Compte rendu

Commission de réflexion et de propositions sur le droit et les libertés à l'âge du numérique

Jeudi
18 décembre 2014
Séance de 8 heures 30

Compte rendu n° 10

SESSION ORDINAIRE DE **2014-2015**

Présidence de Mme Christiane Féral-Schuhl,
coprésidente
Et de
M. Christian Paul,
coprésident

COMMISSION DE RÉFLEXION ET DE PROPOSITIONS SUR LE DROIT ET LES LIBERTÉS À L'ÂGE DU NUMÉRIQUE

Jeudi 18 décembre 2014

La séance est ouverte à huit heures quarante.

(Présidence de Mme Christiane Féral-Schuhl, co-présidente et de M. Christian Paul, co-président)

Table ronde sur les « données personnelles et les activités économiques » avec M. Benoît Tabaka, directeur des politiques publiques de Google France, M. Paul-Olivier Gibert, directeur de Digital & Ethics, et M. Pierre Bellanger, fondateur et président-directeur général du groupe Skyrock

Mme Christiane Féral-Schuhl, coprésidente. Nous avons le plaisir d'accueillir, pour cette table ronde consacrée aux données personnelles M. Benoît Tabaka, directeur des politiques publiques de Google France et, rappelons-le, premier récipiendaire du prix de l'ADIJ, l'Association de juristes passionnés de technologie, M. Paul-Olivier Gibert, directeur de Digital & Ethics, cabinet spécialisé dans la gestion des risques et la protection des données personnelles, et M. Pierre Bellanger, président-directeur-général de Skyrock

Nous souhaitons vous interroger, messieurs, sur les principes directeurs qui doivent, selon vous, guider la protection de la vie privée et l'utilisation des données personnelles et sur l'opportunité d'une transformation du régime de protection actuel.

M. Pierre Bellanger, président-directeur-général du groupe Skyrock. Pour cette intervention liminaire, je m'appuierai sur la contribution que j'ai faite à l'étude du Conseil d'État sur le numérique et les droits fondamentaux.

Les données personnelles sont le plus souvent considérées comme des entités autonomes à la manière des fiches individuelles cartonnées des fichiers du XX^{e} siècle. Cette conception ne correspond pas à la réalité. Il faut se les représenter non plus comme des billes dans un sac de billes mais comme une pelote de laine. Elles ne sont pas isolables en pratique. Intriquées les unes aux autres, elles forment un tout indissociable. Elles ne renseignent pas sur un seul individu mais sur un réseau d'individus. Ainsi lorsqu'en souscrivant à une application, vous donnez l'autorisation d'accéder à votre carnet d'adresses, vous donnez accès aux données personnelles d'autrui. De plus, les algorithmes permettent, à partir des données des uns, de prédire les données des autres.

Nous sommes face à un nouvel objet juridique : les données personnelles ne sont plus granulaires mais réticulaires. Toute la question est de savoir comment traiter ce réseau de données car elles peuvent donner lieu tant à des droits individuels – droit de modification, droit de retrait, droit de consultation – qu'à des droits collectifs. Leur organisation en réseau en fait un bien collectif : elles appartiennent à tous tout en appartenant à chacun, ce que l'on pourrait comparer, d'un point de vue juridique, à l'indivision. Ces caractéristiques appellent une nécessaire intervention de la puissance publique.

Prenons l'exemple des données relatives à la circulation automobile et au stationnement, dont on sait l'importance étant donné qu'un tiers de la consommation d'essence est consacré à la recherche d'une place disponible. À qui appartiennent ces données ? À celui qui les collecte ou à tous, sachant qu'elles ont une utilité collective ?

C'est un piège de raisonner en termes de données personnelles individuelles. Celles-ci ne sont qu'un cas rare faisant partie d'un ensemble plus vaste, comme l'espace newtonien est un cas de figure de l'espace einsteinien. L'un des principes directeurs de notre réflexion est de garder toujours à l'esprit que les données personnelles sont organisées en réseau, notamment dans nos échanges avec les grandes entreprises du numérique. Si nous continuons à réfléchir en termes de données granulaires, nous perdrons notre souveraineté : grain de sable par grain de sable, toute la plage disparaîtra. En revanche, si nous réfléchissons en termes de réseau global, il sera plus difficile de faire disparaître une plage entière d'un coup.

M. Paul-Olivier Gibert, directeur de Digital & Ethics. Je partage en partie la position de Pierre Bellanger. Depuis les années soixante-dix où ont été élaborés les dispositifs de protection des données à caractère personnel, nous avons assisté à un changement radical : d'un découpage de l'information en données, nous sommes passés à un traitement de l'information à partir de gisements de données.

Nous devons repenser la problématique des données personnelles, non que la protection de la vie privée ne constitue plus un enjeu, mais parce que les modalités de la relation entre l'individu et les données qui le concernent sont en train de changer. La loi du 6 janvier 1978, construite en opposition au projet de système automatisé pour les fichiers administratifs et le répertoire des individus, dit projet SAFARI, a été votée par des personnes très marquées par le souvenir de la seconde guerre mondiale. Il était inimaginable à l'époque que les individus contribuent eux-mêmes à la production des données, comme c'est le cas aujourd'hui.

Prenons l'exemple des applications relatives à la circulation automobile en Ile-de-France. Elles utilisent des données produites par les individus. Certes, ils ne les produisent pas seuls : l'évaluation du temps de trajet suppose qu'un opérateur de télécommunication collecte et traite, à travers les algorithmes, les signaux envoyés ; toutefois, sans les déplacements des individus, l'information n'existerait pas.

Ce nouveau schéma conceptuel commence à rentrer de manière timide dans le projet de règlement européen sur la protection des données à caractère personnel.

Faut-il beaucoup de droit pour encadrer et protéger ? Je n'en suis pas si sûr. À la suite du doyen Carbonnier, je dirai qu'il ne faut pas trop de droit. On peut se demander si la société française actuelle n'est pas tombée dans l'excès : une activité sociale semble n'avoir d'existence légitime qu'à condition d'être couverte par une réglementation. Dans ce contexte, on voit tout l'intérêt qu'il peut y avoir à développer la dimension éthique par laquelle un organisme fixe des règles d'usage et veille à assurer leur respect.

M. Benoît Tabaka, directeur des politiques publiques de Google France. Les débats sur l'économie numérique se résument souvent à un acronyme, les GAFAM : Google, Apple, Facebook, Amazon et Microsoft. Or l'économie numérique recouvre un champ bien plus vaste : l'innovation portée par les données, personnelles ou non, la *data-driven innovation*. Il est très difficile d'évaluer les contours de cette économie car elle est de plus en plus ancrée dans l'économie classique. Un rapport publié par le cabinet de conseil McKinsey

a montré qu'une analyse massive des données permettrait aux administrations publiques européennes de réaliser 250 milliards d'économies, fondées principalement sur la lutte contre la fraude et l'optimisation des services publics.

De plus en plus d'entreprises ont pris conscience de cette bascule et du besoin de rentrer dans l'ère numérique, qu'elles appartiennent aux secteurs de l'industrie automobile, de la banque, des assurances, du voyage ou des transports. Les constructeurs automobiles ont ainsi développé des voitures intelligentes munies d'une multiplicité de capteurs ; ils ont même mis au point des systèmes de reconnaissance morphologique du conducteur. Les données relatives à la consommation énergétique sont appelées à connaître un grand nombre d'applications, notamment à travers les outils de gestion mis au point par les start-up spécialisées dans la maison connectée. Citons encore les moteurs d'avion, parmi les premiers instruments producteurs de données, à destination des équipages, des compagnies aériennes ou encore des constructeurs.

Ce secteur économique que représente les données a donné lieu à l'émergence de nouveaux acteurs en France comme *Criteo*, leader mondial de la publicité ciblée, ou *BlaBlaCar*, spécialiste du covoiturage, dont nous avons pu mesurer l'essor au fait qu'il fait partie des dix termes les plus recherchés sur Google.

S'agissant de *Google*, il faut savoir que le moteur de recherche développé par Larry Page et Serguei Brin et leurs équipes a été créé à une période où les données commençaient à peine à être exploitées par *Yahoo* et *AltaVista*. Son activité principale est d'organiser les informations de manière à les faire correspondre aux recherches des internautes. Des robots d'indexation, appelés *spiders*, analysent les contenus présents sur internet pour identifier les mots-clefs et les mettre en relation afin de calculer la pertinence de telle ou telle page par rapport à une requête donnée. En outre, ils détectent les sites infectés ou les activités de *phishing* et diffusent des mises en garde, qui contribuent à renforcer la sécurité. Par ailleurs, *Google* a mis en œuvre un système d'autocomplétion, *Autocomplete*, qui, grâce à l'agrégation des requêtes des internautes couplée aux premiers chiffres de l'adresse IP, complète automatiquement les mots recherchés selon les aires géographiques. Autre illustration de l'intérêt de la conservation des requêtes : *Google Flu Trends*, outil qui permet de suivre la progression de la grippe à partir de l'analyse de certains mots-clefs.

L'économie portée par les données a été à l'origine de divers modèles économiques : la vente de produits et de services, comme sur *GooglePlay*, mais principalement la publicité ciblée, qui se veut la plus pertinente et la moins invasive possible, à l'instar d'*Adwords*, qui fait apparaître des bannières et des annonces en fonction du mot-clef tapé par l'utilisateur dans le moteur de recherche. Citons encore les systèmes de détection d'œuvres protégées sur les sites d'hébergement de vidéos tels que *YouTube* ou *Dailymotion* qui donnent lieu à une monétisation pour les ayants droit par le biais d'un pourcentage sur les revenus publicitaires générés par telle ou telle vidéo, par exemple si une chanson célèbre est identifiée sur une vidéo de fête d'école.

Un des éléments clefs de l'économie numérique est la confiance. Les acteurs de ce secteur ont l'obligation de susciter et de conserver la confiance des utilisateurs car, contrairement à des secteurs plus traditionnels comme la téléphonie ou la banque, il est très facile pour eux de passer d'un service à un autre, qu'il s'agisse d'un moteur de recherche ou d'un réseau social.

Deux éléments contribuent fortement à installer cette confiance. Premièrement, il s'agit d'offrir aux utilisateurs la capacité de contrôler leurs données en leur donnant la possibilité de les exporter dans un format ouvert, par exemple, en déplaçant tous leurs mails vers un autre service de messagerie. Deuxièmement, il s'agit de veiller à la sécurité des données personnelles car sans sécurité des données, la protection de la vie privée ne peut être assurée.

Mme Christiane Féral-Schuhl, coprésidente. L'une de nos interrogations porte sur le renforcement de la protection des internautes et sur les notions de *privacy by design* et de *privacy by default*.

Monsieur Bellanger, Skyrock s'adresse à une population jeune, souvent peu consciente de la portée des informations qu'elle rend publiques. Comment gérez-vous de manière concrète les déréférencements et les demandes de retrait de contenus ?

M. Pierre Bellanger. Skyrock.com est le premier réseau social de blogs en France, l'un des premiers réseaux sociaux français, le premier réseau social francophone. Il touche une population plutôt jeune mais, en matière de numérique, qui dit population jeune ne dit pas forcément population mal informée. Cela dit, indépendamment du niveau d'information des utilisateurs, des impératifs de protection s'imposent. Notre plateforme s'est inscrite dans la continuité de notre ADN : la liberté d'expression populaire. Elle a pour slogan « Ici T libre », liberté que nous avons traduite dans nos conditions générales.

Elle se manifeste d'abord par la possibilité d'exister sous plusieurs identités, de manière multi-contextuelle, grâce à des pseudonymes, quand d'autres réseaux sociaux s'appuient sur le recours à l'identité réelle, avec des techniques d'une précision toujours plus grande, comme l'étiquetage nominatif des photos.

Elle se manifeste ensuite par la possibilité d'effacer le contenu d'un blog, y compris lorsqu'il y a eu des republications sur d'autres blogs, ce qui est un autre avantage, surtout pour les adolescents qui passent par des passions différentes en grandissant. Nous ne solidifions pas le passé en organisant sa persistance. De surcroît, les internautes demeurent intégralement propriétaires des contenus. Nos conditions générales ne comportent pas ces clauses d'acceptation – « *I accept* », « *I agree* » – dont les contrats anglo-saxons sont remplis.

Ce qui nous a permis de demeurer un acteur majeur de l'Internet français et européen, c'est cette liberté que nous offrons par rapport à d'autres acteurs dont les services sont beaucoup plus contraignants.

Mme Christiane Féral-Schuhl, coprésidente. Et qu'en est-il des préjudices que peuvent subir d'autres personnes du fait de la publication de certains contenus ? La violence peut avoir des conséquences désastreuses chez les adolescents, nous le savons. Avez-vous été contraints de retirer certains contenus ? Agissez-vous indépendamment des notifications que vous avez pu recevoir ?

M. Pierre Bellanger. Créé en 2002, Skyrock a été l'un des premiers réseaux sociaux créés dans le monde. En douze ans – autrement dit une longue période en temps Internet –, nous avons été confrontés à toutes sortes de situations. Il y a le cadre légal, auquel nous nous conformons, bien évidemment. Il y a les demandes de la justice, auxquelles nous répondons en révélant certains contenus dans le cadre de commissions rogatoires. Il y a aussi toute cette zone grise que vous évoquez, madame la présidente. Deux services nous aident à traiter les

situations qui en relèvent. La cellule psychologique dirigée par Michael Stora permet de détecter les personnes en souffrance qui, sur le réseau, disent à tout le monde ce qu'elles ne peuvent dire à personne. Son rôle est d'essayer de rentrer en contact avec elles et de les mettre en relation avec les associations. La cellule de modération intervient, avec la cellule psychologique, en cas de disputes. Cela dit, ces cas sont rares. L'usage des pseudonymes change beaucoup la nature des échanges : les conséquences ne sont pas les mêmes qu'avec un cercle d'amis qui apparaissent sous leur identité réelle, même si les réflexes de politesse ont tendance à disparaître sur Internet. Et lorsqu'il y a des demandes de retrait formulées par les parents, nous y répondons bien évidemment, puisque les blogs de mineurs sont placés sous l'autorité parentale.

Mme Christiane Féral-Schuhl, coprésidente. Et quelle est la ligne de conduite retenue pour les technologies liées à la *privacy by design* ou à la *privacy by default* ?

M. Pierre Bellanger. La seule information que nous détenons étant le pseudo, nous nous situons dans une logique de profil. Nous pouvons élaborer des agrégats à partir de certaines caractéristiques et de certains comportements mais nous ne pouvons pas procéder à des rapprochements entre des identités réelles et des comportements.

M. Philippe Aigrain. Vous ne pouvez pas utiliser les emails et les adresses IP ?

M. Pierre Bellanger. Nous disposons d'emails, nécessaires aux inscriptions, mais vous savez avec quelle facilité il est possible d'en créer pour la circonstance. Quant aux adresses IP, elles dépendent des terminaux et n'ont pas la précision d'un identifiant unique – UID – ou d'un numéro de mobile. Notre logique n'est pas de chercher à connaître l'identité réelle de nos utilisateurs.

M. Christian Paul, coprésident. L'un des points saillants de ce premier tour de table est la tension, apparue depuis une dizaine d'années, entre la libre circulation des données personnelles, toujours plus nombreuses, à laquelle peut s'attacher une certaine valeur collective, et la volonté non seulement de protéger la vie privée mais aussi de maîtriser ces données personnelles. Il est de plus en plus question de l'auto-détermination informationnelle, notion qui existe depuis longtemps en Allemagne et qui a été mise en avant dans l'étude du Conseil d'État.

Selon vous, où placer les digues pour garantir une solide protection des données personnelles ?

Mme Laure de La Raudière. Et est-ce possible ?

M. Benoît Tabaka. La principale question à laquelle sont confrontés les services Internet est le contrôle des données personnelles par l'utilisateur, qui peut contribuer à la régulation, à côté des outils de régulation classiques.

Je prendrai des exemples très concrets. Un internaute possédant un compte d'utilisateur Google a la possibilité d'activer l'option de conservation de l'historique dans lequel il peut ensuite procéder à la suppression de certaines recherches. Autre exemple : un utilisateur peut choisir de ne pas télécharger une application sur son téléphone mobile après avoir pris connaissance des messages l'informant qu'elle implique l'accès à son carnet d'adresses ou à ses données de géolocalisation.

Ce renforcement du contrôle de ses données par l'utilisateur suppose le développement de l'information spécifique qui est communiquée produit par produit. Il implique également de rendre effectifs ses choix, par exemple en supprimant intégralement les données de son compte Google s'il en décide ainsi. Du reste, les acteurs du numérique ont tout intérêt à respecter les contrats qui les lient à l'utilisateur, notamment en matière de contrôle effectif des données, s'ils ne veulent pas s'exposer à des conséquences médiatiques.

M. Paul-Olivier Gibert. Notre positionnement est un peu différent puisque notre société ne gère pas de données personnelles mais intervient auprès de structures qui sont amenées à en gérer. J'estime que plus les utilisateurs disposeront de moyens de contrôle sur les traces numériques qu'ils laissent, plus la situation sera satisfaisante.

Il me semble important de prendre en compte le cas des données que les individus contribuent à créer et qui sont utilisées indépendamment d'une prise en considération de leur identité réelle. Par exemple, si, pour retracer les déplacements d'une population, les historiques des connexions à un téléphone mobile sont utilisés, il faut veiller à éviter toute possibilité d'identification individuelle. Nous intervenons pour qu'un outillage adéquat assure une forme de non-réversibilité et d'anonymisation. Nous préconisons une intégration dans le code des applications, au-delà des principes de la *privacy by design* ou de la *privacy by default*.

M. Pierre Bellanger. Le contrôle des données par les utilisateurs m'apparaît comme un subterfuge destiné à masquer un siphonnage global de nos données. Dans le domaine du numérique, l'absence de normes publiques tend à laisser reposer sur l'acceptation individuelle certaines décisions extrêmement graves. Tout se passe comme si un automobiliste pouvait choisir un modèle de voiture avec ou sans ceinture de sécurité.

L'individu dispose en réalité d'un faux pouvoir sur ses données dont l'auto-détermination informationnelle, vers laquelle le Conseil d'État s'est orienté, n'est que le prolongement. Dans le domaine numérique, il faudrait à l'utilisateur un mois et demi pour lire toutes les clauses d'acceptation des contrats auxquels il est lié, contrats, qui, soit dit en passant, dépendent pour leur grande majorité du tribunal de Sacramento. Il se doit d'être un expert du droit de la consommation et du droit d'Internet doublé d'un ingénieur informaticien. Un médicament, un produit alimentaire, une margelle de piscine sont soumis à une somme impressionnante de normes quand, dans le domaine numérique, l'accès à un service ou un produit n'est soumis qu'à une clause d'acceptation.

En outre, j'aimerais savoir de quels moyens de contrôle dispose un internaute pour savoir si sa demande de suppression de données a bel et bien été effectuée car, en réalité, tout dépend de manœuvres menées depuis un territoire qui ne dépend pas de notre souveraineté.

Il importe de sortir de cette mascarade par le haut en imposant une protection de l'intérêt général par le droit. Les citoyens français ne sont pas des super-héros ninjas capables de se protéger contre des sociétés disposant de bataillons d'influenceurs et de lobbyistes et imposant des contrats qui ne dépendent pas de nos tribunaux ; ce sont simplement des gens pressés. En n'assurant pas la défense des utilisateurs face au siphonnage de leurs données personnelles, les pouvoirs publics se rendent coupables de non-assistance à personne en danger.

M. Christian Paul, coprésident. Au-delà de l'indignation, quelles règles et quels outils mettre en place pour assurer cette protection collective ?

Mme Laure de La Raudière. Sans doute y a-t-il un corpus juridique supplémentaire à élaborer pour protéger la vie privée mais il faudrait se demander aussi comment rendre applicable le droit français existant, sachant que les contrats auxquels les utilisateurs français sont liés dépendent du tribunal de Sacramento.

M. Pierre Bellanger. En effet, la question se pose. Les données provenant de citoyens européens ne sont plus protégées par le droit européen lorsqu'elles ont pour support des serveurs situés outre-Atlantique et les citoyens européens ne peuvent être protégés par le droit américain. Autrement dit, ces données échappent à tout contrôle et à tout droit.

La première chose à établir est la localisation des serveurs, de fait et de droit. Il faudrait poser le principe selon lequel les données relèvent du tribunal du lieu où elles sont collectées. De la même manière, en matière fiscale, l'imposition devrait dépendre du lieu de collecte.

Aujourd'hui nos données circulent sur le réseau comme si toute notre correspondance s'étalait à la vue de tous sous forme de cartes postales et était dupliquée par la poste. Le cryptage constituerait une première solution qui permettrait de mettre nos échanges sous enveloppe. À mon sens, la création d'une agence de données est impérative : elle établirait les critères, les normes, les protocoles et les niveaux du cryptage. Prenons l'exemple des données captées par un musée : il y aurait un premier niveau, d'accès public, concernant le nombre de visiteurs ; un deuxième niveau, d'accès plus restreint, concernant le nombre de visiteurs classés selon des profils ; un troisième niveau, d'accès limité aux services du ministère de la justice, concernant l'identité des visiteurs.

Ces données chiffrées seraient ensuite associées à des métadonnées, s'apparentant à des *bitcoins*, qui indiqueraient comment les données sont utilisées.

Ces propositions, que j'ai eu l'occasion de formuler à plusieurs reprises, se heurtent aux résistances des acteurs d'Internet qui les considèrent comme une atteinte à l'innovation et à la fluidité des données. Traduisons : par innovation, il faut entendre la capacité à utiliser vos données sans que vous puissiez en avoir le contrôle et par fluidité, le pillage.

Nous voyons l'effet réseau à l'œuvre. La loi de Metcalfe, selon laquelle la valeur d'une machine est proportionnelle au carré du nombre de machines auxquelles elle est connectée, s'applique aussi aux données. Autrement dit, la valeur d'une donnée est proportionnelle aux nombres de données auxquelles elle s'agrège. Le plus gros détenteur de données est le plus susceptible d'acquérir d'autres données jusqu'à constituer un monopole. Et c'est à la constitution de tels monopoles que nous assistons aujourd'hui.

Face à ces évolutions, une autorité publique doit venir garantir le cryptage des données, la localisation des serveurs et les droits individuels. Ces droits individuels recouvrent le droit au mystère que peut revendiquer chaque individu en tant qu'être en devenir défini par sa liberté de choix et par sa liberté d'évoluer. Chacun a le droit de préserver son intimité du regard d'autrui et de choisir l'image qu'il veut donner de lui-même. Chacun a le droit aussi à la neutralité du monde.

Les informations collectées sur les individus servent à orienter leurs choix et donc à restreindre leurs libertés. Si une personne accède à un site commercial à partir d'un Mac récent, les prix qui lui seront proposés seront plus élevés que s'il y accède avec un vieux PC pourri, tout simplement parce que votre équipement sera considéré comme un indicateur de

votre niveau de vie. Il en va de même pour les moteurs de recherche : en fonction de vos requêtes précédentes, ils orienteront le choix des résultats. Il a été constaté que les résultats des recherches Internet que les enfants faisaient pour élaborer leurs exposés sur l'Affaire Dreyfus variaient en fonction des requêtes passées de leurs parents et donc de leurs orientations politiques.

Le droit à la neutralité du monde implique d'avoir accès à une version standard, non paramétrée en fonction de vos préférences.

De nombreux droits restent à construire, mais, chose positive, tout est possible en la matière.

M. Benoît Tabaka. En effet, nous avons perçu cette volonté d'avoir accès à une information neutre. C'est pourquoi, il y a quatre ou cinq ans, nous avons introduit dans le moteur de recherche un bouton permettant de passer de la version personnalisée des résultats – qui tient compte des centres d'intérêt et des requêtes précédentes de l'utilisateur – à la version universelle, identique pour n'importe quel internaute.

Les révélations des affaires PRISM et Snowden sur l'interception massive d'informations circulant sur les réseaux par les services de renseignement nous ont convaincus qu'il fallait chiffrer les échanges entre nos infrastructures – par exemple entre nos centres de traitement des données aux États-Unis, en Belgique, aux Pays-Bas, en Finlande ou en Irlande – et entre nos utilisateurs et les services Google. Mais nous ne pouvons contrôler les courriels qu'un utilisateur de Gmail envoie à un client d'Orange : à partir du moment où le message sort de nos infrastructures, nous ne sommes plus maîtres de son chiffrement et – sauf si l'opérateur en face a adopté les mêmes standards – l'information circule à découvert. Nous avions d'ailleurs volontairement publié les données relatives au chiffrement de la correspondance électronique échangée sur la toile : chez certains opérateurs et fournisseurs d'accès Internet, le taux de chiffrement avoisine les 100 %, alors que chez d'autres, il n'atteint que 2 à 3 %. Comme nous l'avons annoncé hier sur nos blogs, nous sommes en train de développer des modules en accès libre qui permettront à l'utilisateur de chiffrer lui-même ses courriels, qui seront ensuite à nouveau chiffrés par Gmail. De cette façon, seul l'utilisateur final détiendra la clé du chiffrement – point qui nous paraît capital.

S'agissant des problèmes liés à l'application du droit, les tribunaux français ont aujourd'hui à connaître d'affaires concernant Google sur énormément de sujets et la justice française nous a déjà infligé des condamnations. Dans la récente affaire Costeja, la Cour de justice de l'UE a appliqué à Google Spain la directive européenne de 1995 relative aux données personnelles, reconnaissant aux résidents européens le droit à l'effacement des moteurs de recherche. Cette décision consacre la création du droit à l'oubli.

Deux textes – et prochainement un troisième – couvrent le problème des données situées en dehors du territoire français : la directive de 1995 sur les données personnelles, bientôt remplacée par le règlement européen en cours de discussion, et les mécanismes élaborés par la Commission européenne – *Safe Harbor* ou modèle de clauses contractuelles types –, qui garantissent un niveau de protection équivalent en Europe et dans les autres territoires, et obligent, à défaut, de demander une autorisation auprès des autorités de protection pour exporter les données. Ce dispositif – impliquant les banques et les centres d'appel – lie l'Europe aux États-Unis, mais également à de nombreux autres pays, en particulier du Maghreb et d'Asie. Les mécanismes juridiques existent donc ; la décision

Costeja sur le droit à l'oubli consacre notamment la reconnaissance de l'application du droit européen à tous les acteurs de l'Internet.

M. Paul-Olivier Gibert. Le chiffrement – qui représente une mesure technique destinée à assurer la protection des données sensibles à transmettre – constituera l'un des enjeux de la décennie. Au-delà des aspects juridiques, les droits de la personne sur ses données doivent être inscrits dans la technologie ; en effet, on ne saurait imaginer un code de la route si les véhicules ne sont pas équipés de freins ! Les enjeux apparaissent considérables et renvoient à une variante du *privacy by design*, c'est-à-dire à l'intégration, dans les outils, de dispositifs permettant aux personnes de garder une maîtrise sur leurs données – maîtrise qui doit être adossée à un droit. Certes, trop de droit tue le droit, générant des situations incompréhensibles et des textes contradictoires, voire contreproductifs ; mais cette dimension technico-juridique doit apparaître dans le règlement européen en tant que droit à la portabilité – la possibilité pour l'utilisateur d'amener ses données s'il change de prestataire.

M. le président Christian Paul. Comment faire en sorte que ce soit le droit français ou européen qui s'applique aux données personnelles de citoyens français, même situées sur un serveur hors de France ?

M. Pierre Bellanger. Lorsque des affaires contre des acteurs américains ont été portées devant la justice, le réflexe a le plus souvent consisté à dire que puisque les serveurs sont situés ailleurs, le tribunal est incompétent et le droit français, inapplicable. La jurisprudence sur cette ligne de défense apparaît contradictoire – parfois, cette extraterritorialité est reconnue, parfois non – mais de fait, quelle que soit la localisation du serveur où les données sont hébergées, c'est le lieu de leur collecte qui doit être celui du droit. Je pense pour ma part que certaines informations doivent impérativement être conservées sur des serveurs situés sur notre territoire ; en effet, le dispositif de *Safe Harbor* relève de la cosmétique juridique – fine pellicule de droit qui fait croire aux utilisateurs à un mécanisme légal, alors qu'il ne s'agit que d'une déclaration d'intention échappant à tout contrôle.

S'agissant de la souveraineté numérique – la maîtrise de notre destin sur les réseaux, extension de notre souveraineté nationale – et de l'affaire Snowden, lorsque l'État formule une requête dans le cadre d'une affaire de terrorisme, n'importe quelle entreprise américaine ou française lui ouvre l'accès à ses informations, quels qu'en soient les niveaux de chiffrement. En démocratie, cette soumission à l'intérêt supérieur de l'État apparaît parfaitement légitime. Les acteurs américains offrent des services exceptionnels et font un travail de qualité ; c'est notre droit – qu'ils respectent pour l'essentiel – qui apparaît mal conçu. Placés sous la souveraineté de leur pays, ils obéissent à leur puissance publique ; nous devons disposer de nos propres données, soumises à notre propre souveraineté. On ne peut pas faire autrement : sans le secret des données, il n'existe ni propriété intellectuelle, ni brevet, ni valeur économique, ni diplomatie, ni armée. Organiser un formidable défilé du 14 juillet tout en étant incapable de garantir le secret de la correspondance militaire constitue une anomalie ! Le chiffrement fournit une réponse au problème, mais non une solution en soi ; il ne doit masquer ni les responsabilités ni le droit auquel il devrait s'adosser.

Mme Christiane Féral-Schuhl. Au fil des procès, la jurisprudence s'est chargée de définir les critères de rattachement des sites Internet – la nationalité de l'éditeur, le public visé, la langue utilisée – qui nous permettent aujourd'hui de poursuivre certains d'entre eux en France alors même que l'hébergeur est délocalisé. L'exécution des décisions reste pourtant problématique lorsque celui-ci ne dépend pas des autorités françaises. Comment les acteurs privés – notamment les grandes entreprises internationales comme Google – pourraient-ils

intervenir plus activement dans ce domaine ? La presse a relayé, il y a quelque temps, l'action de Google pour identifier un criminel américain ; s'agissait-il de sa propre initiative ? Quel rôle pouvez-vous jouer dans l'exécution des décisions, au-delà de la procédure des notifications de la loi pour la confiance dans l'économie numérique (LCEN) ? Peut-on songer – comme vous semblez le faire – à la corégulation ?

M. Benoît Tabaka. Les autorités de divers pays adressent régulièrement à Google des demandes de réquisition afin d'identifier des personnes – auteurs d'une chaîne YouTube ou titulaires d'un compte Gmail – soupçonnées d'avoir commis des délits ou des crimes. Dès lors que cette demande obéit au cadre juridique légal, Google s'y conforme ; mais nous ne faisons pas droit aux demandes qui ne respectent pas ce cadre. Il nous est même arrivé de contester en justice la validité de la demande des autorités lorsque le volume des informations sollicitées nous paraissait excessif par rapport au périmètre de la loi.

Au-delà des notifications LCEN visant des contenus hébergés par Google, nous recevons également des demandes de déréférencement d'un contenu, au titre du droit à l'oubli – entré en vigueur depuis le mois de mai dernier – ou parce que des juges en ont prononcé l'illégalité. Sans procéder à une exécution au sens judiciaire du terme – la décision nous est notifiée par simple lettre recommandée –, nous supprimons alors les liens qui mènent vers le contenu déclaré illicite.

Enfin, les contenus pédopornographiques font l'objet de détections systématiques. En effet, il s'agit d'une des rares infractions universelles dont la sanction ne fait débat nulle part au monde – contrairement par exemple à la contrefaçon ou à la diffamation. Lorsque le système de l'algorithme détecte des images pédopornographiques, on transmet ces informations au *National Center for Missing and Exploited Children* (NCMEC) aux États-Unis, qui les envoie ensuite aux autorités des différents pays – à l'Office central de lutte contre la criminalité liée aux technologies de l'information et de la communication (OCLCTIC) à Nanterre pour la France, au FBI pour les États-Unis, etc. Un amendement à la récente loi française sur la lutte contre le terrorisme a étendu le périmètre des retraits de contenus aux sites faisant l'apologie du terrorisme ; lorsque nous recevrons une liste de la part des autorités, nous procéderons à une exécution spontanée en supprimant ces contenus de notre moteur de recherche.

M. Philippe Aigrain. Tous les intervenants auditionnés par cette commission s'accordent à reconnaître que les paradigmes des années 1970 sont aujourd'hui dépassés par la multiplication des données. Le débat a porté sur le degré d'adaptation du droit européen à cette situation, les promoteurs de ce droit et les autorités réglementaires comme la CNIL tendant à défendre leur pré carré, et les observateurs externes se montrant plus réservés. En revanche, malgré le consensus quasi absolu dans les médias et au sein de la classe politique, je continuerai inlassablement à affirmer que les données ne constituent pas le modèle dominant de l'économie numérique. En effet, il ne faudrait pas répéter l'erreur commise à propos de la propriété intellectuelle lorsque l'omniprésence du droit d'auteur avait amené à y réduire toutes les activités économiques qui en faisaient usage pour proclamer que le contenu était roi ou que l'économie européenne de la propriété intellectuelle pesait 49 milliards d'euros. La macro-économie a appris, dans les années 1930, à traiter ce genre de questions : ce n'est pas parce que l'écrasante majorité des activités économiques utilisent l'électricité que toute l'économie s'y ramène ! Certes, certains acteurs tirent des revenus – essentiellement publicitaires – du traitement des données, mais leur taille et leur puissance impressionnante ne doivent pas nous faire croire qu'il s'agit du modèle dominant. Si nous partons de ce postulat, nous ferons – comme par le passé – de mauvaises politiques publiques.

En revanche, la collecte de données est en effet omniprésente, et comme l'a noté Pierre Bellanger, elle apparaît associée au formatage des comportements – que Bernard Stiegler a qualifié de « grammatisation » ; il en va ainsi de l'autocomplétion des requêtes dans les moteurs de recherche ou de la proposition de thèmes dans un blog. Ces phénomènes qui reposent sur des algorithmes de traitement de données dépossèdent les individus de leur autonomie et leur interdisent l'accès à l'autodétermination informationnelle – capacité, en construisant leur identité sur le net, de choisir des outils qui les aident à faire des choses plutôt que des outils qui aident à leur vendre des choses. Plus globalement, ce formatage des comportements prive nos sociétés de la capacité à décider de leur devenir commun. Je remercie Pierre Bellanger d'avoir soulevé ce problème – dont nous avions déjà discuté il y a des dizaines d'années, lorsqu'il s'agissait de savoir s'il fallait permettre des radios qui ne soient pas l'objet de monopole. En revanche, je suis moins enthousiaste quant aux solutions devant nous permettre de nous réapproprier ce destin collectif. Depuis quarante ans, on fait du consentement la racine des législations sur la protection des données ; mais pour qu'il y ait consentement, il faut que l'individu dispose d'une solution de rechange ! Comme disait Ted Turner, nous avons aujourd'hui 400 chaînes de télévision, mais nous n'avons jamais eu aussi peu de choix sur les programmes à regarder. Disposer de douze moteurs de recherche et pouvoir instantanément passer de l'un à l'autre grâce à une icône de la barre de recherche ne sert à rien s'ils fonctionnent tous sur le même modèle et utilisent les mêmes algorithmes. Aussi faut-il sans arrêt évaluer les concepts tels que le consentement, la portabilité des données ou la neutralité du réseau. Les consommateurs ne peuvent choisir librement que s'ils rencontrent une offre réellement neutre et accessible partout ; si les opérateurs de télécommunications se mettent d'accord pour la biaiser, ils n'auront pas le choix.

Je suis favorable non seulement à une vision territoriale en matière de droit applicable, mais également à la relocalisation physique des données ; en revanche, je ne suis pas sûr qu'indexer cette reterritorialisation sur la souveraineté nationale nous amènera les bénéfices que Pierre Bellanger en attend. En effet, le président la commission des lois affirmait ce matin à la radio que la protection de nos ressources économiques nécessitait de mettre en place des dispositifs de capture et de surveillance des données ; malheureusement, les acteurs nationaux ou européens n'inspirent pas plus confiance que ceux des autres pays. La confiance procède du contrôle effectif, aussi exige-t-elle que le droit ne soit pas trop éloigné ; mais c'est la fiscalité qui, comme toujours, décidera de tout. En effet, l'évasion fiscale constitue la motivation première pour contourner la territorialisation ; ces deux problèmes apparaissent étroitement liés et la relocalisation de la protection des données ne se fera que le jour où l'on aura relocalisé la fiscalité. Il est également important que les individus puissent contrôler physiquement leurs données, ce qui suppose de décentraliser les services et les applications. Ainsi, j'attends avec impatience que les *skyblogs* fonctionnent comme WordPress : chaque jour, l'usager doit pouvoir héberger lui-même son blog là où il le souhaite, avec une plateforme et un logiciel libre. C'est cela, la véritable portabilité des données !

M. Pierre Bellanger. Nous avons mis en place ce dispositif.

M. Philippe Aigrain. Je m'en réjouis.

Enfin, pour que le chiffrement apporte une garantie, il faut qu'il accompagne le message de bout en bout, que les clés aient été générées par l'usager lui-même et que les protocoles n'aient pas été corrompus. En effet, la NSA a dépensé 250 millions de dollars par an pendant dix ans pour corrompre certains algorithmes de chiffrement. Des membres de la Quadrature du net travaillent actuellement sur des outils de courriel qui respecteront ces

critères ; un projet similaire est également en cours en Islande. Il existe dans ce domaine un besoin énorme de politiques publiques et de soutien car le courriel ne représente qu'un aspect du problème ; ainsi, il faut également trouver des solutions pour la messagerie instantanée, la communication vidéo ou l'hébergement des données de médias.

M. Daniel Le Métayer. Monsieur Gibert, vous avez bâti une société sur l'éthique et la protection des données. Quelle est aujourd'hui, pour une entreprise, la valeur ajoutée de la vie privée – sans doute variable selon le secteur d'activité et le degré de contact de l'entreprise avec le public ? Si cette valeur ajoutée n'existe pas encore, la sentez-vous venir ? Sinon, comment la susciter ?

Monsieur Tabaka, vous dites vouloir accorder aux utilisateurs le contrôle sur leurs données. Or il s'agit d'un des sujets sur lesquels le groupe de travail Article 29 vous a fait des reproches ; le pack de conformité qu'il a récemment publié vous interpelle sur la transparence, le contrôle et la durée de conservation des données. Comment y répondez-vous ? Mettez-vous en œuvre certaines de ses recommandations ? La loyauté de votre algorithme de classement des pages fait également débat. Une façon de prouver votre bonne foi serait d'accepter qu'une autorité indépendante – par exemple la CNIL – audite régulièrement cet algorithme en constante évolution pour s'assurer qu'il traite les différents acteurs de manière égale. Êtes-vous prêts à entrer dans ce jeu de la responsabilité ? Sinon pourquoi ?

Monsieur Bellanger, le chiffrement constitue une mesure technique nécessaire qu'il faut davantage employer, mais non la solution ultime à tous les problèmes. Vous semblez avoir entremêlé deux aspects : le chiffrement et l'anonymisation. En effet, on rencontre de moins en moins une alternative – pouvoir ou non accéder à une donnée –, et de plus en plus des situations où l'on peut accéder à des données agrégées, simplifiées et regroupées, comme vous l'avez suggéré en filigrane dans votre exemple de la fréquentation du musée. Mais plus que de trois niveaux précis que vous avez distingués – données anonymes, données couvertes pas des pseudonymes ou des profils et données identifiant pleinement une personne –, il s'agit d'un continuum de données plus ou moins susceptibles d'identifier l'individu. Les autorités de protection ont travaillé sur cette question : le groupe Article 29 a récemment publié un guide qui définit les règles du jeu, précisant ce qu'est un bon algorithme d'anonymisation et dans quel cas on peut considérer qu'un jeu de données est réellement anonyme. Cette mesure me semble aller dans le sens de vos préconisations.

M. Paul-Olivier Gibert. La valeur ajoutée de la vie privée se lit à plusieurs niveaux. Tout d'abord, les atteintes à la vie privée des clients peuvent être destructrices de valeur pour les entreprises, à la fois par le biais des sanctions – pour l'instant mineures – et par la perte de confiance qu'elles suscitent. Ces destructions se font par seaux alors que la valeur se construit à la cuillère !

Les effets dépendent évidemment de la sphère d'activité de l'entreprise et de l'intensité de sa relation avec le client, mais aussi de la population à laquelle on s'adresse. Les rares études sur les attentes du public en matière d'utilisation des données personnelles font apparaître quatre groupes distincts. Le premier soit ne s'inquiète pas de laisser des traces numériques, soit le considère nécessaire car il y trouve des contreparties ; le deuxième tient une position plus nuancée, estimant qu'il ne faut pas aller trop loin ; à l'autre bout du spectre, un groupe extrême voit tout élément de traçabilité et de visibilité comme une atteinte intolérable à la liberté ; enfin, un groupe intermédiaire, attentif aux révélations de l'affaire Snowden, considère que pour vivre heureux, il faut vivre caché, mais reconnaît toutefois l'utilité des outils de géolocalisation pour les personnes dépourvues de sens de l'orientation.

Le même service ne jouira pas du même degré de confiance selon qu'il s'adresse à une population de tout-venant ou à sa partie *geek*, favorable à la diffusion d'informations et à une forme de transparence qui renvoie à une conception de la vie privée aux frontières bien plus reculées qu'il y a soixante-dix ou même vingt ans.

M. Benoît Tabaka. Au mois de février 2012, nous avons engagé un processus avec le groupe Article 29 – qui réunit l'ensemble des autorités européennes de protection de données – en annonçant un grand changement de nos règles de confidentialité, l'idée étant de fusionner les règles des différents services en une règle unique, plus claire et plus lisible pour les internautes. Au terme d'une première phase de travail, six autorités du groupe – la CNIL française, mais aussi les autorités italienne, allemande, britannique, hollandaise et espagnole – ont décidé de mener des enquêtes spécifiques sur différents sujets. À la fin du mois de septembre, nous avons reçu les recommandations qui font suite à toutes ces discussions : clarifier mieux encore nos règles de confidentialité, préciser la durée de détention des données et donner davantage de contrôle aux utilisateurs. Nous y avons répondu par courrier et un rendez-vous est prévu au mois de janvier 2015 ; les choses avancent donc.

Quant à la loyauté de l'algorithme, la *Federal Trade Commission* (FTC) américaine et la Commission européenne ont eu accès à beaucoup d'informations. Dans le cadre de la procédure actuellement en cours, la Commission a notamment souhaité vérifier la manière dont notre algorithme agissait sur les contenus.

M. Daniel Le Métayer. L'algorithme évolue régulièrement et ses paramètres changent. Il faudrait un processus continu de contrôle !

M. Pierre Bellanger. Faire du consentement la racine de toutes les procédures de droit revient à affirmer un droit de se faire du mal. Une telle approche – légaliser une action contraire à mes intérêts sous prétexte que j'y consens – serait acceptable si nos données étaient individuelles ; mais comme les données des différents utilisateurs sont toujours intriquées, chaque fois que je révèle quelque chose sur moi, je révèle quelque chose sur autrui. Conférer aux données le statut de bien commun ferait s'effondrer la logique de consentement individuel puisqu'on engage ici non seulement son propre destin, mais également celui des autres.

Christiane Féral-Schuhl a rappelé les décisions mettant en avant la préséance des tribunaux français ; mais cela ne concerne que des actions visibles en France. Dès lors que nos données sont stockées outre-Atlantique, leur destin – la manière dont elles seront cédées, utilisées et agrégées – nous échappe, et jusqu'à la connaissance même des délits dont ils feront éventuellement l'objet.

Enfin, le chiffrement est une application du droit, mais ne peut se substituer à celui-ci. Il faut créer un droit où les données seraient considérées comme un bien commun et où une agence des données – peut-être adossée à la CNIL – garantirait les protocoles de cryptage et de codage. Ce projet reste imparfait et demande encore à être amélioré ; mais je crois qu'il constitue un bon début. En prenant l'exemple du musée, j'identifiais trois niveaux : celui de l'identité, du profil et des données. Certes, il est possible de reconstituer les identités à partir des données ou des profils, mais l'on peut en faire un délit – à l'instar de la tentative de percer un coffre. C'est un nouveau droit qui émerge, à la fois collectif et individuel, à partir de cette vision des données comme un bien commun, sujet à la souveraineté nationale et européenne, et certainement sujet de droit.

M. le président Christian Paul. Messieurs, je vous remercie pour vos interventions et pour vos réponses.

La séance est levée à dix heures trente