

Quels droits, et quel pouvoir pour les individus ?



Entretien avec Philippe AIGRAIN

Informaticien et essayiste, co-fondateur de « La Quadrature du Net »

Dans le droit actuel, les droits des individus face aux traitements informatiques reposent sur deux principes de base : la finalité et le consentement. Chaque traitement a une finalité claire ; chaque individu, informé de cette finalité, donne son consentement pour participer au traitement. Ces principes établis dans les années 1970 pour des recueils de données structurées sont mis en difficulté aujourd'hui où le champ des données s'est considérablement étendu, puisqu'il couvre les traces laissées par les internautes ou les opinions qu'ils expriment. Et l'anonymisation n'est pas la panacée. Suffira-t-il pour autant de définir de nouveaux droits ? On a également besoin d'une attention plus forte portée aux architectures informatiques et aux modèles commerciaux, qui conditionnent la réalité de l'exercice d'un droit comme celui du consentement.

Statistique & Société : Quel est le projet de « La Quadrature du Net » ?

Philippe Aigrain : La Quadrature du Net a été créée en 2008 par cinq individus pour défendre les droits fondamentaux et les libertés dans l'espace numérique. Je le souligne : pour nous, il ne s'agit pas de droits ou de libertés spécifiques à l'espace numérique, il s'agit de la mise en œuvre dans cet espace des droits généraux, tels qu'ils sont énoncés notamment dans la Déclaration universelle des droits de l'Homme. Nous sommes maintenant une association loi de 1901, qui intervient soit en réaction à des projets de textes juridiques, soit spontanément, pour proposer des chartes de droits adaptées aux enjeux du numérique, des politiques publiques porteuses de développement humain dans ce domaine, et des conditions de développement des technologies favorisant leur usage social. Nous sommes financés à plus de 50% par les dons individuels, le reste provenant du soutien de fondations.

S&S : Dans quels domaines intervenez-vous ?

PA : Dans deux grands domaines. D'une part, celui des droits fondamentaux : et c'est ce qui nous concerne essentiellement dans cet entretien. D'autre part, la Quadrature du Net intervient sur tout ce qui concerne les enjeux culturels et sociaux dans le monde numérique.

S&S : Alors vous êtes un groupe de « geeks » ?

PA : Pas du tout. Certes nous avons des compétences en informatique, et nous défendons depuis l'origine le modèle des logiciels libres ; mais nous intervenons surtout sur le terrain juridique, national ou international, à l'aide de membres qui ont des compétences reconnues

en droit et en sciences politiques. Nous sommes consultés par les pouvoirs publics : par exemple, je fais partie des experts membres de la commission de l'Assemblée Nationale qui se penche là-dessus¹. Quant il y a une saisine du Conseil Constitutionnel sur un sujet de notre compétence, nous nous efforçons d'éclairer son jugement.

S&S : Avez-vous des équivalents à l'étranger ?

PA : Dans le monde anglo-saxon, nos équivalents sont des ONG professionnalisées, comme « Open Rights » ou encore la « Foundation for information policy research » (FIPR) au Royaume-Uni. En Allemagne, je peux citer le « Chaos Computer Club » qui est la plus ancienne association dans le domaine, ou « Digitale Gesellschaft ». L'influence anglo-saxonne est sensible au niveau international : ainsi, au niveau européen, l'« EDRI » (« European digital rights »), établie à Bruxelles, est très liée au monde anglo-saxon. Au niveau mondial, l'association « Article 19 »² créée en 1987 se consacre à la défense de la liberté d'expression.

S&S : Venons-en donc aux droits individuels dans l'espace numérique. Par la loi « Informatique et liberté » de 1978³, il est prescrit que « les données sont collectées pour des finalités déterminées, explicites et légitimes » (article 6) et que « un traitement de données à caractère personnel doit avoir reçu le consentement de la personne concernée » (article 7, sous réserve des exceptions énumérées dans cet article). Est-ce que ce cadre légal est toujours pertinent ?

PA : Si - ce qui n'est pas mon cas - on considère les données en tant que « richesse », circulant soit librement selon les modèles d'OpenData, soit par vente comme c'est le cas aux Etats-Unis des données de santé, la finalité ne peut pas être définie au moment de la collecte, car ce sont les futurs exploitants qui vont décider des exploitations possibles. Du coup, la doctrine qui est à l'origine de la CNIL, selon laquelle tout va bien à condition que chaque individu donne son consentement pour une finalité connue, ne fonctionne plus. Pour ma part, je remets en cause cette vision des données : pour moi, l'acceptation des modèles commerciaux fondés essentiellement sur la publicité demande à être critiquée, nous y reviendrons⁴. Mais dans ce contexte, il est clair que le système « finalité-consentement » est affaibli. Par contre, sa pertinence n'a pas disparu dans d'autres contextes : ainsi d'une étude sur le rapport alimentation santé dans laquelle les enquêtés, correctement informés des finalités de l'étude, ont apporté volontairement des informations hautement confidentielles les concernant.

S&S : Alors, est-ce qu'il faut défendre pied à pied cette référence à la finalité ? Ou faut-il imaginer quelque chose de nouveau ?

PA : Il faut protéger les contextes dans lequel le concept de finalité a encore un sens. Mais les finalités sont rarement aussi simples. Souvent, la collecte d'information a plusieurs finalités : et on constate souvent que « les mauvaises chassent les bonnes » : des finalités secondes chassent celles qui avaient été affichées en premier. Ainsi, lorsque dans un service public on s'intéresse aux données pour améliorer la qualité des services aux usagers et pour réduire les coûts, on finit par viser seulement ce dernier objectif. Autre exemple : les données collectées pour accroître la satisfaction du consommateur le sont aussi pour générer des recettes publicitaires, et cet objectif devient prédominant.

Il faut aussi se rendre compte que la notion de « données » s'est élargie à des contextes nouveaux, tout différents de ceux qu'on connaissait en 1978. On appelle maintenant « données »

1. Commission de réflexion sur les droits et libertés à l'âge numérique

2. Ainsi nommée d'après l'article 19 de la déclaration universelle des droits humains.

3. Cette loi a été modifiée à plusieurs reprises depuis, mais reste en vigueur à ce jour ; ses principes sont inchangés.

4. La publicité représente 95% des profits commerciaux de l'exploitation des données ; le reste va aux réductions de coûts par la maîtrise des risques en assurance, par exemple.

des choses très diverses : les « traces » laissées par les individus dans leur parcours d'usage de certains services ; les « informations » capturées par hasard ; ou encore les « communications » entre personnes, et les « expressions » d'opinions laissées sur des forums qui relèvent de tout autres mécanismes de droits... Pour ces catégories de données, le principe de finalité qui a été conçu pour des formulaires précis s'applique malaisément. Il risque même de conduire à protéger des choses qui ne devraient pas l'être, comme des expressions publiques, et à en protéger insuffisamment d'autres.

S&S : L'anonymisation est-elle une solution ?

PA : Le « miracle de l'anonymisation » viendrait nous sauver, et nous autoriser à mettre en circulation soit libre soit commerciale ces données, pourvu qu'elles aient été préalablement « anonymisées » ? Il y a de nombreux spécialistes qui travaillent là-dessus, qui cherchent à identifier les failles de l'anonymat : mais il est difficile de distinguer l'accessoire de l'essentiel. Le problème, c'est qu'une anonymisation parfaite enlève en général tout intérêt aux données ! Le chiffrement est en partie une alternative, mais peut poser lui-même ses problèmes de sécurité. Je ne crois pas que la solution soit : un droit assez contraignant pour les données personnelles, et une libre circulation des données qu'on dit anonymisées.

S&S : Alors, faut-il créer de nouveaux droits ?

PA : Attention : la vraie loi n'est pas toujours où l'on croit. On a tendance à déclarer des droits très forts, et à remettre à des décrets d'application le soin de préciser comment ces droits vont s'appliquer. A ce moment là, c'est le décret d'application, voire le cahier des charges de l'application informatique, qui définit la réalité du droit. Et c'est pourquoi les architectures informatiques et les protocoles d'accès sont cruciaux : le pouvoir des individus passe par là.

S&S : Pouvez-vous donner un exemple ?

PA : Oui, l'exemple du dossier médical personnel (DMP), sur lequel j'ai personnellement travaillé. Je citerai deux aspects. La loi prévoit que le malade aura le pouvoir de contrôler qui aura accès à ses données, de façon précise. Mais, comme c'est impraticable dans le contexte hospitalier, une clause supplémentaire a été introduite : le consentement donné à un membre d'une équipe de soins sera réputé avoir été donné à l'ensemble de cette équipe, y compris les personnels administratifs semble-t-il. Autre exemple : la loi de 2005 formule un principe essentiel : chacun dispose d'un droit d'accès à sa propre information de santé. Mais pour la mise en œuvre de ce droit, à l'ère numérique, en ce qui concerne le DMP, il est pour le moment prévu qu'il faudra envoyer une lettre aux autorités compétentes, pour finalement recevoir une liasse de papier !

S&S : Est-ce que certaines architectures informatiques sont plus propices que d'autres à l'exercice des droits des individus ?

PA : Certainement. Dans le cas du DMP, la décentralisation du stockage des données vers les individus eux-mêmes, avec bien entendu les procédures de sauvegarde nécessaires, donnerait un contrôle bien plus réel. On peut aussi envisager une « décentralisation partielle », selon un modèle intermédiaire qui a déjà été mis en œuvre à grande échelle dans certains pays. Mais il faut réfléchir de façon conjointe aux architectures et aux droits, tant ceux des individus que ceux de la collectivité. Autre exemple : prenez l'hébergement des médias – vos photos, vos films, vos enregistrements musicaux. Il n'y a en réalité aucun problème pour héberger cela soi-même, dans des espaces sous contrôle de l'utilisateur : des dispositifs simples peuvent être développés pour cela. Il existe une société française, OVH, qui est un leader de l'hébergement de sites, mais on préfère citer comme succès des entreprises qui développent des technologies pour la publicité ciblée. L'hébergement de contenus de chacun s'est déplacé vers des sites centralisés

du type Youtube, ou aujourd'hui du cloud d'Apple, dans lesquelles l'utilisateur ne maîtrise plus le devenir de ses objets, en échange de possibilités largement imaginaires dans la plupart des cas, comme de partager ses contenus avec des millions d'autres personnes.

S&S : Alors, qu'est-ce qui pourrait être amélioré pour garantir aux individus un pouvoir effectif sur leurs données ?

PA : La définition juridique des droits, et des sanctions, reste indispensable : si on ne l'a pas, une approche purement technologique et sociétale ne marchera pas, elle restera confinée dans de petits cercles de volontaires désireux d'appliquer des pratiques « vertueuses ». Mais je le répète : la question de l'effectivité des droits est pour moi soumise à celle des architectures informatiques et des modèles commerciaux. Je ne crois pas qu'on puisse progresser par une approche purement juridique. Mais on peut imaginer des architectures et des modèles économiques qui permettent un plus grand contrôle des individus. J'ai déjà parlé des architectures décentralisées : j'ajouterais que le contrôle peut être « logique » sans être nécessairement physique. Ainsi, les créateurs de pages Web chez les fournisseurs d'accès internet gardaient-ils à la fin des années 1990 un contrôle de fait sur leurs contenus, sans rien savoir de leur implémentation physique.

Les conditions du recueil du consentement à l'utilisation des données sont essentielles. Ainsi dans le cas de cookies, il faut souvent donner son consentement pour accéder à la suite du site : l'utilisateur n'a pas vraiment le choix ! De plus, le consentement doit être réversible, et les exigences de contextualisation doivent être respectées (de la même façon que la lumière doit rester allumée au cinéma pendant les publicités ou que les publicités rédactionnelles doivent se différencier clairement des articles dans la presse écrite).

S&S : Et les modèles commerciaux ? Comment jouent-ils sur les droits ?

PA : En ce qui concerne les modèles commerciaux, je considère que c'est la « tache aveugle » de l'action publique, alors que l'on constate ici clairement à quel point certains modèles en vigueur sont pernicioseux tant sur le plan social que sur le plan macro-économique. L'industrie du livre, qui n'a jamais reposé sur le modèle de la publicité, continue à représenter la plus grande part de la valeur ajoutée des industries culturelles, alors même que le temps de lecture décroît au profit du temps consacré aux médias audiovisuels. Pourquoi la publicité est-elle le revenu fondamental des services sur le web ? Parce qu'il est plus facile de prélever des sommes importantes auprès d'un petit nombre d'annonceurs que de prélever de petites sommes auprès d'un grand nombre d'utilisateurs finaux. Mais ce n'est pas sans conséquence sur l'émergence d'oligopoles. Dans tous les domaines, les modèles commerciaux sont gérés par de la réglementation, une politique de la concurrence : c'est d'ailleurs une politique de ce genre qui a permis le succès de l'Internet fixe en France. Tout le débat sur la « neutralité du Net » est en fait un débat sur les modèles commerciaux qu'on autorise. Le politique a renoncé à faire des choix, même lorsqu'il s'agit d'écarter ce qui est le plus nuisible, le plus inacceptable.

Encore un exemple, celui de la géo-localisation. Actuellement, les services offerts en cette matière reposent sur des services centralisés. Ces services ne permettent pas du tout aux gens de savoir où ils sont eux-mêmes ; mais ils permettent aux entreprises commerciales de savoir où les gens sont, leurs itinéraires, etc. de façon à leur proposer des restaurants, des moyens de transport, etc. C'est Google qui sait où vous êtes, pas vous ! On pourrait renverser la logique : faire en sorte que chacun puisse seul savoir où il se trouve, et décider de diffuser cette information à un prestataire de service s'il veut interroger celui-ci. Autre exemple encore, la RATP et les données issues du « Pass Navigo » : la CNIL a prouvé que les analyses de trafic nécessaires au fonctionnement de la RATP pouvaient se faire avec des données anonymes. Cependant, actuellement, le client doit payer 5€ de plus par mois s'il veut une carte anonyme

empêchant que ses données soient exploitables à d'autres fins. Pourquoi ? Parce que la RATP veut être un acteur du marché de la publicité.

S&S : Si l'essentiel se joue au niveau des architectures informatiques et des modèles commerciaux, que peut faire la puissance publique ?

PA : Les pouvoirs publics peuvent agir par des politiques d'innovation, en appuyant les initiatives qui donnent aux usagers un contrôle physique, ou au moins logique, sur leurs données. En ce qui concerne les modèles économiques et les usages, les politiques répugnent beaucoup à interférer ; mais l'évidence de l'évasion fiscale à grande échelle permet d'espérer que les États vont se mobiliser. Au moins, on pourrait ne pas agir dans l'autre sens : ne pas favoriser systématiquement les solutions centralisées ou seulement compatibles avec le modèle de la publicité. En matière de transmission de fichiers et de droits d'auteur, les pouvoirs publics ont fait une guerre sans merci au partage « peer to peer » en tant qu'architecture, pour favoriser le « streaming » ou le « download » centralisé qui se prêtent tout autant à des usages non autorisés et sont nuisibles à la diversité culturelle.

Il faut aussi reconnaître l'existence de « données d'intérêt public » sur lesquelles on ne peut pas raisonner en termes de propriété. A titre d'exemple, il est inimaginable que les accords de partenariat public-privé ne soient pas rendus publics ! Ces informations sont en effet des « biens communs », à rendre utilisables par tous pour autant qu'on ait assuré la sécurité des données individuelles.

S&S : Se pose aussi le problème de la réutilisation de données d'opérateurs privés pour établir des informations d'intérêt général. Par exemple, l'opérateur de téléphonie mobile Orange exploite ses données sur les communications (métadonnées, données de géo localisation) et permet à d'autres de les exploiter. Il s'agit d'un gisement statistique qui est loin d'être anecdotique. La statistique publique pourrait l'utiliser pour mieux connaître le tourisme ou les déplacements. Quel statut pour de telles données ?

PA : Je ne suis pas favorable à rémunérer les individus qui sont à l'origine de ces données, pour ne pas les engager dans des négociations totalement asymétriques avec des opérateurs infiniment plus puissants qu'eux. En revanche, leur consentement devrait être obtenu sur la base d'un système « d'opt-in » : il faut alors un acte positif de l'utilisateur pour que ses données puissent être utilisées.

S&S : Qui dit consentement dit « biais de sélection » : ce n'est pas bon pour la statistique !

PA : Le consentement a toujours un coût ! C'est une question de volonté, de choix politique. Quand on a mis en place le consentement aux essais cliniques, il y a quelques années, cela a entraîné des difficultés sérieuses, des conséquences multiples : mais on est arrivé à y faire face. L'argument de la commodité ou de l'incommodité de faire autrement sert souvent à fermer le choix politique. Les statisticiens doivent trouver des solutions pour atteindre leurs objectifs en considérant le consentement comme un impératif supérieur.