprises à position dominante. Cette perspective semble globalement alignée avec la philosophie du RGPD, puisque l'extension de la portabilité se ferait sous le contrôle de l'individu. De nombreux écosystèmes disposent déjà de services interconnectés, avec des API privées : celles-ci pourraient-elles être ouvertes à la concurrence ? Cette solution peut avoir un impact à condition de s'assurer que l'API et les données sont entièrement accessibles, et pas seulement une petite fraction, mais aussi de s'assurer que l'API continuera à fonctionner dans le futur. En effet, dans le cadre de l'interopérabilité, un accès constant aux données est nécessaire.

Par ailleurs, certains algorithmes sont de plus en plus efficaces et ont moins besoin de données que par le passé pour s'entraîner. Quoi qu'il en soit, un très grand nombre d'algorithmes de *deep learning* continuent à avoir besoin de beaucoup de données, très riches. Dans ce cas, l'accès aux données est plus difficile, et nécessite des solutions techniques adaptées car la majorité de celles utilisées pour entraîner ces algorithmes sont personnelles, d'une manière ou d'une autre. Un récent article publié dans *Nature Communications* montre qu'il suffit de connaître 15 attributs pour identifier une personne de manière unique, dans n'importe quel dataset et dans 99,98 % des cas. Il est donc extrêmement difficile de déconnecter des données d'une personne.

Face à cet enjeu, le design de système, qui permet d'utiliser les données de manière anonyme, présente un fort potentiel. Ses systèmes ne sont pas parfaits comme le montre l'attaque d'un système de questions/réponses (Diffix) protégeant une base de données sensibles : avec 32 requêtes par utilisateur, l'information protégée par le système pouvait être extraite avec une précision de 91 % (*Usenix'19 - When the Signal is in the Noise: Exploiting Diffix's Sticky Noise*).

Les solutions d'interoperabilités sous le contrôle de l'individu présentent aussi un intérêt, dans le cas particulier de l'accès aux données pour l'entraînement des algorithmes d'intelligence artificielle. Des solutions *market-based* peuvent aussi présenter un intérêt.

La régulation de l'accès aux données

Fabienne Siredey-Garnier
Vice-présidente de l'Autorité de la concurrence

L'appréhension par le droit de la concurrence de la question de l'accès aux données détenues par les entreprises n'est pas née avec l'économie numérique et les plateformes et ne se réduit pas à elle. Elle peut donc aussi concerner des secteurs traditionnels.

La jurisprudence de l'Union européenne, notamment au travers des arrêts Magill de 1995, IMS Health de 2004 ou Microsoft de cette même année, témoigne de l'ancienneté de cette question. En France aussi, plusieurs décisions du Conseil puis de l'Autorité de la concurrence montrent que les outils classiques du droit de la concurrence – notamment la notion d'abus de position dominante – peuvent apporter une réponse aux difficultés liées à l'accès aux données. C'est ainsi le cas de la décision NMPP de 2003, (néanmoins cassée par la Cour de cassation), de l'avis du 14 juin 2010 sur l'utilisation croisée des bases de clientèle dans le secteur des télécoms ou encore des décisions Cegedim et GDF Suez de 2014. Plus récemment, l'avis sur l'ouverture à la concurrence des bus franciliens insiste sur la nécessité pour les candidats aux lignes concernées d'avoir accès aux données essentielles détenues par la SNCF et la RATP. La problématique des données était donc bien connue. Mais elle a pris de l'ampleur, d'une part avec l'économie numérique et les avancées technologiques qui ont élargi la nature, les sources et le volume des données, d'autre part avec le développement d'entreprise dont le modèle repose sur la collecte et l'exploitation massive de données, souvent personnelles.

Cet essor des plateformes et de l'économie numérique pose de nouveaux défis pour l'encadrement de l'accès aux données. La collecte massive de données confère-t-elle un avantage concurrentiel durable ? Cette question controversée a été traitée dans de nombreux avis et rapports récents. La Commission européenne et les autorités de concurrence ont d'abord eu une position nuancée. Encore en 2016, Margrethe Vestager considérait qu'il n'était pas vraiment besoin de construire un droit spécifique à l'accès aux données, observant qu'on n'en avait pas élaboré pour les cartes de crédit ou les fax… Il convient également de relativiser les inquiétudes liées à la collecte massive, le principal défi n'étant pas la détention des données mais la possibilité de les exploiter intelligemment. Au demeurant, dans ses avis de 2016 et 2018, l'Autorité de la concurrence a mis en lumière que l'acquisition de données pouvait être complexe – notamment du fait des règles de confidentialité et sans compter les réticences de certains opérateurs à communiquer leurs données – et coûteuse.

Plusieurs décisions, prises ou en cours, s'appuient sur des notions bien connues du droit de la concurrence pour appréhender le comportement des grandes plateformes. Deux volets sont à distinguer : celui des pratiques anticoncurrentielles, par le biais de l'article 102, et celui des fusions – plusieurs d'entre elles ayant retenu l'attention en raison de leur nature et leurs montants, même si elles ont, jusqu'à présent, toutes été autorisées.

Le droit de la concurrence n'est donc pas obsolète et peut saisir le digital. Toutefois, des difficultés demeurent. Le fait que les données n'ont pas de valeur économique stricto sensu et que les marchés sont bifaces et globaux, par exemple, perturbe les grilles d'analyse traditionnelles. Aussi est-il nécessaire que les autorités de concurrence aient une réflexion sur une éventuelle mise à jour de leurs méthodes et outils, à la fois entre elles et avec les autorités sectorielles. Le consensus est fort, en la matière. L'an dernier, d'ailleurs, Margrethe Vestager a considéré que

« *les préoccupations relatives aux données, aux plateformes et aux stratégies d'acquisition pourraient justifier de nouvelles règles de concurrence* » et que « *nous devons dire aux grandes plateformes que les choses vont changer* ». Le ton est plus martial qu'en 2016 !

Faut-il une régulation ex ante ou ex post ? En tout état de cause, il faut une combinaison d'outils procéduraux, au premier rang desquels le recours accru aux mesures conservatoires et à l'article 22 du règlement Concentration. Par ailleurs, nous avons besoin d'ingénieurs. Aussi l'Autorité de la concurrence vient-elle de se doter d'un service de l'économie numérique, dans lequel elle a recruté des data scientists.

Une réflexion devra également être conduite sur l'élargissement possible des concepts classiques (position dominante, marché pertinent, infrastructure essentielle) qui peuvent paraître mal adaptés car trop restrictifs. Le Parlement européen s'est prononcé en ce sens le 21 octobre 2020. L'Europe a fait plus que tout autre continent pour réguler le monde numérique, mais comme l'a déclaré un parlementaire européen, « même ici c'est encore le *Far West* ». Le New Competition Tool, notamment, permettrait de remédier à des problèmes de concurrence structurelle qui ne peuvent pas être résolus avec les outils actuels.

Pour citer la commissaire Jourová, il est temps de « *faire entrer le tigre dans sa cage* ». Nous ne le ferons que si les autorités de régulation, sectorielles ou de concurrence, nationales et européennes voire mondiales, coopèrent entre elles.

Echanges

Jean-Yves Ollier

Que faire lorsqu'une plateforme fait écran à l'accès des vendeurs aux données relatives à leurs clients, ou des annonceurs ou des entreprises de publicité à certains éléments qui leur permettraient d'appréhender l'impact de publicités ciblées ? Le règlement P2B envisage une solution peu contraignante, en prévoyant que les conditions d'utilisation soient explicites sur ce point. Peut-on aller plus loin, par exemple en imposant de demander à l'utilisateur s'il accepte que ses données soient fournies au vendeur ultime ? Ou bien peut-on imaginer des obligations de transparence, à l'instar de l'autorité australienne de la concurrence et de la protection des consommateurs qui envisage un retour systématique des plateformes sur la façon dont les données ont été utilisées ? Comment articuler ces enjeux concurrentiels et l'exigence de consentement ?

Henri Isaac

Le recueil du consentement peut être organisé. Il n'est d'ailleurs un obstacle que s'il n'est pas organisé. Le régulateur peut aussi l'imposer. La transparence vis-à-vis de l'utilisateur est fondamentale. Sinon, ce serait nier la philosophie du RGPD qui cherche à redonner la main aux consommateurs et aux citoyens sur leurs propres données.

Fabienne Siredey-Garnier

Dans sa décision Facebook, le Bundeskartellamt subordonne l'obligation pour la plateforme de communiquer ses données au consentement exprès du consommateur. C'est la voie dans laquelle nous devrons nous engager.

Yves-Alexandre de Montjoye

L'interopérabilité des données serait une extension du droit à la portabilité, avec le consentement de l'utilisateur et sous son contrôle. Les données observées et les données contribuées peuvent et devraient sans doute être rendues disponibles en temps réel. Pour les données inférées, il faudrait en général sans doute privilégier le cas par cas. Les questions de la sécurisation de l'accès et de la répartition des coûts devront aussi être posées.

Du point de vue droit de la concurrence, la principale option à considérer est celle de l'utilisation anonyme des données – lesquelles ne sont alors plus couvertes par le RGPD. Il convient alors de s'assurer, sur le plan technique, qu'elles sont effectivement utilisées de manière anonyme et que des informations privées et sensibles ne peuvent pas être extraites.

Le design du système est primordial dès lors que l'accès aux données est rendu obligatoire.

Henri Isaac

Le risque existe de compliquer l'entraînement des programmes de *machine learning*, en cas d'anonymisation des données par des technologies d'obfuscation. Le service proposé par celui qui est à l'origine des données disponibles ne sera donc peut-être pas tout à fait le même pour les tiers.

Yves-Alexandre de Montjoye

Dans les cas d'entraînement distribué, les algorithmes ont accès aux mêmes données que le système originel.

Jamal Atif

Dans le *machine learning*, il est possible de masquer, crypter ou bruiter les données tout en contrôlant les performances de l'algorithme. Aux États-Unis, la *differential privacy* est utilisée en pratique pour fluidifier la donnée sans compromettre la vie privée ou la précision de l'algorithme. Qu'en pensez-vous ?

Yves-Alexandre de Montjoye

La *differential Privacy* est une garantie très forte de protection de la vie privée. Elle fonctionne très bien pour la protection d'agrégats, mais elle s'avère en général moins utile dans d'autres cas, par exemple d'accès par des systèmes de questions/réponses.

De la salle

Que pensez-vous du projet Gaia-X, qui a pour vocation de créer un cloud souverain?

Henri Isaac

L'initiative est intéressante, mais dans de nombreux secteurs, par exemple celui de la mobilité, les besoins en puissance de calcul en temps réel sont colossaux et les volumétries concernées sont uniquement calculables sur des clouds américains ou chinois, aujourd'hui.

Joëlle Toledano

Quel est le délai nécessaire pour faire évoluer les outils procéduraux et les concepts du droit de la concurrence ? Dans quelle mesure ces évolutions deviendront-elles structurelles, et à quel horizon ?

Fabienne Siredey-Garnier

Les mesures conservatoires sont de plus en plus utilisées. Elles le seront plus encore quand la directive ECN+ sera transposée, puisque nous pourrons nous saisir d'office. Qui plus est, le DSA devrait être adopté rapidement. L'approche des concentrations et l'article 22 devraient également se concrétiser prochainement.

De la salle

Le modèle des tiers de confiance a été moins évoqué. Pourtant, ces acteurs qui fluidifient le marché peuvent aussi être intéressés par l'accès aux données issues de différentes sources.

En France, par exemple, l'autorisation de la publicité télévisée segmentée a soulevé des questions sur la mesure de l'audience. Jusque-là, Médiamétrie avait une sorte de monopole naturel dans la production de ces données qui bénéficiaient ensuite à l'ensemble du marché. Or avec le morcellement lié à la segmentation, chaque fournisseur d'accès à internet pourrait fournir sa partie de mesure de l'audience, réduisant de fait l'optimalité de la mesure collective produite par Médiamétrie. Pour sa part, l'équivalent du CSA canadien impose aux producteurs de ces données de les fournir à un tiers de confiance, qui recueillerait l'ensemble des données du marché dans un but d'optimalité.

Henri Isaac

Ce modèle de notariat n'est pas nouveau. Le tiers de confiance en question effectue, de façon neutre, de la réagrégation de données. C'est une piste intéressante, à condition que cette instance normalise les données. Une métrologie commune est indispensable.

Clôture

Cédric O
Secrétaire d'État chargé de la transition numérique
et des communications électroniques

Le DSA, qui sera présenté en décembre prochain, marquera une étape supplémentaire dans la régulation des grandes plateformes de manière générale et, en particulier, sous l'aspect des données.

Nous nous accordons tous à reconnaître l'importance de la donnée dans l'économie de l'attention, de la gratuité et du service personnalisé. La donnée est même probablement devenue le premier sous-jacent de valeur économique dans l'estimation de la puissance d'une entreprise. Désormais, les entreprises qui valent le plus d'argent – Google, Facebook et Amazon aux États-Unis, Alibaba et Tencent en Chine – sont celles qui ont la capacité de capter et valoriser le plus grand nombre de données.

Au cours des vingt dernières années, ont émergé des entreprises d'une taille et d'une empreinte économique et démocratique inédites. Elles posent des problèmes juridiques, économiques et de souveraineté, mais reposent sur l'économie de la donnée, en partie de manière prospective. En effet, il est fort probable que ce qu'elles savent déjà faire avec la donnée soit assez limité au regard de ce qu'elles pourraient faire. Ainsi, dans la valorisation qui est la leur, la partie prospective n'est pas négligeable.

En corollaire, les chaînes de valeur traditionnelles de certains secteurs comme la banque, l'automobile ou l'agriculture se sont modifiées pour s'aligner sur le modèle de ces nouveaux acteurs – ce qui ne va pas sans poser un certain nombre de questions. Cela, à tel point que les banques redoutent l'arrivée des GAFA qui connaissent finalement mieux leurs clients qu'elles-mêmes. Autre exemple, le secteur automobile entrevoit que grâce à sa capacité à capter de la donnée, la voiture autonome ne sera plus qu'un vecteur de déplacement, donc de temps libre et de loisirs personnalisés sur le marché de la mobilité.

Par ailleurs, bien qu'elles ne constituent qu'une partie du sujet, les données personnelles sont celles qui posent le plus question, en matière juridique mais aussi ontologique. En tout cas, ce sont celles qui ont fait l'objet des premières batailles. Les données industrielles sont un sujet tout aussi essentiel, mais elles demeurent devant nous.

L'évolution des modèles d'affaires et de la structure de notre économie – et même de nos démocraties – nécessite sans aucun doute d'être régulée. Le RGPD a, de l'avis unanime, a marqué un pas essentiel dans la conceptualisation et la mise en œuvre d'une politique publique de régulation de la donnée. Après avoir été bloqué par une partie du monde, il fait quasiment consensus au sein des démocraties. Il est

devenu une sorte de standard international, copié à la fois par des pays asiatiques et des États américains. Espérons qu'il le sera aussi par le niveau fédéral.

Quelle suite lui donner, dans la dialectique entre données et innovations ? Pour commencer, le RGPD est-il suffisant ? Il me semble que non. Certes, il a représenté un pas essentiel, notamment parce qu'il était le premier texte de ce type et parce qu'il a inversé un certain nombre de logiques. Il a permis de donner plus efficacement la main aux citoyens et aux consommateurs sur leurs données personnelles. Chacun peut le constater au quotidien. En revanche, il n'a pas mis un terme à la mainmise de très grandes entreprises et de certains modèles d'affaires particuliers, bien au contraire. La meilleure manière de s'en rendre compte est peut-être de considérer l'évolution ininterrompue de la capitalisation boursière de ces acteurs depuis l'adoption du RGPD. Ainsi, comme toute législation pionnière, le Règlement nécessite d'aller plus loin.

Sur le plan de la recherche et d'un point de vue juridique, des questions de plus en plus essentielles voient le jour autour de la propriété des données. Peut-être fonderont-elles la base d'une politique de la donnée ? Dans les grandes lignes, il s'agit de savoir qui est le propriétaire des données produites. Le droit, qui extrêmement clair sur ce point, postule que le propriétaire de la donnée est le service qui a été à l'origine de sa création. Quand une personne se déplace en utilisant un outil de cartographie via son téléphone, par exemple, celui-ci n'est pas propriétaire de la donnée de déplacement ainsi générée. En revanche, la valorisation de cette donnée est à la main du service utilisé – si tant est que la personne lui en ait donné l'autorisation, conformément au RGPD. De fait, la logique de valorisation étant à la main du service utilisé, les algorithmes peuvent s'entraîner à valoriser, revaloriser et valoriser de nouveau les données.

Certains travaux réfléchissent au renversement de cette logique, en considérant qu'on ne ferait que louer ses données personnelles à une entreprise qui devrait à un moment les rendre. Il s'agit là d'un élément de prospective très poussé, qui mérite d'être approfondie. Il pose, en tout cas, plusieurs questions plus ou moins connexes, à commencer par des questions de faisabilité, notamment dans le cadre de l'utilisation et la réutilisation des données. Par exemple, si une entreprise entraîne un algorithme à partir de données personnelles, l'individu propriétaire de ces données devrait théoriquement percevoir une part de la valeur créée par cet algorithme – cette part étant fonction de l'importance de ses données dans le jeu complet. Mais quand l'algorithme est réutilisé et croisé avec d'autres, eux-mêmes issus d'autres données donnant lieu à une construction extrêmement complexe, comment savoir quelle donnée a servi à quel effet ?

Il serait également intéressant de travailler sur la progression de la notion de données d'intérêt général, dans ce monde oligopolistique, au regard de l'importance de la donnée pour la gouvernance publique ou dans le cadre de la gestion d'une pandémie – soit parce que certains services sont devenus essentiels, soit parce qu'ils sont uniques. Dans des cas de figure très encadrés par la loi, les données considérées d'intérêt général devraient donner lieu à une régulation. En l'occurrence,

certaines plateformes dites structurantes sont devenues tellement importantes qu'elles doivent être considérées comme fournissant une infrastructure essentielle donnant lieu à une régulation spécifique.

Néanmoins, la dialectique est probablement plus compliquée, car les données sont essentielles à l'innovation. Un corollaire peut ainsi être établi entre le nombre de données à disposition d'un service et sa capacité d'innovation et de création de valeur. Cela crée parfois une asymétrie de concurrence entre acteurs de différentes régions du monde. Quand des acteurs chinois ou américains ne subissent pas la même régulation que les acteurs européens sur leurs terres nationales, ils peuvent entraîner leurs algorithmes grâce à l'acquisition de données. Dans certains cas extrêmes, liés à la surveillance, certains procédés seraient inacceptables dans une société occidentale. Aussi la question de la donnée doit-elle conduire à repenser certaines règles internationales et la manière d'approcher l'ouverture des marchés. Si nous souhaitons des réglementations qui se conforment à nos valeurs et n'aient pas pour corollaire un désavantage compétitif écrasant, il faut tout repenser dans le cadre d'une politique commerciale, à l'instar de ce qui se passe en matière d'environnement.

Nous sommes pris dans un dilemme. Etant donné que nos réglementations sont – et c'est heureux – plus dures que celles des pays extra-européens, et que nous ne fermons pas nos marchés à des services développés sur des jeux de données dont nous considérons que le fondement n'est pas éthique, il existe une asymétrie en matière de compétitivité internationale. À l'avenant, plus nous augmentons les barrières et plus nous régulons la donnée, plus nous rendons une partie de l'innovation difficile et plus les acteurs les mieux à même de s'adapter aux règles européennes – donc d'être à la fois compétitifs et innovants, tout en respectant la régulation à la lettre – seront les plus puissants et posséderont la plus grande empreinte. Dans cette dialectique de la « régulation de la régulation », nous savons que mettre des barrières revient souvent à figer la compétition telle qu'elle existe, comme cela a été le cas dans le secteur financier. Probablement le RGPD a-t-il renforcé les très grandes entreprises américaines du numérique. De fait, elles sont les seules à pouvoir assurer sur l'ensemble du territoire européen une conformité à peu près parfaite avec cette réglementation, ne serait-ce que grâce à l'ampleur de leurs services juridiques.

Cette question de la régulation a d'ailleurs été au cœur des discussions européennes sur la directive *ePrivacy*. Il s'agissait, de façon très simple et en toute bonne volonté, de redonner la main aux Européens sur leurs données, en interdisant un certain nombre de pratiques comme celle des murs de cookies. Mais in fine, interdire une approche de captation de la donnée et de développement de services sur la connaissance de l'internaute a renforcé les acteurs qui n'en avaient pas besoin et bénéficiaient déjà d'un écosystème immersif, en l'occurrence les GAFA. En somme, dans cette économie de la donnée, il est nécessaire de conduire une réflexion holistique quant aux conséquences réelles de réglementations qui vont dans le bon sens, mais peuvent également avoir pour effet de renforcer les plus puissants.

C'est d'autant plus vrai que la question de la régulation de la donnée telle que nous la concevons jusqu'ici se fonde sur une approche du consommateur en léger décalage avec la réalité. Au sein d'un même individu, la scission entre le consommateur et le citoyen est de plus en plus forte : autant le citoyen peut émettre des réserves sur le comportement de certains acteurs, autant il est évident que le consommateur apprécie grandement ces acteurs et la qualité de leurs services. Dans les faits, cela conduit à constater un renforcement constant de la compétitivité et de l'empreinte des grandes entreprises du numérique – notamment américaines – sur notre société et notre économie. En caricaturant un peu, le rêve démocratique d'un consommateur ultra-éduqué et lucide qui inverserait la dialectique actuelle ne deviendra pas réalité avant très longtemps.

Dans ce contexte, l'un des apports du RGPD concerne la transparence, même si la plus grande utilité du dispositif n'est peut-être pas pour le citoyen lui-même – car je ne suis pas totalement certain que nous fassions systématiquement une inspection très détaillée des conditions générales d'utilisation de nos données avant d'accepter les pop-ups des sites internet que nous visitons. Je ne pense pas non plus que chaque Français passe un temps infini à étudier les données enregistrées dans son compte Gmail, pour les paramétrer. Néanmoins, l'intérêt évident du RGPD reste celui de la transparence à l''égard de tierces parties, qui peuvent faire ce travail pour la démocratie.

Au total, si la régulation des données est absolument essentielle, elle doit être pensée en coordination avec des réflexions sur la donnée en soi et sur ce qu'elle est, c'est-à-dire sur la transférabilité réelle de la donnée. Il ne suffit pas d'affirmer qu'il faut mettre à disposition de l'utilisateur l'ensemble de ses données personnelles créées par un service. Il faut aussi et avant tout poser la question de l'interopérabilité.

En outre, il importe de non seulement penser la régulation de la donnée en soi, mais en lien avec celle des modèles d'affaires. Je salue, à cet égard, les travaux de Joëlle Toledano en la matière. Le fond du sujet est la donnée, et plus encore le lien entre la donnée et la gratuité. Dans un monde toujours plus financé par la publicité, nous sommes lancés dans une course en avant, quasiment démiurgique, de déterminisme de comportements de consommateurs menant à capter un maximum de données. L'objectif étant d'entrer dans un monde dans lequel la façon dont vous mangez ou vous vous habillez déterminera ce que vous voterez aux élections cantonales. Il y a là une volonté quasiment théologique et très consumériste ! À terme, nous ne nous exonérerons pas d'avoir une réflexion sur le sujet. La gratuité doit être questionnée, car je considère que tout part de là.

Enfin, la réflexion doit aussi porter sur la capacité technique de la puissance publique. Une des difficultés de l'économie de la donnée est sa complexité technique, compte tenu de l'intrication des données et de l'utilisation d'algorithmes de plus en plus performants. C'est pour cela qu'au sein de l'État, nous avons créé un pôle de savoir-faire algorithmique, partagé avec un certain nombre de régulateurs RGPD : le PEReN, pour Pôle d'expertise de la régulation numérique. La question de

la capacité technologique de la puissance publique est essentielle pour décoder et aller vers plus de transparence au sujet des plateformes. Nous ne sommes pas encore au bon niveau sur le plan des capacités technologiques. En tout cas, nous ne sommes pas au même niveau que les acteurs que nous prétendons réguler.

Je ne vous ai parlé que des données personnelles, car il me semblait que c'était l'un des éléments les plus intéressants dans la réflexion sur la dialectique entre régulation, gouvernance et innovation. Mais il est évident que l'un des autres horizons que nous devons aborder est celui des données industrielles. Je sais que cela tient énormément à cœur de la Commission européenne, en particulier de Thierry Breton.

En matière économique, il est probable que la France et l'Europe aient une bataille de retard sur les données personnelles. Ce n'est pas encore le cas s'agissant des données industrielles, compte tenu de la taille des acteurs et de leurs capacités – dans l'environnement, les transports, l'industrie ou la banque. Ce potentiel économique extrêmement important pose, lui aussi, nombre de questions juridiques sur le fondement desquelles tout reste à bâtir et dont il faudra se saisir dans les mois et les années qui viennent. Faute de quoi, ce seront toujours les mêmes qui s'en saisiront !

Echanges

Bruno Lasserre

Pourra-ton progresser en la matière sans le faire sur l'harmonisation fiscale et la construction d'un marché dans lequel les États membres ne seront pas en concurrence sur le moins-disant fiscal pour attirer ces plateformes ? Si la régulation européenne délègue la mise en œuvre de ces règles aux pays d'implantation des sièges de ces plateformes plutôt qu'au pays dans lequel se trouve le public destinataire de ces règles, ne risque-t-on pas d'avoir aussi un moins-disant régulatoire ? Les GAFA s'installeront alors dans le pays où la fiscalité et les règles sont les plus douces.

Cédric O

La régulation dans le monde numérique et la création d'un cadre harmonisé représentent l'un des plus grands défis de l'Europe. Le sujet n'est pas uniquement fiscal. Certes, le volet fiscal est primordial symboliquement et politiquement. Mais la question de la régulation des données personnelles est absolument essentielle dans la réflexion européenne.

Un site d'e-commerce français doit effectuer au minimum deux clics pour respecter l'interprétation que la Cnil fait du RGPD. En Irlande ou au Luxembourg, un clic suffit. Cela peut sembler anecdotique, mais la différence de chiffre d'affaires peut aller jusqu'à 20 ou 30 %. Et ce, alors même que le RGPD est censé assurer un niveau maximal d'harmonisation européenne. Évidemment, il existe des mécanismes d'appel et de mise à niveau. Mais s'il faut deux ou trois ans à l'organisme

de coordination des Cnil européennes pour rappeler à l'ordre ou harmoniser l'interprétation du RGPD, des empires se seront construits entre-temps et des entreprises seront mortes, dans un monde où le gagnant prend tout le marché.

S'agissant de la fiscalité, la situation n'est pas tenable à long terme. C'est une discussion que nous devons avoir avec les Européens. Le corollaire de la création d'une régulation unifiée au niveau européen, dépassant de fait la question du pays d'origine et de celui de destination, est un dessaisissement du régulateur français ou de l'institution française au bénéfice d'une institution supranationale. Cela pose évidemment d'autres questions, qui ne plaisent pas toujours aux États, aux parlementaires ou aux régulateurs, mais qui sont indispensables pour créer un marché unifié.

Bruno Lasserre

Tout ce que vous avez tracé alimentera d'autres colloques dont le Conseil d'État reste candidat à l'organisation, tant ces questions bouleversent nos métiers, mais aussi les affaires que nous avons à juger. Je citais ce matin l'ordonnance du juge des référés sur le *Health Data Hub*. Le juge administratif doit se préparer à recevoir d'autres affaires de ce type.

Chaire Gouvernance et Régulation
Fondation Paris-Dauphine
Place du Maréchal de Lattre de Tassigny - 75016 Paris (France)
http://chairgovreg.fondation-dauphine.fr