



HAL
open science

Enjeux géopolitiques des données, asymétries déterminantes

Ghislaine Chartron, Evelyne Broudoux

► **To cite this version:**

Ghislaine Chartron, Evelyne Broudoux. Enjeux géopolitiques des données, asymétries déterminantes. Document numérique et société, May 2015, Rabat, Maroc. pp.67-83, 10.3917/dbu.chron.2015.01.0065 . sic_01304035

HAL Id: sic_01304035

https://archivesic.ccsd.cnrs.fr/sic_01304035

Submitted on 19 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Enjeux géopolitiques des données, asymétries déterminantes

Ghislaine CHARTRON, Professeur CNAM, Evelyne BROUDOUX, Maître de conférences, laboratoire DICEN-IDF, EA7339, CNAM-Paris

A Paraître Editions de Boeck, actes du Colloque *Open data , big data :quelles valeurs ?*
Colloque DocDoc, 4-5mai 2015, Rabat

Résumé :

Cette communication veut contribuer à une analyse critique du big data et de l'open data en convoquant le concept d'asymétrie pour une lecture géopolitique des données massives, dans la filiation de certains travaux antérieurs sur la géopolitique du Cyberspace. La géopolitique des données (nous adoptons ici une définition extensive de la notion de « données ») est mise en perspective entre les enjeux de l'économie numérique et de l'apparente gratuité et les enjeux de la sécurité, des droits fondamentaux difficilement convergents.

La grille de lecture s'appuie sur l'analyse de plusieurs asymétries installant des déséquilibres mondiaux : l'asymétrie technologique conférant à quelques acteurs un pouvoir central en terme de capacité de stockage, de calculateurs et de savoir-faire pour le traitement informatique des données à l'échelle mondiale ; l'asymétrie de la collecte des données et notamment le pouvoir des plateformes d'intermédiation notamment les GAFAs (Google, Apple, Facebook, Amazon) et les *data brokers* spécialisés dans chaque secteur ; l'asymétrie de cadres législatifs qui confère à certaines zones géographiques des avantages de développement économique au détriment de protections plus attentives à la vie privée et enfin une asymétrie entre les acteurs produisant des contenus et les nouveaux acteurs du numérique revendiquant une ouverture sans barrière de ces contenus à leurs algorithmes dans une vision d'innovations de services.

Mots-clés :

Géopolitique des données ; asymétrie ; économie numérique ; gratuité ; droits fondamentaux

Pourquoi parler de géopolitique ? Le cyberspace de retour...

Si la géopolitique est définie par l'*Encyclopédie Universalis* comme la discipline qui s'interroge sur les rapports entre espace et politique et sur l'influence prise par les réalités géographiques sur les organisations sociales, on peut assurer avec (Robine, Salamatian, 2014) que la disposition des infrastructures de l'internet et la distribution spatiale des routeurs et des centres de calcul jouent sans aucun doute un rôle géopolitique dans les stratégies de pouvoir des états et des acteurs socio-économiques. La définition du géopolitique s'est donc élargie à des facteurs désormais déterminants dans l'économie de nos sociétés contemporaines, à savoir le déploiement des technologies de l'information et de la communication et en premier lieu le réseau Internet et le Web qui ont pour particularité d'inscrire les comportements dans des espaces virtuels plutôt que géographiques.

La notion de cyberspace revient alors à point nommé pour traduire un univers dont la gouvernance, la régulation, les alliances, sont devenus des enjeux majeurs dans un contexte d'économie globalisée mais aussi de déploiement opaque d'une cybercriminalité devenue centrale, mettant en cause la sécurité des Etats. L'affaire Snowden a publiquement mis en lumière ce dessous des cartes (Pétiniaud, 2014). On parle désormais d'armée électronique : « *Comment le journal le Monde a été piraté par l'Armée électronique syrienne* » titre le quotidien le 20 janvier 2015, à la suite d'une tentative d'intrusion dans l'espace éditorial du quotidien à des fins probables de propagande. La guerre électronique est engagée et les Etats y consacrent plus que jamais une part importante de leur budget. Que ce soit le piratage massif de *Sony Pictures*¹ ou les attaques en déni de service, la lutte contre le

¹<http://abonnes.lemonde.fr/pixels/article/2015/01/28/piratage-de-sony-la-france-doute-de-la-piste-nord->

cybercrime s'organise aujourd'hui avec ses conférences et de nouveaux diplômés dans des écoles spécialisées en cybersécurité.

La géopolitique traite donc des relations entre pouvoir et territoire, le territoire est un enjeu pour développer son pouvoir, qu'il soit constitué de ressources naturelles, d'alliances stratégiques, de contrôles d'infrastructures numériques. La conquête du territoire conduit à des affrontements, à des négociations, à des accords multi-parties. Le web est désormais au cœur de cette problématique et s'est éloigné de ses origines marginales et libertaires en l'espace d'un quart de siècle.

L'objectif de cet article est d'une part de contribuer à une lecture critique de l'engouement actuel pour un mouvement généralisé qui prône l'ouverture des données, des contenus dans tous les secteurs au nom de valeurs humanistes mais sans prendre en compte les stratégies à l'œuvre et les déséquilibres potentiels. Nous adopterons volontairement ici une définition extensive de la notion de « données » telle qu'elle tend à se généraliser, promue par des enjeux technologiques, à savoir des contenus qui peuvent nourrir tout type d'algorithmes, source de nouvelles valeurs. D'autre part, il s'agira de structurer cette analyse par le concept d'asymétrie, asymétrie de l'économie des données, mettant en exergue quelques dimensions déterminantes : l'asymétrie technologique, l'asymétrie contenu/contenant, l'asymétrie de la collecte des données et l'asymétrie des cadres juridiques que cible notamment la révision de directives européennes en vue d'une harmonisation pour l'essor du marché numérique européen dans un contexte de compétitivité accrue et de transactions mondialisées.

De la géopolitique du cyberspace à la géopolitique des données

Concernant le cyberspace, les décennies passées ont surtout été marquées par des stratégies de contrôle des infrastructures (les câbles, le nommage, les instances de régulation de l'Internet, plus récemment le Cloud...) et sur les outils et services d'accès et d'échange (les moteurs de recherche, les réseaux sociaux) (Hérodote, 2014).

Aujourd'hui, ce sont les données qui focalisent les enjeux... Elles sont devenues, pour le commerce électronique, la brique majeure de contrôle et d'affinement de la relation-client, convoitées pour nourrir les algorithmes de recommandation (Kembellec, Chartron, Saleh, 2014) (Mercanti-Guérin, 2014). Elles sont aussi, dans le cadre du mouvement open-data, l'expression d'une transparence revendiquée dans un contexte de corruption grandissant, la réponse à une demande de retour sur investissement face à des dépenses publiques critiquées, la contribution à des objectifs d'efficacité accrue pour des appareils d'Etat souvent paralysés par l'inertie de leur organisation. Dans le contexte scientifique, l'accès aux données est associé à la qualité (preuve du résultat scientifique), à une capitalisation plus facile et à un transfert potentiel facilité pour le secteur économique.

Les données, une nouvelle utopie technologique ? Peut-être, elles portent à la fois la croyance d'une économie plus performante, d'un contrôle optimisé de nos ressources (*smart cities*), d'une science vérifiable et mieux partagée, d'un Etat plus efficace... Elles portent aussi le risque d'une société de contrôle en extension et d'une dépendance accrue à la gouvernance de systèmes automatisés (Réseaux, 2013), (Broudoux, 2015).

La donnée, au cœur de l'économie numérique et de l'apparente gratuité des services

L'accès ouvert aux contenus et aux services numériques, la prise en compte de l'avis de l'internaute et de sa contribution, la personnalisation des services par la gestion d'une relation client (CRM) affinée placent désormais l'expérience-utilisateur au premier rang. L'objectif majeur est de retenir l'attention de cet utilisateur dans une profusion de services et une inflation d'informations. Les modèles économiques du numérique sont difficiles à trouver car ils doivent opérer avec cette apparente gratuité toujours plus revendiquée, dans des équilibres très instables : modèles *freemium*, licences générales permettant des accès généralisés, mouvement *open access* avec la recherche de financements amont... L'exploitation des données personnelles est souvent devenue centrale dans la gestation de ces

nouveaux modèles : recommandation, publicité ciblée, services profilés adossés à l'apparente gratuité et monétisant les données collectées. Le récent rapport de *Fabernovel* (Fabernovel, 2014) pose le concept de « GAFAnomics » de façon très éclairante pour décrire les modèles d'affaires de *Google*, *Apple*, *Facebook* et *Amazon* (d'âge moyen de 22 ans) ; l'analyse conforte l'adage déjà bien popularisé selon lequel, « si c'est gratuit, c'est que vous en êtes le produit », débattu de façon critique depuis plusieurs années (Bomsel, 2007).

A en croire ce rapport, la clé du succès serait la suivante, : si, hier, il s'agissait de miser sur la marque et de travailler au mieux sa stratégie publicitaire, aujourd'hui il faut miser sur l'empathie développée avec les utilisateurs et travailler la stratégie de l'expérience utilisateur. Les revenus viendront dans un second temps quand la captation de l'attention sera gagnée, la valeur de la marque renforcée et la valeur-client captée par les multiples données ; la stratégie doit ensuite s'appuyer sur une diversification de services proposés à des clients qui auront été fidélisés (cf. fig.1 et fig. 2).

La stratégie a une base planétaire, les services qui sont développés sont ceux qui intéressent des millions de personnes, la base client est toujours pensée au niveau international et permet d'envisager que le gratuit va finir par s'équilibrer avec un seuil minime de ventes par client doublé par les revenus publicitaires d'une région internationale.



Figure 1 - La recette du succès
source : <http://www.fabernovel.com/fr/gafa/> (Fabernovel, 2014)

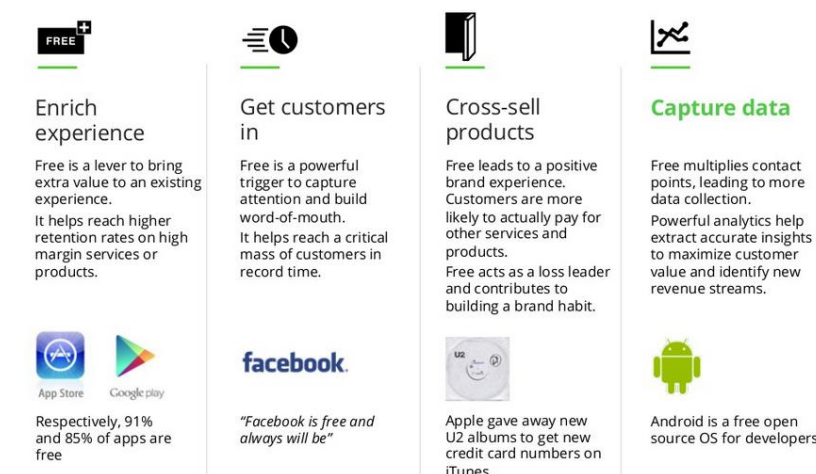


Figure 2 - Données et stratégie de développement des GAFAs
source : <http://www.fabernovel.com/fr/gafa/> (Fabernovel, 2014)

Les GAFAs détiendraient aujourd'hui un marché adressable de 7 milliards d'utilisateurs, une base client équivalente à près de 50% de la population connectée. Ils réinvestissent massivement leurs gains dans l'innovation technologique et l'amélioration de l'expérience utilisateur. Pour autant cette stratégie est-elle déployable par d'autres entreprises de moindre taille opérant sur des secteurs et des zones de marché beaucoup plus réduites ? Est-ce aussi le modèle à suivre pour toutes les start-ups du numérique aujourd'hui ? Le pari paraît risqué au moins pour deux raisons : les GAFAs en place rendent désormais difficile l'émergence de champions numériques tant ils investissent sur tous les domaines porteurs ; d'autre part, le modèle que nous traquent ces géants s'appuie sur une assise internationale et reste peu convaincant pour des acteurs dont la valeur du produit n'intéresse qu'une clientèle limitée, ne serait-ce par la limite linguistique ou la nature-même des services proposés.

Par ailleurs, outre les données personnelles et comportementales, les enjeux sont aujourd'hui projetés plus globalement sur l'économie de la donnée au sens large : « *L'objectif du gouvernement est de hisser la France sur le podium des pays les plus avancés numériquement. Nous proposons de tirer parti au maximum de l'économie de la donnée* », a expliqué le ministre, invitant les parlementaires à « *construire ensemble la République, à l'heure du numérique, la République numérique.* » (Axel Lemaire, 2014)²

De quelles données parle-t-on ? Le *mainstream* est général pour la donnée (MCKINSEY, 2013), (OCDE, 2013) : données produites par l'Etat (open data et open gouvernement), données personnelles des clients, données de la recherche (open sciences), données des capteurs divers comme celles des villes intelligentes (optimisation des consommations d'énergie)... Le calcul que l'on pourra opérer sur ces données est identifié comme source de valeur reléguant au second plan les conditions de la production qualitative de ces données. Cette vision fait évidemment débat quand la donnée est mise en contexte. Les données d'un capteur sont-elles comparables à des textes d'auteurs ? Vraisemblablement pour un algorithme mais pas sans revendication pour des auteurs et des professionnels qui ont, eux, investi pour produire ces contenus.²

Les données au carrefour de l'économie numérique, de la sécurité, des droits fondamentaux et des diversités culturelles

Au sein des instances européennes, la prise de position sur des volets numériques n'est pas toujours convergente entre les différentes Directions Générales, notamment entre la Direction Générale de la concurrence, de la Justice et des consommateurs, de l'Education et de la Culture. La question des données confronte différentes dimensions, à savoir notamment l'économie, la protection des sociétés et des individus, la liberté de chacun. Si l'ouverture de certaines données est bonne pour l'économie, elle peut aussi exposer des informations qui seront utiles au terrorisme ou encore être potentiellement discriminantes pour des individus, par exemple. Les domaines actuels du *big data* et du *cloud computing* incluant des processus d'externalisation, de transfert et d'exploitation de données sont particulièrement assujettis à ce type de tensions. La clause contractuelle ne suffit plus à encadrer les pratiques numériques dans une économie globalisée, l'enjeu sociétal est tel que le passage par la loi devient la voie privilégiée.

Au sujet de l'accord de libre-échange entre l'UE et les Etats-Unis, le Président Juncker déclare sur le site de la commission³ « *La Commission européenne est en faveur du libre-échange, mais pas au point de sacrifier les normes européennes en matière de sécurité alimentaire, de santé, de protection sociale ou des données, ou de diversité culturelle.* ». Mais encore concernant le volet « Justice et droits fondamentaux » : « *J'entends utiliser les prérogatives de la Commission pour défendre, dans notre domaine de compétence, nos*

² La Tribune, « Données personnelles : le sujet polémique du futur projet de loi Numérique », 19/01/2015
<http://www.latribune.fr/technos-medias/20150119tribcac5f6de/donnees-personnelles-le-sujet-polemique-du-futur-projet-de-loi-numerique.html>

³ http://ec.europa.eu/priorities/eu-us-free-trade/index_fr.htm

valeurs communes, l'État de droit et les droits fondamentaux, tout en tenant dûment compte de la diversité des traditions culturelles et constitutionnelles des 28 pays de l'UE »⁴. La tension est toujours latente entre l'objectif d'un alignement unique pour l'efficacité du marché unique européen qu'il « faut » développer et la prise en compte des dimensions sociale, sociétale et culturelle de chacun des pays membres... Le marché et la compétitivité reprennent le dessus : « *Les technologies numériques ne connaissent pas de frontière. Une Union européenne dans laquelle chaque pays possède ses propres règles en matière de télécommunications, de droits d'auteur, de protection des données ou de gestion du spectre radioélectrique n'a donc plus de sens* »⁵.

Des asymétries déterminantes pour la géopolitique des données

Le concept d'asymétrie au cœur des enjeux

L'asymétrie est une notion qui s'emploie en biologie, en logique, mais aussi en économie, en stratégie militaire. En économie, l'asymétrie de l'information est associée à la question du fonctionnement des marchés et de la concurrence dite « parfaite », le concept décrit la situation où des agents disposent d'informations pertinentes que d'autres n'ont pas⁶. L'asymétrie d'information pouvant être destructrice de la rationalité des agents économiques. En stratégie militaire, le conflit asymétrique est un conflit où « *les adversaires n'ont ni le même statut, ni les mêmes critères de victoire ou de défaite, ni les mêmes règles et méthodes, ni n'emploient les mêmes moyens... Terrorismes, guérillas, désordres mafieux conflits dans les zones de non-droit..., sont des conflits asymétriques* » (Huygues).

Dans notre contexte, ce concept va nous aider à éclairer l'économie globalisée de la donnée et à situer les enjeux. L'affaire Snowden nous a déjà pointé la disparité des moyens de surveillance de données entre les Etats-nations, l'avantage concurrentiel des GAFAM précédemment évoqué illustre une autre forme d'asymétrie majeure pour l'avenir du commerce électronique. Nous allons pointer plus finement plusieurs facettes de l'asymétrie de la donnée. En premier lieu, les écosystèmes de son traitement : qui sont les leaders technologiques pour réceptionner ces masses de données et les analyser automatiquement ? En termes de gisement de données, quels sont les flux majeurs de transaction des données, les lieux de captation dominants au niveau international ? En quoi les cadres législatifs contribuent à certaines asymétries avec quels déséquilibres potentiels ? Quel est le devenir des producteurs de la donnée (au sens large) si toute la valeur se porte aujourd'hui sur son traitement ? Le bien commun de la donnée est-il réaliste ?

Données et asymétries technologiques

Les secteurs *a priori* intéressés par le big data sont les télécommunications, les banques, les assurances et les acteurs des marchés financiers avec, entre autres, les services bancaires, l'analyse des transactions, la gestion de patrimoine mais aussi celle des risques et des fraudes. C'est la nature même des activités de ces organisations qui les a conduites à forger ou adopter un ensemble d'algorithmes visant l'analyse des marchés en temps réel, la prédiction événementielle, la détection concurrentielle, etc. Alors que les services informatiques de ces grandes entreprises étaient gérés par des acteurs traditionnels (serveurs EMC, HP, Microsoft, IBM, etc.), la tendance observée en 2014 a été la sélection par ces mêmes entreprises des architectures Cloud dites « web-scale IT », c'est-à-dire en open source et appartenant aux nouvelles entreprises informatiques du web⁷. Le projet « Open Compute » fondé en 2011 par Facebook et rejoint par Microsoft, Intel et Goldman Sachs⁸ en 2014 est éminemment

⁴ http://ec.europa.eu/priorities/justice-fundamental-rights/index_fr.htm

⁵ http://ec.europa.eu/priorities/digital-single-market/index_fr.htm

⁶ voir les travaux du professeur Joseph E. Stiglitz a reçu en 2001 le prix Nobel d'économie, avec George A. Akerlof et A. Michael Spence pour "leurs travaux sur les marchés avec asymétrie d'information"

⁷ <http://www.larevuedudigital.com/2014/03/06/les-grandes-entreprises-imitent-les-geants-du-web-pour-leur-informatique/>

⁸ <http://www.usine-digitale.fr/editorial/pourquoi-la-france-ne-devrait-pas-bouder-l-open-compute-project.N293526>

symbolique de cette nouvelle convergence. La conception de fermes de serveurs est construite par des entreprises comme *Amazon* avec *EC2* ou *AWS* dont une version spéciale a même été conçue pour la CIA⁹.

Les quatre couches fonctionnelles du big data (Aide à la décision et automatisation, Analyse et Découverte, Organisation et Gestion des données, Infrastructure) sont directement dépendantes des performances calculatoires du Cloud (IaaS, SaaS et PaaS). Selon *Orange-Business*, une progression de 36% du marché mondial du cloud computing est attendue pour 2015 avec 104 milliards d'euros pendant que les promesses du big data sont évaluées à 110 milliards d'euros. A titre de comparaison le marché français du Cloud devait atteindre 5 milliards d'euros fin 2014. La répartition mondiale du CA global du Cloud est également révélatrice d'une asymétrie : environ les deux tiers (68%) a été réalisé aux Etats-Unis en 2013, contre un cinquième en Europe (19%) et 13% dans le reste du monde. Le IaaS quant à lui, constituerait environ 27% du marché du cloud total¹⁰.

La nouvelle génération de plateforme « cloud » qui fera converger en 2015 interaction sociale, mobilité, cloud et information ressemble fortement à un « Data-as-a-Service »¹¹ spécialisé pour le big data des... données personnelles.

Le carré magique BI (Business Intelligence and Analytics Platforms) du Cabinet *Gartner* qui décrit chaque année le secteur en termes de bases installées a révélé en 2014 la baisse des leaders du secteur (*SAP*, *Oracle*, *Microsoft*) mais aussi l'arrivée de petits éditeurs spécialisés dans la visualisation de données comme *Tableau*. Le carré 2015 de *Gartner*¹² confirme la tendance à l'installation durable d'entreprises de niche dans la découverte de données (data discovery) comme *Qlik* reposant sur un moteur de recherche à mémoire associative ou *Logi Analytics* qui a lancé en 2014 une plateforme collaborative de découverte de données. La tendance est également aux plateformes « tout-en-un » capables de traiter les données (non)structurées et hétérogènes et d'en tirer des analyses traduites par des infographies pertinentes (*Alteryx*, *Board*, *Datawatch*, *GoodData*, *Information Builders*, *OpenText*, *Prognoz*, *Targit*, etc.) (cf. fig.3).

Il faut noter ici que face à cette diversité se tiennent les géants du web qui rachètent à tour de bras les start-up et investissent lourdement dans les technologies émergentes. L'infographie d'Amy Webb publiée en 2006 et mise à jour en 2008 faisait le compte des opérations de rachat systématique des start-up du web 2.0 par les acteurs majeurs du web : *Google*, *Yahoo!*, *Microsoft*, *AOL*, *IAC* et *News Corporation* se partageaient alors le gâteau en parts inégales¹³. La distribution des applications elles-mêmes est asymétrique (cf. fig. 4). Aujourd'hui, *Apple*, *Facebook*, *Amazon* et *Twitter* ont rejoint le club et les opérations de rachats concernant toute l'innovation liée aux usages de l'informatique se chiffrent en milliards d'euros comme le montre le Tableau 1 établi pour 2013.

⁹ <http://www.entreprisesetech.com/2014/11/14/rare-peek-massive-scale-aws/>

¹⁰ <http://www.orange-business.com/fr/blogs/cloud-computing/actualites/infographie-etat-des-lieux-et-tendances-du-cloud>

¹¹ <http://sdtimes.com/idcs-top-10-technology-predictions-2015/>

¹² <http://www.gartner.com/technology/reprints.do?id=1-2AH4Q85&ct=150224&st=sb>

¹³ https://stilgherian.com/business/who_owns_internet/



Figure 3 - Carré magique de Gartner pour la Business Intelligence et les plateformes analytiques

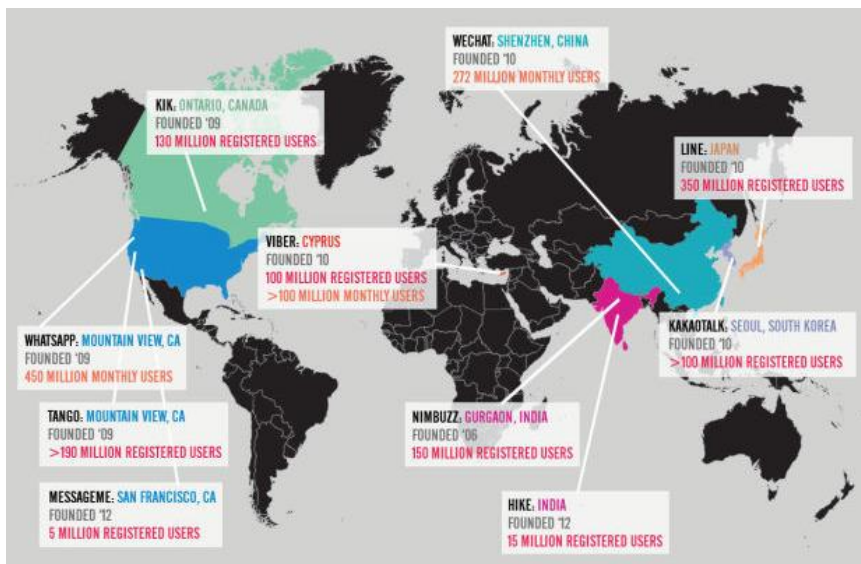


Figure 4 - Exemple d'asymétrie dans la répartition des 10 premières applications de messagerie instantanée dans le monde en 2014 : l'Europe, l'Afrique et l'Amérique du Sud en sont absentes (URL : <http://readwrite.com/2014/03/06/10-biggest-popular-mobile-messaging-apps-world-whatsapp>)

	Acquisitions 2013	Domaine	Valeur d'achat en milliards de \$
Google	Waze	Navigation GPS en crowdsourcing	1,15
	Flutter	Détection de mouvements	0,40
	Nest Labs Inc	Domotique	3,2
	Deep Mind Technologies	Intelligence artificielle	0,65
	Wavii	Résumés en langage naturel	0,30
	DNN Research	Réseaux de neurones	?
Microsoft	Meka	Robotique et robots humanoïdes	?
	Nokia	Téléphonie mobile	7,2
	Yammer	RSE	1,2
Apple	Skype	Télécommunications	8,5
	Cue	Assistant personnel	0,50
	PrimeSense	Semi-conducteurs	0,345
Yahoo!	Topsy	Moteur recherche Twitter	0,20
	Xobni	CRM	0,40
	Qwiki	Production vidéo	0,50
	Tumblr	Plateforme de blogs	1,1
Facebook	Summly	Résumés de news	0,30
	Oculus VR	Réalité virtuelle	2
	Instagram	Partage photos	1
Twitter	WhatsApp	Messagerie instantanée	19
	Bluefin Labs	Analyse réseaux sociaux	0,80
	TweetDeck	Outil client Twitter	0,40
Amazon	Goodreads	Réseau social	?
	IMDB	Base de données de films	?
	Twitch	Jeu vidéo	0,97

Tableau 1 - Acquisitions 2013 par les majors du web (synthèse réalisée)

Asymétrie de la captation des données

C'est la captation des données dites sociales, c'est-à-dire fabriquées par nos interactions en ligne, qui apparaît être une source inépuisable intéressant principalement le marketing des entreprises et les industries liées à la surveillance. La tendance à créer des systèmes rendant captifs les usagers qui échangent indistinctement tout type de messages faisait dire à Olivier Ertzscheid (Ertzscheid, 2009) que « l'homme [était] un document comme les autres ». Cette indexation de l'humain réalisée à la fois par les moteurs de recherche et par les réseaux sociaux n'est pas sans poser de questions sur les objectifs sociopolitiques de ceux qui les ont mis en place.

Stéphane Grumbach et Stéphane Frénot (Frénot, Grumbach, 2014) distinguent les données primaires des données secondaires générées par les activités des usagers sur les plateformes d'« intermédiation » nommées ainsi parce qu'elles mettent en relation des personnes avec des personnes ou des documents (réseau social, moteur de recherche, plateforme web2, etc.). Les données secondaires étant les traces de lecture et d'interactions laissées par les usagers de ces plateformes. Or ces flux de données qui s'échangent lors d'une identification de type « Facebook » ou « Twitter » sur la planète, constitués par les données de profil mais aussi par l'historique des activités, tombent systématiquement dans l'escarcelle de quelques entreprises situées aux Etats-Unis et en Asie. L'Europe fait ici figure de parent pauvre car aucune des données dites « sociales » produites sur son territoire n'y reste et ne peut y être exploitée. Leur analyse des connexions à partir du territoire français est parlante : les vingt premiers sites visités par les français sont constitués par 664 millions de connexions vers treize premiers sites nord-américains et 158 millions vers sept sites français. Il faut ajouter à cela que les connexions aux sites français sont moins riches d'enseignements car plus spécialisées et qu'elles s'effectuent majoritairement à partir des plateformes de type *Google* et *Facebook*. Ces deux entités sont donc à même de mieux connaître les centres d'intérêts des citoyens

que les instituts nationaux chargés de leur analyse. *Google* a par exemple créé le service *Flutrends*¹⁴ qui permet de suivre la propagation de la grippe dans le monde à partir des requêtes entrées par les internautes et qui a donné lieu à publication d'un article dans la revue *Nature*.

Cette suprématie nord-américaine dans le traitement des données dites sociales est dangereuse pour les Etats eux-mêmes car ils n'apparaissent pas en mesure de traiter les données produites par leurs citoyens sur leur propre territoire.

Enjeu crucial dans le contexte de l'économie numérique, les données sociales sont dites « personnelles » lorsqu'elles concernent l'individu et ses caractéristiques objectives et subjectives. (Rochandelet, 2010) en propose une classification et ses exemples de techniques de collecte et d'exploitation donnent à penser qu'une signature unique individuelle peut être obtenue à partir d'un ensemble de variables. C'est bien donc la singularité qui est captée ici.

Asymétrie, données et cadres législatifs

Le développement d'une économie numérique très compétitive associée à la vision d'une créativité qui s'adosserait à un domaine public étendu des contenus s'oppose au cadre législatif actuel du droit d'auteur et à celui des droits fondamentaux de protection des individus. Les défenseurs des potentialités du numérique prônent de nouvelles exceptions en phase avec de nouvelles pratiques de manipulation et d'exploitation des données.

La réforme du droit d'auteur est emblématique des tensions et des postures différentes dans le contexte européen. Du point de vue du développement d'un marché européen unifié, l'enjeu est de faciliter l'essor d'acteurs européens capables de rivaliser avec les plateformes américaines notamment, il s'agit de créer des règles uniques dans l'ensemble de l'Union européenne et donc de remettre en cause les spécificités des 28 réglementations des Etats-membres, ciblées comme une barrière pour le développement des industries culturelles et plus généralement des industries du numérique. La question est sensible et renvoie à des équilibres historiquement élaborés dans chaque Etat membre : le *faire use* anglophone s'oppose à une conception plus protectrice de l'auteur et de ses ayants droits en France notamment, adossée à un modèle spécifique de financement de la création (Sirinelli, 2014). Les défenseurs des potentialités numériques pour des échanges non marchands militent au contraire pour l'adoption de nombreuses exceptions : ramener la durée des droits de 70 à 50, le droit au *remix*, *mashup* et autres usages transformatifs, l'élargissement de l'exception pédagogique actuellement existante (Quadrature, 2014), et certains partisans de l'*open access* pour l'introduction d'une nouvelle exception en faveur de l'exploration de données (*text et data mining*) sans compensation particulière, cette exception est d'ailleurs déjà adoptée au Royaume-Uni¹⁵. Ce sont donc des visions différentes des valeurs à prioriser qui s'opposent, l'une unissant de façon conjoncturelle une posture libérale et une posture libertaire pour le développement du domaine public, l'autre plus héritée et attachée à la défense première du créateur, de ses intérêts mais aussi des industries culturelles fondées sur l'exploitation des droits de l'auteur et de l'œuvre.

Le transfert des données personnelles au centre d'une économie globalisée, est également source d'asymétrie entre les pays pouvant déboucher sur des déséquilibres majeurs où l'on verrait des sociétés étrangères exploiter massivement des données de citoyens de tous pays au profit de leur unique développement, sans limite éthique, sans garantie, sans accord de réciprocité entre pays. Par l'initiative pionnière de la loi « Informatique et Liberté » en 1978¹⁶, puis par l'action soutenue de la CNIL et son rôle moteur au sein du groupe G29 rassemblant les « CNIL européennes », la France veille particulièrement à la protection des données personnelles. Le cadre législatif stipule (CNIL, 2012) :

¹⁴ <http://www.google.org/flutrends>

¹⁵ <https://www.gov.uk/exceptions-to-copyright>

¹⁶ La loi n°78-17 du 6 janvier 1978, Loi informatique et libertés, <http://www.cnil.fr/documentation/textes-fondateurs/loi78-17>

« En principe, les transferts de données à caractère personnel hors du territoire de l'Union européenne sont interdits à moins que le pays ou le destinataire n'assure un niveau de protection suffisant...Plusieurs outils ont été développés pour permettre aux acteurs d'apporter un niveau de protection suffisant : les règles internes d'entreprise (ou BCR), les Clauses Contractuelles Types et l'adhésion aux principes du « Safe Harbor ». La loi prévoit également des exceptions permettant de transférer des données vers des pays tiers sans qu'il n'existe pour autant un niveau de protection suffisant. »

La révision actuelle de la Directive 95/46/CE du Parlement européen, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, veut renforcer les droits des personnes et relever les défis de la mondialisation et des nouvelles technologies¹⁷. Cette priorité à l'encadrement juridique peut s'opposer en bien des points, comme nous l'avons déjà évoqué, à d'autres divisions de la Commission européenne chargées du soutien au marché numérique. Elle s'oppose aussi fondamentalement à la vision britannique où prime l'urgence de permettre aux start-ups du numérique d'accéder à un grand marché sans barrière pour y travailler, innover et faire du commerce.

Asymétrie contenu/contenant : les opérateurs techniques (vs) les producteurs des contenus

La montée en puissance d'une vision positive de la multiplication des échanges non marchands dans la société numérique conduit, comme évoqué précédemment à vouloir renégocier le cadre juridique du droit d'auteur. Les argumentations des *pour* et des *contre* sont hélas souvent trop chargées d'idéologie pour se confronter dans une analyse globale des opportunités et des risques du changement de ces règles.

Le transfert de la valeur de l'amont (la production des contenus) à l'aval de la chaîne (exploitation par des logiciels divers) n'est pas sans conséquence sur le devenir de l'auteur, de sa motivation, ni sur la qualité de ce qui sera produit. Investir sur la production risque de ne plus être attractif car sans garantie de retour sur cet investissement dans un cadre devenu trop permissif sur les règles d'exploitation. L'incertitude des modèles économiques pourrait conduire à réduire l'engagement dans de telles activités, laissant place à une économie horizontale ouverte à tous sans barrière de sélection, également à une économie subventionnée ou totalement publique dont le risque est aussi, outre la diversité et la sélectivité, la pérennité et la garantie d'indépendance. L'enjeu d'une circulation plus ouverte des contenus apparaît indéniablement bénéfique à l'ensemble de la société pour autant qu'elle ne conduise pas à un appauvrissement global de ce qui est lu et débattu dans nos sociétés démocratiques.

On comprend aisément que les avantages des uns ne sont pas ceux des autres, autrement dit que les opérateurs techniques ont tout à gagner d'une ouverture des contenus pour proposer de nouveaux services d'analyses, de ciblage segmentés. Prenons l'exemple de l'*open access* des publications scientifiques produites par une diversité d'éditeurs de toute taille. Imposer une mise à disposition rapide de leur production conduira très certainement à voir disparaître les plus petits d'entre eux car la plupart pourront difficilement rebondir en déployant de nouveaux modèles économiques pensés sur des bases élargies que majoritairement seuls les grands groupes ou les grands acteurs technologiques auront la capacité de perfectionner sur une échelle mondiale. Le transfert de la valeur ne se fera pas sur les mêmes acteurs, le risque étant de voir se creuser le fossé entre ceux qui produisent du contenu et ceux qui l'exploitent. Le scénario prévisible est probablement celui d'une concentration extrême et celui d'un avenir où les opérateurs technologiques fonderont progressivement leurs propres filières de production : Google, Amazon, Apple éditeurs de demain¹⁸...

De façon plus spécifique mais convergente avec ces propos, la vision européenne qui tend à imposer des délais très courts de libre accès aux publications de la recherche (6 mois en STM et 12 mois en SHS) se heurte à la difficile mise en œuvre par les acteurs de l'édition et l'alignement est controversé, notamment dans les pays où une activité éditoriale est portée par des acteurs privés ou d'économie

¹⁷ http://ec.europa.eu/justice/data-protection/index_fr.htm

¹⁸ France : Amazon recrute et prépare l'ouverture de sa maison d'édition

<https://www.actualitte.com/les-maisons/france-amazon-recrute-et-prepare-l-ouverture-de-sa-maison-d-edition-55008.htm>

mixte (Lomazzi, Chartron, 2014) (Chartron, 2014). Il faudra aussi se poser la question de savoir, dans le cadre d'un déséquilibre entre les politiques publiques d'*open access*, à qui profitera ces contraintes d'ouverture dans un contexte d'internationalisation des équipes de recherche et des lieux de publication.

Conclusion : Quelle société pour demain ?

Les différentes asymétries sur les données pointées montrent à quel point les régulations des pouvoirs publics seront déterminantes sur les sociétés de demain. Quelle priorité se fixe-t-on ? Doit-on suivre les sirènes du numérique qui misent prioritairement sur la création de nouvelles valeurs en revendiquant un domaine public de données étendu ? Quelles mesures juridiques équitables avec les producteurs de contenus ? Doit-on laisser les géants du web capter les données de tous les Etats-nations sans retour sur les territoires ? La protection des droits fondamentaux restera-t-elle au centre des décisions qui seront prises ? Comment se préserver des dérives potentielles de la collecte de données personnelles ? Le citoyen peut-il rester maître de ses données personnelles ? Il s'agira de hiérarchiser les priorités.

La réflexion a volontairement été conduite à un niveau très général dans ce texte car le discours dominant tend bien aujourd'hui à amalgamer tous les types de données et à débattre globalement sur l'économie des données considérées comme nouvelle ressource généralisée. Mais il est inéluctable que les raisonnements devront s'affiner et considérer les conditions de production et d'exploitation de ces données dans chaque contexte.

L'encadrement juridique et la régulation à mettre en place seront importants. La révision actuelle de plusieurs cadres juridiques inscrits dans l'agenda numérique de la Commission européenne¹⁹ est un enjeu considérable et les lobbyings des différentes parties se multiplient. L'année 2015 sera marquée en Europe à la fois par les décisions relatives à la révision du droit d'auteurs²⁰, par la réforme de la législation relative à la protection des données²¹, par le déploiement de la stratégie de l'UE en matière de cyber-sécurité. Le marché unique européen qui est une priorité affichée devra concilier compétitivité des services, cyber-sécurité, protection des données personnelles, soutien à la diversité culturelle... En France, la loi sur le numérique devrait être présentée au Parlement au premier semestre 2015, elle devrait renforcer notamment la protection de la vie privée en créant un droit de déréférencement pour les mineurs et un droit à l'auto-détermination informationnelle, tout en s'engageant très certainement sur les mesures prioritaires fixées par l'Europe, à savoir le soutien au marché unique numérique programmé. Aux Etats-Unis, au même moment, les récentes prises de position de l'administration d'Obama pour promouvoir de nouvelles lois afin de protéger la vie privée, les données personnelles, particulièrement dans le cadre éducatif annoncent un virage majeur au sein même de la société américaine (White House, 2015), virage qui ne sera certainement pas sans influence sur les politiques européennes.

Bibliographie

- BELLANGER (P.), *la souveraineté numérique*, 2014, Stock, 264p.
- BOMSEL (O.), *Gratuit ! Du Déploiement De L'économie Numérique*, Gallimard, 2007.
- BROUDOUX (E.), *Autorités calculées in La preuve par l'original*, Editions de l'EHESS (à paraître).
- CHARTRON(G.), Open access et SHS : Controverses, *Revue européenne des sciences sociales*, pp.37-63, vol 52-1, 2014, URL : <http://ress.revues.org/2658> ; DOI : 10.4000/ress.2658

¹⁹ <https://ec.europa.eu/digital-agenda/digital-agenda-europe>

²⁰ http://ec.europa.eu/internal_market/consultations/2013/copyright-rules/index_fr.htm

²¹ http://ec.europa.eu/justice/data-protection/index_fr.htm

- CNIL, *Les transferts de données à caractère personnel hors Union européenne*, novembre 2012, http://www.cnil.fr/fileadmin/documents/Vos_responsabilites/Transferts/GUIDE-transferts-integral.pdf
- ERTZSCHEID (O.), " L'homme est un document comme les autres : du World Wide Web au World Life Web ", *Hermes*, CNRS-Editions, 2009, pp.33-40. <sic 00377457v2>
- FABERNOVEL, *GAFAnomics*, 2104, <http://www.fabernovel.com/fr/gafa/>
- Revue Hérodote, *Cyberspace : enjeux géopolitiques*, n°152-153, 2014, <http://www.cairn.info/revue-herodote-2014-1.htm>
- FRENOT (S.) et GRUMBACH (S.), « Les données sociales, objets de toutes les convoitises », *Hérodote*, 2014/1 n° 152-153, p. 43-66. DOI : 10.3917/her.152.0043
- HUYGUE F.-B., page personnelle <http://www.huyghe.fr/>
- KEMBELLEC, G., CHARTRON, G., SALEH, I. 2014, *Recommender Systems*. WILEY, 229 p. (vf) *Les moteurs et systèmes de recommandation*. ISTE, 228 p.
- LOMAZZI L., CHARTRON G., « The implementation of the European Commission recommendation on open access to scientific information: comparison of national policies », *ELPUB 2014*, Thessalonique, 18-19 juin 2014, *Proceedings of the 18th International Conference on Electronic Publishing Published 2014*, IOS PRESS, pp.23-29 — <http://ebooks.iospress.nl/book/lets-put-data-to-use-digital-scholarship-for-the-next-generation>
- MCKINSEY GLOBAL INSTITUTE (2013), *Open Data. Unlocking innovation and performance with liquid information*, octobre. www.mckinsey.com
- MERCANTI-GUERIN (M.), systèmes de recommandation et réseaux sociaux, quelles implications pour le marketing digital ? in *Les moteurs et systèmes de recommandation*, ISTE, pp.69-84.
- OCDE (2013), "Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data"", *OECD Digital Economy Papers*, No. 222, OECD Publishing. DOI:[10.1787/5k47zw3fcp43-en](https://doi.org/10.1787/5k47zw3fcp43-en)
- PETINIAUD (L.), Cartographie de l'affaire Snowden, *Hérodote* 1/ 2014 (n° 152-153), p. 35-42. URL : www.cairn.info/revue-herodote-2014-1-page-35.htm.
- ROBINE (J.), SALAMATIAN (K.), « Peut-on penser une cybergéographie ? », *Hérodote* 1/ 2014 (n° 152-153), p. 123-139 URL : www.cairn.info/revue-herodote-2014-1-page-123.htm.
- QUADRATURE du net, *Eléments pour la réforme du droit d'auteur et des politiques culturelles liées* <https://www.laquadrature.net/fr/elements-pour-la-reforme-du-droit-dauteur-et-des-politiques-culturelles-liees>
- Revue Réseaux, *Politique des algorithmes*, 2013/1 (n° 177), 284 p, <http://www.cairn.info/revue-reseaux-2013-1.htm>
- ROCHANDELET (F.), *Économie des données personnelles et de la vie privée*, Paris, La Découverte, Repères , 2010, 128 p., <http://www.cairn.info/Economie-des-donnees-personnelles-et-de-la-vie-pri--9782707157652.htm>
- SIRINELLI (P.), *rapport de la mission sur la révision de la directive 2001/29/CE sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information*, CSPLA, Octobre 2014, <http://adbu.fr/actualites/droits-dauteurs-et-droits-voisins-dans-la-societe-de-linformation-de-la-necessite-de-la-revision-de-la-directive-copyright/>
- White House, *Big data: seizing opportunities, preserving values*, 2015, http://www.whitehouse.gov/sites/default/files/docs/20150204_Big_Data_Seizing_Opportunities_Preserving_Values_Memo.pdf