

Industries culturelles et politisation de l'esthétique

AWADA Ali, BAASANDORJ Chinbat et JORANDON Guillaume

PH03 - P18

Université de technologie de Compiègne



Table des matières

1	Qu'est-ce qu'une industrie culturelle ?	4
1.1	Histoire des industries culturelles	4
1.1.1	Aux origines de la culture	4
1.1.2	Émergence des industries culturelles	7
1.1.3	La question de la reproductibilité technique	10
1.2	Les industries culturelles à l'ère du numérique	10
2	Plateformes de diffusion de masse et algorithmes	13
2.1	Comment les algorithmes de profilage sélectionnent le contenu	13
2.2	Conséquences	15
2.2.1	Uniformisation des modes de pensée	15
2.2.2	Mal-être et aliénation	17
2.2.3	Discriminations algorithmiques	18
2.3	Étude de cas : Netflix	20
2.3.1	L'algorithme de suggestion de Netflix	20
2.3.2	Fonctionnement de l'algorithme de suggestion	22
2.3.3	Confidentialité des données	24
2.3.4	Netflix et politisation de l'esthétique	25
3	Vers une autre culture de l'information	28
3.1	Penser l'informatique pour une gouvernance citoyenne : logiciel libre et éthique des données	28
3.2	Les technologies sont un pharmakon	30

Introduction

Le 20^e siècle a été une époque charnière dans la courte Histoire de l'humanité. Siècle des grandes guerres mondiales, il a été le témoin de projets techniques et scientifiques fulgurants, grandement motivés par les conflits et la lutte pour la suprématie. Dans l'après-guerre, on assiste progressivement à une globalisation et une industrialisation de la culture, phénomène étudié notamment par les philosophes de l'école de Francfort, comme Theodor ADORNO. Ils développent ainsi dans les années 50 le concept d'industrie culturelle, et cherchent à étudier le fort pouvoir performatif de ces industries, étroitement liées à des enjeux économiques productivistes. ADORNO postule notamment que les industries culturelles sont un des outils d'un pouvoir autoritaire et totalitaire moderne :

- autoritaire, car il impose une mentalité unique aux individus ;
- totalitaire, car s'exerce à grande échelle et à tous les niveaux ;
- moderne, car il n'a plus la forme des "dictatures d'hier".

Il est important d'insister sur ce dernier point : les industries culturelles sont en effet devenues un instrument de pouvoir notamment à l'oeuvre dans nos belles démocraties du monde libre, au sein des sociétés libérales au sein desquelles la consommation de masse et le productivisme sont profondément ancrés.

Les conséquences sont multiples, et nous essaierons d'y revenir en détails dans ce mémoire. Nous nous attacherons en effet à comprendre comment s'articulent ces industries culturelles, au travers d'un exemple récent : les plateformes de diffusion de contenu en ligne. En effet, le 21^e siècle fait la part belle, avec l'essor de l'informatique moderne, à la collecte et au traitement automatisé¹ de grands volumes de données, que l'on désigne souvent sous le terme très marketing de big data. L'accès à un Internet haut débit pour une part croissante de la population mondiale nous permet de consommer toujours plus de produits culturels : musiques, films, séries, émissions de télévision, livres, vidéos YouTube... YouTube justement, est l'une des plateformes qui règne en maître sur cette ère connectée, aux cô-

1. Ce traitement est automatisé dans le sens où les volumes de données sont traités par des algorithmes qui synthétisent l'information pour lui donner un sens. Nous reviendrons dans ce rapport sur la nature de ce sens.

tés d'autres géants culturels comme Spotify, Amazon ou Netflix, plateforme sur laquelle nous reviendrons plus en détail. Ces plateformes utilisent massivement le big data et l'analyse de données pour différentes raisons : proposer de la publicité ciblée, catégoriser le contenu selon les utilisateurs, et les utilisateurs selon le contenu, extraire des profils types, parfois même faire de la science²...

On peut alors conjecturer que cette masse d'informations sur la population mondiale³, dans les mains de ces quelques acteurs, leur confère un pouvoir politique et économique gigantesque. Elles s'inscrivent dans une logique de biopouvoir au sens de Michel FOUCAULT, c'est à dire l'exercice du pouvoir sur la vie et les corps des individus. Nous nous sommes alors demandé comment ces nouveaux acteurs, avec leurs algorithmes de profilage, peuvent-ils façonner une esthétique politisée ? Comment construisent-ils les goûts et les choix du consommateur, quelle influence ont-ils sur notre perception de l'art, et du "beau" ? En quoi prennent-elles place dans des enjeux de pouvoir ? En quoi ces plateformes sont-elles politiques ? Pour répondre à cette question, nous reviendrons d'abord en détail sur la notion d'industrie culturelle, leur histoire, l'histoire de leur étude, et ses évolutions récentes. Puis nous étudierons les plateformes de diffusion de masse et leur impact : quels problèmes soulèvent-elles, Enfin, nous essayerons de réfléchir à des alternatives pour essayer d'aborder différemment notre rapport à l'information et la culture.

Une bibliographie sélective des documents utilisés pour rédiger ce mémoire est disponible à la fin du document.

2. Des chercheurs se sont montrés intéressés par la masse d'utilisateurs de certaines plateformes en ligne, qui peuvent être un terrain d'expérimentation intéressant. On citera par exemple le cas de "corrupted blood", un bug de World of Warcraft, célèbre MMORPG de l'éditeur Blizzard. Ce bug provoqua en 2005 une pandémie virtuelle, la diffusion massive d'une maladie affectant les avatars des joueurs. Des articles ont été publiés par des chercheurs enthousiastes, comme [RD07].

3. Facebook par exemple revendique dans ses utilisateurs près d'un tiers de la population mondiale.

Chapitre 1

Qu'est-ce qu'une industrie culturelle ?

Pour commencer, il est essentiel de revenir sur la notion d'industrie culturelle, sur ses origines et ses évolutions récentes. Les industries culturelles sont un maillon important de notre monde industrialisé moderne, comme vecteurs de transmission des idées dominantes.

1.1 Histoire des industries culturelles

1.1.1 Aux origines de la culture

L'Homme s'est distingué dès le néolithique par sa capacité à extérioriser du savoir au travers d'objets techniques. Cette exosomatization de la mémoire lui a permis de vivre et de survivre en s'adaptant à des environnements divers et variés, et a donc été un critère essentiel de différenciation (et de conflit) entre les peuples et les individus : c'est la naissance de la culture. Cette évolution ne s'est pas faite sur la génétique, puisqu'Homo Sapiens a relativement peu évolué depuis Néanderthal, mais sur ce que Bernard STIEGLER appelle l'épiphylogénétique : il s'agit de facteurs d'évolution non héréditaires au sens génétique du terme, c'est à dire qui ne se transmettent pas par la reproduction. Ces caractères traversent l'Histoire grâce à des outils techniques, traces mémorielles du savoir du passé qui permettent de poursuivre la vie au delà du trépas individuel. Ces premières traces du passé, silex taillés, peintures rupestres, on les retrouve d'ailleurs aujourd'hui dans les musées d'Histoire de l'Homme. Ces traces se matérialisent au travers de ce que l'on peut appeler des rétentions tertiaires. Les rétentions sont un terme d'HUSSERL, et désignent ce qui est recueilli par la conscience. Husserl en distinguait trois types : rétentions immédiates,

ou primaires ; et secondaires, à savoir les rétentions passées devenues des souvenirs. Mais STIEGLER en introduit une troisième la rétention tertiaire, qui est une accumulation intergénérationnelle des mémoires individuelles, matérialisées donc au travers des outils techniques. On comprend alors la relation entre savoir et outil technique. Ce processus d’homínisation a permis l’émergence de sociétés et de cultures complexes, et les grecs ont bien résumé cette particularité de l’Homme au travers du mythe de Prométhée. Dans celui-ci, Épiméthée, après avoir créé le vivant, a donné toutes les qualités au animaux, ce qui laisse l’Homme incomplet, car dépourvu d’une spécificité. Prométhée décide alors de voler le feu (la technique) à Héphestos pour le donner aux humains (figure 1.1).



FIGURE 1.1 – *Prométhée dérobant le feu*, Jan COSSIERS.

Nous parlions à l’instant de l’émergence des sociétés dans l’histoire de l’Homme. L’agriculture et l’écriture ont joué des rôles centraux dans ce processus. L’agriculture a concentré des individus autour de bassins fertiles, comme le Nil, et l’écriture dote notre espèce d’outils concrets dédiés à la conservation de la mémoire : des outils mnémotechniques. L’écriture a pu ainsi stabiliser les civilisations naissantes : par exemple, sur le Nil, les égyptiens ont pu la mettre à profit pour observer les étoiles, calculer et anticiper les crues du fleuve, et ainsi concevoir un calendrier adapté qui organise la vie des individus. Cette écriture est d’abord iconique (idéogrammes), mais les signes qui la composent se détachent peu à peu de leur relation au signifié pour devenir plus symboliques. L’apparition des systèmes alphabétiques permet une transmission de la pensée avec une exactitude que

ne permettent pas les systèmes idéogrammatiques. Et c'est cette invention technique qui va permettre l'essor de la culture et du savoir sur les millénaires suivant qui constituent l'Histoire humaine. Dans L'origine de la géométrie, HUSSERL affirme même que l'existence de la géométrie serait rigoureusement impossible sans système d'écriture alphabétique, et on peut le croire sans trop de peine. Les systèmes alphabétiques conservent intacts les processus de pensée, ils sont à même de retranscrire exactement un discours, ce que ne peuvent pas faire par exemple les hiéroglyphes égyptiens.

Nous pouvons avancer jusqu'au 15^e siècle, avec l'arrivée de l'imprimerie qui modifie profondément notre rapport à la diffusion du savoir. Auparavant réservé à une élite érudite, notamment de moines copistes, l'imprimerie permet dans une certaine mesure une plus large diffusion des livres, qui sont déjà depuis Alexandrie un mode privilégié de transmission du savoir, tout du moins dans nos sociétés occidentales. Cette invention préfigure d'ailleurs l'essor des médias modernes, écrits d'abord (presse), puis hertziens (télévision, radio) et numériques (Internet). Dans une temporalité plus courte, elle est aussi en partie à l'origine des écrits de Martin LUTHER au début du 16^e siècle, qui dénonça l'Église chrétienne pour ses vices et sa manipulation des masses, que LUTHER jugeait en décalage avec le message affiché¹. Cette scission donna naissance au protestantisme et aux guerres de religion qui ont suivi, et il est intéressant de constater qu'elle n'aurait sans doute jamais eu lieu sans l'imprimerie, qui a permis de répandre la pensée luthérienne en Europe. Nous pouvons d'ailleurs revenir sur le mythe de Prométhée, sur sa fin que nous n'avons volontairement pas évoqué plus haut : en effet, Zeus a puni les Hommes² en amenant la discorde en leur sein. En effet, incapables de s'entendre sur l'utilisation de la technique, de ces prothèses exosomatiques, les Hommes entrèrent alors en guerre. Et en effet, l'unité des groupes humains est sans cesse remise en cause par les évolutions de la technique : la technique dépasse l'Homme, et l'Homme dépasse la technique en permanence. D'où l'importance d'aborder les questions de justice et de sagesse dans leur utilisation, mais nous y reviendrons.

Le 19^e siècle marque ensuite l'entrée dans l'ère industrielle, et l'émergence du capitalisme avec l'organisation scientifique du travail. Cette période pose un terreau essentiel pour l'entrée dans la consommation de masse, qui amène à l'émergence des industries culturelles dont nous allons maintenant pouvoir parler.

1. Il condamnera notamment fermement le commerce des indulgences, qui permettaient aux plus aisés d'acheter le pardon divin auprès de l'Église.

2. Prométhée fût lui aussi puni par Zeus, qui l'enchaîna au Caucase, condamné à avoir éternellement le foie dévoré le jour par un vautour, et réparé la nuit pour être à nouveau consommé le lendemain. Les dieux ont un sens du châtement particulièrement vicieux.

1.1.2 Émergence des industries culturelles

Le terme "industrie culturelle" a été inventé par deux théoriciens de l'École de Francfort : Theodor ADORNO et Max HORKHEIMER (figure 1.2), au milieu du 20^e siècle, dans [WAH47]. Dans ce livre, ils ont proposé que la culture populaire s'apparente à une usine produisant des biens culturels standardisés : films, programmes radio, magazines, etc. Qui sont utilisés pour manipuler la société de masse. Le terme "industrie de la culture" désigne donc la commercialisation de la culture, la branche de l'industrie qui s'occupe spécifiquement de la production de la culture par opposition à la "culture authentique".



FIGURE 1.2 – Theodor ADORNO et Max HORKHEIMER.

Les membres de l'école de Francfort ont été beaucoup influencés par le matérialisme dialectique et le matérialisme historique de Karl MARX. Dans des ouvrages tels que *Dialectique des Lumières* et *Dialectique Négative*, ADORNO et HORKHEIMER ont théorisé que le phénomène de la culture de masse a une implication politique, à savoir que toutes les nombreuses formes de culture populaire font partie d'une seule industrie culturelle dont le but est d'assurer l'obéissance continue des masses aux intérêts du marché. Le terme "industrie" exclut toutefois la peinture, l'architecture, ainsi que le spectacle vivant pour n'englober que les biens pouvant être industrialisables, par la reproduction d'une industrie. Ainsi, le champ des industries culturelles englobe des entreprises d'édition (édition de livres, journaux, musique...) et de production cinématographique ou télévisuelle.

Hollywood justement, est devenue la capitale de l'industrie cinématographique mondiale parce que les États-Unis l'ont exploitée plus tôt que les autres pays. Ils furent suivis de près par l'URSS, l'Italie fasciste et l'Allemagne nazie. Les États-Unis ont été en avance sur leur temps, à la fois

parce qu'ils devaient intégrer des flux permanents de migrants, esclaves ou "intégrés", et parce que ayant tout à construire d'un pays sauvage, sinon vierge, dont ils avaient éliminé les habitants, ils avaient noué un rapport à la technologie tout à fait nouveau. Cette industrie de la culture a pris forme dominante à Hollywood parce que l'Amérique fut le premier pays d'immigration mais aussi et surtout de colonisation et d'esclavagisme moderne (on pense notamment au commerce triangulaire). Les États-Unis comprirent très tôt le pouvoir des objets temporels audiovisuels parce qu'ils furent confrontés à la question de l'adoption comme aucune autre nation. C'est parce qu'il fallait en permanence projeter le modèle américain aux immigrants fraîchement arrivés aussi bien qu'aux États qu'il fallait maintenir unis, notamment après la guerre de Sécession, que les USA devinrent LE pays du cinéma industriel.

La puissance américaine, bien avant sa monnaie et son armée, c'est donc la force de ces images hollywoodiennes, c'est à dire la capacité à produire des symboles nouveaux, des modèles de vie et des programmes comportementaux par la maîtrise des industries du programme au niveau mondial. Les films et les programmes télévisuels remplacent avantageusement des GI's coûteux à entraîner. De plus, les États-Unis découvrirent très rapidement qu'il s'agissait d'une industrie à fort pouvoir rentable. L'audiovisuel est ainsi, après l'aéronautique, le deuxième poste le plus bénéficiaire dans la balance commerciale américaine.

Ce pouvoir culturel, ils en ont tiré parti au cours de la Seconde Guerre Mondiale. Son issue fût décisivement conditionnée par la maîtrise des technologies de transmission, qui avaient déjà joué un grand rôle dans les tranchées. La guerre psychologique menée sur le "deuxieme front" fut une guerre des médias, et la guerre technologique de la cryptologie et des instruments de calcul permit à l'Angleterre et aux États-Unis de gagner la bataille de l'Atlantique, puis de devancer l'Allemagne nazie pour la production de la bombe atomique.

Après la Libération, lorsqu'ils mirent en oeuvre le plan Marshall, les États-Unis menèrent une politique systématique de diffusion de la culture américaine. Les sommes allouées par l'Amérique aux nations qu'ils aidaient à se reconstruire étaient en particulier conditionnées par une large diffusion du cinéma américain dans ces pays. Elle a utilisé le cinéma comme instrument de guerre psychologique, idéologique et commerciale. Cette guerre des images, après le nazisme allemand, lui a aussi permis de s'opposer au communisme soviétique en imposant un modèle hégémonique au monde : l'American Way of Life. En modifiant les habitudes de consommation et les modèles relationnels, et en faisant vibrer le monde entier pour une histoire particulière (de Charlot, à Apollo 13) l'American Way of Life a

imposé une histoire unique : l'aventure des États-Unis. Les images hollywoodienne d'abord, les feuilletons télévisés ensuite, ont fait de ce pays le pays de la modernité par excellence et le rêve de tous les candidats à l'émigration. En fait, ce que la révolution industrielle avait initié en Europe au 19^e siècle, l'Amérique en hérita exemplairement au 20^e siècle. Après la Libération, elle apparaissait comme le pays où tout est possible, où ce bouleversement inouï qu'avait inauguré l'industrialisation (et qui devait se développer comme processus d'innovation permanente) avait trouvé sa véritable patrie.

Les États-Unis forgent l'image de la modernité à travers Charlot, autant en emporte le vent et Mickey, tout autant qu'avec la haute technologie et les gratte-ciels de Wall Street. L'histoire de l'Amérique du Nord et celle de l'appropriation de la mnémotechnologie de l'imaginaire que comme technologie de calcul et de la logistique. Et c'est aussi plus récemment l'organisation industrielle systématique et raisonnée depuis longtemps de leur convergence en un seul et même système technique intégralement numérique. Cette politique de la technologie est indissociable de cette politique de l'adoption, elle-même à la base de sa politique de l'invention et de la création artistique. La culture de l'adoption qui est le ferment de l'histoire des États-Unis constitue une incomparable capacité à attirer et accueillir l'étranger. Les plus grands cinéastes européens sont d'ailleurs allés travailler à Hollywood.

La publicité a aussi joué un rôle important dans ce processus. Dans [Pac57], Vance PACKARD pose un postulat critique sur la moralité des techniques nouvelles techniques de manipulation de masse dans la publicité. Ces techniques, rassemblées sous une nouvelle "science" dite de la "Recherche des Mobiles", permet de vendre et faire vendre des biens de consommation débordant des usines du monde occidental, en créant les besoins dans l'esprit des consommateurs. En faisant appel au subconscient du public, et en exploitant des affects émotionnels nouveaux, la publicité entre dans l'ère du storytelling, et forme une autre méthode pour diffuser l'American Way of Life et les bienfaits de la consommation de masse. Ces nouvelles publicités jouent énormément sur les images et les symboles : images de marques, images de produits pour les individualiser.

À la fin du 20^e siècle, l'unification par l'image est ainsi devenue essentielle dans un système politique et économique globalisée, où la technologie dans un contexte de guerre commerciale permet à la fois :

- en tant que système planétaire de production unifiée par des normes techniques internationales, la mondialisation de la division industrielle du travail pour la production des biens de consommation, et un dispositif de télécommunication grâce auquel se généralisent dé-

localisations et télémanagement ;

- en tant que système technique planétaire de diffusion d'industrie de programmes, principalement américaines, la constitution de marchés d'hypermasses dans les pays solvables.

Ainsi on observe une hyperindustrialisation, qu'implique la numérisation, qui donne un attrait fondamentalement politique à la question de la culture. C'est cette politique industrielle que l'on retrouve dans la société de l'information, et qui a engendré les industries culturelles modernes.

1.1.3 La question de la reproductibilité technique

(introduction du livre + partie de Ali)

1.2 Les industries culturelles à l'ère du numérique

Nous l'avons vu, le 20^e siècle a été une époque charnière pour l'essor des industries culturelles modernes. Mais à partir des années 60-70, une nouvelle révolution prépare le 21^e siècle : l'informatique. Si BABBAGE pose dès le 19^e siècle les bases de la machine à calculer, c'est pendant et après la guerre que tout s'accélère. Citons pêle-mêle John VON NEUMANN qui définit l'architecture de calculateur qui porte son nom, et qui à la base de toutes les architectures informatiques qui suivirent, et Alan TURING, qui avec son calculateur électromécanique³, parvint à décrypter les messages chiffrés d'Énigma, système de chiffrement allemand, faisant prendre à la guerre un tournant décisif.

Avec l'abandon des parties mécaniques dans les calculateurs, ainsi que les progrès croissants sur la miniaturisation des transistors, leur puissance augmente de manière exponentielle, ce qui a conduit à la fameuse Loi de Moore ; en fait une conjecture, elle stipule que la puissance des ordinateurs double tous les deux ans. Jusqu'à maintenant, cette loi empirique s'est avérée, et elle a accompagné la naissance d'Internet, d'abord comme un projet militaire (ARPA), jusqu'aux foyers des particuliers à partir des années 1990.

Cette "révolution numérique"⁴ a permis l'émergence de nouveaux modes

3. Calculateur à qui VON NEUMANN attribue d'ailleurs la véritable parentalité de l'architecture de VON NEUMANN.

4. Les anglophones parleront volontiers de *digital revolution*, car le terme *digital* en anglais renvoie au numérique (*digits* : les chiffres). Il conviendra de ne surtout pas traduire cela par "révolution digitale", révolution qui ne concerne que l'utilisation que l'on a de nos doigts, et dont nous laissons le soin de l'interprétation au lecteur.

de communication, qui auraient encore été considérés comme de la science-fiction 50 ans avant : e-mails, forums, blogs, la genèse d'Internet s'est d'abord caractérisée par une large décentralisation de l'information, tant dans ses modes d'émission que de transmission. Mais plus récemment, Internet a été le terrain d'émergence de nouvelles gigantesques industries culturelles, que l'on tend à désigner en France sous l'acronyme Google, Apple, Facebook, Amazon, Microsoft (GAFAM), les 5 entreprises majeures de l'Internet "contemporain."

Il est légitime de se demander en quoi Facebook ou Apple peuvent être considérés comme des industries culturelles. Facebook n'est-il pas un réseau social, et Apple un fabricant d'ordinateurs⁵ ? Il se trouve que ces entreprises, malgré des secteurs d'activité différents, partagent plusieurs points communs.

Tout d'abord, ce sont des sociétés américaines, pays qui, comme nous l'avons vu plus tôt, a su imposer une hégémonie culturelle dans son histoire récente. Elles ont chacune, à un certain niveau, une manière de diffuser une certaine culture au sens large du terme. Il s'agit d'une culture avant tout consumériste et individualiste, sur un mode de pensée néolibéral classique. Par exemple, Apple diffuse au travers de ses produits et de sa culture d'entreprise une certaine idée de l'humain, émancipé par la toute puissante technologie numérique, et cet idéal se communique notamment au travers de ses publicités. De plus, elles réunissent sous leur coupe une écrasante majorité du marché, chacune dans leur propre domaine. Facebook par exemple est le premier réseau social mondial, accueillant sur ses serveurs près du tiers de la population mondiale. Enfin, elles font un usage massif des techniques de big data afin d'effectuer un profilage de ses utilisateurs. Les raisons de ce profilage sont diverses, mais la principale est que les plateformes que proposent les GAFAM reposent sur un modèle économique de publicité ciblé, où les utilisateurs paient de leurs données personnelles l'utilisation des plateformes, qui revendent ces données à des annonceurs publicitaires. Cela permet de proposer une publicité ciblée, et potentiellement encore plus efficace que la publicité du temps de Vance PACKARD dont nous avons parlé précédemment.

Nous allons avoir l'occasion de revenir plus en détail sur le fonctionnement et les conséquences de telles pratiques dans la suite de ce rapport, qui étudiera en détail un autre type d'industrie culturelle, plus évident :

5. Apple vend aussi des smartphones, mais il est inutile de le mentionner : les smartphones sont des ordinateurs. Des ordinateurs qui se font passer pour des smartphones, mais des ordinateurs quand même. Cette stratégie qui consiste à cacher à l'utilisateur la nature réelle de ce qu'ils ont dans leur poche au quotidien est cruciale dans la stratégie globale d'Apple et des autres fabricants de smartphones, mais nous ne nous étendrons pas plus sur ce sujet. Cependant, la lecture de [Sco13] apporte quelques points d'analyse très intéressants.

les plateformes de diffusion de contenu culturel.

Chapitre 2

Plateformes de diffusion de masse et algorithmes

Nous allons maintenant nous attacher à comprendre comment les plateformes de diffusion culturelle numériques fonctionnent, notamment dans leur manière de proposer le contenu à ses utilisateurs. Afin de définir clairement les termes de l'étude, arrêtons nous un instant sur ce que nous entendons par plateformes de diffusion culturelle numériques, ou plateformes de diffusion de masse : il s'agit de portails en ligne qui mettent à disposition de leurs utilisateurs des produits culturels (livres, musique, films, séries, etc.) selon un modèle de paiement par produit (Amazon, iTunes store, services de VOD d'Orange ou SFR par exemple), ou par abonnement (Netflix, Spotify, Audible¹, Youtube Red...).

2.1 Comment les algorithmes de profilage sélectionnent le contenu

Avant de parler d'algorithmes de profilage, il nous faut revenir sur la notion d'algorithmes. Ils sont au coeur des systèmes modernes d'information et de communication, mais ne datent pas d'hier. le premier à formaliser la notion d'algorithme est Muhammad Ibn Mūsā AL-KHAWARIZMĪ, mathématicien perse du 9^e siècle : son nom a d'ailleurs donné le mot "algorithme." Cependant, la notion d'algorithme a toujours existé, puisqu'un algorithme n'est qu'une suite d'opérations que l'on effectue pour passer d'un état initial à un état final. Au sens le plus large, un algorithme accepte des informations en entrée, et retourne des sorties, déterminées par la suite d'opérations. Par exemple, une recette de cuisine peut être consi-

1. Audible est une plateforme qui met à disposition des livres audio.

dérée comme un algorithme, travaillant sur des ingrédients en entrée, et retournant un plat en sortie.

Dans sa définition plus stricte cependant, on parle d'algorithmes pour décrire une suite d'opérations logiques mises en oeuvre par des systèmes informatiques. C'est cette mise en oeuvre que l'on désigne sous le nom de programme informatique, et le premier de son genre fût écrit par Ada LOVELACE, mathématicienne britannique du 19^e siècle. Destiné à la machine de Babbage, ce programme permettait de calculer la suite des nombres de Bernoulli. Les travaux de LOVELACE ont été fondamentaux pour la science informatique moderne, et ses apports très importants aux travaux de BABBAGE leur ont donné une envergure qui manquait à ce dernier.

Les algorithmes qui nous intéressent sont des algorithmes de profilage : ce sont des algorithmes de *machine learning* qui reposent sur des modèles statistiques pour résumer de grands volumes de données et découvrir des marqueurs synthétiques qui expliquent le jeu de données. Ils sont utilisés dans de nombreux domaines, et notamment par les plateformes culturelles numériques pour sélectionner le contenu présenté aux utilisateurs. Pour pouvoir parvenir à ce résultat, il ne suffit évidemment pas d'avoir des informations sur le contenu que l'on cherche à proposer : il faut surtout disposer d'un maximum d'informations sur chaque utilisateur : historique de consommation, termes de recherche préférés, mots-clefs les plus utilisés dans les commentaires, si la plateforme dispose d'un espace de commentaires, etc. Ces informations sont autant de variables que l'on peut quantifier, et qui définissent les données brutes d'entrée pour nos algorithmes de profilage. Ces données sont difficilement appréhendables par l'esprit humain : les plateformes peuvent avoir des millions d'utilisateurs, caractérisés par des centaines, voire des milliers d'indicateurs bruts. L'objectif d'un algorithme de profilage est alors de réduire l'information utile à l'essentiel, pour essayer d'extraire des profils types de consommateurs dans lesquels on pourra ranger les utilisateurs. Ainsi, ces systèmes algorithmiques fonctionnent en deux temps : synthèse des données et classification (*clustering* en anglais).

Les outils statistiques à disposition des concepteurs d'algorithmes de profilage sont nombreux, et nous allons en présenter rapidement un qui permet d'effectuer la synthèse des données. L'outil en question s'appelle l'analyse en composantes principales (ACP), et il permet de réduire la dimensionnalité d'un espace vectoriel de grande dimension, en passant à un sous-espace vectoriel de dimension moindre, et dont les axes sont les axes principaux du jeu de données, c'est à dire des axes qui résument suffisamment bien les données d'entrée. Pour reprendre l'exemple cité plus haut, nos utilisateurs précédemment introduits sont résumés par de multiples

variables numériques, que l'on peut considérer comme autant d'axes d'un espace vectoriel. Chaque utilisateur est alors un point de cet espace vectoriel. Parmi ces variables, il est possible que certaines soient plus ou moins corrélées, ce qui signifie que l'on pourrait les combiner pour avoir une information résumée. Par exemple, sur la figure 2.1 on est dans un espace de dimension 3, avec un nuage de points qui semble répartir sur un plan ; alors l'ACP permet de passer à un sous-espace de dimension 2, à deux variables qui sont une combinaison des 3 variables de départ. On a ainsi synthétisé l'information, ce qui permet d'en simplifier la classification.

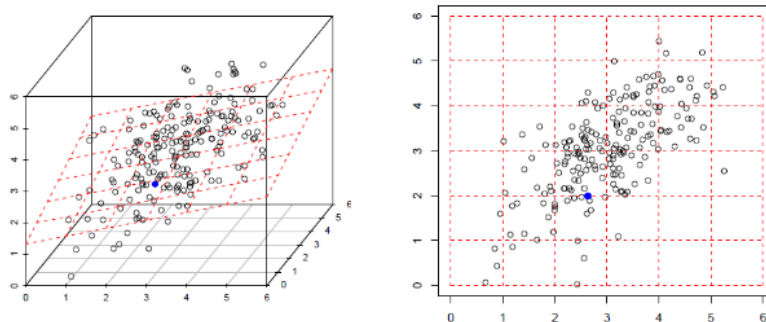


FIGURE 2.1 – Principe de l'ACP.

La classification permet ensuite, à partir de ce nuage de points, de former des catégories dans lesquelles on pourra positionner les utilisateurs, mais aussi le contenu qui leur est proposé. On pourra non seulement suggérer à l'utilisateur de nouveaux contenus qui correspondent à ce que sa catégorie semble apprécier généralement au travers de leurs habitudes de consommation. Ce n'est pas autrement que par l'étude des corrélations qu'Amazon par exemple remplit les rubriques de suggestions d'achat personnalisées.

2.2 Conséquences

Ce traitement automatisé de l'information n'est pas sans conséquences, et nous allons essayer maintenant de présenter les différents problèmes que cela soulève.

2.2.1 Uniformisation des modes de pensée

HORKHEIMER et ADORNO soutiennent que le divertissement de masse vise, de par sa nature même, à attirer un vaste public et, par conséquent, à la libération de base de l'art de bas niveau. Ils ne suggèrent pas que tous les

produits de ce système sont intrinsèquement inférieurs, simplement qu'ils ont remplacés d'autres formes de divertissement. Ils font des comparaisons entre l'Allemagne fasciste et l'industrie cinématographique Américaine, et mettent en évidence la présence d'une culture de masse, créée et diffusée par des institutions exclusives et consommée par un public passif et homogénéisé dans les deux systèmes. Cela illustre la logique de domination dans la société moderne post siècle des Lumières, par le capitalisme monopoliste ou l'État-nation.

Livrée au marché, la culture industrialisée est instrumentalisée pour le développement d'un nouvel esprit, celui de la modernité, plus moderne que jamais, de mode de vie américain où le calcul logistique est désormais devenu tout à fait hégémonique, ce qui se concrétise par la liquidation de l'État, incommensurablement accentuée par la numérisation et par laquelle les industries de l'information et la communication atteignent le stade de l'hyperindustrialisation de la culture, où Internet permet à la fois la production de biens de consommation, de leur promotion (nouvelles industries de programmes issues de la convergence des technologies) et de leur diffusion (commerce électronique), où le récepteur de télévision se transforme en un organe de téléaction, évolutions par lesquelles le système technique devient véritablement impérial et planétaire.

Ainsi, la consommation de masse des oeuvres de la culture populaire rendrait les gens dociles et contents, quelle que soit la situation économique des consommateurs, qu'ils soient riches ou pauvres. Il s'agirait alors d'une manipulation de masse qui vise à contrôler l'opinion publique pour accomplir un certain but ou une certaine idéologie, en l'occurrence le capitalisme.

L'autre danger inhérent à l'industrie de la culture est la culture de faux besoins psychologiques qui ne peuvent être que satisfaits par les produits du capitalisme ; ainsi ADORNO et HORKHEIMER ont surtout établi la culture de masse comme dangereuse pour les arts supérieurs plus difficiles sur le plan technique et intellectuel. En revanche, les véritables besoins psychologiques seraient la liberté, la créativité et le bonheur authentique.

Les objectifs de l'industrie de la culture sont, comme dans toute industrie, de nature économique. Tous les efforts se concentrent sur les bénéfices économiques. La culture "authentique", cependant, n'est pas orientée vers un but, mais est une fin en soi. La culture "authentique" favorise la capacité de l'imagination humaine en présentant des suggestions et des possibilités, mais d'une manière différente de celle de l'industrie de la culture puisqu'elle laisse place à la pensée indépendante. La culture "authentique" n'est pas canalisée dans la réalité régurgitante, mais va au-delà. La culture "authentique" est unique et ne peut être forcée dans aucun schéma préformé. Quant à la découverte des causes du développement de l'industrie

culturelle, HORKHEIMER et ADORNO affirment qu'elle découle de la poursuite par les entreprises de la maximisation du profit, au sens économique du terme.

2.2.2 Mal-être et aliénation

Pour STIEGLER, le cinéma hollywoodien est une "barbarie esthétique" qui gangrène tous les secteurs de la production industrielle, à l'unique fin de "marquer les sens des hommes de leur sortie de l'usine, le soir, jusqu'à leur arrivée à l'horloge de pointage, le lendemain". Pour HORKHEIMER et ADORNO, le cinéma hollywoodien associé à la radio et aux magazines est le produit d'un dispositif d'aliénation où les "autos, les bombes et les films assurent la cohésion du système". Il s'agit en fait d'un divertissement pascalien.

En effet, selon PASCAL, le divertissement fait référence aux activités humaines futiles (recherche du succès ou des biens matériels) pour échapper à notre condition humaine. L'homme est prédestiné à vivre une existence de mortel et de contingence, c'est pourquoi il recherche désespérément et constamment une consolation face à la difficulté d'être soi. Il s'agit de se détourner d'une réalité déplaisante. Cependant cette dernière n'est pas un mal circonstanciel comme un deuil, échec professionnel ou amoureux mais un malheur constitutif de notre existence. Notre condition est celle d'un être faible, mortel, exposé à la maladie, à la solitude. "Les hommes n'ayant pu guérir la mort, la misère, l'ignorance, ils se sont avisés pour se rendre heureux de n'y point penser". Ainsi le cinéma associé aux autres formes d'industrie culturelle est une tentative pour s'échapper de notre condition.

Le cinéma est un moyen de se détourner de l'ennui et la solitude. Le temps de soi est toujours celui des autres. En effet le flux de conscience est contraction du temps, le cinéma peut déclencher ce processus d'adoption où mon temps, durant le temps d'un film devient le temps d'un autre et un autre temps. Mon temps se construit sur le temps qu'il prélève aux autres. C'est pourquoi la solitude est si difficile à supporter. Dans la solitude où l'autre fait défaut, il n'y a plus de temps, "rien ne se passe", "rien n'arrive", je rencontre l'ennui car je ne rencontre que la coquille vide d'un moi que le temps de l'autre ne porte plus. Si dans ces maussades dimanches après-midi, la distraction cinématographique ou télévisuelle peut me procurer un échappatoire, c'est parce qu'il vient m'altérer et me désaltérer (me décontracter), me remonter (c'est une sorte de remontant).

C'est pourquoi HORKHEIMER et ADORNO accusent ainsi le cinéma de paralyser l'imagination et plus généralement le discernement du spectateur au point que celui-ci n'est plus en mesure de distinguer "perception" et "ima-

gination", réalité et fiction, discours qui pourrait s'appliquer aujourd'hui tel quel à la réalité virtuelle ou aux jeux électroniques.

Ainsi plus l'industrie culturelle réussit à donner par ses techniques une reproduction ressemblante des objets de la réalité, plus il est facile de faire croire que le monde extérieur est le simple prolongement de celui que l'on découvre dans le film. L'introduction subite du son a fait passer le processus de reproduction industrielle entièrement au service de ce dessein. Il ne faut pas que la vie réelle puisse se distinguer du film. Ainsi l'industrie est source d'assujettissement et d'aliénation de l'individu.

HORKHEIMER et ADORNO affirment que la culture industrielle prive les gens de leur imagination et prend le contrôle de leur pensée pour eux. L'industrie de la culture livre les "biens" de sorte que les gens n'ont plus qu'à les consommer. Par la production de masse, tout devient homogénéisé et ce qui reste de la diversité est constitué de petites trivialités. Tout se comprime à travers un processus de schémas en partant du principe que le mieux est de refléter le plus fidèlement possible la réalité physique.

Les films en sont un exemple. "Tous les films sont devenus similaires dans leur forme de base. Ils sont façonnés de manière à refléter le plus fidèlement possible les faits de la réalité. Même les films fantastiques, qui prétendent ne pas refléter cette réalité, ne sont pas vraiment à la hauteur de ce qu'ils prétendent être. Même s'ils s'efforcent d'être inhabituels, les fins sont généralement faciles à prévoir en raison de l'existence de films antérieurs qui ont suivi les mêmes schémas."

2.2.3 Discriminations algorithmiques

On parle ici de discrimination au sens où les algorithmes reproduisent des mécanismes de séparation d'un individu du groupe social, sur la base de critères fondamentalement arbitraires. Quand elle est du fait d'algorithmes, cette discrimination est souvent discrète et pernicieuse, et peut se cacher dans des détails d'apparence anodine. Il est facile de se bercer du mythe comme quoi les algorithmes éliminent l'humain du processus de décision, et qu'ils sont donc neutres, objectifs, impartiaux, tels des juges divins. Après tout, ils ne reposent que sur des données brutes, par essence neutre, sans jugement de valeur, non ? Cependant, les algorithmes sont écrits par des développeurs, humains, faillibles, qui lui insufflent une part de subjectivité et d'arbitraire. De plus, les données sont des données créées par d'autres êtres humains, avec des comportements variables et complexes, et ces algorithmes cherchent à apprendre à partir de ces comportements. IL n'est donc pas étonnant de les voir reproduire des défauts de fonctionnement de la pensée humaine, mécaniquement, par réplication, puisque

ces algorithmes ne font que moyenner toutes ces pensées, par corrélation statistique.

On peut penser à de nombreux exemples pour illustrer ce problème. Ainsi, dans [Bar15], Le Wall Street Journal racontait l'histoire de l'algorithme de reconnaissance d'images de Google, qui avait classé la photo d'une femme noire comme étant un "gorille". Dans le même ordre d'idée, des recherches de l'Université de Harvard en 2013 ont montré que les recherches de noms à consonnance africaine dans des moteurs de recherche tendent à montrer plus de faits divers d'arrestation. On peut aussi s'amuser à chercher des noms de métiers très genrés dans un moteur de recherche d'images, et constater les différences de résultats : "CEO" ressort en majorité des photos d'hommes blancs (seulement 4 femmes sur la première page, contre 24 hommes dans notre test), alors que "cashier" (caissier/caissière) montre principalement des femmes (20 femmes contre 8 hommes). Cela s'explique par le fait que ces types d'algorithmes ne travaillent pas à partir de rien, mais qu'il se basent sur un corpus d'images existantes. Dans le cas de la recherche d'images, on trouve bien souvent des photos dites "stock images", créées par des banques d'images destinées généralement à la communication des entreprises. Et ces banques d'images ne créent elles-mêmes pas d'images de nulle part, elles se basent sur des normes sociétales pré-existantes (les caissiers sont souvent des caissières, les dirigeants sont en majorité des hommes...).

On peut regarder du côté des affiches de *The autocomplete Truth*, campagne de l'ONU sur les droits des femmes, qui met en avant les stéréotypes véhiculés par l'algorithme d'autocomplétion de Google (figure 2.2). Ces algorithmes se basent sur les recherches les plus populaires pour proposer des termes de recherche qui complètent les premiers mots que l'on peut saisir dans un moteur de recherche. Si ces recherches ont pu être à la base propulsées par des personnes mal intentionnées (on parle de bombardement google, quand des groupes d'influence propulsent des recherches ou des pages dans le référencement Google pour véhiculer une idéologie, ou pour s'amuser), l'apparition de ces recherches dans l'autocomplétion d'autres utilisateurs peut induire des recherches qui n'auraient jamais été faites sans : on a là un cercle vicieux, où les termes de recherches populaires alimentent les termes de recherche populaires.

On constate ainsi comment les algorithmes peuvent orchestrer ce que Pietro MONTANI a décrit comme une politisation de l'esthétique, dans [Mon13]. Des processus, apparemment neutres, basés sur des données, apparemment neutres, peuvent produire des résultats fondamentalement orientés, vecteurs d'un sens politique profond.

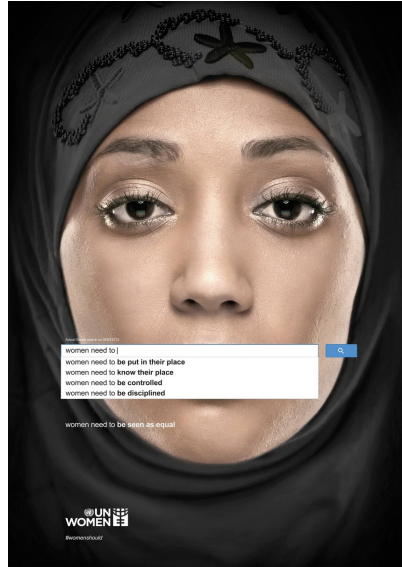


FIGURE 2.2 – Extrait de la campagne *The Autocomplete Truth* de l'ONU.

2.3 Étude de cas : Netflix

Netflix est un service de vidéo à la demande américain qui propose de visionner des vidéos et des films sur internet, en streaming et en illimité. L'entreprise a été fondée à Los Gatos en 1997 aux États-Unis par Reed HASTINGS. Au 31 décembre 2017, Netflix compte 117,6 millions d'abonnés payants à son offre de streaming, dont 54 millions d'abonnés aux États-Unis. Le tarif d'un abonnement est situé entre 8 et 14 euros. Le forfait varie en fonction de la qualité d'image proposée standard, HD ou avec ultra HD et en fonction du nombre d'écrans accessibles simultanément. Netflix peut représenter jusqu'à 33% de bande passante aux États-Unis.

2.3.1 L'algorithme de suggestion de Netflix

L'algorithme de suggestion de Netflix est un produit breveté et unique à Netflix. Ils n'ont jamais cédé les droits à d'autres services.

En 1992, le premier système de recommandation (pour des articles d'Usenet) a été utilisée par 2 chercheurs en informatique (RESNICK et RIEDL) pour collecter des notes données par les utilisateurs lorsqu'ils lisaient des articles. Ces notes étaient utilisées pour évaluer la probabilité que ce dernier soit apprécié par les personnes n'ayant pas encore lu l'article. Cette recommandation automatisée repose, comme hors ligne, sur l'évaluation par les premiers lecteurs : si les amis de Nicolas lui suggèrent de lire un livre qu'ils ont aimé, il est probable qu'il le lise et l'apprécie aussi, davantage qu'un livre que ses proches n'auraient pas lu ou pas aimé.

C'est Neil HUNT (figure 2.3), directeur technique de la plateforme de streaming américaine, qui en charge du fameux algorithme, pièce maîtresse du succès du site, permettant notamment de gérer les recommandations. Son département représente à lui seul 8% des investissements de Netflix.



FIGURE 2.3 – Neil HUNT, directeur technique de Netflix.

Celui-ci déclare à ce sujet : "Si je ne peux pas réduire le catalogue à 50 titres qui vous intéressent, techniquement j'ai perdu la partie et Netflix ne remplit pas son rôle." Dans une interview, il évoque l'idée d'optimiser ce catalogue de 50 titres déterminé à partir du "goût" de ses utilisateurs. En effet, il parle de "faire en sorte que (...) vous vous asseyez dans votre canapé pour trouver quelque chose à regarder rapidement". Il s'agirait alors de consommer une oeuvre, un film, et rapidement si possible.

De plus, lors de l'inscriptions de ses clients, Netflix demande à ces derniers de choisir parmi 3 films ou séries pour établir une base de titres. Par exemple, pour quelqu'un qui choisira un film d'action, des "films d'aventures", "films épiques" et "d'actions" sont proposés et pour un client qui choisira des comédies romantiques, la base de titres ne sera pas la même. La sélection sera notamment des "teen movie", des "séries romantiques" ou "comédies familiales".

Enfin il est évoqué que 75% des programmes sont consommés suite aux

propositions par l'algorithme de recommandation et non par le moteur de recherche. Cela en dit long sur l'efficacité de l'algorithme de suggestion.

2.3.2 Fonctionnement de l'algorithme de suggestion

Plus de 80% des émissions de télévision que les gens regardent sur Netflix sont découvertes grâce au système de recommandation de la plateforme. Cela signifie que la majorité de ce que les gens décident de regarder sur Netflix est le résultat de décisions prises par une mystérieuse boîte noire d'un algorithme.

Il y a deux idées principales en jeu sur le système de recommandation de Netflix, et elles proviennent toutes les deux de ce que Netflix a appris en sondant les données des utilisateurs au fil des ans. Premièrement, ils savent que la plupart de leurs utilisateurs ne veulent pas perdre trop de temps à chercher quelque chose à regarder. "La personne typique ne va pas regarder des milliers de titres, elle va regarder une moyenne de 40 à 50 titres à chaque session ", dit Todd YELLIN, vice-président produit chez Netflix.

Netflix a donc une petite fenêtre de temps dans laquelle il peut piquer l'intérêt du téléspectateur. Leur objectif principal est de s'assurer que les premières choses que voit le téléspectateur lorsque celui-ci se connecte sont des titres que ce dernier veut potentiellement regarder. Deuxièmement, ils ont appris en cours de route que ce que les usagers disent sur la façon dont ils utilisent le service et leur comportement réel ne sont pas toujours corrélés. "Beaucoup de gens nous disent qu'ils regardent souvent des films ou des documentaires étrangers. Mais dans la pratique, cela n'arrive pas beaucoup", a déclaré Carlos GOMEZ-URIBE, ancien vice-président de Netflix en charge de l'innovation produit lors d'un entretien avec Wired en 2013. Cela rejoint d'ailleurs les propos développés par Vance PACKARD dans [Pac57], dont nous avons déjà parlé, et qui affirmait que les affirmations conscientes des consommateurs sont de très mauvais indicateurs de leurs comportements de consommation réelle.

De même, ils savent que le téléspectateur peut choisir d'évaluer un documentaire intelligent que ce dernier regardé une fois, avec 5 étoiles, alors qu'il peut donner une note plus basse, ou aucune note, au film qu'il a regardé quatre fois cette année. C'est probablement l'une des deux raisons pour lesquelles ils ont décidé de supprimer le système de classement par étoiles au profit d'un modèle pouces en haut, pouces en bas. Il n'est plus possible de déterminer ce que les autres utilisateurs de Netflix pensent d'un spectacle (les notes du pouce ne sont pas visibles mais vont vers un classement "match"), ce qui signifie que la probabilité que le téléspectateur

apprécie un titre est basé sur les algorithmes mentionnés. Par conséquent, ce n'est pas un hasard si cela signifie que les projets Netflix d'un million de livres ne risquent plus d'être étiquetés avec un mauvais classement par étoiles pour que tout le monde puisse les voir.

Quant au classement en lui-même, il consiste en termes simples : l'usage d'étiquettes. Un certain nombre d'employés de Netflix sont payés pour regarder tous les titres et noter un certain nombre d'éléments déterminants qui se produisent. Par exemple, un film tel que *Wall-E* est étiqueté comme suit : Des dialogues chaleureux, clairsemés, satiriques, etc. Il peut y avoir n'importe quel nombre d'étiquettes, et plus il y en a, mieux c'est. Puis les algorithmes entrent en jeu. Plus le téléspectateur regarde Netflix, mieux il vise à comprendre ses goûts en compilant un profil basé sur des tags récurrents dans les émissions que ce dernier regarde.

Donc, si le téléspectateur a regardé *Jessica Jones* de Marvel, qui peut être étiqueté comme sombre, avec une forte personnalité féminine entre autres choses, il est tout à fait probable que *Orange Is the New Black*, autre série sombre avec des personnages féminins marquants, arrive dans les suggestions de l'utilisateur.

Ainsi chaque catégorie sur la page d'accueil du téléspectateur est personnalisée en fonction de son comportement de visionnage, en poussant vers l'avant le contenu qui correspond aux modèles qu'il a dessinés sans le savoir. Les algorithmes prennent également en compte des informations spécifiques sur l'utilisateur : le type d'appareil sur lequel vous regardez et les heures auxquelles vous avez tendance à regarder par exemple. "Nous prenons toutes ces étiquettes et les données sur le comportement de l'utilisateur, puis nous utilisons des algorithmes d'apprentissage machine très sophistiqués qui déterminent ce qui est le plus important, ce que nous devrions peser", dit M. YELLIN.

Les téléspectateurs s'inscrivent alors dans de multiples groupes de goûts (et il y en a des milliers sur Netflix), et ce sont ces derniers qui affectent les recommandations qui apparaissent en haut de l'interface des utilisateurs à l'écran. Les étiquettes utilisées pour les algorithmes de recommandation sont les mêmes partout dans le monde. Cependant, un sous-ensemble plus petit d'étiquettes est utilisé en fonction du pays, de la langue et du contexte culturel. "Il faut les localiser de manière à ce qu'elles aient un sens", explique M. YELLIN. Par exemple, le mot *gritty* (comme dans *gritty drama*) ne peut pas se traduire en espagnol ou en français.

Le nouveau système permet aux téléspectateurs de recommander des spectacles qui n'ont été appréciés que par un nombre relativement restreint de personnes en se basant uniquement sur le fait qu'ils portent les mêmes étiquettes que les spectacles que ces derniers ont aimé.

2.3.3 Confidentialité des données

La question de la protection de la vie privée a fait la Une des journaux au moment de l'écriture de ce rapport (avril-juin 2018), depuis la révélation que les données personnelles d'environ 87 millions d'utilisateurs de Facebook ont été recueillies par la société de conseil Cambridge Analytica.

De plus, un certain tweet posté par Netflix (figure 2.4) a fait polémique fin 2017.



FIGURE 2.4 – Tweet polémique de Netflix.

Le tweet s'est moqué de quelques dizaines de clients anonymes qui ont passé les trois dernières semaines à regarder un film de Noël un peu dépassé. Le tweet, qui comptait au moment de la rédaction plus de 100000 retweets, est devenu l'un des tweets les plus populaires du site de streaming en quelques heures seulement après son affichage.

Il y a eu de dizaines d'atteintes à la protection des données mais des millions de personnes touchées. En disant cela, Netflix a non seulement admis que l'entreprise peut déterminer avec précision combien de ses plus

de 100 millions de clients regardent une certaine émission sur une certaine période de temps, mais aussi que certains employés ont accès à ces données de visionnement.

Cependant il n'est pas surprenant que Netflix recueille des données sur ses utilisateurs, car elle utilise l'analyse des données pour recommander de nouveaux spectacles de façon algorithmique et pour aider à améliorer ses services. Mais certains ont fait valoir que Netflix abusait effectivement de sa position privilégiée pour faire des blagues sur ses propres clients. "Mes habitudes d'écoute personnelles ne sont pas du fourrage pour les tweets", a déclaré un utilisateur de Twitter en réponse au tweet.

Un porte-parole de Netflix ne répondait pas à ces questions spécifiques, mais envoyait à ZDNet une déclaration en conserve. "Le respect de la vie privée de nos membres est important pour nous", a déclaré le porte-parole. "Ces informations représentent les tendances générales de visionnement, pas les informations personnelles d'individus spécifiques et identifiés."

En toute honnêteté, Netflix n'est pas la première entreprise à utiliser sa vaste richesse de données à des fins de marketing ou de publicité. L'année dernière, Spotify a lancé une campagne publicitaire qui a repris des statistiques plus bizarres et plus étranges de l'année de streaming de la société, comme, "Chers 3749 personnes qui ont diffusé en streaming 'It's the End of the World as We Know It' le jour du vote Brexit, accrochez-vous là".

L'utilisation de données anonymes de clients en vrac pour la publicité n'est pas illégale, tant que l'information n'est pas attachée publiquement au nom d'un client spécifique. De même, en vertu d'une loi américaine peu connue de 1988, la *Video Privacy Protections Act*, Netflix et d'autres sociétés de streaming sont interdites de divulguer les habitudes de visionnement d'un consommateur sans son consentement.

2.3.4 Netflix et politisation de l'esthétique

Netflix est secrètement très bon pour faire en apprendre plus sur les questions politiques. Il peut faire véhiculer des messages politiques sans qu'on s'en rende compte. D'abord les émissions que le téléspectateur regarde à l'aveuglette racontent des histoires plus importantes. De nombreuses séries sur la politique existent sur Netflix comme "House of Cards", mais même les émissions qui ne parlent pas explicitement de politique aident à en apprendre plus sur l'Amérique et la politique.

"Jane the Virgin" et "Orange is the New Black", par exemple, nous enseignent subtilement des sujets comme l'immigration, l'incarcération de masse, le racisme systémique et les inégalités économiques et sociales. Ainsi, des téléspectateurs qui diraient probablement non si quelqu'un leur deman-

daît de regarder une émission sur la réforme de l'immigration, accepteraient volontiers de regarder la charmante "Jane the Virgin" sans forcément avoir conscience de la portée politique de l'oeuvre.

Les séries se déroulant dans le passé, comme "Mad Men", ne concernent pas seulement l'environnement de la série, ils touchent à des problématiques de masculinité et de féminisme, de luttes pour les droits civiques, et la manière dont les événements d'actualité affectent la vie humaine. Et ils enseignent aux téléspectateurs les aspects les plus quotidiens de l'histoire américaine, les petites nouvelles que ceux-ci n'apprennent pas dans les cours d'histoire, comme les crimes et les scandales.

"Daredevil", distribué par Netflix, série à succès de Marvel, comporte des thématiques politiques dans le sens où le héros, aveugle, apparaissant au premier abord comme faible, a la rare humanité de se concentrer sur son quartier et de chercher à combattre la pauvreté et la criminalité. C'est un croyant qui se refuse à tuer et qui devient avocat, se promettant de lutter contre la corruption au nom de la primauté du droit. La question du droit et l'inefficacité parfois de la justice est régulièrement remise en lumière.

Luke CAGE, qui a aussi vu le jour sur Netflix, est "une tentative de réflexion sur les conditions de vie de la communauté noire face à une intensification des violences" et grâce à l'intensité et charisme de son personnage, la série montre subtilement la situation des violences policières aux Etats-Unis et la question du racisme ambiant.

Lorsque les téléspectateurs regardent un film étranger, même s'il s'agit d'une comédie romantique, ils en apprendront aussi sur la culture sans avoir à en regarder une émission culturelle. On comprend ainsi mieux comment des messages politiques externes peuvent percer dans les suggestions Netflix, suggestions que l'on pensait au début de notre analyse comme purement personnelles et indépendantes de processus autres que le seul spectateur. La politique fait irruption dans le champ du goût.

Évidemment, il n'est pas du tout négatif d'éduquer les gens à des problématiques de racisme, de sexisme, etc. Mais il est important de comprendre que d'autres messages politiques peuvent passer à travers ces plateformes, notamment la diffusion d'un certain idéal de société, fondamentalement libéral et basé sur la consommation.

On peut aussi, avec ce type d'outil, faire l'opération inverse, et identifier des opinions politiques chez des consommateurs. On sait par exemple qu'en France, de manière un peu caricaturale, les gens de droite regardent TF1, et les gens de gauche regardent France 2. Avec Netflix, suivant les séries, on sait reconnaître les personnes par leurs opinions politiques, sociales, sociétales, et faire un profil très précis de chaque personne. De la même manière, Facebook est capable dans une certaine mesure d'identifier

l'orientation sexuelle de ces utilisateurs. On imagine aisément les conséquences que de telles informations pourraient avoir si elles tombaient aux mains de gens ou d'états moins "démocratiques" et ouverts que nos sociétés actuelles. Il ne faut en effet pas prendre nos démocraties comme acquises. Et même sans rentrer dans l'idée d'une dérive totalitaire de nos états occidentaux, on peut aussi penser aux risques de piratage, à la surveillance d'employeurs sur leurs employés (l'homophobie existe encore dans le milieu de l'entreprise !)...

On constate donc un problème avec notre culture de l'information, et nous allons rapidement étudier dans la troisième partie quelques pistes pour revoir notre copie sur la culture de l'information.

Chapitre 3

Vers une autre culture de l'information

Nous venons de dresser un tableau critique de l'utilisation actuelle des algorithmes de profilage par les industries culturelles contemporaines. Avant de conclure ce rapport, proposons quelques pistes de réflexion pour envisager différemment nos usages de l'analyse de données.

3.1 Penser l'informatique pour une gouvernance citoyenne : logiciel libre et éthique des données

Il serait trop facile d'affirmer que nous sommes libres dans le choix de l'utilisation ou non de ces plateformes, et de résumer le problème à une responsabilité individuelle. En effet, nous savons qu'aujourd'hui ces outils sont au centre de nos vies, qu'on le veuille ou non. Il suffit de regarder les étudiants de l'UTC : l'essentiel de notre communication passe par Facebook, et ne pas l'utiliser revient à s'isoler de la communauté. Cependant, il serait possible d'imaginer un autre rapport à ces outils, et c'est cet espace des possibles qu'explore la communauté du logiciel libre depuis sa naissance.

Le logiciel libre est d'abord un logiciel diffusé sous une licence qui facilite sa modification et sa redistribution. Pour ce faire, le code source (la mise en oeuvre de l'algorithme qui régit le fonctionnement du logiciel) est accessible, ce qui permet de contribuer à son amélioration, mais aussi de l'auditer, afin de s'assurer qu'il ne présente aucune faille, et qu'il ne comporte pas de comportements douteux ou irrespectueux des utilisateurs (moyens détournés d'accès à leurs données, publicité cachée, etc.). Il s'op-

pose ainsi aux logiciels propriétaires, dont la diffusion, sans être forcément restreinte, se fait sans possibilité d'accès au code source, qui reste jalousement protégé par l'entité qui édite le logiciel. Cela prive les utilisateurs de garanties sur la bonne foi du développeur quant à la santé de son logiciel.

Le logiciel libre est aussi une philosophie et un mode de diffusion des savoirs, basé sur le principe que les idées ne peuvent être possédées et privatisées, et qu'elles devraient pouvoir circuler librement. Ces principes sont, faut-il le rappeler, des principes fondateurs d'Internet, et permettent notamment de concevoir des systèmes informatiques moins centralisateurs, où l'on ne met pas tous nos oeufs dans le même panier. En effet, la nature même du logiciel libre favorise une pluralité de solutions interopérables, par opposition à l'écosystème propriétaire qui privilégie un oligopole de quelques solutions généralement non-interopérables, ou basées sur des normes non standardisées.

Prenons l'exemple des réseaux sociaux, et quittons un peu Facebook pour l'occasion. Twitter, réseau social de micro-blogging, est un autre GAFAM majeur, bien que loin derrière Facebook en termes d'utilisateurs. Il possède cependant toutes les caractéristiques de l'industrie culturelle moderne : une grande entreprise qui instaure un modèle hégémonique, un modèle économique basé sur l'exploitation des données personnelles de ses utilisateurs, et une architecture propriétaire centralisée. Il est impossible d'auditer le fonctionnement interne de Twitter, et être inscrit à Twitter signifie lui fournir ses données personnelles. On retrouve ainsi tous les problèmes évoqués précédemment. Comparons maintenant Twitter à une alternative libre, Mastodon. Projet lancé par Eugen ROCHKO, développeur allemand, il s'agit d'un réseau social de micro-blogging qui repose sur des normes de communication standardisées : OStatus et ActivityPub. En effet, au contraire de Twitter, Mastodon ne repose pas sur une architecture centralisée, où tous les serveurs sont aux mains d'une seule entité. Le fonctionnement de Mastodon est dit fédéré : n'importe qui peut en installer une instance, qui communique avec d'autres instances avec les protocoles précités. On peut comparer son fonctionnement aux e-mails : de la même manière qu'une adresse en `@protonmail.com` peut sans problème envoyer et recevoir des courriels d'une adresse en `@gmail.com`, un utilisateur de Mastodon sur l'instance `octodon.social` peut communiquer de manière transparente avec un utilisateur sur l'instance `mastodon.social`. Mieux encore, comme les standards utilisés par Mastodon sont totalement ouverts, les utilisateurs de ce dernier peuvent communiquer avec d'autres plateformes implémentant les mêmes protocoles : on peut citer GNUSocial (plateforme de micro-blogging plus ancienne), ou Peertube (une alternative à Youtube basée sur le même principe de fédération). En outre, le code

de Mastodon est bien évidemment accessible et ouvert aux contributions grâce au logiciel de développement collaboratif `git`.

Ce que l'on constate, c'est que ce genre d'initiative remet l'utilisateur au coeur des processus techniques de décision. Toujours sur Mastodon, de nombreuses fonctionnalités ont été implémentées afin de répondre à des critiques récurrentes à l'encontre de Twitter, notamment sur sa modération très laxiste des contenus illégaux (racisme, homophobie, néo-nazisme...). Elles permettent à chaque instance d'adopter sa propre politique de modération, ce qui a fait émerger des instances plus communautaires, centrées autour de points d'intérêt, sans pour autant être coupées du monde extérieur : on peut citer par exemple des instances pour les artistes, pour les personnes LGBTI, etc.¹ Cette initiative permet ainsi de retisser du lien social dans des espaces protégées de l'ingérence des géants d'Internet dont nous avons parlé tout au long de ce rapport. Ils constituent une culture alternative basée sur des valeurs de partage et d'éthique.

Tempérons tout de même nos ardeurs : le logiciel libre ne saurait être une solution complète, et on constatera d'ailleurs non sans ironie que le logiciel libre réussit à convaincre même les GAFAM : le rachat tout récent de Github² par Microsoft et la participation de Google à de nombreux logiciels libres illustrent concrètement cet intérêt croissant du secteur pour une approche qu'ils résument généralement à un moyen de développer plus vite et moins cher, en déléguant le travail à des bénévoles à travers le monde. Non, pour penser efficacement une autre approche des systèmes d'information, il faut se rappeler ce que nous disions au début du rapport sur la destruction des savoirs.

3.2 Les technologies sont un pharmakon

Bernard STIEGLER, dont nous avons parlé plusieurs fois dans ce rapport, est un chercheur en philosophie spécialisé dans les changements de la société face au numérique. Il est aussi enseignant à l'UTC où il présente l'UV IC01. Dans cette matière, il présente l'histoire des industries culturelles à travers une grille de lecture héritée de l'école de Francfort. Il propose l'idée que les systèmes décisionnels actuels sont une conséquence de plus d'un processus qu'il appelle la prolétarianisation : la perte de savoir individuel. En effet, l'automatisation de la prise de décision avec les big datas amène à son paroxysme une situation dans laquelle l'individu n'est plus citoyen dans la société, puisqu'il est privé de connaissances et de sens

1. Il existe même une instance hébergée sur Etalab (mission interministérielle dédiée à coordonner l'*open data* des données publiques), réservée aux détenteurs d'adresses e-mail en `.gouv.fr`.

2. Forge logicielle dédiée au développement collaboratif et basée sur `git`.

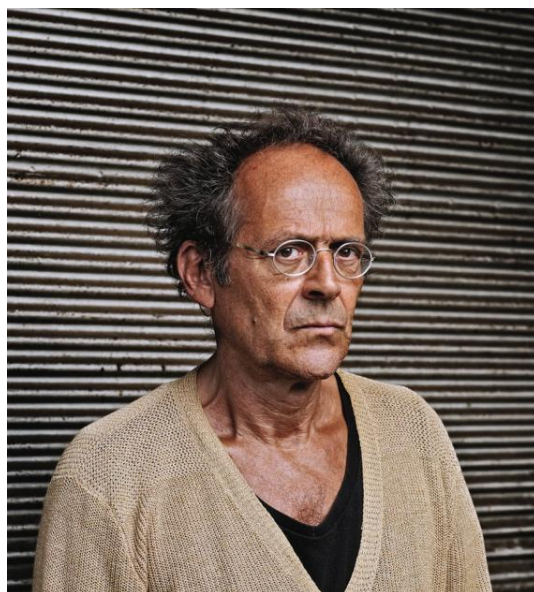


FIGURE 3.1 – Bernard STIEGLER, enseignant à l'UTC et chercheur en philosophie.

qui lui permettent d'agir en tant que citoyen. Dans le paradigme actuel, on tend vers une augmentation de l'entropie (du désordre) en termes de savoirs, et ce que STIEGLER suggère, c'est une autre approche qui permette de passer à une diminution de l'entropie : une néguentropie.

Par exemple, ce que nos décideurs appellent à l'heure actuelle des "smart cities", des villes intelligentes, n'en sont en fait pas. Elles utilisent des procédés de surveillance de masse et des algorithmes décisionnels classiques (comme ceux que nous avons étudiés précédemment) pour ôter aux habitants leur capacité à influencer sur l'orientation de la ville, ce qui les prive de leur pouvoir citoyen. Mais une autre manière d'envisager ces villes "connectées" serait de mettre ces techniques numériques au service des citoyens, et c'est ce que STIEGLER expérimente avec le projet de la Plaine Commune, en Seine-Saint-Denis. En utilisant le décisionnel pour rassembler des informations de capteurs, pour convoquer des réunions publiques, on a alors des citoyens qui sont à la fois chercheurs et acteurs de l'aménagement du territoire, sans que le numérique leur fauche l'herbe sous le pied.

On comprend alors que le problème n'est pas tant le numérique, mais l'usage que l'on en fait. La technique, pour STIEGLER, c'est un pharmakon. Emprunté à Jacques DERRIDA, qui a lui même emprunté ce concept à la Grèce antique, un pharmakon est ce qui est à la fois remède et poison. La technique, selon comment elle est utilisée, peut à la fois faire le mal comme faire le bien, être outil d'émancipation des masses comme l'instrument de

leur propre aliénation.

Conclusion

Nous avons vu dans ce rapport le rôle des industries culturelles dans la société postmoderne. Nous avons étudié leurs articulations au sein d'une société de l'information, hyperconnectée et massivement numérisée. Nous avons pu constater que les plateformes culturelles en ligne exploitent des procédés statistiques de big data pour sélectionner la présentation du contenu aux utilisateurs. Nous avons vu que cette sélection est tout sauf neutre, qu'elle est porteuse d'un sens qui contribue à ce que l'on appelle la politisation de l'esthétique : avec des données d'entrées a priori objectives, Netflix et les GAFAM imposent un modèle de pensée sur les consommateurs, modèle qui est, comme nous l'avons vu, hérite d'une pensée capitaliste et libérale américaine d'après-guerre. Cette politisation de l'esthétique exerce donc une forme de biopouvoir sur les consommateurs, avec de multiples conséquences : pensée unique, reproduction des inégalités sociales, perte de sens et de lien entre les individus. Netflix est une illustration de ce phénomène, son fonctionnement par algorithmes et son poids dans l'économie américaine confirment le rôle fondamental des industries culturelles.

Face à ces problèmes, nous comprenons alors qu'il est nécessaire d'avoir une éthique des données, d'envisager autrement notre traitement de l'information. Si les algorithmes ne peuvent, et ne doivent pas être des analystes du monde, alors qui ? Notre capacité à reprendre le contrôle des processus de décision qui déterminent nos vies sera déterminante pour l'avenir de l'espèce humaine.

Bibliographie

- [Bar15] Alistair Barr. Google mistakenly tags black people as "gorillas," showing limits of algorithms. *The Wall Street Journal*, jul 2015. <https://blogs.wsj.com/digits/2015/07/01/google-mistakenly-tags-black-people-as-gorillas-showing-limits>
- [Ben55] Walter Benjamin. *The Work of Art in the Age of Mechanical Reproduction*. 1955.
- [Mon13] Pietro Montani. *Bioestetica : Senso comune, tecnica e arte nell'età della globalizzazione*. 2013.
- [Pac57] Vance Packard. *The Hidden Persuaders*. 1957.
- [RD07] Balicer Ran D. Modeling infectious diseases dissemination through online role-playing games. *Epidemiology*, pages 260–261, mar 2007.
- [Sco13] Marc Scott. Kids can't use computers... and this is why it should worry you, jul 2013. <http://www.coding2learn.org/blog/2013/07/29/kids-cant-use-computers/>.
- [WAH47] Theodor W. Adorno and Max Horkheimer. *Dialektik der Aufklärung*. 1947.