Diplôme national de master

Domaine - sciences humaines et sociales

Mention – sciences de l'information et des bibliothèques

Parcours – archives numériques

Mémoire de recherche / août 2017

Du *leak* en tant qu'archive, ou comment le *leak* est devenu une archive

Elodie Somé-Blad

Sous la direction de Clément Oury
Enseignant – ENSSIB

Remerciements

Un mémoire, un rapport, ou tout autre travail de recherche ne s'écrit évidemment pas seul. Il convient donc de rendre à César ce qui appartient à César, et de fait de remercier les personnes qui, de près ou de loin, auraient contribuées à l'élaboration de ce travail de recherche.

Ainsi, à Mr Oury qui aura suivi mon travail, ainsi qu'à l'ensemble des enseignants du master Archives Numériques de l'ENSSIB, j'adresse mes plus sincères remerciements. Il et ils auront été précieux, tant par leurs conseils avisés que par leurs expériences respectives.

À Mme Miremont, attentive et impliquée pour nos stages, rassurante et compréhensive pour nos mémoires, je tiens à exprimer toute ma gratitude. Elle nous aura permis de donner le meilleur de nous-mêmes en nous offrant un cadre de discussion stable et apaisant.

Enfin, à ma mère, à mon compagnon, et à ma colocataire, pour m'avoir supportée et soutenue dans les moments les plus importants, ainsi qu'à ma famille et mes amis qui, sans implication directe, ont toujours eu quelques mots motivants, merci.

Résumé : Le leak est un document qui fascine autant qu'il révulse. De sa fuite à sa publication, il entretient les fantasmes les plus incongrus et joue avec les secrets du désir, du besoin, de l'envie jusqu'à en faire oublier son essence. Mais le leak est un document, une donnée, une information qui est pleinement imprégné de la notion d'archive.

Descripteurs : leak(s) ; scandale ; leakers ; lanceurs d'alerte ; hackers ; archive ; Libre ; secret.

Abstract : The Leak is a magnetic and abhorring record. From its leak to its publication, the Leak maintains incongruous fantasies and plays with the secrets of desire, envy or need, and it makes us forget its essence. Only, the Leak is a record, a data, an information which is fully imbued with notion of archive.

Keywords : leak(s); scandal; leakers; whistleblower; hackers; archive; Open; secret

Droits d'auteurs

Sommaire

INTRODUCTION

Vendredi 5 mai 2017, en pleine campagne du second tour de l'élection présidentielle, alors que la violence politique bat son plein et que tous les coups semblent déjà permis, 70 663[1] données et documents sont diffusés par le site de lanceurs d'alerte aux pratiques fortement controversées, Wikileaks. Ces *MacronLeaks* – du nom du futur président français – s'ils sont en partie assurés comme faux, portent principalement sur la « gestion de la campagne au quotidien », documents que l'on juge « de par leur spécificité technique, quasiment infalsifiables ». Et s'ils se distinguent par la manière dont on les a obtenus, ils sont loin d'être les seuls à faire scandale.

Football Leaks, *Lux Leaks*, *Swiss Leaks*, ou encore *Panama Papers*, *Sony Leaks*, ou *Cablegate*, les fuites d'informations et les vols de données semblent aujourd'hui légion. Ils touchent tout autant aux domaines de la finance, de la politique, de l'environnement, du sport et du dopage, des affaires criminelles ou des violences policières, des données personnelles comme les photos ou les numéros de sécurité sociale, du cinéma ou de toute autre œuvres de l'esprit. Apolitiques, anarchistes, militants activistes, amateurs actifs ou encore trolls expérimentés, les pirates, lanceurs d'alerte, leakers et diffuseurs à l'origine de la récupération et de la diffusion semblent être des génies de la manipulation de données et de documents. Ils ne sont certes pas nés du numérique. Pourtant, ils voient leurs moyens démultipliés par le réseau et les communautés qui le composent, s'arment d'une puissance technique et technologique facilement accessible, se confortent dans une visibilité rapide, se lovent dans un cocon numérique auquel on aime prêter une idéologie nouvelle et révolutionnaire.

Au-delà de ce raccourci apprécié, les premières communautés se construisent dans des périodes de l'histoire des sociétés occidentales offrant un regard libertarien et autorégulé à la construction des futurs réseaux humains. Des données ouvertes aux logiciels libres, internet se fabrique par et dans ces tentatives d'évolution vers une société, une autre appréhension de l'économie et de la communication à l'autre. Il relance de grands débats oubliés sur quelle technologie tuera cela[2], impose de se pencher à nouveau sur la notion de document, d'interroger celle de la donnée, de réinventer l'information, et plus que jamais de comprendre notre mémoire. L'archive alors, dont les rapports sont pleinement réinventés par le web, s'offre de nouveaux processus, de nouveaux cycles, emprunte des chemins inédits. On pense son authenticité, sa fiabilité, son intégrité, sa pérennité, sa traçabilité, on tente d'en verrouiller le support tout en la publicisant pour qu'enfin, l'archiviste soit moins considéré comme gardien des secrets que comme médiateur d'une mémoire fragile.

L'archiviste pourtant se voit aujourd'hui confronté à ce qu'on pense, à tort, être une sorte d'usurpation d'identité, ou dans ce cas de rôle. L'internaute, amateur passionné, créateur de son propre terrain de jeu, particulièrement attentif à l'univers dans lequel il évolue et qu'il modèle lui-même, se voit doté d'une capacité à accéder à la connaissance et à agir dont il n'était plus conscient. La compréhension du document, de la donnée, de l'information et de fait, de l'archive, n'est plus réservée à une poignée d'élus ayant la mission de répandre la bonne parole ; aujourd'hui,

[1] *Que contiennent les « Macron Leaks » ?* El Idrissi, 2017.

[2] Inspiré de Victor Hugo.

l'internaute *lambda* sait et réfute, conteste, dément, contredit, corrige, réajuste, prend en main. Le lapsus, l'erreur et le mensonge – ou toute autre vérité alternative[3] – n'ont plus le privilège de l'oubli, de l'affaire étouffée ou de l'indifférence.

De fait, ces archives, dérobées aux coffres-forts, arrachées à leurs retraites, exposées anarchiquement aux regards affamés et attentifs des journalistes, scientifiques, et autres décrypteurs, ne sont jamais silencieuses. Elles parlent de ce que l'on veut leur faire dire, après ce long travail de contextualisation, d'authentification, de sélection, et sont soigneusement mises en scène pour contribuer, justifier, faire éclater un nouveau scandale. On ne peut oublier, en effet, qu'entre *buzz* et instantané, la situation politique et médiatique actuelle dynamise le statut de lanceur d'alerte. Le numérique facilite tout autant que complique l'anonymisation, la diffusion, le développement de rumeurs qui frôlent parfois le complotisme. Les documents et données naissent avec une facilité déconcertante quand leur destruction, elle, s'avère plus complexe que jamais ; les « œuvres de l'esprit » se retrouvent volées, violées, protégées par des lois surannées incapables de rattraper le train en marche du numérique et des technologies du XXI^{ème} siècle. D'un côté, les textes prolifèrent sur la manière de gérer les réseaux numériques, pour protéger, censurer, contrôler ; de l'autre, les pratiques numériques puisent leurs forces dans des règles établies alors que les tous premiers réseaux n'en étaient qu'à leurs balbutiements. D'une part, on tente de réfréner ; de l'autre, on prône l'anarchie.

Archivistes et lanceurs d'alerte se présenteraient, dans cet environnement, comme les deux faces d'une même pièce ; les premiers resteraient gardien de l'archive, fervents protecteurs de son intégrité mais aussi de son secret ; les seconds iraient communiquer cette même archive à tout un chacun, après l'avoir violemment arraché aux mains des premiers. Pire ! journaux, blogueurs et autres médias du web se permettraient de parler des *archives des archivistes* comme des *archives des lanceurs d'alerte* comme si, par ce changement de main – mais non pas de propriétaire – les lanceurs d'alerte eux-mêmes devenaient archivistes. Certes, le lanceur d'alerte ne se contente pas de *voler* et de diffuser l'archive ; il la trie, la sélectionne, la conserve, la classe, la traite. Devient-il pour autant archiviste ?

Si la question peut faire sourire, elle peut tout de même paraître légitime – jusqu'à un certain point. L'archive revêt de multiples formes et englobe une multitude de données et documents aussi divers les uns que les autres ; son cycle n'a rien de figé et de fixe et invite à des modelages multiples en fonction de ce qu'elle mémorise. De fait, si tel document est archive avant qu'il ne soit volé, il paraît légitime de penser qu'il l'est toujours après son vol ; mais est-ce toujours la même archive ? si son propriétaire ne change pas, l'usage qui en est fait est parfois à des lieues de l'usage premier, originel, de cette archive. Qu'est-ce que l'action constituée par la fuite modifie dans l'essence de l'archive qui puisse justifier le refus de faire l'équation *leak = archive* ? comment les documents et données sont-ils employés par les lanceurs d'alerte ? de quelles archives viennent-ils et quel processus leur est appliqué avant et pour qu'ils deviennent *leak* ? les leaks seraient-ils les archives des lanceurs d'alerte ? il semble déjà complexe de définir ce qui tient de l'archive et de ce qui n'en tient pas, mais il conviendra tout de même de comprendre pourquoi ils – ou ceux qui les communiquent – disent leurs documents *archive*.

[3] Selon Donald Trump.

De même, comment définir la notion de possession et de propriété de l'archive lorsqu'elle est diffusée par ceux qui ne l'ont pas produite et qui en ont dépossédé les producteurs ? car si l'on sait que la fuite documentaire, aujourd'hui, entend plus la copie illégale de documents que leur vol, les producteurs d'archives confidentielles, secrètes, privées se voient dépossédés de ce statut très clôturé dont ils ont affublé leurs archives. En quoi le traitement que les lanceurs d'alerte appliquent aux documents dont ils s'emparent peut-il s'approcher du cycle de vie des archives tel qu'on le perçoit au sein de l'espace numérique ? le peut-il, seulement ? en est-ce l'objectif ? car si les lanceurs d'alerte sont sensibles à la question de la mémoire, leur objectif immédiat semble être, souvent, la nécessité de transparence. Ainsi, peut-être les notions de fiabilité, d'authenticité ou d'intégrité de l'archive ne sont-elles pas au centre de leurs préoccupations, et certaines problématiques seraient délaissées en faveur d'autres plus en adéquation avec leurs besoins.

Le leak ne finirait-il pas même par s'imposer comme une forme d'archive ouverte ? on entend, de plus en plus, parler de l'ouverture des archives, de la volonté des gouvernements et des pouvoirs publics à ouvrir à tous et pour tous l'accès aux données et aux documents, sous forme de bases de données lisibles et consultables pour tous ; en France, l'actualisation de la loi *Informatique et libertés* entend découler de ces grandes tendances. Ces dispositions, louables sans aucun doute, posent cependant la question très épineuse de la gratuité de l'accès à des archives qui rapportent et qui, surtout, font partie d'une culture du secret très profondément ancrée dans nos sociétés occidentales. Jusqu'où, alors, pourrait-on aller dans cette notion d'ouverture des archives ? les dispositions prises par les lanceurs d'alerte, leakers et hackers qui diffusent des œuvres musicales, littéraires, cinématographiques, mais également et surtout des alertes sur des sujets politiques, économiques, sociétaux sensibles ne seraient-elles pas pleinement dans cette optique d'*open archive* ? les lanceurs d'alerte ne seraient-ils pas en train de redéfinit la transparence de l'archive ?

En quoi, finalement, **le leak serait-il archive ?**
Cette problématique, relativement provocante, entend réfléchir la place de la fuite documentaire comme ajustement de l'archive au sein des réseaux enchevêtrés de la toile, et d'observer la manière dont les lanceurs d'alerte se sont emparés des mécanismes web pour récolter, authentifier, pérenniser et diffuser leurs leaks et leurs alertes. Qui de mieux pour tenter d'y répondre que des figures maîtres de l'archive, du leak et d'internet ? pour ce faire, ainsi, nous nous appuierons sur les travaux de Benjamin Loveluck, Marie-Anne Chabin, Louise Merzeau ou encore Gabriella Coleman, pour ne citer qu'eux.

Cette problématique entend, avant toute analyse, de se pencher sur la manière dont le leak s'est frayé une place sur la scène numérique en reprenant les codes originels d'internet, et comment le monde du Libre légitimise, d'une certaine manière, la fuite documentaire et le lanceur d'alerte. Nous commencerons donc par nous interroger sur les règles qui régissent l'univers internet et sur le contexte historique qui aura vu naître de telles démarches. Puis, si l'objectif de ce travail n'est certainement pas de comprendre ce qui pousse les lanceurs d'alerte à agir et à diffuser, il sera toutefois nécessaire d'en saisir les contours pour analyser, le plus neutralement possible, les démarches de diffusion et la représentation de ces documents et données.

Dans un second temps, nous poserons le contexte de ces différents scandales les organismes qui les supportent, avant de pouvoir, enfin, observer et analyser le processus du leak qui conduit à nous interroger sur sa dimension d'archive. En

s'essayant d'abord en un parallèle entre les 4C de l'archive et ceux du leak, nous proposerons ensuite une typologie du leak qui offrira le cadre nécessaire à sa définition en tant qu'archive. Enfin, parce que la représentation du leak, tout comme la représentation de l'archive, définissent ce qu'on en fait et comment on les emploie, nous tenterons de présenter le portrait qui nous permettra, peut-être, d'affirmer que le leak constitue une forme d'archive.

COMMENT INTERNET A (RE-)CRÉÉ LE *LEAK*

> *« Champions du partage des savoirs, héritiers de Voltaire et d'un long combat contre la censure et l'obscurantisme, contre la diplomatie secrète et les magouilles entre princes, nous ne pouvons pas ne pas voir, à travers le hacker, l'ouverture virtuelle d'un nouvel espace de délibération collective. […] Mais nous ne pouvons non plus, héritiers de Montaigne et du soupçon critique, ne pas nous demander pourquoi il est recommandé de 'faire sortir l'information' mais interdit de savoir d'où et comment »* (Debray, 2013)

Internet n'a pas créé le leak. Ce dernier, héritier d'une (pré-)histoire qui voudra que la parole de l'homme soit instable et peu fiable, naît le jour où le premier secret est divulgué. Certains diront avant : la création du secret sous-entend que l'information ne doit pas être accessible, donc qu'elle l'a été et qu'elle ne le devait pas. Le secret naît d'une frustration pour en créer une autre, et le leak, meilleur ennemi ultime, entend le justifier tout autant que lui couper l'herbe sous le pied. *Bromance* à l'américaine, déclinée en autant de langues qu'il existe de langages, le duo leak et secret s'installe sur la scène numérique dans un emménagement agité, scandaleux, mais pas illégitime pour autant.

Mais que leur a fait internet ? rien, sinon inoculer sournoisement la transparence et l'esprit libertarien entre les deux entités d'un couple inséparable. Internet n'a pas créé le leak, le leak n'a pas créé internet, mais les deux sont manipulés et façonnés par les mêmes idéaux de liberté, de savoir, d'expression. Internet offre un terrain de jeu fécond et protéiforme au transport du leak et à son évolution ; le leak offre un combat ardent et contesté aux acteurs et détracteurs d'internet.

Car parler de leak, ce n'est pas toujours parler d'alerte ; parler d'alerte, c'est rarement parler de leak. Les motivations du leak ne s'adaptent pas toujours aux conditions de « bonne foi » prônées par les protecteurs de la loi et du règlement qui consentent à considérer l'alerte comme un outil politique utile. L'alerte tente de justifier la fuite, la fuite se présente comme l'unique outil à l'alerte, et l'amateur prend la place du lanceur d'alerte, du journaliste, et même du justicier.

Ces pratiques ne sont pas nées avec l'avènement d'internet, mais existent sans aucun doute depuis la nuit des temps ou, au moins, depuis l'existence du document quelle que soit sa forme. Ainsi, l'histoire nous a appris à de nombreuses reprises le risque inhérent à la constitution de documents scellés, secrets, confidentiels, et la propension au vol et à la récupération de ces données. Ne nous y trompons pas, c'est bien la récupération d'information à des fins contraires aux volontés des sociétés – au sens structurel du terme – qui a conduit à la fabrication du *secret*, et non l'inverse

Complexe, dans cet univers, de définir les frontières de l'information et de l'alerte qui lui est associée. Nous chercherons, ici-même, à entendre ce qui fait d'internet ce terrain de jeu si libre et *anarchique* pour tout élément qui s'y trouve : des internautes à ses productions (traces, informations, documents). Un second chapitre nous permettra d'observer plus en avant le monde Libre prôné par les défenseurs du web, d'internet, de l'informatique, mais également par les journalistes ; enfin, nous observons la fuite documentaire en nous attardant plus particulièrement sur ses acteurs principaux : les lanceurs d'alerte.

INTERNET, BERCEAU DU *LEAK*

Internet, ce *no man's land*

Qu'est-ce qu'internet ?

Internet est un terme abrégé pour *internetting*, éventuellement traduisible en français par *interconnecter des réseaux*, et le réseau, un « ensemble de lignes qui s'entrecroisent plus ou moins régulièrement », soit au sens figuré, un *ensemble de relations*. *Internet*, c'est donc une *interconnexion d'ensembles de relations*. Cette définition, floue s'il en est, offre toutefois une certaine profondeur d'interprétation. En effet, *internet* serait en ce cas l'ensemble *majeur* intégrant une multitude d'ensembles *mineurs*, intégrant eux-mêmes de multiples connexions et relations entre elles. Une toile d'araignée aux fils illimités et à la mise en abyme qui ferait pâlir d'envie les Époux Arnolfini.

Ce que n'est pas internet, comme le dit très justement Dominique Cardon, c'est « un média comme les autres » (2010, p. 7) qui serait simplement constitué d'une amélioration technique des médias traditionnels, une évolution qui suivrait le cours naturel des évolutions technologiques de l'histoire et qui épouserait les moules mille fois visités de la culture communicationnelle et informationnelle de l'humanité. Il y a dans la définition du web, la construction d'internet, la fondation de l'espace numérique, une dimension de *lien* et de *partage* qui n'avait pas été expérimentée jusque-là dans les autres médias.

Mais internet n'en devient pas pour autant une zone de non droit. Il est vrai qu'il peut être assez facile de se laisser bercer par les inquiétudes des technophobes qui reprocheraient notamment au numérique de se faufiler dans une brèche juridique qui n'a pas su, ou trop tard, s'adapter aux évolutions techniques. Toutefois, la principale difficulté qu'ont les autorités à réguler l'espace numérique tient en ce que l'esprit du web ne peut s'accorder ou s'adapter avec la manière dont on régule nos sociétés aujourd'hui.

La raison en est simple, et s'explique par le fait que « l'esprit du web plonge ses racines dans la contre-culture américaine des années 1960 » (Cardon, 2010, p. 21). Ces années (60/70) voient exploser les mouvements hippies et, de fait, une idéologie qui tend à repenser la manière dont l'individu s'investit, prend place, et créé la société. Ils développent « toutes sortes d'outils d'autoproduction afin de préserver leur autonomie » et explorent « la manière dont l'information fait système » (Cardon, 2010, p. 22).

Lorsqu'internet naît il est donc déjà imprégné de cette « mouvance communautaire, écologiste et autarcique » (Cardon, 2010, p. 22) qui fait partie des pratiques de « coopération, de co-conception et de réputation auprès des pairs »(Cardon, 2010, p. 13) dont sont friands les informaticiens. Ils ont « établi un code déontologique qui valorise l'autonomie, la liberté de parole, la gratuité, le consensus, la tolérance » (Cardon, 2010, p. 13) et ce code est ce qui, aujourd'hui encore – voire spécifiquement aujourd'hui – rend toute tentative de réglementation ou de régulation des pratiques de l'internet inefficace à plus ou moins long terme. Parce qu'elle « valorise une culture de l'échange et de la coopération entre égaux, l'architecture de l'Internet accorde peu d'importance aux règles de centralisation, de hiérarchisation et de sélection » (Cardon, 2010, p. 16). Mais cette architecture n'en devient pas pour autant anarchique au sens de chaotique ou d'informe, et subit en vérité des règles strictes de gestion de la communauté.

Pour mieux comprendre ce qu'il en est véritablement de ce « chaos » brandi par les technophobes pour désigner internet, il faut revenir aux définitions du libertarianisme et de l'anarchisme qui sont au cœur même de la création d'internet, entre les hippies et les cyberpunks, les hackers et les informaticiens. Benjamin Loveluck met en lumière dans son chapitre intitulé « Internet face à l'État : *hackers*, crypto-anarchie et droits civils numériques » (2015, p. 162), c'est la difficulté de différencier ce qui tient du *libertarianisme* et de l'anarchisme.

Le Larousse définit la philosophie libertarienne comme reposant sur « la liberté individuelle conçue comme fin et moyen »[4], c'est-à-dire comme laissant à chaque individu la liberté de jouir de son droit de propriété sur ses possessions ainsi que sur lui-même. Héritée du libéralisme, elle conçoit la liberté comme fondamentale au point où la plupart de ses partisans prônent un antiétatisme où l'État serait réduit à ses fonctions régaliennes afin de conserver, notamment, la liberté du marché. L'anarchisme, lui, conçoit l'antiétatisme comme rejet, pur et simple de « toute tutelle gouvernementale, administrative, religieuse et qui privilégie la liberté et l'initiative individuelles »[5].

L'une des raisons justifiant cette levée de boucliers à l'heure des tous premiers *hackers* et des premières formes d'autorégulation et d'autoréglementation sur internet se présente donc, notamment, par ce mélange fait entre libertarianisme et anarchisme, qui conduit à fortement réguler la cryptographie[6] des communications et des données informatiques (Loveluck, 2015, p. 163). Celle-ci, tout d'abord vue comme un outil résolument anarchique, vit sa représentation adoucie : en tant qu'outil, elle fut vue tout autant comme un moyen de « protéger la propriété intellectuelle et les communications des entreprises privées, sécuriser les transactions bancaires, assurer la confidentialité des opérations de maintien de l'ordre de la part des gouvernements, et enfin protéger les communications privées et les données personnelles des individus » (Loveluck, 2015, p. 165), que comme celui de dissimuler des activités illégales. Encore aujourd'hui, le fond des débats concernant les libertés informatiques se retrouvent dans cette confrontation entre « la protection des données personnelles et des communications privées, d'une part, et l'ouverture des informations publiques, d'autre part » (Loveluck, 2015, p. 165) et dans la difficulté, voire l'impossibilité, de formellement distinguer là où finit l'une, ainsi que là où commence l'autre.

Mais ce qui rend les technophobes (et pas seulement) si inquiets à propos d'internet et de l'espace numérique en règle générale, c'est cette faculté qu'a l'internet à s'autoréguler sans intervention de la part des régulateurs *officiels*.

Autoréglementation, autorégulation, autogestion

L'autoréglementation se définit comme un ensemble de règles établies par une entité engagée comme responsable. Les premières, et encore aujourd'hui les plus légitimes, tentatives de régulation des comportements et des structures numériques sont construites par des organismes tels que le World Wide Web Consortium, organisme de standardisation dirigé par son fondateur Tim Berners-Lee, ainsi que

[4] Larousse en ligne : http://www.larousse.fr/dictionnaires/francais/libertarien/10910910

[5] Larousse en ligne : http://www.larousse.fr/dictionnaires/francais/anarchisme/3276

[6] Le Larousse la définit comme l'ensemble « des techniques de chiffrement qui assurent l'inviolabilité de textes, et, en informatique, de données.
En ligne, http://www.larousse.fr/dictionnaires/francais/cryptographie/20864?q=cryptographie#20742)

l'Internet Society, association américaine avec pour vocation la promotion et la coordination des réseaux informatiques. Ces deux structures sont toujours reconnues comme deux des autorités les plus importantes en ce qu'il s'agit du *droit* de l'Internet, et si elles tendent à ne plus s'exclurent des discussions juridiques et politiques engagées par les pouvoirs gouvernementaux, elles gardent une influence importante et non-supervisée par le droit.

Pour Yves Poullet (2003), cette autorégulation ne constitue pas une zone de non-droit au sein des réseaux informatiques mais pourrait, et devrait, être « une forme de droit » à la condition qu'elle « respecte cette triple condition de légitimité des acteurs, de conformité du contenu, d'effectivité des règles » (2003, p. 11). Une autoréglementation ne pourrait se justifier que par la « coexistence et le dialogue entre diverses sources de réglementation » (ibid.), et ainsi rendre effectif un « pluralisme juridique » (ibid.) qui puisse prendre en compte, tant une dimension globale qui régisse l'entièreté des réseaux informatiques, qu'une dimension locale qui prenne en compte la multitude de cultures, de traditions, de droits complémentaires et opposables constitutifs d'un Internet accessible à la totalité des états et des sociétés du monde.

L'équipe de la faculté de droit de l'Université de Montréal explique cette forme de droit par sa « référence aux normes volontairement développées et acceptées par ceux qui prennent part à une activité » (Trudel, Abran, Benyekhlef, & Hein, 2003, p. 22). C'est donc directement l'usager de l'outil qui valide ces normes et cette déontologie, lui accordant de fait le pouvoir d'agir, d'exécuter, de faire respecter voire de punir ceux attentant à ces règles. Ensemble de bonnes pratiques et de chartes d'éthique, l'autoréglementation émane « d'organismes spécialités et [s'impose] le plus souvent en raison de l'expertise que possèdent ces organismes » (ibid., p. 11). De fait, l'adoption de telle ou telle régulation ne dépend que de la légitimité qui lui est accordée par la popularité qu'elle peut avoir auprès des utilisateurs, et non pas d'une légitimité *de droit* qui se ferait confirmer par les pouvoirs publics.

Comme ils l'expliquent notamment, c'est la maîtrise de la technique qui permet d'accorder un tel pouvoir à ces organismes privés, et ce sont majoritairement les utilisateurs eux-mêmes qui tendent à faire respecter ces règles. Au-delà même des organismes privés reconnus comme compétents dans cette vague d'autorégulation, ils reconnaissent que « ceux qui ont la maîtrise d'un lieu dans le réseau ont la possibilité d'adopter des politiques relativement à l'accès au site, aux comportements acceptés et aux autres prohibés » (Trudel et al., 2003, p. 11). Les différentes chartes, éthiques, déontologiques, les codes de bonne conduite et les applications de sanctions que l'on voit fleurir sur les différents sites, forums, réseaux sociaux, correspondent eux aussi à cette autoréglementation qui échappe – en partie – aux politiques juridiques des gouvernements.

Pourquoi seulement en partie ? Les organismes tels quel la CNIL[7], en France, autorité administrative indépendante, veillent à l'application de certaines règles assurant un respect des droits de l'Homme et du Citoyen, à la vie privée et aux libertés en ce qui concerne l'informatique. Si elle n'est pas régie par le gouvernement, elle reste financée par celui-ci et parle au nom d'un pouvoir public légal. Elle est, notamment, garante des comportements en ligne ainsi que de la protection des données personnelles, et s'inscrit en ce sens dans l'application des Conditions Générales d'Utilisation et autres Mentions Légales, ainsi que de la

[7] « Dans l'univers numériques, la Commission Nationale de l'Informatique et des Libertés est le régulateur des données personnelles » – https://www.cnil.fr/fr/la-cnil-en-france.

déclaration des entreprises ou des sites traitant et récoltant des données personnelles. Garante des lois *informatique et libertés* tout comme l'ensemble des CNIL de l'article G29, elle possède ainsi une autorité qui peut éventuellement s'exprimer au-delà du territoire national quand la situation des citoyens français peut l'exiger.

Mais ces organismes sont souvent peu influents lorsque la question de la territorialité s'en mêle. Lorsque la conception de l'espace et de la responsabilité donnée par la loi invite à une juridiction par pays, enracinée sur l'hébergement des sites web et la nationalité des auteurs, celle de l'autorégulation offre une vision plus complexe car inhabituelle, fondée sur les espaces et sous-espaces. Plus simplement, l'autoréglementation se fait par et dans les zones numériques, s'adaptant voire s'écrivant entièrement en fonction des besoins, des cultures, des règles éthiques de chacune des communautés numériques. D'un forum à l'autre, d'un réseau social à l'autre, d'un site web (personnel, professionnel, associatif) à l'autre, les règles ne sont pas les mêmes et n'invitent pas aux mêmes sanctions.

Cette multiplicité des normes et principes en vigueur via l'autorégulation se ressent également au sein même des institutions juridiques, avec les lois *Libertés et Informatique* ou, plus lourd encore, le *Règlement Général sur la Protection des Données personnelles*[8], règlement européen très théorique adaptable à chaque pays, chaque culture, et dont l'application dans chaque état voire même région ne saurait ressembler à aucune autre.

Ainsi, si les tendances viennent à changer[9], on ne peut que constater l'incapacité des textes quant à la gouvernance d'internet, et la faculté qu'ont les multinationales de type GAFAM[10] à monopoliser le marché économique, financier, politique et, plus largement encore, technologique d'internet. Les pays engagés sur ce terrain glissant qu'est la réglementation d'une *société* avec pour fer de lance l'autorégulation et l'autoréglementation se heurtent systématiquement à des problématiques aussi diverses et complexes que la territorialité, la temporalité, la possession ou l'identité des organismes, entreprises, groupes, individus et mouvances peuplant la sphère numérique. Nous ne nous engagerons nous-même que peu sur ce sujet délicat qui, sans connaissance approfondie du domaine juridique notamment, nous serait aussi périlleux qu'hasardeux.

D'autant qu'en mettant la majorité du pouvoir décisionnel, productif, et de sanction entre les mains de ses utilisateurs, internet refuse catégoriquement « toute politique paternaliste qui définirait pour les autres ce qu'il convient de dire ou d'entendre » (Cardon, 2010, p. 42). L'autorité qui importe et à qui l'on rend compte, dans l'espace numérique, c'est donc l'auto-organisation des internautes, c'est-à-dire *l'autre*, sans préoccupation hiérarchique. L'internaute qui semblera posséder plus de *pouvoirs* qu'un autre ne l'aura en vérité que parce que la *communauté* lui prêtera ce pouvoir, et s'en verra dépourvu dès lors qu'il nuira aux intérêts de la communauté ou enfreindra les règles éthiques et déontologiques instaurées par cette même communauté.

Cette gestion permet notamment de *faire taire* sans *supprimer* l'information ou l'existence numérique de l'internaute. Il est en effet régulièrement reproché aux

[8] Ce règlement européen, sorti en 2016, commencera à sanctionner à partir de 2018 les entreprises et organismes qui n'auraient pas mis en place une politique de gestion des données personnelles, notamment.

[9] Nous pensons notamment à la révision de la loi *Informatique et Libertés* qui semble enfin prendre la pleine – ou presque – mesure de la multiplicité d'internet.

[10] Google, Amazon, Facebook, Apple et Microsoft.

nouvelles technologies, et plus spécifiquement à internet, de ne pas poser de garde-fous, de *gate-keepers* susceptibles de fermer l'accès aux publications comportant des informations fausses, à la diffusion illégale, illégitime, interdite, etc. Il est évident que ces informations, aussi outrageuses et aberrantes qu'elles puissent être, possèdent sur la toile un espace de visibilité qui ne saura, ou très difficilement, leur être retiré. Cependant, si ces « propos peuvent être accessibles (c'est-à-dire *visibles*) » ils n'ont pas pour autant la prétention de « se voir reconnaître un caractère public » (Cardon 2010, 40). Ou plutôt, s'ils l'ont, la communauté d'internautes et leur architecture d'auto-organisation permettra que les « informations qui doivent rester dans les bas-fonds du web ne remontent pas les échelles de visibilité » (ibid., 2010, p. 42).

Cette autogestion permet également, à l'inverse, de *faire parler*, de *rendre public*, de *diffuser* sans craindre la suppression ou l'inaccessibilité, ou du moins en dépassant ces contraintes-là. Cette architecture dote en effet « chaque utilisateur du pouvoir d'innover, de rendre visibles ses innovations et de les diffuser à tous ceux qui les jugent pertinentes » (Cardon, 2010, p. 17) et de fait de s'affranchir des interdits légaux et administratifs imposés hors de l'espace numérique[11]. Pour exemple, Wikileaks offre la possibilité, en partie grâce à cette architecture, de rendre « anonymement publiques (et donc disponibles pour des opérations critiques) des 'fuites' issues de sources officielles ou des informations de dissidents vivant dans des pays soumis à un contrôle de l'information » (ibid., p. 76).

Dans ces cas, il est complexe voire impossible de contenir la diffusion de ces informations, quels que soient les moyens dont on peut disposer. Dominique Cardon explique la difficulté de cette régulation par le fait que le réseau est « une plateforme neutre ». Cette assertion n'a rien de candide ou de naïf puisqu'en vérité, elle s'appuie sur l'idée que le réseau est « dépourvu de tout centre » (Cardon, 2010, p. 17), et que donc il n'y a rien à *contenir*, à *retenir*, à *détenir*. Le réseau des réseaux s'attache au contraire à communiquer, diffuser, partager, et ce par deux moyens : « la connectivité, et l'extension continue du réseau » (ibid., p. 16).

Mais au-delà de l'essence même d'internet, c'est également et surtout, peut-être, deux mondes qui s'opposent, s'affrontent, et tentent de dominer l'autre. La multiplication des usages numériques, la démocratisation de l'accès à internet et l'idéologie de pratiques qui y règne ont mis en lumière un aspect élémentaire à l'introduction de tout nouveau média (quand bien même celui-ci serait *particulier*). Edwy Plenel, fondateur de Médiapart, fait état d'ailleurs d'un espace numérique à l'enjeu « clairement politique » qui offre l'opportunité d'un « affrontement de nouvelles émancipations et de vieilles dominations » (Plenel, 2013, p. 136).

Cet affrontement se caractérise principalement par les individus qui se croisent et se répondent dans un espace qui, nous l'avons vu, *leur appartient sans leur appartenir*. Ce paradoxe engendre une dualité identitaire dans l'existence même des internautes au sein de l'espace numérique. Où et comment, de fait, l'individu, ou plus exactement l'internaute, se positionne dans cet espace ? c'est ce que nous allons tenter d'éclaircir.

[11] Ainsi que *dans* l'espace numérique mais, comme nous l'avons dit, les pouvoirs publics se heurtent à un espace qui a été justement créé et développé avec l'idée de s'affranchir de ces règles et règlements.

L'individu, l'internaute, et l'espace numérique

Qu'est-ce que l'identité numérique ?

On entend communément par *identités numériques* ces facettes de notre identité que nous présentons numériquement, en fonction des mondes virtuels auxquels nous participons. Sans remettre en question l'idée commune selon laquelle l'identité est en constante évolution, on considère que l'identité numérique est également en constante évolution, se construisant avec et par *l'autre* comme une « une co-construction négociée entre les inter-actants » (Coutant, 2011), soit entre les autres internautes.

Il faut, pour définir l'identité numérique et comprendre ce qu'elle implique, partir tout d'abord du principe que l'identité de manière générale est un miroir, une représentation de ce que nous sommes à un instant T. Cette représentation de soi dans le monde virtuel, cette « sculpture agissante de soi » que Fanny Georges nomme *hexis numérique* (Georges, 2008) se modèle en fonction de nos interactions, de l'espace dans lequel nous évoluons, et est donc « moins un dévoilement qu'une projection de soi » (Cardon, 2009). Il n'y a donc pas d'identité *réelle* opposée à une identité *virtuelle*, encore moins deux – ou plus – identités particulières et distinctes, se découpant elles-mêmes en de multiples et nombreuses *représentations de soi*, mais il y a des *traces* identitaires, ou du moins laissées par la présence de l'identité, et c'est cette « somme des traces numériques se rapportant à un individu ou à une collectivité » (Ertzscheid, 2013) qui constituent l'*hexis* numérique.

Cette *hexis* distingue trois types d'identités : calculée, agissante et déclarative.

L'*identité déclarative*, cette « description de la personne par elle-même » (Georges, 2008) est la base qui permet aux identités agissantes et calculées d'exister. C'est la première étape de l'action engagée afin d'être présent électroniquement, du nom que nous attribuons à notre ordinateur aux informations renseignées lors de la création d'un profil sur un réseau social, un site d'e-commerce ou un jeu en ligne. Il n'importe pas qu'elle reflète exactement le soi incarné derrière l'écran : comme évoqué, elle peut déclarer un personnage fictif ou une partie spécifique de notre identité, comme un profil professionnel ou militant. Elle nous apporte des données strictes telles que le pseudonyme, l'avatar, la date de naissance ou le sexe, mais également des données plus interprétables telles que des courtes rédactions de descriptions physiques, morales, souvent présentées sous la forme d'un « à propos de moi ». Enfin, c'est celle qui reste la plus contrôlable et qui invite pourtant souvent à des dérives de la part des utilisateurs : on choisit consciemment d'indiquer son adresse postale ou son numéro de téléphone, par exemple, qui risqueraient de se faire pirater et de se retrouver en libre accès sur le web. Mais c'est également celle qui permet le plus de se distinguer, car plus elle est « forte et détaillée, plus la représentation singularise la personne par différenciation » (ibid.).

L'*identité agissante* est constituée de l'ensemble des *traces temporaires* de nos actions. Ainsi les liens hypertextes, images, vidéos, enregistrements, fichiers en tout genre des « objets magiques » aux « amis » Facebook sont autant d'objets rencontrés et avec lesquels nous interagissons qui composent le « relevé explicite [de nos] activités » (Georges, 2008) : cette identité est le « produit immédiat » de l'activité de l'utilisateur. Ces données ne sont pas remarquées pour rien : tantôt elles étoffent une identité déclarative, tantôt elles alimentent les rouages de l'identité calculée.

Enfin, dernier maillon de la chaîne identitaire, l'*identité calculée* correspond aux données créées par le système en fonction des données fournies par l'identité

agissante. Le nombre de *j'aime* et de commentaires, d'amis, de tableaux sont autant de pré-calculs permettant ensuite d'établir des « classements » où le chiffre, sacro-saint du monde numérique, décide de la visibilité, du référencement, de la notoriété de telle identité. On ne parle dès lors plus d'identité mais de « système identitaire ». Ces variables ne sont d'ailleurs pas que quantitatives : des données immédiates déduites par le système telles que « untel est en train d'écrire », « vu par untel », ou « untel est connecté/occupé/absent/en train de… » qui découlent elles aussi directement du « traitement de l'identité agissante [ou non agissante, justement] par le système » (Georges, 2008).

De fait, si de premier abord nous pourrions croire que l'identité déclarative est celle qui *compte* le plus, au sens où elle devrait prendre plus de place que les autres, la réalité s'avère autrement plus complexe et *complète*. En effet, l'*hexis* numérique est bien composée de ces trois identités sans qu'il soit possible de calculer la part de chacune dans la *représentation de soi* dans sa globalité. De plus, une action que l'on pourrait faire appartenir à l'identité déclarative au premier abord peut s'avérer trompeuse : pour exemple, « l'écriture de soi sur les profils vaut moins pour le déclaratif en lui-même que pour les réactions attendues de nos audiences » (Coutant, 2011), et donc, pour l'identité calculée. La seule véritable constante dans cette identité, et dans la création d'icelle, c'est que « l'information dont nous étions les émetteurs ou les récepteurs, est aujourd'hui ce qui façonne notre identité » (Merzeau, 2009), au point où nous sommes nous-mêmes devenus l'information que nous *trans*portons (Ertzscheid, 2009).

Une fois ce constat réalisé, il semble compliqué de combattre ce qu'il nous arrive de vivre comme une intrusion, puisque nous autorisons nous-mêmes ces intrusions à partir de l'instant où nous nous emparons des outils numériques qui sont, s'il faut le rappeler, nativement intrusifs (Casilli, 2010). Pourtant, si nous offrons évidemment une part de notre vie privée, intime, personnelle à l'instant où nous usons d'un outil numérique, cela ne veut pas dire pour autant que nous ne sommes pas capables de contrôler le poids de cette part et son extension à d'autres éléments personnels. En devenant *compétent*[12] à la « gestion de l'accessibilité des traces, de leur pérennité par opposition à la malléabilité des souvenirs » ainsi qu'à l'« évaluation des contenus pertinents » (Coutant, 2011), il devient aisé de garder en main, de contrôler, de maîtriser ces *hexis* numériques.

En considérant « la Toile comme un espace en clair-obscur, plastique et paramétrable », Dominique Cardon note que les utilisateurs « exploitent les propriétés des différentes plateformes pour construire un public adressé tout en se cachant des autres » (Cardon, 2009), c'est-à-dire qu'ils usent de leurs identités déclaratives, agissantes et calculées pour définir quand, à qui, où, et comment le autres utilisateurs peuvent se représenter un individu. Pour ce faire, les internautes usent donc de compétences acquises par et pour la communauté dans laquelle ils évoluent, quel que soit leur degré d'expertise, et en forgeant de fait une légitimité de la compétence qui leur permettra d'être reconnus par leur pairs.

Qui sont les internautes ? le cas de l'amateur

Comment expliquer que l'espace numérique puisse permettre une telle polysémie dans l'orientation donnée à l'identité numérique ? Antonio Casilli note, à ce propos, que « l'architecture même du Web actuel représente la mise en place de

[12] Au sens de « maîtriser des compétences ».

la décentralisation et de l'autonomie prônées par ses pionniers » (Casilli, 2010, p. 87). Pour lui, « la spécificité de la culture numérique réside dans l'inséparabilité de ses valeurs politiques et de ses usages technologiques » (ibid.) et offre donc une nouvelle[13] palette de rôles au sein de la communauté.

Dans ces rôles, l'un attire particulièrement notre attention : l'amateur. Patrice Flichy présente l'amateurisme comme un reflet de « la volonté de l'individu de *construire son identité* » (2010, p. 87) qui s'inscrit dans un « *mouvement de diffusion et d'élargissement des savoirs et des compétences* » (2010, p. 88) et offre une « *société plus démocratique* » (2010, p. 89). Pour lui, l'amateur n'est rien d'autre qu'un expert au sens premier du terme, c'est-à-dire qui a *acquis de l'expérience*, qui *additionne des expériences* dans un domaine et qui, par là-même, obtient de ses pairs une légitimité suffisante pour faire acte d'expertise sur certains sujets.

C'est que l'amateur n'est pas un utilisateur du web passif. En vérité, les « amateurs de savoir partagent des expériences et mettent en forme les connaissances courantes », et dans le même temps « produisent aussi des connaissances par eux-mêmes, soit en collaborant avec les scientifiques, soit en élaborant des contre-expertises » (Flichy, 2010, p. 65). De fait, l'amateur n'a pas pour vocation de se substituer à l'expert-spécialiste, au sens de *professionnel*[14]. Ainsi, internaute par excellence parce qu'il produit autant qu'il reçoit, l'amateur est un individu possédant un pouvoir particulier au sein de la sphère numérique. Ses pratiques « débouchent sur une production d'informations et d'opinions qui comptent » (Flichy, 2010, p. 63).

Flichy va même plus loin :

> « L'amateur de la chose publique est un citoyen qui veut s'informer par lui-même, exprimer ouvertement son opinion, développer de nouveaux modes d'engagement. Il se méfie des experts-spécialistes et n'accorde pas toujours sa confiance aux représentants qu'il a contribué à élire. On est ici au cœur de la démocratie d'interaction » (Flichy, 2010, p. 43)

Pris dans son acception politique, l'amateur, sans pour autant être un militant, a un pouvoir de contestation et de critique qu'on ne peut négliger. En fait, en se permettant d'intervenir « dans le jeu politique en contestant une décision, en dénonçant un choix politique, en proposant une alternative » (Flichy 2010, 43) l'amateur se positionne comme un adversaire redoutable au sein de l'espace public et politique. Tout autant qu'il ne se substitue pas à l'expert-spécialiste, il « ne remplace ni le journaliste ni l'élu, mais les interroge, les surveille et les bouscule », et de fait « oppose […] au contrôle centralisé des pouvoirs politique et médiatique, un contrôle partagé des citoyens, facilité par Internet » (ibid., p. 63). En ce sens, « l'activité politique amateur étend le domaine de la citoyenneté » (ibid., p. 43).

Mais l'amateur ne s'adresse pas à tous. Parce qu'il a un public adressé, choisi, parce qu'il fait partie d'une communauté et qu'il se situe « dans une sphère autonome de production d'informations et d'opinions », les opinions et informations qu'il construit et partage « sont destinés à une communauté restreinte » (Flichy, 2010, p. 45). Lorsqu'elles ne le sont pas, c'est qu'ils « revendiquent leur position de citoyen concerné par un évènement ou une question précise » et que, par cette position et ce pouvoir citoyen ils « souhaitent contester le discours des experts-

[13] Entendu que l'idée de *nouveauté* réside ici dans la réorganisation et la réappropriation de ces rôles, et non pas dans l'apparition de rôles qui ne seraient pas préexistants aux évolutions technologiques de ces dernières années.

[14] Ou au sens *d'expert* comme on a l'usage de l'utiliser.

spécialistes qui les ont ignorés et n'ont pas pris en compte leur point de vue : ils veulent dénoncer des projets politiques, chercher à convaincre, rallier à une cause » (ibid.). De là, on note deux formes d'amateurs sur la question politique qui nous intéresse : l'amateur *de politique*, au sens de *passionné* par la politique, et l'*engagement en amateur* au sens militant de l'expression (ibid., p. 44).

Et si le premier cherche à s'exprimer, à entrer dans le débat public et à se faire entendre sans forcément convaincre – il a d'ailleurs souvent un rôle de dénonciateur – les seconds cherchent à créer un véritable espace numérique politique :

> « Les militants amateurs n'appartenant pas toujours à une organisation pérenne et structurée, leur action est souvent coordonnée selon un mode réticulaire. Internet est un outil idéal pour l'organisation de ces réseaux politiques. Ceux-ci n'apparaissent pas seulement dans des espaces politiques minoritaires, mais aussi au sein des grandes organisations. Leur objectif est alors de *convaincre et de constituer un réseau social du politique* » (Flichy, 2010, p. 61)

Les premiers n'ont pas plus d'espace que les seconds, et inversement. En vérité, comme le dit Flichy, internet « permet d'étendre la citoyenneté en facilitant l'expression publique de tous les citoyens » (Flichy, 2010, p. 43) sans qu'une acception en prenne le pas sur l'autre, et peut se permettre d'être « un dispositif d'expression et de débat public » et en même temps « une nouvelle configuration d'action » (ibid., p. 44). Et de la même manière que chaque communauté a et trouve sa place, quels que soient leurs objectifs, chaque citoyen est « autorisé et capable d'être mis en contact avec les responsables publics » (Casilli, 2010, p. 268) sans que sa légitimité ne puisse être questionnée par rapport à sa popularité ou sa réputation.

En vérité, tout internaute a prétention à être amateur à un moment donné de son existence, au sein d'une communauté particulière. Dès lors que l'individu peut se permettre d'intervenir dans une communauté afin d'aider un membre, de partager une expérience et de discuter à propos d'un sujet qui n'est pas en rapport direct avec le contenu de son activité professionnelle et d'expertise-spécialisation, il revêt le rôle d'amateur avec tout le pouvoir qui peut lui être associé.

Espace et expression

Ce pouvoir particulier vient du fait qu'internet peut être lu comme « un *espace* social où des *corps* interagissent pour créer des *liens* de coexistence » (Casilli, 2010, p. 11), c'est-à-dire une société du lien où la prétention hiérarchique n'a pas lieu d'être et d'exister. Cet abattement des barrières hiérarchiques conduit à une redéfinition de la notion de *responsabilité*, d'où l'émergence particulièrement importante de celle d'*amateur* qui sans être *officiel*, est responsable. Dans l'espace public qu'est notamment la sphère numérique, Patrice Flichy nous dit que « celui qui s'exprime s'engage […], il y a une responsabilité de l'énonciateur » (Flichy, 2010, p. 51).

Cet engagement, qui se traduit notamment par le militantisme développé sur internet par certains internautes, par l'engagement amateur, par l'amateur de politique et bien d'autres encore, s'explique par la forme *extime* que prend l'expression numérique. Cette notion, notamment développée par Serge Tisseron, sous-tend le « processus par lequel des fragments du soi intime sont proposés au regard d'autrui avant d'être validés » (Tisseron, 2011, p. 84), soit la nécessaire approbation de l'autre, et donc sa propre responsabilité. Ce mode d'expression qui substitue la parole individuelle à « la parole collective exprimée par des porte-parole

dûment mandatés » (Flichy, 2010, p. 55), puisque la parole individuelle se légitimise par le regard de la communauté, prend des formes diverses et variées : « la participation à l'écriture collective sur des sites encyclopédiques » (Flichy 2010, 48) en est une ou, moins connue, la participation au dépouillage, au tri, au classement de *leaks* effectuée sur des sites tels que *4chan* ou *Reddit*.

Ces deux exemples illustrent une volonté participative exacerbée par la possibilité d'être plus facilement reconnus par ses pairs, d'une part, et de faire partie d'un tout organisé et puissant de l'autre. Si l'on prend l'exemple de *Wikipédia*, on s'aperçoit assez vite que ces deux dimensions sont représentées. D'un côté, l'expert ou l'amateur participant à la rédaction, à la collecte d'informations, à la correction d'un article sera, si tant est qu'il s'en soit donné les moyens, aisément reconnu et respecté par les autres internautes qui mettront en avant sa parole, la porteront et l'amélioreront ; de l'autre, la publication d'un article ne fera l'objet d'aucune signature et le tout-un-chacun aura accès à la synthèse d'une réflexion de groupe qui constituera l'article en question. Bien sûr, il sera toujours possible d'observer le déroulé de la réflexion grâce à l'historique, aux débats, etc. mais au premier niveau de lecture, l'utilisateur de Wikipédia *lambda* reconnaîtra l'article comme un 'tout organisé et puissant'.

Le deuxième exemple donné, celui du tri et du classement de *leaks* sur les sites *4chan* et *Reddit*, met en avant un autre niveau de structuration. Dans l'exemple de *Wikipédia*, les participants sont pour la plupart d'entre eux dans un rôle de synthèse, où l'objectif est de résumer toute la complétude d'un sujet en un article lisible par tous, donc vulgarisé. Dans le second exemple, les participants ont pour objectif de *collecter* une information et d'en reconstituer le contexte. Ici, le collecteur « est un médiateur » (Flichy, 2010, p. 57) qui cherche moins à dénoncer qu'à « *donner un autre sens* à un évènement, à une conjoncture, à une séquence de fait » (ibid., p. 60). De plus, « plus le traitement de l'information est sophistiqué, plus il engage de ressources intellectuelles et techniques. Il se créé alors une nouvelle expertise, une élite qui se distingue du commun des militants » (Flichy, 2010, p. 57), et un nouveau mouvement cyclique se met en marche pour dénicher de nouveaux experts-amateurs qui feront circuler l'information jusqu'à sortir du « collectif d'origine » (Flichy, 2010, p. 57) et offrir à tous les utilisateurs une information traitée.

Ces activités de collecte et de veille sont capitales pour tout amateur, pour tout internaute qui se respecte. Si elles prennent des formes différentes selon les objectifs finaux de lecture et de compréhension, elles sont au cœur du traitement de l'information numérique. Dans le cadre de l'activisme électronique, ou du moins de l'amateur de politique ou de l'engagement amateur, ces activités permettent de « surveiller l'action des pouvoirs publics et des entreprises […], de dénoncer leurs agissements, de donner un sens à des phénomènes sociaux et, finalement, d'exprimer une opinion minoritaire en lien avec d'autres citoyens » (Flichy, 2010, p. 56). On est loin, ici-même, d'une action au long-court ou d'une professionnalisation de la pratique puisque ces démarches « relèvent plutôt d'une action ponctuelle liée à des situations, des évènements, des configurations particulières » (Flichy, 2010, p. 62) et n'ont pas prétention à être pérennes. Toutefois, ces « nouvelles formes démocratiques […] s'opposent aux règles de la démocratie représentative » notamment parce qu'elles n'ont pas spécifiquement besoin d'un représentant, et qu'elles évoluent hors du champ de l'action politique habituelle et, de ce fait, « elles sont donc souvent contestées par les élus eux-mêmes » comme une attaque ou un risque encouru à leur profession et leur pouvoir (Flichy, 2010, p. 62).

La crainte envers ces mouvements et cette dynamique numérique communautaire va même beaucoup plus loin, puisque nous l'avons vu, les pouvoirs publics peinent à s'imposer face à une architecture qui ne leur est pas favorable :

> « Malgré les restrictions du trafic des données, les filtres aux contenus et les dispositifs de surveillance imposés constamment par les pouvoirs étatiques, l'effort pour normaliser ce bourdonnement d'informations et d'opinions antagoniques s'avère vain » (Casilli, 2010, p. 87)

Traces, empreintes, données

Antonio Casilli distingue trois espaces auxquels nous sommes confrontés au sein du *méta*-espace que peut constituer internet : l'espace physique, technologique, et enfin l'espace social (Casilli, 2010, p. 29). S'il les sépare, c'est qu'un peu à la manière des identités (déclarative, agissante, calculée) la gestion de l'espace numérique est constitutive de l'existence de l'individu et de l'information sur le web. En fait, comme il l'explique, nous « associons volontiers l'information et la communication à la notion d'*espace* » (Casilli, 2010, p. 19) comme si ces deux notions étaient des véhicules, des objets qu'il fallait déplacer d'un point A à un point B et, de ce fait, nécessitaient pour exister une représentation consciente de l'espace.

On peut donc logiquement supposer qu'il existe des espaces ouverts et fermés, des espaces publics et privés, des espaces visibles et non-visibles, et ce tout en acceptant qu'ils ne soient pas tous strictement définissables[15]. Pour exemple, Casilli note que « les espaces privés peuvent s'ouvrir, accueillir des membres externes et solliciter à travers le don un comportement coopératif et pacifique qui ne rechigne pas à une certaine forme d'affectivité et d'identité commune » (Casilli, 2010, p. 50) et de fait ne pas se laisser enfermer par la stricte définition de la notion. Un espace, tout comme une identité, est profondément malléable et en constante évolution.

Mais s'il y a présence[16] (ou du moins *identité*), s'il y a espace, il y a donc aussi et surtout *empreinte*, *trace*, *résidu*. Agissantes, calculées ou déclarées, tout comme les identités (puisqu'après tout ce sont bien nos données qui constituent sur internet notre identité), ces traces sont inhérentes à notre existence en ligne. Tout comme l'école de Palo Alto nous a appris qu'*on ne peut pas ne pas communiquer*, Louise Merzeau nous fait remarquer qu'aujourd'hui, grâce ou à cause du numérique, « *on ne peut pas ne pas laisser de traces* » (Merzeau, 2009). Elle va plus loin d'ailleurs en précisant que ces traces, ces données qui viennent de nos activités et qui sont donc des données personnelles « ne sont donc plus seulement l'envers invisible de notre présence numérique. Elles sont devenues l'espace où nous naviguons » (ibid.). Ces traces qui visent à « calibrer au plus près des renseignements sur mesure (signature) » (ibid.) sont notamment traitées, classées, classifiées, indexées et constituent des bases de données exploitables par n'importe quel logiciel pour, notamment, affiner l'expérience utilisateur.

C'est cette expérience utilisateur qui est brandie pour justifier, aux yeux des consommateurs et utilisateurs du web, les usages toujours plus avancés et abusifs des données personnelles, parce que « le marquage des données participe d'une

[15] On a vu plus tôt que la construction de l'identité construit également l'espace dans lequel on pratique cette identité, et que donc un espace qu'un individu jugera public ne le sera pas pour un autre. Pour définir la notion d'espace, il faudrait donc choisir, en premier lieu, si on les définit selon l'individu qui s'en est emparé ou selon l'objectif de la communauté qui y est née, etc.

[16] Définir les notions de *présence* et *d'absence*, au cœur de la notion numérique, est pourtant un sujet et un débat houleux sur lequel nous ne nous aventurerons pas.

restructuration en amont des informations, qui a pour logique la personnalisation » (Merzeau, 2009). La trace seule ne peut suffire, il lui faut un contexte, un univers dans lequel elle évolue et de fait une possibilité de la *tracer*. De cette traçabilité, Merzeau dit qu'elle consiste « à pister les singularités pour cibler toujours plus finement l'information », et que de fait, « chaque acteur n'existe que dans son rapport discriminant aux autres usages, préférences ou opinions » (ibid.). Là encore, on voit apparaître en substance la question des différents niveaux, des différentes couches d'identités et de l'effet miroir appliqué à chacune d'entre elles.

Pour contrer cet effet de traçabilité, l'utilisateur peut avoir recours à deux types particuliers de *ré*-identification, le pseudonymat et l'anonymat. Le premier tente de rendre « inviolable l'accès aux données personnelles confiées aux tiers de confiance, tout en aussi faisant sauter les restrictions de l'anonymisation, qui freinent le développement du commerce électronique », quand le second garde « les traces d'une personne (caractéristiques, comportements, etc.) sans avoir la moindre possibilité de connaître sa véritable identité » (Arnaud, 2009). Ainsi l'anonymat garde l'identité *originelle*[17] mais supprime, efface, fait disparaître les caractères qui la rende identifiable, remarquable, quand le pseudonymat offre une « autre identité qui ne pourra pas être facilement rattachée à sa véritable identité » (Arnaud, 2009), et donc créé une facette supplémentaire.

Ainsi, « face à l'impossibilité de se soustraire aux systèmes de surveillance, c'est la mise en œuvre d'une *sousveillance* où l'acteur enregistre lui-même les indices de sa présence, qui peut l'aider à préserver l'intégrité de son identité » (Merzeau, 2009), qu'il choisisse le principe d'anonymat, de pseudonymat, ou tout autre type d'enregistrement, de contrôle et de surveillance des données qu'il diffuse. L'utilisateur numérique n'est pas un simple consommateur, il est acteur de par sa nature productive – *je suis (dans l'espace numérique) donc je produis* – et vecteur de sa propre information.

L'information, le document, l'archive

L'information numérique

Nous sommes, dit-on à tout-va, entrés dans la société de l'information. Parce que l'information ne se laisse plus aussi bien déformée par la mémoire et l'oubli, parce qu'elle se laisse traiter, enregistrer, maîtriser, organiser, archiver, elle contribue à construire une société, un monde gouverné par elle-même. La surcharge informationnelle à laquelle la plupart des chercheurs et journalistes tentent de sensibiliser les utilisateurs n'est pas un mythe et constitue sans doute l'un des plus grands fléaux[18] de notre ère. Avoir accès à cette information n'est aujourd'hui plus un problème, quand bien même dans de rares cas devrions-nous rechercher un peu plus longtemps, un peu plus ardemment l'information désirée ; le problème réside dans la capacité de l'utilisateur du Web « à traiter cette information (c'est-à-dire à la qualifier, à l'analyser et à faire des recoupements) » (Arnaud, 2009).

Louise Merzeau nous fait remarquer que « désormais, non seulement tout communique, mais tout informe » :

[17] Ou du moins déclarée comme telle sur l'espace numérique.

[18] Et si non l'un des plus grands, l'un des plus médiatisés.

> « Public, intime et privé se rejoignent dans une même sphère, où la mise en commun relève autant de la collecte et de l'indexation que de l'échange et de la publication. Le Web incite à la spontanéité du mode conversationnel, tout en documentant de façon pérenne chaque source ou 'prise de parole' » (Merzeau, 2009)

En effet, il ne suffit pas d'informer sur une action, l'espace numérique nous impose également de documenter l'information de l'action, et de faire de cette documentation une information traçant le processus de l'action en question, et ce de manière totalement illimitée. Chaque donnée, métadonnée, trace, empreinte, information, document, seuls ou justifiant un autre élément, est enregistré, consigné, marqué. Cette justification s'explique par l'« instabilité des dispositifs » qui « impose d'identifier l'état du document en même temps que son contenu, démultipliant les informations sur l'information. [Cette logique] indexe la valeur sur l'actualité (une information périmée n'en est plus une), tout en provoquant un développement sans précédent des métadonnées » (Merzeau, 2009). On ploie, aujourd'hui, sous la diversité et le poids de ces données qui fleurissent et coulent à flot quoi que l'on fasse, où que l'on soit. L'information se fragilise en perdant la linéarité à laquelle elle nous avait habitué, et dans le même temps se pérennise en s'étoffant de ces multiples traces de son contexte, de son historique, de son origine et des éléments qui ont permis à un moment *T* de la lire, de la comprendre, de la construire. Pour Merzeau, c'est la « valorisation des empreintes » qui a permis, seule, de révéler « la double nature de l'information, à la fois instable et pérenne » :

> « Pendant longtemps, on a voulu croire que les messages circulaient dans les réseaux sans sillage, et on ne s'est pas privé de dénoncer cette volatilité des contenus. De plus en plus de documents sont de fait des structures dynamiques dont les composantes se renouvellent continuellement. L'organisation des traces en hypertextes accentue la fragilité des échafaudages sémantiques ou logiques – la bifurcation devenant le pivot de toute navigation. La rotation rapide des standards et des formats rend pour finir tout système numérique en lui-même éphémère. » (Merzeau, 2009)

Exit la *linéarité*, bonjour la *transversalité*, la réflexion réticulaire – puisque nous sommes en réseau – l'absence de système central et fédérant : on s'organise désormais les uns par rapport aux autres, non plus selon un représentant qui centraliserait les informations. De fait, elle note à juste titre que la traçabilité dont il est question pour cette information « appelle une science du document élargie aux raisons pragmatiques, et une science de la communication tournée vers les processus documentaires » (Merzeau, 2009). On ne peut plus se référer à « un document maître et des copies », il faut accepter l'idée d'une « cascade d'états où l'information s'adapte à chaque condition de lecture et d'écriture ». On parle bien ici de web sémantique, d'ontologies, d'une *radicalisation* de « cette fragmentation des contenus en documents virtuels personnalisables, recalculés dynamiquement à partir d'un moteur et d'un ensemble d'ontologies » (Merzeau, 2009).

Le document numérique

Jean-Michel Salaün nous dit, très justement, que « tout objet peut devenir document à partir du moment où il permettra, en relation avec d'autres, d'interpréter l'époque de sa circulation » (Salaün, 2012, p. 47). Plus familièrement – mais de manière non moins scientifique – la métaphore de l'antilope nous explique qu'une antilope, prise dans son environnement naturel – soit la savane africaine – n'est en

rien un document mais s'apparente plutôt à un sujet. Seulement, une fois capturée, une fois désignée comme un *objet* d'étude, elle ne devient rien de moins qu'une *preuve* physique, un document. De là, nous-mêmes êtres humains, sommes soumis à cette métaphore existentielle dès lors que nous choisissons d'entrer dans l'espace numérique. Olivier Ertzscheid nous explique, avec sa célèbre assertion (*l'homme est un document comme les autres*) que :

> « L'homme est devenu un document comme les autres, disposant d'une identité dont il n'est plus 'propriétaire', dont il ne contrôle que peu la visibilité (ouverture des profils à l'indexation par les moteurs de recherche), et dont il sous-estime la finalité marchande » (Ertzscheid, 2009)

Mais qu'est-ce qu'un document ? L'Organisation Internationale de Normalisation définit le document comme « un ensemble formé par un support et une information, généralement enregistrée de façon permanente, et tel qu'il puisse être lu par l'homme et la machine ». De son côté, Salaün analyse plus critiquement la notion de document, et plus spécifiquement le fragmente en trois éléments essentiels à la définition du document :

- Le *vu* définit la forme, le contenant, le support. Il est l'objet que nous tenons entre nos mains et que nous sommes capables d'appeler objet, car nous sommes aptes à le reconnaître simplement en le voyant. Ainsi l'objet livre est le *vu*, c'est ce qui permet de lire, de *saisir*[19] le document, c'est sa réalité physique et matérielle, c'est le « rapport de notre corps et de nos sens à l'objet document » (Salaün, 2012, p. 49). On ne parle pas de sens, ici, mais bien de matière, de perception matérielle et physique ;
- Le *lu*, quant à lui, concerne le texte, le contenu. C'est ce qu'on lit, ce qu'on a appris à l'école et qui nous permet de déchiffrer les codes d'un texte et les symboles de l'écriture. La question du *lu* est intellectuelle, elle ne concerne plus seulement la reconnaissance mais la compréhension et l'interprétation – qui en découle logiquement – de ce qui est inscrit sur le support ;
- Le *su*, enfin, vient compléter le tableau en venant jouer le rôle de médiateur entre le *vu* et le *lu*. Le *su* est en vérité le rapport du document à notre humanité, à nos sociétés, et concerne la question de la transmission par le partage, l'échange, la diffusion. Intégré à notre culture, le document devient un objet *social*.

Le document est l'ensemble articulé de ces trois notions, sans lesquelles il ne serait rien de plus qu'un prototype, un « protodocument », une « trace permettant d'interpréter un évènement passé à travers un contrat de lecture » (Salaün, 2012, p. 60). Articulé au *vu*, au *lu* et au *su*, le document devient la « représentation d'un protodocument sur un support, pour une manipulation physique facile, un transport dans l'espace et une préservation dans le temps » (Salaün, 2012, p. 60). Et nous autres, documents parmi d'autres, subissons le même sort.

L'archive numérique

Comment, alors, se place l'archive dans la réticularité de ce *processus* documentaire ? il semble clair, désormais ces premières définitions posées, que l'on ne peut guère parler de *où* ou de *quand* placer l'archive, tant elle apparaît tout au

[19] Au sens premier du terme.

long de l'existence du document, de l'information, de la donnée. Autant dire qu'une fois qu'elle naît sur l'espace numérique, elle ne disparaît jamais, sous une forme ou une autre. En fait, « l'archive voit son rôle renforcé. À la fois consignation, publication et normalisation, elle vise aussi à anticiper » (Merzeau, 2009) mais également, au même titre que la trace (puisque l'archive est constituée de traces et se constitue, elle-même, en tant que trace) à identifier.

Le code du patrimoine définit encore, s'il faut le rappeler, les archives comme « l'ensemble des documents, y compris les données, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité ». Mais au-delà de cette définition très large, l'archive est aussi cette notion de preuve qui inclut les questions d'authenticité, d'intégrité et de traçabilité inhérentes à tout document ou donnée numérique.

Il semble donc bien complexe de définir à quel moment l'archive intervient dans le processus documentaire de la trace, de la donnée, du document numériques, tout autant qu'il est simple et pourtant juste de dire qu'elle intervient *partout, tout le temps*. Rappelons-nous : le document, en constante et permanente évolution, n'est jamais *validé*, au sens où lui est arraché le privilège de dire *je suis terminé, je ne changerai plus*. À chaque seconde, possiblement, il est un *autre* document, différent de la seconde d'avant, pas encore le même que la seconde d'après, et en même temps unique puisque si sa forme, son contenu, son support peuvent changer son essence reste la même.

Difficile, ici-même, de démêler strictement l'archive de ces multiples existences et évolutions sans entrer dans un débat autant philosophique que documentaire, technologique, sociologique. Ce qu'il faut retenir c'est qu'il y a deux niveaux de l'archive qui nous intéressent ici-même ; d'une part, la notion stricte définit par le code du patrimoine qui entend l'archive comme un document d'activité au sens *professionnel* du terme ; de l'autre, une notion moins tangible qui assimile l'archive numérique à cet ensemble de traces, d'empreintes et de métadonnées qui constituent le document tout en s'y substituant.

D'UN MONDE *LIBRE*

Qu'est-ce que le *Libre* ?

On observe depuis plus de vingt ans maintenant, l'expansion d'une tendance à la donnée ouverte, au libre-accès, au partage de tout pour tous. Nous l'avons vu, cette mode tenace est en vérité indissoluble de l'architecture même d'internet qui, par son organisation réticulaire, en tant rétablir le *réseautage* sans notion hiérarchique particulière ou définie. En fait, Antonio Casilli va plus loin et nous dit que « la seule raison pour laquelle le Web a prospéré et acquis sa forme actuelle est la quantité impressionnante de contenus en libre-accès – et en libre échange » (Casilli, 2010, p. 43). C'est qu'en plus d'être inhérent à son architecture, l'esprit *Open* serait l'essence, le carburant du monde numérique.

Open access, open archive, open data, open source on ne compte plus le nombre de termes affublés du suffixe *open* comme une arme brandie contre la marchandisation du web ou la fermeture des frontières. Pourtant, la réalité est plus complexe, et la notion d'ouverture n'entend pas toujours gratuité ou libre circulation, et va jusqu'à, parfois, rimer avec capitalisation ou anti-mondialisation.

Mais qu'est-ce que l'*open*, tout d'abord ? difficile de définir cette notion sans éclaircir tout d'abord quelques-unes de ses occurrences.

L'*open archive*, traduisible par *archive ouverte*, appartient, contrairement à ce que son nom pourrait laisser penser, strictement au domaine de la publication scientifique et technique. À l'origine en vérité, une initiative – *The Open Archives Initiative* (OAI) – qui « *develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content* »[20]. Comme il est noté sur la page d'accueil de leur site, l'OAI trouve ses racines dans le mouvement d'*open access* et développe notamment trois autres projets qui découlent de cette première initiative : *eScholarship*, *eLearning* et *eScience*.
Mais le succès de l'OAI a aussi donné naissance à d'autres initiatives et acceptions de la notion d'*open archive*. D'abord, dans le domaine de la publication scientifique, des initiatives telles que HAL-archives ouvertes ou Open Edition ont donné à l'expression *archive ouverte* un sens plus large : désormais, on appelle archive ouverte l'espace dans lequel sont déposées des données issues de la recherche et de l'enseignement, et dont on se refuse à instaurer des barrières (imposées notamment par les éditeurs scientifiques ou les universités), et ce grâce à des protocoles découlant du projet originel OAI.
On nomme, également, *open archives* les archives en ligne accessibles à tous déposées par certaines entreprises. La SNCF, par exemple, a nommé *Open Archives* l'un de ses projets intégrés à sa démarche de *Transparence Archives* et d'*Open Data*.

L'*open access* ou *libre accès* concerne, lui aussi, la publication scientifique et technique, ou plus exactement le domaine de l'édition de revues scientifiques et techniques. L'idée est de mettre en libre accès – d'où son nom – une partie ou la totalité des articles et des numéros d'une revue, c'est-à-dire que les revues sont gratuitement et librement accessibles. En se basant sur les protocoles de l'OAI, l'*open access* entend redéfinir les contours de l'édition scientifique et s'attaque directement à des questions économiques.

L'*open data* entend donner accès aux données produites par une entreprise à ses usagers. Sur le principe employé par la SNCF pour donner accès à ses données, donc, l'entreprise déposer en ligne ses données, structurées, et les laisse à disposition des internautes sans restriction économique, juridique ou politique. En France, le mouvement autour de la donnée ouverte est intégré dans la Déclaration des Droits de l'Homme et du Citoyen (droit d'accès aux informations publiques) mais le mouvement n'a eu droit à un souffle nouveau que depuis 2011, avec la parution d'un décret posant la gratuité du droit à la réutilisation des documents et données publiques.

L'*open source* fait référence, la majorité du temps, au logiciel libre. Pourtant, la différence entre ces deux notions est fondamentale, puisque le premier est une méthodologie de codage, un type de développement, quand le second est un type de logiciel qui s'oppose au logiciel propriétaire. Dans les deux cas, la notion de gratuité n'est pas inhérente à ces mouvements, loin de là.

Du logiciel libre découle une autre notion, la *culture du libre*, qui semble-il peut être étendue à ce mouvement *open* qui est indissociable de notre environnement d'aujourd'hui. À ce propos, Camille Paloque-Berges et Christophe Masutti rappellent qu'il faut « réfléchir au Libre en termes de régime technologique, qui se

[20] L'OAI développe et promeut des normes et standards d'interopérabilité qui visent à faciliter la diffusion efficace de contenus. Site de l'OAI - http://www.openarchives.org/

différencie du régime de la technique (dans son rapport raisonné à l'action) en tant qu'il considère l'action non pas en soi, mais en tant qu'environnement » (Paloque-Berges & Masutti, 2013), c'est-à-dire que le libre ne doit pas seulement s'entendre comme un mouvement, tel que nous l'avons fait jusqu'à présent, mais comme un espace, un monde social, économique et politique. En reprenant l'idée selon laquelle les mouvements de partage, de diffusion, d'ouverture des données sont « inscrits dans les codes mêmes d'Internet (des protocoles aux applications) » ils ajoutent que ces « principes d'ouverture de techniques pour mieux faire circuler l'information » sont au cœur de l'esprit informatique.

L'idée d'un monde *open* n'est rien de plus que cette « éthique du Libre » qui est à l'origine d'internet, mais qui se constitue également comme son tuteur, son squelette, son essence. Cette volonté, éminemment politique, d'amplifier les phénomènes du *libre* ne s'arrête évidemment pas à la question logicielle et constitue en fait la pierre angulaire du débat public actuel autour de la donnée.

L'*open data*

Plusieurs politiques d'ouverture des données sont à l'œuvre. On sait, par exemple, que certains lots de données seront déposés tels quels, *bruts*, sans traitement, sans tri, sans sélection, et qu'il sera offert à tout un chacun de faire son propre 'marché' pour trouver ce qui l'intéresse ; à l'opposé, on trouvera également des données traitées, classifiées, contextualisées et pérennisées qui auront fait l'objet d'un soin tout particulier. Afin de les distinguer, Tim Berners-Lee a noté les données ouvertes sur cinq niveaux[21] :

★	Available on the web (whatever format) *but with an open licence, to be Open Data*
★★	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
★★★	as (2) plus non-proprietary format (e.g. CSV instead of excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
★★★★★	All the above, plus: Link your data to other people's data to provide context

Tableau de classification des *Linked Open Data*[22] par Tim Berners-Lee

En plus d'offrir une grille d'appréciation des différentes politiques d'ouverture des données existantes, ce tableau permet également de lister les éléments principaux nécessaires à la définition de l'*Open Data*. Ainsi, on y lit qu'une donnée ouverte doit, par essence même, *être disponible sur le web grâce à une licence ouverte*, puis seulement ensuite être *structurée*, dans un format *non-propriétaire*, en standard *ouvert*, et enfin, pour obtenir ses cinq étoiles, être *liée aux données d'autres personnes afin d'être placée en contexte*.

Car la question de l'*open data* et l'esprit du Libre de manière plus générale sont portés, de bout en bout, par la notion de contextualisation. N'est-ce pas la base même du web sémantique ? la recherche de sens, le besoin de *faire sens*, ne sont-elles pas les principales problématiques des individus de la génération numérique ? Plus une donnée sera complète sur l'échelle offerte par Berners-Lee et plus,

[21] Site du W3C : https://www.w3.org/DesignIssues/LinkedData.html

[22] Traduisible par « données ouvertes *liées* ».

évidemment, elle sera compréhensible, exploitable, et pérenne. Pour Tim Berners-Lee « *it is the unexpected re-use of information which is the value added by the web* », et la donnée ouverte qui prétend à l'interconnexion est une information éminemment réutilisable (Berners-Lee, 2006).

Comme nous l'avons déjà évoqué, on trouve plusieurs formes de dépôt de données qui s'entrecroisent avec le tableau présenté ci-dessus. Si l'on prend l'exemple d'une plateforme comme Wikileaks, on s'aperçoit que malgré la volonté affichée de diffuser des masses de données, aucune d'entre elles n'est diffusée sans traitement. En fait, les données sont triées, sélectionnées, classées et indexées pour permettre aux utilisateurs de les manipuler avec le moindre effort. D'autres plateformes supportent une diffusion plus brute, comme *4chan* ou *Reddit* qui, sans avoir prétention à diffuser des masses de données, sont utilisés comme espace de *stockage* ou de *traitement*, où les internautes vont dépouiller l'arrivage de données et les traiter par eux-mêmes. Dans ce cas, le lot sera logiquement déposé dans son intégralité ; dans le premier cas, certaines données seront mises de côté *en attendant*. Mais nous y reviendrons.

Les lanceurs d'alerte eux-mêmes sont, pour la majorité d'entre eux, grands défenseurs du mouvement d'ouverture de données et d'internet de manière générale. En effet, ces mouvements promeuvent une certaine forme d'honnêteté puisque toutes les données sont susceptibles d'être analysées par n'importe qui. Si les données sont accessibles, il est plus délicat de détourner, de cacher, de tromper. En fait, Antonio Casilli note qu'entre en compte ici une véritable « stratégie collective » visant à « répondre à la fermeture progressive des lieux de confrontation démocratique et à l'opacité des modalités de participation à l'*espace public* » (Casilli, 2010, p. 73), et qu'il y a de fait une forme d'activisme, de militantisme, de réaction dans la défense du monde numérique Libre aujourd'hui.

Liberté de savoir, liberté d'expression

Pourtant, Edwy Plenel note que « la Déclaration des droits de l'homme de 1789 énonçait déjà, en son article 15, que 'la société a le droit de demander compte à tout agent public de son administration'. » (Plenel, 2013, p. 73). La France a, par rapport à d'autres pays occidentaux comme les États-Unis et leur *FOAI*[23], un retard considérable quant à sa prise en compte de la liberté d'expression et du droit de savoir. Aux États-Unis donc, on « considère que le droit d'accès des citoyens aux informations doit être la règle et le refus l'exception » et qu'il « revient à l'administration de justifier ses refus et non pas aux citoyens de justifier leurs demandes » (ibid., p. 74). Comme il le note, les pays scandinaves ne sont pas en reste sur ses questions de droit de la presse, droit de savoir, droit à l'information, et de fait sur les textes *régissant* l'espace numérique, quand la France est, là aussi, toujours à la traîne.

Les États-Unis, pour les prendre en exemple, ajoutent à la protection de la liberté de savoir et d'expression une « obligation positive imposée aux administrations à l'heure des archives électroniques : désormais, nombre d'informations doivent être spontanément mises en ligne, même sans aucune demande venue du public » (Plenel, 2013, p. 74) dès 1996. C'est ce qui, près de

[23] *Freedom of International Act*, adopté en 1966.

vingt ans après, donnera naissance à l'initiative Open Data de la SNCF ou à l'ouverture des données publiques initiée par le gouvernement français, notamment.

Journalistes et lanceurs d'alerte

Les acteurs principaux au sein de ces mouvements liés à la protection de la liberté de savoir et d'expression sont bien évidemment les journalistes. Médiateurs de la première heure, ils voient leur rôle réajusté au sein de l'espace numérique et de fait, plus largement, au sein de la société moderne. En effet, auparavant « pour s'exprimer dans l'espace public, pour défendre son point de vue, pour faire connaître sa cause, le citoyen devait passer par un journaliste ». Nous l'avons vu, l'expansion du numérique, la démocratisation d'internet dans les foyers et les mouvements politiques et sociaux du web ont conduit l'internaute à prendre en main lui-même son traitement de l'information. Aujourd'hui, Edwy Plenel note que « potentiellement, n'importe quel citoyen peut désormais s'exprimer directement, briser l'indifférence médiatique, dévoiler l'ignorance journalistique » (Plenel, 2013, p. 110) et la presse alors se doit de revoir sa position médiatrice. Les premiers à diffuser, et surtout à manipuler les données numériques ne sont pas les gouvernements ou les journalistes, mais les internautes. Casilli le note, que ce soit « sous forme de blogs, de sites Web ou de simples chaînes de mails, cela fait désormais 15 ans que les réseaux jouent un rôle clé dans le soutien des causes les plus disparates » (Casilli, 2010, p. 90), et ce justement en usant des données à leur disposition pour dénoncer ce qu'ils considèrent comme illégal, illégitime, immoral. Le pouvoir disponible, à savoir le sentiment que le « comportement d'un seul individu puisse changer la donne ou 'compter pour quelque chose' dans le contexte des communautés en ligne » (Casilli, 2010, p. 54) arrive en troisième position des sources de motivation à la participation de la vie communautaire du web. De plus, lorsqu'on regarde de plus près le processus de diffusion de l'information, Plenel note que l'énonciateur et le récepteur sont essentiels, et que c'est le niveau d'interactivité de l'information qui juge « la qualité de la démocratie » (Plenel, 2013, p. 126).

Rappelons un instant, à ce propos, les missions du journaliste :

> « Si l'on s'en tient à l'artisanat du métier, produire des vérités de fait, c'est d'abord respecter toutes ces règles et opérations qui font les informations rigoureuses, honnêtes et loyales : vérifier, sourcer, recouper, confronter, contextualiser » (Plenel, 2013, p. 119)

La montée d'un mouvement sociétal hyperactif, de la nécessité du *buzz*, de l'instantané, d'une exigence de l'information laconique développe, tant en parallèle qu'à contre-courant, un intérêt croissant des internautes et des citoyens de manière générale vers un journalisme d'investigation qui s'attacherait plus à collecter et traiter les données (et donc l'information) qu'à la diffuser. Sans pour autant s'inscrire dans la lignée des hackers ou des lanceurs d'alerte, les journalistes (*d'investigation*, entendons-nous bien) offrent une médiation tournée aujourd'hui vers la constitution de corpus de documents et données offerts par la fuite d'information.

La notion de transparence qui s'intègre à l'esprit du web et à l'ouverture de la donnée permet ainsi un double traitement. D'une part, les journalistes s'attachent à un véritable travail de détective et, misant sur une éthique journalistique qu'on ose dire entérinée par l'ère du *fast*, travaillent la fuite et la donnée ouverte pour re-contextualiser et offrir au public une histoire aux ramifications plus gigantesques

que jamais ; d'autre part, les *cyber-citoyens* ou plutôt les internautes et amateurs du politique œuvrent, parfois en parallèle des journalistes, parfois de concert, à dépouiller les lots d'informations, de données et de documents mis à disposition. De là, deux assertions : la première étant que jamais dans l'histoire de l'humanité a-t-on pu voir autant d'individus attachés à la gestion de l'information – et ce sans qu'ils soient forcément intéressés par le contenu même de l'information – et la seconde, que la notion de secret vit sans doute l'un de ses plus grands moments de décadence.

La société du secret

La notion de *secret* agite l'histoire depuis, sans doute, les prémices de l'humanité. Du latin *secernere* qui signifie *sécréter*, il intègre la dimension d'objet *à cacher, honteux,* et de fait la *mise au rebut*, la *frontière*, la *séparation*. Le secret, « suspect, reste sous le tapis des origines » (Renucci, 2013, paragr. 11), c'est un objet dérangeant qu'il semble nécessaire de taire, et pourtant : « un secret n'existe que s'il est connu de quelqu'un » (*ibid*.). Parce qu'il inclut la *séparation*, le secret inclut la *frontière* :

> « Et qui dit frontières dit conflit. Quelqu'un a une raison de vouloir s'emparer du secret. Ce peut être l'adversaire, le concurrent, le représentant de la loi… Quelqu'un a une raison de le défendre. L'histoire du secret est une histoire d'épée et de bouclier : elle suppose un antagonisme envers un adversaire précis ou un péril plus général qui se nommerait la Loi, l'Opinion, la Société… » (Huyghe, 2013b, paragr. 28)

Le secret a donc besoin, pour exister, d'un défenseur (si personne ne cherche à conserver le secret, c'est qu'il n'y a rien à cacher, donc pas de secret) et d'un adversaire (si personne ne cherche à s'emparer du secret, c'est qu'il n'existe pas aux yeux des autres, et que donc il n'y a pas, là aussi, de raison de le cacher : on ne cache que ce que d'autres voudraient chercher à dévoiler). De fait, « le secret lui-même, le fait qu'il y a un secret ou le « coffre » qui le recèle, matériel ou immatériel mais toujours médiologique, doit être exhibé pour être efficace » (Soriano, 2013, paragr. 19). S'il n'est pas exhibé, si l'on ne sait pas qu'il y a un secret, ce dernier n'a pas de raison d'être.

C'est donc un objet profondément paradoxal et amusant. Pour qu'il soit tu, il faut qu'on cherche à le faire parler, pour qu'il existe et ne soit pas divulgué, il faut qu'on sache qu'il existe (et que son contenu ne doit pas être divulgué).Il y a une dimension très philosophique de désir, d'attente, de fantasme derrière la notion de secret, et ce que l'on parle du secret de l'enfant cachant sa bêtise ou bien des documents attestant d'essais nucléaires et autres affaires gouvernementales.

Dès lors qu'il est divulgué, le secret n'en est plus un : il devient *secret dévoilé*, et donc n'existe plus que pour prouver qu'un autre secret existe. C'est l'un des paradoxes du *leak* que nous observerons plus loin : le secret divulgué s'apparente à une fuite, mais la fuite s'entend comme faisant partie d'un ensemble plus large (une fuite de quelques centilitres d'eau de canalisation suppose qu'il y a des litres et des litres d'eau qui passent dans la canalisation, qu'il y a un espace de stockage de milliers de m^3 d'eau, qui viennent eux-mêmes de nappes phréatiques, etc.). De fait, Tout ce qui est dit suggère aussitôt que quelque chose d'autre ne l'est pas, le secret hante tout discours » (Soriano, 2013, paragr. 9). La communication du secret se révèle alors problématique : « l'une ne va pas sans l'autre, comme l'ombre et la lumière, comme le fond ou l'écran nécessairement opaque sur lequel se détache le message » (*ibid*.).

Et qu'importe que le contenu du secret, l'information qui porte sa marque soit « inexacte, fantaisiste ou mensongère : c'est le sceau qui fait le secret et rend l'information agissante » (Soriano, 2013, p. 24). Ainsi le contenu même n'a aucune importance, la valeur du secret se calcule par les moyens mis à disposition pour le préserver. Paul Soriano nous offre d'ailleurs l'ébauche d'une loi économique « de portée générale » savoureuse : « *un secret vaut très exactement le coût des moyens requis pour le protéger ; ou encore, mais c'est la même chose, le coût des moyens requis pour le percer* » (Soriano, 2013, paragr. 22).

François-Bernard Huyghe conteste ce qui pourrait nous paraître comme la plus évidente des raisons d'exister du secret. Pour lui, « le secret ne sert pas qu'à dissimuler actions ou intentions coupables. Il protège des richesses volatiles, affranchies de leur support matériel : des données. L'économie dite de l'accès implique un contrôle de l'accès » (Huyghe, 2013b, paragr. 11). Le secret comme protecteur de ces données peut être vu, notamment par Françoise Gaillard, comme « le principe structurant du politique et la clé de voûte qui maintient l'édifice du pouvoir » (Gaillard, 2013, paragr. 16).

Le secret, c'est le pouvoir. Du moins, de Platon à Machiavel, l'un brandissant le bien commun au détriment de l'intérêt individuel[24], l'autre arguant que le pouvoir du *prince* nécessite de tromper tous ceux qui pourraient le menacer, « la pratique du secret dans les démocraties est courante et serait justifiée d'abord par l'intérêt général » (Conesa, 2013, paragr. 22). Pourtant, Marc Conesa, à juste titre, nous rappelle qu'avec Rousseau « les individus autonomes, libres et égaux en droit, participant directement aux affaires de l'État, forment le peuple » et que de fait le secret serait « inutile, voire interdit dans une démocratie directe » (Conesa, 2013, paragr. 22).

Quoi qu'il en soit, le secret existe bel et bien, et sans doute l'ère qui le verra s'éteindre n'est-elle pas encore à portée de main. Pour preuve de son enracinement sociétal, on le retrouve défendu par l'administration elle-même où « sa possibilité est dissimulée dans nos valeurs les plus anciennes » (Renucci, 2013, paragr. 10).

En France, contrairement à d'autres démocraties qui contrôlent la procédure et le bien-fondé ou non du secret, les « règles dérogatoires sont très nombreuses et peu – voire pas – contrôlées ». En fait, Conesa note que c'est par le « non-droit » que « le droit à l'information du citoyen est protégé » (Conesa, 2013, paragr. 12) en France, puisque la juridiction s'échine à garder la maîtrise du secret du côté du pouvoir exécutif, et dans le même temps totalement opaque. On peut noter, paradoxalement, que la France se présente comme fervente défenseuse de la transparence et de la liberté de savoir de ses citoyens tout en tenant d'une main de fer les secrets qu'elle estime être les siens. Ainsi la presse resterait « le seul réel contre-pouvoir à l'utilisation abusive du secret par le pouvoir exécutif ». Cette dernière assertion est, à notre sens, à tempérer : en effet, les lois relativement récentes quant au statut du lanceur d'alerte accordent des droits, certes minces et à la liberté d'action réduite, à ceux qui prétendent divulguer et dénoncer.

Toutefois, et nous le verrons un peu plus loin, ces droits accordés aux lanceurs d'alerte et le fait même que la législation reconnaisse aujourd'hui leurs statuts ne les rendent pas obligatoirement plus légitimes ; de plus, et ce sans tomber dans une quelconque méfiance ou paranoïa, l'histoire nous a appris que « s'il y a des choses secrètes, comme le donnent à voir, entendre et lire tous les pirateurs d'informations

[24] « Pour faire droit en gros, il est permis de faire tort dans le détail », Platon.

classifiées ou confidentielles, c'est qu'il y a (encore) un pouvoir du politique » (Gaillard, 2013, paragr. 27) et que s'il y a pouvoir du politique, il y a des secrets encore bien mieux enterrés.

Mais oserait-on prétendre à l'inviolabilité de certains secrets ? pas le moins du monde. Le secret est aujourd'hui confronté à « l'impossibilité technologique de renfermer » qui justifie, comme un effet balance, le « droit politique de savoir » (Huyghe, 2013a). Cette impossibilité s'explique sur deux niveaux distincts.
D'une part, l'idée selon laquelle « l'information protège de l'information, mais pas éternellement » est un fait avéré. Quel que soit le dispositif utilisé pour protéger l'information, il finira inéluctablement par baisser les armes « soit face à une technologie nouvelle (un logiciel espion qui a découvert une vulnérabilité), soit face au facteur humain (un citoyen ne supporte plus la contradiction entre le discours sur la démocratie et ce qu'il voit et fait) » (ibid.).
D'autre part, une fois que la fuite est là, il « suffit d'une connexion Internet pour que la planète ait accès à des documents originaux prouvant l'espionnage, la bavure ou le déni de droit » (ibid.).

Impossible donc de protéger éternellement le secret ; et dans le même temps, impossible de l'empêcher d'exister. L'espace numérique offre un panel plus large que jamais aux individus désirant protéger leurs secrets et espionner ceux des autres, et les capacités techniques se superposent aux compétences technologiques. Le secret numérique est l'apanage de tout internaute qui se respecte, et le fait de protéger les siens c'est aussi, aujourd'hui, « garder l'accès à ses bases de données, s'assurer contre le risque de l'intoxication ou de l'information désorganisatrice, tels les virus » (Huyghe, 2013b, paragr. 3).

Ainsi le secret n'a-t-il jamais véritablement concerné le contenu, et ne se repose plus sur la dissimulation d'une information : la technologie numérique impose de « contrôler des passages et des flux plutôt que des contenus » (Huyghe, 2013b, paragr. 3) au point où certaines entreprises en viennent à dire à leurs informaticiens et gestionnaires de l'information : *qu'importe ce que l'on protège, tant que c'est bien protégé*.

LA FUITE DOCUMENTAIRE

> *"*[…] *a leak is taken to be a targeted disclosure by a government insider (employee, former employee, contractor) to a member of the media of confidential information the divulgence of which is generally proscribed by law, policy, or convention outside of any formal process with an expectation of anonymity"* (Pozen, 2013, p. 521)

Lanceurs d'alerte

La question du *leak*, peu définie, amène en premier lieu à penser la place de l'individu à l'origine de la fuite, c'est-à-dire le *lanceur d'alerte*. Sa définition est aujourd'hui, en France, relativement bien documentée dans les textes de loi tout en restant suffisamment vague pour être régulièrement remise en question.

La formule *lanceur d'alerte* apparaît en janvier 1996 lors de travaux scientifiques intitulés « Risques collectifs et situations de crise » (Chateauraynaud, 2013). L'idée alors est de mettre une dénomination sur une notion bien connue afin de désigner « les personnages qui, en s'évertuant à faire reconnaître un danger ou

un risque émergent, rencontraient toutes sortes de difficultés cognitives et politiques » (Chateauraynaud, 2013). Souvent associé au *dénonciateur*, terme à la connotation plus négative, il se confronte également à celui de *whistleblower* (qui donne un coup de sifflet) préféré par les anglo-saxons, qui lui n'alerte pas, ne prévient pas, mais condamne, prend sur le fait accompli.

On peut également faire un parallèle étroit avec « l'acte de vigilance active » que constitue l'action d'alerter, en se rapprochant alors du terme *vigilanty* en anglais, traduit par *justicier* en français et qui offre une facette presque punitive à la manière dont on va user de l'alerte. On s'aperçoit alors que le porteur de l'alerte peut revêtir de nombreuses casquettes, avec des intentions et des perceptions de ses actes parfois très différentes. C'est qu'entre le « prophète de malheur » et le « capteur », il y a toute une palette de graduation de l'agissement et de la représentation à faire.

À ce sujet, Chateauraynaud nous rappelle justement que :

> « la place du lanceur d'alerte étant avant tout définie par le processus critique qui relie des actes de vigilance et des formes de prise de parole publique, elle ne peut être réifiée sous la forme d'un statut, et encore moins d'une identité, individuelle ou collective, dont ne pourraient bénéficier que quelques figures héroïsées ou préalablement organisées » (Chateauraynaud, 2013)

Ainsi, la loi du 9 décembre 2016 relative à la transparence, à la lutte contre la corruption et à la modernisation de la vie économique (dite Sapin II) définit le lanceur d'alerte selon ses intentions, d'une part, et en fonction de l'acte et de l'entité dénoncés, d'autre part :

> « Un lanceur d'alerte est une personne physique qui révèle ou signale, de manière désintéressée et de bonne foi, un crime ou un délit, une violation grave et manifeste d'un engagement international régulièrement ratifié ou approuvé par la France, d'un acte unilatéral d'une organisation internationale pris sur le fondement d'un tel engagement, de la loi ou du règlement, ou une menace ou un préjudice graves pour l'intérêt général, dont elle a eu personnellement connaissance. »
> *LOI n° 2016-1691 du 9 décembre 2016 relative à la transparence, à la lutte contre la corruption et à la modernisation de la vie économique.*

En complément, on peut citer l'étude adoptée en février 2016 par le Conseil d'État qui définit le lanceur d'alerte en opposition avec ces fameuses intentions que l'on juge personnelles et donc condamnables :

> « Le lanceur d'alerte n'est ni un dissident, qui s'opposerait radicalement à une collectivité, ni un partisan de la désobéissance civile, qui revendiquerait une « contre-légitimité ». Il n'est pas non plus un délateur ou un sycophante, qui agirait dans son intérêt personnel, ni un calomniateur, qui chercherait à nuire ou à jeter l'opprobre. »
> *Le droit d'alerte : signaler, traiter, protéger*, Étude adoptée le 25 février 2016 par l'assemblée générale plénière du Conseil d'État

Une « exigence de séparation de l'alerte et de la dénonciation » (Chateauraynaud, 2013) s'impose dans les textes, et la Commission nationale de la déontologie et des alertes en matière de santé publique et d'environnement (CNDASE), appuyée par la loi de 2013, d'ajouter que :

> « Toute personne physique ou morale a le droit de rendre publique ou de diffuser de bonne foi une information concernant un fait, une donnée ou une action, dès lors que la méconnaissance de ce fait, de cette donnée ou de cette action lui paraît faire peser un risque grave sur la santé publique ou sur l'environnement.
> L'information qu'elle rend publique ou diffuse doit s'abstenir de toute imputation diffamatoire ou injurieuse. »
> *LOI n° 2013-316 du 16 avril 2013 relative à l'indépendance de l'expertise en matière de santé et d'environnement et à la protection des lanceurs d'alerte*

Ces définitions mettent en lumière plusieurs visions sur lesquelles il convient de s'arrêter un instant. Tout d'abord, on s'accorde à dire que le lanceur d'alerte agit « de bonne foi », dans « l'intérêt général », de « manière désintéressée » et si possible « librement », ce qui revient à dire que la personne lançant l'alerte ne doit pas être sous une quelconque contrainte de diffusion de l'information. Personne physique ou morale, le lanceur d'alerte « signale », « diffuse », « révèle » ou « rend publique » une information dont il a pris directement connaissance et dont il n'est pas auteur. Jusque-là, les trois définitions proposées ci-dessus concordent et s'accordent relativement bien. L'interprétation se corse néanmoins lorsqu'il s'agit de positionner un curseur de *gravité* définissant le fait, la donnée ou l'action présentés par le lanceur d'alerte. Les trois définitions ne proposent aucune grille d'évaluation ou d'appréhension, se contentant d'évoquer des menaces, des manquements, des risques toujours considérés comme « graves » au regard de la loi, des règlements internationaux ou encore d'un point de vue éthique. C'est d'ailleurs la question éthique qui pose spécifiquement problème, envisagée sous l'angle de « l'intérêt général », des « intérêts publics ou privés », ou clairement ajustée en fonction de l'interprétation du lanceur d'alerte si l'information « lui paraît faire peser un risque grave sur la santé publique ou sur l'environnement » où à ce propos, Chateauraynaud précise très justement : « la charge de la preuve de l'existence d'un risque et de sa gravité ne porte pas sur le lanceur, puisque, pour être protégé, il suffit qu'il pense, de bonne foi, que la méconnaissance des faits qu'il révèle est constitutive du risque que se produise un dommage pour la santé ou l'environnement » (Chateauraynaud, 2013).

Mais être convaincu que « la méconnaissance des faits […] est constitutive du risque que se produise un dommage » ne suffit souvent pas au lancement d'une alerte, notamment parce que les risques encourus peuvent sembler dépasser le fameux risque engendré par la méconnaissance.

Pour prendre l'exemple parlant du lanceur d'alerte usant de son statut d'employé Pauline Abadie s'interroge sur « comment minimiser le conflit entre loyalisme et loyauté pour faire place à cette forme particulière de l'employé-citoyen du 21ème siècle qu'est le lanceur d'alerte » (Abadie, 2016, paragr. 46). Cette dualité très particulière se heurte notamment à ce qu'en France, dans l'espace salarial, le lanceur d'alerte jouisse d'un statut intimement lié à la liberté d'expression, espace où justement cette liberté d'expression est régulièrement remise en cause par le secret professionnel, les clauses de confidentialité et l'idée même d'être, en tant qu'employé, défenseur des valeurs de son entreprise ; espace où « l'idée de responsabilités professionnelles se pose bien souvent comme obstacle à la prise de parole salariale » (Abadie, 2016, paragr. 46). En vérité, la liberté d'expression « constitue la première protection offerte par le droit à la personne du lanceur d'alerte » (Abadie, 2016, paragr. 26) et s'appuie sur le code du travail pour définir ses limites :

> « En effet, l'article L.1121-1 du code du travail énonce, pour le principe, que 'Nul ne peut porter aux droits des personnes et des libertés individuelles et collectives des restrictions' » (Abadie, 2016, paragr. 37)

Tout en ajoutant :

> « et pour l'exception, 'qui ne seraient pas justifiées par la nature de la tâche à accomplir ni proportionnée au but'. Ainsi, en vertu de l'exception de l'article L. 1121-1, l'exercice d'une liberté peut se voir restreinte si 'la nature de la tâche à accomplir' l'impose. Dès lors, en fonction de la tâche et du poste occupé, eux-mêmes inhérents au contrat de travail, certains salariés devront faire preuve de plus de retenue et de discrétion que d'autres. Par conséquent, en creux de l'exception de l'article L.1121-1 se loge la bonne foi du salarié, ce qui revient à réintroduire l'exigence de loyauté salariale, mais cette fois-ci, comme élément du contrôle de l'exercice d'une liberté fondamentale. » (Abadie, 2016, paragr. 37)

Ainsi donc, la « bonne foi » se définit ici en partie par le contrat de travail, plus spécifiquement par la nature des tâches, des missions et des postes confiés au salarié. L'idée même que l'employeur puisse retirer « un bénéfice direct du signalement d'une défectuosité des machines qui menacerait la vie et la santé de son personnel » (Abadie, 2016, paragr. 45) là où sa responsabilité est en jeu peut être remise en cause par celle de « l'abus de liberté d'expression » (Abadie, 2016, paragr. 46). Toutefois, si l'obligation de loyauté à son entreprise/employeur/hiérarchique n'est pas légalement et directement reconnue comme une restriction à la liberté d'un lancement d'alerte, « reste que les salariés situés en haut de l'organigramme doivent faire preuve d'un engagement loyal plus complet, et que, ce sont aussi ceux-là qui sont les plus à mêmes de signaler les pratiques illégales, irrégulières ou immorales », et que sans restreindre, donc, « le droit français peine encore à protéger ceux les plus dans un conflit de loyautés » (Abadie, 2016, paragr. 46).

Soumis à de nombreuses interrogations, de multiples conflits et un risque élevé de poursuites et de dommages professionnels et personnels, le lanceur d'alerte se révèle être, en prime, un véritable caméléon, au sens où il revêt de nombreuses casquettes et de multiples facettes qui peuvent le rendre, parfois, difficilement reconnaissable. Jean-Philippe Foegle souligne que ce qui détermine le choix entre la dénomination de « fuiteur »[25] ou de « lanceur d'alerte » tient au « caractère irrégulier et illégal de ces divulgations » (2016, paragr. 14). Ainsi, il note que pour les anglo-saxons le terme *fuiteur* ou *leaker* serait plus adapté à la situation des « agents des services de renseignement, de la diplomatie et des armées qui révèlent publiquement des faits contraires à l'intérêt général » quand le *lanceur d'alerte* serait envisagé comme l'individu avertissant en tant que citoyen et non pas en tant qu'employé, professionnel, ou agent. Dans le même ordre d'esprit, le *leaker* aurait comme caractéristique d'avoir « créé une fuite » et ce faisant, serait dans l'incapacité de la réguler – ou tout du moins ne voudrait stopper le flot des fuites d'information qu'il aurait entamé ; quand le *whistleblower* ne « viserait que de manière ponctuelle à mettre à jour des faits contraires à l'intérêt et au bien-être des sociétés pour sommer les pouvoirs publics d'y mettre fin » (Foegle, 2016, paragr. 17). Il y aurait donc un *bon* et un *mauvais* lanceur d'alerte, un citoyen modèle et un

[25] L'auteur entend le terme « fuiteur » au sens de « *leaker* », soit celui qui *fait fuiter*, et le distingue du « lanceur d'alerte » par ce que le *leaker* serait plus un « désobéissant » qu'un citoyen modèle. L'article tend justement à recadrer ces dénominations (Foegle, 2016).

détracteur des intérêts publics et des sociétés, un bon samaritain et un oiseau de mauvais augure.

Dans le même temps, on considère (particulièrement aux États-Unis, en partie dans l'Union Européenne) que la fuite, notamment dans le cadre d'une fuite « d'information gouvernementale » n'est plus un droit mais un devoir. Ainsi aux États-Unis la divulgation « d'information classifiée peut apparaître comme une obligation constitutionnelle dès lors qu'un droit consacré par la Constitution américaine apparaît violé » (Foegle, 2016, paragr. 22), ce qui pourrait légitimer les *lanceurs d'alerte* comme rien de moins que des défenseurs de la « légitimité démocratique ».

C'est en ce sens que Foegle introduit la notion de « professionnels de la transparence » en distinction du *lanceur d'alerte*, que nous envisagerons comme un niveau de plus dans la définition du lanceur d'alerte plutôt que comme une catégorie à part. Il considère en effet que les lanceurs d'alerte sont « des acteurs de la transparence […] par accident », estimant que la divulgation d'informations dans l'objectif d'alerter ou d'obtenir une punition ne peut constituer une bataille contre l'opacité ou l'obscurantisme de manière générale, quand ces fameux « professionnels de la transparence » n'auraient comme seul et unique objectif la guerre à « l'opacité gouvernementale ». Foegle distingue d'ailleurs ces derniers en une catégorie « soft » et une « hard », arguant que les premiers, majoritairement journalistes et médias, n'érigent pas « le combat contre le secret gouvernemental au rang d'impératif catégorique » quand les seconds seraient « marqué[s] par une idéologie de la transparence » et se « confronte[raient] de manière plus radicale au secret gouvernemental en visant à mettre à bas celui-ci » (Foegle, 2016, paragr. 27-28).

Alerte, fuite, *leak*

La fuite

La fuite, action de *fuir*, est définie par le CNRTL comme la « mise à jour, divulgation de documents qui auraient dû rester secrets », illustrée par l'exemple parlant – puisqu'il met en lumière la difficulté de réguler une fuite une fois qu'elle existe :

> On s'aperçoit au ministère de la Guerre que, malgré la condamnation de Dreyfus et son internement à l'île du Diable, les fuites [it. ds le texte] continuent. Clemenceau, Iniquité, 1899, p. 113

L'étymologie ajuste cette définition en précisant qu'en parlant de fuite, on parle de la « disparition de documents destinés à demeurer secrets ». Ainsi la fuite d'un élément ne se cantonne pas à sa mise en lumière, à sa communication au-delà du secret dans lequel il est sensé évoluer, mais se définit également au moment de disparition, de son absence. Cette dualité d'absence et de présence, que nous expérimenterons au cours de ce mémoire, fait partie intégrante de l'existence des documents et des données.

Une fuite documentaire, c'est donc l'action qui ôte le statut de *secret*, de *privé* ou de *confidentiel* d'un document ou d'une donnée, en le communicant, en le mettant à jour, en le transmettant à un tiers ou à un public, quel qu'il soit ; ou bien plus étrangement encore en le faisant disparaître, en le déplaçant du lieu ou de l'espace qui garantit le secret de son contenu ou de sa forme.

Car la fuite ne concerne pas uniquement le contenu de l'élément, mais également sa forme, son fond, son statut, la communication autour de son existence seule. Nous l'avons vu, on peut parler de fuite lorsque l'existence d'un document nous est communiquée, et que cette existence même était soumise au secret. Le contenu du document nous sera encore inconnu, disposera toujours du cachet du secret, quand la réalité et la présence de ce document sera divulguée : ce qui, pour certains documents, ouvre déjà une brèche conséquente et suffisante pour faire fuiter d'autres documents. On pourra opposer alors l'idée selon laquelle un secret qui n'a pas un peu fuité n'est pas un secret, puisque personne ne connaît son existence ; ce à quoi il faudra sans doute répondre : le secret et la fuite ont ça de comparable qu'ils sont finalement les deux faces d'une même pièce.

L'alerte

Si secret et fuite peuvent être similaires, ou du moins emprunter les mêmes processus, l'alerte choisit un chemin légèrement différent. En France, la loi Sapin II exclut, très clairement, certaines fuites de la typologie de l'alerte, occultant ainsi un type de dénonciation qui pourrait porter atteinte à des secrets dont l'inviolabilité est protégée par les textes :

> « Les faits, informations ou documents, quel que soit leur forme ou leur support, couverts par le secret de la défense nationale, le secret médical ou le secret des relations entre un avocat et son client sont exclus du régime de l'alerte défini par le présent chapitre. »
> *Article 6 de la LOI n° 2016-1691 du 9 décembre 2016 relative à la transparence, à la lutte contre la corruption et à la modernisation de la vie économique*

Ce que laisse entendre ce texte, en dehors de la discrimination faite à l'encontre de certains documents quant à leur reconnaissance en tant que constitutif de l'alerte, c'est que cette dernière est *légale*, du moins reconnue par la loi et de fait, légitime. Plébiscitée aux États-Unis, elle offre une liberté d'action à la dénonciation de faits illégitimes ou illégaux concernant l'administration publique ou privée et les affaires touchant à la vie publique de manière générale. Elle est une arme contre la corruption et le pouvoir absolu, se love dans les concepts de la liberté d'expression et l'action citoyenne, épouse la légitimité du pouvoir étatique et ce, jusqu'à tant qu'elle ne franchit pas les barrières qui lui ont été posés, aussi floues puissent-être ces dernières.

Au-delà, l'alerte n'est plus considérée comme légitime, elle est d'ailleurs fortement condamnée, discriminée, vilipendée, et fait l'objet d'une véritable omerta par les politiques en place qui la jugent dangereuse pour la stabilité du pouvoir. À ce stage, le document ou la donnée n'est plus considéré comme *alertant* mais comme *fuitant,* et ce avec toute la trame péjorative que ce terme entraîne.

Le nom du scandale

Le *leak*

Leak est le terme utilisé en anglais pour désigner la *fuite*, le *lieu de la fuite* ou, en tant que verbe, l'action de *fuiter*. C'est par extension qu'on l'utilise pour signifier le scandale d'une fuite documentaire, ou le document ou la donnée qui a fuité.

Ainsi, apposé à la suite d'un nom propre, généralement un pays ou un nom de famille, il est aujourd'hui utilisé par la presse pour désigner une affaire, un scandale, mis en lumière par une fuite d'information – *MacronLeaks*, *LuxembourgLeaks*, *FootballLeaks*. On l'utilise également, dans la même logique, pour désigner le document ou la donnée ayant fuité qui alimentera l'affaire, le scandale en question.

Terme anglophone donc, utilisé pour désigner le document ou la donnée n'ayant pas vocation à être divulgué, qui aurait été communiqué dans l'objectif de mettre en lumière une action, une information déclarée secrète ou simplement non communiquée par ses auteurs et/ou producteurs, il est parfois remplacé par *paper*[26], ou se voit encore abandonné au privilège du terme *gate*.

Le *gate*

Gate en anglais signifie *portail, barrière, vanne*, et si nous pourrions trouver un lien entre cette définition et l'utilisation qui en est faite dans la dénomination du scandale, ce n'est pas là qu'il faut chercher l'explication. En vérité, l'utilisation du suffixe *gate* tient au scandale du *Watergate* en 1972, affaire d'espionnage politique, sous forme d'écoutes téléphoniques à l'hôtel du Watergate, qui conduisit à la démission du président Nixon deux ans plus tard, et qui fut à ce point spectaculaire qu'elle offrit ses deux dernières syllabes pour désigner des affaires politiques et financières.

On note toutefois que l'usage s'étend aujourd'hui aux scandales les plus disparates et variés, comme le *Nipple Gate*, affaire impliquant la révélation d'un téton de Janet Jackson à la télé, le *Crash Gate* qui envoya volontairement un pilote automobile se crasher contre un mur pour favoriser son équipier, le *Antenna Gate* qui faillit conduire les utilisateurs d'iPhone à poursuivre la marque pour un défaut de réception de réseau, ou encore le *Pasta Gate* lorsque l'OQLF[27] exigea qu'un restaurant italien abandonne sa pratique d'écrire sur ses menus *pasta* au lieu de *pâtes*.

De fait, on utilise le *gate* pour « rendre compte d'une polémique qu'ils [les journalistes] espèrent ainsi voir accéder au rang de scandale majeur »[28] et ainsi désigner tout scandale médiatisé dont le nom principal retenu pour le désigner se prête à l'exercice de l'apposition d'un suffixe : *Penelope Gate*, *Coin Gate*, *Foreclosure Gate*.

Le *paper*

Paper signifie tout simplement *papier* et est employé au pluriel pour désigner certaines fuites documentaires conséquentes. Un peu à la manière du *gate*, le *paper* tire son nom d'une première affaire, à laquelle cette fois le suffixe était déjà apposé : les *Pentagon Papers* (littéralement, les *papiers du Pentagone*). La diffusion de ce document classé secret-défense (c'est en réalité un unique rapport contenant pas moins de 7 000 pages et 47 volumes) en 1971 secoue la classe politique et militaire américaine et offrira son nom à la diffusion d'une autre quantité astronomique de documents, en rapport avec l'évasion fiscale cette fois : les *Panama Papers*.

---

[26] Toujours employé au pluriel dans ce sens.

[27] Office Québécois de la Langue Française.

[28] Anglais du journalisme : comprendre et traduire, pp. 54-55

APPRÉHENDER LE LEAK EN TANT QU'ARCHIVE

Quels sont les types de documents ayant vocation à être des *leaks* ? difficile de se représenter le *leak* et son fonctionnement sans s'attarder un instant sur ce qu'il est véritablement. Il n'existe pas de définition à proprement parler du *leak* en tant qu'état du document ou nom du scandale, qui est apparenté à la fuite sans être véritablement questionné. Toutefois, les médias étant friands de ce terme dont ils usent et abusent dès qu'un lanceur d'alerte diffuse de nouveaux documents, nous tenterons ici de proposer une définition qui puisse englober la multiplicité de son existence.

On utilise le terme *leak* pour désigner différents éléments.
Directement apposé à un nom propre, il définit la grande majorité du temps une *affaire* ou un *scandale*, tel que nous l'avons présenté ci-dessus.
Présentés au pluriel, les *leaks* définissent l'ensemble des documents, données, fichiers, dossiers associés à une affaire ou un scandale.
Proposé au singulier, le *leak*, en tant que donnée ou document, est un l'élément créé/modifié par la fuite sans que sa nature ne soit altérée. On use de ce terme notamment pour parler d'informations divulguées dans le domaine du *geek*, concernant des jeux vidéo, des sorties d'appareils technologies (smartphones, ordinateurs, etc.)

Cette dernière proposition, qui nous intéresse tout particulièrement, n'entend pas prendre le *leak* comme un unique élément arraché à son contexte mais bien comme un concept complexe définissant un type de document ou de donnée à un certain stade de son existence, ou plutôt en fonction du canal de collecte, de conservation, de traitement et de diffusion qu'il emprunte.

Pour être plus spécifique, il s'agit ici de concevoir le *leak* comme un élément modifié par son utilisation et/ou sa transformation. Toute donnée peut prétendre à cette appellation dès lors qu'elle possède un caractère privé, personnel, confidentiel ou secret et que celui-ci est violé. Il suffira qu'elle soit lue ou utilisée par des individus n'ayant aucun droit sur sa manipulation pour qu'elle soit considérée comme une fuite. Le document ou la donnée que l'on nommera *leak* pourra donc être de tout type, l'utilisation de cette dénomination sera fonction de son état d'existence, et le support n'importera que dans la mesure où, pour cette analyse, nous nous pencherons uniquement sur les canaux numériques empruntés par l'élément en question.

De fait, les fuites qui conduisent au *leak* peuvent être de nature très diverses. Afin d'en saisir toute la variabilité, il semble nécessaire de se pencher, tout d'abord, sur les scandales de fuites documentaires qui ont secoué les médias ces dernières années, et les organisations affiliées au leak, avant de pouvoir envisager les processus qui construisent le leak en tant qu'objet, et qui en font, ou n'en font plus, une archive. Enfin, nous observerons comment les représentations du leak contribuent à le faire exister en tant qu'archive.

DE SCANDALES EN SCANDALES

"Leaks are as old as secrets" (Kwoka, 2014, p. 1394)[29]

Petites histoires de grands scandales

Les années 2000 ont été profondément marquées par la révélation de scandales politiques, financiers, environnementaux et tant d'autres, alimentés et enrichis par la facilité d'accès aux données que peut offrir internet. Si la fuite existe depuis la nuit des temps, internet a offert un large terrain de jeu à sa diffusion, mais pas seulement.

On se souvient, sans doute, de l'affaire qui mena la nudité d'un grand nombre de célébrités féminines américaines à l'écran de tout un chacun. Ce *Celebgate* survenu en août 2014 offrit à la presse mondiale une matière savoureuse pour dénoncer la non-fiabilité des outils de sécurité mis en place par les fournisseurs et hébergeurs comme Apple, et en dehors de la responsabilité de ces firmes internationales quant à la protection des données de leurs utilisateurs, l'intéressant se trouve dans la manière dont ces données ont été diffusées. D'abord, parce que la première publication sera suivie par deux autres, toutes aussi massives, et les photos publiées sur de nombreux sites et plateformes. Ensuite, parce que malgré la réplique immédiate des célébrités exposées et la demande faite de supprimer les photos incriminées, une grande partie des photos (si ce n'est la totalité) sont, encore aujourd'hui, disponibles sur le web.

Comment est-ce possible ? le schéma de diffusion de cette affaire est particulièrement intéressant. La première diffusion officielle et remarquée fut effectuée sur le forum *4chan*, connu pour ses publications douteuses et son absence de modération. Ce dernier est un forum *imageboard*, c'est-à-dire un forum où l'on poste des images, et ses *boards* reposent sur un principe assez simple : tout sujet n'étant plus alimenté[30] est purement et simplement délesté. 4chan n'a pas d'archives ou d'archivage, et la plupart des sujets ont une durée de vie de quelques heures, voire, pour les plus 'chanceux', quelques jours. Comment, alors, les images ont-elles pu se diffuser aussi massivement ?

Il a suffi, pour cela, que quelques individus se chargent d'enregistrer sur leurs propres appareils les photos déposées. De là, ceux-ci postèrent à nouveau sur *4chan* ou une autre plateforme, et l'effet de groupe faisant, les photos incriminées furent déposées, redéposées, mises en ligne sans cesse jusqu'à épuisement. Le fait est que *4chan* étant un nid chaleureux pour la grande majorité des *trolls* existant sur la toile, il n'y a eu de leur part aucun épuisement. Sur *YouTube*, *Facebook*, *imgur* ou *Tumblr*, les photos furent censurées à *posteriori*, ces plateformes ne permettant pas la censure à *priori* ce qui, forcément, a laissé le temps aux photos de se retrouver hébergées ailleurs. Sur des plateformes de téléchargement illégal comme *ThePirateBay* ou *celebjihad*, la censure fut tout bonnement impossible. En 2016, on compte trois arrestations et condamnations pour *intrusion dans la vie privée*, mais aucune pour publication sur internet. De même, si les sites reconnus comme diffusant les photos firent l'objet d'un déréférencement de la part de Google, ils restaient accessibles via d'autres moteurs de recherche, quand les administrateurs ne choisissaient pas tout

[29] *Les fuites sont aussi vieilles que les secrets*.

[30] « *Since most boards are limited to ten pages, content is usually available for only a few hours or days before it is removed* » FAQ 4chan - http://www.4chan.org/faq#prunedelete

simplement de changer d'hébergement, et donc de nom de domaine, pour réapparaître sur Google quelques heures ou jours plus tard.

Cet exemple illustre parfaitement la difficulté de maîtriser un document ou une donnée une fois que ceux-ci ont fuité. On dit communément qu'internet *n'oublie pas*, et au-delà d'une quelconque forme de mémoire 'technologique' d'internet, ce sont les utilisateurs eux-mêmes qui sont et font la mémoire du web.

L'intérêt de cette affaire ne s'arrête pas là, et offre une singulière approche de la question du mouvement de *foule*. La publication de ces photos volées ouvrit en effet la porte à la publication d'autres photos dont la collecte était bien plus ancienne, et qui n'avaient jamais été publiées ou n'avaient pas reçues un accueil suffisant de la part du public – selon leurs diffuseurs. Ainsi à la masse de photo arrachées aux *clouds* des célébrités se mêlèrent de *vieilles* photos compromettantes, comme si la diffusion des premières encourageait la publication des secondes.

Ces effets dominos sont en fait constitutifs de la fuite, voire même plus largement de la dénonciation. Si le Celebgate n'est pas en soi une dénonciation[31] et n'a pas été effectué par des individus assimilables de près ou de loin à des lanceurs d'alerte, ce n'est pas le cas des Offshore Leaks, des Panama Papers ou du Cable Gate. Observons d'un peu plus près comment quelques-uns de ces grands scandales ont pu se dérouler.

Cable Gate

Le 2 septembre 2011, 251 287 câbles diplomatiques[32] sont diffusés par la plateforme et le site de lanceurs d'alerte Wikileaks. Les documents couvrent une période allant de 2004 à 2010 pour la grande majorité d'entre eux, quelques-uns d'entre eux vont jusqu'aux années 60 ; 40% d'entre eux sont classés confidentiels et 7% secrets[33]. En fait, la diffusion de ces documents a commencé près d'un an auparavant ; en partenariat avec cinq grands titres de presse (*Der Spiegel*, *El Pais*, *The Guardian*, *Le Monde*, *The New York Times*), Wikileaks a traité, trié, contextualisé et anonymisé pour certains, ces câbles diplomatiques, au compte-goutte. En effet, il faut à leurs équipes un temps considérable pour *nettoyer* les documents, et ainsi éviter d'exposer des informateurs, notamment. Au-delà de cette attention portée à la protection des individus pouvant être exposés par ces données, c'est aussi la recherche d'attention de la part du public qui justifie la lente diffusion des informations. Par un procédé emprunté aux grands titres de presse, justement, Wikileaks s'assure un public en attente et attentif à la totalité des informations qui lui sont proposées. En août 2011 pourtant, un fichier contenant l'intégralité des câbles est retrouvé sur internet, visiblement publié par des membres de Wikileaks plusieurs mois auparavant et faiblement protégé par un mot de passe ; Wikileaks est contraint d'expurger et de publier à la hâte les documents concernés avant qu'ils ne soient diffusés à tout va. L'intégralité de ces documents a été fournie à Wikileaks par Chelsea Manning, militaire américaine ayant accès, dans le cadre de ses

[31] Cette affirmation reste à démontrer. En effet, en publiant ces photos, les hackers ont également mis en avant la nécessité de se protéger du *phishing* (technique d'hameçonnage utilisée pour récupérer les mots de passe des célébrités) et l'impératif de réfléchir à comment réguler, maîtriser la diffusion une fois les données récupérées. La question n'est pas de savoir si les données vont être piratées, mais *quand* elles le seront.

[32] Les câbles diplomatiques sont des télégrammes envoyés entre les gouvernements et les ambassades. Dans le cadre du Cable Gate, ce sont des échanges effectués entre le gouvernement américain et les ambassades américaines à l'étranger dont il s'agit.

[33] Chiffres : Slate - http://www.slate.fr/story/31001/wikileaks-10-questions-cablegate

missions, à ces documents, et qui sera condamnée (avant d'être graciée en 2016) pour espionnage, notamment en raison de la confidentialité des données transmises à Wikileaks. Chelsea Manning est également à l'origine de la transmission à Wikileaks des *Afghan War Logs* et *Irak War Logs* (Marsac, 2015).

Offshore Leaks

Une vaste enquête sur des sociétés offshore est menée en 2013 par l'ICIJ et ses multiples médias partenaires. Pas moins de 2,5 millions de documents, soit 260 GO de données, ont été fourni par d'anciens employés de Portcullis Trustnet et Commonwealth Trust Limited[34] via l'envoi d'un disque dur au journaliste Gerard Ryle, qui traitera avec l'ICIJ pour les gérer. Les documents, qui révèlent l'existence de près de 120 000 sociétés offshore dans des paradis fiscaux, feront l'objet d'une base de données qui permettra aux journalistes de « télécharger et de chercher les documents », associée à un « forum pour faciliter la communication »[35]. L'affaire concerne des entreprises de tout type aux quatre coins du monde, et mènera notamment à l'affaire des *Princes Rouges* qui conduira la Chine à un scandale sans précédent. Les documents vont d'une période s'étendant des années 90 à 2010, et révèlent les agissements de nombreuses et diverses personnalités. Sur la plateforme spécifique aux Offshore Leaks de l'ICIJ, l'ONG précise :

> *"There are legitimate uses for offshore companies and trusts. We do not intend to suggest or imply that any persons, companies or other entities included in the ICIJ Offshore Leaks Database have broken the law or otherwise acted improperly"*

Dans une grande majorité des cas toutefois, les documents révèlent des affaires frauduleuses et conduiront à de vastes campagnes de procédures judiciaires ; on notera notamment, dans les personnalités françaises visées par ces enquêtes, l'ex-ministre délégué chargé du Budget Jérôme Cahuzac.

Révélations d'Edward Snowden

Edward Snowden est un ancien employé de la CIA et de la NSA[36] qui s'est fait connaître en 2013 pour avoir transmis à des médias des informations confidentielles et secrètes concernant la surveillance électronique effectuée par les grandes agences de sécurité dans lesquelles il avait travaillé. Il fera l'objet de lourdes condamnations et s'exilera en Chine et en Russie pour échapper aux représailles que ses révélations pourront engendrer. Les documents fournis, traités en premier lieu par *The Guardian* et *The Washington Post*, seront notamment portés par le journaliste d'investigation Glenn Greenwald et la documentariste Laura Poitras. Le parcours des documents et données est complexe à retracer. On compterait 1,7 million de documents transmis, mais nombre d'entre eux ne seront pas dévoilés au public à cause des risques pouvant peser sur la sécurité de certains individus dont les noms apparaissent dans ces informations. Les affirmations et démentis constants des deux agences visées, plus

[34] Portcullis Trustnet et Commonwealth Trust Limited sont deux entreprises de services financiers offshore, c'est-à-dire qu'elles fournissent des sociétés enregistrées à l'étranger (par rapport au pays de résidence de l'entreprise en demande) afin, en général, de favoriser l'extraction de flux financiers. Cette activité n'est pas illégale en soi, mais profite majoritairement à des entreprises désireuses de ne pas payer d'impôts, de taxes, etc.

[35] Le Figaro - http://www.lefigaro.fr/international/2013/04/04/01003-20130404ARTFIG00405-offshore-leaks-une-enquete-choc-aux-multiples-ressorts.php

[36] Respectivement *Center Intelligence Agency* et *National Security Agency*.

particulièrement de la NSA, rendent impossibles l'émergence de la vérité sur le contenu de la totalité des documents et sur leur importance ; ou plutôt, on ne peut savoir avec certitude si les *archives* Snowden, comme on les appelle communément, concernent les documents les plus importants de la surveillance électronique américaine ou non.

Luxembourg Leaks

Les Luxembourg Leaks, ou *Luxleaks*, sont des documents ayant fuité en deux temps. Les deux fuites sont l'œuvre de deux employés de PricewaterhouseCoopers (PwC), un cabinet d'audit, la première étant réalisée par Antoine Deltour, la seconde par Raphaël Halet. Les documents révèlent les agissements de multinationales, qui recouraient à la politique judiciaire du Luxembourg pour s'évader fiscalement. Comme pour les affaires impliquant Chelsea Manning ou Edward Snowden, ce sont les lanceurs d'alerte qui sont poursuivis, car jugés responsables de la fuite, ainsi que le journaliste à l'origine de la révélation et de la contextualisation des informations, Edouard Perrin. L'ICIJ et ses partenaires, principaux acteurs du traitement des documents et données, accèderont grâce à ces fuites à près de 4 GO de documents concernant une période s'étendant de 2003 à 2011.

Football Leaks

Toujours en Europe, les Football Leaks ont fait en 2016 trembler le monde du sport avec un nouveau scandale d'évasion fiscale. En vérité, l'affaire remonte à 2015 avec l'apparition d'un site, Football Leaks, qui dévoile des documents confidentiels. Ne trouvant pas l'écho suffisant, les lanceurs d'alerte feront appel à des médias. Ainsi, plus de 18,6 millions de documents, soit 1,9 TO de données, sont obtenues par le média allemand *Der Spiegel*, grâce à au lanceur d'alerte portugais pseudonymé John. L'histoire ne dit pas si ces documents ont été récupéré dans le cadre d'un emploi, si *John* et son équipe ont hacké les serveurs conservant ces données, ou encore s'ils se sont procurés ces documents grâce à d'autres sources. Grâce à l'écho trouvé via les médias s'étant emparés de l'affaire et l'European Investigative Collaborations (EIC), une vaste enquête a été menée sur les agissements fiscaux de joueurs de football de renommée mondiale, ainsi que sur des clubs, des entraineurs, etc.

Panama Papers

L'affaire des Panama Papers est l'un des derniers et des plus grands scandales ayant secoué les affaires financières et politiques du monde entier. En avril 2016, pas moins de 109 médias internationaux se mettent, de concert, à divulguer les liens entretenus par divers organismes et personnalités de tout horizon avec le cabinet panaméen spécialisé dans la domiciliation de sociétés offshore Mossack Fonseca. Pas moins de 11,5 millions de documents, soit 2,3 téraoctets de données, constituent la base de donnée mise à disposition par l'équipe de l'ICIJ. Comme pour l'affaire des Offshore Leaks, les sources de ces leaks resteront anonymes. Les données, initialement transmises au média allemand Süddeutsche Zeitung qui se gardera de dévoiler l'identité du ou des lanceur(s) d'alerte, auraient été préalablement vendues, par lots incomplets, aux « autorités fiscales allemandes, américaines et

britanniques »[37], appelant à une réaction de leur part : la vente sera fructueuse, puisqu'en 2015 une série de procès et d'amendes seront imposés à des établissements allemands concernés par ces Panama Papers. Le déroulé du traitement de ces leaks, documenté par les médias partenaires de l'ICIJ et notamment Le Monde, montre que la question de l'authenticité des données présentées est restée au cœur des préoccupations des médiateurs. Comme pour une grande majorité de ces affaires, l'intégralité des fichiers n'a pas été diffusée pour éviter que des données personnelles se retrouvent en pâture à l'univers sans pitié du web.

Sony Leaks

S'il faut parler d'univers sans pitié, sans doute l'affaire des Sony Leaks est-elle un exemple parlant. Non content de s'être fait pirater ses services de PlayStation Network en 2011, Sony se voit l'objet d'une des attaques informatiques les plus massives en 2014. Un groupe de hackers, *Guardians of Peace*, s'introduit sur les serveurs de la firme internationale et s'emparent de près de 38 000 000 documents, soit l'équivalent de 100 téraoctets de données allant des emails aux scenarii de productions futures, des contrats d'acteurs aux films prêts à sortir tels que l'*Interview qui tue !*. La coïncidence temporelle laissera croire que l'attaque vise à empêcher la diffusion de cette dernière œuvre qui s'attaque directement au régime dictatorial de la Corée du Nord. Si cette hypothèse ne sera jamais fermement confirmée, le chantage imposé, puis la diffusion des données volées et la destruction informatique des serveurs de Sony immobiliseront la société durant des semaines. Les pertes financières, inédites suite à une attaque de ce type, et l'absence totale de « bonne foi » dans la motivation des hackers ne permettent évidemment pas d'associer cette action à celles des lanceurs d'alerte. Toutefois, les canaux de diffusion des données (sites de téléchargement en *peer-to-peer*, *streaming*, forums tels que 4chan et Reddit), l'incapacité pour les équipes informatiques de Sony d'endiguer la fuite et la destruction de leurs matériels et les revendications paradoxales du groupe de hackers laissent perplexe :

> « Nous sommes une organisation internationale, qui inclut des célébrités issues des milieux politiques et de plusieurs pays comme les Etats-Unis, la Grande-Bretagne et la France. Nous ne sommes pas contrôlés par un Etat. Notre cible n'est pas le film 'The Interview', comme le laisse entendre Sony Pictures. » Szadkowski, Le Monde, 2014.

Les organisations

Lanceurs d'alerte et autres *fuiteurs* disposent, notamment dans l'espace numérique, d'un environnement étoffé d'organisations et de structures gérant l'alerte et la fuite. Des médias aux organisations de journalistes, en passant par les groupements de hackers, d'idéologues de tous univers ou d'internautes divers et variés, les organisations qui peuvent aider à la fuite et à l'alerte ne manquent pas. Certains sont plus officiels que d'autres, plus reconnus, avec des moyens plus imposants mais tous, à leur échelle, contribuent à faire de la fuite ou de l'alerte des éléments probants et solides.

[37] *« Panama Papers » : comment « Le Monde » a travaillé sur plus de 11 millions de fichiers*, Vaudano et Baruch, Le Monde.

ICIJ
International Consortium for Investigative Journalism

Dans les organisations officielles, reconnues, et aisément défendues et défendables, on trouve l'*International Consortium for Investigative Journalism* (ICIJ) qui, comme son nom l'indique, se place comme défenseur du journalisme d'investigation. Il a pour origine le *Center for Public Integrity*, qui est une organisation de journalisme d'enquête centrée sur les États-Unis. En créant l'ICIJ, ses fondateurs se permettent d'ouvrir leurs recherches et leurs intérêts à l'international. Sur son site, l'ICIJ se présente comme un « *global network* » fondé en 1997 par Chuck Lewis qui a comme objectif « *to bring journalists from different countries together in teams - eliminating rivalty and promoting collaboration* » et par-là même, de constituer une équipe de journalisme d'investigation transfrontalière. En s'associant ainsi à des journaux et journalistes étrangers, l'ICIJ peut compter de nombreuses affaires dévoilées, dénoncées voire condamnées grâce à ses travaux : « *Over the years, our teams have exposed smuggling by multinational tobacco companies and by organized crime syndicates; investigated private military cartels, asbestos companies, and climate change lobbyists; and broke new ground by publicizing details of Iraq and Afghanistan war contracts* »[38]. Ses équipes sont notamment à l'origine des révélations autour des *Offshore Leaks*, des *Luxembourg Leaks*, des *Swiss Leaks* ou encore des *Panama Papers*.

EIC
European International Consortium

Dans la même trame, mais plus spécifiquement orienté sur l'Europe, on trouve l'*European International Consortium*. On compte dans leurs *projets* les *Malta Files*, les Football Leaks ou encore une cartographie des armes de la terreur (*Mapping the weapons of terror*). Coordonnée par Stefan Candea, l'équipe de l'EIC compte des partenaires tels que Médiapart, El Mundo, Politiken, Le Soir ou Der Spiegel, médias européens connus et reconnus pour leur implication dans le traitement et la révélation de scandales financiers, politiques, etc. Lancé en 2016, il reste relativement jeune mais s'inscrit dans la juste lignée des organismes concentrés sur le journalisme d'investigation et la défense de la transparence numérique.

GlobalLeaks et le Centre Hermès

Le Centre Hermès pour la transparence et les droits humains numériques n'a rien, comme l'officialisme de son nom pourrait le laisser croire, d'une initiative portée par les pouvoirs publics. Présidé par Fabio Pietrosanti, l'organisme regroupe en vérité des militants de l'Internet libre considérés comme radicaux, et se fait connaître au lancement de sa plateforme destinée à la collecte de leaks, Global Leaks. Cette dernière, fonctionnant presque comme un réseau social du leak, permet aux journalistes et aux lanceurs d'alerte de communiquer en ligne de manière anonyme, et s'est imposée comme une référence pour de nombreux médias du monde entier. Soutenu par des ONG attentives aux questions de liberté d'expression, de l'information, de transparence et de droits numériques telles qu'Amnesty International ou Transparency International, le Centre Hermès s'inscrit dans la droite ligne des initiatives lancées par le jeune prodige à la courte carrière Aaron

[38] *About the ICIJ*, https://www.icij.org/.

Swartz, l'un des plus fervents défenseurs de l'internet Libre des années 2000 et créateur de Tor2Web.

Wikileaks

Du côté des organisations contestées, on trouve notamment la célèbre ONG *Wikileaks*, dirigée par le non moins contesté Julian Assange. Outre la réputation de son fondateur, Wikileaks a, depuis 2006, profondément changé les rapports aux archives numériques et papiers pour les entreprises, les gouvernements et les internautes. Wikileaks est le symptôme d'une génération numérique en recherche de transparence et de vérité, en dehors des notions de buzz, de scoops, de scandales. Les méthodes de sélection de l'information sont aussi discutables qu'un autre média (Julian Assange est par exemple accusé d'avoir volontairement saboté la candidature d'Hillary Clinton, ou d'avoir voulu « voter » contre Emmanuel Macron), mais la différence se trouve dans l'organisation de l'information et le type d'information proposé : un journal présentera des données contextualisées, narrées pour expliquer l'information à son lecteur ; Wikileaks offre des données presque brutes, certes organisées et indexées (pour la plupart des lots, un moteur de recherche est disponible) mais généralement non contées. Des initiatives pour contrer la puissance unique de Wikileaks n'ont cessé de fleurir suite à l'extension de son influence, et ont permis, au-delà même d'une tentative de concurrence, à une plus large diversité quant aux sujets divulgués par les leaks, ainsi qu'une plus grande prise en charge des pouvoirs publics.

Anonymous

Du côté des organisations qui œuvrent moins en tant que lanceurs d'alerte qu'en tant que *vigilanty*, on trouve les Anonymous. Le collectif *hacktiviste*, emprunte le masque de Guy Fawkes (et de fait, la révolte qu'il représente) et s'arme d'un emblème à la symbolique forte et très représentative de ce qu'ils font, de ce qu'ils combattent, et de ce qu'ils sont[39] : la corruption, l'anonymat, l'omniscience et le triomphe (ou le pouvoir). Sous cet étendard se regroupent des individus de divers horizons et qui n'ont pas toujours les mêmes intentions. En effet, n'importe quel internaute peut revêtir la bannière du collectif, l'idée d'une reconnaissance individuelle altérant fondamentalement les valeurs de cette culture *troll*. De fait, entre sous-groupe et sous-réseaux, absence de hiérarchie et organisation sous forme de multiples labyrinthes, ils se placent moins en tant que dénonciateurs qu'en tant que juges et bourreaux, en diffusant notamment des données personnelles et privées ou en attaquant des sites et plateformes. Ardents défenseurs de la liberté d'expression qu'ils considèrent inattaquable, inaltérable et indiscutable, ils sont la représentation d'une nouvelle forme de communautés auxquelles seuls l'espace numérique, internet et le web ont pu donner naissance.

On trouve encore bien d'autres organismes et organisations œuvrant, directement ou indirectement, à la collecte, le traitement ou la diffusion du *leak*. En France notamment, on repère French Leaks, initiative de Médiapart, EUleaks pour le parti politique Europe Écologie les Verts, Source Sûre pour le journal Le Monde

[39] L'emblème est constitué d'un costume avec un *white collar* (les *cols-blancs* qui désignent tant la bureaucratie que la fraude criminelle et comptable) ; d'un point d'interrogation à la place du visage (synonyme d'anonymat, notamment) ; d'un globe terrestre simplifié par les latitudes et longitudes ; d'une couronne triomphale.

ou encore Wildleaks (plateforme pour lanceurs d'alerte « écolos », destinée aux catastrophes écologiques et à la défense de la cause animale et soutenue par Global Leaks). À Dakar, la PPLAAF[40] est une initiative associative lancée en mars 2017, disposant d'une plateforme internet ainsi que d'une ligne téléphonique protégés, et Open Leaks tenta, un temps, d'endiguer l'engouement autour de Wikileaks.

DU PROCESSUS À L'ARCHIVE

Les 4 *C* du leak ?

La règle des 4 C (pour *collecter*, *conserver*, *classer*, *communiquer*) peut paraître étrange à appliquer aux leaks, pourtant elle semble particulièrement adaptée. Elle prend en compte les éléments du processus qui constituent la notion du leak, de sa fuite à sa diffusion qui lui permettent d'adopter la dénomination de *leak*.

Utilisée habituellement dans le cadre des archives, ou du moins par les archivistes en poste pour définir leurs missions, cette règle est parfois étoffée des notions de *conseil* et de *contrôle*. Ces deux notions supplémentaires ne sont pas pertinentes pour notre analyse.

Collecter

La collecte du leak s'effectue rarement, voire jamais, dans un cadre légal et réglementée. En effet, la notion même du leak impose que le document visé n'est pas sensé sortir de son lieu de conservation initial. La collecte se fait donc presque obligatoirement à l'encontre des règles imposées par l'organisme producteur de l'archive, puisque le salarié (lorsqu'il s'agit d'un salarié) viole le contrat qui le lie à son employeur. Lorsqu'il s'agit d'un piratage, le document est *volé*, puisque le lieu de conservation de celui-ci est forcé. Ce vol ne signifie pas forcément que le document n'apparaît plus dans le lieu original, le numérique permettant la copie de documents sans que cela n'entache l'authenticité du document *original*. Ainsi, comme nous avons pu l'observer grâce aux brefs retours sur quelques grands scandales, la collecte revêt des formes bien différentes.

Lorsqu'il s'agit d'une récolte effectuée dans le cadre des activités d'un employé (ou grâce à ce cadre), les documents sont en général transférés sur un support mobile, tels qu'une clé USB ou un disque dur. De là, plusieurs solutions s'offrent pour transmettre les documents à ceux qui les traiteront :

- Dans le cadre des Offshore Leaks, les lanceurs d'alerte envoie un disque dur contenant les leaks à un journaliste. Le support peut être une clé USB ou n'importe quel autre outil matériel et physique ;
- Les documents peuvent être transmis via une plateforme spécialisée pour les lanceurs d'alerte, comme French Leaks, GlobalLeaks, Wikileaks.

Ce que l'on sait des affaires, et notamment des lanceurs d'alerte impliqués (notamment Chelsea Manning et Edward Snowden), c'est que deux étapes de collecte se distinguent et font appel à des besoins et des objectifs bien différents. En effet, la première étape consiste à récupérer les documents qui paraissent nécessaire à la constitution de l'alerte. À cette étape, le lanceur d'alerte est seul face aux

[40] Plateforme de Protection des Lanceurs d'Alerte en Afrique - https://pplaaf.org/fr/

documents auxquels il a accès dans le cadre de ses activités, il agit en quelque sorte comme un espion qui sait plus ou moins ce qu'il doit récupérer, ou au contraire préfère récupérer tout ce à quoi il a accès. Une fois les documents en sa possession, le lanceur d'alerte transmet ces documents à un tiers : médias, administrations, individu haut placé dans la hiérarchie de son entreprise, activistes, militants, etc. En effet, le lanceur d'alerte, comme l'indique son nom, « lance une alerte », et ce avec l'idée d'*envoyer loin*, de *mettre en mouvement*, d'*engager* quelque chose[41], c'est-à-dire qu'il ne prend pas la responsabilité de traduire ce qu'il découvre, mais offre l'information à ceux qui peuvent s'en faire les traducteurs.

Le choix de ce premier récepteur, ce traducteur, celui à qui le lanceur d'alerte transmet les documents, est capital pour la définition juridique du lanceur d'alerte et de fait, pour sa protection, sa légitimation, sa *bonne foi* pour reprendre ce terme. En France par exemple, l'alerte diffusée aux médias ne peut survenir que dans le cadre d'une urgence immédiate qui exigerait de ne pas passer par les deux premiers interlocuteurs à alerter : l'interne (c'est-à-dire le supérieur hiérarchique ou le déontologue) et l'externe (soit le défenseur des droits, l'Agence française anti-corruption ou encore les délégués du personnel). Si le lanceur d'alerte ne respecte pas ces différentes étapes, il est aisément catégorisé comme délateur et de fait perd toute légitimité vis-à-vis des pouvoirs publics. Cependant, ça ne l'empêche pas d'être légitime pour certaines communautés, notamment numériques.

L'un des exemples les plus parlants à ce sujet concerne les actions menées par des pirates informatiques. Les scenarii de récolte peuvent être particulièrement nombreux, en fonction des motivations qui animent les hackers à l'origine des fuites, des protections dont se sont parés les organismes attaqués, etc.

Dans l'affaire des Sony Leaks, les hackers se sont emparés des données de l'entreprise et ont bloqué les accès, de sorte que Sony s'est retrouvé dans une paralysie presque totale, et l'incapacité de faire travailler ses employés. Les données diffusées visaient à nuire (toute hypothèse de *bonne foi* est ici écartée) directement l'entreprise, mais il est intéressant de noter que les *lanceurs d'alertes* ne se sont pas contentés de diffuser des documents et données appartenant à l'entreprise et à ses employés, ni de publier allègrement des productions de celle-ci (films et scenarii), mais a également détruit une grande partie des serveurs de l'entreprise.

Toutes les attaques informatiques et piratages n'ont pas une ampleur si dramatique. On peut prendre pour exemple les Macron Leaks, dont l'affaire, apparue à quelques jours de l'investiture du président Emmanuel Macron, a secoué la fin de la campagne présidentielle 2017. Les hackers se sont infiltrés sur les réseaux du parti politique En Marche! et n'ont pas cherché à supprimer, bloquer, détruire quoi que ce soit, se contentant de copier 9 giga-octets de données flirtant entre le privé, le confidentiel et le personnel. Les données, diffusées sur le forum 4chan, seront ensuite reprises par la plateforme Wikileaks qui permettra une navigation à travers l'intégralité de ces données non triées, et qui se chargera surtout de les mettre à disposition du public.

Conserver

Cette mise à disposition est effectuée selon des processus, parfois, très complexes. Les leaks passent en effet entre de nombreuses mains avant d'arriver au public, ou même parfois avant d'arriver au journaliste qui se chargera de faire le lien

[41] Définition « *lancer* », CNRTL - http://www.cnrtl.fr/definition/lancer

entre la grande quantité de documents et le public. L'un des éléments qui oblige les acteurs du leak à se multiplier est la question de la conservation. En fait, comme nous l'avons vu plus tôt, la nature même d'internet oblige ces acteurs à se multiplier, puisque la centralisation d'information autour d'un seul individu ne suit pas les grands idéaux qui jalonnent la politique d'internet, et de fait peine à fonctionner.

En dehors même d'un aspect idéologique ou éthique, le risque de piratage des données[42] constituant les leaks, ou plus exactement des serveurs les hébergeant, est plus que jamais présent. Pour pallier à ce risque – et bien que ce ne soit pas la raison principale à cette activité – les leakers usent d'une pratique qui fait plus que jamais ses preuves : la diffusion immédiate. On s'aperçoit en effet que pour certains des scandales qui ont jalonné l'histoire de l'internet ces quinze dernières années, les acteurs chargés du traitement des leaks n'hésitent pas à diffuser massivement les documents avant un quelconque traitement, en passant plus spécifiquement par des forums tels que 4chan ou Reddit, pour les plus connus. Si ces publications sont, pour la majorité d'entre elles, éphémères, elles permettent un maximum de visibilité, et l'action d'un grand nombre d'internautes qui pourront récupérer ces données et les conserver sur leurs disques durs ou leurs propres serveurs. Il n'y a pas tant une question de conservation telle qu'on l'entend dans le domaine archivistique, puisque l'idée ici est spécifiquement de pallier au risque de perte de données à court terme. Cependant, on peut noter que cette méthode fait ses preuves quant à la notion de perte ou d'oubli. L'affaire du Celebgate nous l'a par exemple prouvé, en faisant ressortir des photos volées quand bien même une chasse aux hébergeurs de ces photos avait été menée. Qu'importe le nombre de fois où l'on supprimera des données, et l'ampleur des moyens à disposition pour effacer ces traces sur l'espace numérique ; nous l'avons dit, et il convient peut-être de l'appuyer encore une fois, *internet n'oublie jamais*.

De fait, cette méthode permet également de diffuser des documents de manière relativement anonyme et de s'offrir une visibilité suffisante pour que l'action soit remarquée par des plateformes plus au fait du traitement des leaks. Wikileaks, par exemple, aura récupéré l'intégralité des Macron Leaks via 4chan, sans avoir, à priori, de contacts avec les pirates informatiques directement.

Mais au-delà de l'assurance de ne pas perdre de documents, comment sont véritablement conservés les documents ? le traitement qui est fait sur Wikileaks, GlobalLeaks, ou sur quelques autres scandales nous offre quelques pistes de réflexions qui ne peuvent que rester au stade d'hypothèses. Les documents diffusés peuvent être de supports très divers, mais sont, sur Wikileaks par exemple, présentés sous PDF afin, sans doute, de faciliter le téléchargement, le partage et le visionnage des informations. On trouve toutefois tout type de documents : texte, image, vidéo, base de données, email ou encore des télégrammes diplomatiques[43]. Les documents sont autant des scans de dossiers papiers que des données nativement numériques. Sur *Cryptome*, connu pour sa diffusion bien moins médiatisée que Wikileaks (mais également très peu commentée) de documents à caractère confidentiel, on trouve aussi bien du PDF que du JPG (qui correspond soit à des captures d'écran, soit à des scans de documents manuscrits ou non, soit encore à des photographies), des liens vers d'autres sites contenant des leaks, etc. La conservation du leak se fait ainsi

[42] Données parfois elles-mêmes piratées en premier lieu, lors de la collecte, ce qui offre une mise en abîme amusante.

[43] Les télégrammes diplomatiques sont des textes chiffrés, et de fait peuvent bénéficier d'un traitement particulier.

moins par la sécurisation des données que par la multiplication des lieux de stockage et de diffusion.

Quelques exceptions tout de même viennent contrebalancer l'absence de sécurisation des données en prenant garde à la préservation du contenu (au sens de non-communicabilité). On note que selon le type de plateforme, et de fait ses motivations et son positionnement sur la question éthique, les documents ne sont pas publiés dans leur intégralité, et surtout plus ou moins dépouillés des informations sensibles. L'équipe de Wikileaks a par exemple prouvé, avec le Cable Gate, qu'elle était particulièrement attentive à l'anonymisation de données pouvant s'avérer dangereuses pour la sécurité d'individus.

Classer

Le classement des leaks se rapproche déjà beaucoup plus du classement des archives que de leur conservation. Et à nouveau, en fonction des mains dans lesquelles les documents sont passés, on trouve différentes manières et processus de classement qui n'ont rien à voir les uns avec les autres.

Parmi les leakers particulièrement attentifs au classement de leurs leaks, on trouve, encore une fois, Wikileaks. En effet, les fuites massives font l'objet d'un traitement particulier qui permet aux lecteurs de naviguer dans les bases de données grâce à un moteur de recherche avancé, offrant la possibilité de se contenter d'une simple recherche par mots-clefs ou de plonger dans les champs booléens. Plus loin encore, le site propose des moteurs recherche différents selon les affaires. Pour exemple :

- L'affaire *Berat's Box*[44] : le moteur de recherche dédié à ces leaks propose trois types de filtres : textuelle (dans le contenu des emails), par nom de pièce-jointe, ou par identifiant mail (adresse email des correspondants) ;
- Les emails d'Hillary Clinton : cette fois, le moteur de recherche ne propose de filtrer que par identifiant mail ou par recherche textuelle : cependant, cette dernière permet d'utiliser des opérateurs booléens ;
- Les *War Diaries* qui rassemblent les leaks des affaires *Iraq War Logs* et *Afghan War Logs* proposent une simple recherche par mots-clefs, qu'il est possible d'affiner par période (de 2004 à 2010), par type (*Air Mission*, *Explosive Hazard* ou *Suspicious Incident*), par région, affiliation, classification, catégorie, par nombre d'ennemis tués ou encore par nombre de civils blessés. Il y a, en totalité, pas moins de 20 filtres différents.

Ces exemples ne montrent qu'une partie de l'étendue du type d'indexation que les équipes de Wikileaks sont capables de réaliser. La classification des documents (au sens archivistique, ici), est particulièrement élaborée et méticuleuse, autant dans une optique de traçabilité de l'information que de facilitation de l'accès à l'information. En effet, l'indexation dans ce cadre passe un processus méticuleux d'ajout de métadonnées qui concernent deux niveaux :

- Le premier niveau concerne l'état originel du document, ce sont les métadonnées qui lui ont été attribué lorsque le document n'était pas encore un leak, et n'existait que sur les serveurs de l'organisme à qui il a été dérobé. Ces

[44] 57 934 emails provenant de la boîte mail personnel du ministre de l'énergie turc Berat Albayrak, également gendre du président Erdogan, sont publiés par Wikileaks en décembre 2016.

métadonnées révèlent le niveau de confidentialité du document, le nom du producteur, les dates qui lui sont spécifiques, le format, le poids, etc. ;
- Le second niveau de métadonnées est ajouté par les lanceurs d'alerte eux-mêmes, ou du moins les médiateurs qui traitent le leak. Il peut s'agir de précisions comme le format et le poids (qui peuvent différer du premier niveau de métadonnées), de l'équipe qui a traité le document (les noms des lanceurs d'alerte ne sont jamais révélés, sauf quand ils sont déjà célèbres comme Edward Snowden, Chelsea Manning ou Antoine Deltour), et d'autres métadonnées nécessaires au fonctionnement des moteurs de recherche mis en place, et qui concernent donc le contenu et le sens du document.

Il en effet intéressant de voir que ce second niveau d'information à propos du document propose de revenir à l'essence même de l'information. Le contenu par le leak est, pour Wikileaks par exemple, très précautionneusement manipulé pour faciliter l'accès au moindre détail d'information. Ce n'est pas le cas de tous les médiateurs. Cryptome ne propose par exemple qu'un index sous forme de liste, avec la date de dépôt, le nom du document (un nom pertinent qui permette de savoir quel type de contenu on peut trouver), et une colonne permettant de savoir si le document est disponible sur un autre site, s'il a été reposté, mis à jour, réparé, etc.

Mais Wikileaks et Cryptome ne sont pas les seuls à être acteurs du classement des documents. L'ICIJ propose par exemple deux sous-sites spécifiques aux affaires des Offshore Leaks et des Panama Papers. Pour le cas des Offshore Leaks, le moteur de recherche propose un filtrage par pays ou par juridiction, et propose ensuite d'affiner en fonction du type de mot-clé entré dans la barre de recherche (société offshore, adresse, etc.). Ce qui est particulièrement intéressant dans ce cas, c'est que l'ICIJ ne propose pas l'accès au document en lui-même (la base de données complète est disponible en téléchargement mais les données brutes ne sont pas visualisables directement sur le site) mais aux données traitées et mises en contexte par rapport à d'autres données. Le site utilise un outil de *data visualisation*, Linkurious, qui permet selon sa page d'accueil de « découvrir ce qui est caché à l'intérieur de vos données »[45]. L'ICIJ se sert de cet outil pour permettre aux internautes de visualiser directement les informations de la base de données des Offshore Leaks de manière intelligible, en fonction de l'indexation préalablement effectuée. Ici, l'ICIJ est déjà dans la notion de communication.

Communiquer

On trouve aujourd'hui deux types de communicants dans le cadre du leak : les leakers et les journalistes. Les premiers, éternels acteurs de la fuite, sont les premiers médiateurs entre le leak et le public, l'interface de compréhension et de contextualisation de l'information. Ils sont également des garants de l'authenticité de l'information, avec la mission, de plus, de vérifier leurs sources. Mais si cette mission est la plus évidente, elle n'est néanmoins pas la principale dans le cadre du leak : en effet, le travail de communication s'effectue en amont, pendant le traitement de l'information où les journalistes tiennent une place prédominante.

Reprenons l'exemple de l'ICIJ. La manière dont l'organisme classe et traite les données qu'il reçoit montre un souci permanent de rendre l'information intelligible. Lorsque l'ICIJ classe une base de données, par exemple, le résultat de

[45] Selon la page d'accueil de Linkurious - https://linkurio.us/

son indexation sera déjà visualisable de manière intelligible. Le but alors n'est pas seulement de retrouver l'information mais de la comprendre, et d'avoir un aperçu des liens que les données entretiennent entre elles. Cette première contextualisation n'est certes pas définitive : les données proposées seront reprises, notamment par des journalistes, et contextualisées sous forme d'article en fonction des recherches effectuées et des éléments pertinents repérés. L'ICIJ, qui travaille avec près de 60 partenaires médias réguliers, sans compter les journalistes détachés, ne se charge pas personnellement de la communication au public, bien que son site soit accessible à tous et que les données et articles présentés soient disponibles pour tout lecteur : son but n'est pas tant de communiquer que de traiter les données reçues et de réaliser un travail d'enquête en profondeur en associant des partenaires d'horizons divers (et de fait avec des angles de recherches différents et complémentaires). De même que GlobalLeaks, ce sont les journalistes travaillant en parallèle qui racontent et diffusent l'information, et non pas ceux qui l'organisent (ou du moins, ces deux étapes sont distinctes l'une de l'autre).

L'organisation de l'information ne diffère pas beaucoup du côté des leakers tels que Wikileaks ou Open Leaks, mais ne regroupe pas les mêmes enjeux. Si l'ICIJ ne cherche pas à se faire entendre des médias (puisque leurs partenaires sont eux-mêmes des médias, la communication au public est constitutive de leurs travaux) Wikileaks, par exemple, doit parvenir à attirer l'attention des médias. Bien que sa réputation ne soit plus à faire, l'attention portée à l'explication des données proposées, à leur lisibilité, à l'organisation visant à rendre les informations compréhensibles pour tout un chacun s'est considérablement renforcée. Wikileaks propose, en plus de ses moteurs de recherche adaptés aux différentes affaires et lot de leaks, de nombreux articles en anglais, en français, en turc (en fonction des principaux pays concernés par le contenu des leaks). Il ne s'agit clairement pas ici d'un travail d'enquête et de journalisme, les études n'étant pas aussi poussées que celles de l'ICIJ, par exemple, qui travaille spécifiquement avec l'idée de faire remonter des informations et des scandales quand Wikileaks se positionne plus du côté de la transmission, du flux. Cependant, les articles visent à donner le contexte non pas du contenu du scandale, mais du leak en lui-même. De fait, on pourra trouver dans ces textes l'histoire qui précède la diffusion du leak, la manière dont les lots ont été traité, l'indexation qui en a été faite, les modifications qui ont pu être effectuées sur les documents (changement de formats, anonymisation de certaines données personnelles, suppression d'informations sensibles, etc.) et bien évidement le sujet auquel ils touchent. En fonction des cas, on pourra également trouver les noms, pseudonymes ou équipes de lanceurs d'alerte ayant participé à la divulgation des leaks, les affaires précédentes auxquelles ces leaks sont reliés, etc. Pour des leaks diffusés en plusieurs fois (l'exemple des *Vault* est particulièrement parlante), chaque publication fait l'objet d'un article recoupant entre un et une dizaine de documents fuités : le traitement de l'information est ainsi plus précis, mieux découpé, et permet de ne pas perdre le fil de l'information.

Ainsi Wikileaks choisit, comme l'ICIJ par exemple, de type de transmission de l'information ; d'une part, les bases de données presque complètes sont publiées indexées et rendues navigables par un moteur de recherche : c'est aux lecteurs, journalistes, de faire les recherches et le tri en fonction de l'information qu'ils recherchent ; de l'autre, les documents sont triés en fonction de leurs liens et publiés par petits lots, et contextualisés de sorte que le lecteur sait, avant de lire le document, ce qu'il va y trouver. De plus, comme pour l'ICIJ, Wikileaks permet aux journalistes de tous médias de s'emparer des données afin de les contextualiser et de les publier

pour le grand public : les données sont ouvertes, leur exploitation libre, et la propriété du leak n'importe pas.

Des acteurs aux processus

Cet essai de parallèle entre les 4 C de l'archive et ceux du leak est particulièrement parlant du point de vue du placement des acteurs par rapport au traitement du leak. En effet, aucun acteur ne se contente d'un rôle ; les tâches ne sont pas strictement divisées et ne tendent pas à l'être. On reconnaît bien évidemment le rôle prépondérant des journalistes dans la diffusion de l'information, ou la fonction évidente du lanceur d'alerte dans la divulgation du document ; cependant, les premiers auront tout autant un rôle dans le traitement de l'information que les seconds, et entre les deux, les leakers seront acteurs d'une médiation qui ne leur est pas pour autant spécifique.

La place de chaque acteur et le suivi de chaque processus ne sont en effet pas strictement divisibles. Ils se complètent en permanence dans des allers-retours qui peuvent parfois paraître confus mais qui répondent en vérité à une nécessité profonde de ne pas être le détenteur unique d'une information ; rappelons-le, le monde numérique naît dans un mouvement d'autorégulation qui refuse le principe de hiérarchie verticale, et le leak d'aujourd'hui ne se développant pas à la marge mais en plein dans cette expansion numérique, son traitement ne pourrait se faire selon des règles qui ne correspondent pas à l'*esprit* du web. Si Wikileaks et l'ICIJ sont prépondérant sur la scène du leak, ils sont très loin d'être les seuls : d'OpenLeaks[46], créé justement pour contrer la puissance de contrôle de l'information que pouvait posséder Wikileaks au Centre Hermès pour la transparence numérique et sa plateforme GlobalLeaks, les initiatives de toute sorte se multiplient et offrent, en plus d'une grande diversité concernant l'accessibilité et le traitement des sources, de nombreux processus, multiples, variés, et pour la majorité d'entre eux complémentaires.

Toutefois, il apparaît qu'une autre forme d'actions du leak ressort de cette analyse, non pas en quatre point comme le propose l'archive, mais en trois. Si la collecte reste un élément stable et indiscutable dans le processus du leak, la question de la conservation est plus complexe et n'apparaît pas comme un objectif pour les plateformes de lanceurs d'alerte : leur but est moins de présenter le leak en tant que preuve aux autorités que de le divulguer en masse au public. Encore une fois, le leak contourne les limites légales pour emprunter des chemins définis par l'espace numérique dans lequel il évolue. La notion de classement correspond mieux à un traitement qui sélectionne, trie, indexe, et les possibilités de communication et de diffusion de l'information apparaissent à différents niveaux, et de la part de différents acteurs. De fait, il apparaît que trois grands types d'acteurs du leak sont récurrents, bien que difficilement contenus par des frontières indistinctes et non pertinentes : ceux qui collectent, ceux qui traitent, et ceux qui diffusent.

[46] La plateforme, créé par Daniel Domscheit-Berg, ancien de Wikileaks

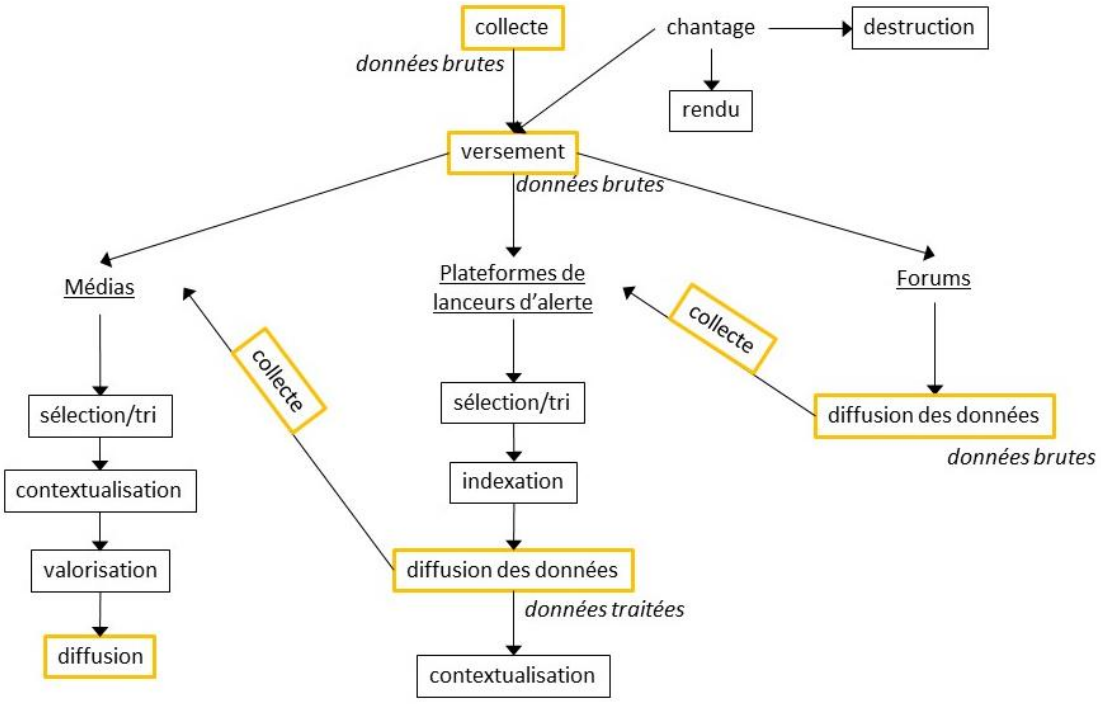


Figure 1 - Les processus du leak

L'alerte en France

L'alerte peut toutefois n'être ne pas être tout à fait de l'ordre de la fuite et entrer dans un cadre légal et réglementaire qui n'offrira pas le même processus que décrit ci-dessus. Rappelons qu'en France, dans le cadre de ses activités professionnelles[47], un salarié devra[48] avoir la possibilité de signaler toute activité qu'il juge illégale, immorale ou illégitime à « des personnes morales de droit public ou de droit privé ou des administrations de l'État » (décret n° 2017-564 du 19 avril 2017). Ainsi les organismes se devront de s'équiper d'un représentant dit « déontologique » qui sera garant d'une « procédure de recueil des signalements » à laquelle devra se conformer le lanceur d'alerte pour être considéré comme tel et, de fait, jouir des protections qui sont accordées à son statut. Le référent sera, dans cette même logique, « susceptible de recevoir les alertes », et aura donc, de fait, un rôle de *leaker* au sens où il prétendra au recueil des documents. Il ne sera toutefois pas le seul acteur de ce processus puisqu'il aura à charge de faire connaître les faits sur lesquels on tente d'alerter aux sphères hiérarchiques ainsi qu'aux autorités compétentes. De plus, la création de cette procédure fera également état de la « stricte confidentialité de l'auteur du signalement, des faits objets du signalement et des personnes visées » et permettra notamment au lanceur d'alerte de « fournir les faits, informations ou documents quel que soit leur forme ou leur support de nature à étayer son signalement ». Mais ce cas de figure ne peut se présenter que dans le cas où le lanceur d'alerte choisit de passer par ce processus qui, en plus d'être relativement nouveau, ne concerne que la France et ne peut prétendre au recueil et au bon traitement de toutes les alertes.

[47] Les textes se gardent de préciser si les activités associatives sont encadrées par ce processus, clairement orienté en vue de protéger les lanceurs d'alerte au statut d'employé des représailles de leur entreprise, notamment.

[48] Le décret, datant du 19 avril 2017, entrera en vigueur au 1er janvier 2018.

Le processus de définition et de traitement de l'archive qui apparaît en France avec les modification réglementaires de 2017 pourrait ressembler à ce schéma :

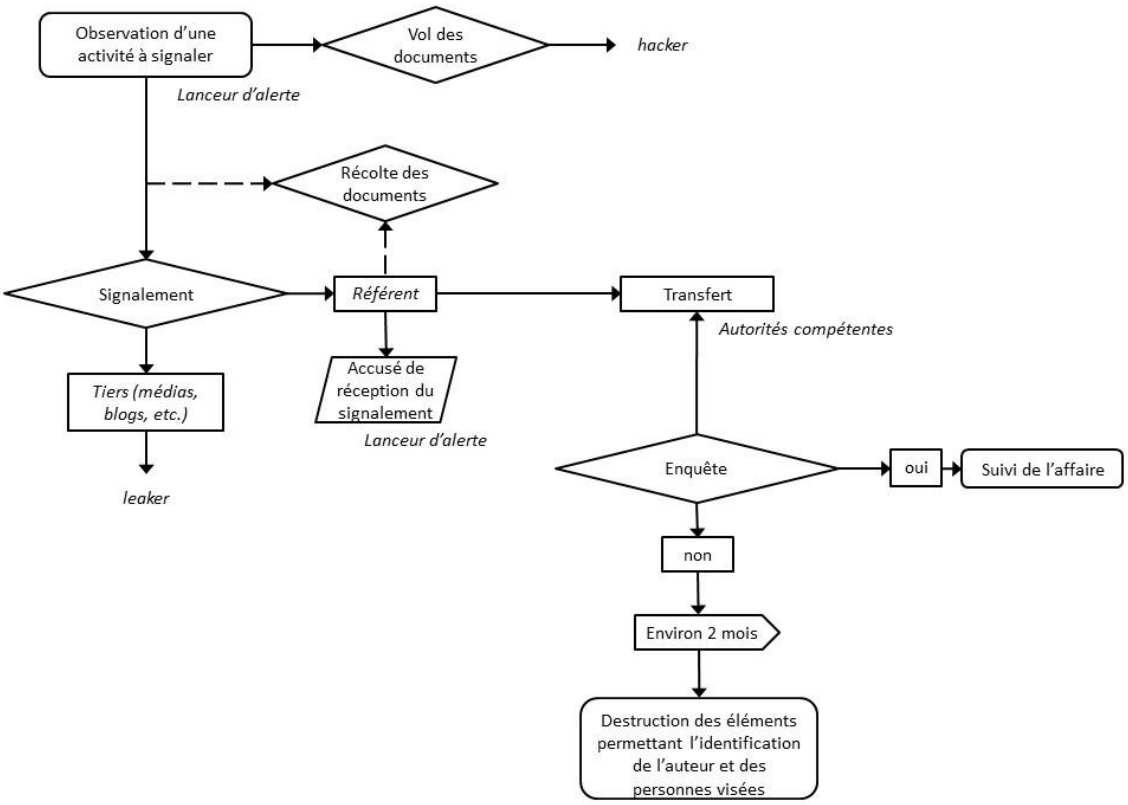


Figure 2 - L'alerte réglementée en France

D'une typologie du leak

Typologies existantes

La fuite, le leak et l'alerte, on le sait, ne datent pas d'hier. Ce qui change profondément et invite à une réflexion plus approfondie sur la nature et l'essence de ces documents, c'est la massivité avec laquelle ils sont dévoilés. Jamais on n'a vu une telle quantité de données récoltées et diffusées en si peu de temps, avec une communication aussi large et incontrôlée, et autant d'individus se penchant sur le contenu de ces données fuitées.

L'expansion de ces fuites laisse entrevoir de nouveaux processus et *cycles de vie*, si l'on peut les nommer ainsi, de la fuite, du leak et de l'alerte. Pour Margaret Kwoka, c'est éminemment le numérique qui change la nature de la fuite :

> *"Technology has changed the access to information lower-level government officials and contractors have, thereby enabling them to deluge leak. It has also vastly increased the ease of distributing leaked information."* (Kwoka, 2014, p. 1402)

En effet, un employé, aujourd'hui, n'aura pas accès uniquement aux documents sur lesquels il travaille puisqu'il aura accès à un serveur interne regroupant toute ou partie des informations relatives à son entreprise. Les règles établies par la hiérarchie et les services informatiques ne peuvent suffire à maîtriser l'accessibilité aux données. Pour exemple, il suffit d'un bug, d'une erreur d'attribution de rôle, d'un oubli de fermeture de session pour que les données soient accessibles à tout un

chacun ; de fait, à partir du moment où un employé possède un ordinateur ou tout autre appareil numérique de travail et que, pour la réalisation de ce dernier, il a accès à une session sur les serveurs internes à l'entreprise, tous les éléments techniques sont réunis pour qu'une fuite se produise. De plus, une fuite massive entraîne indubitablement la découverte de plusieurs faits pouvant prétendre à l'alerte sans même que le lanceur d'alerte en soit conscient :

> *"Recent leaks encompass vast quantities of records that the leaker likely knows nothing about, if he or she has even read them"* (Kwoka, 2014, p. 1390)

Il suffit, en effet, d'un seul point d'entrée pour que de vastes opérations soient dévoilées. La fuite aujourd'hui ne peut se contenter de quelques documents disséminés ; les Football Leaks et leurs 18,6 millions de documents ou les Offshore Leaks et leurs 260 giga-octets ont prouvé, s'il le fallait, que la masse compte presque autant que le contenu dans la popularité du scandale.

David Pozen, Margaret Kwoka proposent tous deux différentes manières d'appréhender la typologie du leak. Nous allons tenter d'en présenter les principaux traits avant de proposer notre propre typologie.

David Pozen tout d'abord distingue deux types de fuite en fonction de la quantité de matière récoltée, ou plus exactement la précision de la récolte :

> *"An additional distinction (really, a spectrum) worthy of note relates to the quantum and scope of material disclosed and the difference between what we might call specific leaks, which convey a limited amount of content about a discrete matter, and general leaks, which disclose vast swaths of information more or less indiscriminately"* (Pozen, 2013, p. 533)

On observe ici deux procédures distinctes de collecte du leak. D'une part, on choisit de faire une récolte dite *spécifique*, avec des facteurs discriminants afin de ne récolter *que* l'ensemble de documents et données désiré : le lanceur d'alerte sait très exactement ce sur quoi il veut alerter et quels documents sont nécessaires à la validation de son alerte. De l'autre, la récolte est faite de manière très large et avec des critères peu discriminants soit parce que les moyens techniques le permettent, soit parce que le lanceur d'alerte ne sait pas exactement quels documents feront office de preuve, soit encore parce qu'il désire alerter sur les agissements d'un organisme de manière générale[49].

Ces *general leaks*, Margaret Kwoka les nomme *deluge leaks*, en se concentrant plus spécifiquement sur la quantité finalement récoltée plutôt que sur la méthode de récolte (Kwoka, 2014, p. 1401) sans pour autant ignorer *le périmètre de recherche*, *l'identité du lanceur d'alerte* ou les *motivations derrière la fuite* :

> *"Deluge leaks are distinct from past leaks in their scope, the identity of the leaker, and the motivations behind the leak. Their typology, however, is not their only difference. This new type of leak also comes with new types of risks, ones that are likely of concern to a broad cross-section of society"* (Kwoka, 2014, p. 1402)
> *"Deluge leaks are[…] of a different nature than the leaking that had occurred before"*(Kwoka, 2014, p. 1404)

[49] Ces suppositions quant aux raisons qui justifient le choix d'une récolte spécifique ou générale ne cherchent pas à être exhaustives ; les motivations des lanceurs d'alerte ne font pas l'objet de cette recherche.

Elle considère les *deluge leaks* comme un nouveau type de fuite inhérent aux capacités techniques et technologiques du numérique, et dont le nom prend la pleine mesure du *déluge* que peuvent représenter ces récoltes et, de fait, la diffusion d'icelles. De plus, par les risques de poursuites encourus par les lanceurs d'alerte, par les motivations de ces derniers, par la nature des révélations réalisées par ces fuites, par leur réception par les médias et, surtout, par le public, la nature des « fuites-déluges »[50] varie considérablement. En ce sens, Pozen distingue trois sortes de fuite, distinction réalisée notamment d'un point de vue juridique. D'abord, il distingue les *leaks* des *plants* :

> *"[…] although legal scholars generally have not done so, some journalists have distinguished between leaks and plants. Plants are taken to be 'authorized' disclosures designed to advance administration interests and goals. Leaks are 'unauthorized' disclosures"* (Pozen, 2013, p. 534)

À ces fuites allant à l'encontre des intérêts du gouvernement et à celles les favorisant (Kwoka 2014, 1395), il note que certaines, non catégorisables par ces deux types, flottent dans une zone grise ou neutre :

> *"[…] much of what we call leaking occurs in a gray area between full authorization and no authorization, so that it is neither 'leaks' nor 'plants' but what I will term pleaks that dominate this discursive space"* (Pozen, 2013, p. 515)

Margaret Kwoka ajoute :

> *"Merely distinguishing between leaks and plants, however, still lumps together a wide variety of activity under the label "leaking" without accounting for important differences between various kinds of leaking activities. In fact, leaks vary principally by the motivation of the leaker, the identity of the leaker, and the scope of leaked material. Leaks may be motivated by a variety of concerns"*(Kwoka, 2014, p. 1395)

Pour elle, cette catégorisation fait cruellement défaut aux questions de l'*activité* de la fuite, c'est-à-dire de l'*identité du lanceur d'alerte*, de sa *motivation* et de la *portée des documents divulgués*. Le concept *d'autorisation* (Pozen 2013, 566) ne suffit pas à créer une typologie de la fuite. On peut y ajouter la hiérarchie de la personne à l'origine du leak (Kwoka 2014, 1395), le type de révélation, le support documentaire, la motivation du médiateur (tels que Wikileaks, GlobalLeaks ou toute autre plateforme n'étant ni à l'origine du *leak*, ni véritablement actrice de sa contextualisation[51]. De plus, ce choix de catégorisation est fortement remis en question par la difficulté pour les journalistes de tracer l'origine de la fuite, et les motivations sous-jacentes :

> *"Plants and pleaks may be camouflaged as leaks. Journalists at the top of their profession say they often do not know, and could not realistically ascertain, the precise level or type of authorization their sources have received. The distinctions among leaks, pleaks, and plants become messier still in cases that involve multiple disclosures. News stories not infrequently combine elements of leak and plant. Journalists may acquire some initial piece*

[50] Traduction proposée par Foegle.

[51] Par opposition aux journalistes, par exemple.

of classified information through a leak or pleak or an off-the-record hint[…]" (Pozen, 2013, p. 571)

De plus, ces propositions ne prennent pas en compte le processus, non pas celui qui conduit au leak, mais celui qui *le* conduit, c'est-à-dire le *cycle de vie* du leak.

Pour une typologie du leak

Nous ne nous attarderons pas sur les motivations d'exercice de la fuite, qui ne sont pas l'objet de ce mémoire, et de fait ne nous pencherons pas plus en avant sur les *plants* et les *pleaks* définis par Pozen. Cependant, le premier niveau qui nous intéresse se définit en partie par les raisons qui poussent les lanceurs d'alerte à le devenir. L'alerte, puisqu'il ne s'agit pas encore de fuite ici, peut être distinguée en deux axes : l'alerte réglementée et celle qui ne l'est pas. Comme nous l'avons observé lors de la présentation des processus du leak et de l'alerte, il est possible de faire une première séparation à ce niveau puisque l'alerte réglementée et la fuite ne suivent pas le même chemin. Tant qu'elle reste dans le cadre réglementaire et légal – en France tout du moins – l'alerte ne peut pas être considérée comme un leak puisqu'elle ne *fuit* pas, elle ne se déverse pas en dehors de son activité. Le document ne sort pas de son espace, on ne fait que prévenir et alerter qu'il existe, et qu'il contient des faits à alerter, ou qu'il fait partie d'un ensemble d'informations qui nécessitent une alerte. L'alerte vise autant à améliorer la gestion de l'entreprise ou de tout autre organisme qu'elle concerne, tout en protégeant ceux qui préviennent. Si, à ce premier niveau, l'alerte ne reçoit pas l'accueil espéré par le lanceur d'alerte, il peut saisir d'autres autorités, plus compétentes et surtout plus impartiales ; ce n'est qu'une fois toutes les ressources prévues par le cadre légal épuisées qu'il lui est permis de sortir de ce fameux cadre.

- **Alerte réglementée** : document ou donnée produit ou reçu dans le cadre des activités professionnelles du lanceur d'alerte et faisant l'objet d'un signalement dans le cadre légal et réglementaire des lois sur l'alerte (comme en France ou en Belgique, par exemple).

Ainsi il existe une alerte réglementée, et une non réglementée. La seconde, de par son absence de cadre, recoupe un ensemble assez vaste de processus et d'objets différents. Si elle n'est pas réglementée, elle n'est est pas pourtant illégale : ce n'est pas l'action d'alerter qui contrevient aux lois, mais le fait de diffuser les documents qui lui sont associés, de les voler, de les manipuler. On pourra prétendre que la différence est d'ordre sémantique et que, puisque toute action effectuée sur les documents découlant de l'alerte peut être jugée illégale, l'alerte elle-même l'est. Il semble toutefois que la nuance soit plus fine.

- **Alerte non réglementée** : document ou donnée à caractère confidentiel, privé, personnel ou secret, concernant des personnalités ou tout organisme public ou privé, faisant l'objet d'un signalement pour un fait considéré comme immoral, illégal ou illégitime, et ce sans entrer dans le cadre légal et réglementaire des lois sur l'alerte.

À partir de ce premier niveau, et plus particulièrement de cette *alerte non réglementée*, il est possible de détailler plus en avant les différents types de leaks qui en découlent. Il semble qu'on puisse distinguer deux types de fuite, en fonction de la manière dont ils ont été récolté. Il s'agit du second niveau :

- **L'élément piraté** : document ou donnée résultant d'un piratage informatique visant à exposer des informations relatant des faits considérés comme immoraux, illégaux ou illégitime ;
- **L'élément leaké** : document ou donnée visant à exposer des informations relatant des faits considérés comme immoraux, illégaux ou illégitimes, collecté :
  - dans le cadre des activités professionnelles du lanceur d'alerte ;
  - hors du champ des activités professionnelles du lanceur d'alerte mais ne résultant pas d'un piratage informatique.

La fuite quant à elle entend regrouper l'ensemble de ce deuxième niveau (piratage et leak), puisque le principe même de la fuite est de faire passer des informations d'un lieu A, qui leur est réservé, à un lieu B, qui n'est pas sensé recevoir ces informations. Dans l'alerte réglementée, les textes prévoient justement un récolteur de ces informations qui est spécifiquement présent pour les recevoir, pour les traiter. De plus, il est à noter que *l'élément piraté* tel qu'il est entendu dans cette analyse ne cherche pas à recouvrir les piratages initiés par la seule volonté de nuire à l'organisme attaqué sans véhiculer un autre message que le pouvoir de destruction possédé par les hackers. Les Anonymous, par exemple, dans une majorité des piratages qui leur sont associés, comportent une motivation d'alerte ou, le cas échéant, de jugement (rappelons à cet effet que les Anonymous peuvent aisément être associés à des *vigilanty*).

REPRÉSENTATION DU *LEAK*

L'idée du *leak* sous-entend qu'il y a eu une brèche à un moment donné, ouverte par un élément perturbateur d'ordre informatique (bug, piratage, mauvaise manipulation) ou non (conservation de documents d'activité, récolte, constitution de preuves) et qu'un vaste champ d'information a, par ce biais, pu être exploité d'une manière ou d'une autre par des personnes non-habilitées. On appelle « personnes *non-habilitées* » des individus ayant eu accès à ces documents sans autorisation, soit par piratage, soit par récolte de documents envoyés par des personnes elles-mêmes habilitées. De fait, cette idée de brèche appelle les responsables de nombreuses entreprises à faire preuve de prudence ainsi qu'à mettre en place des mesures destinées à dissuader ou à minimiser la fuite. Ces *data-loss prevention* comme on les nomme font surtout appel à une représentation négative du leak, qui conduit à des barrières érigées pour l'empêcher, barrières augmentant l'intérêt que peuvent porter les lanceurs d'alerte, les leakers ou les hackers à ces données.

Car le leak est une revendication, au droit de savoir, de dénoncer, d'expression. Tenter de le contrer, c'est justifier, aux yeux de ceux qui convoitent la donnée cachée, la nécessité de l'obtenir et de le divulguer. C'est également, et surtout, prouver que des actions effectuées par les organismes protégeant leurs données peuvent être condamnables et de fait laisser entendre que les documents en question portent en eux la preuve de leurs mauvaises actions. Car le leak est un outil de justification, de preuve, au même titre sans doute que l'archive.

L'obstination qu'ont les médias à vouloir considérer le leak comme une archive (les Archives Snowden, Manning, Wikileaks, etc.) ne répond pas seulement à la nécessité de trouver un synonyme vague du document ou de la donnée. Au-delà d'une simple traduction de l'anglais *archive*, cette dénomination répond à des

critères de l'archive associables au leak tel qu'il est manipulé par ceux qui le divulgue.

Le *leak* en tant que preuve

Une archive pour les producteurs

L'une des caractéristiques primaires de l'archive tient en sa capacité de preuve, à « servir de preuve à l'action qu'elle supporte » (Chabin, 1999). En fait, qu'on la cantonne à un besoin d'historicité, d'écrire l'histoire, de la traduire, ou qu'on considère l'archive comme un outil de justification face à des risques juridiques (pour une entreprise, un organisme public ou encore un particulier), elle est profondément marquée par la nécessité de *faire preuve*. Une archive qui n'a pas la possibilité de prouver n'est pas une archive. Cette affirmation ne veut tout de même pas dire que la preuve se doit d'énoncer une vérité ; du moins elle justifie bien une vérité, expérimentable et contestable, qui, quoi qu'il en soit, ne pourra s'affirmer comme unique.

La notion de preuve est l'une des premières – si ce n'est pas la première – raison qui pousse les entreprises à constituer une politique archivistique. La crainte de ne pouvoir fournir un document en tant que pièce justificative ou probatoire et l'expérience de la perte de données menaçant l'économie de l'entreprise, par exemple, sont autant d'éléments appuyant l'effort fait vers les archives, et vers la préservation de leur capacité à faire preuve. Car les archives, comme l'explique Marie-Anne Chabin, ne sont « produites non pour elles-mêmes mais pour prouver un droit, prévenir un risque, tracer par écrit une réalité à laquelle on aura besoin de se référer demain (le cas qui correspondant aux données NSA), ou encore enregistrer une mémoire » (Chabin, 2016). L'explosion numérique, la dématérialisation et l'accélération de certains processus de travail ont demandé, de fait, une attention de plus en plus portée vers la question de l'archive, de sa traçabilité, de son authenticité ou de son intégrité.

Les documents des Offshore Leaks, des Football Leaks ou des Panama Papers naissent du besoin de valider, de justifier, de prouver la création de structures ou d'échanges et d'accords commerciaux entre différentes instances. Pour les données fournies par Chelsea Manning, il s'agit plus de la nécessité de retracer la progression des actions militaires américaines à l'étranger, ainsi que de rendre compte des activités et des faits survenus durant la durée des opérations, et de tenir des journaux de bord utiles à la compréhension des choix stratégiques et organisationnels effectués sur le terrain. Edward Snowden, lui, nous offre notamment l'accès à des réflexions et des avant-projets concernant la surveillance informatique de la CIA ou de la NSA, mais également des guides d'utilisation d'outils destinés à l'espionnage électronique aux États-Unis et dans les pays du monde entier. Ces archives sont nécessaires à la construction des organismes qui les produisent et servent de trace de leurs avancements, de support à leurs activités, et au-delà encore :

> « Ces archives-là sont des documents récents, vivants, sensibles, stratégiques. Ce sont des documents émanant directement du pouvoir. C'est bien le sens originel du mot archives (arkheia, archivum) vu sous l'angle de la valeur du contenu et non du lieu (les documents entreposés et non l'entrepôt) ou, plus exactement sous l'angle de la valeur du contenu qui justifie un lieu spécifique de conservation » (Chabin, 2016)

Les documents qui constituent le leak sont des éléments de preuve pour leurs producteurs. Documents d'activité retraçant les faits datant parfois de plus de cinquante ans d'organismes divers, ils n'existent que pour faire preuve au sein de ces mêmes organismes. Parce qu'ils sont confidentiels ou secrets, pour la plupart d'entre eux, ils n'ont toutefois pas vocation à faire preuve auprès d'un public large mais auprès des hiérarchies productrices, et éventuellement auprès les autorités qui en feraient la demande.

Ces archives dites *sensibles* parce qu'elles ne doivent pas sortir de leur environnement premier constituent de fait les secrets des organismes producteurs et détenteurs de celles-ci. Elles sont *chasse-gardée*, et sont vouées à rester silencieuses si ce n'est pour les individus les produisant, et ceux qui, prenant leurs fonctions par exemple, en auraient l'utilité dans le cadre de leur activité. Elles peuvent être particulièrement détaillées et, pour l'exemple des *Afghan* et *Iraq War Logs*, être rédigées au fil de l'eau des activités en question. Dans ce cas en effet, il s'agit des journaux et rapports réguliers des militaires au cœur de la guerre, et sont de fait des documents de travail utilisés par leurs hiérarchies. Dans le cas des documents divulgués par Snowden, il s'agit « de données, de documents, de courriers qui sont les traces écrites des opérations de surveillance, illicites voire illégitimes, effectuées par l'Agence de sécurité d'une des plus grandes puissances de la planète, sources de tensions entre les États du globe » (Chabin, 2016). Ce ne sont pas des documents qui ont pu faire l'objet d'une validation, par exemple ; ils constituent un patrimoine de guerre délicat qui renseigne et informe plus qu'il n'officialise, et qui en tant qu'archive « permet d'affirmer que tel fait s'est passé et qu'il s'est bien passé de telle façon, sans qu'il y ait forcément contestation. C'est ainsi que les archives constituent la source de l'Histoire par excellence. Leur valeur première est de montrer à quoi elles ont servi » (Chabin, 1999).

Si leur contenu peut être particulièrement bien documenté, il en est de même pour les données qui les décrivent. Les métadonnées qui les accompagnent sont une véritable mine d'information qui permet de lire « les motifs de son élaboration, c'est-à-dire la poursuite d'une action donnée et le contexte dans lequel [l'archive] prend place » (Chabin, 1999). Elles contiennent des informations qui vont du nom de l'identité du producteur (nom, statut, hiérarchie, etc.) au niveau de classification du contenu (secret, top-secret, confidentiel, etc.) en passant par toute donnée temporelle relative à la création, la manipulation et l'expiration du support documentaire et de son contenu, et sont commentées, augmentées, liées à d'autres données et documents qui les complètent. De fait, elles sont relativement solides en tant que preuve (traçables, authentiques, intègres), et se parent parfois des garanties (signatures, tampons, etc.) qui assurent la fiabilité et la recevabilité de l'archive en tant que preuve.

Une preuve documentaire pour les lanceurs d'alerte

Mais pour que l'archive soit considérée véritablement considérée comme une preuve, il faut, selon l'historien Charles Reagan, « formuler une question, une hypothèse » (Reagan, 2008, paragr. 21). Cette action, si elle n'est pas réalisée par le producteur, est effectuée par celui qui va traiter l'archive. Dans le cadre du leak, et donc de la manipulation du document par des *personnes non-habilitées*, ce sont les lanceurs d'alerte, les plateformes de lanceurs d'alerte, les journalistes ou encore les internautes qui traitent l'archive, qui s'interrogent et formulent ces fameuses questions ou hypothèses.

La question de l'alerte laisse en effet inévitablement entrevoir la notion de preuve. Alerter sur un fait, une action sans fournir de justification, sans présenter de support qui appuie les propos avancés, équivaudrait à l'histoire de l'enfant qui criait au loup : l'alerte serait aisément réfutable et ne serait pas prise au sérieux. L'archive produite dans le cadre des activités de l'organisme public ou privé qui est concerné par l'alerte revêt alors un aspect capital pour le lanceur d'alerte, puisqu'elle atteste que l'action qui a été réalisée possédait un caractère illégal, immoral ou illégitime. Elle n'en est pas pour autant une « trace destinée à justifier sa bonne exécution vis-à-vis d'un tiers » puisqu'elle est « la cible même de l'opération ». Chabin ajoute :

> « Les 'Archives Snowden' constituent le bénéfice de la mission d'information que s'est attribuée Snowden ; cette documentation est son tableau de chasse, et non ses archives à proprement parler » (Chabin, 2016)

De manière générale, le leak ne met pas à jour des pratiques nouvelles et inédites, mais vient au contraire corroborer ce que l'on sait déjà ou, éventuellement, ce dont on se doutait. L'archive recueillie devient une preuve en la défaveur de son producteur, justifie la raison d'exister de l'alerte, et donc rend l'alerte recevable. Dans le cas de la fuite et/ou de l'alerte, l'archive ne sert pas de preuve à l'action relatée par son contenu, mais à l'action du lanceur d'alerte : c'est-à-dire qu'elle justifie la manipulation qu'elle-même a engendré. Elle prouve que l'alerte devait être lancée et, parfois même en dehors des sentiers réglementaires, que la fuite avait raison d'être. Le cas échéant, elle (*re-*)deviendra ensuite un élément de preuve pour un éventuel procès qui serait ordonné à l'encontre des producteurs de l'archive. L'arme se retourne contre son possesseur.

Toutefois, dans le cadre du leak, sa reconnaissance en tant que preuve peut être biaisée par la nécessité d'endiguer la fuite, d'étouffer le scandale de documents volés ou piratés, de condamner les leakers pour leurs actions illégales. Dans les cas d'Edward Snowden, Chelsea Manning ou Antoine Deltour, on peut noter que la violation de confidentialité entre l'employeur et l'employé peut aisément passer outre la gravité de l'alerte lancée et des faits divulgués. Dans d'autres cas, notamment là où les leakers sont restés anonymes, les accusations se concentrent sur l'alerte en elle-même et mènent à des procès engagés contre les organismes producteurs de l'archive – ou du moins directement concernés par celle-ci – et elle devient alors preuve de l'illégalité, l'immoralité ou l'illégitimité des actions de ces derniers.

Cette réappropriation du document donne la possibilité de faire preuve par l'archive ainsi que par le leak, soit à deux niveaux du cycle de vie du document. Car on parle bien d'un seul et même document qui, selon qu'il soit considéré comme archive ou comme leak possède une valeur probatoire qui ne vise pas à la même justification. On peut notamment expliquer cette dualité par les motivations qui impulsent les manipulateurs du document, ou bien par le fait que ces manipulateurs, justement, ne soient pas de même nature. D'une part, en effet, on trouve les producteurs du document, et de fait de l'archive ; d'autre part, le lanceur d'alerte n'est pas producteur de l'archive, mais du leak, c'est-à-dire qu'il collecte l'archive en lui offrant un nouveau cycle, un nouveau processus :

> « Il convient de faire remarquer que, si Edouard Snowden détient ces documents (une copie de ces documents car la NSA dispose toujours des données originales – il en eût été autrement peut-être dans un environnement papier…), donc si Snowden est le détenteur ces archives, il n'en est pas le

> producteur et il n'en est pas le propriétaire au sens juridique du terme » (Chabin, 2016)

Il détient mais n'est pas propriétaire, il s'approprie mais n'est pas producteur, mais il possède des documents qui peuvent aboutir aux condamnations des organismes et personnes que le contenu de ces archives finit par accuser. Qu'en est-il pour les journalistes, par exemple ? il s'avère qu'en réalité, dès lors que le document premier est manipulé par les journalistes, ou les leakers, ou même encore les lanceurs d'alerte, la contextualisation qui en résulte et de fait son information et son support deviennent l'archive de leurs détenteurs. Le leak tel qu'il est récolté n'existe plus en tant que tel ; ou du moins, il ne s'agit que d'une *source* et non pas d'une production de celui qui divulgue. Mais le *nouveau* document qui en découle est une production unique qui appartient, non pas aux premiers producteurs du document, mais aux producteurs de l'information contextualisée qui en résulte.

Le leak, archive ou *archive* ?

> « Les « Archives Snowden » sont tout bêtement une transposition journalistique de l'expression anglaise « Snowden Archive », expression où le mot Archive, au singulier, signifie « fonds documentaire », « collection de documents », « accumulation de témoignages écrits », avec une acception bien plus large que les archives organiques d'un producteur-propriétaire, avec une primauté donnée au lieu de conservation et à la constitution délibérée d'une collection, bien au-delà de la notion archivistique de « fonds d'archives ». » (Chabin, 2016)

Cette affirmation de Marie-Anne Chabin pourrait clôturer ici le débat et, pourtant, elle nous invite à aller plus en profondeur dans la reconnaissance du leak en tant qu'archive. Car la définition de l'archive – ou toute autre définition – se construit et évolue en fonction de l'utilisation qui en est faite. La langue s'amusant d'une utilisation parfois incongrue de ses mots, on pourrait croire que la notion d'archive ici n'a pas lieu d'être ; il faut convenir tout de même que le leak est bien plus qu'empreint de la notion d'archive. Car, avouons-le sans hésiter : le leak ne peut exister sans archive. Sa raison d'être ne tient qu'en la production de documents d'activité par les organismes dont il est issu, et qu'est-ce donc que l'archive sinon *tout document produit ou reçu dans le cadre des activités de l'organisme* ?

Les médias usent et abusent notamment du terme d'archive pour désigner les documents et données fournies par Edward Snowden. Celles-ci dénotent d'autres fuites alimentant l'histoire numérique de ces dernières années par la grande précision de leur récolte. En vérité, Snowden ne s'est pas contenté de faire une collecte à large spectre des documents auxquels il avait accès. Il se distingue d'autres lanceurs d'alerte par l'attention qu'il porte à chaque lot divulgué, par son implication dans le traitement de chacune des données qu'il fournit. Il est pleinement acteur du processus appliqué aux fichiers qu'il apporte, et de fait travaille avec les médias chargés de traiter les documents fournis. On parle, bien évidemment, de sa « collection de documents », de son « fonds documentaire », et non pas de son fonds d'archive en tant qu'employé de la NSA ou de la CIA.

Mais Edward Snowden, aujourd'hui, est surtout considéré comme un lanceur d'alerte et ce, pas seulement parce qu'il a divulgué à un moment donné les données qu'il avait recueilli ; il est lanceur d'alerte parce qu'il continue, des années après, à alerter, à suivre les documents qu'il a fourni, à aider à l'interprétation et à la découverte de nouveaux scandales. Il se place en tant que fervent défenseur de la

transparence numérique et du droit de savoir, et est primairement désigné, par la grande majorité des journaux que nous avons pu consulter, comme *lanceur d'alerte*. De fait, ce titre s'apparente à son activité actuelle ; non pas parce que le lancement de l'alerte pourrait correspondre à un métier ; mais parce que son activité principale consiste, encore aujourd'hui, à alerter. La production des documents qui ont fuité ne peut lui être attribuée, certes ; mais le leak, en tant qu'élément, est produit par le lanceur d'alerte. C'est l'action de fuite, réalisée par les lanceurs d'alerte, qui *produisent* le leak, qui créé des lots de documents volés, fuités, traités et enfin divulgués. Edward Snowden est donc producteur d'un ensemble de leaks dans le cadre de son activité de lanceur d'alerte.

La même logique peut être appliquée au cas de Julien Assange. Ce dernier, fondateur de Wikileaks, n'est pas accusé de produire la fuite mais de faciliter sa diffusion, sa divulgation. En tant que *leaker* – il semble difficile de lui attribuer le titre de lanceur d'alerte puisqu'il est du côté de ceux qui reçoivent les lots fournis par les lanceurs d'alerte – Assange met en place les conditions de bonne réception, de transmission des documents qu'on cherche à lui transmettre. *Dans le cadre de son activité* – là encore la notion est importante – il reçoit des documents – via les lanceurs d'alerte donc – et produit, à partir de ceux-ci, des bases de données et des documents de contextualisation qui permettent de comprendre les documents transmis. De même, les moteurs de recherche et les bases d'indexation construits pour accueillir les documents sont des productions des équipes de Wikileaks et d'Assange lui-même. Si ces derniers ne peuvent être directement considérés comme des leaks, ils n'en sont pas moins les supports et les outils de lecture qui constituent le leak.

En effet, un leak ne se contente pas de résulter d'une fuite documentaire. Le document volé seul ne suffit pas à créer un leak. Ce dernier est une production d'un ensemble d'éléments liés qui le rendent compréhensibles : le fichier, les données qui y sont associées, l'outil de classement, d'indexation, parfois de conservation, les documents qui le contextualisent, les canaux de diffusion qu'il emprunte, etc. Ces éléments, assemblés, sont ce qui constituent le leak. Un leak n'est pas un *leak* tant qu'il n'est pas divulgué, transmis, traités, contextualisé, communiqué. L'entièreté du processus (quel que soit le processus emprunté par le document) est nécessaire à l'attribution du terme *leak* à l'élément désigné.

Ainsi, lorsque les médias francophones usent et abusent du terme de *leak* et l'associent à celui d'*archive*, il n'y pas tant une simple transposition de l'*archive* anglophone. La possession de cette fameuse archive n'est pas attribuée au hasard. D'une part, ils peuvent utiliser le terme d'archive pour désigner les documents de Mossack Fonseca – on parle alors des *archives de Mossack Fonseca* ; de l'autre, ils peuvent attribuer l'archive directement au lanceur d'alerte – le cas célèbre des *Archives de Snowden*. La distinction est faite pour deux raisons :

- La majorité des documents attribués comme les archives ***de l'organisme à qui l'on a volé les archives*** le sont parce que le lanceur d'alerte n'est pas défini. Ainsi les leaks divulgués par des anonymes, ou bien par des groupes de hackers ou d'activistes qui n'ont pas d'acteur principal connu sont-ils toujours présentés comme les archives de l'organisme qui subit la fuite ;
- Les documents attribués comme les archives ***du lanceur d'alerte/du leaker*** le sont parce que ce dernier s'implique pleinement dans le traitement et le contextualisation des documents qu'il transmet et, de fait, devient producteur (ou receveur actif) des données en question.

La dénomination de l'archive, sans doute, tient en la nécessité de redéfinir ses frontières. Pourtant, il convient de noter que le statut du lanceur d'alerte ou du leaker, bien au-delà des textes réglementaires et des définitions difficilement offertes par les pouvoirs publics va bien au-delà de la simple appellation d'une forme de criminalité ou de *vigilantisme*. L'expansion numérique et, de fait, l'expansion des concepts inhérents à la création d'internet et du web, offre une palette d'activités – considérées par leurs actifs comme constitutifs de leurs métiers – en marge des standards professionnels que l'on connait si bien. Tout comme les activités de piratages, à un certain point, laissent voir l'apparition d'une nouvelle catégorie de professionnels de l'informatique nommés *hackers professionnels* – au point où ils sont aujourd'hui embauchés dans certaines multinationales spécialisées dans le numérique qui font appel à eux, non pas pour leur sécurité informatique, mais bien pour leurs compétences en piratage informatique ; tout comme eux, disions-nous, les lanceurs d'alerte se voient aujourd'hui supportés par des organismes divers, des structures complexes et reconnues, et l'on voit quelques-uns d'entre eux se concentrer uniquement sur le traitement des leaks dont ils sont la source.

Certes, on peut concevoir qu'il soit difficile pour le lanceur d'alerte de continuer à alerter lorsque son identité est divulguée, et si quelques-uns y parviennent grâce, notamment, à la quantité astronomique de leaks qu'ils ont produite, ils ne sont pas légion. Toutefois, dans la même lignée, et tout aussi légitime, le leaker se présente comme un acteur capital dans la conception du leak. Ces médiateurs entre les premiers producteurs du leak (les lanceurs d'alerte, donc) et les médias (ou parfois directement le public) ne se présentent plus comme des amateurs-politiques, mais comme de véritables professionnels, producteurs d'une forme d'archive inattendue.

Et les œuvres de l'esprit ?

> « […] le droit de copie a depuis longtemps cessé d'être l'apanage exclusif des maisons de disques, des géants de la distribution cinématographique, des éditeurs de presse et de multimédias » (Casilli, 2010, p. 45)

Il serait prétentieux et dangereux de croire que les notions abordées précédemment ne concernent que les documents d'activité ou relatifs à l'activité d'une entreprise ou d'une structure quelle que soit sa forme. Profondément liées au numérique et à l'Internet, ces problématiques touchent également de plein fouet les œuvres de l'esprit et les aspects patrimoniaux et culturels des données et documents.

Sans trop s'avancer sur ce terrain, qui n'est pas au centre de la problématique de ce mémoire, il est intéressant de noter que ces pratiques revêtent parfois une dimension militante qui va bien au-delà du simple acte de piratage. Ne nous y trompons pas : il serait difficile et peu légitime de parler de lanceurs d'alerte là où la notion de secret et/ou d'urgence n'a que peu – voire pas – de sens. Cependant, le masque de *justiciers* endossé par certains défenseurs du libre accès au patrimoine ou à la culture, diffusant des œuvres de l'esprit librement et gratuitement après les avoir obtenu illégalement, mérite qu'on s'y attarde.

Dans l'univers scientifique autant que dans celui de la culture, la question numérique agite les foules depuis plus de 20 ans, alternant entre les défenseurs d'un accès libre et désintéressé aux informations et la volonté de transposer numériquement les réglementations existantes. Étiqueté d'abusif voire de *préhistorique*, le fonctionnement actuel de la diffusion des documents multimédias des industries du livre, du cinéma, de la musique pour ne citer qu'elles, est

régulièrement attaqué sur son incapacité à accepter et à intégrer une économique numérique qui ne peut pas lui être bénéfique. Car au-delà des questions de droit d'auteur, de copie, de paternité, la grande question agitant le domaine des œuvres de l'esprit tient bien en la problématique économique.

> « Le secret garantit ainsi de la rareté. La propriété des œuvres numériques – qui peuvent être copiées à distance, à l'identique et sans frais – engendre en réaction des dispositifs législatifs ou techniques destinés à en maîtriser la prolifération gratuite. » (Huyghe, 2013b, paragr. 12)

Nous expliquions en quoi l'Internet ne pouvait être considéré comme une zone de non-droit, bien que son aspect anarchique puisse inviter à ce genre de croyance. *L'idéologie* même de la création d'Internet, et sa prise en main par les internautes, traduit la volonté d'une économie et une politique aux antipodes des économies et politiques des sociétés soumises à la notion de territorialité, et ce aux prémices même de son existence. Il ne faut tout de même pas croire que l'ouverture des données est la solution à tout ; il y a des données qui s'avèrent plus que délicates à diffuser, voire dangereuses (politique intérieure, défense, etc.). De plus, la diffusion des données nécessite des moyens, matériels et humains, non négligeables qui vont bien au-delà du stockage ou de la conservation de celles-ci. Toutefois, la représentation du leak lorsqu'il concerne des œuvres de l'esprit peut parfois s'apparenter à une forme de défense des droits de communication et de libre circulation de l'information et de la culture, qui s'inscrit directement, bien qu'en parallèle dans les revendications des lanceurs d'alerte.

CONCLUSION

« La technologie change les facilités, les finalités et les fragilités du secret comme elle bouleverse celles de la communication. » (Huyghe, 2013b)

Et l'espace numérique nous prouve encore combien les frontières d'hier ne peuvent être les frontières d'aujourd'hui. D'abord, parce que le numérique n'est pas un monstre intenable tapi sous les bureaux des grandes entreprises ou des gouvernements prêt à détruire toute la société économique et politique construite depuis de si nombreuses années ; de même qu'il n'est pas celui prêt à écraser toute volonté de liberté d'expression ou de savoir de la part de ses internautes. S'il possède une nature qui laisse plus volontiers place à l'autorégulation et à l'autoréglementation qu'au calque de l'architecture de nos sociétés actuelles, il n'en est pas moins un outil manipulé par ceux qui y évoluent. Ceux-là même, les internautes, contribuent à perpétuer l'esprit des pionniers de l'internet, des hippies aux cyberpunks, des informaticiens militaires aux amateurs du codage.

Espace d'expression par excellence, le web modifie tant notre rapport à la machine qu'il redéfinit, même, notre rapport à l'autre et de fait, notre propre identité. L'internaute n'est pas un homme comme les autres ; il s'offre des facettes d'identité qui augmentent autant qu'elles répriment les grandes aspirations de l'être humain. Là où certains se sentent effrayés de la vitesse du buzz et la violence des rapports entre communautés, d'autres s'emparent de ces identités pour laisser libre cours à des passions inédites et incongrues. De plus, là où l'on tentait de définir l'humain en une identité pleine et entière, on peint des internautes aux capacités techniques, technologiques, intellectuelles démultipliées par les communautés dans lesquelles ils évoluent. L'intelligence collective, qu'elle soit mythe ou réalité, nourrit un engouement qui construit, qui produit, et qui laisse des traces.

Cette mémoire, d'ailleurs, n'a jamais été aussi pérenne et fugace dans le même temps. D'une part, elle se diversifie tant qu'elle assure sa préservation par la multitude de point d'accueil : internet n'oublie pas, parce qu'il disperse et multiplie ses sources. D'autre part, elle semble aussi fragile que l'oralité, effaçant, perdant, réinventant à outrance tout élément à sa portée. Cette dualité forme une instabilité difficilement contrôlable qui met en péril l'information, le document et l'archive tels qu'on les connaissait, tels qu'on les maîtrisait. L'authenticité d'une information n'a jamais été aussi aisément démontrable et, dans le même temps, difficilement assurable. Le document n'a jamais été aussi multiple, et en même temps aussi creux et vide lorsqu'il est pris seul, sans sa multiplicité de données qui lui est associée. Les capacités de conservation, de traçabilité, d'intégrité n'ont jamais été aussi diverses et variées et, dans le même temps, le devenir de l'archive n'a jamais été aussi flou.

Pour certains, la réponse à ses problématiques angoissantes tant elles changent notre perception du monde se trouvent du côté du Libre, avec un grand L. Là où l'on tentait de verrouiller, il faut ouvrir ; là où l'on taisait, il faut dire. Ce n'est pas que la communication qu'internet interroge, mais toute la société : la politique, l'économie, l'autre et sa communauté, la sécurité et la défense, l'intimité et la vie privée. Internet, objet *inverrouillable* par excellence, requiert la transparence sous toutes ses formes, la libre circulation de ses ressources, la liberté d'expression, de savoir et ce, parfois, quel qu'en soit le prix. Le secret n'a plus lieu d'être ; alors le secret est, plus que jamais. Peut-être peut-on penser, en effet, que si les leaks n'ont

jamais été si nombreux, c'est qu'il n'y a jamais eu tant à dévoiler. Quoi qu'il en soit, l'affirmation suivante semble confirmée par ce que l'on sait d'internet : parce qu'il n'y jamais eu autant de connaissances à la portée de tous, il n'y a jamais eu un si grand besoin de tout savoir.

La fuite documentaire ou d'information, alors, n'a jamais été aussi légitime. Soutien de l'alerte, parfois, elle se redéfinit en fonction des outils numériques qu'elle a désormais à disposition et, de fait, se présente beaucoup plus massive qu'on ne l'a connue jusqu'à présent. Une fuite, aujourd'hui, ne concerne plus une poignée de documents glissés dans une mallette à l'occasion d'une journée calme dans les bureaux de son entreprise ; elle concerne des centaines de mégaoctets, des dizaines de téraoctets copiés sur des disques durs, transmis via des plateformes garantissant l'anonymat, et offre à sa diffusion une véritable dimension de *déluge*. Le lanceur d'alerte, vecteur primaire de cette fuite (et de l'alerte, bien évidemment) ne revêt plus seulement le masque du salarié ; leaker, hacker, lanceur d'alerte, il est multiple, divers, difficilement définis par les textes réglementaires qui tentent de le rattraper pour mieux l'encapsuler.

C'est que les scandales qui jalonnent ces vingt dernières années ne se ressemblent jamais vraiment. Parfois piratages dont les motivations s'embellissent de *bonne foi*, parfois véritables fuites effectuées par seule volonté de nuire, le scandale est protéiforme et sa ligne de conduite, insaisissable. Loin d'être l'œuvre unique d'un seul individu, il est désormais porté et soutenu par des communautés imposantes et sur-armées. De la défense de la transparence à la liberté d'expression, de la méfiance envers les médias traditionnels mêlée à la déviance envers le système de justice de nos pays, ces organisations plus ou moins officielles, plus ou moins légitimes, peuvent difficilement être classées selon leurs véritables motivations. Quoi qu'il en soit, elles s'avèrent être un appui indispensable à l'alerte de notre ère, couvant les lanceurs tant dans la défense de leurs droits que dans le traitement et la diffusion de leurs alertes.

Difficile de considérer, en effet, que les leakers tels que Julien Assange ou les lanceurs d'alerte tels qu'Edward Snowden, Chelsea Manning ou Antoine Deltour ne sont que des délateurs. Leurs motivations, si elles ne sont peut-être pas pures – ce que nous ne jugeons pas ici – sont suffisamment saluées par une partie de l'opinion publique pour paraître légitimes. De plus, ils contribuent, associés à ces organisations de *leakers*, à collecter, conserver, classer et communiquer les leaks dont ils disposent, soit à constituer une mémoire de l'alerte indispensable à la construction de l'expérience. Leurs travaux, attentifs et minutieux, offrent même la possibilité de construire une typologie du leak : car le leak, tout comme le lanceur d'alerte, n'a rien d'unique. Loin d'être défini par les motivations qui l'impulsent, il se distingue par les processus qu'il emprunte, comme de multiples chemins de construction d'une information, ou plutôt de divulgation de cette information. De l'alerte réglementée à celle qui viole les lois, du piratage au leaking, le leak, dans son acception la plus large, revêt de multiples aspects.

Et à ses multiples formes s'ajoutent ses multiples représentations. Preuve, d'abord, parce qu'il prouve ce que l'alerte ose dire, ensuite parce qu'il justifie l'existence même de l'alerte, le leak est une preuve indiscutable pour le lanceur d'alerte. Aussi, parce qu'avant d'être leak il est archive, document d'activité d'un organisme public ou privé, il est la preuve de la production de ses premiers détenteurs, la justification des actions menées. Qu'on le nomme archive par simple transposition de l'anglais n'importe que peu ; il est au moins archive avant d'être leak.

Mais mérite-t-il pour autant d'être nommé archive ? la question est plus complexe qu'elle n'y paraît. La transformation du document en tant que leak ne lui ôte pas son existence en tant qu'archive ; il reste en effet un document d'archives provenant des fonds d'archives des producteurs volés. Toutefois, le document qui fuite n'est pas l'archive des leakers ou des lanceurs d'alerte. Mais parce qu'ils produisent en contextualisant, en analysant et en diffusant, les lanceurs d'alerte, leakers et autres acteurs du leak se voient pourvus du statut de producteurs d'un élément : le leak, justement. Et parce qu'il se retrouve en tant que document produit dans le cadre des activités de ceux qui le traitent et le diffusent, le leak ose se présenter comme une forme d'archive.

Cette hypothèse offre un panel de questions et de réflexions à explorer particulièrement riche, d'autant plus qu'elle se place précisément dans l'univers numérique qui ne cesse – et ne cessera sans doute pas avant longtemps – de faire parler de lui. Car si le leak est une archive, ses producteurs, peut-être, pourraient se voir affubler du titre d'archiviste. Pire ! le processus qui construit le leak et le fait exister laissera certains se poser la question de savoir s'il n'y a pas là, tout de même, une forme d'archivage.

SOURCES

CNIL. Consulté à l'adresse : https://www.cnil.fr/

Cryptome. Consulté à l'adresse : https://cryptome.org/cryptome-2016-1996.htm

Documents Wikileaks - Dossier Le Monde. Consulté le 2 juin 2017, à l'adresse : http://www.lemonde.fr/documents-wikileaks/

Edward Snowden et la surveillance des données. Dossier ARTE. Consulté le 2 juin 2017, à l'adresse : http://info.arte.tv/fr/edward-snowden-et-la-surveillance-des-donnees

EUDES, Y., & CHECOLA, L. (2016, octobre 26). *Comment fonctionne Wikileaks ? - Chat avec les internautes.* Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/technologies/chat/2010/10/24/comment-fonctionne-wikileaks_1430107_651865.html#IIj36EsgTdEoDpVx.99

GlobalLeaks. Consulté à l'adresse : https://www.globaleaks.org/fr/

HASSAN, G. (2017, mars 15). *États-Unis. Une feuille d'impôts de Trump révélée au public.* Consulté 2 juin 2017, à l'adresse : http://www.courrierinternational.com/article/etats-unis-une-feuille-dimpots-de-trump-revelee-au-public

HERMES - Center for Transparency and Digital Human Rights. Consulté à l'adresse : https://www.hermescenter.org/

International Consortium of Investigative Journalists. Consulté à l'adresse : http://www.icij.org/

Offshore Leaks Database - ICIJ. Consulté à l'adresse : https://offshoreleaks.cloud.icij.org/

Open Archives Initiative. Consulté à l'adresse : https://www.openarchives.org/

POUCHARD, A., LAURENT, S., DAMGÉ, M., & BRETEAU, P. (2015, juin 25). *Des « Pentagon papers » aux « Frenchleaks », 40 ans de « fuites ».* Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/evasion-fiscale/article/2015/06/25/des-pentagon-papers-aux-frenchleaks-quarante-ans-de-fuites-et-de-revelations_4661272_4862750.html

SNCF Open Data. Consulté à l'adresse : https://data.sncf.com/

The Panama Papers - ICIJ. Consulté à l'adresse : https://panamapapers.icij.org/

The Santa Fe Convention by the Open Archives Initiative. (2000, février 15). Consulté le 2 juin 2017, à l'adresse : http://www.openarchives.org/sfc/sfc_entry.htm

The Top 100 Most Damaging WikiLeaks. Consulté à l'adresse : http://www.mostdamagingwikileaks.com/

UNTERSINGER, M. (2015, août 5). *3 questions pour comprendre le scandale Netzpolitik en Allemagne.* Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/pixels/article/2015/08/05/allemagne-le-scandale-netzpolitik-en-trois-questions_4713051_4408996.html

War Diaries - Wikileaks. Consulté à l'adresse : https://wardiaries.wikileaks.org/

WikiLeaks. Consulté à l'adresse : https://wikileaks.org/

WikiLeaks : naviguez dans les mémos diplomatiques. Consulté 2 juin 2017, à l'adresse : http://www.lemonde.fr/international/visuel/2010/12/06/wikileaks-lire-les-memos-diplomatiques_1449709_3210.html

World Wide Web Consortium (W3C). Consulté à l'adresse : https://www.w3.org/

Football Leaks

Ce que l'on sait des « Football Leaks ». (2016, décembre 3). Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/football/article/2016/12/03/ce-que-l-on-sait-des-football-leaks_5042921_1616938.html

*Football Leaks : «El Mundo» a été sommé par la justice de ne pas publier. (*2016, décembre 3). Consulté 2 juin 2017, à l'adresse : http://www.rfi.fr/europe/20161203-football-leaks-el-mundo-justice-football-ronaldo-seen-ferrero-eic

« Football Leaks » : si vous avez manqué le début. (2016, décembre 9). Consulté 2 juin 2017, à l'adresse : http://www.liberation.fr/sports/2016/12/09/football-leaks-si-vous-avez-manque-le-debut_1534098

Football Leaks. Un juge espagnol demande l'interdiction de la publication dans toute la presse européenne. (2016, décembre 6). Consulté 2 juin 2017, à l'adresse : http://www.courrierinternational.com/article/football-leaks-un-juge-espagnol-demande-linterdiction-de-la-publication-dans-toute-la-presse

LuxLeaks

GEOFFROY, R. (2016, décembre 12). *LuxLeaks : « Ce procès est un message envoyé contre les lanceurs d'alerte ».* Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/evasion-fiscale/article/2016/12/12/luxleaks-ce-proces-est-un-message-envoye-contre-les-lanceurs-d-alerte-et-les-journalistes_5047308_4862750.html

MICHEL, S. (2014, novembre 6). *LuxLeaks : 28 000 pages de documents secrets, 548 accords confidentiels.* Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/evasion-fiscale/article/2014/11/06/luxleaks-28-000-pages-de-documents-secrets-548-accords-confidentiels_4519428_4862750.html

Macron Leaks

EL IDRISSI, A. (2017, mai 10). *Que contiennent les « Macron Leaks » ?* France Culture.fr. Consulté à l'adresse : https://www.franceculture.fr/politique/que-contiennent-les-macron-leaks

MATHIOT, C., & PHILIPPIN, Y. (2017, mai 6). *« Ce leak est fait pour créer le chaos »*. Libération.fr. Consulté à l'adresse : http://www.liberation.fr/politiques/2017/05/06/ce-leak-est-fait-pour-creer-le-chaos_1567717

UNTERSINGER, M., & LELOUP, D. (2017, mai 11). *Qu'y a-t-il dans les « MacronLeaks » ?* Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/pixels/article/2017/05/11/qu-y-a-t-il-dans-les-macronleaks_5126397_4408996.html

WikiLeaks publie l'intégralité des « MacronLeaks ». (2017, juillet 31). Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/pixels/article/2017/07/31/wikileaks-publie-l-integralite-des-macronleaks_5167002_4408996.html

Offshore Leaks

Le public est prié d'enquêter sur les paradis fiscaux. (2013, juin 15). Le Matin. Consulté à l'adresse : www.lematin.ch/news/standard/public-prie-enqueter-paradis-fiscaux/story/12544731

« L'Offshore Leaks » ouvre sa base de données. (2013, juin 15). Le Matin. Consulté à l'adresse : www.lematin.ch/monde/offshore-leaks-ouvre-base-donnees/story/29327026

Offshoreleaks : la presse priée de collaborer avec la justice. (2013, avril 9). Consulté 2 juin 2017, à l'adresse : http://www.lalibre.be/economie/libre-entreprise/offshoreleaks-la-presse-priee-de-collaborer-avec-la-justice-51b7357be4b0de6db975b3fa

Panama Papers

Dossier « Panama Papers ». Consulté 2 juin 2017, à l'adresse : http://www.lessentiel.lu/fr/news/dossier/panamapapers/

Les « Panama Papers » en sept définitions. (2016, avril 4). Consulté 2 juin 2017, à l'adresse : http://www.liberation.fr/planete/2016/04/04/les-panama-papers-en-sept-definitions_1443872

VAUDANO, M., & BARUCH, J. (2016, avril 3). *« Panama papers » : comment « Le Monde » a travaillé sur plus de 11 millions de fichiers.* Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/panama-papers/article/2016/04/03/panama-papers-comment-le-monde-a-travaille-sur-plus-de-11-millions-de-fichiers_4894836_4890278.html

SarkoLeaks

CASSELY, J.-L. (2014, mars 5). *« SarkoLeaks » ou « Buissongate » : comment nommer un scandale ?* Consulté 2 juin 2017, à l'adresse : http://www.slate.fr/france/84235/sarkoleaks-buissongate-nommer-scandale

L'Elysée dément avoir effectué des recherches illégales dans les archives de Nicolas Sarkozy. (2013, octobre 30). Le Monde.fr. Consulté à l'adresse : http://www.lemonde.fr/politique/article/2013/10/30/l-elysee-dement-avoir-effectue-des-recherches-illegales-dans-les-archives-de-nicolas-sarkozy_3505550_823448.html

Recherches illégales dans les archives protégées de Sarkozy selon Valeurs actuelles, l'Elysée dément. (2013, octobre 30). Consulté 2 juin 2017, à l'adresse : http://www.leparisien.fr/lyon-69000/recherches-illegales-dans-les-archives-protegees-de-sarkozy-selon-valeurs-actuelles-l-elysee-dement-30-10-2013-3272301.php

Sony Leaks

AUERBACH, D. (2014, décembre 19). *Les responsables du piratage de Sony sont des cyberterroristes.* Consulté le 2 juin 2017, à l'adresse : http://www.slate.fr/story/95949/piratage-sony-terrorisme

Elodie. (2014, décembre 15). *Quand on la hack, Sony Pictures contre-atthack.* Consulté le 2 juin 2017, à l'adresse : http://www.journaldugeek.com/2014/12/15/hack-sony-pictures-contre-attaque/

PANFILI, R. (2015, novembre 23). *Un an après le piratage, la vie bouleversée des employés de Sony.* Consulté le 2 juin 2017, à l'adresse : http://www.slate.fr/story/110435/sony-pictures-piratage-chantage

PETERS, J. (2014, décembre 18*). Il faut continuer d'utiliser le piratage de Sony dans des articles.* Consulté le 2 juin 2017, à l'adresse : http://www.slate.fr/story/95929/continuer-utiliser-sony-leaks-medias

SZADKOWSKI, M. (2014, décembre 18). *Que sait-on des hackeurs de Sony Pictures ?* Consulté le 2 juin 2017, à l'adresse : http://www.lemonde.fr/pixels/article/2014/12/18/ce-que-l-on-sait-des-pirates-qui-ont-attaque-sony-et-conduit-a-deprogrammer-the-interview_4542475_4408996.html

BIBLIOGRAPHIE

ALERTE, LEAKING, HACKING

ABADIE, P. (2016). *Le salarié lanceur d'alerte aux États-Unis et en France : pour une articulation harmonieuse entre dissidence et loyauté.* La revue des droits de l'homme, (10). Consulté à l'adresse : https://revdh.revues.org/2649#tocto2n2

CARDON, D., & GRANJON, F. (2013). *Médiactivistes*. Presses de Sciences Po.

CHABIN, M.-A. (2016, décembre 11). *Réflexions sur les « Archives Snowden »*. Consulté le 15 mai 2017, à l'adresse : http://www.marieannechabin.fr/2016/12/reflexions-sur-les-archives-snowden/

CHATEAURAYNAUD, F. (2013). *Lanceur d'alerte.* Dictionnaire critique et interdisciplinaire de la participation. Consulté à l'adresse : http://www.dicopart.fr/fr/dico/lanceur-dalerte

COLEMAN, G. (2016). *Anonymous. Hacker, activiste, faussaire, mouchard, lanceur d'alerte*. Lux.

FOEGLE, J.-P. (2016). *Lanceur d'alerte ou « leaker » ? Réflexions critiques sur les enjeux d'une distinction.* La Revue des droits de l'homme. Revue du Centre de recherches et d'études sur les droits fondamentaux, (10). Consulté à l'adresse : https://revdh.revues.org/2367

GAILLARD, F. (2013). *Merci Assange*. Médium, 37-38(4-2014/1), 118-125.

HUYGHE, F.-B. (2013a). *Alerte en sept leçons*. Médium, 37-38(4-2014/1), 126-127.

KWOKA, M. B. (2014). *Leaking and Legitimacy*. UC Davis Law Review, 1387-1456.

LOVELUCK, B. (2016). *Le vigilantisme numérique, entre dénonciation et sanction. Auto-justice en ligne et agencements de la visibilité*. Politix, 115(3), 127-153.

MARSAC, L. (2015). *The Cablegate : l'Affaire diplomatique du XXIe siècle.* Les médias : approches géohistoriques et géopolitiques, (6-7). Consulté à l'adresse : http://rgh.univ-lorraine.fr/articles/view/55/The_Cablegate_l_Affaire_diplomatique_du_XXIe_siecle

MASUTTI, C. (2013). *Ingénieurs, hackers : naissance d'une culture*. Histoires et cultures du Libre. Des logiciels partagés aux licences échangées, 31-65.

POZEN, D. (2013). *The Leaky Leviathan: Why the Government Condemns and Condones Unlawful Disclosures of Information.* Harvard Law Review, 127, 512-635.

TRÉGUER, F. (2015). *Hackers vs States: Subversion, Repression and Resistance in the Online Public Sphere.* Droit et société, (91), 639-652.

ARCHIVES ET NUMÉRIQUE

ASSOCIATION DES ARCHIVISTES DE FRANCE (2012). *Abrégé d'archivistique. Principes et pratiques du métier d'archiviste* (3e édition revue et augmentée).

BANAT-BERGER, F. (2015). *De l'écrit à internet : comment archive-t-on l'immatériel ?* Pouvoirs, 153(2), 109-124.

CHABIN, M.-A. (1999). *Je pense donc j'archive. L'archive dans la société de l'information*. L'Harmattan.

CHABIN, M.-A. (2005, mai). *Le succès du mot « archives » dans les médias : une opportunité pour les archivistes !* Présenté aux 10èmes Journées d'archivistique catalanes et 20ème anniversaire de l'Association des Archivistes de Catalogne, Terrassa. Consulté à l'adresse : http://www.arxivers.com/index.php/documents/publicacions/revista-lligall-1/lligall-23-1/109-06-le-succes-du-mot-archives-dans-les-medias-une-opportunite-pour-les-archivistes-1/file

CHABIN, M.-A. (2013). *Archives*. Médium, 37-38(4-2014/1), 102-117.

CŒURÉ, S., & DUCLERT, V. (2011). *Les archives « mémoire du monde » : l'internationalisation des enjeux* (p. 93-104). Paris : La Découverte. Consulté à l'adresse : http://www.cairn.info/les-archives--9782707167811-p-93.htm

COHEN, É., & VERLAINE, J. (2013). *Le dépôt légal de l'internet français à la Bibliothèque nationale de France*. Sociétés & Représentations, 35(1), 209-218.

GUYON, C. (2015). *La pratique archivistique publique en France, entre adaptation et négociation. Expériences et réflexions d'une archiviste.* Les Cahiers du numérique, 11(2), 77-114.

MUSSOU, C. (2012). *Et le Web devint archive : enjeux et défis*. Le Temps des médias, 19(2), 259-266.

DONNÉES, DOCUMENTS, INFORMATIONS

BERNERS-LEE, T. (2006). *Linked Data - Design Issues*. Consulté à l'adresse : https://www.w3.org/DesignIssues/LinkedData.html

CHABIN, M.-A. (2004). *Document trace et document source. La technologie numérique change-telle la notion de document ?* Revue I3 - Information Interaction Intelligence, Cépaduès. Consulté à l'adresse : https://archivesic.ccsd.cnrs.fr/sic_00001020/document

ERTZSCHEID, O. (2009). *L'homme, un document comme les autres*. Hermès, La Revue, 53(1), 33-40.

GIRAUD, G. (2010). *Comment réagir face au vol ou à la perte d'informations ? La DLP comme point de départ d'une nouvelle approche en matière de protection de l'information*. Sécurité et stratégie, 4(2), 81-89.

MAURY, Y. (2013). *Classements et classifications comme problème anthropologique : entre savoir, pouvoir et ordre*. Hermès, La Revue, 66(2), 23-29.

MERZEAU, L. (2009). *Du signe à la trace : l'information sur mesure*. Hermès, La Revue, 53(1), 21-29.

RÉGIMBEAU, G. (2008). *Pour une typologie documentaire de l'information en art contemporain.* L'information dans les organisations : dynamique et complexité. Consulté à l'adresse : http://books.openedition.org/pufr/864

RÉGIMBEAU, G. (2013). *Classer, c'est penser*. Hermès, La Revue, 66(2), 16-17.

RENUCCI, F. (2013). *Les origines*. Médium, 37-38(4-2014/1), 139-152.

SALAÜN, J.-M. (2012). *Vu, Lu, Su. Les architectes de l'information face à l'oligopole du Web* (La Découverte). Paris.

HISTOIRES ET CULTURES D'INTERNET

CARDON, D. (2010). *La démocratie Internet. Promesses et limites* (Le Seuil). Paris.

CASILLI, A. (2010). *Les liaisons numériques. Vers une nouvelle sociabilité ?* (Le Seuil). Paris.

COLEMAN, G., & HILL, M. (2004). *How Free Became Open and Everything Else Under the Sun*. M/C Journal, 7. Consulté à l'adresse : http://journal.media-culture.org.au/0406/02_Coleman-Hill.php

GRANJON, F., & Cardon, D. (2005). *Mouvement altermondialiste et militantisme informationnel.* Consulté à l'adresse : https://archivesic.ccsd.cnrs.fr/sic_00001336

LOVELUCK, B. (2008). *Internet, vers la démocratie radicale ?* Le Débat, 151(4), 150-166.

LOVELUCK, B. (2015a). *Internet, une société contre l'État ? Libéralisme informationnel et économies politiques de l'auto-organisation en régime numérique.* Réseaux, *192*(4), 235-270.

LOVELUCK, B. (2015b). *Réseaux, libertés et contrôle : une généalogie politique d'internet* (Armand Colin). Paris.

PALOQUE-BERGES, C., & MASUTTI, C. (2013). *Introduction générale*. Histoires et cultures du Libre. Des logiciels partagés aux licences échangées.

POULLET, Y. (2003). *Légitimité démocratique versus autorégulation ?* Les Enjeux Juridiques de l'Internet.

SAJUS, B., CARDON, D., LEVREL, J., VATANT, B., BERMÈS, E., OURY, C., & SUSSAN, R. (2009). *Web 2.0, et après ? Critique et prospective*. Documentaliste-Sciences de l'Information, 46(1), 54-66.

SCHAFER, V. (2013). *(Pré-)histoire*. Histoires et cultures du Libre. Des logiciels partagés aux licences échangées, 3-29.

TRUDEL, P., ABRAN, F., BENYEKHLEF, K., & HEIN, S. (2003). *L'autoréglementation : fondements, formes et limites.* Les Enjeux Juridiques de l'Internet.

IDENTITÉS ET RÉSEAUX

ARNAUD, M. (2009). *Authentification, identification et tiers de confiance*. Hermès, La Revue, 53(1), 127-136.

BEAU, F., & DESEILLIGNY, O. (2009). *Une figure du double numérique : l'avatar. Entretien*. Hermès, La Revue, 53(1), 41-47.

CARDON, D. (2009). *L'identité comme stratégie relationnelle.* Hermès, La Revue, 53(1), 61-66.

COUTANT, A. (2011). *Des techniques de soi ambivalentes*. Hermès, La Revue, 59(1), 53-58.

ERTZSCHEID, O. (2013). *Qu'est-ce que l'identité numérique ? Enjeux, outils, méthodologies* (Open Edition Press). Marseille.

FLICHY, P. (2010). *Le sacre de l'amateur. Sociologie des passions ordinaires à l'ère numérique* (Le Seuil).

GEORGES, F. (2008). *L'identité numérique dans le web 2.0.* Le mensuel de l'Université, (27).

KESSOUS, E., & REY, B. (2009). *Économie numérique et vie privée*. Hermès, La Revue, 53(1), 49-54.

MERZEAU, L. (2013). *Partager ses secrets en public.* Médium, 37-38(4-2014/1), 153-172.

TISSERON, S. (2011). *Intimité et extimité*. Communications, 88(1), 83-91.

SOCIÉTÉ DU SECRET

CONESA, P. (2013). *Hypocrites démocraties*. Médium, 37-38(4-2014/1), 42-63.

DEBRAY, R. (2013). *Rien de nouveau ?* Médium, 37-38(4-2014/1), 258-268.

HUYGHE, F.-B. (2013b). *Ce que nous cache le numérique*. Médium, 37-38(4-2014/1), 26-40.

PLENEL, E. (2013). *Le droit de savoir* (Le Seuil). Paris.

SCHEHR, S. (2016). *Mutations du renseignement, métamorphoses de la trahison*. Hermès, La Revue, 76(3), 98-105.

SORIANO, P. (2013). *Le secret, c'est le médium.* Médium, 37-38(4-2014/1), 5-19.

VAN PUYVELDE, D. (2011). *Médias, responsabilité gouvernementale et secret d'État : l'affaire WikiLeaks.* Le Temps des médias, 16(1), 161-172.

TEXTES RÉGLEMENTAIRES

Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés :
https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068624&dateTexte=20080609

Loi n° 2009-669 du 12 juin 2009 favorisant la diffusion et la protection de la création sur internet :
https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000020735432&categorieLien=id

Loi n° 2013-1168 du 18 décembre 2013 relative à la programmation militaire pour les années 2014 à 2019 et portant diverses dispositions concernant la défense

et la sécurité nationale :
https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000028338825&categorieLien=id

LOI n° 2015-912 du 24 juillet 2015 relative au renseignement :
https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000030931899

Projet de loi relatif à la transparence, à la lutte contre la corruption et à la modernisation de la vie économique (texte définitif) :
http://www.assemblee-nationale.fr/14/ta/ta0830.asp

ANNEXES

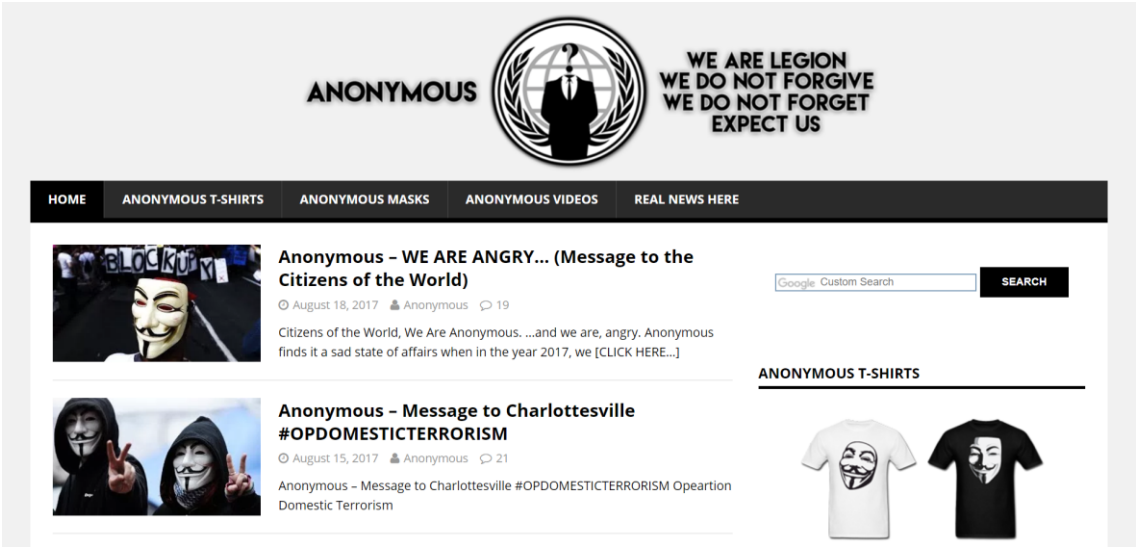
Table des annexes

PLATEFORMES UTILISÉES PAR/POUR LE LEAK

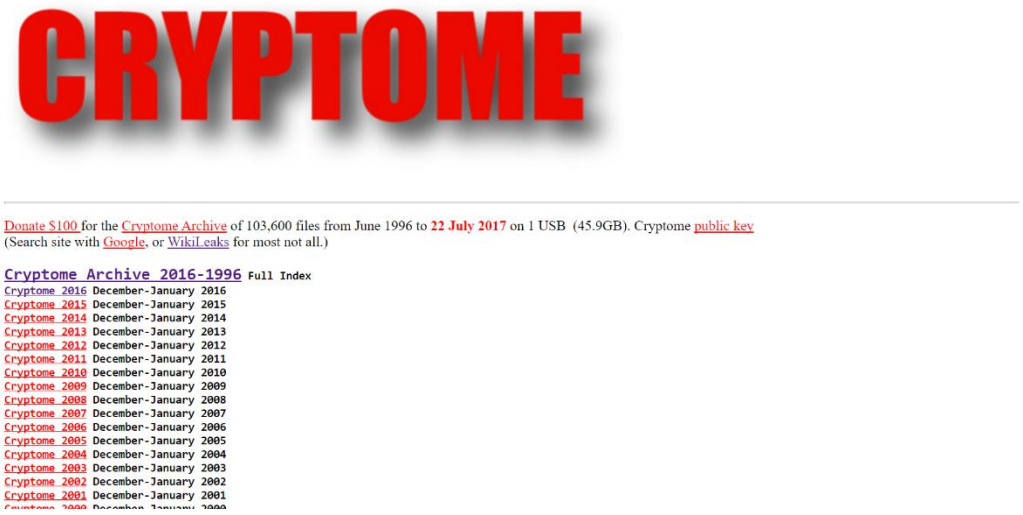
SITES DE LANCEURS D'ALERTE

Anonymous France

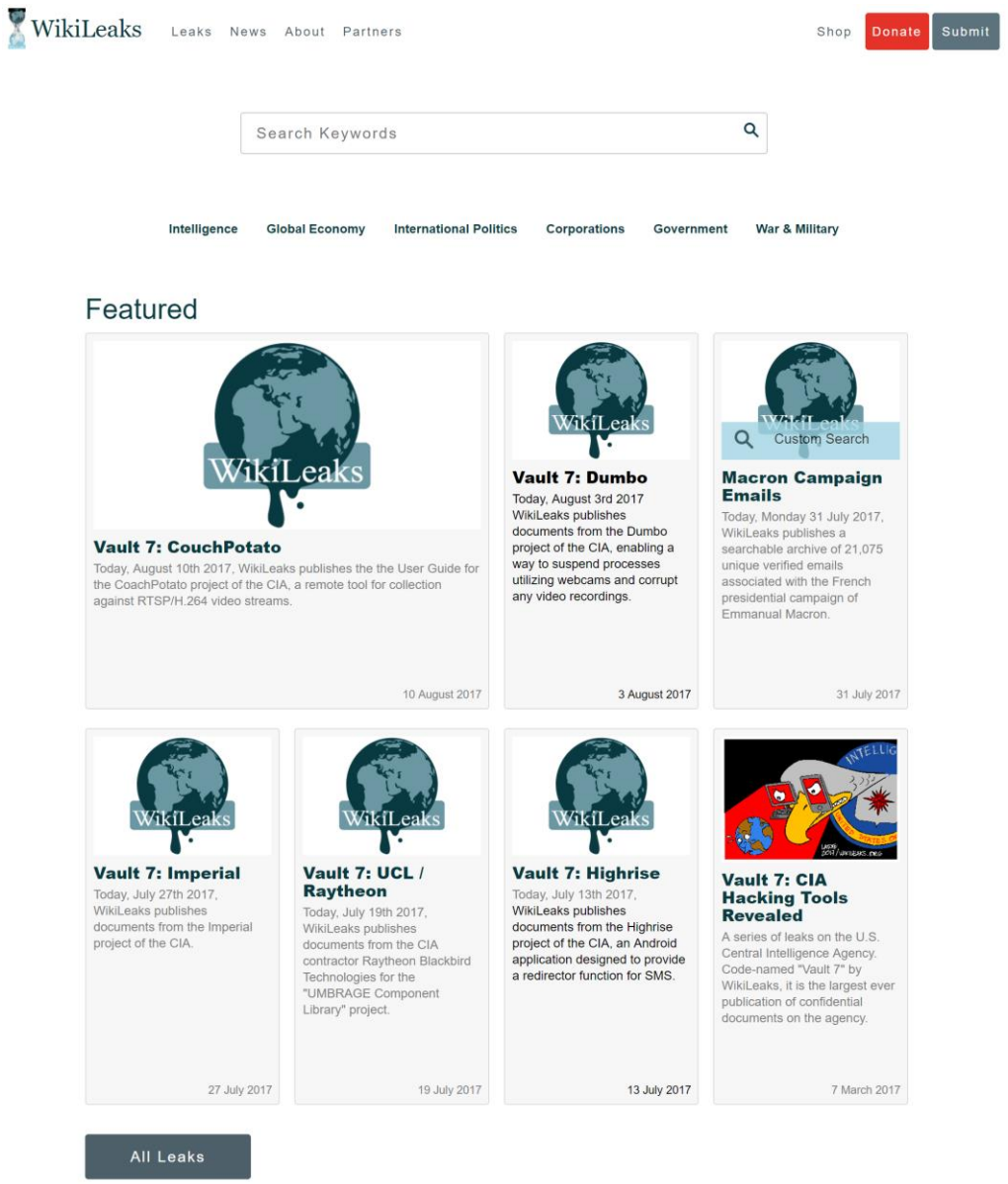
AnonOfficiel



Cryptome



Wikileaks



FrenchLeaks

FrenchLeaks

un site MEDIAPART.fr

NOS OBJECTIFS | NOUS SOUTENIR | CONTACT | ENVOYER UN DOCUMENT | ENGLISH VERSION

« Tout individu a droit à la liberté d'opinion et d'expression, ce qui implique (...) le droit de chercher, de recevoir et de répandre, sans considérations de frontières, les informations et les idées par quelque moyen d'expression que ce soit. » (Article 19 de la Déclaration universelle des droits de l'homme, adoptée à Paris le 10 décembre 1948.)

FrenchLeaks est un site dédié à la diffusion de documents d'intérêt public concernant notamment la France et l'Europe. Edité par le journal d'information en ligne Mediapart, il est au service du droit à l'information et du débat démocratique, dans une indépendance totale vis-à-vis des pouvoirs politiques et économiques.

FrenchLeaks est un outil documentaire et un instrument d'alerte. D'une part, il met à la libre disposition du public des documents ayant fait l'objet d'investigations des journalistes de Mediapart. D'autre part, il permet à des sources de nous transmettre, en toute sécurité et confidentialité, des documents d'intérêt public qui seront mis en ligne après une enquête préalable répondant aux règles professionnelles du journalisme.

Dossier mis en ligne le 11.09.2013 — **L'argent Libyen de Nicolas Sarkozy**

Nicolas Sarkozy a-t-il bénéficié du financement du régime libyen de Mouammar Kadhafi lors de sa campagne électorale victorieuse de 2007 ? Les documents que nous avons publiés le suggèrent. De même que les déclarations des proches de l'ancien dictateur libyen.

Dossier mis en ligne le 11.09.2013 — **Le compte suisse de Jérôme Cahuzac**

Après quatre mois de démentis acharnés, l'ancien ministre du budget est passé aux aveux devant les juges dans l'affaire de son compte suisse le 2 avril 2013. Notre enquête a permis de montrer que Jérôme Cahuzac possédait un compte en suisse depuis les années 90. Mais pas seulement. Voici les principaux documents qui ont appuyé nos investigations.

Dossier mis en ligne le 25.08.2011 — **Les documents Takieddine**

Depuis la mi-juillet 2011, Mediapart a commencé la publication d'une vaste enquête sur le marchand d'armes Ziad Takieddine, principal suspect dans le volet financier de l'affaire Karachi. Ses liens avec le clan Sarkozy, son rôle de diplomate occulte...

Dossier mis en ligne le 01.06.2011 — **L'affaire des quotas à la Fédération française de football**

Plusieurs dirigeants de la Direction technique nationale (DTN) de la Fédération française de football (FFF), dont le sélectionneur des Bleus en personne, Laurent Blanc, ont approuvé dans le plus grand secret, fin 2010, le principe de quotas discriminatoires officieux dans les centres de formation de la fédération, les écoles de foot du pays, selon une enquête de Mediapart.

Dossier mis en ligne le 01.06.2011 — **L'affaire Tapie/Lagarde**

La ministre de l'Economie, Christine Lagarde, est soupçonnée d'abus de pouvoir dans le cadre du règlement du litige opposant Bernard Tapie à l'Etat. Mediapart suit le dossier depuis le début. Nos documents.

Dossier mis en ligne le 29.04.2011 — **France Télécom-Orange : la lettre de Rémy L.**

En septembre 2009, alors que l'entreprise fait la une des médias à cause d'une série de suicides, Rémy L., l'employé de France Télécom-Orange de 57 ans qui s'est suicidé le 26 mars, avait envoyé un courrier de six pages signé de sa main à la direction du groupe. Mediapart s'est procuré ce document.

Dossier mis en ligne le 19.04.2011 — **L'ordonnance de renvoi de l'affaire de la Fondation Hamon**

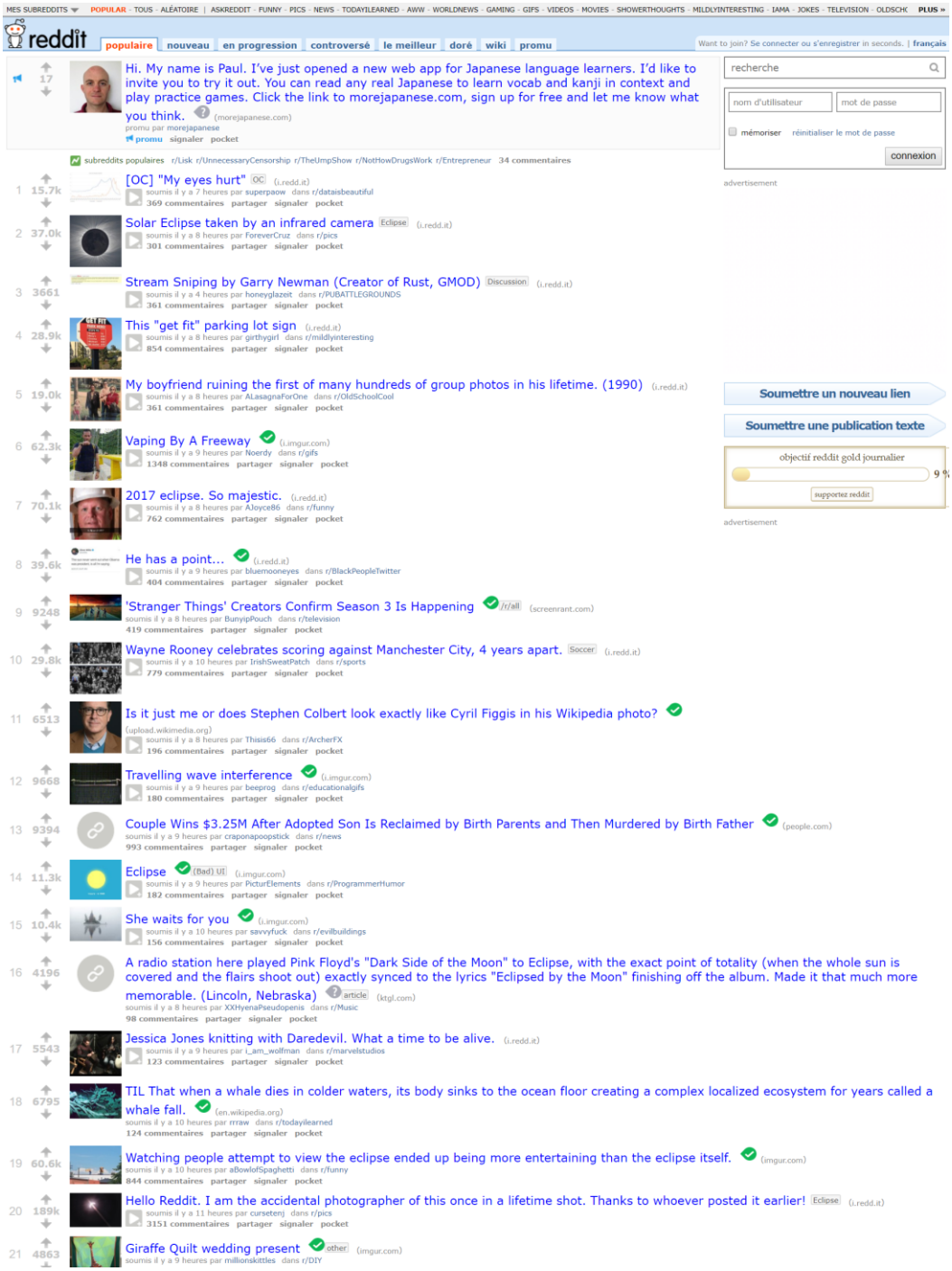
Charles Pasqua et André Santini ont été renvoyés en correctionnelle le 6 avril 2011 dans l'affaire de la Foudation Hamon, un ruineux projet de musée d'art contemporain dans le département des Hauts-de-Seine. Voici l'ordonnance de renvoi devant le tribunal correctionnel signée par la juge d'instruction du tribunal de grande instance de Versailles Nathalie Andreassian.

Dossier mis en ligne le 13.04.2011 — **L'affaire des moines de Tibéhirine**

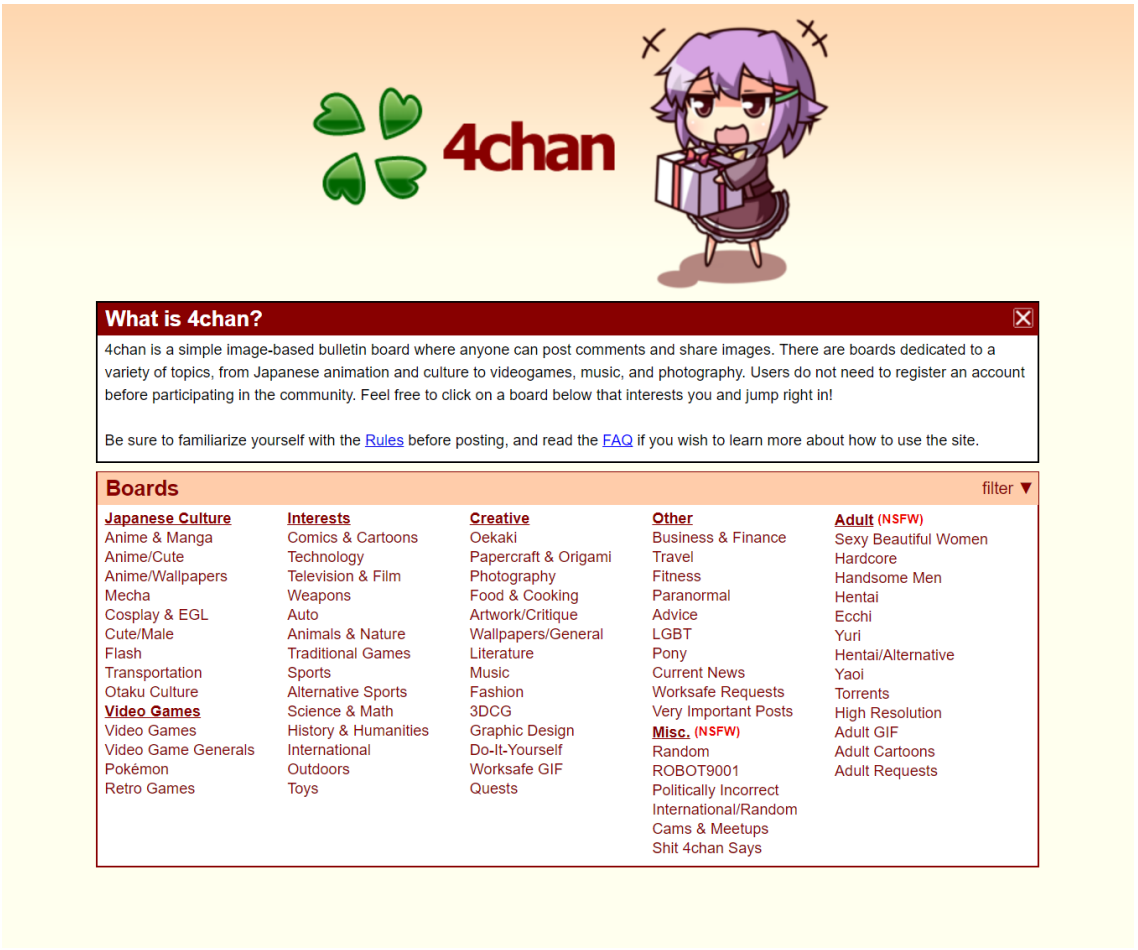
Dans la nuit du 26 au 27 mars 1996, sept moines trappistes du Monastère de Tibhirine, en Algérie, sont enlevés lors de la guerre

FORUMS

Reddit



4chan



SCANDALES

NOM	SUJET	ANNEE	POIDS	NBRE DE DOCS	TYPE DE DOC	NATURE DES DOCS	PERIODE	PROVENANCE DES ARCHIVES	MIS EN CAUSE	NOM(S) LANCEUR(S) D'ALERTE	STATUT LANCEUR(S) D'ALERTE	CONSERVE PAR/SUR	COMMUNIQUE PAR	EDITEURS CLEFS	SITE DE REFERENCE
United Nations Confidential Reports	Surveillance d'Internet	2008		70	TXT	confidentiels						Wikileaks	Wikileaks		https://wikileaks.org/wiki/United_Nations_confidential_reports
Afghan War Logs	Militaire	2010		91 000	TXT - vidéo - audio	confidentiels - secrets	2004 - 2010			Chelsea Manning	militaire	Wikileaks	Wikileaks		
Cable Gate	Diplomatie	2010		251 287	câbles diplomatiques	confidentiels - secrets	(1966) 2004 - 2010	Département d'Etat (Etats-Unis)		Chelsea Manning	militaire	Wikileaks	Wikileaks	Wikileaks - Le Monde - The New York Times - The Guardian - El Pais - Der Spiegel	
Irak War Logs	Militaire	2010		391 832		confidentiels - secrets	2004 - 2010				militaires				
Syria Files	Miliaire	2012		2 434 899	emails	confidentiels	2006 - 2012	emails de figures politiques syriennes	Gouvernement syrien et opposants	Anonymous	pirates	Wikileaks	Wikileaks	Al Akhbar - Al-Masry Al-Youm - L'espresso - Norddeutscher Rundfunk - OWNI - Publico	
Detainee Policies	Conditions de detention	2012		100	TXT	confidentiels - secrets		Département de la Défense des Etats-Unis	Politiques de détention des Etats-Unis						
Global Intelligence Files	Surveillance d'Internet	2012		5 millions	emails	confidentiels	2004 - 2011								
Offshore Leaks	Fiscalité	2013	260 GO	2,5 millions	BDD - TXT	confidentiels	années 90/2000	Portcullis Trustnet - Commonwealth Trust Limited	environ 120 000 sociétés offshores	anonyme	employé	ICIJ	Médias divers	ICIJ	https://offshoreleaks.icij.org/
Révélations d'Edward Snowden	Surveillance d'Internet	2013		1,7 million		confidentiels - secrets		NSA - CIA	NSA - CIA	Edward Snowden	employé		Laura Poitras - Glenn Greenwald		
Sony Leaks	Cinéma	2014	100 TO	37 937 159	emails - BDD - TXT- vidéo - audio	confidentiels - secrets		Sony Pictures	Sony Pictures	Guardians of Peace	hackers	WikiLeaks	Sites de streaming - Wikileaks	Wikileaks	https://wikileaks.org/sony/

NOM	SUJET	ANNEE	POIDS	NBRE DE DOCS	TYPE DE DOC	NATURE DES DOCS	PERIODE	PROVENANCE DES ARCHIVES	MIS EN CAUSE	NOM(S) LANCEUR(S) D'ALERTE	STATUT LANCEUR(S) D'ALERTE	CONSERVE PAR/SUR	COMMUNIQUE PAR	EDITEURS CLEFS	SITE DE REFERENCE
Swiss Leaks	Fiscalité	2014	3,3 GO	60 000	BDD	confidentiels	2005 - 2007	HSBC	HSBC - 106 000 clients	Hervé Falciani	employé	Le Monde	Médias divers	Le Monde - ICIJ	https://www.icij.org/project/swiss-leaks/explore-swiss-leaks-data
Luxembourg Leaks	Fiscalité	2014	4 GO		BDD - PDF - TXT	confidentiels	2003 - 2011	Fisc luxembourgeois	Fisc luxembourgeois - cabinets d'audit - multinationales	Antoine Deltour - Raphaël Halet - Edouard Perrin	employé - employé - journaliste	ICIJ	Médias divers	ICIJ - Le Monde - The Guardian - Süddeutsche Zeitung	https://www.icij.org/project/luxembourg-leaks
Buisson Leaks	Politique	2014				privés - confidentiels		écoute de l'élysée	Sarkozy - collaborateurs	Patrick Buisson	employé - pirate		Le Canard Enchainé - Atlantico		
Football Leaks	Fiscalité	2016	1,9 TO	18,6 millions	emails - BDD - images - TXT	confidentiels			industrie du football	John	anonyme	Der Spiegel	Médias divers	Der Spiegel - EIC	https://eic.network/projects/football-leaks
Panama Papers	Fiscalité	2016	2,3 TO	11,5 millions	emails - BDD - images - TXT - autres	confidentiels	1970-2016	Cabinet d'avocats panaméen Mossack Fonseca	214 488 sociétés offshores	anonyme	anonyme	ICIJ	Médias divers	Süddeutsche Zeitung - ICIJ	https://panamapapers.icij.org/
Email Gate		2016						boites mails des collaborateurs d'Hillary Clinton	Hillary Clinton - collaborateurs						
Pizza Gate	Pédophilie	2016		149	emails	privés - confidentiels	2015	emails John Podesta	John Podesta - entourage de Hillary Clinton			Wikileaks - 4Chan - Reddit	Wikileaks - 4Chan - Reddit	Wikileaks - 4Chan - Reddit	Wikileaks
Macron Leaks	Politique	2017	9 GO	71 848	emails - TXT	confidentiels - privés - personnels		boites mails professionnelles En Marche!	En Marche!	WikiLeaks	hackers	WikiLeaks	WikiLeaks	WikiLeaks	https://wikileaks.org/macron-emails/
Poison Papers	Environnement	2017		100 000	TXT		1920 - 2017	industries chimiques américaines et administrations	industries chimiques américaines et administrations	Carole Van Strum		Carole Van Strum	Biosience Resource Project - Center for Media and Democracy	Mr Mondialisation ?	https://www.poisonpapers.org/the-poison-papers/
Kissinger Cables				1,8 million	emails - câbles diplomatiques - TXT	secrets			Henry Kissinger						
Guantanamo Files	Conditions de detention					confidentiels	2002 - 2008								

NOM	SUJET	ANNEE	POIDS	NBRE DE DOCS	TYPE DE DOC	NATURE DES DOCS	PERIODE	PROVENANCE DES ARCHIVES	MIS EN CAUSE	NOM(S) LANCEUR(S) D'ALERTE	STATUT LANCEUR(S) D'ALERTE	CONSERVE PAR/SUR	COMMUNIQUE PAR	EDITEURS CLEFS	SITE DE REFERENCE
Berat's Box		2016		57 934	emails	confidentiels - privés - personnels	2000 - 2016	Boite mail du ministre de l'énergie turc Berat Albayrak	Berat Albayrak et son entourage	Redhak (groupe hacktiviste turc)	activistes	Google Drive - Dropbox - Wikileaks	Wikileaks	Wikileaks	https://wikileaks.org/berats-box/

TABLE DES MATIÈRES